

CVCV-WG特別報告：コンピュータビジョンにおける技術評論と将来展望 (VII)

— コンピュータビジョンのための画像センサ —

佐藤 宏介

奈良先端科学技術大学院大学 情報科学研究科

コンピュータビジョンにおける画像センサはこれまで、認識の対象となる三次元シーンをピンホールカメラモデルで二次元射影し、その投影像に対して空間サンプリングと濃淡値の量子化を行って得られる二次元デジタルデータを出力するブラックボックスとして扱われてきた。しかし、工学としてのコンピュータビジョンを正しく理解するためには、画像センサの原理やその限界を知るにしくはない。本稿では、放送用画像センサとの設計思想の違いや、コンピュータビジョンの主要アルゴリズムの動作を左右する空間分解能や濃淡分解能、ノイズレベル等について概観し、コンピュータビジョン用画像センサの将来を展望する。

CVCV-WG Special Report: Technical Review and View in Computer Vision (VII)

- Image Sensors for Computer Vision -

Kosuke SATO

Graduate School of Information Science
Nara Institute of Science and Technology

So far, only few attempts have so far been made at image sensors in Computer Vision researches. Although a novel image sensor has been regarded as the Pin-hole camera model with ideal spatial sampling and quantization of intensity, practical ones have many kinds of limitations, ex. poor resolution, non-linearity on intensity, lens blur, color coherency electric noises, and so on. In this report, the principles of practical image sensors and their opto-electronic characteristics are reviewed by comparing broadcasting TV cameras, so that researchers could solve practical problems in Computer Vision successfully.

1. はじめに

広義のコンピュータビジョンは、人間や高等動物の持つ高度な視覚機能を心理学的観点からではなく情報学的観点から解明することと、それと同じ視覚機能をコンピュータを用いて計算機工学的に実現することを目的とした学問分野である。それに対して狭義のコンピュータビジョンは、立体物の記述や認識、それに加えて三次元シーンの理解が目的とされる。画像データも、日常の室内シーンや屋外シーンを対象とすることがほとんどである。本稿では、狭義の意味で用いることにする。

残念ながらこれまでのコンピュータビジョン研究では、投影変換における不変量や立体幾何の記述など数学的なアプローチが主であり、最終的な目的である3次元シーンの理解を目指して、そのプロセス全体の入力となるべき画像データの質やそのセンシングに対しての議論はほとんどなされてこなかった。それは、画像センシングがもっぱら計測工学やテレビジョン工学、画像通信工学において研究されてきたという歴史のためであろう。

だからといって、コンピュータビジョンにおいて画像センサを二次元デジタルデータを出力するブラックボックスとして扱うべきではなく、正しくその原理や限界を知るべきである[1]。現実の画像センサの持つ空間分解能や濃淡分解能、ノイズレベルといったものが、領域分割やステレオ対応、Shape from Shadingといったコンピュータビジョンの主要な手法の死命を決するからである。

本稿では、既存の放送用画像センサとの差異や、その性能について概観し、コンピュータビジョン用画像センサの将来を展望する。

2. 人の視覚とテレビ放送用画像センサ

電子的な光電センサと異なり、バイオセンサである人の視覚には独特な特性がある[2]。まず第一に、他の感覚器と同じように、光エネルギーの物理強度と知覚強度との関係は直線ではなく、対数特性となる。第二に、約60Hz以上のフレームレートがあればフリッカを感じずに（周囲が明るい場合）、連続現象として知覚してしまう残像現象がある。第三に、どこまで細かい模様が知覚できるかという空間分解能がスペクトルによって一定ではない特性がある。白色の時が最も細かく見え、赤や青の細かい模様は変化がないように見える。その他、網膜にはマッハ効果と呼ばれるエッジ強調効果があり、微妙なステップエッジが強調

され知覚される。

放送用など肉眼による実見を目的とした画像センサは、人間の視覚特性を巧妙に利用し、より簡便な技術で実現できるよう数多くの工夫がなされている[3]。例えば、アメリカのテレビジョン方式委員会NTSC (National Television System Committee) のカラー放送規格では、色毎に分解能を変化させ、明暗信号に帯域幅4.2MHz (テレビ放送における帯域幅は静止画の横画素数に相当する。4.2MHzは横方向有効画素数にして約447となる)、オレンジ-シアン系I信号に1.5MHz (約160画素)、黄緑-マゼンタ系Q信号0.5MHz (約53画素) を割り当てている。人には見えない細かい黄緑やマゼンタの模様をいくら頑張っても撮像をし、それを放送しても意味がないのである。

このためテレビ放送用画像センサは、緑系を細かく撮像できるような工夫がされている。モザイク型カラーフィルタによる単板RGB固体撮像素子では、総画素の半分をGに、1/4ずつをRとGに割いている。3板タイプのもので、RGBにそれぞれ各1枚ではなく、2板をGに1枚を赤と青共用にしているものさえある。

さらにNTSC規格では、低いフレームレートでも像の早い動きを伝えるために、走査線を飛び越して走査し、一つのフレームを奇数フィールドと偶数フィールドに分けて伝送するというインタレース走査が行われている。これも、早い動きは認知できても、細かい形状は認知できないという視覚特性を利用している。しかし、奇数走査線と偶数走査線ではサンプリング時刻が異なり、動画像解析にはまったくやっかいである。

コンピュータ画像では、画素という概念で分解能を表し画素の縦横比は1である。しかし、テレビ放送では垂直方向には走査線、水平方向には映像信号という概念で設計され、分解能が等方性ではない。このため撮像素子、特に固体撮像素子の受光セルは正方格子状に配列されておらず、色搬送波周波数 (3.57MHz) の都合から水平方向のセル数が決定され、縦横比が1とは通常ならない。

その他、見かけの映像をくっきり見せるために、高域強調フィルタを映像信号出力部に入れ、エッジを強調しているカメラが、民生用や監視用として発売されている。しかもこれは縦方向のエッジのみに作用するので、エッジ方向によって鮮鋭度が異なるという問題も引き起こす。

表1 放送用カメラと理想カメラとの差異

	放送用カメラ	理想カメラ
フレームレート	30Hz	任意
走査	飛び越し走査	無走査
感度特性	ガンマ特性	直線
空間分解能	白色は高分解能 有色は低分解能	スペクトル 非依存
空間周波数特性	高域強調	フラット

以上の点などから、テレビ放送用に設計された画像センサをコンピュータビジョン用に流用する際には細心の注意を怠ってはならない。表1に放送用カメラと理想カメラとの差異を列挙する。

3. 画像センサ開発の歴史

画像入力には銀塩写真の発明から今日のハイビジョンカメラまで長い技術史を持っている。ここでは現状の画像センサを正しく理解するため、簡単にその歴史を振り返ることにする[4-8]。

3.1 点センサと二次元機械式走査

・銀塩写真のドラムスキャナ入力 (図1(a))

感度、空間分解能いずれを観点からも、銀塩写真(写真フィルム、乾板)は最も優れた画像センサである。しかし、入射光は感材の化学変化をもたらすだけで2次元の電子情報とはなりえない。そこで利用されるのが、ファクシミリから発展したドラムスキャナである。写真フィルムや印画紙を円筒ドラムに巻き付け回転させ(主走査)、ピンポイントの光量を計測するホトセンサを回転軸方向に平行移動させる(副走査)。ホトセンサの出力をAD変換すれば、ラスタ走査されたデータ列が得られる。

ドラムスキャナは画像入力の最も初期のものであるが、幾何学歪みがない、輝度むらがないなどの利点があり、印刷用デジタルデータの入力として今日でも重要性は高い。

・スキャニングミラー方式 (図1(b))

三次元シーンを直接計測する。一つのホトセンサに機械式の偏向ミラー系を組み合わせたもので、リモートセンシングなど物理計測の分野では今日でも重要である。しかし、機械走査のむらによりジッタと呼ばれる走査線ずれが発生しやすく、計測時間が長い点がコンピュータビジョン用センサ

としては不向きである。

3.2 点センサと二次元電子式走査

・フライングスポットスキャナ (FSS、図1(c))

ドラムスキャナと同じように、銀塩写真から画像を入力するものであるが、写真、ホトセンサとも固定されており、光源を電子的に走査させることで画像を得る。まず暗箱の中に写真を入れ、CRT管面上で輝点をラスタ走査させる。管面上の輝点はレンズ系で写真上に結像され、スポット光源となる。ホトセンサは写真全体を観測しているが、輝度情報はこのスポット光源上の点のものが得られる。

可動部がないことやビデオレートで画像が得られる点、撮像管よりノイズが少ない等の利点が多くあり、放送スタジオでのテロップ撮影にしばしば用いられてきた。

・比例型撮像管

これまで紹介してきた撮像デバイスの受光部の多くは、光電子増倍管である。光電子増倍管は感度が高くしかも幅広いダイナミックレンジを持つ優れた点計測ホトセンサである。この光電子増倍管の中に電子的にラスタ走査するデバイスを封入し組み合わせたものが、イメージディセクタ撮像管である。三次元シーンをレンズ系で光電面に結像させ得られる電子像を電磁偏向系で平行移動させる。

・蓄積型撮像管

3次元シーンを直接電子的に撮像するデバイスとして撮像管は古い歴史を持つ。その中でも、光の蓄積効果の応用は、撮像管の感度を飛躍的に向上させる画期的なものであった。その代表であるイメージオルシコンは、光電面からの電子像を高絶縁膜に叩き込み二次電子を放出させることで、光量と露光時間を乗じたものに比例した正電荷を蓄積してゆく。この正電荷は高絶縁膜上で像をなし、電磁偏向系でラスタ走査された電子ビームによって読み出される。

安価な撮像管として知られるビジコン (図2(a))は焦電効果を利用したものであるが、焦電膜内で蓄積効果をなすため感度は十分実用的である。

撮像管はいずれも電磁偏向系を必要とするので、幾何学的な非線形性(注:歪曲収差など、本来正方形の像が糸巻形や樽形に見えたりする)は避けられない。

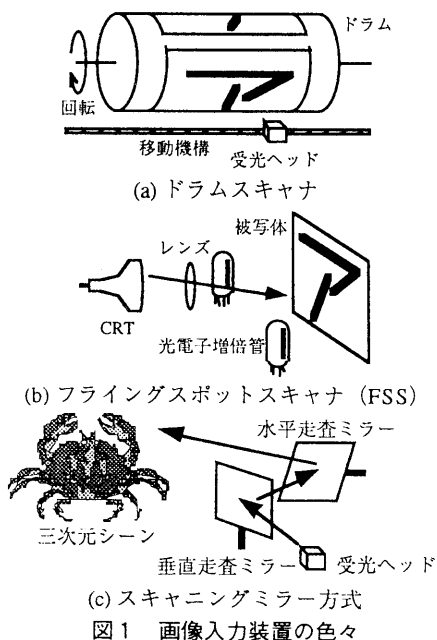
3. 3 二次元センサと二次元電子走査 ・固体撮像素子(蓄積型) [9-11]

ホトダイオードは光子の数に比例した電子・正孔対を発生する比例型センサであるが、コンデンサを並列接続すれば蓄積型ホトセンサとなる。このホトダイオードをチップ上に多数並列したものが固体撮像素子である。しかし、すべてのホトダイオードから出力線を引き出すことは不可能であるため(注:完全並列読出型は、16×16画素のものが商用で最高である)、蓄積電荷の二次元電子走査による読み出しが必要になる。CMOSマトリクススイッチによるMOS型とCCD電荷転送素子を用いたCCD型(図2(b))の二種がある(注:固体撮像素子一般をCCDと略称するのは正しくない)。

小型軽量で堅牢、低電圧で動作するなどの利点があり、しかもホトダイオードが正確に二次元配列されているため幾何学歪が原理上ない点が重視され、コンピュータビジョン用センサとして隆盛を極めていく。しかし、各画素の感度がばらつく、熱雑音が多いなどの欠点も多々ある。

4. 現実の画像センサの問題点

コンピュータビジョン研究ではカメラを理想ピンホールカメラモデルとすることが多い[12]。同次座標系表現を行えば、透視変換と平行移動を線



形表現できるからである。しかしながら、現実の画像センサはピンホールカメラモデルとは程遠いものである。

<理想ピンホールカメラモデル>

- ・幾何学的な歪はいっさいなし
- ・結像面上一点の光量をセンサする
- ・画素の感度は一定で線形
- ・出力にノイズはなく、飽和もしない
- ・正方格子で同時刻サンプリングする
- ・分光感度はフラットである

以下では、現実の画像センサに見られる各種アーチファクトを紹介する。

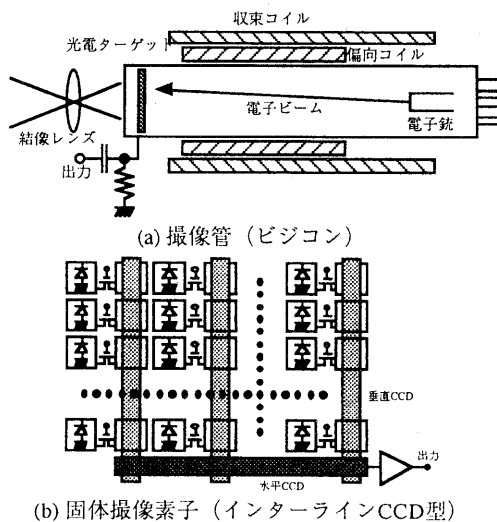
4. 1 レンズ系によるアーチファクト

光学レンズで結像させた場合、レンズが有限の厚みと球面(注:非球面レンズは作製困難)を持つため、ピンホールカメラでは現れなかった各種アーチファクトが出現する。最もコンピュータビジョンにとって不適切なものは被写界深度(ピントの合焦範囲)の制約による像のぼけ、つまりピントぼけである。その他に、レンズの各種収差が像の質を劣化させる。その中でも歪曲収差(図3(a))については事前較正を行えば補正することができる[13]。

<レンズのアーチファクト>

- 歪曲収差:幾何学的に像が歪む
- 色収差:赤と青ではピント位置が異なる
- 球面収差、非点収差:焦点位置でも像がぼける
- シェーディング:像の周辺で暗くなる

図3(b)に示すシェーディングは、像の中心部と周辺部で輝度が一樣にならない現象である。撮



像素子のホトセルが有限の面積を有しているため、光電流はその面積に照射される光量の積分となる。したがって、周辺の画素ほど入射光が傾斜し、ホトセル内に降り注ぐ光エネルギーが減少してしまう。つまりシェーディングは画角が広角ほど出現する（図3(c)）。

コンピュータビジョンではカメラに正対する明るさが一定の平面を撮像した場合に、画素値も一定になるとの仮定を置くことがほとんどであるが、これは焦点面上の理想点でサンプリングした場合のみに成立し、現実の撮像デバイスではシェーディングを必ず起こす。

また、ピント位置によってレンズ長が変化したり、ズームレンズでズームするとピント位置がずれる等、現実のレンズ系には問題が多い。

4. 2 受光素子によるアーチファクト

固体撮像素子の受光部であるシリコンホトダイオードには、PN接合の膜厚や保護膜の作りなどで特性が大きく変わる問題点がある。さらに、熱雑音や暗電流（注：暗黒でも光電流が流れ、出

力が零にならない現象。図4(b))等の問題も多い。常温のシリコンは熱雑音が多く、通常7ビット程度しかダイナミックレンジが得られず、8ビット画像と言われるものでもそのLSBの多くは有意ではない。

また、テレビカメラは、図4(a)に示すように人間の視覚特性に合わせてわざと受光特性を直線ではないように設定しているため、コンピュータビジョン研究には必ずガンマ値設定があるカメラを用い、その設定値を1（直線）にすることを忘れてはならない。

4. 3 AD変換器によるアーチファクト

ホトダイオードが正しく光電変換したとしても、それをデジタル値に変換するAD変換器の性能が悪ければ意味がない。今日のAD変換器（注：ビデオ用には高速フラッシュ型が通常用いる）ではそれ自身で非直線性やオフセットといった問題点は少なく、前置されるビデオアンプの非直線性やオフセットの方が問題となる。特に、暗黒が零とならないオフセットずれがしばしば発生するので、注意が必要である。

4. 4 サンプリングによるアーチファクト

理想カメラでは、結像面を正方の直交格子を被せ、ある時刻で一斉に空間サンプリングする。しかし現実の画像センサを用いた場合には、ある時刻で一斉という仮定も、正方直交格子という仮定のいずれをも正しくないことがある。

撮像管では、電子ビームを垂直と水平の鋸波信号で駆動するため、走査線は水平ではなく、走査線の左端と右端では走査線一本分の傾きがある。このように像が菱形に横ずれする現象をスキュー

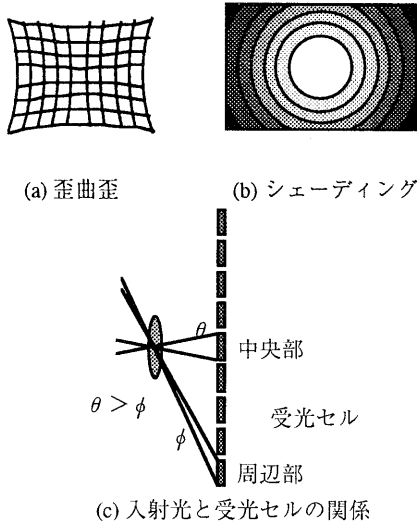


図3 現実の結像レンズ

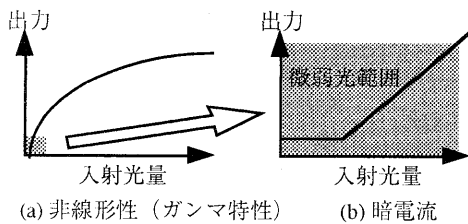


図4 現実の感度特性

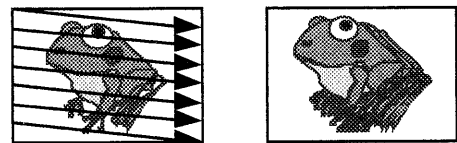


図5 現実の画像センサにおけるスキュー歪

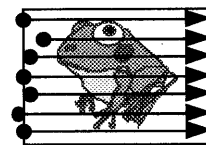


図6 ジッタ

歪（図5(a)）という。さらに撮像管では、画面の左上部と右下部ではサンプリング時刻が異なっている。このため動きのある物体を撮像した場合にも、画像が歪んでしまう（図5(b)）。このスキュー歪は点計測をラスタ走査するという画像センサ共通のものであり、ドラムスカナなどでも生じる。

さらにドラムスカナなど機械式画像センサには、走査線の左端の位置がずれるというジッタと呼ばれる歪みが生じることがある（図6）。ジッタには、画像センサの同期信号とフレームメモリの同期信号が同期していない場合に、±0.5画素分だけ画素位置がずれるものもある。

4.5 蓄積によるアーチファクト

画像センサには、入射光量に比例した出力を出す比例型センサ（ラムスカナ、フライングスポットスカ、イメージディセクタ撮像管など）と、露光量（＝入射光量×露光時間。通常フレーム時間が露光時間となる）を出力する蓄積型センサ（銀塩写真、撮像管一般、固体撮像素子など）がある。このため、蓄積型センサはある瞬間の時刻で入射光量をサンプリングした値を出力するのではないため、動きによって像がぼやけるモーションブラ（Motion Blur）が発生する。モーションブラを少なくするためには、感度が低下が支障にならない限り、できるだけ露光時間を短くするようなシャッタ回路が必要になる。

4.6 画像センサの分光感度[14,15]

撮像とは、光源→物体表面→画像センサという光の物理プロセスであり、撮像した画像の分光強度がそのまま対象物体の分光強度とはならない。白色光源のように見えても、線スペクトルからなる光源とフラットなスペクトルからなる光源とでは、センサは異なったカラー値を出力する。すなわち、光源の分光特性と画像センサの分光特性を同時に考えなければ正しい分光情報を得られない。しかし、本稿では紙面の関係で画像センサについてのみ言及する。

さらに分光について注意しなくてはならないこととして、光の分光は物理現象の一つであるが、色知覚は心理現象でしかないことを再確認してはならない。我々がリングを見て「赤い」と思っているのは、心理として「緑」木の前に「赤い」実を見ているのであって、必ずしも「赤い」ことが可視光の長波長帯の強度分布に対応しているわけではない。

表2 画像センサの特徴

		ドラム式	FSS	スキャン式	撮像管	固体撮像素子
レンズ系	幾何学歪	◎	○	△1	△	△
	シェーディング	◎	△	◎	△	△
ホトセンサ系	線形性	○	○	○	○2	△
	雑音	○	○	○	○	△
	感度むら	◎	◎	◎	○	△
走査系	縦横比	○	◎	○	◎	△
	幾何学歪	◎	○	◎	○	◎
	スキュー歪	△	△	△	△	◎
	ジッタ	△	○	△	○	◎
	ビデオレート	×	○	×	◎	○
スペクトル系	可視帯外	○	○	◎	○	○
	カラー	◎	◎	◎	×3	△

注1：回転ミラーは等角走査となる。注2：ビジコン系は非線形。注3：3管方式が必要。

しかしそれでは工学的に不便なので、色彩工学の分野では国際照明委員会CIE（Commission, International d'Eclairage）がRGBのスペクトル特性を規定し、商用のカラー画像センサはほぼすべてそれに準拠している。だが、このRGB各チャネル毎のスペクトル特性は視覚心理実験で策定されたものであり、カラーフィルタ（注：ゼラチンフィルタやダイクロイックミラー、干渉フィルタなど）の製作の容易さを考慮したものではない。したがって、現実のカラー画像センサの分光特性は機種毎に異なっている。

しかも、カラー画像センサにはホワイトバランスという調整箇所があり、この設定次第でいくらかでも分光特性が変化してしまう非常に大きな問題がある。分光特性の既知の白色光源でホワイトバランスを取らない限り、カラー画像センサの出力値を鵜呑みにしてはいけない。

5. 画像センサの選択と利用法

コンピュータビジョン用として現在利用できる画像センサの種別毎の特性を表2に示す。

この表と4章で述べたことから、コンピュータビジョンの研究用として現状の画像センサを用いる場合、次の点を考慮することを勧める。

<カメラの選択>

- ・ノンインタレース方式（ビデオカメラ業界では、プログレッシブスキャン方式と称されている）であること。
- ・画素が正方の固体撮像素子であること。
- ・カメラコントローラにガンマ値設定があり、それを1に設定する。
- ・シャッタ機能があり、感度低下が支障のない範囲でできるだけ高速シャッタとする。
- ・素子間で画素ずらしのないRGBに各一枚を割り当てた3板方式のカラーカメラであること。
- ・デジタル式インタフェースが望ましい。（現状では少数）

<フレームメモリ>

- ・カメラとフレームメモリは同期を取る。

<レンズ>

- ・幾何学歪の少ない計測用レンズを利用する。
- ・感度低下が支障のない範囲でできるだけレンズを絞り、被写体深度を深くし、結像ボケを少なくする。

<エイリアシング対策>

その他、画像センサは空間と時間のサンプリングセンサであるから、サンプリング周波数の2倍以上の濃淡変化と時間変動はエイリアシングノイズ（Aliasing）として現れる。

高精度レンズで精細なテキストを有する対象を撮像する場合には、画素ピッチの2倍以上の高い空間周波数を低減する空間ローパスフィルタを装着しなくてはならない。そのためわざとレンズの結像ぼかすことがある。画像センサ自体に、微妙に凹凸で散乱する透明フィルタあるいは特殊な光学結晶膜が貼り付けられていることもある。

時間軸方向にもエイリアシングノイズは考えられるが、一般に固体撮像素子が蓄積型であるため、蓄積によるローパスフィルタ効果で現れることはない。

6. コンピュータビジョン用画像センサの将来展望

現状では、コンピュータビジョン専用の画像センサというものは存在していないと言っても過言ではない[16,17]。電子印刷やコンピュータにおける画像技術では正方画素が基準であり、放送映像とは分野を異にするものであった。しかしこれまで画像センサ量産の背景をなしていた走査線と映像信号を中心としたテレビ放送が、近い将来には画素を中心とするデジタル放送に移行する。次期CS放送やDVDではMPEG2映像圧縮技術を採用するなど、放送とコンピュータが融合する真のマルチメディア時代が到来する。そこで、画像セン

サがコンピュータビジョン用に新規設計される可能性は少ないが、コンピュータビジョン適したものがより容易に入手できるようになるであろう。

例えば、1) 512×512正方画素ノンインタレース方式で、2) 画素ずらしなし3板固体撮像素子方式、3) ガンマ値1、4) 高速シャッタ付可変フレームレート（最高240Hz程度まで）、5) AD変換器内蔵デジタル出力インタフェース、6) 撮像素子冷却機構付のカメラが出現するであろう。

7. おわりに

これまでのコンピュータビジョン研究において、画像センサは、ピンホールカメラモデルとサンプリング定理を満たす完全なサンプリングを仮定してきた。しかし、現実の画像センサはいずれをも満足していない。

本稿では、画像センサをブラックボックスとして利用のみに徹するのではなく、コンピュータビジョンの手法を正しく適用するために、画像センサの撮像原理やその限界を紹介した。特に、既存の放送用画像センサの設計思想はコンピュータビジョン用とは大きく異なり、その利用には注意が必要である。

また、ある視覚機能が要求されているとき、人間の視覚を模倣して解決するというアプローチではなく、工学的に利用できる手段があるかぎり制限なく用いて解決を図るというマシンビジョンの立場がある。そこでは、人間の視覚が奥行きを両眼視差やT型エッジを手掛かりにしているのに対し、直接奥行きを計測するというアプローチが許される。例えば、レーザ光を投射することで正確に奥行き計測を行うレンジファインダや超音波の発振しその反射波の時間遅れを測るソナーなどである。本稿では、これらレンジファインダはコンピュータビジョン用画像センサの本流ではないものとして説明を割愛した。しかし、これらのセンサを過小評価することはできず、コンピュータビジョンの将来を担う重要な画像センサの一つであり、その動向にも注意を向ける必要がある。

参考文献

- [1] 情報処理学会コンピュータビジョン研究会IUE-WG編：画像データベースの構築調査報告書、(財)イメージ情報科学研究所(1995)。
- [2] 原島博監修：先端技術の手ほどきシリーズテレビジョン学会編画像情報圧縮、オーム社(1991)。
- [3] 長坂進夫他：現代テレビ・放送技術、オーム社(1989)。
- [4] 佐柳和男編：光応用計測の基礎、計測自動制御学会

- (1983).
- [5] 高橋清編；センサの事典、朝倉書店(1994).
- [6] 河田聡、南茂夫；科学計測のための画像データ処理、CQ出版(1994).
- [7] センシング技術応用研究会編；センサの集積化・知能化技術、日刊工業新聞社(1987).
- [8] 相田貞蔵他；電子計測基礎と応用、培風館(1989).
- [9] 木内雄二、長谷川伸；固体撮像デバイス、昭晃堂(1986).
- [10] 特殊カメラの利用動向、映像情報INDUSTRIAL、Vol.26、No.2(1994).
- [11] CCDカメラの新展開、映像情報INDUSTRIAL、Vol.27、No.5(1995).

- [12] 出口光一朗；カメラキャリブレーション手法の最新の動向、情報処理学会研究会報告CV82-1、pp.1-8(1993).
- [13] R.Y. Tsai, An efficient and accurate camera calibration technique for 3D machine vision, CVPR'86, pp.364-374(1986).
- [14] 大田登；色彩工学、電機大出版局(1993).
- [15] 日本色彩学会編；新版色彩科学ハンドブック、東京大学出版会(1980).
- [16] 丸谷洋二他；画像処理装置とその使い方、日刊工業新聞社(1989).
- [17] インテリジェント画像センシング、計測自動制御学会第15回応用計測部会講演会資料(1995).

CVCV-WGメーリングリストから、

(大阪大学久野さん)

>レンジファインダはコンピュータビジョン用画像センサの本流ではないものとして説明を割愛した。しかし、これらのセンサを過小評価することはできず、コンピュータビジョンの将来を担う重要な画像センサの一つであり、その動向にも注意を向ける必要がある。

できればレンジセンサについても触れていただければ良かったのですが、ページ数的にも無理とのことでしたので、次の機会(CVCVの成果は何らかの形でまとめる予定)をお願いします。なお、レンジセンサに対しては「過小評価することはできず」とありますが、発表の際にはこの辺についての御意見をもう少し詳しくお聞かせ願えればと思います。

(岡山大学和田さん)

コンピュータビジョンの研究者は、センサに関して、どんなセンサを用いたら、自分のアルゴリズムが正しく動作するか？(センサをアルゴリズムに合わせる)・どのようにアルゴリズムを修正すれば、あるセンサからの入力を正しく解析できるか。(アルゴリズムをセンサに合わせる)

という興味を持っていると思います。これらの疑問は、各センサはどのようなモデルで表現できるのか？という問いに集約できるはずですが、

こう思って原稿を読ませて頂きましたが、4章現実の画像センサの問題点で紹介されている各種アーチファクトが漠然としていてちょっと残念です。個体撮像素子にしばって、各種アーチファクトの数理的モデルを紹介して頂ければ、と思います。例えば、どういったレンズを用いるとピンホールカメラモデルとのずれが大きくなるのかといった話などを具体例を織り混ぜて説明して頂けるとありがたいです。

また、アーチファクトを問題視してキャンセルしようとするのではなく、逆に利用している例もあると思います。例えば、depth from defocus等では、ぼけを利用していますし、蓄積によるアーチファクト(motion blur)はoptical flowを計算するためにはあった方がいいはずですが、そういう意味では、6. 画像センサの選択と利用法では、「タスクに応じたセンサの選択肢がある」という議論があるべきではないでしょうか？例えば、プログレッシブスキャンは、動画を撮る時には必要ですが、静止画のときには必ずしも必要ないし、個体撮像素子のプログレッシブスキャンでは電荷を運ぶ際のロスが大きくなるという話を聞いたことがあります(チャンネルが長くなるから?)。

以上、勝手な感想で申し訳ありませんが、聞き入れて貰えるとありがたいです。

(筆者)

<阪大久野さんに対して> レンジセンサについては専門としますので、別の機会あるいはこのCVCV-WGの中でも別の順番をいただけるのであれば、是非

非技術評論してみたいと思います。

しかしレンジセンサもそうなのですが、CVCV-WGのメンバの皆さんは、どこまでの特殊センサをCVの範疇とお考えなのでしょうか。CV研究会ではこれまで、種々のセンサで撮像された2次元画像を扱った研究が確かに発表されてきましたが、CVという学問分野はどこまでを含めるべきなのでしょうか疑問です。

- ・可視光カメラ(ビジョントラック)
- ・赤外線カメラ(人物抽出するとか)
- ・紫外線カメラ(モンシロチョウの雄雌を画像識別するとか)
- ・X線カメラ(金属物を抽出するとか)
- ・光学式レンジセンサ(オクルージョン判定するとか)
- ・光学式傾斜センサ(法線ベクトルを出す)
- ・超音波センサ(医用画像認識)
- ・MRIセンサ(医用画像認識)
- ・X線CTセンサ(医用画像認識)

<岡山大和田さん> 第1の項目ですが、例えば濃淡分解能が16ビットあれば私のShape from Shadingがうまく動作すると言われても、そのようなカメラは製作できません。常温シリコンを使う限り7ビットが限界だからとかですね、やはり第2の項目が中心となるのではないのでしょうか。

当方の意見は、「タスクに応じたセンサの選択肢がある」はこのメールの前半部分で挙げたようなセンサから選択することであって、可視光センサの中から色々選ぶ必要はないと思います(コストの面は無視して)。もし、私が7章で示したようなカメラがあれば、motion blurは連続画像の時間積分をすればいいですし、depth from defocusも素性のよく分かったプリズムを前置してdefocusを起こせばいいわけで、コントロールのできないアーチファクトは無いほうがよく、アーチファクトが無ければアーチファクトと同じ現象を再現するのは容易、というのが意見です。

(岡山大和田さん)

分かりました。ただ、全てのアーチファクトを無くすることは不可能であり、あるアーチファクトを無くすることにより、別のアーチファクトが顕在化することはあるはずですが、ですから、「全てのアーチファクトを無くす」という方向でセンサの高精度化を考えるのは無理なのではないかと思えます。

そもそも、ここでアーチファクトと呼んでいるのは、理想化されたピンホールカメラからのずれであり、単純なセンサモデルを前提にしているからアーチファクトがあると云えるのだと思います。むしろ、現実に存在するセンサを前提にしてそのモデルを高精度化することにより、得られる画像の性質をきちんと把握する。そして、そのセンサを用いた場合の処理を考える(アルゴリズムの改良、タスクの変更の両方を考える)という方向の方が現実的ではないかと思えます。