



会議レポート

ACCV2020 会議報告

ACCV が 2020 年 11 月 30 日～12 月 4 日までオンライン開催された。筆者も目まぐるしく発展するコンピュータビジョン関連の動向と、オンライン国際会議の実態を肌で感じるべく参加してきたので、その紹介をしたい。

ACCV について

Asian Conference on Computer Vision (ACCV) とは、アジア最大のコンピュータビジョン (CV) 関連分野の国際会議である。1993 年から隔年でアジア地域において開催され、世界中の研究者、開発者をはじめとした有識者たちが集まり、画像・映像の処理、認識、理解など計算機視覚にかかわるさまざまなトピックについて議論が行われる。コンピュータビジョンについての歴史や詳細は省略するが、身近なところでは、たとえば近年私たちがスマートフォンのアプリで顔認識を用いた画像の加工や AR (Augmented Reality) などを楽しめるのもこういった技術の発展が深くかかわっており、ますます注目を集めている分野である。論文については Springer による出版のほか、Computer Vision Foundation^{☆1} によるオープンアクセスを利用して閲覧が可能である。

第 14 回はオーストラリアのパース、そして第 15 回の開催となる今回の ACCV は京都での開催予定であったが、新型コロナウイルス感染症の拡大防止のためにオンライン開催となった。オンライン会議の在り方が注目を集める中、ACCV2020 で私が感じたのは、世界中のより多くの人々が参加できるようなバーチャル国際会議場としての工夫が凝らされていたことであり、会議場では昼夜を問わず活発な議論が繰り広げられていた。

ACCV2020 の採択率、開催形式

ACCV2020 では計 768 件の論文が投稿され、そのうち 255 件が採択された (採択率 33.2%)。さらにその中の

63 件が Oral 発表 (8.2%)、192 件 (25.0%) が Poster 発表となっており、例年と似た傾向であったと言える。査読は論文 1 本につき 3 人の査読者が付き、加えて 3 人のエリアチェアの協議によって採録の可否が決定された。またこれらの発表に加え、チュートリアルが 2 件と、4 つのワークショップが開催された。

開催形式としては、まず Sococo というプラットフォームを利用したインタラクティブな場としてのバーチャル会議場が非常に印象的であり、詳細は改めて後述させていただくが、全体が俯瞰でき移動できるような会議場の中で、エントランス、Oral 発表用の部屋、Poster 発表用の部屋、スポンサー企業のブース、休憩所などが用意されており、各々好きな場所で過ごしている様子が見られた。また、個々の論文の詳細な情報などについては、参加登録者は専用の Web ページにアクセスすることでいつでも閲覧できるようになっていた。こちらは Mini-Conf とよばれる Web システムをベースとしており、各 Paper についてあらかじめ録画されたプレゼンテーションビデオ (SlidesLive) と質疑応答用のチャット (Rocket.Chat) が一体となったページになっているのだ。プレゼンテーションビデオについては、10 分以内のスライドと発表者が話している様子を録画したビデオが同時に見られるようになっており、個性的な録画も多かった。このようにどのような人が発表者なのか、またその発表内容について事前に窺い知ることも可能であるため、バーチャル会議場でのディスカッションに臨みやすかったというのが私の所感である。

Sococo をベースとしたオンライン開催

前述のように、ACCV2020 では Sococo をプラットフォームとしたバーチャル会議場で行われた (図-1)。参加者はそれぞれが丸っこいアイコンで表示され、Oral 発表の部屋、ポスター発表の部屋などの状況が一目で分かり、それぞれの部屋の間を簡単に移動できる仕組みだった (ちなみに部屋内の席も移動できるが、人が増えすぎると席はもはや関係ない)。つまり実際のポスター発表を回る感覚に近く、人だかりができていときに議論してもよいし、誰かと一緒に回ってもよい。そして全体の様子が見えると、「これだけは聞いておこう」といった発表だけでなく、「これも面白そうだ」といった感覚が自然と生まれてくるように感じた。ほかの参加者についても、誰かが話していると (会話があるとアイコンが明滅するので判断できる) 自然とその部屋に集まっていくような傾向が見られた。

そして驚くことに休憩場所も用意されており、参加者は発表のない時間は大文字山や寺社、庭園などのエリアで休憩し、再び発表のある部屋へ戻っていく様子が見ら

☆1 <https://openaccess.thecvf.com/ACCV2020>

れるなど、常に人が絶えない状態であったのが印象的だった(図-2)。また、部屋内ではZoomに切り替えることもでき、その場合は自動的に招待されるようになっていた。ちなみに私は興味を持った発表などについて軽く予習してから Sococo の会議場の部屋にお邪魔して質問などを行い、発表のない時間などはエントランスに移動して休憩していたが、エントランスで突然 Zoom に招待され、フレンドリーに話しかけられて驚くなどといったこともあった。このような交流の機会も Sococo の会場ならではのことであったと思う。

発表時間帯は4日間共通して4通り(日本時間で11:00～13:00, 15:00～17:00, 18:00～20:00, 22:00～24:00)の枠が設けられており、各発表者について合計2回の発表が、現地の時間帯として無理のない範囲で割り振られていったようだ。

ACCV2020での発表論文について

今回の ACCV から感じる研究動向としては、深層学習に関する論文は実社会への応用をより意識して、大規模ではないデータから深層学習を行うようなものが見られた。また推定対象として、もしくはネットワーク構造の中で3Dに関するものが少しずつ増えている。その他に構造としては、Attention^{☆3}やTransformer^{☆4}に関連する論文が多く見られた。



図-1 Sococoによるバーチャル会議場の様子^{☆2}

実際に参加した範囲でも印象に残った発表はいくつもあったが、本稿では受賞した論文のうちの2つを簡単に紹介したい。1つ目は、Manuel Stoiberらの“A Sparse Gaussian Approach to Region-Based 6DoF Object Tracking”である。この論文が対象とするのは単一のRGBカメラでシーン中の形状既知のオブジェクトの位置と向き(6D)をトラッキングする問題であり、2D画像中でオブジェクトの境界らしさをオブジェクトの法線に沿って確率的に定式化し、さらにそれがガウス分布に従うことを利用して効率の良いトラッキングを実現している。この論文はBest Paper Saburo Tsuji Awardを受賞している。

2つ目の論文は、Hirokatsu Kataokaらの“Pre-training without Natural Images”である。この論文では深層学習における事前学習において、自然画像とその正解ラベルではなく、フラクタル構造を利用して自動的に生成された画像と正解ラベルを用いた学習でも高い精度が得られることを示している。この論文はBest Paper Honorable Mentionを受賞している。

ACCV2022

オンライン会議ならではの良さもあると感じられた ACCV2020 であったが、じっくり腰を据えて議論するにはやはり現地が望ましい。次回の ACCV は2022年の12月4日～8日にマカオで開催される。昨今の新型コロナウイルス感染症の拡大がいち早く収束し、現地で開催されることを切に願う。



図-2 バーチャル会議場で休憩する人々

☆2 ACCV2020 Twitter 公式アカウントより。
 ☆3 自然言語処理の深層学習の手法に由来し、入力と出力シーケンス対して、広範囲にそれぞれの部分に関連しているかも学習できるネットワーク構造。
 ☆4 Attentionと多層パーセプトロンを組み合わせた構造で、エンコーダ部分の構造が画像処理に応用されている。



川原 僚
 (京都大学大学院情報学研究所)