

ICMI2007 参加報告

神山 祐一[†], 米澤 朋子[‡]

[†] 名古屋大学, [‡] ATR 知能ロボティクス研究所

概要

ICMI(International Conference on Multimodal Interfaces) は 1996 年にはじめて開催された国際会議で, 1997 年以降に UIST97 (User Interface Software and Technology) のワークショップとして開催されてきた PUI(Perceptual User Interface) が 2003 年に加わり, 以降 ACM の毎年の国際会議として開催することとなり, これまでに採択難易度が徐々に上がっている. 視覚情報・音声情報処理やそれらに基づいたマルチモーダルインタフェースに焦点を当てた会議である. 今年は 11 月 12 日～15 日 (4 日目はワークショップ) に日本国内, 名古屋能楽堂にて開催された. 参加者は 150 名程で, 大規模な国際会議ではないが SIGCHI などのスポンサーがついた多数の先進的な研究が集まる会議である. また, HCI 研究会の協賛を得ている. 特にマルチモーダル情報処理に関しては, 視線, 音声, 顔画像, ジェスチャ画像分析といった個々の従来の技術から統合的な情報処理へと発展させた様々な研究が見られた.

100 件を超える投稿の中で, オーラルセッションには 18 件, ポスター発表は 37 件, そのほか Doctral Spotlight セッションやデモセッションがあった. 国内開催のため, デモセッションは日本からの出展が目立った.



図 1: 会場 (名古屋能楽堂) 向かいの名古屋城

Keynote

Keynote には 3 名のスピーカーが招かれた. Yuri A. Ivanov 氏はビルのフロアに巡らせた人感センサと少数のカメラによる監視を紹介し, 行動の軌跡による人物映像の検索等のデモを行った. Norihiro Hagita 氏はマルチモーダルインタフェースの観点からネットワークロボットの可能性を語った. Dominic Massaro 氏は CG チューターを用いた言語教育を紹介した.



図 2: Keynote の様子 (Yuri A. Ivanov 氏)

オーラルセッション

1 日目は Spontaneous Behavior に関する 2 セッションが設けられ, 臨床場面における痛み (Ahmed Bilal Ashraf 氏や Gwen C.Littlewort 氏) や微笑み (Michel Francois Valstar 氏) の定量化, 自閉症の発見 (Rana el Kaliouby 氏) 等が取り上げられた. Michel Francois Valstar 氏のスライドで, 怒りの (に見える) 顔表情が示された後に, この人物がテニスコート上でガッツポーズをしている全体像から至上の喜びの表現であることが示され, “マルチモーダル” を端的に表していた.

2 日目は Cross-modarity のセッション, 及び Meeting Application のセッションが設けられた. Outstanding Paper Award に選ばれた Kazuhiro Otsuka 氏の発表 “Automatic

Inference of Cross-modal Nonverbal Interactions in Multiparty Conversations - Who Responds to Whom, When, and How? from Gaze, Head Gestures, and Utterance” では、会議状況における複数話者の頭部方向、頭部ジェスチャ、発話から、マルコフ連鎖モデルを用いて会話レジーム（集中/2者対話/発散の状況）、視線、インタラクション構造（行為間の因果関係）を同時に推定し、無意識の会話の構造に迫った。非言語情報が会話のダイナミクスにより直接影響を与える情動やパラメッセン（複数の感情表出チャンネルから同時に発せられる2つ以上のメッセージ）の伝達への展開も含めて、社会相互作用の解析に有用なアプローチに思われる。

3日目は Interactive Systems の2セッションがあり、Boris Brandherm 氏の発表では、バーチャルショールームで2人のCGプレゼンターが商品の説明を行う infotainment (information+entertainment) システムが紹介された。CGプレゼンターの身振り、及び視線に対するベイジアンネットワークを用いたユーザの興味（プレゼンテーションへの興味、オブジェクトへの興味、注視）推定により、入力と出力の両面でマルチモーダルなシステムとしていた。

ポスター/デモセッション

ポスターセッションは1日目から3日目にかけ、デモセッションは2日目と3日目に、それぞれ2時間の枠が設けられ、充実した議論が交わされた。Rony Kubat 氏による長時間映像の可視化とアノテーションを行った Total Recall、Hiroko Tochigi 氏によるユーザ、ぬいぐるみロボット、インタラクティブアニメーションの3者関係を扱った研究、Susumu Harada 氏の音声の同時入力によりペン入力を拡張した Voice Pen、Tobias Klug 氏のウェアラブルマルチモーダルシステムの設計等、人間の経験に関わる幅広い領域の研究が見られた。米澤らはポスターセッションおよびデモセッションにて視線計測技術を用いたぬいぐるみロボットの視線行動について発表し、画像処理などの技術者や心理学よりの研究者など多方面の参加者から多くの反響を得ることができた。

併設ワークショップ

今年は本会議以外にも、4日目にワークショップが並列で3件（1. Multimodal Interfaces in Semantic Interaction, 2. Tagging, Mining and Retrieval of Human Related Activity Information, 3. Massive Datasets）併設され、80名程度の参加があり大いに盛り上がった。米澤は1.の Seman-



図 3: ポスターセッションの様子



図 4: 2日目夜のバンケット

tic Interaction のワークショップに参加した。Keynote の、Jordan Zlatev 氏による Mimesis Hierarchy Model のセッションでは、繰り返しの言葉など擬音・擬態語や人間の発達・学習、そしてその過程の表現などに触れ、マルチモーダルインタフェース研究に携わる研究者や技術者に興味深い基礎研究が紹介された。Kate Saenko 氏によるショートペーパーの発表では、ロボットが意味学習をする際に大量の画像を学習し、実際にユーザが指し示したものの意味を理解するために用いる実験を紹介していた。また、Tanaka Kazuaki 氏の発表は、Aibo に正しい動作を学習させる過程で、実際に動作してみてもユーザの反応を Reward として学習するという興味深いものであった。このように、本会議に劣らない興味深い発表を聴講できるワークショップが併設されたことは、今後の ICMI の発展に大きく貢献すると考えられる。

次回開催予定

次回の ICMI2008 はギリシャ・クレタ島にて、10月13日～15日に開催される。