

研究速報

超解像処理による自由視点画像の画質改善

福嶋 慶繁^{†a)} 石橋 豊[†](正員)

Improvement of Free Viewpoint Image Quality via Super Resolution

Norishige FUKUSHIMA^{†a)}, Nonmember and Yutaka ISHIBASHI[†], Member

[†]名古屋工業大学大学院工学研究科, 名古屋市 Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya-shi, 466-8555 Japan
a) E-mail: fukushima@nitech.ac.jp

あらまし 本論文では, 自由視点画像合成と, 超解像処理を同時に行うことで, 前進した視点位置での仮想視点画像の高周波情報が失われる問題を解決し, その画質向上を図る. シミュレーション実験の結果, 超解像による手法が常に画像品質が最も高く, 線形補間法に比べて最大 5.1 dB の PSNR の向上を確認した.

キーワード 自由視点画像合成, 超解像, 光線空間法, カメラアレー

1. まえがき

多眼カメラを規則的に配置したものはカメラアレーと呼ばれ, そこから任意の視点の画像を合成する自由視点画像合成の研究が盛んに行われている. 文献 [1] では, 視点に応じた奥行を計算し, 必要な光線を補間することで合成を行っている. この手法で, カメラアレーから前進した視点位置の画像合成をすると, 線形補間等の影響で画像の高周波情報が失われる.

この問題を超解像処理 [2] を用いて解決する. 超解像とは, 同一物体を写す複数枚の画像を統合することで, 単体の解像度よりも高い解像度をもつ画像を合成する手法である. そのため, 超解像処理は線形補間等の拡大処理とは異なり高周波成分の復元も可能とする.

本論文では, 低解像度画像から高解像度画像を再構成する超解像ではなく, 前進位置での自由視点画像合成と超解像処理を同時に行う手法を提案する.

2. 自由視点画像合成とその解像度

ある位置 $p = (x, y, z)$ の仮想視点画像 I_p 上の光線 $r = (u, v)$ は, 近傍のカメラアレー上の画像 I_i (i は対応するカメラ番号) の重み付線形補間で表され, 所望の光線の画素の位置 r から近傍カメラの画像 i 中の対応する画素への視差ベクトル $d_{p,i}(r)$ を用いて以下で表される.

$$I_p(r) = \sum_{i \in V_p(r)} w_i(r) I_i(r + d_{p,i}(r))$$

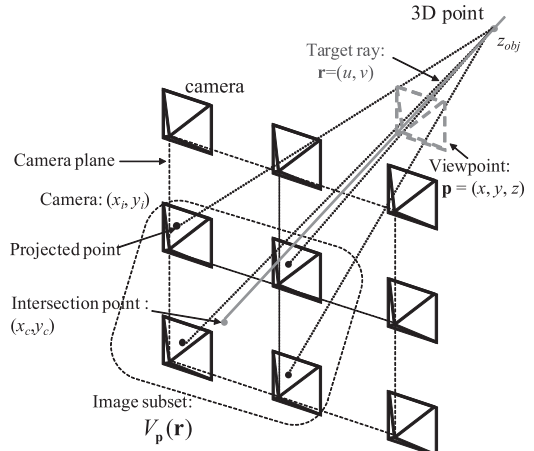


図 1 カメラアレーによる光線の参照

Fig. 1 The reference of ray from the camera array.

画像 i の光線に対する重み w_i は, 光線と多視点画像の距離に比例する関数であり, $\sum_{i \in V_p(r)} w_i(r) = 1$ となるように定義している. この各カメラへの視差 $d_{p,i}$ は仮想視点上の奥行画像 $zmap_p$ より導出可能である. ある光線 r の奥行値が $zmap_p(r) = zobj$ であるとき, $d_{p,i}$ は以下の式で求められる.

$$d_{p,i} = \begin{bmatrix} f_u(x_c - x_i)/zobj \\ f_v(y_c - y_i)/zobj \end{bmatrix}$$

ここで, f_u, f_v はカメラ焦点距離であり, x_c, y_c は光線 r とカメラアレー平面が交わる点の x, y 座標である. また, x_i, y_i は i 番目の実カメラの x, y 座標である. 図 1 は 3×3 のカメラアレーから仮想視点上の光線を参照する様子を示したものである. 本論文では, すべてのカメラの焦点距離が等しく, また u, v 各軸の焦点距離も等しいとして $f = f_u = f_v$ として扱う. $d_{p,i}$ は実数精度のベクトルである一方, 画素は整数精度であるため, 本論文では, 最近傍法または線形補間により, 画素参照を行う.

必要な光線が含まれる多視点画像を表す集合 $V_p(r)$ を, 光線空間 [3] の関係から, 例えば 4 近傍で表せば,

$$X = \lfloor x + uz \rfloor, Y = \lfloor y + vz \rfloor$$

$$V_p(r) = \left\{ \begin{pmatrix} X \\ Y \end{pmatrix}, \begin{pmatrix} X+1 \\ Y \end{pmatrix}, \begin{pmatrix} X \\ Y+1 \end{pmatrix}, \begin{pmatrix} X+1 \\ Y+1 \end{pmatrix} \right\}$$

となる. X, Y はカメラアレーの縦横のインデックスに相当する. これらの詳細や仮想視点上の奥行画像 $zmap_p$ の推定法は文献 [1] で述べられている.

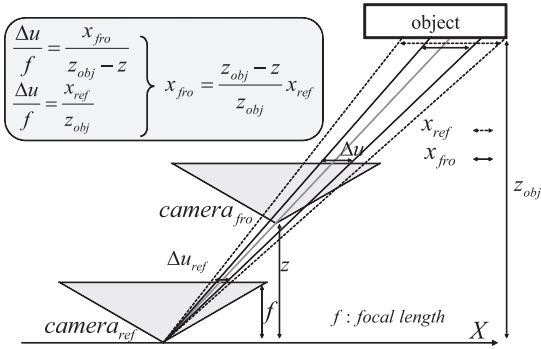


図 2 前進した仮想カメラの空間分解能
Fig. 2 The sampling pitch of the forwarding virtual camera.

さて、光線空間法では、光は幅のない線として扱われる。しかし実際のカメラでは、個々の受光素子が一定の面積をもつため、光線は幅をもつ。そのため、カメラアレーよりも前進した視点で画像合成したとき、入力となるカメラアレー上の画像解像度により、その仮想視点上の画像の空間分解能は制限される。ここでは、図 2 を用いて具体的に説明する。この図は、位置 z にある仮想カメラとカメラアレー平面中のカメラの空間分解能の違いを $X-Z$ 平面上で図示したものである。一つの画素が表す光線は、カメラの焦点を頂点、被写体の微小区間を底辺とする三角錐となり、その値は三角錐中の光の積分値となる。そのため、カメラから被写体が遠いほど、その積分する体積も大きくなる。

まず、画素間隔が Δu のとき、距離 z_{obj} にある物体は、カメラアレー上のカメラからは、区間 x_{ref} でサンプルされ、前進位置 z の仮想カメラからは区間 x_{fro} でサンプルされる。もし、同一の区間で物体をサンプリングするためには、理想的な多視点画像の画素間隔 Δu_{ref} は、図 2 中の幾何学的関係から次のようになる。

$$\Delta u_{ref} = \frac{1}{M} \Delta u = \frac{(z_{obj} - z)}{z_{obj}} \Delta u$$

つまり、前進した視点での理想的な自由視点画像は、カメラアレー上の実カメラの M 倍の解像度を必要とする。 v 軸に関して、焦点距離が等しければ u 軸と同様である。

3. 超解像

本論文では、MAP (Maximum A Posteriori) 推定である BTV (Bilateral Total Variation) 法 [2] で超解像を行う。MAP 法による超解像とは、低解像度の

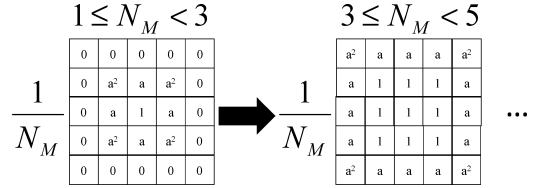


図 3 適応的な Point Spread Function
Fig. 3 Adaptive Point Spread Function.

観測画像を条件とし、事後確率を最大化する高解像度画像を求めるものである。この問題は、推定された高解像度画像の画素値と観測画素値の距離である誤差項と、高解像度画像の事前確率をもとにした拘束項の同時最適化問題として定式化される。これを、自由視点画像合成の場合に適用すると、以下のようになる。

$$\hat{I}_p(r) = \underset{I_p(r)}{\operatorname{argmin}} \sum_{i \in \mathbf{V}_p(r)} w_i(r) \|h_p(z_{map}(r)) * I_p(r) - I_i(r + \mathbf{d}_{p_i}(r))\|^2 + \lambda \gamma(I_p(r))$$

これは、合成画像 I_p に PSF (Point Spread Function) の h_p を畳み込んだものと、奥行に応じて対応させた近傍の入力画像 I_i との差が小さいかつ、画像の事前情報に一致するものを、推定合成画像 \hat{I}_p とするものである。この PSF は、レンズによる焦点ぼけを無視すれば、受光素子の単位面積当りの空間分解能から定義できる。目的の光線の拡大率を M とすると、その拡大率に応じて輝度値を積分すればよいが、取得画像は既に画素単位に量子化されているため、ボックスフィルタを変形したカーネルを用いる。まず、ボックスフィルタは、拡大率がちょうど奇数倍 (1 倍, 3 倍, 5 倍, ...) となると、該当する面積当りの積分値を表すことができるカーネルであると考えられる。これらの各カーネル間の中間の状態は、核を中心部と周辺部に分けて考える。まず、中心部は輝度値の 100% を積分に用いる。また周辺部は、対応する面積に応じて重みを変える。図 3 にこのカーネルを示す。図中の N_M は PSF カーネルの各要素の総和を示しており、また a は下記で定義される。

$$a = 0.5 * (M - \text{odd}(M)), (0 \leq a < 1)$$

odd は実数である M を奇数に切り捨てる関数である。この画素空間での表現を行列表現を用いて簡単化する。

$$\hat{I}_p = \underset{I_p}{\operatorname{argmin}} \sum_{i \in \mathbf{V}} \mathbf{W}_i \|\mathbf{H} \mathbf{F}_i I_p - I_i\|^2 + \lambda \gamma(I_p)$$

ここで、 F_i は画像 i の視差による移動を表す行列であり、 H は PSF を表している．また、 $\|\cdot\|$ はノルムを表し、 W_i は画像 i の重み行列である．更に γ は画像の先見情報を表す関数であり、今回は画像の滑らかさを拘束とする BTV 法を用い、パラメータ λ により二つの項をバランスする．なお、ここでの集合 V は多視点画像すべての集合となるが、実際の計算では重み行列が 0 となる画像では計算を省略する等の処理を行っている．上式を次式の最急降下法による繰返し計算で求める．

$$I_p^{k+1} = I_p^k + \beta \left\{ \sum_{i \in V} W_i F_i^T H^T (H F_i I_p^k - I_i) + \lambda \gamma(I_p^k) \right\}$$

$$\gamma(I) = \sum_{l=0}^P \sum_{m=0}^P \alpha^{m+l} |I - S_u^l S_v^m I|$$

ここで、 β は最急降下法での漸近係数を表し、各行列の T は逆変換を表す．拘束項は、注目画素との、輝度、距離の違いによりペナルティを与えるものであり、 P は注目するウィンドウのサイズ、 $\alpha (0 < \alpha < 1)$ は距離に応じた重みを表し、 $|\cdot|$ は絶対値を表す． S_u^l, S_v^m は u 軸、 v 軸への平行移動を表し、この場合、画像 x を l pixel だけ u 軸へ移動し、 m pixel v 軸へ移動する．また、繰返し数 $k = 0$ の初期値は、2. の自由視点画像合成結果（線形補間）を用いる．

4. 実験及び結果

実験は、POV-Ray [4] で生成した CG を用いて行い、評価関数には PSNR (Peak Signal to Noise Ratio) を用いた．シーンには、傾いた平面に画像 Mandrill を貼り付けたものの上に、画像 Lenna をテクスチャとしてもつ半球を乗せたものを用いた．カメラアレーは 14×14 台の二次元配置を用い、カメラ間隔は 0.5 とし、最も近い物体までの距離は 20 とした．各カメラは、画角 25 度、解像度 640×480 である．実験では、カメラアレーの中央の座標である、 $(x, y) = (6.5, 6.5)$ に位置する仮想カメラを $z = 0$ から $z = 17.5$ まで前進させて自由視点画像合成を行い、POV-Ray にて生成した理想画像と比較した．画像合成は、最近傍補間、線形補間、超解像で行い、超解像の場合は、16 近傍の光線を用いて行った．また PSF として、 1×1 (何も行わないフィルタ)、 3×3 、 5×5 のボックスフィルタ及び、空間分解能に応じたカーネルの 4 種類を用いた．各パラメータは、 $\beta = 0.2$ 、 $\lambda = 0.1$ 、 $\alpha = 0.6$ 、繰返

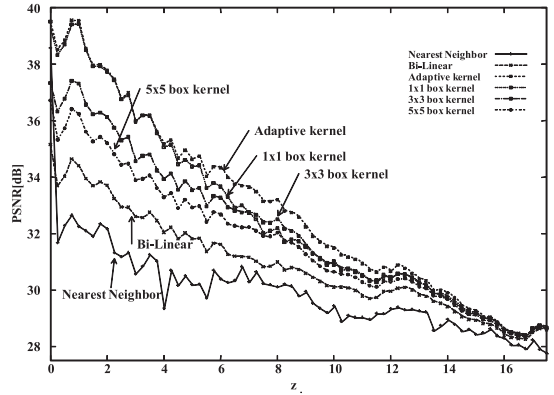


図 4 従来手法と超解像による合成画像の PSNR
Fig.4 PSNR of the conventional and the super-resolution methods.

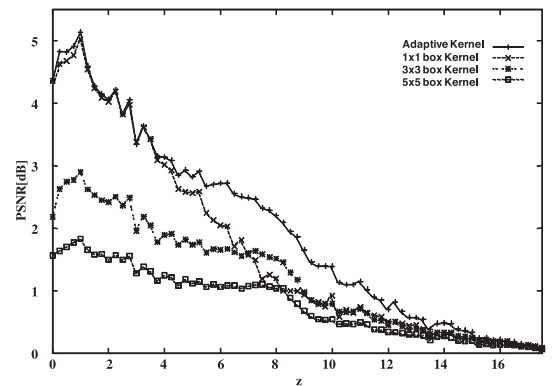


図 5 提案手法の線形補間法からの PSNR の増分
Fig.5 Incremental PSNR of the proposed method from the bi-linear method.

し回数は 20 回である．なお、仮想視点上の奥行画像 z_{map} は既知として CG での計算結果を与えた．

図 4 は各 z での PSNR であり、図 5 は超解像による画像合成法の PSNR 値の線形補間法からの増分を示している．超解像による合成手法が、従来の最近傍法や線形補間法に比べて PSNR 値が高く、線形補間法に比べて最大で 5.1dB 高い．更に、4 種類の PSF の異なる超解像手法を比べると、空間分解能に応じて、適切な PSF を設定した手法が最も高い PSNR 値が得られることが分かる．これは、固定の PSF では、一つの奥行平面しか対応できないため、その平面から物体が離れるほど正しく処理できないためである．また図 6 は、 $z = 6$ のときの結果画像の一部を切り出したものである．主観的に判断すると、エッジが復元され、より高い解像感が得られたことが確認できる．

提案手法は前進するほど効果が小さくなる．図 5



図 6 レンダリング結果
Fig.6 Rendering results.

から、 z が 15 の付近、つまり最も手前にある物体 ($z = 20$ の位置) の拡大率が 4 倍になる付近で、線形補間法と提案法の差がほぼなくなることが分かる。文献 [5] によれば、超解像による分解能の向上は、入力画像の枚数と輝度値の量子化に依存し、 4×4 の近傍光線を用いる今回の場合、理想的には 4 倍の拡大率までしか対応していない。この実験のようにランパート反射するような入力画像からは、より多くの近傍光線を使えば、前進可能な範囲が広がると考えられるが、実際は、光の反射輝度は指向性をもつ。そのため、より広範囲の光線を使うことを考えるのなら、放射輝度の反射特性を考える必要があるだろう。

5. む す び

本論文では、前進位置での自由視点画像を超解像により高画質化し、どのような前進位置でも最も高品質な画像合成が可能であることを確認した。今後は実写画像への適用するため、ノイズにロバストな高解像度化手法及び、本論文では既知とした奥行画像のサブピ

クセル精度での推定法を検討する予定である。

日ごろ、御討論頂く本学菅原真司准教授に深謝する。本研究は、一部、平成 22 年度科学研究費補助金若手研究 (B) (課題番号 22700174) によって行われた。

文 献

- [1] 福嶋慶繁, 圓道知博, 藤井俊彰, 谷本正幸, “Multi-Pass Dynamic Programming による光線空間補間,” 信学論 (D), vol.J90-D, no.7, pp.1721–1725, July 2007.
- [2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super resolution,” IEEE Trans. Image Process., vol.13, no.10, pp.1327–1344, 2004.
- [3] 藤井俊彰, 金子正秀, 原島 博, “光線群による 3 次元空間情報の表現とその応用,” テレビジョン学会誌, vol.50, no.9, pp.1312–1318, 1996.
- [4] <http://www.povray.org/>
- [5] S. Baker and T. Kanade, “Limits on super-resolution and how to break them,” IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.9, pp1167–1183, 2002. (平成 21 年 11 月 30 日受付, 22 年 3 月 31 日再受付)