

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 27 日現在

機関番号：13901

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24560450

研究課題名(和文)自由視点映像の新しい表現形式と取得原理の研究

研究課題名(英文)A study on a new representation and acquisition principle of a free-viewpoint video

研究代表者

藤井 俊彰 (FUJII, Toshiaki)

名古屋大学・工学(系)研究科(研究院)・教授

研究者番号：30273262

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：本研究では、自由視点映像の生成においてViewとDepthを用いるという従来のCG手法からのアナロジーに基づく手法を根本的に見直し、自由視点映像を生成するための新しい表現形式(Representation)を研究した。一つ目は、EPIブロックの表現法として、基底ブロックの線形結合や辞書による光線空間の効率的な表現法を検討し、有効性を示した。二つ目として、符号化開口の考え方をを用いて実際の光学系で実現可能な圧縮取得について検討した。4次元DCT領域における重み付きL1ノルム最小化法を提案し、復元品質について検討した。

研究成果の概要(英文)："View plus Depth" data has been considered as a promising data format for future 3-D visual communication. However, this format is not feasible for real situation where no accurate depth information is available. We addressed the problem by proposing a new representation format that requires no depth information. We focused on a ray space representation, which can generate free-viewpoint images without depth data. First, we proposed to adopt compressed sensing to sparsely sense and reconstruct a ray space. We focused on an Epipolar Plane Image (EPI) block and approximated it by using common orthonormal basis in compressed sensing procedure. Second, we proposed a statistically weighted model in the reconstruction of compressively sampled ray space. For both cases, in the experiment, we confirmed that the proposed method can achieve much better reconstruction quality for both 2D image patch and 3D image cube cases.

研究分野：情報通信工学

キーワード：自由視点映像 奥行きマップ 光線空間 疎表現 エピポーラ平面画像 圧縮センシング

1. 研究開始当初の背景

3次元映像や自由視点映像などが注目を集めており、最近では裸眼で観察できる多視点のディスプレイも市販されるまでに至っている。このような自由視点映像や多視点映像として表示されるコンテンツとしては、CG映像が用いられていることが多い。これは自由視点/多視点映像の作成が容易である上、テクスチャとポリゴンというコンパクトなデータで表現できるという事情による。一方で、実写映像の自由視点や多視点のコンテンツを作ろうとすると、撮影が大変であることはもとより、情報量が2次元映像の数倍～数百倍になるため、その取り扱いや蓄積、伝送が困難という問題に直面しており、実際の利用が少ないのが現状である。

この問題に対し、実写映像に対して多数の視点位置からの映像(View)と奥行情報(Depth)とを有したデータ形式が考えられている。この両データは、CGにおけるテクスチャとポリゴンデータに相当するため、レンダリングによって自由な視点の映像が作ることが可能であり、従って3次元ディスプレイ表示に必要な多視点映像の生成も可能になる。映像・音声の国際標準化組織であるMPEG(Moving Picture Experts Group)においてもこのデータ形式が次世代の3次元テレビのデータ形式の有力候補として挙げられている。

ところが、このデータ形式を元にして通信・放送への応用を考える場合、実際の物体のDepthデータをいかに取得するかという本質的な問題が残されている。この問題への一つの研究の方向性は、コンピュータビジョン(CV)の技術を用いるものがある。しかし、2次元映像からDepthを復元することは不良設定問題であるため安定に正確な奥行きを求めることはできず、物体表面の滑らかさ拘束など高次の知識を利用して近似的に解いているのが現状である。他の研究の方向性は、光を照射するなどの能動的手法により奥行きを直接計測しようとするものであるが、取得できるDepthの範囲や得られたDepthの品質が低く、自由視点映像生成に用いることは困難である。

そもそも、多視点映像とは実シーンの光線空間のサンプリングであり、自由視点映像を生成とは限られた標本値からの信号の再構成・補間問題に他ならない。CG手法の利用を念頭にViewとDepthを用いる方法は自由視点映像生成の一手法に過ぎず、信号の再構成・補間問題という広い観点から自由視点映像の表現法(Representation)を開拓し、またその情報を得るための取得原理について研究することにより、この問題が解決可能なのではないかとこの着想に至った。

このように、研究開始当初には、ViewとDepthを用いるという手法が一般化しつつあるものの、高精度なDepthを直接取得する

手法がないという背景があり、そのためDepthを使わずに3次元映像の情報を表現する手法が強く求められていた。

2. 研究の目的

本研究では、自由視点映像の生成においてViewとDepthを用いるという従来のCG手法からのアナロジーに基づく手法を根本的に見直し、自由視点映像を生成するための新しい表現形式(Representation)を研究する。さらに、開発した表現形式に基づくデータを得るために必要最低限の情報は何かについて研究し、その取得原理の開拓と検証を行う。

これまでの自由視点映像の生成・伝送の研究は、主にDepthをいかに求めるかという点に関心が集まっていた。コンピュータビジョン(CV)の分野では、視差推定の問題を視差の滑らかさ拘束を含めたエネルギー関数の最小化問題として定式化し、Graph-CutやBelief Propagationを用いて問題を解くことが一つの研究の方向性となっている。また、光照射などの能動的手法によってTime-of-Flight方式に基づいて奥行きを直接計測手法も試みられている。映像・音声の国際標準化会合であるMPEGでは、多数の視点位置からのViewとDepthとを有したデータ形式が有力候補として挙げられており、N-view N-depth方式と呼ばれている。しかしながら、このようなViewとDepthに基づく自由視点映像生成手法は、Depth取得の手法の安定性、得られたDepthの品質に問題を抱えており、実用化に当たってはこの問題を解決することが強く求められていた。本研究は、全く新しいアプローチでこの問題の解決を試みるものであり、疎な光線空間の標本化・再構成という観点から自由視点映像を生成するための必要最低限の情報を明らかにしていこうとするものである。さらに、本研究成果を実システムへと実装してことにより、自由視点映像の実用化と普及へ道を開くものである。MPEGなどの標準化作業がDepthベースを中心に動きつつある現状を鑑み、世界に先駆けて新たな自由視点映像原理の開発とシステム構築を行うことを目的とする。

3. 研究の方法

ユーザ指定の自由な視点の画像を生成する手法として、光線空間法がある。これは規則的に配置された多数のカメラ画像を並べたものから生成され、自由視点画像の生成は光線空間からの切り出しによって実現できるという性質を持っている。従って、すべての自由視点映像生成のアルゴリズムは何らかの形で光線空間を再構成する手法であると考えることができる。例えば、3次元ポリゴンモデルを用いた映像生成や3Dワーピングを用いた映像生成では、形状や奥行きとそ

ここに付与されたテクスチャ情報をもとに、2次元カメラで投影された映像を生成する。これは、形状・奥行きとテクスチャ情報をもとに光線空間を再構成していることに相当する。

本研究では、密な間隔で取得された光線空間での理論を拡張し、比較的疎な光線空間についての標準化・再構成の理論へと発展させる。次に、得られる標本値の情報を表現形式 (Representation) として定義し、自由視点映像の生成はこの表現形式からの原光線情報の再構成に相当するものとして定式化する。さらに、取得における標準化のプロセスを実現するものとして取得原理を探求する。最後に、これらの知見を統合し、自由視点映像の新しい表現形式と取得原理の開拓のための基礎資料として整備していく。

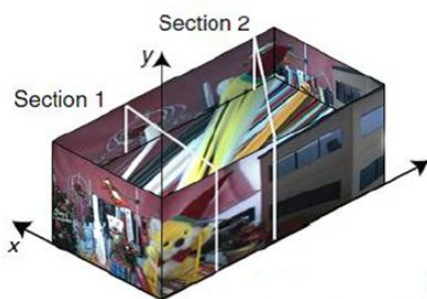


図 1：光線空間データの例

4. 研究成果

(1) 多視点画像を並べて水平断面をとった時に現れるエピソード平面画像 (EPI) を効率的に表現する手法を検討した。EPI をブロックに分割し、各ブロックのデータを少ない数の要素画像の線形結合で表すことができる手法を開発した。図 2 にその結果を示す。図 2 左の光線空間ブロックは 64 次元であるが、図 2 右に示した 4 つの基底の線形結合でよく近似ができることを示した。これは画像符号化でよく知られている変換符号化と密接に関わりがあるが、DCT のような固定の基底ではなく、EPI の性質を考慮した基底である必要がある。

また、このような基底の画像としてどのようなものを用いるのがよいかについて検討を行った。まずフーリエ変換基底や DCT, Matching Pursuit に基づく基底を検討した後、多数の EPI をトレーニングデータとして用いて学習により辞書を修正していくアルゴリズムを開発した。図 3 に学習により生成された辞書の一例を示す。

このように、Depth を用いずに光線空間を表現する手法として、EPI ブロックの表現法を検討した。基底ブロックの線形結合や辞書による光線空間の効率的な表現法を開発し、その有効性を確認した。



図 2：光線空間ブロックの基底画像への分解

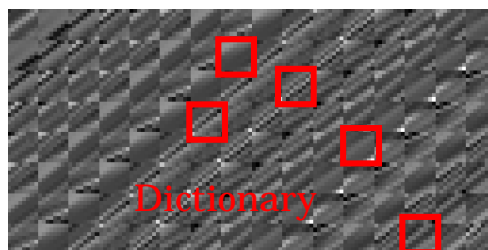


図 3：学習により生成された辞書画像

(2) 光線空間データの取得方法として、符号化開口により圧縮された形で光線情報をセンシングしておき、計算により非圧縮の光線空間データを復元する手法がある。本研究では、その復元手法として、4次元 DCT 領域における L1 ノルム最小化を用いて復元する手法について検討し、さらに重み付き L1 ノルム最小化を用いた場合の復元品質について検討した。

まず、符号化開口による撮像過程について簡単に述べる。符号化開口法では、光線空間は、開口面と撮像面の 2 つの平面で定義され、開口面 (s, t) を通り撮像面 (u, v) に入射する光線は $F(s, t, u, v)$ で表される。開口面のある 1 点 (s_1, t_1) を開けると、その点を視点とする像を撮像面で得られる。別の点 (s_2, t_2) を開けると、別の視点の像を得ることができる。 (s, t) を 1 点ずつ順次開けて撮影すると多眼撮影となり、複数の点を同時に開けると符号化開口撮影となる。前者による光線空間の取得も可能だが、後者を用いると S/N 比の改善や、圧縮センシングの導入による効率的な光線空間の撮影が可能になる。圧縮センシングとは、所望のデータを圧縮された形 (限られた観測点) で取得し、計算により非圧縮データを復元する手法である。

本手法において、まず最初に検討したのは L1 ノルム最小化である。4次元 DCT 領域において L1 ノルムを最小化することにより、少ない観測点数のデータから原データを復元する実験を行った。次には、重み付き L1 ノルム最小化を検討した。さらに、実際の光学系を考慮した検討へと進めていき、これらの手法を符号化開口の光学系にあてはめ、その有効性を確かめた。

光線空間の圧縮センシングと復元過程を計算機上でシミュレーションした。光線空間として Stanford 大学のシーケンス (Knights) のうち、 4×4 の 16 視点を 256×256 画素にリサイズし、これらを開口面上の異なる点からの観測画像とみなした。撮像面上のブロックサイズは 8×8 の 64 画素とした。観測の回数は $M = \{2, 3, \dots, 15\}$ とした。開口

のパターンは、で 256 段階の値を一様分布によって生成した乱数とした。ただし、各列の平均値を揃え、撮影全体で開口の各点に入射する光量が等しくなるように調節した。また、乱数の実現値による影響を除くため、各 M に対し 10 回の試行を行い結果を平均した。復元手法として最小二乗ノルム復元法 (Least Square 法), $L1$ ノルム最小化法, 重み付き $L1$ ノルム最小化法の 3 つを用い, PSNR および SSIM で復元品質を評価した。重み付き $L1$ ノルム最小化復元における重み行列 W は、同大学の別シーケンス (Truck) より学習した。

図 4 に、観測回数 M を変化させたときの 3 つの手法による復元品質を示す。全体的な傾向として、 M の増加に伴い PSNR および SSIM が向上することがわかる。3 つの復元手法を比較すると、PSNR, SSIM のいずれの尺度 (どちらも大きいほど良い) においても、重み付き $L1$ ノルム最小化復元がほぼ一貫して優れた性能を示している。図 5 は、ある一視点の (a) 正解画像, (b) 最小二乗ノルム復元による画像, (c) $L1$ ノルム最小化復元による画像、および (d) 重み付き $L1$ ノルム最小化復元による画像である。なお、(b), (c), (d) は、 $M = 5$ として同一の観測行列を用いた。最小二乗ノルム復元では画像がぼけており、 $L1$ ノルム最小化復元でもぼけが除去できていない。一方、重み付き $L1$ ノルム最小化復元では、背景の壁のブロックの境目の線がより鮮明になるなど、画質が改善された。これらの結果より、符号化開口法で圧縮センシングされた光線データの復元には 4 次元 DCT 領域における重み付き $L1$ ノルム最小化が有効であることがわかる。

このように、実際の光学系での実現可能性を考慮した上での結果を比較したところ、重み付き $L1$ 最小化法が最も高い性能を示した。今後は、この手法に基づいて実際の光学系を構築し、実データへと適用することである。

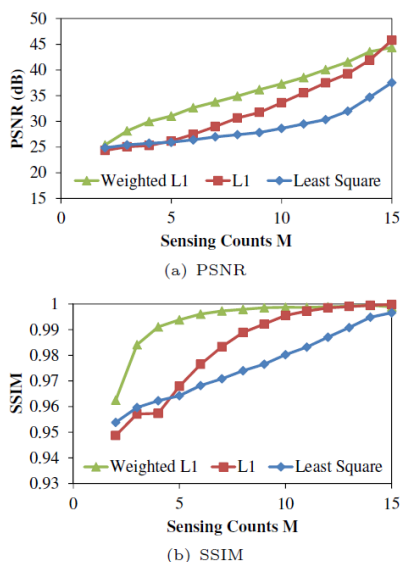


図 4 : 復元品質の評価結果

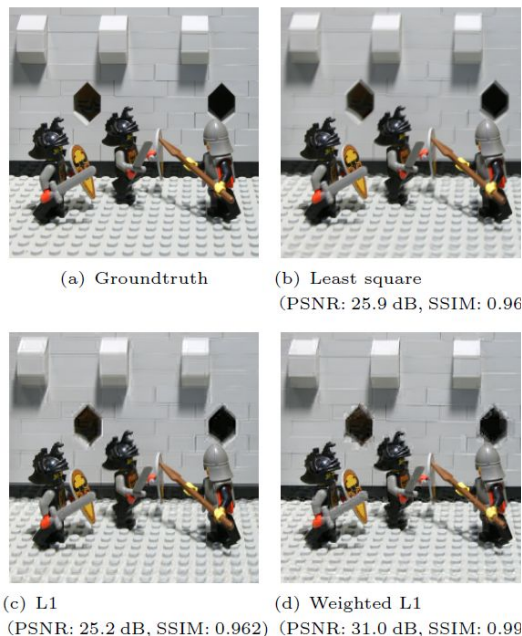


図 5 : 正解画像と復元画像

5. 主な発表論文等

[雑誌論文] (計 3 件)

宮城 雄介, 高橋 桂太, 要 強, パナヒプル テヘラニ メヒルダド, 藤井 俊彰, "符号化開口により圧縮センシングした光線空間データの重み付き $L1$ ノルム最小化による復元手法の検討," 電子情報通信学会論文誌, Vol. J98-D, No.9, Sep. 2015. (査読有) (採録決定)

Qiang Yao, Keita Takahashi, Toshiaki Fujii, "Reconstruction of Compressively Sampled Ray Space by Statistically Weighted Model," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E97-A, No. 10, pp. 2064-2073, 2014.10.1. (査読有) (https://www.jstage.jst.go.jp/article/transfun/E97.A/10/E97.A_2064/_article)

Qiang Yao, Keita Takahashi, Toshiaki Fujii, "Compressed Sensing of Ray Space for Free Viewpoint Image (FVI) Generation," ITE Transactions on Media Technology and Applications, vol. 2, no. 1, pp. 23-32, 2014. (査読有) (<http://dx.doi.org/10.3169/mta.2.23>)

〔学会発表〕(計 24件)

Qiang Yao, Keita Takahashi, Mehrdad Panahpour Tehrani, Toshiaki Fujii, "Joint Directional-Positional Multiplexing for Light Field Acquisition by Kronecker Compressed Sensing," 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2015, 19-24 April 2015, Brisbane, Australia.

宮城 雄介, 高橋 桂太, 要 強, パナヒプル テヘラニ メヒルダド, 藤井 俊彰: "符号化開口により圧縮センシングされた光線空間の重み付き L1 正則化による復元手法の検討," 画像符号化シンポジウム, 静岡県伊豆市, P-4-13 (2014.11.12-14).
(ベストポスター賞、学生論文賞受賞)

Qiang Yao, Keita Takahashi, Toshiaki Fujii, "Reconstruction of compressively sampled ray space by using DCT basis and statistically-weighted L1 norm optimization," IS&T/ Electronic Imaging 2014, Vol. 9020-34, Hilton San Francisco, California, USA (2014.2.2-6)

Qiang Yao, Keita Takahashi, Toshiaki Fujii, "Overcomplete Compressed Sensing of Ray Space for Generating Free Viewpoint Images," Asian-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2013, OS.30-SPS.4, 85 Sky Tower Hotel, Kaohsiung, Taiwan (2013.10.29-11.1)

Qiang Yao, Keita Takahashi, Toshiaki Fujii, "Three Dimensional Compressed Sensing of Ray Space," 電気関係学会東海支部連合大会, M1-9, 静岡県浜松市/静岡大学(2013.9.24-25).

Qiang Yao, Toshiaki Fujii, "Sparse Acquisition of Ray Space by Sensing EPI," 3次元画像コンファレンス2013, pp.149-152, 東京 (2013.7.4-5).

Qiang Yao, Toshiaki Fujii, "Highly Efficient Representation of Ray Space for FTV System," Three Dimensional Systems and Applications 2013, S5-1, Osaka, (2013.6.26-28).

〔その他〕

ホームページ

<http://www.fujii.nuee.nagoya-u.ac.jp/>

6. 研究組織

(1) 研究代表者

藤井 俊彰 (FUJII, Toshiaki)

名古屋大学・大学院工学研究科・教授

研究者番号: 30273262