

パネルディスカッション：大量カメラとネットワークは本当に必要か？

谷口倫一郎^{†1}, 大田友一^{†2}, 美濃導彦^{†3}, 石黒浩^{†4}, 桑島茂純^{†5}, 和田俊和^{†3}

「大量カメラを用いなければ得られない情報とは何か」、「大量カメラを用いたシステムにおいて通信が果たす役割」、「大量カメラを用いたシステムの今後の展望」などのテーマについて各パネラーの意見を述べ、これをもとにして「大量カメラとネットワーク」の今後について展望する。

Panel Discussion: Is mass-camera system necessary ?

RIN-ICHIRO TANIGUCHI^{†1}, YUICHI OHTA^{†2}, MICHIIHIKO MINOH^{†3},
HIROSHI ISHIGURO^{†4}, SHIGESUMI KUWASHIMA^{†5}, TOSHIKAZU
WADA^{†3†3}

In this panel, we will discuss varieties of topics, e.g., "What is the intrinsic information obtained only by mass-camera system?", "What is the role of 'communication' in mass-camera system?", "Future systems based on mass-camera". Through the discussions, we hope to have a vision of the future in this research field.

1. はじめに

「大量カメラ」とは言っても、この範疇にあてはめることができるシステムは、対象追跡、形状復元、全方位撮影、距離測定、視線誘導など、実に多様である。

これまでコンピュータビジョンの分野で研究されてきた内容の多くは、2つの目しか持たない人間の知覚機能を念頭において、それを模倣するという立場から考え出されてきたものが多い。しかし、3つ以上のカメラを用いたシステムを構築しようとした時点で、内省のみによっては何の研究の指針も得られないという状況が発生する。このとき、研究者が複数のカメラをどのように見なすかという点については、以下に示す2通りの捉え方がある。

- 知覚を行う主体が複数あり、それらが分散して存在し、通信によって互いの知覚結果を交換しているという「マルチエージェントモデル」。
- 知覚を行う主体はあくまで1つであり、昆虫のようにそれが大量の視覚センサを有しているという「複眼モデル」。

また、この大量のカメラが1) 環境に埋め込まれていると考えるか、2) 環境に埋め込まれていないと考えるかで、もう一つの分類の軸が得られる。

さらに、大量カメラを用いる理由が、1) 広い観測空間をカバーするためか、2) 精度や安定性を確保するためか、によっても分類の軸が得られる。

これらの観点から眺めてみると、上記の多様なシステムの違いが見えてくるはずである。

また、大量カメラを扱う場合、特に視野が共有されている複眼モデルの場合に必要な問題として、空間的・時間的なキャリブレーションが挙げられる。前者はエピソード幾何の復元問題であり、後者は撮影および処理の同期の問題である。

エピソード幾何の復元問題に対しては、2台のカメラの組のキャリブレーション結果を複数持ち寄る形で大量カメラのキャリブレーションを行うか、ワールド座標を規定するランドマークを基準にしてキャリブレーションを行うかで、大きな差が出る。具体的には、前者の方式を採用すれば、所謂 bundle adjustment を行う必要が生じ、カメラを1台追加しただけで、全体

†1 九州大学大学院システム情報科学研究科知能システム学部門
Department of Intelligent Systems, Faculty of Information Science and Electrical Engineering, Kyushu University

†2 筑波大学 機能工学系
Institute of Engineering Mechanics and Systems, University of Tsukuba

†3 京都大学 大学院情報学専攻 知能情報学専攻
Department of Intelligent Science and Technology, Graduate School of Informatics, Kyoto University

†4 和歌山大学 システム工学部 情報通信システム学科
Department of Computer and Communication Sciences, Wakayama University

†5 株式会社 ビュープラス
ViewPLUS Inc

のキャリブレーションを再度やり直す必要が生じる。また、後者の方法を採用した場合、そのランドマークが十分な空間的広がりをも有し、しかも常に位置が変わらないものを採用しなければならない。

同期の問題に関しては、カメラに同期信号を入れることはたやすいが、処理の段階でタイミングが合わなくなる可能性もあるため、この問題をどのように解決するかが問題になる。

こういった問題以外にも様々な問題が山積している状況ではあるが、以下のパネリストの講演を参考にして、議論に参加して頂ければ幸いである。

2. パネリストの講演タイトルと内容

大量カメラと実時間処理

谷口倫一郎 (九州大学)

大量カメラによる三次元世界の観測は、コンピュータビジョンの不良設定性を緩和できる、大規模な環境を観測できる、などの利点があり、様々な応用に用いられるようになってきている。大量カメラを利用する上で、もっとも厄介な問題の1つは、視点数の増加に伴う計算量の増大である。実時間処理を必要としない応用では画像を蓄積した上でオフラインでゆっくり計算すればよいが、実時間処理を必要とする場合は、実時間処理を実現する高速な通信・計算の仕組みを考える必要がある。特に、大量カメラから発生するバンド幅の大きなデータをいかに処理するかが重要であり、性能の高い並列・分散処理の技術を利用することが不可欠である。ここでは、このような観点から、大量カメラを用いた場合の実時間処理について論じることとする。

大規模空間を対象とした多視点映像の撮影と利用

大田友一 (筑波大学)

スタジアム規模の大規模空間において完全に同期し、キャリブレーションされた多視点映像を撮影する技術について概観し、今後の展開を考察する。

講義撮影における大量カメラの利用

美濃導彦 (京都大学)

講義撮影は、講義状況の認識とそれに基づくカメラ制御を基本としている。このため、講義室に設置するカメラを講義状況の認識に用いるカメラと撮影するカメラに分けて処理を単純化している。講義状況の認識は、カメラだけでなく他のセンサも活用する。講

義室はそのサイズが大きいので、センサとしてのカメラは、講義室すべてを万遍なく見渡せるように設定する。個々のカメラは基本的に動かさず全体として講義室全体の状況を画像的に取得する。カメラ間の通信は行わず、全てのカメラがサーバに対して処理結果を送信する方式を取る。サーバに存在するプロセスがさまざまなセンサ情報を統合し講義状況を認識する。

撮影は、認識された講義状況により、予め定められた規則に従ってカメラを制御する。このとき、遠隔講義のようなりアルタイム性が必要な撮影と講義のアーカイブのような要約目的での撮影は根本的に異なる。すなわち、遠隔講義は失敗しない撮影が要求されるのに対して、講義のアーカイブ化は冗長性のある多様な映像を得ることが要求される。遠隔講義の自動撮影システムを3年間にわたりUCLAとの遠隔講義に実利用した結果、映像取得や映像切り替えに関してもそれほど大きな問題は指摘されなかった。

人を見るカメラネットワークと人をつなぐカメラネットワーク

石黒浩 (和歌山大学)

カメラネットワークには、多数のカメラで局所領域を観測する形態と、広い領域を覆う形態があるが、後者を目指している。特に、街や建物内において、人間の活動を観測しモデル化する目的と、カメラを通してコミュニケーションする人間をつなぐ目的で研究を行っている。人間の活動の観測においては、人間の行動を緻密に追跡し、行動要素を取り出し、さらには行動要素間の構造を発見、記述することで、人間の行動モデルを獲得しようとしている。この人間の行動モデルは、情報基盤からの情報提供など、さまざまな目的に用いることができる。コミュニケーションを目的としたカメラネットワークでは、遠隔地にいる人々を動的に結びつけることを目的にしている。単なるテレビ会議ではなく、ネットワーク上にカメラから得られる映像を基に、カメラ間を動的に切り替えながら、遠隔地の人々がネットワーク上で出会う、コミュニティを形成することが目的である。

おまけのカメラと効果

桑島茂純 (株式会社ビュープラス)

対象の本質的認識は難しい。大きさを測る、たとえば背の高さを測ることは、難しい。1台のカメラでは画像の中の美人が背が高いのか、それとも近くにいるのか、それを判断することができないからである。だから、複数台のカメラで3次元計測となる。CPUパ

ワーは、あまっている。ネットワークは、流すものに困るほど太くなってきている。カメラの価格は、システムの価格からすれば、無視できるくらいに小さい。「カメラはおまけ」である。1個がおまけなら100個?もおまけのうちである。これを使って3次元画像処理をするのは当然の流れである。この理屈に納得できないなら、費用対効果の「効果」を探るしかないと思っている。微小な効果（と我々が思っているもの）に本当の価値があるかもしれないのではないか。

大量カメラのコンフィギュレーションと通信・処理

和田俊和(京都大学)

大量カメラで3次元空間を観測する際に様々なカメラの設置方法が考えられる。1) カメラの視野を共有させないように設置する場合、2) カメラの視野を共有させるように設置する場合の2つに大別されるが、1), 2) とともにさらに細かく以下のように分類することができる。1-1) 対象の移動経路を把握するために、対象の移動可能な通路にカメラ視野を設定し、それぞれのカメラの観測結果を対応付ける、1-2) 視野を広げるために外向きにカメラを設置し、モザイクングを行う、2-1) 複数のアクティブカメラの視線誘導を行いながら対象追跡を行う、2-2) 物体の3次元形状復元を行う、などである。これらそれぞれの概要と、必要となる通信のパターンや処理について述べ、単に大量カメラというのではなく、それを構造化して、体系的に扱う方法について議論する。
