

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 26 日現在

機関番号：57101

研究種目：基盤研究(C) (一般)

研究期間：2012～2014

課題番号：24500231

研究課題名(和文) スパース表現の並列演算を用いた多視点動画の実時間符号化に関する研究

研究課題名(英文) On real-time multi-view video coding using parallel processing of sparse representation

研究代表者

黒木 祥光 (Kuroki, Yoshimitsu)

久留米工業高等専門学校・制御情報工学科・准教授

研究者番号：60290847

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：信号のスパース表現は与えられた信号をできるだけ少数の基底ベクトルの線形結合で表現する方法であり、信号処理やパターン認識など、多方面での応用が期待されている。本研究ではスパース表現を用いて多視点動画をより効果的に符号化する研究を行った。また、スパース問題を実時間で解くためにGPUおよびCPUの複数コアを用いた並列演算を行った。実験結果からスパース解法の適用は有効であり、提案法はCPUのみの場合に比べ約2.7倍の高速化を実現できた。多視点画像の幾何情報を得るために基礎行列の高精度な推定法も実現できた。

研究成果の概要(英文)： Sparse representation is one of the hottest topics in signal processing, and approximates a given signal by a linear combination of less numbers of basis vectors. In this research, we applied sparse representation to multi-view video coding to achieve higher coding performance than the conventional video coding framework used in MPEGs, and tried faster implementation using multi-cores in GPU and CPU. Research results show the effectiveness of sparse representation in video coding, and also tell us that the proposed parallel processing is about 2.7 times faster than the execution using CPU. This research includes accurate estimation of Fundamental matrix to obtain geometric information of objects from multi-view images.

研究分野：知覚情報処理

キーワード：画像符号化 スパース表現 GPGPU 多視点画像

1. 研究開始当初の背景

信号のスパース表現は与えられた信号をできるだけ少数の基底ベクトルの線形結合で表現するものであり、理論的な観点からコンプレッド・センシングと呼ばれるサンプリング理論の研究がされてきた。信号のスパース性は線形結合係数の l_0 ノルムにて表現され、その解法は NP 困難とされている。しかし、一定の条件下での l_0 最小化は l_1 最小化にて代替できると証明されている。 l_1 ノルムは凸関数であるため、多数の解法が知られており、プログラムも公開されている。この l_1 最小化問題は、機械学習の分野では線形回帰問題における回帰係数の l_1 正則化に対応しているため、画像や音声の認識、修復など幅広い分野で応用されている。しかしながら、公開されているソフトウェアを画像修復などデータ量の多い問題に適用した場合、処理時間が長いため、動画を実時間で処理することは不可能である。

一方、複数のカメラを用いて映像を獲得する多視点動画は、単に人間が娯楽として観る、いわゆる TV としての役割のみならず、例えば美術品や工芸品などのデジタルアーカイブ、制御工学における視覚サーボ、人物などの動作解析、などへの応用が期待されている。使用するカメラ数が増加するほど、デジタル化された動画のデータ量は必然的に増加するため、高効率に圧縮する必要がある。多視点でない動画符号化の国際標準では、まず、動き補償予測により動画の時間的冗長性を削除した後、動き補償予測誤差の空間的冗長性を離散コサイン変換 DCT (Discrete Cosine Transform) などの直交変換により削除している。一方、多視点動画では、同一時刻に複数のカメラで 3 次元空間中にある同一物体を撮影しており、カメラ間の冗長性も存在する。動画符号化の国際標準 MPEG でも Multiview Video Coding (MVC)、あるいは Free viewpoint TV (FTV) の符号化について 3DAV Ad Hoc による幾つかの議論がなされてきた。しかし、採用された方法は従来の時間軸で冗長性削減を行う動き補償予測を単にカメラ間に拡張しただけであり、基本的な原理は MPEG-1 の延長線上にある。他の興味深い方法として光線空間法が提案されているが、これは符号化よりも復号側で有限個のカメラ画像から任意視点の画像を得る方法として有益と思われる。

2. 研究の目的

動画符号化では実時間処理が不可欠であるが、近年のコンピュータの発展状況を見るとクロック周波数の点では飽和傾向にあり、マルチコア化に移行している。更に LSI の製造プロセスの微細化も伴い、タブレット PC やスマートフォンにも複数コアを持つ CPU や GPU (Graphical Processing Unit) が搭載されるに至っている。研究代表者は数年来、スパース表現を用いた顔画像の認識、CSIFT

(Colored Scale-Invariant Feature Transform) によるカメラ間対応点を用いた基礎行列の最尤推定法、高精細画像を実時間にて符号化するための GPU の応用などについて検討してきた。本研究では上記のうち、圧縮の高効率化として前者 2 つの知見を、実時間処理に向けた並列演算による高速化として後者 1 つの知見を多視点動画の符号化に適用し、実時間による多視点動画の符号化を目的に設定した。画像の符号化では冗長性を削減した後、エントロピー符号化を適用するが、本研究では既存の方法を流用するに留め、エントロピー符号化の高効率化については研究の対象外とした。

Web カメラの低価格化にみられるように、複数のカメラを様々な処理に応用する機会が増えつつある。本研究では画像圧縮を主な目的とするが、採用する手法は撮影された物体の幾何情報を使用するため、得られる知見は制御工学における視覚サーボなど、様々な分野への応用が期待できる。また、信号処理および機械学習の分野にて近年特に顕著な成果を上げているスパース表現を GPU で高速化する研究も含むため、比較的処理能力が低い携帯型端末でもスパース表現を実現可能となり、より高性能なアプリケーションを開発することも可能になるとと思われる。

3. 研究の方法

計算機シミュレーションは平成 22 年度に知的クラスタの予算で購入した既存の計算用サーバを用いる。これは当時最高速であった NVIDIA 社の GPU2 枚を使用しているため、平成 24 年度は引き続きこのサーバを利用し、平成 25 年度に計算機サーバを更新した。研究協力者として MPEG のテストシーケンスを作成した名古屋大学の藤井俊彰教授、画像符号化および画像処理一般につき早稲田大学の鎌田清一郎教授に助言をいただき、以下の方法にて研究を行った。

(1) 平成 24 年度

最新の動画符号化である HEVC (High Efficiency Video Coding) の LIC (Local Intensity Compensation) にスパース表現を適用し、予測精度の向上を確認した。LIC のスパース表現は次式で定式化される。

$$b \approx Aw + o \quad (1)$$

ここで b は着目するマクロブロックの画素をラスタスキャン順に並べて生成したベクトルである。以下、参照するマクロブロックは全てラスタスキャンによってベクトル化されている。 A は参照マクロブロックを各列に並べた行列、 w は b を参照マクロブロックの線形結合で表すための結合係数、 o はオフセットである。スパース問題の解法には、インターネット上で公開されている Sparsity を用いて実験を行った。既存の方法として、通常の動き補償予測で 2 乗誤差の少ないブロックをあらかじめ選定し、それらを用いた最小

2 乗法で予測係数を求める方法を採用した。

また、多視点画像間における基礎行列を高精度に推定する方法を提案した。基礎行列は多視点画像における同一物体の画面座標の対応関係を表すため、これにより、LIC の高速化も期待される。提案方法は画像の特徴点を信頼度関数で評価し、信頼度の高い特徴点のみを用いて対応点を検索する方法、および特徴点の抽出に色情報を付加した方法である。

(2) 平成 25 年度以降

平成 24 年度に得られた成果をより発展させるため、スパース表現を符号量 - 歪特性を制御する方法、従来の動画像符号化では個別の処理と考えられていた動き補償予測や DCT など一つのスパース問題として統合し、符号化効率の向上を目指す方法、多視点画像の基礎行列推定の更なる高精度化、GPU などを用いた並列処理による高速化、の 4 点について研究を進めた。以下に詳細を説明する。なお、平成 25 年度よりスパース解法には AIHT (Accelerated Iterative Hard Thresholding) を用いた。

符号量 - 歪特性の制御

一般のスパース解法では着目するマクロブロックの近似誤差と結合係数ベクトル w のスパース性を基準に最適解を計算する。しかし、実際の動画像符号化では近似精度、すなわち再生画像の画質の他に動きベクトルや DCT 係数の符号量も考慮する必要がある。動きベクトルの符号化では、既に符号化された近傍マクロブロックとの差ベクトルを符号化対象とするため、符号量 - 歪特性の制御において、例えば画質を犠牲にしてもできるだけ符号量を少なくするために隣接するマクロブロックとの差が小さい動きベクトルの参照ブロックを優先的に選ぶ、といった工夫が必要である。そこで、一旦全参照候補ブロックのノルムを 1 に正規化した後、差ベクトルが小さいブロックほど大きな定数を乗じる関数を提案し、符号量 - 歪特性の制御を試みた。

動き補償予測と DCT などの統合

従来、動画像符号化は動き補償予測で時間的冗長性を削減し、予測誤差に対して DCT を施して空間的冗長性を削減する 2 段階で構成されており、これがハイブリッド符号化と呼ばれる所以である。提案法はこの 2 段階を 1 つのスパース問題として解く。つまり、式(1)の行列 A に DCT の基底ベクトルを加え、着目マクロブロック b を参照マクロブロックと DCT の基底の線形結合で表現する。更にこれを発展させ、H.264/MPEG-4 AVC で行われているフレーム内予測を加え、更なる残余誤差の低減を試みた。

基礎行列推定の更なる高精度化

基礎行列の推定には一般に最尤推定が用いられるが、対応点の座標の誤差分布が正規分布に従うと仮定し、誤差分布におけるマハラノビス距離を最小化する。しかし、基礎行列の精度は対応点の精度に依存するため、誤対応点等の座標誤差が非常に大きな対応点を用いた場合、基礎行列の精度が著しく低下する。従って、基礎行列を高精度に推定するためには、正しい対応点を多く残しつつ、誤対応点を除去することが重要である。この研究ではエピソード方程式の誤差が正規分布に従うことを活用し、F 検定を用いた反復処理により誤対応点の除去を行う手法を提案した。

GPU などを用いた高速化

スパース解法の AIHT を並列化することにより高速化を行った。AIHT のプログラムは Matlab で書かれていたが、これを C++ に移植し、GPU の開発環境である CUDA および CPU の複数コアで並列処理する OpenMP を活用して並列化した。

4. 研究成果

3. 研究方法 (2) 平成 25 年度以降の から に示した項目に従って説明する。

(1) 符号量 - 歪特性の制御

提案法に従い、参照マクロブロック、つまり、行列 A の列ベクトル a_i に掛ける定数 L_i を、 a_i に必要な符号量 c_i によって定める関数を提案した。その関数の概観を図 1 に示す。

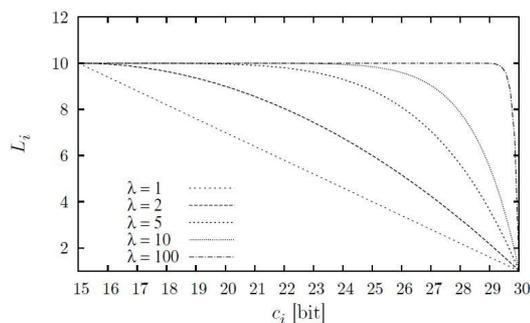


図 1: 符号量 C_i と乗数 L_i の関係

図 1 より、この関数はパラメータ λ によってその形状が変化する。つまり、 λ が小さいほど符号量の少ない参照ブロックが選択されやすくなる。H.264/MPEG-4 AVC などの最新の国際標準ではエントロピー符号化に算術符号を使っており、動きベクトルに割り当てられる符号量を厳密に求めることができないため、ここでは MPEG-2 のハフマン符号によるビット数を用いた。参照ブロック数 K を 5 に固定し、 λ を変化させた時の符号量 - 歪特性を図 2 に示す。Method 1 と 2 の方法が書いてあるが、これは隣接マクロブロックが複数存在するときの選択方法の違いであり、大きな意味の違いはない。

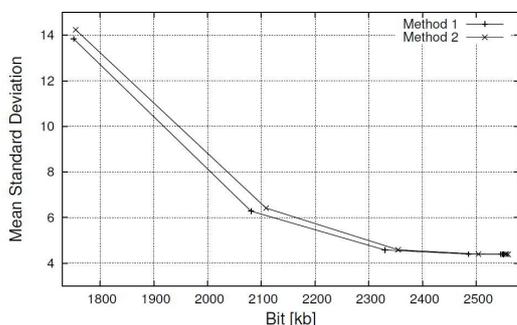


図 2: $K=5$ における符号量 - 歪特性

図 2 より, 提案法にて符号量 - 歪特性を制御できることがわかる.

(2) 動き補償予測と DCT などの統合

スパース問題の対象として, 動き補償予測と DCT を用いた場合 (Method 1), 動き補償予測とイントラ予測を用いた場合 (Method 2), 動き補償予測と DCT とイントラ予測全てを用いた場合 (Method 3) における着目ブロックの PSNR を示す. PSNR 大きな値ほど高画質であることを示す.

表 1: 誤差の比較

foreman				
K	Conventional Method	Method 1	Method 2	Method 3
1	32.559	32.559	32.573	32.573
2	34.651	34.652	34.679	34.679
3	35.203	35.687	35.706	35.716
4	35.608	36.248	36.258	36.277
5	35.947	36.668	36.671	36.701
6	36.248	37.016	37.012	37.046
7	36.521	37.331	37.302	37.360
8	36.773	37.599	37.584	37.631
9	37.010	37.863	37.821	37.890
10	37.234	38.094	38.047	38.116

K は用いる参照ブロック, DCT, イントラ予測画素ベクトルの総和である. 従来法では動き補償予測のみにスパース解法を適用した. 表 1 より提案法 Method 1 から Method 3 すべてで従来法より PSNR が向上しており, Method 3, つまり動き補償予測, DCT, イントラ予測全てを用いる方法が最善との結果を得た.

(3) 基礎行列推定の更なる高精度化

平成 24 年度に実施した, 信頼度関数を導入した結果を図 3 に示す. 図 3(a) が従来法による 2 視点間の対応点検索結果, 図 3(b) が提案法で除去された対応点である. 提案法では信頼度の低い特徴点を除去して対応点検索を行うため, 検索候補が少なくなり, 対応点検索が高速化されて基礎行列の精度も向上させることができた. また, F 検定を最尤推定に用いた結果を図 4 に示す. 図 4 の左右は同じ物体で視点の異なる画像である. 青線が除去された対応点, 白線が正しいと判定された対応点である. 図 4(a) の従来法では OpenCV などを用いられる RANSAC 法を用いた. 図 4(b) の提案法は効率よく誤対応点を除去し, 基礎

行列を高精度に推定することが示された. また, 提案法は繰り返し法であり, 誤対応点除去に関する閾値の依存性が改善された

(4) GPU などを用いた高速化

使用した計算機サーバは 2 枚の GPU を保持しているため, これらを CPU の 2 つのコアで並列処理した. 5 つのカメラで構成される多視点画像にて提案法を実行し, 通常の CPU に比べて 2.7 倍の高速化を実現した. 提案法はより多数枚の GPU ボードおよび CPU コアに容易に応用できるため, 更なる高速化が可能である.



(a) 従来法による対応点

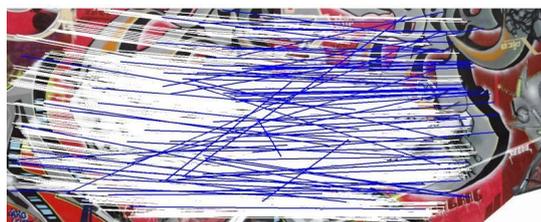


(b) 提案法で除去された対応点

図 3: 信頼度関数による対応点の除去



(a)



(b)

図 4: F 検定を用いた誤対応点の抽出

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11件)

K. Isechi, and Y. Kuroki, "Local Intensity Compensation Using Sparse Representation with Additional Sparse Dictionary," APSIPA ASC 2014, 査読有, 2014年, pp.1-4, DOI: 10.1109/APSIPA.2014.7041737.

K. Isechi, S. Motomatsu, and Y. Kuroki, "Parallel Processing of Motion Compensation Using Sparse Representation for Multi-view Video Coding," SISA 2014, 査読有, 2014年, pp.140-143, http://www.ieice-sisa.org/?page_id=402.

K. Isechi and Y. Kuroki, "Acceleration of Local Intensity Compensation Using Sparse Representation with Parallel Processing," GCCE 2014, 査読有, 2014年, pp.153-156 DOI: 10.1109/GCCE.2014.7031203.

K. Isechi, and Y. Kuroki, "Motion Compensation Using Sparse Representation Considering Rate-distortion Characteristics", IWAIT 2014, 査読有, pp. 398-401, <http://webstaff.knutt.ac.th/~iauaroen/iwait2014/>.

K. Inoue, H. Saito, K. Isechi, and Y. Kuroki, "An Inter-Prediction Method Using Sparse Representation for High Efficiency Video Coding," IEICE Trans. on Fundamentals, 査読有, vol.E96-A, no.11, 2013年, pp. 2191-2193, Online ISSN: 1745-1337.

R. Okutani and Y. Kuroki, "An Estimation of the Fundamental Matrix Using Hybrid Statistics," VCIP 2013, 査読有, 2013年, pp.1-6, ISBN: 978-1-4799-0288-0.

K. Isechi, and Y. Kuroki, "Motion Vector Prediction for Local Intensity Compensation Using Sparse Representation," ISTS 2013, 査読有, 2013年, no.163, <http://ists2013.kumamoto-nct.ac.jp/>.

K. Inoue, H. Saito, and Y. Kuroki, "Local Intensity Compensation Using Sparse Representation", ICPR 2012, 査読有, 2012年, pp.951-954, ISBN: 978-1-4673-2216-4.

K. Inoue, H. Saito, and Y. Kuroki, "Motion Compensation Using Sparse Representation," SISA 2012, 査読有, 2012年, pp.181-184, http://www.ieice-sis.org/?page_id=31.

R. Okutani and Y. Kuroki, "Accurate Estimation of the Fundamental Matrix by Limiting the Number of Corresponding Keypoints", SISA 2012, 査読有, 2012年, pp.162-166, http://www.ieice-sis.org/?page_id=31.
R. Okutani and Y. Kuroki, "Accurate Estimation of the Fundamental Matrix by Statistical Hypothesis Test," ISTS 2012, 査読有, 2012年, pp.449-452, <http://www.ists2012.kmitl.ac.th/>

[学会発表](計2件)

奥谷遼, 黒木祥光, "複数の統計量を用いた誤対応点除去による基礎行列の最尤推定," 信学技報, vol. 113, no. 318, IE2013-67, pp. 27-32, 2013年11月26日, 久留米高専(久留米市).

伊瀬知光平, 黒木祥光, "スパース表現を用いた動画像符号化," 信学技報, vol. 113, no. 318, IE2013-61, pp.1-4, 2013年11月25日, 久留米高専(久留米市).

[図書](計0件)

[産業財産権]

出願状況(計0件)

取得状況(計0件)

[その他]

ホームページ等

久留米高専産学民連携テクノセンター研究者プロフィール 黒木祥光

<http://www101.cc.kurume-nct.ac.jp/profile/CIS/kuroki.html>

6. 研究組織

(1)研究代表者

黒木 祥光 (KUROKI YOSHIMITSU)

久留米高専・制御情報工学科・准教授

研究者番号: 60290847

(2)研究分担者なし

(3)連携研究者なし