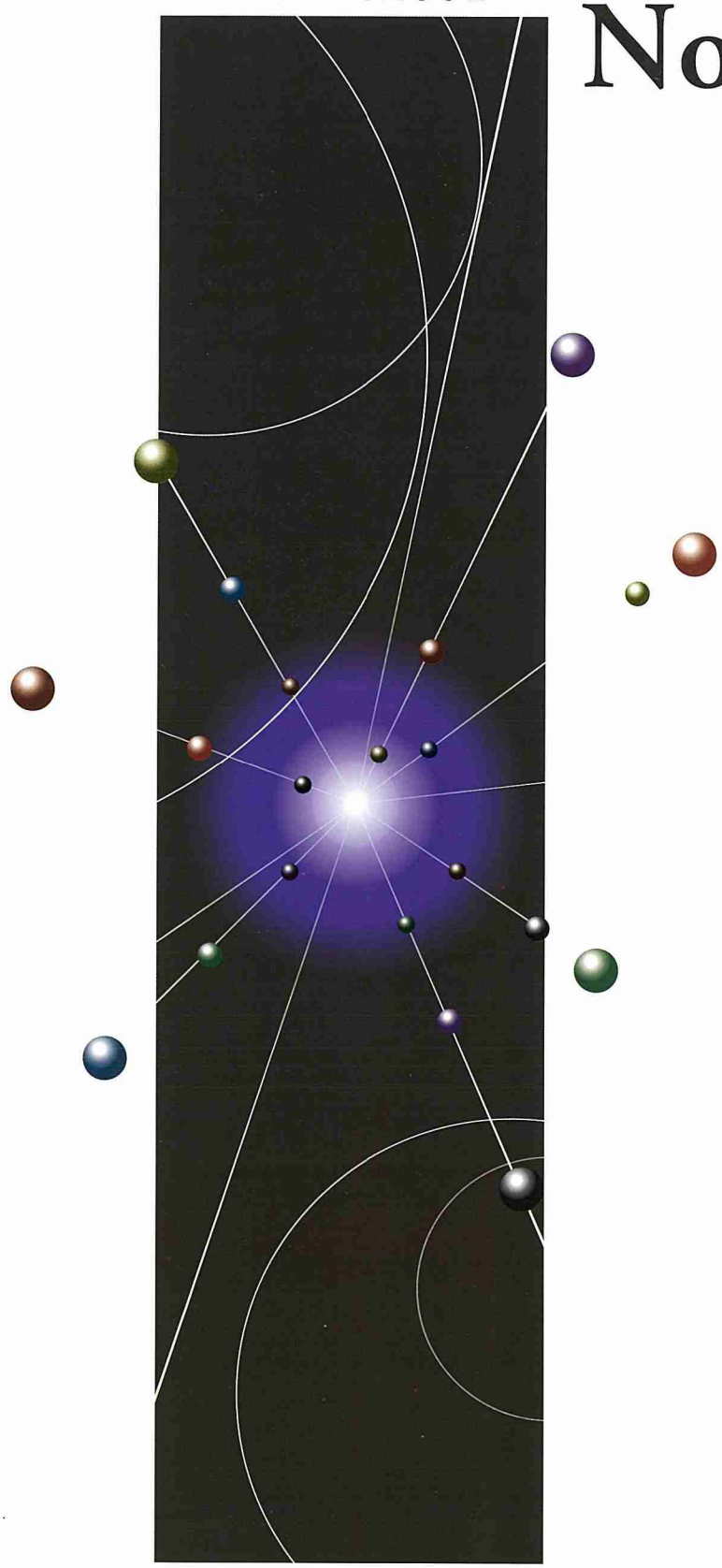
ISSN 0918-6573

ATR Journal

SUMMER 2001

No.44



ATR

株式会社 国際電気通信基礎技術研究所

●研究動向紹介 Current Research Topics at ATR	「顔を発見！ 以後、動きを追跡します」 Face Detection and Tracking － A first step to face gesture recognition －	2	川戸慎二郎 鉄谷　信二
	日中間の円滑な音声コミュニケーションを目指して －ATR音声言語研究所における中国語音声認識の試み－ Towards Chinese-Japanese Speech Communication - ATR Approach to Recognizing Chinese Mandarin Speech -	4	Shuwu Zhang Jinsong Zhang
	適応アンテナを用いたアドホックネットワーク Wireless Ad-Hoc Community Network utilizing adaptive antenna	6	植田　哲郎 S. Bandyopadhyay
	発話運動を“見る” －3次元MRI動画撮影（3D MRI movie）による発話運動の観測－ New Motion Imaging Technique for Observation of Articulatory Movements using a 3D MRI movie	8	竹本　浩典
●研究所だより	エイ・ティ・アール知能映像通信研究所	10	
	エイ・ティ・アール音声言語通信研究所	11	
	エイ・ティ・アール環境適応通信研究所	12	
	国際電気通信基礎技術研究所 先端情報科学研究部	13	
● Letters from Alumni	最近とりとめなく考えること	14	吉田　匡雄
	竜宮城を現実のものに	15	田中　昭二
●トピックス What's new with ATR	山口きらら博NTT西日本館展示にATRの成果が採用される	16	
	ATR科学技術セミナーの開催状況	16	
	最近のATR成果の展示状況	17	
●成果展開 Fruits of ATR's Research	受賞 等	17	
● SHIEN日記	パンダちゃんの入学手続き	18	
●所員往来 Changes in The Staff	平成13年4月～6月	19	
●イベントカレンダー		20	
●編集後記 Editor's Notes		20	

「顔を発見！ 以後、動きを追跡します」

Face Detection and Tracking

― A first step to face gesture recognition ―

私たちは人の話を聞くとき、話し手の顔を見ます。顔は言葉にならない情報をたくさん発信しているからです。この顔表情から発信される言葉にならない情報をコンピュータでキャッチするために、動いている顔画像を実時間で分析する必要があります。その第一歩として、実時間で動画像の中の顔を検出し、その中の特徴点を追跡する研究を進めています。

When we listen to talks, we look at his/her face. This is because the face expresses much information other than words. To catch such information by a computer, a live video sequence must be analyzed in real time. Toward such research, we are now studying detection and tracking of faces in live video.

（株）ATR知能映像通信研究所
第一研究室
川戸　慎二郎、鉄谷　信二

1. はじめに

人は言葉以外にも、表情やジェスチャーでさまざまな情報を発信しており、コンピュータがそれを認識できるようになれば、さまざまな分野でより自然なコンピュータとのインターフェースが可能になると思われます。なかでも顔から発せられる情報量は多く、「目は口ほどにものを言い」と言われるように人間どうしのコミュニケーションでも重要な役割を果たしています。国会で証人喚問の動画中継が復活しましたが、口では嘘をつけても、表情ではごまかせないことも、理由の一つでしょう。手話でコミュニケーションする人たちの視線は、対話者の顔を常に注視しており、顔から視線が外れることはほとんどないと言われています[1]。手話でも手のジェスチャーによる情報と同程度かそれ以上に顔表情からの情報が重要なのでしょう。解剖学的にも、顔には多くの表情筋が複雑に張られていて、表現できる情報量の大きいことが知られています。

情報処理の分野で顔に関する研究は十年ほど前から盛んになり、表情認識についても、多くの成果が報告されています。しかし、ほぼ固定した顔が対象であり、表情も喜び、怒り、嫌悪、など基本6表情と呼ばれる表情の識別に関するものでした。

しかし私たちは、喜びや怒りなど自然に表れる表情ではなく、人がジェスチャーで意図する情報を抽出する必要があると考えています。そのためには、そういった情報が表現される顔の特徴点を自動的に抽出し、時間経過に沿って追跡する必要があると考えました[1]。また、顔の位置についても束縛を感じさせない動きを許容できることが必要です。

2. あなたな～ら、どうする～♪

一般に動画像中の顔領域を抽出することは、背景から肌色領域を分離抽出することでできます。しかし顔領域の中から一定の特徴点をリアルタイムで抽出することはまだ確立されているとは言えません。これまでの多くの研究は、顔部品と言われる目とか、鼻や口を直接抽出しようとしています。しかし、ジェスチャーの認識を目的としているとき、目や口は大きな形状変形を予想しておかなくてはなりません。ウィンクで片目をつむることもあります。人間にとって容易でも、目や口をその変形に惑わされず、リアルタイムに検出することは現在のコンピュータにとっては難しそうです。鼻孔はパターンとしては非常に安定していると言えますが、観測できるカメラアングルがかなり限定されます。

もうずいぶん前のことですが、「あなたな～ら、どうする～♪」という歌がヒットしていたことがありました。あなたなら、どうしますか？

私たちは、顔部品とは言えないけれど、表情が変化してもパターンとしてあまり変化しない眉間に着目しました。眉間は、顔の向きが変化しても観測できる範囲がかなり広いという特長があります。いったん眉間を見つければ、それを手掛りに目や口の形状を抽出することは、少しは容易になりそうです。

そこで、まず眉間をリアルタイムに検出し、追跡してみることにしました。

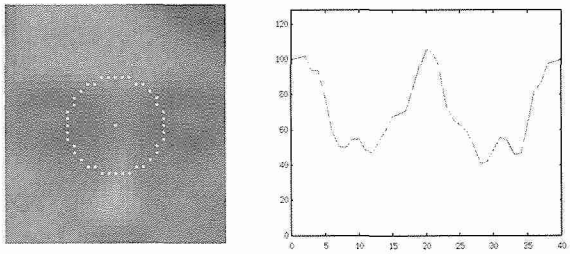


図1　眉間周囲の濃淡パターン

[1] 顔の全体的特徴から情報を抽出しようという考えもあります。

3. 眉間抽出フィルター

眉間は上方が額、下方が鼻筋で比較的明るく、両サイドは目や眉で比較的暗いという特徴があり、これは多くの人に共通していると考えられます。このような特徴点を軽い計算負荷で抽出するために、リング周波数フィルターと称するフィルターを考案しました。これは、注目画素を中心とする数画素の半径の円を描いて、その円周上の画素値を標本値とするフーリエ変換を計算し、周波数2のスペクトルパワーを出力するものです。図1にそのような円周上の画素値のプロット例を示します。2周期のコサイン波形に似ています。注目点が眉間から離れるにしたがって、このような波形パターンがくずれていくため、眉間ではこのフィルターの出力が極大値となります。リング周波数フィルターの特徴は円周上のデータを使って計算されるので顔が傾いても影響を受けないことです。一方同じフーリエ変換によって、周波数2の位相角が計算されますが、これは顔の傾き角に相当します。顔領域にリング周波数フィルターを適用すると、眉間以外にもパワー出力が極大値となる点が出てきますが、その中から本当の眉間を選択するのに、この位相角は有効な情報となります。

4. どうやって選ぶ、本当の眉間？

フィルター出力が極大値をとる点の中から本当の眉間を選ぶのに、平均的な眉間パターンと比較して判断します。その目的で、データベース[2]の398の顔画像から図2のような平均パターンを作成しました。

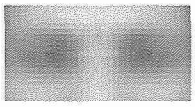


図2　眉間の平均パターン

しかし、いつでも目が水平に並んでいる顔画像が相手とは限りません。そこで、リング周波数フィルターから得られる位相角分だけ、調べたい点の回りに入力画像を逆回転させてから、比較するパターンを切り出すようにします。このようにして、パターン比較方式が持つ回転ズレに対する弱点を克服しました。

この方式をデータベースの400の顔画像でテストしたところ、眉間の検出率98.5%が得られました。

5. 以後、追跡します

いったん、眉間を検出できたら、以後は画像内での動きを追跡します。新しい画像が入力されるたびに検出してもよいのですが、前の画像で検出した結果を利用した方が処理が簡単で済みます。まず、新しい画像内で眉間のありそうな位置を予測します。前の画像で初めて検出したのであれば、同じ位置にあると予測しますが、もう一つ前の画像でも検出していれば、動きの速度を推定して、現在位置を予測します。そして予測した点の周辺のみリング周波数フィルターを適用し、出力が極大値となる点を検出します。

しかしこのままでは、一度眉間以外の点を取り違えて検出すると、その間違った点を追跡し続けることになりかねません。そこで眉間を検出、追跡するたびに、それを中心とする小さな矩形パターンを記憶し直します。そして、新しい画像で眉間と思われる点が検出されると、前の画像で記憶しておいたパターンと比較し、大きく違っていなければ正しく眉間が見つかったものとみなします。大きく違っていた場合は履歴データをクリアし最初からやり直します。

このような追跡処理により、最初の眉間検出処理に比べて、肌色領域の抽出やパターンの回転処理が不要となり、より高速な処理が実現できます。

6. パソコンでデモ

プログラムをPentium III 866MHzのパソコン上に組んで実験してみました。検出処理で27フレーム/秒が可能で、追跡処理では30フレーム／秒の処理をして少し余裕があります。図3の実験モニター画面の例では、左端上の5つのパターンが眉間候補画像で、その下に選ばれたパターンが示してあります。顔の傾きが修正されているのがわかります。入力画像上には眉間候補点の位置が小さい点で示されています。

図3　実験モニター画面

7. 今後の予定

できたのは眉間を見つけて追跡するまでで、ジェスチャー認識にはまだ遠い道のりがあります。今後、目と口の形状をリアルタイムで検出、追跡する方向で研究を進める予定です。

参考文献

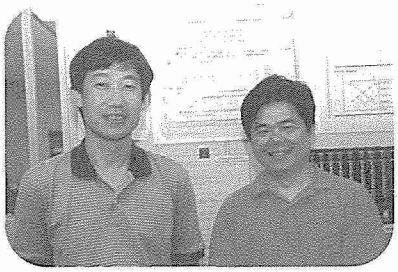
[1] 長嶋祐二: 手話情報学の現状と課題; 信学技報 PRMU99-141 (99.11)

[2] AT&T Laboratories Cambridge: The ORL Database of Faces; http://www.cam-orl.co.uk/facedatabase.html

日中間の円滑な音声コミュニケーションを目指して
ーATR音声言語研究所における中国語音声認識の試みー
Towards Chinese-Japanese Speech Communication
- ATR Approach to Recognizing Chinese Mandarin Speech -

ATR音声言語通信研究所では現在、中国語と日本語との間の音声コミュニケーションの研究に取り組んでいます。その第一段階として「ホテル予約に関する会話」に対して優れた性能を持つ中国語音声認識のプロトタイプを作成しました。本稿では中国語の音声認識システムを開発する際の課題とそれらに対する我々のアプローチを紹介します。

ATR Spoken Language Translation Research Laboratories is currently working on Chinese and Japanese speech communication project. After a period of work, we have developed a prototype Chinese recognition system with good performance for ATR task. This paper gives a brief introduction of the problems in developing Chinese speech recognition system and the approaches we have implemented.



（株）ATR音声言語通信研究所
第二研究室
Shuwu Zhang、Jinsong Zhang

1．はじめに

中国語、特に普通話と呼ばれる言語は中国大陸内外で公用語として使用されており、世界最大の話者人口を持っています。中国語を話す地域は近年世界的にも有数の経済的な発展を遂げており、日本語と中国語の間のコミュニケーションもますます頻繁かつ重要になっています。こうした背景のもと、ATR音声言語通信研究所では2000年3月に中国語の音声翻訳をATRの多言語音声翻訳システムに組み込む新しいプロジェクトを立ち上げました。本稿では、このうち中国語音声認識についてご紹介します。

従来の音声認識技術は日本語や西欧言語を中心に開発されてきました。音声認識では対象言語の音響的、言語的特徴を計算機内でうまくモデル化しなければなりませんが、中国語は日本語や英語とは音声的、言語的に大きく異なるため、既存の技術がそのまま適用できるとは限りません。我々は中国語の特性に対する検討を行い、ATRの音声認識エンジンSPRECをもとに1万語規模の中国語音声認識を行うことに成功しました。現在のシステムは丁寧に（「朗読調」で）発話されたホテルの予約に関する会話を高い性能で認識できます。

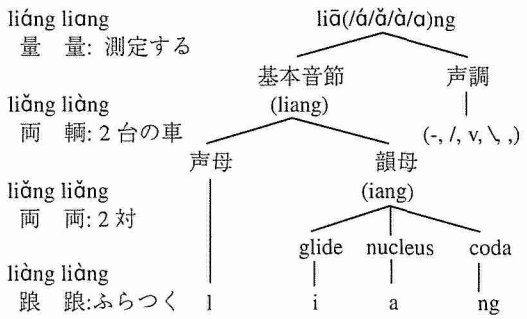


図1　中国語の単語の例と音節の構造

2．中国語音声の性質

中国語の文字は表意文字である漢字です。各漢字の発音（読み）はたった一つで、その特徴は必ず一つの音節（単音節）として発声されることと、固有の音の高さのパターン（声調）を持つことです。ここで、音節とは母音を中心とする音のかたまりのことで、大雑把には仮名1文字の音に対応します（ただし「ゃ」や「ん」などは特別扱い）。このような漢字が1個から数個つながって単語になります。図1の左側に2文字の単語を4つ示します。それぞれの漢字の上にあるのはピンインと呼ばれる一種の発音記号で、アクセント記号のようなものは声調を表します。例えば、「両」は低い声でliang（リャン）、「輛」は高い声から始めて下がり調子でliang（リャン）と発音します。お気づきの通り、図中の4つの単語は声調を無視するとどれも同じ音（“liang liang”）になります。図1の右側は中国語漢字の発音の構成要素です。1つの漢字の発音は「声調」と声調を除いた「基本音節」に分けられ、基本音節はさらに声母（shengmu）と呼ばれる最初の子音の部分と韻母（yunmu）と呼ばれる残りの部分に分けられます。普通話では22の声母と38の韻母があります。なお、韻母は最大3つの部分からなりそれぞれは従来の音声認識の基本単位である「音素」によく似た単位になります。中国語音声認識の難しさは以上で述べた中国語の性質に起因します。まず、現在の技術では声調の認識が困難なため図1に示したような単語は音響的に区別できません。また、中国語では音節ごとに比較的独立して発声され、隣接する音節の間に単語境界があってもなくても音響的にあまり変化しません（他の言語の場合この限りではありません）。このことは単語の境界を検出する信頼度を下げるもとになります。

このほかの重要な問題として中国語では日本語と同様に単語の区切りが明確でないことが挙げられます。中国語テキストは漢字のべた書きであり、英語のように単語の間にスペース等を入れません。従って、どういう単語の並びが中国語としてもっもらしいかという言語モデルを大量のテキストから自動学習する際に単語の区切り方に関する曖昧性が生じます。この曖昧性と上で述べた音響的に生じる曖昧性のいわば掛け算によって中国語の音声認識は非常に困難なものになります。図2に示すように各音節に対する候補が比較的少ない場合でも非常に複雑なグラフになります。

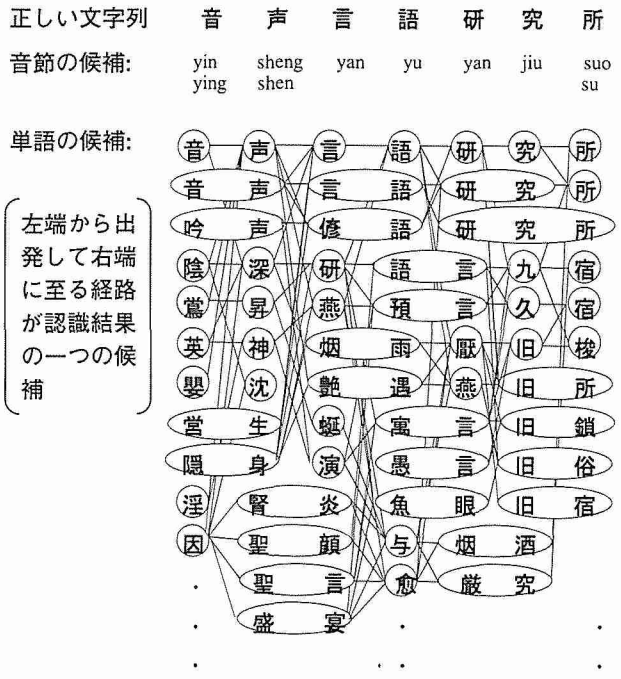


図2 “音声言語研究所”に対する認識候補の例

3. ATR中国語音声認識プロトタイプシステム

ATRの中国語音声認識の研究は中国語の音声データベースの構築から始まりました。我々は中国科学院（China Academia Sinica）の協力を得て、ホテル予約に関する会話を100人の話者によって延べ10時間収録したCAS00という音声データベースを作成しました。さらに、基本となる音響モデルを作成するためにいくつかの音声データを入手しました。

我々はATRで開発された音声認識エンジンSPRECをもとに中国語用のプロトタイプシステムを構築しました。SPRECは日本語用に開発されたものですが、英語にもうまく移植されています。しかし、元来声調言語用ではないため、本プロトタイプでは声調の認識を直接行うことはできません。

日本語や英語の音声認識においては「音素」が認識の基本単位として利用されますが、我々は音素の代わりに2で述べた22の声母と38の韻母を基本単位として用いました。この方が単音節言語の音声を適切にモデル化できると考えたからです。実験の結果、我々の方法は音素を使う方法と比べて認識精度が高く、また、既に提案されている他の方法と比べて計算量が少ないことがわかっています。

音声認識では図2で示したような多くの同音異義語の中から適切なものを選ばなければなりません。本プロトタイプでは頻出する熟語をシステムの辞書に追加することでこの問題の一部を回避しています。また、中国語の言語モデルも有効です。同音異義語のうちどれが正しいかはその前後関係から判断できることが多いからです。

プロトタイプシステムの音響モデルは、外部から入手したデータベースを使って構築した基本的なモデルをもとにCAS00を使って調整（適応化）しました。この新しい音響モデルはATRのホテル予約会話、および、比較的広い範囲の新聞記事読み上げ音声の認識において優れた性能を示しました。

4. 今後の展望

今回開発した中国語音声認識プロトタイプにはまだまだ改善の余地があります。1つ目は声調の処理を取り入れることです。これによって語彙やタスクの拡大が容易になります。2つ目は話者の種類、特に方言のアクセントに対応することです。中国は広大であり、地方ごとに発音が大きく異なります。北部の人々の音声データで構築された認識システムで南部の人の声を認識すると精度が大きく下がります。3つ目の課題は構文の変形や多義語の識別、固有名詞の同定など中国語の言語的な処理を高度化することです。最後の課題はより自然な発話を扱えるようにすることです。このためには単にそのような音声のデータを収集するだけでなく、間投詞（「えーっと」のような）などの自然な発話固有の現象を処理する必要があります。

5. おわりに

本記事では中国語音声認識における中国語特有の問題とそれらに対するATRのアプローチを紹介しました。将来、ATRの音声翻訳技術によって日中間の円滑な音声コミュニケーションができるように、我々はシステムをより現実的で頑健にするための課題に取り組んでいます。

適応アンテナを用いたアドホックネットワーク

Wireless Ad-Hoc Community Network utilizing adaptive antenna

第一研究室では、環境に適応したネットワーク構成法として、基地局などのインフラに依存せず端末のみで構成されるアドホックネットワークについて研究を行っています。ここでは、WACNetというコンセプトの下に、アンテナやプロトコルを端末自身が自律分散的に制御することにより通信環境の動的変化に対応する適応的な通信システムの研究状況について紹介します。

The primary focus of our research is to implement the concept of a Wireless Ad-Hoc Community Network (WACNet) with the aim of supporting data communications among a group of people within an area, where no installed infrastructure is available. We are conducting research on new architectures and protocols based on adaptive antenna, in order to realize a highly adaptive communication system that may autonomously behave to a dynamic environment.

（株）ATR環境適応通信研究所
第一研究室
植田　哲郎、S. Bandyopadhyay

1. はじめに

近年の移動無線通信システムの発展によりいつでもどこでも通信が可能になってきました。しかし、これらの無線通信は、携帯電話の基地局や無線LANのアクセスポイントなど通信を制御するインフラを必要とし、カバーエリアにも制限があります。今後更に無線通信を便利なものにするには、インフラが整備されていない場合でも複数ユーザ間の通信をサポートすることが必要で、そのニーズに答える一つの形態が、アドホックネットワークと呼ばれるものです。

アドホックネットワークは、一時的に集まった端末同士で構成され、基地局などのインフラによる集中制御が不要なシステムです[1]。端末自体がルーティング機能を持ち、端末間で直接通信（シングルホップ）するだけでなく、他の端末を経由して通信（マルチホップ）を行うことにより任意の端末分布状況に応じた自律分散ネットワークが構成されます（図1）。

私たちの研究室では、ある特定の目的の通信を行うために集まったユーザ同士が、伝搬環境やユーザの移動・参加・離脱などの様々な動的環境変化に自律的に対応できる適応的な通信システムの実現を目指して研究しています。

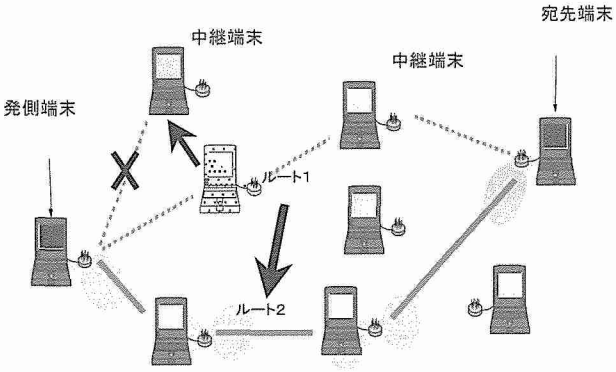


図1　アドホックネットワークにおけるルーティング

2. WACNet

私たちは、アドホックネットワークの考え方を一歩進めて、何らかの共通のベースをもって一時的に集まった不特定多数の人達で形成されるコミュニティの中でデータ通信をサポートするネットワークをワイヤレスアドホックコミュニティネットワーク（WACNet:Wireless Ad-Hoc Community Network）と名付け、環境に適応した通信ネットワークのプラットフォームとして提案してきました[2]。このようなネットワークは、今後のユーザ用携帯データ端末の進展に伴って、例えば、国際会議、大学キャンパス、競技場やシアターなどの娯楽施設など、同種の目的を持って集まった人達の間での各種コミュニケーションのトリガーになるものと考えられ、以下の特徴を持っています。

(a) スケーラビリティ（端末数の増減に対してもネットワークが対応可能）が高い
(b) 資源（周波数、電力）の利用効率が高い
(c) ネットワークへの自由な参加と離脱

これらを、空間・電力効率の良い端末用適応アンテナ、またアンテナの特性を生かしたアクセス制御やルーティングを搭載することにより達成することを目指しています。

3. 適応アンテナ（ESPARアンテナ）

適応アンテナは、ビームとヌルの方向を適応的に制御し空間的な周波数利用効率を高める技術として現在注目を集めていますが、アンテナ素子毎にRF回路を設けビーム制御を行うため回路規模や消費電力などに問題があります。

ATRで検討しているESPARアンテナ（Electronically Steerable Passive Array Radiator）[3]は指向性制御が可能でRF系回路は一つで済むため低コスト低消費電力であり小型化も容易で、アドホックネットワークを構成する各端末が伝搬環境に適応的に対処することができます。

4. アクセス制御とルーティング

適応アンテナをアドホックネットワークで活用するためには、上記の端末への適用の他に、アクセス制御とルーティングが課題です。固定基地局などのインフラ側が通信先へのルート設定を行うセルラー網と異なり、アドホックネットワークはルーティング機能を持つ端末自体が適応アンテナを使いながらルートを決定し更に中継端末などの移動にも対処する必要があります（図1）。

そこで、可変ビームアンテナをユーザ端末に搭載した場合に、その特性を生かしたルーティング方式としてATR（Angle-SINR Table Routing）方式を提案しました[4]。まず各端末が周囲の端末分布状況を知るため360度ビーム走査し、どの方向に端末が存在するかをSINR（Signal to Interference and Noise Ratio）により把握すると同時に（Angle-SINR Table作成）、宛先端末・隣接端末・ホップ数で構成されるルーティングテーブル（RT）を更新します。データ送信・中継時には、まず自分のRTから宛先端末名をキーとして隣接端末名を検索し、Angle-SINR Tableから隣接端末名をキーとして最大SINRとなる方位角を検索し、その方位角にビームを形成して送信します。Angle-SINR Tableに基づいて常にホップ先のみにビームを向けてルーティングを行うため他の端末への干渉量が減り、システム全体のスループットが向上します。また、テーブルは定期的に更新し、端末の移動・参加・離脱に対応します。

このAngle-SINR Tableを活用し、現在通信中の端末を妨害することがないよう新たな通信時の指向性を制御する適応的なアクセス制御方式[5]や端末間の通信品質を考慮したルーティングプロトコルについても検討を進めています。

5. アプリケーション

アドホックネットワークにおけるアプリケーションとして、物理的な端末の振る舞いをサービスと連携させるために接続履歴からサービス履歴を構築することが考えられ、CoCoNUT（Contiguous Communication Network on Ubiquitous Transmission）フレームワークとして提唱しています（図2）[6]。

CoCoNUTは、ノード同士の近接がトリガーとなり過去の接続履歴に応じて通信を自律的に行う接触型通信で、スポット的に散在するアドホックネットワーク間の通信にも活用できるものです。アドホックネットワークは一時的に構成されるネットワークですが、会議場や展示会場などでの再現性は高く、それらの場所に応じたノードを設置すれば、例えば会議場への参加履歴（接続履歴）に応じて参加者が会議場へ近接したときに自発的に位置に応じた通信を行うことができます。更に、

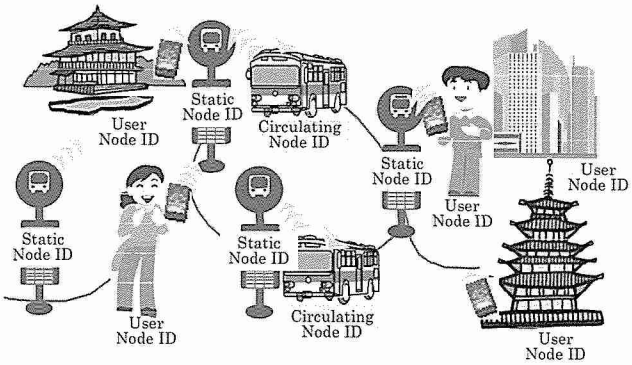


図2　アプリケーション例（CoCoNUT）

あるノードが会議場への訪問履歴のある別ノードに情報伝達を委託する、あるいは、既定ルートを巡回するバスなどが所定ノード（バス停）との接近により情報流通・委託を行う、などによりアドホックネットワーク間の通信を実現します。

6. むすび

適応アンテナを活用したアドホックネットワークと、アプリケーション例について述べました。現在、要素技術の実証に向け装置を開発中であり、適応アンテナを活用したアクセス制御とルーティング方式について更に検討を進める予定です。

参考文献

[1] B. Johnson, D. A. Maltz: Dynamic Source Routing in Ad Hoc Wireless Networks; Mobile Computing, Kluwer, (96)

[2] M. Kawai, M.Nozaki, K.Gyoda: A wireless Ad Hoc Community Network with Reconfigurable Topology Architecture; Proc. Of the GLOBECOMM '98 (98)

[3] T. Ohira, K.Gyoda: Electronically Steerable Passive Array Radiator (ESPAR) Antennas for Low-cost Adaptive Beam forming; IEEE International Conference on Phased Array Systems (00.5)

[4] 大野, 門, 行田, 大平: アンテナ走査角対SINR情報に基づくルーティング方式の性能評価; 信学ソ大, SB-3-1 (00.9)

[5] S. Bandyopadhyay, K. Gyoda, K. Hasuike, S. Horisawa, Y. Kado, S. Tawara: An Adaptive MAC Protocol for Ad Hoc Networks with Directional Antenna; The 7th Asia-Pacific Conference on Communications (01.9予定)

[6] 門, 小菅: 巡回ノードを活用した無線アドホックネットワークの一構成; 信学技報, RCS2000-44 (00.6)

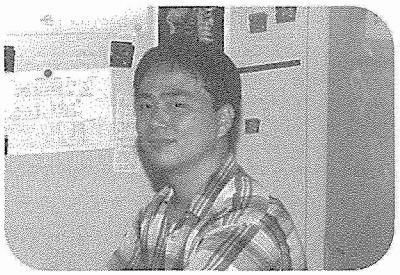
発話運動を“見る”

ー3次元MRI動画撮影（3D MRI movie）による発話運動の観測ー

New Motion Imaging Technique for Observation of Articulatory Movements using a 3D MRI movie

3D MRI movieという新しいMRI撮像法を用いることにより、発話運動を3次元的に観測することに成功しました。この観測法により、声道の立体形状の時間変化を観測し、声道断面積関数の時間パタンを抽出することができます。これらの知見によって、人間の複雑な発話のメカニズムを解明することが期待されます。

A new MRI motion imaging technique called 3D MRI movie has been successfully adopted to 3D recording of articulatory movements. This method enables us to observe the temporal pattern of vocal tract movement and extract the area function. The results will contribute to the understanding of complex mechanisms of human speech production.



（株）国際電気通信基礎技術研究所
先端情報科学研究部
音声生物科学プロジェクト
竹本　浩典

1．はじめに

音声生物科学プロジェクトでは、音声の生成と認識の生物学的な機構を解明し、人間の機能を模擬する新しい音声通信技術の開発を目指しています。研究は、音声の形態的基盤、機能モデル、脳機能の3つのテーマで進めています。ここでは、形態的基盤についての研究動向を紹介します。

音源フィルタ理論[1]によれば、音声は、喉頭で生成された音源波が声道という音響フィルタを通過し、音響的な修飾を受けて口唇から放射される現象として説明されます。音源と音響フィルタは相対的に独立しているので、私たちは高いピッチや低いピッチ、ささやき声など、音源の性質を変えることによりさまざまな声で話すことができます。この理論によれば、音声波と、音響フィルタまたは音源のいずれかが観測できれば、音響現象のすべてを解明することができるので、音声をこの2つに分けて分析し、それぞれの特性を調べることは、音声の合成にも認識にも役に立つと考えられます。

しかし、声道音響フィルタについての分析はこれまできわめて不十分でした。声道音響フィルタの特性は、声道断面積関数と呼ばれる、音声波の進行方向に垂直な声道の断面積と声帯からの距離の関数によって決定されます。X線映画では、実時間で発話運動を観測できますが、2次元の透過画像しか得られないため、断面積を実測することができず、断面積は推定に頼っていました。一方、MRI（磁気共鳴画像法）では、静止時であれば3次元のデータを得ることが可能で声道断面積関数を実測することができますが、時間的な変化を実測することはできませんでした。

私たちは、脳機能イメージングセンターと協力して、3次元MRI動画撮影（3D MRI movie）という手法を用いて、発話運動を空間的・時間的に記録し、声道断面積関数の時間変化を抽出することに初めて成功しました。以下では、この手法と研究成果について紹介します。

2．3D MRI movieによる発話運動の観測とは？

3D MRI movieは、同期サンプリング方式のマルチスライスMRI movie[2]を応用して、被験者の体内の様子を3次元的に記録するだけでなく、その時間変化も記録する方法です。

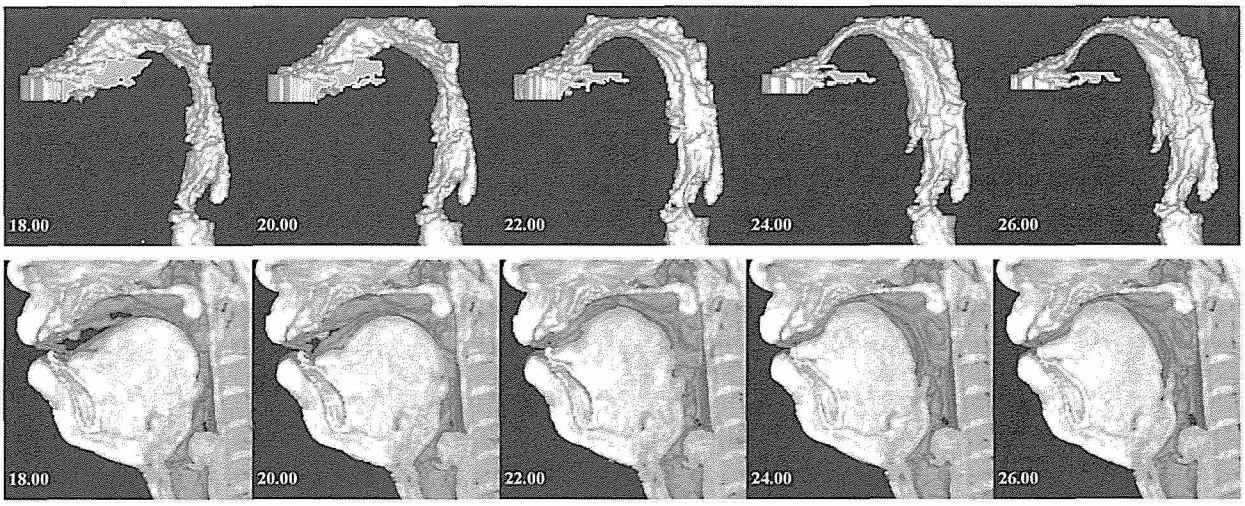
3D MRI movieは、繰り返し規則正しく行われる運動しか記録することができません。それは、すべてのデータを一回の運動中に採取することができないので、何回にも分けて採取する必要があるからです。つまり、毎回の運動がほぼ同じであると仮定されるときのみ、この方法を用いることができます。

3D MRI movieによる発話運動の観測とは、被験者に発話タスクを正確に何度も繰り返し行わせることによって、発話運動を時系列に沿ったボリュームデータとして記録し、観測することです。残念ながら現在では、わずか2秒弱のデータを得るために、被験者は約1000回も同じ発話を繰り返す必要があり、被験者の負担が大きいことが難点です。

3．どのような観測が可能か？

3D MRI movieによって、発話運動を空間的・時間的に観測することが可能です。3D MRI movieにより得られたデータは、“発話中のバーチャル・ヒューマン”とでも言うべきもので、話者の発話運動は、あらゆる角度、あらゆる断面、あらゆる時点で観測することが可能です。

その例を図1に示しています。図1は、日本語5母音/aiueo/を連続発話しているときの、/a/から/i/へ移行する様子を66msごとに表示したものです。本



/a/　　+66ms　　+132ms　　+198ms　　/i/

図1　/a/から/i/への声道形状の変化

来は動画なので、もっとたくさんのフレームがありますが、ここでは静止画としてごく一部だけを表示しています。下段は、正中矢状断面より右側だけを可視化し、矢状面から口腔部分を拡大して表示したものです。上段は同じデータから声道だけを抽出して左前方60度から表示したものです。このように、従来不可能であった、/a/から/i/へ口の構えや声道の形が移行してゆく様子を立体的に観測することができます。

しかも、ただ単に発話運動が“見える”だけではありません。フレームごとに3次元計測を行い、これを時間的な変化として観察することも可能です。図2は声道断面積関数の時間パタンを示しています。つまり、図2は声道断面積関数が/aiueo/の連続発話の間にどのように変化してゆくかを示したグラフです。このように、これまで、推定関数を用いて予測することしかできなかった声道の立体形状の時間変化を実測することもできます。

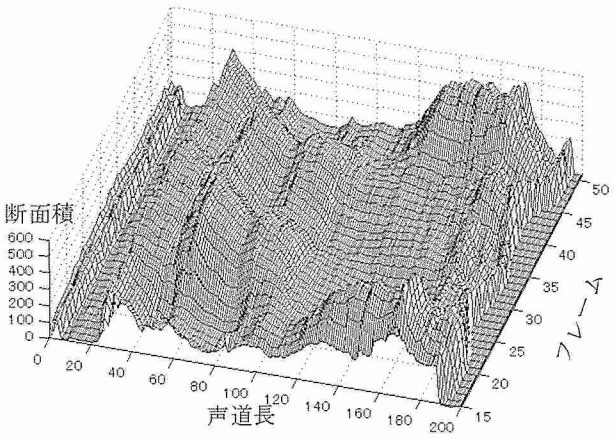


図2　声道断面積関数の時間パタン

4. おわりに

ここで紹介したように、3D MRI movieでは発話運動を空間的・時間的に観測することができる新しい手法です。この手法によって、どのようなことが明らかになるのでしょうか？

まず、発話器官全体の協調した運動を観測することができます。単に発話運動を視覚的に理解する上で役に立つだけではなく、発話器官各部の運動の時間的な関係を計測することができます。

また、音声の個人性の要因として重要である梨状窩や歯冠間隙などの声道分岐管の形や大きさが、調音によって変化するかどうかを実測することができます。

さらに、連続した音素間で声道形状がどのように変化するかを実測することができます。これにより、量子説と呼ばれる母音から母音への不連続性を検証することも期待できます。

なお、この観測法の最大の問題点である、繰り返し発話による被験者の負担は軽減される見通しがあります。それは、複数のコイルを並列して使う撮像法を用いることで、これにより発話回数を大幅に減らすことが期待されています。

参考文献

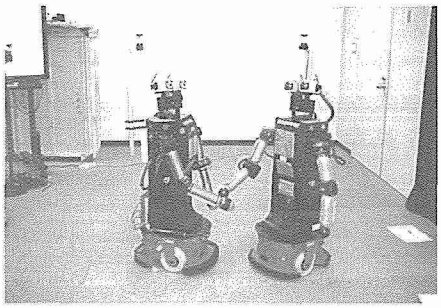
[1] レイ・D・ケント，チャールズ・リード: 音声の音響分析; 海文堂 (96)

[2] モリエル・ネスエイバー: 図解 原理からわかるMRI; 医学書院 (98)

日常活動型ロボット「ロボビー」の展示動向

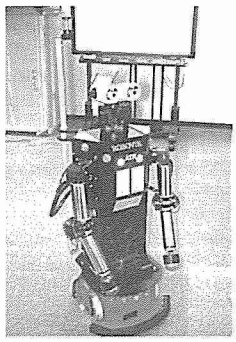
昨年開催されたロボフェスタ関西、神奈川のプレ大会に始まりＲＯＢＯＤＥＸ２０００、ザロボット博とＡＴＲでは、知能映像通信研究所で開発した「ロボビー」を出展し、好評を博してきました。いずれの展示会でも、近年注目を浴びつつあるパートナー型ロボットが一堂に会し、将来のロボットとの共存を予感させるものでした。

ロボットはこれからも注目度が高く、ＡＴＲでは、日常活動型ロボット「ロボビー」を中心としたコミュニケーションロボットの研究を展開していきたいと考えております。また、研究成果を広く社会一般に発表するため、今後も展示活動を行なっていく予定です。また、一般からのコメントを反映させる形で「日常活動型ロボット」という我々のコンセプトを実現していきたいと考えています。



今後の展示スケジュール

ロボフェスタ生駒 奈良先端科学技術大学院大学・高山サイエンスプラザ	２００１年７月３１日
２１世紀 みらい体験博（神戸） 神戸国際展示場	２００１年８月１５日～２２日
ロボフェスタ神奈川（神奈川） 横須賀市南体育会館	２００１年８月２５日～９月２日
うつくしま未来博（福島） 福島県須賀川市	２００１年９月１日～２日
ロボフェスタ神奈川（神奈川） 川崎市とどろきアリーナ	２００１年９月８日～１６日
ロボフェスタ神奈川（神奈川） 相模原市銀河アリーナ	２００１年１０月６日～１４日
ロボフェスタ神奈川（神奈川） パシフィコ横浜	２００１年１１月１６日～２５日



ハンズフリー音声通信に関する国際ワークショップ（HSC2001）を開催

ATR音声言語通信研究所では、4月9日から4月11日まで、International Speech Communication Association （ISCA）、IEEE Signal Processing Society、日本音響学会、電子情報通信学会の協賛を得て、ハンズフリー音声通信技術に関するワークショップを主催しました。

本ワークショップでは、日々増大するハンズフリー音声通信への期待と需要に応えるべく、ハンズフリー音声通信に関連する科学あるいは工学のあらゆる分野を対象とし、さまざまな視点から、現在解決すべき課題のみならず、グローバルなコミニュケーションを目指した新たな課題を提案することを目的としました。

山本社長によるWelcome Speech

本ワークショップでは、板倉 文忠 名古屋大学教授、ベル研究所 B. H. Juang 博士による基調講演を皮切りに、10件の招待講演、39件の一般講演、そして最終日にはパネル討論も行われました。

アメリカ、ドイツ、韓国、中国、イギリス、イタリア、ベルギー、そして日本から延べ106名の参加者が、ハンズフリー音声通信を実現するために必要とされるエコーキャンセラー、雑音除去、頑健な音声認識、マルチモーダル音声認識などの各テーマに関して、基礎的研究課題から現実的な問題に対して、さまざまな視点から白熱した議論が交わされました。

Panel Discussionの様子

特に、古井 貞熙 委員長の司会によるパネル討論では、米国パナソニック社のJ-C. Junqua博士、エルランゲン・ニュルブルク大学のWalter Kellermann 教授、ベル研究所のChin-Hui Lee 博士、ダイムラークライスラー社のHelmut Mangold 博士、IRSTのMaurizio Omologo 博士、ベル研究所のFrank K. Soong博士、名古屋大学／ATR音声言語研究所の武田一哉 博士の各パネリストの討論に一般参加者の活発な議論も加わり、本ワークショップは大盛況のうちに幕を閉じました。

プログラムなどの詳細については、本ワークショップのWEBページ（http://www.slt.atr.co.jp/hsc2001/）を御参照ください。

ハンズフリー音声通信における要素技術のいくつかは、長い歴史を持つ成熟した研究分野です。しかし、ハンズフリー音声通信という包括的な観点ではまだ萌芽的研究分野であり、これまで総括的な議論が交わされる機会はほとんどありませんでした。

このような現状において、本ワークショップの成功は大きな意味を持ち、ハンズフリー音声通信に関するワークショップは、早くも次回の開催についての検討が進んでいます。

我々は、本ワークショップは、将来的には当該分野の重要な会議として位置付けられ、個々の研究を加速させるうえでも、また研究者あるいは開発者の情報交換の場として、非常に重要な役割を果たすであろうと考えています。

エイ・ティ・アール環境適応通信研究所

― やわらかいマルチメディア移動通信をめざして ―

環境適応通信研究所では、多様化する通信環境に対して、適応技術を用いることにより、その環境からの制約を緩和したり、時には意識しないで通信を行えるようなネットワークの構成法や制御法、アンテナやデバイスの研究をしています。この中で第二研究室では、ネットワークを始めとするさまざまなシステムの設計・制御の研究に取り組んでいます。ここでは、それらの研究の概要と進捗状況を報告します。

最適設計・適応制御技術の研究

情報通信ネットワークやさまざまなシステムを最適設計・適応制御するための手法を開発することが目標です。力学系の運動が持つ性質をシステムの設計・制御に利用するというアプローチを行っている点が特徴です。

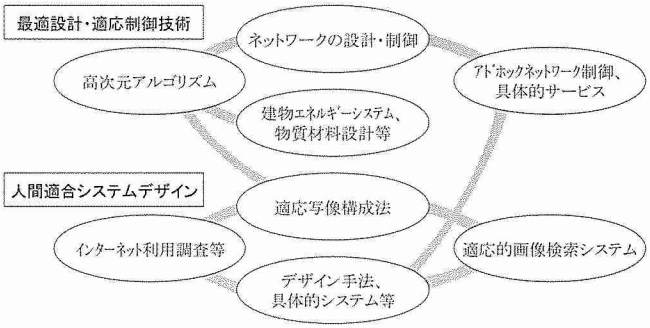
システムの設計・制御においては、与えられた制約条件のもとで、その性能を最大化することが望まれます。そのためには一般に、多くの変数を含む最適化問題を解く必要があります。その解法の一つとして、高次元空間で構成された力学系の自律的運動を利用する「高次元アルゴリズム」を提案しました。この方法は、変数の多い問題に有効である、広い領域で解を探索できるなどの特徴を持っています。この方法を基に、システム性能をほぼ最大にして、かつ、実装可能なアルゴリズムを得るための考え方を明らかにしました。

高次元アルゴリズムをネットワークに用いることで、複数の経路へのパケット振分け率（静的な経路配分率）を最適化でき、ノードサービス速度、リンク帯域などのリソース配分法、さらには、ネットワークトポロジーの設計法も提案できました。一方、個々のパケットの経路を決定する動的な経路選択も検討し、平均遅延時間の短縮を図るためには、トラヒック負荷に応じたルーティング政策を採用するのが良いことを明らかにしました。

ネットワーク以外のシステムへの適用例としては、建物エネルギーシステムの設計、物質材料の設計などが挙げられます。電力のみならずガスをエネルギー源とするコージェネレーションシステムの設計では、投資回収年数を最小にすることをねらい、装置の容量や運用を最適化する実務設計ツールを開発しました。本ツールは、共同研究先であるＮＴＴファシリティーズにおいて実用に供される予定になっています。物質材料設計では、ある性能を極大化する物質の探索において、任意の構造から安定な構造を計算できる手法を考案しました。

ネットワーク制御に関しては、高次元アルゴリズムから離れた検討も行っています。例えば、蟻コロニーをアナロジーとして、エージェントを利用した、ネットワーク障害に強い適応型ルーティングを提案しています。また、ルーティングアルゴリズムに対し、設計者側が指標としている遅延時間などとユーザの満足度を関連づける研究も行っています。

最適設計・適応制御技術の研究では、車車間通信を含むアドホックなネットワークに展開していくこと、性能が改善された材料物質を具体的に提示することを今後の目標としています。



人間適合システムデザインの研究

システムにとって人間（ユーザ）は環境変化の主要な要因の一つと考えられます。そこで、人間の応答や評価、行動に適合したシステムを設計するための考え方と方法の検討を進めています。

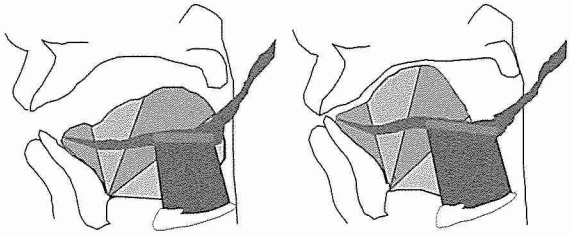
若い世代がインターネットや携帯電話などのコミュニケーションツールを利用するきっかけのアンケート調査・分析、総合スーパーマーケットの経営戦略の調査分析等を通して、消費者に受け入れられるモノ（システム、ツール、サービス等）を創り出すための考え方を検討しました。人がモノを利用するときの特徴として、例えば、性差による違い、すなわち、男性はモノと深くかかわり合ってモノの機能の中で動き回るのに対し、女性はモノ本来の機能の中に留まらず、日常生活で普通に起こる事柄と結びつけて使用するという傾向があることが明らかになりました。また、モノがユーザに受け入れられるためには、「驚き」、「鳥瞰」、「プライベート性」、「儀式性」といった要素が重要であると考えられます。

モノがユーザに受け入れられるかどうか、ある画像データが個人の判断基準に合致するかどうかという問題において、入力（モノや画像）と出力（ユーザの判断）を関係づける写像を具体的に構成することができれば、モノのデザインを考える上で強力な武器となります。そこで、適応的に写像を構成する手段として階層型ニューラルネットワークを採用し、その構造を高次元アルゴリズムで最適化できることを示しました。

個人個人の要求に合った画像をインターネット上で探し出してくる適応的なコンテンツベース画像検索システムとして、ニューラルネットワークを用いた構成を提案しました。このシステムはユーザおよびネットワーク上の不均一な計算機環境の双方に適応可能となっています。

人間適合システムデザインの研究では、デザイン手法の提案、人に受け入れられる具体的なツール、サービスの提案、特にワイヤレスアドホックネットワークなどでの具体的なサービスの提案をめざしていきます。

人間の生物機能に基づいた音声コミュニケーションの研究



実在する音声

「人間の音声は、思考や通信の道具として、あるいは情緒表現や個人識別の媒体として、人間の生活や社会の形成に大きな役割を果たしており、これらの諸機能は人間の生物としての働きに支えられている」、という意味のことをよく書くようになりました。最近はインターネットばかりで、あまり音声を使わなくなりましたが、日常生活を音声なしで済ますのはまだ抵抗があるでしょうね。自前の音声合成装置を持っているのですから、当然のことでしょう。

音声が人間の体で作られる、これも当たり前のことですが、体の方は電気通信では扱いにくいので、音声を録音したものを対象にします。こうして音声を人間の体から分離しておくと、音声は物理的に実在する、再現性があり科学の方法を適用できる、ということになります。この方面の研究はずいぶん進みましたが、反面、音声と人間との対応関係が希薄になりました。例えば、特定の個人の声質で音声を作るにはどうするのか、音声によって個人を識別するにはどうするのか、これらの問題は、体との対応関係の薄れた音声からは解決できないでしょう。

ことばを話す自動機械

前置きが長くなりましたが、ATR 先端情報科学研究部の音声生物科学プロジェクトでは、人間の音声の生物的側面を扱う研究を始めました。音声がどのようにして作られるかを調べ、音声と体との対応関係を理解した上で、人間の体の仕組をまねた音声合成や音声認識の方法を考えることを目標にしています。もともと、音声合成や音声認識は人間の機能を代行する技術なので、本来の姿に戻そうとしているだけのことです。今から 30 年前の音声研究の展望は、「人間のように話したり理解したりできる機械が 20 世紀末までに実現されるであろう」、でした。もしこのような機械ができたならば、機械に人格を認めることができるので、倫理問題にもなったはずですが、幸か不幸かその心配はまだありません。

発話器官の形態を計測して個人に適合する音響モデルをつくる、発話器官の運動を計測して発話運動モデルをつくる、発話時の脳の働きを計測して脳機能のモデルをつくる、これが音声生物科学プロジェクトの戦略です。時間のかかる作業ですが、完成すると人間の働きをまねた機械をつくることができます。動く人間の模型は、西欧社会ではやや不適合ですが、日本人には制限がないのでロボットの研究が盛んです。考えてみると、今から 60 年前に日本で行われた母音の研究は、声道の形から音声の特徴を計算できることを証明した画期的なものでした。これも、音声を計算するという発想が日本の思想と矛盾することがなかったからではないかと思います。

音声合成装置を獲得したサル

人間はことばが使えるのに、サルはなぜことばを使わないのか、という質問が時々あります。人間の体が二足歩行に適した形に変わり、発話器官の形も音声に適したものになると、大脳の機能がこの形の変化に適応して、ことばを話すことができるようになったのではないかと思いますが、物的証拠があるわけではありません。質問の答えとしては、大脳連合野の発達、随意運動機能の発達、発話器官形態の発達の３つを挙げることにしています。逆に、この３つが分かると、ことばを話す人間の模型を作ることができます。幸い、ATR 脳活動イメージングセンタが発足して、最新の磁気共鳴画像(MRI)装置と優秀なスタッフがそろいました。MRI による形態計測や３次元動画撮像により、発話器官や声道の３次元形状の時間変化を実測することができるようになりました。３次元動画撮像につてはこの巻に紹介記事がありますのでご覧ください。MRI による発話時の脳機能の観測も進んでおり、近々発表できる模様です。上の図は、NTT 誉田 CREST プロジェクトと共同で進めている MRI による生理計測の一例です。

[音声生物科学プロジェクト　本多清志]

最近とりとめなく考えること

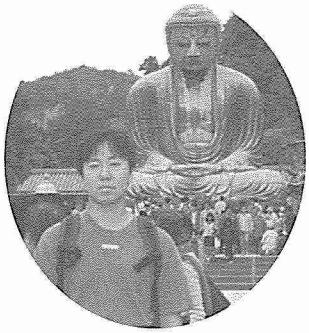
（株）国際電気通信基礎技術研究所 顧問
（前 国際電気通信基礎技術研究所 代表取締役 副社長）
吉田　匡雄

ATR-Iの副社長を退任して、はや3年が過ぎようとしています。ATRも創立以来15年が経過し、今までの仕組みを見直して、新しい制度の下で再スタートすることになりました。今までの研究資金提供の方法としての出資制度を改め、国からの研究委託費によって研究活動を行ってゆく制度になりました。昨今の国の構造改革による諸制度の見直しが進むなかにあって、ATRの研究活動の継続が認められたことは、ひとえにこの15年間にATRに在籍された方々がすばらしい研究成果を挙げられ、それらが世に認められ高く評価されたということであります。設立以来12年間ATRに携わってきた者として、この誌面をお借りしてこれらの方々に厚くお礼と感謝を申上げるものです。

設立時からATRの研究は「人間の研究」あるいは「人に学ぶ」ということを原点に出発しています。「人に学ぶ」というのは、最終ユーザである人の機能の偉大さを認識し、それをユーザ側の視点から研究を進めていくという人間尊重の立場に立つものであります。こうした立場から、最近のものすごい通信インフラの発展を見てみますと、例えばATRの発足時にキャッチフレーズとして言われた、どこでも・いつでも・誰とでも話せる腕時計レベルの超小型携帯機、その当時夢のようであった話が今や携帯電話のものすごい普及から、現実のものになっています。こうした使い勝手の良さを目指した科学技術の進歩発展が人間に大きな福利をもたらしているのは事実ですが、その反面、最近世間を賑わしているゲーム感覚の殺人事件や出会い系サイトのメル友に絡む事件など、使い方次第では個人や社会に大きな損害を与えるという面も目立ってきています。こうした影の面にも視点を合わせて、人間にとってどうあるべきか、すなわち人間尊重の立場からの配慮が、研究を進めて行く上でも必要になってきているのではないだろうかと思われます。

次にもう一つ。齢70を越えてみますと、身辺に心筋梗塞・脳梗塞で倒れられる先輩や同僚などよく見られるようになり、その人たちの闘病の話を伺ったり、闘病記を読む機会が多くなっています。先日新聞でつくば市でジュロンテクノロジーに関する国際会議が開かれたという記事を目にしました。ジュロンテクノロジーというのは10年程前に生まれた概念で、ジュロンとは加齢を意味し、ジュロンテクノロジーとは高齢者にとって快適な製品の技術を指すものです。従来の高齢者向けに設計された製品や機械が高齢者からすると、使い勝手が良くなく、ストレスが溜まるものが少なくない、もっと高齢者に配慮した設計が必要というのがこの発想です。実際脳梗塞で左手左足の機能障害を起こしてリハビリに励んでおられる先輩の闘病記を読むと、人間の機能がいかにすばらしいものであり、失って始めてそのすばらしさに気付くものだと書かれています。しかしここでも考えてみなければならないのは、人間の機能をまったく代替するものではなく、リハビリによる機能の再獲得を手助けするような技術の開発を目指すべきではないだろうかということです。最近の医学界ではホリステイック・ヘルス（心身総合健康運動）が唱えられ出されています。これは人間の精神力が絶大な病気治癒力を持ち、リハビリに対する効果にも深い関係を持っていることを主張するものです。前出の闘病記に紹介されているノーマン・カズンズの「笑いと治癒力」は、不治に近く回復の可能性は500分の1と宣告された膠原病の著者が、生きる意欲の結晶とも言うべき笑いとユーモアを総動員して「笑って」治した闘病体験記であります。人間の生きたいという意欲と然るべき動機や刺激を与えることにより、人間の持っている自然治癒力の驚くべき可能性を主張しています。人間の生きる意欲とそれを実現するためのリハビリへの意欲を喪失させるようなあまりにも便利なジュロンテクノロジーであってはならないし、そうした意欲を助長するテクノロジーであるべきだと思われます。それが人間尊重であり、人にやさしい技術であり、高齢化を迎えている日本社会のみならず世界社会への貢献につながるものだと考えています。

竜宮城を現実のものに



三菱電機株式会社 情報技術総合研究所 主務
（前 ATR 知能映像通信研究所 第三研究室 研究員）
田中　昭二

ATR を離れてからはや一年が経とうとしています。今思えば、ATR での四年間はまるでフラッシュライトのような一瞬の出来事のように思えます。ATR 着任時は27歳だった私も今や32歳とすっかり「おっさん」の域に達しようとしています。 ATR のすばらしい環境と優秀な人たちに囲まれて思う存分研究できたことは、私の生涯においてかけがえのない出来事になったことは言うまでもありません。本当にすばらしい経験をさせていただき感謝しております。

ところで、今回このコラムを執筆するに当たり、以前の ATR ジャーナルを読み返していたところ、ATR を竜宮城にたとえ、自らを浦島太郎にたとえて ATR の思い出をつづられた方のコラムに目が止まり、そのコラムを再度拝読させていただきました。そこで、自分が現在置かれている状況を踏まえて感じたことは、竜宮城はあくまでも夢物語であり、竜宮城から帰ってきた浦島太郎は、玉手箱を開けたとたんに現実とのギャップを痛感させられたことです。私も出向元に復帰し、現在製品開発に携わっているのですが、研究成果を本当に「使えるもの」にする難しさを痛感しています。

研究においては、研究者が主体となり、研究者自身が設定した制約条件の上で自らのアイディアの正当性を立証していくことができます。しかし、製品開発では、あくまでも主体は利用する側にあり、利用者が利用したいとき、利用したい場所で必ず動くものをつくらなければなりません。開発する製品が特定用途のものであり、研究時に設定した制約条件が許されるものであれば問題ないのですが、ほとんどの場合そうはいきません。言うなれば実用化を目指した研究開発というのは、そこからが本当の出発点となるのです。結構自分では良くできたと思っていても、実際に利用者の手に渡り使用してもらうと、予め想定していなかった状況で使われることが多く、実験環境で得られた精度が達成できないことが多々あります。そこから、さまざまな状況でも正しく動作するように何度もアルゴリズムを練り直し、それが終わると、実際に動作させるプラットフォームの仕様に合わせてさらに修正を加える。このような、いわゆる、「ドロ臭い」作業を納期という時間の制約と格闘しながら行い、やっとの思いで製品を出荷します。いわゆる「生みの苦しみ」というものでしょうか。でも、苦労した分、実際に製品を使ってもらい、「これいいね」という言葉を聞いたときの達成感は、格別のものがあります。

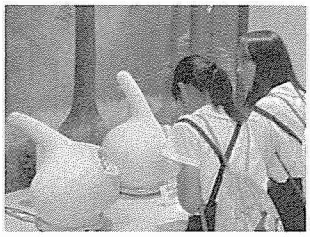
私は、研究も製品も「使ってもらってなんぼ」の世界だと思います。ですから、「良い研究」というのは、その研究において発明された新たな手法や、実験を通して得られた知見等が多くの人に利用されるものだと思います。

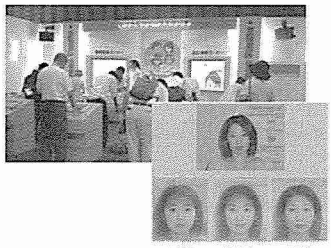
ATR から発信される研究成果はレベルが高く、我々にとって非常に魅力的ですが。どのように利用すれば多くの人に使ってもらえるものになるかが少々分かりにくいのではないかと感じます。言い換えれば、ATR 市場には非常に良い食材がそろっていて、後は、それらをどう料理すれば皆がおいしくいただけるかを考えることができれば世界一のレストランをつくれるのではないでしょうか。

ですから、良い研究成果が出た暁には、それを皆が広く利用できるような道筋を最後のお仕事としてつけていただき、竜宮城のお話を現実のものに近づけておいて欲しいです。そうすれば、私のような浦島太郎が安心して玉手箱を開けることができます。

山口きらら博NTT西日本館展示にATRの成果が採用される

2001年の夏休み、山口県は「元気」を世界と未来に発信する「きらら博」を開きます。NTT西日本（山口支店）では21世紀型パビリオンの形として各個人の顔にスポットをあてた展示をしています。そのNTT西日本パビリオン（「わくわく・どきどき ひまわり大魔神」館）にATRの研究成果である、映像研の「Muu」と人間研の「顔画像合成システム」を採用いただきました。



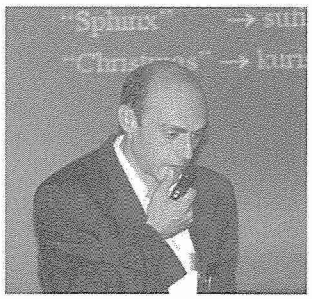


- 「Muu」：ATR知能映像通信研究所の研究成果。
  私たちの何気ない行動は、いつも周囲との関係に支えながらデザインされています。こうした何気ない行為とそれによって顕在化するコミュニケーションなどを構成的に議論するためのテストベッドとして開発された、社会的な相互行為の過程や幼児の発達過程、社会的な学習過程などを研究するためのクリーチャ（ロボット）です。
- 「顔画像合成システム」：ATR人間情報通信研究所の研究成果の応用。
  人間の認知機能の研究用ツールである顔画像合成システム「FUTON」をベースに、きらら博用に簡易なインタフェースでモーフィングを楽しめるように手を加えました。
  会場ではパビリオンの来場者の写真を撮り、来場者自らの操作で、自分自身の若い頃や老けた時の顔を推定合成します。さらに、自分の顔と動物の顔との合成もできるようになっています。

会　　場：山口きらら博（山口県阿知須町きらら浜）
期　　間：2001年7月14日～9月30日
展示場所：NTT西日本パビリオン（街のゾーン「わくわく・どきどき ひまわり大魔神」館）

― ATR科学技術セミナーの開催状況 ―

第87回 2001年6月29日（先端情報科学シリーズ 第5回）
音声知覚における音韻知識の役割：言語処理、言語習得、脳イメージ　エマニュエル・デュプー(CNRS)


Emanuel Dupoux教授

フランスCNRSのエマニュエル・デュプー教授に、人間の脳が音声を理解する過程に関してご講演いただきました。氏の研究室では、言語音の知覚過程に関して、母語の場合と外国語の場合の比較、乳児の研究、脳活動の計測などの研究方法を通した多角的な検討をされています。また、韻律に関連した題材を研究対象とされているのも特徴です。日本語話者がフランス語の音を聞く場合の研究もされており、“ebzo”という発音は、私たちには“ebuzo”と聞こえてしまうといった身近な例を通して内容を理解することができました。

講演の中で氏はいろいろな実験結果を通して、人間が生後1年までの間に自分の母語にあった“Phonetic Decoder”（音声情報処理システム）を確立すること、そしていったん確立されると母語を聞く場合のみならず外国語の音声を聞く場合にも使用されてしまうためさまざまな困難が生じていることなどを示されました。また、このPhonetic Decoderは音韻（母音や子音）ではなく、シラブル（音節）を単位として処理が行われるという仮説に基づいた実験結果も示されました。先端情報科学研究部でも、音声情報処理過程の理解は研究課題の一つとなっており、体系的に勉強する良い機会となりました。

最近のATR成果の展示状況

山口きらら博以外でもATR成果が展示されています。最近の展示および今後の予定は以下のとおりです。機会がございましたら是非ご覧ください。

イベント名	開催場所	期　間	展　示　物 （期間表示がないのは通期展示）
ザ・ロボット博21	長島町 長島温泉スパーランド特設会場	3/20～5/6	ロボビー　3/20～4/26展示 4/27～5/6 展示・実演
デジタル コミュニケーション2001	福岡市 博多スターレーン展示場	6/5～6/7	ロボビー
ロボフェスタ関西2001	大阪市 グランキューブ大阪	7/20～7/29	ロボビー Muu
21世紀☆みらい体験博	神戸市 神戸国際展示場	7/20～9/2	ロボビー　（ロボスクエア） 8/15～8/22展示 Muu　（ふしぎジャングル） インタラクティブ万華鏡 （ふしぎジャングル）
うつくしま未来博	須賀川市 須賀川テクニカルリサーチガーデン	7/7～9/30	ロボビー　（虹の台地 企業未来館） 9/1～9/2展示
ロボフェスタ生駒	生駒市 奈良先端科学技術大学院大学 ・高山サイエンスプラザ	7/31	ロボビー Muu
ロボフェスタ神奈川2001	横須賀市 南体育館	8/25～9/2	ロボビー
	川崎市 とどろきアリーナ	9/8～9/16	
	相模原市 銀河アリーナ	10/6～10/14	
	横浜市 パシフィコ横浜	11/16～11/25	

受賞等

年　月	学会・賞名等	受　賞　者	タイトル
2001/4/18	文部科学省2001年度文部科学大臣賞	中津　良平	感情による対話の可能なコンピュータシステムの開発

※詳細についてはhttp://www.atr.co.jp/journal/index.html を参照願います。

6月18日(月)

パンダちゃんの入学手続き

「ぼくのお名前はなに？」「パンダちゃーん！」これがパンダちゃんと交わした最初の会話だ。あっ、そうか中国から来たからパンダちゃんなんだ、と思った。どうやらテレビを見ながら片言の日本語を覚えたらしい。満面の笑顔のその男の子はポケモンのTシャツにウルトラマンの靴を履き、しっかりとお母さんの手を握りながら上を見上げていた。お父さんはニコニコとしながら本当の名前を口にしてくれたが、私には発音が難しく、うまく呼べなかった。それから私はその6歳の男の子をパンダちゃんと呼ぶようになる。

この日はお父さんが、留学していた東京の大学を卒業し、はるばる埼玉から電車や新幹線を乗り継いでATRの最寄り駅がある京都と奈良の県境に来た日だった。四月からここで研究員として働くことになり、その数日前に到着したのだ。三月下旬のよい天気のなか、社宅まで一家三人と手荷物をゴロゴロと引きながら歩いた。お父さんの一番の心配はパンダちゃんの学校のことだ。社宅に到着するまでに自然と小学校の話になった。「いまからでも入学式に間に合いますか？」「大丈夫ですよ。こちらの教育委員会に連絡してありますから。明日、さっそく手続きに行きましょう」お父さんはまたニコッとした。引いている荷物が少し軽くなったように見えた。

翌日、家族三人と駅で待ち合わせをする。駅前のバスターミナルから一日数本しかないバスに乗り、教育委員会に向かう。30分くらいの道のりなのだが、子供には長い時間だ。一つ二つバス停を過ぎてパンダちゃんは降車する人が押しているものに興味を持ったようだ。そして間もなく、エイッとボタンを押した。「ピンポーン！」「すみません！　間違えました！」と焦った私が大きな声を出す。「はーい」と運転手さん。一息ついて思った。こんな活発なパンダちゃんなら、学校でもすぐにお友達ができるだろう。

目的のバス停で降りると、まずは近くの役場で外国人登録証明書の住所変更を済ませ、それから教育委員会に向かった。カウンターで就学届けを記入し、健康診断書を手渡す。すぐに担当者は校区を調べ、小学校に電話をしてくれた。「明日、学校で教頭先生と面談して下さいね」と言いながら地図を渡してくれた。感謝、感謝。

次の日、今度は小学校に向かった。なぜかパンダちゃんはお母さんの手をぐいぐいと引っ張ってお父さんと私の前を歩いていた。あれ？　なんで道を知っているんだろう？　そういう顔をした私にお父さんは言った。「昨日うれしくて、うれしくて、先に学校を見に行っちゃったんです！」そう聞くと私も俄然気合いが入る。よしっ、しっかりとパンダちゃんを紹介するぞ。

学校に入るとさすがのみんなも少し緊張していた。教頭先生や他の先生方は入学手続きの説明を丁寧にしてくれた。給食の時に必要なナプキンのサイズや、体操着袋を机の横につるした時に床に付かない紐の長さ等、通訳する内容は多岐にわたる。私が英訳すると、お父さんが中国語に直してお母さんに一生懸命説明する。お母さんはそのつど深くうなずく。二時間ほどで面談が終了、あとは家族三人で入学式に行ってもらうとパンダちゃんは立派な一年生になる。「入学式では校長先生がお名前を呼びますから、大きな声で返事ができるようにしておいて下さいね」と教頭先生は言った。帰り道、パンダちゃんの「ハーイ！」の声が団地にこだました。

こんな具合にSHIENのサポートは行われている。家族が新しい環境にいち早く慣れ、安心して生活できることにより、研究員が仕事に専念できるようにすることが目的だ。今回は入学手続きの紹介をしたが、一般的に企業内業務と家族に関する事柄は切り離すべき、と思われる人も少なくないだろう。しかし、実際にはそんなことを言っていては世界の優秀な人材を集めることはできない。なぜなら家族を大切にする気持ちは万国共通だからである。

SHIEN 三神　恵

所員往来

平成13年4月1日より、6月30日までの間の採用および退職の方々は以下のとおりです。
（ただし、6ヵ月以上滞在の方のみ掲載）

採用年月日	ATR所属	氏　名	出　向　元　等
H13.4.1	（国）経営企画部 担当課長	岡留 剛	NTT
H13.4.1	（国）総務部 秘書室 係長	北野 清治	NTTインフラネット
H13.4.1	（国）経理部 担当課長	天野 了一	住友金属工業
H13.4.1	（国）先端情報科学研究部 研究技術員	上田 達也	サンモアテック
H13.4.1	（国）先端情報科学研究部 研究員	内部 英治	大阪大学
H13.4.1	（言）企画課 係長	若嶋 和政	NTT西日本
H13.4.1	（言）第一研究室 研究技術員	伊藤 玄	
H13.4.1	（言）第二研究室 室長	菊井 玄一郎	NTT
H13.4.1	（言）第三研究室 研究員	渡辺 太郎	Carnegie Mellon University
H13.4.1	（言）第四研究室 研究員	Jinfu Ni	東京大学
H13.4.1	（環）第一研究室 研究員	小菅 昌克	富士通関西
H13.4.1	（環）第一研究室 研究員	植田 哲郎	KDDI
H13.4.1	（環）第三研究室 研究員	橋口 正哉	三菱電機
H13.4.4	（環）第一研究室 研究員	山崎 達也	CRL
H13.5.1	（環）第三研究室 研究員	平田 明史	三洋電機
H13.5.7	（国）先端情報科学研究部 研究員	相馬 亘	京都大学
H13.6.6	（映）第四研究室 研究員	Christoph Bartneck	オランダIPO
H13.6.22	（国）代表取締役副社長	畚野 信義	テレコム先端技術研究支援センタ

退職年月日	転　出　先	氏　名	ATR所属
H13.4.13		中野 恵理	（国）先端情報科学研究部
H13.4.20	米国	劉　雲	（環）第四研究室
H13.4.30	住友金属工業	長原 基裕	（国）経理部 資金課
H13.4.30	富士ゼロックス	坂本 彰司	（映）第四研究室
H13.5.31	中国	Shuwu Zhang	（言）第二研究室
H13.6.13	神戸大学	阪田 真己子	（映）第三研究室
H13.6.22	CRL	酒井 保良	（国）代表取締役副社長
H13.6.30	NTT西日本	永野 清	（国）企画部 担当課長

正誤表

ATRジャーナル43号掲載内容に一部誤りがありましたので、以下のとおり訂正させていただくとともにお詫び申し上げます。

頁	誤　り	訂　正
19　受賞		
ARS ELECTRONICA Interactive Art部門	Honoraty Mention Prize	Honorary Mention Prize
（財）電気通信普及財団 第16回電気通信普及財団賞	杉原 俊昭	杉原 敏昭

イベントカレンダー

ATR研究成果の展示予定

山口きらら博（7月14日～9月30日）山口県阿知須町きらら浜
　NTT西日本パビリオンの展示に「Muu」、「顔画像合成システム」が採用されています
21世紀☆みらい体験博（7月20日～9月2日）神戸国際展示場
　ふしぎジャングルで「Muu」、「インタラクティブ万華鏡」が、展示されます
　また、ロボスクェアにおいて8月15日～22日、「ロボビー」が展示予定です
うつくしま未来博（7月20日～9月30日）福島 須賀川
　企業未来館において9月1日、2日、「ロボビー」が展示予定です
ロボフェスタ神奈川2001（8月25日～11月25日、横須賀、川崎、相模原、横浜で順次開催）の各会場で「ロボビー」が展示予定です

ATR研究発表会2001の延期について

平成13年11月29日、30日に開催を予定しておりましたATR研究発表会2001は平成14年2月中旬に延期いたします。詳しい日程につきましては次号でお知らせいたします。
また、http://www.atr.co.jp/expo2001/ でも逐次ご案内を行いますのでご参照ください。

お問い合わせは事務局〔tel (0774)95 1177〕まで

編集後記

速いものでATRに来てから2度目の夏を迎えようとしています。この間に多くの研究員を迎え、また送り出しました。多くの研究員が入れ替わりましたが、研究そのものは継続して行われています。平成8年に始まった7年間の「環境適応通信の基礎研究」プロジェクトも残すところあと2年を切りました。

昨年の夏は、「21世紀夢の技術展」に研究成果のひとつを出展し、多くの方に見ていただき、体験していただきました。

最近、世の中ではイヤなこと、信じられないことが次々と起こっています。ついこの間も多くの小学生が殺傷されるという事件がありました。でもATRの研究は続いています。

この夏、世の中で、ATRで、皆様そして私に、何が起こり、何を体験するのでしょうか。これまでの夏とは違うことがきっと起こり、違うことを体験することになるでしょう。

でも今、これを読んでいる皆様は、暑い暑い夏を今年もまた体験していることだけは確かでしょう。

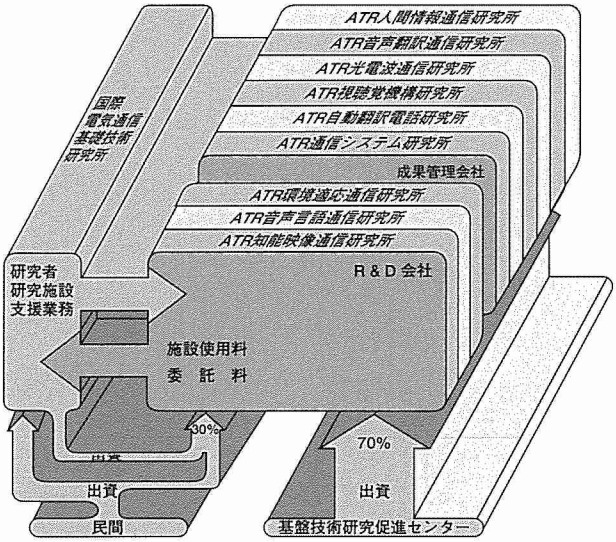
（ATR環境適応通信研究所 企画課長　熊谷　真）

ATR Journal 第44号　　**2001年8月1日発行**

●発行・編集　株式会社　国際電気通信基礎技術研究所
〒619-0288
京都府相楽郡精華町光台二丁目2番地2
（0774）95 1111（大代表）
●製　作　学会センター関西（財団法人 日本学会事務センター大阪事務所）
●定　価　800円（送料・消費税込み）

ご購読をご希望の方は、（財）日本学会事務センター o-publ@bcasj.or.jp へお申込み下さい。
本誌記事の無断転載を禁じます。
©2001　（株）国際電気通信基礎技術研究所

ATR グループのご紹介



ATR グループは電気通信分野における基礎的・独創的研究の一大拠点として内外に開かれた研究所を設立する構想のもとに産・学・官の幅広いご支援をいただき1986年3月に設立しました。

ATR グループは研究活動を行っている3つの研究会社(R&D 会社)と、既に研究を終了し成果の普及活動などを行っている6つの成果管理会社、およびこれらを支援する国際電気通信基礎技術研究所の10の株式会社の総称です。

R&D の研究費は基盤技術研究促進センターからの出資70％、民間からの出資30％で構成されています。

国際電気通信基礎技術研究所はR&D 会社に対し、建物スペース・研究施設の貸与・研究者の確保・派遣、研究資金の出資、研究企画の支援、各種事務の援助など、総合的な支援を行うとともに成果管理会社に対する研究成果の管理・販売などの各種の支援を行っています。

ATR ホームページ
http://www.atr.co.jp

役に立つ様々な情報を公開しています。
今後も随時拡充予定です。皆様のアクセスをお待ちしております。

ATR 知能映像通信研究所	http://www.mic.atr.co.jp
ATR 音声言語通信研究所	http://www.slt.atr.co.jp
ATR 人間情報通信研究所	http://www.hip.atr.co.jp
ATR 環境適応通信研究所	http://www.acr.atr.co.jp
特許と成果物	http://results.atr.co.jp

ATR ジャーナル担当宛

TEL ： (0774) 95 1177
FAX ： (0774) 95 1178
E-mail ： editor@ctr.co.jp

ご連絡内容（いずれかに印をお願いします。）

□ATR Journal 新規購読申込
□テクニカルレポート購入申込
【テクニカルレポート　番号：TR－　　－　　】
□ご意見、ご要望等
□送付先変更連絡
□研究用ソフトウェア購入申込
【ソフトウェア名整理番号：　　】

		変更後	変更前	変更事由
送付先	お名前（フリガナ）			□人事異動 □住所変更 □その他
	送り先			
	会社名			
	部署名			
	役職名			
	Tel / Fax			
	E-mail			
ご意見ご要望				

●ATR ジャーナルのご購入希望、送付先変更等をお寄せ下さる場合には、上記にご記入の上、FAX 等でご送付下さい。
●送付先変更以外については、変更後の欄に必要事項をご記入願います。

株式会社 国際電気通信基礎技術研究所

ATR Journal No.44