

映像表現&コンピュータグラフィックスの研究開発動向

名手久貴^{†1}, 久保尋之^{†2}, 白石路雄^{†3}, 岡市直人^{†4}

1. まえがき

本稿では、映像表現&コンピュータグラフィックスに関する研究動向について、①映像制作、②CGと画像処理、③画像生成技術における深層学習の利用動向、④3D立体映像の視点から報告する。「映像制作」ではコロナ禍における映像制作、「CGと画像処理」ではACM SIGGRAPH 2020を中心に報告されている。また、「画像生成技術における深層学習の利用動向」ではキャラクターアニメーションや人物の顔画像生成への深層学習利用、「3D立体映像」ではホログラフィやライトフィールド技術の最近の動向について報告されている。

(名手)

2. 映像制作

新型コロナウイルス蔓延防止の影響が映像制作の方法に大きく変化をもたらした期間であった。映像制作には、多数の人数が関わる。特に収録には多数の人が密集することが多い。そのような感染リスクの高い状況を少しでも改善するため、さまざまな工夫を凝らして映像制作が行われた。

密閉空間に多数の人が集まるスタジオ収録において、感染リスクを低減させる工夫として、スタジオに入る人数を減少させることが挙げられる。そこで、スタジオに入る出演者の数を減少させるため、簡易スタジオや出演者の自宅など収録スタジオとは別の場所からリモート出演する例が数多く見られた¹⁾。新型コロナウイルスが広まる前から制作されていた100人の解答者が出演するクイズ番組(フジテレビ「超逆境クイズバトル99人の壁」)では、99人の解答者が、左右前後4方に配置したLED25面マルチ画面に出演した例があった²⁾。スタジオに入る人の数を一挙に減少さ

せた例であった。また、セット内のLED画面に75人の観客をリモート出演させ、リモート出演すること自体を積極的に演出として利用した歌番組(TBSテレビ「音楽の日2020」2020年7月18日)もあった³⁾。

スタジオ撮影においてカメラマンを減らした事例も見られた。2020年6月、テレビ東京は歌番組でロボットアームに乗せたカメラで歌手を至近距離から撮影する検証を音楽番組で行った⁴⁾。ロボットアーム自体の動きとカメラのパン、チルト、フォーカスをスタジオ外の部屋から遠隔操作した。

制作スタッフを分散させる手段として映像信号IP化が一役買った。映像信号のIP化は、2000年代中頃から構想され、2012年SMPTE ST2022、2017年SMPTE ST2110と規格化された。IP化のメリットとして高速に信号を送送できるということもあるが、制作スタッフの分散という観点からは伝送距離の延長が挙げられる。従来のSDI(Serial Digital Interface)の伝送距離が100m程度であるのに対して、IP化された信号はLAN技術によりその制約が取り払われる⁵⁾⁶⁾。NHKが2020年に行った平和祈念式中継では、現場の各カメラの映像、音声を局内の副調に集め、現場スタッフの数を減少させた⁴⁾。このような信号のIP化による拠点分散を行った例は他にもあった。テレビ朝日とソフトバンクが東京6大学野球春季リーグ第8日においてクラウドサーバへ伝送した映像を遠隔でスイッチングした事例⁷⁾やWOWOWによる演奏者の拠点数が2、制作側の拠点が3と合計5拠点を設けたりリモートプロダクション⁴⁾、フジテレビによる出雲駅伝、富士山女子駅伝、フジサンケイクラシックなどのスポーツ中継によるIPリモートプロダクションの実証実験など⁸⁾多くの事例が報告された。映像信号のIP化は以前から行われていたが、コロナ禍の中、スタッフの分散化という新たな目的も加わり、多くの実証実験やそれを利用した映像制作が行われた。

撮影、収録、編集を1台で完結可能な装置「Pixellot」を用いた実証実験も多数行われた^{9)~11)}。「Pixellot」は自動撮影、自動編集機能を持つため、撮影スタッフの数を大幅に減少させることが可能である。サッカー、ラグビーやバレーボールなどのスポーツ撮影に対応可能であり、それらの実

†1 東京工芸大学 芸術学部

†2 東海大学 情報通信学部

†3 東邦大学 理学部

†4 NHK放送技術研究所

"Annual Report on Artistic Image Technology and Computer Graphics" by Hisaki Nate (Faculty of Arts, Tokyo Polytechnic University, Tokyo), Hiroyuki Kubo (School of information and telecommunication engineering, Tokai University, Kanagawa), Michio Shiraishi (Faculty of Science, Toho University, Chiba) and Naoto Okaichi (Science & Technology Research Laboratories, NHK, Tokyo)

証実験が報告された。新型コロナウイルス感染拡大防止のために用いるということもあるが、視聴者が限定されているスポーツの配信や記録用映像の撮影等、小規模撮影に向いていることから、コロナ禍が収束した後も幅広く利用される可能性がある。

コロナ禍関連以外の映像制作関連技術では、本格的にAIが利用されてきていることが挙げられる。日本テレビは駅伝中継においては、選手のユニフォームと顔情報を学習し、ラップタイムデータの自動化や選手間の距離推定に活用した¹²⁾。AmabeTVは、出演者のコメントや天気予報を音声認識し、画面上に表示させた¹³⁾。NHKは、白黒映像のカラー化にAIを利用し、従来数日かかっていた作業が30秒から5分に短縮できたと報告した¹⁴⁾。

TBSは「第61回輝く！日本レコード大賞」において、LEDスクリーンに出演者の動きに対応した映像を表示させるジェネラティブな映像表現による演出を行った³⁾。インタラクティブアートで使用される手法が、一般的な演出手法として認知されてきたことが確認された例といえよう。

コロナ禍の中、密閉空間に多数の人が集まる状況避けるさまざまな工夫が行われた。リモートプロダクションなど一部の手法は、コロナ禍が終焉した後も残り、発展していくことが予想される。将来、コロナ禍が映像制作手法の転換点であったと言われる時期が来るかもしれない。（名手）

3. CGと画像処理

CGと画像処理については、世界最大にして最高峰であるACM SIGGRAPH 2020 Technical Paperプログラム¹⁵⁾での発表をもとに技術動向を紹介する。まず、表1が同プログラムにおけるセッションの一覧であり、合計163件の発表から構成されている。

次に、各発表論文を分野毎に分類するとその内訳は表2の通りである。なお、分類は論文誌¹⁶⁾に掲載されたものを用いた。

各分野が扱う問題の大きさは一律ではないため数値を見て分野間の単純な比較はできないが、個人的には比較的バランスの取れた件数だと感じた。SIGGRAPH 2018では深層学習を用いたImage Processingに関する論文が多い¹⁷⁾との報告もあったが、2020年には深層学習はその適用範囲を大きく広げ、結果的に採択論文数の分野間のばらつきが小さくなったのではないかと推察する。（久保）

4. 画像生成技術における深層学習の利用動向

深層学習は、計算機科学の諸分野において広く利用されているが、前節で述べられたようにコンピュータグラフィックス分野においても例外ではない。例えば、深層学習を用いて、照明条件、カメラパラメータ、キャラクタの姿勢、形状、材質のようなコンピュータグラフィックスで利用されてきた性質を制御して画像や映像を生成する技術の総称としてNeural Renderingという概念が生まれている¹⁸⁾。本

表1 SIGGRAPH 2020 Technical Paperプログラムのセッション一覧
(括弧内は発表件数)

1	RGB++: Depth, 3D and Light Fields (4)
2	Contact and Collisions (4)
3	Modeling and Synthesis (4)
4	Geometric Deep Learning (4)
5	Deformation and Cloth (4)
6	Vector Graphics (4)
7	Capturing and Editing Faces (4)
8	Geometric Foundations (4)
9	Geometry Processing (4)
10	Fluids (4)
11	Physically-Based Character Animation (4)
12	Optimization (4)
13	Tracking (4)
14	Bubbles and Surface Tension (4)
15	Characters and Rigid Bodies (4)
16	Mapping and Parameterization (4)
17	Real-Time Rendering (4)
18	Face and Full Body Motion (4)
19	Fabrication (4)
20	Appearance Acquisition and Inverse Rendering (4)
21	Cinematography (4)
22	Pattern and Color (4)
23	Simulating Natural Phenomena (4)
24	Appearance Modeling (4)
25	Haptics and Touch (4)
26	Developing Geometry (4)
27	Crowds, Paths and Puzzles (4)
28	Image and Video Processing (4)
29	Monte Carlo and Perception (4)
30	Smart Sampling (4)
31	Motion Matching and Retargeting (4)
32	VR Hardware (4)
33	Simulating with Points (4)
34	Shape Modeling (4)
35	Creative Fabrication (4)
36	Making Delaunay and Voronoi Proud (4)
37	Artistic Imaging (4)
38	Differentiable Rendering and Applications (4)
39	Computational Imaging (4)
40	Motion and Matching (4)
41	Systems and Software (3)

表2 分野別掲載件数と割合*

分野名	件数	割合 [%]
Animation / Simulation	37	30.1
Hardware	3	2.4
Imaging / Video	14	11.4
Interaction / VR	9	7.3
Methods & Applications	9	7.3
Modeling / Geometry	32	26.0
Rendering / Visualization	19	15.4

節では、コンピュータグラフィックスに関する主要な国際会議であるACM SIGGRAPH 2020、および、ACM

* SIGGRAPH2020 Technical Paperプログラムで発表された論文のうち Transaction on Graphics招待論文を除くため、総数は表1と異なる。

SIGGRAPH Asia 2020で発表された論文を取り上げて、深層学習を利用した画像生成技術の研究例を紹介する。

キャラクターアニメーションは、その応用分野も広いことから、コンピュータグラフィックスに関する研究の中でも重要な位置を占める。キャラクターアニメーションでは、骨に相当するスケルトンをモーションキャプチャ装置から取得するなどして設定することが多いが、これを映像から学習し、映像中のモーションのスタイルを持ったキャラクターアニメーションを生成する手法が提案されている¹⁹⁾。また、あるキャラクターが持っているモーションを、そのキャラクターとは異なる幾何情報を持つキャラクターに適用させるモーションリターゲットングと呼ばれる処理に、深層学習を利用した手法が提案されている²⁰⁾。それ以外にも、バスケットボールをプレイするキャラクターのように、他のオブジェクトとの接触があるキャラクターのモーションの生成に深層学習を用いる手法が提案されている²¹⁾。

また、キャラクターアニメーション以外にも、多くの分野で深層学習を利用した手法が提案されている。例えば、人物の顔画像をGAN (Generative Adversarial Network) によって生成する技術はすでに多く利用されているが、MichiGAN (Multi-Input-Conditioned Hair Image GAN) という手法では、特に頭髪に着目し、ユーザが指定した頭髪の構造や外観がリファレンスと同じようになるような頭髪画像の生成や、ユーザが描いた形状に合わせた頭髪画像の生成を行える²²⁾。また、モデリングの分野においても、ディスプレイメントマッピングで表現されてきた幾何モデル上の凹凸の生成にGANを利用した手法²³⁾や、平面上に多角形を敷き詰めるタイリング問題にニューラルネットワークを適用した手法²⁴⁾、ユーザのスケッチを入力として、それに合うように液体の飛沫の画像を生成するためにConditional GANを使用した手法²⁵⁾などが提案されている。

このように、コンピュータグラフィックス分野におけるさまざまな問題の解決法として深層学習が取り入れられている。また、3次元グラフィックス向けの深層学習ライブラリーであるPyTorch3D²⁶⁾の提供が開始されるなど深層学習を利用した研究や開発のための環境も整いつつあり、今後も深層学習を利用したコンピュータグラフィックス分野での研究がより活発になることが期待されている。(白石)

5. 3D立体映像

近年、高画素密度ディスプレイやさまざまな光学素子の発展により、自然な3次元映像を表示できるホログラフィやライトフィールドディスプレイ技術は、着実に進歩してきている。本節では、3Dディスプレイ分野に関する研究事例および産業界における最近の動向について報告する。

ホログラフィ技術においては、東京農工大学の高木教授の研究グループが、「ホログラフィック・コンタクトレンズディスプレイ」を開発した²⁷⁾。物体から発せられる波面をコンタクトレンズ内の表示デバイスで発生させて、遠方に

立体像を生成することで、立体像に対する目のピント調節を可能にしている。位相型空間光変調器をレーザ照明するバックライトとして、HOE (ホログラフィック光学素子) を用いることで0.1 mm程度の薄さにしている。本技術による再生像は実物体に重ねて表示できるため、AR分野での応用が期待されている。NTTでは、フィールド全体を見渡せるスポーツ観戦や、複数人で製品をモデリングする協調作業などの用途を想定して、テーブル上の円形スクリーン全周囲から視聴できる360度テーブルトップ3D表示技術の開発を行っている。隣り合う視点映像の輝度を視点位置に応じた輝度比率で合成し提示することで、中間視点の映像を視覚的に補間して知覚させる「リニアブレンドング」という視覚系による知覚メカニズムを表示原理に用いて、従来の1/5のプロジェクタ数でなめらかな運動視差を実現している。本システムにおいて、スクリーンの不要反射面を削減するスクリーン構造も提案されており、迷光の軽減による画質向上を実現している²⁸⁾。また、NHKが研究開発を進めているインテグラル3Dディスプレイでは、直視型ディスプレイに要素画像群を表示し、そこにレンズアレーを密着させることで、薄型の3次元表示系を実現している。しかし、直視型ディスプレイではサブ画素にカラーフィルタが配置されているが、レンズアレーでサブ画素を標本化することによって色モアレ (不要な縞模様) が生じ、3次元映像の品質が低下する。そこで、光線をシフトさせることができる偏光回折素子を用いて、2段の光学的な時分割ウォブリングによりRGB各色を等価的に重ね合わせて一つの画素として表示することで、色モアレを低減する手法を提案した²⁹⁾。これにより、光学系を薄型に保ちながら色モアレの強度を25%まで低減することに成功している。

産業界では、ソニーが高精細な3DCG映像を裸眼で見ることが可能な空間再現ディスプレイ (ELF-SR1) を発表した³⁰⁾。ゲームやCGクリエイター、建築家や車のデザイナーなどをターゲットにしている。また、本装置向けのコンテンツ制作を行う専用SDK (ソフトウェア開発キット) も提供している。本装置は、15.6インチの4Kディスプレイの表面に独自のマイクロオプティカルレンズを貼り付け、観察者の左右の目に視差のある映像を提示する。さらに、高速ビジョンセンサと視線認識技術を用いて観察者の目の位置を検出し、その位置に応じた映像をリアルタイムで生成するため、視点を上下左右に変化させると運動視差が得られる。観察可能な範囲は、ディスプレイからの距離が30~75 cm、視域角が水平50°、垂直60°である。また、ジャパンディスプレイは、5.5型の「ライトフィールドディスプレイ」を開発者向けに販売している³¹⁾。ライトフィールドディスプレイは、見る位置に応じて物体からの光線を再現することで、特別な眼鏡を用いることなく、3次元映像を表示できる。開発キットでは、3次元CGデータをライトフィールドディスプレイ映像に変換・表示するソフトも提供している。用途として、美術品などのデジタルアーカ

イブ、教育・医療分野など、長時間視聴するさまざまな場
面を想定している。バリア方式を採用しており、広い視域
で水平視差を生成している。有効表示画素数は1440×2560
画素、視域は100°となっている。また、Looking Glass
Factory社のLooking Glass³²⁾、Leia社のLume Pad³³⁾、
Dimenco社のSimulated Reality Display³⁴⁾など、海外の企業
からもさまざまなタイプの3Dディスプレイが次々と製品化
されており、今後も動向を注視したい。（岡市）

6. むすび

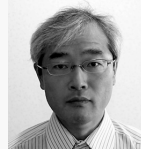
本稿では、映像表現&コンピュータグラフィックスに関
する研究動向について、①映像制作、②CGと画像処理、
③画像生成技術における深層学習の利用動向、④3D立体
映像の視点から報告した。映像制作技術のコロナ禍への対
応、深層学習のCGへの応用、ホログラフィやライト
フィールド技術の着実な進化と映像表現&コンピュータグ
ラフィックスに関する技術が高度化しており、今後も注目
する必要がある。（名手）

(2021年7月26日受付)

〔文 献〕

- 1) 放送技術, 73, 10 (2020)
- 2) 放送技術, 74, 2 (2021)
- 3) 放送技術, 73, 9 (2020)
- 4) 放送技術, 73, 12 (2020)
- 5) 倉掛卓也: “番組制作のための光・IP伝送技術の研究開発動向”, NHK技研R&D, 79, pp.4-16 (2020)
- 6) 川本潤一郎: “映像・音声のIPトランスポート方式”, 映情学誌, 73, 4, pp.79-85 (2019)
- 7) https://www.nikkei.com/article/DGXLRSF539791_T00C20A9000000/
- 8) 真崎晋哉, 大崎雅典, 林和喜, 藤本剛, 小山智史, 北島正司, 鈴木高幸: “IPリモートプロダクションの取り組み実例”, 映情学誌, 74, 3, pp.52-58 (2020)
- 9) 望月貴裕: “映像自動要約技術の最新動向”, NHK技研R&D, 182, pp.4-15 (2020)
- 10) 川本龍文, 川崎拓真: “AIカメラ「Pixellot」によるスポーツ配信継と今後の可能性”, 映情学誌, 74, 3, pp.48-51 (2020)
- 11) <https://pixellot.jp/>
- 12) 佐藤誠, 加藤大樹, 鈴木寿見, 篠田貴之: “放送局の番組制作における画像認識AI技術の活用例”, 映情学誌, 74, 1, pp.35-39 (2020)
- 13) 藤井祐介: “音声認識を使用したAIによる自動リアルタイム字幕システムの開発”, 映情学誌, 74, 3, pp.39-42 (2020)
- 14) 小森智康, 三島剛, 宮崎太郎, 遠藤怜, 河合吉彦: “放送におけるAI活用の取り組み—スマートプロダクション—”, 映情学誌, 74, 3, pp.32-38 (2020)
- 15) ACM SIGGRAPH 2020 Proceedings, Technical Papers (2020)
- 16) S. Rusinkiewicz (editor), ACM Trans. Graph., 39, 4, (2020)
- 17) 向井信彦, 名手久貴, 杉本志織, 山口健: “映像表現およびコンピュータグラフィックスの研究動向”, 映情学誌, 73, 6, p.97-101 (2019)
- 18) A. Tewari, O. Fried, J. Thies, V. Sitzmann, S. Lombardi, K. Sunkavalli, R. Martin-Brualla, T. Simon, J. Saragih, M. Nießner, R. Pandey, S. Fanello, G. Wetzstein, J.-Y. Zhu, C. Theobalt, M. Agrawal, E. Shechtman, D.B. Goldman and M. Zollhöfer: "State of the Art on Neural Rendering", Computer Graphics Forum 39, pp.701-727 (July 2020)
- 19) K. Aberman, Y. Weng, D. Lischinski, D. Cohen-Or and B. Chen: "Unpaired motion style transfer from video to animation", ACM Trans. Graph. 39, 4, Article 64 (July 2020)

- 20) K. Aberman, P. Li, D. Lischinski, O. Sorkine-Hornung, D. Cohen-Or and B. Chen: "Skeleton-aware networks for deep motion retargeting", ACM Trans. Graph. 39, 4, Article 62 (July 2020)
- 21) S. Starke, Y. Zhao, T. Komura and K. Zaman: "Local motion phases for learning multi-contact character movements", ACM Trans. Graph. 39, 4, Article 54 (July 2020)
- 22) Z. Tan, M. Chai, D. Chen, J. Liao, Q. Chu, L. Yuan, S. Tulyakov and N. Yu: "MichiGAN: multi-input-conditioned hair image generation for portrait editing", ACM Trans. Graph. 39, 4, Article 95 (July 2020)
- 23) A. Hertz, R. Hanocka, R. Giryes and D. Cohen-Or: "Deep geometric texture synthesis", ACM Trans. Graph. 39, 4, Article 108 (July 2020)
- 24) H. Xu, K.-H. Hui, C.-W. Fu and H. Zhang: "TilinGNN: learning to tile with self-supervised graph neural network", ACM Trans. Graph. 39, 4, Article 129 (July 2020)
- 25) G. Yan, Z. Chen, J. Yang and H. Wang: "Interactive liquid splash modeling by user sketches", ACM Trans. Graph. 39, 6, Article 165 (Dec. 2020)
- 26) Facebook, Inc., PyTorch3D, <https://pytorch3d.org>
- 27) J. Sano and Y. Takaki: "Holographic contact lens display that provides focusable images for eyes", Opt. Express, 29, 7, pp.10568-10579 (2021)
- 28) 巻口誉宗, 高田英明: “360度テーブルトップ型3D映像表示技術における画質向上手法の提案”, 3次元画像コンファレンス講演論文集, 2-4 (2020)
- 29) H. Sasaki, N. Okaichi, H. Watanabe, T. Omura, M. Kano and M. Kawakita: "Color Moiré Reduction and Resolution Improvement for Integral 3D Displays Using Multiple Wobbling Optics", 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp.109-116 (2020)
- 30) Spatial Reality Display, <https://www.sony.jp/CorporateCruise/Press/202010/20-1016/>
- 31) https://www.j-display.com/technology/jdinew/light_field.html
- 32) Looking Glass, <https://lookingglassfactory.com/>
- 33) Lume Pad, <https://www.lume-pad.com/>
- 34) Simulated Reality Display, <https://www.dimenco.eu/>



名手 久貴 2001年、大阪大学大学院人間科学研究科博士後期課程修了。同年、通信・放送機構高度三次元動画像遠隔表示プロジェクト国内招聘研究員。2003年、東京農工大学産官学連携研究員。同年、東京工芸大学芸術学部助手。2006年、同大学専任講師。2011年、同大学准教授を経て、2016年より、同大学教授。立体視知覚の研究に従事。博士(人間科学)。正会員。



久保 尋之 2011年、早稲田大学大学院先進理工学研究科博士後期課程単位取得退学。同年、同大学理工学術院助手。2012年、キヤノン(株)。2014年、奈良先端科学技術大学院大学助教。2020年、東海大学情報通信学部情報メディア学科特任講師となり、現在に至る。コンピュータグラフィックス、コンピュータビジョンに関する研究に従事。博士(工学)。正会員。



白石 路雄 2003年、東京大学大学院総合文化研究科博士課程修了。2005年、東邦大学理学部情報科学科講師。2013年から、准教授。コンピュータグラフィックスの研究に従事。博士(学術)。正会員。



岡市 直人 2008年、東京大学大学院新領域創成科学研究科修士課程修了。同年、NHKに入局。山口放送局を経て、2012年より、NHK放送技術研究所にて、3次元映像システムに関する研究に従事。正会員。