

分散協調視覚プロジェクト

一分散協調視覚研究、システム開発の概要－

松山隆司（京都大学） 浅田 稔（大阪大学）

美濃導彦（京都大学） 和田俊和（岡山大学）

本論文では、筆者らが中心となり、本年度より5年間の計画で開始した分散協調視覚に関する研究プロジェクトについて、その背景、目的を述べたのち、具体的な研究課題とアプローチの方法を示す。

分散協調視覚システムでは、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（多自由度カメラ雲台を備えた実時間3次元画像・映像処理装置）や視覚機能を備え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、以下の2つの機能を実現することを目的としている。

1. 分散協調型状況理解：観測ステーション同士、観測ステーションと移動ロボット、移動ロボット同士のコミュニケーション、協調によって、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. 対話的実時間映像生成：理解の結果得られた状況記述やネットワークを介して得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表現・生成・編集する。

こうした分散協調視覚システムを利用すれば、実時間広域監視・交通管制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、対話型3次元立体テレビシステム・知的テレビスタジオ、手術、芸術、スポーツなど高度な身体技能の詳細な映像記録の作成、さらには、移動ロボットや身体障害者の対話的誘導システムや、サッカーなど移動ロボット群によるチームプレイが実現できる。本研究プロジェクトでは、これらの個別的応用システムの構築を目指すのではなく、それらのプラットフォームとなる分散協調型実時間3次元視覚情報処理機構を実現するための基盤技術、特に、ネットワーク接続されたマルチカメラによる多角的映像撮影・処理手法、分散協調視覚システム内部およびシステムと外的世界との協調的相互作用・コミュニケーション方式を重点的に研究する。

Cooperative Distributed Vision Project

－ Overview of the Project －

T. Matsuyama (Kyoto Univ.) M. Asada (Osaka Univ.)

M. Minoh (Kyoto Univ.) T. Wada (Okayama Univ.)

This paper gives an overview of our 5 years project on *Cooperative Distributed Vision* (CDV, in short). The goal of CDV is summarized as follows. Embed in the real world *Observation Stations* (real time image processors with versatile camera control capabilities) and *mobile robots* with vision, and realize

1. *robust, flexible, and real time* real world scene understanding
2. *versatile* image media capturing, generation, and editing.

Applications of CDV include Real Time Wide Area Surveillance, Remote Conference and Lecture Systems, Interactive 3D TV and Intelligent TV Stadium, Navigation and Guidance of Mobile Robots and Disabled People, and Cooperative Soccer Robots. The aim of the project is not to develop these specific application systems but to establish scientific and engineering foundations to realize CDV systems enough capable to work persistently in the real world. We focus on such topics as versatile imaging methods using multiple network-connected cameras, cooperative interaction and communication methods between observation stations as well as between a CDV system and the real world.

1 分散協調視覚研究の目的

一般に知能システムは、知覚・推論・行動の3つの基本機能を持つと考えられる(図1)が、従来の人工知能は記号処理に基づいた推論が中心となっていた。これに対し最近では、知覚、行動を介した外的世界との相互作用を基本とした知能システムの実現に大きな関心が寄せられている。こうした動きは、記号処理的計算理論に基づいた従来の情報処理から、**sensing&action**に基づく相互作用やコミュニケーションを基本とした新たな情報処理への展開を目指したものと考えられ、情報処理システムのヒューマン・インターフェイスの向上という観点からも重要なものであり、重点的な研究が望まれている。

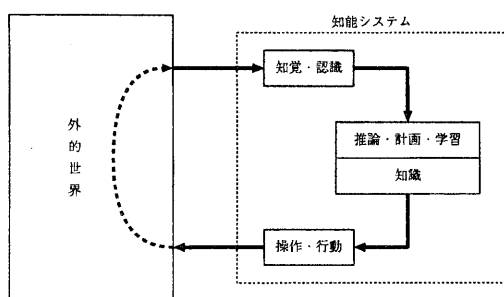


図 1: 知能システムの構成

一方、21世紀における社会は、政治、経済、文化、教育、娯楽などあらゆる個人的、社会的活動の分野においてマルチメディア情報処理システムがその活動基盤を支えるようになっていわれている。しかし、インターネットに代表される情報ネットワークシステム上に作られた「情報メディア社会」が急速に膨張することにより、それと我々生身の人間が生活する「実世界」との乖離が生じ、両者の狭間で様々な深刻な社会問題が生じる危険性を指摘する声もある。

我々は、視覚を通じた、情報システムと実世界との双方実時間相互作用によって、

1. 実世界の多様な状況において柔軟かつ頑健に機能する知能システムの実現および、
2. ネットワーク上に作られた情報メディア社会と実世界との有機的な結合

を図るための方式として、図2に示すような分散協調視覚システムを取り上げ、そのためのハードウェア、ソフトウェアを研究することを目的として本年度より5年間の計画でプロジェクトを開始した。

具体的には、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(多自由度カメラ雲台を備えた

実時間3次元画像・映像処理装置)や視覚機能を備えた移動ロボットにより、動的に変化する世界の状況を多角的に観測し、

1. 分散協調型状況理解: 観測ステーション同士、観測ステーションと移動ロボット、移動ロボット同士のコミュニケーション、協調によって、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状況記述やネットワークを介して得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表現・生成・編集する。

ことを目的としている。

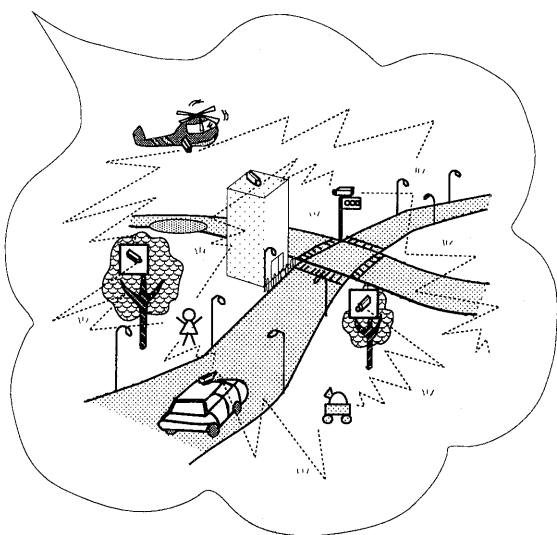


図 2: 分散協調視覚システム

こうした分散協調視覚システムを利用すれば、
・実時間広域監視・交通管制システム
など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
・対話型立体テレビシステム・知的テレビスタジオ
・商品・展示関心度モニタリングシステム
・手術、芸術、スポーツなど高度な身体技能の詳細な映像記録の作成

さらには、

・移動ロボットや身体障害者の対話的誘導システム
・サッカーなど移動ロボットによるチームプレイの実現が可能となる。

本研究プロジェクトでは、上記のような個別的应用システムの構築を目指すのではなく、それらのプラット

フォームとなる分散協調型実時間3次元視覚情報処理機構を実現するための基盤技術、特に、ネットワーク接続されたマルチカメラによる多角的映像撮影・処理手法、分散協調視覚システム内部およびシステムと外的世界との協調的相互作用・コミュニケーション方式を重点的に研究する。

2 研究の背景と特徴

2.1 視覚認識研究から見た位置づけ

視覚認識研究の歴史を振り返ってみると次のようになる(図3)。1970年代:デジタル画像処理(画像→画像への変換)、1980年代:コンピュータビジョン(2D画像→3Dシーンへの変換)、そして、1990年代になり、次の2つの流れが生じた。
アクティブビジョン:視覚と行動の統合
イメージメディア:画像の認識と生成の統合 [1]

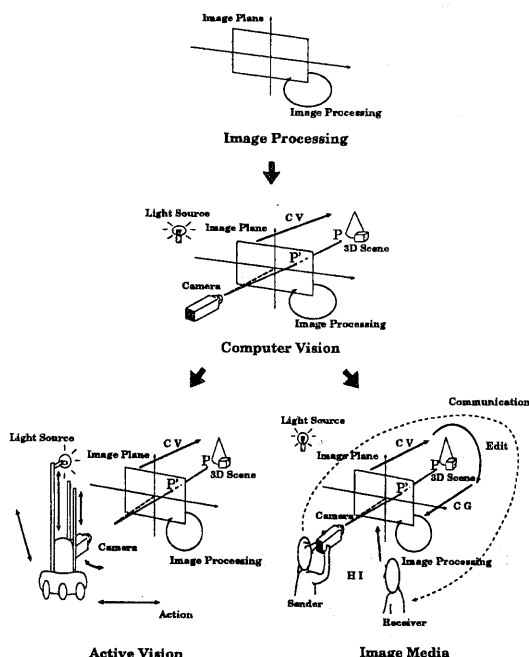


図3: 視覚認識研究の展開

本研究でいう分散協調視覚は、ネットワークで結ばれた多数の観測ステーションや視覚機能を備えた移動ロボット間の協調により、

1. アクティブビジョンの柔軟性、頑健性の向上
および

2. イメージメディア世界の実現

の共通基盤となる分散型視覚情報処理システムを実現しようとするもので、21世紀に向けての新たなビジョンシステムの展望を開くものであると考えている。

2.2 実世界で機能するシステムとしての意義

本プロジェクトで考えている分散協調視覚システムに対する最も基本的な要求仕様として、「実世界で有効に機能すること」がある。これは、いわゆる「積み木の世界」のような単純な規則によって記述できる世界や、雑音や歪み、情報の誤りや欠落のない理想的な論理・数学世界を対象とするのではなく、我々人間が暮らしている「実世界」そのものを対象として視覚情報処理の問題を考えることを意味する。

実世界を対象とした視覚認識問題に対する基本的アプローチの仕方として、プロジェクト・リーダーである第1筆者は次のように考えている。

『従来の還元主義に基づく自然科学は、「もの」を構成する根源的要素の追求を論理的、客観的に行うことを基本原理としており、従来のビジョン研究、特に computational vision は、その考えに忠実に従ってきた。しかし、Shannonの情報理論が送信者と受信者の間の通信によって情報を定義していることから明らかに、情報処理、情報科学の問題は「もの」に関する研究ではなく、「こと」に関する研究であると言え、「もの」を科学するための方法が限界に突き当たるのは当然であると言える。』

我々は、『「こと」は「もの」と「もの」との相互作用(interaction)(従来の「関係」といった静的なつながりではなく、ダイナミックスを伴った動的係わり合い。)と考えるのがまずは妥当な出発点であり、「こと」の1つの様相として「統合」がある』と考え、数年前から「多角的情報の統合」というスローガンのもとに、実世界で機能するビジョンシステムを実現するための情報統合方式やアルゴリズムの研究を行ってきた。

その一方、最近では我々以外にも、より一般的な立場から情報統合を考えようとする動きや、マルチメディア情報処理におけるメディア統合、アクティブビジョンや subsumption architecture のように異なった機能を持つモジュールの有機的結合を目指す機能統合、さらには多数のエージェントを集団あるいは社会として統合することを目指すマルチエージェントシステムに関する研究も行われており、この機会に何とかして「統合」を科学的基盤を持つ新たな情報処理方式として確立したいと考えている。』[2]

分散協調視覚システムは、こうした考えに基づき、実世界において有効に機能するために必要な基本特性として次の3つのものを取り上げ、以下に示すアプローチによってそれらの実現を目指している。

実時間性：ビデオレート¹での画像・映像の認識・生成処理、超高速ネットワーク、並列・分散ソフトウェアの開発

頑健性：多角的情報の統合による対雑音性、信頼性の向上および有用情報の獲得 [3]、分散協調処理による fail-safe、smooth-degradation 機能の実現

柔軟性：分散協調処理による実世界の多様な状況への柔軟な対応。すなわち、従来の単一の内部状態、制御フローしか持たないシステムでは、実世界の多様な状況に対応しようとすると、事前に可能な状況をすべて場合分けしておく必要があるが、実際上そうした場合は困難である。これに対して分散協調処理システムでは、各エージェントが独自に状態や制御フローを持ち、それらの相互作用の結果として創発される全体システムが持つ状態数、制御フローの分岐数は組み合わせ的に増大し、実世界の多様な状況に十分対応できると考えられる。言い換えると、エージェント間の相互作用によって、システムの状態、制御フローに組み合わせ爆発を生じさせようというわけである。

2.3 マルチエージェントシステムから見た位置づけ

「分散協調」の名前が表すように、本研究はいわゆるマルチエージェントシステムに関する研究の1つであると言える。しかし、本研究が目指す分散協調視覚は、分散計算機システム上での基本ソフトウェア・ミドルウェア・応用ソフトウェア実現のためのソフトウェア・エージェントや、新たな計算モデルとしてのエージェント指向プログラミングに関する研究とは異なり、次のような特徴を持つ。

物理的身体性を持ったエージェント：分散協調視覚システムにおけるエージェントは、観測ステーションや移動ロボットといった物理的身体を持った実在物であり、視覚を通じた sensing およびカメラ制御や身体運動といった action によって、実世界との直接的相互作用機能を持つ。我々は、こうした機能により、実世界で有効に機能する知能システムの実現およびネットワーク上に作られた情報メディア社会と実世界との有機的な結合が可能となると考えている。

メッセージ通信なしのコミュニケーション・協調：ソフトウェアとしてのマルチエージェントシステムでは、エージェント間のコミュニケーションや協調はメッセージ通信を基本としている。一方、分散協調視覚システムが対象とする実世界では、観

測ステーションや移動ロボットの意志とは無関係にそれらの物理的身体に対して直接的に反応する音や光というメディアが存在し、そうしたメディアを用いた sensing によって、他の観測ステーションや移動ロボットの存在や行動を認識することができる。言い換えると、sensing 機能の利用によって、メッセージ通信を用いないコミュニケーションや協調が実現できる。本研究では、いわゆる「アイ・コンタクト」のように、視覚という sensing 機能に基づいたエージェント間のコミュニケーションや協調を実現するための基本機構の実現を目指している。

3 研究課題とアプローチの方法

分散協調視覚システムを実現するための研究課題を大別すると次のようになる。

1. 分散協調視覚システムのためのハードウェア、ソフトウェア基盤の構築
2. 多機能高精度視覚センサの開発
3. 観測ステーション群による移動対象の協調的監視・誘導方式の考案
4. 動的3次元状況の柔軟な映像撮影・生成・表示方式の考案
5. 視覚に基づくロボットの多種行動の獲得及び統合方式の開発

分散協調視覚システムについては、既に第1、4著者がこれまで数年間基礎研究を行ってきており[4][7][8][9][10][11]、以下では各研究課題の具体的内容と本プロジェクトでのアプローチの仕方、これまでに得られた成果の概要を述べる。

3.1 分散協調視覚システムのためのハードウェア、ソフトウェア基盤

現在本プロジェクトでは、分散協調視覚システムのためのハードウェア、ソフトウェア基盤として図4のようなシステムを設計、開発している。

観測ステーション：本プロジェクトでは、多自由度カメラ雲台を備えた実時間3次元画像・映像処理装置を1つの観測ステーションと考え、それらをネットワークで結合したものを分散協調視覚システムのハードウェア基盤とする。観測ステーションのハードウェア構成は次のようにする。

- カメラ：フォーカス、ズーム、アイリスが計算機制御可能な3 CCDの計測用TVカメラを基本とするが、後に述べる独自方式の実時間3次元距離計測カメラの搭載も考える。
- 多自由度カメラ雲台：後に述べる独自方式のものを設計・開発する。

¹ 一般には、実時間 ≠ ビデオレートであるが、技術的实现容易性の観点から、本プロジェクトではビデオレートを実時間性の基準として考える。

案し、その有効性を実験によって示してきた [3][7][8]。

これまでに実世界での有効性が確認されたものとしては、次の2つのものがある。

1. 多重フォーカスカメラ：多重フォーカス画像を用いた実時間3次元距離計測用カメラ [7]
2. Appearance Sphere (APS)：首振りカメラを用いた全方位画像の生成 [8]

本プロジェクトでは、多重画像の統合という考え方に基づいた多機能高精度視覚センサの開発を目指しており、このことは本プロジェクトの大きな特徴の1つとなっている。

具体的には、1については、通常の3CCD TVカメラを改良して実時間で多重フォーカス画像が撮影できるカメラを試作し、基本性能の評価を行った。その結果、実用的な実時間距離センサを実現するにはまだ幾つかの点についての研究・開発が必要であるが、十分期待が持てることが分かっている [7]。

また、2のAPSについては、APS画像が撮影できるカメラ雲台の試作が終了し（図5）、現在はその性能評価、APS画像から任意のパン、チルト方向の画像生成を実時間で行うソフトウェアの開発を進めている。

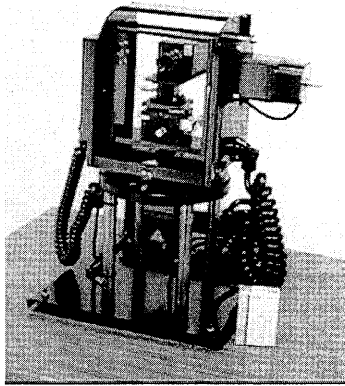


図 5: 試作した APS カメラ

APSは、首振りカメラを利用した視覚認識や3次元シーンに関する動的な映像生成など様々な問題に利用可能な映像撮影法・視覚環境モデルである。本プロジェクトでは、APSを観測ステーションにおける基本撮像機構と考え、現在APSを利用した様々な画像認識・生成アルゴリズムの開発を進めている。これらのアルゴリズムについては、成果がまとまり次第、本研究会で発表する予定である。

3.3 観測ステーション群による移動対象の協調的監視・誘導

実世界の状況は非常に多様であり、観測ステーション群による移動対象の協調的監視・誘導のための一般的方法を一気に考えるのはむずかしい。そのため、本研究では、以下のような部分問題を設定し、基本問題から一般的な複雑な問題へと順次解決法を考える。

1. 単一観測ステーションによる単一移動対象の追跡
2. 複数の観測ステーションによる単一移動対象の共同注視
3. 観測ステーション間のコミュニケーションのための共通的情報表現基盤の確立
4. 複数の観測ステーションによる複数の移動対象の協調的追跡
5. 視野に重なりを持たない観測ステーション群による広域監視

なお、ここで考える移動対象の協調的監視（検出・追跡・認識）法は、次に述べる「動的3次元状況の柔軟な映像生成・表現」のための映像撮影法にもなっている。さらに、観測ステーションで得られた情報とロボットなどの移動対象が得た情報を相互に交換することにより、対象の柔軟な誘導が実現できる。この意味で、本研究は、本プロジェクトにおける中心的な研究課題となっていると言える。

3.3.1 単一観測ステーションによる単一移動対象の追跡

ここでは、最も基本的な問題として、単一の観測ステーションによって1つの移動対象の実時間追跡を行うアルゴリズムおよびカメラ制御方式を考える。

先に述べたように、各観測ステーションにはAPSカメラが備えられ、それを用いて移動対象の検出と追跡を行う。そのためのアルゴリズムとして、現在次のようなものを研究している。

頑健な背景差分実現のための動的背景モデル：任意のパン・チルト（・ズーム）で撮影した観測画像と、APSから生成した背景画像との差を取るという背景差分によって、既知の背景とは見え方が異なる「変化領域」として移動対象を検出することができる。また、カメラの視線制御と背景差分を繰り返すことにより、対象の追跡も実現できる。しかし、実世界におけるシーンでは、様々な要因によって背景の映像はたえず変動しており、1枚の静止画像によって背景モデルを表現した場合、変動する背景（風にそよぐ木々、CRTのフリッカなど）も移動対象として検出されてしまう。我々は、こうした問題を解決し、頑健な背景差分を実現するための方法として、背景の変動をモデル中に取り込んだ動的背景モデルの研究を行っている。この考え

方の有効性は、すでに [9] において示したが、現在はより頑健な動的背景モデルの構築を進めている。

選択的注視に基づく動作認識：これは、固定カメラを用いて撮影した動画に背景差分を適用し、得られた変化領域の時系列を基に、対象の動作モデルの学習とモデルに基づいた動作認識を行おうとするもので、その詳細は [10] を参照して頂きたい。

カメラ運動を利用した移動対象の検出と運動情報の獲得：

APS カメラの場合、カメラ運動によって生じる画像中のオプティカルフローは、シーンの 3D 構造に拘らず画面内の各点において常に一定となる。このため、他の場所とは異なったオプティカルフローを示す場所を求めることによって容易に移動対象が検出できる。また、画面全体のオプティカルフローからカメラの正確な運動速度が計算でき、それから対象の運動情報も求めることができる。現在、このアイデアに基づいた移動対象の検出と運動情報の獲得アルゴリズムの開発を進めており、結果が得られ次第報告する。

3.3.2 複数の観測ステーションによる単一移動対象の共同注視

複数の観測ステーションが関わる協調処理の最も基本的、かつ重要な機能としては、それらの協調によって、ある 1 つの移動対象を実時間で追跡し、対象に関する多視点観測情報を統合することにより、その対象に関する詳細な（3 次元、運動）情報を獲得することがある。

こうした共同注視問題では、各観測ステーションで捉えた対象が同一のものであるかどうかの認識が問題となり、そのために使う特徴および検証のためのコミュニケーションをいかに行うかをまず考えなければならない（3.3.3 参照）。また、複数のカメラによって、異なった視点から見た対象の像が多数得られるので、それらから対象の（3 次元）モデルを作成することも考えられる。さらに、一方のカメラの視野から対象が出た場合や新たに入ってきた場合など、対象の多様な運動をうまく追跡するための観測ステーション間のコミュニケーション法も重要となる。

これらの技術課題を解決するための 1 つの方法として、先に述べた選択的能動注視に基づく動作認識法を拡張することが考えられる。すなわち、

1. まず、ある同一の対象動作を複数の観測ステーションで観測し、各観測ステーションにおいてその動作のモデルを事前に学習させる。
2. ある観測ステーションでその動作クラスに対応する状態への遷移が起きると、他のステーションにそれを通知する。

3. 通知を受けた観測ステーションは、その動作を認識するためのカメラアクションを行い、

4. 複数の観測ステーションが互いの状態を参照しながら共同で認識処理を行う。

このように、イベント—アクションループを複数の観測ステーションのグループに拡張することにより、シーン中の対象を多角的に観測し、より信頼性の高い動作認識を行う共同注視による動作認識の機能を実現することができる（図 6）。

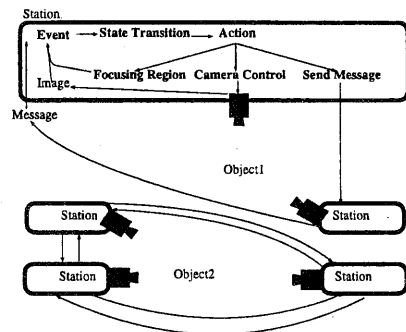


図 6: 共同注視による動作認識

3.3.3 観測ステーション間のコミュニケーションのための共通的情報表現基盤の確立

分散配置された複数の観測ステーションの間で、それぞれが観測した情報を交換するには何らかの共通の枠組み（基準、言語）が必要である。先に述べた共同注視による動作認識では、同一の対象の動作を同時に学習させることによって、観測ステーション間のコミュニケーションの共通基盤が作られることになるが、一般的にはそうした学習を行わせることは困難である。また、各観測ステーションにおけるカメラの位置や動きを共通の世界座標系によって記述しておくことも考えられるが、そのためのキャリブレーションは容易ではない。特に、観測ステーション間の距離が大きく離れている場合には、土木測量的キャリブレーションも必要となる。

本研究では、実世界において物理的に裏付けられた水平・垂直軸といった幾何学的基準軸や同時刻性、さらには幾何学的、光学的不変特徴を利用した、観測ステーション間のコミュニケーションのための情報の共通表現基盤の設定および、それらを利用したメッセージの記述法について研究する。

具体的な研究課題としては、次のようなものがある。

複数カメラの同時キャリブレーション：一般に固定カメラの場合でも、そのカメラパラメータを正確にキャリブレーションするのは容易なことではない。これに加えて首振りカメラの場合は、首振り運動と幾何学的基準軸との間の関係を求めるキャリブレーションをどのように行えばよいか、など多くの問題がある。

不変特徴：各観測ステーションで撮られた画像は、ある特定の位置、視線方向からの観測データであり、それをそのまま他の位置にある観測ステーションに伝えても意味をなさない。両者の間のコミュニケーションを実現するには、何らかの幾何学的、光学的不変特徴を求める必要がある。

カメラ制御プロトコル：カメラパラメータには、光学的なもの（フォーカス、焦点距離、絞り、シャッタースピード、レンズ特性）と幾何学的なもの（投影中心、視線方向、倍率）、さらにはデジタル化に関するもの（サンプリング間隔、量子化レベル）がある。カメラの状態はこれらのパラメータの組として表され、その状態変化と撮影される画像・映像の間の写像関係は非常に複雑なものになる。カメラを制御する、特にある観測ステーションが他の観測ステーションのカメラを制御する場合には、そうした写像関係に基づきカメラをコントロールする必要があり、そのための制御プロトコルをどのように表現、記述するかが問題となる。

3.3.4 複数の観測ステーションによる複数の移動対象の追跡

一般的状況では、シーンには複数の移動対象があり、システム全体としては、それらの運動状況をすべて把握し追跡することが要求される。しかし、個々の観測ステーションの視野や処理能力には自ずと限界があり、ある時点である対象を追跡するのに適した観測ステーションは動的に変化する。また、前述のように対象の詳細な形状、運動情報を獲得するには複数の観測ステーションによる共同注視が必要となる。さらに、複数の利用者が異なった目的で同一シーンの状況を把握しようとする場合には、システムはそうした多数のタスクを同時に実行しなければならない。

これらの問題を解決するには、複数の観測ステーションが動的にうまく作業分担をして、複数の移動対象を適応的に追跡するための Dynamic Task Sharing が必要となる。こうした機能を図4に示したシステム上で実現し、実世界での有効性を実証することが、分散協調視覚研究の最終ゴールである。

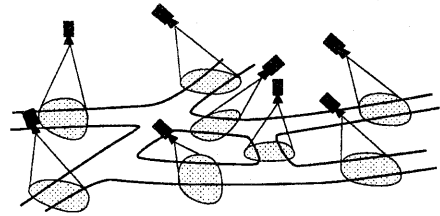


図 7: 観測範囲の重なり

3.3.5 視野に重なりを持たない観測ステーション群による広域監視

これまでの議論は、観測ステーションの持つ視野に重なりがあり、移動対象を複数の観測ステーションによって同時に観測できることを前提としていた。しかし、各観測ステーションの観測範囲が互いに重なりを持つことをシステムの前提条件とした場合、

1. 広い地域全体をカバーするには膨大な数の観測ステーションが必要になる。
2. 観測ステーションの故障や通信路の障害等によって観測範囲の重なりが損なわれた場合、システム全体が動作しなくなる。

という致命的な問題が生じる。このため、道路網や広い工場などを対象とした広域分散監視システムを構成するには、図7に示すように観測範囲が重なりを持たない場合の移動対象の追跡方式が必要になる。

こうした視野に重なりを持たない観測ステーション群による対象の広域追跡のための基本的アルゴリズムをすでに [11] で提案した。現在は、このアルゴリズムの改良、高度化を進めている。

3.4 動的3次元状況の柔軟な映像生成・表現

本研究では、予め撮影・蓄積された映像を表示するだけでなく、利用者の意思に基づき自由な視点から実世界中の対象の映像を取得し、それに関連した付随情報をネットワークを介して柔軟に検索・生成・表示できる情報メディア環境の構築を目的としている。言い換えると、本研究は、実世界と情報メディア社会の統合化を実現するために分散協調視覚システムを利用するものと言え、観測ステーション間の協調によって注視対象となる物体映像を最適に取得する方法および、その映像を利用者に効果的に提示する方法を開発する。

このような機能が要求される問題の具体例としては、遠隔講義や会議・ミーティングなど、空間（場）を共有する複数の人間のグループがいくつかあり、それらのグループをネットワークで結び共同作業を実現することがある（図8）。これらの問題では、臨場感を醸し出すためにいかに映像を撮影、表示するか、テキストや

資料、データといった様々な付加情報を映像といかに融合させて表示するかが重要となる。

本研究では、上記のような問題における映像の撮影・生成法に焦点を当て、以下の課題について研究を行う。

対象モデルを利用した不完全映像の完全化

予め決められた有限の観測環境の下で、利用者による自由な視点位置や視線方向の選択を可能とするため、本研究では対象に関するモデルを知識として用意し、観測ステーションからの映像情報が不完全な場合でもモデルを利用した映像生成・合成により、利用者には対象の完全な映像や情報を提示し続けることができるシステムを構築する（図9）。

映像撮影資源の協調的割り当て方式の開発

多数の利用者が同じ場に関心を持ちその状況を見ようとする場合（図8）、同一の場に属す複数の異なった対象をそれぞれ注視対象とした映像の獲得が必要となる。この問題は、本質的に3.3.4で述べたものと同一であり、そのための映像撮影資源の協調的割り当て方式の開発は本プロジェクトにおける中心的課題となる。

カメラ群の最適自動配置アルゴリズムの考案

TVスタジオや会議・講義室といった場では、場の幾何学的、光学的構造が予め分かっており、そこで行われる行為、アクションには一定のシナリオがある。本研究では、こうした場に関する知識と、撮影に使用するカメラの特性、台数に関する情報を利用して、場において生じる様々な動的状況を最も効果的に撮影するためのカメラ群の配置を自動決定するアルゴリズムの考案を目的としている。現在は、計算幾何学的アプローチに基づいた最適なAPSカメラの自動配置アルゴリズムの開発を進めている。

カメラワークの記述方式の考案

TVスタジオや映画撮影では、複数のカメラで撮られた映像を動的に切り替えたり、編集したりすることによって、「見飽きない映像の生成」や「制作者からのメッセージや意図の伝達」、さらには「芸術的価値の付与」がなされている。本研究では、分散協調視覚システムを用いて、こうした付加価値を持った映像を生成するための、カメラワーク記述方式の考案を目指している。現在は、研究の第一歩として、ネットワークに接続された複数の首振りカメラを対話的に制御し、効果的な映像獲得・生成を行う分散型ネットワークカメラシステムおよび、一般のニュース、ドラマ映像で用いられているカメラワークを属性文法に基づいて記述する研究を進めている。

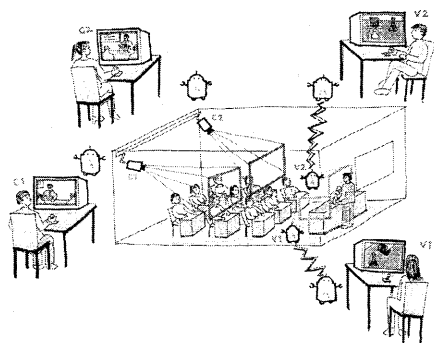


図8: 分散協調視覚システムを備えた講義室

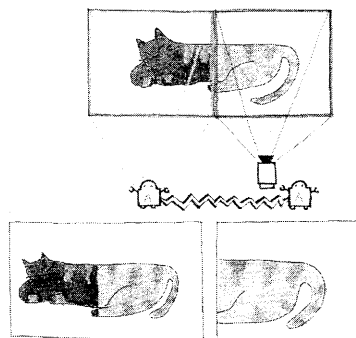


図9: 実写映像とモデル映像の合成による対象の表示

3.5 視覚に基づくロボットの多種行動の獲得及び統合方式の開発

本研究では、視覚を備えた移動ロボット群に統一の取れたチームプレイ動作を行わせるためのロボット間の協調及び競合行動を視覚に基づいた強化学習[12]によって獲得する手法を考案し、その有効性を実ロボットによる実験によって検証する。特に、本研究では、物理的身体性を持ったロボットに協調動作を行わせるには、いわゆる「アイ・コンタクト」のようなメッセージ通信なしのコミュニケーションを実現することが重要であると考え、そのための視覚機能や行動制御方式の研究を行う。

一般に認識・理解のこれまでの手法は、認識対象のモデルをプログラマが用意し、実データとモデルとの照合を行うことによって実現されてきた。しかし、モデルの妥当性や照合結果の評価を行うための基準が明確でなく、比較的簡単な認識問題しか扱えなかった。本

研究では、単独あるいはチームとしての他のロボットの行動理解という高度な認識機能を実現するための方式として、環境のモデルをロボット自身が学習を通して構築し、それに基づいて行動するというものを探る。こうした方式では、学習されたモデルの妥当性や、いわゆる認識・理解の深さが、ロボットの物理的行動として現れ、行動を評価することによってロボットの持つ認識能力を測ることができる。

具体的な問題設定としては、ロボットによるサッカー競技を取り上げ、ドリブル、シュートなどの個々のロボットの行動、パス、センタリングなどの複数ロボット間の協調行動、さらにはブロックなどの競争行動を視覚に基づいた学習によって獲得するための方式の開発を目指す。この問題は、敵、味方に分かれた多くのロボット群が存在するという環境の中で、他の個々のロボットの行動理解及びチームとしての行動戦略パターンの理解などといった高度な視覚認識の問題を含む。図10に現在整備中の実験環境を示す。

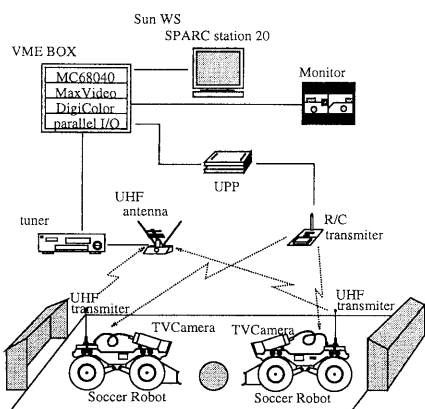


図 10: 視覚に基づくロボットの多種行動の獲得研究の実験環境

4 おわりに

本報告では、日本学術振興会未来開拓学術研究推進事業の一環として1996年から5年間の計画で進めている分散協調視覚プロジェクトの概要を述べた。本プロジェクトには、著者のほか、池内克史（東大）、谷口倫一郎（九大）、吉田紀彦（九大）、尺長 健（岡大）および各研究室のメンバが参加している。

本プロジェクトでは、ほぼ毎月研究会を開催しメンバ間での研究討論、検討を重ねており、その成果は CVIM 研究会およびプロジェクトのホームページ (<http://vision.kuee.kyoto-u.ac.jp/CDVPRJ/>) を通じて公表する予定である。また、今後公開の研究会や国際シンポジウムなどの開催も計画している。さらに、本プロジェクトでは、成果として得られたハードウェア、

ソフトウェアを企業と共同して実用化することも検討している。これらに関する情報は適宜ホームページを参照して頂きたい。

我々は、本稿で述べた分散協調視覚の考え方やそこで開発されるハードウェア、ソフトウェアは、21世紀の社会を築く上で大きな役割を果たすものと考えており、精力的な研究を進める計画である。

参考文献

- [1] 松山、久野、谷口、和田：特別企画：わが国における IP、CV 研究の軌跡と現状、情報処理学会研究会資料、CVIM100-1, 1996
- [2] 松山：AI マップビジョン研究から見た統合アーキテクチャ、人工知能学会誌, Vol.10, No.6, pp.888-894, 1995
- [3] 松山：多重画像の統合-高精度画像計測と多機能画像生成-, 電子情報通信学会誌, Vol.79, No.5, pp.490-499, 1996
- [4] 松山：分散協調処理による画像理解、計測と制御, Vol.31, No.11, pp.1149-1154, 1992
- [5] 手塚ほか：PC とギガビット LAN による PC クラスターの構築、情報処理学会研究会資料、ARC-96-119, 1996
- [6] Y.Ishikawa, et al: MPC++, In *Parallel Programming Using C++* (G.V.Wilson and P.Lu eds), MIT Press, pp.429-464, 1996
- [7] 竹村、松山：多重フォーカス画像を用いた実時間 3 次元距離計測-分散協調視覚システムのための多機能高精度センサの開発-, 情報処理学会研究会資料、CVIM103-5, 1996
- [8] 和田、浮田、松山：Appearance Sphere -パン・チルト・ズームカメラのための背景モデル-, 画像の認識・理解シンポジウム (MIRU'96), pp.II-103-II-108, 1996
- [9] 和田、松山：動的背景モデルを用いた移動領域の抽出, 情報処理学会第 49 回全国大会, pp.2-141-2-142, 1994
- [10] 和田、加藤：選択的注視に基づく動作識別-分散協調視覚システムにおける対象の動作認識法-, 情報処理学会研究会資料、CVIM103-6, 1996
- [11] 和田、田村、松山：広域分散監視システムにおける分散協調型対象同定法, 画像の認識・理解シンポジウム (MIRU'96), pp.I-103-I-108, 1996
- [12] 浅田、野田、俵積田、細田：視覚に基づく強化学習によるロボットの行動獲得, 日本ロボット学会誌, Vol.13, No.1, pp.68-74, 1995