

Cooperative Distribute Vision for Dynamic Three Dimensional Scene Understanding 分散協調視覚による動的3次元状況理解



プロジェクトリーダー 松山 隆 司

京都大学 大学院情報学研究科 教授

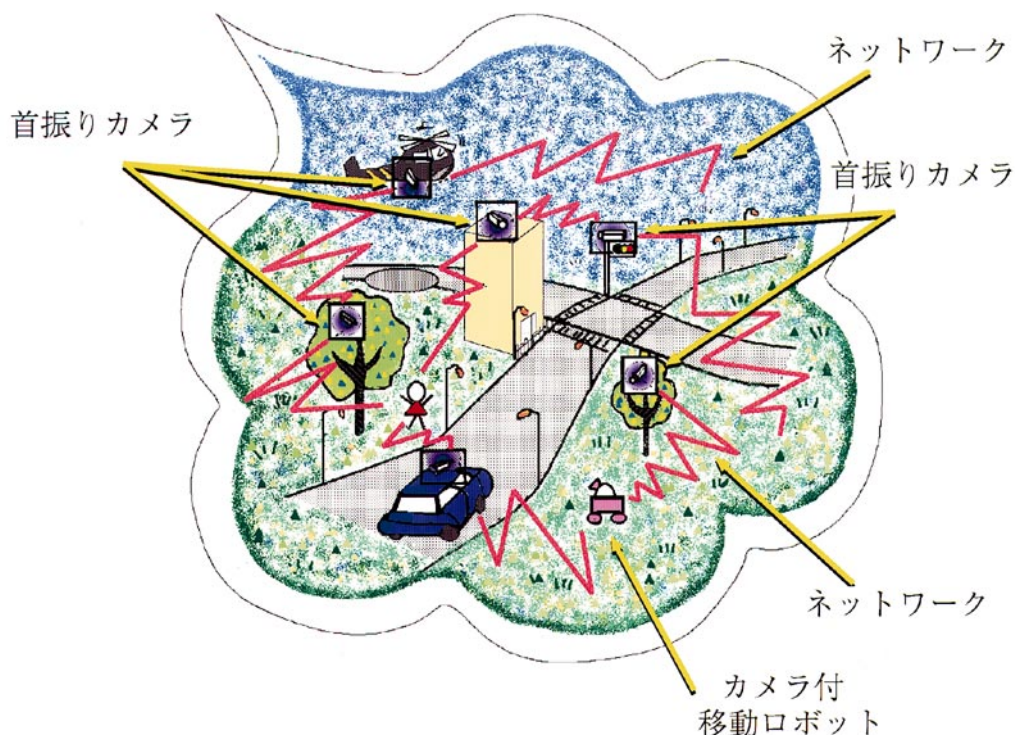


写真1：分散協調視覚システム

1. 研究の目的

人間は、感覚器を通して物事の様子や変化を認識し、それに基づいて思考を深めたり状況に即した適切な行動を取るといった優れた知的能力を持っています。このプロジェクトでは、視覚情報処理の立場から、人間の持つ知能のモデル化、柔軟で高い能力を備えた知能システムの実現を目的としています。

我々は、知能システムを構成する基本機能として、知覚、運動、コミュニケーションがあり、これら3つの機能の統合によって「知能」が発現すると考えており、その方法について研究するのがプロジェクトの中心的研究課題となっています。

具体的に言うと、我々が提案している分散協調視覚システム（写真1）では、有線・無線ネットワークで結ばれた多数の観測ステーション（首振りカメラを備えた実時間映像処理装置）や、視覚機能を持った移動ロボットにより、動的に変化する実世界の状況を多角的に観測し、それらの間の協調処理・動作によって以下の2つ機能を実現することを目指しています。

- (1) 分散協調型状況理解：観測ステーションと移動ロボットのコミュニケーション、協調処理・動作によって、動的に変化する実世界の多様な状況を実時間で把握する。
- (2) 対話の実時間映像生成：理解の結果得られた実世界の構造・変化に関する情報やネットワークを介して得られた情報を人間に分かりやすい多様な形態の映像情報として実時間で対話的に生成・編集・表示する。

分散協調視覚システムが実現できれば、交通管制やビル監視システムといった広域シーンを対象とした視覚認識システムのほか、対話型遠隔会議・講義システムや対話型3次元立体テレビシステム、手術・芸術・スポーツなどにおける高度な身体技能の詳細な映像記録、さらには移動ロボット群によるサッカー競技などが可能となります。すなわち、分散協調視覚システムは、政治・経済・学術・文化・教育から趣味・娯楽に至る、あらゆる社会的、個人的活動を支える情報システム基盤として役立つことが期待できます。

2. 研究の内容

このプロジェクトでは、分散協調視覚システムを構築する上で必要となる基本技術として、高機能視覚センサの開発、視覚と運動機能の統合メカニズムの考案、コミュニケーション機能を利用した協調処理・動作の実現、さらには実写映像とCG映像の実時間合成法の開発を取り上げ、理論研究から応用指向の技術開発まで幅広く研究しています。具体的な研究内容は以下の通りです。

- (1) 多重画像の統合による多機能高精度カメラシステムの開発: ビデオカメラ（以下では単にカメラと言います）の持つパラメータ（絞り、フォーカス、視点、視線など）を変えながら撮影した多数枚の画像（多重画像）を統合することにより、通常のカメラでは得られない情報の獲得や計測精度の向上を目指しています。これまでに高精細カラー画像や3次元距離画像、360°全方位パノラマ画像（写真2）を撮影できるカメラシステムを開発しました。
- (2) 視覚と運動機能の統合: 動的に変化する実世界の状況を捉え続けるために、視覚と運動機能を統合した対象の自動追跡システムを開発しています。具体的には、パン・チルト・ズームを変えても視点が変わらない視点固定型カメラを利用して、人間や車といった移動対象を実時間で追跡するシステムを開発しました。これまでの研究により、対象の動き予測に加え、システム自身の処理・運動能力（画像処理やカメラ制御に要する時間）を考慮することによって、対象追跡の追従性・安定性が大幅に向上することが分かりました。
- (3) 実世界環境における人間の動作認識: 異なる視点から同時に撮った複数のビデオ映像を解析し、人間の動作（部屋からの入退出動作など）の識別と、その人数の認識を行うシステムを開発しています。この研究では、複数の非決定性オートマトンを用いた「選択的注視に基づく協調的動作認識法」という新たな考え方を提案し、その有効性を実験によって実証しました。今後は、自動車の追越し動作やジェスチャなど、より複雑な動作の認識を目指しています。
- (4) コミュニケーションを利用した協調処理・動作の実現: 首振りカメラを備えた計算機群をネットワークで接続し、それらの間の協調動作（たとえば、あるカメラが対象を捉えた場合に他のカメラの視線をその対象に誘導したり、対象が死角に入った場合には他のカメラに対象追跡を引き継ぐなど）によって、障害物が多数置かれた広い空間内を移動する、特定の対象を実時間で安定に追跡するシステムを開発しました。今後は、多数の移動対象が存在する場合に計算機間でうまく役割分担をして追跡を行う機能を実現する予定です。
- (5) 強化学習による協調行動の獲得: 人は言葉を交さなくても身振りなどで相手の意図を理解することができます。つまり、位置・姿勢が変えられる物理的身体と視覚機能を持つロボットでは、ネットワークを介した通信に頼ることなくコミュニケーションを図ることが可能だと考えられます。この研究では、ロボットによるサッカー競技（写真3）を例として取り上げ、パスなどの協調動作やブロックなどの競合動作をトレーニングにより学習させる研究を行っています。
- (6) 対話的実時間映像生成・編集: 様々な方法で得られた情報も、最終的に人に伝えられてはじめて意味を持ちます。この研究では、視点の異なるカメラで撮った映像やCGによって合成した映像に効果的なカメラワークを付加することによって「見飽きない映像」や「分かりやすい映像」を実時間で生成・編集するシステムを開発しています。また、既存の映像からカメラワークを自動的に抽出したり、カメラワークが人間に与える心理的影響を分析するといった基礎的研究も行っています。具体的なシステム例としては、臨場感豊かな遠隔講義の実現を目指して、講師の動きや黒板の映像を実時間で撮影し、グラフィック表示された教材とともに分かりやすく生徒に示すシステム（写真4）の開発を進めています。

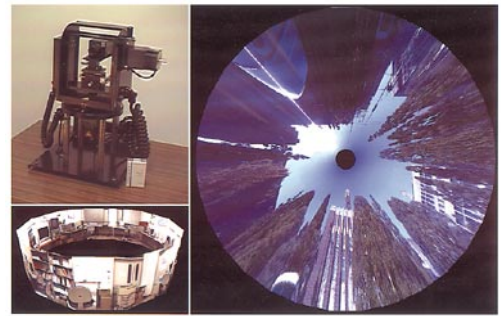


写真2：360°全方位パノラマカメラ（左上）とその映像（左下：計算機室、右：京都大学時計台）



写真3：ロボットによるサッカー競技

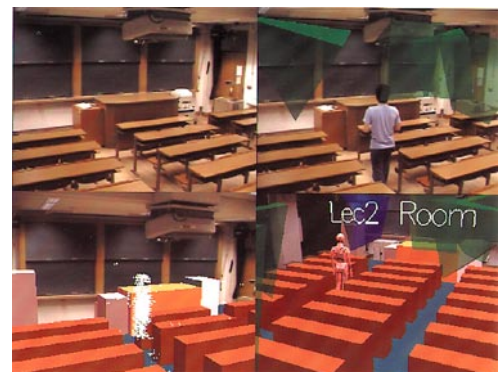


写真4：遠隔講義システム（左上：教室の実写映像、右上：人物とカメラの視野、左下、右下：CGと実写映像の合成）

3. 研究の体制

期 間：1996年10月～2001年3月

構 成：研究組織は、プロジェクトリーダー1名、コアメンバー3名、研究協力者10名（日本学術振興会研究員1名を含む）から構成されています。また、各メンバーの研究室に所属する大学院生が多数研究に参加、協力しています。

実施場所：主拠点は京都大学の松山研究室ですが、京都大学美濃研究室と大阪大学浅田研究室が研究テーマを分担、実施しています。このほか、東北大学出口研究室、東京大学池内研究室、九州大学谷口研究室、長崎大学吉田研究室、岡山大学尺長研究室が、月例のプロジェクト研究会や毎年開催されるプロジェクト主催の国際ワークショップに加わり、一体となって研究を推進しています。