

研究速報

奥行画像符号化のための重み付きジョイントバイラテラルアップサンプリングによる超解像

松尾 琢也[†] (学生員) 福嶋 慶繁[†] (正員)石橋 豊[†] (正員)

Depth Map Cording via Super Resolution by Using Weighted Joint Bilateral Up-Sampling

Takuya MATSUO[†], Student Member,Norishige FUKUSHIMA[†], and Yutaka ISHIBASHI[†], Members[†]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, Nagoya-shi, 466-8555 Japan

あらまし 本論文では、奥行画像符号化を目的とした奥行画像の超解像法を提案する。本手法は、奥行画像の輪郭が維持できるジョイントバイラテラルアップサンプリング (JBU) の拡張である。実験では、本手法が既存超解像手法や JPEG に対して高い性能を示した。

キーワード 奥行画像, ジョイントバイラテラル, 符号化, 超解像, Depth Image Based Rendering

1. ま え が き

自由視点映像を生成する研究が盛んに行われている。その中でも、Depth Image Based Rendering (DIBR) [1] による手法は、描画時の計算コストが少ないため、自由視点映像を伝送するためのフォーマットとして注目されている。DIBR では、通常の画像に加え、その画像に写る被写体までの距離情報となる奥行画像が必要となる。つまり、画像に加えて奥行画像も符号化しなければならない。H.264/AVC などにより、奥行画像を低ビットレートで符号化すると、自由視点映像の品質が大幅に劣化することが分かっている。文献 [8] では、H.264/AVC で符号化した画像と奥行画像を用いて、自由視点映像合成を行っている。実験では、奥行画像が低ビットレートで符号化されている場合は、画像に多くのピットを割り当てたとしても画像品質が向上しないことを示している。また、文献 [9] では、H.264/AVC の多視点映像用の拡張である MVC (Multiview Video Coding) [10] を用いて複数の奥行画像を符号化した場合、低ビットレート時には自由視点映像の品質が大きく劣化し、逆に一定以上のピットを奥行画像に割り当てたとしても符号化利得がほとんどなくなるが示されている。

筆者らは、この劣化がポストフィルタにより解決することを、JPEG 符号化を例にして示した [2]。こ

で、フィルタにより情報が復元できるということは、サブサンプルした情報を可逆圧縮して送信し、復号側で情報を復元することも可能である。この復元アルゴリズムとして、奥行画像超解像手法 [3]~[6] がある。各手法は、精度は高いが計算時間がかかる [3]、輪郭がぼける [4], [5]、平面部分の滑らかな補間が不可能 [6] などの欠点がある。

本論文では、輪郭がぼけず、それ以外の部分は滑らかな補間が可能となり、計算コストの低い奥行画像の超解像手法を提案する。また、JPEG による符号化よりも、サブサンプルしたものを可逆圧縮し、受信側で超解像した方が、より効率的な符号化方式となることを示す。

2. ポストフィルタリングによる符号化ひずみ除去

ポストフィルタリングによる符号化ひずみ除去手法として文献 [2] を説明する。この文献では、奥行画像の JPEG 符号化によって発生するノイズやひずみが自由視点画像合成の際に、合成画像の品質を著しく低下させていることを指摘している。そして、それらはポストフィルタリングによって除去することができ、品質の低下を抑制できるとしている。流れとしては、奥行画像の圧縮、伝送、伸張、ポストフィルタリング、合成といった流れである。そして、ポストフィルタリングとして、マルチラテラルフィルタ (MLF) を用いている。

3. 奥行画像超解像による圧縮

提案手法では、JPEG 符号化ではなく、奥行画像をサブサンプルと可逆圧縮により圧縮し、超解像により復元することで効率的な伝送を行う。以下の節では超解像手法の従来手法と提案手法を述べる。

3.1 従来の超解像手法

文献 [4] で提案されている JBU を説明する。この手法では、周辺画素までの距離と参照画像の色の類似度を重みとした線形和で画素を補間する。注目画素位置を p 、そのサブサンプルされた画素位置を p' とし、補間に用いる参照画素を q 、 q' とするとき、補間される奥行値 d_p^{jbu} は、

$$d_p^{jbu} = \frac{\sum_{q' \in M} s(p', q') c(I_p, I_q) d_{q'}'}{\sum_{q \in M} s(p', q') c(I_p, I_q)} \quad (1)$$

$$c(x, y) = \exp\left(\frac{|x - y|}{\sigma_c}\right), s(x, y) = \exp\left(\frac{|x - y|}{\sigma_s}\right) \quad (2)$$

である。ここで、 M は参照画素集合、 $c()$ 、 $s()$ 、 σ_c 、

σ_s はガウス重みとパラメータ, I_p, I_q は参照画素値, d'_q は奥行画像の奥行値である. 文献 [6] では, 線形和ではなく, 最も重みが大きくなる値を採用することで, 混色を防ぎ, JBU の輪郭がぼける問題を解決している. その値 d_p^{bnn} は,

$$\begin{aligned} d_p^{bnn} &= d'_v \\ v &= \arg \max_{q' \in N} s(p', q') c(I_p, I_q) \end{aligned} \quad (3)$$

となり, 最も重みが大きくなる画素位置 v 上の, サブサンプルされた奥行画像の奥行値 d'_v の値を採用する. また補間に用いる画素集合 N は, 補間画素位置 p' の最近傍 4 画素である.

3.2 提案する超解像手法

提案手法では, 線形和で画素を混合することで, 平面部分で滑らかな補間ができる. ぼけを防ぐため, 輪郭部分では大きな差がある奥行値同士を混色しないように, 重み w_q による重み付き線形和により奥行値 d_p^{pro} を決定する.

$$d_p^{pro} = \frac{1}{\sum_{q \in N} w_q} \sum_{q \in N} w_q d'_q \quad (4)$$

重み w_q は, 文献 [6] の手法で求めた奥行値と周辺の奥行値 d'_q との差がしきい値より小さければ, 一般的な双線形補間の距離に応じた重み $li(p', q')$ と色重みの積を重みとして与え, それ以外は 0 としてその奥行値を使わないように補間する.

具体的には, まず注目画素位置のサブサンプルされた座標を $p' = (m, n)$ (m, n は小数点座標) としたとき, その最近傍 4 画素のうちの左上の画素は, $q' = (\lfloor m \rfloor, \lfloor n \rfloor) = (u, v)$ ($\lfloor \cdot \rfloor$ は床関数) となり, つまり, 近傍画素の集合は, $N \in (m, n), (m+1, n), (m, n+1), (m+1, n+1)$ となる. この座標の様子を図 1 に示す. この近傍画素から双線形補間する重み $li(p', q')$ は,

$$li(p', q') = \begin{cases} (u+1-m)(v+1-n) & q'=(u, v) \\ (m-u)(v+1-n) & q'=(u+1, v) \\ (u+1-m)(n-v) & q'=(u, v+1) \\ (m-u)(n-v) & q'=(u+1, v+1) \end{cases} \quad (5)$$

となり, 選択される近傍画素に応じて決定される. そして, 線形和の重み w_q は, 画素ごとのしきい値 $th_d(q')$

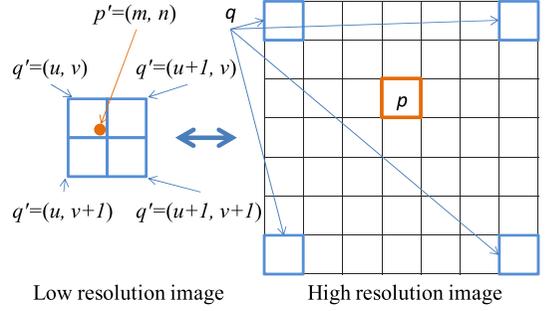


図 1 高解像度画像座標と低解像度画像座標

Fig.1 Coordinate of high resolution image and low resolution image.

に応じて下記で決定される.

$$w_q = \begin{cases} li(p', q') c(I_p, I_q) & |d_p^{bnn} - d'_q| < th_d(q') \\ 0 & \text{else} \end{cases} \quad (6)$$

しきい値 th_d は, 物体境界では小さく, それ以外では大きく設定することで, より輪郭のぼけを抑制しつつ平面部分の滑らかな補間が可能になる. そこで, 縮小された奥行画像でラプラシアンフィルタ $Laplacian()$ によるエッジ検出を行うために, 下記のラプラシアンフィルタカーネルを設定し,

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (7)$$

縮小された奥行画像 D' 上で, その応答値の絶対値をしきい値処理して, エッジだと判定された領域には低いしきい値 th_e , それ以外の平坦な領域だと判定された領域には, 高いしきい値 th_f を設定する.

$$th_d(q') = \begin{cases} th_e & |Laplacian(D'_q)| > th_L \\ th_f & \text{else} \end{cases} \quad (8)$$

このようにして, 重みを適応的に変化させた. 各しきい値は, 経験的に $th_e = 3, th_f = 9, th_L = 6$ と設定した.

この方法では, エッジではないと判定された領域のしきい値が低く, その領域ではずれ値が存在した場合に, 望まない奥行値との線形和になり得る. しかしながら, そのようなはずれ値をもつ領域は, 色情報が異なっている可能性が高いため, そのような場合に,

式 (6) の色重み $c(I_p, I_q)$ が有効に働く. この色重みは d_p^{bmn} での途中計算結果を再利用することで求めた.

4. 実験及び結果

実験では, 奥行画像の復元精度を評価した. また, 奥行画像と画像を圧縮・伸張し, 文献 [1] の DIBR による自由視点画像合成を行った. 実験には, 多視点画像データセット Teddy [7] を用い, 奥行画像には真値を用いた. この真値には欠損があるため, 欠損部は周囲の最小奥行値でパディングした. 奥行超解像の評価では, 文献 [7] のマスク「ALL」を用い, 真値との差がしきい値以上のピクセルが何%あるかで比較した. 符号化評価では, 自由視点画像を, 多視点画像データセットの中心にある原画像の位置に合成し, PSNR で評価した. そのとき, 本手法と文献 [2] による JPEG とポストフィルタによる手法と比較した. サブサンプルされた奥行画像は PNG により可逆圧縮した. JPEG 圧縮の場合は, 画像と奥行画像を総当りで品質を変え圧縮し, 最も効率の良いときを結果とした. また, 原画像の圧縮は全て JPEG で符号化した.

表 1 に, 各手法での奥行画像の復元結果を示す. 表中の文献 [2] の結果に関しては, 提案手法と同サイズまで小さく圧縮できなかったため, データサイズが 3 倍程度でありクオリティ 15 での評価結果である. 誤差が 1 画素以内のときも, サブピクセル精度である 0.5 画素以内のときも提案手法が最もよく復元可能であった. また, 図 2 は, 各手法で (a) の画像, 奥行画像を

用いて 8 倍の解像度に超解像した結果である. 提案手法が平面部分も輪郭部分もよく復元できていることが分かる. 文献 [3] の手法は精度が高いとされているが今回のようなノイズが少ない場合には効果が小さく, また, 奥行画像を超解像するのに 1 枚当たり約 3 分かかった. 文献 [2] の手法では, カーネルサイズ 7 の際にポストフィルタに 44 ms, 符号化に 3 ms の計 47 ms を要した. 提案手法は, 約 1.5 ms で計算可能であり, 計算時間と復元画像の精度の観点から効率的な手法であった. 使用した計算機は Intel Core i7 2.93 GHz である. また, 図 3 に JPEG, 文献 [2], 提案手法の 2, 4, 8, 16 倍で超解像した符号化結果を示す. このグラフは x 軸は自由視点画像合成に用いた入力画像と入力奥行画像の和である総 bit 量である. 文献 [2] の手法に対して, 提案手法で 4 倍に拡大した場合は全てのビットレートで高効率であり, また, 低ビットレート

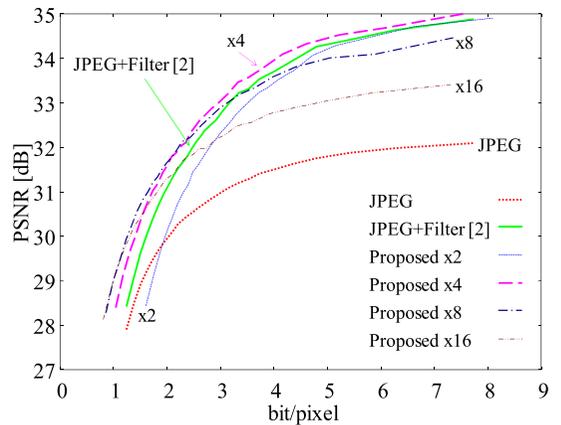


図 3 JPEG と超解像をもとにした圧縮手法の符号化結果
Fig. 3 Coding results of JPEG based and super resolution based compression methods.

表 1 各手法の奥行画像の復元効果 (リサイズ 8 倍)
Table 1 Reconstruction performance of each depth super resolution method (Resize factor is 8).

Method	[3]	[4]	[6]	Proposed	[2]
Th = 0.5 [%]	6.86	7.56	7.26	6.23	34.57
Th = 1.0 [%]	4.56	4.99	5.27	3.98	5.50

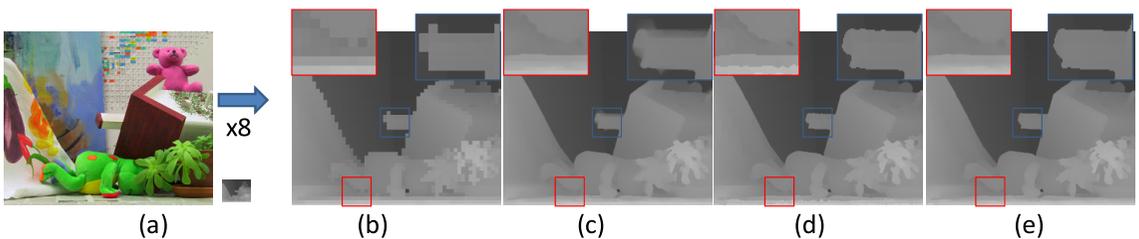


図 2 各奥行超解像手法の結果. (a) 入力画像と奥行画像, (b) NN upsample, (c) JBU, (d) JBF-NN, (e) 提案手法, 緩やか, 急激な奥行変化による輪郭の拡大

Fig. 2 Results of each method of depth super resolution. (a) input image and depth map, (b) NN upsample, (c) JBU, (d) JBF-NN, (e) proposed. Specified areas of slant edges and sharp edges are enlarged.

では 8 倍, 16 倍といった倍率でも有効であった。

5. むすび

本論文では, 奥行画像の超解像を用いることでサブサンプルされた奥行画像を可逆圧縮する方が, JPEG で奥行画像を圧縮するよりも符号化効率が高いことを示した。また提案した超解像手法は高速かつ高精度であることを示した。今後の課題として, より高解像度なデータへの適用やノイズの影響などの調査, 他符号化形式との比較が挙げられる。

謝辞 日ごろ, 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)(22700174), 科学技術振興機構研究成果展開事業研究成果最適展開支援プログラム (A-STEP) 「奥行き画像超解像によるリアルタイム自由視点映像伝送の実現」及びに, 総務省戦略的情報通信研究開発推進制度 (SCOPE) によって行われた。

文 献

- [1] 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.
- [2] 福嶋慶繁, 石橋 豊, "Depth Image Based Rendering のための奥行画像へのポストフィルタリング," *信学論 (D)*, vol.J94-D, no.12, pp.1992–1995, Dec. 2011.
- [3] Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," *Proc. IEEE Comput. Vis. Pattern Recognit.*, pp.1–8, June 2007.
- [4] J. Kopf, M.F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," *ACM Trans. Graphics*, vol.26, no.3, July 2007.
- [5] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, "A noise-aware filter for real-time depth upsampling," *Proc. Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*, 2008.
- [6] M.O. Wildeboer, T. Yendo, M.P. Tehrani, T. Fujii, and M. Tanimoto, "Depth up-sampling for depth coding using view information," *Proc. 3DTV Conference*, pp.1–4, May 2011.
- [7] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV*, vol.47, no.1-3, pp.7–42, May 2002.
- [8] S. Zinger, L. Do, and P.H.N. de With, "Free view-point depth image based rendering," *J. Visual Communication and Image Representation*, vol.21, no.5-6, pp.533–541, July 2010.
- [9] K. Klimaszewski, K. Wegner, and M. Domanski, "Distortions of synthesized views caused by compression of views and depth maps," *Proc. 3DTV Conference*, pp.1–4, May 2009.
- [10] P. Pandit, Y. Chen, and S. Yea, "Text of ISO/IEC 144965:2001/PDAM 15 reference software for multiview video coding," *ISO/IEC JTC1/SC29/WG11 MPEG2008/W9974 Hannover, Germany*, July 2008. (平成 23 年 12 月 15 日受付, 24 年 4 月 13 日再受付)