ACM Multimedia (ACMMM) 2016 参加報告

(正会員) 井手 一郎[†], 柳井 啓 司^{††}

本稿では、2016年10月15日から19日まで、オランダ・アムステルダムで開催された第24回 ACM マルチメディア国際会議 (ACM Multimedia 2016) の様子を報告する。

キーワード：ACMマルチメディア国際会議、アムステルダム、マルチメディアシステム、マルチメディア体験、マルチメディア理解、マルチメディアエンゲージメント

1. まえがき

今回で24回目を迎える ACM マルチメディア国際会議は、米国計算機械学会 (ACM) のマルチメディア分科会 (SIGMM) が主催する、マルチメディア分野のトップカンファレンスである。我が国でも2012年に奈良で開催されたので、その際に参加された読者もおられるだろう。

今回、同会議は、オランダの首都アムステルダムで開催された。主催者の発表によると、参加者は約550名で、そのうち日本からは、中国、オランダ、米国、ドイツに次いで多い30名であった。6位以降は、台湾、イタリア、シンガポール、フランス、韓国と続く。なお、ここ数年続いているが、今年も予稿集が配布されず、代わりに会議開催期間前後のみ ACM Digital Library が無料開放され、参加者は Digital Library 経由で予稿を読むというスタイルであった。ネット環境がないと予稿が読めないのは大変不便であるので、個人的には USB メモリー配布を復活させてほしいものである。なお、直前に開催された ECCV では USB メモリーの予稿集が配布されていた。

開催国のオランダは、ライン川下流域の低湿地帯に位置し、国土の3割は海面下にある。また、国土の2割以上が干拓によって造成されたポルダ (polder) と呼ばれる埋め立て地である。アムステルダムも例にもれず、今でも中心を流れるアムステル川の河口をダムでせき止めて造成された街であり、全長100 km を越える運河と90を越える細長い島々、それらを結ぶ1,500を越える橋からなる水上都市である。名物の風車も、排水用に建てられたものが多く、アムステルダム中心部にもまだ数基残っている。そのうちワー

図1 チュートリアル・ワークショップ会場になったアムステルダム大学ロータースアイランド校舎

クショップ・チュートリアル会場に近い1基は、今では醸造所 (Brouwerij 't ij) になっており、美味しい地ビールを堪能した参加者もいたようだ。

そのチュートリアル・ワークショップは、本会議に先だって、街中にあるアムステルダム大学の人文系の校舎 (図1) で開催された (ここでも理工系の校舎は街外れの自然豊かだが、周りに何もない不便なキャンパスに追いやられている)。今回のチュートリアル・ワークショップの特色は、

[†] 名古屋大学 大学院情報科学研究科

^{††} 電気通信大学 大学院情報理工学研究科

"ACM Multimedia (ACMMM) 2016 Report" by Ichiro Ide (Graduate School of Information Science, Nagoya University, Nagoya) and Keiji Yanai (Department of Informatics, the University of Electro-Communications, Tokyo)

図2 ワークショップ会場は、ECCVのワークショップ会場と同じ場所に設定され、双方の参加者はどちらのワークショップにも参加することができた。

前週に同じくアムステルダムで開催されていた欧州コンピュータビジョン会議 (ECCV) と合同で、両方の参加者が相互参加可能な形で開催されたことだ(図2)。ただし、2週間に渡り国際会議に連続参加することはいろいろな意味で大変なことであり、ACMマルチメディアの参加登録者551名の約1割の53名のみがECCVとの(双方の参加費が1割引になる)同時参加登録をしたとのことであった。

一方、本会議は、繁華街にある「トゥシンスキー劇場 (Pathé Tuschinski)」という、1920年代に建てられた、アールヌーボーやアールデコを含むさまざまな建築様式が混在した、やや奇妙でありながらも壮麗な映画館で開催された(図3、図4)。映画館ということで、椅子の座り心地が非常に良く、~~よく眠れる~~長く座っていても疲れないと評判であった。

招待講演では、チューリッヒ連邦工科大のDirk Helbing教授によるIoT時代における技術のあり方に関する問題提起、アイントホーフェン工科大のJack van Wijk教授による大量データの可視化とインタラクション、いわゆるヴィジュアル＝アナリティクスに関する話題提供があった。また、パネル討論では、深層学習の問題点と今後の発展性に

図3 本会議場になったトゥシンスキー劇場の外観

図4 本会議場になったトゥシンスキー劇場の内部

ついて議論がなされた。

さて、本会議が映画館で開催されたために、通常は口頭発表と並行して数日にわたって開催されるデモ・ポスタ・ビデオ展示は、本会議初日の午後に集約され、近くの「バーゼル館 (De Bazel)」と呼ばれる別の建物(現在はアムステルダム史料館)で開催された。この建物も、日本人になじみ深いオランダ東インド会社の後継会社オランダ貿易協会の本店として、同じく1920年代に建てられたアールデ

図5 デモ・ポスタ・ビデオ会場となったパーゼル館(旧オランダ貿易協会本社)の入口

図6 ポスタ・ビデオ展示会場で熱心に議論する参加者たち

コ様式の重厚な建物であった。吹き抜けの天井が高い1階ホールでポスタおよびビデオ展示が、4階の重厚な会議室でデモ展示が行われた(図5, 図6)。

他の国際会議にない、ACMマルチメディアの大きな特徴として、いわゆる学術発表とは別に、メディアアートに関する展示会が併設される点が挙げられる。今回も例に漏れず、現地のワグ協会と呼ばれる芸術・科学・技術に関

図7 アート展示会場となったアムステルダム公共図書館中央館(右から二つ目の建物)

する団体と共催で、アムステルダム公共図書館中央館という近代的な建物で開催された(図7)。今回のテーマは「データの美学 (DATA AESTHETICS)」で、ビッグデータの可視化のほか、ビッグデータに関する倫理や可能性について表現したメディアアートが展示されていた。

以下、主要なプログラムについて、その概要を紹介する。

2. プログラムの概要

以下、主要なプログラムについて、その様子を報告する。

2.1 本会議

今回も2010年に始まったエリア制が採用され、システム (System)、体験 (Experiences)、理解 (Understanding)、エンゲージメント (Engagement) の4テーマの下に設定された15のエリア毎に投稿および査読が行われた(表1に、エリア毎の投稿数を示す)。エリア毎に任命された、全体で46名のエリアチェアが査読割当および採否判定作業に携わり、2016年6月10日に米国ニューヨークのコロンビア大で採否判定会議が開かれた。世界中から集まったエリアチェアが1日ばかりで、時には原文にあたりながら、多角的な視点から真剣な議論を重ねた結果、フルペーパーについては52件(22.2%)、ショートペーパーについては127件(30.3%)の論文が採択された。日本からは、残念ながらフルペーパーの採択はなかったものの、東大、NTT、電通大、東工大、北大、名大の研究グループから、合計9件のショートペーパーが採択された。

今回、2015年のSIGMM 公開運営会議における議論の結果導入された新たな試みとして、フルペーパーにおけるページ制限と、ショートペーパーにおける追加ページの導入があった。前者は、従来10ページ以内としか規定されていなかったのに対して、最低ページ数が8ページに設定された。後者は、従来4ページ以内に制限されていたのに対して、参考文献に限って5ページ目を使用できるようになった。

表1 エリア別投稿件数

テーマ	エリア	フルペーパー	ショートペーパー
システム (System)	Multimedia Transport and Delivery	17	10
	Multimedia Systems and Middleware	13	12
体験 (Experiences)	Multimedia Telepresence and Virtual/Augmented Reality	10	12
	Multimedia Scalability and Management	5	5
	Mobile Multimedia	6	12
理解 (Understanding)	Multimedia HCI and Quality of Experience	12	21
	Music, Speech and Audio Processing in Multimedia	8	17
エンゲージメント (Engagement)	Multimedia Art, Entertainment and Culture	9	7
	Multimedia for Collaboration in Education & Distributed Environments	2	3
	Deep Learning for Multimedia	20	57
	Multimodal Analysis and Description	26	35
	Multimedia and Vision	45	125
	Emotional and Social Signals in Multimedia	7	26
	Social Multimedia	16	19
	Multimedia Search and Recommendation	41	52
その他		11	12
合 計		248	425

査読過程の特徴として、今回も2015年から始まった反論(rebuttal)制度と擁護(advocate)制度が採用された。前者は、査読者からの指摘や質問に対して、著者らが反論・回答する機会を与え、最終的な査読結果に反映させる制度である。後者は、査読過程に疑義がある著者による苦情受付窓口のようなものである。時間的猶予が少ないなかで実施される制度であるため、実効性には若干疑問があるが、査読過程の透明性を高める試みとしては評価できるであろう。

最優秀論文賞の選定は、まず判定会議の場で各エリアから推薦されたものから8件が候補として選ばれた。その後、別途設けられた賞選定委員会による慎重な審査を経て4件に絞り込まれたものが、本会議の冒頭に設けられた最優秀論文セッションで講演された。これをうけて最終審査がなされ、本会議2日目に開催されたSIGMM公開運営会議の場で発表された。今回の受賞論文は次の通りであった：

- ・最優秀論文賞：「Multi-modal multi-view topic-opinion mining for social event analysis」(中国科学院自動化研究所のグループ)
- ・最優秀学生論文賞：「Deep-based ingredient recognition for cooking recipe retrieval」(香港城市大のグループ)

本稿の著者らが食事画像に関する研究に携わっていることもあり、後者の受賞は個人的にも大変喜ばしいことであった。

全体的な傾向としては、コンピュータビジョン(CV)分野とのオーバーラップが年々増加し、マルチメディア分野独自の研究というものがかなり少なくなっているという印象を受けた。これは深層学習の登場によって画像認識の精度が飛躍的に向上したおかげで、CV分野の研究が多様化して、マルチメディア化していることが大いに影響している。なお、直前に開催されたECCVではほぼすべての画

像認識の研究は深層学習を利用したもので、3次元復元や光学解析にも深層学習を応用した研究が多くみられるようになり、まさに深層学習一色であった。それに対して、ACMマルチメディアではそこまで比率は高くないものの、それでも半数近くが何らかの形で深層学習を利用していると思われ、深層学習の利用率は昨年よりもさらに向上していた。

2.2 チュートリアル

チュートリアルについては、本会議独自のものと、前週から開催されていた欧州コンピュータビジョン会議(ECCV)と合同開催のものの2種類が開催された。以下にそれらの一覧を載せる。

【ACM Multimedia独自のチュートリアル】

- ・Technology & Art in Stimulating Creative Placemaking in Public-Use Spaces
- ・Social and Affective Robotics Tutorial
- ・Emerging Topics in Learning from Noisy and Missing Data
- ・Multimedia Privacy
- ・Situation Recognition from Multimodal Data
- ・the Lifecycle of Geotagged Multimedia Data

【ECCVと共同開催のチュートリアル】

- ・Segmentation and Tracking
- ・Developments in Deep Learning
- ・Multimodal Human-Robot Interaction
- ・Output Embedding for Large-Scale Visual Recognition
- ・Video Event Detection, Localization and Cross-Media Linking
- ・Human Body Shape Modelling
- ・A Guide to Creating and Managing Lifelog Collections

- ・ Deep Visual Understanding from Deep Learning
- ・ Bridging Video and Language with Deep Learning
- ・ Color and Spectral Information in Modern Computer Vision
- ・ Developing Home-based Stroke Rehabilitation Systems using Interactive Cognitive Vision

2.3 ワークショップ

ワークショップについても、ECCVの一部のものと合同で開催された。以下にその一覧を載せる。

- ・ Multimedia COMMONS Workshop 2016
- ・ 6th Int. Audio / Visual Emotion Challenge and Workshop (AVEC 2016)
- ・ 1st Workshop on Lifelogging Tools and Applications (LTA 2016)
- ・ 1st Int. Workshop on Multimedia for Personal Health and Health Care (MMHealth 2016)
- ・ Workshop on Vision & Language Integration Meets Multimedia Fusion
- ・ 2nd Int. Workshop on Multimedia Assisted Dietary Management (MADiMa 2016)
- ・ Int. Workshop on Multimedia Alternate Realities (AltMM 2016)
- ・ 7th Int. Workshop on Human Behavior Understanding (HBU 2016)
- ・ [ECCV] Workshop on Computer Vision for Audio-Visual Media (CVAVM)
- ・ [ECCV] 3rd Workshop on Computer Vision for Art Analysis (VISART)
- ・ [ECCV] 2nd Workshop on Virtual / Augmented Reality for Visual Artificial Intelligence (VARVAI)
- ・ [ECCV] Joint Workshop on Storytelling with Images and Video (VisStory) / Large Scale Movie Description and Understanding Challenge (LSMDC)

ECCVのワークショップとの合同開催は良い試みだったと思うが、内容の重複に関する調整がなされていなかったため、関心があっても聴講できない講演があったのは、残念であった。

2.4 デモ・ビデオ展示

31件のデモ展示と3件のビデオ展示があった。日本からは、国立情報学研究所と日本電気から2件のデモ展示と日本電気から1件のビデオ展示があった。

デモセッションは、ポスタセッションと同時に開催され、3時間の枠があったが、127件のショートペーパーに加えて、52件のフルペーパーの大部分もポスタ発表を行ったため、発表件数が多く、それに加えて31件のデモ発表を見て回るにはとても時間が足りない状況であった。1つのセッションですべて同時に発表すると、参加者の多くが発表者も兼ねることになってしまい、互いに他の発表をゆっくり見て回

図8 本会議場前の街並
突き当たりに見えるのは、中世の市壁の一部として残っているムント塔。

ることが難しくなってしまうことも問題であった。

ビデオ展示は、デモ展示やポスタ展示と大差なく、しかも投稿は3本のみで3本とも採択される状況であった。今後の工夫もしくは廃止も含めた検討が求められる。

2.5 その他の企画

以上に紹介したもの以外に、以下の恒例の企画もあった。

- ・ Brave New Ideas：マルチメディア研究の新たな方向性を示す斬新な研究を集める企画。今回は、「マルチメディア研究の社会的影響」がテーマで、社会的影響が大きいとみなされた、危機管理、環境、疾病対策に関する4つの研究が選ばれた。
- ・ Multimedia Grand Challenge：最先端の企業・研究機関（今回は、マイクロソフト、ヤフー、クイーンズランド工科大）から提案された4つの課題の解法が競われた。
- ・ Open Source Software Competition：新たに開発されたツールやライブラリーなどのソフトウェアについて、その汎用性や有効性に基づいて競う企画。また、男女共同参画を目的とした昼食会「Women in SIGMM Lunch」や、若手研究者の奨励を目的とした以下の企画もあった。
- ・ SIGMM Rising Star シンポジウム：2015年から始まった、最近活躍している若手研究者（ポスドク、助教層）による講演会。研究紹介だけでなく、来歴や着想に

至った経緯などを聴けて面白かった。

- ・博士シンポジウム：博士課程学生のメンタリングを目的にした事前審査制の企画。

3. 次回以降の開催予定

ACM Multimediaは、毎年、米国→アジア大洋州→欧州と巡回して開催されることになっている。次回、記念すべき第25回のACM Multimediaは、2017年10月23日から27日にかけて、発祥地である米国カリフォルニア州に戻り、マウンテンビューにあるコンピュータ歴史博物館で開催される予定である。それ以降は、2018年に韓国のソウル、2019年にフランスのニースで開催されることが決まっている。

なお、今回はフルペーパーとショートペーパーの区別をなくし、6ページ程度の統一フォーマットで査読を経た論文が、査読結果および内容に応じて口頭発表かポスター発表に振り分けられることになりそうである。詳細は、ウェブページ(<http://www.acmmm.org/2017/>)を参照されたい。

4. むすび

【井手の感想】2010年から2011年にかけて1年間在外研究で過ごした思い出深い街ということもあり、連日自転車を借り、久しぶりのアムステルダム滞在を満喫した。実は今回で連続参加20回目を迎えたのだが、参加者が減少し低迷していた2000年代前半に比べれば、昨今は盛り上がっているように思う。しかし、CVPR, CIKMといった他分野のトップカンファレンスでも、従来マルチメディア分野で取組まれていた課題に関する面白くかつ優れた研究が発表されるようになったことを考えると、分野および国際会議としてのマルチメディアの将来は必ずしも楽観視できない。今後はますます、同分野ならではのことは見極めながら研究を進めていく必要があると考えている。

【柳井の感想】アムステルダムには2週間滞在し、ECCVから連続参加したが、CV系会議のマルチメディア化は相

当進んでおり、むしろECCVの方が「マルチメディア」に関して先を行っている印象を強くもった。CVでは、CNNとLSTMの利用による画像と言語の融合が高いレベルで進んでおり、またWebly-supervisedなどWebからの自動画像知識獲得など、以前はマルチメディア系会議の専売特許であった研究もECCVで多く発表されており、そのレベルも極めて高くなっている。動作画像認識などは、ほぼ同じような手法・評価方法の論文がECCVとACMMMの両方で発表されており、しかもECCVの方が、精度が高い研究が多いということもあった。今後、ACMマルチメディアがCVPR, ICCV, ECCVのsecond-tier conferenceにならないためにも、独自性を出して存在意義を明確にする工夫をしていく時期にきていると強く感じた。実際に、ACMマルチメディアに参加している多くの主要研究グループはECCVでも発表しており、技術的な内容はCVの会議、応用的な要素を含んだ研究はマルチメディアの会議、と使い分けている傾向があり、より応用的・実践的な研究を重視するというのも今後の1つの方向であろう。

謝辞 本稿執筆にあたり、快く資料を提供してくださったACM Multimedia 2016実行委員長およびプログラム委員長に感謝する。

(2016年12月22日受付)

井手 一郎 2000年、東京大学大学院工学系研究科博士課程修了。同年、国立情報学研究所助手。2004年、名古屋大学大学院情報科学研究科助教授。2007年、同准教授となり、現在に至る。博士(工学)。正会員。

柳井 啓司 1997年、東京大学大学院工学系研究科修士課程修了。同年、電気通信大学情報工学科助手。2006年、同大学助教授、2007年、同大学准教授、2015年、同大学大学院情報理工学研究科情報学専攻教授となり、現在に至る。博士(工学)。