

研究速報

Depth Image Based Rendering のための奥行画像へのポストフィルタリング

福嶋 慶繁[†] (正員) 石橋 豊[†] (正員:フェロー)

Depth Map Post-Filtering for Depth Image Based Rendering

Norishige FUKUSHIMA[†], Member
and Yutaka ISHIBASHI[†], Fellow

[†]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

あらまし 本論文では, 自由視点画像合成への符号化ひずみの影響を抑えるため, 3種類のポストフィルタを適用する. 結果, 低ビットレートでも画像品質を高く保つことが可能となることを示した. 特に, マルチラテラルフィルタが最も性能が良いことを示した.

キーワード Depth Image Based Rendering, ポストフィルタ, バイラテラルフィルタ, マルチラテラルフィルタ

1. まえがき

近年, 自由視点映像合成の研究が盛んに行われている. 特に, 画像とともにその視点の奥行画像を用いる Depth Image Based Rendering (DIBR) の研究が活発に行われている. これらのデータの通信・記録を行うには符号化は必須である. 通常, 画像符号化による劣化は, 圧縮率にもよるが, 元の画素値に誤差が付加される形となるものの, 元の画素値に近い値をとる. しかし, DIBR で合成画像を生成するときの, 奥行画像の符号化ひずみの影響は, 合成画像の画素の配置がシャッフルされたように出力されるため, 大きな画質の劣化を起こす (詳細は 3. の実験及び結果で述べる). そこで, 本論文では奥行画像の符号化ひずみをポストフィルタで軽減することで, 合成画像品質を高く維持する手法を提案する.

本論文では, 符号化形式として JPEG を用いる. JPEG 画像は, モスキートノイズやブロックノイズ, 階調数の変化・減少 (特に低ビットレート時) によりひずむ. これらのひずみが奥行画像に現れたときの DIBR への影響として, モスキートノイズは奥行境界でスパイクノイズとして現れ, ブロックノイズは輪郭部のブロック状のぼけに, 階調数の変化は画像全体の先鋭さの喪失につながる事が挙げられる. これらのノイズは, 輝度変化が大きいエッジ部分に発生する.

エッジを保持しながらノイズ除去が達成可能なフィルタには, メジアンフィルタ (MF) とバイラテラルフィ

ルタ (BLF [1]) がある. 加えて, バイラテラルフィルタを拡張したマルチラテラルフィルタ (MLF [2]) は, ブロックノイズにより失われたエッジの復元が可能となる. 本論文では上記三つのフィルタを符号化ひずみをもつ奥行画像に適用し, DIBR による自由視点画像合成のためのひずみ除去を行う.

2. 自由視点画像合成と各種フィルタ

2.1 DIBR による自由視点画像合成

DIBR による自由視点画像合成には様々な手法がある [3], [4]. これらの合成の流れを図 1 に示し, 概略を説明する.

(1) 左右の画像・奥行画像が与えられたとき, 画像・奥行画像を左右別々にそれぞれの視点から目的の視点にワーピングする. このとき, 奥行値は量子化されているため, たとえ滑らかに奥行が変化していても, ワーピングにより小さなホールが発生する.

(2) 次に, 奥行画像にメジアンフィルタ, バイラテラルフィルタの順で適用することで, この量子化により生じたホールを埋める. このフィルタでは, オクルージョンなどの大きなホールは変更されない. 以下, 本論文では, この工程の (2) にあたるフィルタを抑制フィルタと呼ぶ.

(3) フィルタによって変更された奥行画像部だけ, (1) とは逆に, ワープ先の画像からワープ元の画像へと画像を逆参照し, 量子化によりホールが生じていた画像部分を求める.

(4) ワーピングされた左右の画像を, 距離に応じた重みでブレンドする. この重みは, 合成視点から, 各々の左右のカメラまでの距離の比で決定し, 距離が近い画像ほど重みが高くなる. ただし, オクルージョンなどによりどちらか一方の画素しか存在しない場合はその画素を優先する.

(5) オクルージョン部分や誤対応により, 求まら

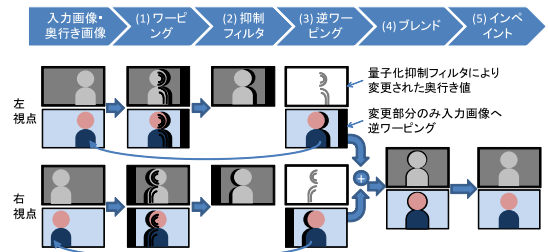


図 1 DIBR による自由視点画像合成の流れ
Fig. 1 Flowchar of free viewpoint image synthesis by using DIBR.

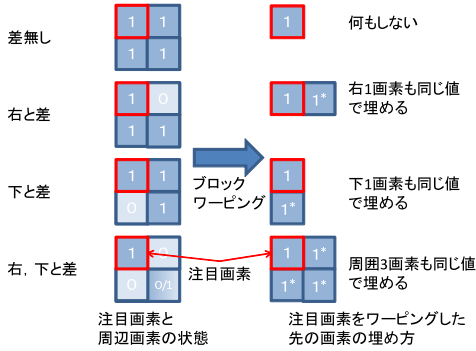


図 2 ブロックワーピング

Fig. 2 Block warping.

なかった部分をインペイントにより、合成画像の周囲から予測することで決定する。

しかし、奥行き画像にノイズがあるときにワーピングを行うと、そのノイズは拡散する。加えて、その拡散したノイズの含まれる奥行き画像に、抑制フィルタを通すと誤差が更に広がる。抑制フィルタは非圧縮時の奥行き画像に適用することを念頭に置いているため、符号化ひずみが大きい場合は効果が望めない。

その他、フィルタを用いない量子化抑制手法としては、(1)のステップで画像を2倍など拡大したのちにワーピングする、オーバーサンプリングによる方法や、ワープする画素をブロックとして数個の画素をまとめてワープさせる手法がある。そこで本論文では、注目画素に3画素の周辺画素含めた 2×2 の画素をブロックごとにワーピングすることで量子化誤差を抑制する手法も比較に加える。ただし、量子化は奥行き値の不連続部で発生することに着目し、周囲の画素との奥行き値の差が1以上のときのみ適用する。図2にワーピングさせるケースを示す。枠内の数字は奥行き値に相当する。まず、周囲画素と注目画素の奥行き値が同一ならば、ブロックではなくピクセルでワープさせる。次に、注目画素と右、下に位置する画素との奥行き値の差が1以上の場合、差のある方向に沿って、注目画素の値を右へ、下へコピーする。また、右、下に位置する画素のうち両方とも1以上差がある場合は、 2×2 のブロック状に値をコピーする。*の記号は、注目画素がコピーされたことを表している。この手法により、奥行き画像をワープさせた例を図3に示す。図3中の右上はブロックを用いずにワープさせたもの、右下がブロックワープをさせたものとなる。奥行き値が0から1に変化する場所で、量子化の影響であるホールが抑制できている

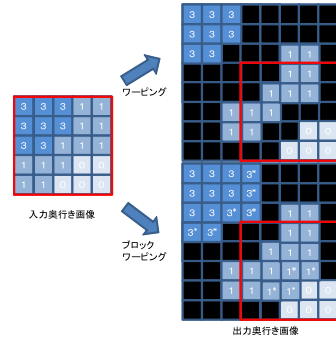


図 3 ブロックワーピングの例

Fig. 3 Example of block warping.



図 4 撮影から自由視点映像出力までの流れ

Fig. 4 Flowchart from capturing to output free viewpoint image.

ことが分かる。また、1から3に変化する場所など大きな変化は、オクルージョンの可能性が高く、この手法ではホールを抑制する必要がない。

本論文では[4]の合成手法を基本に行う。また奥行き画像はSemi-Global Matching [5]により求められた。

2.2 ポストフィルタリング

図4に、撮影から自由視点映像出力までの流れを示す。本論文では、従来の手法とは違い、奥行き画像の伸張後に、奥行き画像へのポストフィルタを行い、図1のDIBRを行っている。本論文ではエッジキープ型のフィルタをポストフィルタとして用いている。

まず、MFはカーネル内の値の中央値を用いる非線形フィルタである。次に、BLFはカーネル中心から離れるほど、またカーネル中心の色から離れるほど重みが低くなる非線形フィルタである。そして、MLFは、BLFに参照画像を加えて重みを計算する方法である。MLFは、奥行き画像をフィルタリング対象とし、奥行き画像の視点に対応する画像を参照画像として扱うと、

$$D_o(\mathbf{p}) = \frac{\sum_{\mathbf{r} \in R} s(\mathbf{p}, \mathbf{r}) d(\mathbf{p}, \mathbf{r}) c(\mathbf{p}, \mathbf{r}) D_i(\mathbf{p} + \mathbf{r})}{\sum_{\mathbf{r} \in R} s(\mathbf{p}, \mathbf{r}) d(\mathbf{p}, \mathbf{r}) c(\mathbf{p}, \mathbf{r})}$$

$s(\mathbf{p}, \mathbf{r}) = \exp(-\frac{\|\mathbf{p} - \mathbf{r}\|_2}{2\sigma_s})$, 位置重み

$d(\mathbf{p}, \mathbf{r}) = \exp(-\frac{\|D_i(\mathbf{p}) - D_i(\mathbf{p} + \mathbf{r})\|_1}{2\sigma_d})$, 奥行き値重み

$c(\mathbf{p}, \mathbf{r}) = \exp(-\frac{\|I_i(\mathbf{p}) - I_i(\mathbf{p} + \mathbf{r})\|_1}{2\sigma_c})$, 参照画素値重み

で表される。ここで、位置 \mathbf{p} の画素値及び入出力の奥行き値は $I(\mathbf{p}), D_{i,o}(\mathbf{p})$ であり、 \mathbf{r} はカーネル内の周辺画素、 R は周辺を表す集合である。 $\|\cdot\|_{1,2}$ はL1, L2

ノルムである。なお、参照画素値重みを一定にした場合、MLFはBLFと同一となる。これらのフィルタをDIBRのプロセスの直前に適用する。BLFとMLFは初めにMFを通した後に適用する。これは、はずれ値が除外され、効果が向上するためである。

3. 実験及び結果

画像解像度が 640×480 、最大視差137画素であるステレオ画像を用いて^(注1)、中間視点の画像合成を行いPSNRを計測した。画像合成には文献[3],[4]の抑制フィルタのオン/オフと3種類のフィルタを組み合わせ6種類の手法を適用した。抑制フィルタがオフの場合はブロックワーピングを代わりに用いている^(注2)。最大視差の範囲には対応点が含まれない可能性があるためPSNR測定領域は水平方向に左右から140画素ずつクリップした 360×480 の領域を設定した。MFでは 3×3 のフィルタを設定し、BLF, MLFはカーネルサイズ 13×13 、 $\sigma_s, \sigma_d, \sigma_c = (9.5, 6.3, 5.4)$ とした。ただしBLFのときは $\sigma_c = \infty$ である。

図5に左視点の奥行画像と合成画像、合成画像の奥行画像を示す。画像合成は抑制フィルタを用いた場合である。ポストフィルタを通さない場合、(a)の奥行画像にはJPEG特有のノイズが目立つ。結果、(b)の奥行画像にスパイクノイズとして現れ、(c)の合成画

像の品質は劣化する。しかし、ポストフィルタを通した場合、(d)の奥行画像には、JPEGノイズが軽減され、(e)の奥行画像にはノイズが少なく、(f)の合成画像の品質が高く保たれていることが分かる。表1に各手法で画像合成した場合のPSNRを示す。抑制フィルタはポストフィルタを通した場合のみ効果があり、PSNRが向上することが分かる。また、MLFが最もPSNRを高く維持できる。図6に仮想視点の奥行画像を、抑制フィルタの有無、ポストフィルタの有無に分けて示す。なお、参考までに、非圧縮時のものも示した。なお、奥行画像のビットレートは奥行画像1枚当たり0.16bppである。図6より、ポストフィルタがない場合、物体の境界部分にスパイクノイズが現れていることが分かる。更にこのノイズ下で抑制フィルタを用いるとノイズが拡大していることが分かる。一方、ポストフィルタを通した場合は、スパイクノイズそのものが少ないので、抑制フィルタによりスパイクノイズが拡大することが少ない。また、文献[6]の奥行画像の一致度を示す尺度である、非圧縮時の仮想視点の奥行画像を基準とした、各手法の合成語の奥行画像が ± 1 階調よりも大きく外れた画素の割合を示す^(注3)。(ポストフィルタ/抑制フィルタ)の組合せが(なし/なし)

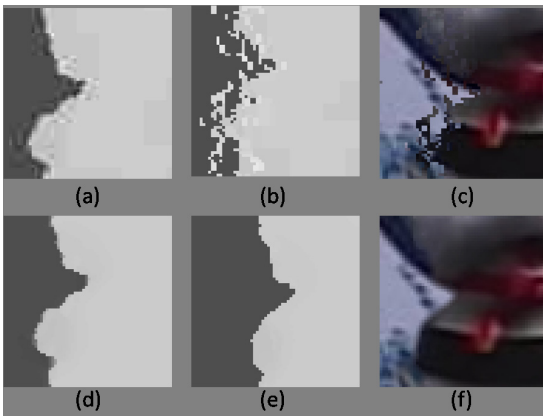


図5 奥行画像と合成画像（符号量は奥行画像1枚当たり0.15bpp 画像1枚当たり0.78bpp）：上段、ポストフィルタなし、下段、ポストフィルタMLF。(a, d) 左画像の奥行画像、(b, e) 合成視点の奥行画像、(c, f) 合成画像

Fig. 5 Depth map and synthesized image (0.15 bpp per a depth map and 0.78 bpp per a image): upper) without post-filter, lower) with MLF. (a, d), Depth map of left view (b, e), Depth map of synthesized view, (c, f) synthesized view.

表1 各手法のPSNR：単位 [dB]
Table 1 PSNR [dB] of each method.

手法	No Filter	MF	BLF	MLF
抑制フィルタ無	28.09	28.65	28.70	29.11
抑制フィルタ有	27.51	28.77	28.89	29.28

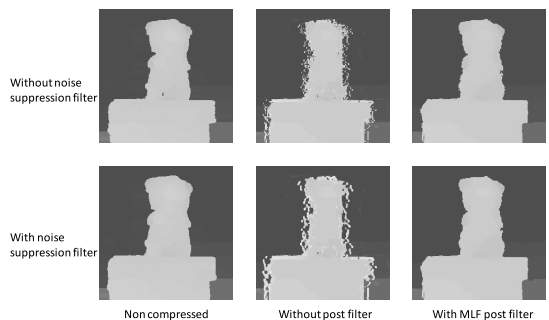


図6 各手法の合成視点での奥行画像
Fig. 6 Depth map of each method on synthesis view.

(注1)：使用したデータセットは <http://nma.web.nitech.ac.jp/fukushima/multiview/2darray.html> よりダウンロード可能である。

(注2)：抑制フィルタもブロックワーピングも用いない場合、量子化によるホールに背景部分に相当する画素が混入し画質が劣化する[3],[4]。

(注3)：文献[6]では非圧縮時の奥行画像ではなく、レンジセンサや手動入力した奥行画像を正解として比較している。

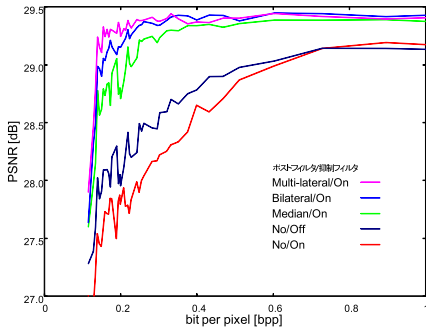


図7 奥行画像1枚当りの符号量と合成画像のPSNR (入力画像の1枚当りの符号量は0.78 bpp)

Fig. 7 Bit per a map versus PSNR (input image bit per one image is 0.78 bpp).

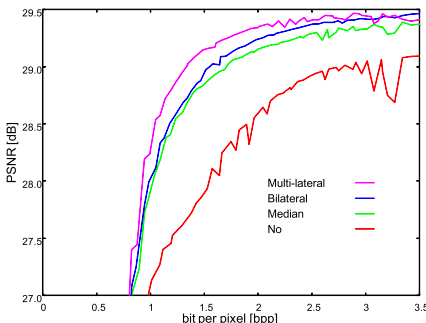


図8 総ビット量(左右の画像と左右の奥行画像:計4枚の画像)と合成画像のPSNR

Fig. 8 Total bit, which includes left-right image and depth map, versus PSNR.

が10.5%, (なし/あり)が12.7%となり, (あり/なし)が5.21%, (あり/あり)が5.16%となり, ポストフィルタがあるときのみ, 非圧縮時の奥行画像に近くなった。つまり, 抑制フィルタが有効に働くときは, ポストフィルタを通したときのみである。

図7に奥行画像へ与える符号量を変化させた場合の合成画像のPSNRを示す。ポストフィルタを行うことで低ビット時でも高く画像品質を維持でき, またMLFが最も効果が高いことが確認できる。高ビット時はオーバーフィッティングとなるため, MLFよりもBLFのほうが効果が高い。また, ポストフィルタがない場合は抑制フィルタがうまく働かないことが分かる。図8に総ビット量(左右の画像と左右の奥行画像:計4枚の画像)を変化させた場合の合成画像のPSNRを示す。なお図8ではすべて抑制フィルタを用いてい

る。ポストフィルタとしてMLFを用いると高い符号化効率を得られたことが確認できる。なお, 図8は左右の奥行画像は同じ品質であり, 左右の画像も同じ品質で圧縮するとして, 画像, 奥行画像の品質を総当りで変化させて, 最大の品質をもつ組合せをプロットした。

4. むすび

DIBRによる自由視点画像合成への符号化ひずみの影響を抑えるため, 3種類のポストフィルタリングを適用した。実験の結果, 提案手法は量子化抑制フィルタによる合成画像品質の劣化の効果をなくすことに成功するとともに, 低ビットレートでも十分な画像品質を保つことが可能となることを示した。特に, 輪郭のブロックノイズの影響まで復元することが可能であるマルチラテラルフィルタの性能が最も性能が良いことを示した。今後の課題として, 画素ごとの適切なパラメータ推定・学習方法の検討や, H.264などその他符号化形式への適用が挙げられる。

謝辞 日頃, 御討論頂く本学菅原真司准教授に感謝する。本研究の一部は, 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] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Proc. ICCV1998, pp.839–846, Aug. 1998.
- [2] P. Lai, D. Tian, and P. Lopez, "Depth map processing with iterative joint multilateral filtering," Proc. PCS2010, pp.9–12, Dec. 2010.
- [3] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV," Signal Process. Image Commun., vol.24, no.1-2, pp.65–72, Jan. 2009.
- [4] S. Zinger, L. Do, and P.H.N. de With, "Free-viewpoint depth image based rendering," J. Visual Communication and Image Representation, vol.21, Issue 5-6, pp.533–541, July 2010.
- [5] H. Hirschmuller, "Stereo processing by semi-global matching and mutual information," IEEE Trans. Pattern Anal. Mach. Intell., vol.30, no.2, pp.328–341, Feb. 2008.
- [6] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," Int. J. Comput. Vis., vol.47, nos.1/2/3, pp.7–42, April-June 2002.

(平成23年2月28日受付, 7月1日再受付)