

科学研究費助成事業 研究成果報告書

平成 28 年 9 月 13 日現在

機関番号：12612

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330126

研究課題名(和文) 超多数ユーザへの同時提供のための高効率高機能映像符号結合方式とメディア間連携方式

研究課題名(英文) High efficient video transcoding and media orchestration for scalable provisioning

研究代表者

笠井 裕之 (Kasai, Hiroyuki)

電気通信大学・その他の研究科・准教授

研究者番号：40312079

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究では、大規模多数ユーザへの高品質多数映像情報の同時提供を実現することを目的として、提供映像の高画質化・低遅延化を実現するための高効率な符号変換方式の研究、複数映像視聴をより効果的に提供するための複数映像間及び他メディア間連携方式の研究を行った。
その結果、生成される映像の高画質化、また生成のための遅延化・高速化・低処理化を達成した。また、映像や音響信号の解析・認識技術の検討により、より高精度な認識、高機能化を実現し、メディア連携方式の高度化を実現した。

研究成果の概要(英文)：We face a big issue when a highly interactive multiple-video is provided to a huge number of users simultaneously across a network. The dominant technical problem is the processing speed and delay. In order to reduce this issues, the degradation of its video quality is critical. In this research project, we tackle this trade-off problem. In addition, it is essential to provide high functional media orchestration. To this end, we proposed a video control scheme to avoid the degradation of the video quality while keeping the processing time and delay lower. We also proposed some media recognition schemes and visualization schemes. The results show that the proposed methods are feasible and outperform than the existing methods.

研究分野：メディア信号処理

キーワード：映像符号化 映像符号結合 多数映像 メディア認識 メディア連携

1. 研究開始当初の背景

近年、7680×4320画素/60fps等の超高精細映像配信に向けた取り組みが、放送分野等で活発に行われ、H.264/AVCやJPEG2000を基軸方式とした高効率符号化及び配信方式の研究が進められている[A][B]。また高効率化だけでなく、全く新しい映像体験を提供するための挑戦的な映像配信の検討が進められ、例えば[C]では複数映像や複数視点の映像提供方式が提案されている。

[A] ITU-T Recommendation H.264 “Advanced video coding for generic audiovisual services,” May 2003.

[B] T. Yamaguchi et al. “SHD movie distribution system using image container with 4096x2160 pixel resolution and 36 bit color,” Proc. IEEE ISCAS 2005, pp. 5918-5921, 2005.

[C] 森他, “多視点型映像オンデマンド配信方法”, 信学論誌 Vol.J87-D1 No.5 pp.610-623, 2004.

一方、申請者はこれまで、圧縮映像符号に対する高効率符号変換技術の基礎及び応用研究を行ってきた。そして2009年、大量の超高精細映像を多くのユーザに同時に提供する映像符号結合技術を発表した。本技術は「タイル(領域)ベース符号化」と「タイルストリーム符号結合」の連携を基軸とし、後に一連の研究成果はIEEE国際論文誌に掲載された。これまで、動的な圧縮映像符号結合技術に集中し、初期版と比較して100倍程度の高速化処理を実現するに至り、現市場の一般サーバ機器を用いて、数千ユーザへの同時提供が可能なまで性能を向上させることに成功した。このように学術研究・知財化・産業化を進めているものの、一般ユーザ投稿映像等を含むインターネット映像の指数関数的増加に対応するためには、数十万クラスのユーザへの同時提供を実現するための結合速度の向上、結合画質の品質向上、また複数映像の視聴体験の向上を促進するための他メディアとの連携方式の実現が必要であり、信号処理上の抜本的な基礎アルゴリズムの検討が急務である。

2. 研究の目的

本研究では、さらなる高画質化・低遅延化を実現するための高効率な符号変換方式の研究、複数映像視聴をより効果的に提供するための複数映像間及び他メディア間連携方式の研究が必要である。

超大量の圧縮映像符号の高効率符号結合方式(課題A)

これまで、画素領域まで復号する符号数ができる限り低減するため、結合速度を高速化する画面内画素予測に起因する画質劣化低減方式、及びタイルベース符号化時とタイルストリーム符号結合時に隣接ブロックの係数個数値の不一致による画質劣化が生じない係数個数符号予測に起因する画質劣化低減方式について研究してきた。しかしながら、a.係数個数固定に伴う画質劣化、b.結合時の部分復号処理量の増大、c.映像多重化に関わる処理遅延の増大、の問題が生じる。以下、

各課題に対する方式について研究する。

I.係数個数固定に伴う画質劣化を抑制可能な符号化量子化制御方式と高速量子化パラメータ符号変換方式【符号化器側、結合器側】(課題A-a)

予測参照情報固定を行うパラメータの内、量子化後ブロック係数のパラメータである『非ゼロ係数個数の固定』に起因して画質劣化が発生すること判明した。当該劣化について、符号化時の量子化制御方式により画質向上を実現しながら、高速な量子化パラメータ符号の補正方式により、結合時の結合処理負荷の増大を抑制する方式について研究する。II.ランレングス符号同時復号による高速復号・結合方式【結合器側】(課題A-b)

H.264/AVC採用の可変長符号化方式では、軽量の復号処理が可能なコンテキスト適応可変長符号化CAVLCが採用されている。しかしながら、各符号は順次復号されることを前提としているため、処理の並列化が不可能であり、H.264/AVC復号器で処理の特性上、最大のボトルネックとなっている。そこで、申請者らによってアルゴリズムの並列化により高速化を実現したレベルスキップ復号方式に加え、複数の符号を同時に復号可能な方式を提案する。これにより、符号テーブルのアクセス回数を大幅に低減することにより復号処理の高速化を実現する。

III.低遅延なりアルタイム順次多重化方式【結合器側】(課題A-c)

最新のインターネット映像ストリーミングサービスで使用される伝送多重化フォーマットMP4では、既に蓄積された映像や音響ストリームを用いて多重化を行う。従って、申請者が提唱する動的な符号結合処理後に多重化して伝送する際には、結合映像ストリームを生成完了後、多重化処理を開始する必要がある。この場合、多重化処理の遅延、映像ストリームの一時保存のためのディスクアクセス量や、使用メモリ量の増大が課題である。そこで、タイルストリームの解析による低遅延なりアルタイム順次映像多重化方式を提案する。これにより、1時間長の映像に対して遅延低減とディスクアクセス負荷低減を達成する。

映像間連携及び他メディア間連携方式(課題B)

複数映像視聴では、各映像に付随する音響データを単純に再生することは意味をなさない。そこで、a.各映像に付随する音響データから視覚化データを作成し映像データと連携させることで、音響信号の視覚的な同時視聴も可能となる。さらに、複数映像間の視覚的な影響度合いにより、視聴者の各映像に対する認知度は大きく異なる。そこで、b.複数映像に対する認知特性に適した提示映像位置の制御のため認識モデルについて研究する。

I.映像に付随する音響情報可視化のための音響信号分離・解析・認識方式(課題B-a)

映像に付随する音響データは、一般的に様々な特徴を有する音響データが複数含まれた混合信号である。そこで、混合信号からの音響データ分離技術と、それらの音響信号認識・クラスタリング技術について研究する。特に、混合音の分離技術として、イベント音と背景音の特徴に着目した分離方式について研究し、分離後の音響SNRを向上させる。また分離された音響データのうち調波構造に着目した楽曲信号の抽出方式の認識方式について研究し、抽出率を向上させる。さらにそれ以外の混合音響データについては、独立成分基底と音響解析に特化し且つ時間遷移をモデル化した信号辞書の構築に基づく認識方式について研究し認識率を向上する。

II. 映像に付随する音響情報可視化のためのシーン知識構築方式と動的画像合成方式(課題 B-b)

前記 I. で得られた音響認識結果に基づき可視化を実現する。具体的には、可視化対象となるシーンと可視化情報との関係を記述するため、画像一般オブジェクト認識技術に基づくシーン知識の自動構築方式と、知識に基づく画像・映像素材の合成による可視化情報生成方式について研究する。

III. 映像間の効率的な提供を可能とする大量映像認知モデル化方式(課題 B-c)

複数映像の提示では、同時に再生される映像間の視覚的な影響度合いにより、視聴者の各映像に対する認知度が異なることから、従来検討が行われてこなかった複数映像に対する認知モデルの構築を行う。これらの成果は、複数映像のどの場所にどの程度の時間、広告映像を提示することで効果があるかを制御するための装置などへの応用が期待される。

3. 研究の方法

(1) 係数個数固定に伴う画質劣化を抑制可能な符号化時量子化制御方式と高速量子化パラメータ符号変換方式(課題 A-a)

符号化時の量子化パラメータ制御方式では、非ゼロ係数個数固定による画質への影響に着目し、二つの方式を提案する。第一は、符号化単位小領域中のブロック係数特性が等しいという仮定に基づき、制御前の入力非ゼロ係数個数と制御目標の係数個数との関係から量子化パラメータを導出する方式を検討する。第二は、前記仮定を前提としない推定誤差ベース量子化パラメータ導出方式を提案する。両方式において、非ゼロ係数個数固定による誤差モデルを理論的に確立した。一方、本制御により結合時の量子化パラメータ(符号)の補正が必要となる。結合処理速度の高速性を維持しながら補正処理を実現するため、本処理をバイト単位で行う手法を検討する。

(2) ランレングス符号同時復号による高速復号・結合方式(課題 A-b)

前記 CAVLC の高速復号・結合処理にむけて、

通常、各符号を逐次復号する処理から、複数符号を同時に復号するための符号テーブルの構成と処理アルゴリズムを提案する。

(3) 低遅延なりアルタイム順次多重化方式(課題 A-c)

伝送対象の MP4 多重化ファイル先頭に、全映像フレームの構成・符号量情報等を挿入するため、全映像ストリームの生成後でなければヘッダ情報を構成できない。従って、前記の動的ストリーム結合方式では、結合完了後にヘッダ情報の生成を開始せざるを得ない。そこで「結合処理前に結合対象タイルストリームを解析　ヘッダ情報を作成　結合処理　ネットワークへの順次伝送」が可能な 2 パス多重化方式を提案する。

(4) 映像に付随する音響情報可視化のための音響信号分離・解析・認識方式(課題 B-a)

混合音響信号からの音響分離方式については、非負値行列因子分解(NMF)をベースとした背景音と複数イベント音の分離方式に着目して進める。次に、楽器やピッチを認識する事で環境中を流れる楽曲を認識する方式として、従来、同一楽器のもつ全ての音が同一調波成分を持つと仮定した多重音解析手法が提案されたが、実際は 1 音ごとに調波成分は異なることに着目してモデル構築する手法を検討する。最後に、音響認識方式としては、独立成分分析(ICA)と Matching Pursuit 法(MP)を複合的に用いた音響認識技術を研究する。

(4) 映像に付随する音響情報可視化のためのシーン知識構築方式と動的画像合成方式(課題 B-b)

あるシーン(例えばクリスマス等)に関連付けられた画像群に対する一般オブジェクト画像認識技術を用いることで、シーンを構成するオブジェクト群を抽出し、それらの包括関係からシーン毎の構成オブジェクトグループを抽出する手法を検討する。これらのグループにより、動的に画像を合成する手法を検討する。

(5) 映像間の効率的な提供を可能とする映像間認知モデル化方式(課題 B-c)

複数映像に対する認知モデルの構築を行う。具体的には、映像内の動き量やその領域、映像全体の色や色変化など、映像を構成する各信号特徴量と画面内の位置との関係に基づくモデル化を行う。

4. 研究成果

(1) 係数個数固定に伴う画質劣化を抑制可能な符号化時量子化制御方式と高速量子化パラメータ符号変換方式(課題 A-a)については二つの方式を提案した。第一は、符号化単位小領域中のブロック係数特性が等しいという仮定に基づき、制御前の入力非ゼロ係数個数と制御目標の係数個数との関係から量子化パラメータを導出する方式を提案した。第二は、前記仮定を前提としない推定誤差ベース量子化パラメータ導出方式を提案した。

本方式は非ゼロ係数個数固定に伴う誤差を推定し、誤差が最小となる量子化パラメータを導出する。両方式において、非ゼロ係数個数固定による誤差モデルを理論的に確立した。一方、結合処理速度の高速性を維持しながら補正処理を実現するため、バイト単位での量子化パラメータ補正処理方式を提案し高速化を達成した。

(2) ランレングス符号同時復号による高速復号・結合方式(課題 A-b)については、これまで有効な高速復号方法が提案されていない「各非ゼロ係数の直前のゼロ係数の連続個数の符号」を対象とし、1つのテーブル要素に複数の符号と復号結果を有する新たなテーブルを構成し、テーブルアクセス回数を大幅に低減した。但し、単純な適応により分岐処理が増加してしまうことから、テーブルに含める符号とその結果に、残ゼロ係数の個数情報を用いることで余分な分岐処理を回避するアルゴリズムを提案した。

(3) 低遅延リアルタイム順次多重化方式(課題 A-c)については、ストリームの高速結合のための各種ヘッダ情報を活用することで、1パス目の解析時に可変長復号を行わずストリーム解析を実現し、低処理負荷で各種フレームサイズやストリーム内位置を算出する方式を提案した。これにより、処理負荷だけでなく結合後ストリームを一時蓄積するためのディスクアクセスも回避可能となった。

(4) 映像に付随する音響情報可視化のための音響信号分離・解析・認識方式(課題 B-a)については、非負値行列因子分解(NMF)による背景音とイベント音の分離について、振幅値分散値等の特徴情報に基づく方式を提案した。次に、楽器やピッチを認識については、音数とそれぞれに対応する調波パターンを自動で求めることで多重音解析の精度を向上する方式を提案した。その他、オクターブを含む和音における音数の推定、NMFによる頻出スペクトルパターンを用いた非調波成分の判別についても研究した。最後に、音響認識方式としては、独立成分分析(ICA)と Matching Pursuit 法(MP)を複合的に用いた音響認識技術を提案し、特徴ベクトルにより認識率向上を実現した。

(4) 映像に付随する音響情報可視化のためのシーン知識構築方式と動的画像合成方式(課題 B-b)については、構成オブジェクトグループの抽出グループ情報と抽出過程から得られる評価値から構成されるシーン知識を自動構築する方式を提案した。そして、映像に付随する音響情報から認識されるシーンやシチュエーションを構成するオブジェクトを本知識から取得し、予め用意する画像・映像素材を動的に組み合わせ可視化する方式について研究した。

(5) 映像間の効率的な提供を可能とする映像間認知モデル化方式(課題 B-c)については、数千個の映像ストリームに対する個々の

前記特徴量をデータベース化し、様々な設定パターンに応じて、時空間で制御可能な評価環境を構築した。そして、それらを用いて被験者実験を行い、様々な仮説のもとで、確率統計的アプローチを用いてモデル化した。

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 3 件)

H.Kasai, “Online Low-rank Tensor Subspace Tracking from Incomplete Data by CP Decomposition using RLS,” IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2016), 査読あり, MLSP-P4.3, 2016.

H.Kasai and B.Mishra, “Riemannian preconditioning for tensor completion,” 8th NIPS workshop on optimization for machine learning (OPT2015), 査読あり, 2015.

R.Shinkuma, Y.Sawada, Y.Omori, K.Yamaguchi, H.Kasai, T.Takahashi, “A Socialize System for Enabling the Extraction of Potential Values from Natural and Social Sensing,” Springer, Modeling and Processing for Next-Generation Big-Data Technologies, 査読あり, Vol.4, pp.385-404, 2015.

〔学会発表〕(計 0 件)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

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

http://kjk.office.uec.ac.jp/Profiles/53/0005267/prof_e.html

6 . 研究組織

(1)研究代表者

笠井 裕之 (KASAI, Hiroyuki)

電気通信大学・大学院情報システム学研究

科・准教授

研究者番号：40312079

(2)研究分担者

()

研究者番号：

(3)連携研究者

()

研究者番号：