

実世界情報システムプロジェクト ～ ネオサイバネティクス研究グループ ～

安藤 繁 石川 正俊 嵯峨山 茂樹 酒向 慎司 篠田 裕之 並木 明夫 眞溪 歩
情報理工学系研究科システム情報学専攻

廣瀬 啓吉 峯松 信明
情報理工学系研究科電子情報学専攻

概要

実世界情報システムプロジェクトの中の「ネオサイバネティクス」(略称:NCグループ、別名:視聴覚グループ)について活動経過を報告する。人間に代わって働き、人間と対話し、人間に見せて聞かせる機械あるいは環境を構成するために、音響、音声、聴覚、視覚、触覚情報などの物理量・感覚量のセンシング、認識と理解、制御と呈示の技術を開発する。

1 本グループの方向性

本グループでは、視聴覚情報などの人間情報処理を中心に、信号処理、パターン情報処理、画像処理、音響処理などの分野で先進的な研究を進めることを目的としている。

本グループは「ネオサイバネティクス」(Neo-Cybernetics Group) と称して研究教育活動を行っている。これは、N. Wiener によって提唱されたサイバネティクスに対し、現代では Cyberspace などの語のように cyber- に仮想空間や仮想現実などの意味も付与されていることから、より情報論的な新サイバネティクスと言う意味で Neo-Cybernetics (NC) と称することにした。広く感覚量や物理量を情報処理の対象として捉えて、技術開発を行なうものである。研究内容は多岐に渡っており、人間の視聴触覚の知覚と認識から、行動制御、感覚行動統合までを含む。

人間は、実世界の情報に囲まれてその中で生きている。いわゆる五感と呼ばれる視覚・聴覚・味覚・嗅覚・触覚や温感、湿感、力覚などの各情報を取り込み、認識理解し、それに基づいて思考行動する。これに関して、研究の方向は二つに分かれる。一方は、このような人間と同様に実世界情報を知覚・認識・理解する能力を機械に与えることである。もう一方は、このような情報チャネルと能力を持つ人間に対して機械から情報を呈示し、あるいは機械自ら行動制御することである。これらにより、人間と対話する機械や環境、人間に代わって仕事をする機械、人間のための生活空間で自律的に機能する機械などが生み出される。

近年、人間型ロボットに代表されるような人間の身体や小脳に対応する機械工学が注目されているが、当グループは人間の感覚や大脳に対応する情報工学を志向している。両者のバランスの良い発達が無ければ、有用な人間的機械の出現は望めないことは明らかだろう。

実世界情報処理プロジェクトでは、未来の人間の生活および作業環境において、人間と対話・協調する機械・環境を提供する、そのための技術開発研究を目標の一つとして目指している。その中で、上記のような物理量・感覚量のセンシング、認識と理解、制御と呈示などは、このような機械の将来像において極めて重要な意味を持つ。

本グループのゴールは、視覚・聴覚・触覚などの機能と能力を機械に持たせることである。本 COE は、そのようなゴールに向かって、グループメン

バーがいままで培って来た研究のポテンシャルを
発展させ統合して行く過程である。現在は、それ
を現実的な可能性と制約の中で具体的な研究推進
の構想を議論し、新たにスタートさせる研究のた
めの研究設備等を導入している段階である。研究
推進の詳細は遂行中に逐次決定し、最終的なゴール
を目指す。

2 本グループのメンバー構成

現時点では、本 COE メンバーの中で、7名の
メンバーがこのグループに属して活動している。
以下に代表的な専門分野とともに示す。

- 安藤 繁 教授 (システム情報学専攻)
視覚, 聴覚情報のセンシング
- 石川 正俊 教授 (システム情報学専攻)
ビジョンチップ、センサフュージョン
- 嵯峨山 茂樹 教授 [主査] (システム情報学専攻)
音声認識・対話、音楽情報処理、文字認識
- 酒向 慎司 特任助手 (システム情報学専攻)
音声合成・対話、エージェントシステム統合
- 篠田 裕之 助教授 (システム情報学専攻)
触覚センサ、二次元通信
- 並木 明夫 講師 (システム情報学専攻)
ビジョンサーボ、ロボットハンド
- 眞溪 歩 助教授 (新領域創成科学研究科 複雑
理工学専攻)
多チャンネル信号処理、音響信号処理
- 廣瀬 啓吉 教授 (電子情報工学専攻) (2005 年
より)
音声情報処理、音声合成、音響信号処理
- 峯松 信明 助教授 (電子情報工学専攻) (2004
年まで)
音声情報処理、音声合成、言語学習・音声学

この他に、RA(Research Assistant) 1名:

- 米田 隆一 (システム情報学専攻 博士課程学
生)

が、本グループで活動した。

3 本グループの統合研究構想

本グループは、高いレベルの特徴ある個々の研
究の集合体でもあるが、同時に本 COE プログラ
ムにおいてそれらの技術を融合して具体的なシス
テムとして実現し、個々の技術の有用性を立証す
ることも目標にしている。個々の技術の組合せか
ら生じるシステム構想は数多くあるが、以下にそ
の一例を示す。



図 1: 案内ロボット構想 (融合システム例)

上の図に示すのは、たとえば博物館内を案内す
るシステムである。音声対話擬人化エージェント
を搭載した大型のディスプレイが自走機能を持ち、
ユーザと対話をする。この機器や環境に埋め込ま
れたマイクロフォンアレイによる音響処理、音源
定位、高速視覚機能により、来客を先導したり追
従したりする。

本グループでは、グループ内で保有する技術を
生かしつつ、今後の実世界情報システムの方向性
を打ち出して、未来志向の研究開発を行うべく、
対話型案内ロボットを統合技術のターゲットに選
び、その設計と要素技術の開発を行っている。そ
の主要技術要素は

- 音声認識、音声合成、対話制御、顔画像生成
- 全方位移動台車、プロジェクタとスクリーン

- 音源定位センサ、音源分離システム、位置センサ
- 高速ビジョンシステム、ジェスチャ認識

などである。これらを組み合わせて、人々を引率案内し、展示物の紹介説明を行う、感覚・対話機能を志向したセンサとソフトウェア中心のロボットを実現する。

本グループでは、昨年度に引続きこの統合構想に向けて、構想の具体化の議論、詳細な実現法の検討、設計、必要な要素技術の開発、必要部品の購入とシステムの組み立てを進めた。

4 本グループの本年度の研究成果

メンバーらによる各研究ユニット(研究活動単位)からの報告は、本報告書の後の節に掲載する。

本COEプログラムでは、21世紀の新しい情報学を構築するためには、人間のまわりに遍在するヒューマノイド、エージェント、ユビキタスデバイスが、人間とともに生き、人間を支える情報システム(実世界情報システム)の構築が必要であると考えている。その中で、実世界情報システムプロジェクトは、その実現を目指して、実世界情報学を展開することを念頭において、人と新しいインタラクションをする知的環境の構築を進めている。

4.1 音声認識・音声言語対話統合・音楽の研究

(1) 対話型案内ロボットの開発

全方位移動車に搭載したディスプレイに表示された擬人化エージェントと人間との間で音声対話を行うシステムを開発中である。そのためのセンサ情報、言語情報統合アーキテクチャの検討を行い、基本設計を行った。

(2) 音声対話擬人化エージェントの身体動作の精密化

擬人化エージェントの表現力向上を目指して、従来開発してきたツールキット Galatea の

エージェントの心的状態を視線移動などの挙動に結びつける力学モデルを提唱し、構築・評価を行った。

(3) 残響や雑音に頑健な音声認識に関する研究

擬人化エージェントとの対話を前提とし、特に実環境における音声認識の性能向上を目指し、雑音あるいは残響に対する音響モデルの適応アルゴリズムを考案した。また、新しいマイクロホンアレイ信号処理方式 CSCC(Complex Spectrum Circle Centroid)法を考案し、その高い性能を実証した。

(4) 音楽信号解析、音楽情報処理

複数楽器で演奏された音楽信号など、複数のピッチを持つ信号の解析法の研究が大きく進展した。Harmonic Clustering 法は、複数ピッチ信号のスペクトルを調波構造を持つクラスタへクラスタリングする手法である。また、Specmurt 法は spectrum の周波数軸を対数に変換してそのフーリエ変換領域での演算により高調波を抑圧する手法である。また、時空間クラスタリング法は、時空間に分散するエネルギーを、音響オブジェクトを調波構造と時間継続性を持つ個々音へとクラスタリングする手法である。これら3種の新手法はともに従来を大きく上回る解析能力を示した。また、音楽の自動編曲のための確率モデルに基づく自動対位法が大きく進展した。

4.2 音源定位・音源分離に関する研究

本グループでは、実世界で活用可能な音源定位センサと音源分離システムの実現を目指している。

(1) マイクロフォンアレイ計測の研究

マイクロフォンアレイ計測による音源分離の現実的な応用例として、特定の場所に存在する話者の音声抽出を検討した。問題設定として2音源の2マイクロフォン計測による音源分離を取り上げた。この条件設定では最適化(評価関数設定と収束演算)を伴わないパー

チャルなICAによって音源分離が可能であることを示した。

(2) 実時間3次元音源定位システムの開発

観測点で観測される時空間勾配信号が作る線型空間の階数と音源配置の関係を明らかにし、観測信号から未知の音源の方向や距離を求めるアルゴリズムを構築した。6個のマイクロフォンを近接配置したアレイを用いて、提案アルゴリズムの有効性特に高速性を確認した。

(3) 時空間勾配計測による指向性制御・音源分離

時空間勾配の荷重和は、音源信号とその時間微分に対して異なる指向特性を持つフィルタの和として働くことを明らかにし、指向特性を具体的に求め、荷重特性の設計法を明らかにした。実環境で音源分離が可能であることを明らかにした。

4.3 高速視覚センサに関する研究

(1) 高速ビジョンネットワークの開発

毎秒1000フレームの画像処理能力を持つ高速ビジョンシステムを複数台統合して、人間の動きや意図を人間が行動を終了する前に察知するようなセンシングシステムを開発した。人を超える性能を提示するデモンストラーションを実現した。

4.4 音声合成・音声学の研究

(1) 構造不変の定理に基づく音声の構造的表象

昨年度提唱した音声の音響的普遍構造の表象の数学的解釈を解き明かした。相対論・量子論・情報論を融合させることで構造不変の定理を導出し、不可避的に混入する非言語情報を表現する次元を消失させることで音声言語ゲシュタルトを定義した。

5 結び

—昨年度に開始された21世紀COE「情報科学技術戦略コア」のもとで実世界情報システムプロジェクト中の一グループとして活動した「視聴覚研究グループ」の状況を報告した。メンバーの専門性と独創性を生かしながら、それらを統合したシステムの構築に邁進中である。今後は、統合システムの構築とともに、それをプラットフォームとした要素技術の新規開発と評価を進める方針である。

(文責: 嵯峨山茂樹)