

ライトフィールドカメラを導入した焦点調節模擬システムの設計

The design of simulation system for accommodation introducing a light field camera

梅田 修一†
Shuichi Umeda

比留間 伸行†
Nobuyuki Hiruma

1. まえがき

我々は現実の3次元空間において物体の奥行き形状を把握する際に、両眼視差・輻輳・運動視差・焦点調節を主要な視覚の手がかりにする（奥行き手がかり）。現在までに実用化されている2眼式立体テレビでは、4つの主要な奥行き手がかりのうち両眼視差と輻輳を再現することによって奥行き感を与えている。しかしながら視点移動による運動視差（以下、単に「運動視差」）および焦点調節は再現できず、これが視覚疲労や、現実と異なる奥行き感を生む原因となっている可能性が高い[1]。この問題に対して、マイクロレンズアレーを用いて撮影することで原理的には焦点調節までも再現できるインテグラルフォトグラフィ（IP）立体テレビの研究が続けられている[2]。より自然な立体テレビ方式の実現に期待がかかるが、光学的な問題やデータ量が膨大になることに起因する技術的な問題により、実用化には至っていない。

テレビで奥行き感を伝えるためには、視聴者に奥行き感を与えるための奥行き手がかりを忠実に記録したあとで再現する必要がある。問題の所在を明らかにするために、テレビを見るとときと実物を見るときの概念上の比較において、図1のように情報伝送の過程を分割する。この分割により、立体テレビを視聴するときと現実の物体を見るときの間に違和感がある場合に、ディスプレイの性能が悪いのか、それとも奥行き感を与えるために十分な光線情報を取得できていないのかを判断できる。撮影・記録系で十分に奥行き手がかりを取得していない状態では、いくら再生・表示系の性能を向上させても奥行き感の向上を見込むことはできない。

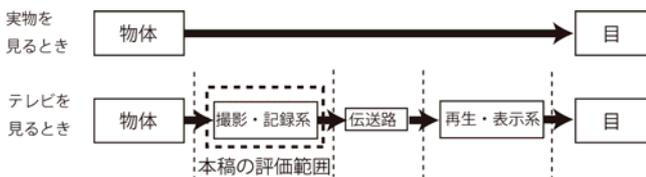


図1. 実物を見ることとテレビを見ることの比較

本稿では、特に撮影・記録系について、主に「焦点調節」の入力となる光線情報を十分に取得できているかどうかを評価するモデルを提案する。ヒトの眼の構造においてもカメラの構造においても、焦点調節の結果は、ピントが合っている箇所とボケている箇所が存在する2次元画像（ボケ画像）で表されることに着目した。取り込んだ光線情報を用いて自然なボケ画像を生成できるかを評価の基準にすることで、再生・表示系とは独立して評価することを目指した。

評価モデルの実現のために、IP立体テレビと同様の撮影手法を導入した。マイクロレンズアレーを使って光線情報を取り込んだ後にコンピュータの処理でレンズの働きを計算する考え方は、コンピューショナルフォトグラフィの一例であるライトフィールドカメラに近い。ただし、一般的なライトフィールドカメラは実時間処理することが重要であり[3]、取得する光線の数を制限するため、ボケ具合とボケた箇所を同時に調整できない[4]など、再現性に優れているとは言い難い。本稿では入力を静止画に限定して処理にかかるコストを考慮せずに、焦点調節時の光線の進行を忠実に追跡するモデルの構築を図った。

2. 研究の背景

2.1 焦点調節とライトフィールドカメラ

通常のカメラでは、撮影時にピントを調整して画像を記録してしまうと、後からピントが合っている場所を変えることができない。しかし3次元空間において物体を見ている眼球の調節機能を測定すると、水晶体の表面曲率を変えて焦点距離を調節している。また、絞りに相当する瞳孔径も、主に光強度の要因で変化する（情動によっても変化する）とされる。固定した単眼で見て物体の前後関係を判別することができるのは焦点調節の要因も寄与しているが、撮影後にピントを調整できない通常のカメラは焦点調節によって得られる全ての光線情報を取得していないといえる。一方、ライトフィールドカメラは通常のカメラと異なり撮影後に焦点位置を変えることができる。したがってライトフィールドカメラが記録する画像情報によれば現実の物体を見るときに眼球が時系列的に取り込んでいる情報の状態を再現できる可能性がある（図2）。原理的には、通常のカメラでは撮像素子上に入射する光線情報のうち、光線の強さのみを記録するのに対して、ライトフィールドカメラではマイクロレンズアレーを介して光線の強さと方向を記録する。その後コンピュータで複数枚の画像に分解して焦点距離に応じたシフト加算後、平均画素値を求めることにより、焦点位置を変えた2次元画像を得ることができる[4]。

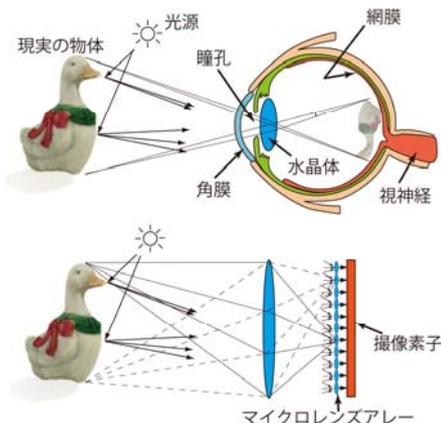


図2. 眼球(上)とライトフィールドカメラ(下)の比較

2.2. 多眼式とインテグラルフォトグラフィ方式

3.1節で焦点調節を模擬する手法を提案するために、要素画像が通常のカメラの画像に近いことから理解しやすい多視点カメラ技術を導入する。視点を複数持って被写体の立体情報を得る面からライトフィールドカメラも一種の多視点カメラであると考えられるが、本稿では複数視点映像技術の種類として両者を区別する。また、複数視点映像技術と光線取得方式の混同を避けるために各方式名称を表1の通りに分類する。

表1. 各方式名称

複数視点映像技術	光線取得方式
多視点カメラ	多眼式
ライトフィールドカメラ	インテグラルフォトグラフィ(IP)

第3章で評価モデルを実装する上で注意すべき、多眼式と比較した際のIP方式画像の特徴を以下に挙げる。

- (A) 要素レンズ1単位あたりの撮像素子面積が小さいため、画面全体にボケのない画像が得られる。
- (B) 視点数が比較的多い。
- (C) 要素レンズの歪曲収差は全周魚眼レンズに近い。

3. 提案システム

3.1 多眼式画像群を入力とする焦点調節模擬

(第I段階)

第3章では具体的に撮影・記録系の評価モデルを提案する。撮影・記録系において情報が欠損する原因を詳しく見ると(ア)マイクロレンズアレー等の設計に基づくサンプリングによる欠損と(イ)レンズの収差による光学的な限界に分類できる。原因の混在を避けるために、評価モデルの構築にあたり3つの段階を経ることとした。第I段階ではマイクロレンズアレーを用いない概念的な多眼式入力で焦点調節を模擬する。第II段階ではマイクロレンズアレーを模したシミュレーション入力に拡張して、(ア)の評価モデルを実装する。第III段階ではマイクロレンズアレーを用いて撮影した実写を入力として、(ア)および(イ)の評価モデルを実装する。第II段階と第III段階の結果を比較することで(イ)の収差による原因のみを評価することを図った。

はじめに第I段階として、多眼式画像群を入力とする焦点調節模擬システムを試作する。ここでは提案システムへの入力画像を、CGソフトを使ったシミュレーションによって作成する(図3.)。まず3DCGソフトのシーン内で、3次元幾何学モデルを製作し、このモデルを被写体とする仮想カメラの位置をずらしながらレンダリングして、視差のある複数の画像群を生成する。次に、レンダリングした際の仮想カメラの相対的な位置関係を保持したまま、出力した画像群をタイル状に並べて1枚の画像を作成する。作成した画像は、仮想カメラの位置座標情報(u, v)と、カメラ座標系の座標値(s, t)からなる多眼式画像 F とする。このとき、多眼式画像 F におけるある一点の画素値は $f(s, t, u, v)$ で表され、4次元配列に格納される。

作成した多眼式画像を入力して焦点調節を模擬するアルゴリズムを図4.で説明する。調整項目は通常のカメラでボケ具合を変えるための「絞り」と、ボケている箇所を変えるための「焦点距離」および「視点位置」とした。調整によって変化する明るさや画角の違いは反映しない。このア

ルゴリズムでは3つの調整項目を同時に変更できる。一般のライトフィールドカメラでは「焦点距離」または「視点位置」のどちらかしか変更できない点对照的である[4]。

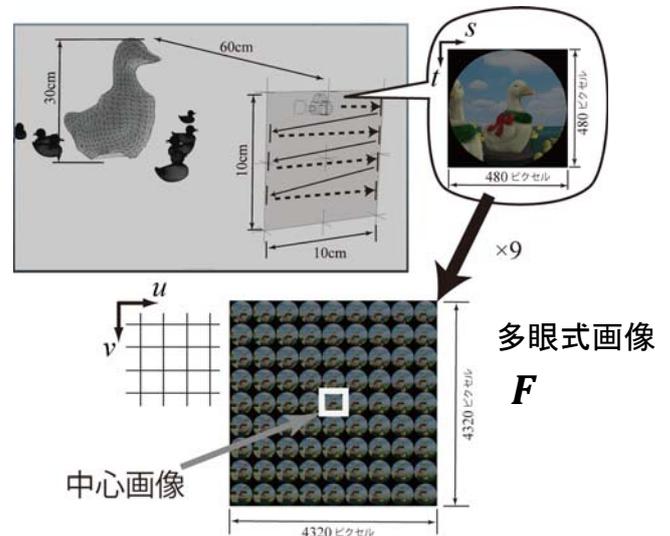


図3. 3DCGのシミュレーションによる多眼式画像作成

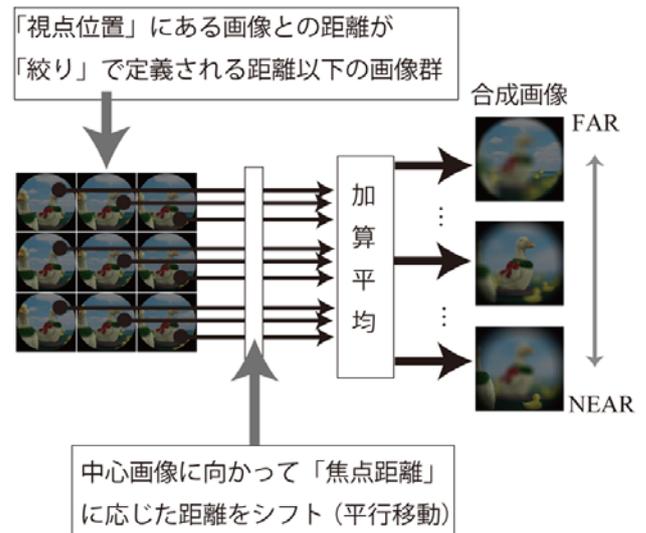


図4. 多眼式画像の焦点調節模擬アルゴリズム概要

図4.に示すアルゴリズムを実装して、リアルタイムに3つの調整項目修正を行うインターフェースを試作した(図5.)。画面右側に表示されるカメラアレー中で、 \times で示した部分が「視点位置」であり、マウスのクリックによって指定する。また、 \square で示した部分は「絞り」を開閉することで面積が変化する。これは合成に加わる要素画像を表しており、面積が大きくなるほど合成画像のボケ具合は大きくなる。またレンズの絞りを開いた状態で、「焦点距離」に応じた距離だけ要素画像がシフトする結果、合成画像のピントが合っている箇所が変わる。こうして「絞り」と「焦点距離」の組み合わせにより、あたかもカメラのファインダーを覗いてレンズの絞り値と焦点距離を変える時のような操作感を確認することができる。また、「視点位置」をリアルタイムに変えることができるため、運動視差と焦点調節を同時に模擬することができる。



図5. 多眼式画像の焦点調節システム (第I段階)

入力画像に用いたCGのレンダリング画像は理想的に歪みの無いパンフォーカスな画像である。しかし実写の多眼式画像に同じアルゴリズムを適用した場合には、全体にボケが残って解像感の低い合成画像になることが予測される。そこで第II段階以降では、2.2節で挙げた特徴(A)を有することから、入力をIP方式の画像に拡張する。

3.2 焦点調節模擬のIP方式画像入力への拡張 (第II段階)

第II段階として、IP方式による入力画像をシミュレーションにより生成し、第I段階で試作した模擬システムに適用する(図6)。

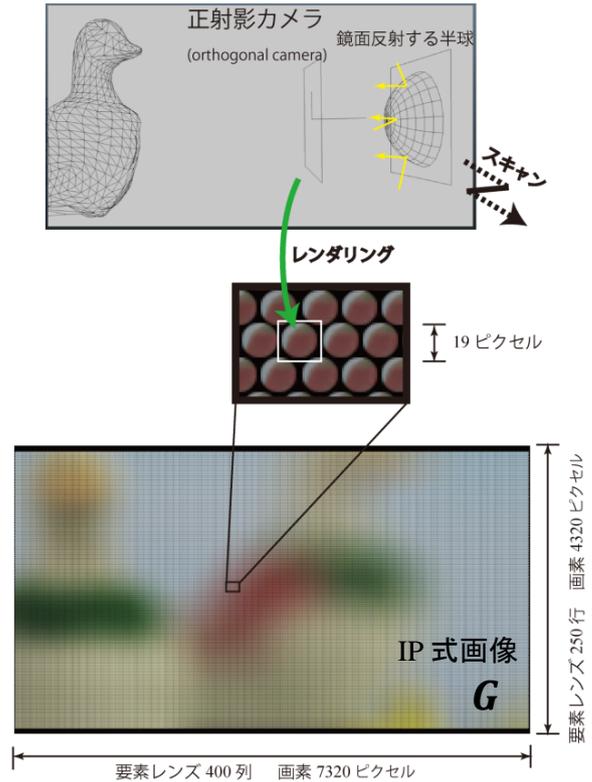


図7. IP方式画像のCGシミュレーションによる作成



図6. IP方式シミュレーション画像への拡張 (第II段階)

まず第I段階と同様にCGシミュレーションにより入力画像を生成する。ただし、今回のCG画像には2.2節で挙げたIP方式の特徴(B)と(C)を持たせる。(B)に対応するために、実在するマイクロレンズアレイに合わせて、要素レンズを水平方向400列、垂直方向に250行のハニカム状に配置したマイクロレンズアレイを模して仮想カメラの位置を定めた。記録する画素数の合計も実在するSHV用撮像素子に合わせて7320×4320ピクセルとした。また、(C)に対応するために、CGソフトで魚眼レンズを模擬した。通常のCGソフトでは、ほとんどの魚眼レンズで採用される等距離射影方式の仮想カメラを実装していない。そこで、CGのシーン内で、鏡面反射する半球形のオブジェクトを配置し、天頂方向に配置した正射影カメラに至る反射像をレイトレーシング法によってレンダリングした。シミュレーションの概要を図7に示す。

次に、第I段階で提案したアルゴリズムに入力するために、IP方式画像を多眼式画像に変換する。ただし、多眼式画像では視点位置を定義できるのに対して、IP方式画像は視点位置を定義できない[5]ため、変換時に視点位置を仮定する必要がある。ここでは、射影する視点は正射影カメラであるとする。正射影カメラとすることで、視点の奥行方向の距離を考慮する必要が無く、画像と視点位置がつくる相対位置関係から射影を計算できる(図8)。

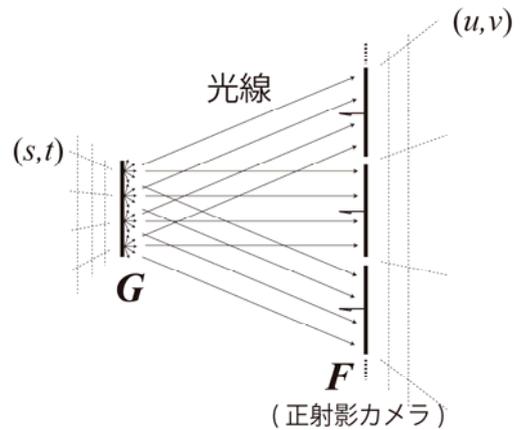


図8. IP方式画像から多眼式画像への変換

このとき画像Gと画像Fの間には

$$G \rightarrow F \quad (1)$$

$$f(s, t, u, v) = g(u, v, \max(s) - s, \max(t) - t) \quad (2)$$

f, g はF, G上の点

が成り立つ。シミュレーション画像の変換結果を図9に示す。注意が必要なのは、CGによるシミュレーションで元画像を作成しているために収差が無く、ボケの無い理想的な要素画像となることである。仮想カメラと実写の撮影条件が同一である仮定のもとでは、収差の影響を除いた光線サンプリングのための設計を評価する基準にできる。

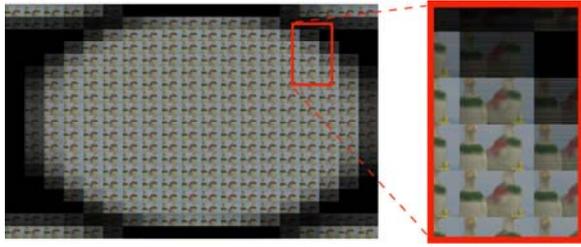


図9. IP式画像から多眼式画像への変換画像

3.3. レンズアレーを介した撮影画像入力の適用 (第III段階)

第III段階として、実際に物体をIP撮影した実写画像を入力とする。撮影に用いたマイクロレンズアレーはシミュレーションと同じ400列×250行、使用した撮像素子もシミュレーションと同じ7320×4320ピクセルである。得られる画像を第II段階で示した画像変換後、焦点調節模擬システムに入力する(図10.)。ここで実写画像を入力して第II段階のときと同じように焦点調節が模擬できない場合には、レンズの収差に問題があると判断することができる。

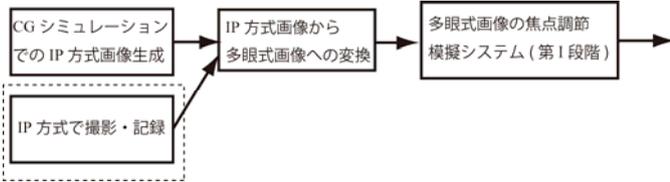


図10. IP式実写画像入力への拡張(第III段階)

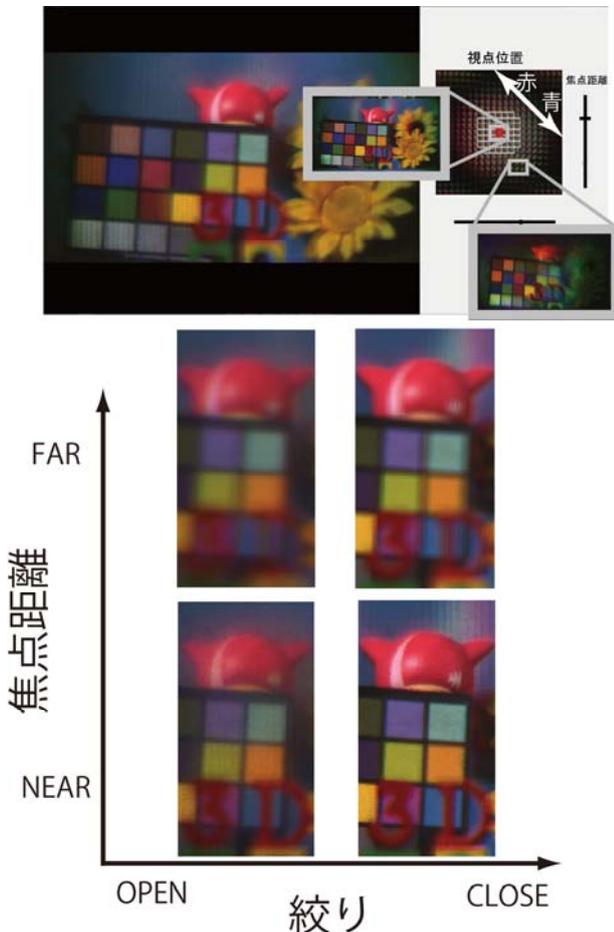


図11. 実写画像を入力とした焦点調節模擬システム

図11.に示した焦点調節模擬システムの例で、画面右に表示する多眼式画像へ変換後の入力画像を見ると、収差の影響が明らかとなる。色収差のために左上の要素画像に赤みがかかり、右下の要素画像には青みがかっている。また、周辺に位置する要素画像は中央の画像に比較してモアレ等で劣化している。変換した多眼式画像を焦点調節模擬システムに入力して、絞りを開いた状態で焦点位置を遠方にすると、ピントが合っていた近景(図中、「3D」の文字)のみがボケる。これは試作システムが焦点調節を模擬していることを示すが、シミュレーションによる入力するときほど顕著ではない。レンズの収差により入力の要素画像が劣化した結果、合成画像においてピントがあっている箇所とボケている箇所の画質の差が小さくなったためと考えられる。

4 まとめと考察

4.1 試作システムと奥行き手がかりとの関連

ライトフィールドカメラの機能を拡張して、光線の数を変える「絞り」と光線の屈折率を変える「焦点距離」、および「視点位置」をインタラクティブに制御する画像処理システムを試作した。試作システムを用いて奥行き手がかりの焦点調節や運動視差を模擬すれば、インテグラル立体テレビの撮影・記録系におけるマイクロレンズアレーの設計や、収差の補正技術に役立てられる可能性がある。

現在のところは収差の補正を行っていないために焦点調節の模擬は顕著ではないが、視点位置を変えて運動視差を模擬するとはっきりと視差を与えることができる。運動視差の方が焦点調節よりも感度が高いことは弁別閾を測定した実験結果[6]が示しており、試作システムが模擬する結果に近いと言える。

4.2 今後の課題

本稿の主な目的は立体情報を取得する撮影・記録系の評価モデルの提案であり、シミュレーションと実写での実験条件を統一していないが、本来はCG空間と現実空間を同じ条件にして、同じ形状モデルを撮影して比較するべきである。今後は双方の条件を可能な限り一致させて比較しながら、収差の補正技術を向上したい。さらには、奥行き感を与えるために最適な撮影システムの設計等に応用したい。

参考文献

[1] 伊藤崇之, “技研における立体テレビの研究成果”, NHK 技研 R&D, No. 123, pp. 48-55 (Sep. 2010)
 [2] 日浦人誌, 岩館祐一ほか, “インテグラル立体映像観察時の調節応答に関する一検討”, 映像情報メディア学会 立体映像技術研究会資料 HI2013-34 3DIT2013-2, pp. 3-6 (Mar. 2013)
 [3] A. Isaksen, L. McMillan and S. J. Gortler, “Dynamically reparameterized light fields”, SIGGRAPH 2000, 297-306.
 [4] 蚊野浩, “コンピュータショナルフォトグラフィーライトフィールドカメラ Lytro の動作原理とアルゴリズム”, 日本学術振興会, 光エレクトロニクス第130委員会第286回研究会公開シンポジウム資料 (Mar. 2013)
 [5] 小池崇文, “光線空間のサンプリングに基づいた多眼式立体ディスプレイ及びインテグラルフォトグラフィ立体ディスプレイの測定”, フラットパネルディスプレイの人間工学シンポジウム資料, (Mar. 2009)
 [6] Stephen R. Elis, Mary K. Kaiser, “Pictorial communication in virtual and real environments”, CRC Press, ISBN 0748400826, 9780748400829 (Jul. 1993)