

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 3 月 31 日現在

機関番号：32660

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23700217

研究課題名（和文）スパースサンプリングによる任意視点画像合成の画質向上の実現

研究課題名（英文）Image quality enhancement in arbitrary viewpoint image synthesis by the sparse sampling

研究代表者

保坂 忠明（HOSAKA TADAAKI）

東京理科大学・工学部・助教

研究者番号：60516235

研究成果の概要（和文）：複数の視点から撮像された画像を基に、カメラの存在しない仮想視点からの映像を推定する任意視点画像合成が盛んに研究されている。多くの従来法における問題は、任意視点画像に生じる欠損領域を精度よく補完することが困難であること、画像合成処理に必要なメモリ量や演算量が増大することである。本研究では、スパースサンプリングや圧縮センシングといった信号の再構成手法を利用して、これらの問題の解決を試み、有意な結果を得た。

研究成果の概要（英文）：Major problems of traditional methods to synthesize virtual-view images are that it is difficult to precisely complete missing pixels in synthesized images and that it requires huge amount of data storage and computational efforts. In this research, we propose methods to overcome these difficulties by using techniques of the sparse sampling and the compressed sensing. Experimental results show the validity of the proposed methods qualitatively and quantitatively.

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	1,700,000	510,000	2,210,000

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：画像情報処理，任意視点画像合成，スパースサンプリング，圧縮センシング

1. 研究開始当初の背景

複数の視点から撮像される画像を基にして、カメラの存在しない視点からの画像を推定する任意視点画像合成技術が、コンピュータビジョンの分野において盛んに研究されている。この技術は、ユーザの望む視点でスポーツ中継などを提供する自由視点テレビやテレビ会議における視線の一致などをはじめとする様々な仮想現実・複合現実応用の基盤技術として注目されている。

多くの任意視点画像合成技術では、まず多視点画像を利用し何らかの方法でカメラから被写体までの距離を推定する。そして、得られた距離情報と三次元幾何を基にして、仮

想視点カメラの各画素に入射する光線を計算により求めることで所望の視点からの映像を得る。従来手法の問題点として、1) 任意視点画像に生じる欠損を高精度に補完することが困難である、2) 精度よく被写体の距離情報を推定するために要するメモリ量や演算量が大きくなる、ことが挙げられる。

一方で研究の開始当初は、信号処理・機械学習・情報理論などの分野でスパースサンプリングと呼ばれる技術が盛んに研究されていた。これは、少数の標本から原信号を復元するための技術である。近年では特に、圧縮センシングと呼ばれる手法が注目されている。圧縮センシングは、スパース性を持つ信

号（少数の基底によって表現されるベクトルとして定式化される）に対して、原信号の次元数よりも少ない数の線形観測から、原信号の完全な再構成を可能にするための方法と条件に関する理論である。

2. 研究の目的

前節で述べた任意視点画像合成の二つの問題点を解決することを本研究の目的とする。具体的には、1) 高周波成分を保存しつつ、画素値の連続性を保つように任意視点画像に含まれる欠損画素を補完すること、2) 合成品質の劣化を抑えつつデータ量と演算量を削減すること、である。結果として、同程度の演算量において従来手法と比較した場合に、任意視点画像の画質向上がもたらされることを期待している。

以上の目的を達成するために、本研究では、前節で述べたスパースサンプリング・圧縮センシングを利用する。そして、実画像を用いた評価実験により提案手法の有効性を検証する。

3. 研究の方法

本研究では、複数のカメラを光軸が平行になるように配置する。そして、各カメラ位置を空間上の一点で表した場合に、それらの重心に相当する位置を仮想視点とする。

以下では、1) 高周波成分を保存しつつ、画素値の連続性を保つように任意視点画像に含まれる欠損画素を補完する方法、2) 合成品質の劣化を抑えつつデータ量と演算量を削減する方法、に分けて述べることとする。

(1) 欠損画素を補完する方法

まず、手法の基本となるスパースサンプリングについて述べる。 M 次元の観測ベクトル \mathbf{y} を過完備基底で表現することを考える。 $N(>M)$ 個の基底を並べた $M \times N$ 行列を A で表し、 N 次元の基底係数ベクトルを \mathbf{x} と表すと、ベクトル \mathbf{y} は

$$\mathbf{y} = A\mathbf{x} \quad (1)$$

と表現される。そして、ベクトル \mathbf{y} と行列 A が与えられた下でベクトル \mathbf{x} を求めることが目的となる。ここで、 $N > M$ であるため、一般には式(1)を満たすベクトル \mathbf{x} は無限に存在する。そこで、式(1)を制約条件とした l_1 ノルム最小化問題:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - A\mathbf{x}\|_2 \leq \varepsilon \quad (2)$$

を解くのがスパースサンプリングの基本的な考え方の一つである。ここで、 $\|\cdot\|_p$ は l_p

ノルムを表し、 $\varepsilon (\geq 0)$ は許容誤差を表す。この問題に対しては、線形計画法による厳密解の求解以外にも、LARS (least angle regression) などの貪欲法を基本とした高速な近似解法が知られている。

本研究では、任意視点画像に含まれる欠損画素をスパースサンプリングに基づき補完する。まず、各カメラ間でステレオマッチングを行い、各カメラを基準とした奥行きマップを得る。次に、オクルージョン領域や奥行き推定精度の悪い画素を除外して、空間幾何を基に仮想視点画像を推定する。この段階で得られる仮想視点画像には欠損画素が含まれている。この欠損画素を以下の手順により補完する。

1. 仮想視点画像を走査し、欠損画素 i を中心とした $u \times u$ 画素のブロック K_i を設定する。
2. 仮想視点画像内の画素 i に関して実画像上に規定されるエピポーラ線に位置する実カメラの画素集合を J とおく。画素 $j \in J$ を中心とする $u \times u$ 画素のブロック C_j を考える。以下のベクトルと行列を定義する。
 - ブロック K_i における確定画素（非欠損画素）の RGB 値を並べたベクトルを \mathbf{y} とする。
 - ブロック C_j に対して、(ブロック K_i の) 確定画素位置にある RGB 値を上と同様の順序で成分としたベクトルを基底として捉え、得られるすべての基底ベクトルを並べた行列を A とする。
 - ブロック C_j に対して、(ブロック K_i の) 欠損画素位置にある RGB 値を一定の順序で成分としたベクトルを同様に並べた行列を D とする。
3. 上で定義したベクトル \mathbf{y} と行列 A を用いて式(2)の最適化問題を LARS により解く。
4. 求めた係数ベクトル \mathbf{x}^* により、欠損画素は $D\mathbf{x}^*$ と推定される。以後、推定された欠損画素は、確定画素として扱われる。
5. 欠損画素がなくなるまで、手順 1 へ戻り処理を繰り返す。

(2) データ量と演算量を削減する方法

まず、手法の基本となる圧縮センシング

(スパースサンプリングの一種と捉えることができる)について述べる. N 次元ベクトル \mathbf{x} を原信号として, これを N 次元空間の正規直交基底で表現する. N 個の基底ベクトルを並べた $N \times N$ 行列を $\boldsymbol{\psi}$ とした時, 表現係数ベクトル $\mathbf{s} = \boldsymbol{\psi}^T \mathbf{x}$ がスパースであると仮定する. そして, ベクトル \mathbf{x} に対する M 個の線形観測 (成分の線形結合によってモデル化される観測) を考える. 線形結合係数を並べた $M \times N$ 観測行列を $\boldsymbol{\phi}$ とした時, M 次元観測ベクトル \mathbf{y} は,

$$\mathbf{y} = \boldsymbol{\phi} \mathbf{x} = \boldsymbol{\phi} \boldsymbol{\psi} \mathbf{s} \quad (3)$$

と表現される. $M < N$ の時にデータは圧縮されたことになる. 一方で, \mathbf{y} が与えられた時に \mathbf{x} を再構成することを考える. 圧縮状況である $M < N$ の場合には, 式(1)を満たす \mathbf{x} は無限に存在する. しかし, \mathbf{s} がスパースである時, l_1 ノルム最小化問題:

$$\mathbf{s}^* = \underset{\mathbf{s}}{\operatorname{argmin}} \|\mathbf{s}\|_1 \quad \text{subject to } \mathbf{y} = \boldsymbol{\phi} \boldsymbol{\psi} \mathbf{s} \quad (4)$$

によって得られる表現係数ベクトル \mathbf{s}^* がある条件の下で, 真の係数ベクトル \mathbf{s} に高確率で一致する. なお, 式(4)の最適化問題は線形計画法によって厳密に解くことができる.

第1節で述べたように, 多くの任意視点画像合成手法では, 1) 複数視点の実画像に基づいた被写体までの距離の推定, 2) 取得距離情報と三次元幾何を基にした仮想視点画像の推定, という二つの処理からなる. 本研究では, 実画像は未知であるが, その線形観測が可能な状況を扱う. 提案手法では, 線形観測ベクトルを用いて距離推定を行い, 再構成過程を経て仮想視点画像を得る.

画素値を一定の順序で並べてベクトル \mathbf{x} として表現する. 画像ベクトル \mathbf{x} は未知であるが, \mathbf{x} に対する線形観測ベクトル \mathbf{y} ($M < N$) は得られるとする. $M < N$ であるので, 画像 \mathbf{x} に対して処理を行う場合に比べて, 提案手法ではデータ量と演算量を削減することができる.

① ブロック圧縮センシングによる画素位置の保持

画像には位置に関する情報が含まれているが, 式(3)で表現される線形観測では, その情報が陽には失われる. そこで, 画像の位置情報を保持するためにブロック圧縮センシングを利用する. 具体的には, 以下のようにする:

- 1) 画像を $W \times W$ 画素からなる小領域に重複を許容せずに分割する. 各小領域に対し,

ラストスキャンの順にインデックス $b(=1, 2, \dots, B; W^2 B = N)$ を割り当てる. また, 小領域 b に含まれる画素を $\hat{\mathbf{x}}^b$ で表す. このとき, 画像ベクトル $\mathbf{x} = [\hat{\mathbf{x}}^{1T}, \hat{\mathbf{x}}^{2T}, \dots, \hat{\mathbf{x}}^{BT}]^T$ と構成する.

- 2) 小領域ごとの基底表現を考える. すなわち, 基底行列 $\boldsymbol{\psi}$ は, $W^2 \times W^2$ のサイズの小行列 $\hat{\boldsymbol{\psi}}$ を用いて

$$\boldsymbol{\psi} = \begin{bmatrix} \hat{\boldsymbol{\psi}} & & & \\ & \hat{\boldsymbol{\psi}} & & \\ & & \ddots & \\ & & & \hat{\boldsymbol{\psi}} \end{bmatrix} \quad (5)$$

と定義される. 本研究では, 小行列 $\hat{\boldsymbol{\psi}}$ を W^2 次元空間における離散コサイン変換の基底行列によって定義する. 小行列の配置されていない成分はすべて 0 である.

- 3) 小領域ごとの線形観測を考える. 各小領域から m 個の線形観測 ($m = M/B$) が得られるとすると行列 $\boldsymbol{\phi}$ は

$$\boldsymbol{\phi} = \begin{bmatrix} \hat{\boldsymbol{\phi}} & & & \\ & \hat{\boldsymbol{\phi}} & & \\ & & \ddots & \\ & & & \hat{\boldsymbol{\phi}} \end{bmatrix} \quad (6)$$

と定義される. ここで, $m \times W^2$ のサイズの小行列 $\hat{\boldsymbol{\phi}}$ の各要素は $\{a, 0\}$ の二値乱数によって決められる. ただし, 線形観測の値が画素値の値域である $[0, 255]$ を越えないように a を定める. 小行列の配置されていない成分はすべて 0 である.

- 4) $b(=1, 2, \dots, B)$ 番目の小領域から得られる観測ベクトルを $\hat{\mathbf{y}}^b$ とすると, 全体の観測ベクトルは, $\mathbf{y} = [\hat{\mathbf{y}}^{1T}, \hat{\mathbf{y}}^{2T}, \dots, \hat{\mathbf{y}}^{BT}]^T$ と表される. ただし, 画素値が離散値であることを考慮して, 線形観測値も整数値になるように端数処理を施す.

② 線形観測ベクトルに基づく任意視点画像合成処理

ここでは特に, 左右に並べて配置された二眼カメラを考える. 上記の定式化に基づき, 左右カメラより線形観測ベクトル $\mathbf{y}_l, \mathbf{y}_r \in \{0, 1, \dots, 255\}$ をそれぞれ取得する. そして, 左右画像の線形観測値を特徴量として, 小領域単位でステレオマッチングを行い視差を推定する. ただし, 各小領域内では位置情報が失われ, 同一の視差が割り当てられることになる.

得られた視差を基に仮想視点画像の各小領域に対する線形観測ベクトル \mathbf{y}_v を推定する.

この処理は、左右の実画像の各小領域に対して得られた視差を基に線形観測ベクトル \hat{y}_l^b, \hat{y}_r^b ($b = 1, 2, \dots, B$) を仮想視点へ移動させることで行われる。ただし、双方向マッチングを考慮して信頼性の高い視差値のみを用いる。この段階では線形観測ベクトル y_b に欠損が含まれるため、周辺の画素値を用いて補完を行う。

得られた線形観測ベクトル y_b を線形計画法 (シンプレックス法) を用いて再構成することで仮想視点画像を得ることができる。最後に、距離推定および再構成における誤りを軽減させる目的で、バイラテラルフィルタによる平滑化処理を施し画像を出力する。

4. 研究成果

本節においても、1) 高周波成分を保存しつつ、画素値の連続性を保つように任意視点画像に含まれる欠損画素を補完する方法、2) 合成品質の劣化を抑えつつデータ量と演算量を削減する方法、により得られた成果を分けて述べることとする。

(1) 欠損画素を補完する方法により得られた成果

図 1 に示した四視点のカラー画像 (解像度 640×480 , 筑波大学多視点画像データベースより) を用いた評価実験について述べる。

欠損を含む任意視点画像を図 2(a) に示す (欠損画素を黒で表示している)。さらに、 $u=13$ として欠損画素を前節の手順にしたがって補完した結果を図 2(b) に示し、被写体の目の周囲を拡大した画像を図 2(c) に示す。図 2(d) に周辺画素値を用いた単純な補間による結果を示す。提案手法により、欠損領域 (目の輪郭) がボケずに自然に復元されていることが確認できる。また、本手法と単純な補間により得られた合成画像の PSNR は、それぞれ 33.77dB , 29.05dB (補完部分のみで評価すると、それぞれ 29.96dB , 22.63dB) であり、定量的にも提案手法の有効性が示された。

今後は、より複雑な形状をした被写体の場合に有効となる手法を検討していく。

(2) データ量と演算量を削減する方法により得られた成果

① データ量削減の成果

図 3 に示す二視点のグレイスケール画像 (解像度 640×480 , 筑波大学多視点画像データベースより) に対する線形観測ベクトル y_l, y_r が得られるとして、提案手法の有効性を検証した。

圧縮率 R を画素数 N と観測数 M の比によって $R=M/N (\leq 1)$ と定義して、 R および W を変化させたときに推定される仮想視点画像の PSNR を図 4 に示す。また、 $W=4$ において得られる仮想視点画像を図 5 に示す。圧縮率 R が小さくなるにつれて、PSNR が低下するものの $W=2, W=4$ では小領域あたりの線形観測数を 1 としても 30dB を越える画質を維持し、図 5(a) に示すように視覚的にも良好な画像が得られている。これらの結果から、提案手法によるデータ削減の有効性を確認できる。なお、 $W=6, W=8$ の場合には距離推定精度が低下して、合成画像の品質も低下してしまう。

現状では、小領域ごとにステレオマッチングを行うため、距離推定の精度が高いとは言えない。そこで、今後は視差推定の精度向上を検討する予定である。

② 演算量削減の成果

図 6 に、ステレオマッチング、 l_1 ノルム最小化、提案手法に含まれる全ての処理に要した演算時間を示す。なお、使用した PC の性能は Core i7-920 2.67GHz (シングルスレッド処理)、メモリ 12.0GB (RAM) である。提案手法におけるステレオマッチングの演算量は $O(W^{-1}N^{3/2}R)$ であり、図 6(a) に示される計算時間は、これに合致している。 l_1 ノルム最小化の計算時間については、図 6(b) に示すように、 $W=2$ と $W=4$ を比較すると計算時間がおよそ 4 倍になっている。これは小領域総数が 4 倍となることで、試行回数も 4 倍になるからであり、 $W=2$ と $W=4$ のような小規模なサイズのシンプレックス法を 1 回当たりに解くことに要する時間はほぼ変わらないことを意味している。 $W=8$ では、圧縮率 R が上昇すると急激に計算量が増大する。シンプレックス法では制約式と未知変数の数が増えるに従い、その演算量は急激に増大し、多項式オーダーでの求解は保証されていない。このことから、 $W=8$ における急激な計算時間の増大が説明できると考える。また、全行程において (線形観測値ではなく) 実視点画像を用いて処理を行った場合の計算時間は 528sec であり、図 6(c) に示す結果と比較すると、提案手法ではいずれのパラメータ設定においても大幅に計算時間を削減できていることが分かる。以上より、演算量の観点からも提案手法の有効性が示されたと言える。

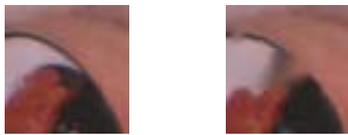
本研究では、欠損画素を補完する手法とデータ量・演算量を削減する手法を別々に提案しており、現状では両者を組み合わせて利用することはできていない。今後は、これら二つの手法を組み合わせることも検討していく。



図 1: 実視点画像



(a) 欠損画素 (b) 提案手法による結果



(c) 拡大画像(提案手法) (d) 拡大画像(単純補間)
図 2: 実験結果



図 3: 実視点ステレオ画像

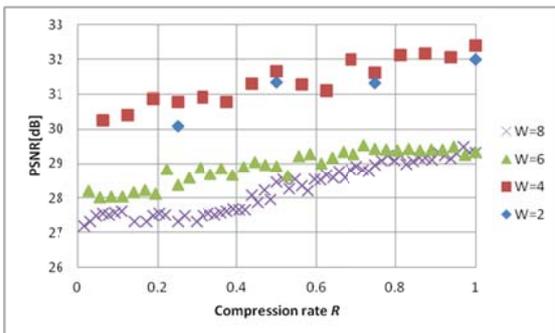
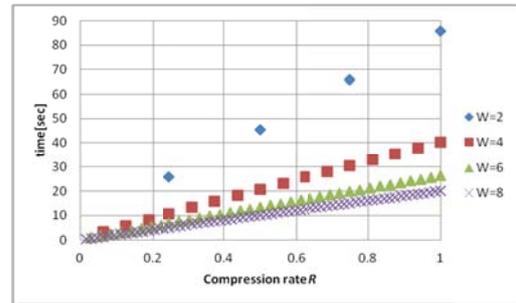


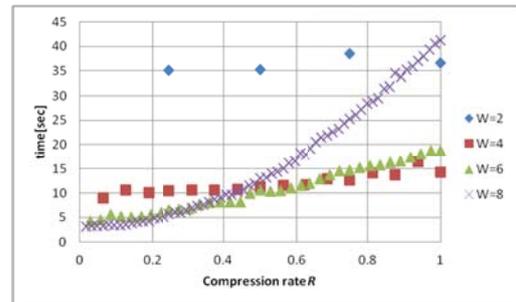
図 4: 圧縮率に対する PSNR の変化



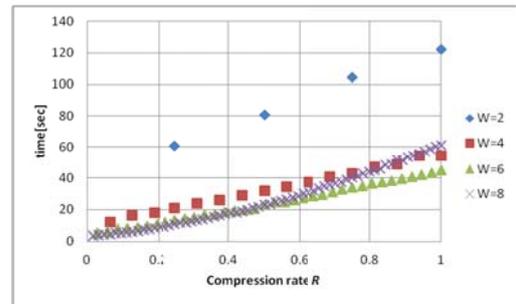
(a) $R=0.0625$ PSNR:30.25dB
(b) $R=0.5$ PSNR:31.66dB
図 5: 合成された仮想視点画像($W=4$)



(a) ステレオマッチング



(b) l_1 ノルム最小化



(c) 全行程

図 6: 演算時間の評価

5. 主な発表論文等

〔学会発表〕(計4件)

- ① 苫米地大, 保坂忠明, 浜本隆之, 「圧縮センシングに基づく自由視点画像合成の効率化」, 第15回情報論的学習理論ワークショップ (IBIS2012) (2012年11月8日, 筑波大学文京校舎).
- ② 苫米地大, 保坂忠明, 浜本隆之, 「任意視点画像合成における圧縮センシングを用いたデータ量の削減」, 2012年映像メディア処理シンポジウム(2012年10月26日, 静岡県ニューウェルシティ湯河原).
ベストポスター賞を受賞.
- ③ 保坂忠明, 浜本隆之, 「Compressed Sensingを利用した欠損画素補完による仮想視点画像の画質改善」, 第14回情報論的学習理論ワークショップ (IBIS2011) (2011年11月10日, 奈良女子大学).
- ④ 保坂忠明, 浜本隆之, 「Compressed Sensingによる任意視点画像の補完」, 2011年映像メディア処理シンポジウム(2011年10月29日, 静岡県ニューウェルシティ湯河原).
ベストポスター賞を受賞.

6. 研究組織

(1) 研究代表者

保坂 忠明 (HOSAKA TADAAKI)

東京理科大学・工学部・助教

研究者番号: 60516235