

合焦判定を利用した中心視点画像の生成

Generating center-view image by using Depth-from-focus

大澤 克俊†
Katsutoshi Oosawa

久保田 彰†
Akira Kubota

羽鳥 好律†
Yoshinori Hatori

1. まえがき

現在、インターネット、デジタル放送の発展から双方技術の需要が高まっている。そのような技術の1つとして自由視点画像がある。自由視点画像とは画像を受け取る側(クライアント)が任意に視点を選択可能な画像であり、主にスポーツ中継や、無形文化物の保存などへの応用が期待できる。しかし、ユーザが求める視点全てにカメラを配置することは現実的に不可能なので、仮想視点の生成を行い、実視点の画像からの補間することで自由視点画像を生成する手法が一般的である。

多数の画像を元に自由視点画像を合成する手法の1つとしてLight field rendering(LFR)[1]がある。この手法は空間全体を様な平面(focal plane)と仮定して扱うので、物体の三次元構造を用いることなく、簡単な処理で対応点の計算が可能であるが、focal plane から離れた物体の場合、対応点の不一致が生じ、画像に劣化が生じてしまう。この問題を解決するための手法としてステレオマッチングとdepth-from-focusがある。ステレオマッチングは各画素の設定した仮想平面における対応点を計算し、マッチングの度合いが一番高いものを正解の奥行きとする方法である。この手法は全ての画素、全ての奥行きとの対応点を計算するため、検索が面倒な部分がある。depth-from-focusはLFRによる画像合成を複数のfocal planeにおいて行い、それにより得られた画像を画素ごとに最も焦点が合っている画像を合焦判定(焦点があっているかどうかの判定)により選択し、その画像を読み出すことによって全体的に鮮鋭な画像を求める方法である[2]。この手法の利点は合焦判定に対応点の計算が必要ないことである。しかし、LFRによる合成画像の場合、一般の画像で用いられる、高周波成分を利用した合焦判定[3]が機能しない(figure1)。これは、合成画像のぼけは高周波成分を持つ二重像(ゴースト)を含んでいるからである。よって、本稿ではLFRの中間視点画像生成に特化し、中間視点画像におけるLFRのぼけが正しい奥行きの前後において同じになること(ぼけの対称性)を利用した新しい合焦判定法を提案し、その判定法を用いた全焦点画像の品質の定量的評価を行った。

2. 提案手法

2.1 合焦判定の原理

LFRで合成した画像群のある画素におけるぼけの様子をfigure2に示す。ここで、 z_l はfocal planeの位置であり、物体が z_l の物体の奥行きにあるときを示している。再構成画像のぼけ方が、中間視点画像の場合、正しい奥行きの前後で同じ劣化を示す。よって、ある奥行きに対してその前後の奥行きの画像を比較し、その差分値を評価する。提案手



Figure 1 the synthesized image by using general focus measure

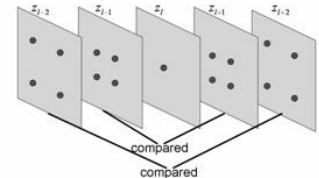


Figure 2 Shift-variant LFR images

法では差分値をグラフにあらわしたときに最も前後の値と変化の大きい極小の値を持つ奥行きが合焦していると判断する。ここで、最小値でなく極小値を用いる理由は、正しい奥行きから大きく離れた奥行きでは再構成画像の変化が極小であるため、正解付近の奥行きにおける差分より小さくなるためである。

2.2 全焦点画像合成手順

1. まず、あらかじめ候補となるfocal planeの奥行き z_n ($n=0,1,\dots,N-1$)を設定する。 z_n は、視差空間(奥行きの逆数)を等分割にするように、次式で設定する。このように配置することにより、中間視点画像において、前述したぼけの対称性を満たすことが確認されている。

$$\frac{1}{z_n} = \frac{1}{z_{\max}} + \left(\frac{1}{z_{\min}} - \frac{1}{z_{\max}} \right) \times \frac{n}{N-1} \quad (1)$$

z_{\max}, z_{\min} は、配置するfocal planeの奥行きの最大値と最小値である。

2. それぞれにおいてLFRによる中心からの画像合成を行う。ここで、 z_n において合成された画像を $G_n(x, y)$ とあらわすことにする。 (x, y) は画素の位置を表す。

3. 評価関数を式(2)に示す。ここで、 W は比較するブロックである。 z_n の前後の s 番目の奥行きまでの合成画像の差分を取り、その結果を足し合わせる。

$$E_0^{(n)}(x, y) = \sum_{i=1}^s \sum_{(x,y) \in W} |G_{n-i}(x, y) - G_{n+i}(x, y)| \quad (2)$$

4. 式(2)を n の関数とみたとき、極小になるような n' を求める。そのような n' の中で、最もその前後の値との変化が大きい奥行き $n_{x,y}$ を求めることにより、最も焦点があう奥行きとする。

$$E^{(n)}(x, y) = E_0^{(n-1)}(x, y) + E_0^{(n+1)}(x, y) - 2E_0^{(n)}(x, y) \quad (3)$$

$$n_{x,y} = \arg \max_{n'} \{E^{(n)}(x, y)\} \quad (4)$$

5. 下記の式に従い、全焦点画像 $G(x, y)$ を得る。

$$G(x, y) = G_{n_{x,y}}(x, y) \quad (5)$$

† 東京工業大学大学院総合理工学研究科
物理情報システム専攻羽鳥研究室

3. シミュレーション

3.1 実験条件

実験の設定を figure3 に示す。ここで、本実験は実際の画像ではなく、CG で作った画像を用いており、対象シーンの存在する奥行き (カメラからの距離) は 9~20 とした。Figure3 の位置関係において、カメラ間隔 2 で正方形に配置された 4 つのカメラの入力画像から得る。各入力画像は水平画角 106° 、 320×240 画素とする。配置する仮想平面は $N = 20$ とし、比較するブロック (W) のサイズを $s = 1, 2, 3, 4$ と変化させ、PSNR による画像品質の定量的評価を行った。ここで、focal plane の奥行きは式(1)を用いて設定し、 $Z_{\min} = 7$ 、 $Z_{\max} = 70$ とした。配置する focal plane を実際のシーンの奥行きより広くとった理由は、前後の奥行きの差分によって評価をするため、実際のシーンより広い範囲の画像が必要になるためである。

3.2 比較する枚数、ブロックサイズによる影響

比較する focal plane の数(s)と、ブロックサイズ(BS)、合成画像の PSNR の関係を表した図を figure4 に示す。また再構成画像を figure6(a)~(d)に示す。BS、 s が大きいほど、PSNR の値は大きくなった。これは s, BS を大きくすると、参照箇所が増え、オクルージョン (対応点が存在しない部分) を含む割合が減少し、影響が低下することによって、画質が向上していると思われる。しかし、欠点として、 s を大きくすると、対象の奥行き前後に必要な画像の数が増えるため、候補となる奥行きの数が減少し、実際より広い範囲の奥行きをとる必要性が生じる。また BS を大きくすると、figure6(g)(h)のように、全体的に実際のシーンの奥行きよりも広がった奥行きを示した。しかし、PSNR は BS が大きいほど良い数値を示した。これは奥行きの広がった部分はそもそもオクルージョン部分でありどんな条件下でも正しく再構成できないため、画質に大きな影響を与えていないものと思われる。

3.3 極小値を用いる正当性

合焦判定に評価関数の極小値を用いる正当性を示すため、提案手法と評価関数 $E_0^{(n)}(x, y)$ の最小値をとる n を最も焦点のあう奥行きとする手法と比較実験を行った。後者の手法による再構成画像を figure6(e)に示す。ここで BS は 5×5 、 $s = 4$ とした。Figure6(b)と(e)を比較すると、最小値を用いた画像は提案手法に比べエッジ部分に大きな劣化が生じており、提案手法の方が視覚的にも鮮鋭で PSNR 値も高かった。エッジ部分の画素における $E_0^{(n)}(x, y)$ のグラフを figure5 に示す。この画素の正解の奥行きを検証したところ、 $n = 8$ であった。しかし最小値を用いた方法では、奥行き $n = 17$ が最小値なので、誤った部分を選択してしまっており、それが劣化の原因になっている。これは、正解の奥行きから大きく離れてしまうと、大きくぼけたもの同士を比較することになり、ぼけたもの同士の比較は、一般的に差分は小さくなる傾向があるため、正解の奥行きの前後の差分より小さくなってしまふことが原因である。それに比べ、式(3)を利用して評価した提案手法は、極小値で前後と最も値の変化が大きい奥行きが選択されるので、正解の奥行きが選ばれ、エッジ部分の劣化が改善されている。

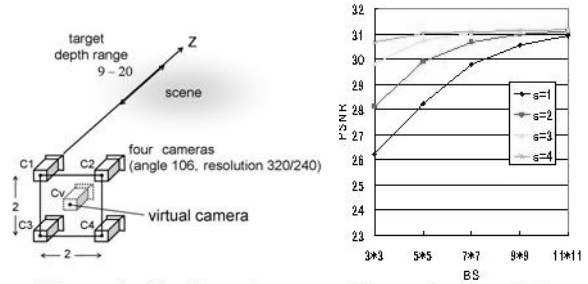


Figure 3 Configuration of the experience

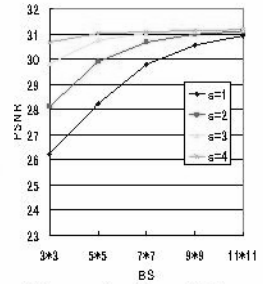


Figure 4 the relation of PSNR, s and BS

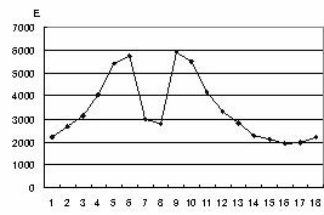


Figure 5 the evaluation function value of a certain pixel

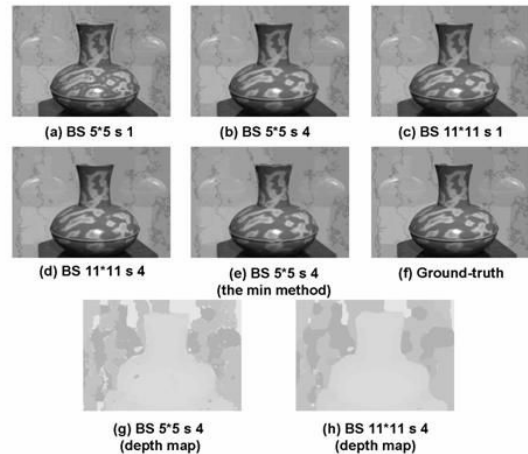


Figure 6 simulation results

4. 終わりに

本稿では、多数の focal plane の奥行きにおいて LFR によって合成した画像を用いて、ぼけ方の対称性を利用し全焦点画像を生成する方法を提案し、その有効性を示した。しかし、このぼけの対称性は、中心視点画像のみに成立する性質で、任意視点画像に応用できないことが現状の課題である。これに対し、参照するブロックの位置を変えることで、任意視点画像に対応する方法を現在検討中である。

5. 参考文献

[1] M.Levoy and P.Hanrahan. "Light Field Rendering", Proc.ACM SIGGRAPH, pp 31-42. 1996.
 [2] 高橋 桂太 久保田 彰 苗村 健 "Light field Rendering における合焦判定と全焦点画像合成", 電子情報通信学会誌 D-II Vol.J88-D-II No.3 pp.573-584,2005
 [3] S.K. Nayer and Y. Nakagawa. "Shape from Focus", IEEE Trans. PAMI, 16, 8, pp. 824-831, 1994.