



## 「環境知能」プロジェクトの進展

NTTコミュニケーション科学基礎研究所

みなみ やすひろ どうさか こうじ もり あきら まえだ えいさく  
南 泰浩 / 堂坂 浩二 / 森 啓 / 前田 英作

NTTコミュニケーション科学基礎研究所では、「環境知能」をテーマとした研究プロジェクトを進めています。このプロジェクトの目的は、音声処理、音響処理、言語処理、対話、視覚情報処理、探索、学習、ネットワークのコミュニケーションのための情報処理技術を有機的に統合することであり、それによって実現される新たな生活様式の提案も視野に入れていきます。本誌2005年11月号で、プロジェクトの立ち上げから昨年6月のオープンハウス2005までの取り組みについて紹介しました。ここでは、環境知能プロジェクトのその後の進展について報告します。

### これまでの経緯

NTTコミュニケーション科学基礎研究所では、昨年の本誌11月号で紹介<sup>(1)</sup>したように、「環境知能」をテーマとした研究プロジェクトを進めています。このプロジェクトの目的は、音声処理、音響処理、言語処理、対話、視覚情報処理、探索、学習、ネットワークなどのコミュニケーションのための情報処理技術を有機的に統合することであり、それによって実現される新たな生活様式の提案も視野に入れていきます。このプロジェクトを推進するにあたり、研究開発だけでなく、コンセプトの策定・デザイン・広報活動までも包括的、かつ戦略的に進めることに留意しています。

環境知能の基本コンセプトイメージを進化する妖精・妖怪と重ね合わせ、昨年6月のオープンハウスでは、そのコンセプトに基づいた環境知能のデモンストレーションを実施しました。また『まっしゅるーむの世界2005』と題する小冊子<sup>(2)・(3)</sup>を作成、配布し、情報系基礎研究としてのコミュニケーション科学が目

指すべき方向性について、提案を行いました。

その後、環境知能に関するさまざまな問題点について議論を重ね、目指すべき将来像を提示するとともに、その実現に向けて取り組むべき課題について整理し、一部を「妖精・妖怪の復権——新しい環境知能像の提案」としてまとめました<sup>(4)</sup>。そして、本年6月に開催した、研究所の新たな公開イベントである「未来想論2006」において、環境知能のデモンストレーション「まっしゅるーむの世界2006」を実施しました<sup>(5)</sup>。

ここでは、オープンハウス2005から現在までの環境知能プロジェクトの進展について紹介します。

### 環境知能とは

環境知能プロジェクトでは、「環境知能」「妖精・妖怪」「まっしゅるーむ」という3つの言葉を使っています。研究すべき新しい分野を技術用語として表現したのが環境知能、環境知能によって実現される世界をメタファ（隠喩）として特徴的に表現したのが妖精・妖怪の世

界、そして、試みにつくった妖精・妖怪のニックネームがまっしゅるーむです。まっしゅるーむの世界は、現在の情報処理技術を使って環境知能の1つの形態を実現したものです。まずはじめに、その根幹となる環境知能について説明します。

私たちは、環境知能の世界観を「たくさんいる」「あまり主張しない」「呼べば答える」「隠れている」などの言葉で表現しました。これには、利便性と合理性ばかりが強調されてきたこれまでの技術のあり方を問い直したいという思いを込めています。そしてさらに、環境知能は、ちょっとした人との触れ合いや小さな発見の中に、やすらぎ・驚き・愉しみを感じ、ささやかな幸福を得ているのと同様なことを、コミュニケーション科学の力を使って実現しようという挑戦的な取り組みでもあります。

このような言葉で表わされる環境知能の世界には、実は妖精・妖怪の世界と数多くの類似点があることに気がきます。古今東西の妖精・妖怪に共通した役割は、身近にいつも寄り添い、見守

り、そっと支えてくれるというものでした。ですから、環境知能とは、妖精・妖怪を復権させるための技術であるといえます。そこには、論理的思考はもちろんのこと、感情や感性、コミュニケーション能力など人の知性をかたちづくる高度な知的活動がかかわってきます。

妖精・妖怪の世界は、ただ単に情報処理技術を統合しただけでは決して実現できるものではありません。妖精・妖怪がかつての人間社会において果たしてきた役割に着目し、人の持つ精神性や知性に対して情報処理技術の側から近づくことが必要です。これこそが環境知能の大きな目的です。

この目的を実現するために、私たちは、最新の技術で妖精・妖怪の棲む世界の一例を毎年つくっています。これがまっしゅるーむの世界です。まっしゅるーむは技術の進歩に伴い進化するので、まっしゅるーむの進化の系統樹は研究開発のロードマップそのものとなることをねらっています。つまり、この系統樹を見れば、環境知能を実現するためにどんな技術がすでに実現されていて、将来どんな技術を実現すべきなのかということが一目で分かるようにしていきたいと考えています。

## 環境知能の将来像

環境知能が目指す将来像をより具体的なかたちで提示するために、生活シーンのいくつかを情景シナリオとして書き起こしてみました<sup>(5)</sup>。現在の技術で実現できることをオープンハウス2005、および未来想論2006の展示、まっしゅるーむの世界2005/2006などで示したわけですが、それらと環境知能の理想像との間には、もちろん大きな落差があります。

しかし、目指すべき情景シナリオのいくつかを共有しておくことで、各技術分

野におけるマクロな視点からの方向性を合わせておくことができ、また情景シナリオから解決すべき具体的な技術課題が見えてくることもあります。

ここでは、これらの情景の中からまっしゅるーむの世界2006に関連が深い2つの情景について述べます。

### ■情景1：会議室

複数の人が1つの部屋の中で会議をしていて、その中に混じっていくつかのまっしゅるーむがいます。

まっしゅるーむA：人に気付いて、話の内容から議題が何であるかを知ります。

まっしゅるーむB：「…って何だっけ」といった人の発言に反応して、世界各地のまっしゅるーむから情報を集め答えやヒントを提示します。

まっしゅるーむC、D：ある話題について議論を始めます。人はそれを興味深く聞いています。

この情景において、会議におけるまっしゅるーむの第1の役割は環境理解です。誰が何をいつ話したかを、画像、音声などありとあらゆる情報から理解し、分析します。第2の役割は、物知り博士です。世界中のまっしゅるーむから情報を瞬時に得て提示します。第3の役割は、発想の支援者です。人の記憶力や推論能力では思いつかないような事実をさりげなく提示します。

### ■情景2：孫

お祖父さんとお祖母さんとは離れて住んでいる孫の家です。

まっしゅるーむA：孫がバッハのヴァイオリン協奏曲を聴いていることを認識します。

まっしゅるーむB：孫に対して、お祖父さん、お祖母さんもその曲が好きだったことを教え、彼らがその曲を楽しんでいた様子を過去の映像から再生します。

まっしゅるーむがいることにより、長

い年月を超えて家族の間で記憶が共有されていきます。孫が暮らす家に棲むまっしゅるーむは、有名なアメリカ民謡に出てくる「大きな古時計」のように、家族の記憶を担っています。何げなくかけた音楽から、孫は、まっしゅるーむを通して、自分が孤立した存在ではなく、家族の長い歴史の中に位置付けられる存在であることに気がきます。

## まっしゅるーむの世界2006

まっしゅるーむの世界2006では、3つのデモンストレーションを実施しました。これらのデモンストレーションのシーンと具体的なシステム構成を図1に示します。まっしゅるーむの世界2006では、昨年から進んだ新しい研究成果を取り入れており、まっしゅるーむは図2（色つきの部分）のように進化しています。

### ■展示A：「見守る・気付く・話を聞いて答える」

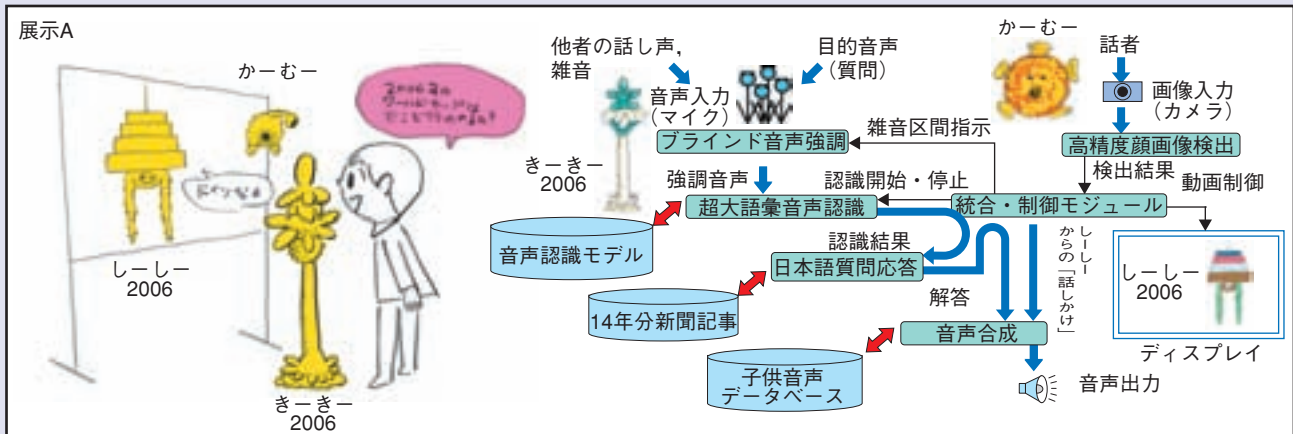
展示Aは、情景1においてまっしゅるーむが人に気づき、人の質問に答えるというところを実現しています。

まず、「かーむー」が、人の話したげな様子に気がきます。そして、「シーレー」が話しかけます。人が「2006ワールドカップはどこで行われるの?」と質問すると、シーレーは「ドイツ」と答えます。

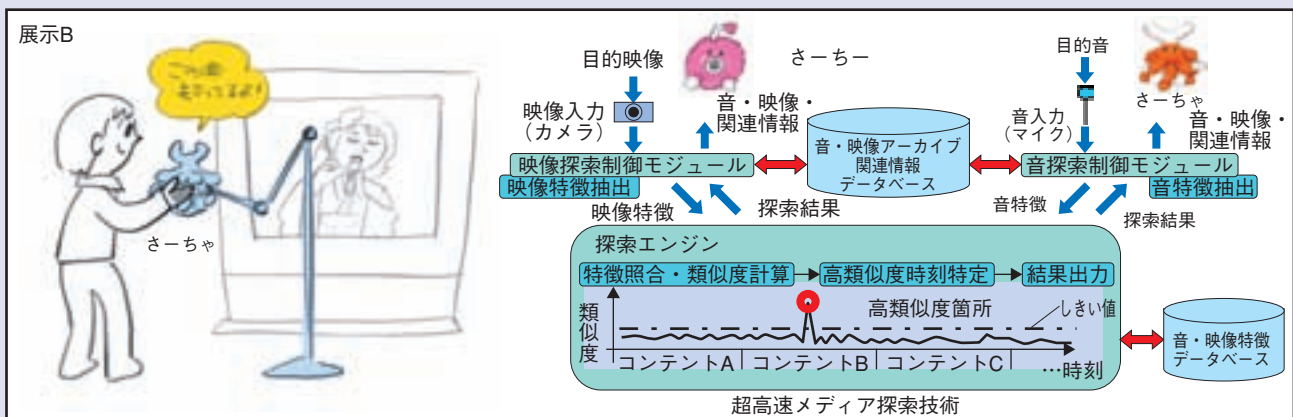
「気付く」ことには高精度顔画像検出を用いています。ブラインド音声強調技術<sup>(6)</sup>を用いて人の声を聞き取ります。超大語彙音声認識技術<sup>(7)</sup>を使って人の質問を理解します。そして、日本語質問応答技術<sup>(8)</sup>を使って、14年分の新聞記事からその質問の答えを探し出し、合成音声<sup>(9)</sup>で答えます。

### ■展示B：「見る・聞く・探す」

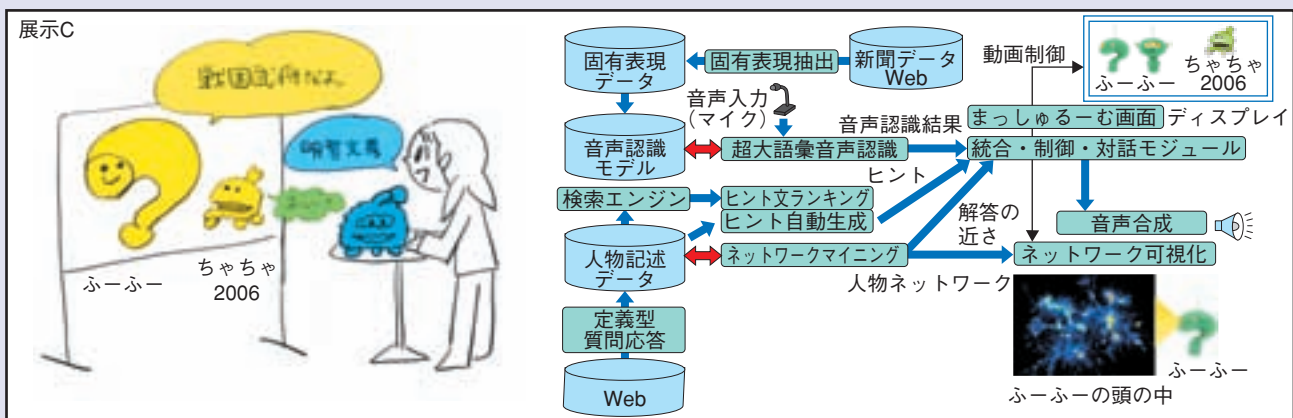
展示Bは、情景2においてまっしゅるーむが音楽を聞いて、過去の映像を探



(a) 見守る・気付く・話を聞いて答える



(b) 見る・聞く・探す



(c) 隠れていて、さりげなく、なぞなぞをする

→ : 制御フロー  
→ : 信号フロー  
↔ : データ参照

図1 デモンストレーションとシステム構成

し出すところを実現しています。

「さーちー」と「さーちゃ」は、周りの様子をずっと見て聞いています。そして、人がさーちーとさーちゃにTVを見せたり聞かせたりすると、その部分を含む過去の映像を探し出してくれます。

蓄積された膨大な映像・音特徴データベースから所望の映像を探してくる超高速メディア探索技術を使っています。

### ■展示C：「隠れていて、さりげなく、なぞなぞをする」

展示Cは、情景1においてまっしゅるーむが人の知らない知識をそれとなく提示するところを実現しています。「ふーふー」と「ちゃちゃ」は隠れてい

て、突然なぞなぞ遊びを始めます。私たちは、なぞなぞのやり取りを見ているだけですが、知らず知らずのうちに賢くなります。さらに、ふーふーはなぞなぞで人と対話することもできます。

定義型質問応答技術とヒント自動生成技術により、ある人物に関するヒントを生成します。ちゃちゃや人がそのヒントから人名を答えると、ネットワークマイニング技術<sup>(9)</sup>でその人名がどれくらい正解に近いかという情報を調べます。その情報を音声合成とディスプレイでちゃちゃや人に提示します。

まっしゅるーむの世界2006は、昨年の展示で利用した超大語彙音声認識技

術、日本語質問応答技術をさらに高性能化すると同時に、次の7つの要素技術を新たに開発、利用しました。

- ① 高精度顔画像検出技術：二段階に分けて処理を行います。一段階目で濃淡の差を利用して画像の大半の顔以外の領域を取り除き、二段階目でより詳細な検出処理を実行しています。
- ② 画像情報を利用するブラインド音声強調：ブラインド音声強調は周囲の雑音やほかの人の話し声を抑圧します。今回は、これに加え、画像情報を使って人の顔が見えないときに、雑音を抑制する処理も行ってい

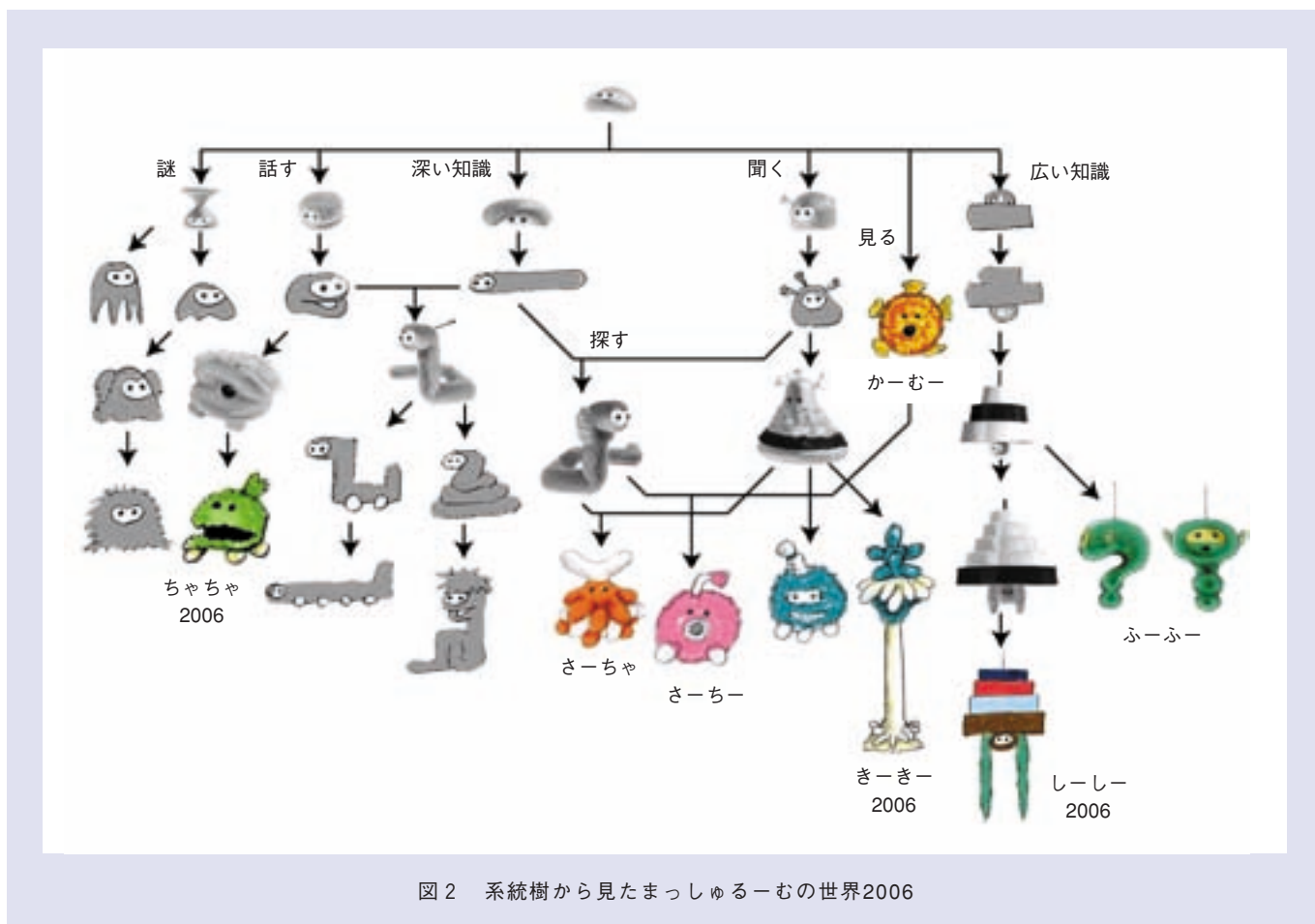


図2 系統樹から見たまっしゅるーむの世界2006



- ます。
- ③ 超高速メディア探索：1年分の映像や音の中から、必要な部分を高速に見つけ出します。
  - ④ 定義型質問応答技術：大量の文書集合から、ある特定の人と事物に関する説明をしている文を収集します。
  - ⑤ ヒント自動生成技術：集めた情報がどれくらいよく知られている情報か、またその人を特定するのにどれくらい役立つ情報か、などの観点からヒントを点数付けています。この点数の順序によってヒントを生成します。
  - ⑥ ネットワークマイニング技術：人名についての知識を、ネットワーク構造に変換します。そのネットワーク上で距離を測り、人の答えと正解との近さを調べます。
  - ⑦ 音声合成：子供をイメージした音声合成音を実現しました。

## 今後の取り組み

まっしゅるーむの世界2006では環境知能で提案する2つの生活シーンを目標として、その一部を現在の情報処理技術で実現しました。しかし、理想とする生活シーンを実現するためには、まっしゅるーむのより一層の進化が必要です。また環境知能という言葉も非常に大きな概念であり、環境知能にかかわる問題は、科学技術全般から、人類、社会にまでおよぶものと考えています。そこで、より広い視点から環境知能を論じるために、下記のシンポジウムを開催することにしました。できるだけ多くの方々のご来場をお待ちしています。

### 「環境知能シンポジウム2006」開催のお知らせ

9月に環境知能シンポジウム2006を開催します。このシンポジウムでは、環境知能というキーワードを軸に、日本社会あるいは世界のあるべき将来像と、それに向けて私たちが今何をなすべきかについて、広くかつ深く議論したいと考えています。詳細は下記のホームページなどを通じてお知らせいたします。

日時：2006年9月22日（金）

13：30～17：30

場所：原宿クエストホール

司会：外村佳伸（NTT）

パネリスト

： 東 浩紀（哲学者・批評家）

石黒 浩（大阪大学）

竹内郁雄（東京大学）

前田英作（NTT）

ビデオ出演

： 下條信輔

（カリフォルニア工科大）

URL：http://www.kecl.ntt.co.jp/

KCS2006/

### ■参考文献

- (1) 前田・南：“「環境知能」の実現に向けて,” NTT技術ジャーナル, Vol.17, No.11, pp.52-55, 2005.
- (2) NTTコミュニケーション科学基礎研究所編：“まっしゅるーむの世界2005,” 私家版, 2005.
- (3) http://www.brl.ntt.co.jp/cs/kikang/index-j.html
- (4) 前田・南・堂坂：“妖精・妖怪の復権—新しい「環境知能」像の提案—,” 情報処理, Vol.47, No.6, pp.624-640, 2006.
- (5) NTTコミュニケーション科学基礎研究所編：“まっしゅるーむの世界2006,” 私家版, 2006.
- (6) 中村・南・McDermott：“次世代の音声認識技術,” NTT技術ジャーナル, Vol.15, No.12, pp.13-18, 2003.
- (7) 前田・磯崎・佐々木・賀沢・平尾・鈴木：“質問応答システム：SAIQA,” NTT R&D, Vol.52, No.2, pp.122-133, 2003.
- (8) T. Kurozumi, H. Nagano, and K. Kashino：“特許出願PCT/JP2005/012667,” (WO/2006/006528), 2006.
- (9) 水野・磯貝・長谷部・浅野・阿部：“コーパスベースアプローチによるテキストからの音声合成,” NTT技術ジャーナル, Vol.16, No.1, pp.23-26, 2004.
- (10) 澤田・向井・荒木・牧野：“多音源に対する周波数領域ブラインド音源分離,” 人工知能学会研究会資料, JSAI Technical Report,

SIG-Challenge-0522, AIチャレンジ研究会第22回, pp.17-22, 2005.

- (11) 齊藤・山田・風間：“k-dense 法によるネットワークのコア部抽出,” 第2回 ネットワーク生態学シンポジウム, 2006.

(参考文献(2), (5)の私家版冊子, および(1), (4)の別刷りには残部があります。ご希望の方はご連絡ください。)



(左から) 南 泰浩/ 堂坂 浩二/  
森 啓/ 前田 英作

音声処理, 音響処理, 言語処理, 対話, 視覚情報処理, 探索, 学習など人間の知能や知性に関係のある新しい技術が次々と提案されています。ところが、これらの技術を融合化し、人間の社会生活にどう役立てていくかについては、まだまだ十分な議論がなされていません。私たちは環境知能というコンセプトに基づいて、人の生活を豊かにする知能や知性の実現を目指しています。

### ◆問い合わせ先

NTTコミュニケーション科学基礎研究所  
メディア情報研究部 環境知能研究グループ  
TEL 0774-93-5323  
FAX 0774-93-5305  
E-mail minami@cslab.kecl.ntt.co.jp