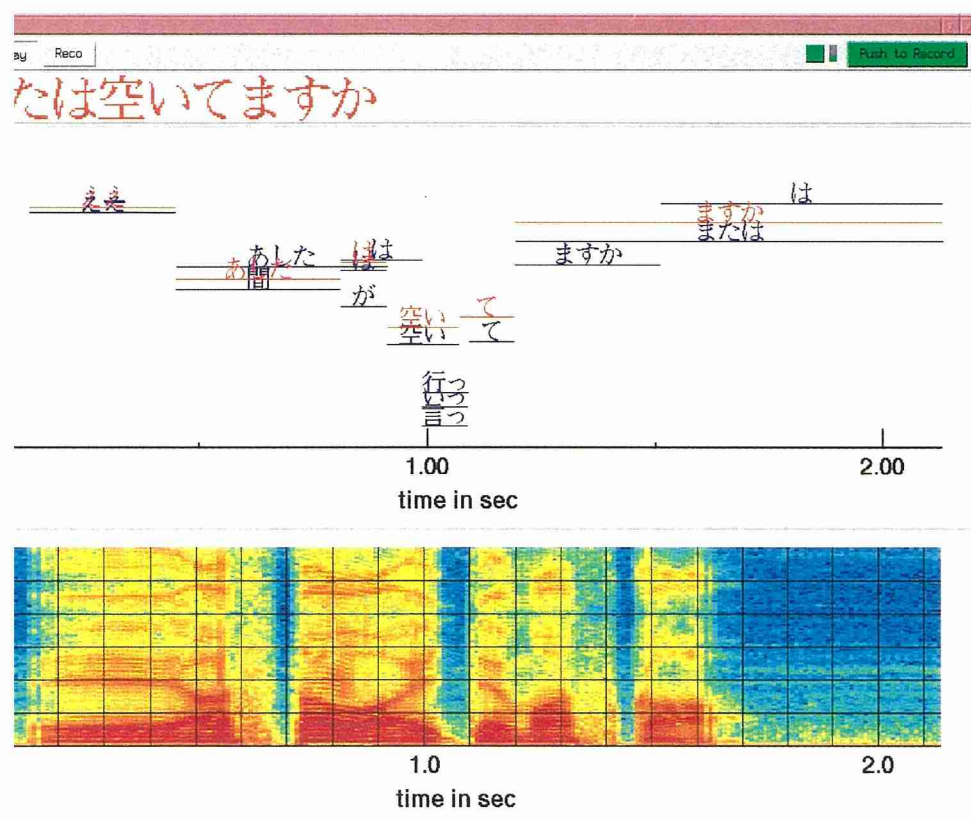


ATR Journal

WINTER 1998

30



〈表紙解説〉

これは、「ええ、あしたは空いてますか?」という自然な話し言葉の音声を、ATRが開発した音声認識システム、ATRSPRECに入力したときの出力結果の一例です。この出力結果は、時々刻々（横軸）の認識単語候補を、各々の音響的及び言語的確からしさのスコア（縦軸）と共に効率よく表現した「単語ラティス」と呼ばれる認識結果情報と、入力音声の周波数成分分布を可視化したスペクトル情報からなっています。これらの情報を併用することで、認識誤り箇所の特定や、誤り原因の解析などが可能です。

ATRSPRECは音声翻訳システム、ATR-MATRIX (Multilingual Automatic Translation System for Information Exchange) の音声認識部として用いられており、ATR-MATRIXが実現する「自然な話し言葉での音声翻訳通信」に大きく貢献しています。

●巻頭言 Foreword	「以心伝心」への期待	1	石渡 隆人
●研究動向紹介 Current research topics at ATR	プロカメラマンの感性を探る - プロカメラマンのカメラワーク解析とその利用 - A method for generating camera operations to appeal one's kansei	2	灰塚 凡樹
	自然な話し言葉を聞き取り、リアルタイムで英語の 音声に：ATR-MATRIX ATR MATRIX Translates Spontaneous Utterances into English Speech in Real Time	4	森元 遼 竹澤 寿幸 山本 誠一
	ことばのリズムが理解をたすける How Do Humans Perceive the Temporal Structures in Spoken Language?	6	加藤 宏明
	アンテナはどこをみているか - 光空間信号処理マルチビーム受信アンテナ - How to Create a Multibeam Antenna Radiation Pattern - Spatial Optical Signal Processing Antenna -	8	柴田 治
●A T R Monologue	二人称の勧め（その1）	10	葉原 耕平
●短信 letter	新鮮な経験 - 「ええと」と「あの～」のニュアンス -	12	山崎 泰弘
●トピックス What's new with ATR	科学技術セミナー	13	
	V.R. E X P O ' 97に出展	14	
	UNU/NTTマルチメディアと地球環境フォーラムに出展	14	
	けいはんな産学官エキシビション' 97に参加して	15	
	国際シンポジウム Info-Tech'97へのA T R 成果の出展	15	
	第10回研究発表会を好評開催	16	
●A T R 成果展開 Fruits of ATR ATR' s research	特許紹介	16	
	受賞	17	
●学研都市あれこれ Walking around Kansai sience sity		18	
●所員往来 Changes in the staff	・ 10月～12月	20	
●編集後記 Editor's		20	

「以心伝心」への期待



(株) 国際電気通信基礎技術研究所
取締役経理部長 石渡 隆人

年が改まり20世紀も残すところあと3年となりましたが、正直なところ最近では年が変わっても「今年こそは」と誓いを立てるような新鮮な気分にはなれず、むしろ1年間という時間の早さを痛感し、将来に対する焦燥感のようなものを感じることもの方が多くなってしまいました。

人口論的に見ると日本では1995年をピークに生産年齢人口(15～64才)は減少を始めており、昨年(1997年)には年少人口(0～15才)と老年人口(65才～)が逆転したそうです。21世紀を担う若者よりも20世紀での生産活動を終えた高齢者の数の方が多い時代に転じたということを示しています。厚生省人口問題研究所の予測によると、10年後の2008年頃には日本の総人口そのものが減少し始めるとのことです。また、マスコミ等で最近2020年という年が注目されていますが、これは高齢化のピークが来ると言われてきた年です。しかし最近の少子化傾向が続けばそれ以後もさらに老年人口の比率は上昇を続けるということであり、高齢化はまさに急ピッチで進んでいることを再認識する必要があるのではないのでしょうか。

高齢化社会の抱える問題というのは社会保障制度や国民負担の問題、あるいは国力維持といったマクロ的な観点から様々な指摘がなされていますが、人間一個人にとってみても高齢化による身体機能の低下というのは何かとやっかいなことをもたらすものです。私の身近にも老化とともに成人病からきた脳の障害により人とのコミュニケーションが不能に近い状態になり、人生の楽しみを失ったまま生き続けている人が何人もいます。身近にそういう人を見ていると、明日はわが身という不安感が自然と強くなってくるものです。

そこで期待されるのはそういう老化等により人体の機能の一部が衰えたときに、人生に対する楽しみをサポートしてくれる技術、すなわちバリア・フリー&人生エンカレッジを実現する技術です。現在でもおそらくその方面の研究開発は各方面で進められているのですが、明日はわが身という不安に駆られている身としては、よりいっそう重点的に取り組んで頂くことを願っています。エージェントとか電子秘書という概念が既に応用技術の中に用いられてきていますが、ビジネスシーンだけでなく、社会的弱者の生活シーンでの実用化を重視した開発を是非期待したいものです。事業としてもマーケット規模は結構大きいものがあるのではないのでしょうか。

たとえば新聞や本を好きな人の声でなめらかに代読してくれるロボット、電話番号などを覚えていなくても声でつながるテレビ電話などは、既に実用化に近いところまで来ているようですが、人の体調や脳の活動状態を察知して最も心地よい気分させるような映像や音楽を提供する装置等、勝手に書けばきりがいいほど多種多様なニーズが存在するでしょう。そして誰しもが期待する究極のニーズは文字どおりの「以心伝心」を可能にする技術と言えるのではないのでしょうか。たとえ脳の病気で言語機能に障害が残っても、特別の努力なしに人とのコミュニケーションができるとすれば、はるかに生き甲斐を感じることができるであろうと思われます。

21世紀のキーワードの一つは「人に優しい技術」と言われていますが、まさに高齢化社会にはそれが次々と実用化されていかねばなりません。ATRが設立の時から掲げてきた「人に学ぶ」という研究コンセプトから得られるいくつもの知見が21世紀には見事に花開き、人に優しい技術となって世の中に大きく貢献することを期待し確信しています。

プロカメラマンの感性を探る

ー プロカメラマンのカメラワーク解析とその利用 ー

A method for generating camera operations to appeal one's kansei

効率良く映像制作を支援するためのツールは既に多くのものが実用化されてきた。しかし、映像制作の専門家でない人にとっては、専門家が作成するような感性に訴える映像をどうすれば制作できるのか困難な問題である。そこで、我々は前記のような感性に訴える映像表現技術を非専門家が容易に利用できる事を目指して研究を進める事とした。まず、多彩な映像表現技術のうち、カメラワーク（ズーム、パン等）による映像表現方法に絞って、人の基本的動作とカメラワークの関係を整理する事を試みた。

We are studying a method that would allow non-professional cameramen to easily produce video to appeal one's kansei. We are focusing on camera operations among some professional techniques to produce videos. At first, we extracted camera operations from the video shots produced by professional cameramen using a 3D camera model. Following this step, information extracted from the scenario and a picture was used in order to describe each movement. We then proposed a method to generate camera operations for video shots produced by non-professional cameramen.



(株) ATR知能映像通信研究所

第三研究室

灰塚 凡樹

1. はじめに

最近の映画ではSF物に限らず、多くの場面にCG処理が用いられ、効率良く、効果的な映像が作られるようになってきました。我々に身近なビデオカメラを例に考えてみても、色々な機能が実現され、提供されるようになってきました。例えば、ディゾルブ機能（撮影シーンを切り替える時に、ある場面と次の最初の場面の絵を重ね合わせる）やセピア撮影機能（映像の色がセピア調になる）等。しかし、それらのツールを使いこなして、専門家が作成するような感性に訴える映像をどうすれば制作できるのか映像制作の専門家でない人にとっては、想像もつかないほど困難な問題です。そこで、我々は前記のような感性に訴える映像表現技術を専門家でない人が容易に利用して、映像制作を可能にする事を目指して研究を進めています。映像表現技術といっても多種・多彩なのですが、身近なビデオカメラによる撮影等に適用できる事を考えて、カメラワーク（ズーム、パン等）による映像表現方法に絞る事としました。以下に、感性に訴えるカメラワークを自動生成する方法について説明します。

2. 専門家のカメラワーク抽出

まず、専門家のカメラワークを解析する必要があります。このためには、分析を容易にするために素材を選択する必要があります。その条件として、あまり専門家が様々な技を使わずに制作し、かつ、撮影対象とカメラワークとの意味的な対応関係を知るためにシナリオに基づいている映像である必要があります。また、これらの映像やシナリオを容易に入手できる物があれば好都合です。

(1) 素材選択の条件

カメラワーク以外の映像表現技術を使うことな

く、また、シナリオに沿って制作された映像として思い付くのが演劇の中継番組です。劇のシナリオを理解する事により、悲しい場面とかうれしい場面とか、見せ場とカメラワークの対応付けは容易になります。また、踊りであれば動作手順が決まっているので、カメラワークと動作との対応関係を整理する事が容易となる事が期待できます。その様な条件を考慮した結果、テレビ番組として放送された「能」の舞いを素材として解析する事としました。

(2) 「能」の舞いの特長

「能」の舞いの主な特長を以下に示します。①歌詞（台詞、演技方、歌い方の指示等も含む）、舞いの動作、舞台上の演者の動作軌跡がシナリオに示されている。②前記各情報は各々が関連付けて示されている。③舞いは60種類程の動作により構成される。④能舞台は大きさ、構造が基本的に同じである。

(3) カメラワーク算出

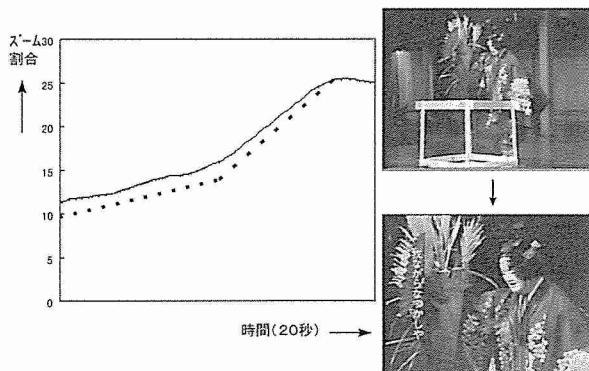
カメラワークは3次元カメラモデルを使用して自動算出します。このモデルではカメラ位置は固定されており、1台のカメラを使用する事を前提としています。そこで、専門家が制作した複数カメラによる映像を、手動で分割し各カット毎にカメラワークを算出する事としました。

(4) 動作解析データベース

1動作の映像を1単位として、シナリオ上の情報を用い、演者の動作名、役柄、感情の劇的変化の有無を映像情報と連携してデータベースに登録しました。また、映像情報としては、映像の開始時間・終了時間、カメラワーク算出結果から、カメラワーク（ズーム、パン）、及び、演者の映像上のサイズ情報（上半身だけの映像とか全身像の映像等）を登録して動作や、場面に対応したカメラワークの傾向を調査しました。

3. カメラワーク解析結果

能映像の効果的表現方法として、専門家は主にズーム操作を用いている事がわかりました。また、ズーム操作は一直線に行なうのではなく何らかのアクセントのついたものとなっています。



実線は実測値、破線は近似線

図-1 プロのカメラワーク（ズーム）の例

A, B, Cの3種類の映像のカメラワークを自動算出し、動作解析データベースに各情報を登録し、映像表現の解析に利用した例を示します。ズームイン操作の例（サイズが10%以上変わった場合のみ）を表-1に示します。

表-1 ズームイン操作の例

No.	動作名	映像	サイズ (注1)
1	扇を広げ	A	BS→CU
2	角取り	A	FF→FF
3	サシ込	A	WS→WS
4	大左右 (右)	A	WS→WS
5	開	B	FF→FF
6	正へ出	B	BS→CU
7	開	B	BS→BS
8	開	C	FF→FF
9	左へ回り	C	FF→FF
10	角へ行く	C	FF→FF
11	正へ出	C	FF→FF

注1) CU: 頭部だけの映像、BS: 胸から上の映像、WS: 上半身だけの映像、FF: 全身像の映像

これらの情報を整理して舞を舞う方に提示すると、見せ場と考えている所が強調したカメラワークを与えられているとの見解が得られました。そこで、非専門家がカメラワーク無しで撮影した映像に後付けで、動作解析データベースの情報を基にし、また、強調情報をユーザ設定情報として与える事によりカ

メラワーク用情報を生成し、専門家が撮影した映像の様に感性に訴える映像を得られる方法について提案する事としました。

4. カメラワーク生成方法

(1) 強調部分の指定

舞いの舞った人、あるいは映像を見る人が自分の好みで、どこが強調すべき動作か入力させます。動作対応のカメラワーク情報として、強調する場合のズーム操作方法と画面に対する人のサイズを役柄対応に生成します。動作の実施前か、実施中に感情の変化がある場合はその状態に相当するカメラワークを選択/生成します。

(2) 単位カットの決定

強調する動作を含む句・含まない句を明らかにし、単位カットとする。

(3) カットの決定

見せ場のある単位カット同士のマージ、見せ場の前後の句間の画角の整合性から、各句対応のカメラワークを決定し、合わせて、カットチェンジ個所を選択します。

(4) カメラワークの与え方

映像に特殊効果をかける装置を用い、ズームイン/アウト操作を映像の拡大/縮小機能で擬似し、また、パン操作は映像を上下、左右方向に移動させる事により擬似する事としました。

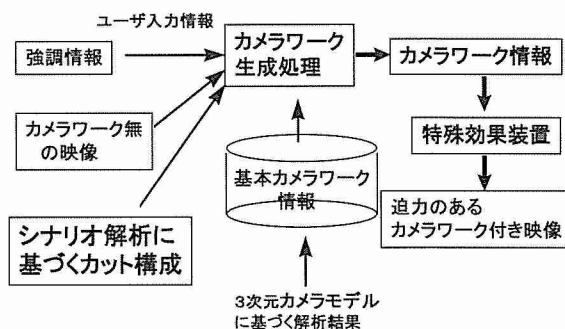


図-2カメラワーク生成方法

5. まとめ

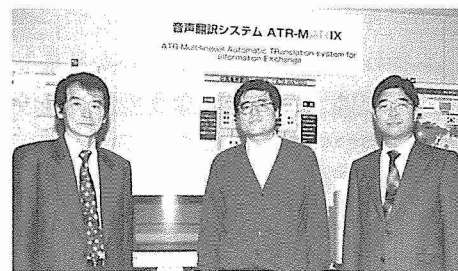
非専門家でも感性に訴える映像を、容易に制作できる方法について提案し、プロトタイプを試作しました。カメラワークに絞ってみても動作解析データベースを用いる事により、専門家の制作した映像に秘められる各種の技がある事が見えてくるのですが、上手く整理するのは難しく、全自動化には更に解析が必要です。今後、自動化を進めるための諸条件を明らかにするとともに、能の舞以外についての適用方法を提案して行きたいと考えています。

自然な話言葉を聞き取り、リアルタイムで英語の音声に：ATR-MATRIX

ATR-MATRIX Translates Spontaneous Utterances into English Speech in Real Time

ATR音声翻訳通信研究所では、これまで音声翻訳に必要な要素技術の研究を進めてきましたが、今回これらを有機的に統合した日本語から英語への音声翻訳システムを開発しました。普段我々が使うような自然な音声でもこれを認識し、英語へ翻訳し、合成音声で出力することができます。

ATR Interpreting Telecommunications Research Laboratories has been pursuing researches on componential technologies necessary for developing an advanced speech translation system. This time, it has successfully developed an innovative speech translation system. This system can recognize natural utterances such as spoken in daily life, can translate them in English and output synthesized voice.



(株) ATR 音声翻訳通信研究所

森元 逞、竹澤 寿幸、山本 誠一

1. はじめに

国際化の進展に伴い、音声翻訳システムの実現に対する期待が高まっています。外国の人と日本語で自由に話しができれば、どんなに素晴らしいことで



しょう。ATR音声翻訳通信研究所では、自然な話し言葉を対象とした音声翻訳システム ATR-MATRIX (Multilingual Automatic Translation system for Information eXchange)を開発しました。このシステムの構成を図に示します。入力された音声はSPRECで認識され、認識結果が文として出力されます。次にCHAT-TRANSLATION(以下、CHAT-Tと呼びます)で認識結果の日本語を英語に翻訳します。最後に、CHATRで英語の音声を合成して出力します。全体の処理の流れは制御部で制御されています。これらの処理はすべて1台のワークステーションで動作し、またほぼ実時間で処理を行うことができます。以下では、このシステムの特徴をご紹介します。

2. 自然な音声の認識

音声認識を行うには、音のモデル(これを音素モデルと呼びます)を用意しなければなりません。人が実際に喋った声を分析し、/a/、/i/、/u/、/e/、/o/などの母音や、/k/、/s/、/t/、/n/などの子音のモデルを作ります。しかし記号で書けば同じ音素でも、その特徴は多種多様に変化します。同じ/k/でも、/aka/(赤)のものと/aki/(秋)のものではかなり異なります。また話し方(口の開け方や話す速さ)などによっても大きく変わります。最も大きい変動は、人の違いです。男性、女性で違いがあることはもちろんですが、個人によっても大きな違いがあります。SPRECでは、色々な人の音のモデルを用意しておき、話し手の声の質にあわせて、どのモデルを用いるのが最も良いかを自動的に判断し、そのモデルを用いて認識を行っています[1]。この判断は一文ごとに行います。これにより、色々な人の声でも高い精度で認識できます。音声認識では、上記のような音のモデルとともに、認識すべき単語を定義した辞書

と単語の並び方に関する情報を利用します。これを言語モデルと呼びますが、これをどう作るかが問題となります。なぜなら、音声認識精度を良くするためには、単語の並び方に関する制約をきつくするのが良いのですが、あまりきつくと、ほんのちょっと崩れた表現をしても、システムが認識してくれなくなるからです。SPRECでは、実際に人の喋った多量の文から、自動的にもっとも良い制約を取り出して言語モデルを作成しています[2]。また、自然な話し言葉では、「あのー」、「えーと」といった言葉がかなり頻繁に現れることがあります。このような表現があっても正しく認識できます。

上で述べたように、声は人によって様々に異なり、また同じ人でも喋るたびに変わるため、なかなか100%正確に認識することは出来ません。なるべく高い精度を得るため、音声認識では同時に多数の候補について並列にチェックしています。しかし、候補の数を増やせば増やすほど処理が重くなって実時間で処理することが難しくなります。この問題を解決するため、SPRECでは、発話が始まったただちに認識処理を開始します。この方法をとることにより、発話が終わってほんの少しの時間遅れで認識結果を出力することができます。

3. 自然な発話の翻訳

SPRECから出力された認識結果を受け取り、英語への翻訳を行います。現在市販されている多くの翻訳ソフトでは、翻訳規則を手で定義し、これに基づいて翻訳を行う方法が採用されています。例えば、「～したい」→「want to～」のような規則が定義されています。このような規則を組み合わせることにより、日本語の文を英語の文に翻訳するわけですが、しかし、このような方法を話し言葉に適用しようとすると、慣用的な表現や文法的に多少崩れた表現を翻訳することができなかつたり、翻訳できたとしても固い訳しか作り出せない、などの問題が生じてしまいます。CHAT-Tではこのような問題を解決するため、用例を用いた翻訳機能を実現しています[3]。これは、種々の翻訳用例を参照することによって、柔軟な翻訳を行おうとするものです。例で説明しましょう。日本語の「AのB」という表現に対して、英語への翻訳を考えると「B of A」、「B in A」

など多数の表現があり、そのうちどれが最も適切かを判断するのはなかなか困難です。このような問題に関し、CHAT-Tでは、まず以下のような種々の対訳用例をシステムに格納しておきます。

「僕の本」→`a book of mine`
「英語の本」→`a book in English`
:
「イギリスの毛織物」→`fabric made in England`
:

そして、例えば「中国の絹織物」という文を翻訳する場合には、これらの用例のなかから最も類似した用例を使って翻訳を行います。この例では、「中国」と「絹織物」に最も似た単語が使用されている「イギリスの毛織物」の用例が選択されます。次に、「イギリス」、「毛織物」をそれぞれ「中国」、「絹織物」に置き換えることによって、最終的に`silk made in China`という翻訳結果を得ます。

音声認識では、文の一部のみを間違えて認識することが時々あります。つまり、ある部分は正しい単語が認識され、人間が聞けばそれだけでも十分に意味が通じる場合があります。CHAT-Tではこのような誤りに対処するため、部分翻訳機能を実現しています。文全体として翻訳できない場合でも、文法的、意味的にまとまりのある翻訳結果が得られればそれを出力します[4]。

4. 自然な音声の合成

CHATRでは、滑らかな音声で、かつ話し手の声に似た声で合成することができます。まずある人が実際に喋った音声をデータベース化し、種々の音声単位を切り出しておきます。CHATRでは、合成すべき文が入力されると、どのような音声単位をつなぎ合わせるをもっとも自然で滑らかな音声合成できるかを判断します。合成すべき音声との差が小さく、またつなぎの部分の不連続度になるべく小さいものを選択するようにします。ある程度の量のデータベースを用意していれば、極めて自然な音声合成することができます[5]。また前に述べたように、SPRECでは話し手の声に最も適した音声モデルを選択しますが、どの話者のモデルを使用したかはCHATRにも知らされます。CHATRでも複数話者の音声データを用意していますので、話し手の声に最も近い話者データベースを使用することにより、話し手の声に似た声で合成ができるわけです。

5. システム制御

システム制御は、システム全体の制御を行います。特にMATRIXでは、各コンポーネントのデータの入出力を全て制御部を通して行う方式とし、またこの時に制御部で各コンポーネント特有のデータのフォーマットから標準フォーマットに変換する方式を実現しています。これにより、あるコンポーネント

のプログラムが多少変更されても、その影響を最小限に抑えることができます。このような制御を行うと同時に、システム全体に係わる機能の一部を分担しています。例えば、自然な会話では、文ごとに区切らず、2つの文をつないで喋ってしまうことがありますが、このような場合は発話分割機能により2つの文に分け、1文ごとに翻訳や合成の処理を行います。また、「部屋はありません?」というように文末を上げることによって疑問を表すことがありますが、このような文には疑問文のマークを付けてCHAT-Tに渡します。

6. まとめ

今回試作したシステムでは、「ホテルの予約」の場面を対象としていますが、辞書内容を置き換えることにより、他の場面にも適用することが出来ます。

今後、さらに精度向上のための各種改良を進めるとともに、英日方向の同様なシステムを開発し、最終的には日英、英日双方向での会話が可能システムを実現する予定です。特に対話では、話しの流れから省略された語やあいまいな表現の意味を理解しなければならない場合がありますが、このような問題にも積極的に取り組んで行く予定です。さらには、多言語間の音声翻訳をめざし、C-STAR IIという国際コンソーシアムを通じて世界各国の研究機関と研究協力を進めています。プロジェクト最終年度(1999年度)には、これらの機関と共同で、日本語、英語、ドイツ語、韓国語間での多言語音声翻訳国際実験を行う予定です。

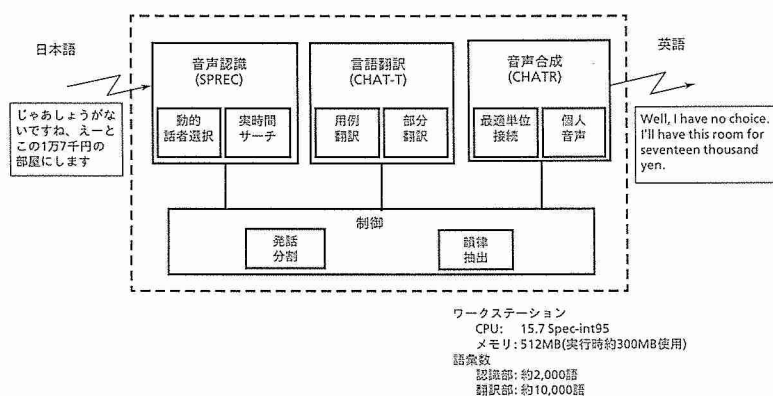


図 ATR-MATRIXの構成

参考文献

- [1] 小坂、勾坂、松永: 「人見知りをしない音声認識 -誰の声でも認識できる技術を目指して-」、ATRジャーナル、No. 16, 1994
- [2] 清水、山本、政瀧、勾坂: 「単語のつながりで言葉を聴きとる -自然な話し言葉の音声認識をめざして-」、ATRジャーナル、No. 24, 1997
- [3] 隅田、古瀬: 「経験的知識を活用する新しい言語翻訳手法」、ATRジャーナル、No. 12, 1992
- [4] 脇田、飯田: 「人の声を十分に聞きとれなくても翻訳出来ますか? -音声認識の英会話を特定しながら部分的に翻訳する-」、ATRジャーナル、No. 28, 1997
- [5] キャンベル、樋口: 「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-」、ATRジャーナル、No. 26, 1997

ことばのリズムが理解をたすける

How Do Humans Perceive Temporal Structures in Spoken Language?

話しことばにとってリズムは、情報の受け渡しを円滑に行なうための基本的な約束事のようなものです。私達のグループは、今、このことばのリズム、もっと一般的に言えば音声の時間構造、を耳で聞いて把握するための手がかりに興味を持っています。機械が人と同じことばのリズムを共有できるような人にやさしい通信技術への貢献を狙って、今まで直感的に説明されることの多かったリズムの手がかりの究明に、実証的な研究方法を持ち込むことにより、従来は知られていなかった新しい説明が可能であることを発見しました。

Rhythms or, more generally, the temporal structure of speech sounds serve an essential role as a form of protocol for oral communication. In our search for clues that will allow us to share such protocols between humans and machines, we extensively examined the perception of speech rhythms by introducing empirical approaches instead of the intuitive ones that have been commonly used in this area. Psychoacoustical experiments revealed that the perceptual salience of the temporal cues embedded in speech sounds can be described more effectively in terms of their fundamental auditory base, i.e., loudness, than by linguistic factors such as syllables or morae.



(株) ATR 人間情報通信研究所

第一研究室

加藤 宏明

1. ことばのリズム

「かっぱかっぱらった
かっぱらっぱかっぱらった
とってちってた」

(谷川俊太郎『ことばあそびうた』福音館書店より)

ふだん、あまり意識することはありませんが、うまくできた詩を耳にすると、なるほど確かにことばにはリズムがあるのだ、と納得せずにはいられません。

話しことばにとってリズムは、情報の受け渡しを円滑に行なうための基本的な約束事のようなものです。話し手(情報の送り手)と聞き手(情報の受け手)で同じリズムを共有しているからこそ情報伝達はいくつもの点でうまくいくのです。もし、これが双方で食い違っていたりすると、とたんにコミュニケーションに支障をきたすことでしょう。このような事情は、話しことばに限らず、時間とともに変化する媒体を介した情報伝達手段(例えばモデムなど)には広く共通のものであります。

ことばのリズムの共有は、コミュニケーションが人と人との間で行なわれている場合には、特に意識されることなく自然に実現されています。ところが、時代は既に人が機械と話しことばを使ってコミュニケーションをとろうというところまで来ています。人と機械の場合には、機械は人のように直感的にリズムを感じとってはくれませんから、それを補うためにどうしても人の側が機械に合わせることになり、余分な負担を強いられるようになりがちです。こちらが急いでいるかどうかとはお構いなしに、いつも一定のリズムで繰り返される自動応答システムに、不自然さを超えてある種のいらだちを覚えたことのある方も少なくないと思います。これを避ける

ためには、例えば、機械の話すことばを人に聞かせる場合には、人が使う手がかりを使って人が聞きやすいリズムを作ってやることが求められます。逆に、機械で人の話すことばを聞きとる場合には、人が使っている手がかりでリズムを把握した方がより自然なしゃべり方に対応できると期待されます。つまり、「人の感じるリズムが何に基づいているのか」を機械に教えてやる必要が生じてきたのです。

2. 何を聞いているか

ところが、これまでことばのリズムの手がかりについての研究は、どちらかといえば直感的な論考が多く、機械に教えてやれるような具体的なデータを伴う実証的な調査はあまり行なわれてきませんでした。例えば、拍あるいはモーラと呼ばれるのは基本的にひらがな1文字に一致するリズムの単位として直感的に定義された用語です。「刻」というタイミングのポイントも提唱されました。これは、ほぼ母音の開始時間に相当するタイミングであると説明されていますが、やはり直感から得られたものです[1]

そこで、私たちのグループでは、何をリズムの手がかりとして人はことばを聞いているのか、実証的なアプローチを試みることにしました。まずは、手がかりの候補を絞ることから始めるのが早道です。とは言うものの、発話によって大気中に放射された音声は、音楽(特に西洋音楽)のように楽器のアタックとか音程の離散的な変化で拍節構造が明示されているわけではありませんから簡単には決められません。しかし、リズムを示すための手がかりは、音声の中の音響特性が定常な部分やぼんやりとした変化しか示さない部分ではなく急激に変化する部分であるということは確かなように思われます。そして、その手がかりは発話された音声の中に普遍的に(それ相応の頻度で)しかも繰り返して存在しているもので

なければなりません。

上記2つの条件をもっともうまく満たす候補として、私たちは“a, i, u”などの母音と“k, s, t, n”などの子音との境界をリズムの手がかりとして想定し、ここから調査を始めることにしました。

実は、ここまでは、おもしろいことに先達により直感的に得られた解と同じです。子音から母音への境界(C-to-V境界)は「刻」が想定したタイミング点と基本的に一致しますし、母音から子音への境界(V-to-C境界)は拍あるいはモーラの切れ目に常に一致します。しかしながら、C-to-V境界とV-to-C境界とどちらが手がかりとして重要であるのか、となると人によって見解が分かれます。双方に言い分があり、議論は尽きません。

3. V-to-CかC-to-Vか

そこで私たちは、この議論に決着をつけるため、一連の聴取実験を行ないました。図1に示すように、単語の中のある一つの音韻(母音か子音)の継続時間長を少し縮めてやり、隣の音韻を同じだけ伸ばしてやります。伸ばしたり縮めたりする部分をうまく定常的な性質をもつところだけに絞れば、変化の激しい境界部分だけが移動することになります。このようにして操作した音声进行をたくさんの人に聞いてもらい、元の音声に比べて時間構造がどれくらい違っているかあるいはどれくらい歪んでいるかを判定してもらえば、境界部分が時間の手がかりとしてどれくらい強く働くかがわかります。つまり、元のものの違いがより大きいと判定されるほど、その際移動させられた境界部分が手がかりとしてより強力である、ということになります。

さて結果ですが、実は、V-to-C説、C-to-V説どちらにも軍配は上がりませんでした。そのかわり、手がかりの強さを示す別の要因のあることが分かりました。それは、図2に示すような境界の前後におけるラウドネス¹の落差で、この落差が大きいほど、境界の移動の影響が大きいという結果でした[1]。

結局、直感によって片方が有利に見えたりしたが、おしなべて見ればV-to-C境界もC-to-V境界も変わりなかったということです。それよりも、急激な大きさの変化が生じた点を「節目」として捉えようとする聴覚の基本的な特性が手がかりの重要性を支配するのだ、と解釈できます。さらに、これまでC-to-VかV-to-Cかという見方でリズムの聞こえの説明にある程度成功していたのは、それらが心理音響的にも意味のある大きなラウドネスの変化部分にたまたま一致していたから、とも云えるわけです。つまり、研究にあたって用いられたサンプルによって2通りの主張が導かれていたということになります。

この結果は、ある意味では当たり前ではあるのですが、直感に基づく論考を行なった人々、それを常識と考えていた人々には少なからぬ驚きをもって迎えられました。

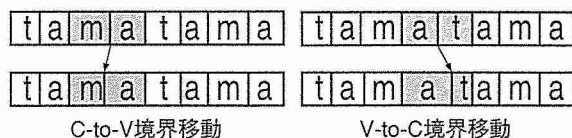


図1 境界部分の移動操作

大きな落差のある境界の方が移動による影響が大きい

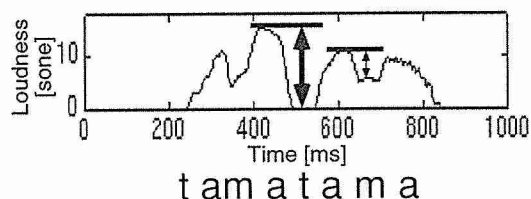


図2 境界部分でのラウドネス落差

4. これから

以上お話しした成果はまだ基礎的なレベルのもので、これにより人と機械との話しことばによるコミュニケーションが飛躍的に向上するというわけではありません。C-to-V境界とV-to-C境界との関係についても、手がかりとしての強さの点では両者に差はないものの、機能的な違いはあるらしいことが最近分かってきており、まだ詰めなければならない点が多く残されています。しかし、視点の転換、あるいはことばのリズムについての保守的な考えからの解放、を促すという意味で意義のある成果であると考えています。このような知識の積み上げが本当の意味で人に負担を強くないコミュニケーション技術の実現につながるものと信じています。

5. 謝辞

聴覚による時間構造の知覚について助言をいただいた津崎主任研究員に感謝いたします。

参考文献

- [1] 川上: 日本語のリズムの原理: 国学院雑誌, 82(9):48-55, 1981.
- [2] H.Kato, M.Tsuzaki, and Y.Sagisaka: newblock Acceptability for temporal modification of consecutive segments in isolated words., J. Acoust. Soc. Am, 101(4):2311-2322, 1997.

¹ 平たく言えば聞こえの大きさ。聴覚心理学的に具体的に言えば、人間の周波数感度特性、時間積分特性を考慮した重みづけをして計算したパワーということになる。ただし、時間的な積分については、ここでは考慮していない。

アンテナはどこをみているか

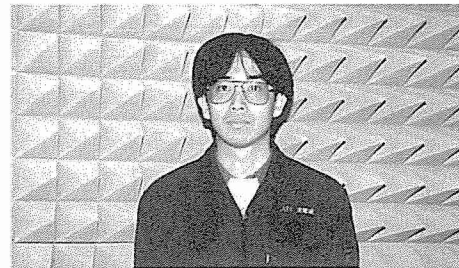
How to Create a Multibeam Antenna Radiation Pattern

— 光空間信号処理マルチビーム受信アンテナ —

Spatial Optical Signal Processing Antenna

マルチメディアというキーワードのもと、有線はもちろん無線通信に求められる機能も非常に複雑になってきております。無線通信では、その情報の出入り口にアンテナが使われており、そのため多くの機能がアンテナに求められるようになってきました。私たちの研究室では、将来のアンテナに求められる姿を想定して、さまざまなアンテナの研究を行っています。ここでは、その1つである光空間信号処理を用いるマルチビーム受信アンテナについて紹介します。

As wireless telecommunication systems advance, antennas are expected to have many functions and high performances. In ATR, we are studying a many types of antenna for future wireless communication systems. Spatial optical signal processing antenna is the one of these researches. In this paper, we introduce a receiving mode spatial optical signal processing antenna for multibeam reception.



(株) ATR環境適応通信研究所

第三研究室

柴田 治

1 はじめに

最近の通信分野の進歩にともない、アンテナにも多くの機能が求められるようになってきました。このようなアンテナの1つに、1つのアンテナでたくさんの方向から来る信号を受信出来るアンテナが考えられています。この様なアンテナがあれば、従来たくさんの方向から来る信号を受信するためには方向毎にたくさんのアンテナが必要だったのが、1つのアンテナでカバーすることができるようになります。このようなアンテナをマルチビームアンテナとよび、図1のように、たとえば移動通信の基地局や、衛星通信などに使えるのではないかと考えられています。この場合、アンテナはいろいろな方向から来た電波を同時に受信するため、アンテナには様々な方向から来る信号を、到来方向毎に分離する機能が必要になります。ATRでは、このような機能を光空間信号処理によって実現する方法を提案しています。

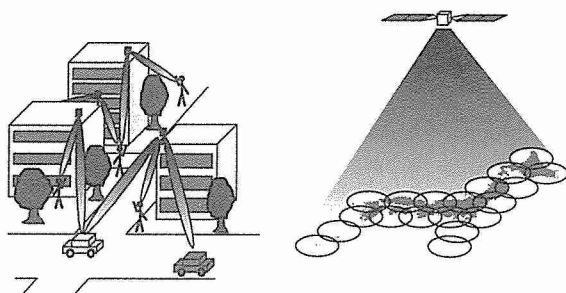


図1 マルチビームアンテナの一例

2 光空間信号処理ってなに？

図2のように真ん中にレンズを立て、左側に物体を、右側にスクリーンを立てれば、物体のイメージがスクリーンに映ることはよく知られています。レンズの代わりにピンホールの開いた板を立てた場合でも同じです。説明を簡単にするために、左側に置

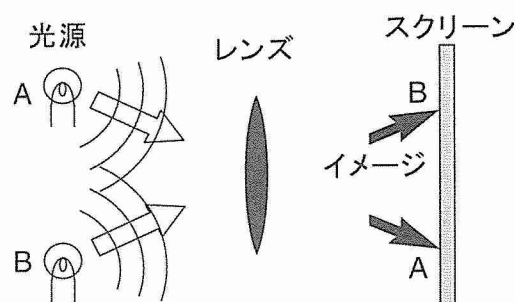


図2 光による並列空間信号処理

く物体を光源として説明を続けます。光源の位置が上であればスクリーン上のイメージは下に映り、光源が下であればイメージは上に映ります。ここで光源が沢山ある場合を考えてみましょう。もしレンズがなければ、スクリーンにはそれぞれの光源からでた光がすべて混ざったものが照射されます。一方間にレンズを立てた場合、レンズの左側ではそれぞれの光がすべて混ざっていても、スクリーン上にはそれぞれの光が、光源の位置によって別々のところにイメージを結びます。ATRでは、この原理がマルチビームアンテナに必要な、様々な方向から来る信号を到来した方向毎に分離する機能に応用できるのではないかと考え、研究を始めました。

3 なぜ光？

実はこのように、マルチビームアンテナの受信信号を到来方向毎に信号を分離する回路は、光を用いる方法以外でもいくつか方法が提案されており、すでに実現されているものもあります。ではなぜATRでは光を使うことを考えたのでしょうか。1つの理由として、構造が比較的シンプルで小型に出来るのではないかとということが考えられます。アンテナの素子の数が非常に多い場合、従来技術では回路規模が、おおざっぱに言って素子数の2乗に比例して

非常に複雑になっていました。一方光空間信号処理では、基本構成はそのまま一部回路を増やすだけで対応できます。また、光を使うもう1つのメリットは、広帯域な信号を同時にあつかうことができるという事です。光の信号は電波の周波数と比べると、 10^{15} 以上もの非常に高い周波数のため、電気の領域で非常に広帯域な信号でも、光の領域での信号処理では、ほとんど無視できるほどの帯域として扱う事ができます。

4 実際のアンテナ系

実際にATRが提案している、光空間信号処理を用いた受信系のアンテナを説明しましょう。図3のように、複数の方向から到来する電波の信号を1つのアンテナの中にある複数の素子で受信します。ここで、アンテナ面で受信した信号は、レンズ面に到達する光といっしょで、いろいろな方向から到来した電波がすべて混ざった状態となっています。次にこれら各素子で受信した電気の信号を光の信号に変換(E/O変換)します。そして変換された光を空間に放射すると、その光は、レンズから出た光がスクリーンに焦点を結ぶように、アンテナで受信した電波の到来方向毎に、それぞれ異なった方向にイメージを結び、空間的に分離されます。これらをそれぞれ検波して光信号から電気信号に戻せば(O/E変換)、アンテナで受信した信号をそれぞれ到来した方向毎に別々に取り出す事ができます。

ATRでは実際にE/O変換部分を試作して、機能実証を行いました。上述のように原理は簡単です

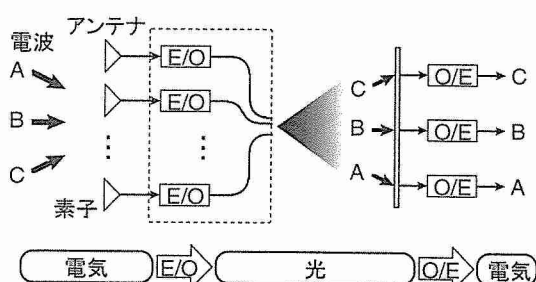


図3 光空間信号処理アンテナの構成

が、実際に回路を実現するには多くの技術課題があります。たとえば光の安定性をいかにして確保するかということです。光の波長は非常に短いため、材質のちょっとした伸び縮みなどにより、その状態が簡単に変わってしまいます。今回の回路で特に重要となるのは、図3に点線で示す、電気-光変換(E/O)された光が空間に放射されるまでの間です。この部分で光の状態が変化してしまうと、放射された光がうまく右側にイメージを結ばなくなってしまいます。ATRでは点線部分を1枚の基板上にまとめて作り込み、となりの経路との相対的な安定性を確保することでこの問題を解決しました。また、どうやって電波の情報を光の情報に変換するか、といったことも課題の1つです。これについては、あらかじめレーザー光を用意しておき、光変調器を用いて電気の信号を光に乗せるといった方法で解決しました。このため、実際の構成では図3に示す構成とはE/OとO/Eの変換部分が若干異なっていますが、電波を光に変換している関係から、実際には電波の到来角度と光の分離する角度が異なりますが、その関係は比例関係になっており、アンテナで受信した信号は、すべて光の領域でそのイメージを形成することができます。

試作コンポーネントを用いた実験の結果、実際にアンテナ受信信号を模擬した電気信号を入力すると、信号到来方向毎に光が異なった方向に分離することが実証され、電波の到来方向が複数になった場合でも、それぞれが分離されることを確認しました。

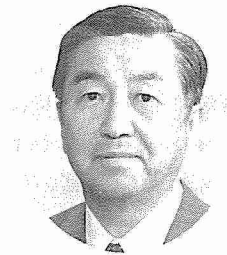
5 まとめ

このようにアンテナの信号処理に光を用いることで、複雑な処理を、簡単な構成で、同時に実現することができます。今回ここで紹介したのは、光空間信号処理による受信側のアンテナですが、ATRでは同様に送信側のアンテナについても研究しております。また光空間信号処理以外にも、将来の無線通信に必要なアンテナの姿を想定して、さまざまな機能をもったアンテナの研究を行なっております。

二人称の勧め（その1）

(株) 国際電気通信基礎技術研究所

顧問 葉原 耕平



今から十数年前、真藤さんが当時の電電公社総裁に就任された折り「おまえ達の話しているのは電電話だ」とか「一人称で行動しろ」など多くの真藤語録が生まれました。それ以来旧電電公社、NTTでは猫も杓子も「一人称、一人称」と口にするようになり、現在に至っているように思えます。私は割に天の邪鬼なところがありますので、それ以来ことさら「二人称で考えよう」と言ってきております。そして「君たちの設計している装置機器は奥さんやおじいさん、おばあさんでも使えるかね」と言ってきました。私は「二人称で考える」ことが「一人称で行動する」ことの原点だと思っています。注意して頂きたいのは「考える」と「行動する」の使い分けです。「二人称で考える」ことなしに行動すると一人よがりになり兼ねません。今回と今回はこのテーマについて私の考えをいくつか述べてみます。

① ユーザの立場で考える

ATRでは「通信の最終ユーザは人である」という立場で、そのユーザである“人”がどのように外界を認識し、行動しているか、などの視点で研究を行っている、という思想をことある度に強調してきました。その発想の原点は次のような当たり前のことにあります。

電話交換の初期いわゆる手動式の交換台が使われていた頃、ユーザは交換手に「角のうどん屋」というだけで繋いでくれました。ところが、いわゆる“自動交換機”が導入されて以来ユーザは手でダイヤルを廻さねばなくなり、ユーザにとって“自動式”は“手動式”ということになりました。このように“自動”という言葉はネットワークの運用側（キャリア側）の見方であって、ユーザの見方であるとは言えません（もっとも、ずっと昔は“自動”ではなく“自動”という字が当てられていました）。同じようなことは自動販売機、自動改札機などにも言えます。サービスの提供側とユーザの立場を変えると“自動”とか“手動”とかいう言葉は再検討が必要でしょう。ただし、私はいわゆる“自動化”が間違っていたと言いたいのではありません。しかし、“自動化”によって得たものも多い反面、「角のうどん屋」という人間味のあるやりとりのように失われたものもあります。情報社会に向けては、サービスの提供側からユーザ側に視点を移して考えてみるのが大切です。余談ですが、最近そういうことに気がついて“売り場”を“お買い場”に変えたデパートがあるとも聞いています。ついでですが、私は関東で話しをする時は「うどん屋」を「そば屋」に変えています。

② 相手の立場で考える

私は信条を聞かれたとき「相手の立場で考える」と言っております。これは一見大変良心的なようですが、本当は辛辣な発想でもあるのです。人と話したり交渉をしたりする時、大抵の人はまず自分の考えを主張するでしょう。私も大抵はそうです。ですが、時によっては同時に自分を相手の立場に置き換えることもやってみます。そうすると、どういうことを言い、どういう行動をするかということの想像が付きまします。大袈裟に言えば「敵を知る」ということです。相手の発言が傍目には異様であったり、けしからんというようなことがあっても、自分が相手の立場ならやはりそう言い、行動するだろう、ということであれば、それは異様でも何でもありません。その想像と異なっていればその方がよほど大変な情報です。よく裁判で判決が出たとき敗れた方は「判決文を精査して、云々」と言います。庶民には「木で鼻をくくったようで往生際が悪い」と映ることもままありましょうし、私も大抵同感です。しかし、当事者の立場では多くの場合それ以上のことが言えるでしょうか。もし、庶民の期待のようにはじめから「悪うございました」と言えばこれは大変なことで、私なら「何かあったのかな」と思ってしまいます。そして、それならはじめから裁判など起こすべきでなかったことにもなります。

少し生々しいですが、身近な例を述べます。ATRでは外部からしばしば資料提出や説明を求められます。ときには「なぜこんなことまで」と思うことも有り得ます。しかし、相手の担当の立場に立てば上司に説明する時どんな質問があるかも知れず、準備万端整えるのは至極当然で、そのためにはしつこい程聞いておくのは当たり前です。さらに上司はまたその上司あるいは別の機関に説明する、ということを思えば出来るだけ適切な情報を提供するのはわれわれのためであり、われわれに代わっていろいろと外部に説明して頂いている、ありがたいことだ、という単純明快なことが理解できるわけです。

また、ちょっと具合の悪いことを何とかしなければならぬ、というようなことを想像すると、出来るだけ手短かに最小限の説明でさっと逃げ帰りたい、と思うのは人情でしょう。そういう態度や説明に遭遇す

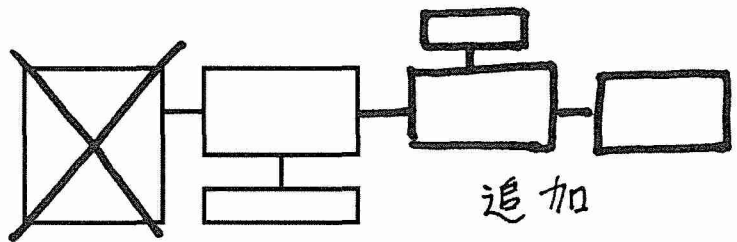
ると「ああ、やっぱりね」という印象が残ります。逆効果です。私には、こうして知らず知らずの内に人の行動を品定めしている悪い癖があるようです。

③ 特許の書き方

A T Rでは特許出願の事務手続きは各研究開発会社の自由裁量です。ですから、アイデアだけ伝えて後は弁理士さん任せのところもあれば、担当者が私に説明に来るところもありました。担当者は緊張気味だったかも知れませんが、これはお互いにいい機会ですから私は出来るだけよく話を聞くようにした積もりです。ところが技術内容がよくても、特許そのものの書き方が上手でない、という場合がしばしばありました。それは一人称で書くからです。

具体例（仮想）を説明します。最近は大抵のことはデジタル処理されます。音声もそうです。従って担当者はまずA/D変換器で音声をデジタル化するでしょう。そして“音声信号をデジタル化するA/D変換器と・・・”と自分が実験でやった通りに請求範囲を書いてきます。そこで私が質問します「仮に君がライバル会社の担当で、この特許の逃げ道を必死に考えるという立場だったらどうだろう。君の請求範囲にはA/D変換器が構成要素に入っているけど、遠隔地から直接デジタル信号で入ってきてあとは同じ装置で処理するとうとういうことになるかな。A/D変換器は要らない訳だから君の特許は逃げられて仕舞うよ」と。「だから“デジタル化された音声信号を・・・”だけでいいんだよ。ついでに同じ原理で映像情報も処理できるなら、音声に限るのはまずいよ。音声にしか使えないかどうか検討してある？」というわけです。このように特許逃げ、特許つぶしという相手の立場で吟味すると不必要な制限事項が入っていたり、必要なものが抜けていたりすることに気がつきます。

私はよく「今日の午後は A T Rの社員を辞めて〇〇会社の社員になってご覧。そして必死でこの特許を逃げる方法を考えてご覧」と言ったものでした。そうすることで脇の締まったものになります。そしてより大切なことは、本当にオリジナルなことは何か、もっと一般化できないか、など自分の研究の本質がはっきり見えてくることです。エッセンスの抽出は本人がやるべきだと思っています。そしてそれは自信につながります。ですから、そういう自信がつくまでは私は本当は弁理士任せは反対です。ついでに言えば、こういうことが私のところでしか議論されない場合があるのは不思議なことでした（次回参照）。



④ 仮想変身の勧め

二人称で考えるということは、頭の中でいろんな人に変身し、その人の置かれている環境、条件を想像して仮想行動を試みることです。お金をかけないでいろんな人、例えば大統領にでもなれるのですから面白いことです。世の中では良かれと思ってやったことが反って相手を傷つけることもあります。これなども行動の前に相手に変身してみれば未然に防げることが多いのではないのでしょうか。

別の例を話します。A T Rには外国人の研究者など来訪者が大勢見えます。私は（かな漢字の読めない）外国人になった積もりで英語（ローマ字）表記だけを頼りに関西国際空港から公共交通機関で A T Rまで来るシミュレーションをやってみました。難物の一つは近鉄西大寺駅での乗り換えです。大阪難波行きと京都市行きが同じホームから発着します。折角難波から西大寺に来て京都行きに乗らないで難波に逆戻りするようなどことはないだろうか、果たして A T Rにたどり着けるだろうか。そういうことを近鉄の方にお話したことがあります。私のシミュレーション結果のことはここでは割愛します。

⑥ うまく使おう二人称

さて、この A T Rジャーナルもお客様つまりユーザは読者です。記事というプロダクトを読者というユーザにいかにかいい品質でかつタイムリーに提供するかという点では世の中の商品と何ら変わりません。

実は A T Rでは品質向上のため手近で実行可能な方策として、既に二人称を活用しているのです。私は学会誌の記事などでも、対象が大学院生レベルだとすると、二人称の立場にある大学院生とか同等の新入社員などに実際に原稿を読んで貰って具体的に意見を聞き、推敲するというをやってきました。どの雑誌でも大抵査読が行われますが、往々にして同じ分野の専門家が見ます。その結果、非専門家への初歩的な説明不足などがとかく見落とされ勝ちです。A T Rジャーナルの場合は専門的なチェックと合せて非専門家、具体的には事務系の社員などに読んで貰う仕組みにしています。そして、わからないことは遠慮なく指摘して貰うことで進めてきました。ですから、これらのたとえば経理部の若手社員は、その時だけは A T R社員ではなくて読者のひとりに変身して貰っているのです。そしてユーザとしての厳しい意見を出すのが務めなのです。遠慮は悪です。私はそれがいつまでも機能するよう願っています。



新鮮な経験

—「ええと」と「あの一」のニュアンス—

国際電信電話株式会社 研究所 主席研究員
(前ATR音声翻訳通信研究所 代表取締役社長)

山崎 泰弘



日常の会話で、言葉がうまく出てこなかったり、考え込むときに、繋ぎの言葉としてよく「ええと」とか「あの一」といった言葉が無意識に使います。しかしこの「ええと」と「あの一」にもニュアンスの違いがあり、それぞれ使い方があるようです。職場の上司から「例の件はどうなっているか」と問われ、咄嗟の返答に窮した場合、先ずどんな言葉を発しますか。迂闊にも「ええと」と応えると、「考えさえも未だ纏まっていけないのではないか」と上司に見透かされてしまいます。「あの一」の方が少しは罪が軽くなるかも知れません。

5年前(1993年春)ATR音声翻訳通信研究所に赴任して間もなく、「言語処理」に関する雑談をしていた時、上述のような「ええと」と「あの一」の使い方や違いが話題となりました。「言語処理」が専門でない私は、はてなと思い「ええと」と言って、暫くしてある考えを答えようとしたましたが適当な言葉が見つからず、「あの一」と言ってしまいました。その時「言語処理」の専門家でもある室長が「ええと」という言葉と「あの一」という言葉の使い方はそれでいいのですと助け舟(?)を出してくれました。訳も分からず、ぼかんとしていると、くだんの室長が「ええと」という言葉は、頭のなかで考えや答えを探し出そうとして思考し、検索している最中に発する言葉であり、「あの一」という言葉は考えがある程度まとまり、聞き手に伝わるために、どう表現しようか言葉を探している時によく発する言葉とのことでした。(詳しくは、音声文法研究会編「文法と音声」13章参照)

今まで全く意識せず、「ええと」と「あの一」を使っていたましたが、言われてみると確かに違いがあり、日本人は無意識に使い分けをしているようです。専門が違うとはいえ、今まで全く気が付かなかったことに大変びっくりし、内心、大変な研究所に来てしまったと観念してしまいました。新鮮な経験で、興味深いことではありましたが、日本語が気になりはじめ、少なくともその後2〜3日間、私の日本語の話す速さが普段の倍以上遅くなってしまったことを覚えています。

ATR音声翻訳通信研究所の研究目標は、日本語と外国語の話し言葉をコンピュータを用いて翻訳する技術を研究開発することです。前述の「ええと」と「あの一」に違いがあるとすれば、英語ではどう表現するのでしょうか。“Well”、“Let me see”などが咄嗟に頭に浮かびますが、その違いとなると定かではありません。そのうち英語を母国語にしている人に尋ねてみたいと思っています。もちろん現在の音声翻訳技術の研究にこのような訳し分けまでも追求する必要はありません。しかし話し言葉を対象に音声翻訳技術を研究開発する研究者は、人と人の会話に現れる言語現象や音声現象に対し、研ぎ澄まされた「感性」を常にもつことが大切です。「人に学ぶ」という基本理念はやはりATRのバックボーンなのです。深い専門的知識と広い視野を持った多くの研究者にATRで出会い、新鮮な経験を数多く積むことができました。私の貴重な財産です。

第58回（人間情報科学第48回） 97年11月10日

フランスにおける計算論的神経科学

Simon Thorpe 教授 Nicolas Franceschini 教授 Yves Burnod 教授

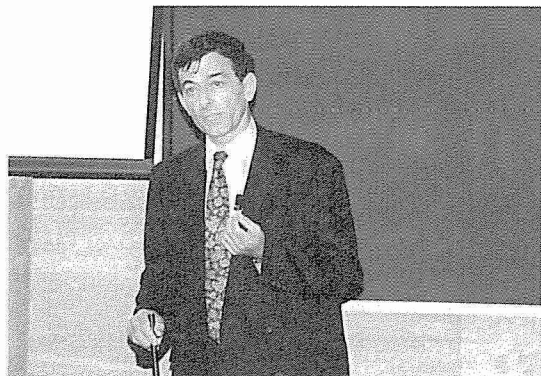
今回は、認知脳研究所のサイモン・ソープ教授、LNB3のニコラ・フランチェスキーニ教授、ピエール&マリー・キューリー大学のイブ・ブルーノ教授をお招きし、フランス国内におけるそれぞれの研究分野の最新研究動向・成果をご紹介頂きました。

ソープ教授には、迅速な視覚情報処理を可能にする非同期スパイク計算モデルについて、ビデオやOHPを交えて解説頂きました。

フランチェスキーニ教授には、ハエの対象物動き検出ニューロンについて、電気生理学的研究から、その結果を応用したロボットの障害物回避の研究に至るまでを幅広くご紹介頂きました。

ブルーノ教授には、fMRI（機能的核磁気共鳴画像装置）と計算論的神経科学を組み合わせた手法による、視覚運動プログラムを学習する大脳皮質における神経特性について様々なデータ等を交えてご紹介頂きました。

途中、参加者からの多数の質問があるなど、長時間にもかかわらず熱気にあふれた活発な討論が繰り広げられました。



Nicolas Franceschini 教授



Yves Burnod 教授

第59回（人間情報科学第48回） 97年12月1日

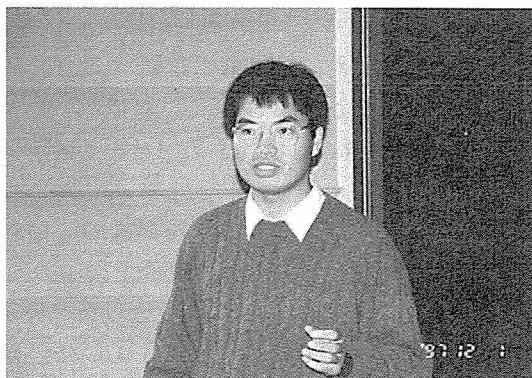
コンピュータビジョンのイメージ生成への応用に関する集中講義

Zhengyou Zhang 教授

ATR人間情報通信研究所にサバティカルを利用して1年間滞在された仏INRIAの研究員Zhengyou Zhang 教授に、ATRでの滞在研究成果の報告も兼ねて、

- (1) Multiple view analysis,
- (2) Photorealistic Modeling,
- (3) Facial image,

という3つの話題についての集中講義を行っていただきました。射影幾何学の数学的原理についての基礎的解説から出発し、3次元CADモデルによらない写実的な画像生成など、コンピュータビジョンのイメージ処理への様々な応用の可能性が豊富な実験例とともに示されました。約5時間に及ぶ長時間の集中講義でしたが、本分野に専門的に取り組んでいる研究者が多数聴講したため、最後まで活発な質疑応答が繰り広げられました。



Zhengyou Zhang 教授

－ V.R. EXPO '97に出展 －

(株) ATR知能映像通信研究所

(株) 国際電気通信基礎技術研究所

Virtual Reality Exposition 1997は、1992年、1994年に引き続き第三回目の開催で、10月9日～12日の日程で名古屋国際会議場において行われました。「V.R.が開く新しい世界」をテーマにV.R.デモンストレーション、V.R.国際シンポジウム、V.R.ワークショップが開催され、V.R.デモンストレーションに株式会社電通テック殿と共同出展でATR知能映像通信研究所より「Talking Eye」を出展しました。主催者の発表によるとExpo全体での入場者数は2万人を超えるなど大変盛況でした。今回初めて外部へ出展した「Talking Eye」は、CGの関係者や若いプログラマに関心を持たれ、「何に使えるのか」という質問も多くの方からいただきました。中には「スクリーン・セーバーに是非欲しい」、「家にもって帰りたい」といった意見もあるなど一般の方や子供達の反応が、今後の研究を進める上で大変参考となりました。



－ UNU/NTTマルチメディアと地球環境フォーラムに出展 －

(株) ATR知能映像通信研究所

(株) 国際電気通信基礎技術研究所

UNU/NTTマルチメディアと地球環境フォーラムは、マルチメディアの進展によって、今後は「物流」から「情流」へというパラダイムシフトを促し、地球環境に大きく貢献できるという基本コンセプトのもと、10月30、31日の2日間、東京大手町で開催されました。このフォーラムのコンセプトに対応して、ATR知能映像通信研究所の来場者体験型の二つのシステム「Talking Eye」と「フラクタルによる自然なCG作成技術」を出展しました。会場には、ホームゾーン、サテライトオフィスゾーン、個別展示ゾーンがあり各ゾーンにそれぞれ展示しました。来場者には研究者の方が多いこともあり、活発な議論が交されました。



Talking Eyeは、「映像」と「声」がホームゾーンの雰囲気作りに大いに貢献しました。「情報の伝達」主体のコミュニケーションに対し、「人と人とを繋ぐ」側面の重要性を強調しながら、従来の情報伝達、行為遂行的なコミュニケーションモデルの問題点を指摘しました。また、本研究の環境問題との関わりにおけるポイントをアピールし、大変好評を得ました。フラクタルによる自然なCG作成技術は、ホームゾーン、個別展示ゾーンに体験型の展示をしました。展示では、「情報環境」を保全・補完する、という問題意識について興味深く関心を持っていただきました。また、サテライトオフィスゾーンにおけるNTT入出力研究所殿の超高精細ディスプレイへの作品デモを併わせて行い、従来コンテンツとは一線を画したもののとの評価をいただきました。

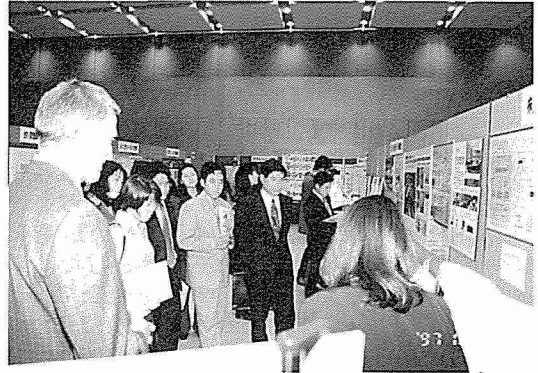
けいはんな産学官エキシビジョン97に参加して

ATR 音声翻訳通信研究所 第2研究室 Nick Campbell

「けいはんな産学官エキシビジョン97」が10月15日、16日の2日間、けいはんなプラザイベントホールで開催されました。

21世紀の未来都市づくりをめざす「けいはんな学研都市では」およそ50の研究機関・大学等に約25,000人の研究者、学生等が活動し約51,000人の人々が住んでいます。

この展示会は、都市内の研究機関の施設概要・研究内容の紹介と京都や、けいはんな学研都市での大学等の産学共同の取組みを紹介し、広く産学官や研究機関の相互交流、地域の方々との交流を図ることを目的に開催され、各研究機関・大学等から多数出展されました。ATRからは音声翻訳通信研究所、第2研究室の多言語音声合成システム「CHATR」を出展しデモを行なった。多くの方々が訪れ合成音声の自然性に驚かれる参加者も多く盛況でした。



国際シンポジウム Info-Tech'97へのATR成果の出展

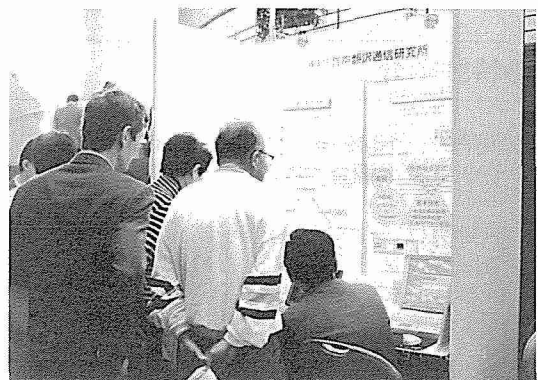
英語の聴き取り学習 / 多言語音声合成 / 多言語対話翻訳

ATR 音声翻訳通信研究所 第3研究室 飯田 仁

10月23日、24日にけいはんなプラザにおいて、通産省と京都府の後援で(財)関西情報センターが国際シンポジウムInfo-Tech'97を主催しました。同時に「インフォメーションテクノロジー紹介」というコーナーが併設され、関西学研都市を中心として誕生した最新のテクノロジーが紹介されました。ATRから、ATRヒアリングスクール、多言語音声合成CHATR、多言語対話翻訳Chat Translation Systemを出展しました。

Info-Tech'97の同展示コーナーでは、ATRのほか奈良先端科学技術大学院大学、NTT、(財)関西情報センターから技術紹介がありました。同大学院大学からは、利用が始まっている電子図書館の実演や、サッカーをするロボットの紹介、全方位の画像センサを使った実時間の画像提示装置の実演など盛りだくさんのデモンストラーションの展示がありました。(財)関西情報センターは電子決済などを含む電子商取引の各種実証実験を紹介し、NTTは日本語の株式市況速報を英語に翻訳するサービス実験を紹介しました。

ATRから紹介した3件の技術は、いずれもソフトウェアの販売実績があり、これまでも機会あるたびに実演を行なってきました。すでにパンフレットなども整備されており、参加者の方々から馴染みやすい技術内容であるという感想が聞かれました。豊富な話題で賑わった招待講演と共に、同コーナーでも活発な意見交換が行なわれ盛況でした。



第10回 研究発表会を好評開催

昨年11月6、7日の両日に第10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、ご好評を頂きました。

ATR研究発表会は、グループ各社の最新の研究成果及び動向をご紹介する場並びに研究成果のご活用をお考え頂く契機として毎年開催致しております。

今年は第2世代の研究所が揃ってから2年目であり、より一層充実した内容でお客様をお迎え致しました。

会場内は日頃からお世話になっている企業、大学、行政及び各種団体等の関係機関から、2日間にわたって延約1700名の方々のご来場を頂きました。

今回は、総括講演6件、テーマ発表7件、ポスターセッション78件とこれまでの最大級の規模となりました。また、前回は1日目が半日開催でしたが、今回は両日も全日開催とし、特に2日目は大学等の研究機関に広く公開致しました。開催当日は朝早くからお客様が来訪され、いずれの全会場も議論の花が咲き、熱気に包まれた2日間でした。

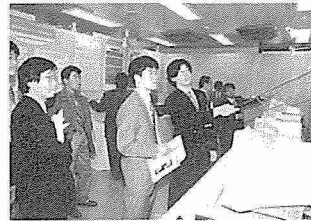
また、研究発表会の内容をより良くするため皆様にアンケートにご協力頂いておりますが、今回は約半数の方々からご回答を頂戴致しました。ご協力心から御礼申し上げます。その結果と致しましては大半の皆様方から概ね良好とのご意見を頂きました。しかし、改善案等貴重なご意見も頂戴致しておりますので、参考にさせて頂き、一層の改善を検討していきたいと考えております。



酒井副社長総括講演



テーマ発表の一例



ポスターセッションの一例

●特許紹介

特願平07-186915 No.2677534

マイクロ波半導体装置とその発振周波数を変化するための方法

本発明は、半導体超格子構造を有し、マイクロ波信号を発振するために好適なマイクロ波半導体装置とその発振周波数を変化するための方法を提供する。具体的には、2つの電極間に障壁層と量子井戸層とを交互に積層されている超格子構造を有する真性半導体i層を挟設したp i n型ダイオード素子を備えた装置で、隣接する障壁層のX点の準位と量子井戸層のΓ点の準位を互いに共鳴するようにし、かつ逆バイアス電圧を印加するようにした。また、光を入射するための光発生手段をさらに備えることにより、高周波発振を発生させることを可能とした。本発明により、極めて簡単な構成で、より広帯域な高周波発振周波数を変化させることができるマイクロ波半導体装置を提供すると共に、高周波可変発振器を実現できる。

特願平07-169623 No.2677532

可動プレート支持装置

本発明は、可動プレートを広い範囲で移動することができるだけでなく、任意の位置姿勢で固定することができ、しかも、小型・軽量化を実現した装置を提供する。具体的には、上部プレートを少なくとも3本の支持機構を用いて下部プレートに連結し、各支持機構はそれぞれ、2つのリンクとその両リンク間を回転可能に連結するための自由回転型関節で構成し、かつ負動作型電磁ディスクブレーキと回転型スプリングダンパーを設置している装置である。本発明により、例えば本可動プレート支持装置を2個用いて、可動式肘支持装置として使用することにより、脳神経外科手術に適用することができる。その際、ブレーキ解放スイッチをフットスイッチにすれば、術者は手術中の手先を使わず、足を使ってブレーキ解放スイッチを操作することにより、極めて容易な操作で手術中の肘の位置姿勢を変更しかつ固定することができる。

●受賞

★情報処理学会 マルチメディア通信と分散処理研究会／
モバイルコンピューティング研究会 Best Presentation賞（1997年7月4日）

受賞功績	受賞者	所属	内容
即興演奏における協調作業を支援する音楽CSCWシステムの提案	西本 一志	A T R 知能映像通信研究所 第二研究室 客員研究員	音楽演奏は高度な協調作業の一つである。本稿では、統計手法を用いて演奏データから演奏意図を抽出し提示することにより、特にジャズの即興演奏における演奏者同士の協調作業をリアルタイムに支援し、さらには音楽演奏上の創造性を触発する音楽CSCWシステム" MUSIC-AIDE" を提案した。

★日本神経回路学会 論文賞（1997年11月6日）

受賞功績	受賞者	所属	内容
A Kendama Learning Robot Based on Bi-directional Theory	宮本 弘之 五味 裕章 小池 康晴 和田 安弘 川人 光男	A T R 人間情報通信研究所 第三研究室 滞在研究員 第三研究室 研究員 第三研究室 研究員 第三研究室 研究員 第三研究室 室長	最適化原理に基づく運動パターン認識を用いて、見まねによる作業レベルのけん玉学習ロボットの数値実験結果について報告したもの

★大川情報通信基金 大川出版賞（1997年11月26日）

受賞功績	受賞者	所属	内容
脳の計算理論	川人 光男	A T R 人間情報通信研究所 第三研究室 室長	脳情報処理とコミュニケーションの機能を解き明かそうとする計算論的神経科学に関する基礎から最新の成果をまとめた専門書

★人工知能学会 研究奨励賞（1997年12月3日）

受賞功績	受賞者	所属	内容
個人の視点を伝え合うことによる協同発想支援	角 康之 西本 一志 間瀬 健二	A T R 知能映像通信研究所 第二研究室 研究員 第二研究室 客員研究員 第二研究室 室長	研究活動やビジネスにおいて日常的になされる電子メールやニュースシステムを介した対話に着目し、グループ内の協同発想や情報共有を支援する対話支援環境を構築した。特に本論文では、グループ内での相互理解を促進するために、各対話参加者の主観的な視点を定量化し、それらの関連を可視化する方法を提案した。
決定木学習による形態素解析	柏岡 秀紀 Ezra Black Stephen Eubank	A T R 音声翻訳通信研究所 第三研究室 研究員 第三研究室 客員研究員 第三研究室 客員研究員	形態素解析のための知識を辞書を用いずコーパスから得る手法を提案する。本手法は、決定木学習アルゴリズムにより、コーパスから形態素解析に有効な知識を学習する。決定木で用いられる属性は言語的な特徴やコーパスから得られる統計的な特徴を用いる。英語において、複数の品詞体系での品詞付与の実験について報告するとともに、日本語を対象とした場合の課題についても考察する。

●所員往来

平成9年10月1日より、12月31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
H 9 . 10 . 1	(音) 第4研究室 主任研究員	菅谷 史昭	K D D
H 9 . 10 . 1	(音) 第2研究室 研究員	樋口 宜男	K D D
H 9 . 10 . 1	(音) 第3研究室 研究技術員	鷹尾 和享	東洋情報システム
H 9 . 10 . 1	(音) 第3研究室 研究員	Julia Elisabeth Heine	ドイツ ザールラント州立大学
H 9 . 10 . 1	(音) 第3研究室 研究員	金城 由美子	九州大学文学部
H 9 . 10 . 1	(人) 第5研究室 研究員	蘆田 宏	京都大学
H 9 . 10 . 7	(音) 第2研究室 研究員	Wei Zhang	カナダ ウォータールー大学工学部
H 9 . 10 . 9	(国) 企画部 担当部長	伊藤 博	N T T アドバンステクノロジー
H 9 . 10 . 13	(音) 第1研究室 研究員	Petra Philips	ドイツ カールスルーエ大学
H 9 . 11 . 1	(環) 第3研究室 研究員	秋山 智浩	三菱電機
H 9 . 11 . 20	(音) 第3研究室 研究員	高木 一広	九州工業大学情報工学部
H 9 . 12 . 1	(音) 第1研究室 研究員	高野 優	日本電気
H 9 . 12 . 1	(音) 第3研究室 研究員	石川 開	日本電気
H 9 . 12 . 1	(音) 第2研究室 研究員	David Sann	フランス フランステレコムCNET

退職 年月日	復 帰 先	氏 名	A T R 所属
H 9 . 10 . 31	なし	Annika Berthold	(人) 第5研究室
H 9 . 11 . 28	フランス	Ming-Yue Xie-Zhang	(音) 第2研究室
H 9 . 11 . 30	Univ. of Glasgow	Frank Pollick	(人) 第3研究室
H 9 . 11 . 30	アルジェリア	Abdesselam Klouche-Djedid	(環) 第3研究室
H 9 . 12 . 8	フランス INRIA	Zhengyou Zhang	(人) 第2研究室
H 9 . 12 . 19	フランス ENST	Nicolas Pelletier	(音) 第1研究室
H 9 . 12 . 31	松下電器産業	脇田 由実	(音) 第3研究室
H 9 . 12 . 31	ジャストシステム	美馬 秀樹	(音) 第3研究室

A T R ジャーナル29号掲載記事の訂正

A T R ジャーナル29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きます。

A T R Monologue

頁 誤	正
10 山田伶子	山田玲子
受賞	
頁	
19 アーツ アンド サイエンス ファンデーション ロレアル賞	
誤	正
受賞者 所属	受賞者 所属
土佐 尚子 第六研究室 客員研究員	土佐 尚子 第五研究室 客員研究員

編集後記

A T R ジャーナルも30号を迎えました。B 5 版で半年に1回刊だったA T R ジャーナルがA 4 版の季刊誌に生まれ変わったのが第15号ですから、今号で新版が旧版を上回ったことになります。読みやすく内容のあるものを、ということをは心がけてはいるのですが、言うは易く行ふは難しで、正直言って季刊は事務方や執筆者にはきついです・・・。

そんな具合ですが、本誌に関する読者の皆様のご意見やご希望が関係者の励みにもなりますので、どんなことでも結構ですからコメントを頂戴したいところです。是非よろしく願いいたします。

(国際電気通信基礎技術研究所 秘書室長 阪田正和)

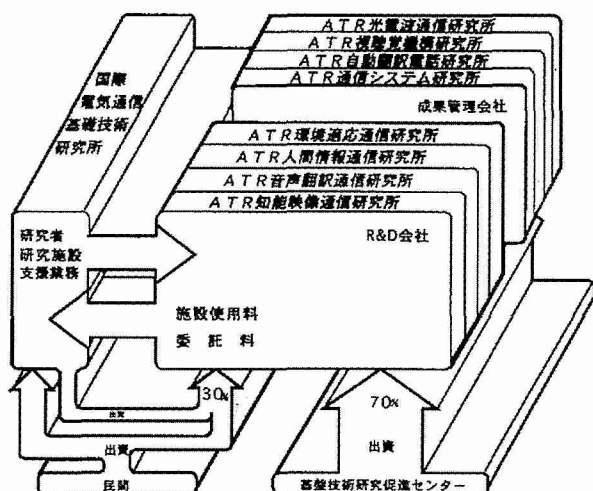
ATR Journal 第30号 1998年2月1日発行

- 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 Tel Wel 京都営業所
- 定 価 300円 (税込・送料別)

本紙記事の無断転載を禁じます。

© 1996 (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4つの研究会社（4 R & D会社）、既に研究を終了し成果の普及活動などを行っている4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9つの株式会社の総称です。

4 R & Dに研究費は基盤技術研究促進センターからの出資70%、民間約140社からの出資30%で構成されています。

国際電気通信基礎技術研究所は4 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4成果管理会社に対する研究成果の管理・販売などの各種支援を行っています。

ATRのWWWホームページのご案内

アドレス <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ジャーナル担当宛

TEL : (0774) 95 1177
FAX : (0774) 95 1178

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

☐ 送付先変更連絡

☐ テクニカルレポート購入申込

☐ 研究用ソフトウェア購入申込

【テクニカルレポート 番号：TR- - 】

【ソフトウェア名整理番号： 】

☐ ご意見、ご要望等

		変 更 後	変 更 前	変更事由
送 付 先	フリガナ お 名 前			☐ 人事異動 ☐ 住所変更 ☐ その他
	送 り 先			
	会 社 名			
	部 署 名			
	T E L			
	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。

