

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 6月20日現在

機関番号：62615

研究種目：挑戦的萌芽研究

研究期間：2011～2012

課題番号：23650083

研究課題名（和文） 振幅スペクトログラム量子化と位相復元復号化に基づく音響信号符号化の研究

研究課題名（英文） Audio Coding Based on Amplitude Spectrogram Quantization and Phase-reconstruction Decoding

研究代表者

小野 順貴 (Ono Nobutaka)

国立情報学研究所・情報学プリンシプル研究系・准教授

研究者番号：80334259

研究成果の概要（和文）：本研究では、新しい高効率音響信号符号化を目指し、以下の成果を得た。1) 量子化された振幅・位相スペクトログラムから高音質に波形を合成するために、範囲制限位相復元という新しい手法を考案した。2) 音楽信号などに典型的にみられるスペクトログラム繰り返し構造を符号長最小規準で自動的に推定する手法を考案した。3) 人間の聴覚が位相に対して感度が低いことを利用し、ステレオ音響信号の位相スペクトログラムに情報を埋め込む高耐性音響電子透かしアルゴリズムを考案した。

研究成果の概要（英文）：In this project, aiming to a new high-quality audio coding, our achievements can be summarized as follows. 1) We developed “the range-constrained phase reconstruction” technique, which is a method to synthesize well-quality time-domain signal from quantized amplitude and phase spectrogram. 2) We developed a method to automatically estimate repetitive structure of audio signals, which can be typically observed in music spectrogram, with minimum coding-length criterion, and 3) We developed a robust audio watermarking method, which embeds the information to phase spectrogram of stereo audio signals, utilizing the fact that human auditory system is not much sensitive to phase.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	2,800,000	840,000	3,640,000

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：音声情報処理・音響信号符号化

1. 研究開始当初の背景

ポータブル音楽再生機器の普及やインターネット音楽配信の広がりなどを背景に、音響信号の符号化技術には常に大きな需要があり、高音質で高い圧縮率をもつ符号化手法が常に求められている。音響信号の符号化に用いられている AAC、mp3 などの符号化技術は非可逆符号化であり、人間の聴覚特性を利用し、聴感上できるだけ目立たないような情報損失を許すことで、音質の劣化を防ぎながら高い圧縮率を実現している。現状の符号化

方式の多くは、粗いサブバンド分割と変形離散コサイン変換を組み合わせた領域において、聴覚特性を考慮しながらビット割り当てを行うものであるが、これらの技術はすでにかなり成熟・飽和しており、この方式での圧縮率はほぼ限界にきていると考えられている。

2. 研究の目的

本研究の目的は、振幅スペクトログラムの量子化・符号化と、位相スペクトログラムの復

元に基づく波形復号化を原理とする新しい音響信号符号化技術を構築することである。本研究の核となる、1) 振幅スペクトログラムのみから位相スペクトログラムを復元する位相復元手法と、2) 振幅スペクトログラムを非負値行列とみなし、これを少数の基底スペクトルと振幅時間パターンに分解する非負値行列分解法、の2つは、近年発展した最新の音響信号処理技術であり、本研究はこれらを音響信号化の分野に導入することで、新しい音響信号化方式を開拓する。この方式は位相の符号化を省略できることに加え、振幅スペクトログラムの長期的冗長性を利用できる利点を有しており、これらの特長に基づき音響信号符号化の圧縮率を大幅に向上させることを目指す。

3. 研究の方法

研究計画は大きく、1) 音響信号符号化に適した位相復元技術の構築、2) 非負値行列分解の最適符号化、の2項目からなる。1) では、聴覚的な周波数分解に対する位相復元の適用、振幅の量子化を前提とした位相復元手法の改良、位相復元を前提とした最適な時間周波数分解特性の導出、などを行う。2) では、非負値行列分解により得られる基底スペクトル、振幅時間パターン、残差スペクトル、それぞれの要素に対する最適ビット割り当て、基底数の最適化、非負値行列分解における距離尺度の最適選択、等を行う。これらの2項目は相互に関連しているため、平行して進めながら、符号化の音質と圧縮率を評価していく。具体的にはビットレートを固定した上で、主観音響評価を行いながら、パラメータのチューニングや手法の選択などを行っていく。方針としては、アーカイブ用の符号化を想定し、計算量よりは音質に重点を置いて研究を進める。

4. 研究成果

本研究の目的は、振幅スペクトログラムの量子化・符号化と、位相スペクトログラムの復元に基づく波形復号化を原理とする新しい音響信号符号化技術を構築することである。本研究の成果は主に以下の3つである。

1つ目は量子化された振幅/位相スペクトログラムからの高音質な波形合成する手法である。従来、振幅スペクトログラムのみから位相スペクトログラムを推定するために用いられてきた位相復元手法を、量子化された振幅・位相スペクトログラムからの波形合成部分に応用し、範囲制限つき位相復元という手法を考案した。図1に示すように、この提案法では、量子化された値は量子化前にはある量子化範囲に含まれなければならないことを制約条件として用いて、位相復元のた

めの反復時にこの範囲の外に出てしまった場合には、値を範囲内に引き戻す操作を行う。図2のブロック線図に示すように、これはSTFTと逆STFTを反復する従来の位相復元法の一部に位相のクリッピング操作を加えるだけで実現できる。図3に示すように、通常逆量子化、従来の位相復元よりも、範囲制限位相復元を適用する方が音質が改善されることが示された。

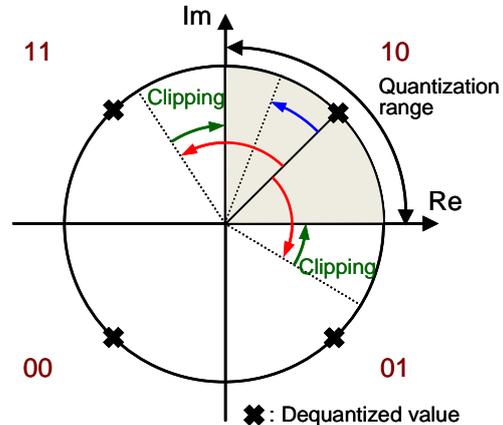


図1：複素平面上での範囲制限位相復元

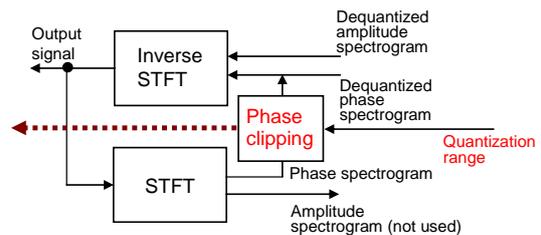


図2：範囲制限位相復元のブロック線図

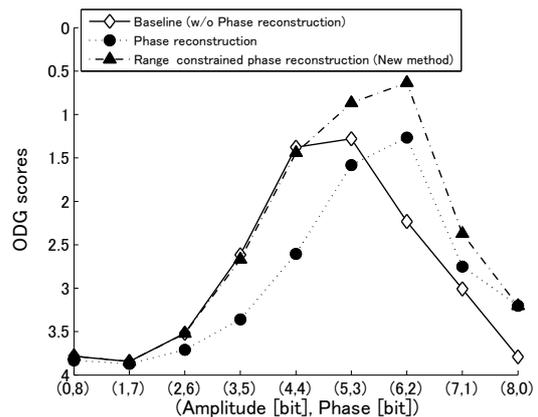


図3：音質比較

2つ目は音楽構造を利用した音響信号の符号化である。多くの音楽楽曲は繰り返し構造を持ち、その繰り返し構造は振幅スペクトログラムによく現れる。我々は、図4に示すように2回目以降の繰り返し部分は、最初に現

れた部分との差分のみを符号化することで、全体の情報量を削減する手法に取り組んだ。これを行うためには、与えられた振幅スペクトログラムの類似箇所を自動的に推定する手法が必要になる。我々は、符号長最小化の規準で動的計画法を用いることにより、図5のように最適な繰り返し構造を検出するアルゴリズムを考案した。

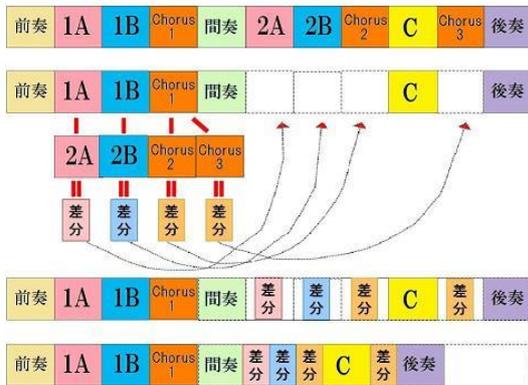


図4：繰り返し構造を用いた符号化

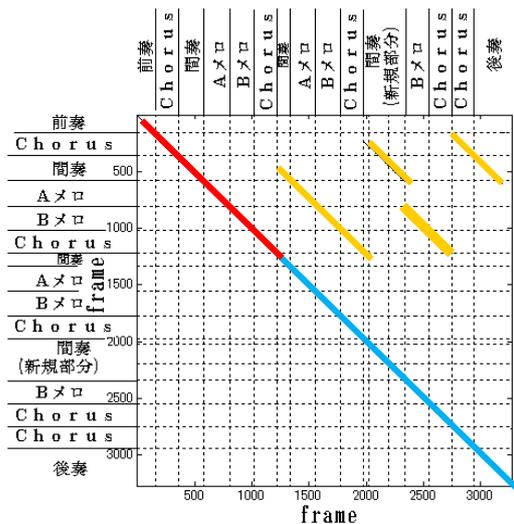


図5：検出された繰り返し構造

3つ目は、振幅位相スペクトログラム表現を電子透かしに応用したものである。人間の聴覚が位相に対して感度が低いことを利用し、ステレオ音響信号の位相スペクトログラムに情報を埋め込む高耐性音響電子透かしアルゴリズムを考案した。具体的には、位相スペクトログラムのある点に位相を埋め込む際には、図6に示すように、ステレオ間位相差が、埋め込む位相が0であるか1であるかによって、異なる点に量子化されるように位相を修正する。また、各種攻撃への耐性を高めるために、図7に示すように

位相スペクトログラムの様々な箇所にも、同じ情報を冗長に分散して埋め込むアルゴリズムを構築した。本アルゴリズムは第1回音響電子透かしコンテストへ応募し、評価基準をクリアしたことが認定された。

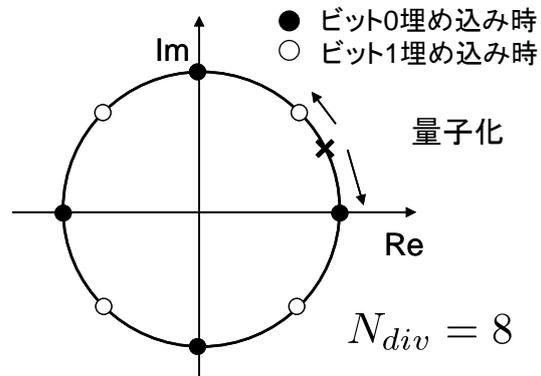


図6：位相差への情報埋め込み

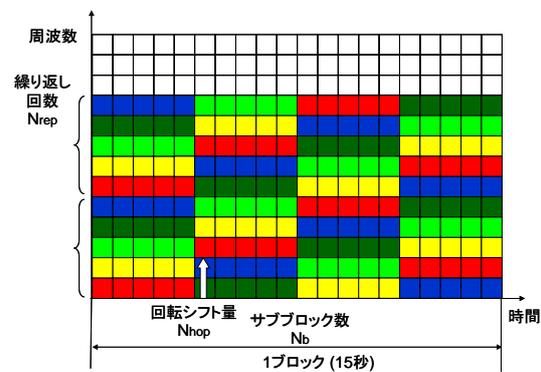


図7：位相スペクトログラムへの冗長分散埋め込み

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4件)

①小野 順貴, "時間周波数領域におけるステレオ間位相差への分散埋め込みに基づく音響情報ハイディング," 電子情報通信学会技術研究報告, vol. 112, pp. 75-80, 2012. (査読無)

②Sho Sato, Nobutaka Ono, Yutaka Kamamoto, and Shigeki Sagayama, "Range-constrained Phase Reconstruction for Recovering Time-domain Signal from Quantized Amplitude and Phase Spectrogram," Proc. DAFx, 2012. (査読有)

③藏 悠子, 小野 順貴, 鎌本 優, 嵯峨山

茂樹, “動的計画法に基づく音楽構造解析とその音楽信号符号化への応用,” 日本音響学会春季研究発表会講演集, pp. 1049-1050, 3月, 2012. (査読無)

④佐藤 匠, 小野 順貴, 鎌本 優, 嵯峨山茂樹, “スペクトログラムの振幅・位相量子化と範囲制限位相復元に基づく音響信号符号化の検討,” 日本音響学会秋季研究発表会講演集, pp. 229-230, 9月, 2011. (査読無)

[産業財産権]

○出願状況 (計 1 件)

名称: 音声再生装置、音声再生方法及びコンピュータプログラム

発明者: 小野順貴、嵯峨山茂樹、佐藤匠

権利者: 情報・システム研究機構

種類: 特許

番号: 特願 2011-197404

出願年月日: 2011年09月09日

国内外の別: 国内

6. 研究組織

(1) 研究代表者

小野 順貴 (Ono Nobutaka)

国立情報学研究所・情報学プリンシプル研究系・准教授

研究者番号: 80334259

(2) 研究分担者

嵯峨山 茂樹 (Sagayama Shigeki)

東京大学・情報理工学(系)研究科・教授
研究者番号: 00303321

鎌本 優 (Kamamoto Yutaka)

日本電信電話株式会社NTTコミュニケーション科学基礎研究所・守谷特別研究室・研究員

研究者番号: 00418550

(3) 連携研究者

なし