Editor-in-Chief
Dr. Saif alZahir

SIGNAL PROCESSING (SPIJ)

AN INTERNATIONAL JOURNAL

ISSN : 1985-2339

Volume 9 • Issue 5 • November / December 2015
Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
http://www.cscjournals.org

SIGNAL PROCESSING: AN INTERNATIONAL JOURNAL (SPIJ)

VOLUME 9, ISSUE 5, 2015

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-2339

International Journal of Computer Science and Security is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

SPIJ Journal is a part of CSC Publishers

Computer Science Journals

http://www.cscjournals.org

SIGNAL PROCESSING: AN INTERNATIONAL JOURNAL (SPIJ)

Book: Volume 9, Issue 5, November / December 2015

Publishing Date: 31-12-2015

ISSN (Online): 1985-2339

SPIJ Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2015

EDITORIAL PREFACE

This is *Fifth* Issue of Volume *Nine* of the Signal Processing: An International Journal (SPIJ). SPIJ is an International refereed journal for publication of current research in signal processing technologies. SPIJ publishes research papers dealing primarily with the technological aspects of signal processing (analogue and digital) in new and emerging technologies. Publications of SPIJ are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics covers by SPIJ are Signal Filtering, Signal Processing Systems, Signal Processing Technology and Signal Theory etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 10, 2016, SPIJ will be appearing with more focused issues related to signal processing studies. Besides normal publications, SPIJ intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of SPIJ is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position SPIJ as one of the top International journal in signal processing, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to signal processing fields.

SPIJ editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. SPIJ provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
Signal Processing: An International Journal (SPIJ)

TABLE OF CONTENTS

Volume 9, Issue 5, November / December 2015

Pages

Iterative Soft Decision Based Complex K-best MIMO Decoder

Mehnaz Rahman *mehnaz@tamu.edu*
Department of ECE
Texas A&M University
College Station, Tx- 77840, USA

Gwan S. Choi *gchoi@ece.tamu.edu*
Department of ECE
Texas A&M University
College Station, Tx- 77840, USA

Abstract

This paper presents an iterative soft decision based complex multiple input multiple output (MIMO) decoding algorithm, which reduces the complexity of Maximum Likelihood (ML) detector. We develop a novel iterative complex K-best decoder exploiting the techniques of lattice reduction for 8×8 MIMO. Besides list size, a new adjustable variable has been introduced in order to control the on-demand child expansion. Following this method, we obtain 6.9 to 8.0 dB improvement over real domain K-best decoder and 1.4 to 2.5 dB better performance compared to iterative conventional complex decoder for 4th iteration and 64-QAM modulation scheme. We also demonstrate the significance of new parameter on bit error rate. The proposed decoder not only increases the performance, but also reduces the computational complexity to a certain level.

Keywords: Complex K-best Algorithm, MIMO, Lattice Reduction, Iterative Soft Decoding, SE Enumeration.

1. INTRODUCTION

With the advancement of wireless system, MIMO has been acclaimed by different wireless standards such as IEEE 802.11n, IEEE 802.16e to achieve high data rates and performance with ML or near-ML algorithms. Most of these standards have a specified minimum error rate to guarantee quality of service (QoS), which is either in bit error rate (BER) or packet error rate (PER) (e.g., 10^{-6} is specified as maximum tolerable BER according to IEEE 802.11n standard [1]).

The main challenge behind MIMO system is maintaining the performance of the receiver with low complexity. Several algorithms have been proposed to address the issue, offering different tradeoffs between complexity and performance. The ML detector minimizes BER performance through exhaustive search. However, with increased number of transmitting and receiving antennas, and bits in modulation, the complexity grows exponentially [2, 3]. In contrast, sub-optimal detectors with polynomial complexity such as zero forcing (ZF), minimum mean square error (MMSE) detectors etc. have been developed with significant performance loss.

Recently, lattice reduction (LR) has been proposed in order to achieve high performance, yielding much less complexity than the conventional K-best decoder [4, 5, 6]. LR-aided detector can achieve the same diversity as of ML at the cost of some performance loss [7, 8]. Later, it is implemented in complex domain [9]. All of these suboptimal detectors mentioned above were based on hard decision. Therefore, soft input-soft output (SISO) detectors, suitable for subsequent iterative decoding are introduced in [10].

Researchers further improved these SISO detectors with low density parity check (LDPC) decoder [10, 11]. The output of LDPC decoder is fed back to the detector for updating the soft value in order to achieve better performance. This is called iterative decoding. It can achieve near Shannon performance with less computational complexity compared to other near Shannon decoders [12].

This paper presents an iterative soft decision based complex K-best decoder, which enables the utility of lattice reduction and complex SE enumeration in MIMO decoder. For soft decoding, the log likelihood ratio (LLR) values for LDPC decoder are first computed from the K best candidates and then, they are fed back to LLR update unit as inputs to the next iteration. This process of iterations is continued until the gain of subsequent iteration becomes saturated. Then, the last updated LLR values are forwarded to the LDPC decoder for final detection. Besides list size K, a new tunable parameter Rlimit is introduced in order to enable adaption of the computation of on-demand child expansion for choosing the list candidates.

We compare the results of our proposed decoder with those of iterative conventional complex decoder in [11] and LR-aided real decoder in [13]. For 8×8 MIMO, it achieves 6.9 to 8.0 dB improvement over real domain K-best decoder and 1.4 to 2.5 dB better performance comparing to conventional complex K-best decoder for 4th iteration and 64 QAM modulation scheme. If we consider only 1st iteration, the gain increases to more than 9.0 dB and 2.9 dB comparing with iterative real and complex decoder respectively. This provides significant gain in terms of practical execution. The effect of Rlimit is also analyzed to achieve the maximum performance. The introduction of Rlimit also leads to complexity reduction significantly.

The rest of the paper is organized as follow. In Section II we introduce soft decision based complex MIMO decoding algorithm. Then, Section III presents the results of our studied cases and Section IV concludes this paper with a brief overview.

2. SYSTEM MODEL

Let us consider a MIMO system operating in M-QAM modulation scheme and having $\mathrm{N_T}$ transmit antenna and $\mathrm{N_R}$ receiving antenna as:

$$y = Hs + n, \tag{1}$$

where $\mathrm{s} = [\mathrm{s}_1, \mathrm{s}_2, \ldots.\ \mathrm{s}_{\mathrm{N_T}}]^{\mathrm{T}}$ is the transmitted complex vector, H is complex channel matrix and $\mathrm{y} = [\mathrm{y}_1, \mathrm{y}_2, \ldots.\ \mathrm{y}_{\mathrm{N_R}}]^{\mathrm{T}}$ is the symbol of $\mathrm{N_R}$ dimensional received complex vector. Noise is a $\mathrm{N_R}$ dimensional complex additive white Gaussian noise (AWGN) with variance and power σ^2 and N_0 respectively. Noise can be represented by $n = [\mathrm{n}_1, \mathrm{n}_2, \ldots.\ \mathrm{n}_{\mathrm{N_R}}]^{\mathrm{T}}$.

The detector solves for the transmitted signal by solving non-deterministic hard problem:

$$\hat{\mathrm{s}} = \arg_{\tilde{\mathrm{s}} \in \mathrm{S}^{\mathrm{N_T}}} \min \|\mathrm{y} - \mathrm{H}\tilde{\mathrm{s}}\|^2 . \tag{2}$$

Here, $\tilde{\mathrm{s}}$ is the candidate complex vector, and $\hat{\mathrm{s}}$ is the estimated transmitted vector [8]. In the expression, $\|.\|$ denotes 2-norm. This MIMO detection problem can be represented as the closest point problem in [14], which is an exhaustive tree search through all the set of all possible lattice points in $\tilde{s} \in \mathrm{S}^{\mathrm{N_T}}$ for the global best in terms of Euclidean distance between y and $H\tilde{s}$. Each transmit antenna offers two level of search for real-domain MIMO detection: one for real and the other for imaginary part. However, in complex domain detection method, only one level of search is required for each antenna.

ML detector performs a tree search through the set of all possible branches from root to node, thereby achieves the best performance. However, its complexity increases exponentially with the number of antennas and constellation bits. Therefore, suboptimal detectors such as LR-aided detector comes into consideration.

2.1 LR-aided Decoder

Lattice reduction provides more orthogonal basis with short basis vector from a given integer lattice. Hence, it effectively reduces the effects of noise and mitigates error propagation in MIMO detection. Since lattice reduction is most effective for unconstrained boundary, the following change is made to (2) to obtain a relaxed search.

$$\hat{s} = \arg_{\tilde{s}\in\mathcal{U}^{N_T}} \min\|y - H\tilde{s}\|^2 , \tag{3}$$

where $\mathcal{U}$ is unconstrained complex constellation set $\{\ldots,\ -3+j,\ -1-j,\ -1+j,\ 1-j,\ldots\}$. Hence, $\hat{s}$ may not be a valid constellation point. This is resolved by quantizing $\hat{s} = \mathcal{Q}\ (\hat{s})$, where $\mathcal{Q}(.)$ is the symbol wise quantizer to the constellation set S.

However, this type of naive lattice reduction (NLD) does not acquire good diversity multiplexing tradeoff (DMT) optimality. Hence, MMSE regularization is employed as proposed in [15, 16], where the channel matrix and received vector are extended as $\overline{H}$ and $\bar{y}$:

$$\overline{H} = \begin{bmatrix} H \\ \sqrt{\dfrac{N_0}{2\sigma_2^2}}I_{N_T} \end{bmatrix}, \qquad \bar{y} = \begin{bmatrix} y \\ 0_{N_T\times 1} \end{bmatrix}, \tag{4}$$

where $0_{N_T\times 1}$ is a $N_T \times 1$ zero matrix and I_{N_T} is a $N_T \times N_T$ complex identity matrix [17, 18]. Then, Eq. (3) can be represented as:

$$\hat{s} = \arg_{\tilde{s}\in\mathcal{U}^{N_T}} \min\|\bar{y} - \overline{H}\tilde{s}\|^2 . \tag{5}$$

Hence, lattice reduction is applied to $\overline{H}$ to obtain $\widetilde{H} = \overline{H}T$, where T is a unimodular matrix. Eq. (5) then become:

$$\hat{s} = T\arg\ \min_{\tilde{z}\in\mathcal{U}^{N_T}} \left(\left\|\tilde{y} - \widetilde{H}\tilde{z}\right\|^2 + (1+j)_{N_T\times 1}\right), \tag{6}$$

where $\tilde{y} = (\bar{y} - \overline{H}(1+j)_{N_{T\times 1}})/2$ is the complex received signal vector and $1_{2N_T\times 1}$ is a $2N_T \times 1$ one matrix. After shifting and scaling, (6) became the following one.

$$\hat{s} = T\tilde{z} + (1+j)_{N_T\times 1} . \tag{7}$$

Lattice reduction is an NP complete problem. However, polynomial time algorithms such as Lenstra-Lenstra-Lovasz (LLL) algorithm in [19] can find near orthogonal short basis vectors.

2.2 Complex K-Best LR-Aided MIMO Detection

Complex K-best LR-aided detection offers a breadth first tree search algorithm, which is performed sequentially starting at N_{th}-level. First, it requires QR decomposition on $\widetilde{H} = QR$, where Q is a $(N_R + N_T) \times (N_R + N_T)$ orthonormal matrix and R is a $(N_R + N_T) \times N_T$ upper triangular matrix. Then (6) is reformulated as

$$\hat{s} = T \arg\min_{\tilde{z} \in \mathcal{U}^{N_T}} \left(\|\breve{y} - R\tilde{z}\|^2 + (1+j)_{2N_T \times 1} \right), \tag{8}$$

where $\breve{y} = Q^T\tilde{y}$. The error at each step is measured by the partial Euclidean distance (PED), which is an accumulated error at a given level of the tree. For each level, the K best nodes are selected and passed to the next level for consideration. At the end, all the K paths through the tree are evaluated to find the one with minimum PED. The number of valid children for each parent in LR-aided K-best algorithm is infinite. Hence, in our proposed algorithm, the infinite children issue is addressed by calculating K best candidates using complex on-demand child expansion.

2.3 Complex On-demand Expansion

Complex on-demand expansion exploits the principle of Schnorr-Euchner (SE) enumeration [9, 20]. The strategy employs expanding of a node (child) if and only if all of its better siblings have already been expanded and chosen as the partial candidates [21, 22]. Hence, in an order of strict non-decreasing error, K candidates are selected. In conventional complex SE enumeration, expansion of a child can be of two types: Type I, where the expanded child has same imaginary part as its parent, i.e. enumerating along the real axis; and Type II for all other cases. The example of conventional complex on-demand SE enumeration is shown in Fig. 1.

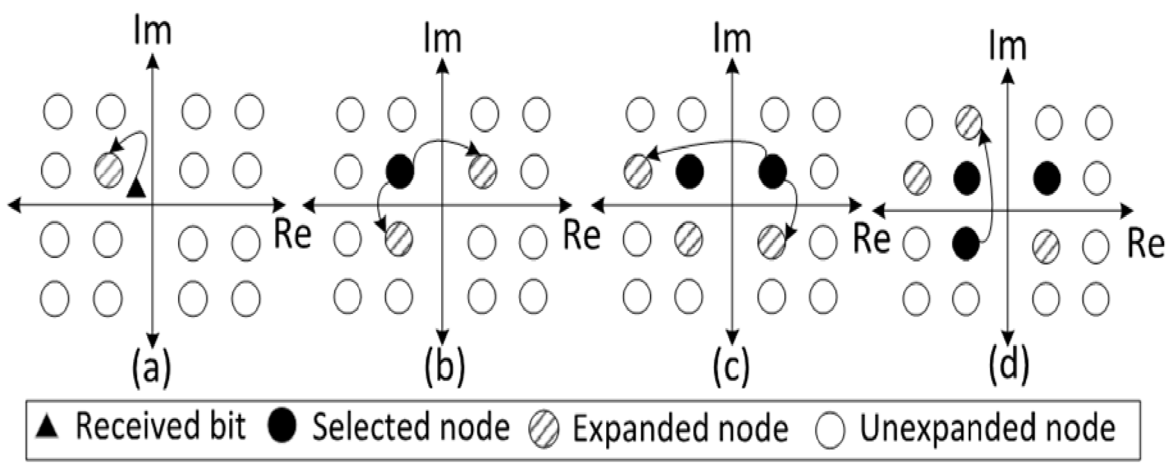


FIGURE 1: Complex SE Enumeration.

First received symbol is rounded to the nearest integer as shown in Fig. 1(a), which includes quantizing of both real and imaginary components of the signal to the nearest integer. Type-I candidate will be expanded two times along real and imaginary axis using SE enumeration, and the two expanded nodes are considered candidates, as demonstrated in Fig. 1(b). Then, the one with the minimum PED is chosen, and expanded for further calculation depending on the type. As in Fig. 1(c), the chosen node is of type I, it will be expanded to 2 more nodes. If the chosen node is of Type II, as shown in Fig. 1(d), it will be expanded only along imaginary axis.

The number of nodes needs expanded at any level of the tree is considered as measurement of complexity analysis. The worst case scenario will be if all the nodes chosen are of type I. Then, at an arbitrary level of tree, the number of expanded nodes is bounded by $K + 2(K-1)$. Taken over the entire tree, the complexity for the search becomes $3N_TK - 2N_T$. Comparing with the real domain detection algorithm in [13], the number of the expanded nodes is $4N_TK - 2N_T$. For instance, with K as 4 and N_T equal to 8, the number of expanded node is 80 and 112 considering complex and real decoder respectively. Hence, complex SE enumeration requires less calculation, thereby reduces hardware complexity.

In this paper, we introduce another parameter, Rlimit while performing the complex on demand child expansion. In contrast with the conventional one, the type of a child is not considered for further expansion. The example of improved complex SE enumeration with Rlimit as 3 is given in Fig. 2.

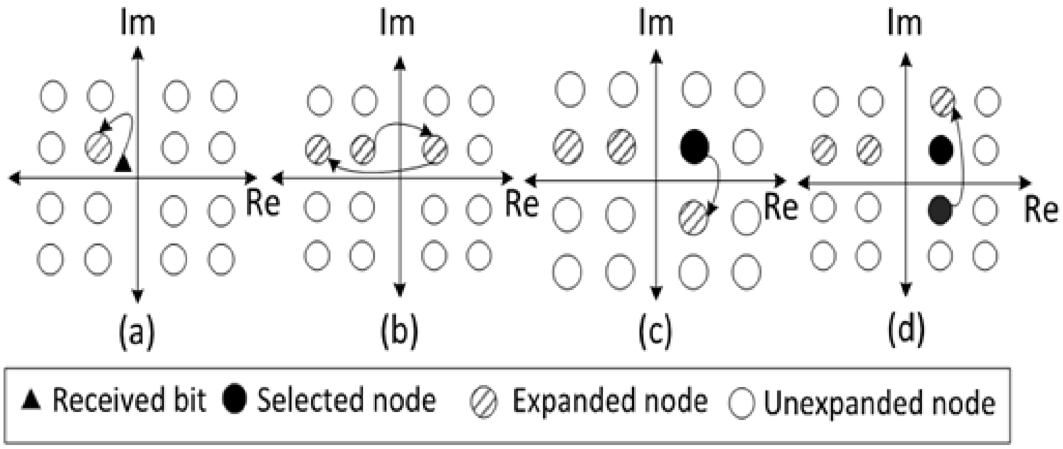


FIGURE 2: Improved Complex SE Enumeration with Rlimit as 3.

As shown in Fig. 2, after rounding the received symbol to the nearest integer, first real SE enumeration is performed to calculate Rlimit candidates. Hence, it means that, all the calculated nodes up to Rlimit will have same imaginary value, as demonstrated in Fig. 2(b). Then, the one with minimum PED is selected and expanded only along the imaginary axis using imaginary domain SE enumeration. This process is continued till K nodes are selected at that level of tree as presented in Fig. 2(c)-(d).

The complexity analysis of the improved child expansion proceeds as follows. At any level of tree search, first $\mathrm{KRlimit}$ nodes need to be expanded. After that, only imaginary domain SE enumeration will be performed. Hence, considering the worst case, the total number of nodes calculated at each level is $\mathrm{KRlimit} + (K - 1)$. For N_T levels, the complexity becomes $N_TK\,(\mathrm{Rlimit} + 1) - N_T$. Therefore, introduction of Rlimit may increase the complexity as evidenced in result section, although offers better BER performance comparing to the conventional one. However, comparing with the real domain detection, the total complexity is still less. We have used improved complex on demand expansion to perform the list calculation and then the chosen K paths are passed to the iterative soft input soft output (SISO) decoder.

2.4 Iterative Soft Decoding

LDPC decoder in [12] calculates approximate LLR from the list of possible candidates using (9).

$$L_E(x_k|Y) \approx \frac{1}{2} \max_{x \in X_{k,\ +1}} \left\{-\frac{1}{\sigma^2} \parallel y - Hs \parallel^2 + x_{[k]}^T \,.\, L_{A,[k]}\right\} - \frac{1}{2} \max_{x \in X_{k,\ -1}} \left\{\frac{1}{\sigma^2} \parallel y - Hs \parallel^2 + x_{[k]}^T \,.\, L_{A,[k]}\right\}, \tag{9}$$

here $\mathbf{x}_{[\mathbf{k}]}^{\mathbf{T}}$ and $\mathbf{L}_{\mathbf{A},[\mathbf{k}]}$ are the candidates values {-1 or 1} and LLR values except k-th candidate respectively. In order to perform the soft decoding, the LLR values are first computed at the last layer of K-best search. Then, the soft values are fed into the iterative decoder for the subsequent iteration. This process continues until the difference in error levels between the last two iterations becomes negligible. Lastly, the updated LLR values are used for hard decision.

From the perspective of hardware design as proposed in [16, 23], the LLR calculation unit takes one of the candidates at a given time and computes the LLR value. Then, the new LLR is compared to the maximum of previous LLRs. Hence, this unit has to keep track of 2 values for each LLR; one for those whose k-th bit of the candidate list is 1 (Lambda-ML), and the other for 0 (Lamda-ML-bar). After that, the LLR values are calculated as the subtraction of Lambda-ML and Lambda-ML-bar divided by 2.

3. SETUP AND RESULTS

This section demonstrates the performance of the proposed iterative soft decision based complex K-best decoder. The test and simulation environment has been implemented using IEEE 802.11n standard. All the simulations are for 8×8 MIMO with different modulation schemes. The ratio of the signal and noise power is considered as signal to noise ratio (SNR).

We first analyze the performance of four iterations of our proposed decoder for different modulation scheme. Then, the effect of Rlimit on BER performance is shown for 64QAM modulation scheme. Finally, we demonstrate the comparison of performance of our proposed work with that of iterative conventional complex decoder and real decoder for 64QAM modulation scheme.

The total number of the nodes expanded for 8×8 MIMO is considered as measurement of the complexity analysis. For iterative real decoder, as shown in [13] the improvement gained from 3rd to 4th iteration is limited and negligible for iterations beyond that. Hence, we consider BER versus SNR curve up to four iterations in order to perform comparison among maximum performance.

3.1 Simulation and Analysis

The performance of four iterations of our proposed soft decision based complex decoder for QPSK modulation scheme is presented in Fig. 3.

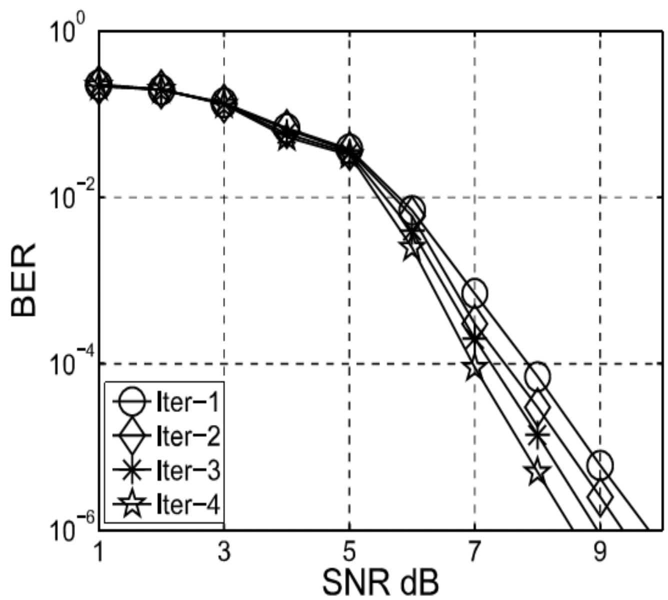
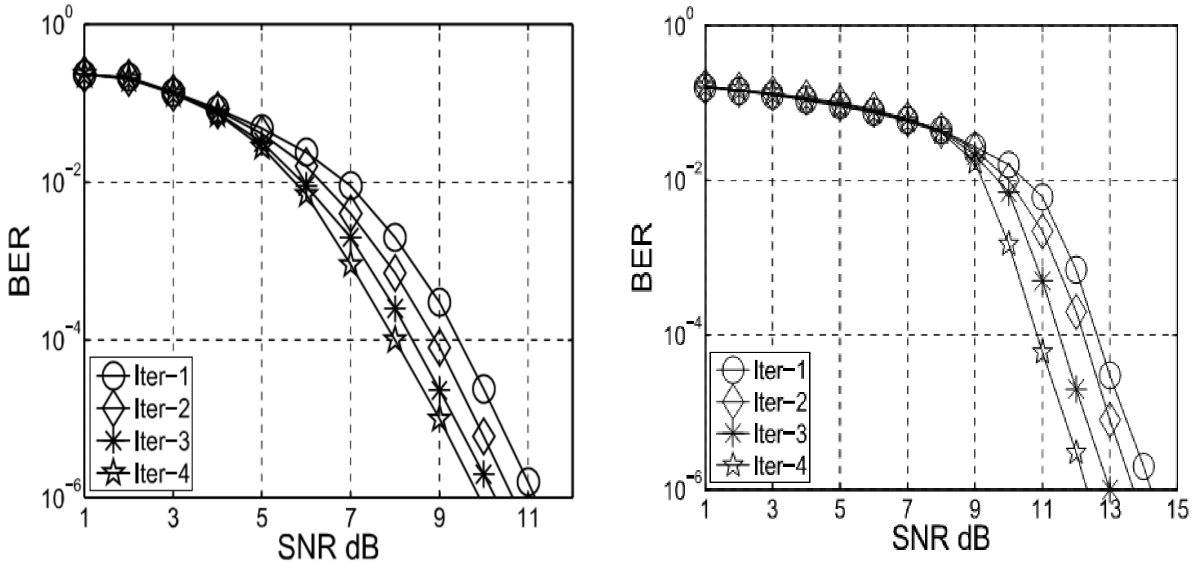


FIGURE 3: BER vs SNR curve of the first 4 iterations of iterative complex decoder for 8 x 8 MIMO system with K as 4 and QPSK modulation scheme.

As shown in Fig. 3, for QPSK modulation with list size, K of 4 and Rlimit of 4, we observe 0.4 dB improvement in BER due to the 2nd iteration at the BER of 10^{-6}. When we compare the performance of 1st iteration with 3rd and 4th one, the improvement increases to 0.7 dB and 1.0 dB respectively.



(a) 16 QAM (b) 64 QAM

FIGURE 4: BER vs SNR curve of the first 4 iterations of iterative complex decoder for 8 x 8 MIMO system with K as 4.

Next, Fig. 4 represents the performance curve of 4th iteration for 16 QAM and 64 QAM modulation scheme. As demonstrated in Fig. 4(a), the performance of 2nd iteration is approximately 0.4 dB better than the 1st one with K as 4 and Rlimit set to 4 for 16 QAM modulation scheme. When increasing the iteration, the performance improves by 0.8 dB for the 3rd and 1.1 dB for the 4th iteration compared to the 1st one.

For 64QAM having same K as 16QAM, the improvement due to the 2nd iteration is 0.4 dB, as shown in Fig. 4(b). If we then compare the 3rd and 4th iteration with respect to the 1st one, the improvements are 0.8 dB and 1.0 dB respectively. By extensive simulation, we observe that the performance does not improve beyond 4th iteration. Therefore, with iteration number, the performance between i-th and (i+1)-th iteration gets saturated.

3.2 Effect of Rlimit on BER

The effect of Rlimit, as discussed in section 3.2 for proposed complex on demand child expansion is shown in Fig. 5. It represents BER performance for the 4th iteration over different SNR considering 8×8 MIMO and 64 QAM modulation scheme with list size, K as 4.

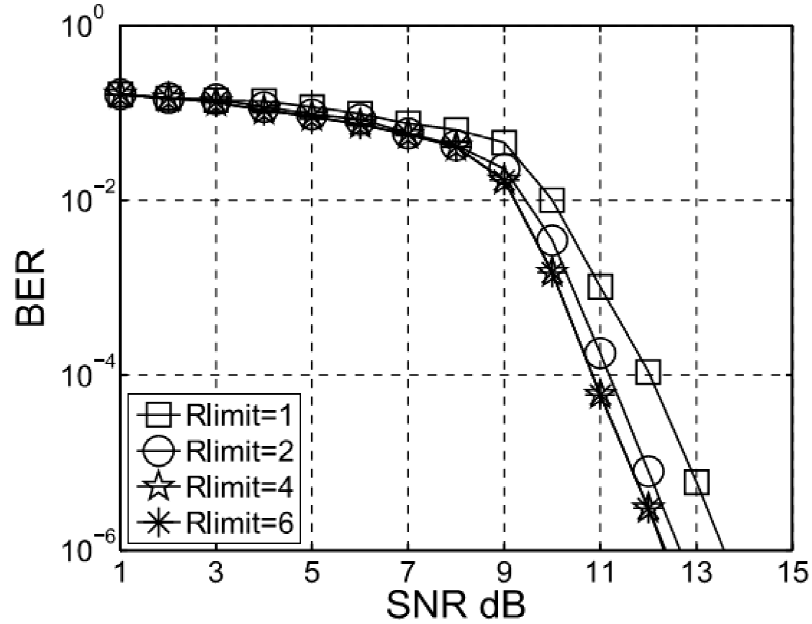


FIGURE 5: BER vs SNR curve of the 4th iteration of iterative complex decoder for 8 x 8 MIMO with 64QAM modulation scheme having K as 4.

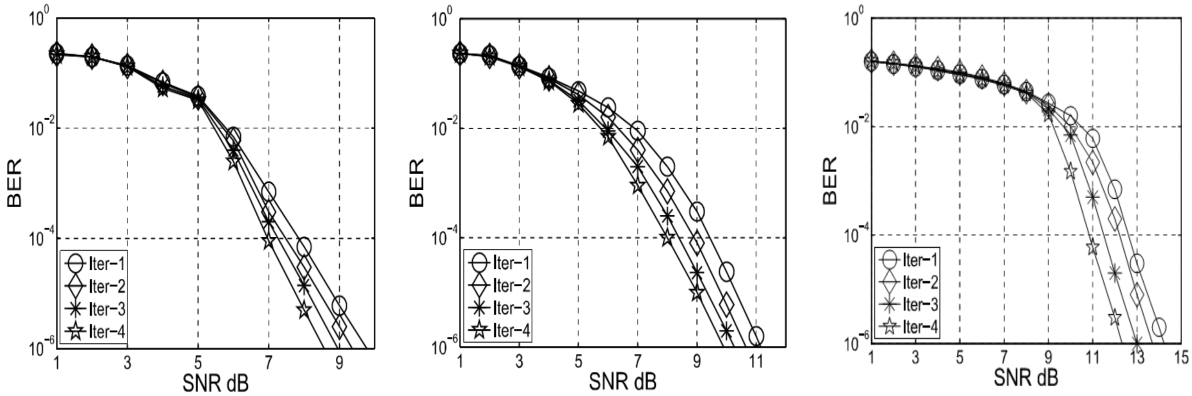
It is evident that if the value of Rlimit is increased, the performance improves and then, it saturates with Rlimit. On the other hand, decreasing Rlimit will degrade BER. Hence, as shown in Fig. 4, when Rlimit increases from 4 to 6, the performance get saturated. However, decreasing the Rlimit to 2 and then 1, degrades the performance 0.3 dB and 1.1 dB respectively.

Similar curves can be obtained considering 1st, 2nd and 3rd iteration of proposed iterative decoder for different Rlimit. By extensive simulation, we also observe that, for QPSK and 16 QAM modulation schemes, Rlimit set to 4 can obtain the maximum performance. Even if the value of Rlimit is increased, the performance does not improve.

3.3 Comparison of Performance

The comparison of the performance of different iterations of our proposed work with those of iterative conventional complex decoder and real decoder for 64QAM modulation scheme with K as 4 is presented in Fig. 6.

For proposed iterative complex decoder, we have considered Rlimit as 1, 2 and 4 for performance evaluation. Simulation with Rlimit higher than 4 is not considered, since it is the minimum value required to achieve the maximum performance. We consider BER versus SNR curve up to four iterations in order to perform comparison among maximum performance, as shown in [14] .



(a) (b) (c)

FIGURE 6: BER vs SNR curves of proposed iterative complex decoder, conventional complex and real decoder for 8 x 8 MIMO with 64QAM modulation scheme having K as 4. For proposed decoder, Rlimit is set to 1, 2, and 4.

As demonstrated in Fig. 6(a), a 3.4 dB improvement in performance can be achieved comparing the 1st iteration of proposed decoder with that of conventional iterative complex decoder with Rlimit as 4 at the BER of 10^{-6}. When Rlimit is changed to 2 and 1, the improvements become 3 dB and 2.9 dB respectively. We also compare the performance of proposed decoder with that of the iterative real decoder for 1st iteration [14]. As presented in Fig.6(a), 9.0 dB to 9.5 dB improvement can be achieved using Rlimit as 1 to 4.

Next, as shown in 6(b), a 1.5 dB improvement can be obtained if we consider the performance of 1st iteration of proposed decoder with the 4th iteration of conventional complex one using Rlimit as 4. Decreasing Rlimit to 2 and 1 results in 1 dB and 0.8 dB improvement. Comparing to the 4th iteration of iterative real decoder, 6.1 dB to 6.8 dB SNR gain can be achieved using Rlimit set to 1 to 4 respectively.

Fig. 6(c) presents the comparison curves considering the 4th iteration of iterative decoders. As demonstrated in the figure, a 2.4 dB improvement can be obtained using Rlimit as 4 at the BER of 10^{-6} comparing the conventional iterative complex decoder. In addition, when simulating for Rlimit as 2 and 1, the gain becomes 2.2 dB and 1.4 dB respectively. Similar analysis can be performed comparing to the 4th iteration of iterative real decoder. A gain of 6.9 dB to 8.0 dB can be achieved for Rlimit set to 1 to 4.

Then, we have performed the computational complexity analysis for the presented work. The total number of the nodes expanded for 8 x 8 MIMO is considered as measurement of the analysis. Complexity analysis of proposed and conventional complex decoder is shown in Table 1.

K	Proposed		Conv. Complex	Proposed vs Conv. (in dB)		
	Rlimit	Node	Node	1st vs 1st	4th vs 4th	1st vs 4th
4	1	56	80	2.9	1.4	0.8
4	2	88	80	3.0	2.2	1.0
4	4	152	80	3.4	2.5	1.5

TABLE 1: Complexity Analysis of Conventional and Proposed Complex Decoder.

As tabulated in Table 1, for iterative conventional complex decoder, we need to perform 80 calculations for K equal to 4. Although our proposed decoder calculates 56, 88 and 152 nodes using same list size and Rlimit set to 1, 2, and 4 respectively. Hence, with less computational complexity, the proposed decoder can achieve 1.4 dB better performance than that of conventional one for the 4th iteration. However, 2.2 to 2.5 dB gain can be achieved by tolerating higher computational complexity using proposed complex decoder. Considering 1st iteration with same level of complexity, 2.9 dB to 3.4 dB gain can be achieved using proposed decoder. Next, complexity analysis of proposed and iterative real decoder is presented in Table 2.

K	Proposed		Real	Proposed vs Real (in dB)		
	Rlimit	Node	Node	1st vs 1st	4th vs 4th	1st vs 4th
4	1	56	112	9.0	6.9	6.1
4	2	88	112	9.1	7.7	6.5
4	4	152	112	9.5	8.0	6.8

TABLE 2: Complexity Analysis of Iterative Real and Proposed Complex Decoder.

As shown in Table 2, the number of the nodes need to be expanded for LR-aided real decoder [14] for list size 4 is equal to 112. Considering the same list size, proposed complex decoder requires 56, 88 and 152 node expansion for Rlimit set to 1, 2 and 4 respectively. Hence, proposed decoder can achieve 6.9 dB to 7.7 dB better performance even with less computational complexity comparing with the iterative real one. Allowing more complexity, can increase the performance to 8.0 dB. If we consider the performance of only 1st iteration, with same level of complexity the proposed decoder can attain 9.0 to 9.5 dB improvement comparing with the real one.

Therefore, our iterative soft complex decoder with Rlimit offers a trade-off between performance and complexity for different iteration. It not only increases the performance, but also can reduce complexity to a certain level.

4. CONCLUSION

In this paper, an iterative soft decision based complex domain K-best decoder is proposed exploiting the improved complex on-demand child expansion. It includes the use of LR algorithm in order to achieve orthogonality among the constellation points reducing the effect of noise. An additional parameter, Rlimit is introduced to tune the complexity of computation with improvement in BER performance. Reduction of computational complexity directly results to less power consumption of the decoder as well.

We also compare the result of 4th iteration of our proposed decoder with iterative conventional complex decoder and obtain 1.4 to 2.5 dB improvement at the BER of 10^{-6} for 8×8 MIMO and 64 QAM modulation scheme with comparable complexity. Comparing with iterative LR-aided real domain decoder, the improvement increases more than 7.0 dB with less computational complexity. Although more than 2.9 dB and 9.0 dB gain can be achieved with same level of complexity comparing 1st iteration of proposed decoder with that of conventional iterative complex and real decoder respectively.

Future work of this proposed decoder includes evaluating the detector performance using additional channel and simulation environment and also implementing the algorithm on FPGA and so on.

5. REFERENCES

[1] "IEEE Standard for Information Technology- Local and Metropolitan Area Networks-Specific Requirements- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput." IEEE Standard 802.11n-2009 (Amendment to IEEE Standard 802.11-2007 as amended by IEEE Standard 802.11k-2008, IEEE Standard 802.11r-2008, IEEE Standard 802.11y-2008, and IEEE Standard 802.11w-2009), pp. 1-565, Oct. 2009.

[2] J. Jalden and B. Otterston. “On the Complexity of Sphere Decoding in Digital Communications.” IEEE Transaction on Signal Processing, vol. 53, no. 4, pp. 1474-1484, Apr. 2005.

[3] I. Lai, G. Ascheid, H. Meyr and T.-D. Chiueh. "Low-Complexity Channel-Adaptive MIMO Detection with Just-Acceptable Error Rate." IEEE 69th Vehicular Technology Conference: VTC-2009 Spring, Apr. 2009, pp. 1-5.

[4] M. Shabany and P. Glenn Gulak. “The Application of Lattice-Reduction to the K-Best Algorithm for Near-Optimal MIMO Detection.” IEEE International Symposium on Circuits and Systems (ISCAS), May 2008, pp. 316-319.

[5] C. Windpassinger and R. Fischer. “Low-Complexity Near-Maximum Likelihood Detection and Precoding for MIMO Systems Using Lattice Reduction.” Proceeding IEEE Information Theory Workshop, Mar. 2003, pp. 345-348.

[6] Q. Zhou and X. Ma. “An Improved LR-aided K-Best Algorithm for MIMO Detection.” Proceeding IEEE International Conference on Wireless Communication and Signal Processing (WCSP), Oct. 2012, pp. 1-5.

[7] X. Qi and K. Holt. “A Lattice-Reduction-Aided Soft Demapper for High-Rate Coded MIMO-OFDM Systems.” IEEE Signal Processing Letters, vol. 14, no. 5, pp. 305-308, May 2007.

[8] B. M. Hochwald and S. Ten Brink. "Achieving Near-Capacity on a Multiple-Antenna Channel." IEEE Transactions on Communications, vol. 51, no. 3, pp. 389-399, Mar. 2003.

[9] F. Sheikh, E. Wexler, M. Rahman, W. Wang, B. Alexandrov, D. Yoon, A. Chun and A. Hossein. "Channel-Adaptive Complex K-Best MIMO Detection Using Lattice Reduction." IEEE Workshop on Signal Processing Systems (SiPS), pp. 1-6, Oct. 2014.

[10] M. Rahman, E. Rohani and G. Choi. "An Iterative Soft Decision Based Adaptive K-Best Decoder Without SNR Estimation." Asilomer Conference on Signals, Systems and Computers, Nov. 2014, pp. 1016-1020.

[11] Q. Wen, Q. Zhou, C. Zhao and X. Ma. "Fixed-Point Realization of Lattice-Reduction Aided MIMO Receivers With Complex K-Best Algorithm." IEEE International Conference on Acoustics, Speech and Signal Processing, May 2013, pp. 5031-5035.

[12] K. Gunnam, G. Choi, W. Weihuang and M. Yeary. "Multi-Rate Layered Decoder Architecture for Block LDPC Codes of the IEEE 802.11n Wireless Standard." IEEE International Symposium on Circuits and Systems (ISCAS), May 2007, pp. 1645-1648.

[13] M. Rahman, E. Rohani and G. Choi. "An Iterative LR-Aided MMSE Extended Soft MIMO Decoding Algorithm." International Conference on Computing, Networking and Communications, California, Feb. 2015.

[14] E. Agrell, T. Eirksson, A. Vardy and K. Zeger. "Closest Point Search in Lattices." IEEE Transaction on Information Theory, vol. 48, no. 8, pp. 2201-2214, Aug. 2002.

[15] J. Jalden and P. Elia. "DMT Optimality of LR-Aided Linear Decoders for a General Class of Channels, Lattice Designs, and System Models." IEEE Transaction on Information Theory, vol. 56, no. 10, pp. 4765-4780, Oct 2010.

[16] M. Rahman, E. Rohani, J. Xu and G. Choi. "An Improved Soft Decision Based MIMO Detection Using Lattice Reduction." International Journal of Computer and Communication Engineering, vol. 3, no. 4, pp. 264-268, Apr. 2014.

[17] J. Jalden and P. Elia. "DMT Optimality of LR-Aided Linear Decoders for a General Class of Channels, Lattice Designs, and System Models." IEEE Transaction on Information Theory, vol. 56, no. 10, pp. 4765-4780, Oct. 2010.

[18] M. Taherzadeh and A. Khandani. "On the Limitations of the Naive Lattice Decoding." IEEE Transaction on Information Theory, vol. 56, no. 10, pp. 4820-4826, Oct. 2010.

[19] A. K. Lenstra, H. W. Lenstra and L. Lovasz. "Factoring Polynomials with Rational Coefficients." Mathematische Annalen, vol. 261, no. 4, pp. 515-534, Dec. 1982.

[20] M. Mahdavi and M. Shabany. "Novel MIMO Detection Algorithm for High-Order Constellations in the Complex Domain." IEEE Transaction on VLSI Systems, vol. 21, no. 5, pp. 834-847, May 2013.

[21] M. Shabany and P. Glenn Gulak. "The Application of Lattice-Reduction to the K-Best Algorithm for Near-Optimal MIMO Detection." IEEE International Symposium on Circuits and Systems (ISCAS), May 2008, pp. 316-319.

[22] C. P. Schnorr and M. Euchner. "Lattice basis reduction: Improved practical algorithms and solving subset sum problems." Mathematical Programming, vol. 66, pp. 181–191, Aug. 1994.

[23] M. Rahman and Gwan S. Choi. "Fixed Point Realization of Iterative LR-Aided Soft MIMO Decoding Algorithm." An International Journal on Signal Processing, vol. 9, issue 2, pp. 14-24, May 2015.

Spatialization Parameter Estimation in MDCT Domain for Stereo Audio

K. Suresh *sureshk@cet.ac.in*
Department of Electronics & Communication
Government Engineering College
Wayanad, Kerala, India, 670644

Akhil Raj R. *akhilraj.89@gmail.com*
Department of Electronics & Communication
College of Engineering, Thiruvananthapuram
Kerala, India, 695016

Abstract

For representing multi-channel audio at low bit rate parametric coding techniques are used in many audio coding standards. An MDCT domain parametric stereo coding algorithm which represents the stereo channels as the linear combination of the 'sum' channel derived from the stereo channels and a reverberated channel generated from the 'sum 'channel has been reported in literature. This model is inefficient in capturing the stereo image since only four parameters per sub-band is used as spatialization parameters. In this work we improve this MDCT domain parametric coder with an augmented parameter extraction scheme using an additional reverberated channel. We further modify the scheme by using orthogonalized de-correlated channels for analysis and synthesis of parametric stereo. A synthesis scheme with perceptually scaled parameter set is also introduced. Finally we present, subjective evaluation of the different parametric stereo schemes using MUSHRA test and the increased the perceptual audio quality of the synthesized signals are evident from these test results.

Keywords: Parametric Audio Coding, MDCT, Parametric Stereo.

1. INTRODUCTION

For multichannel audio compression with reasonable quality at low bit-rates, parametric coding has emerged as a suitable method with many potential applications [1,4,11]. In multichannel audio, significant amount of inter channel redundancies present along with perceptual irrelevancies and statistical redundancies. Effective removal of the inter channel redundancy and perceptual irrelevancy is required for low bit rate compression of multichannel audio. Considering the constituent channel data individually, we can apply mono audio compression methods to remove perceptual irrelevancies and statistical redundancies. To remove inter channel redundancies, cross channel prediction based methods were suggested [2]. However, in most of the cases, the correlation between the channels is low and the cross channel prediction based methods will not result in significant compression [3]. Another strategy for suppressing the inter channel redundancy is through parametric coding. In parametric coding, the encoded data consists of a mono 'sum' channel derived from the individual channels and model parameters representing the spatialization cues. Binaural cue coding (BCC) introduced in [8,9,10] and parametric stereo coding method introduced in [11] are examples for parametric multichannel audio coding. BCC uses inter channel level difference (ICLD), inter channel time difference (ICTD) and inter channel coherence (ICC) as parameters for spatial audio modeling. In the encoder, a 'sum' channel is derived by adding the individual channels followed by energy equalization. The 'sum' channel is compressed using any of the existing audio coding algorithms. The spatialization parameters are extracted in a frame by frame basis and quantized for compact representation. In the decoder, the multichannel audio is synthesized using the 'sum' channel signal and the spatialization parameters; to synthesis the required level of correlated side channels, decorrelated signals are generated from the ‘sum’ channel. In the human auditory system, the processing of binaural cues is performed on a non-uniform frequency scale [12,13]. Hence, in order to estimate the spatial parameters from the given input signal, it is desirable to transform its time domain representations to a representation that resembles this non uniform frequency scale. This transformation is achieved either by using a hybrid Quadrature Mirror Filter (QMF) bank or by grouping a number of bands of a uniform transform such as DFT [9,11]. However, in practice, for audio coding purpose, spectral representations such as Modified Discrete Cosine Transform (MDCT) are used which has the advantage of time domain alias cancellation and better energy compaction. Hence additional filter bank analysis or transform is needed for the parameter extraction in the encoder and for the synthesis in the decoder. The spatialization parameter extraction and ‘sum’ channel formation is done as a pre-processing step in the encoder; conversely, the stereo synthesis is a post-processing step in the decoder. Similarly, the de-correlated channel generation in the decoder is done either by time domain convolution or equivalent DFT domain multiplication [9,11]. MDCT domain analysis and synthesis of reverberation for parametric coding of stereo audio has been proposed in [14]. Spatialization parameter extraction and stereo synthesis from the ‘sum’ channel are done in the MDCT domain. For parameter estimation, the MDCT coefficients are divided in to twenty two non-uniform blocks and an analysis by synthesis scheme in the MDCT domain is used. The stereo channels are approximated to the linear combination of the ‘sum’ channel and a reverberated channel derived from the ‘sum’ channel. Four parameters are extracted from each block and encoded as the side information. The spatialization parameters such as ICTD, ICLD and ICC are not estimated directly. Instead, the de-correlated channel used for stereo synthesis in the decoder is generated in the encoder and it is used to estimate the synthesis coefficients through least square approximation method. The parametric coder realized using this parameter extraction method is capable of achieving reasonably good quality stereo audio. In this paper, we propose an improved parametric extraction scheme in the MDCT domain using three different approaches. In the first scheme we use two reverberated channel instead of the single reverberated channel as proposed in [14]. This results in better modeling of spatialization cues which is reflected in the perceptual evaluation. In the second scheme, we use sub-band wise mutually orthogonalized sum channel and reverberated channels for parameter extraction and synthesis. In the third scheme, psychoacoustically weighted parameter extraction scheme is introduced. Performance evaluations of proposed methods are conducted through listening tests.

2. SPATIALIZATION PARAMETER EXTRACTION AND STEREO SYNTHESIS USING MULTIPLE REVERBERATED CHANNELS

Formation of ’sum’ channel through down-mixing, generating reverberated channels from the ’sum’ channel and parameter extraction are the functions of a parametric stereo encoder. We follow the methods as used in [14] for in our encoder as described below. The block diagram for the MDCT domain parametric stereo encoding section is shown in Figure 1. In the first step, MDCT

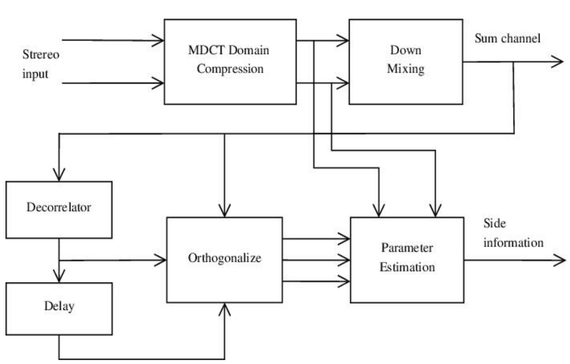


Figure 1: Block Diagram of MDCT Domain Parametric Stereo Encoder.

of the stereo channels $(x_1(n)$ and $x_2(n))$ are computed. The stereo signals in the MDCT domain are then down-mixed to form a ‘sum’ channel $S(k)$. A de-correlated channel $S_r(k)$, is derived from the sum channel. Additional de-correlated channel is generated by delaying the de-correlated channel. The frequency domain signals $S(k), S_r(k)$ and a delayed version of the de-correlated channel $S_{rd}(k)$ are used to extract the spatialization parameters.

2.1 Down Mixing to Sum Signal

For the analysis of stereo signal 1024 point MDCT of the input stereo channels are computed. In order to imitate the features of human auditory system which is more sensitive to lower frequency bands than higher bands, the MDCT coefficients are grouped into bands with different spectral resolution. We use a partitioning method same as that suggested by C. Faller in [1] in which the MDCT coefficients are grouped into 22 non-overlapping frames of different lengths. The partition details are given in table 1. A sub-band by sub-band energy equalized ‘sum’ signal is obtained from stereo transforms by down mixing. The MDCT of the j^{th} sub-block of the energy equalized sum signal is given by Equation 1.

$$S^j(k) = c_j\{X_1^j(k) + X_2^j(k)\}, \quad for\ all\ k \tag{1}$$

$$c_j = \sqrt{\frac{\frac{1}{2}\sum_k((X_1^j(k))^2 + (X_2^j(k))^2)}{\sum_k(X_1^j(k) + X_2^j(k))^2}} \tag{2}$$

$X_1^j(k)$ and $X_2^j(k)$ are the MDCT coefficients of j^{th} sub band. The energy equalization factor c_j given by equation 2 is introduced to make the energy of the ‘sum’ sub-band signal equal to the average energy of the corresponding stereo sub-bands.

Partition	B_0	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}	B_{11}
Boundary	0	8	16	24	32	48	64	80	96	128	160	192
Partition	B_{13}	B_{13}	B_{14}	B_{15}	B_{16}	B_{17}	B_{18}	B_{19}	B_{20}	B_{21}	B_{22}	
Boundary	224	256	288	320	384	448	512	576	640	768	1024	

TABLE 1: Partition boundaries of sub-bands in MDCT domain.

2.2 Synthesis of De-correlated Channels

For parametric coding in MDCT domain, the synthesis parameters are not derived directly from spatialization parameters like ICLD, ICTD and ICC. Instead, we estimate them using the sum signal, a de-correlated ‘sum’ signal and a delayed de-correlated signal. The same de-correlated signals are created in the decoder to synthesize the stereo signal. The de-correlated signal is computed using the method suggested by J. Breebaart et al., in [11], i.e., by convolving the ‘sum’ signal with an all pass de-correlation filter whose impulse response is given by equation 3.

$$h(n) = \sum_{m=0}^{\frac{N}{2}} \frac{2}{N} \cos\left(\frac{2\pi mn}{N} + \frac{2\pi m(m-1)}{N}\right) \tag{3}$$

De-correlated signal is generated by the convolution of 'sum' signal and $h(n)$. The convolution is implemented in the MDCT domain following the method described in [16]. The MDCT of the de-correlated 'sum 'is given by

$$S_r^{mdct}(k) = |H_{u,8N}^{dft}(2k+1)| \{S^{mdct}(k)\cos(\theta_{2k+1}) + S^{mdct}(k)\sin(\theta_{2k+1})\}, \quad 0 \le k \le N-1 \quad (4)$$

where $H_{u,8N}^{dft}(2k+1)$ is the 8N point DFT of up-sampled $h(n)$. The length of the de-correlation filter is selected such that L << N for which the error due to aliasing is irrelevant [16]. We use de-correlation filter of length L=40 for N=1024.

2.3 Parameter Estimation in MDCT Domain

For the parameter extraction, we use 'sum' signal $S(k)$, de-correlated signal $S_r(k)$ and delayed de-correlated signal $S_{rd}(k)$ which is obtained by shifting $S_r(k)$ by a certain number of samples. The MDCT coefficients representing the synthesized stereo channels $(X_1^j(k), X_2^j(k))$ are given by the linear combination

$$X_i^j(k) = a_i^j S^j(k) + b_i^j S_r^j(k) + S_{rd}^j(k), \qquad (5)$$

where *i=1,2* and *1 ≤ j ≤ 22.*

To estimate the synthesis parameters a_i^j, b_i^j and c_i^j the inner product between the transforms of stereo signals $S(k)$, $S_r(k)$ and $S_{rd}(k)$ are computed.

$$\begin{aligned} d_i^j &= \langle X_i^j(k) S^j(k) \rangle \\ e_i^j &= \langle X_i^j(k)\, S_r^j(k) \rangle \\ f_i^j &= \langle X_i^j(k)\, S_{rd}^j(k) \rangle \end{aligned} \qquad (6)$$

An approximation of the stereo signal is obtained by adding the projections of $X_1^j(k)$ and $X_2^j(k)$ on $S^j(k), S_r^j(k)$ and $S_{rd}^j(k)$ as

$$P_i^j(k) = d_i^j S^j(k) + e_i^j S_r^j(k) + f_i^j S_{rd}^j(k) \qquad (7)$$

Further amplitude scaling of $P_i^j(k)$ is done to equalize the energy in the synthesized sub-bands to that of the original sub-band energy by multiplying with the scaling factors computed as below

$$X_i^j(k) = s_i^j P_i^j(k) \qquad (8)$$

where the scale factors s_i^j i=1,2 are given by

$$s_i^j = \frac{\sum_k (X_i^j(k))^2}{\sum_k (P_i^j(k))^2} \qquad (9)$$

The synthesis parameters are obtained by multiplying the inner products obtain in the first step with the corresponding scale factors.

$$a_i^j = s_i^j d_i^j \qquad b_i^j = s_i^j e_i^j \qquad c_i^j = s_i^j f_i^j \qquad (10)$$

The synthesis parameters are compressed using uniform quantizer and encoded as side information.

2.4 Stereo Decoding

The block diagram for the decoder is shown in Figure 2. The MDCT coefficients of the synthesized stereo are reconstructed from the linear combination of the equalized 'sum' signal $S(k)$, and de-correlated signals $S_r(k)$ and $S_{rd}(k)$. The de-correlated and its delayed signals are obtained using the same procedure followed in the encoder. Side information forms the weights for the linear combination.

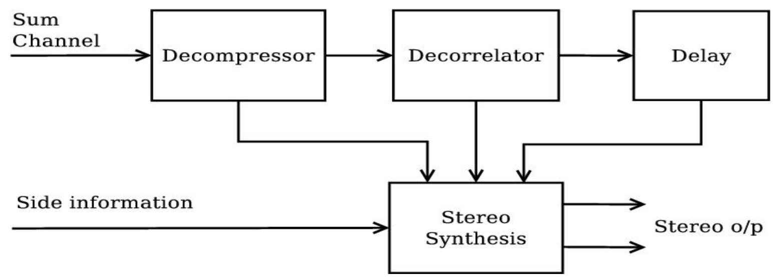


FIGURE 2: Block diagram of parametric stereo decoder.

2.5 Spatialization Parameter Estimation using Orthogonalized De-correlated Channels

De-correlated channels derived from the 'sum' channel through filtering are not perfectly correlated. To further reduce the correlation between them we performed sub-band wise orthogonalization of the channels in the MDCT domain. The 'sum' signal $S(k)$, de-correlated signal $S_r(k)$, and delayed de-correlated signal $S_{rd}(k)$, are modified through Gram-Schmidt orthogonalization procedure. Orthogonalization of these signals is performed in every sub-band to make

$$\begin{aligned} \langle S^j(k)\ S_r^j(k)\rangle &= 0 \\ \langle S^j(k)\ S_{rd}^j(k)\rangle &= 0 \\ \langle S_r^j(k)\ S_{rd}^j(k)\rangle &= 0 \end{aligned} \tag{11}$$

where j = 1,2,..,22 denotes the sub-band indices $S^j(k), S_r^j(k)$ and $S_{rd}^j(k)$ the corresponding orthogonalized 'sum', de-correlated and delayed de-correlated signals. We used these signals for parameter extraction. The decoder is also modified to include the orthogonalization steps performed in the encoder.

2.6 Stereo Decoding with Orthogonalized De-correlated Channels

The block diagram of the stereo decoder is shown in Figure 3. The decoder receives $S^j(k)$ and spatialization parameters $a_i^j(k), b_i^j(k)$ and $c_i^j(k)$. From this, first we obtains the de-correlated signals $S_r^j(k)$ and $S_{rd}^j(k)$. Then sub-band wise orthogonalization is done and the decoder synthesizes back the stereo channels as a linear combination of these orthogonalized signals using the spatialization parameters.

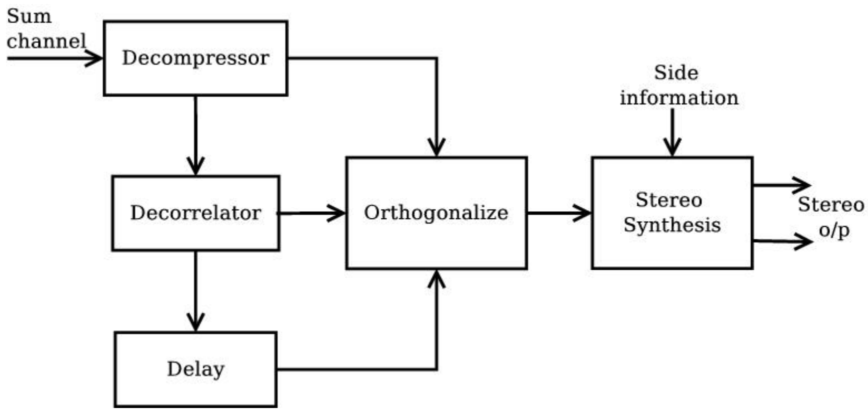


FIGURE 3: Stereo decoder for synthesizing stereo from orthogonalized 'sum' and reverberated channels.

2.7 Psychoacoustically Weighted Parameter Estimation Scheme

Masking threshold estimated through psychoacoustic analysis is used for determining perceptually irrelevant components of audio signals and extensively used in perceptual audio compression algorithms [17]. A higher masking threshold indicates higher noise masking capability of the frequency component. Components with lower masking threshold are more sensitive to quantization noise. We used masking threshold estimate to scale the spatialization parameters. Low weightage is given for spatialization parameters representing frequency components with higher masking thresholds since they are less sensitive components. Coefficients corresponding to lower masking threshold are amplified to give them more weightage. We used MDCT domain masking threshold estimation scheme presented in [18] for computing the masking threshold estimate $T_i(k)$ of the every audio frame. Scaling factors $m_i(k)$ for frequency components are obtained by uniform linear interpolation of $[\min(T_i(k))\ \max(T_i(k))]$ into the range [1.5 0.5] such that minimum weightage is given for frequency components having maximum masking threshold as given below.

$$N_i(k) = \left\lceil 1023\left(\frac{\max(T_i(k)) - T_i(k)}{\max(T_i(k)) - \min(T_i(k))} \right) \right\rceil \tag{12}$$

$$m_i(k) = 0.5 + \frac{N_i(k)}{1023} \tag{13}$$

This scaling factor $m_i(k)$ is used to scale the spatialization parameters in the stereo analysis stage as given in equation

$$a_i^j(k) = s_i^j(k)\, m_i(k) \langle X_i^j(k)\, S^j(k) \rangle$$

$$b_i^j(k) = s_i^j(k)\, m_i(k) \langle X_i^j(k)\, S_r^j(k) \rangle$$

$$c_i^j(k) = s_i^j(k)\, m_i(k) \langle X_i^j(k)\, S_{rd}^j(k) \rangle \tag{14}$$

where *j=1,2,..,22* denotes the sub-band indices, i the channel number, $s_i^j(k)$ the sub-band wise energy equalizing factor and $m_i(k)$ scaling function obtained from the masking threshold estimate. The method for stereo synthesis from parameters is same as in the case of stereo synthesis using multiple reverberated channels.

3. SUBJECTIVE EVALUATION RESULTS

To evaluate the perceptual quality of the encoded audio signal using the proposed algorithm, listening test was conducted. Six listeners participated in the test. The listeners are asked to evaluate both the spatial audio quality as well as other audible artifacts. In a MUSHRA test [19], the listeners had to rate the relative perceptual quality of ten processed items against original excerpts in a 100-point scale with 5 anchors. Tests are conducted with high quality headphone in a quiet room. The following items are included in the test.

- The original as the hidden reference.
- A low-pass filtered (cut off frequency of 7 kHz) mono channel derived from the original.
- Stereo audio compressed by MPEG-2 AAC with TNS and M/S stereo enabled at a rate of 64 kbps.
- Stereo signal synthesized using uncompressed 'sum 'signal and synthesis parameters estimated with single reverberated channel.
- Stereo signal synthesized using uncompressed 'sum 'signal and synthesis parameters estimated with two reverberated channel.

3.1 Perceptual evaluation of parametric stereo generated using multiple reverberated channels

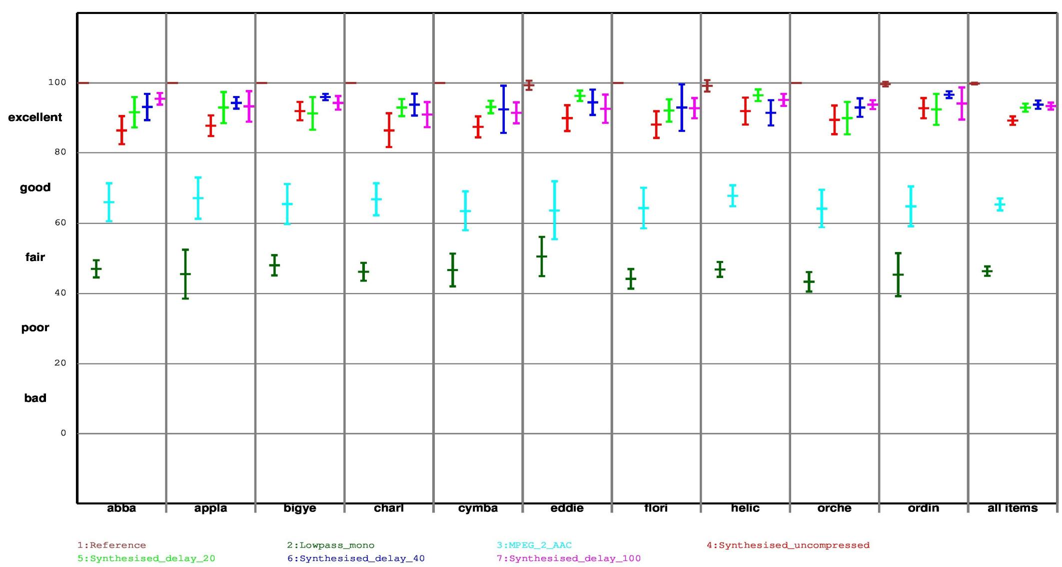


FIGURE 4: MUSHRA Scores for Average and 95% Confidence Intervals for stereo synthesis.

In the case of stereo synthesized using two reverberated channel, we have considered three different delay conditions (20 ms, 40 ms and 100 ms) for the second delay channel. Thus, there are seven clips including the hidden references for each channel in the subjective evaluation. All clips had a resolution of 16 bits per sample and sampling rate of 44.1 kHz. The list of stereo clips used for the subjective evaluation is shown in Table 2.

The average MUSHRA scores obtained for the seven version of each clip are shown in Table 3. The average score for all the clips combined is also given. The synthesized audio with two reverberated signals performed better than the synthesized signal with single reverberated signal. The average MUSHRA score for the synthesized signal with reverberated channels having no delay and 40 ms delay is 93.8 while that of single reverberated channel is 89.3. The performance of 20 ms delayed reverberated channel and 100 ms delayed channel are also very close to that of the 40 ms delayed case.

Sl. No.	Index	Name	Origin/Artist
1	abba	Abba	SQAM Database
2	appla	Applause	SQAM Database
3	bigye	Big Yellow	Counting Crows
4	charl	Charlies	Danny O'Keefe
5	cymba	Cymbal	Radiohead
6	eddie	Eddie Rabbit	SQAM Database
7	flori	Florida Sequence	Pre-echo test case
8	helic	Helicopter	bAdDuDeX
9	orche	Orchestra	Dave Mathews band
10	ordin	Ordinary World	Duran Duran

TABLE 2: List of excerpts used in the subjective listening test.

Name	Original	Low pass Mono	MPEG 2 AAC	Synth Uncomp-ressed	Synth. Delay 20	Synth. Delay 40	Synth. Delay 40
Abba	100	47	66	86.5	91.7	93.2	95.5
Applause	100	45.5	67.2	87.8	93	94.3	93.3
Big Yellow	100	48	65.5	92	91.3	96	94.3
Charlies	100	46.2	66.8	86.5	93	93.8	91
Cymbal	100	46.7	63.5	87.5	93.2	92.5	91.5
Eddie Rabbit	99.3	50.5	63.7	90	96.3	94.5	92.7
Florida Sequence	100	44.2	64.3	88.2	92.2	93	92.8
Helicopter	99.2	46.8	67.8	92	96.5	91.5	95.2
Orchestra	100	43.3	64.2	89.5	90	93	93.8
Ordinary World	99.7	45.3	64.8	92.8	92.5	96.7	94.2
All Items	**99.8**	**46.4**	**65.4**	**89.3**	**93**	**93.8**	**93.4**

TABLE 3: Average MUSHRA scores for test clips generated using multiple reverberated channels.

3.2 Perceptual evaluation of parametric stereo generated using orthogonalized signals

To evaluate the performance of stereo synthesized using orthogonalized de-correlated channels subjective evaluation was conducted with following audio clips.

- The original as the hidden reference.
- A low-pass filtered (cut off frequency of 7 kHz) mono channel derived from the original.
- Stereo audio compressed by MPEG-2 AAC at a rate of 32 kbps.
- Stereo signal synthesized using energy equalized 'sum' signal, de-correlated signal and its delayed version (with delay of 40 samples) weighted by synthesis parameters.
- Stereo signal synthesized using orthogonalized and energy equalized 'sum' signal, de-correlated signal and its delayed version (with delay of 40 samples) weighted by synthesis parameters.

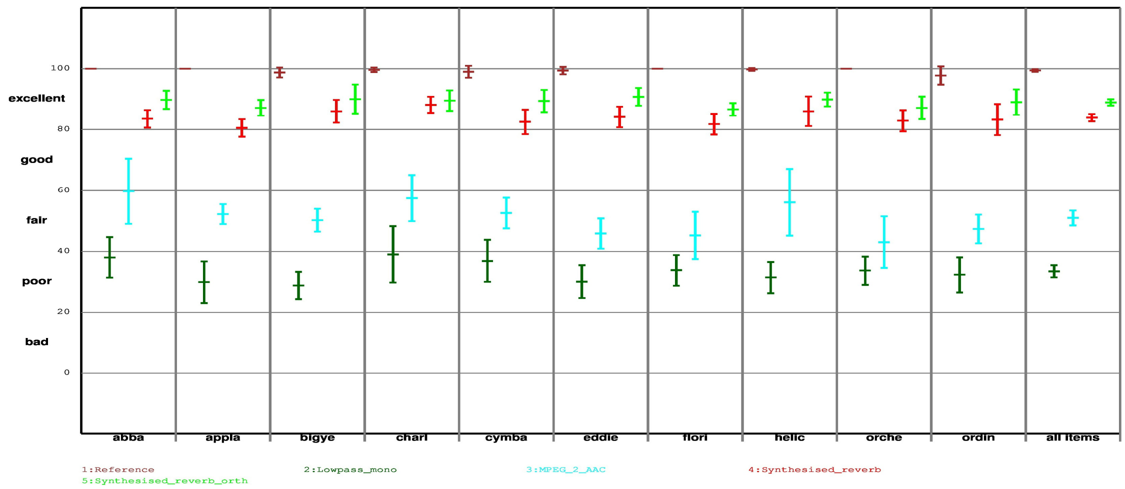


FIGURE 5: MUSHRA Scores for Average and 95% Confidence Intervals for audio clips.

Name	Original	Low pass mono	MPEG 2 AAC	Reverb. delay	Orthogonal Delay
Abba	100	38.9	59.8	83.5	89.8
Applause	100	29.9	52.2	80.5	87.1
Big Yellow	98.8	28.8	50.2	86	90
Charlies	99.6	39	57.5	88.1	89.5
Cymbal	99	36.9	52.6	82.5	89.4
Eddie Rabbit	99.4	30	45.9	84.1	90.8
Florida Sequence	100	33.8	45.2	81.8	86.6
Helicopter	99.8	31.4	56.1	86	89.9
Orchestra	100	33.6	43	82.9	87.1
Ordinary World	97.8	32.2	47.4	83.2	89
All Items	**99.4**	**33.4**	**51**	**83.9**	**89.9**

TABLE 4: Average MUSHRA scores of audio clips synthesized using orthogonalized signals.

Stereo signal synthesized using energy equalized ‘sum’ signal, de-correlated signal and its delayed version(with delay of 40 samples) weighted by synthesis parameters. Stereo signal synthesized using orthogonalized and energy equalized ‘sum’ signal, de-correlated signal and its delayed version(with delay of 40 samples) weighted by synthesis parameters. The average MUSHRA scores assigned by the listeners for these clips are shown in Table 4. The average score for all the clips combined is also given. MUSHRA scores for average and 95% confidence intervals are plotted in Figure 5.

Results clearly show that orthogonalizing the ‘sum’ and reverberated channels for spatialization parameter extraction and stereo synthesis have resulted in improving the perceptual quality of the synthesized audio. The average MUSHRA score for the synthesized stereo has increased from

83.9 to 89.9 when orthogonalized signals were used for parameter extraction and stereo synthesis.

3.3 Perceptual evaluation of parametric stereo generated using scaled parameters

Test audio clips were generated by scaling spatialization parameters extracted using estimated masking threshold values. This perceptual weighting of parameters is done for the stereo synthesis using multiple reverberated channels with and without orthogonalization. We used the following items for listening test:

- The original as the hidden reference.
- Stereo audio compressed by MPEG-2 AAC at a rate of 32 kbps.
- Stereo signal synthesized using energy equalized ‘sum’ signal, de-correlated signal and synthesis parameters.
- Stereo signal synthesized using energy equalized ‘sum’ signal, de-correlated signal and its delayed version (with delay of 40 samples) with scaled parameters.
- Stereo signal synthesized using orthogonalized and energy equalized ‘sum’ signal, de-correlated signal and its delayed version (with delay of 40 samples) with scaled parameters.

The average MUSHRA scores obtained for different test audio clips and average score obtained for all clips are shown in Table 5. MUSHRA scores for average and 95% confidence intervals are

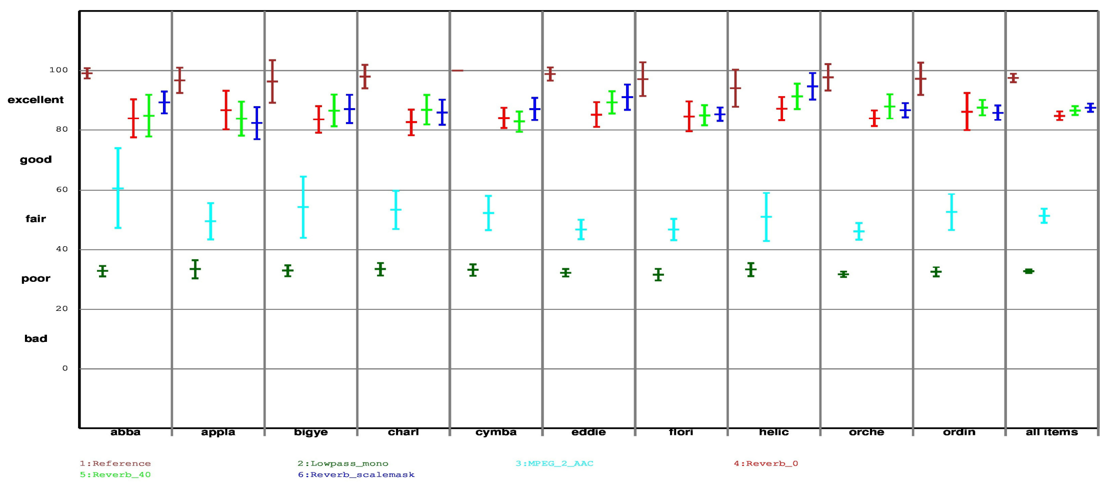


FIGURE 6: MUSHRA Scores for average and 95% confidence Intervals obtained for synthesized audio clips using scaled spatialization parameters.

plotted in Figure 6 The results of the listening test shows that the perceptual quality of the synthesized stereo using two reverberated channels was 86.7 whereas including masking threshold also to scale the spatialization parameters resulted in an average score of 87.6.

3.4 Discussion

It is not surprising that the quality of the encoder increases as the number of parameters is increased. Use of an additional delayed reverberated channel for stereo analysis and synthesis increases the perceptual quality at the cost of increased side information rate. The delayed signal helps the parametric model to capture the late reverberated part of the original signal. The reverberation filter length is 40 samples which is approximately equal to 1 ms and a delay of 40 samples effectively extends the filter length to 80 samples and that results in better modeling of the spatial cues. The number of parameters representing spatialization cues is increased to three per sub-band for each channel and this result in a better approximation for the stereo audio.

Name	Original	Low pass mono	MPEG 2 AAC	Synth. Uncompres sed	Synth. Delay 40	Perceptually Weighted
Abba	99.1	32.8	60.6	84	84.9	89.4
Applause	96.8	33.4	49.5	86.8	83.9	82.4
Big Yellow	96.4	32.9	54.2	83.6	86.6	87.1
Charlies	98	33.4	53.4	82.6	86.9	86
Cymbal	100	33.1	52.2	84.1	82.9	87.1
Eddie Rabbit	98.9	32.1	46.8	85.2	89.4	91.1
Florida Sequence	97.1	31.5	46.8	84.6	85	85.4
Helicopter	94.1	33.2	51	87.2	91.4	94.8
Orchestra	97.8	31.6	46.1	84	88	86.8
Ordinary World	97.2	32.5	52.6	86.2	87.6	85.9
All Items	**97.5**	**32.6**	**51.3**	**84.8**	**86.7**	**87.6**

TABLE 5: Average MUSHRA scores of synthesized audio clips using scaled Spatialization parameters.

Perceptual tests shows better quality for stereo synthesized using additional reverberated channel delayed by 40 samples with an average MUSHRA score of 93.4 whereas that for stereo synthesized using single reverberated channel was 89.3. When we orthogonalize the de-correlated signals we expect better spatial modeling at the cost of additional computation. MUSHRA test reveals a better performance of stereo synthesized using orthogonalized signals with an average score of 89.9 against the average score of 83.3 obtained for stereo synthesized non orthogonalized reverberated channels. Stereo synthesis using perceptually scaled spatialization parameters produce marginal improvement in perceived quality at the cost of increased complexity. When evaluated simultaneously, the average MUSHRA score is 87.6 for scaled parameter synthesis while that of multiple de-correlated channels is 86.7.

4. CONCLUSION

Methods for parametric stereo coding in MDCT domain using sum and two de-correlated channels have been introduced. Three methods are used for estimating spatialization parameters. It can be seen that stereo synthesized using ‘sum’ and two de-correlated channels has a better perceptual quality than that synthesized using ‘sum’ and a single de-correlated channel. The quality of the encoder has increased when orthogonalized signals were used for parameter extraction and stereo synthesis. Orthogonalization makes the three signals used in parameter extraction and stereo synthesis independent of each other. But the computational complexity of encoder as well as decoder will be increased due to the additional orthogonalization process. In method 3, spatialization parameters were further modified using scaling functions obtained from masking threshold and results in marginal improvement in the perceptual quality of the synthesized audio.

5. REFERENCES

[1] C. Faller, “Parametric Coding of Spatial Audio,” *Swiss Federal Institute of Technology Lausanne (EPFL), PhD Thesis, No. 3062,* 2004.

[2] D. Yang, H. Ai, C. Kyriakakis, ans C.C. J. Kuo, “An inter channel redundancy removal approach for high quality multichannel audio compression,” in *AES convention,* Los Angeles, CA, Sept 2000.

[3] S. Kuo and J.D. Johnston, “A Study of Why Cross Channel Prediction is Not Applicable to Perceptual Audio Coding," IEEE Sig. Proc. Letters, vol. 8, No. 9, pp 245-247, Sep. 2001.

[4] J. Herre, et.al, “The reference Model Architecture for MPEG Spatial Audio Coding," in 118th AES convention, Barcelona, Spain May 2005, Preprint 6447.

[5] J.D. Johnston, and A.J. Ferreira, “Sum Difference Stereo Transform Coding,” in Proc. IEEE ICASSP-92, San Francisco, vol. 2, pp. 569-572, March 1992.

[6] Christian R. Helmrich, Pontus Carlsson, Sascha Disch, Bernd Edler, Johannes Hilpert, Matthias Neusinger, Heiko Purnhagen, Nikolaus Rettelbach, Julien Robilliard, and Lars Villemoes, “Efficient Transform Coding Of Two-Channel Audio Signals By Means Of Complex-Valued Stereo Prediction,” in Proc. IEEE ICASSP-2011, pp. 497-500, 2011.

[7] Christof Faller, and Frank Baumgarte, “Binaural Cue Coding: A Novel and Efficient Representation of Spatial Audio,” in Proc. IEEE ICASSP-2002, vol: 2, pp. II-1841 - II-1844, 2002.

[8] F. Baumgarte, and C. Faller,“Binaural Cue Coding-part I : Psychoacoustic fundamentals and Design Principles,” in IEEE Trans. on Speech and Audio Proc., vol. 11, No. 6, pp. 509-519, June 2003.

[9] F. Baumgarte, and C. Faller,“Binaural cue coding-part II : Schemes and applications,” in IEEE Trans. on Speech and Audio Proc., vol. 11, No. 6, pp. 520-531, June 2003.

[10] C. Faller, “Parametric Multichannel Audio Coding: Synthesis of Coherence Cues," IEEE Trans. Speech and Audio Proc., vol. 14, No. 1, pp. 1-12, Jan. 2006.

[11] J. Breebaart, et al.,“Parametric Coding of Stereo Audio,” in EURASIP Journal on Applied Signal Processing, vol 2005, No. 9, pp 1305 - 1322, June 2005.

[12] A. Kohlrausch, “Auditory filter shape derived from binaural masking experiments," J. Acous. Soc. America, vol. 84, no. 2, pp. 573-583, 1988. 16

[13] B. R. Glasberg and B.C.J. Moore, “Derivation of auditory filter shapes from notched-noise data," Hearing Research, vol. 47, no. 1-2, pp . 103-138, 1990.

[14] K. Suresh, and T. V. Sreenivas, “MDCT Domain Analysis and Synthesis of Reverberation for Parametric Stereo Audio,” in AES 123th Convention, 2007 October 5-8, New York.

[15] K. Suresh, and T. V. Sreenivas, “Parametric stereo coder with only MDCT domain computations,” IEEE International Symposium on Signal Processing and Information Technology, pp. 61-64, December 2009.

[16] K. Suresh and T. V. Sreenivas, “Linear Filtering in DCT-IV/DST-IV and MDCT/MDST Domain”, Signal Processing, vol 89, Issue 6, pp 1081-1089, June 2009.

[17] T. Painter, and A. Spanias, “Perceptual Coding of Digital Audio", Proc. IEEE, vol. 88, no 4, pp. 451-513, 2000.

[18] K Suresh and T. V. Sreenivas, “Direct MDCT Domain Psychoacoustic Modeling”, IEEE International Symposium on Signal Processing and Information Technology, pp. 742-747, December 2007.

[19] ITU/ITU-R BS 1534. Method for subjective assessment of intermediate quality level of coding systems, 2001.

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Signal Processing (SPIJ)* lays emphasis on all aspects of the theory and practice of signal processing (analogue and digital) in new and emerging technologies. It features original research work, review articles, and accounts of practical developments. It is intended for a rapid dissemination of knowledge and experience to engineers and scientists working in the research, development, practical application or design and analysis of signal processing, algorithms and architecture performance analysis (including measurement, modeling, and simulation) of signal processing systems.

As SPIJ is directed as much at the practicing engineer as at the academic researcher, we encourage practicing electronic, electrical, mechanical, systems, sensor, instrumentation, chemical engineers, researchers in advanced control systems and signal processing, applied mathematicians, computer scientists among others, to express their views and ideas on the current trends, challenges, implementation problems and state of the art technologies.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 10, 2016, SPIJ will be appearing with more focused issues related to signal processing studies. Besides normal publications, SPIJ intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

SPIJ LIST OF TOPICS

The realm of Signal Processing: An International Journal (SPIJ) extends, but not limited, to the following:

- Biomedical Signal Processing
- Communication Signal Processing
- Detection and Estimation
- Earth Resources Signal Processing
- Industrial Applications
- Optical Signal Processing
- Radar Signal Processing
- Signal Filtering
- Signal Processing Technology
- Software Developments
- Spectral Analysis
- Stochastic Processes
- Acoustic and Vibration Signal Processing
- Data Processing
- Digital Signal Processing
- Geophysical and Astrophysical Signal Processing
- Multi-dimensional Signal Processing
- Pattern Recognition
- Remote Sensing
- Signal Processing Systems
- Signal Theory
- Sonar Signal Processing
- Speech Processing

CALL FOR PAPERS

Volume: 10 - **Issue: 1**

i. Paper Submission: January 31, 2016 **ii. Author Notification:** February 28, 2016

iii. Issue Publication: March 2016

CONTACT INFORMATION

Computer Science Journals Sdn BhD
B-5-8 Plaza Mont Kiara, Mont Kiara
50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627
Fax: 006 03 6204 5628

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2015
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6204 5627
FAX: 006 03 6204 5628
EMAIL: cscpress@cscjournals.org