

情報バリアフリーとVR

一聴覚障害者のコミュニケーション支援技術一

井野秀一

東京大学



1. はじめに

現代の医学ではどうしても治すことのできないために、手足や感覚に障害が残ったままになっている人々は全国に約325万人おり、そのうち聴覚に障害を持つ人々は約30万人と推定されている[1]。難聴は障害の部位により、伝音性難聴、感音性難聴、混合性難聴に分けられる。感音性難聴のように補聴器の効果があまり期待できない難聴者の場合、手話や読話というコミュニケーション手段を習得しなければならないが、それらの習得は外国語の学習と同じように容易ではない。特に、音声によるコミュニケーションに慣れ親しんできた中途失聴者の場合には大変な努力を伴う。日本の場合、聴覚障害者で手話ができる人の割合は17%ほどである[2]。

また、私たちの身の回りでのコミュニケーションは音声に頼る場合が多い。そのため、耳の不自由な人たちは、手話を獲得しても、学校や職場などの日常生活の場面で社会的な情報障害を被りやすい。情報科学技術が急速に発達した昨今でも、携帯電話等による電子メールが広く普及したとはいえ、聴覚障害者のコミュニケーション支援技術の選択肢は決して多くない。さらに、社会の急速な高齢化に伴い、聞こえに衰えのある老人性難聴者も確

実に増加の一途をたどっている。

聴覚障害者のための情報バリアフリー、そして、高齢になっても余生を豊かに過ごすためのコミュニケーション手段の確保は、現代における重要課題の一つである。

本稿では、まず、コミュニケーション支援技術と密接な関係にある感覚代行について、人間情報処理の視点から説明する。そして、VRを応用した聴覚障害者の情報バリアフリー技術については、コミュニケーション環境の違いに着目し、私たちの取り組んでいる研究例をいくつか挙げながら、具体的に紹介する。

2. 感覚代行

VRを支える技術体系の一つに感覚ディスプレイがある。福祉工学においても感覚ディスプレイは大切であり、その研究は感覚代行に関わる分野のなかで進められている[3]。ここでは、VR技術から生み出された感覚ディスプレイを聴覚障害者のコミュニケーション支援装置に応用する際に考えておきたいことについて触れてみたい。

2.1 伝達すべき感覚情報

社会生活を送る上で必要な情報は、図1のように言語

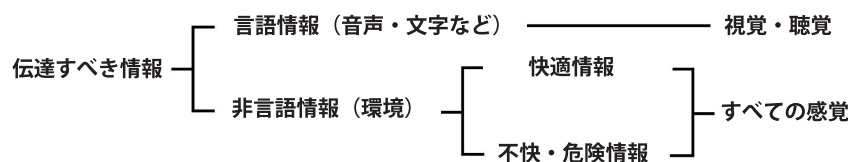


図1 私たちの社会生活に必要な情報と感覚器の関係

情報と非言語情報に大別される。言語情報には音声や文字が含まれ、これらの受け手は主に聴覚と視覚である。非言語情報には環境に関する多様な情報が含まれている。快適や不快あるいは危険がそれである。前者には快い音楽や映像などがあり、視聴覚以外にも心地よい肌触りや爽やかな香りなどの触覚や嗅覚などに関する情報が含まれる。後者に関しては、痛覚はもちろんのこと腐った臭いや異様な味などがある。これらは、個体生存のために忘れてはならない情報である。しかし、非言語情報はあまりにも多様で把握が難しく、現在のところ、感覚代行の研究では、残存感覚を通して言語情報をいかに言語中枢に送り込むかに、その力点が置かれている。

2.2 感覚代行インタフェースの構築

聴覚障害者にとって、触覚と視覚が聴覚情報を代替して受け取る貴重な窓口になる。また、重複障害の盲ろう者にとって残された感覚とは、言うまでもなく触覚である。このような人たちの感覚代行ディスプレイを設計する上で留意すべき点を整理すると、次の3点に要約できる。

第1点は、本来の感覚と代用感覚における伝達情報量の違いについてである。これは、どんな情報をどこまで伝達するのかという問題と結びついている。第2点は、感覚受容器から中枢までの経路および感覚中枢における特徴抽出の違いについてである。これは、代用する感覚器で受容できるように変換した刺激と本来の音声との整合性をいかに確保するかということに関わっている。第3点は、代用感覚を経由した情報の場合、果たして本来と同様の感覚概念を惹起できるのかということである。これは、大脳における複雑な情報処理の問題であり、最も判断の難しい問題である。図2に聴覚代行に関わる感覚器と大脳の領域との関係を示す。

以上のように、感覚代行では、聴覚領野・視覚領野・触覚（体性感覚）領野などに、どのような刺激パターンとルートで情報を送るかということを慎重に考える必要

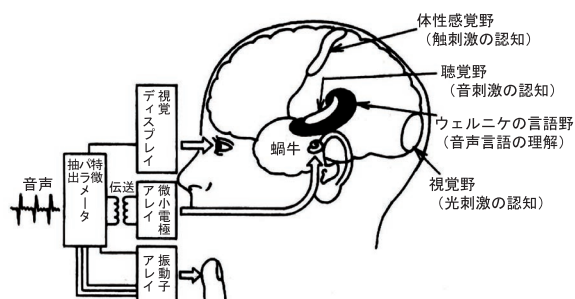


図2 聴覚障害者のための感覚代行の方法

がある。また、脳の可塑性・概念形成・学習などの高次中枢機能について考慮し、障害をもった時期の違い（先天性と後天性）や残存機能の状況にも留意すべきである。それと同時に、感覚代行装置を使いこなすための教育トレーニングプログラムの開発も求められる。

3. 聴覚障害者のためのバリアフリー技術

私たちが生活する日常のコミュニケーション環境には、家族や友人同士での会話が主であるパーソナルな環境と役所や病院あるいは学校などのパブリックな環境がある。ここでは、コミュニケーション環境の違いに応じて、聴覚障害者のための情報バリアフリー技術について述べる。

3.1 パーソナルなコミュニケーション環境

パーソナルなコミュニケーション環境で使う装置は、いつでもどこでも利用できることが前提になるため、小型・軽量かつ携帯可能であることが求められる。また、一歩進んで、気軽に身につけるという観点からはウェアラブルであることが望ましい。これは、ウェアラブル性の確保というVRと共通した技術テーマを内包していることになる。そこで、私たちの研究室で開発に取り組んでいる、多チャンネルの触覚ディスプレイを利用した手のひらサイズの触知ボコーダと透過型HMDを利用したウェアラブル自動筆談装置について紹介する。

【触知ボコーダ】

触知ボコーダとは、ヒトの耳の機能を電子回路で置き換え、指先の振動で音声ができるようにするための「指で聴く」装置である[4]。いわゆる、タクタイルエイド（Vibro-Tactile Aids）の一種である。この基本構成を図3に示す。

まず、マイクロホンから入った音声は、外耳道の音響フィルタの役目を果たすイコライザを通る。外耳は一種のハイパスフィルタであり、音声帯域の音を強める働きを持っている。その回路を通った信号は、内耳（聴覚系）の働きをする臨界帯域フィルタで16分割される。16個

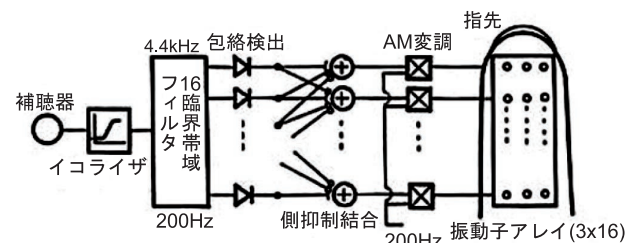


図3 触知ボコーダの基本構成

の臨界帯域フィルタからの出力波形は包絡検出器を通過し、側抑制回路に送られる。ヒトの感覚情報処理メカニズムと同様に、この抑制回路は空間的なシャープニングの処理を行い、ホルマントや子音の弱い成分を強調する働きをもつ。さらに、その出力は触覚のパチニ小体が敏感に反応する 200Hz の振動に AM 変調され、縦 16 列 × 横 3 行で 48 本配列されている触覚ディスプレイで皮膚表面に展開される。高い音を含む音声の場合はディスプレイの上部ピンが振動し、低い音の場合は下部が振動する。例えば、/a/ の音声の場合には真ん中あたりが強く振動し、/i/ の場合には両端に分かれる。すなわち、音声の周波数的な特徴パターン(ホルマント周波数)を時々刻々と指腹面に呈示し、点字を読むように指先で音を感じとらせる仕組みである。

この研究過程では、札幌の聾学校の協力を得て、1 年間ほど聾児たちの音声認識訓練に実際に利用した。そのときの経過は NHK で「指で聴いたアイウエオ」というドキュメンタリー番組に取り上げられ、1970 年代の感覚代行研究を知るための貴重な映像資料になっている。また、この装置は、肩掛け式で携帯できるようにポータブル化して製品化された。しかし、価格などの問題で普及するまでにはいたらなかった。

その後も、「触知ボコーダ」の改良研究は続き、/ku/ や /tsu/ などの短くてすぐ消えてしまう子音を指先で知覚できるように工夫を加えた。これは、触覚の振動パターンを電光掲示板のように指先で動くようにスワイプする方式であり、麻雀の盲牌がヒントとなっている [5]。また、生体力学的な観点から、振動刺激を効率よく、かつ確実に触覚受容器へ伝達できる触覚ディスプレイの構造設計のために、指先皮膚の詳細な機械インピーダンス特性も調べている [6]。

現在の「触知ボコーダ」は「タッチボイス」という名称で、図 4 のような新しい手のひらタイプの大きさとなっている。また、触覚ディスプレイ部は、本体からの分離と USB 制御を可能にしてあり、視覚障害者の情報

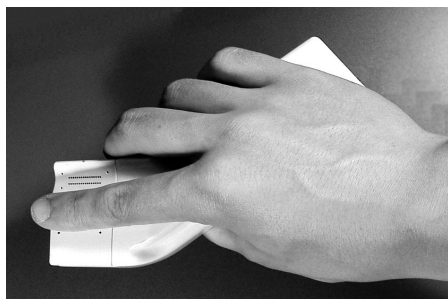


図4 手のひらサイズの触知ボコーダ

アクセスのための触覚ジョグダイヤル(例えば、インターネットのホームページなどのリッチテキスト情報の呈示手段)としての応用研究も行っている [7]。

【ウェアラブル自動筆談装置】

触知ボコーダを中途失聴の人たちに使用してもらったところ、指先で触ったときの振動パターンの違いは分かるが、それは、直接、音声として聞こえないといって、落胆する人たちがいた。既に、彼らは文字という概念を獲得しているので、日常生活では筆談を使っている。特に、高齢になってからの失聴では、新たに手話や読話を覚えるのは難しい。もし、音声情報を文字に自在に変換することができれば、声を文字として聞き取れる(読み取れる)はずである。そこで、音声認識技術を利用したビジブルスピーチ (Visible Speech) としての聴覚補助装置の開発に取り組むこととした。

その先導研究は、触知ボコーダの時間スペクトルパターンの出力を利用して、聴覚での情報処理機構を踏まえながらコンピュータで音声を文字に置き換えるという、1980 年代の音声タイプライタの開発である [8]。これは、「ア」「カ」「サ」「タ」「ナ」・・・「ン」と 1 音ずつ発声した 68 個の単音節音声をカナ文字に変換するシンプルな自動筆談装置であった。しかし、当時のエレクトロニクス技術では、まだまだコンピュータや視覚ディスプレイが大きく、小型化は難しかった。ウェアラブルにするには、これらの装置の小型化技術と歩調を合わせる必要があった。

さて、ここ数年間、複合現実感 (MR) の分野において、CG と現実映像を臨場感あふれる画質と速さで融合させて呈示する技術開発が盛んに進められている。メガネくらいの大きさの透過型 HMD も製品化されるようになってきた。さらに、ウェアラブル PC の開発も進んでいる。

このような背景のなかで、改めて、図 5 に示すように、音声認識装置と透過型 HMD を組み合わせ、認識された

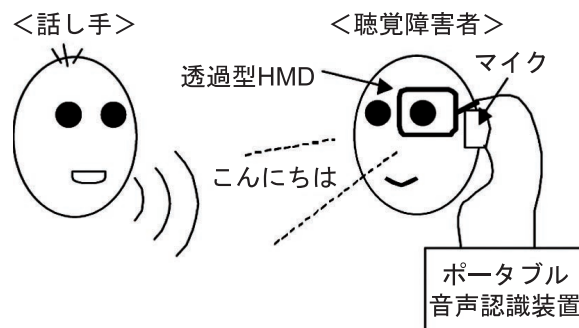


図5 透過型HMDを利用した自動筆談装置の概念



図6 ウェアラブル自動筆談装置の試作器

音声が対話の相手の口元あたりに文字としてマンガの吹き出しのように現れるコンセプトの聴覚障害者用コミュニケーション装置の開発をはじめた。このような方式だと、話者の表情や口の動きも文字と同時に見えるため、完全な音声認識装置を必要としない。音声認識の誤りが機械側にあっても、利用者側の推論能力の助けを借りて、それを、ある程度、訂正することが可能になるからである。この場合の音声認識装置は、人間と機械の仲立ちをするマン・マシン・インタフェースではなく、聴覚障害者と健聴者をつなぐマン・マン・インタフェースである。

そこで、図6に示すようなポータブル PC とメガネ式の軽量の透過型 HMD を利用したウェアラブル自動筆談装置を試作した。本装置の基本構成は、携帯型音声認識装置の本体にポータブル PC を利用し、その音声認識エンジンには日本 IBM 社の ViaVoice を利用した。また、ViaVoice の認識結果の出力部分には、音声認識の確からしさ（尤度）の閾値によって出力文字の表示形態（認識結果の間違って確率が高い場合には「***」とするなど）が調節可能になるように、独自の改良を加えた。透過型 HMD には、メガネのフレームに取り付けられる Clip-On Display を利用した。総重量は約 540g であり、携帯向けの装置として利用可能な範囲の重さであった。また、音声認識エンジンにより音声を文字列に変換して、話者の喉元に呈示するまでの時間はおよそ 1 秒前後であった。若干の遅れがあるものの、人間の会話に間があることを考慮すれば、コミュニケーション補助装置としてのリアルタイム性は概ね確保されていると言える。音声認識による文字は、横書き・1 行表示で話者の喉元に表示されるが、不完全だと予測された認識文字は「***」のように伏せて、同音異義語や意味不明な文字列が利用者を惑わさない仕組みにした。なお、これら

の表示系に関わるパラメータの設定値は、不完全な音声認識結果による文の意味理解や呈示位置や行数の違いによる文理解精度の心理実験の結果に基づいて決めた[9]。

3.2 パブリックなコミュニケーション環境

多くの人たちが同時に利用する公共施設等での情報バリアフリー機器は、パーソナル向けの装置とは異なる開発コンセプトが必要になる。すなわち、情報伝達時に、曖昧性が極力少なく、誰もが利用できるユニバーサルデザインの発想を持つことが求められる。

【音声認識自動字幕システム】

聴覚障害者がパブリックな環境で利用できる情報保証の手段に、手話通訳・ノートテイク・要約筆記（パソコン要約筆記）がある。しかし、手話通訳は、手話を身につけていないと利用することができず、ノートテイクや要約筆記の場合は、手作業で文字を作成するために、情報呈示までに時間がかかってしまう、あるいは、全ての内容を伝えることができないという状況が生じる。そこで、音声を入力手段として用いることができれば、入力速度の限界という問題が大幅に解消され、内容を要約することもなく、全ての情報を聴覚障害者に伝えることが可能となる。このような背景のもとで、私たちはウェアラブル自動筆談装置を改良して、公共施設での会議や学校の授業等での情報保証を目的とした音声字幕変換システムを開発することにした。

システムの構成図を図7に示す。本システムでは、講演者と音声認識装置の間に同時復唱者を配置し、その復

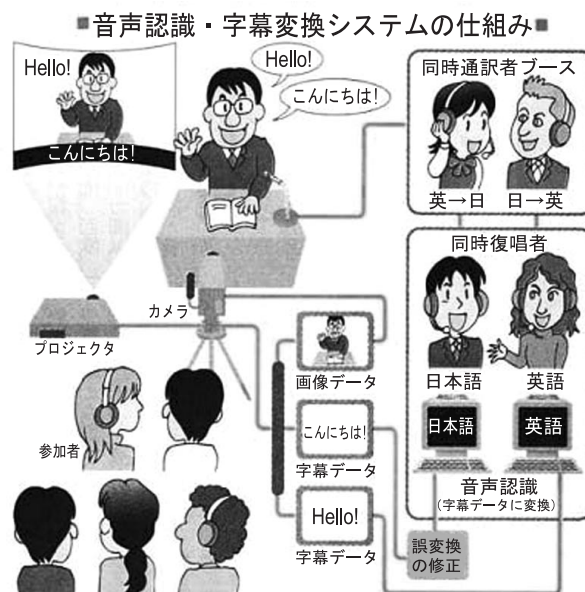


図7 音声認識自動字幕システムの基本構成



図8 DPI2002 会場での字幕スクリーンの様子

唱者が講演者の話した内容を復唱して音声認識装置へ入力する方式を採用した。この方法により、不特定話者の音声を、復唱者、すなわち特定話者の音声に変換することで、認識率の優れる特定話者認識の技術を利用できることになる。また、国際会議にも対応できるように、英語の音声認識エンジンを組み入れた。復唱者は、日本語は民放アナウンサー、英語は北米英語を母語とするネイティブスピーカーとした。同時復唱者は、講演者の話す言語に応じて、講演者または同時通訳者からの音声を復唱する。このようにして、同時復唱された音声はそれぞれの音声認識装置で字幕データに変換される。同時復唱音声による音声字幕変換の正答率は約92%（復唱者は民放アナウンサー）であった。なお、字幕データのより一層の正確性の観点から、本システムと連動して誤変換部分を外部オペレータがキーボード入力で迅速に修正できるインタフェースも開発した [10]。

字幕データはミキサーに送られ、カメラからの講演者の映像に重ね合わせて会場のスクリーンに映される（図

8）。このとき、講演者の表情と字幕にズレがあると違和感を利用者に与えることになる。音声字幕になるまでの時間は、講演者→（同時通訳者）→同時復唱者→音声認識装置のルートを経由するために、4～5秒くらいである。そのため、画像情報と字幕情報の時間的ずれを吸収する画像遅延呈示装置（BUG製：VideoBOX）を製作して、重ね合わせの時間補正に利用した。

本システムを障害者インターナショナル世界会議札幌大会（DPI2002）で運用したところ、高い精度（約96%）で字幕情報を会議参加者に提供することができた。大会参加者からの評価も高かった [11]。また、このシステムをネットワーク対応型に改良して、横浜での国際ユニバーサルデザイン会議（UD2002）における字幕提供を、会場から遠く離れた札幌のオペレーションルームで行った [12]。このようなシステムにすることで、復唱者や修正オペレータが一つの場所に集まることなく、本システムがあたかも会場にあるかのごとくに運用できることを確認した。

このような遠隔字幕システムをさらに発展させていくと、例えば、歩行機能に障害のある人たちや地方に住んでいる人たちが、同時復唱者として居場所を問わずに参加することができ、それが新しい雇用創出につながる可能性がある。また、字幕のディスプレイを携帯電話やPDAにすると、利用者はスクリーンの場所に左右されずに会場の何処にいてもよいことになる。教育分野では、東京大学バリアフリー支援準備室の聴覚障害を持つ大学生への支援事業の一環として試験運用されている（図9） [13]。

また、この字幕システムやウェアラブル自動筆談装置は、話し手の映像を現実世界、字幕映像を仮想世界とするならば、一種の複合現実感の環境となり、これらのシ

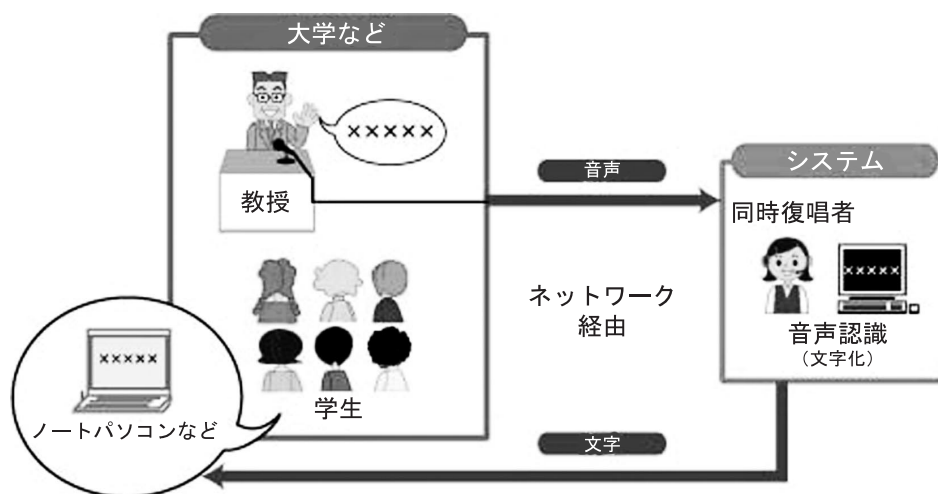


図9 遠隔操作タイプの自動字幕システム（授業参加支援の場合）

ステムは MR 技術や高臨場感ディスプレイの発展に応じて一層使い易いものになっていくであろう。

4. まとめ

聴覚障害者の情報バリアフリー技術と VR の係わりについて、感覚代行におけるインタフェース設計の視点を交えながら、タクトイルエイドとしての触知ボコーダ、ビジブルスピーチとしての自動筆談装置、そして会議参加支援のための自動字幕システムについて紹介した。これらは、人間と機械の最適な関係はどうあるべきかを第一に考えなければならないものであり、その過程には、心理学や生理学を基礎とする知見と実験、ユーザからの評価、そして、目的に応じた機能を持つ感覚ディスプレイの開発が求められる。

また、聴覚障害者の情報バリアフリー機器は、ちょっと見方を変えると、例えば、騒音で音声の聞き取りにくい作業環境でのコミュニケーションにも役立つ。つまり、誰にとっても使い得のあるユニバーサルデザイン指向の機器になる。同じことは、他の障害者支援技術にもあてはまる。そして、人間との関係性が強い科学技術である VR は、必然的に人間支援技術であり、これらの障害者支援技術と底流で結びついているのではないだろうか。

今後、VR が、聴覚に障害のある人たちの自立生活をサポートし、社会参加を円滑にするアシスティブテクノロジー (Assistive Technology) の一翼を担いつつ、それらが互いに共鳴しながら発展していくことを期待したい。

参考文献

- [1] 厚生労働省統計表データベース (厚生統計要覧), <http://www.dbtk.mhlw.go.jp/toukei/youran/index-kousei.html>
- [2] 岩淵紀雄: 聴覚障害者用コミュニケーション関連用具, OT ジャーナル, Vol.36, No.6, pp.677-683 (2002)
- [3] 井野秀一: 人工現実感の評価 - VR の生理・心理・社会的影響 - (伊福部達監修), (1.4「福祉への VR 技術応用の現状と展望 - 概説 -」), 培風館, pp.42-62 (2000)
- [4] 伊福部達, 湊博, 吉本千禎: 心理物理実験によるタクトイルボコーダの基礎的研究, 日本音響学会誌, Vol.31, No.3, pp.170-178 (1975)
- [5] 和田親宗, 井野秀一, 伊福部達: 高度難聴者用タクトイルボコーダにおける音声スペクトルのスイープ呈示方式の提案と評価, 電子情報通信学会論文誌 A, Vol. J78-A, No.3, pp.305-313 (1995)
- [6] T. Homma, S. Ino, T. Izumi, H. Kuroki and T. Ifukube: Measurement of Mechanical Characteristics of a Fingerprint Surface in the Design of a Tactile Display, Journal of Robotics and Mechatronics, Vol.15, No.2, pp.153-163 (2003)
- [7] C. Asakawa, H. Takagi, S. Ino and T. Ifukube: Auditory and Tactile Interfaces for Representing the Visual Effects on the Web, Proceedings of the 5th International ACM SIGCAPH Conference on Assistive Technologies, pp.65-72 (2002)
- [8] 伊福部達: 音声タイプライタの設計, CQ 出版 (1983)
- [9] 井野秀一, 小坂井敦, 奈良博之, 服部裕之, 伊福部達: 透過型 HMD と音声認識技術を利用した聴覚補助方式の提案, 日本バーチャルリアリティ学会論文誌, Vol.6, No.3, pp.177-184 (2001)
- [10] 加藤士雄, 井野秀一, 永井謙芝, 渡邊括行, 堀耕太郎, 似鳥寧信, 高杉弘子, 服部裕之, 伊福部達: 国際会議における聴覚障害者支援を目的とした音声字幕変換システムの設計, ヒューマンインタフェース学会研究報告集, Vol.4, No.4, pp.65-70 (2002)
- [11] <http://www.mainichi.co.jp/digital/solution/archive/200210/15/5.html>
- [12] <http://www.bug.co.jp/topics/ud2002.html>
- [13] <http://www.bug.co.jp/topics/todai.html>

【略歴】

井野 秀一 (INO Shuichi)

東京大学 先端科学技術研究センター 助教授

1993 年北海道大学大学院博士後期課程修了, 1991 年日本学術振興会特別研究員 (DC), 1993 年北海道大学電子科学研究所助手, 1995 年同大講師, 2003 年より現職。工学博士。専門は、生体工学・福祉工学・ロボット工学。人間情報システム論に基づくヒューマンインタフェースの研究開発に従事。電子情報通信学会, 日本機械学会, バイオメカニズム学会, IEEE などの会員。著書 (共著) 『人工現実感の評価』 (培風館), 『健康・福祉工学ガイドブック』 (工業調査会), 『新版ロボット工学ハンドブック』 (コロナ社) など。