

インタラクションに基づく体験共有コミュニケーション

間 瀬 健 二^{†1,†2} 萩 田 紀 博^{†2,†3} 角 康 之^{†2,†4}
 小 暮 潔^{†5} 片 桐 恭 弘^{†6} 伊 藤 禎 宣^{†5,†7}
 岩 澤 昭 一郎^{†3} 鳥 山 朋 二^{†5} 土 川 仁^{†8}

数々のメディア技術の発展の歴史をみるときに、個人の体験を記録として残し、自己を振り返ったり他者と共有したりしたいという欲求は古来から人類に備わったものであることが分かる。超高速ユビキタスネットワークを前提としたウェアラブルコンピューティングやユビキタスコンピューティング技術の発展により、個人やグループの体験や行動を自動的に記録することが可能となりつつある。我々は、ユビキタス体験メディアと総称する、インタラクションを手がかりとするウェアラブル体験記録デバイスや体験共有デバイスを開発し、体験記録を実践し、体験日記や体験共有のアプリケーションの効果を確認する実験をすすめている。本稿では、ユビキタス体験メディアのとしての体験記録デバイス、コーパス型データベース構築技術、体験日記応用などの概略を解説する。

Experience Communication Based on Interactions Captured by Ubiquitous Experience Media

KENJI MASE,^{†1,†2} NORIHIRO HAGITA,^{†2,†3} YASUYUKI SUMI,^{†2,†4}
 KIYOSHI KOGURE,^{†5} YASUHIRO KATAGIRI,^{†6} SADANORI ITO,^{†5,†7}
 SHOICHIRO IWASAWA,^{†3} TOMOJI TORIYAMA^{†5}
 and MEGUMU TSUCHIKAWA^{†8}

Human desire of recording personal and group experience has a long history since ancient period, which is discerned by the development of numerous kinds of media that is used to record and share experiences. A ubiquitous computing environment will become a solid social infrastructure for recording human experiences in the real world and facilitating human communication. The recorded activity can be used as a good source of novel communication for sharing experiences on the infrastructure. We have developed wearable and ubiquitous experience recording systems and experience sharing systems which we call the ubiquitous experience media. This paper focuses on the non-cumbersome recording system, a corpus-based experience database system and abstracting and summarizing application in the form of experience diary.

1. はじめに

人類は長い歴史のなかで、紙、ペン、写真、ビデオといったさまざまなメディア技術を発明し、メディアを使って体験したことや思考を表現し人と交換したり共有したりして、社会活動を行ってきた。メディア技術は、自分の行動や体験を記録して伝えることによって、他者がそれを利用できるという利便性がある。

一方、インターネットと Web の発展と、それらの経済活動や生活への浸透により、大量の情報や知識が地球規模で入手・共有・発信できるような情報基盤が構築されつつある。次には、多くの人が多様な情報や知識を流通しあう異分野や異文化のコミュニケーションを可能にするグローバル・メディアが望まれる時代

†1 名古屋大学情報連携基盤センター

Information Technology Center, Nagoya University

†2 ATR 知能ロボティクス研究所

ATR Intelligent Robotics and Communication Laboratories

†3 ATR メディア情報科学研究所

ATR Media Information Science Laboratories

†4 京都大学大学院情報学研究科

Graduate School of Informatics, Kyoto University

†5 ATR 知識科学研究所

ATR Knowledge Science Laboratories

†6 公立はこだて未来大学

Future University Hakodate

†7 東京農工大学

Tokyo University of Agriculture and Technology

†8 日本電信電話株式会社

Nippon Telegraph and Telephone Corporation

に突入すると予測される。情報革命を通じた政治・経済・文化のグローバル化は、異文化の対立を先鋭化し、デジタル・デバイド、大衆の分衆化、世代間断絶の拡大などさまざまな問題をも引き起こしている。グローバル・メディアの時代には、世代、地域、職業、文化、社会を超えて、情報を入手・共有・発信できるコミュニケーション形態を支えられるような、革新的なメディア技術が必要不可欠である。

既存のメディア技術を使っても、大規模超高速のインターネットを背景に、Web や SNS (Social Networking Services) にみられるように、多様な分野にまたがったコミュニケーションが進展していくことであろう。しかしながら、分野や文化が異なると、教科書的に文字や写真で知識を表現しただけでは、相手にはなかなか伝わらない。知識の 1 方向の伝達は表面的な理解をもたらすが、感動や共感にはつながらないからである。コミュニケーションによって相手に伝えたい感動を伝えたり、コミュニケーションを通して育まれる相手との共感が伝達できなければ、異分野の知識の流通は困難である。これはヒューマノイドロボットが生活シーンに現れる未来社会において、人間とロボットという異質な主体が共生する場合にも人間の認知や思考とロボットの認知や思考のすりあわせによる円滑なコミュニケーションのためにも必要となる。

知識の流通をつなぐ鍵は、電話や電子メールのやりとりのような気軽な双方向コミュニケーションの上に、経験や体験のような具体例を、感動とともに伝えることである。他者の体験の再現と共有によって、他者の感情や感動の理解がなされ、異なる背景・文化を持つ人々の理解を促進する。ここで、体験を双方向に伝えあうことが可能な「体験共有コミュニケーション」を実現するメディア技術が重要となってくる。

また、体験の共有は教科書レベルの知識の伝達を超えて、「わざ」や「目利き力」の伝承・習得・創出など、深いレベルでのコミュニケーションを通じた人々の協調的な創造力発揮にとっても鍵となる。たとえば、個人の活動の暗黙的な知識やルールを意識下におくことで、生活を改善できる道を開くことができる。さらに、いろいろな場所や場面で個人や組織が体験したことを共有することで、価値判断の材料を相互に提供しあうことが可能となる。すなわち、経済活動や社会活動のウェアラネスを適度に提供することで、コンプライアンスを高め、安心で安全な社会の構築に貢献できると思われる。

そこで我々は、インターネット・Web を介して体験を伝えられる新しいインタラクション・メディア技

術の 1 つとして、「ユビキタス体験メディア」を提案し、その実現にむけた研究開発を行ってきた。具体的には、体験を記録する装置としてのウェアラブル・ユビキタスデバイス、体験を演出するヒューマノイドロボット、記録した体験データを構造化して蓄積するインタラクション・コーパス技術、記録した体験データを検索しブラウジングする技術、体験データを要約して日記としてまとめる技術、さらに、体験データから新たな体験を生み出す体験共有デバイスの研究開発を進めてきた。本稿では、これらの成果をまとめて概要を報告する。

2. 体験共有コミュニケーション

我々は実世界における各個人の体験を適切な方法で記録し、蓄積し、共有する枠組みを開発し提供することで「体験共有コミュニケーション」を実現することにした。その足がかりとして、「インタラクション・コーパス」と呼ぶ、人の体験や行動に関わる、人と人、人とモノの相互作用を体系的に記録した事典を構築することを提案する。

このような事典やコーパスについては、すでに、人工知能の分野においては Cyc¹⁾、Wordnet²⁾ など言語的な常識 (Common Sense) を体系的に整備する研究プロジェクトが進められ、また、音声認識・合成の分野では音声認識率の向上や合成品質の向上のために大量の音声コーパスデータが収集されている。これらの例にならえば、実世界の相互作用の常識や行動パターンの体系的収録こそが、知的なユビキタス情報環境の実現に必要な不可欠のものである。いったん、コーパスの形式で体験を蓄積すると、特定の体験を検索し再現 (合成) することは容易となり、それを使って体験共有コミュニケーションが可能となる。

体験共有コミュニケーションの実現において、我々はインタラクション・コーパスを中心に置き、4 つの観点からアプローチした。第 1 のアプローチは、個人や集団の体験の収集技術の追究である。実世界の自然な人間行動から人間同士の相互作用や人間とモノの相互作用を観察して、体系化して収録しなければ、コーパスとしての価値がない。実世界の行動を阻害せず、現実的な時間と手間でコーパス化できる観測・収集技術が必要である。第 2 のアプローチは、体験データベースの構築技術である。収集した行動をどのような形式で構造化・蓄積すれば、コンピュータが理解でき (machine readable)、効率的・効果的に検索・再利用できるか明らかにする必要がある。信号処理やメディア認識技術と、人間の相互作用・社会活動理解の橋渡

しをする接点であり、両者の知見と統合が不可欠である。第3は、体験データベースの利用技術である。記録した体験をもとに、記憶想起支援、体験要約、体験日記、体験物語による体験共有のほか、行動計画支援、行動予測による情報個人化などのサービスが考えられる。ここでは、ある種の追体験のための特別な体験デバイスも必要となる。最後は、体験共有コミュニケーションの社会的環境整備である。実世界の個人の行動や体験を記録し利活用することは、その便利さの反面、プライバシー保護やセキュリティ保全の点から社会的なインパクトが強い。コストとメリットのバランスの問題を考えつつ、社会に受容される技術や方策の検討が必要である。本稿では、技術的側面に焦点を当てることにして最初の3つのアプローチについて具体例を示して議論を深める。

3. 背景・関連技術

Bush は、第二次世界大戦が終わる 1945 年の時点で、Memex という全文検索やテキスト音声の注釈付けと情報のハイパーリンクに関するビジョンを提示した。そこでは、カメラを額に装着するなどして自分の見たものをすべて記録するという将来像を示している³⁾。すでに、さまざまなセンサがモバイル・ネットワーク化されることが現実化し、実世界での体験をいろいろなモダリティで、そのときにその場所で記録することが容易になってきている。たとえば、Gordon Bell らは Bush のビジョンを引用して、MyLifeBits というプロジェクトで Bell が読み書きした文書や電子メール、撮影した写真など個人の体験をすべて電子化するための技術開発と実践をしている⁴⁾。

体験を記録する研究はさかんであるが、体験共有まで踏み込んだ議論は少ない。体験記録については次のような研究がある。DARPA (Defense Advanced Research Projects Agency, 高等研究計画局) は 2003 年 5 月に “LifeLog” と呼ぶプログラムをスタートさせ研究テーマを公募した⁵⁾。LifeLog は人間が自然で簡単にインタラクションできるコンピュータを開発することをめざし、PDA (personal digital assistant) が将来は PDP (personal digital partner) になるとした。残念ながら、プライバシー問題を契機にいったん中止され、現在は ASSIST (Advanced Soldier Sensor Information System and Technology)⁶⁾ と呼ばれる軍のプロジェクトに引き継がれている。

ウェアラブル・コンピュータを使った体験記録には、心拍や湿度をとって主観的な体験記録をねらった上岡ら⁷⁾、脳波をとって心的状態と体験との相関を調べた

相澤ら⁸⁾の研究がある。Clarkson ら¹³⁾は、画像、音声、加速度計のデータから HMM を使い状況を監視し、教えた状況を検出する手法を提案している。エピソード記憶とその想起支援の研究には Lamming らの Forget-me-not¹⁰⁾がある。彼らはユビキタスコンピューティング環境の中で、携帯端末を使って個人の活動を記録するシステムを作り、その記録が想起の手がかりになることを示した。川嶋¹¹⁾および Kawamura ら¹²⁾は、画像処理により特定の人物や動作を検出し、エピソードとなる場面を記憶する方法を提案している。

人の行動を環境側から観察、センシングし、補助しようという研究は、坂村の TRON 電脳住宅をはじめとして、MIT の Smart Rooms¹⁴⁾、ジョージア工科大学の AwareHome、Microsoft 社の EasyLiving など国内外の大学や企業で多数のプロジェクトの報告がある。2005 年からスタートした、MIT の Roy らの Human Speechome プロジェクト⁹⁾は、Roy の自宅にカメラとマイクを多数設置し、幼児の 5 年間の生活を観察記録し、言語獲得のプロセスを解明しようとするものである。

これらに対して 2001 年度に開始した我々の研究プロジェクトでは、体験記録のデバイスやシステムを構築することにより、インタラクション・コーパスを構築する枠組み¹⁵⁾を提示している点が新しい。ウェアラブルセンサと環境センサをそれぞれ第 1 人称視点と第 3 者視点からの体験記録を実現するものとして、さらにパートナーロボットを導入して第 2 人称視点によるはたらきかけと体験記録をし、それらを統合してとらえてユビキタス体験メディアとしている。このはたらきかけは、後述するように体験を演出する機能として考えている¹⁶⁾。記録した体験を要約したり物語的な提示により、体験を日記化¹⁷⁾したり共有する新しい応用分野の開発も進めている。体験記録を人間にとっての人工記憶として扱う枠組みの研究¹⁸⁾も重要である。インタラクション・コーパスの構造の研究は、記憶の概念化やアクセス手段の計算論的なモデルを提示し、人工記憶を組織化するための新しい理論の提案につながる期待がある。

4. ユビキタス体験メディア (UEM)

まず体験記録のデバイスやシステムを構築して、体験共有のもとになるデータを取得する手段を開発した。その際、環境センサとウェアラブルセンサをそれぞれ第三者視点と一人称視点を提供するものとして用意することとし、それらを統合してとらえて、ユビキタス体験メディア (Ubiquitous Experience Media,

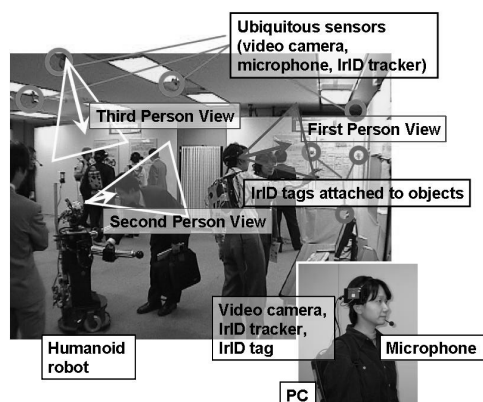


図 1 ユビキタス体験メディア実験室の設置状況と多重視点の融合
Fig.1 Setup of UEM sensor room and multiple viewpoint integration.

UEM) と呼ぶ。

まず展示会などにおける多数の展示者と訪問者のインタラクションを記録することをターゲットとした。ATR の研究発表会の機会を使って、システムをデモ展示し、外部からの見学者と発表者に実際に体験記録に協力してもらいデータを収集することを想定した。図 1 は展示室の様子である。図中のロボットは、二人称の視点を提供するパートナーである。ここで、ロボットは、センサだけでなく、アクチュエータも備えて、体験を協働で創る役目を果たす。UEM はこのようなロボットなどのアクチュエータも含み、ユーザとインタラクションするメディアとして拡張して考える。そこで我々は、このロボットのように二人称視点を有する UEM を協創パートナーと呼ぶことにした。

ユビキタス体験メディアとして、我々は、ビデオカメラ、マイクロフォン、赤外線 ID (Infrared ID, IrID) タグとトラッカ (図 2) からなるクライアントシステムを試作した (第 1 世代と呼ぶ)。IrID タグとトラッカは、人のインタラクションを理解・計測する際に、視線が手がかりになるという知見を利用している¹⁹⁾。IrID トラッカを使うことで、モノや人に装着された IrID タグがトラッカの視野に入れば、その位置と識別番号を手でできるようにした。ここで、位置とは、トラッカ中の画像座標であり、視野中の相対位置である。トラッカのタグ位置検出分解能は、装着光学レンズの画角と対象物距離によって決まる。タグの移動への追従は 400 fps での各撮像ごとに 2 pixel まで行う。このため、90° レンズ時に、3m の距離でのタグ水平移動への追従速度は 25 m/sec である。

空間中に存在する人物やモノの空間的な座標算出の手法がかりとして、設置された環境センサの位置と向き

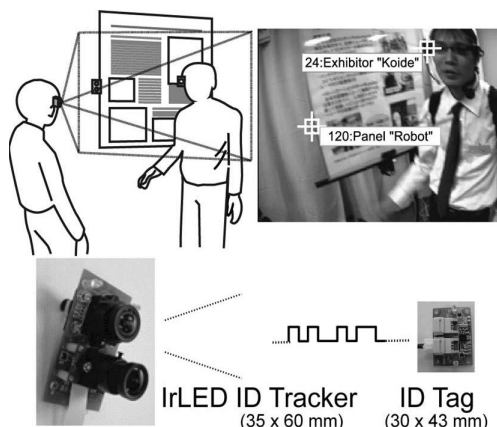


図 2 IrID トラッカと IrID タグによる視線の抽出 (赤外光で点滅するタグを IrID トラッカが検出、追跡する。トラッカの視野にタグが入れば、そちらに視線が向いていたことになる)
Fig.2 Gaze extraction by IrID tracker and Ir-LED ID tag.

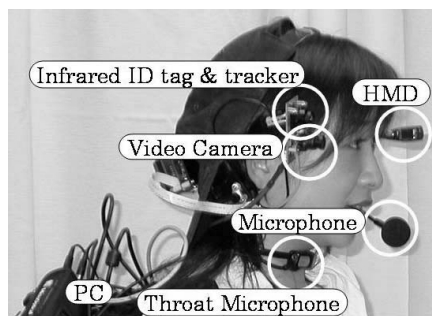


図 3 UEM 機能評価用ウェアラブルクライアント (第 2 世代、Head Mounted Display (HMD) や喉マイクなどが追加された)
Fig.3 Wearable sensors for evaluating functions of UEM (2nd Generation: an HMD and a throat microphone are attached).

が使えるが、本稿の範囲では、インタラクションは相対的なものとして考え、絶対的な位置の計算はとくに行っていない。

クライアントは、設置型と装着型があり、設置型はポスター会場の各場所に設置し、装着型は、ウェアラブルクライアントとしてユーザに着用してもらう。発表者のウェアラブルクライアントには、第 2 世代では喉マイクとヘッドマウントディスプレイ (HMD) を追加し、発話区間の直接センシングをしたり、環境からのフィードバックの提示効果を検証した (図 3)。

こうして、第 2 世代は、頭部に装着するヘッドセットは重量 500 グラム、バックパックで背負う PC は 2,720 グラムとなった。外部の見学者にもウェアラブルクライアントを貸し出し、体験記録に協力してもらった。しかしながら重さやケーブルの取り回しに対する

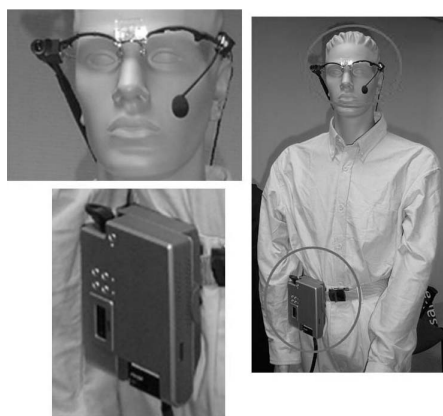


図4 ウェアラブルクライアントの軽量化（第3世代，専用のクライアントPCと超軽小型のヘッドセットカメラ部）

Fig.4 Lightweight wearable client (The third generation client with a capture PC and a very small head set).

不満の声が多かった．そこで，第3世代では，機能を体験記録に絞って軽量化をめざして設計をなおした．図4は軽量化したシステムである．ヘッドセットはカメラ部39g（ケーブルを除く），新たに開発した記録用モジュールは，145mm×113mm×26mm，245グラム（電池部はさらに300グラム）となった．ヘッドセットは，1つのレンズで集光し，ホットミラーを使ってビデオセンサとIDトラッカセンサに可視光と赤外光を分光して光軸合わせを実現している²³⁾．レンズの画角は第2世代までの60度に対し，90度で設計した．この設計値は，人がだれかと対話しているときに頭部の向きが相手とどのような位置関係にあるかをいろいろな場で測定して求めた²⁴⁾．たとえば座った状況では，眼球を動かすことによって，頭部の正中方向に対し左右各約35度の位置範囲の人と話をしていることが観察された．そこで，物体の幅を20度程度に見積もり，90度とした．また，CPUとしてMPEG-4エンコーダ/デコーダ内蔵システムLSIであるLC690132Aを使用することで，部品点数や消費電力を抑えるとともに，シリアルポートや汎用のA/Dポートなども確保した．取得した体験データの通信用には，CFタイプのIEEE 802.11bの無線LANアダプタを用いている．これで，第1世代に比べウェアラブルクライアントの使用条件をある程度改善できたが，依然，バッテリーとケーブルの問題が残っている．

5. インタラクション・コーパス

体験共有コミュニケーションの核となるインタラクション・コーパスについて述べる．図5に示すように，UEMで観測したデータをセグメンテーションし

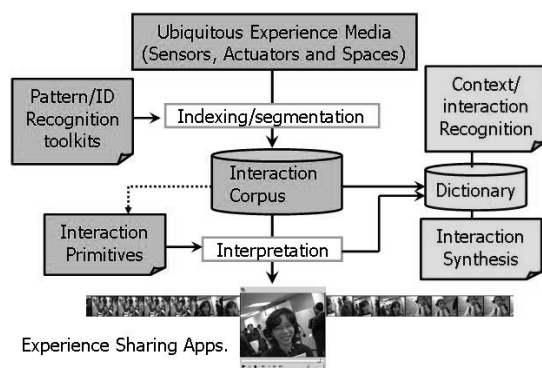


図5 インタラクション・コーパス（UEMで記録されたインタラクションは，インデックスをつけてコーパス化され，各種アプリケーションで利用される．さらに一般化されたコーパスはインタラクションの辞書として体系化される）

Fig.5 Interaction corpus (Recorded interactions are segmented and stored as a corpora. Interaction corpus is used for various applications).

て，インデックスを振り，それをもとにインタラクション・コーパスを構成する．いったんインタラクションがコーパス化されれば，それをインデックスを適当に翻訳してアプリケーションで利用したり，一般化によりインタラクションに関する辞書を構築し，認識や合成に利用できよう．

5.1 UEMとコンテキスト処理

体験の事象は，UEMを使って記録する．記録された事象や相互作用（インタラクション）は，対象やタスクに依存して自動的にあるいは半自動的にインデックスをつけて記録保管する．具体的にはIrIDタグとIrIDトラッカを使うことによって，自動的に（近接）位置情報を抽出する技術を使って，人間と対象物のインタラクションを記録できる．知的なビジョンシステムなどを使えば，物体の種別や人の氏名を認識するなども将来的に可能になると考えられるが，当面はIrIDを用いた識別技術を使って，コーパス構築を進める．これによってインタラクションをコンテキストを手がかりに分析できるので，コンテキストに基づくインタラクション・コーパスを整備するところから，アプローチした．会議などの，内容や意味がインタラクションの方法に影響を及ぼすものは，コンテンツ解析が必要になる．これについてはメディア処理のさらなる進歩が必要であり，将来の課題と考えている．まずコンテキストの理解を進めることが，コンテンツ処理技術の発展に寄与するという立場である．

5.2 コーパス構造のトップダウン設計

前述したATRの研究発表会における公開実験で記録されたインタラクションの中には，ポスターの見学，

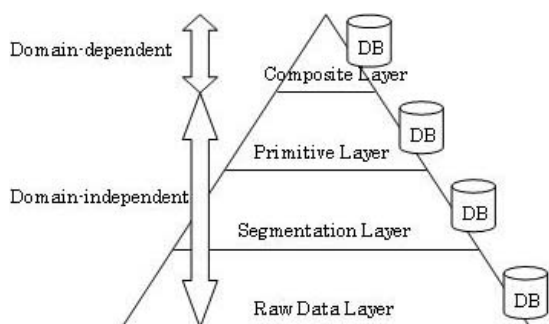


図 6 インタクション・コーパスの構造 (トップダウンに構成した 4 階層構造: 下から, 生データ層, セグメンテーション層, プリミティブ層, および複合層)

Fig. 6 Structure of interaction corpus (Four layer structure: From top, raw-data, segmentation, Primitive, and Composite layers).

離脱, 説明, 発表の聴取, モノへの注目, 他者との対話, 会議のほか, あたりを見回すなどの行動がみられた。これらはすべてデータベースサーバに時刻情報と一緒に図 6 のような階層構造で蓄積されている。

我々はインタクション・コーパスの構造化においてインタクション構造のプリミティブな単位として, 少なくとも発話 (SPEAK) と視認 (LOOK) の 2 種類の行動が取得可能であるという前提にたつ。そこで, これらのプリミティブを第 3 層とする, 4 階層モデルを提案する²⁷⁾。この階層化には, 言語の階層化のアナロジを援用して説明することができる。後述するように, 階層構造にすることによって, 多様なアプリケーション構築に非依存の汎用構造と依存したアプリケーション指向構造を分離して, 効率的な開発が可能である。センサ技術が進歩するにつれて, 接触, 姿勢, 位置など, プリミティブを増して, インタクションの表現力を高めることができる。

まず第 1 層 (図 6 の最下層) は, 生データ層と呼び, センサデータとタイムスタンプの組を格納している。これらはセンサから直接, 実時間で記録される。現在は IrID トラックの ID と座標, ビデオ, マイクのデータを別々に格納している。この層は音声でいえば, 音素にあたるようなものである。

第 2 層は, セグメンテーション層と呼ぶ。ID と発話のモダリティのデータを信号処理により, ある程度意味のある長さにまとめる。また, 短いセグメントはノイズとして消去される。音声では発話単位が切り出される。ID は同じ ID が規定時間長以上連続するとセグメントとされる。言語でいえば, ことばのレベルである。

第 3 層はプリミティブ層で, インタクションの単

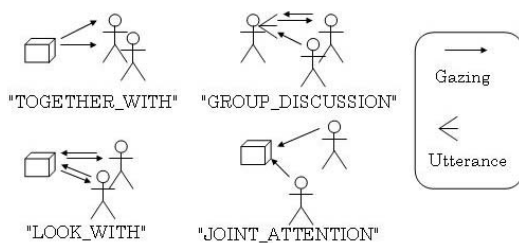


図 7 複合インタクションの例 (注視 (LOOK) と発話 (TALK) のプリミティブ組合せで構成)

Fig. 7 Interaction composites: compositions by LOOK and TALK primitives).

位としてのプリミティブを定義し格納する。行動の基本的な区間として定義する。1 つのインタクションの意味づけが行われる。これは, 言語の形態素解析を行うことに類推できる。たとえば, IrID トラックデータから注視状況をプリミティブとして抽出できるし, あるいは, マイクデータから発話状況を 1 つのプリミティブとして抽出することもできる。前者は, 主体と客体のある 2 者間インタクションのプリミティブであり, 後者は, 主体のみの自己インタクションの行動単位のプリミティブである。このように, マルチモーダルなセグメントを単体であるいは複合させて, 数々のプリミティブを形成できる。我々は, 視線と発話が人間同士あるいは人間—物体間のインタクションで非常に重要であると考えている。たとえば, ユーザ A の IrID トラックが別の人のタグを見ている状況のプリミティブを “LOOK” と定義する。また, トラックが物体を検出し同時に発話が検出されれば “TALK” と定義する。

最上位は複合層であり, 社会的なあるいはアプリケーションに依存してプリミティブを複合的に組み合わせることによって, 複合的なインタクションを記述する。この第 4 階層のデザインにおいては, トップダウンとボトムアップの 2 アプローチを考えている。トップダウンアプローチでは, プリミティブの組合せをアプリケーションのデザイナーがトップダウンに指定する。図 7 に示した複数人の主体があるような状況は, トップダウンにデザイナーが設計した例である。たとえば LOOK_WITH は, 2 名がポスターの前にいてポスターを眺めている状況である。それぞれの視野にポスターがあり, ポスターの環境クライアントからも両者をとらえている。この様子は, 図 8 のように抽出し表示できる。なお, このときには, GUI ベースのコーパスビューワを用いている。ビューワは, MySQL データベースの形式で格納しているインタクション・コーパスの Query インタフェースである。その画面



図 8 複合インタラクション LOOK_WITH の 1 シーン (ポスターのカメラ (下映像) が 2 名の来客をとらえ, 2 名の来客のカメラ (上中央と上右) がポスター (と説明者) をとらえている. 説明者のカメラ (左上) は中空を写している)

Fig. 8 An example of interaction composites: LOOK_WITH (Poster camera (bottom) captures two visitors, while two visitors' cameras (top center and top right) capture the poster and a presenter).

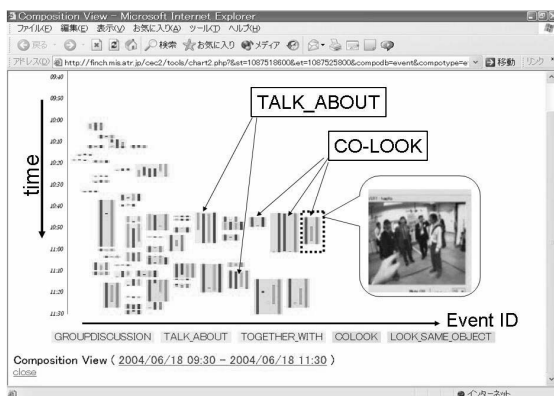


図 9 インタラクション・コーパスビューワの表示画面 [GUI ビューワにより, 時間 (縦軸) と ID (横軸) の 2 次元空間からマウスで色分けされた任意のイベントを選択し, ビデオを閲覧できる]

Fig. 9 Interaction corpus viewer [A GUI viewer helps users to select any event on 2D display. Chronological axes (vertical) and ID number axes (horizontal)].

例を図 9 に示す. 個々のインタラクションが線状オブジェクトで表現され, 複合インタラクションは複数をとりまとめた領域で表現されている²⁸⁾.

5.3 ボトムアップによるパターン抽出

ボトムアップアプローチでは, 第 3 階層のプリミティブの生成パターンからデータマイニング的手法により, アプリケーションドメインごとにインタラクション事象を効率的・効果的に記述できるものを抽出して定義する. 我々は, あるパターンの実際の生起回数を, ランダムに起こりうる生起確率で正規化した値を尺度と

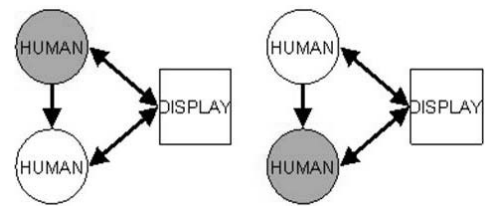


図 10 ボトムアップ抽出された複合インタラクション (CO_LOOK) (有向矢印は LOOK イベントを, ◦ は人を, □ はモノを表す. とくに • の色つき丸はその人が発話している TALK イベントであることを表す. たとえば右図はディスプレイの前に見学者 上 と説明者 下 があり, 説明者がポスターを見ながら説明している様子を, 見学者がディスプレイと説明者の両方を同時に見ている状況と解釈できる. 説明者と見学者の割当てが逆の場合もこれで記述される)

Fig. 10 Composites extracted by a bottom-up method (CO_LOOK, arrow: LOOK event, •: TALK event. Right figure is interpreted as a snapshot that a presenter is explaining poster looking at a display while a visitor is looking at both the presenter and the display).

して, その値が高い組を抽出する方法を提案した²⁵⁾. 図 10 は, この方法により抽出された一例で, 「ディスプレイ (ポスター) の内容の説明者 (左上) と見学者 (左下) の関係」をグラフ構造で示したものである. 前述の LOOK_WITH のスーパーセットとして, 人間同士の視線交換をともなうインタラクションが実際に頻繁に行われている主要なパターンとして抽出できた. このインタラクション (CO_LOOK と命名した) はトップダウンの設計時には思いつかなかったものである.

この事態は次のように考察される. トップダウン設計は, アプリケーションを前提に設計者の経験により主要と思われるものを用意する方法である. したがって, 目的志向でインタラクションの語彙が決められるが, それをもとに設計するため想定外のイベントが準備されないという欠点がある. 今回は, 体験を要約して日記を自動生成するアプリケーションを前提とし, ポスター見学シナリオの中で使われる語彙から複合インタラクションを設計した. その際, 見学者同士のインタラクションを中心に, 図 7 に示すような役割分担のないシナリオを前提に設計した. しかし実際には, ポスターの説明者と見学者という役割分担のある状況で利用されたため, それを的確に説明するインタラクションを見落としていたものである. このようにトップダウン設計においては, 設計者の語彙が複合インタラクションの記述能力を制限することがあるので, 注意が必要である. 今回の抽出例は, 実験室シナリオで, 組合せが小さい空間での探索例であったため, 効果は顕著とはいえないが, 多人数・多物体のインタラクションなどで, シナリオをすべて事前にチェックして, 複

雑なインタラクションもあらかじめ用意することは現実的でなく、このようなデータマイニング手法の有効性がでてくるものと考えられる。

6. 体験共有コミュニケーションの応用

体験共有コミュニケーションの応用から、ビデオ要約、ビデオ日記編集、体験ものがたり支援を紹介する。紙面の都合で詳細は割愛するが、これらのほかに、人間型ロボットなど自律的で物理的なエージェントを空間に放ち、社会的行為者・演出者 (social actor) として、主体的に人に近づきインタラクションをしながら人間の行動を記録する仕掛けも開発した。たとえば小学校や博物館にロボットを置き、ユーザとのインタラクションによって、発話や案内内容を変えるとどのようにユーザが反応するかを調べ、生徒らの人間関係のセンシングが可能であるなどの成果がある²²⁾。

6.1 体験の要約

インタラクション・コーパス構造を使って、特定の複合インタラクションの箇所を自動抽出して接続すると、容易にビデオ要約を作ることができる。会議録やメンテナンス作業記録など、さまざまな応用が考えられる。たとえば2時間の訪問・見学が5時間に要約されれば便利である。しかし、どの複合インタラクションを抽出するかは、その体験における利用者の意図や目的と関わる。

意図や目的から、個別インタラクションの重要度を計算する自動選択・選別のアルゴリズムを検討した。まず、滞在時間を重要度とする手法にヒントを得て、インタラクション時間の相対的長さを重要度とするアルゴリズムを作り、上位のインタラクション区間から各々15秒程度を抽出した。そして、1つの抽出区間における複数のクライアントで記録した体験は、素材を順に切り替えて再生提示するようにし、次の抽出区間へつなぐようにした。こうしてビデオ要約を自動生成したところ、それぞれの見学で、だれとどこで話したかというエピソード記憶を想起支援させるのに適した要約が生成された (図11)。利用者からは、それに満足する意見と、体験における対話の内容に立ち入った要約ではないという意見が寄せられた。この反応は、我々の初期の目的を満足するものである。対話の内容を要約するには、インタラクションの分析からさらにコンテンツ理解に踏み込んだ視点での研究開発が必要である。そこで我々は、意味レベルでの自動要約は現時点では困難と考え、利用者が主観的に体験日記を編集するのを支援する下記のシステムを試作した。

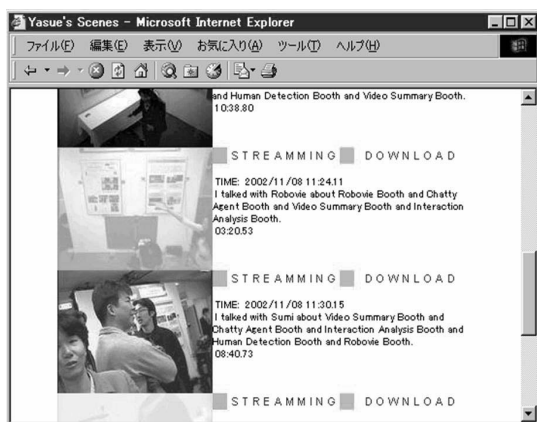


図 11 ビデオ要約 (インタラクション・コーパスから自動生成した体験の要約を Web ブラウザで閲覧する。主要なイベントを抽出し、そのイベントを “I talked with Adam and Bob about Booth-C.” のように説明して時刻順に並べている。説明の左には再生可能なビデオのサムネイルがある)

Fig. 11 Video summary (A summary is automatically generated and viewed with a web browser. Important events are extracted and listed in chronological way with video thumbnail and description such as “I talked with Adam and Bob about booth-C.”).

6.2 ビデオ日記

日記を構成するのはあくまでも日記制作者に任せ、その制作過程を支援する枠組みを考案した。体験キャプチャシステムから得られる一人称視点を含むビデオリソースの中から、日記のために選択しようとする際に手掛かりとなりうる各シーンのスコア群を提示する。具体的には、インタラクション・コーパスのデータベースから TALK の頻度などの特定のインタラクションの時間変化や、総インタラクションに対する割合など数種の統計的スコアを算出して提示する。日記制作者は数種類提示される統計的スコアを参考することで、原体験におけるさまざまなインタラクションの生起状況を目安として、膨大なビデオリソースの中から効率的にシーンを選択して日記様に構成することができる。

図12に、その考えに従って半自動的に日記を作成するために作成した GUI を示す。この GUI では、統計的スコアはグラフなどの形式を用いて左側にグラフィカルに提示される。時間軸上の特定時間を指定すると当該時間を含むシーンの体験インタラクションの素材が提示され、これらの中から日記に掲載したい候補を選び、右列で編集と文章記入をすることが可能である。これにより、日記制作者が主体的にビデオ日記を生成することができるようになった。

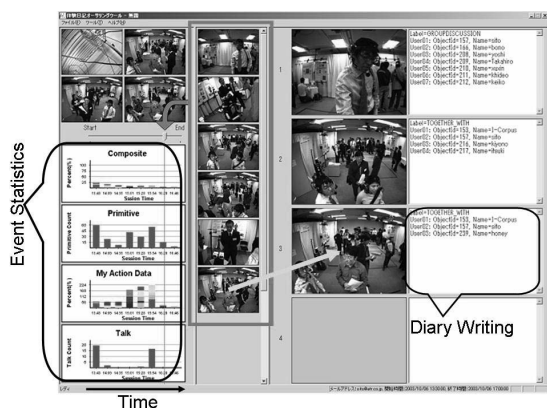


図 12 ビデオ日記（インタラクション・コーパスから自分に関連した素材を選び、日記を作成するシステム。素材の選択にあたって、インタラクションの統計データを提示したり、関連するイベント一覧を提示して、選択プロセスを支援する）

Fig. 12 Interaction corpus assisted video diary (Statistical data of interactions are used to assist for user to select diary events).

6.3 インタラクション・コーパスによる体験物語支援

上記の要約や日記は、それを相手に送付・呈示することによって、ある程度、体験を共有することができる。しかし、体験の共有には、背景の説明、イベントの連関の説明、理解と共感などのステップがインタラクティブに行われる必要があると考える。写真やビデオを見せながら体験を語るように、インタラクション・コーパスを利用して経験した体験の「ものがたり」を支援する道具を、模索した²¹⁾。図 13 は、Augmented Reality (AR) 技術を使って、コーパス中の体験シーンを操作しながら、自分の体験を相手に説明するツールである。GUI で説明するより、実空間を使って説明するほうが、相手との視線一致を得やすい、両手が使える、などの観察を得ている。本手法による体験共有への貢献の評価はこれからである。

7. おわりに

本稿では、体験共有コミュニケーションのためのユビキタス体験メディアの技術開発について報告した。体験を取得するウェアラブルデバイス、取得した体験を構造的に記録するインタラクション・コーパス、そして、体験を共有するための応用について、概要を紹介した。底流の考え方に焦点を当てて述べたが、紙面の都合で詳細については書ききれなかった部分が多々ある。それらは、参考文献に詳細があるのでそれらを参照してほしい。

本プロジェクトでは、2 章で述べたとおり、体験の



図 13 AR を使ったものがたりツール（あらかじめインタラクション・コーパスデータから、シーン、プレイヤー、小道具などのビデオクリップを抽出しておき、ARtoolkit を使ってビデオクリップを操作しながら、ものがたりをする）

Fig. 13 A storytelling tool with augmented reality (Video clips of scenes, players, and props are extracted from the interaction corpus. They are used for storytelling).

収集技術、体験 DB の構築技術、利用技術、社会的環境整備の 4 つの観点からアプローチし、本稿では前 3 者について具体的に概説した。これらについて、簡単に総括する。

まず、第 1 の体験の収集技術については、IrID を主たる手掛かりにしたインタラクション情報の記録システムを試作し、これと同時に記録する音声映像により実世界の体験記録をするという斬新なパラダイムを呈示できた。音声や画像のパターン処理の上にある、高次のインタラクション処理に目を向けたことは意義深いと考えている。装置については 3 世代の研究試作により、1 つの形を呈示することができたが、まだ改良の余地が多く残されている。また、体験を音声、映像、インタラクションだけで十分記録して他者と共有できるか、あるいはにおいや体制感覚が必要かどうかなどの疑問に答えるだけの材料は十分でない。しかし、コンテキストの記録により、状況説明の材料が収録できたことは特筆したい。

次に、第 2 の体験 DB の構築技術について、非言語常識としてのインタラクション・コーパスの提唱は、先端的であったと自負している。本研究では、インタラクションが階層化構造で記述できるとして議論を進めることにより、インタラクションを分析的にアプローチすることを目指し、ある程度成功した。しかし、各階層での処理を分析して性能を十分に吟味するには至っていない。また、現在はボトムアップに情報が統合されるモデルとなっているが、下層の解釈の信頼度は、

センサや状況によってつねに変化する．処理を単純化した IrID トラッカでさえも，その画像処理の性能に，出力が依存しているという事実がある．したがって，決定論的に情報統合してよいかどうか，メディア処理技術との整合性をとりながらさらに研究する必要がある．また，コーパス収集を拡大し，インタラクション語彙の拡充をしていく必要があるが，ドメインに応じたデバイスの開発が必要となろう．

さらに，第 3 の利用技術については，要約・日記などの例を呈示できた．当初の非言語常識の体系的収録と利用という，高邁な目的からすれば，より積極的な知識の記述・伝達の一助となる例を示せるという期待があったが，端緒を示したにすぎない．また，こつや技能の伝達といったエキスパートの暗黙知を収集・表現することも，いまだ困難であることが分かった．しかしながら，研究開始当初には考えられなかった blog と SNS の世界的流行などをみると，音声ビデオ日記は，今後利用技術の 1 つとして重要度が増すかもしれない．もの作りや教育・生活の現場の安心・安全と効率化のためには，インタラクションの分析が不可欠であり，利用技術について潜在的ニーズは非常に高い．子供が家族らとの体験共有を支援するという当初のビジョンには，まだ装置の手軽さ，インタフェースの改良など課題は多い．

最後に第 4 の社会的環境整備については，紙面の都合で本稿ではふれなかったが，社会的に国家安全保障とプライバシーへの関心が高まり，体験記録装置がもたらすプライバシー侵害の危険性について，当初の想定以上に深い議論が必要となった．我々はこれをユビキタス情報環境の「コストとメリットの天秤」問題の 1 つであるとして，プライバシー侵害コストを低減させる技術の追求と，対コスト・メリットのきわめて高い環境での利用の導入が必要であるとしたが，整理して論を展開するには至っていない．しかしながら，前者については，センシング時に顔領域をモザイク化する技術などを追究した²⁹⁾．後者について病院や工場はよい例であり，看護師の行動分析などへの展開を検討した³⁰⁾．体験データの暗号化や共有ポリシー制御などは，インターネット上の SNS などから学ぶところが多い．

ユビキタス体験メディアとインタラクション・コーパスの研究はまだ緒にいたばかりである．解決すべき技術課題，社会的課題は多々あるが，よりよいグローバル社会の時代に貢献するメディア技術を追求していきたい．

謝辞 本稿は，情報通信研究機構の委託研究「超高

速知能ネットワーク社会に向けた新しいインタラクション・メディアの研究開発」により ATR にて実施された成果を中心に，一部に文部科学省 ee-Society 委託事業により実施したものを含めてまとめた．本稿をまとめるにあたって，協力・支援をいただいた ATR メディア情報科学研究所，ATR 知能ロボティクス研究所の所員の皆様に感謝する．

参 考 文 献

- 1) <http://www.cyc.com/>
- 2) Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. and Miller, K.J.: Introduction to WordNet: An on-line lexical database, *International Journal of Lexicography*, Vol.3, No.4, pp.235–312 (1990).
- 3) Bush, V.: As We May Think, *The Atlantic Monthly*, Vol.176, No.1, pp.101–108 (July 1945).
- 4) Gemmell, J., Bell, G. and Lueder, R.: MyLifeBits: A personal database for everything, *Comm. ACM*, Vol.49, Issue 1, pp.88–95 (2006).
- 5) <http://www.darpa.mil/ipto/Programs/lifelog/>
- 6) http://www.darpa.mil/ipto/solicitations/open/04-38_PIP.htm
- 7) 上岡玲子，広田光一，廣瀬通孝：ウェアラブルコンピュータによる主観的体験とその展開，ヒューマンインタフェース学会研究報告集，Vol.5, No.4, pp.65–68 (2003).
- 8) 相澤清晴，石島健一郎，椎名 誠：ウェアラブル映像の構造化と要約：個人の主観を考慮した要約生成の試み，信学論 D-II，Vol.J86-D-II, No.6, pp.807–815 (2003).
- 9) Roy, D., et al.: The Human Speechome Project, *Proc. 28th Annual Cognitive Science Conference (CogSci2006)*, pp.2059–2064 (July 2006).
- 10) Lamming, M. and Flynn, M.: “Forget-me-not” Intimate computing in support of human memory, *Proc. International Symposium on Next Generation Human Interface '94FRIEND21*, pp.150–158 (1994).
- 11) 川嶋稔夫：記憶想起支援とウェアラブルコンピュータ，第 5 回知能情報メディアシンポジウム論文集，pp.107–108 (1999).
- 12) Kawamura, T., Kono, Y. and Kidode, M.: Wearable interfaces for a video diary: Towards memory retrieval, exchange and transportation, *The 6th International Symposium on Wearable Computers (ISWC2002)*, pp.31–38 (2002).
- 13) Clarkson, B., Mase, K. and Pentland, A.:

- The Familiar: A living diary and companion, *CHI2001 extended abstracts*, Seattle, pp.271–272 (Apr. 2001).
- 14) Pentland, A.P.: Smart Rooms, *Scientific American*, pp.54–62 (Apr. 1996).
 - 15) Sumi, Y., Matsuguchi, T., Ito, S., Fels, S. and Mase, K.: Collaborative Capturing of Interactions by Multiple Sensors, *Adjunct Proc. 5th International Conference on Ubiquitous Computing (UbiComp2003)*, Seattle, pp.193–194 (Oct. 2003).
 - 16) Kogure, K., Hagita, N., Sumi, Y., Kuwahara, N. and Ishiguro, H.: Toward Ubiquitous Intelligent Robotics, *Proc. 2003 IEEE International Conference on Intelligent Robots and Systems*, pp.1826–1831 (Oct. 2003).
 - 17) Sumi, Y., Sakamoto, R., Nakao, K. and Mase, K.: ComicDiary: Representing individual experiences in a comics style, *Ubicomp 2002*, Borriello, G. and Holmquist, L.E. (Eds.), LNCS 2498, pp.16–32, Springer (Sep. 2002).
 - 18) Mase, K., Sumi, Y., Tsuchikawa, M. and Kogure, K.: Interaction Corpus and Memory for Experience Sharing, *International Symposium on Large-scale Knowledge Resources* (Mar. 2004).
 - 19) Stiefelwagen, R., Yang, J. and Waibel, A.: Towards tracking interaction between people, *Proc. Intelligent Environments AAAI Spring Symposium*, Stanford Univ., Calif. (1998).
 - 20) 角 康之, 間瀬 健二, 土川 片桐, 萩田 伊藤, 岩澤 中原, 神田: ユビキタス環境における体験の記録と共有, システム制御情報学会誌「システム/制御/情報」, Vol.48, No.11, pp.458–463 (2004).
 - 21) Lin, N., Mase, K., Sumi, Y. and Toriyama, T.: Enabling communication-based interactive storytelling through a tangible mapping approach, *International Conference on Virtual Storytelling 2005*, LNCS3805, pp.229–238, Springer (Nov.–Dec. 2005).
 - 22) 神田 崇行, 佐藤 留美, 才 脇直樹, 石黒 浩: 対話型ロボットによる小学校での長期相互作用の試み, ヒューマンインタフェース学会論文誌, Vol.7, No.1, pp.27–37 (2005).
 - 23) 土川, 岩澤, 伊藤, 鳥山, 小暮, 萩田, 間瀬, 角: 装着型共有体験記録装置の試作, VR 学会論文誌, Vol.10, No.4, pp.533–542 (Dec. 2005).
 - 24) 伊藤, 岩澤, 土川, 角, 間瀬, 片桐, 小暮, 萩田: 装着型体験記録装置による対話インタラクションの判別機能実装と評価, ヒューマンインタフェース学会論文誌, Vol.7, No.1, pp.167–178 (2005).
 - 25) 森田 友幸, 平野 靖, 角 康之, 梶田 将司, 間瀬 健二, 萩田 紀博: マルチモーダルインタラクション記録からのパターン発見手法, 情報処理学会論文誌, Vol.47, No.1, pp.121–130 (2006).
 - 26) 角 康之, 伊藤 慎宣, 松口 哲也, シドニーフェルス, 間瀬 健二: 協調的なインタラクションの記録と解釈, 情報処理学会論文誌, Vol.44, No.11, pp.2628–2637 (2003).
 - 27) Takahashi, M., Ito, S., Sumi, Y., Tsuchikawa, M., Kogure, K., Mase, K. and Nishida, T.: A layered interpretation of human interaction captured by ubiquitous sensors, *The 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE 2004)*, pp.32–38 (Oct. 2004).
 - 28) 角 康之, 岩澤 昭一郎, 間瀬 健二: インタラクション・コーパス分析ツールの試作, 情報処理学会研究報告 (ヒューマンインタフェース), HI104-7 (2003).
 - 29) Kitahara, I., Kogure, K. and Hagita, N.: Stealth Vision for Protecting Privacy, *Proc. 17th International Conference on Pattern Recognition (ICPR 2004)*, Vol.4, pp.404–407 (2004).
 - 30) 桑原, 野間, 鉄谷, 萩田, 小暮, 伊関: ウェアラブルセンサによる看護業務の自動行動計測手法, 情報処理学会論文誌, Vol.44, No.11, pp.2638–2648 (2003).

(平成 18 年 6 月 12 日受付)
(平成 18 年 11 月 10 日採録)

(担当編集委員 浮田 宗伯)



間瀬 健二 (正会員)

1981 年名古屋大学大学院工学研究科情報工学専攻修士課程修了。同年 NTT 入社。1995～2002 年 (株) 国際電気通信基礎技術研究所研究室長。2002 年より名古屋大学情報連携基盤センター教授。コンピュータによるコミュニケーション支援の研究を推進している。IEEE, ACM, 電子情報通信学会, VR 学会, 画像電子学会各会員。博士 (工学)。



萩田 紀博（正会員）

1978 年慶應義塾大学大学院工学研究科電気工学専攻修士課程修了。同年電電公社（現 NTT）武蔵野電気通信研究所入所。文字認識、画像認識、コミュニケーションロボット、ネットワークロボット等の研究に従事。現在 ATR 知能ロボティクス研究所長、ATR メディア情報科学研究所長兼務、工学博士。IEEE、電子情報通信学会、人工知能学会、日本ロボット学会各会員。



角 康之（正会員）

1990 年早稲田大学理工学部電子通信学科卒業。1995 年東京大学大学院工学系研究科情報工学専攻修了。同年（株）国際電気通信基礎技術研究所（ATR）入所。2003 年より京都大学大学院情報学研究科助教授。博士（工学）。研究の興味は知識処理システムとヒューマンインタフェース。



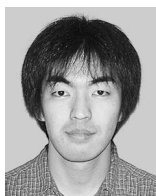
小暮 潔（正会員）

1981 年慶應義塾大学大学院工学研究科電気工学専攻修士課程修了。同年日本電信電話公社に入社。現在は ATR 知識科学研究所所長。博士（工学）。自然言語処理、エージェント、ロボット、知的環境等の研究に従事。電子情報通信学会、人工知能学会、言語処理学会、日本認知科学会、日本音響学会各会員。



片桐 恭弘（正会員）

1981 年 3 月東京大学大学院工学系研究科情報工学専攻修了。工学博士。NTT 基礎研究所、ATR メディア情報科学研究所を経て現在公立はこだて未来大学教授。自然言語処理、社会的インタフェース、対話インタラクションの認知科学の研究に従事。日本認知科学会、日本人工知能学会、社会言語科学会、自然言語処理学会、Cognitive Science Society、ACM、AAAI、IEEE 各会員。



伊藤 禎宣（正会員）

2003 年北陸先端科学技術大学院大学知識科学研究科博士後期課程修了。同年（株）国際電気通信基礎技術研究所（ATR）メディア情報科学研究所研究員。2006 年より、東京農工大学大学院工学府特任講師。博士（知識科学）。



岩澤 昭一郎

2000 年成蹊大学大学院工学研究科博士後期課程単位取得退学。1999 年通信・放送機構招へい研究員。2002 年 ATR メディア情報科学研究所客員研究員を経て、現在 ATR 認知情報科学研究所より情報通信研究機構へ専門研究員として在籍出向中。2000 年電気通信普及財団賞受賞。CG や画像処理の研究に従事。ACM 会員。博士（工学）。



鳥山 朋二

1987 年富山大学大学院工学研究科修了。同年 NTT 厚木電気通信研究所入所。以来、LSI の設計方式、ヒューマンインタフェースおよび画像処理に関する研究に従事。2004 年富山県立大学大学院工学研究科社会人後期博士課程単位取得退学。現在は ATR 知識科学研究所研究室長。電子情報通信学会、ヒューマンインタフェース学会、画像電子学会各正会員。博士（工学）。



土川 仁

1990 年早稲田大学大学院（機械工学）修了。同年日本電信電話（株）入社。ヒューマンインタフェース研究所、サイバーソリューション研究所等において、動画像認識、医用画像伝送の研究に従事。2003 年 7 月より（株）国際電気通信基礎技術研究所（ATR）に出向。2005 年 11 月より NTT 知的財産センタ勤務。