

実世界情報システム 視聴覚グループ

グループ主査: 嵯峨山茂樹

(情報理工学系研究科システム情報学専攻)

概要

実世界情報システムプロジェクトの中の「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 はじめに

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。

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ンネルと能力を持つ人間に対して機械から情報を呈示することである。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。

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。

本グループは「視聴覚グループ」と称して研究教育活動を行なうが、これは略称・仮称であり、上記のような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。

2 グループ構成

現時点では、本 COE メンバーの中で、以下のメンバーがこのグループに属して活動している。

- 安藤 繁 教授 (システム情報学専攻)
- 石川 正俊 教授 (システム情報学専攻)
- 嵯峨山 茂樹 教授 (システム情報学専攻)
- 篠田 浩一 助教授 (システム情報学専攻)
- 篠田 裕之 助教授 (システム情報学専攻)
- 橋本 浩一 助教授 (システム情報学専攻)
- 眞溪 歩 助教授 (新領域創成科学研究科 複雑理工学専攻)
- 峯松 信明 助教授 (電子情報工学専攻)

このメンバー構成および名称は、今後、よりよい研究推進のために変更する可能性もある。

これらのメンバーの所属する研究室およびその研究内容を以下に簡単に示す。

- システム情報第1研究室 (嵯峨山、篠田 (浩))

– 音声・音楽

音声では、話者や雑音環境への適応などを含む音声認識アルゴリズムや、音声変換・加工などを研究している。音楽では、確率モデルと統計学習を用いて、

- 音楽の自動採譜・解析などの新しい情報処理分野を開拓している。
- ヒューマンインタフェース
ユーザと音声対話する擬人化エージェント(バーチャル人間)を、音声認識・音声合成・顔画像合成の有機的な組み合わせにより実現している。
- システム情報第3研究室(安藤、篠田(裕))
 - 視覚, 聴覚, 触覚のセンシング
人間や生物の感覚にヒントを得たセンサと知覚システムの研究. センサの構造や信号検出方法を工夫し, 高機能なセンサや, 人間のような知覚特性をもつセンサを開発.
 - 時間相関型イメージセンサとその応用
全ての画素が, 受光強度と参照信号との時間相関を出力可能なイメージセンサ. 照明と参照信号の同期を工夫することで, 三次元形状から, 振動, 反射率のスペクトル, 偏光特性のパターンなどまで計測できる.
 - 人工皮膚と二次元の通信
センサや機能素子を2次元面内に分布させ, それらを個別配線なしに自在に通信させる技術. 皮膚のように莫大な数のセンサや機能素子が分布したデバイスを実現.
 - システム情報第4研究室(眞溪)
 - 計測信号処理の応用
脳波を用いたブレインコンピュータインタフェース, マイクロフォンアレイによる多重音声の分離, 非線形音響を利用した組織性状計測など.
 - システム情報第6研究室(石川、橋本)
 - センサフュージョン
複数の感覚情報に対して統合的・融合的な処理を施すことにより, 単一の感覚のみでは得られない新たな認識機能を付与し, 柔軟な構造を持つ階層的分散処理系を実現する.
 - ビジュアルサーボ
フィードバックループにビジョンを組み込んだサーボ系のアーキテクチャと情報処理手法を研究することにより, 認識と行動の高度な統合を実現する.
 - ビジョンチップ
光検出器と汎用処理回路を画素ごとに直結した集積回路を開発し, 従来のビデオフレームレートをはるかに越える超高速ビジョンシステムを実現する.
- 電子情報 広瀬・峯松研究室(峯松)
 - 音声分析
対話調音声の韻律的特徴(持続時間)の分析, 音程モデルによる中国語基本周波数パターンの定式化, など.
 - 音声合成
波形編集型高品質音声合成の自動化と性能向上など.
 - 音声認識
モーラ遷移確率モデルを用いた句境界検出と連続音声認識への応用, 中国語連続音声認識など.
 - 対話・教育システム
文献検索システム, 発音教育システムなど.
 - 音楽情報処理
歌声とピアノ音の分離など.

3 グループの研究の方向性

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。

20世紀の情報学は、コンピュータやネットワー

クやソフトウェアなどの情報世界に閉じた情報学を中心にしてきた。しかし、21世紀の情報学は、人間が生活する実世界に深く関るようになるだろう。そのためには、情報世界と実世界(物理世界)をつなぐ技術要素が極めて重要になる。

実世界の情報処理においては、実世界から情報を獲得することは不可欠の研究分野である。例えば、自律動作するロボットのような機械を考えると、視覚・聴覚・触覚などの知覚および認識機能、を持つことが必須である。さらにそれが、人間とコミュニケーションをする能力を持つためには、音声による対話の能力が必要である。

これらの研究には人間の視覚・聴覚・触覚の研究も必要であり、その応用として人間側にも有用な応用技術が得られる。

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。

本グループの研究構想として、現在以下のような項目を考えている。これらは、研究の進展とともにさらに拡大し、統合する方針である。

- 音源方向の検出、音源分離

複数マイクロフォンによる音源方向の検出と音源分離、およびヤドリバエ型小型音源方向定位を用いる。標語的には「呼べば振り返るロボット」と表現できる。将来の知能的な実世界情報システムでは、人間が両耳を用いて音源方向を検知でき、それによってコミュニケーションや行動が円滑に行なわれることに注目して、そのような機能を探り入れる必要があると考えられる。

また、音声認識の入力として、雑音環境中から対象とする音声を抽出することが望ましい。この目的のためにも、上記の技術が有用

である。

- 音場の制御

バーチャルリアリティの一つの目標は臨場感であるが、音響的には古くからステレオ音響再生技術により実現されて来た。さらに多チャンネルの音源と信号処理を組み合わせれば、より自由に音場を制御し、複数の聴取者あるいは移動する聴取者に臨場感のある音響出力を提供できる。

- 音声対話

機械が人間の生活に今まで以上に深く関るであろう21世紀の生活では、人間と機械が円滑にコミュニケーションを行なうことが不可欠である。従来の機械のようにスイッチ、従来のコンピュータのようにキーボードとマウスを主体としたヒューマンインタフェースを超えて、人間同士のコミュニケーションに近付いたコミュニケーション方式が必要となる。その中で音声対話能力は最重要な機能である。

- 擬人化エージェント

機械が音声対話機能を持ち、人間的になれば、その姿も人間を模擬することが円滑な人間機械コミュニケーションに向けて有効であろう。それを機械的に実現すれば人間型ロボットであり、画面上に実現すればコンピュータグラフィクスによる擬人化エージェントである。本グループでは、ここに挙げる各機能の研究を擬人化エージェントをプラットフォームとして統合することを想定している。その研究結果は、人間型ロボットに搭載する機能として有用である。

- 高速の視覚機能

従来のロボットビジョンなどの視覚情報処理の研究では、通常のビデオカメラを視覚入力として用いる場合が多く、そのフレームレートの制約から、遅い動きしか捉えられなかった。人間の生活空間を含む実世界では、これより高速の視覚情報処理が要求される。そのためには、従来の入力デバイスの制約を超え

る入力手段と処理手段が必要であり、それにより初めて、高速の運動機能を持つロボットが十分に動作できると考えられる。

- 触覚センシング機能

人間の生活では触覚が重要な位置を占める。人間の生活空間に共存・同居する機械(ロボットなど)にも人間と同等の触覚が必要となろう。同時に、人間に対して触覚を通して情報を伝えることも必要となる。特にバーチャルリアリティにおいてバーチャルな触覚および力覚の呈示は、待望される場所である。

4 本年度のグループの活動

以上に述べたような観点から、本年度の本グループの研究の方向として、本COEの本プロジェクトの活動として以下のような研究テーマを取り上げた。

- 音声対話擬人化エージェント
- 音響センサ(ヤドリバエ模倣型小型音源定位センサ、空間2階微分を用いた複数音源同時定位センサシステム)
- カクテルパーティ効果の工学的利用
- 視覚情報処理(ビジョンチップ、高速ビジョンシステム)
- 人間の聴覚(聴覚インタフェース、人間の言語学習支援、伝統歌唱分析)

これらについて、各研究ユニット(研究活動単位)からの報告を本報告書に掲載する。

5 本年度の予算使用状況

本21世紀COEの研究予算については、本COEの予算のみでなく、個別に外部資金を積極的に獲得することによって賄う方針である。そのため、各研究ユニットの持つ外部資金の多寡により、相補的な方針でCOE予算を使用した。それらの主な用途は以下の3点である。

- 音声対話エージェントCGソフト製作、装置

- 音源定位視線移動装置
- マイクロフォンアレイ信号処理装置

すでにこれらの設備を利用した研究が進んでおり、その研究の構想、研究状況などについては個別の研究報告を参照頂きたい。

6 結び

本年度に開始された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「視聴覚グループ」の状況を報告した。グループを編成してやっと3ヵ月が経過したところであり、その期間は研究構想と設備の導入と使用開始が主要な活動であったが、いままでの研究の蓄積の上に進展を見ることができた。今後は、導入設備を生かし、本文中に述べた大きな統合目標に向かって一層の研究を進める。