

修士論文要旨 (2017 年度)

符号化開口を用いたぼけを含むステレオ画像からの視点内挿 View Interpolation from Defocused Stereo Images with Coded Aperture

16N5100043C 生清 雄太

電気電子情報通信工学専攻 久保田研究室

1. 研究目的

ステレオ画像から中間視点画像を生成する視点内挿問題に対しては、ステレオ画像から対応点（視差画像あるいは奥行画像）を推定し、その対応点の画素値を混合する方法が多く検討されている [1]。しかしながら、高精度に対応点を推定することは一般的に難しく、推定誤差より生成画像の品質が低下する問題がある。

この問題に対して、近年、焦点を変えたステレオ画像を用いることによって、対応点の推定精度を向上させる方法 [2] が提案されている。視差だけでなく焦点ぼけの違いも利用できるため、ステレオ法と Depth from defocus (DFD) を組み合わせることによって高精度な対応点推定が可能である。中間視点画像を生成する場合、ステレオ画像に生じた焦点ぼけを復元する必要があるが、ぼけ復元に適した符号化開口が提案された [3]。

これらの方法とは異なり、対象シーンが2層である場合は、対応点を推定することなく、焦点を変えたステレオ画像から視点内挿を行う方法もある [4]。この方法では、左右の各ステレオ画像に対して空間不変フィルタを施すことによって、中間視点画像の再構成が可能である。しかし、シーンの奥行が2層に限定されているため、汎用性に乏しい問題がある。

本論文では、多数の層から構成される一般のシーンに対して、対応点の推定が不要な視点内挿法を提案する。提案手法では、ステレオ画像として、最も手前に焦点を合わせた左視点画像と最も後ろに焦点を合わせた右視点画像を取得する。中間視点画像が多数のレイヤ画像の和で構成されるモデルを導入し、ステレオ画像からレイヤ画像を最小二乗法を用いて反復的に推定する。推定したレイヤ画像を加算して中間視点画像を再構成する。

2. 提案手法

2.1 問題の記述

奥行を Z で表すと、対象シーンの被写体は、 $Z_{\min} \leq Z \leq Z_{\max}$ に存在するとする。平行ステレオカメラにより、左視点において Z_{\min} に焦点を合わせた画像 g^L と右視点において逆に Z_{\max} に焦点を合わせた画像 g^R を取得する。

目的は、基線上の中間視点における完全合焦画像 f （以後、中間視点画像）を生成することである。中間視点の位置は、基線を $\alpha : (1-\alpha)$ に内分した点で表す。ただし、 $0 \leq \alpha \leq 1$ である。

2.2 画像のモデル

対象シーンが N 層の奥行レイヤからなると仮定する。各レイヤの奥行を手前から Z_1, Z_2, \dots, Z_N とし、各レイヤの視差が一定間隔で変化するように設定する。ただし、 $Z_1 \leq Z_{\min}$, $Z_N \geq Z_{\max}$ である。

奥行 Z_n ($n = 1, 2, \dots, N$) のレイヤが所望の中間視点において観測される画像を f_n と表す。これらをレイヤ画像と呼ぶこととする。レイヤ画像は、視点 α に依存した画像であり、未知である。

所望の中間視点画像 f をすべてのレイヤ画像の和

$$f(x, y) = f_1(x, y) + f_2(x, y) \cdots + f_N(x, y) \quad (1)$$

としてモデル化する。 (x, y) は画像座標を表す（以後、同様）。

ステレオ画像は、視点 α に応じて水平移動した後に焦点ぼけが生じたレイヤ画像の和を用いて次式のようにモデル化する。

$$g^L(x, y) = \sum_{n=1}^N h^L(x, y; r_n^L) * f_n(x - d_n^L, y), \quad (2)$$

ただし、 $*$ は 2 次元の畳み込み積分を表す． g^R も同様である．

関数 $h^L(x, y; r_n^L)$ と $h^R(x, y; r_n^R)$ は、左右のステレオ画像の各レイヤ画像の点広がり関数である． r_n^L と r_n^R は各点広がり関数の広がり程度を表す量（ぼけ量）であり、

$$r_n^L = \frac{R}{N-1}(n-1), \quad (3)$$

$$r_n^R = \frac{R}{N-1}(N-n) \quad (4)$$

と表される．ただし、 R は、最大のぼけ量であり、左視点画像の f_N および右視点画像の f_1 に生じるぼけ量を表す．ぼけ量が零の場合、各点広がり関数はデルタ関数となる．

d_n^L と d_n^R は、左右のステレオ画像の各レイヤ画像の水平移動量である．各視差に関しては、視差の幅を $\Delta = d_1 - d_N$ と定義すると

$$d_n^L = \alpha \left[d_N + \frac{\Delta}{N-1}(N-n) \right], \quad (5)$$

$$d_n^R = -(1-\alpha) \left[d_N + \frac{\Delta}{N-1}(N-n) \right] \quad (6)$$

と表される．したがって、 d_1 と d_N が既知であれば、 d_n ($n \neq 1, N$) が決定される．

2.3 反復再構成

実際に取得したステレオ画像とモデル（式 (2) の右辺）との二乗誤差が最小となるように反復的に各レイヤ画像を求め、その結果を式 (1) の右辺に代入して中間視点画像を再構成する．

まず、最小化すべき目的関数を次式で定める．

$$J(f_1, \dots, f_N) = \frac{1}{2} \iint [e^L(x, y)]^2 + [e^R(x, y)]^2 dx dy \quad (7)$$

ここに、 $e^L(x, y)$ は取得した左視点画像と式 (2) のモデルとの誤差を表す． $e^R(x, y)$ も同様とする．

次に、最急降下法を用いて、レイヤ画像 f_n の k 回目の反復解 $f_n^{(k)}$ を次式で求める．

$$f_n^{(k)} = f_n^{(k-1)} - \rho \frac{\partial J}{\partial f_n}(f_1^{(k-1)}, \dots, f_N^{(k-1)}) \quad (8)$$

ここに、 ρ はステップパラメータである．各レイヤ画像の初期解は $f_n^{(0)}(x, y) = 0$ とした．

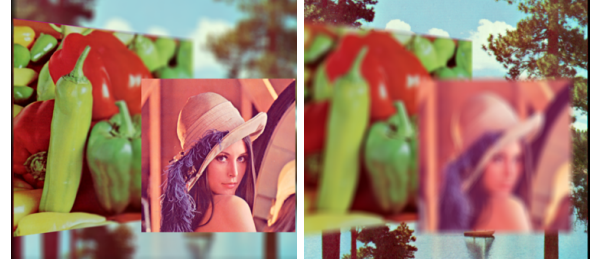


図 1: Zhou コードを用いたステレオ画像の例 ($\Delta = 5.48, R = 5.48$)

最後に、 k 回反復後の中間視点画像 $f^{(k)}$ を各レイヤ画像の反復解 $f_n^{(k)}$ の和

$$f^{(k)} = f_1^{(k)} + f_2^{(k)} + \dots + f_N^{(k)} \quad (9)$$

として求める．

3. シミュレーション実験

3.1 ステレオ画像の準備

シミュレーション実験に用いるステレオ画像は、CG ソフト（POV-ray）で作成した多視点画像から合成開口法により合成した． $Z=300$ に画像 Lena と画像 Peppers、 $Z=400$ に画像 Sailboat を配置し、滑らかな奥行きを表現するために、Peppers を垂直軸を中心に 30° 回転させて対象シーンを構成した．

ステレオ画像の例として、視差の幅が $\Delta = 5.48$ [pixels]、最大ぼけ量が $R = 5.48$ [pixels] のときの画像を図 1 に示す．開口には、ガウス関数と Zhou らが提案した符号化開口（以後、Zhou コード）[3] の二種類を用いた．図 1 にそれぞれのステレオ画像を示している．左視点画像では近景の Lena に焦点が合っており、右視点画像では遠景の Sailboat に焦点が合っていることがわかる．左右で視差の変化があることも確認できる．

Zhou コードについては、取得画像に加算される雑音の電力によって復元に最適なパターンが提案されている（図 2）．本シミュレーションでは、最も結果が良かった $\sigma' = 0.0001$ のパターンを採用した．なお、ガウス関数のぼけ量には、その標準偏差を用い、Zhou コードは多視点画像の採用視点数がガウス関数とほぼ同じになるサイズに縮小したものを用いる．

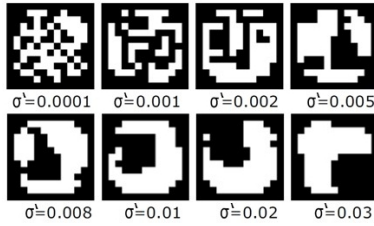


図 2: Zhou らの提案した符号化開口パターン [3]. σ' は取得画像の雑音電力の正の平方根を表す.

3.2 再構成画像の結果

中心視点 ($\alpha = 0.5$) における中間視点画像の再構成を行った. 反復は, 反復解の差の二乗和が 0.001 以下になるまで繰り返した. レイヤは, 280 から 420 の間に 12 層を設定し, シーン全体を覆うように配置した.

$\Delta = 5.48$ [pixels], $R = 5.48$ [pixels] の場合において, 開口にガウス関数と Zhou コードを用いたときの再構成画像をそれぞれ図 3 (a) と (b) に示す. 中間に配置した Peppers の焦点ぼけの復元が十分ではないが, 近景と遠景については鮮鋭に復元ができています. ガウス関数の結果と比べると, Zhou コードの結果の方がやや鮮鋭であることがわかる. PSNR (Peak Signal to Noise Ratio) は, ガウス関数の場合は 28.3 [dB] であったが, Zhou コードの場合は 31.0 [dB] であった.

Zhou コードを用いた場合において, 最大ぼけ量 R を 1.84 [pixels] と 5.48 [pixels] に変えたときの再構成画像を比較した. 図 4 (a) と (b) にそれぞれの結果を示す. ただし, $\Delta = 10.96$ [pixels] (前述の 2 倍) に設定した. ぼけ量が小さい $R = 1.84$ の結果では, 視差が正確に付与されず, 品質が低下している. PSNR は 24.3 [dB] であった. これに対して, ぼけ量が多い $R = 5.48$ の結果では, 品質が改善されていることがわかる. PSNR は 3.0 [dB] 改善し, 27.3 [dB] であった. この結果は, 焦点ぼけ (開口) を大きくすることによって, 再構成の精度が向上することを示している.

ガウス関数と Zhou コードを用いた場合において, Δ と R を変化させたときの再構成画像の PSNR を図 5 (a) と (b) にそれぞれ示す. どちらの場合も Δ が大きくなると PSNR は低下するが, Zhou コードの場合は, ぼけ量



(a) ガウス関数 (b) Zhou コード

図 3: 再構成した中間視点画像 ($\Delta = 5.48, R = 5.48$)



(a) $R = 1.84$ (b) $R = 5.48$

図 4: 最大ぼけ量 R を変えたときの再構成画像の比較 ($\alpha = 0.5, \Delta = 10.96$)

を増やすことで PSNR の低下を抑えることができています. Zhou コードがぼけの復元に最適に設計されているためであると考えられる.

4. 実写画像を用いた実験

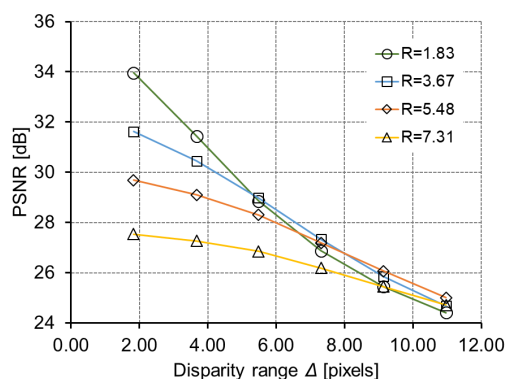
4.1 ステレオ画像の準備

1 台のデジタルカメラ (Nikon D7200) を水平方向に 10 mm 移動させてステレオ画像を撮影した. レンズには SIGMA 35mm F1.4 DG HSM を用い, 開口部 (後玉) に普通紙に印刷した Zhou コードを貼り付けた. 最も結果の良かった図 2 の $\sigma' = 0.001$ のパターンを採用した.

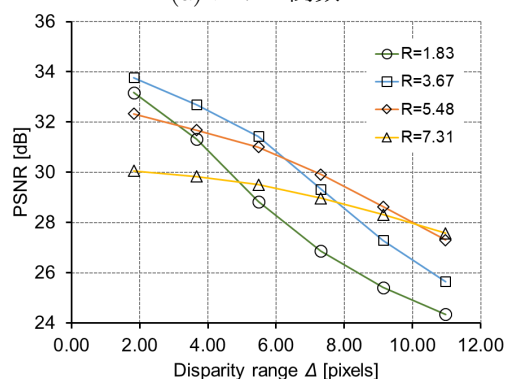
対象シーンは, 手前から順に, クマのぬいぐるみ, ペン入れと花, 本からなる. それぞれの奥行は, 1000 [mm], 1100 [mm], 1500 [mm] である.

撮影したステレオ画像を図 6 に示す. 焦点を合わせた奥行は, 左視点画像で 1000 [mm], 右視点画像で 1500 [mm] である. また, 文献 [4] の方法に基づいて, 左視点画像は 0.9919 倍に縮小して焦点による像倍率を補正した. 本実験の条件では, 視差の幅 Δ は 7.45 [pixels] となった.

左右の点広がり関数は, 点光源 (光を通した光ファイ



(a) ガウス関数



(b) Zhou コード

図 5: 視差の幅 Δ と最大ぼけ量 R を変えたときの再構成画像の PSNR



左視点画像 g^L 右視点画像 g^R

図 6: 取得したステレオ画像

バ) を撮影することによって取得した。他のレイヤ画像における点広がり関数は、これを縮小することによって作成した。右視点画像の点広がり関数も同様にして取得した。最大ぼけ量は、8.0 [pixels] となった。

4.2 再構成画像の結果

中心視点 ($\alpha = 0.5$) における中間視点画像を提案手法により再構成した。反復は、反復解の差の二乗和が 0.001 以下になる 24 回まで繰り返した。レイヤは、1000 [mm] から 1500 [mm] の間に 12 層を設定した。

再構成画像を図 7 に示す。シーン全体が鮮鋭であり、ぼけが復元されていることがわかる。ただし、強いエッ



図 7: 中心視点で再構成した中間視点画像

ジの周辺では、焦点ぼけの輪郭部が除去できずに残っている。これは、点広がり関数の推定誤差が原因であると考えられる。また、背景が前景に透けて見えている領域があるが、これは、モデル式において隠れを考慮していないことによる。

5. むすび

本論文では、焦点を変えたステレオ画像から反復的に中間視点画像を再構成する方法を提案した。提案手法は、対応点の推定を行うことなく、各奥行に領域に対して視差の付与とぼけの除去が最小二乗誤差の範囲で可能である。

参考文献

- [1] S.C. Chan, H.Y. Shum, and K.T. Ng, "Image-based rendering and synthesis," IEEE Signal Process. Mag., vol.24, no.7, pp.22–33, 2007.
- [2] C. Frese and I. Gheeta, "Robust Depth Estimation by Fusion of Stereo and Focus Series Acquired with a Camera Array," IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems, pp.243–248, 2006.
- [3] C.Zhou and S.Nayar, "What are Good Apertures for Defocus Deblurring?" Proc. Int. Conf. on Computational Photography, pp.1–8, 2009.
- [4] 久保田 彰, 児玉 和也, 羽鳥 好律, "焦点の異なるステレオ画像からの線形フィルタによる視点内挿法," 信学論, Vol.J98-D, No.10, pp.1325–1334, 2015.