

## 非接触 3次元計測技術とライトフィールドカメラ -カメラにおける測距・AF との関係-

蚊野 浩<sup>†</sup><sup>†</sup>京都産業大学コンピュータ理工学部

### Non-Contact 3D Measurement Technologies and Light Field Camera - Relationship to Ranging for Auto Focus -

Hiroshi KANO<sup>†</sup><sup>†</sup>Kyoto Sangyo University, Faculty of Computer Science and Engineering

#### 1. はじめに

デジタルカメラにはさまざまな画像処理技術が利用されている。代表的なものは、色補正やノイズ除去などの画像信号処理技術、電子式手ぶれ補正や高ダイナミックレンジ画像生成のための画像計測技術、顔検出やシーン判別のための画像認識技術である。これら以外にもあらゆる画像処理技術が搭載されており、デジタルカメラは視覚情報処理専用のコンピュータといっても過言ではない<sup>1)</sup>。このようなデジタルカメラの重要な機能の一つに、被写体までの距離を高速・正確に計測し、ピントを合わせる測距・AF機能がある。

数年前に製品化され注目を集めたライトフィールドカメラは、カメラ装置におけるデジタル技術の有効性をさらに明確にした。この装置は、カメラに入射する光線集合(ライトフィールド)をデジタルデータとして記録し、光線集合に対する演算によって写真画像を生成するカメラである。例えば、撮影後の後処理によって任意の位置にピントを合わせた画像を生成することが可能である<sup>2),3)</sup>。ライトフィールドカメラは、撮影時に被写体にピントを合わせる必要がないカメラとして有名である。この解説では、3次元空間における光線の状態を記録するライトフィールドカメラの原理と、通常のデジタルカメラにおける測距・オートフォーカス(AF)の関係を、非接触3次元計測技術を仲介として説明する。

#### 2. 3次元計測技術と測距・AF

カメラ的な装置を用いることで、被測定物に接触することなく3次元空間や実物体の位置・形状を測定する技術を非接触3次元計測技術とよぶ。図1にその分類を示す。ここで、

能動法は何らかのエネルギーを被写体に照射した観察結果から位置・形状を測定し、受動法はあるがままの被写体の像から測定する。また、ステレオ法は三角測量に基づく方法、Time-of-Flightは照射エネルギーの反射時間に基づく方法、Shape from Xは被写体像の鮮鋭度や陰影、カメラを動かした時の像の動きなど、単眼カメラの画像を手がかりにする方法である。

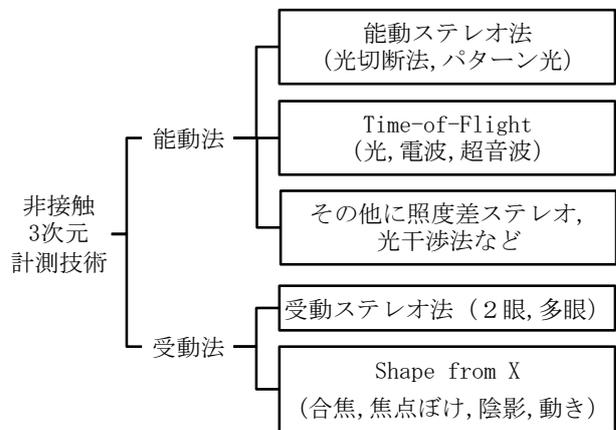


図1 非接触3次元計測技術

Fig.1 Non-contact 3D measurement technologies

能動法は比較的安定に動作するが、装置のコストが高い。実用化されている例として、工業製品の形状入力に用いる3Dスキャナ、自動車の障害物検知に用いるミリ波レーダーやレーザースキャナなどの車載3Dセンサ、ゲーム機用のヒューマンインタフェースデバイスがある。いずれの方式も技術開発が急速に進んでおり、利用可能な分野が拡大している。例えば、ゲーム機用のデバイスであるKinectは能動ステレオ

法を原理としているが、同じ原理で動作する3D スキャナと比較して画期的に安価である。しかしそれでも、デジタルカメラに組込むことはコスト的に難しい。

過去には、能動法による3次元計測技術をAFに利用したカメラが存在した。コンパクトフィルムカメラの多くは、被写体に赤外LEDのスポット光を投影し、その反射像をPSD (Position Sensitive Device) とよばれるアナログセンサで位置検出する方式を用いた。超音波の反射時間から距離を測定する超音波センサを用いたカメラもある。いずれも、画角中央の1点までの距離を測定した。

次に、図2を用いて受動ステレオ法の原理を説明する。焦点距離 $f$ の2台のカメラを、水平方向に基線長 $B$ の間隔で設置する。この時、被写体の2枚の投影像には、距離 $z$ に応じた位置ずれ $d$ が生じる。この位置ずれ量を視差とよび、 $z=Bf/d$ の関係がある。したがって、被写体距離は視差に反比例し、視差がゼロであれば被写体は無限遠に存在する。受動ステレオ法では、図2のように撮影した2枚の画像から被写体の対応点を推定し、被写体の位置・形状を自動的に計算する。

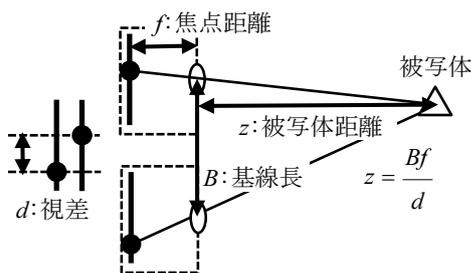


図2 受動ステレオ法

Fig.2 Passive stereo method

AFの測距に受動ステレオ法を用いるデジタルカメラも存在するが、より興味深いのは、ライカなどが開発したレンジファインダーカメラである。これは、撮影レンズとは別に測距用に2つの窓を持つカメラである。それぞれの窓から得る光像を可動プリズムで平行移動させた後ハーフミラーで重畳し、目視確認できる構造になっている。さらに、光像の平行移動量とピント合わせのレンズ繰り出し量が機械的に連動している。したがって、ピント調整リングを操作する事でレンズの撮影距離を設定すると同時に、その撮影距離にある被写体像が完全に重なり、二重像が解消される。カメラマンはファインダ内の被写体像が鮮明になったことを確認すると、即座にシャッターを切ることができる。

受動ステレオ法で、3台以上のカメラを用いるものを多眼ステレオとよぶ。多眼ステレオではマルチベースライン法<sup>4)</sup>などのアルゴリズムを用いることで、測定精度が向上し誤計測が減少する。多眼ステレオの極限として、図3左のように、小型カメラを縦横に規則的に配置した装置を考

える。このような装置をカメラアレイ<sup>5)</sup>とよぶ。カメラアレイも多眼ステレオであるから、受動ステレオの原理を用いて被写体までの距離・形状を測定することが可能である。一方、カメラをピンホールカメラでモデル化すると、図3右のように、それぞれのカメラが撮影する写真画像をピンホールの位置を通過する2次元構造の光線集合に対応させることができる。そしてカメラアレイが記録する光線集合が、ピンホールに対応する2次元光線集合を、さらに2次元サンプリングしたものであることが理解できる。したがって、カメラアレイは光線空間を4次元サンプリングする装置であるともいえる。カメラアレイは、後で述べるライトフィールドカメラの一つの実装例になっている。

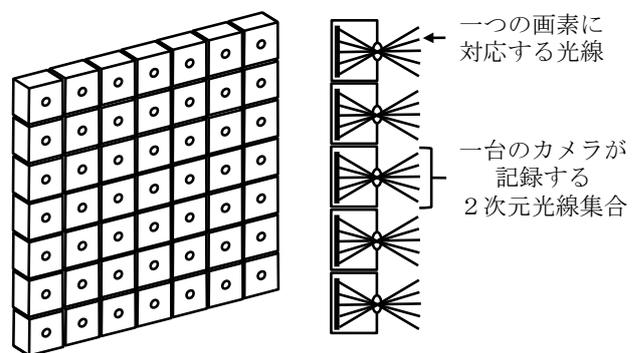


図3 カメラアレイと光線空間の標準化

Fig.3 Camera array and sampling of ray space

受動法におけるShape from Xは、単眼で撮影した画像から被写体までの距離を測定する。被写体に合焦しておれば、レンズの公式から被写体までの距離を計算できる。像がぼけている状態であっても、撮影距離が異なる2枚のぼけ画像から、被写体までの距離を計算することができる。この方法をShape from Defocusとよぶ。また、一台のカメラを移動させて被写体像を複数回撮影し、複数の被写体像の位置関係を正しく対応づけることができれば、カメラの移動・回転量と被写体の位置・形状を同時に計算することができる。この技術をShape from Motionとよぶ。現在では、Shape from Motionの技術を大規模に用いることで、市街地全体の詳細な立体形状を計算することなどが可能である。

撮影した画像から被写体像の合焦判定を行う技術は、デジタルカメラのコントラストAFに用いられる技術である。また最近では、Shape from DefocusをAFのための測距に用いるカメラがある<sup>6)</sup>。

### 3. コンピューショナルフォトグラフィ

デジタル技術を駆使した写真撮影技術であって、単なる画像処理にとどまらず、レンズ・撮像素子・画像処理のシステムを再構築し、従来不可能であった撮影手法を可能とする

技術分野をコンピュータショナルフォトグラフィ (CP) とよぶ<sup>7)</sup>。CP を実用化した装置として良く知られているものにライトフィールドカメラがある。これについては4章で詳しく説明し、ここでは、他の手法を簡単に紹介する。

CP の中で、比較的、古くから知られているものに Wavefront コーディング<sup>8),9)</sup>がある。これは、図4に示すように、撮像レンズの一部に位相板とよばれる光学素子を挿入する。この位相板の働きによって、レンズを通過した光線はカメラの焦点位置近傍で像のぼけが一定の状態を長く保つようになる。この像をデコンボリューション (画像復元) することで、シャープな像を得る。この手法によって、通常よりもはるかに被写界深度が深い写真撮影が可能になる。

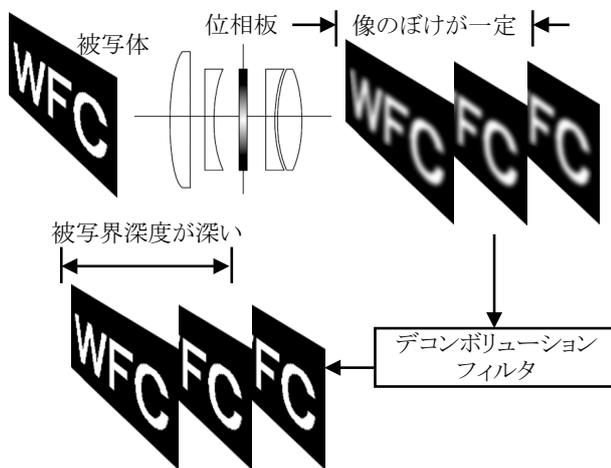


図4 Wavefront コーディングによる被写界深度の拡大

Fig.4 Extended depth of field by Wavefront coding

写真撮影における失敗の多くは、ピンぼけと手ぶれである。Wavefront コーディングはピンぼけを防ぐ技術である。一方、手ぶれを防ぐための CP 手法として符号化露光<sup>10)</sup>が知られている。最も性質が悪いぶれは、撮像面が等速で移動するぶれである。これによって、画像中の高周波成分が急速に失われる。逆に性質が良いぶれというものがある。これは、チャープ信号のように高周波から低周波までを含むぶれである。実際にこのような機械的な動きを撮像素子に加えるのではなく、ぶれが発生する撮影期間中の露光をチャープ信号などで変調する、という考え方である。このように露光を変調して撮影すると、一見、手ぶれ写真に見えるが、実は、元の信号が十分に残っている。そこで、画像復元を行うと、きれいにぼけが修復される。

#### 4. ライトフィールドカメラ

ライトフィールドカメラは3次元空間中の密な光線集合 (ライトフィールド) を取得し、光線集合に対する一種の画像処理によって最終的な写真画像を生成する。ここでの画像

処理は、概念的には、光線に対するレンズの働きをシミュレーションすることである。ライトフィールドを取得する方法として、図3のカメラアレイを用いる方法と、次に説明するマイクロレンズアレイを用いる方法がある。本章では、市販されているライトフィールドカメラである Lytro の構造・機能を説明することで、ライトフィールドカメラの動作原理を明らかにする。

通常のコンパクトデジタルカメラと比較すると、Lytro の構造上の特徴は撮像素子の直前にマイクロレンズアレイが配置されていることである。図5に Lytro の撮像素子部の拡大写真を示す。撮像素子はデジタルカメラ用 CMOS センサであり、その画素ピッチは  $1.4\mu\text{m}$  である。撮像素子表面をハニカム構造のマイクロレンズが覆っている。そのピッチは  $14\mu\text{m}$  である。撮像素子の  $3280 \times 3280$  画素の領域が  $330 \times 380$  個のレンズによってカバーされており、一つのマイクロレンズの直径は、おおむね 10 画素分である。

マイクロレンズは保護ガラスと一体化しており、撮像素子上に若干の空間をおいて配置されている。保護ガラスの上面側が平坦で、底面側がマイクロレンズアレイになっている。保護ガラスとマイクロレンズアレイの厚さは約  $430\mu\text{m}$  である。また、CMOS センサ表面にはベイヤー型のカラーフィルタアレイが配置されている。

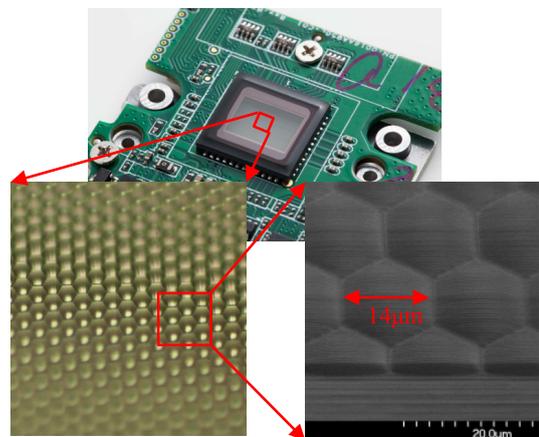


図5 Lytro の撮像素子部の拡大写真

(写真: 日経 BP 社, 豊通エレクトロニクス ヴァン・パートナーズ)

Fig.5 Closeup picture of image sensor portion

図6に Lytro が取得する生画像の一部とその部分拡大を示す。生画像にはマイクロレンズアレイの構造が強く反映されている。拡大画像に見られる円構造が一つのマイクロレンズに対応する。その直径が約 10 画素であることを、この画像からも確認することができる。円構造の中のそれぞれの画素が、後述するように、一本の光線に対応する。マイクロレンズの境界付近に位置する画素は暗くなっており、その領域では光線の取得が難しいことを示す。

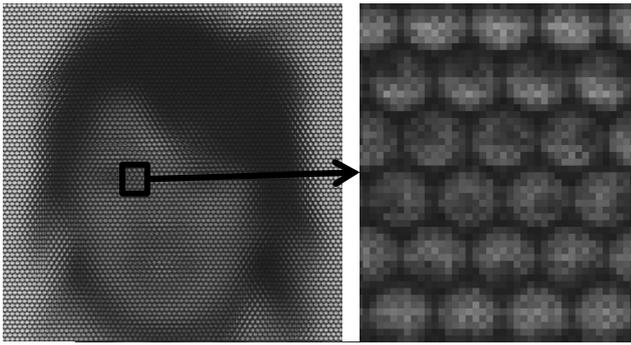


図6 Lytro が取得する生画像の一部とその部分拡大

Fig.6 Closeup picture of acquired image

撮像素子部と主レンズによってライトフィールドを取得する様子を図7に示す。図において、Aの位置にある被写体が主レンズによってマイクロレンズアレイの位置に焦点を結ぶとする。このとき、被写体から発してマイクロレンズに到達した光線は、その下に位置する画素によって方向と明るさが記録される。一つのマイクロレンズに対応する全ての画素を平均化することで、Aの位置にピントを合わせた粗い写真画像が生成される。これは、1の位置に撮像素子を置いた画像を計算したことに相当する。

図7でBの位置にある灰色マークに注目すると、この位置を通過する3本の光線は、異なるマイクロレンズを介して撮像素子に記録される。それらの画素値を平均化すると、Bにピントを合わせた写真画像の画素を生成できる。これは、Bの位置にピントを合わせるとAの像がぼけるという現象を、ライトフィールドを用いて計算したことに相当する。このとき、計算上の撮像面は2の位置になる。同様の手順で任意の距離にピントを合わせた画像を計算することが可能であり、この機能をリフォーカスとよぶ。

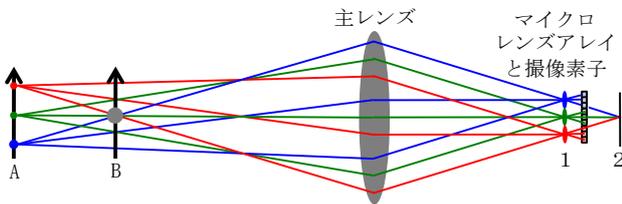


図7 ライトフィールドの取得とリフォーカス

Fig.7 Acquisition of light field and refocus

図7のマイクロレンズがカバーする3つの画素で、最も下に位置する画素に対応する光線を主レンズまで追跡すると、主レンズの上部に達する。したがって、これらの画素だけを再配列して生成される画像は、主レンズ上部を通過した光線による像である。同様に、マイクロレンズがカバーする画素で、最も上に位置する画素だけを再配列して生成される画像

は、主レンズ下部を通過した光線による像である。このようにマイクロレンズがN個の画素をカバーするとき、マイクロレンズに対して同じ位置にある画素だけを再配列してできるN個の小画像は、主レンズをN個の部分に分割して取得されるN個のステレオ画像群を形成する。Lytroは一つのマイクロレンズが、75画素(=3080×3080÷330÷380)をカバーする。したがって、撮像素子部は75画素の超微小カメラを330×380個配列したカメラアレイのように見える。しかし、上で述べたマイクロレンズと主レンズの働きにより、実際には、330×380画素の小画像を撮影する微小カメラが、主レンズ口径上に75個配置されたカメラアレイと考えるのが適切である。

カメラアレイで撮影した画像群からリフォーカス画像を生成する手順を説明する。まず、入力された画像上で、ピントを合わせる被写体を指示する。次いで、多眼受動ステレオ法のアルゴリズムを用いてその被写体の視差を計算する。以下、図8に示すように、画像群を視差に応じて平行移動し、重ね合わせ、全ての画像を平均化することでリフォーカス画像を得る。視差が大きければ手前にピントが合い、視差が小さければ奥にピントが合ったリフォーカス画像が生成される。

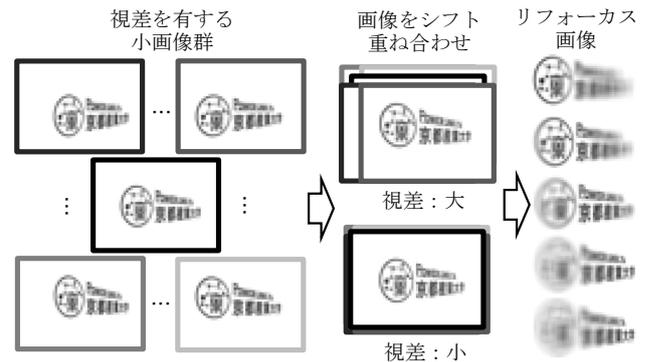


図8 カメラアレイによるリフォーカス

Fig.8 Refocus using camera array

ライトフィールドカメラにおけるリフォーカス画像の生成をこのように説明すると、そこで行われている処理が、2章で説明したレンジファインダーカメラにおける測距の手順と同じであることに気がつく。レンジファインダーカメラは、光学機械と人間の合焦判断によってピント合わせが容易なカメラを実現したが、ライトフィールドカメラは、光学部品とコンピュータによってピント合わせが不要なカメラを実現したと言える。

## 5. 位相差 AF とライトフィールドカメラ

ライトフィールドカメラの説明で明らかのように、単眼カメラを使ったワンショットの撮影であっても、主レンズに入射する光線を分離することで、被写体の位置・形状を測定することが可能である。これは位相差 AF と単眼立体カメラ<sup>11)</sup>

にも利用されていた考え方である。ただし、これらの装置では、ライトフィールドカメラほどに細かく光線を分離していない。主レンズ全体に入射する光束を左半分と右半分に分離する程度のことである。

図9に位相差AFの原理図を示す。主レンズの両端付近を通過する光線をセパレータレンズで分離し、それぞれの像をラインセンサで検出する。被写体までの距離によって像の位置に違いが生じるので、位置の差を距離に換算することができる。カメラ業界ではこの方式を位相差方式とよんでいるが、主レンズ口径を基線長とする2眼受動ステレオのバリエーションということもできる。

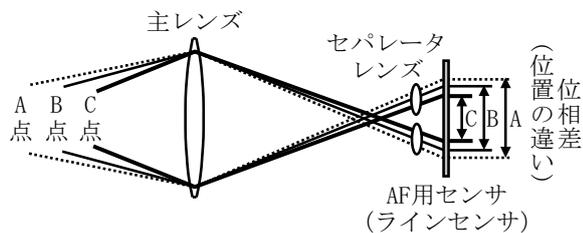


図9 位相差AF

Fig.9 Phase Detection AF

位相差AFのセンサ部は、一眼レフカメラのミラーボックス下部に配置することが普通である。そこに像を導くためにメインミラーの一部をハーフミラーとし、透過した光線をサブミラーで下部に導く。このような構成のため、メインミラーが跳ね上がるライブビューでは、位相差AFを動作させることができない。また、一眼レフ以外のカメラに位相差AFを搭載することも難しい。

位相差AFの課題を解決する一つの方法として、近年、像面位相差AFが実用化されている。その基本的な考え方は、位相差AF用のセンサを通常の撮像素子に埋め込むことである。像面位相差AFの実現方法にはいくつかのバリエーションがあるが、キャノン社が開発したディアルピクセルCMOS AFの構成は、図7に示すライトフィールドカメラモデルにおいて、一つのマイクロレンズがカバーする画素を2個にした場合と全く同じである。マイクロレンズに対して左側に位置する画素が作る像は主レンズの右半分を通過する光束による像、右側に位置する画素が作る像は主レンズの左半分を通過する光束による像になる。したがって主レンズの口径を基線長とする2枚のステレオ画像となり、これを用いて画像中の全ての被写体までの距離を測定することが可能になる。

## 6. おわりに

この解説では、非接触3次元計測技術とカメラに実用化されているAF・測距技術、およびライトフィールドカメラの関係について説明した。いずれの技術においても受動ステレ

オ法が重要な役割を果たしていた。視差画像の撮影と受動ステレオ法を基盤技術として、技術の進歩に応じて、実現方法や構成に非常に巧みな改良が加えられてきたことがわかる。

ところで、ライトフィールドカメラはあまりに未来的なカメラであると考えられてきた。また、その有用性には若干の疑問が持たれている。しかし、最新の像面位相差技術がライトフィールドカメラの要素技術と共通しているのを見ると、ライトフィールドカメラは、やはりカメラの未来像の一つを予言しているように思える。

## 参考文献

- 1) 蚊野 浩監修, 映像情報メディア学会編:“デジカメの画像処理”, オーム社, (2011).
- 2) R. Ng: “Digital light field photography”, Dissertation of Stanford University, (July 2006).
- 3) 蚊野 浩:“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, 「光の日」公開シンポジウム講演予稿集, pp.1-6, (2013).
- 4) 奥富 正敏, 金出 武雄: “複数の基線長を利用したステレオマッチング”, 電子情報通信学会誌, Vol. J75-D-II, No. 8, pp.1317-1327 (1992).
- 5) B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz and M. Levoy: “High performance imaging using large camera arrays”, ACM Trans. Graphics, Vol. 24, No. 3, pp.767-776 (July 2005).
- 6) 河村 岳, 大原 正満: “単眼カメラで空間認識, 「ぼけ」から距離を推定”, 日経エレクトロニクス, No. 1137, pp.59-67 (2014).
- 7) 日浦 慎作: “コンピュータショナルフォトグラフィ”, 電子情報通信学会誌, Vol. 95, No. 9, pp.823-828 (2012).
- 8) E. R. Dowski and W. T. Cathey: “Extended depth of field through wave-front coding”, Applied Optics, Vol. 34, Issue 11, pp.1859-1866 (1995).
- 9) 小松 進一: “Wavefront coding による被写界深度の拡大”, 情報処理学会研究報告, Vol. 2010-CVIM-174, No. 5, pp.1-8 (2010).
- 10) 西 一樹: “符号化露光法と超解像”, 映像情報メディア学会誌, Vol. 67, No. 8, pp.655-660 (2013).
- 11) 寺西 信一監修, 電子情報通信学会編: “画像入力とカメラ”, オーム社, (2012).



蚊野 浩

1984 年京都大学大学院情報工学専攻修了。同年三洋電機株式会社入社。カラー記録装置、画像処理技術、コンピュータビジョン技術などの技術開発に従事。2010 年京都産業大学コンピュータ理工学部教授。工学博士。