様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 31 日現在

機関番号：12612
研究種目：基盤研究(C)
研究期間：2009 ～ 2011
課題番号： 21500097
研究課題名（和文）　圧縮ドメインにおける楽曲アノテーションに関する研究

研究課題名（英文）A study on music annotations based on the compressed domain of MPEG audio

研究代表者
星　守（HOSHI　MAMORU）
電気通信大学・名誉教授
研究者番号：80125955

研究成果の概要（和文）：本研究課題では，MPEG-4 audio の圧縮過程から算出されるパラメータ，あるいは楽曲圧縮データのビットストリームに格納されているパラメータを用いた楽曲アノテーションに関する研究を実施した．本研究課題の成果として，１）圧縮過程で算出される自己相関係数系列を用いた楽曲の構造分析手法，２）ビットストリームに格納された LSP パラメータから算出される LPC ケプストラム系列を用いた楽曲のジャンル分類手法，３）ビットストリームに格納された LSP パラメータ系列を用いた楽曲に対する印象語付与手法の提案を行った．

研究成果の概要（英文）：In this study, we focused on music annotation using a sequence computed in encoding step of TwinVQ compression or from a bit-stream in compressed music data. We proposed three methods for, 1) music structure analysis using a sequence of auto-correlation coefficients, 2) music genre classification using a sequence of LPC cepstrum computed from bit-stream, 3) impression word annotation using LSP parameter directly extracted from bit-stream.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2009 年度	1,100,000	330,000	1,430,000
2010 年度	1,000,000	300,000	1,300,000
2011 年度	1,300,000	390,000	1,690,000
年度			
年度			
総　計	3,400,000	1,020,000	4,420,000

研究分野：総合領域
科研費の分科・細目：情報学、メディア情報学・データベース
キーワード：マルチメディア

１．研究開始当初の背景

人が日常の生活で何か情報を処理するときには，まず，処理の対象となる情報を探し，取り出すことから始まる．この情報を取りだす操作，すなわち検索操作は情報を処理するための基本的な操作の 1 つである. ネットワークを通じて大量のマルチメディアデータが活用される今日，マルチメディアデータの情報処理においても例外ではない．大量のマルチメディアデータを蓄積できるマルチメディアデータベースが構築出来ても有効な検索機能がなければ，宝の持ち腐れとなる．このようにマルチメディアデータベースを有効に活用するためには，マル

チメディアデータを有効かつ効率的に検索することのできる検索システムが必須となる.

研究代表者らは，マルチメディアデータ検索における内容に基づいた類似検索機能の研究を 1992 年より継続して研究を行ってきた. さらに 2000 年からは，圧縮ドメインにおける楽曲検索手法に関する研究を実施した.

はじめに，MPEG-4 TwinVQ 圧縮過程における線形予測分析部に着目し，線形予測分析部で算出される自己相関係数が，ビットレートの変化に対して頑強な特徴量を生成可能であることを理論的に示した(ICME2001 発表). さらに，市販 CD から取得した 1,023 曲に対して検索実験を実行し,実験的にもビットレート変化に頑強な検索が可能であることを示した(AMC MM2005, IEEE MDDM2007 で発表).

また，TwinVQ 圧縮過程で算出さる自己相関係数列から楽曲のビート周期を抽出するアルゴリズム（IEEE PCM2003 で発表），節長を抽出するアルゴリズム（ICME2005 で発表）を提案した。

このように，研究代表者らは圧縮ドメインにおける楽曲検索・分析に関する技術を蓄積してきた.

２．研究の目的

本研究課題の目的は，楽曲圧縮ドメインにおける楽曲内容に基づいた楽曲へのアノテーション技術の開発である. 本研究課題では MPEG-4 Audio 規格である TwinVQ 楽曲圧縮に着目し，楽曲圧縮過程中に算出される，あるいは圧縮データのビットストリームに格納されているパラメータを用いた楽曲へのアノテーション，およびアノテーションに基づいた楽曲検索手法に関する研究を行う. 具体的な研究内容は(1)～(3)の通りである.

(1) 楽曲の部分にタグ付加するための楽曲構造抽出の開発
(2) 楽曲へのジャンルタグ付加するためのアノテーション技術の開発
(3) 楽曲への印象語タグ付加のためのアノテーション技術の開発

３．研究の方法

本研究では， MPEG-4 TwinVQ 圧縮における線形予測分析に着目する. MPEG-4 TwinVQ 圧縮の枠組みを図 1 に示す. 圧縮過程で実行される線形予測分析では、自己相関係数，線形予測係数，LPC ケプストラム，LSP パラメータ，LSP スペクトルが算出され，LSP パラメータ

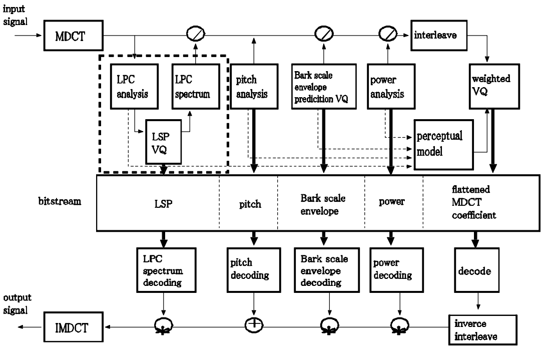

図　1. TwinVQ 圧縮の枠組み

をビットストリームに格納する. 本研究では，線形予測分析で算出される各パラメータは，信号における基本的な特徴量であることに着目し，これらパラメータを用いた楽曲アノテーション手法を開発する.

(1)楽曲の部分にタグを付与するための楽曲構造抽出に関する開発では，圧縮過程で算出される自己相関係数系列を用いて楽曲構造の分析を行う. 楽曲構造分析は，以下の３つの手順からなる.

①楽曲の節長の抽出：自己相関系列から音楽的に意味のある分割単位長を抽出する.
②楽曲分割：自己相関係数列から，分割単位長を用いて楽曲を部分列へと分割する.
③構造分析：分割された部分列間の類似性をもちいて，楽曲の構造を抽出する.

(2) ジャンルタグ付加技術の開発では，ビットストリームに格納されている LSP パラメータ系列，LSP パラメータ系列から算出可能な LSP ケプストラム系列，線形予測係数系列に着目する.

各系列に対して，ウェーブレット変換による多重解像度解析を実行し，各ウェーブレットレベルにおける差分係数を算出する. 算出された各レベルの差分係数の平均，標準偏差を算出し，それらの組み合わせを楽曲特徴量として用いた. 算出した楽曲特徴量のジャンル分類性能を評価するため，楽曲から算出した楽曲特徴量を説明変数，ジャンル情報を目的変数として，判別分析を実行した.

(3)印象語付加技術の開発では，ビットストリームに格納されている LSP パラメータ系列に着目する.

LSP パラメータ系列に対して，ウェーブレット変換による多重解像度解析を実行し，各ウェーブレットレベルにおける分解係数を

算出する．算出された各レベルの近似係数，ウェーブレット係数のそれぞれの平均，標準偏差を算出し，各レベルの平均，標準偏差の組み合わせを楽曲特徴量とて用いた．算出した楽曲特徴量の印象語分類性能を評価するため，(2)と同様に判別分析を行う．

４．研究成果

(1) TwinVQ 楽曲圧縮過程において算出される自己相関係数系列を，節を単位として区間に分割を行った(節抽出アルゴリズムは ICME 2005 で発表済み)．分割された各節間の類似度を用いて，Ward 法によるクラスタ分析を行い，デンドログラムを作成した．デンドログラムをある閾値 De により分割することにより，各節をクラスに分類する(図 2)．各クラスにラベルを割り当てることにより，自己相関係数系列を，節に基づいたラベル系列に変換した(図 3)．

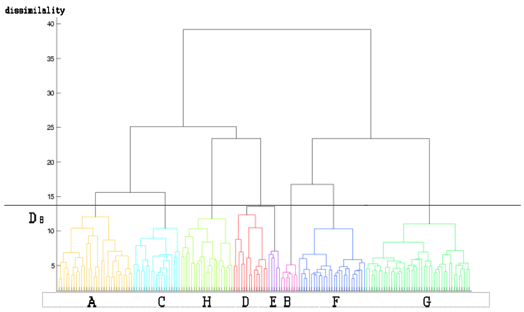

図 2.　クラスタリング結果

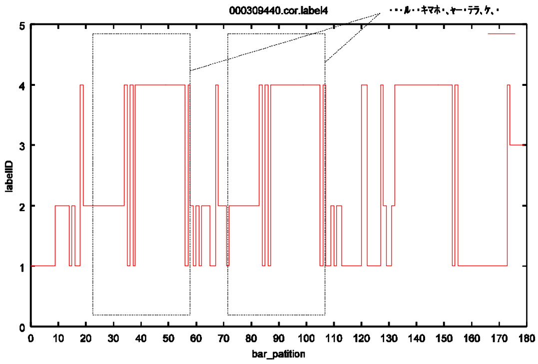

図 3. ラベル系列

ここで，De を変化させることにより，長短さまざまな階層的なラベル系列に変換可能であることに注意されたい．ラベル系列の最長一致区間を繰り返し区間として検出することで，楽曲の長短さまざまな繰り返し区間からなる階層的な構造抽出が可能となった(図　4)．

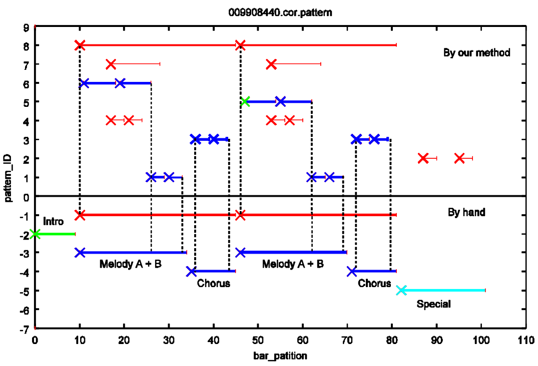

図　4. 構造分析の結果

(2) ジャンルアノテーション技術の開発として，楽曲を短時間（数十秒程度）の部分に分割し，分割された部分楽曲に対するジャンル分類を試みた．用いた楽曲データは，バロック，ボサノバ，ダンス，ヒップホップ，ジャズ，マーチ，オールディーズ，ロック，タンゴ，ワルツの 10 種類のジャンルから，計 2,196 曲を用意した（表　1）．これらの楽曲に対して TwinVQ 圧縮を実行し，圧縮データを得た．

表　1．各ジャンルの楽曲数

ジャンル	楽曲数
バロック	225
ボサノバ	119
ダンス	176
ヒップホップ	196
ジャズ	295
マーチ	91
オールディーズ	435
ロック	200
タンゴ	220
ワルツ	99
合計	2,196

楽曲全体に対するジャンル分類に関する研究（DBSJ 2008 で発表済み）より LPC ケプストラム系列が，ジャンル分類に有効であることがわかっているため，圧縮データ中の LSP パラメータから算出されるに LPC ケプストラム系列に着目し，楽曲の一部（部分曲）の LPC ケプストラム系列を用いてもジャンル分類が可能であるかどうかを検証した．ここで，部分曲へのジャンル分類実験のため，ビットストリームに格納された LSP パラメータから LPC ケプストラムを計算し．2048 フレームからなる LPC ケプストラム系列(48 秒相当）を作成した．この LPC ケプストラム系列に対して，ウェーブレット変換による多重解像度解析を実行し、ウェーブレット変換を用いた特

徴量を算出した．算出した楽曲特徴量のジャンル分類性能を評価するため，各楽曲から算出した楽曲特徴量を説明変数，ジャンル情報を目的変数として判別分析を行った．

本実験では，トレーニングデータの平均ジャンル分類正答率 88.5%(表 2)，テストデータの平均ジャンル分類正答率 82.3%を得た（表 3)．すなわち，提案アルゴリズムは楽曲ジャンル分類手法として有効であることが分かった．

表　２．トレーニングデータにおける各ジャンルの正答率

ジャンル	正答率(%)
バロック	81.25
ボサノバ	71.67
ダンス	85.23
ヒップホップ	95.00
ジャズ	93.06
マーチ	91.30
オールディーズ	91.23
ロック	98.00
タンゴ	84.00
ワルツ	72.00
平均正答率	88.45

表　３．テストデータにおける各ジャンルの正答率

ジャンル	正答率(%)
バロック	80.91
ボサノバ	64.41
ダンス	69.32
ヒップホップ	89.00
ジャズ	90.23
マーチ	91.11
オールディーズ	84.14
ロック	90.00
タンゴ	77.00
ワルツ	61.22
平均正答率	82.34

(3) 楽曲への印象語タグ付加のためのアノテーション技術の開発では，圧縮楽曲データのビットストリームに格納されている LSP パラメータに着目した．圧縮楽曲データに格納されている LSP パラメータを直接抽出し，抽出した LSP パラメータ系列に対して，ウェーブレット変換を実行し，ウェーブレット分解の各レベルにおいて，近似係数列の平均と標準偏差，差分係数の平均と分散を算出した．算出された近似係数の平均と分散，差分係数の平均と分散を組み合わせることにより，楽曲特徴量ベクトルを生成した．10 対の印象語「陽気な―陰気な」，「うきうきした―しんみりした」，「楽しい―悲しい」，「コミカルな―泣ける」，「激しい―穏やかな」，「にぎやかな―静かな」，「騒がしい―落ち着いた」，「軽い―重い」，「力強い―弱々しい」，「明るい―暗い」を用いて，被験者 5 人が 1416 曲に対して印象語を付与した．楽曲特徴ベクトルの印象語アノテーションに対する有効性を示すため，楽曲特徴ベクトルを説明変数，それぞれの被験者が付与した印象語を目的変数として，被験者ごとに判別分析を行った．ただし，判別分析に用いるトレーニングデータの数を各印象語に対して，20 曲とし，それ以外のデータをテストデータとした．本実験では，被験者により判別率に多少の差があるが，判別率として被験者 5 人の平均として 84.6% (最悪 80.1%，最高 86.1%) を得た．このことは，MPEG　TwinVQ 圧縮ドメインにおいて楽曲に印象語を付与できたことを示している．

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）

〔学会発表〕（計　2 件）

① Michihiro Kobayakawa，Mamoru Hoshi，“Musical genre classification of MPEG-4 TwinVQ audio data,” Proc of Int. Conf. Multimedia and Expo 2011 (ICME2011)，査読あり，USB，Barcelona，July 15，2011.

② Michihiro Kobayakawa，Mamoru Hoshi，“A Method for Music Structure Analysis Using MPEG-4 TwinVQ Audio Compression,” Advances in Multimedia Information Processing-PCM2010, Part1，査読あり，pp. 100-111，Shanghai，September 22，2010

６．研究組織

(1)研究代表者
星　守（HOSHI MAMORU）
電気通信大学・名誉教授
研究者番号：80125955

(2)研究分担者
小早川　倫広（KOBAYAKAWA MICHIHIRO）
東京都立産業技術高等専門学校・准教授
研究者番号：00334582