



会議レポート

2035年のマルチメディアの姿を予想 — ICME 2015 会議レポート —

ICME とは

IEEE International Conference on Multimedia and Expo (ICME) はマルチメディアに関する技術やそのアプリケーションなどについての国際会議で、画像や映像に加えて音声やデプス画像、ソーシャルメディア、さらにはマルチメディアデータの伝送など幅広い研究発表を対象としている。マルチメディアに関する限りでは ACM Multimedia に次ぐ規模の国際会議であろう。16 回目となる今回はイタリア トリノでの開催であった。本会議に加えて 13 のワークショップが併設され、特定の分野に関する専門的な議論がなされていた。会議全体の参加者は 420 名以上で、アジアから 51%、ヨーロッパから 35%、北アメリカから 13% であった (図-1)。コンピュータサイエンスの分野に共通して見られる特徴ではあると思うが、特に中国からの参加者が多く、アジアからの参加者の割合を押し上げている。今年は日本の研究者らの参加が多いワークショップ (Workshop on Multimedia for Cooking and Eating Activities) が併設されたためか、昨年に比べて日本からの参加者も多いようであった。

投稿された論文はメタ査読者に加えて平均 4.4 人の査読者からのコメントに基づき採録が判定され、投稿されたすべての論文 (532 件) のうちの 30% (157 件、うち口頭発表 78 件) が本会議に採録された。また、評価の高い論文は Best Paper Candidate、もしくは Best Student Paper Candidate に選出され、これらからそれぞれ Best Paper Award と Best Student Paper Award が 1 件ずつ選ばれた。

Multimedia 2035

今年の ICME で特徴的だったのは、今から 20 年後の 2035 年のマルチメディア関連の研究がどのようなものかに関して、著名な研究者 4 名がパネルディスカッションを行う Multimedia 2035 というセッションであった。

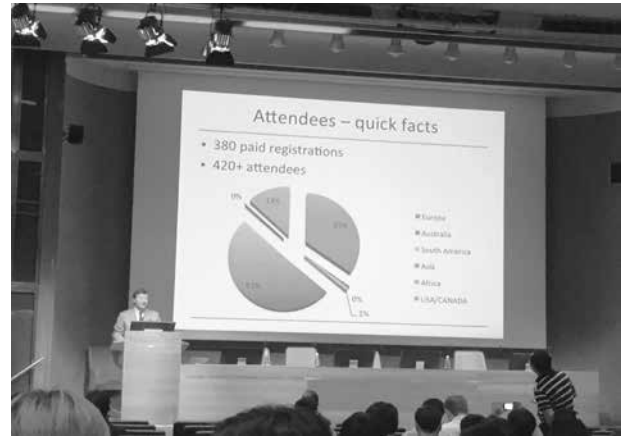


図-1 参加者の分布

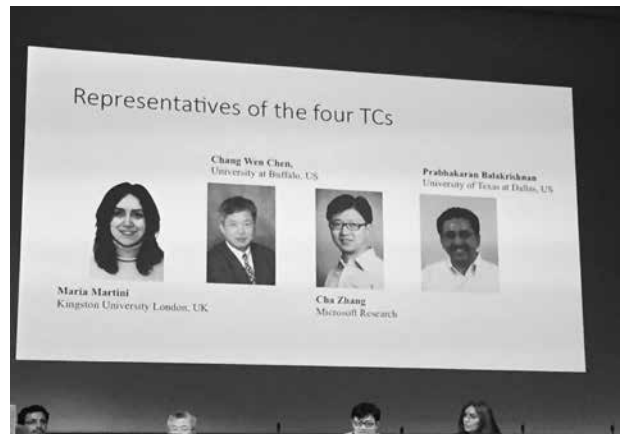


図-2 Multimedia 2035 のパネリスト

パネリストは IEEE の関連するソサイエティの専門委員 (図-2) で、自身の所属する専門委員会の委員の投票や聞き取りなどに基づいて 2035 年のマルチメディアの姿を予想していた。それぞれ未来の予想は難しいとしつつも、没入型テレプレゼンスやヘルスケア、ロボットなどはほぼすべてのパネリストに共通したキーワードとなっており、これらを実現するためのセンシングやそこから得られるマルチメディアデータの処理、伝送などの技術に関する研究開発が進められているだろうとする意見が多いようであった。

たとえば、Circuit and Systems Society, Multimedia Systems and Applications Technical Committee の Chang Wen Chen 教授は、現在のトレンドとして多様なセンサとクラウドを介したデータへのアクセス、大型のタッチディスプレイや頭部の動きに合わせた表示が可能なヘッドマウントディスプレイである Oculus Rift などの表示デバイスを挙げた上で、2020 年から 2035 年までに発展する可能性のある分野として、各種センサからのデータを、クラウドを介して活用 (高齢者やこどもの見守りなど) する監視システムや、リアルタイムに 3 次元データを取得する没入型テレプレゼンスなどを挙げた。

また、Communications Society, Multimedia Communications Technical Committee の Maria G.



図-3 Milanfar 教授の Plenary Talk の様子

Martini 教授は、携帯電話の通信規格である LTE などの伝送技術で伝送可能なマルチメディアデータを例示し、今までの伝送スピードの進化を示した上で、無線ロボットによる立体視可能な映像での内視鏡検査など、大容量データの伝送やリアルタイム性が要求されるアプリケーションの実現を 2035 年までに達成すべき目標として挙げた。

途中、聴講者からの意見として、今までの革新的技術は当初「Crazy」だと言われていたものばかりだったのでパネリストも「Crazy」な予測をしてほしい、というものがあった。この意見に対してほとんどのパネリストは「この予測が Crazy かどうか分からないが……」との反応であった。実際、今回登場したものはすべて現時点で提案されている技術の進化版のような印象で、どちらかという堅実な予測が多かった。もう少し突飛な予測があると面白かったのではないかと思う。いずれにしても、このようなセッションは聴講者にとってこれからの研究方針を考える 1 つの機会になり得るので、今後も同様の試みを期待したい。

Plenary Talk

Plenary Talk では 3 名の著名な研究者が講演した。Imperial College London の Maja Pantic 教授は、自身の研究グループが実施した顔の検出や追跡、認識、および顔の特徴点の検出、追跡に関する研究について、過去から現在にいたる成果を紹介した。Google Research, University of California Santa Cruz の Peyman Milanfar 教授はカメラやデジカメの普及から Computational Photography に至る歴史と Computational Photography の基礎について講演を行った(図-3)。ShanghaiTech University の Yi Ma 教授の講演では、凸最適化の手法を用いた行列の低ランク近似とその画像処理などへの応用について解説した。

受賞論文

今年度の Best Paper Award は、Academia Sinica の Zsu-Yu Chou らの論文“Evaluating music recommendation

in a real-world setting and evaluation metrics”が受賞した。この論文では音楽推薦システムを現実の利用に即して評価する方法を提案し、実際に商用利用されている音楽推薦システムのデータを用いてさまざまな評価指標により推薦手法を評価している。

Best Student Paper Award は、Yuntao Li ら (Dalian University of Technology など) による“Characteristic Number Regression for Facial Feature Extraction”が受賞した。この研究では、Characteristic Number と呼ばれる射影不変量を利用した顔特徴点 (口の両端、目の両端など) の位置の回帰を実現している。その他の賞については、ICME の Web サイトの受賞に関するページ^{☆1}をご参照いただきたい。

ICME に関する所感

コンピュータビジョンや画像処理にターゲットを絞った会議に比べて (画像・映像に関する研究は多いものの) ICME はトピックが幅広く、またアプリケーションに関する研究発表が多いようである。このためか、現在トレンドとなっている手法などに関する研究で特に目立つ印象のものはなかった。たとえば、深層学習に関連する研究はタイトルだけを見ると数件であり、深層学習の新たな利用法などに関する提案に主眼を置いたものがほとんどである。

アプリケーションに注目すると、目新しいものとして Multimedia 2035 でもキーワードとして頻出したヘルスケアに関するオーラルセッションが組まれていた。さまざまなセンサから得られたデータを利用して健康状態のモニタリングなどを行うという点で、マルチメディアの会議として望ましいトピックであると思われる。Multimedia 2035 の予測と併せて、マルチメディアコミュニティの今後の方向性を示しているように感じた。

ICME はポリシーとして採択率を 30% 前後としているとのことで、採択率としては低いものであるが、筆者自身の経験として ACM Multimedia で不採録となった論文が翌年の ICME で採録されるなど、比較的採録されやすいように感じる。また、マルチメディア関連の会議の特色かもしれないが、さまざまなバックグラウンドを持つ参加者が集まる傾向があるように思う。専門分野の異なる聴講者との議論を通じて提案手法の新しい側面に気づかされることもあるのではないだろうか。企画セッションや併設ワークショップもさまざまなものがあり、幅広い分野の研究を知る機会にもなるので、研究発表の場として魅力的な会議であると思う。

(中島悠太/奈良先端科学技術大学院大学)

☆1 <http://www.icme2015.ieee-icme.org/awards.php>