

コンピュータショナルフォトグラフィ： 画像獲得の再定義

日 浦 慎 作^{†1}

画像に関する様々な情報処理手法が提案され利用されているが、そのほとんどは既存のカメラにより得た画像を前提とし、カメラそのものの技術的發展とは独立したものであった。それに対しコンピュータショナルフォトグラフィ（以下、CP とする）は、変換や復号などの処理を施すことを前提に、光の記録段階から画像獲得の手法そのものを捉え直し、それにより高度な機能や性能の達成を目指すものであると言うことが出来る。本稿では CP に関する近年の研究成果をもとに、画像獲得手法が今後どのように再定義されていくのかについて考察したい。

Computational Photography : Redefinition of Imaging

SHINSAKU HIURA^{†1}

Originally, image processing is regarded as a post process of taken images captured by conventional cameras. Contrary, the computational photography is a concept to reconsider the way of imaging from the phase of gathering light with following dedicated signal processing such as decoding. In this paper, I will try to discuss about the upcoming redefinitions of imaging technology based on the recent research activities.

1. はじめに

コンピュータショナルフォトグラフィ(CP)とは計算機による演算によりシーンの画像・映像を獲得する手法である。もちろん、既存のデジタルカメラでもデモザイキングやノイズ除去など、画像に対し様々な演算が行われている。しかし本質的に現在のデジタルカメラは、

従来のカメラの写真フィルムを撮像素子に置き換えただけのものであり、そのため、画像を形成する主役として「記録デバイス上に鮮明な像を形成する」というレンズの役割は変わっていない。

一方、カメラを「被写体の像を画素値として出力する」という機能面から捉えたとき、レンズが記録デバイス上に形成する光学像はカメラ内部の単なる中間表現に過ぎないと言うことができる。つまりフィルム時代とは異なり、レンズはもはや鑑賞に堪えうる美しい像を平面上に形成する必要はなく、センサから出力されたデータを何らかの計算により「復号」することで画像が出てくれば、光学系はなんでもよい、ということになる。それでは、そのような演算の介在を前提としたとき、現在のカメラは最適な構成と言えるのだろうか。従来とは異なる光学系と、それに合致した信号処理により、従来は実現できなかった性能や機能をカメラに与えることが出来るのではないだろうか。これがコンピュータショナルフォトグラフィを支える基本的な動機である。半導体デバイスや計算機技術の発達により、フィルムが撮像素子に置き換えられることで、デジタルカメラが実現した。それに続くカメラ進化の第2幕は、情報技術によるカメラ構造そのものの改革であり、カメラや画像獲得手法の再定義であると言えるだろう。

本稿では近年の CP に関する研究に加え、それに関連する演算イメージング技術を振り返り、これからの CP 研究の展開と、それによるカメラや画像獲得の将来を展望したい。

2. 光線を記録する機械としてのカメラ

本章ではまず、カメラによる撮影とはどのような現象であるのかについて再考したい。普通我々はカメラによる対象物体の撮影というと、対象物体の輝度や形と、画素値との間の直接的な関係について考えがちである。しかしここではシーンを「光の分布を生み出すもの」として考えから切り離し、シーンが生み出した光線の分布と、その記録とについて考えることにする。

2.1 ライトフィールド

カメラが置かれた実環境は図1に示すように、光源から発せられた光や、それが対象物体表面で反射した光で縦横無尽に埋め尽くされている。そのような光線のうちごく一部がカメラのレンズを通過し、それが画像としてカメラから出力される。一般的に、環境を満たす光の強さ（分光放射輝度） I は以下のような数式で表すことが出来る。

$$I = P(X, Y, Z, \theta, \phi, \lambda, t) \quad (1)$$

ここで (X, Y, Z) は空間中の1点を、 (θ, ϕ) は光線の方向を表す。また λ は波長、 t は時刻

^{†1} 広島市立大学 大学院情報科学研究科 Graduate School of Information Sciences, Hiroshima City University

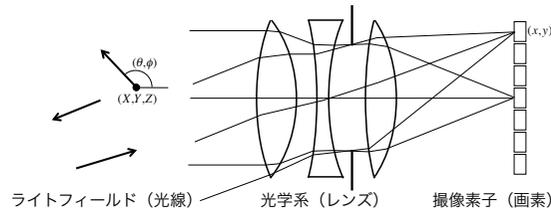


図1 カメラによる光線の記録。カメラは、実空間を埋め尽くす光の分布（ライトフィールド）のうち、レンズ口径内へ飛び込んだ光のみについて、 (θ, ϕ) の分布を記録する。レンズは方位 (θ, ϕ) を画素座標 (x, y) に変換するデバイスである。

であるので、この関数 P により環境中の光の状況を幾何光学的*1)に完全に記述することが出来る。この7変数の関数 P をプレノプティック関数と呼び、またこのような光線で埋められた3次元空間をライトフィールドと言う。

図1から明らかなように、カメラは環境中の光線のうち、レンズへ飛び込んだごく一部を内部に取り込み画像を出力する。その範囲はレンズの有効口径の大きさで決まる。また波長 λ （色彩）についてはカメラが感度分布を持つ範囲、さらに時刻 t についてはシャッターが開いてから閉じるまでの間（露光時間）の光だけを記録することになる。

方位 (θ, ϕ) についてはどうだろうか。これは画素に相当し、画素の面積が小さいほど、その画素が取り込む光の方位の範囲が狭くなる。つまり角分解能が向上する。ただし画素は撮像素子上に多数並んでいて、同時に複数のサンプルがカメラから得られる。同様に連写することで時刻 t について、また画素にカラーフィルタを掛けることによって波長 λ についても複数のサンプルが得られる。

では、位置 (X, Y, Z) についてはどうだろうか。これについては、カメラ単体では複数のサンプルを同時に得ることが出来ない。つまりカメラとは $(\theta, \phi, \lambda, t)$ に関する光の分布を記録する装置であるが、視点位置 (X, Y, Z) については単一のサンプルしか得られないような装置である。

2.2 ライトフィールドの獲得

前節で述べたように、カメラは単一の開口を通過する光だけを記録する。そうではなく様々な位置を通過する光線の輝度を得たいときには、多数のカメラを並べた装置を用いて

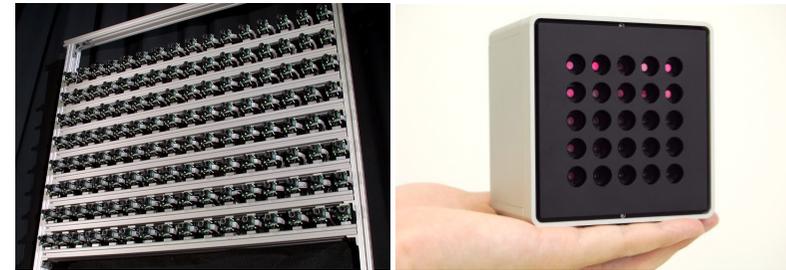


図2 ライトフィールド取得のためのカメラアレイの例。左：The Stanford Multi-Camera Array¹⁾、右：ViewPLUS ProFUSION25²⁾。

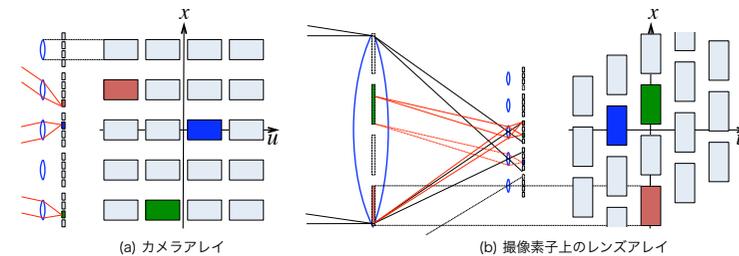


図3 ライトフィールドカメラの2つの構成。(a) 多数のカメラを並べたもの、図2に挙げたシステムはこのタイプである。カメラの個数が光線の位置分解能に、個々のカメラの画素数が角分解能に対応する。(b) 撮像素子上にレンズアレイを載せた構成。レンズアレイのレンズ個数が角分解能に、個々のレンズアレイに対応した小領域中の画素数が位置分解能に対応する。

光を記録する方法が考えられる。実際にこのような方法はイメージバーストレンダリング (IBR) において広く行われており、カメラアレイと呼ばれている¹⁾²⁾。ただし、式(1)において位置パラメータは X, Y, Z の3つだが、光は空気中や真空中を進む間にはほとんど屈折も減衰もしないため、カメラを3次元空間を埋めるように配置する必要はない。図2にカメラアレイの例を示すが、このようにカメラは普通、平面的に並べられる。

ライトフィールドを獲得するもう1つの方法に、レンズは単一であるが、カメラ内部に工夫を施すことにより、レンズうちの部分を通過した光であるのかを求める方法がある。Adelson ら³⁾は図3(b)に示すように、もともと撮像素子が設置されるべき焦点面にレンズアレイを設置し、さらにそこから一定間隔だけ離れたところに撮像素子を設置する方式のライトフィールドカメラを提案した。レンズアレイは主レンズの焦点面に置いてあるため、それ

*1 光には 1) 波動性：干渉や回折を起こす原因となる。波としての性質、2) 粒子性：光の強度は連続値でないなど、粒（光子）としての性質があり、それぞれ波動光学、量子光学として幾何光学と区別される。



図4 カメラ内部に工夫を施すことによりライトフィールドを取得可能としたカメラの例⁶⁾.

ぞれの小レンズはもとあった撮像素子の画素に対応する。それに対してそれぞれの画素は小レンズを介して主レンズの口径の一部に対応しているため、隣り合った画素は、それぞれ異なる部分口径からの入射光に対応する。図4はレンズアレイの代わりにマスクを撮像素子の直前に配置したカメラの例である。

図2や図3(a)に示したような、カメラを多数並べたカメラアレイでは、それぞれの小さなレンズが光線の通過位置(図では x)、またそれぞれのカメラの内部の画素が光線の方位(図では u)に対応する。それに対し、図3(b)や図4の方法では逆に、撮像素子の直前の小さなレンズはそれぞれ光線の方位に、また隣り合った画素は光線の通過位置に対応している。いずれの場合も、光線の通過位置と通過方位を多くの画素に分担して取り込んでいることは変わらない。

2.3 ライトフィールドカメラとコンピュータシミュレーション

先に述べたように、ライトフィールドカメラはIBR分野で広く研究が行われてきている。また立体テレビ方式の1つである、インテグラルフォトグラフィ方式立体テレビの入力手段としても用いられる。インテグラルフォトグラフィ方式立体テレビは、その別名である光線再生方式からも分かるように、カメラアレイによって獲得した光線を立体ディスプレイから完全に再生する方式である(図5)。この立体テレビは以下のように例えられる。例えば室内から窓ガラスを通して屋外を見たときに、屋外の風景から窓ガラスに到達した光線は、窓ガラスを通過してほぼ同じ位置から同じ方位に再び射出される。そこでこれと同様に、ライトフィールドカメラにより到達した光の通過位置と方位を計測し、それをそのまま立体ディスプレイにより表示すれば、理想的にはディスプレイによって表示される像は実物を直

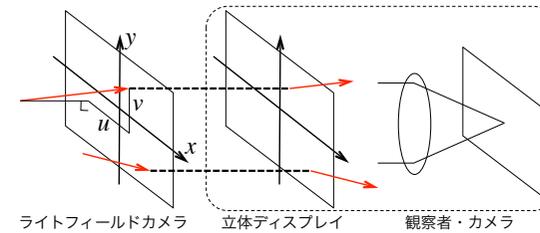


図5 立体テレビとコンピュータシミュレーション。理想的な光線再生型立体テレビをカメラで撮影すると、被写体へのピントの合わせ直しや撮影角度の変更が可能となる。これは計算機内での光線シミュレーションでも可能である。



図6 図4に示したカメラによるデジタルリフォーカスの例⁶⁾.

接視認するのと区別が出来なくなる。両眼用に2枚の画像を提示するだけの立体ディスプレイとは異なり、インテグラルフォトグラフィ方式では視点の移動による隠れ関係の変化が再現され、また頭部姿勢によらず常に立体感が提示される。さらには図5に示したように、カメラによるピント調整と、それによる前後のぼけも再現される。つまり、ディスプレイに表示された像に対し、カメラによる「撮り直し」が可能であるということになる。

コンピュータシミュレーションがもたらす機能の1つに、この撮り直しが挙げられる。図5に挙げた3つの要素のうち、立体ディスプレイとそれを撮影する第2のカメラ(図中の点線内)は実在の機器である必要はなく、ライトフィールドカメラにより取り込んだ光線情報を元にした計算機シミュレーションによっても同じことが出来る。特にピントの合わせ直しはデジタルリフォーカスなどと呼ばれ、コンピュータシミュレーションの代表的な応用例の1つである(図6)。

デジタルリフォーカスにはライトフィールドの取得が必要で、一般のデジタルカメラで撮影した1枚の画像からリフォーカスを行うことは出来ない。対象物体が静止している場合、カメラを少しずつ上下・左右に平行移動しながら多数の画像を撮影することでライトフィー



図7 手持ちカメラで撮影した画像からのデジタルリフォーカスの例⁵⁾。左から順に、原画像（5枚）中の1枚，前景に合焦した例，背景に合焦した例。

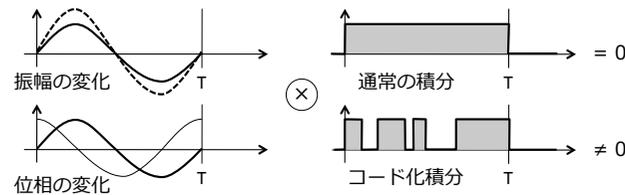


図8 積分による情報の損失と，コード化の効果。周期の整数倍が積分範囲に一致する信号の情報は失われるが，コード化によりそれを回避することが出来る。

ルドを取得することが出来るが，カメラを精密に移動させる機械はやはり一般的とは言えない。そこで筆者らは，手持ちで連写した複数の画像からデジタルリフォーカスを行う手法を開発した⁵⁾。この手法ではコンピュータビジョン分野において発展したカメラの自己較正法により各手持ちカメラの位置・姿勢を算出し，また視点間の画像を奥行き推定に基づく補間により生成することでなめらかなぼけを得ている（図7）。

3. 符号化撮像

2.1節で述べたように，カメラはライトフィールドを構成する光の分布のうち，その一部を各画素の値として出力する。ライトフィールドは連続量であるのに対し，各画素の出力はスカラー値であるため，ここで積分が行われる。例えば，カメラはシャッターが開いてから閉じるまでの間（露光時間）の光の強度を積分するため，露光時間が短いほど出力値は減

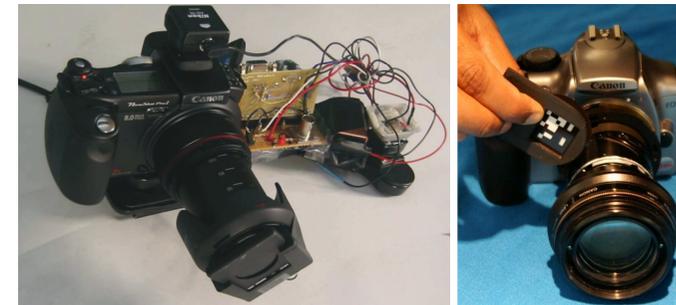


図9 Ramesh Raskarのグループによるコード化撮像の例。左：液晶シャッターによる符号化露光（時間方向の符号化）⁷⁾，右：マスクによる符号化開口（空間の符号化）⁶⁾の例。

少してしまう。しかし露光時間を長くすると，今度は動体がある間に移動する距離だけブレが生じてしまう。同様にピントのボケも絞り値によって決まるレンズの有効径のために生じ，絞りを絞るほど被写界深度が深くなるかわりに画像は暗くなり，SN比が悪化してしまう。このようにブレやボケなどのアーティファクトとSN比はトレードオフの関係にある。

近年，ぼけ除去（デプラー）や超解像などの技術が発展しており，これらのアーティファクトを事後的に改善することが出来るようになりつつあるが，その多くでは原画像に関する何らかの仮定を設けている。なぜなら図8に示すように，特定の周波数成分の情報は積分により完全に失われてしまうため，その値を仮定に基づき推定する必要があるからである。それに対しCPでは，この積分そのものに工夫を施し，情報の損失を最低限に抑える工夫をする。図8のように，積分区間全体に渡り入力値を均一に積分するかわりに，何らかのコードを乗じた値を取り込むことにする。コードとして例えばM系列のような広帯域符号を用いると，高周波成分に渡り情報の損失を抑えることが出来る。

この符号化の考え方は，2.1節で述べたライトフィールドの積分のすべての変数について行うことが出来る。時間変数 t に関する符号化は符号化露光と呼ばれ，Raskarらは図9左に示すように液晶シャッターをレンズの前に設置した装置により効果的にブレ除去を行うことが出来ることを示した⁷⁾。また光線の通過位置に関する符号化は符号化開口と呼ばれ，図9右に示したような装置を用いることで，奥行きと開口に起因するボケの除去に効果的である⁶⁾。筆者はこれらの研究に先立ち，ボケを用いた奥行き推定（depth from defocus）の性能改善法として符号化開口を用いた手法を提案した⁸⁾。この研究ではマルチフォーカスカメラと呼ぶ，異なる3種類の距離に合焦した画像が同時に得られるカメラ（図10）を用いたが，

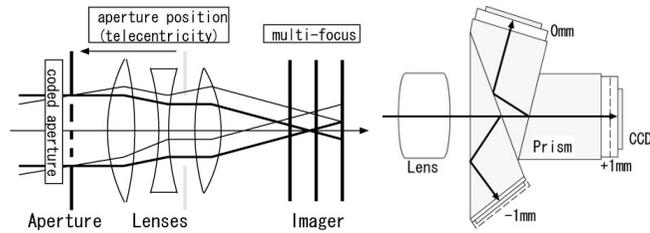


図 10 符号化開口を備えたマルチフォーカスカメラ⁸⁾。左：マルチフォーカスカメラの構成要素，右：3CCD カメラの改造による合焦距離の変更。

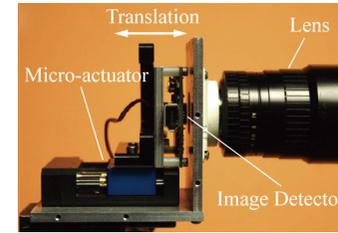


図 12 露光中の撮像素子の移動によるボケの不変量化¹⁰⁾。

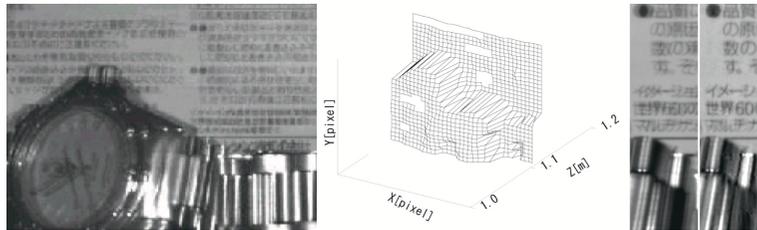


図 11 符号化開口を備えたマルチフォーカスカメラ⁸⁾ による距離計測とぼけ除去。左から入力画像のうちの 1 枚，推定された奥行きマップ，入力画像とぼけ除去画像の部分拡大。

このカメラのレンズの前に符号化開口マスクを付与することでボケ解析を安定化した。これにより、図 11 中央に示すような奥行きマップを得ることが出来、さらに図 11 右に示したようにボケを除去した画像を復元することが可能であることも示した。

このような符号化はライトフィールドを構成する他の変数についても可能である。実際に波長軸 λ については、アダマール変換分光法と呼ばれる方法が分光計測の分野で用いられている。方位 (θ, ϕ) には画素の大きさによる積分が相当するが、これについても筆者らは、画素形状をランダムにコード化することでマルチフレーム超解像の性能が向上することをシミュレーションにより示した⁹⁾。

3.1 積分の不変量化

前節で述べたように、積分に起因する情報の損失は符号化により抑制することが出来る。しかし、得られた画像から元の画像を復元するには、その復元パラメータを知る必要がある。例えばピントのボケは距離によってその大きさが変化し、また動体のブレ量は被写体の移動速度によって変化する。そこで、これらのボケやブレを、距離や移動速度によって変化

しない不変なものとするための工夫についても研究されている。

図 12 はボケの不変量化に関する研究¹⁰⁾ で用いられた装置であり、露光時間中に撮像素子を光軸方向に移動することが出来る。これにより、被写体はそれがどの距離にあっても、それがボケた像と、合焦した像の重ねあわせとして観察されることになり、事実上奥行きによってボケ方がほとんど変化しない。そのため画像全体に渡って同一の復元パラメータにより原画像を復元することが出来る。また、これを撮像素子の移動ではなく、静止した光学系により近似的に達成する手法として wavefront coding がある。また動体のブレについても、露光中にカメラの撮影方位を変化させることによって不変量化する研究¹¹⁾ が行われている。

4. コンピュータショナルフォトグラフィの周辺的话题

「コンピュータショナルフォトグラフィ」はいささか長い用語であるため、その訳語はないのか、と問われることがある。冒頭で述べたように CP は演算により画像を得る手法であるため、「演算写真」が直訳としても良いように思われる。しかしこれではどうも CG のように、純粋な演算のみにより像を作り出すものが第一義的に頭に浮かぶ。CP は逆に、シーンからの情報獲得のために、光学やデバイスにも踏み込んだ研究領域であり、これが CG との大きな違いであると言うことができよう。イメージベースレンダリング (IBR) との違いもここにあると言え、IBR はあくまでそれぞれ単体として完成した「画像」を入力とするのに対し、CP はその画像そのものを得る手法に強くコミットした研究であるという違いがある。

センサから得られた信号は画像ではなく、演算を経て初めて画像が形成されるものが CP であるとする、例えば CT (computed tomography) はそれに含まれるのかという議論もある。CP はあくまで photography、つまり鑑賞される写真のように、被写体表面の反射光の輝度分布を得るということが目的であるのに対し、CT で得る断層像は X 線等の透過率に対応しており、それとは異なるということと言える。しかし演算により像を得るという考え方

には相通じるものがあり、実用面で大いに活躍している CT 発展の歴史に学ぶべきところは多い。

CP に含まれると称する研究には、ブレ除去（デブラー）やノイズ除去（デノイズ）、超解像や HDR 合成なども挙げられる。これらならまだしも、さらには前景の切り出し（マッティング）や穴埋め（インペインティング）なども見受けられる*1。個人的には、これらのようにあまりに範囲を広げると画像処理の大半が含まれることになりそうに思え、抵抗感を持っている。つまり CP は、第一義的には、光を記録する過程に対する工夫に信号処理が組み合わされたものであると考えている。

冒頭の問いである CP の訳語であるが、いまのところもっともしっくり来るものは「演算撮像」であろうか。しかしこの語は現状としては使われていないし、英語には computational imaging に対応するだろう。結局、CG に日本語訳がないまま定着しているように、とりあえずそのままコンピュータショナルフォトグラフィと呼ぶことを続けている。

本稿タイトルの副題につけた「画像獲得の再定義」に絡み、では、CP 技術が発展すると現在のレンズは消えて行くのかという疑問もあるだろう。しかし、私はそうは考えていない。波動光学的な観点からも、多くの用途において従来型のカメラは最適解であると思われる。ただ、極端に広い被写界深度を得たり、事後にピント位置を変更した画像を作り出したりとといった、新たな機能や特殊な性能バランスを達成するためには、CP 技術の有効性が見いだせる場合があるということは無視できない。今後は写真を撮る道具としてのカメラ以上に、機械の眼としてのカメラの利用が盛んになると考えられ、その時に CP 技術は画像の認識・理解技術とより深く関わることで発展することになるだろう。

5. おわりに

本稿では CP 分野を構成する研究例を紹介しながら、CP 研究を支えるその動機や、それらの研究に通底する基本的なコンセプトについて紹介した。これまでの研究により現在は、様々な基本的アイデアの提案と、その効果の実証までが行われた段階であるといえる。しかし、個々の手法における光学系の最適構成やアルゴリズムの探求は十分に行われたとは言えないし、またこれらの手法をカメラ等の製品や工業応用等に展開するための開発もまた緒についたばかりである。

*1 R. Szeliski がまもなく刊行予定の教科書 "Computer Vision: Algorithm and Applications" には CP の章が設けられているが、現在のドラフト稿に挙げられているのはこの種の研究ばかりであり、西海岸流の、カメラに工夫を施すタイプの研究にはほとんど触れられていないようだ。

なお、紙幅の都合で個々の研究の詳細な解説や、その背景となる光学理論については割愛した。より詳しい解説として、2010 年 3 月に CVIM 研究会において開催されたチュートリアル^{12),13)} の原稿^{12),13)} を紹介しておく。

参 考 文 献

- 1) B. Wilburn, N. Joshi, V. Vaish, E. V. Talvala, E. R. Antunez, A. Barth, A. Adams, M. Horowitz and M. Levoy, High performance imaging using large camera arrays, ACM Trans. Graph. (SIGGRAPH), Vol. 24, No. 3, pp. 765-776, 2005.
- 2) ProFUSION25, ViewPLUS, <http://www.viewplus.co.jp/products/profusion25/index.html>
- 3) E. H. Adelson and J.Y.A. Wang, Single Lens Stereo with a Plenoptic Camera, IEEE Trans. PAMI, Vol14, No.2, pp. 99-106, 1992.
- 4) R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan, Light Field Photography with a Hand-Held Plenoptic Camera, Stanford Tech Report CTSR 2005-02, 2005.
- 5) N. Kusumoto, S. Hiura and K. Sato, Uncalibrated Synthetic Aperture Photography for Defocus Control, Journal of the Institute of Image Information and Television Engineers, Vol.63, No.6, pp. 857-865, 2009.
- 6) A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing, ACM Trans. Graph. (SIGGRAPH), Vol. 26, No. 3, 2007.
- 7) R. Raskar, A. Agrawal and J. Tumblin, "Coded Exposure Photography: Motion Deblurring using Fluttered Shutter", SIGGRAPH2006.
- 8) S. Hiura and T. Matsuyama, "Depth Measurement by the Multi-Focus Camera", Proc. of CVPR'98, 953-959.
- 9) 笹尾朋貴, 日浦慎作, 佐藤宏介, 超解像のための画素形状のコード化に関する研究, MIRU2010 画像の認識・理解シンポジウム論文集, pp. 1701-1708, 2010.
- 10) H. Nagahara, S. Kuthirummal, C. Zhou and S. Nayar: Flexible Depth of Field Photography, European Conference on Computer Vision, 2008.
- 11) A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman: Motion-Invariant Photography, SIGGRAPH, ACM Transactions on Graphics, 2008.
- 12) 日浦慎作, コンピュータショナルフォトグラフィ理解のための光学系入門, 情報処理学会研究報告, Vol.2010-CVIM-171, No.13, pp. 1-12, 2010.
- 13) 長原一, 符号化撮像, 情報処理学会研究報告, Vol.2010-CVIM-171, No.14, pp. 1-9, 2010.