# ACM Multimedia（ACMMM）2015参加報告

（学生会員）松 井 勇 佑†，崔　セ ミ††，（正会員）山 崎 俊 彦†

本稿は，2015年10月26日から30日まで，オーストラリア・クイーンズランド州・ブリズベンで開催されたThe 23rd ACM International Conference on Multimedia（ACM Multimedia 2015）の様子を報告するものである．ACM Multimediaはマルチメディア分野では最高峰の国際会議の1つであり，最先端技術の研究発表はもちろん，さまざまなマルチメディアに関するイベントがあることも特徴である．

キーワード：ACM Multimedia，マルチメディア，コンピュータビジョン，画像処理，アート

## 1．まえがき

ACM Multimediaは，Association of Computing Machinery（ACM）の分科会Special Interest Group in Multimedia（SIGMM）が主催するフラッグシップ会議であり，1993年開催以来，今回は23回目を迎えた．

ACM Multimediaは採択率が低く，非常に競争が激しい国際会議の1つである．例えば，Google Scholarで“multimedia”という項目を調べると，ACM MultimediaはIEEE，ACMの論文誌等を抑え，堂々のトップスコアである（図1）．

ACM Multimediaは基本的に北米，アジア，ヨーロッパを順番に廻って開催されている．2015年はオーストラリア・クイーンズランド州・ブリズベンのBrisbane Convention and Exhibition Centreで開催された（図2，図3）．ブリズベンはシドニー，メルボルンに次ぐオーストラリア第三の都市であり，航空会社によっては日本から直行便も運航している．会場周辺（サウス・バンク）は1988年世界万博の跡地で，ブリズベン川に沿って開放感あふれる公園やレストラン街がある．学会会場を出て川辺を歩くと，ジョギングや散歩をしている人にあふれ，レストラン街は熱気たっぷりで，ブリズベンの活気を感じることができるエリアである．また，観光地として名高いゴールドコーストも近隣エリアにあり，空港から会場最寄り駅まで電車で移動する際には，ゴールドコースト行の電車に乗ることになる（できることならそのままゴールドコーストまで行ってしまいたいくらいである）．一方，食事もオージー・ビーフやカンガルー肉など現地ならではのものが楽しめる．

multimedia　Scholarを検索

「*multimedia*」に一致する出版物トップ20　詳細

| 出版物 | h5-指標 | h5-中央値 |
|---|---|---|
| 1. ACM International Conference on Multimedia | 45 | 74 |
| 2. IEEE Transactions on Multimedia | 45 | 58 |
| 3. Multimedia Tools and Applications | 33 | 44 |
| 4. ACM Multimedia Systems Conference (MMSys) | 24 | 48 |
| 5. IEEE International Conference on Multimedia and Expo | 24 | 28 |
| 6. IEEE MultiMedia | 22 | 31 |
| 7. IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks | 22 | 27 |
| 8. ACM International Conference on Multimedia Retrieval | 22 | 26 |
| 9. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) | 21 | 32 |
| 10. arXiv Multimedia (cs.MM) | 18 | 26 |
| 11. Mobile and Ubiquitous Multimedia | 17 | 28 |
| 12. IEEE Workshop on Multimedia Signal Processing | 17 | 23 |
| 13. International Workshop on Quality of Multimedia Experience | 16 | 29 |
| 14. Multimedia Systems | 16 | 20 |
| 15. International Conference on Intelligent Information Hiding and Multimedia Signal Processing | 15 | 22 |
| 16. Conference on Multimedia Modeling | 15 | 20 |

図1　Google Scholarランキング

本稿では，2章で会議の概要，3章では来年以降の開催地について紹介し，4章はまとめについて述べる．

## 2．会議概要

会場はコンパクトにまとまり，ホールエリアにレジストレーション・デスクとデモ・ポスタのためのスペースがあり，コーヒーもそこでふるまわれた．また，キーノートはそこから直接つながるホールで，それ以外のセッションは小部屋にてそれぞれ行われた（図4～図6）．

・キーノート

キーノートは本会議初日と3日目の午前に行われた．初

†東京大学 大学院情報理工学系研究科 電子情報学専攻
†東京大学 大学院学際情報学府 学際情報学専攻
"ACM Multimedia (ACMMM) 2015 Report" by Yusuke Matsui, Toshihiko Yamasaki (Department of Information and Communication Engineering, Graduate School of Information Science and Technology, the University of Tokyo, Tokyo) and Saemi Choi (Graduate School of Interdisciplinary Information Studies, the University of Tokyo, Tokyo)



図2　会場からの風景



図4　コーヒーブレイクの様子




図3　会場の入り口



図5　メインホールの様子

日はシドニー大学Judy Kay教授によるビックパーソナルデータからのユーザモデリングに関する研究事例の紹介だった．Kay教授の研究室は，最近Human Computer Interaction（HCI）のための技術，特に個人データを用いた“Know thyself”（汝自身を知れ）を実現するためのインタフェースについて研究している．3日目はMicrosoft Research Redmondの主席研究員であるZhengyou Zhang氏による相互インタラクションが可能なタッチディスプレイ，ViiBoardプロジェクトの紹介が行われた．ViiBoardは，Zhang氏が研究統括をしているMultimedia, Interaction and Experiencesグループの成果である．ViiBoardはRGBDセンサを用いて多数のユーザを認識することができ，各ユーザの過去のディスプレイとのインタラクションを記憶・理解させることで，タッチディスプレイにおけるユーザ体験をより良いものにすることができる．百聞は一見にしかずで，デモビデオは

https://www.youtube.com/watch?v=-xM1rEhOw00

https://www.youtube.com/watch?v=pBny6KfGreU

などから見ることができる．

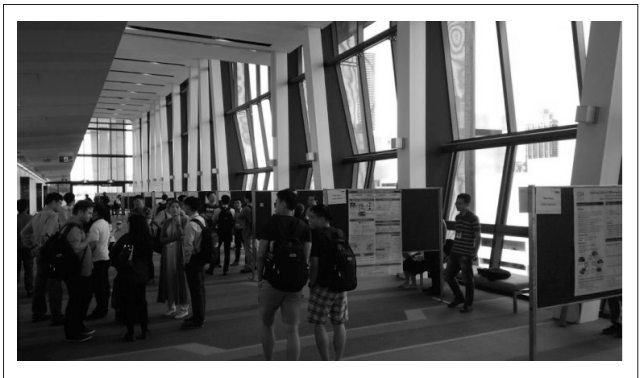


図6　ポスタエリアの様子

・チュートリアル

チュートリアルは，マルチメディア全般における最新の研究を紹介するものである．今年は，10月26日に半日・全日のもの，合わせて下記の8つが提供された．

(1) VM Hub: Building Cloud Service and Mobile Application for Image/Video/Multimedia Services
(2) Interactive Video Search
(3) Learning Knowledge Bases for Multimedia in 2015

(4) An Introduction to Arts and Digital Culture inside Multimedia
(5) Human-centric Images and Videos Analysis
(6) User-centric Cross-OSN Multimedia Computing
(7) Image Tag Assignment, Refinement and Retrieval
(8) Emotional and Social Signals for Multimedia Research

学会初日（本会議前）だったからだろうか，どの部屋も10人を若干越える程度の聴講者数で寂しい感があった．また，発表者を含んで4，5人の規模のテーブルディスカッションを行うところもあった(4)．一般的にチュートリアルでは講演者が研究動向を紹介することが多いが，(8)では，幾つかの論文を聴講者同士で読み合って，論文紹介および改善点などをディスカッションする珍しい形式で行われた．

・オーラルセッション

オーラルセッションは採録されたフルペーパの内容が口頭発表される，ACM Multimediaの花形のセッションである．特に優れたものが選ばれるベストペーパセッションに加え，以下の13のセッションが開かれた．

(1) Multimedia Indexing and Search
(2) Social Multimedia
(3) Emotional and Social Signals in Multimedia
(4) Multimedia and Vision
(5) Multimedia Art, Entertainment and Culture
(6) Telepresence, Virtual and Augmented Reality
(7) Actions and Events
(8) Video Systems
(9) Deep Learning and Multimedia
(10) Multimedia Quality Perception
(11) Multimedia Networking
(12) Data Imperfectness for Multimedia
(13) Multimedia Experiences and Expectations

各セッションでそれぞれ4本程度の論文が発表された．

今年のACM Multimediaのフルペーパの採択率は56/252＝22.2%とのことであり，例年通り極めて競争が激しかったと言える．隣接分野である信号処理系（ICIPやICASSP，4P）やコンピュータビジョン（CVPRやICCV，8P）と比べてページ数が多い（最大10ページ）こともあり，論文中で徹底的な実験による手法比較が求められる，という印象がある．また，Multimediaという名が示す通り，トピックが広範であることもACM Multimediaの特徴である．特に今年のベストペーパセッションでは，群衆解析，ヘッドマウントディスプレイ，ディープラーニング，ディープラーニングのフレームワーク，とさまざまなテーマの論文が取り上げられていた．その中でも群衆解析の内容[1]がベストペーパに，ヘッドマウントディスプレイ[2]がベストステューデントペーパに選ばれた．

・ポスタセッション

ポスタセッションは，採録された4ページのショートペーパのポスタ展示のためのセッションである．ショートペーパは，完成度ではフルペーパほどではないものの，よりチャレンジングな課題が選ばれているとのことであった．とは言うものの，今年のショートペーパの採録率は117/386＝30.3%と非常に難しく，要求されるレベルは高い．今年は1人あたりのポスタの割当て面積が大きく，ポスタ前での議論がしやすかった．

・オープンソースソフトウェアコンペティション

ある意味で一番注目度が高いのがこのオープンソースソフトウェアコンペティションである．これはオープンソースのプロジェクトの開発者がその概要を数ページの論文形式にして投稿し，採録された場合に開発者がその内容を本会議中に発表するものである．その中から最優秀のものが選ばれる．内容としてはマルチメディア処理に関連するライブライリー（例えば，Structure from Motion（SfM），ランダムフォレスト，ディープラーニング，など）のものが多い．開発者は大学院生の個人プロジェクトや，研究室単位での発表が多かったと思う．今年は9件の発表があった．

これまでにオープンソースソフトウェアコンペティションで発表されたライブラリーが，現在各種研究の中心的な役割を果たしている例は少なくない．そのことからも，このコンペティションは非常に注目されていると言える．例えば，Google Scholar上から確認できるACM Multimediaからの引用文献上位三件はこのコンペティションのものである（コンピュータビジョンの包括的なライブラリーであるVLfeat[3]，音の特徴量を抽出するOpenSMILE[4]，ディープラーニングのフレームワークとして，現在最も広く使われているCaffe[5]）．

今年もさまざまなライブラリーが争う激戦であったように感じた．特に，Matlab上からディープラーニングによる処理を非常に簡単に扱えるMatConvNet[6]，およびディープラーニングの分散処理に関するライブラリーであるSINGA[7]の2つは，完成度・有用性ともに高かったと思う．結局優勝したのはSfMのライブラリーであるTheia[8]であった．これはSfMのさまざまな処理をステップステップ切り離して実行できるように工夫されているライブラリーであった．それにより，研究者はSfM全体のうちある部分に着目して研究を行うことができる．私も試してみたが，SfMに関する細かいパラメータ調整などについて親切に初期値が準備されており，新参者がSfMを始めるにはとても便利なものだと感じた．

・アートエキシビジョン

今年もACM Multimediaの特徴の1つであるメディアアート展示が行われた．Art Exhibitでは5件の作品が展示された．より多くの参加者を招くためであろうか，展示はWelcome Receptionを兼ねて本会議初日に開場し，オーラル発表会場のすぐ隣の部屋で本会議の最終日まで続いた．

すべての作品について，本会議最終日に口頭で発表されるセッションがあった（図7）．

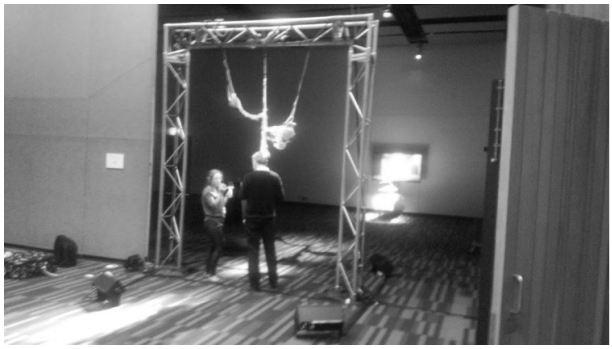


図7　アートエキシビジョンの様子

・グランドチャレンジ

グランドチャレンジは，ACM Multimedia 2009から連続開催されているマルチメディア処理のコンペティションである．参加チームは，各企業から提供されるデータセットとテーマに対して問題を設定し，解法を提案する．今年は，Microsoftはイメージ検索課題を，Yahooはイベント要約の課題を出した．IBMは他のチャレンジとは違い，参加チームが課題設定まで行う形式で，さまざまな問題と解決法が提案された．

今年は以下に挙げる3つのテーマが設定された（括弧中の数字は発表の数を示す）．

1. IBM - New York City 360˚ (2)
2. Microsoft - MSR-Bing Image Retrieval Challenge (2)
3. Yahoo - Yahoo-flickr Event Summarization Challenge (4)

各テーマに参加したチームは，テーマごとに評価を受け，選別されたチャレンジが本会議で発表することになる．評価基準としては，テーマごとに異なっており，Microsoftは評価関数での客観的な評価を行う一方，IBMはデモの質・興味度・拡張性などで評価していた．テーマごとに受賞チームが決まり，また，グランドチャレンジ全体でもGrand Challenge Winnerが選定される．今年の受賞者はYahooのイベント要約チャレンジに参加したEvento 360[9]であった．

・ドクトラルシンポジウム

博士課程学生が博士の間の研究を4ページにまとめ投稿し，採録されたものについてポスタ発表を行うのがドクトラルシンポジウムである．その中で最優秀者が選ばれる．今年の最優秀者はY. Wang（National University of Singapore）であった[10]．

筆者松井も本セッションにて発表した．ACM Multimediaは本セッションを非常に大事にしているようで，投稿した際のレビューコメントが非常に親切かつ有益なものであり，そこでのアドバイスが直接自分の今後の研究に関係してくるほどのものであった．また，セッション中では，シニアな教授がメンターにつき議論するといったイベントもあり，勉強になった．今現在博士課程学生の方々には，ぜひドクトラルシンポジウムへの投稿をオススメしたい．

・Women in SIGMM Lunch

去年に続いて今年も開催された，女性中心のソーシャルプログラムである．女性参加者はもちろん，男性も参加ができる（今年は15%ぐらいの参加者が男性であった）．2014年のSIGMM賞受賞者であるKlara Nahrstedt教授がオーガナイザーを務めた．プログラムの流れは，ランチをしながら，挙げられたトピックに対し意見を交わす形であった．今年のトピックは「あらゆるコンピュータ・サイエンスの学会における女性研究者の参加・活躍の向上」で，参席者の活発な議論のおかげでかなり盛り上がった．本セッションで指摘された多くの問題が解決され，近いうちにWomen in SIGMM Lunchを開催する必要すらなくなれば良いと思う．

・Workshops

ワークショップは，本会議のものと比べて，まだ粗削りで充分な実験結果を伴わないアイデアレベルの発表が多くなされる場である．本年は下記8件のワークショップが開催された．

(1) Multimedia COMMONS - Community-Organized Multimodal Mining: Opportunities for Novel Solutions
(2) 3rd International Workshop on Immersive Media Experiences
(3) International ACM Workshop on Crowdsourcing for Multimedia (CrowdMM 2015)
(4) The 5 th International Audio/Visual Emotion Challenge and Workshop - AVEC 2015
(5) 2 nd Workshop on Computational Models of Social Interactions: Human-Computer-Media Communication
(6) 2 nd ACM International Workshop on Human-centered Event Understanding from Multimedia (HuEvent15)
(7) Workshop on Speech, Language and Audio in Multimedia
(8) The 1 st International Workshop on Affect and Sentiment in Multimedia

初日のチュートリアルが閑散としていたのとは対照的に，ワークショップは粗削りでもこれから伸びてきそうな研究分野の魅力的な発表が集まるからだろうか，聴講者も多く質疑も活発に行われていた（図8）．

・レセプション・バンケット

レセプションは本会議初日に，バンケットは学会最終日の前日にそれぞれ開催された．レセプションは学会会場に隣接するテラスで開催され，軽食と飲み物がふるまわれた．また，ヘビやワニなどの動物がゲストとして登場し，触れ合うこともできた．本当はコアラを呼んでいたのだが予約の都合上でダメになった，という裏話も耳にした．

バンケットは学会会場から歩いて10分程度の場所にあるCity Hallで開催された．荘厳な外見とは裏腹に，内装は最近改装を終えたこともあって，現代風の洒落た雰囲気であった（図9，図10）．




図8　ワークショップの様子



図9　シティーホール



図10　バンケットの様子

## 3．2016年以降の開催について

2016年のACM Multimediaはオランダ・アムステルダムのTheater Tuschinskiでの開催が予定されている．現時点ではまだフルペーパのスケジュールのみが発表となっているが，2016年3月27日が投稿締め切りで，会議自体は2016年10月15日から19日の予定で開催される．ちなみに，コンピュータビジョンでトップクラスの会議であるThe 14th European Conference on Computer Vision（ECCV）が同じくアムステルダムで開催される予定で，関連の研究者が2会議連続で参加できるよう工夫されている．

また，今年の会議中に2017年のBidも行われた．韓国と台湾が立候補し，投票の結果韓国での開催が決まった．これまでのACM Multimediaの参加者を見ると，台湾からは大勢の参加があるのに対して，韓国からの参加者はそんなに多いとは言えない．今回Bidの代表者もコンピュータビジョン分野で著名なProf. Kyoung Mu Leeであった．これをきっかけに新たな研究者がマルチメディア分野に参入してくれればよいと思う．

## 4．むすび

本稿ではACM Multimediaの参加報告を行った．ディープラーニングなどの登場によりモーダル間での垣根が低くなってきており，真の意味でマルチメディア処理がやりやすくなってきた時代にもかかわらず，残念ながらこの分野での日本のプレゼンスは低い．上記でご紹介したようにフルペーパ，ショートペーパ以外にも魅力的なイベントがある珍しい会議でもあるので，ぜひ読者の皆様の参加を促したい．

（2015年12月3日受付）

## 〔文　献〕

1）X. Alameda-Pineda et al.: "Analyzing Free-standing Conversational Groups: A Multimodal Approach", ACM Multimedia（2015）
2）M. Stengel et al.: "An Affordable Solution for Binocular Eye Tracking and Calibration in Head-mounted Displays", ACM Multimedia（2015）
3）A. Vedaldi and B. Fulkerson: "Vlfeat: an open and portable library of computer vision algorithms", ACM Multimedia（2010）
4）F. Eyben et al.: "openSMILE : the Munich Versatile and Fast Open-Source Audio Feature Extractor", ACM Multimedia（2010）
5）Y. Jia et al.: "Caffe: Convolutional Architecture for Fast Feature Embedding", ACM Multimedia（2014）
6）A. Vedaldi and K. Lenc: "MatConvNet: Convolutional Neural Networks for MATLAB", ACM Multimedia（2015）
7）B.C. Ooi et al.: "SINGA: A Distributed Deep Learning Platform", ACM Multimedia（2015）
8）C. Sweeney et al.: "Theia: A Fast and Scalable Structure-from-Motion Library", ACM Multimedia（2015）
9）J. Choi et al.: "Evento 360: Social Event Discovery from Web-scale Multimedia Collection", ACM Multimedia（2015）
10）Y. Wang: "Socializing Multimodal Sensors for Information Fusion", ACM Multimedia（2015）



**松井　勇佑**（まつい　ゆうすけ）　2011年，東京大学工学部電子情報工学科．2013年，東京大学大学院学際情報学府卒業．現在，東京大学大学院情報理工学系研究科博士課程に所属．コンピュータビジョン，マルチメディア処理，特に漫画・線画に対する研究に従事．学生会員．



**崔　セミ**（チェ）　2012年，韓国仁荷大学電子工学卒業，現在，東京大学大学院学際情報学学際情報学専攻修士課程に所属．マルチメディア処理，メディアアート，特にイメージ検索の研究に従事．



**山崎　俊彦**（やまさき　としひこ）　東京大学工学系研究科電子工学専攻博士課程修了．現在，同大学情報理工学系研究科電子情報学専攻准教授．2011年～2013年，コーネル大学訪問研究員．魅力工学，物体認識・機械学習，高速化・最適化処理などの研究に従事．博士（工学）．正会員．