

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 8 日現在

機関番号：13201

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26420348

研究課題名(和文)チャンネルスケラブルな3Dオーディオ符号化のためのアンマスキング雑音抑圧

研究課題名(英文)Suppression of unmasked quantization noise for channel scalable 3D audio codec

研究代表者

安藤 彰男 (Ando, Akio)

富山大学・大学院理工学研究部(工学)・教授

研究者番号：00545668

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：昨年からはまった8Kスーパーハイビジョン試験放送では、22.2チャンネル音響が採用されている。このような多チャンネル音響を家庭で再生することは難しいため、我々は本来の臨場感を保ちながらより少ないチャンネルで再生する8.1チャンネル方式を開発した。
伝送チャンネル数を増やさずに22.2チャンネル信号と8.1チャンネル信号を伝送するスケラブル伝送を行うためには、信号変換が必要である。ところが、変換した信号を圧縮符号化すると、復号して逆変換を行う際に、他の音でマスキングされるはずの量子化雑音が顕在化する。そこで、伝送量を増やさずにこの雑音を抑圧する方法を研究し、処理遅延340msで雑音を25dB以上抑圧した。

研究成果の概要(英文)：In a 8K super hi-vision test broadcasting, which started from 2016 in Japan, a 22.2 channel audio was adopted. However, such a multichannel audio is difficult to reproduce in a home and a new system that reproduces the original sound impression with a lesser number of channels is necessary. We developed a 8.1 channel system as such a method.
A scalable transmission in which a signal of the home reproduction system is embedded in a 22.2 channel signal usually requires a matrix conversion. However, if the converted signal is coded by an audio codec, the quantization noise that should be masked by the other sounds becomes audible after decoded and inversely converted. In this research, a noise compression method was developed and the audible noise was reduced by more than 20dB within a time delay of 340 ms.

研究分野：感性情報学、通信・ネットワーク工学 (音響工学)

キーワード：マルチチャンネルオーディオ 圧縮符号化 スケラブル伝送 空間音響 音響信号処理

1. 研究開始当初の背景

(1) 映像、オーディオ符号化に関する標準化機関である MPEG (ISO/IEC JTC1 SC29 WG11) は、2013年1月に MPEG-H 3D Audio の Call for Proposal を発行し、三次元マルチチャンネルオーディオ信号の符号化に関する標準化活動が開始された。これは、我が国が次世代のテレビジョンであるスーパーハイビジョンの研究開発を進めるにあたり、22.2チャンネルオーディオの放送を検討していたこともその背景にある。それまでのオーディオ符号化方法である MPEG-4 AAC は、2チャンネル信号の符号化方法であったため、これを多次元に拡張できるかが課題であった。

(2) 一方、前述の 22.2チャンネルオーディオは、高さ方向に三層のチャンネル構造を持ち、22個の全帯域チャンネルと、2個の低音効果音用チャンネルを有するものである。一般的には、家庭で 24 個のスピーカを配置することは困難であるため、22.2チャンネルオーディオの放送が始まって、これを家庭でどのように再生するかが課題であった。そこで、その本来の臨場感を損なわずにチャンネル数を減らす研究が行われ、最小のチャンネル数を有する三次元音響方式として、8.1チャンネルオーディオが開発されていた。

(3) ところで、家庭で送られて来た 22.2チャンネル信号を 8.1チャンネル信号に変換するのは受信機に負担がかかるため、8.1チャンネル信号が送信されることが望ましい。そのような方法として、22.2チャンネル信号と 8.1チャンネル信号を同時に放送するサイマルキャスト放送がある。しかしながら、サイマルキャスト放送では、国の重要な資源である電波の帯域を多く占有するため、伝送チャンネル数は 22.2チャンネルと同じだが、その一部分を取り出すと 8.1チャンネル信号になっているようなチャンネルスケラブルな放送が望まれていた。

(4) チャンネルスケラビリティを実現するには、もとの 22.2チャンネル信号を 8.1チャンネル信号とその他の信号(以下、補助信号と呼ぶ)に変換する必要がある。8.1チャンネルオーディオを再生する簡易再生の場合には、伝送された信号から 8.1チャンネルの部分だけを復号し、22.2チャンネルオーディオを再生するフルスペックの場合には、全チャンネル信号を逆変換することによって元の信号を復元するというものである。ところが、多チャンネル信号に線形変換を施して、AAC などを用いて圧縮符号化して放送し、受信側で復号後に逆変換して元の信号を復元する際には、厄介な問題があった。アンマスキング雑音である。これは、伝送された信号では、量子化雑音はオーディオ信号によって聴覚的にマスクされて聞こえないが、逆変換を施すと、オーディオ信号によるマスクが効かなくなって、量子化

雑音が顕在化する現象である。この問題の扱いは、前述の MPEG-H 3D Audio の中でも議論になったが、“nightmare”と呼ばれて取り組む研究者は見つからなかった。

2. 研究の目的

(1) 本研究の目的は、22.2チャンネルオーディオ信号を圧縮符号化してスケラブル伝送を行うことにある。この目的を実現するため、伝送容量をほとんど増やさずに、かつ低遅延でアンマスキング雑音を抑圧することを目的して、本研究を開始した。

3. 研究の方法

(1) アンマスキング雑音の発生は、伝送時の変換方法に依存する。そこで、まずは、変換行列の最適な設計法について検討した。その際に、8.1チャンネル信号は復号後そのまま家庭再生で利用すべきものであるため、十分なビットレートを割り当てる必要がある。その結果、補助信号に割り当てるビットレートが削減されるため、補助信号はエネルギーが抑制されることが望ましい。一方、行列が特異行列に近づく、逆変換行列の要素が大きくなり量子化雑音が増幅される恐れがある。行列の安定度を示す指標として条件数があるが、変換行列は条件数が小さい方が望ましい。ところで、補助信号のエネルギー削減と、変換行列の条件数の現象は、ある意味でトレードオフの関係にある。たとえば、補助信号のエネルギーを削減する究極の方法は、対応する行列の行をすべて 0 にすることであるが、この場合には行列は特異行列となり、条件数は無限大となる。そこで、変換/符号化/復号/逆変換のプロセスによって生じた誤差信号のエネルギーによって、原信号のエネルギーを除いて得られる SN 比を指標として、最適な条件数を検討した。その結果、条件数 20 程度の変換行列が、少ない条件数と補助信号エネルギーの削減を両立させるとの結論を得た。

(2) 8.1チャンネル信号と補助信号へのビットレート配分を検討するため、全ビットレートを 1.2Mbps とした上で、3種類の配分方法を前述の SN 比で評価した。音声コーデックとして、独フラウンホーファー IIS 研究所の MPEG-4 AAC ソフトウェアを用いた。このソフトウェアは、2チャンネル符号化の組み合わせで 22.2チャンネル信号を符号化するものである。実験素材は、10種類の 22.2チャンネル素材(各 10秒)を用いた。同素材は、標準化周波数 48kHz、量子化ビット数 24ビットでデジタル化されたものである。表 1に、教会の鐘の音を集音した素材に対する結果を示す。その他の素材についても、同様の結果が得られた。この結果、以下の検討では、8.1ch に 720kbps、補助信号に 480kbps を配分することにした。

表1 ビットレート配分とSN比の関係

ビットレート配分		SN比
8.1ch	補助信号	(dB)
800kbps	400kbps	17.9
720kbps	480kbps	21.7
640kbps	560kbps	21.1

(3) アンマスキング雑音抑圧法の検討を行った。この方法では、符号化側にもローカルデコーダを用意し、復号音を逆変換した復元音と原音との差分を取ることで量子化雑音を求める。この量子化雑音を復元音の多項式で近似して多項式係数を求め、この係数を伝送して、受信側でアンマスキング雑音を抑圧する。2.2.2 チャンネル素材を用いた実験により、この方法の評価を行った。実験では、各チャンネル信号を帯域幅 1kHz の 24 帯域に分割し、これらの帯域ごとに量子化雑音を多項式近似した。この多項式近似は、バッチ処理、すなわち各素材の全体の時間区間(10秒)に対して行った。アンマスキング雑音の抑圧効果は、復元音のSN比と雑音抑圧後のSN比との比較により評価した。その結果、雑音抑圧処理により、全素材の全チャンネルに対するSN比で、1.3dB 改善した。各素材ともアンマスキング雑音が観測されるのはごく少数のチャンネルであり、その結果、平均改善量は僅かであった。そこで、アンマスキング雑音が観測されたチャンネル、すなわち、原音の信号レベルが低いチャンネルについて調べたところ、25dB から 40dB の改善が得られていた。また、実際に復元音を試聴したところ、アンマスキング雑音は十分抑圧されていた。

(4) アンマスキング雑音抑圧のリアルタイム化を検討した。すなわち、上述の方法では、素材音全体でアンマスキング雑音を復元音の多項式で近似したが、素材音を短い時間区間に分割し、分割されたブロックごとに、アンマスキング雑音を復元音の多項式で近似した。なお、バッチ処理で行っていた周波数帯域分割は行わず、多項式近似は全帯域信号に対して行った。まず、ブロック長とSN比との関係を調べた。その結果、ブロック長が短い程SN比が向上した。しかしながら、例えばブロック長を 20ms 程度にすると、多項式係数を伝送するための伝送量が増大し、100kbps 程度まで達する。そこで、データ量の増大が無視できる程度のブロック長として 340ms を選択した。この結果、前述のバッチ処理と同等の付加情報量で、しかもバッチ処理と同等のSN比の改善を得た。

(5) 今後の課題として、同等のSN比と付加情報量を保ちながら、ブロック長を短くすることが挙げられる。アンマスキング雑音が観測されないチャンネルについて多項式係数の伝送を行わなければ、この課題は解決されるとの感触を得たが、本研究期間でそのシミュレーションまで行うことができなかった。こ

の件については、引き続き検討を行うことにしたい。

4. 研究成果

(1) 前項で述べた通り、バッチ処理では、アンマスキング雑音が観測されたチャンネルについて、そのSN比を 25dB 以上改善することに成功した。また、リアルタイム化のため、元信号を 340ms の時間ブロックに分割することにより、付加情報の伝送量を無視できる程度に抑えながら、同等のSN比改善量を得た。しかしながら、このままでは原理的に 340ms の遅延が発生するため、その解決が課題として残された。

(2) 本研究のパイプロダクトとして、チャンネル間コヒーレント成分を抽出する方法を開発した。この方法により、元のオーディオ信号は、チャンネル間で相関の高い直接音成分と、背景雑音や残響音などの拡散音成分に分離することができる。この方法を併用することにより、効果的なアンマスキング雑音抑圧を行うことが期待されるが、本研究期間ではそこまで検討することができなかった。この件も、今後の研究課題としたい。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 0 件)

現在、Acoustic Science and Technology 誌の Spatial Hearing 特集号に投稿準備中。

[学会発表](計 5 件)

藤田崇史、安藤彰男、" スケーラブル伝送符号化のための量子化雑音を軽減する変換行列の検討 ", 日本音響学会 2015 年秋期研究発表会講演論文、2-P-6、2015

Akio Ando, Hiroki Tanaka and Hiro Furuya, " Extraction of Interchannel Coherent Component from Multichannel Audio, " Proc. Audio Engineering Soc. 140th Convention, 査読有, Paper no. 9545, 2016

田中宏樹、橋本裕太、古屋紘、安藤彰男、" マルチチャンネル音響信号におけるチャンネル間コヒーレント成分の抽出 ", 日本音響学会 2016 年秋期研究発表会講演論文、3-7-1, 2016

橋本裕太、田中宏樹、古屋紘、安藤彰男、" マルチチャンネル音響信号における直接音声分の抽出 ", 日本音響学会 2016 年秋期研究発表会講演論文、3-7-2, 2016

安藤彰男、田中宏樹、橋本裕太、古屋紘、" 多チャンネル音響信号からのコヒーレント

成分の抽出”，電子情報通信学会技術研究報告 EA2016-60，2016

〔図書〕(計 1 件)

安藤彰男，“音場再現”，コロナ社，2014、212

〔産業財産権〕

出願状況(計 1 件)

名称：オーディオ信号処理装置、オーディオ信号処理方法、およびオーディオ信号処理プログラム

発明者：安藤彰男

権利者：富山大学

種類：特許

番号：特願 2016-89417

出願年月日：平成 26 年 4 月 27 日

国内外の別：国内

取得状況(計 0 件)

名称：

発明者：

権利者：

種類：

番号：

取得年月日：

国内外の別：

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

安藤 彰男 (ANDO, Akio)

富山大学・大学院理工学研究部(工学)教授

研究者番号：00545668

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：

(4) 研究協力者

()