

ATR Journal

WINTER 1997

26



〈表紙解説〉

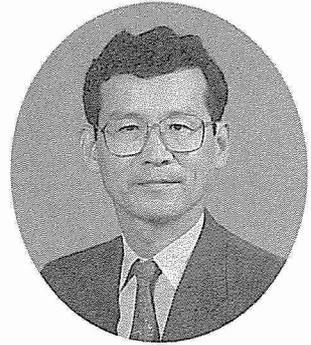
ATRでは、ある話し手の声を音声データベース化しておき、出力したい文に応じて、音韻の並び方や声の高さ、音韻の長さなどの条件が良く適合し、しかも滑らかにつながる音韻を選び出し、それらの音声波形をつなぎ合わせて出力するマルチ言語マルチ話者音声合成システム CHATR (チャター) を開発しました。CHATR の合成音声は <http://www.itl.atr.co.jp/chatr/> で聞いていただけます。詳しくは、「マルチ言語マルチ話者音声合成システム CHATR —お気に入りの声をコンピュータで合成—」をご覧ください。

●巻頭言	言語の壁に風穴を	1	酒井 保良
●報告	中間時試験研究報告を終えて —異なる言語間のグローバルコミュニケーションを 目指す音声翻訳通信の研究—	2	山崎 泰弘
●研究動向紹介	仮想変身システム —自分の姿を好みの姿に変えられるシステムを目指して—	6	大谷 淳 海老原一之 榎沢 順
	マルチ言語マルチ話者音声合成システム CHATR —お気に入りの声をコンピュータで合成—	8	ニック・キャンベル 樋口 宜男
	考えることと喋ること	10	正木 信夫
	ソフトウェアアンテナ —知能と変身の術を備えたアンテナ—	12	唐沢 好男
●ATR Monologue	ゼロからのスタート (その1)	14	葉原 耕平
●短信	大和の国の物語 —けいはんなの丘に寄せて—	16	高橋 保
●報道発表		17	
●トピックス	第9回研究発表会を開催	18	
	ATR 科学技術セミナーの開催状況	19	
●学研都市あれこれ	福寿園 CHA 研究センター訪問	20	
●登録特許の紹介		22	
●イベントカレンダー		23	
●所員往来		23	

言語の壁に風穴を

(株) 国際電気通信基礎技術研究所

代表取締役副社長 酒井 保良



年が変わって20世紀も残り4年足らずとなった。今世紀を代表する諸製品は、自動車、家電、衣類、薬品等々いずれも我々の生活に密着した「物」であり、今世紀は言わば「物の時代」と位置付けられています。併せて、世界規模での武力戦争（資源と市場の争奪戦）の時代であり、地域的にはアメリカの時代だったともいえます。

それでは21世紀はどうなるのであろうか。よく言われる言葉として、「情報化の時代」「地域紛争の時代」「ボーダレスの時代」などがあり、これらに対してそれなりの準備を進める必要があると考えます。

我国は、自動車、家電製品等では、世界市場においてトップシェアに輝きました。そのお陰で円が高騰し、多くの日本人が気軽に海外旅行や海外製品を楽しめるようになり、いろんなレベルでの国際化が進展しています。しかし、大多数の日本人は外国語を不得手とし、本音での意思の疎通に至る真の国際化にはほど遠い状況にあります。自動車や家電製品は言語や生活様式に依存せず、ほとんど世界共通であるため国内用の技術がそのまま世界市場に通用しましたが、情報化時代はどうでしょうか？ 情報化時代が具体的にどのような時代になるのかはよく見通せませんが、少なくとも、情報の基本は言語であり、日本語で作成された情報のサーキュレーションは、英語に比べて1/10以下という状況は変わらないでしょう。

情報化時代においては、「情報」そのものが製品として市販され、流通し、人々によって使用され、改良されていくと考えられます。ゲームやアニメーションやパソコンソフトはその先駆けでしょう。世界中の経済活動の主流が、手工業から機械工業に移った18～19世紀の産業革命と同等またはそれ以上の変革が、すでに進捗しつつあると言われる情報革命で起こるとすれば、一体どのような世界ができてくるのでしょうか。真のグローバル化の結果として、国家が消滅し貨幣も一元化され電子化されるのでしょうか？ 万一そのような事態になったとしても、人々が日常使用する言語は、そう簡単には変えられないでしょう。自然言語は、人類の歴史の中で自然発生的に生まれ、時代と共に進化してきたものであり、よほどの事情がない限り、それぞれの言語圏において、生活必需品としてのみならず、思考や対話の最強の手段として活用され続けるでしょう。

日本語を基本とした「情報」活動で、現在の自動車や家電製品と同等の経済効果を達成するのは、極めて困難であろうということは容易に想像できます。しからば皆で英語を勉強するのか？ 英語だけで良いのか？ この問題に対するひとつの解が、「言語翻訳」と考えます。「情報」は、いずれは文字化・文章化されますが、オリジナルは音声での対話・議論から生まれてくる場合が多いと思われる。かのソクラテスはプラトンと、デカルトはパスカル達と、道元禅師はその師如浄との対話を通じてかねてからの疑問を解きほぐし明確な考え方・思想として整理するに至ったと聞いています。

相矛盾するかまたはお互いに不完全ないくつかの「情報」がぶつかりあい、議論を通じてのコンセンサスで新たな「情報」に変質する時点が最も重要であり、ここに参画していた者のみが、その「情報」の発信者として認知され処遇される時代においては、少なくともお互いが殆どリアルタイムで音声情報を交換しあえる環境が必要になります。

このレベルの「音声翻訳通信」が簡単に実現できるとは思えませんが、21世紀にかけて我国が率先し、できるだけ多くの言語圏の研究機関と協力して、取り組むべき重要な技術課題であることは間違いありません。幸いにも、ネットワークの高度化・低廉化・パーソナル化や、表情・動作・イメージを含む映像通信を活用したマルチモーダル化技術も着実に進展しつつあり、またゲームやアニメーションが国境を越えて楽しられている昨今、これらを総合して少なくとも若者同士は、言語の壁に風穴を明けながら議論し、新たな「情報」を発信できるようにしたいものです。

中間時試験研究報告を終えて

—異なる言語間のグローバルコミュニケーション
を目指す音声翻訳通信の研究—

(株)ATR 音声翻訳通信研究所
代表取締役社長 山崎 泰弘



① はじめに

ATR 音声翻訳通信研究所は「高度音声翻訳通信技術の基礎研究」をテーマに、1993年から2000年の7年間の研究を進めています。ATR 自動翻訳電話研究所では朗読調の話し言葉を対象に自動翻訳電話技術の研究を遂行しました。当研究所の目標は、この先駆的技術レベルをバネにして、普通の会話に現れる会話調の自然な話し言葉を対象として、音声翻訳通信の基礎技術を確立することです。すなわち、連続的に喋られた音声を認識し、相手方の言語に翻訳し、個性豊かな合成音声を生成するための個々の要素技術とそれらを統合化する技術の確立を目指しています。1996年9月、研究期間の中間点である3年半を経過したこの時期に、本試験プロジェクトの最大出資機関である基盤技術研究促進センター（KTC）の規定に基づき、これまでの研究進捗状況を中間時試験研究報告としてまとめ提出しました。また、今回からは将来の成果展開の見通しも報告が義務づけられ、中間時評価報告（経済性評価）として提出しました。

報告会では、基礎基盤研究に取り組む研究体制および研究の進展は高く評価されました。一方、本試験研究の中核技術である音声言語処理技術は、21世紀に最も期待される技術分野の一つであり、研究成果が具体的なアプリケーションとして展開されることを想定して研究に取り組むことも必要であるとのコメントも寄せられました。以下に報告概要を述べます。

② 試験研究の概要

通常の話し言葉には特有のくだけた発声があり、時には言葉が省略されたり、順序が倒置されたりします。また、話し手と聞き手の常識、習慣に基づき、敢えて言葉で表現されない意図もあります。もう少し詳しく見てみると、図1に示すように認識を難しくする音声現象として、話者によって

異なる声色、抑揚・強調のある発音などがあげられます。また、翻訳を難しくする言語現象として、会話の状況に依存した多様な表現、省略、倒置などのくだけた表現などがあります。これらの問題点を解決するために以下の4つのサブテーマを設定し、音声翻訳通信の基礎技術の確立を図っています。

(1) 自然発話音声認識技術

話す早さや音の大きさの変動に対して音響モデルや言語モデルを動的に適応して、不特定話者の自然な話し言葉を認識する技術を確立します。

(2) 発話韻律処理技術

人間が喋る音声には、文字で表現できる狭義の言語情報以外にアクセントやイントネーションの韻律情報や発話の意図などを暗に示す汎言語情報が含まれています。この狭義の言語情報に汎言語情報を加味し、自然で個性豊かな合成音声の生成技術を確立します。

(3) 協調融合翻訳技術

自然な会話には文法の枠にとらわれない多様な表現が現れます。このため規則に基づく翻訳だけでは限界があり、対訳用例を適宜選択し翻訳する手法と規則主導の翻訳とを融合化し、精度の高い翻訳技術を確立します。

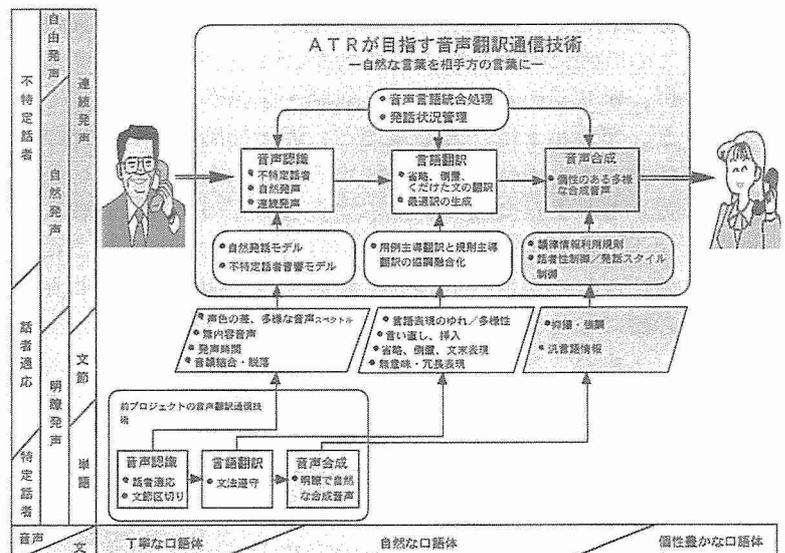


図1 研究の目標と位置づけ

(4) 音声言語統合処理技術

音声情報と言語情報を互いに利用し、その相乗効果により、自然な会話の中に生ずる曖昧さの解消を図ります。この結果、音声認識、言語翻訳、音声合成の精度を高め、音声翻訳システム全体の性能向上が可能となります。

3 研究活動と成果

7年の研究期間のうち、前期2年間（平成5～6年度）は自然な会話に現れる音声・言語データの分析と基本アルゴリズムの確立を図りました。この期間の具体的成果として、例えば話し始めてから4～5秒で話者の特性を把握し、不特定話者音声認識へ適用できる話者クラスタリング方式の提案、対訳用例翻訳の中核技術である高速類似検索法の考案、などがあげられます。特に高速類似検索法は、10万件の用例言語データベースから最適用例を1ミリ秒で検索できるもので、対訳用例翻訳だけでなく曖昧文献検索など幅広く利用できる技術です。

中期2年間（平成7年～8年度）は機能モデルの構築と要素技術の検証に研究の重点をおきました。具体的成果としては、大語彙音声を実時間で認識処理が可能となる会話調音声認識技術の提案、日本語から英、韓、仏語への言語翻訳技術の実証、自然で個性豊かな音声合成システムCHATRの構築、などをあげることができます。

研究成果は直ちに特許出願するとともに適宜、学会、国際会議、学術誌へと積極的に発表しました。この積極性は特許出願70件、学会発表521件という件数で裏付けられるとともに、科学技術庁長官賞、日本科学技術情報センター賞、電子情報通信学会論文賞など外部から10件（31名）の表彰を受け、質の面でも高い評価を受けることができました。

音声翻訳通信の研究は、外国語を対象とする研究の性格上、国際研究協力が必須です。1993年1月の自動翻訳電話国際実験の成功はC-STAR（Consortium for Speech Translation Advanced Research）による研究協力の成果です。この実験の成功を契機に、この組織はATRを初め従来の4機関を中核にC-STAR IIと改編され、現在、図2に示す世界の21の主要研究機関が参加する程に拡充され、この研究分野の世界的隆盛の先導役を果たしています。ATRは設立時より中核メンバーとして参画し、1996年の要素技術の検証、1999年の多言語音声翻訳国際実験を提唱するなど、研究計画でも国際的リーダーシップを発揮するとともに、議長を務め、その運営、取りまとめにも努めています。

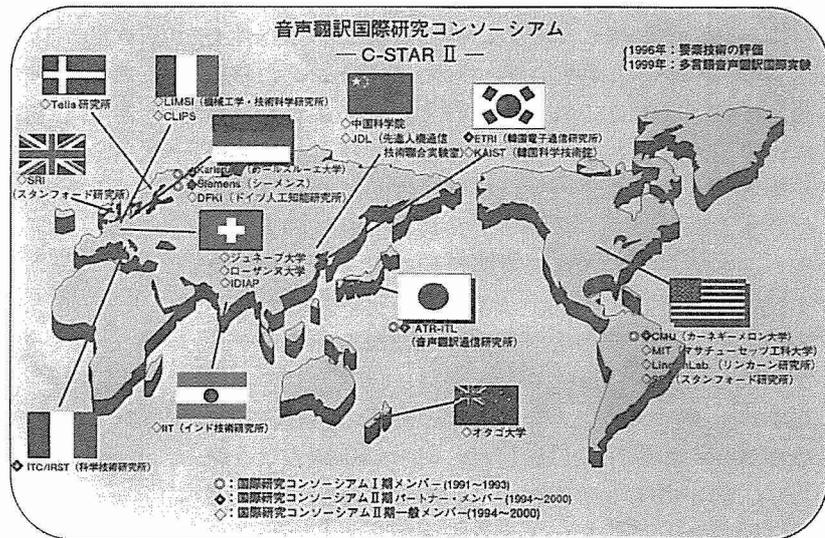


図2 国際研究コンソーシアムのメンバー

1996年9月にはC-STAR II会合をATRで開催し、各要素技術の技術レベルを検証し、1999年の国際実験の見通しを立てることができました。研究成果は専門家に対してだけでなく、一般の方々への周知を図りました。3年半の報道発表として、新聞93件、雑誌31件に掲載され、TV等10件で放送されました。時にはTV放送直後に視聴者から直接「音声処理技術は身体障害者にとっても頼りになるもので、一日も早く実用化してほしい」との激励のコメントが寄せられることもありました。また、けいはんなフェスティバル、APEC大阪会議展示、無線100年記念展示での展示を始め、年間1万人を超えるATR来訪者に音声翻訳システムの現状を紹介しました。入力された日本語が英語に翻訳され、合成音声で出力されると、頷いたり、にっこりするなど音声翻訳通信への関心の高さを伺い知ることができました。

4 要素技術の取組み

(1) 自然発話音声認識の研究

通常の会話のように自然に発声された音声では、発話速度の遅速、話し方の違い等によって引き起こされ種々の音響的な変化や、話し言葉が持つ語順の自由度、「あー」「えー」といった間投詞の挿入等、書き言葉には見られない種々の音声言語現象がみられます。自然発話音声認識の研究では、これら自然発話音声の本来持つ種々の変動にも対処可能な頑健な音声認識技術の確立を目指し、音声認識用音響パラメータ表現・音響モデル構築技術、言語情報利用と音声パーキング手法の研究を行いました。

また、年齢、性別、出身地など話者個人の違いに起因する音声スペクトル特徴のばらつきは、自然発話の変動と合わせて、音声の自動認識を困難にしている最も大きな要因です。このような不特定話者に対する音声認識性能を向上するため、話者が発する

音声を用いて認識システムの持つ音響モデル特性を話者のスペクトル特性に近づけてマッチングの精度を上げる「話者適応技術」の研究を進めました。特に、学習サンプル量に合わせた種々の統計的適応方法を考案し、それらの複合技術でさらに高精度化を図りました。一方、音響モデル自体の性能向上を目指し、多数話者の大量の音声データベースを用いた不特定話者モデル構築手法の研究を行いました。

加えて、これらの技術確立に必要な自然音声データ収集を精力的に進めるとともに、自然発話音声認識プロトタイプ・システム構築のための各種音声認識モジュールを作成しました。以上の研究成果を統合し、図3に示す単語グラフを出力する自然発話音声認識プロトタイプシステムを完成しました。この音声認識エンジンにより、「旅行に関する打合せ」のタスクを対象とした数千単語の自然発話音声に対する準実時間処理の見通しが得られました。

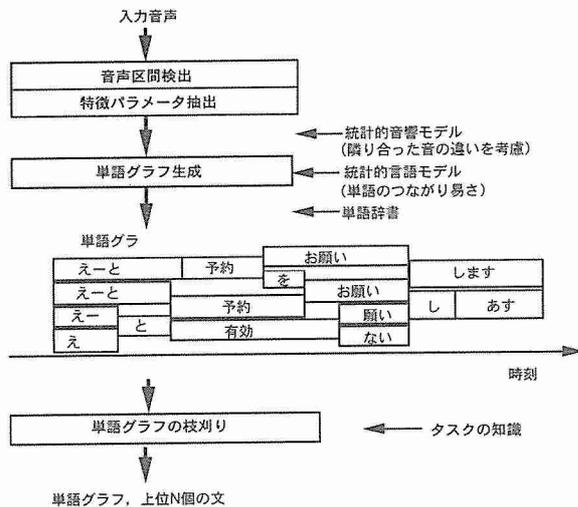


図3 単語グラフを用いた自然発話音声認識

(2) 発話韻律処理の研究

本サブテーマでは音声翻訳システムにおける音声合成部の高度化の研究と、韻律情報の利用の研究を進めています。(注：ここで言う韻律情報とは、音声の高さ・強さ・言葉のリズムなど、音声に含まれる音韻と個人性以外の情報のことです。)

音声合成部の高度化の研究では、音声データベース中から、出力したい文に応じて、音韻の並び方や声の高さ、音韻の長さなどの条件が良く適合し、しかも滑らかにつながる音韻を選び出し、それらの音声波形をつなぎ合わせるという方法を開発しました。

(注：詳細は本号の研究動向紹介8-9頁をご参照下さい) この結果、特定人物の30分~60分程度の生音声があれば、任意の文章をその人物の声に近い合成音声で出力できることが可能となりました。また、音声翻訳システムの利用者の声質や話し方の特徴を少ない量の音声データから学習し再現するために、

人間が話者を判断するときの声のどの特徴にどれだけ着目するかを予測するモデルを提案するとともに、図4に示す発話特徴模擬システムSSSS (Speech-Style Simulation System)を試作しました。

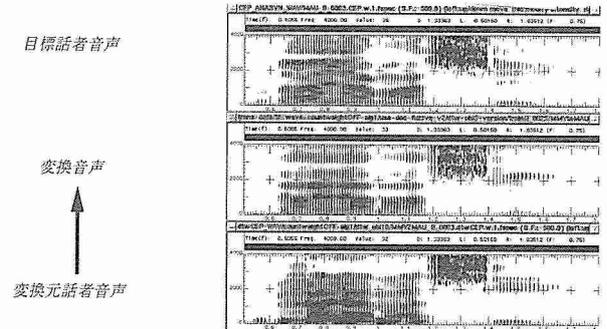


図4 発話特徴模擬システムのスペクトル

さらに、入力された音声の韻律特徴から疑問の終助詞がなくても平叙文か疑問文かを識別したり、文中に現われる躊躇の意図を検出したり、強調箇所を検出したりする手法の研究を行ない、いずれも約8割の正確さでこれらの特徴を識別することができ、言語翻訳精度の向上に見通しを立てることができました。これと併せて、ToBI (Tones and Break Indices) やピッチ周波数生成モデル・パラメータを用いて各国語の韻律特徴を記述した多言語韻律データベースの構築を行い、多言語音声合成の研究に資しています。

(3) 協調融合翻訳の研究

音声翻訳通信のための翻訳技術の研究を進めるに当たり、まず既存手法である文法遵守の翻訳手法や仮想的な中間言語を利用する中間言語方式の翻訳手法などと用例主導翻訳手法との比較分析を進めました。自然な話し言葉の特徴である多様で豊富な言い回しや文法の枠を超えた表現を取り扱う点で、用例主導翻訳手法の効果が大きいことを示しました。さらに、複雑な句や複文などをも翻訳するために、句や文を構成する表現の依存関係をボトムアップに解析しつつ、同時に対訳用例を適用して、最も好ましい用例を選択するメカニズムを実現しました。この機構は、用例と規則という両端に位置する情報を統合的に取り扱うとともに、両者を共通の枠組みで取り扱うことができる融合方式となっており、当初目標である文法に基づく言語解析手法と対訳用例を活用する用例主導翻訳手法とを融合する「協調融合翻訳」技術の骨格を作ることに成功しました。さらに、このメカニズムを核として、翻訳のための知識である表現例の言語パターンや対訳例などを翻訳の言語対ごとに用意することにより、多言語翻訳のプロトタイプ・システムを構築しました。具体的には、日英双方向、日韓双方向、ならびに日独方向の会話翻

訳システムを作成し、キーボードからの自由入力による各種の翻訳実験ができるシステムを築きました。日英双方向翻訳実験においては旅行アレンジに関する収録会話データの10万語規模の表現例から約1,000種の対訳用例パターンを用意し、平均約0.5秒の処理時間で最適翻訳結果を一意に提示することを可能としました。また、今後の用例数増加にも対応できる技術として、10万件の用例データから最適な用例を1ミリ秒で検索する高速用例検索技術を実現しました。

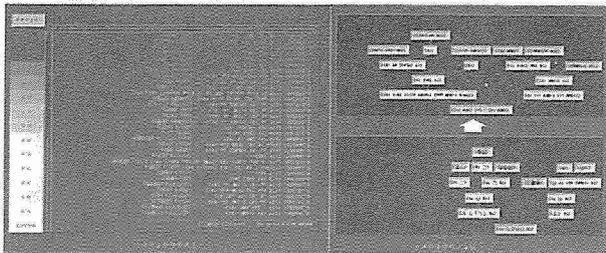


図5 協調融合翻訳の実験

(4) 音声言語統合処理の研究

従来の音声翻訳システムでは、音声処理、言語処理がそれぞれ独立にかつ一方に処理されていました。しかし、このような方式では自然な話し言葉を処理することは困難であり、両者を有機的に統合可能とする新たな方式を実現しなければなりません。このため、以下に述べるような発話状況管理技術、マルチモーダル・インタフェース技術について研究を進めました。発話状況を管理し、これを利用することにより、文脈に依存した表現の適切な翻訳や、文脈的に整合性を欠いた音声認識結果を排除することができます。発話状況として、対話における発話相互の関係、すなわち対話構造と、ある時点における話題を認識することが重要です。このため各発話の文末表現や手掛かり語を用いて発話間の関係を認識する手法を開発しました。また、このような対話構造と、発話に現れた語彙間の関係から、話題を認識する手法の研究を進めました。さらに、これらの発話状況情報を音声認識の曖昧性の解消や、言語処理での照応処理に利用する技術についても研究を進めました(図6)。特に、音声認識の曖昧性解消方式として、対話構造の一種である発話タイプの統計量を用いて次発話タイプを予測する方式を提案し、その有効性を示しました。

将来の音声翻訳通信システムにおいて、音声とともに他の情報伝達手段(モーダル)を利用可能とすることにより、コミュニケーションをさらに豊かで効率的なものにすることができる。このようなマルチモーダル・インタフェースの具体的な実現方法や、その有効性を明らかにしました。

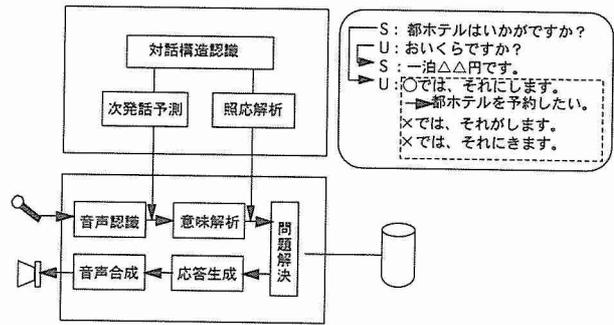


図6 音声認識や意味の曖昧性を解消する発話状況管理技術

5 今後の研究計画

(1) 研究方針

本試験研究の目標とする音声翻訳通信技術に対する社会の期待が高まりつつあり、国内外の研究も活発化してきています。このような傾向を先取りする形での本試験研究の設定は概ね妥当であったと思われます。しかし、音声、言語という研究対象は極めて人間的要素が強く、しかも7年間という研究期間を考慮するとコンピュータで扱える範囲には自ずと限界があります。中間時までには得られた要素技術の成果を踏まえ、プロジェクトの最終技術目標を具体的に見定め、効率的に研究を進めます。

音声翻訳国際研究協力コンソーシアムC-STAR IIに中核メンバー(Partner member)として積極的に参画し、1999年の多言語音声翻訳国際実験に向け最大の努力をします。また、研究成果は適宜、特許、論文等にまとめるとともに、成果普及に努めます。さらに、報道発表、各種展示などを介し、本試験研究に対し一般の方々の理解を得よう努めます。

(2) 研究計画

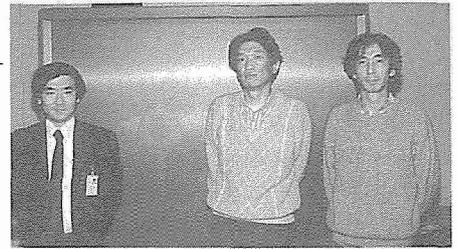
前期(平成5年~6年)に収録した音声的、言語的データの特徴を分析し、基本アルゴリズムを確立しました。この成果をベースに中期(平成7年~8年)には機能モデルを構築し、各要素技術を検証し、プロジェクトの最終技術レベルを見定めます。後期(平成9年~11年)の前半には各要素技術の規模を拡大しつつ、これらを有機的に結合し、システム化を図ります。後半には音声翻訳統合実験システムとして統合化し、最終評価を行います。国際研究協力においても本試験研究計画と同一歩調で進んでいるC-STAR IIのマイルストーンを堅持し、進展を図ります。1999年には上記統合実験システムを利用し、多言語音声翻訳通信の国際実験を実施します。2010年~2020年に実用化が期待されている音声翻訳通信を十分意識し、上述のように研究を進め、本試験研究の目標である「音声翻訳通信の要素技術と統合化技術」の確立を目指します。

仮想変身システム

—自分の姿を好みの姿に変えられるシステムを目指して—

(株)ATR 知能映像通信研究所
第一研究室

大谷 淳、海老原 一之、榎 順



誰もが自分の姿を、任意の姿に仮想的に変えることが可能なシステムの実現を目指して筆者らが行っている研究を紹介します。ここでは、このような仮想変身システムへの第一歩として開発された、バーチャル歌舞伎システムについて述べます。本システムは、歌舞伎役者の3次元モデルの作成処理、歌舞伎役者に変身しようとする人物の表情と全身の姿勢の非接触方式による実時間推定、推定された表情と姿勢の役者モデルにおける実時間再現、の各処理モジュールから構成されています。バーチャル歌舞伎システムを実際に構築し、実演デモを行い、本システムの有効性を確認しました。

① 変身願望

変身は、おそらく誰もが、程度の差こそあれ、いだいている願望だと言えます。例えば、女性はある年齢に達すれば、ほとんど全員が化粧という変身技術を自分自身に施します。(もっとも、最近では、男性でも化粧をする人がいるようですが)。化粧は顔を中心に行いますが、アクセサリ類や服装等を変えることで、気分や雰囲気を変える、というのもしばしば多くの人が行うことであり、自分のありのままの姿を変えたい、という気持ちの表れと言えるでしょう。また、もっと積極的に、「自分の姿を(自分以外の)別の姿に変えたい」という気持ちを反映した行動もしばしば見られます。遊園地や観光地には、図1のように、顔の部分だけが丸く切り抜かれて穴になっている、怪獣やアニメーションのキャラクターの立て看板がよく置いてあります。子供だけでなく大人さえも喜んで立て看板の裏から自分の顔を穴から見せて、家族や友人のカメラのファインダーに収まっている風景を目にします。言うまでもなく、彼らは自分の姿を別の姿に変えて楽しんでいるわけです。似たようなこととしては、誰もが子供の頃、「忍者ごっこ」、「怪獣ごっこ」といった「○○ごっこ」を経験していると思いますが、これも「自分の好きなキャラクターに変身したい」という子供心の表れでしょう。「○○ごっこ」は子供だけのものかということ、アメリカのユニバーサル

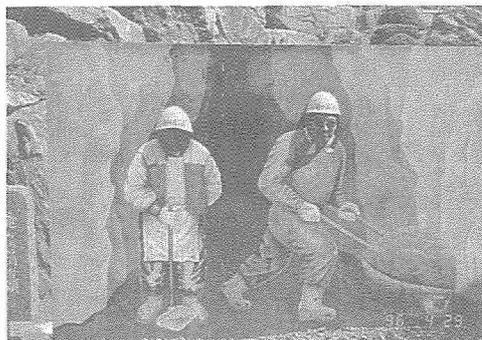


図1 立て看板による変身(昔の鉱山労働者に変身)

スタジオには、テレビドラマ等の番組の場面に、希望するお客さん(ほとんどが大の大人)が番組の役柄として登場できるというアトラクションがあり、大変な人気です。このように、変身願望というのは、世の老若男女を問わず、世界共通の欲求とすら思えてきます。本稿では、誰もが任意の姿に仮想的に変身可能なシステムの実現を目指して行っている筆者らの研究について紹介します。

② 新しいコミュニケーション手段としての仮想変身

離れた場所の人物同士のコミュニケーションの手段は現在でも電話が中心で、映像情報を用いたコミュニケーション手段は、テレビ電話等が研究されているものの、必ずしも積極的に利用されているとは言えません。その理由の一つに、「自分の生の顔が(相手のテレビに)映るのが嫌だ」という意見が根強く存在するようです。と言うことは、映像を中心としたコミュニケーション手段の普及のためには、①で述べたように、人間の変身願望を満足させる技術の実現がポイントの一つと言えるのではないのでしょうか。

第一研究室では、離れた場所にいる人物同士の、映像情報に基づく仮想的な空間を介したコミュニケーション環境の創出に取り組んでいます。このようなコミュニケーション環境のスコープの一つとして、①で述べた「仮想的な変身」が位置づけられます。仮想変身システムを実現するためには、変身しようとする人物の表情と体の動きを検出し、別の姿において実時間で再現する必要があります。その一例として筆者らは、誰もが歌舞伎役者に変身できるバーチャル歌舞伎システムを開発しました。以下にその技術的な内容について述べることにします。

③ バーチャル歌舞伎システム

バーチャル歌舞伎システムは、(1)歌舞伎役者の3次元モデルを作成しておく処理、(2)変身しようとする人物の表情と全身の姿勢の、非接触方式による実時間推定、(3)推定された(2)のデータの歌舞伎役者モデルにおける再現、の3つの主要な処理モジュールから構成されています(図2)。

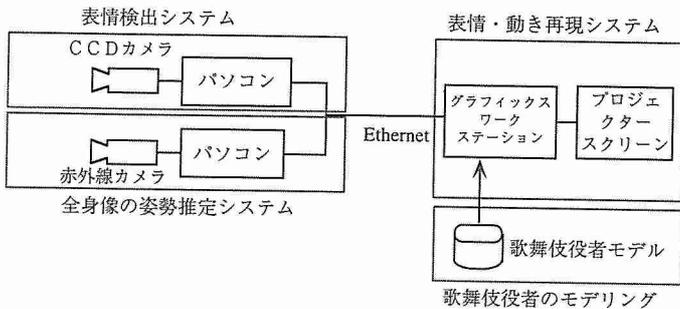


図2 パーチャル歌舞伎システムの構成

(1) 歌舞伎役者のモデリング

歌舞伎役者のモデルを、ワイヤフレームモデル、即ち、小さな三角形のパッチの集合によってあらかじめ作成しておきます。ここで、衣装や皮膚の色彩情報であるカラーテクスチャをワイヤフレームモデルにマッピングしておきます。歌舞伎役者のモデルは、人物の動き情報に従って、適宜変形することができます。図2では、歌舞伎役者モデルは、グラフィックスワークステーション Onyx Reality Engine 2 (Silicon Graphics 社製) に蓄積されます。

(2) 人物の表情と全身の姿勢の実時間推定

歌舞伎役者に変身しようとする人物の表情と全身の姿勢の推定は、従来センサ等を装着することが多かったのですが、動きの妨げになる等の問題があり、非接触な方式により実時間で推定することが望まれます。

顔の実時間表情検出は、筆者らが開発した周波数領域変換に基づく手法 [1] を用いて行います。歌舞伎を演じる人物は、カメラが顔に対して常に一定の位置にあるように、小型 CCD カメラを固定したヘルメットを被ります。カメラから得られた顔画像における目、口等の顔要素の領域に対して周波数領域変換の一つである DCT (Discrete Cosine Transform) を施し、表情特徴として、縦、横、斜め方向の DCT エネルギーを求めます。このような DCT 特徴は、顔要素の形状変化に対応していると考えられます。次に、DCT 特徴を、あらかじめ学習により求めておいた変換式により、各顔要素の周囲の代表点の顔画像中における動きベクトルに変換します。一方、人物の全身の姿勢は、赤外線カメラにより獲得される熱画像 (図3左) から推定します [2]。ここで、熱画像を用いたのは、人物の体に対応する部分が背景から容易に識別でき、照明条件や衣服の色等のノイズ要因によらず安定に人物のシルエット像が抽出できるからです。人物のシルエット像から、頭頂部や手先等の特徴点をまず抽出します。人物の全身の姿勢を推定しようとする場合、これだけでは十分ではなく、肘、膝といった主要な関節の位置を推定する必要があります。このために、前述の特徴点の抽出結果から、学習的に推定する手法を開発しました (図3中。四角が各点の位置の推定結果を示します)。

(3) 人物の表情と姿勢の歌舞伎役者モデルにおける再現

(2) で得られた表情検出結果 (代表点の動きベクトル) は、歌舞伎役者モデルにおける顔部分の変形に用いられます。顔の変形法としては、種々の表情表出時における顔の形の 3 次元計測結果を用いる手法 [1] を用います。また、全身の姿勢再現に関しては、(2) で得られた熱画像における人体の特徴点の 2 次元座標を、歌舞伎役者の 3 次元モデルを変形する 3 次元データに変換する式をあらかじめ学習的に求めて

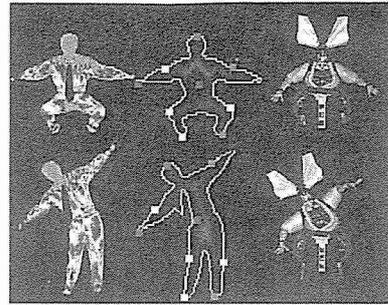


図3 熱画像からの人物の姿勢推定と歌舞伎役者モデルにおける再現

おき、再現時に利用します。図3右に歌舞伎役者モデルにおける動き再現例を示します。

以上述べたように、歌舞伎役者に変身しようとする人物の表情と全身の姿勢が歌舞伎役者モデルで再現されるので、これを図2に示したように、大型のプロジェクタースクリーンに歌舞伎の舞台背景とともに表示します。現在、(2) で述べた表情と姿勢推定処理は、パソコンで20フレーム/秒程度の速度を達成しています。また、(3) の表示処理は、10フレーム/秒程度であり、改善の余地を残しています。しかし、平成8年8月に米国 New Orleans で開催されたコンピュータグラフィックスの分野で世界最大の国際会議 SIGGRAPH 96 [3] の Digital Bayou (インタラクティブなデジタルデモ展示) で実演デモを行い、極めて安定な動作を実現するとともに、マスコミに注目を浴びる等、好評を博しました。

4 むすび

本稿では、仮想変身システムの一例としてバーチャル歌舞伎システムを紹介しましたが、本仮想変身システムでは、歌舞伎役者以外のモデルも作成しておくことにより、任意の姿に変身可能です。例えば、既に筆者らが開発した図4の怪獣モデルを用いれば、怪獣に変身できます。今後は、遠隔地の人々がそれぞれの役柄に姿を変えてシーンの中に登場し、インタラクティブに映画や演劇が作成できるバーチャルシアターの実現を目指します。また、本稿で述べた仮想変身技術は映像情報を始めとするマルチメディア通信の発展・普及に貢献できるものと期待されます。

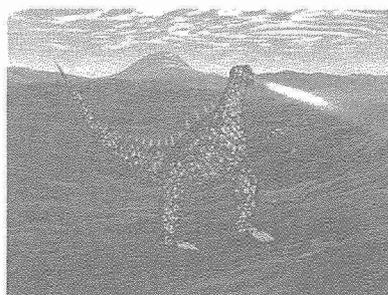


図4 怪獣モデル

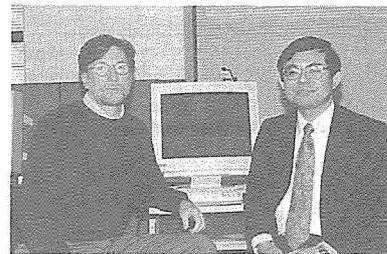
参考文献

- [1] 海老原、大谷：“あなたの表情を理解します—実時間表情検出・再現—”、ATR ジャーナル、No.22, pp.10-11 (1996冬)。
- [2] 岩澤、海老原、大谷：“熱画像からの実時間全身像の姿勢推定法の検討”、信学技報、IE96-51, pp.25-30 (1996.9)。
- [3] 間瀬：“SIGGRAPH96に参加して”、ATR ジャーナル、No.25, pp.12-13 (1996秋)。

マルチ言語マルチ話者音声合成システム CHATR —お気に入りの声をコンピュータで合成—

(株)ATR 音声翻訳通信研究所
第二研究室

ニック・キャンベル、樋口 宜男



当研究所では、ある話し手の声を一定量以上（例えば30分程度）、コンピュータに記憶して音声データベース化しておき、その中に多数ある同じ種類の音韻（子音や母音）の中から、出力したい文に応じて、音韻の並び方や声の高さ、音韻の長さなどの条件が良く適合し、しかも滑らかにつながる音韻を選び出し、それらの音声波形をつなぎ合わせるという方法を開発し、短期間でその人の声を合成することを可能にしました。

① はじめに

現在の日常生活ではコンピュータから出力される音声を耳にする機会が多くなりました。例えば、駅のアナウンスでは「ひかり123号が12番線に参ります。」と言ったり、高速道路の料金所では「料金は1200円です」という声が聞こえたりします。また、文章を読み上げる機能が付いているパソコンもあり、電子メールを音声で聞くこともできます。

これらの音声出力の方法は

- (1) 単語単位の音声を予めアナウンサーに読んでもらい、それらをつなぎあわせて出力する方式（通常、録音編集方式と言う）と、
- (2) 音韻（子音や母音）や仮名のような比較的小さな単位をつなぎあわせて出力する方式（通常、規則合成方式と言う）

の2つに大別されます。規則合成方式はどんな内容の文でも音声に変えられるという利点を持っていますが、どうしても合成音特有の響きがあるため、これまで高い品質を要求される公共サービスへの利用は録音編集方式のみに限られてきました。

一方、録音編集方式を用いた場合、利用する単語すべてを予め録音しておかなければならないために、新幹線に「のぞみ」が登場したときや高速道路に新しいインターチェンジができたときのように新しい単語を増やす必要が生じたときに、既にあるものと同じ声の質で新しい単語を録音することがかなり難しい問題になってきます。

この点を解決するために、声の高さや音韻の長さを変える処理を最小限に抑え、自然な音声の響きをそのまま残した合成音声出力するシステム CHATR（チャター：おしゃべりを意味する chat と ATR の合成語）を開発しました。このシステムは単に自然な音声の響きを残すというだけでなく、その人の個人性やそのときの雰囲気を実在に再現できるので、音声合成器のカスタマイズという点から大いに注目されます。

② 目的文に合った音声波形データの選択

これまでの音声合成システムでは、まず出力する文の内容に従って音韻の特徴を表わすスペクトル特徴を生成し、その後で抑揚を表わす基本周波数（ピッチ周波数とも言う）を制御する方式が採られていました。具体的には、いったん音声波形をスペクトル特徴パラメータに変換してから音声波形に戻したり、音声波形を1周期毎に分離してピッチ周波数を変えたりしてきました。このため、音声データにはいろいろな信号処理が施されることになり、元の音声を持っていた自然な響きが損なわれてしまい、合成音特有のこもった響きの原因となっていました。

この点を解決するために当研究所では、ある話し手の声を一定量以上（例えば30分程度）、コンピュータに記憶して音声データベース化しておき、その中に多数ある同じ種類の音韻の中から、出力したい文に応じて、音韻の並び方や声の高さ、音韻の長さなどの条件が良く適合し、しかも滑らかにつながる音韻を選び出し、それらの音声波形をつなぎ合わせるという方法を開発しました [1]。

音声単位の選び方は図1に示すように、目標とする音韻の持つべき特徴（ t_{i-1} , t_i , t_{i+1} ）とのずれを表わす C^t と、接続される音声単位（ u_{i-1} , u_i , u_{i+1} ）間での不連続さを表わす C^c との和を最小にするような音韻の波形データを音声データベースの中から選んでいきます。

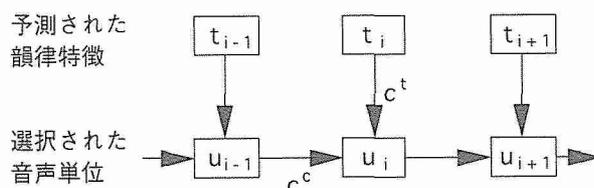


図1 音声単位の選択規準

C^t は目標とする音韻とのずれ

C^c は隣接単位間での不連続さの度合い

信号処理という観点から見ると音韻単位の波形データを相互に接続するだけですから、以前の方式に比べてずっと簡単になっていると考えられますが、最適な音声単位を音声データベースの中から探し出すという観点から見ると膨大な計算を行なっていることがわかります。例えば、30分の音声データには約15,000の音韻が含まれており、その中で特に生起頻度の高い/a/（ア）の音について見ると音声データベース中に約1,200個の波形データが含まれています。従って、その中から最適のものを選んでつなぎあわせるという作業がいかに大変かは容易に想像できると思います。この点を解決するために音声認識で用いられているViterbiアルゴリズムを利用して高速処理を行なっています。

このシステムで実際に音声を出力するときには、図2に示すように音声データベース中の波形データを逐次読み出していきます。この方式では音声波形を変形したり新たに作り出したりせず、音声データベースの中から最適なものを選び出すという処理をしていますので、目標に近い音韻が音声データベース中にない場合には隣り合った波形データ間で不連続感が生じたり、抑揚の不自然さが生じたりすることがあります。この点は音声データベースの規模を拡大すれば解決しますが、小規模でも効率良く多様な音韻を含むような音声データベースをどのようにして設計するかは今後の課題となっています。

③ マルチ言語・マルチ話者

CHATRで音声を出力する場合、ピッチ周波数やそれぞれの音韻の時間長を文の内容に従ってどう決めるかについては言語に対する依存性があります

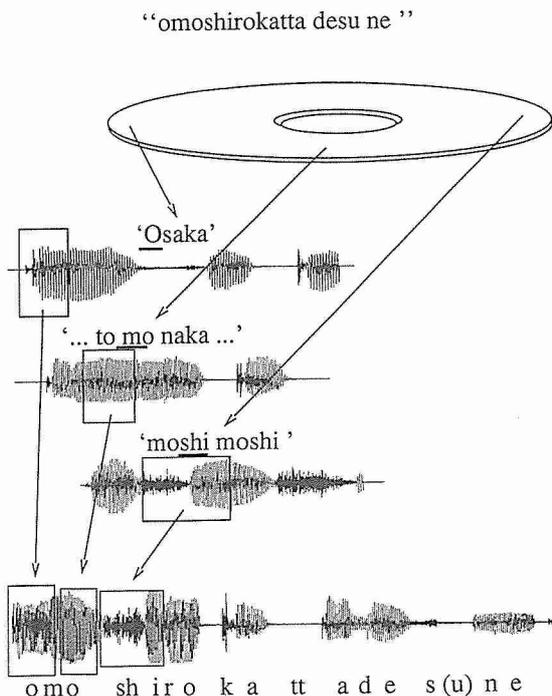


図2 CHATRによる合成音声の生成

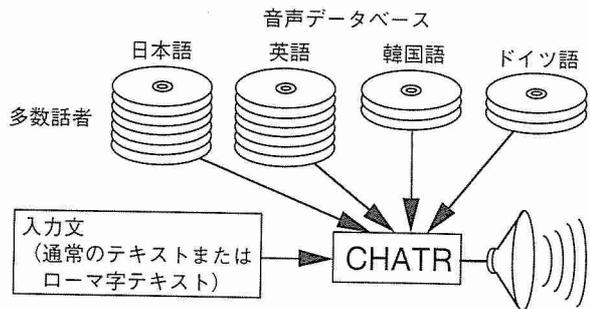


図3 マルチ言語化・マルチ話者化

が、それらが何らかの方法で決められたとすると、音声データベースの中から最適な波形データを選択する過程は言語や話し手に依存しません。このため、CHATRを多言語化することは比較的容易であり、図3に模式的に示したように、現在日本語・英語・韓国語・ドイツ語の4カ国語の合成が可能となっています。また、いろいろな話し手の声を利用することも極めて容易であり、例えば日本語については現在16名の声を出し分けられます。

新しい話し手の音声データベースを作る場合には、大きく分けて2つの方法があります。すなわち、
 (1) 録音されている音声データの内容を書き起こし、それによって音韻ラベルを付ける方法と、
 (2) 決められた文を話し手に読んでもらい、既存の音韻ラベルを自動的に割り振っていく方法です。

前者の場合、市販されているカセット・テープを利用することも可能であり、その一例として黒柳徹子さんが2つの小説を読んでいる新潮カセットブックの約1時間の音声を用いて黒柳さんの声に非常に近い合成音声が出せることを確認しています。（ただし、黒柳さんご本人からは声は非常に似ていますが、もう少し黒柳さんのアクセントを研究して欲しいというコメントをいただいております。）

一方、後者の方法を用いた場合、ほぼ自動処理が可能であり、録音から約3時間という極めて短い時間で合成音声を出力することができます。

④ おわりに

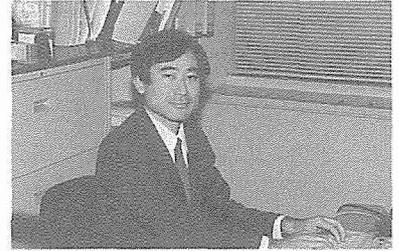
従来の音声合成システムと異なり、音声データベース中から、出力したい文に応じて、音韻の並び方や声の高さ、音韻の長さなどの条件が良く適合し、しかも滑らかにつながる音韻を選び出し、それらの音声波形をつなぎ合わせるという方法を用いて高品質の音声合成システムを実現しました。これらの合成音声はすべて<http://www.itl.atr.co.jp/chatr/>で聞いていただけますので、お試しいただければ幸いです。

参考文献

- [1] N. Campbell, A. Black: "CHATR: 自然音声波形接続型任意音声合成システム、"電子情報通信学会技術研究報告, SP 96 - 7 (1996. 5).

考えることと喋ること

(株)ATR 人間情報通信研究所
 第四研究室
 正木 信夫



私が「音声の研究をしています」と言うと、「コンピュータの音声認識や音声合成の研究ですか」と尋ねられることがよくあります。その時私は、「コンピュータがもっと自然に喋ることができるように、人間がどのようなメカニズムで喋っているかを研究しているのです」と答えることにしています。ここでは、人間が「喋る」という行為に関する研究の現状と展望について述べることにします。

① 音声言語の位置づけ

人間は「音声言語」「文字言語」という大きく分けて2種類のコミュニケーション手段を持ち、それらは、相補的な役割を果たしています。その中で「音声言語」は、身近にいる他者との間で、特に筆記具等の道具を必要とせず、情報伝達を行なえる簡便な手段であるばかりか、獲得時期が文字言語よりも早く、言語の概念を形成するためにも重要な役割を果たしています。さらに音声言語は、「音声生成系」（考えを「喋る」という行為により音声波として生成する）と、「音声理解系」（その音声波を「聞く」という行為を通して理解する）の2つの系を持ち、それぞれが音声波という媒体によって有機的に結合することにより、初めて成立するのです。ここでは、このような音声言語を構成する2つの系のうち、「音声生成系」に焦点を当て、何が問題とされ、それを解明するためにどのような研究が進められているかを紹介しましょう。

② 音声生成の成立過程

音声が出されるまでには、どのような過程を経るのでしょうか。おおよそ、図1に示すような4つの段階に分けられます。以下各段階での処理内容を解説します。

〔第1段階〕：仮に話者の「考え」は「愛の告白」としましょう。その目的の為に話者は脳内辞書から「わたし」「あなた」「好き」という「単語」を引きだします。さらに、これを適切な語順に並べ、日本語に必要な助詞と助動詞を付け加えるという作業が行なわれます。これにより「わたしはあなたが好きです」という「文」が成立します。

〔第2段階〕：この文を発話するためにはもう一度「単語」のレベルに立ち戻り、発話の実行に必要な情報を得るためにさらに細かい構造に分解していく必要があります。例えば、文頭の単語「わたし」はまず3つの「拍」のレベルに分解されます。「拍」という概念は、基本的には仮名1文字の持つ音の長さに相当する単位です（俳句や短歌の「字数」と言われるものは実際はこの「拍数」です）。そしてさらにこの「拍」は「音素」に分解されます。「音素」は発話に用いられる「語音」の最小単位であり、「子音」や「母音」に相当するものです。

〔第3段階〕：次に、各「音素」を発音するために、舌・唇等の発話器官の運動（調音運動）を司る筋肉の収縮の程度やタイミング（運動指令）が計算されます。

〔第4段階〕：最後に各調音器官への運動指令が実行されます。その際、調音運動に伴う感覚や、生成された音声の聴覚的なフィードバック情報も、適切な調音の達成のために利用されます。

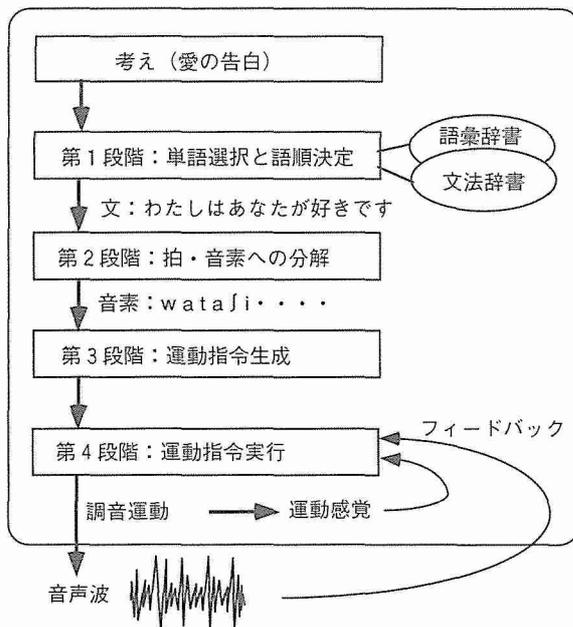


図1 音声生成過程

③ 運動指令生成メカニズムの研究

さて、このような各段階で、具体的にどのような

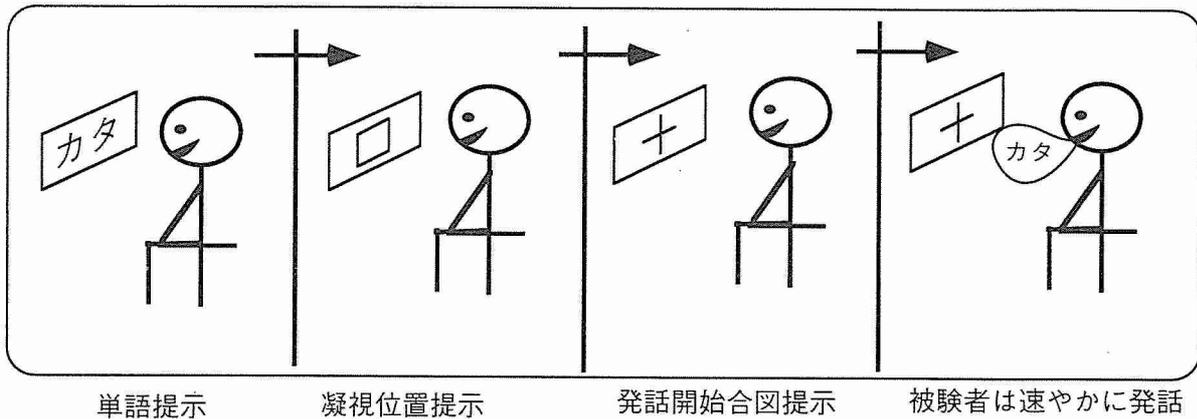


図2 発話潜時測定の方法

処理がいかなるメカニズムで行なわれているのかを解明するために、さまざまな問題が提起され研究されています。例えば、「第1段階の処理に必要な脳内辞書にはどのぐらいの単語が格納されているのか? [1]」、また、「第4段階の処理に関連して、音声波の聴覚へのフィードバックが音声生成を正常に遂行するためにどのような役割を果しているのか? [2]」等です。これらの問題に加えて、音声生成の研究のなかの最重要課題の一つは、第2・3段階の処理に関連した、「発話のために単語がどのように分解され、運動指令の連鎖を構築していくのか?」という問題です。私自信は、この点に興味を持って研究しています。

例えば、「カ」「カタ」「カタカ」「カタカタ」というような単純な音節系列からなるいろいろな拍数の「単語」を発話することを考えましょう。20人程の被験者にこれらの単語を発話してもらいました。ただし、あらかじめ単語は提示しておき、その後で発話開始の合図が表示されたらできるだけ早く、正確に発話するよう指示しました(図2)。あらかじめ単語を提示し、合図が出てから語頭の「か」が発話されるまでの時間(発話潜時)を測定することにより、「読み」の処理にかかる時間を除外し、発話のために必要な運動指令生成に関わる処理時間の変動だけが測定できます。図3は、発話潜時の平均値を単語の拍数毎に示したものです。発話潜時は発話する語の拍数の増加に伴って増加しています。これは、発話に必要な運動が多くなったり複雑になったりするほど、運動指令生成に時間がかかるためと考えられます [3]。

現在、運動指令生成にかかる時間の増加をもたらす要素は何かを検討しています。このような研究から得られるデータは、運動指令の生成メカニズムの解明に必要な基礎資料となります。

④ 音声生成研究における脳研究の必要性

音声は、脳のさまざまな部位における複雑な処理の結果として生成されます。その「部位」と「処理内容」の関係を調べる研究手法は、「言語病理学」として確立されました。それは、脳内出血や脳血栓等を原因として脳に損傷を受けた患者の損傷部位をMRI等で確認し、患者の言語活動にみられる異常との対応を考察するものです。今後、脳磁界測定を始めとする脳内活性部位をモニターする手法を用いることにより、これまで言語病理学の中で蓄積された知見が健常者でも確認されていくでしょう。その過程で、音声生成系の各段階の処理がどこでどのように行なわれているのかが明らかになると期待されます。したがって、今後音声生成のメカニズムを明らかにするために、脳研究が重要な役割を果していくことは間違いのないと思われます。

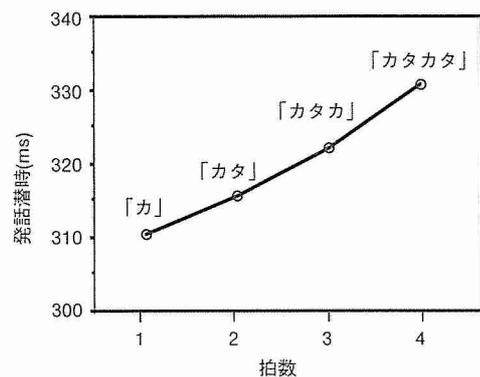


図3 単語の拍数と発話潜時の関係

参考文献

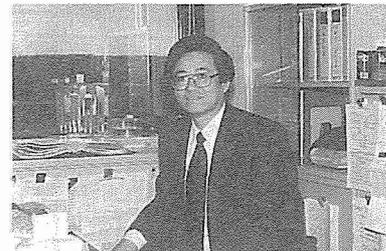
- [1] 近藤, 天野: 単語親密度を利用した心的辞書中の文字/音声単語数の推定. 日本認知科学会第13回大会論文集, 40-41, 1996.
- [2] 河原: 文章朗読時における聴覚フィードバックの影響について. 日本音響学会平成7年度春期研究発表会論文集, 359-360, 1995.
- [3] Masaki, Honda: Control of speech command generation for Japanese words - Estimation from reaction time measurement -, Proceedings of 1st ESCATutorial and Research Workshop on Speech Production Modeling, 105-108, 1996.

ソフトウェアアンテナ

— 知能と変身の術を備えたアンテナ —

(株) ATR 環境適応通信研究所
第三研究室

唐沢 好男



新しいアンテナのコンセプトを紹介します。このアンテナをソフトウェアアンテナと呼び、環境（電波環境、通信環境）の変化に追従して、アンテナシステムがソフト（アルゴリズム）・ハード（論理結線回路）の両面でダイナミックに知的な変身を遂げて行くイメージを示します。ソフトウェアアンテナの実現にはアンテナ技術、デジタル信号処理技術、マイクロプロセッサ技術等幅広い技術を背景とした研究が必要です。マルチメディア移動通信・モバイルコンピューティング時代に生きるアンテナであり、力を入れて研究に取り組んで行きます。

固いアンテナから柔らかいアンテナへ

インターネットで世界中の情報源に自由にアクセスしたり、音声や映像で会話を楽しむことができるようになってきました。また、ISDNに代表される高度な情報通信網も着実な広がりを見せています。このように便利な環境をデスクの上に留めず、いつでも・どこでもその恩恵が享受できるよう、屋外や屋内のあらゆる環境にサービスを提供する移動通信の発達も目覚ましいものがあります。

移動通信では、利用する人の動きに制約を与えないよう、情報の運搬には電波が用いられます。規則正しく振動する波をわずかに乱して乗せられる情報量には限りがあります。特に画像の伝送や大規模なデータベースへのアクセスをたくさんの人が同時に行えば、瞬時の情報量が膨大になり、電波（周波数）はいくらあっても足りません。そこで、限られた電波を有効に使い、欲しい情報だけを正確にとり出す技術がこれからますます重要になります。情報の出入り口であるアンテナに求められるものは、雑音や干渉波の中に埋もれた信号を的確にキャッチする能力（＝ソフト）です。そしてその能力は、様々なことができる「技能」と、それを、複雑な電波環境に適応して発揮させる「知能」よりなります。

本稿では、このような新しい時代（＝マルチメディア移動通信）に求められる知的変身の術を備えたアンテナ「ソフトウェアアンテナ」(1)を紹介いたします。なお、このアンテナ概念はまだ定着しているものではなく、まさにこれから研究として立ち上げたいと思っている概念です。

ソフトウェアアンテナとは？

ATRでは、知能を持ったアンテナとして「ディ

ジタルビームフォーミング（DBF）アンテナ」の研究を行ってきました。DBFアンテナはアンテナに求められる大部分の機能をデジタル信号処理によって実現するアンテナです。たくさんの小さなアンテナ（素子アンテナ）からなり、各素子の情報をそのままAD変換器を介してデジタル信号処理部に取り込む。計算処理によって、一つのアンテナで様々な方向から到来する電波を同時に受信したり、環境の変化に応じてアンテナの特性を制御することができるので、複雑な電波環境を持つ移動通信分野のアンテナに適しています。

ところで、これまでのDBFアンテナはアンテナの機能をデジタル信号処理によって実現するものの、ロジックのハードウェア（論理結線回路）は、少なくともその動作中に変わることはありません。その場合でも、そこには適応的に働くアルゴリズムが実装されており、能力の範囲内で精一杯環境に適応します。しかし、ユーザの移動に伴って、時々刻々変化する電波環境は、あまりにも複雑でダイナミックです。一つのアルゴリズムでは、ある環境で強力な干渉波除去能力を発揮するとしても、環境が変わって劣化要因まで変わるともう対応できません。例えば、同一の源からの散乱波でも、到来時間のばらつきが小さいときは散乱波を全て取り込み、位相を合わせて合成すればよいですが、遅延の差が大きい散乱波をこのルールで集めると伝送波形の歪が生じて、符号間干渉による誤りが発生します。この場合には、遅延の大きい波を除去したり、歪んだ波をもとの形にもどす機能が必要です。これまでに考えられているどんなアルゴリズムでも環境に対して得意不得意があり、万能なものはありません。その場合、必要と思われるアルゴリズムを全て実装してお

いて、切り替えて使う手がありますが、処理部のハードウェア規模に際限が無くなり、結果として信号処理のスピードも落ちてきます。

その様なとき、自らの論理回路を即時に書き換えることができるリコンフィギュラブル・ロジックがあれば、環境の変化に対して最適なアルゴリズムを次々と登場させることができます。リレーのランナーがスムーズにボタンを渡して速やかに入れ替わるように華麗に変身して行く。これは、環境（電波環境、通信環境）の変化に適応して、アンテナシステムがその動作中にソフト（アルゴリズム）・ハード（論理結線回路）の両面でダイナミックに自己変革（＝変身）を遂げて行くイメージです。個々のアルゴリズムが有する制約を越えた信号処理（＝アルゴリズムダイバーシチ）によって環境に適応し、カメレオンのように（カメレオンは見かけだけであるがもっと本質的に）自分自身を変えて行くアンテナ、これを「ソフトウェアアンテナ」と呼んでいます [1]。

図は移動通信の基地局アンテナを想定したソフトウェアアンテナの構成例です。(a) 環境の変化を正確に認識して瞬時に最適なアルゴリズムを判断する信号処理部、(b) 様々な環境を想定してあらかじめ用意されている複数のアルゴリズムソフト群、(c) 選ばれたアルゴリズムを瞬時に実装して機能を発現するリコンフィギュラブルロジック、が基本要素です。上述の DBF アンテナは(b)と(c)の一部に含まれています。(a)は環境に対する目となり耳となる部分、またその結果から環境の構造を正確に認識する機能が求められ、アンテナの知能を担う部分です。(b)はスペシャリスト（＝アルゴリズム）を投入するソフトをプールしておくところ、環境の変化に応じて求められる異質のアルゴリズムを次々と送りだ

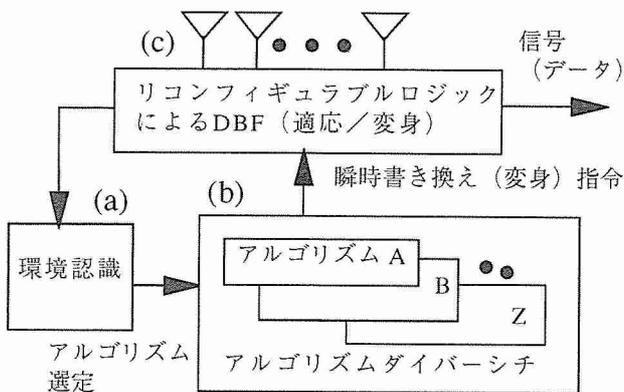


図 ソフトウェアアンテナの基本構成
(移動通信基地局受信アンテナを想定)

し、(c)に流れている情報を乱すことなく、華麗にハードを変身させて行く。

新しい概念としてアルゴリズムダイバーシチを述べましたが、認識した結果にしたがって単に切り替えるだけでなく、アルゴリズム中の要素間の相互作用によって新たな機能を持つアルゴリズムが生まれるような、アルゴリズムの進化が始まる仕掛けができればおもしろいと思っています。

Ⅳ 機は熟してきた

上記ソフトウェアアンテナの構成の中で、唐突に「自らの論理構成を即時に書き換えることができるリコンフィギュラブル・ロジックがあれば」と書きました。ATRがDBFアンテナの信号処理部として開発したFPGA（書換可能なゲートアレー）ボードにはまだその機能はありません。書き換えが動作中にはできないからです。運用開始時あるいは、タイミングをみて論理回路を書き換えるのはインシステムプログラミングと呼ばれていますが、この段階です。リコンフィギュラブル・ロジックの開発動向については、文献 [2] に詳細な解説があります。そこでは、FPGA の本来の機能を活かす使い方として「書き換え可能」な点に着目し、書き換え時間を大幅に短縮した（100 μ sec という数値が挙げられている）リコンフィギュラブル・ロジック対応FPGAの時代が近づいたことを告げています。夢（＝ソフトウェアアンテナ）を実現する手段もまた、生まれつつあります。

Ⅴ マルチメディア社会を一層豊かなものにする

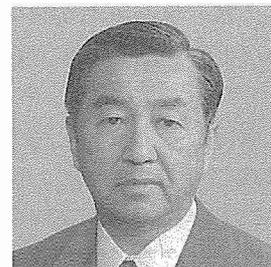
新しい時代のアンテナのコンセプトを紹介しました。このアンテナをソフトウェアアンテナと呼び、環境（電波環境、通信環境）の変化に追随して、アンテナシステムがソフト（アルゴリズム）・ハード（論理結線回路）の両面でダイナミックに知的な変身を遂げて行くイメージを示しました。ソフトウェアアンテナの実現にはアンテナ技術、デジタル信号処理技術、マイクロプロセッサ技術等様々な技術を背景とした研究が必要です。マルチメディア移動通信・モバイルコンピューティング時代に生きるアンテナであり、力を入れて研究に取り組んでいきます。

参考文献

- [1] 唐沢、関口、井上：“ソフトウェアアンテナ”、信学技報、A.P96-103, 1996.
- [2] “プログラムデバイス第3の波：FPGA” 日経マイクロデバイス、July 1996.

ゼロからのスタート（その1）

(株)国際電気通信基礎技術研究所
顧問 葉原 耕平



このシリーズのいきさつ

ATR 発足10年の機会に「10周年記念特集」をお手元にお届けした。その編集委員会和佐野委員長から「あれは言わば正史で、周辺の逸話や裏話など残しておいた方がいいことが沢山あるのではないか」という、忘れていくことを見透かされたような話がありました。

それもそうだ。私自身、1993年春のジャーナルの巻頭言で『いまは、まさに初心に戻る絶好の時である』と自戒の意味を込めて書いたではないか。しかし、初心とは何なのか。自分は昔のことを知っているからいいようなものの、昔を知らない人に初心に返れと言ってもそれは無理というものだ。『人にやさしく』を標榜しながら自分勝手な論理矛盾ではないか。という単純なことに気がついた。

時代が移り人が変わり組織が変わる中で、あまりにも過去にとらわれると、進歩・変革の妨げになりかねない。しかし、それはその時々の方々の判断と良識に委ねるべき筋合いのものではなからうか。それに自分自身の記憶も段々と薄れることも事実だ。たしかに今を置いてチャンスはないかも知れない。一方、ありていに言えば裏話というのは書けないからこそ裏話なのではないか。さあ困った。とまあ、研究者くずれの悪い癖で、御託を並べて考えあぐねた末、「まあいいか。恥さらしかもしれないが、やってみるか。」ということでOKしたのが、このシリーズです。はてさてどういうことに相なりますか。読み物として面白いかどうかは別として、曲がりなりにも悩んだ挙句の筆者の心境をお察しの上、お読み頂ければ幸いです。

① 基礎研究とは？

ATRは基礎研究を標榜して発足しました。大上段に振りかぶるつもりはありませんが、当然、それでは「基礎研究とは？」ということ議論したくなるのは人情です。導入がひどく固くて恐縮ですが、この話題にちょっとだけ触れておきたいと思います。

はやい話が基礎研究の「定義」。私はそれは不可能に近いと思っています。何故なら基礎研究は自由闊達な人間活動（これは単に頭脳活動だけではない）の中から新しい知見を得ることが何よりも大切なことで、一旦「基礎研究とはかくかくしかじかで」と定義をした途端、ある種の枠をはめることとなり、それは研究を窮屈にすることに繋がり兼ねないし、研究者を萎縮させる危険さははらみまます。ある人は「あえて言えば」という注釈付きで「基礎研究とはそれまでの非常識を常識の変えることだ」と言ったのが私には印象的でした。これは逆に「常識を非常識に変える」ことにも繋がります。コペルニクスの例を引くまでもなく、常識的定義をすること自身矛盾かもしれません。私もあえて言わせていただければ（論理矛盾ではありますが）「基礎研究とは一義には定義できないものである」と定義します。あるいは「研究開発活動の中で実用化、商品化など特定の目的を持つものに属さない一切の活動」が対象であるといってもいいかもしれません。

くどくどと基礎研究の定義論をしたのは、このような性格のものですから、「かくかくしかじかの」というマネージの常道もまた考えにくいということと裏腹だからです。だから、基礎研究のマネージは各人各様で、そこにはパーソナリティが強くにじみ出ます。つまり「基礎研究はパーソナリティによって成り立つ」と言えると思います。そういう意味で「基礎研究は人なり」とは言えると考えます。これについては、また触れる機会がありましょう。

② 原則自由、例外規制

幸いATRは世界に類を見ない全く新しい構想でスタートしました。私はひそかに（理想論ではありますが）「すべてのしがらみや制約のない研究とその組織、運営」に強い関心を持ちました。いわば全くの白紙の上に絵を描くのに似ています。その壮大な社会実験に挑戦して見ようと考えました。なぜなら、これは今世紀中二度と起こらないかもしれない程のことで、今を措いて機会はないと思ったからでもあります。そして、その指導理念を「原則自由、例外規制」に据えました。それが具体的にどういうことであったかはこの

シリーズの中でいくつかの事例を述べてみたいと思います。

その前に誤解を招くと困りますので、一つだけ大切なことを述べておきますと、自由と野放図とは峻別しなければならぬということ、自由には総て自己責任が伴うということ、つまりツケの持って行き場所は自分以外にはない、というつい忘れられがちですが自明のことが大前提だということです。

③ 外国出張のやりとり

以下はATR発足直後の、(それまで大組織でまず上司の意向を尊重しようというカルチャーに慣れてきた)担当者Aと私のやりとり、

A: (多少恐る恐る) あの一、調査で外国出張したいのですが、どうでしょうか。

私: 自分で考えてご覧。調査に行けば多分いろいろ収穫があるだろう。けど今は立ち上げの時期だから、君がいない間研究計画の作業が遅れるかも知れないし、他の人に負担がかかるかも知れない。それにお金だってかかる。それらを総合的に考えてトータルでプラスだと思うなら行けばいいんじゃないの。私に相談に来るのはいつでも歓迎だけど、その前にまずは自分の責任で考えてご覧。



(ややあって)、

A: やはり、調査は今の時期が大切だと思うので、行ってきます。

私: いいよ。よく準備してね。だけど、万一結果が思わしくなくても、それは仕方がないことだよ。分からないから、調べに行くんだから。分かってないことが多ければ多い程、お土産も多いかも知れないけど、反対の場合もあるからね。

(帰着後)

A: お陰で、大収穫でした。ところで、報告書要りますか。

私: 君はどう思う？

A: ああ、要りますね。何書きましょうか。

私: 考えてご覧。

A: 分かりました。

私: ちょっとだけヒントをいうと、報告書の目的は何か、よく考えてね。添付資料はあってもいいけど、骨子は1枚に纏めるように。

蛇足ですが、何枚にもわたって書くのはむしろ簡単で、エッセンスを1枚に凝縮するにはよほど要領よくポイントを整理しなければなりません。その過程で本質が見えてくる、というのが私の持論でもあります。ついでに言わせていただければ、折角の資料や報告をなるべく読んで欲しくないと思ったら長々と書くことです。読む立場で考えればすぐ分かると思います。

④ 規定類の整備

こんなことを契機に、各種の様式の整備がスタートしました。何しろわずかな人員でしたので、手っ取り早いのは各社から雛型を取り寄せることでした。そして改めて見てみますと、例えば外国出張報告について言えばNTTの様式は長年の経験が折り込まれていてよくできていました。ATRの特殊性を若干折り込む程度で、最初の様式ができました(もちろんその後逐次改善されています)。これは一例に過ぎず、ご支援いただいた各社には改めてお礼申し上げます。

このエピソードはほんの一例で、ATRでは何事によらず、このようにして必要なものから、自分たちの問題として、そして多くは切羽詰まって作り上げてきました。まだ伝票が一切ない最も最初の時期、最初の購入伝票はどうやって買ったか。これは私も聞いていません。ATRの規定類(これらは私の所管ではありませんが)の多くはそのような苦労の所産です。ここで個人名をあげるのはいささか抵抗がありますが、初代の伊藤人事課長は、日々の業務だけでも人一倍多忙の中、合間をぬって多くの規定を作成して去って行きました。最後に「あれがやっとできてほっとしました」と言い残して。

ATRも10年余を経た今、いろいろと見直すのにいい時期かも知れません。後から組織に入った人には、ぜひ新鮮な目でそれをお願いいたします。その際、言わずもがなのことではありますが、心に留めていただきたいのは常に自分達で自分達のために必要なものを作るのだという意識だと思えます。



(つづく)

「大和の国の物語」

— けいはんなの丘に寄せて —

NTT 信越ネットワーク技術センタ 所長
(元国際電気通信基礎技術研究所企画部 次長)

高橋 保



犬養孝先生著「万葉の人びと」の舞台である飛鳥や藤原京への探訪を夢んでいた小生に、ATR への転勤の命が下ったのは平成4年2月のことです。支店経営をやってみたい、という希望とはおよそかけ離れた研究所への出向でしたが、NTT 研究所でダイヤルQ₂の開発を手がけたこともあり、仕事には不安はありません。ATR の名声はかねてより聞こえていました。

爾来約5年もの長き間ATRに勤務しました。我が会社人生20年の凡そ四半期、ATR十年の歴史のうち約半分です。

この間、葉原前副社長（現顧問）には筆舌に尽くしがたいほどお世話になりました。第一級の見識をいつも問われ、時にはずいぶんおしかりを頂戴しました。折にふれ適切なアドバイスをいただき、感銘を受けたものです。常に高い視点、基本的な考え方をもと諭され、はたしてどこまで応えることができたのか忸怩たる思いです。在任期間中に3つの会社の誕生と4つの会社の研究終了に関与し多用を極めました。「研究は人」といいます。ATRには素晴らしい研究者が一堂に会しています。出資企業からの研究員の受け入れに当たっては必ず候補者と面談し、レベルに達しない場合は書類選考でお断りもしました。小生も及ばずながら研究スタッフとして、いい仕事をさせていただきました。

今振り返ってもATRは研究者にとって天国です。しかし、研究者を支える縁の下の力持ちの人達がいます。また、いくら立派な論文でも具現化しなければただの紙切れです。研究成果を世の中に普及する、それが使命なのだと言者は肝に銘じてほしい。

さて、万葉の人びとへの探訪を紹介しましょう。法隆寺の秘仏公開や、正倉院展、鑑真和上像の公開などがあり、一通り見ることにしました。また界隈の主要な古道は必ず踏破することに決めました。

「てくてくマップ」いう優れたものが近鉄の主要駅にあり、ずいぶん便利。南は吉野、御手洗溪谷、天川神社、葛城山まで踏破しました。熊楠がかつて粘菌を求めた奥吉野や熊野へはとうとういけませんでした。

ちなみに2千メートル級の山は関西以西にはかの石槌山程度しかないと、かつて企画部0社員が悔しがっていたので、「苗場山とか戸隠山とかの象徴的な山はあげられないが、鳥甲とか、鳥帽子とか、赤石などであれば好きなのを持っていけ」といってあるのだが未だに取りに来てないようです。

後半の2年では琵琶湖周辺や、京都府北部等にも足を伸ばすことができました。ある年に伊勢神宮に2年詣りを敢行しましたが、寒いのと二見が浦の初日の出が夫婦岩側ではなく反対の山際から出るのを悟っただけでした。西は暗峠、郡山の矢田山から南生駒へ、生駒山縦走、北は北山辺の道から奈良公園、秋篠寺まで、更に東は伊賀、笠置、柳生等といった具合で、いわゆる大和盆地は隅々まで歩きました。大和郡山から東西南北どこへ行っても、やがてはたたなづく青垣に突き当たるのです。

ATRの皆さんもぜひ土日等を利用し、お出かけになってはいかがでしょうか。漫画家の里中満智子さんは「東京には江戸以降の歴史だけだけど、近畿には古代の歴史が至る所にあるの」といっておられ、住んで実感しました。耳成、香具山、二上山、生駒山、三輪山等皆懐かしい。仕事も散策も充実した本当に夢のような大和の国でした。

さて、かような夢から醒めてNTTに復帰してみれば、マルチメディア時代の到来が現実となっています。マルチメディア技術者の育成が急務です。昨年暮れには経営形態の再編成案が具体化し、大いなる激動の年の幕開けです。大きな潮流のうねりの中、我々も本物の改革を迫られています。ATRで培った進取の精神と多くの経験を活かせれば幸いです。

Palmtop Display for Dextrous Manipulation

－ Palmtop Display を用いた力覚提示装置 －

ATR 通信システム研究所では小型ディスプレイを用いた仮想世界オブジェクトの操作デバイス PDDM (Palmtop Display for Dextrous Manipulation) を開発し報道発表しました。PDDM は、特に臨場感通信会議システムでのオブジェクト組み立てや変形作業のための操作を念頭において設計したものです。従来は個別であった視覚情報表示と操作入力、さらに力覚提示機能を単一のデバイスで実現し、それらの異なる機能モード間の変更操作に伴う作業の流れの中断を意識させない手法を採用しました。

試作機は小型可搬型ディスプレイに力覚提示装置を接続した構造となり、ユーザーはディスプレイ部を把持して情報と操作の入出力を行います。仮想空間内で PDDM を用いる事により、ユーザーは仮想物体の重量感や他の物体との衝突感覚を感じながら物体を操作でき、操作における直感性と操作効率向上に効果があり、主要全国紙および地方紙に大きく掲載されました。

個性のある音声を生成する多言語音声合成システム CHATR (チャター) を開発

－ 例えば黒柳徹子さんの声でコンピュータが話す －

コンピュータによって人間に近い声を出力する音声合成システムは、従来はきまった声しか出力できませんでした。ATR 音声翻訳通信研究所は、ある人の声のサンプルが数十分程度あればその人に近い声を音声をつくりだせるシステム (CHATR) の開発に成功しました。このシステムは日本語のみならず他の言語も合成が可能で、音声翻訳機など幅広い利用が期待されます。CHATR システムは NHK 総合など TV ニュースで放映されると共に主要全国紙及び地方紙に大きく掲載されました。

CHATR システムの詳細は本 ATR Journal 26 号の研究動向で紹介しています。

視覚心理学が可能にした自然な立体表示法

－ 立体映像の立体感を自然なものにするバーチャルフレーム方式を開発 －

これまでの立体表示においては「画枠歪み」と呼ばれる画面の端で立体感が失われるという現象がありました。ATR 人間情報通信研究所は、立体画像の自然さを損ねている画枠歪みの原因を視覚心理学的に究明するとともに、その解決法としてバーチャルフレーム方式を開発しました。

この新方式は両眼視差を利用するすべての立体表示に簡単に適用でき、従来、立体映像ソフトの制作時に課せられていた制約を無くすることができるなど、応用面での展開が期待され、主要全国紙および地方紙に大きく掲載されました。

特殊な結晶面を利用した高性能面発光レーザー

－ 並列・大容量光通信の実現に向けて －

ATR 環境適応通信研究所と東京工業大学精密工学研究所は共同で、簡単な構造で、安定な偏波モードが得られ、かつ低消費電力な面発光レーザーの開発に世界で初めて成功しました。

面発光レーザーは、平面上に多数の素子を配列することが容易なため、並列・大容量光通信に適した光源として期待されています。しかし、従来の面発光レーザーでは、レーザー光の電界振動方向 (偏波モード) が不安定なため、光の干渉性を利用した大容量伝送方式であるコヒーレント通信等への応用が困難でした。この問題を解決するために幾つかの構造が提案されていますが、いずれも複雑な加工プロセスを必要とするため、製造コストが高くなるという欠点がありました。また、多数の素子を駆動するには消費電力を低減する必要があり、より小さな駆動電流でレーザー発振を実現することが求められていました。

今回開発に成功した面発光レーザーは、光学異方性や大きな光学利得を有する GaAs の特殊な結晶面を基板として用いることで、複雑な加工プロセスを追加することなく、これらの課題を克服したものです。GaAs (311)A 面基板を用いた試作面発光レーザーでは、基板の結晶軸で決まる一定の方向に光電界がきれいに揃ったレーザー光が確認でき、またレーザー駆動に要するしきい値電流密度も従来の約 1/4 に低減することができました。

この面発光レーザーの開発は、ATR 光電波通信研究所で長年培ってきた GaAs の特殊な結晶面に関する物性や結晶成長技術の基礎研究をベースにしたもので、今後のデバイス開発の大きな手がかりになるとともに、並列・大容量光通信の実現にも貢献することが期待され、主要全国紙および地方紙に大きく掲載されました。

第9回 研究発表会を開催

昨年11月7、8日の両日に第9回研究発表会を開催し、ATRグループ各研究所の最新の研究動向および成果を総括講演、テーマ発表、ポスターセッションにより紹介しました。当日は多数のお客様のご来場を賜り意見交換等も活発に行われ、ご好評を博しました。



酒井副社長総括講演

ATR研究発表会は、最新の研究成果および動向をご紹介する場ならびに研究成果のご活用をお考えいただく場として毎年開催いたしております。

今年から新たに環境適応通信研究所が加わり、ATRにおける新たな研究体制が確立して初めての研究発表会となりましたが、日頃からお世話になっている企業および大学等の関係機関から、2日間にわたって延約1300名の方々のご来場をいただき会場内は終日賑わいました。

今回は、総括講演6件、テーマ発表6件、ポスターセッション68件とこれまでの最大級の規模となりました。また、今回は初めて2日間にわたって開催し、特に1日目は大学等の研究機関に広く公開いたしました。当日は朝早くからお客様が来訪され、午後のテーマ発表開始時にはほぼ全会場にお客様があふれ、熱気が2日間にわたり続きました。



大橋先生の講演



ポスターセッション見学

この研究発表会の内容をご案内するため同時に開設いたしましたWWW上の研究発表会のホームページに対しても1000名を超えるの方々からアクセスをいただきました。また、皆様からいただいたアンケートでは概ね良かったとのご意見をいただき、特に2日間開催につきましては大半の方々から良かったとの評価をいただきました。今後はアンケートの結果も考慮して、より充実した内容になるよう検討していきたいと考えております。

ATR 科学技術セミナーの開催状況

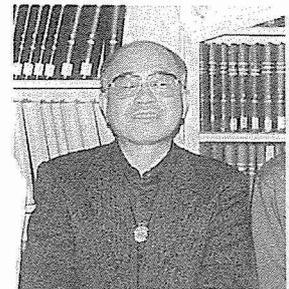
ATRでは、先端分野の第一線の研究者の方をお招きし、その分野の最新の動向等をご講演いただく ATR 科学技術セミナーを開催しています。毎回多数の参加をいただき、講師との活発な討論も行われ、それぞれの先端分野の研究情報交流の場としての役割を果たしています。

第46回 96年9月20日（人間情報科学 第37回）

丸山欣哉（宮城学院女子大学）

マルチ・メディアと異種感覚 -協応・競合・統合-

日本の実験心理学の分野で、異種感覚間の相互作用について、早い時期から問題意識を持ち、取り組んでこられた、宮城学院女子大学の丸山欣哉教授をお迎えしてご講演をいただきました。異種感覚の関係をどのような観点から整理すべきかについて、研究の歴史を振り返りながら、様々な実験現象を交えてご紹介いただき、今日のマルチ・メディア環境に対する重要な指針について改めて考え直す機会を持つことができました。



丸山欣哉 教授

第47回 96年10月7日（人間情報科学 第38回）

David Perrett（セントアンドリュース大学）

コンピュータグラフィックスによって顔の心理的なイメージを探る

英国セントアンドリュース大学のD.ペレット博士に、個人性、年齢、性別、魅力など種々の次元で認知される顔の印象という感性的イメージが、顔の持つどのような物理的な要因によって形成されているのかを探ろうとする研究について講演していただきました。最新のコンピュータグラフィックス技術を用いて顔の3次元的な形態やその特徴を操作した合成顔画像を作成し、これを人間に評定させて行なった様々な心理実験からの知見をご紹介いただきました。



David Perrett 博士

第48回 96年11月13日（人間情報科学 第39回）

Ryszard S.Michalski（ジョージメイソン大学）

機械学習の新しい方向

コンピュータ科学における機械学習分野の創設者のひとりである、ジョージメイソン大学のR.S.ミカルスキー教授に、機械学習の新しい方向についてご講演いただきました。学習の過程を統一的かつ理論的に扱うことができる推論理論の枠組みから、多重推論や計算戦略を用いるマルチ戦略学習、そして、推定的帰納法などについて、世界経済やコンピュータ・ビジョンを例にご紹介いただきました。



Ryszard S.Michalski 教授

【お問い合わせ先】

ATR 科学技術セミナー事務局

人間情報科学担当 FAX (0774) 95 1008, E-mail mieko@hip.atr.co.jp

福寿園CHA研究センター訪問

学研都市・平城相楽地区にある異業種企業13社から構成された研究所村「ハイタッチ・リサーチパーク」。今回は、その一角にある福寿園CHA研究センターにお伺いし、若原義道センター長に案内をしていただきました。

★設立の目的は。

当センターでは、「茶」をCHAととらえて、C：Culture（文化）、H：Health（健康）、A：Amenity（快適）の創造を目指し、また、そのために「人と人」「人と文化」「文化と文化」の出会いのためのコミュニケーション・ツールである茶により親しんでもらうように種々の研究活動を行っています。

★学研都市に進出したのは。

学研都市、ハイタッチ・リサーチパークに進出したのは、異業種交流を通じて情報を交換できるという立地条件の良さを考えたからです。茶園で茶の葉を育てて製茶するだけでなく、いろいろな方と接し、情報を交換することから商品開発に結びつくヒントをもらったりできます。茶というのは、生活に結びついており非常に身近なものですから、多くの方にここへ来ていただきたいと思っています。

★以前、異業種の企業との共同研究をされたと伺っていますが。

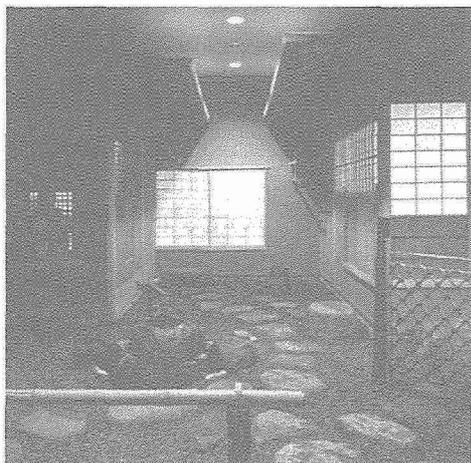
オープンして1年目にYAMANAKAアート研究センター、クロイハイタッチ研究所、二条丸八セレモニー研究所と自動給茶器共同プロジェクトを進めました。試作品はできたのですが、まだ商品化されていません。投資とメンテナンスという点を考えると商品化はなかなか難しいのです。

★見学者の数はどれくらいですか。

多くて年間1万人くらいです。学研都市に研修等で来られた方が、帰りに立ち寄られたり、主婦の方たちの見学が主流です。

★センター内の施設等について。

センター内では次の5つの分野に分けてCHAを追求しています。



*茶と農芸

—世界の茶樹を通じて、よりよい茶づくりの研究—

研究茶園では日本茶だけでなく、中国茶や紅茶など約80もの茶樹、品種を集めて栽培しています。隣にある温室では、土の中にパイプを埋め込み肥料が流れ出るようになっています。このような養液栽培の一種の隙耕栽培で周年摘採を研究しています。また、緑茶・半発酵茶・紅茶等の全ての製茶実験を行う実験室や、茶壺での貯蔵方法を研究している茶壺庫があります。

***茶と化学 ー高品質でより新しい茶を研究ー**

化学実験室、官能審査室、試作実験室と難しい研究をしているように感じられる部屋では、お茶の成分分析や組織培養等の研究、また品質検査や茶の新製品の試作研究、新しい商品開発を行っています。茶のふりかけや茶のガムなど様々なものに茶を取り入れています。

***茶と文化 ー茶の心・日本の心を次世代へ継承ー**

日本の茶研究室には、東大寺西聖坊八窓庵を模した「龍松庵」他4つのお茶席があり、これらを使って伝統をふまえ、かつ新しい茶の文化を研究しています。一般の方にお茶会を開いたり、子供たちにもお茶の作法を体験してもらう場を作っています。こちらの茶室を借りて自分達のお茶会を開くこともできます。

***茶と生活 ーCHAによる豊かな暮らしを提案ー**

世界の茶研究室では、チベット・アラビア・中国・ロシア・イギリス・日本の各々のコーナーを設け、各国での茶と生活の関わり、また茶器、調度品を配しビデオによるお茶の淹れ方などを紹介しています。実際にそれぞれのお茶を試飲することもできます。また、英国式のお茶のマナーの講習等も行っています。チベットではバター、ロシアではジャム等をいれることにより寒い冬を少しでも暖かく過ごすという生活の知恵が活かされています。このように、茶は人々の生活とは切り離せないものであるということがわかります。

***茶と工芸 ーティーライフを豊かに演出する器の創造ー**

工芸工房では、茶器等の試作研究を行っています。従来の茶器のほかに、新しい素材を使い新しい茶器や道具も創作しています。また、ここでは粘土サロンを設け一般の方々にも茶器を作ってもらう体験の場を提供しています。

平成3年に天皇・皇后両陛下がATRへご視察にお越しになられた際、CHA研究センターでお茶の淹れ方、出し方を教えていただきお出ししました。お茶は淹れ方によって、よりおいしくなると思っています。毎日いただくお茶だからこそ、いつもおいしく飲みたいものです。

福寿園CHA研究センター
関西文化学術研究都市/
ハイタッチ・リサーチパーク内
京都府相楽郡木津町相楽台3丁目1-2
TEL: (0774) 73-1200
高の原下車徒歩15分
山田川下車徒歩7分

月曜日休館・見学は予約要



特願平05-063605 No.2047036

自然言語解析方式

本発明は、話し言葉などで頻出する非文法的な自然言語表現にも対処できる頑強な言語解析方式です。具体的には、自然言語入力に対して表層パターンを組み合わせた言語構造を出力する解析方式です。表層パターンは、表層語句と変数部により記述され、文や名詞句など言語的単位で分類されます。入力に形態素解析を行なった後、必要に応じて助詞の補完など入力の修正を行ないながら、入力と照合する言語構造の候補を、表層パターンを組み合わせて導出します。複数の構造候補がある場合、表層パターンの変数を具体化する語について用例と入力の間の意味的距離を計算し、入力に対する最尤の構造を決定して出力します。

特願平02-326282 No.2032639

画像の高次元復元方法

一般に3次元運動をその1次元時系列データから再構成することは運動情報が少なく、困難とされています。本発明は、2次元または1次元の低次元空間へ投射された3次元運動物体の軌跡から3次元運動軌跡をトポロジカルに再構成することを可能にし、次元運動物体の軌跡からトポロジカルな3次元運動軌跡を再構成することを可能にしました。一般に測定される時系列の低次元データは他の座標の運動成分に相関があり、他の座標の運動成分の情報を含みます。その3次元運動が測定された低次元座標と直交する場合を避け、投射された軌跡から得られるフラクタル次元が最大となるように投射すべき低次元座標を選定することにより高い精度で再構成することを特徴としています。

特願平03-226943 No.2022738

非線形発振を用いる情報の記憶方法及び装置

非線形リング共振器の高調波2次分岐発振には、波形の変調パターンが異なる非常に多くの発振モードが存在します。本発明は、光パルスなどの単極性2値パルス列を共振器に注入して、これら多数の発振モードの一つを選択的に励起することにより、この2値パルス列のビットパターンを発振波形に記憶する方法と装置について提案しています。非線形発振の自律的な安定性を利用しているため、記憶情報を保持するための、特別な再生装置が不要であり、非線形媒質をループ状にした簡単な構造の共振器に多数の情報を記憶することが可能となります。

特願平05-182744 No.2534617

人物像の実時間認識合成方法

本発明は、人物の3次元モデルをCG技術により生成した仮想空間に配置し、送信側の人物の動きを検出して、前述の仮想空間中の人物モデルにおいて検出された動きを再現して、立体ディスプレイに立体表示する通信方法を提案しています。具体的には、表情検出のために、顔にマーカを貼付して顔画像中で高速に追跡可能とし、人物の角膜に光を照射し得られた角膜反射像を撮像することにより、視線と瞬きを検出し、併せて人物の頭、体、指等の動き情報を実時間で検出します。これら検出情報を用いて人物モデルを変形するとともに、視線と瞬き情報により人物モデルにおける眼球の黒目の位置とまぶたの開閉を行うことにより、人物の動きを実時間で再現します。臨場感通信会議のように、実時間性が必要な通信に有効です。

特願平05-171264 No.2557789

ヒューマンインターフェース装置

本発明は、筋電信号と身体筋肉骨格系の内部モデルに基づいたヒューマンインターフェース装置です。具体的には、オペレータの複数の筋肉の活動を筋電信号として検出し、これをローパスフィルタを通すことによって筋肉が発生する擬似張力に変換し、非線形神経回路モデルと身体非線形ダイナミクスモデルに基づいて擬似張力から運動軌道や力軌道を推定する装置です。これによりオペレータは自身の身体非線形ダイナミクスに基づいて計算機の世界で自由に活動できます。

●イベントカレンダー

開催日	名 称	場 所	問 い 合 わ せ 先
3月5日(木)	ATR 科学技術セミナー “Foward Models for Physiological Motor Control” (仮題) オックスフォード大学 Dr. R.C.Miall	A T R	ATR 人間情報通信研究所 ☎(0774)95 1011
3月20日(木) ～21日(金)	ATR 社会エージェントワークショップ	A T R	ATR 知能映像通信研究所 ☎(0774)95 1401

●所員往来

平成8年10月1日より12月27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	A T R 所属	氏 名	出向元等
H 8 . 10 . 1	(映) 第二研究室 研究員	小林 薫	東京大学大学院
10 . 21	(映) 第一研究室 研究員	岩澤 昭一郎	成蹊大学大学院
10 . 22	(人) 感性脳機能特別研究室 研究員	仁科 エミ	放送教育開発センター
11 . 1	(映) 第四研究室 研究員	竹内 勇剛	名古屋大学大学院
11 . 1	(音) 第一研究室 研究員	Conrad Sanderson	オーストラリア (Griffith 大学)
11 . 1	(人) 第三研究室 研究員	道免 和久	慶応義塾大学
11 . 18	(音) 第一研究室 研究員	Kuldip Paliwal	オーストラリア (Griffith 大学)
11 . 21	(音) 第一研究室 研究員	R. Viswanathan	インド (タタ研究所)
12 . 9	(人) 第二研究室 研究員	Zhengyou Zhang	フランス (INRIA)

退職 年月日	復帰先等	氏 名	A T R 所属
H 8 . 10 . 14	アメリカ(南カリフォルニア大学)	Nicolas Schweighofer	(人) 第三研究室
10 . 16	住友銀行	池谷 善章	(国) 経理部
10 . 31	ドイツ (フラウンホーファー研究所)	Dirk Reiners	(映) 第五研究室
12 . 24	ドイツ (ハノーバー大学)	Felix Gers	(人) 第六研究室
12 . 26	N T T ソフトウェア	大橋 善卓	(国) 企画部
12 . 27	カナダ(プリティッシュ・コロンビア大学)	Pao-Chen Hwang	(音) 第二研究室

ATR ジャーナル25号掲載記事の訂正

ATR ジャーナル25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共に
お詫び申し上げます。

訂正箇所

所員往来 (19頁)

誤	正
氏名	氏名
Nicolas Rob	Nicolas Robin
Frans Verst	Frans Verstraten
J.C. Terril	J.C. Terrillon
和久本 雅	和久本 雅彦
Nicolas Schweigho	Nicolas Schweighofer

外部発表状況

(21頁) ATR知能映像通信研究所

項	誤	正
2.	土佐、橋本、塩崎、国井、 佐部	土佐、橋本(東大)、塩崎(東大)、国井(東大)、 佐部(東大)
5.~6.	土佐、橋本、塩崎、国井、 佐部、原島	土佐、橋本(東大)、塩崎(東大)、国井(東大)、 佐部(東大)、原島(東大)
8.	土佐、塩崎	土佐、塩崎(東大)
10.	Using.....	抹消
11.	岩室、門林、塚本 人工知能学会論文誌(96.01)	岩室(大阪大)、門林、塚本(大阪大) 人工知能学会論文誌Vol. 11, No. 5(96.09)
12.	仮想環境を.....	抹消

(25頁) ATR人間情報通信研究所

項	誤	正
110.	日本光学会光学	日本心理学会第60回大会
~113.	連合ツボクム福岡96	
113.	山田(大阪大)、山田、 Strange(Univ. South Florida) 中国語話者による・・	山田(放送教育開発センター)、山田、 Strange(Univ. South Florida) 中国語話者による日本語短母音、長母音、促 音の知覚学習

ATR ジャーナル成果展開特集号掲載記事の訂正

ATR ジャーナル成果展開特集号掲載内容に一部不適切な箇所がありましたので、以下のとおり追記さ
せていただくと共にお詫び申し上げます。

13頁に掲載いたしました写真は、NTT基礎研究所様が株式会社ナック様へ試作発注し開発された装置
です。

編集後記

先般、ATR音声翻訳通信研究所は中間期を迎え、基盤技
術研究促進センターにその技術評価を受けました。従って今
回は、その報告記事を掲載しました。ATR音声翻訳通信研
究所の研究成果を充分にお伝えすることができれば幸いです。

また、この後ATR人間情報通信研究所、ATR知能映像
通信研究所と続きます。日頃積み重ねてきた研究の成果が認
められるよう、更に精進を重ねていく所存ですので、今後と
も皆様方の暖かいご支援をどうぞよろしく願います。

(人間情報通信研究所 企画課長 市川 保男)

ATR Journal 第26号 1997年2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
〒619-02
京都府相楽郡精華町光台2丁目2番地
(0774) 95 1111 (大代表)
 - 製作 学会センター関西
 - 定価 300円 (税込・送料別)
-

本誌記事の無断転載を禁じます。

©1997 (株)国際電気通信基礎技術研究所

How a