

[招待論文] コンピュータショナルフォトグラフィ概観

日浦 慎作[†]

[†] 広島市立大学 〒731-3194 広島市安佐南区大塚東 3-4-1

E-mail: [†] hiura@hiroshima-cu.ac.jp

あらまし デジタルカメラは従来のカメラに装填するフィルムを撮像素子に置き換えたものであり、レンズが形成する像をそのまま撮像素子により電子化することが基本となっている。それに対しコンピュータショナルフォトグラフィ(Computational Photography)では、光学系が形成する像はあくまで対象の情報の中間表現であると考え、その後の画像処理との組み合わせにより美しい像を出力しようとするものである。この自由度の付与によって新たな機能や性能をカメラに与えることが出来ると期待されている。本稿では ISSCC2012 のフォーラム”Computational Imaging”における自身の講演の流れを再現する形式で、この技術について概観する。

キーワード コンピュータショナルフォトグラフィ, ライトフィールド, 符号化撮像

Overview of Computational Photography and Imaging

Shinsaku HIURA[†]

[†] Hiroshima City University 3-4-1 Ozuka-Higashi, Asa-Minami-Ku, Hiroshima 731-3194 Japan

E-mail: [†] hiura@hiroshima-cu.ac.jp

Abstract The optical image on a sensor formed by a conventional lens mostly determines the quality of final image. Computational Photography is a concept to change the game of taking photograph by offering attractive functions such as blur reduction or posterior focus adjustment. In this framework, optical device is assumed as an encoder of sampled incoming rays, and the final image is reconstructed by decoding of the captured intensity distribution on the sensor. In this article I will introduce several basic concepts and achievements in this field by reproducing my talk in ISSCC2012.

Keyword Computational Photography, Light Field, Coded Imaging

1. はじめに

本稿では ISSCC における講演の大まかな流れを再現する形で、スライドに説明を加えながら解説を行います。そのためこの原稿を敬体で記述しますが、あらかじめご了承ください。

講演では、まずコンピュータショナルフォトグラフィとは何か？について簡単に述べます。続いて、我々が住む3次元空間を埋め尽くす光線と、それを記録するための機械であるカメラの関係について説明し、光線を完全に記録するにはどのような装置が必要か、またそれが出来るとどのようなことが可能になるかについて2章で説明します。続いて、カメラが光を取り込むときに必ず発生する「積分」の効果と、それによる情報の損失について述べ、その情報の損失を抑制するための方法を紹介します。最後に私自身が近年取り組んでいる、コンピュータショナルフォトグラフィに関する研究事例を紹介します。

1.1. Computational Photography とは何か？

図1は従来のカメラと、コンピュータショナルフォトグラフィとの違いを端的に表しています。デジタル

カメラは、従来のカメラのフィルムを撮像素子と画像処理回路に置き換えたものと考えられます。これらの電子部品は、基本的にはレンズによって形成された画像をなるべくそのまま再現すべきであると考えられており、その意味で画像を形作る主役はレンズです。しかし近年の計算機性能の発展により、画像処理の自由

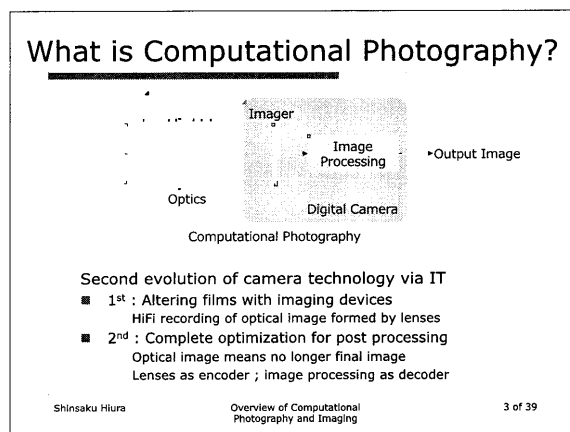


図 1. Computational Photography とは？

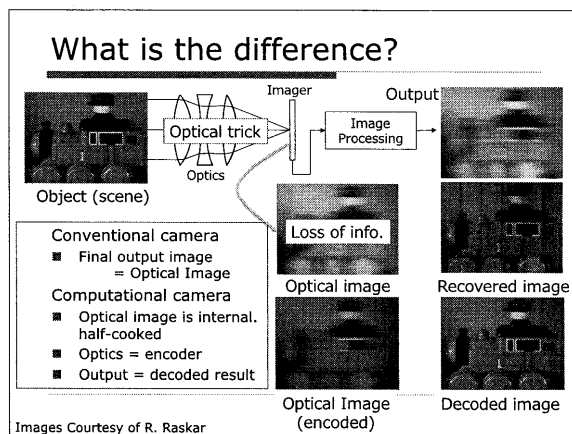


図 2. 違いは何か？

度は飛躍的に向上してきています。ならば、もはや単に光学像を再現するための受身の画像処理ではなく、画像処理の介在を前提として光学系を最適化することでより高性能な、または高機能なカメラを提供できるのではないかと考えられます。これがコンピュータショナルフォトグラフィを動機づける考え方となっています。具体的な例を図 2 に示します。シャッターが開いている間に物体が動くとその像はぶれてしまいます。ブレ除去技術によりシャープな画像を得ることは出来ませんが、どうしても原画像のノイズの影響を受けて画質が劣化してしまいます。なぜなら、画像が撮像素子上でぶれて観測された段階で、既に原画像の情報の一部が失われてしまっているからです。そこで光学系に何らかの工夫を施すことにより、画像上の情報の損失を抑制します。その結果、図 2 の右下の画像のように美しい像を復元することが出来ます。つまり、従来型のカメラは、光学像がほぼそのまま最終出力であったのに対し、コンピュータショナルカメラでは光学像は単なる中間表現の 1 段階に過ぎず、それを「復号」することで最終的な結果を得る点が大きな違いです。

2. ライトフィールドの取得と利用

ここではカメラについて述べる前に、まず空間を埋め尽くす光の分布の表現について考え、次にそれを取り込むための装置としてカメラを定義します。次に、光の分布を余すことなく取り込む方法、さらにその利用方法について述べます。

2.1. ライトフィールドとその習得

我々が生活する 3 次元空間は太陽や照明からの光で溢れています。これを表現するにはいくつの変数が必要でしょうか。図 3 のように光線はシーン中を縦横無尽に飛び交っていますが、その光線の通過位置 (X, Y, Z) と方向 (θ, ϕ) がまず思い浮かびます。さらに光

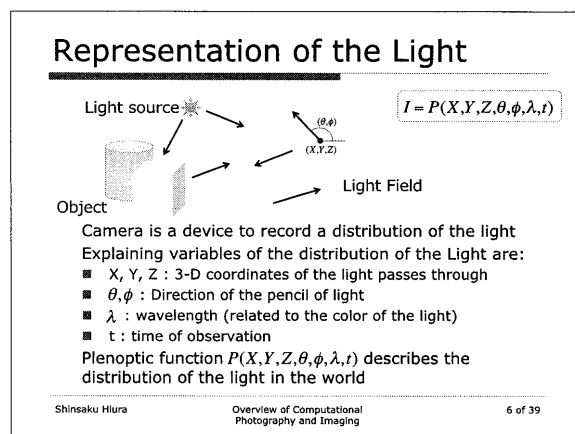


図 3. ライトフィールドの表現

の波長 λ 、光を記録した時刻 t の変数を加えた 7 つの変数で、シーン中の光の状況を（光の波動性や粒子性を度外視すれば）完全に表現することが出来ます。これら 7 変数を引数に持つ関数 P をプレノプティック関数と呼び、表現される光の分布をライトフィールドと呼びます。

次に、このシーン中にカメラを置いたときのことを考えます。図 4 に示すように、カメラは空間を埋める光線のうちごく一部を取り込み記録します。このとき、記録される値はプレノプティック関数 P の値を直接記録するのではなく、関数 P の一定区間を積分した値を記録します。具体的には、位置 (X, Y, Z) にはレンズが光を集める口径が対応し、絞りによって積分範囲を調整することが出来ます。同様に方位 (θ, ϕ) には各画素の面積、 λ には画素の分光感度分布、 t には露光時間が対応します。これらのどれか 1 つでも、積分区間の幅がゼロになると、カメラが取り込む光エネルギー量がゼロとなり、画像を得ることが出来ません。この積分が引き起こす問題とその解決策は、次の 3 章で詳しく論じます。

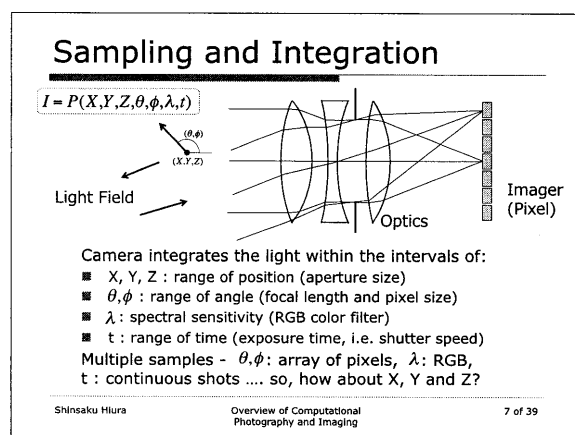


図 4. カメラによる標本化と積分

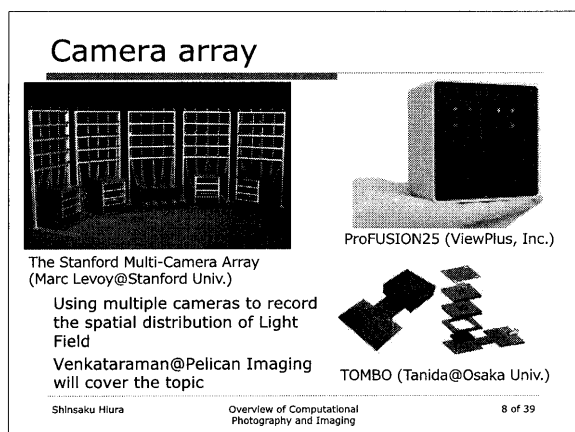


図 5. カメラアレイ

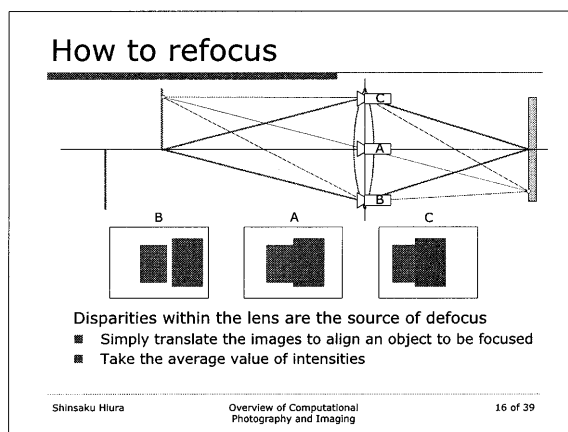


図 7. リフォーカシング

ライトフィールドの記録という立場からカメラを見ると、もう一つ気づくことがあります。それは多数のサンプル列の取り込みです。撮像素子上の多数の画素は、カメラが捉える光の入射方位(θ, ϕ)に関し多数のサンプルを同時に取得するためのものです。同様に時刻 t については連写することで、また波長 λ については赤・緑・青の光の三原色に対応した3つのサンプルが取得されています。しかし、位置(X, Y, Z)については、1台のカメラでは単一のサンプルしか取り込むことが出来ません。そこでこの問題を解決するために、いくつかの機器構成が提案されています。図5はカメラアレイと呼ばれるもので、複数のカメラを並べることでライトフィールドの位置に関するサンプリングを行うものです[1]。このような装置は自由視点映像生成や光線再生式立体テレビのために用いられてきましたが、これまでに述べたように、ライトフィールドのすべての変数についてサンプリングを行う装置として捉えることも可能です。

もう一つの機器構成として、単一のカメラでありながら、その内部に工夫を施したものがあります。普通

の撮像素子は、それぞれの画素へ到達した光の量を捉えることが出来ますが、その光が画素へ入射する角度を記録することは出来ません（もちろん入射位置については、複数の画素を並べることで記録されます）。これが、ライトフィールド情報がカメラによって失われる理由となっています。そこで撮像素子付近に工夫を凝らし、撮像素子へ入射する光の位置と方位の双方を記録する方法が提案されています[2]。図6はそのような装置の例で、撮像素子のすぐそばにマイクロレンズアレイやマスクを置くことで、カメラへと到達した光の方位と位置の双方を記録することが出来るものです。

2.2. ライトフィールドの利用

取得したライトフィールドの利用として代表的なものに、リフォーカシングがあります。これは、撮影後に被写体の任意の位置にピントを合わせ直すような操作を指します。図7ではレンズが手前の緑色の板にピントを合わせている様子を示しています。ここで、レンズの中央と左右に3台のカメラ A, B, C を配置した時を考えます。この時、A からは緑の板に対し、少し奥にある青い板がちょうど接するよう見えます。それに対し視点 B からは2つの板の間に隙間が見え、C からは重なって見えます。このように奥行き異なる物体の見かけの位置が相対的にずれることを視差と呼びます。

次に、これら3枚の画像を重ね合わせることを考えます。緑色の板のずれがなくなるように3つの画像を重ね合わせて画素値の平均を取ると、青の板は左右に少しずれているため、左右に色の薄い部分が生じます。これがぼけの実体です。実際にはレンズは円形ですから、カメラもレンズを埋め尽くすように円形に配置したとすると、視差も円形状に分布するため、画像のボケも円形になります。もし3枚の画像を青の板が一致するようにずらして重ね合わせると、今度は緑の板の

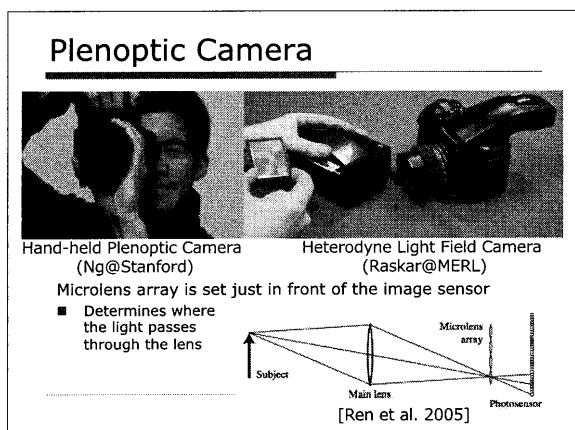


図 6. プレノプティックカメラ

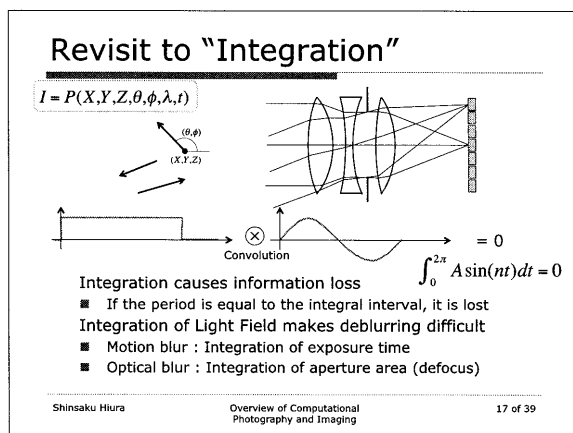


図 8. 積分による情報損失

部分にぼけが発生します。

このように、レンズによるぼけは、視点の異なるカメラにより捉えた像の視差が原因です。またレンズは、ある一定の距離にある物体について、視差をキャンセルして重ね合わせる働きを持っているということが出来ます。そのため、カメラアレイやプレノプティックカメラで様々な視点からの像を得ることが出来ると、レンズの光学的働きをシミュレーションすることによって、事後的なピント合わせが可能になります。もしライトフィールドを余すことなく細かく取り込むことが出来る装置があれば、得られたデータからいかなる光学系もシミュレーションすることが出来ます。

3. 積分と符号化撮像

この章では、2.1 節で述べた積分がもたらす問題について述べ、次にそれを解決する方法の1つである符号化撮像について紹介します。

3.1. 積分と情報損失

さきに述べたように、カメラはライトフィールドを積分する装置です。しかしこの積分が、シーンの情報を失わせる原因となっています。図 8 に示すように、ある正弦波を、その周期と一致する幅で積分（またはその幅を持つ矩形関数で畳み込み）すると、その値は常にゼロになります。つまり積分後の値からは、元の正弦波の位相や振幅がまったく分からなくなります。

このような積分の効果はライトフィールドの各変数の値について生じます。露光時間 t については高速で点滅する物体の輝度変化が分からなくなるだけでなく、動いている物体の像がぶれることで原画像が分からなくなる効果を生じます。位置 (X,Y,Z) にはレンズのぼけが対応し、ピントが外れた部分の原画像の復元に困難を来します。そこで考えられたのが符号化撮像です。

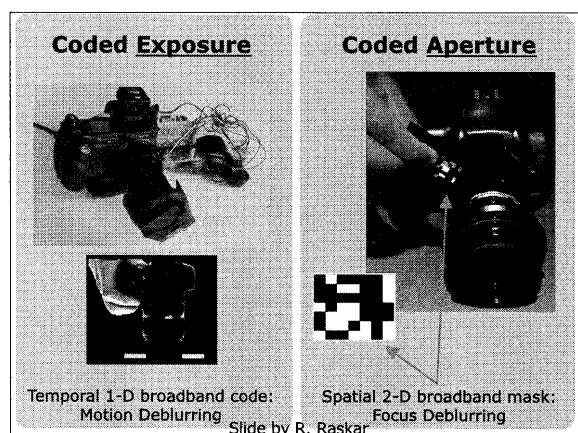


図 9. 符号化露光と符号化開口

3.2. 符号化撮像

符号化撮像は、図 8 にある積分区間の形状を工夫することで積分による情報損失を抑制する方法です。撮影時刻 t の積分に関しては、露光時間中にシャッターを何度も開いたり閉じたりし、積分区間の形状を変化させます。Raskar らは図 9 左に示したような高速液晶シャッターを備えたカメラにより、画像からぶれを良好に除去できることを示しました[3]。図 2 の写真はその結果の1つです。またレンズによるピンぼけに対しては、図 9 右に示したように絞りの部分に複雑な開口形状を持つマスクを設置した符号化開口が提案されています[4][5]。これによりボケを含んだ画像からシャープな画像を復元する性能が向上することが様々な研究者により示されています。

4. 研究事例

最後にこの章では、最近私自身が進めている研究の紹介を通して、コンピュータショナルフォトグラフィ技術やその考え方の広がりについて俯瞰します。

4.1. 画素形状のランダムコード化による超解像

3.2 節で述べた符号化撮像技術は、センシングにつきもののエネルギーの積分の多くに適用が可能な考え方です。実は分光（光スペクトル分布計測）の分野では古くから、アダマール分光法と呼ばれる方法で符号化の考え方が利用されてきました[6]。そうなるに残るは方位 (θ, ϕ) の符号化で、これは画素の面積による積分に相当します。そこで我々は、撮像素子の上に黒色粉末を散布し、各画素の受光感度分布をランダムに符号化することを考えました[7]。図 10 ではその考え方を示しています。普通、撮像素子は矩形の受光感度分布を持ちますが、それに対する最悪のケースとして、撮影対象シーンが縞模様で、その縞の周期が画素の受光感度分布の幅に一致する場合を考えます。この時全

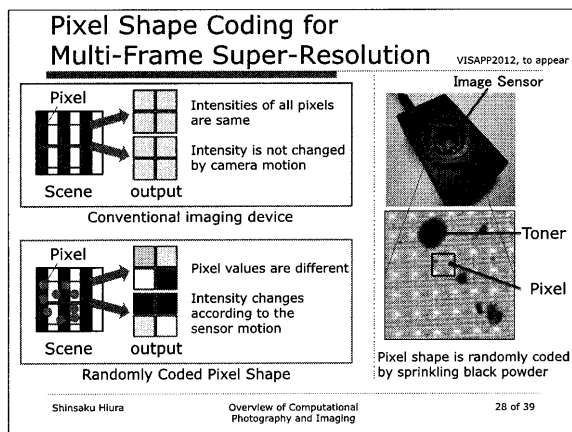


図 10. 画素形状のランダム符号化による超解像

での画素について、縞模様の黒領域が 50%，白領域が 50%観測され、画素値はグレーとなります。これは画素とシーンの位置関係が変化しても一定なので、画素値は常に中間調となり、元のシーンが縞模様であるということを知る手掛かりはありません。それに対し、撮像素子上に黒色粉末を散布するなどして各画素の形状を変化させると、それぞれの画素から得られる輝度値には違いが見られ、またシーンとカメラの位置関係が変化すると画素値にも変化が現れます。この変化を解析することで原画像の復元が可能になります。我々は実際に撮像素子に黒色粉末（レーザプリンタの黒色トナー）を散布し、その位置を画像から推定することにより、散布前の矩形画素のカメラよりも超解像の性能が向上することを示しました。

4.2. 符号化開口ステレオ

符号化開口はシーンの奥行きとレンズの口径に起因する光学的なぼけによる情報損失を抑制します。これをステレオカメラに導入することで、奥行き推定の安定化と、入力画像からのぼけの除去を可能とする研究を行いました[8]。図 11 はその処理の流れを示しています。ステレオカメラの双方のレンズに同一の符号化開口を与えます。これによりぼけの性質が変化しますが、両方の符号が同一であるため、ぼけた部分に生じる変化は左右のカメラで同じとなります。そのため、ステレオマッチング（視差の算出）では従来のステレオ計測と同じ方法を用いることが出来ます。さらに視差が一旦求められれば、画像に含まれるぼけの度合いを得ることが出来ます。ぼけの除去においては、対象物体までの距離によって変化するぼけの大きさを知る必要があります。単一のカメラでは容易ではありませんでした。本手法ではステレオカメラから精度良く距離が得られるので、これを用いて画像からぼけを除去することが出来ます。さらに奥行きマップとぼけ除去画像

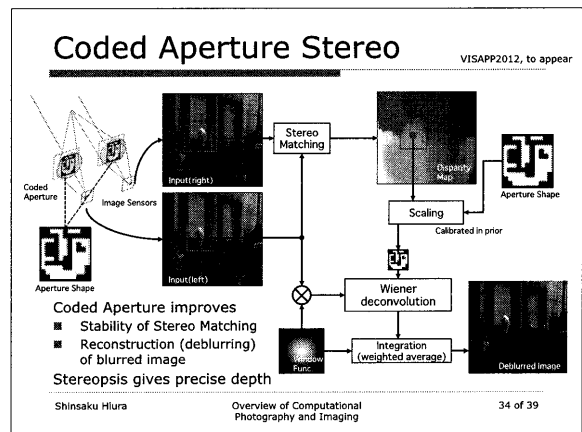


図 11. 符号化開口ステレオ

から、ピントの合う位置を変化させた画像を生成するリフォーカシングも可能です。

5. おわりに

コンピュータショナルフォトグラフィの根幹をなす考え方である、ライトフィールドのサンプリング・積分の問題と、それを解決する方法、さらに用途について述べました。今後は検査や自動化などの工業分野等への適用も期待されています。

文 献

- [1] The Stanford Multi-Camera Array, <http://www-graphics.stanford.edu/projects/array/>
- [2] R. Ng et al, "Light Field Photography with a Hand-Held Plenoptic Camera", Stanford Tech Report CTSR 2005-02, 2005.
- [3] R. Raskar, A. Agrawal and J. Tumblin, "Coded Exposure Photography: Motion Deblurring using Fluttered Shutter", SIGGRAPH2006.
- [4] A. Levin et al, "Image and Depth from a Conventional Camera with a Coded Aperture", SIGGRAPH2007.
- [5] S. Hiura and T. Matsuyama, "Depth Measurement by the Multi-Focus Camera", Proc. of CVPR'98, 953-959.
- [6] E. D. Nelson and M. L. Fredman, "Hadamard Spectroscopy," J. Opt. Soc. Am. 60, 1664-1669, 1970.
- [7] T. Sasao, S. Hiura and K. Sato, Coded Pixels : Random Coding of Pixel Shape for Super-Resolution, Int. Conf. on Computer Vision Theory and Applications, pp.168-175, 2012.
- [8] Y. Takeda, S. Hiura and K. Sato, "Coded Aperture Stereo for Extension of Depth of Field and Refocusing", Int. Conf. on Computer Vision Theory and Applications, pp. 103-111, 2012.