

## [ フェロー記念講演 ] 「人間とは？知能とは？」 - 視覚情報処理からのアプローチ -

松山 隆司

京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町

E-mail: tm@i.kyoto-u.ac.jp

**あらまし** フェローの称号を頂いたことを機会に、筆者の30年間に渡る研究の歩みを振り返るとともに、研究の現状と今後のビジョンを述べる。

**キーワード** 画像処理、画像理解、コンピュータビジョン、3次元ビデオ、ハイブリッド・ダイナミカルシステム

## What's Human? What's Intelligence? - An Approach from Visual Information Processing -

Takashi MATSUYAMA

† Graduate School of Informatics, Kyoto University Yoshida-Honmachi, Sakyo, Kyoto, 606-8501 Japan

E-mail: tm@i.kyoto-u.ac.jp

**Abstract** As the memorial note of being given IEICE Fellowship, the paper gives an overview of research activities of the author for these 30 years and addresses current research topics and future vision.

**Keyword** Image Processing, Image Understanding, Computer Vision, 3D Video, Hybrid Dynamical System

本稿では、学生時代から現在までの研究の歩みを振り返り、最後に今後のビジョンを述べてみたいと思います。(なお、この文章を書くに当たって、過去の資料を色々調べ、できる限り正確を期すようにしましたが、思い違いや記述ミスなどが含まれているかもしれません。その際はご指摘、ご容赦頂ければ幸いです。)

### 1. ザリガニの神経系の分析

最初に研究という活動に触れたのは、大学4回生の卒業論文研究でした。学生時代は、工学部電気系に所属していながら、生物の生態や心理・生理学に興味があり、これらの分野の入門書を読み漁っていました。このため、研究室選びでは、近藤文治教授(制御工学)の研究室で行われていたザリガニの神経系の研究に惹かれ、近藤研究室を希望しましたが、希望者多数によるじゃんけんで負け、宇治の小川徹教授(電離層実験施設)の研究室に配属されました。

小川研究室では量子力学、レーザに関する理論、実験的研究が行われており、卒業研究としては発振波長を変化させることができる色素レーザ装置の開発を行うことになりました。レーザ発振は実現できませんでしたが、卒業研究は、それまでの講義や実験と違って、英語論文の講読や実験装置の試作、性能評価実験などすべてのことが目新しく、宇治に通うことも全く苦にならず楽しい生活を過ごしました。

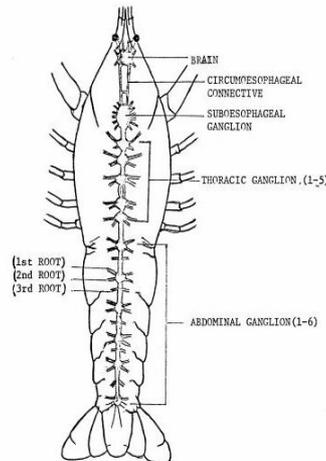


図1 ザリガニの神経ネットワーク

卒論中も生物に関する興味はなくなることなく、1973年夏の大学院入試では、迷うことなく近藤研究室を希望し、1974年4月からザリガニの研究を行うことになりました。

ザリガニの研究は宇山親雄先生が行っておられ、すでに実験環境がほぼ整備されていたので、卒業式の後直ちに、実験の技能向上(ザリガニの解剖やガラス電極の作成法)とザリガニの神経系に関する文献調査(医学部、薬学部の図書館で)を行いました。

修士2回生の夏前までの約1年半の間は、色々な実験を繰り返しましたが新たな知見を得ることができず、宇山先生からは修士論文を書くにはテーマを変えた方がよいのではないかと示唆を頂きました。しかし、折角始めたことなのでもう少し続けたかったため、夏休みまでは続けてよいということで実験を続けていたところ、興味ある現象を発見することができ、一気に研究成果が得られました。後に、この成果をまとめたものが学術研究論文の第一号となりました。

この研究は、ザリガニの尾部には6つの体節があり、体節ごとに小さな神経節が存在し、各神経節がその体節の運動制御を行っていると考えられ、逃避行動を発現させる際にこれら6つの神経節がどのようなタイミングで協調しているのかを解明することが目的でした(図1)。今から考えると、これは正に「**神経系における分散協調メカニズムの解明**」を目指していたことになり、現在、今後の重点的研究課題として取り組んでいる「**コミュニケーションにおける動的インタラクション機構のモデル化・システム化**」にも通じるものであり、私の研究は学生時代から一貫して、「**分散協調**」、「**動的インタラクション**」が中核的キーワードになっていると言えます。

生物に興味を持っていた私は、当時の最先端科学技術であった情報工学、コンピュータについては違和感を持っており、多くの同級生が新設の情報工学科の講義

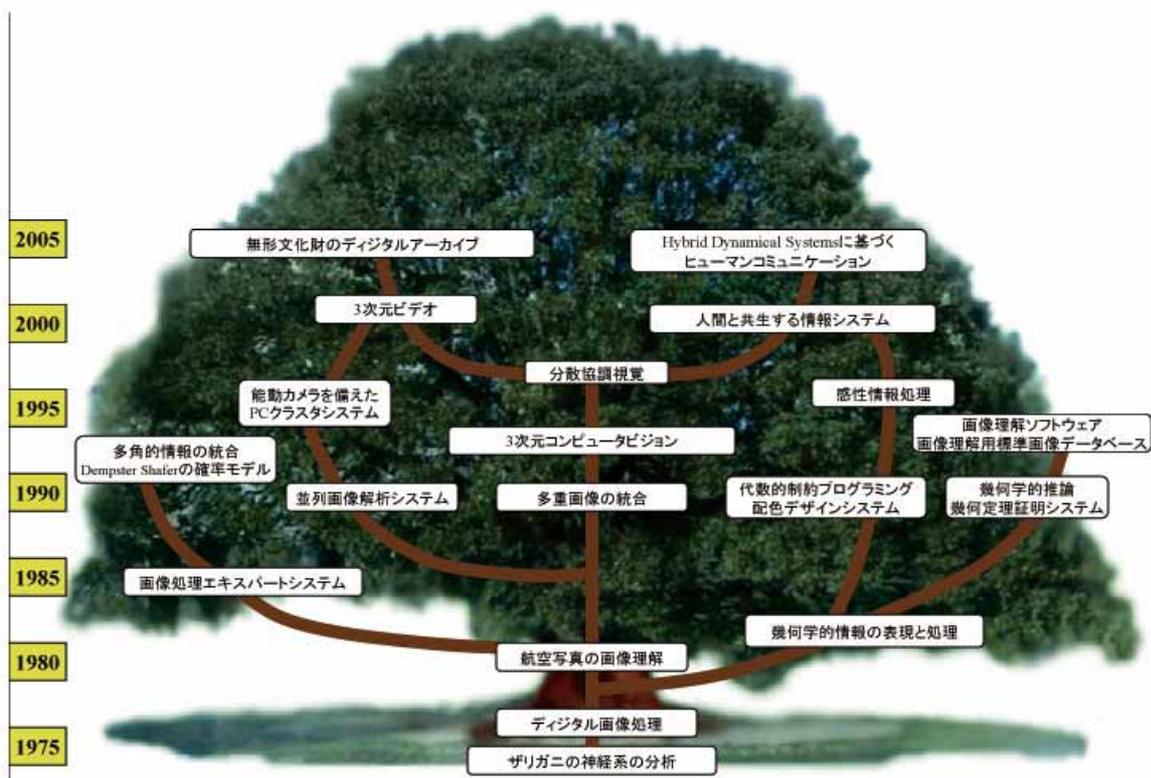
を聴講していたにも関わらず、全く勉強する気にはなれませんでした(私が京都大学に入学した1970年に国内ではじめて情報工学科が設置されました。)こうした中で、長尾真先生から修士修了後長尾研究室の助手にならないかとのお話が突然ありました。長尾先生が大学院で講義されていたパターン認識は、生物の知能につながるものがあり興味があったことと、研究者になりたいかったため、喜んでお話をお受けすることにしました。しかし、今から考えると情報工学に関する専門知識、ハードウェア技術、プログラミング技能がほとんどない状態で情報系の研究室の助手になるということはかなり無謀なことだったと思います。

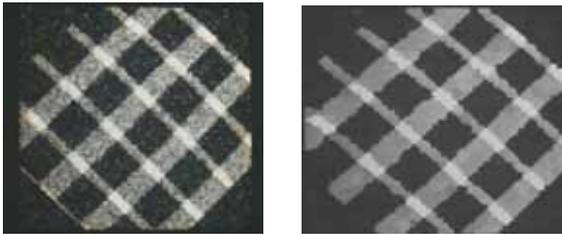
## 2. デジタル画像処理

1976年に長尾研究室助手になった当初は、とにかく情報工学、画像処理、パターン認識の勉強や研究室にあったミニコンの使用法の習得に励みましたが、当然のことながら、修士や博士の学生の方が知識、経験、技能のいずれにおいても先行していましたので、先生というより仲間として一緒に研究をしました。

研究テーマとしては、画像処理アルゴリズムの開発と当時注目を集めていたリモートセンシング画像、特に航空写真の解析に取り組みました。最初の研究成果は、「**エッジを保ったスムージング**」という画像処理アルゴリズムで、画像をぼかすことなくシャープな工

## 研究ロードマップ





入力画像 処理結果

図2 エッジを保ったスムージングによる  
雑音除去とエッジ強調

エッジを保ったまま、雑音を除去できるというものでした(図2)。これは、伝統的な線形フィルタリングでは実現できないような性能(低域強調, 高域強調フィルタを同時に実現する)を達成するには、非線形な(論理的判断を含んだ)演算が有効であることを実証し、国際会議や国際学術雑誌に掲載された論文は世界的に高い評価を得ましたし、今でもこのアルゴリズムを使って画像処理をすることがあります。

### 3. 航空写真の画像理解

リモートセンシング画像解析では、当初はLANDSAT 衛星写真や航空写真を対象として統計的なパターン分類手法の改良や画像に含まれる陰影や幾何学的歪の補正を行っていましたが、特にこれといった成果が得られたわけではなく、長尾先生からも、そろそろ別のテーマを考えてはどうかと言われていました。ザリガニの時もそうでしたが、これまでせっかく努力して来たのだからもう少し続けさせて下さいとお願いして、当時修士2回生だった池田義雄君と一緒に、航空写真から家や道路、畑、森などといった対象を自動認識する手法として、エッジを保ったスムージング領域分割 領域を単位とした解析(図3)を試みました。当時は画素を単位とした解析が中心であり、領域を単位とした解析を行えば、大きさや形状、空間的位置関係といった幾何学的情報が利用できるようになり、複雑な航空写真の認識が実現できるのではないかと考えたわけです。

この研究を1978年に京都で開催された第4回パターン認識国際会議で発表したところ、高い評価を頂き、会議に来られていたフランスのM. Nadler教授から、研究叢書として出版しないかとの申し出が長尾先生にあり、1979年末に原稿を出版社に出すことになりました。叢書の出版に向けて検討したのは、システムの制御メカニズムでした。つまり、国際会議で発表した論文は、画像解析のアルゴリズムが中心で、システムとしての構成や制御機構はまだ明確になっ



マルチスペクトル航空写真 領域分割結果

図3 領域に基づく航空写真の構造解析

いていませんでした。この検討の中で長尾先生から、米国のCMUが音声理解システムを開発しており、そのシステム構成として「黒板モデル」というアーキテクチャを提案している。我々のシステムの制御構造も「黒板モデル」として体系化できるのではないかと、このコメントを頂き、その考えに基づいてシステムの再設計を行いました。

このシステムでは、図4左図のように、黒板と呼ばれる共有メモリを介して、多種多様な解析・認識モジュールが協調して複雑な航空写真を解析するようになっており、図右のような認識結果が自動的に得られるようになっていきます。(白は未認識部分、色は対象物の種類を表しています。)

「黒板モデルに基づく航空写真の解析」に関する研究によって、情報処理学会から論文賞を頂いただけでなく、1980年には米国Plenum社から研究叢書が出版され、国際的にも大きな反響が得られ、1980年11月に京都大学から工学博士の学位を頂くことになりました。(この研究でひとつ残念だったことは、システムに名前を付けなかったことです。)

この研究は、私の研究分野を従来の画像処理、パターン認識から、人工知能の立場から画像認識を研究する「画像理解」へと展開させることとなり、「人間とは何か」、「知能とは何か」という課題に対して視覚情

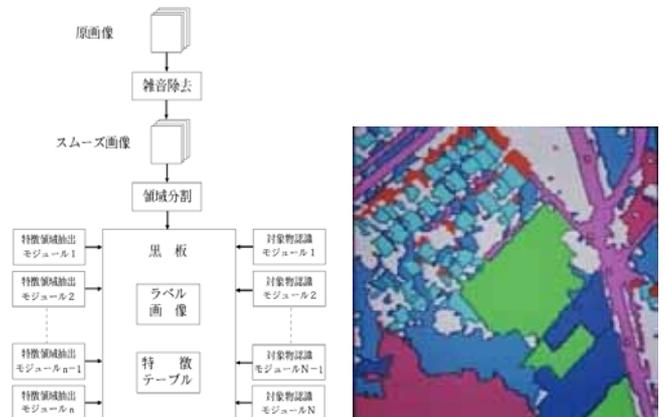


図4 黒板モデルと航空写真から認識された対象物

報処理の立場からアプローチすることが自分の研究スタンスとなりました。

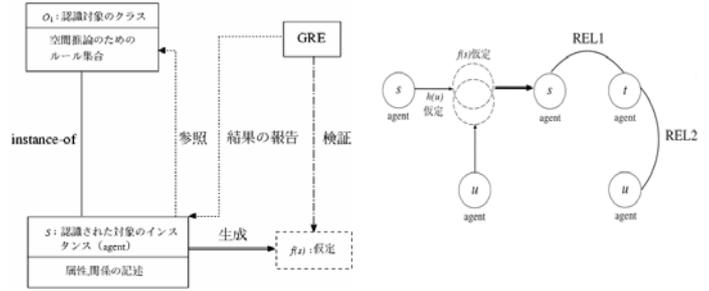
今でも博士課程の学生に言うことですが、研究者としての評価は博士論文の研究を終えた後でどのような研究を行うかということにあると思っており、1980年以降二、三年は、独自の研究テーマの開拓を色々試みました。その中心課題としては、空間的、幾何学的関係をデータ構造として表現し、処理を行うという**幾何学的情報処理**で、画像処理ではテキストの構造解析、地理情報システム関係では線図形のマッチングや幾何学的対象（領域や線）に関する情報を効率的に記録、検索するためのファイル構造、計算幾何学では線分群を対象としたデジタルボロノイ線図の計算アルゴリズムなどを試みました。残念ながら、これらの分野では大きな成果は得られませんでした。幾何学的情報の表現と処理という問題は現在に至るまで絶えず興味ある問題として研究を継続しています。

こうした中で1982年秋から米国 Maryland 大学の A. Rosenfeld 教授の研究室に客員研究員として滞在することになりました。ほぼ同時期に、CMUには大田友一さん、MITには池内克史さんがおられ、アメリカの地で家族を含めたお付き合いができたのはよい思い出となっています。

当時の米国では DARPA が Image Understanding Project を推進しており、Stanford 大学の R. Brooks が複雑な3次元対象の認識システム ACRONYM を開発し、注目を集めていました。また、人工知能の分野では、Stanford 大学の Feigenbaum 教授が提唱したエキスパートシステムやプロダクションシステムが全世界的に注目され、一般社会においても人工知能研究がもてはやされていました。

Maryland 大学では、人工知能の立場から幾何学的情報処理に取り組みようと考え、**空間推論機能を備えた航空写真理解システム SIGMA** および**画像処理エキスパートシステム LLVE** の開発を博士の学生だった V. Hwang 君と一緒に行いました。

博士論文で開発した航空写真解析システムは主として領域の色や形状情報を利用して対象の検出・認識を行っていました。一方、SIGMA では、家同士や家と道路の間の空間的関係に関する知識を利用して、家や道路が存在すると推定される位置に関する仮説を生成し、図5右のように、複数の異なった対象から同じ位置に仮説が生成されるとその信頼度が向上し、高い信頼度を持った仮説領域をトップダウン的に解析することによって、ボトムアップ的な画像解析が失敗したような場合でもうまく対象検出ができる機能を実現しました。こうした推論方式は、仮説推論の一手法とも考えられ、論理式によってその推論メカニズムの定式化



システムの構成 仮説生成に基づく空間推論  
図5 SIGMAにおける空間推論機構

を行いました。

また、SIGMA は Lisp 上でのオブジェクト指向言語 Flavor によって書かれていましたが、図5左にあるように、認識された対象を表す各オブジェクトは、知識に基づいて自律的に仮説を生成したり、仮説の検証結果に応じて他のオブジェクトとの関係の設定や未検出対象の発見を行う機能を備えており、今で言うエージェント指向システムとなっていました。

一方、画像処理エキスパートシステム LLVE は、1984年に米国から帰国したあとで修士2回生の尾崎正治君、4回生の清水周一君と一緒に開発したのですが、画像処理オペレータに関する知識（図6）を利用して試行錯誤的に線や領域といった画像特徴の抽出を行う機能を備えたシステムで、SIGMAのトップダウン画像解析を実現するだけでなく、蓄積された画像処理プログラムライブラリを活用するための利用者支援ソフトウェアとしても有効なものでした。画像処理に関する知識やノウハウを利用した画像処理エキスパートシステムの開発というアイデアは、我々のほかに電総研の田村秀行氏のグループや名古屋大学の鳥脇純一郎教授のグループでも試みられており、1986年電子通信学会にエキスパートビジョン研究会を作って研究討論を行いました。この研究会およびその前身の私的研究会（1984年）では、若手の研究者が泊り込みをして議論を交わすことを通例としており、色々な面で大きな刺激を受けました。

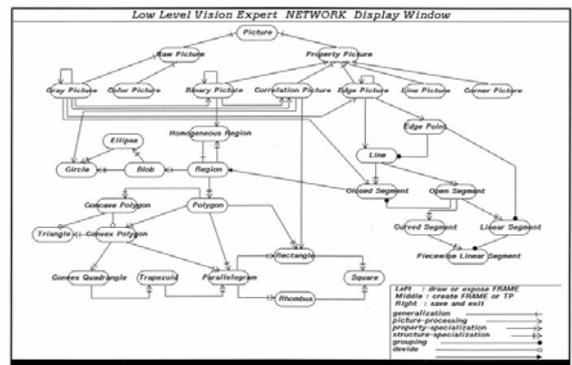


図6 LLVEにおける知識表現

画像理解システムや画像処理エキスパートシステムに関する研究は、国際的にも注目を集め、1984年モントリオールで開催された第7回パターン認識国際会議、1988年ローマで開催された第9回パターン認識国際会議にてそれぞれ招待講演を行いました。また、SIGMAに関する研究成果は1990年に再び米国 Plenum 社から研究叢書として、画像処理エキスパートシステムについては Academic Press の研究叢書の1部としてそれぞれ出版されました。

#### 4. ソフトウェア基礎論

米国から帰国して約1年半が過ぎた1985年8月には、長年住んだ京都を離れ、東北大学工学部情報工学科の伊藤貴康教授の研究室の助教授として仙台に赴任しました。伊藤研究室では、それまでの画像理解研究の取りまとめを行うとともに、伊藤先生が研究されていたソフトウェア基礎論やそのための論理学、推論システムの勉強、さらにはマイクロプロセッサを使った並列計算機、並列 Lisp 処理系の開発を学生と一緒に行いました。

これらの研究は、それまでのシステム開発を目指したものと異なり、ソフトウェア科学の基礎理論に関するものであり、当初はかなり戸惑いを覚えました。しかし、理解が深まるに連れて、計算機科学が科学として確立されたものであり、その基礎が論理学や数学基礎論にあること、ソフトウェア科学のための論理、数学として直観主義論理や構成的数学があることが分かり、研究者として一皮剥けた気がしました。

また、東北大学では、構成的数学に基づいたソフトウェア基礎論を研究されていた佐藤雅彦教授や半導体デバイス関係の優れた研究者の先生方とも交流が持て、電気通信研究所を含めた東北大学電気情報系の層の厚い優れた研究体制に触れることができたのは大変よい経験になりました。この思いから、東北大学のスローガンの1つである「研究第一主義」を研究室のスローガンとさせて頂いています。

#### 5. 多角的情報の統合

岡山大学に赴任した1989年は、情報工学科創設3年目であり、1年間は講義と学科の立ち上げ準備に費やしました。京都大学や東北大学と比べ、当時の岡山大学は、情報系の研究者、設備とも十分ではなく、新設の情報工学科が先頭を切って研究教育体制を構築しなければなりませんでした。具体的には、学科建物の建設に加え、当時大学関係で普及していた電子メールシステムの立ち上げ、学科の教育用計算機システムや学科PBXの導入、情報工学科に相応しいカリキュラムの構築などの作業に加え、最も腐心したのは情報

工学がコンピュータの使い方を教えるだけの技術教育分野ではなく、科学的基盤を持った学術的研究分野であることを学内外の人々に理解してもらうことでした。

1990年4月には、4回生が配属され、松山研究室がスタートすることになりました。また、東京工業大学博士課程を修了した和田俊和君を助手、京都大学助手であった浅田尚紀君を助教授として迎え、夏に竣工した学科新棟の広々とした研究室とともに、新研究室の体制が整いました。

研究面では、単に従来の研究を継続するだけでなく、新たな研究分野の開拓を目指した試みも積極的に展開しようという考えに基づいて、10名の4回生には多様な研究テーマを担当してもらいました。特に、東北大学で学んだ**ソフトウェア科学や並列計算に関する知見を視覚情報処理研究に活かす**ことを目指して、幾何学的推論、代数的制約プログラミング、再帰トラス結合アーキテクチャと名付けた並列計算機といった全く新しいテーマに4回生とともに取り組むことにしました。こうした新たなテーマは、4回生ではこなせないのではないかと心配でしたが、予想に反して彼らは着実に成果を挙げてくれました。このことによって、研究に対する熱意が更に高まっただけでなく、研究室の将来に大きな希望を持つことができました。

研究室ではスローガンとして「**研究第一主義**」と「**自立した社会人としての自覚**」を掲げ、教官、学生が一体となって活発な研究活動を展開しました。また、学生には「大学入試のための偏差値では東大や京大の学生には及ばなかったかもしれないが、研究に関しては今からがスタートであり、学会の論文賞を取れるぐらいの研究成果を目指して頑張ってください。」と檄を飛ばしていました。

当時、画像、人工知能関係の研究スローガンとしては、「**多重画像の統合による多機能高精度画像計測**」、「**多角的情報の統合による柔軟かつ頑健な認識**」を掲げており、これらの分野でほぼ毎年のように論文賞を受けることができ、また科学研究費も途切れることなく採択され、新設の研究室は順調に軌道に乗っていきました。

岡山大学では、個別的研究テーマとして、以下のようなものを行いました。

まず、「多角的情報の統合」を実現するための計算機構として東北大学の最後の年から **Dempster-Shafer の確率モデル**の研究を行っており、画像理解における部分全体関係に基づいた推論や統計的パターン分類の拡張に関する理論を提案し、それぞれ人工知能学会、電子情報通信学会から論文賞を頂きました。当時人工知能の分野では、不完全情報に基づいた推論として、閉世界仮説、circumscription, negation as failure などと

いった論理的推論の拡張や、Assumption-Based Truth Maintenance System といった推論方式などが活発に研究されており、そうした中で Dempster-Shafer の確率モデルの研究を行ったわけです。しかし、結局は外界の状況を能動的に観測し絶えず新たな情報を得る手段を持たない限り、幾ら高度な推論機能があっても不完全情報問題は解決できないとの結論に至り、以下に述べるようなセンサ開発、能動的観測の研究を新たに始めることにしました。つまり、2つの論文賞によって、助手時代から進めてきた画像理解の研究にピリオドを打ったわけです。

投票と多数決を基礎とした計算によって図形の検出、対象認識を行う Hough 変換については、助手時代から興味があったため、軽い気持ちで一期生の藤井高広君の卒論のテーマとして取り上げました。藤井君にプログラムを作成してもらいましたが、その結果がおかしく、和田君にチェックを頼んだところ、プログラムのバグではなく今まで知られていなかった Hough 変換の特徴であることが分かり、その後 Hough 変換関係の研究を継続的に進めていくことになりました。

幾何学的推論に関しては、一期生の新田知明君の卒論として研究を始め、代数的証明と論理的証明のそれぞれが持つ証明能力、限界を理論的に明らかにし、両者の長所を併せ持つ融合推論システムを構築し、補助線の生成プロセスを仮説推論として扱うことなどの機能拡張も試みました。こうした一連の研究によって、幾何学的対象およびそれらの間の幾何学的関係が理論的にどのようにモデル化されるのかについての理解が深まり、成果は AI Journal に掲載されましたが、京都大学への異動によって研究が中断されたままになっています。

代数的制約プログラミングについては、一期生の秋山壮太郎、田中智之君の卒論として始め、二期生の山口修君の修論で壁紙やプリントパターンのデザインを行うシステムを開発しました(図7)。この研究は、大阪大学の辻三郎教授が代表となって進められた感性情報処理に関する重点領域研究の中で、感性情報を代数的に表現、処理する方式の提案へと発展しました。

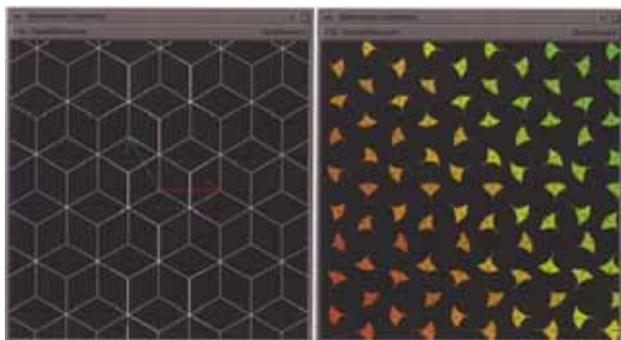


図7 代数的制約記述に基づく配色デザインの例

感性情報処理に関する重点領域研究は、1990年ごろから大阪大学の辻三郎、井口征士教授が中心となり、従来の理性的、論理的な判断に基づいたパターン認識において切り捨てられていた個性や揺らぎ、感覚的、感性的な情報に焦点を当て、その表現と処理手法を開発しようという野心的な試みで、1992年-1994年の3年間に渡って行われました。この重点領域研究では、情報メディア系と心理学系の先生方が一同に会して学際的な研究が展開され、私も計画班の一員として感性情報の意味や感性情報処理独自の処理手法の開発に取り組みました。心理学の先生方との緊密な研究交流によって、学生時代から興味を持っていた「人間とは何か」に関する問いかけを専門家の方々と直接議論することができ、大きな刺激になりました。また、そこで得た知見が現在取り組んでいる「情報学」特定領域研究へと繋がっていくこととなります。

一方、並列計算機および並列プログラミング言語については、東北大学では Lisp を対象としていたのに対して、岡山大学では画像解析の観点からアーキテクチャの考案と言語仕様の設計を進めました。アーキテクチャについては、一期生の青山正人君と一緒に再帰トラス結合アーキテクチャという、MPU 間の通信ネットワークの構造を動的に切り替えるアーキテクチャを提案し、三期生の山下敦也君がトランスピュータを使ったハードウェアの実装まで行いました(図8)。また、並列プログラミングに関しては、オブジェクト指向を拡張して並列データ構造という概念に基づいた言語を設計しました。これらの研究は、青山君の博士論文としてまとめられた後途絶えてしまいましたが、現在行っている3次元ビデオプロジェクトで開発した PC クラスタシステムおよびその上での並列3次元形状復元システムへとその流れは受け継がれています。

最近の情報系の研究はソフトウェアが中心となっ

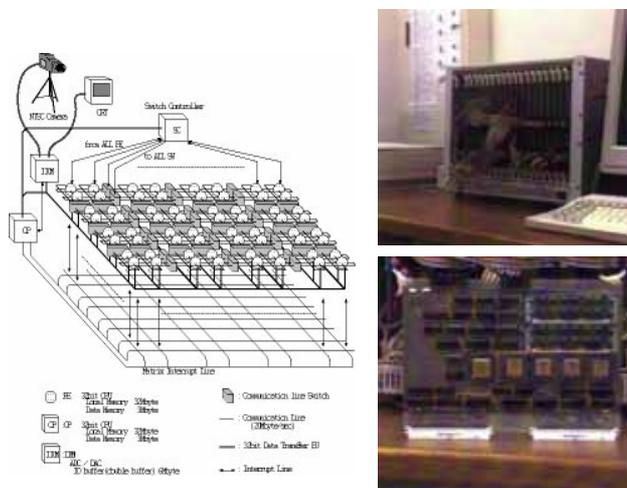


図8 再帰トラス結合アーキテクチャと RTA/I 並列計算機

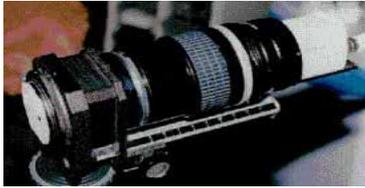
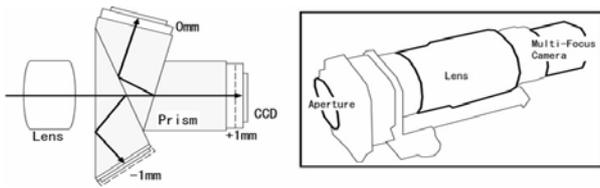


図 9 多重フォーカスカメラの撮像部と構造化瞳付き多重フォーカスカメラ

ていますが、研究室では岡山大学、京都大学を通じて、ハードウェアやその基本ソフトウェアを実際に設計、試作するという活動を継続的に行って来ています。これは、視覚情報処理においては、実世界と情報システムとの関わり合い方が重要であり、そのためのセンサやネットワーク、制御システムを自作できなければ真に意味のある研究はできないと考えているためです。

こうした考えから、「多重画像の統合」というアイデアを提案し、多重絞り画像による広ダイナミックレンジ画像の生成、多重フォーカスカメラ（図 9）による 3 次元距離画像の計測、視点固定型パン・チルトカメラ（図 10）による全方位パノラマ画像の撮影といった従来にはない機能を備えたセンサシステムを開発して来ました。前者は浅田助教授、2 番目は日浦慎作研究員、後者は和田助手が中心となって開発を進め、それぞれ情報処理学会、画像センシングシンポジウム、電子情報通信学会から論文賞を授与されました。

岡山大学以前は、画像理解や人工知能といった「ハイレベル処理」を中心に研究を行ってきた私がカメラや雲台といった「ローレベル処理」のための機器の開発に取り組んだことは、他の人からは奇異に映ったようです。しかし、先にも述べましたが、「ハイレベル処理」を長年行ってきた結果、**実世界の状況を多角的かつ能動的に観測する機能がない限り、「考える脳」だ**

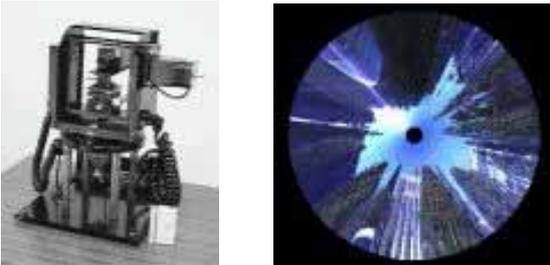


図 10 視点固定型パン・チルトカメラと京大時計台のパノラマ画像



書籍コピー 入力画像 復元画像

図 11 湾曲した書籍表面の歪のないコピー

けのシステムでは真の知能システムは構成できないという考えに至り、そのためにセンサ開発を行ったわけです。この考えは今でも変わらず、今後もこの考えに基づいて新たな研究を展開しようと考えています。

岡山大学時代の研究で最も高い評価を得たものに書籍表面の 3 次元形状復元に関するものがあります。これは、コンピュータビジョンにおける Shape from Shading の原理を拡張し、湾曲した書籍表面の形状復元（図 11）という実世界環境の下で有効に機能するアルゴリズムを考案したもので、「スマートではあるが役に立たないコンピュータビジョンアルゴリズム」という批判に答えたことによって、1995年に同分野で最も権威ある The Marr Prize を受賞しました。

## 6. 画像理解ソフトウェア、標準画像データベース

岡山大学時代の後半から京都大学時代の当初に掛けて行った活動として、画像理解ソフトウェアと標準画像データベースの構築があります。

画像処理ソフトウェアについては、長尾研究室関係の先輩で当時電子技術総合研究所の研究員であった田村秀行氏が 1980 年に開発された FORTRAN の画像処理サブルーチンライブラリ SPIDER にソフトを提供し、その後 1986 年に公開された SPIDERII の開発では仕様策定やサブルーチン作成を行いました。

岡山大学時代には、1989 年から国際パターン認識連合の TC5(Benchmarking and Software)の Chairman として、ワークステーション上での画像処理ソフトウェア環境に関する調査研究や、その一環としてキャノンに移られた田村氏が開発されていた View-Station の仕様の検討や評価を行いました。後にこの活動や画像理解研究が評価され、1996年に国際パターン認識連合からフェローの称号を頂きました。

1993年の夏か秋だったと思いますが、米国 DARPA の Image Understanding プロジェクトの一環として、IUE(Image Understanding Environments)という画像理解のための統合ソフトウェアの開発が行われており、日本にも協力してほしいとの申し入れが日本の画

像処理関係者にありました。相談の結果、TC5の活動経験もあったため、私に対応することになりました。

IUEはオブジェクト指向言語C++によって、画像理解、コンピュータビジョンにおける各種のデータ構造の標準化を図り、それに基づいて解析・認識アルゴリズムを開発することによって、ソフトウェアの可搬性、再利用性を向上させようとするものでした。こうしたソフトウェアの設計や仕様策定については、東北大学で得た知見が大いに役立ちました。

IUEの議論はImage Understanding Workshopの場で行われていましたので、九州大学の谷口倫一郎助教授とともに参加し、開発されたソフトを日本でも使えるようにしてテストや評価を行いました。

議論が進む中で、日本独自の貢献をどうするかを検討した結果、ソフトウェア開発は米国に任せ、画像理解用の標準画像データベース(Calibrated Image Database)を作ることになり、1992年から国のプロジェクトとして進められていたRWC(Real World Computing)の支援を受けて1994年からデータベースを開発することにしました。

画像処理用の標準画像データベースとしては、1970年代後半に東京大学生産技術研究所の尾上守夫教授が開発されたSIDBAがありました。SIDBAは画像データのコレクションとして広く利用されましたが、ステレオ解析などの3次元形状復元や動画像処理、色彩解析といったコンピュータビジョン研究のためには、画像が撮影された際のカメラキャリブレーションのデータや撮影対象の正確な位置、大きさといった情報が必要となります。このため、Calibrated Image Databaseでは、キャリブレーション情報付の由緒正しい画像データを撮影し、データベース化することにしました。完成した画像データベースは、CD-ROM数十枚の膨大なものになりましたが、1999年、2000年に配布希望を募ったところ100を越える機関から希望があり、増刷をして配りました。

こうしたソフトウェアやデータベース作りは、直ちに研究成果に繋がるものではありませんが、研究分野の発展には不可欠な活動であり、我々の活動が画像理解、コンピュータビジョン分野の発展に少しでも貢献できていれば幸いです。

## 7. 分散協調視覚プロジェクト

1995年に京都大学に戻った直後の1996年から、政府による新たな研究推進制度の1つとして日本学術振興会未来開拓学術研究推進事業が始められることになり、理工系分野のプロジェクトとして「分散協調視覚による動的3次元状況理解」を開始することになりました。「分散協調視覚」、すなわち、能動的な



図12 分散協調視覚システム

カメラ群をネットワークで結合したシステム(今で言うセンサネットワーク)によって実世界の動的状況をリアルタイムで計測・認識し、得られた結果を分かりやすく映像化する研究(図12)については、岡山大学のころから構想を暖めていたもので、提案書の作成や研究計画の立案などはごく短い時間でまとめることができ、10月からプロジェクトが開始されました。

このプロジェクトは年間予算1億円で5年間という、それまでにはなかった規模のプロジェクトで、現在では広く行われているプロジェクト研究制度のさきがけであり、ポスドクの雇用や外部委員会による中間、最終評価など従来にはない制度が導入されました。我々のプロジェクトでは、1997年から、和田君に助教授として京都大学に来てもらうとともに、大阪大学の博士課程を短期修了した日浦慎作君をポスドク研究員として雇用して研究体制の充実を図りました。さらに、国際的学术交流を推進し研究成果を広く社会に知ってもらうために、1997年10月には第一回国際ワークショップを開催し、2001年3月のプロジェクト終了まで合計4回の国際ワークショップを開催しました。こうした研究室主催のワークショップは、準備が大変でしたが、最先端の研究に触れ、著名な研究者の前で自分の研究のデモを行う機会が毎年のようにあったことは、若手研究者や学生達の成長に大いに役立ったと思います。

研究面では、照明変化に頑健な背景差分による対象検出法、視点固定型パン・チルト・ズームカメラによるリアルタイム人物追跡、岡山大学から開発を続けてきた多重フォーカスカメラに構造化瞳を付加して高精度かつリアルタイムに3次元距離が計測できるシステム(図9)などの要素技術が開発され、特許や論文賞を頂きました。

分散協調視覚プロジェクトで最も力を入れたのは、視覚、行動(カメラ制御)、ネットワーク通信の3つの機能をリアルタイムで統合し、実世界で有効に機能す

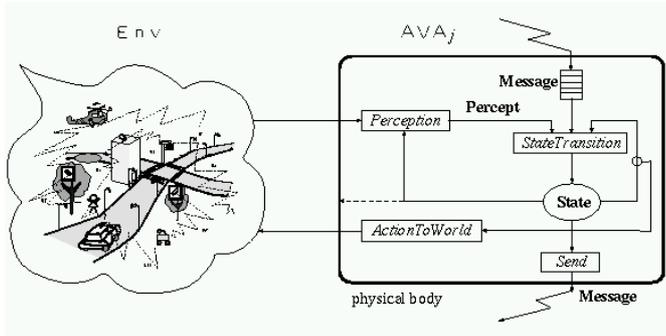


図 13 能動視覚エージェントのモデル

る「能動視覚エージェント」を設計，開発することでした(図13)。これは，先に述べた，実世界に対して能動的に働き掛けを行いリアルタイムに情報を獲得することによって初めて知能システムが実現できるという考えの妥当性を示すことを狙ったものでした。つまり，従来の人工知能研究が計算論に基づいた知能のモデル化として

$$\text{知能} = \text{知識} + \text{推論}$$

というスキーム (Computational Intelligence) を基礎としていたのに対し，分散協調視覚では，実世界や他のエージェントとのインタラクションに基づいた

$$\text{知能} = \text{知覚} \quad \text{行動} \quad \text{コミュニケーション}$$

というスキーム (Interactive Intelligence) を基礎として知能システムを構築することを目指しています。ここで，はリアルタイムの動的インタラクションを表しています。

プロジェクトでは，まず首振りカメラによる人物のリアルタイム追跡システムを設計する中で，**ダイナミックメモリ** (図14) という並列プロセス間のリアルタイム情報交換用共有メモリを考案し，その有効性を実証しました。これは，簡単なメカニズムですが非常に有効なアイデアであるとして，プロジェクトの評価委員会でも高い評価を頂きました。

次に，分散配置された能動視覚エージェント群によって，複数人物のリアルタイム追跡を行うシステム

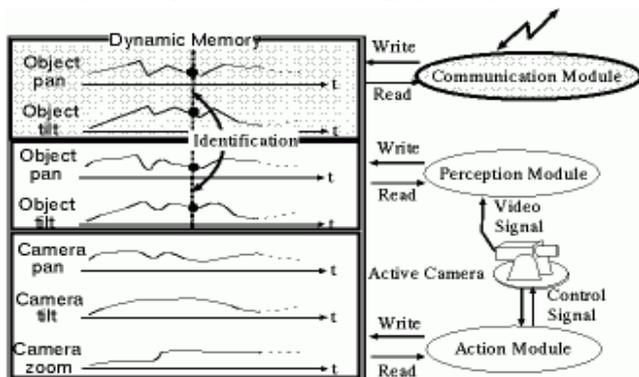


図 14 ダイナミックメモリを用いた視覚，行動，通信モジュールのリアルタイム統合

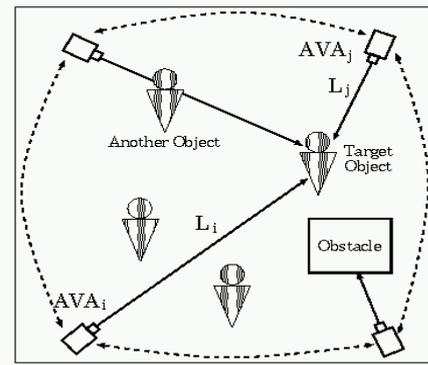


図 15 能動視覚エージェント群の協調による複数人物のリアルタイム追跡

の設計に取り組み，エージェント間の協調プロトコルを考案しました(図15)。この研究は，浮田宗伯君が博士論文の研究として取り組んでくれたもので，国際学術雑誌でのサーベイ論文の依頼を受けるなど高い評価を受けることができ，同君が奈良先端科学技術大学院大学の助手となった後も同君がリーダーとなってプロジェクト研究を進める基盤となりました。

2001年3月にプロジェクトは終了し，事後評価では高い評価を得ることができました。また，研究室ではプロジェクトで得た成果を基に更なる研究の展開を進めることになりました。

### 8.3 次元ビデオ

分散協調視覚プロジェクトでは，そのシステム基盤として，能動カメラを備えた PC クラスタの開発を行ってきました(図16)。これは，岡山大学で開発した並列画像解析システムの後継機として設計，開発したもので，Gbit レベルの高速ネットワーク Myrinet によって首振りカメラを備えた PC 群を結合するという，当時としては他に類をみない最先端システムでした。

1990年代後半，米国のカーネギーメロン大学や

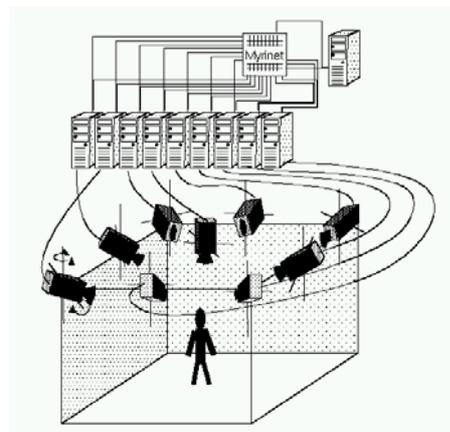


図 16 能動カメラを備えた PC クラスタシステム

カリフォルニア大学サンディエゴ校では、多視点から人物の動作をビデオ撮影し、人物の3次元映像を生成する研究が行われていましたが、処理はオフラインで膨大な計算時間を要していました。そこで、我々の研究室ではPCクラスタを用いてリアルタイムに3次元形状を復元することができれば、人物の詳細な3次元動作解析やその3次元映像化を効率よく行うことができ、またスターウォーズのような立体映像通信が可能となると考え、1998年ごろから3次元ビデオに関する研究をスタートさせました。

当初は、カメラの正確なキャリブレーション法の開発や並列3次元形状復元ソフトウェアの開発に苦労しましたが、研究を重ねるうちに徐々に精度や計算スピードが向上され、2004年度にウ君の博士論文研究として開発された現在稼働中の3代目PCクラスタでは、8mm角の解像度で毎秒30フレームのスピードで3次元動作が復元できるようになりました。また、延原君の博士論文研究では、凹凸のある人体の3次元形状を正確に復元するための弾性メッシュモデルが開発され、舞妓さんの踊りのように袖や裾、帯が複雑に揺れ動くような場合でも、正確に3次元動作を復元できるようになりました。

一方、3次元の映像化については、視点固定型パン・チルトカメラを用いた高解像度全方位パノラマ画像(図10)や6眼ビデオカメラによる全方位ビデオの撮影とドーム型ディスプレイ(図17)を使ったインタラクティブな没入感表示、復元された3次元形状への高精度テクスチャマッピング、さらには全方位画像と3次元ビデオを素材とした映像コンテンツの編集(図18)、立体ディスプレイを用いた3次元ビデオのインタラクティブ表示といった各種のソフトウェアが高井君の博士論文研究で開発されました。図19は、3次元ビデオ映像化技術の進展過程を表しています。

これらの基盤技術を基に、2004年から5年間の計画で、文化財の高精度デジタル化を目標とするプロジェクト研究(文部科学省プロジェクト「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」)



図17 ドーム型ディスプレイ



図18 全方位画像と3次元ビデオを編集した映像コンテンツ



2000年2月 2000年夏 2001年旧方式  
2001年新方式 2004年(宮崎スタジオ)  
2005年(宮崎スタジオ) 2005年(西陣スタジオ)

図19 3次元ビデオの進化過程

を池内克史東大教授と一緒にすることになりました。このプロジェクトでは、研究開発に加え、3次元ビデオの技術を広く社会に普及させるため、デモやソフトウェアの企業へのライセンスを行う一方、波部斉助手が中心となって全方位ビデオや3次元ビデオの符号化に関する国際標準化活動を行っており、2006年の中間評価では高い評価を受けました。

## 9. 人間と共生する情報システム

3次元ビデオ研究と並行して、2001年度からは、5年間の計画で、慶応大学の安西祐一郎教授を領域代表者とした大型の特定領域研究「ITの深化の基盤を拓く情報学研究」が開始され、6つある柱の内の1つであるA03柱「人間の情報処理の理解とその応用に関する研究」の柱長をすることになりました。この特定領域研究では、分散協調視覚プロジェクトによって開発されたセンサやシステム基盤を踏まえて、岡山大学時代から思い続けてきた「人間とは何か」や「人間と情報システムとの関わりあい方」に関する研究に本格的に取り組むことになりました。

A03柱には4つの計画研究があり、20前後の公募研究のグループが採択され、また専門分野も情報学だけでなく神経生理学や心理学、ロボット工学など幅広い学際的なグループを形成しています。こうした大規模かつ広範な研究分野の研究者の方々をうまくまとめ共通した問題意識を持って頂くにはどうしたらよいかを考えた末、A03柱の研究ゴールを「人間と共生

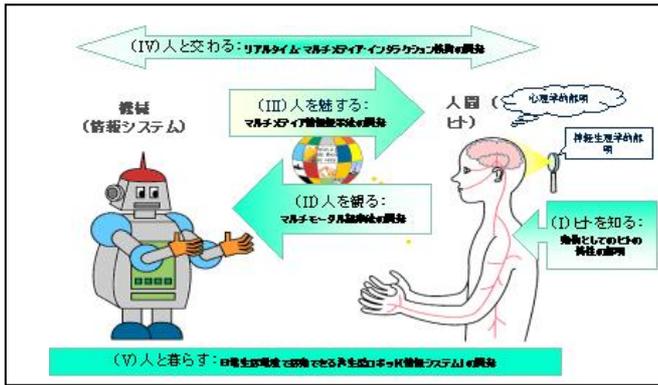


図 20 人間と共生する情報システム実現のための 5つの研究課題

する情報システム」の実現とし、ヒトを知る、人を観る、人を魅する、人と暮らす、人と交わる、人と暮らす、という5つの研究課題を立ててグループ研究を進めることにしました(図20)。

研究室では、まずヒトを知る、人を観る、人を魅するに関する研究として、Ubiquitous and Wearable Vision Systems という視点から、分散協調視覚、3次元ビデオ研究で開発したPCクラスタを用いた環境埋め込み型ビジョンシステム(図16)と図21のような視線検出装置と能動ステレオカメラを備えた装着型能動視覚センサを開発し、様々な状況における人間の3次元的行動・仕草の計測とその意図理解に関する研究を展開しました。



図 21 装着型能動視覚センサ

### 10. Hybrid Dynamical System

この特定領域研究を推進する中で得られた大きな知見として、人間同士、人間と情報システムとの間で円滑なコミュニケーション(人と交わる)を実現するには、従来のような意味情報やマルチメディア情報の利用だけでなく、間の取り方や相槌のタイミングといったインタラクションのダイナミクスが重要であるということがあります。A03柱には、数理モデルやロボットの行動生成さらには人間の時間感覚に関する心理学的研究をされている先生方がおられ、そうした方々も皆ダイナミクスの重要性を指摘されており、次第にそれをA03柱の中核的研究課題として研究を

展開することとなりました。

ダイナミクスを研究課題とすることについては、かなり以前から考えていたことでした。つまり、従来のコンピュータビジョンでは、幾何学と光学を基本として2次元画像からの3次元情報の復元を行っていましたが、対象の持つダイナミクスやシステムのダイナミクスについてはあまり深い検討はなされていませんでした。このため、先の分散協調視覚プロジェクトでは、実世界でリアルタイムに機能する能動視覚エージェントを実現するための視覚、行動、通信モジュール間のダイナミクスという観点からの研究を行いました。しかし、プロジェクトを推進する中で、この研究は情報システムの設計論としてダイナミクスを取り上げたもので、メディアやロボット、人間自身が持つダイナミクスのモデル化や、人間同士、人間と情報システムとのコミュニケーションにおけるダイナミクスの分析といったより深いレベルでの研究が必要であると考えていました。

そこで、1998年度の4回生であった川嶋宏彰君と一緒に研究を始め、現在では、Hybrid Dynamical System という新たな数理モデルの構築と、コミュニケーションにおけるダイナミクスの分析とモデル化・システム化に重点をおいた研究を展開しており、2006年度から新たに始められた特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」(領域代表：喜連川優東京大学教授)の計画研究として研究を進めています。

Hybrid Dynamical System は、物理的現象記述に適した力学系モデル(連続的計量空間における状態遷移を記述する微分方程式系:N.Wiener が提唱したサイバネティクス)と、人間の心的・知的活動の記述に適した情報系モデル(離散的状态遷移を記述する記憶書き換え系:A. Turing の考案したチューリング機械)を統合した計算・制御モデルの実現を目指したもので、

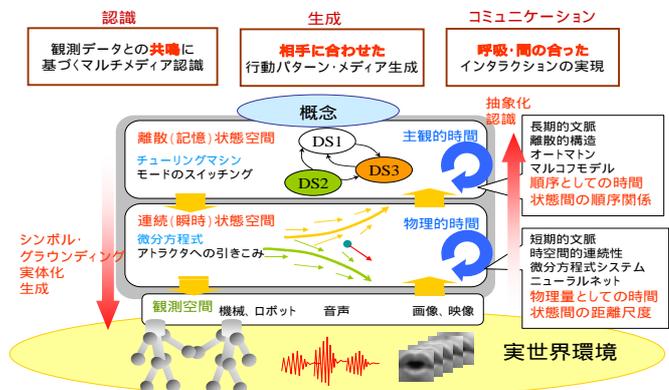


図 22 Hybrid Dynamical System の構成

新たな機能・特性を持ったインタラクティブシステムを生み出すことができると考えています。

これまでに得られた研究成果としては、以下のものがあります。

- (1) 有限の時区間における音声、映像を表す時系列信号のダイナミクスを表現する力学系モデルとして線形システム、人間の発話行動やジェスチャの全体的構造を記述する情報モデルとして確率オートマトンを用いた **Linear Interval-based Hybrid Dynamical System (LIHDS)** を考案し、発話中の口の動きのモデル化およびモデルに基づく映像生成を行った。
- (2) 顔を構成する各部品動きを LIHDS によって表現し、部品間の動きの相互タイミング構造に基づいて自然な笑顔と作り笑いが識別できることを示した。
- (3) 発話における音声と口の動きをそれぞれ LIHDS によって表現し、両メディア間に存在する相互依存的タイミング構造をモデル化することにより、音声信号からそれに同期した口の動きを映像として生成できることを示した。
- (4) 漫才における発話交代のタイミングを分析し、テンポのよい対話や話の盛り上がり感を表現するには、相手の発話終了前に発話を開始するという「負の間合い」が重要であることを明らかにした。

今後は、LIHDS の理論的拡張を進めるとともに、情報システムが人間とスムーズかつ意味のあるインタラクションを継続できるようにするために必要と考えられる相互適応機能の実現を目指して研究を深化させたいと考えています。

## 11. 「こと」の科学への路

以上30年間の研究の歩みを述べて来ましたが、今を含め常に何かを成し遂げたという達成感はなく、絶えず新たな課題が見つかりそれに挑戦し続けてきたというのが実感です。

数年前からは、「こと」の科学を立ち上げたいというのが最も大きな目標となっています。つまり、20世紀までの自然科学は物理学、化学に代表されるように「もの」とは何かを探求する「もの」の科学であったということができると思います。この考え方は、宇宙や生物をも「もの」として表現、分析できることを実証し、その普遍性、有効性を確立してきました。そして、**コンピュータビジョンも幾何学、光学を基に、対象世界を「もの」として理解することを目指してきたと言えます。**

一方、情報とは何か、知能とは何か、あるいは生命と

は何かといった現代的課題に対しては「もの」の科学ではなく、「こと」の科学といった新たな科学哲学が必要なのではないかというのが私の見解です。

「もの」を表現するモデルとして分子や原子、量子があり、それらの組み合わせによって多種多様な「もの」が表現でき、要素分解や合成の技術が次々と作られてきたというのが「もの」の科学の歴史だと考えられます。

これに対し、情報や知能、生命といった対象を表現するには、要素分解・合成といった考え方ではなく、要素間の関係性やそれらの動的係わり合いに焦点を当てたモデル化が必要なのではないかと考えられます。

先の述べた Hybrid Dynamical System に基づいたインタラクション、コミュニケーションの研究は、「こと」の科学への路を切り拓く1つのアプローチとして行っていますが、「こと」とは何か、「こと」と「もの」との関係については、まだ十分に理解すらできておらず、私としては今後の研究生活の中心に「こと」の科学の創成を置き、研究を積み重ねて行きたいと思っています。

(ここで紹介しました研究については、ホームページ <http://vision.kuee.kyoto-u.ac.jp/japanese/happyou/ronbun.html> に論文リストが載せてありますので、詳細な内容はそれらの論文を参照下さい。)