

合焦判定に基づく全焦点 Light Field Viewer

All in-Focus Light Field Viewer Based on Focus Measurement

高橋 桂太[†]
Keita Takahashi

苗村 健[†]
Takeshi Naemura

1. まえがき

あらかじめ取得された多視点画像を利用して、自由な視点位置から見た画像を合成する手法は、イメージベースレンダリングと呼ばれる。本稿で着目する light field rendering [1] は、その最も基本的な手法のひとつである。この手法の特徴は、被写体の幾何学的な構造情報を必要とせず、シーン全体を 1 枚の平面 (focal plane) で近似して扱う点にある。そのため、手間がかかる構造推定処理が不要であり、また描画処理が簡単で実時間化が可能であるという利点がある。しかし、鮮明に合成できるのは、focal plane 付近の物体だけで、focal plane から離れた物体には、ぼけや二重像が発生してしまうという問題がある。このような現象は、light field rendering における“焦点”の効果 [2, 3, 4] と呼ばれるものである。

筆者らは、上記の問題を解決するため、インタラクティブに自由視点かつ“全焦点”の画像を合成する light field 描画システム (light field viewer) について検討を進めてきた [5, 6, 7]。提案システムでは、まず、与えられた視点において、focal plane の位置 (“合焦”距離) を変えながら画像合成を行うことにより、複数枚の“多焦点”画像を得る。その後、それら“多焦点”画像の“焦点”が合っている領域を、リアルタイムで最終的な 1 枚の画像に統合する。提案システムの特徴は、被写体の明示的な幾何形状を利用するのではなく、与えられた視点位置から見た画像における各画素の奥行きを、フレームごとに推定するところにある。本稿では、最終的な合成画像の品質を向上させるために、推定された奥行きを、信頼度を用いて平滑化する手法について報告する。

2. 研究の背景

図 1 を用いて、light field rendering のアルゴリズムを示す。入力多視点画像は、等間隔に配置されたカメラによって取得される。本来カメラは 2 次元アレイ上に配置されるが、図では簡単のため垂直方向の配列のみを示した。画像合成は、レンダリングカメラを任意の位置に置き、その位置に到達する個々の光線の色を求めることによって実現される。このとき、入力カメラアレイのカメラ間を通過する光線について、適切な補間処理を行うことが必要である。

例えば、図 1 に示した光線 r_d を求める場合を考える。まず、focal plane を定義し、すべての光線がこの面上から発していると仮定する。これは、対象空間全体の奥行きを、focal plane によって近似していることと等価である。次に、所望の光線と、入力カメラアレイ面との交点を求める。所望の光線は、その交点の近傍のカメラ Cam_1 , Cam_2 を参照して、focal plane 上の同一の点から発している光線 r_1 , r_2 の重み付け和として求められる。

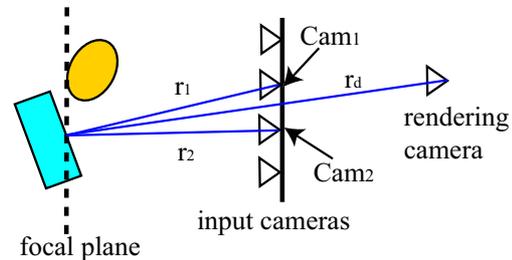


図 1: Light field rendering の原理

被写体が実際に focal plane に近い位置にある場合、その物体は上記の手法によって鮮明に合成される。しかし、focal plane から離れた位置にある物体には、合成の際にぼけや二重像が発生してしまう。これが、light field rendering における“焦点ぼけ”である (図 3 参照)。

この“焦点”が合う奥行き範囲は、多視点画像の取得密度に依存する [8, 9]。入力カメラの間隔が密であるほど、その奥行き範囲は広がる。しかし、多くの場合、この取得密度が不十分であり、単一の focal plane では、シーン全体を鮮明に描画するのは不可能である。Isaksen らは、この問題を解決するため、複数の focal plane を用いる手法 [2] を提案した。これは、“合焦”距離が異なる複数枚の画像を合成し、それらの画像の“焦点”が合っている領域を統合することで、“全焦点”画像を合成する、というものである。

しかし、問題となっていたのは、“焦点”が合っている領域を、どのように実時間で安定に検出するか、という点である。light field rendering で合成された画像に対しては、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 [10] が機能しない。なぜなら、合成画像における“焦点”ぼけは、文字通りのぼけだけではなく、高周波成分を持つ二重像 (ゴースト) を含むためである。そのため、筆者らは、light field rendering に特化した新しい“合焦”判定法を提案し [5]、その判定法を用いた“全焦点”画像の品質の定量的な評価 [6] や、インタラクティブな描画システムの実装 [7] を行ってきた。

3. アルゴリズム

3.1 合焦判定の原理

提案システムでは、文献 [5] に基づき、異なる方法で合成された画像間の差分を用いて“合焦”判定を行う。画像合成の際の補間処理は、周波数領域では、再構成フィルタとしてモデル化されるが [8]、本稿ではその議論を割愛し、空間領域でより直感的に説明する。

本システムで用いる 2 つの合成手法を図 2 に示す。一方では、すべての入力多視点画像を用いて合成を行う (モード A)。他方は入力多視点画像を 1 つおきに間引いたものを用いる (モード B)。双方のモードで、視点

[†]東京大学大学院情報理工学系研究科,
School of Inform. Science & Technology, the Univ. of Tokyo

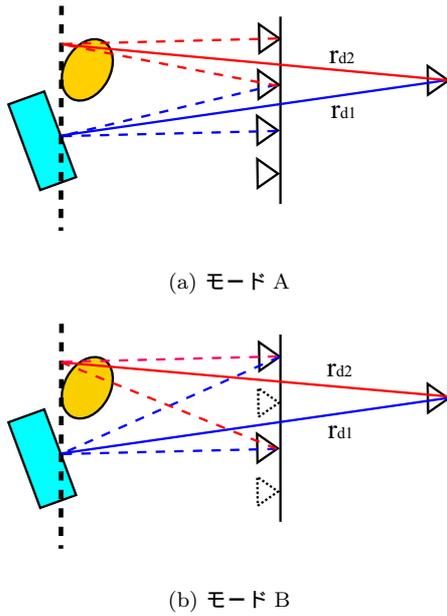


図 2: 2つのモードによる画像合成法

位置および focal plane の奥行き (“合焦” 距離) は同一であるとする。

ここでは、物体の反射特性がランバート (ある一点の色は見る方向に依存しない) であり、かつオクルージョンの問題が無視できると仮定する。物体の表面が focal plane 上にある場合、どちらのモードでも、合成される光線の色は同じである (例として r_{d1})。これは、どちらのモードでも、同一の点から発する光線が合成に使用され、ランバート仮定により、それらの光線の色は同じだからである。しかし、物体が focal plane 上にない場合、両者において合成される光線 (r_{d2}) の色には違いが生じ得る。

そのため、双方のモードによる合成画像を比較すると、“焦点が合っている” 領域はほとんど同じように合成されているのに対して、“焦点が合っていない” 領域は差異が生じ得ることになる (図 3 参照)。したがって、これら 2 枚の画像の差分値を評価することによって、“焦点” が合っている領域を検出することができる。

3.2 全焦点画像合成

上記の “合焦” 判定を用いて、与えられた視点位置に対して、“全焦点” 画像を生成する手法を示す。focal plane を配置する N 個の奥行き候補 $z_n (n = 0, 1, \dots, N - 1)$ は、次の式にしたがってあらかじめ決めておく。

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

Z_{\max}, Z_{\min} は、対象とするシーンの奥行き (入力カメラシステムからの距離) の最大値および最小値である。

1. “多焦点” 画像の合成:

与えられた視点位置で、2つのモードで “多焦点” 画像を合成する。“合焦” 距離 z_n 、モード A で合成された画像を $I_A^{(n)}(x, y)$ と表記する。



図 3: Light field rendering による合成画像。これらの合成画像において視点位置は同じであり、“合焦” 距離を近く (上) から遠く (下) へ変化させている。左段がモード A、右段がモード B による合成である。

2. “合焦” 判定:

まず、同一の “合焦” 距離、異なるモードで合成された画像間の差分を取る。

$$Sub^{(n)}(x, y) = |I_A^{(n)}(x, y) - I_B^{(n)}(x, y)| \quad (2)$$

差分画像を、ブロック平滑化する (M は正の整数)。

$$F^{(n)}(x, y) = \sum_{-M \leq i, j \leq M} \frac{Sub^{(n)}(x+i, y+j)}{(2M+1)^2} \quad (3)$$

各画素に対して、最適な奥行き $n(x, y)$ を次のように定める。

$$n(x, y) = \arg \min_{0 \leq n \leq N-1} F^{(n)}(x, y) \quad (4)$$

$n(x, y)$ は、与えられた視点位置から見た奥行き画像になっている。

3. “全焦点” 画像合成:

すべての点 (x, y) について、最も “焦点” が合っ

いる画像 $I_A^{(n(x,y))}(x,y)$ から画素値を読み出すことにより, 全“焦点”画像 $I(x,y)$ を得る.

$$I(x,y) = I_A^{(n(x,y))}(x,y) \quad (5)$$

3.3 奥行き画像の平滑化

(4) 式で与えられる奥行き画像には, しばしば, 推定誤差によって, 奥行きが周囲と不連続になる部分が生じる. 特に, 近くの被写体と遠くの被写体が重なり合っている場合に, その境界付近にノイズが生じやすいという現象が観察される[‡]. これらの部分は, (5) 式によって得られる最終的な合成画像では, 「ごみ」が浮いているように見える. 特に, インタラクティブに視点位置を動かしたときに, これらの「ごみ」がちらつくように見えるため, 視覚的な品質が損なわれる.

そのため, 実時間性を損なわない範囲で, 奥行き画像に対する平滑化を行うことにする. このとき, ノイズ成分は抑制されることが望ましいが, 奥行きの細かな変化も保持したい. これらの条件を満たす平滑化手法として, 筆者らは次のような手順を採用する.

まず, 各点に対する“合焦”尺度 F の最小値を, その点の奥行きの信頼度の評価値 $R(x,y)$ と定義する.

$$R(x,y) = \min_{0 \leq n \leq N-1} F^{(n)}(x,y) \quad (6)$$

$R(x,y)$ が小さいほど, その点の奥行き情報の信頼性は高いと考える.

次に, (4) 式で得られた $n(x,y)$ を単純にブロック平滑化したものを $n'(x,y)$ と表す. すなわち, W を正の整数として

$$n'(x,y) = \sum_{-W \leq i,j \leq W} \frac{n(x+i,y+j)}{(2W+1)^2}. \quad (7)$$

平滑化の処理は, (8) 式にしたがって行う. R_{th} は信頼度評価値に対する閾値である. 各画素について, 信頼度評価が閾値以上の場合, 奥行きの値を, ブロック平滑化された値で置き換える.

$$n(x,y) = \begin{cases} n(x,y) & R(x,y) \leq R_{th} \\ n'(x,y) & \text{otherwise} \end{cases} \quad (8)$$

上記の処理では, “信頼度” の高い画素は, 平滑化の影響を受けないため, 奥行きの細かい変化が保存される.

4. 実験

4.1 実験の概要

実験環境は PC (CPU: Pentium 4 3.2 GHz, Memory: 2.0 GB, GPU: GeforceFX 5800, VRAM: 128MB) である. ソフトウェアはすべて, C 言語と OpenGL で記述し, gcc でコンパイルした. 画像を合成する際の視点位置や視線方向は, マウスによってインタラクティブに操作できるようにした.

[‡]筆者らの“合焦”判定法は, オクルージョンの影響が無視できることを前提にしているが, 物体同士が重なり合う境界の部分では, その前提が成り立たない.



図 4: 提案手法による“全焦点”自由視点画像合成. (左): 図 3 と同一の視点位置, (右): 異なる視点位置で.



図 5: 提案手法による合成結果. 人形, 植物のシーン.

入力多視点画像として「筑波大学多視点画像データベース」を使用した. 被写体は街のミニチュア模型である. 入力画像は 81 枚 (9×9) で, それぞれの解像度を 256×192 画素とした. 合成画像の解像度は 256×256 画素, “合焦”距離の数 $N = 10$, (3) 式で $M = 3$, (7) 式で $W = 7$, (8) 式で $R_{th} = 10$ とした.

図 3 に, 入力画像が取得された位置とは異なるある視点で, “合焦”距離を変化させながら合成した画像を示す. 左列はモード A, 右列はモード B で合成された画像である. これらの画像には“焦点”ぼけの現象が現れており, モード A とモード B では, “焦点”ぼけの発生の仕方に差異が生じていることが分かる. 提案システムでは, この性質を利用して, “合焦”領域を判定する.

図 4 は, 提案手法 (奥行き画像の平滑化を含む) による“全焦点”画像である. 左には図 3 と同一の視点位置, 右には別の視点位置で合成した画像を示した. このように提案システムでは, 複数の画像の“合焦”領域が選択的に統合されるため, シーン全体を鮮明に捉えた, 視覚的に良好な“全焦点”・自由視点画像が合成できることが分かる. 図 5 は, 同データベースに含まれる人形および植物のシーンについての合成結果である.

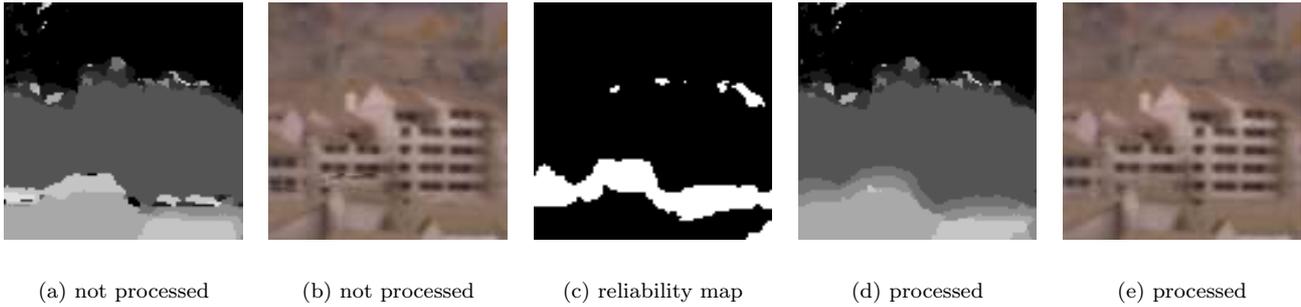


図 6: 奥行き画像平滑化の効果 . (a)(b) : 平滑化しない場合の奥行き画像および合成画像 . (c) 信頼度画像 . (d)(e): 平滑化した場合の奥行き画像および合成画像

表 1: 処理ステップの定義

- Step.(A) 多焦点画像群の合成, およびそれらの差分計算 (2) 式)
- Step.(B) 差分画像のブロック平滑化 (3) 式)
- Step.(C) 奥行き画像および “全焦点” 画像の合成 ((4),(5) 式) . 奥行き画像に対する平滑化を行う場合は, その処理を含む .

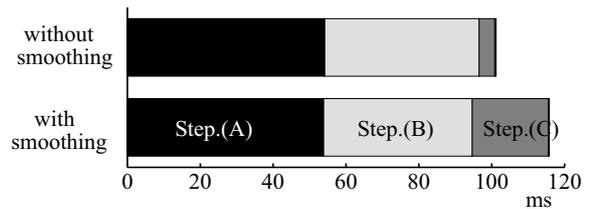


図 7: 描画速度の比較 .

4.2 奥行き画像の平滑化の影響

まず, 図 6 を用いて, 平滑化処理が合成画像の見えに及ぼす影響を説明する . (a),(b) は, 平滑化を行わない場合の奥行き画像と合成画像である . 手前の建物の屋根の部分と, 奥の高層ビルの境界付近に, 不自然なノイズが生じていることが分かる . (c) は, 奥行き推定の信頼度の評価値 $R(x, y)$ が閾値以上になる画素を白で, それ以外の画素を黒で示した画像である . (b) と (c) を比較すると, ノイズが出ている領域は, 信頼度も低くなっているので, 信頼度を考慮した平滑化が有効であることが分かる . (d),(e) は, 信頼度を考慮した平滑化を行った場合の, 奥行き画像および合成画像である . (b) と (e) を比較すると, 平滑化処理によって, 最終的な合成画像がより自然に見えるようになっていることが分かる . 特にインタラクティブに視点位置を移動させる場合に, この効果はより知覚されやすくなる .

次に, 平滑化処理が計算時間に及ぼす影響について議論する . 図 7 は, 平滑化を行わない場合 (上) と行う場合 (下) の計算時間の測定結果 (100 回描画した平均値) である . ここでは表 1 に示したように, 実装の形態に合わせて, 描画処理の全プロセスを 3 つのステップに分けて表記する . ステップ (A), (B) の処理は両者で共通である[§], ステップ (C) の処理時間は, 平滑化をした場合, 平滑化をしない場合と比較して 17ms 増加した . 結果として, 全体の処理時間は約 14% 増加したが, これは, 合成画像の品質の向上に対して妥当な範囲であると思われる .

5. むすび

本稿では, 多視点画像を用いて, インタラクティブに自由視点・“全焦点” 画像を合成する light field viewer における, 奥行き画像の平滑化処理について検討した . 信頼

[§] (A) (B) において, 両者で処理時間にやばらつきが見られるが, ここには本質的な違いはないと思われる .

度を考慮した簡易な平滑化アルゴリズムを導入することにより, 筆者らの条件下では, 計算量の増加を 14% にとどめ, 合成画像の不自然なノイズを低減することに成功した .

謝辞 : 有益な議論をしてくださった東京大学 原島博教授に感謝します .

参考文献

- [1] M. Levoy and P. Hanrahan. “Light Field Rendering”, Proc. ACM SIGGRAPH, pp. 31 – 42. 1996.
- [2] A. Isaksen, M. Leonard, and S. J. Gortler. “Dynamically Reparameterized Light Fields”, Technical Report MIT-LCS-TR-778, 1999.
- [3] A. Isaksen, L. McMillan, and S. J. Gortler. “Dynamically Reparameterized Light Fields”, Proc. ACM SIGGRAPH, pp. 297 – 306, 2000.
- [4] J. Stewart, J. Yu, S. J. Gortler, and L. McMillan. “A New Reconstruction Filter for Undersampled Light Fields”, Proc. EGSR, pp. 150 – 156, 2003.
- [5] 高橋, 久保田, 苗村. “Light Field Rendering における全焦点画像合成に向けた合焦判定に関する基礎検討”, 信学技報 103, 296, pp. 43 – 48. 2003 .
- [6] K. Takahashi, A. Kubota, and T. Naemura. “A Focus Measure for Light Field Rendering”, To appear in Proc. IEEE ICIP, Oct. 2004.
- [7] K. Takahashi and T. Naemura. “All in-Focus Light Field Viewer”, To appear in Proc. ACM SIGGRAPH Posters, Aug. 2004.
- [8] J.-X. Chai, X. Tong, S. -C. Chany, and H. -Y. Shum. “Plenoptic Sampling”, Proc. ACM SIGGRAPH, pp. 307 – 318, 2000.
- [9] 高橋, 苗村, 原島. “自由視点画像合成における光線情報のサンプリングと被写界深度の理論”, 映情学誌. 57, 10, pp. 1292 – 1299, 2003.
- [10] S. K. Nayer and Y. Nakagawa. “Shape from Focus”, IEEE Trans. PAMI, 16, 8, pp. 824 – 831, 1994.