

メディア工学の研究動向

長谷山美紀^{†1}, 河村 圭^{†2}, 田良島周平^{†3}, 新井啓之^{†4}

1. まえがき

本稿では、メディア工学について最近の注目技術を中心に下の三つのトピックを紹介する。

- (1) 拡張現実・仮想現実の研究動向
- (2) 深層学習の研究動向と実用展開
- (3) IoTとメディア処理

(1)については、デプスセンサやヘッドマウントディスプレイ等のデバイスの登場と性能向上を踏まえつつ、これらを活用した研究開発の進展と動向を紹介する。また、(2)については、加速度的に進む、画像映像メディアへの深層学習技術の適用について、具体的に画像分類や物体検出、さらには画像検索から教師データの検討、実用展開に至るまでを解説する。(3)では、(2)を受けてメディア工学の分野でAI技術と相乗効果を生み出すIoT技術について解説する。実用化に向けた動向として、インフラ構造物の状態監視や人の活動状態のモニタリング、ITSのための画像処理に関する研究動向を紹介する。

2. 拡張現実・仮想現実の研究動向

本章では、拡張現実、仮想現実および複合現実(AR/VR/MR)と呼ばれる分野を紹介する。

2.1 デプスセンサと人体骨格モデル

深度情報を取得可能なデプスカメラとして廉価なKinectが発売され、これらを既知の情報としてさらなる映像表現が検討されている。

青木らは立体物にテクスチャを重畳するプロジェクションマッピングにおいて、対象形状に応じた投影位置の調整

への応用を報告した¹⁾。ただし、単純にセンサ値を信頼するとずれが大きいため、求める精度に応じた改善が必要である。また小林らは、バーチャルスタジオにおける演技者の3次元位置取得への利用を報告した²⁾。ここでの報告の趣旨は合成対象との距離感を振動により演技者に伝えることにある。

さらにKinectでは、人体骨格モデルがSDK経由で容易に取得できる。人体骨格モデルは既知のものとしてさらなる映像表現が検討されている。西浜らはジェスチャ認識に利用して、空気塊を仮想的に操作して視覚的にフィードバックするシステムを提案した³⁾。また野村らは、高価なモーションキャプチャの代替としてKinectを用い、ダンスデータの創作支援を実現した⁴⁾。さらに、Kinectを用いて、ユーザの映像視聴中の動作を取得することで、ユーザの好みを推定可能とする手法も実現されている^{5)~8)}。

ここでデプスセンサは、自動運転への利用のためさらなる高性能化が進む一方で、人体骨格モデルはOpenPoseと呼ばれる映像ベースの認識に移行している^{9)~11)}。これはいずれも深層学習による認識性能の著しい改善が背景にある。またKinectは製造中止が発表されている。

2.2 ヘッドマウントディスプレイ

ヘッドマウントディスプレイ(HMD)の一例として、スマートフォンと専用レンズを用いた廉価な方式が提案され、専用のVRゴーグルが多数発売された。この背景として、スマートフォンの画素密度、解像度、GPU性能が著しく向上し、VRとして実用に耐える性能となったことが挙げられる。また、PC接続可能なヘッドマウントディスプレイとしてOculus社Oculus RiftやHTC社Viveが市販されている。隅田らは、伝統芸能や野外行事の記録・再現の一例としてVRによる練り供養行列の視覚化を報告した¹²⁾。

張らは3枚のディスプレイにより視野を覆い、さらにKinectにより身体全体の移動をジェスチャとして捉える、バーチャル茶室制作を報告した¹³⁾。なお、ディスプレイ応用の研究としては、インテグラル立体表示が精力的に続けられている^{14) 15)}。

HMDの外部情報として位置や方向をトラッキングすることはできても、内部情報として視聴者の視線や表情を読

^{†1} 北海道大学 大学院情報科学研究科

^{†2} 株式会社KDDI総合研究所

^{†3} NTTメディアインテリジェンス研究所/東京大学 大学院情報理工学系研究科

^{†4} 日本工業大学 工学部 情報工学科

"Media Engineering" by Miki Haseyama (Graduate School of Information Science and Technology, Hokkaido University, Sapporo), Kei Kawamura (KDDI Research, Inc., Saitama), Shuhei Tarashima (NTT Media Intelligence Laboratories, Kanagawa/Graduate School of Information Science and Technology, The University of Tokyo, Tokyo) and Hiroyuki Arai (Nippon Institute of Technology, Saitama)

取ることは難しく、現在もデバイスとして研究開発が進められている。視線情報はHMDのみならず、熟練者の技術継承にも活用が模索されている。三改木らと高橋らは堤防点検における利用を報告している^{16) 17)}。また、心理物理分野で知られている定量化手法を対象に被験者依存性を除去する方法を提案している¹⁸⁾。

2.3 拡張現実感 (AR: Augmented Reality)

拡張現実感とは、現実世界の映像にコンピュータグラフィックスなどを重畳する技術であり、その結果として現実世界のナビゲーションやエンタテインメントが実現される。ARもVRと同様にスマートフォンの高性能化を契機として急速に実用化が進んだが、現在ではその発展系としてシースルー型HMD利用による実用化が研究されている。

浦らは、実世界における特定文字列を検出・強調する方式を報告している¹⁹⁾。ARにおけるインタラクションとして、蛭原らは重畳されたオブジェクトの把持に現実感を持たせる方式を報告している²⁰⁾。また、空間中の指の動きをキー入力とするアイデアにおいて、ARにより手指映像にキーボードして現実感を高める方式を報告している²¹⁾。

エンタテインメントとしては、スポーツ映像に対する利用が広く検討・利用されている。例えば、大久保らはバレーボールを対象に、動きが早く選手によるオクルージョンが発生しやすいボールの3次元位置追跡を行い、軌跡を重畳する方式を報告している²²⁾。加えて、サッカー映像を対象として、パス可能な領域や重要選手を視聴者に提示する手法も提案されている²³⁾。

2.4 プロジェクションマッピング

重畳先をディスプレイではなく実空間とする空間型拡張現実感、もしくはプロジェクションマッピングも提案されている。プロジェクタを舞台演出として利用することは古くからあるが、中村らは劇場や舞台のような広域な空間においても装置を大掛かりにすることなく空間を計測する方式を報告している²⁴⁾。

さらに静止物体だけでなく動物体に対する射影方式が、特にリアルタイム性実現の観点から報告されており、庭田らは剛体に対して、宮崎らは非剛体に対する方式を報告している^{25) 26)}。

2.5 国際標準化

VR映像について、動画像符号化方式の国際標準化団体であるMPEGにおいて各種の標準化が検討されている。VR映像の射影方式が各社から提案されており、実際のサービスに利用されている。現在広く利用されていると考えられるのはEquirectangular方式であり、別の例としてはFacebook社Cubic射影方式などが挙げられる。これらの射影方式をシグナリングする方式として、ISO based media file format (いわゆるMP4ファイル形式)におけるOMAF (Omnidirectional Media Application Format)がISO/IEC 23000-20として規格化された²⁷⁾。また、射影方式を考慮し

た動画像符号化方式自体も検討されている。HEVCの次世代符号化方式 (Future Video Coding) として提案募集が2017年10月に開始された²⁸⁾。FVCの規格化完了は2020年後半を予定しており、HEVCに対して30%程度の性能改善が見込まれている。

3. 深層学習の研究動向と実用展開

ILSVRC²⁹⁾における畳み込みニューラルネットワーク (CNN) の成功を端緒として、画像映像メディアへの深層学習技術適用は加速度的に進んでいる。当学会 (学会誌、全国大会、研究会) においても、直近2年の間に30件以上もの深層学習に関する研究報告があり、当該技術への期待と関心の高さが窺い知れる。本章では、当学会のフォーカスに含まれるいくつかの分野における深層学習技術の研究動向と、同技術の実用展開状況を紹介する。

3.1 画像分類

2015年以降、CNNの多層化は飛躍的な発展を遂げた。Deep Residual Networks (ResNet)³⁰⁾はその先駆けといえるモデルであり、目的出力の入力に対する残差学習を実現するResidual Blockの導入により、それまでのモデル (例えば、VGG16/19³¹⁾、GoogLeNet³²⁾) を大幅に上回る、152層もの畳み込み層からなるCNN学習の成功が報告された。ResNetの衝撃は大きく、今日に至るまでさまざまな観点 (例えば、広幅化³³⁾、Feature Mapのダウンサンプリング方法の改良³⁴⁾、Residual Blockの改良^{35) 36)}、Block間の接続方法の改良³⁷⁾) で発展研究が進められている。最新のアーキテクチャ^{34) 36)}を用いたImageNet 1000カテゴリー分類のTop 5分類誤差は現状4%台であり、その性能はすでに人間 (5.1%²⁹⁾) を凌駕している。

さらに、CNNに基づく胃がんリスク推定³⁸⁾や、従来のCNNと比べて少量のデータから学習が可能なDeep Extreme Learning Machine³⁹⁾に基づく火星のダストストーム検出⁴⁰⁾など、広い分野における応用も実現されている。

3.2 物体検出

Region-based CNN (R-CNN)⁴¹⁾以降、物体検出はその発展手法^{42) 43)}の登場により処理が大幅に高速化された。Fast R-CNN⁴²⁾で発明されたRoI Poolingと、Faster R-CNNで導入されたRegion Proposal Networks (RPN)により、R-CNNの冗長性は大きく改善されるとともに、物体検出処理を構成するすべてのパイプラインがニューラルネットで表現され、end-to-endにその処理が最適化できるようになった。文献^{42) 43)}以降の研究の主眼は、大きく、①さらなる高速化 (例えば、文献^{44) ~ 46)}) と、②関連タスクとの融合 (例えば、文献^{47) 48)}) とに分岐しつつある。①の流れでは、出力に近い層の順伝播を簡素化し、一度の順伝播で全処理を完結させるSSD⁴⁴⁾やYOLO^{45) 46)}といった技術が、リアルタイムを大きく上回る処理速度 (～90FPS⁴⁶⁾) で現実的な性能の物体検出を実現し、大きな話題となった。一方②の流

れでは、Faster R-CNN⁴³⁾を拡張することでインスタンス毎のセグメンテーションやポーズを推定する手法^{47) 48)}が提案されている。特にMask R-CNN⁴⁸⁾で提案されたRoI Alignは、RoI Poolingの位置量子化問題を解決した手法として注目されており、今後、精緻な位置特定が求められる諸タスクへの応用が期待される。

3.3 画像検索

ImageNet等で学習したclassification CNN (例えば、文献^{30)~32)})の一部を特徴抽出器とし、距離計量学習でファインチューニングするアプローチが主流である。近年では、データの二つ組や三つ組に定義される損失関数(Contrastive Loss, Triplet Loss)に代わる、データ全体の大域的な構造をより明示的に捉えた損失関数(例えば、Structured Embedding Loss⁴⁹⁾, N-pair Loss⁵⁰⁾, Clustering Loss⁵¹⁾)を用いた手法が提案され、商品⁴⁹⁾、車⁵²⁾、鳥⁵³⁾等の大規模画像データセットにおける大幅な検索性能の向上が報告されている。メディア工学研究会においても、堀口ら⁵⁴⁾によって、特徴抽出器のファインチューニング手法として距離計量学習と分類器学習とを比較評価した報告がなされた。

一方で、SIFT⁵⁵⁾に代表される局所特徴ベースの手法が得意とするランドマーク(例えば、文献⁵⁶⁾)等の検索タスクでは、深層学習の適用による性能改善は現状限定的である。数ある取組みの中でもR-MAC⁵⁷⁾は、classification CNNの畳み込み層出力を領域毎に空間プーリングすることで画像特徴を抽出する手法であり、追加学習なしに局所特徴ベースの手法に匹敵する性能が得られることが報告され話題となった。当学会誌においても、Razavianら⁵⁸⁾によって、畳み込み層の出力から得られた特徴集合のマッチングによって検索を行う手法が提案されている。最近では、R-MACをRPN⁴³⁾と組合せファインチューニングすることで、その性能がさらに向上するといった報告⁵⁹⁾もあり、今後の進展に期待が持てる。

3.4 教師データ

深層学習モデルの訓練には膨大な教師データを要するが、人手でそれを構築するには膨大なコストがかかる。教師データを低コストに構築する手法の研究も活発であり、そのアプローチは大きく、①教師ラベルがない/欠損した実データを用いるものと、②教師データを人工的に生成するものとに分類される。①では、近年ではコンピュータビジョンの知見を活用する研究(例えば、SfM⁶⁰⁾、幾何変換モデル推定⁶¹⁾)が目立つ。メディア工学研究会でも、田良島らによって、画像群中に繰り返し出現する見えのパターンを自動抽出することで、局所特徴記述子学習データセットを構築する方法が提案された⁶²⁾。一方②では、Generative Adversarial Networksを用いて人工画像を実画像風に変換する手法⁶³⁾や、実背景画像にCGモデルを違和感なく重畳した画像を生成する手法⁶⁴⁾、オープンワールドコンピュー

タゲームの画面をハッキングすることで教師データを生成する方法⁶⁵⁾等が提案された。いずれのアプローチにおいても、教師データの構築を軸として、深層学習と周辺研究分野(例えば、コンピュータビジョン、コンピュータグラフィックス、拡張現実)との融合が進みつつあることは大変興味深い。

3.5 実用展開

この2年間で、国内外問わず多くの企業が各社独自のAIプラットフォームを発表し、画像や音声等のメディア認識に基づくソリューションを開始した。画像映像メディアを入力としたものに限定しても、現時点で、顔認証⁶⁶⁾、監視カメラからの人物検出⁶⁷⁾、道路の不具合検出^{68) 69)}といった実課題に対して、深層学習の有効活用がすでに実現している。深層学習技術の発展とともに、この流れは今後より一層加速するであろう。

また、オープンソースの深層学習ライブラリー(例えば、文献^{70)~73)})も数多く発表された。国内企業からも、Chainer⁷¹⁾やNeural Network Console⁷²⁾といったオープンソースが発表されている。特にChainer⁷¹⁾は、その設計思想がTensorFlowやPyTorchといったライブラリーにも影響を与えていると言われている^{73) 74)}。国産の深層学習ライブラリーの今後の更なる発展に期待したい。

さらに近年では、深層学習演算に特化したハードウェア(チップ^{75)~78)}やモジュール⁷⁹⁾の発表も相次いでいる。例えばMovidius⁷⁹⁾は、任意のコンピュータとUSB接続するだけで深層学習演算の高速化が可能な小型モジュールであり、クライアントPCで安価かつ低消費電力に深層学習ベースの認識処理を実行できることで話題となった。また最新のiPhoneシリーズ(iPhone 8, 8 Plus, X)のSoCには、テンソル演算専用のモジュール(Neural Engine)がすでに組み込まれており⁷⁷⁾、モバイル端末上で完結する深層学習ベースの認識処理を実現する基盤も急速に整備されつつある。このような、深層学習ベースのシステムの可用性を拡げるハードウェア群も、深層学習の実用展開をより一層推し進める原動力の一つとなるであろう。

4. IoTとメディア処理

ネットワークサービスの低廉化やクラウドサービスの普及が後押しする形でIoT(Internet of Things:モノのインターネット)が進展しつつある。インターネットに各種センサを始めとしたさまざまな機能を持ったデバイス(モノ)が繋がり、相互に情報を交換したり制御したりすることで、個人の生活、企業活動、社会インフラなどさまざまな場面において変革がもたらされると期待されている。また大量のセンサデータ、音声データ、画像・映像データが自動的に収集されることは深層学習を始めとするAI技術を活用しやすい環境を作ることにもなり、IoT技術とAI技術の相乗効果にも注目が集まってきている。

IoTに関連する技術は多岐にわたるが、画像・映像によるセンシング、および画像・映像と各種センサとの連携等が本学会のメディア工学の研究分野と関係が深い。近年、活発な研究が行われている応用分野として、道路やトンネルなどのインフラ構造物の状態監視、人の活動状態のモニタリング、ITS (Intelligent Transport Systems：高度道路交通システム) のための画像処理などをあげることができる。以下、これらの応用分野毎に主な研究例を紹介する。

4.1 インフラ構造物の状態監視

発電所や空港など高い安全性やセキュリティが必要とされる施設での監視や、多くの人命に直結する河川状況の監視などはニーズが切実であり、カメラをネットワークにつないで撮影された画像・映像を活用することは従来から行われてきた。定点カメラを用いて遠隔地の映像を監視者が目視確認するだけであれば、IoTの範疇には含まれないように思われるが、画像・映像を解析することで異常や危険な状況を自動的に検知する技術や、各種点検作業を効率化または自動化する技術の研究開発が進展しつつあり、カメラはIoTにおける重要なセンシングデバイスの一つになってきていると考えることができる。

近年、従来の定点カメラに加え、車載カメラ、ドローン搭載カメラ等の利用を視野に入れたインフラ構造物の状態監視に関連する研究が盛んに行われている。経済成長とともに整備されてきた道路、橋梁、トンネル等の多くは老朽化しつつあり、国や地方自治体ではこれらインフラ構造物の状態監視と健全性維持が重要な課題となっている。特に道路や各種コンクリート構造物の状態監視のニーズは大きく、数多くの研究が行われている。前田らは、橋梁やトンネルなどの道路構造物のさまざまな変状（ひび割れ、剥落、腐食等）を複数の識別器を用いて高精度に分類する手法を提案した^{80) 81)}。石原らは、地下鉄の車両から撮影した画像からトンネル内面の変状を、CNNを用いて検出、分類する方法を提案した⁸²⁾。林らは、視覚センサによる表面形状計測と打音時の聴覚データ計測を統合することでひび割れの位置と方向を検知する技術を提案した⁸³⁾。小林らは、車載カメラの画像を利用して道路の路面標示の劣化度合いを評価する手法を提案した⁸⁴⁾。これらインフラ構造物の状態監視においては、まだ多くの技術課題が残されているが、実際に現場で監視・点検業務を行っているベテラン技術者のノウハウを積極的に取り込みながら現実解を模索する形で研究開発が行われていることは興味深い。

4.2 人の活動状態のモニタリング

ウェアラブルデバイスを利用したヘルスケアや家庭におけるお年寄りの見守り、カメラやセンサを活用した各種施設の安全管理等、人の活動状態をモニタリングするIoTには多様な可能性が存在している。

永井らはウェアラブルセンサを用いて看護師の勤務形態、勤務状態を推定し、看護師の幸福感との関係について

調べ、勤務時間帯の組合せの変更によって看護師の幸福感が向上する可能性があることを示した⁸⁵⁾。辻らは、ウェアラブルセンサから抽出した身体リズムの統計的分布に着目することで職場等の組織における平均ストレス度を推定する技術について、これを複数の職種に適用するための方法を提案した⁸⁶⁾。いずれの研究も幸福や快適性といった人の内面を豊かにすることを目指しているが、これは今後のメディア工学分野の一つの方向性を示唆しているように思われる。

カメラを活用した人の活動状態モニタリングについては、従来から多くの研究が行われてきたが、最近ではプライバシー保護の問題に注目が集まっている。IoT分野全般に言えることであるが、収集されるデータには個人情報またはこれに準ずるパーソナルデータが含まれることがある。これを活用していくためには個人特定ができないことの担保、つまり匿名化が必要となる。画像・映像については、理想的に言えば、知っている人が見ても誰が写っているかわからない形にすることが必要となる。越前は、カメラの高解像度化により、被撮影者の顔や指紋などといった生体情報が当事者に無断で撮影されたり、画像解析技術を利用することで意図せず他者に読取られたりしてしまう脅威を指摘するとともに対策技術を提案した⁸⁷⁾。山本らは、リアルタイムの映像監視において、監視者に対しても適切なスクランブル処理を施しつつ、画像中人物の動作を検出できるようにすることを目指した手法を提案した⁸⁸⁾。合田らは、プライバシー保護のため画像中の人物の風貌を変換する手法を提案した⁸⁹⁾。IoT分野におけるプライバシー保護は画像・映像メディアが最も遅れているものの一つであり、個別の技術のみならずデータの管理方法や活用のための制度設計も含め早期に具体化し、社会的なコンセンサスを得られるようにすることが必要である。

4.3 ITSのための画像処理

ITSに関しては、カメラやセンサを利用して自動車が周囲の状況を自動的に判定し、運転者や制御にフィードバックをかける技術がすでに実用化されつつある。しかしながら現状では精度やロバスト性に課題が残されており、引き続き活発な研究が行われている。渡辺らは、画像からの人物検出において広く利用されているHOG特徴量を改良し、より誤検出の少ないCoHOG特徴量を開発し、その有効性を確認した⁹⁰⁾。長南らは、交差点における歩行者の検出に基づき左折巻き込み事故を防ぎ、かつ渋滞を発生させにくくするような車速制御方法を提案した⁹¹⁾。なお人物検出については深層学習技術も有効性が高いと考えられ、実用化に向けた今後の動向が注目される。

5. むすび

本稿では、メディア工学について最近の注目技術として拡張現実・仮想現実、深層学習、IoTにフォーカスして研究動

向を紹介した。現在，メディア工学に対する社会の期待は非常に大きくなっており，今後も活発な研究開発および実用化が進展していくものと思われる。（2017年12月13日受付）

〔文 献〕

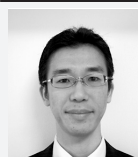
- 1) 青木耀平, 丸山啓介, 橋本直己: “安価な深度センサによる高精度プロジェクトマッピングの実現”, 映情学技報, 40, 5, ME2016-24, pp.89-92 (2016)
- 2) 小林大梯, 千葉匠, 八木伸行: “バーチャルスタジオにおけるインタラクション法の検討”, 映情学技報, 41, 5, ME2017-25, pp.345-348 (2017)
- 3) 西浜正人, 外村佳伸: “ForceIt: 仮想中間媒体を介した領域作用インタラクションの提案”, 映情学技報, 41, 14, ME2017-71, pp.9-12 (2017)
- 4) 野村哲聖, 曾我麻佐子: “Kinectを用いた身体動作入力による振付合成システム”, 映情学技報, 40, 5, ME2016-16, pp.61-62 (2016)
- 5) T. Ogawa et al.: "Human-Centered Video Feature Selection via mRMR-SCMMCCA for Preference Extraction", IEICE Transactions on Information and Systems, E100-D, 2, pp.409-412 (2017)
- 6) Y. Sasaka, et al.: "Multimodal Interest Level Estimation via Variational Bayesian Mixture of Robust CCA", ACM Multimedia Conference, pp.387-391 (2016)
- 7) Y. Ito et al.: "Accurate Estimation of Personalized Video Preference Using Multiple Users' Viewing Behavior", IEICE Transactions on Information and Systems (Accepted for publication)
- 8) 長谷山美紀: “ビッグデータからの価値創出を実現する次世代情報検索～AI・IoT技術の発想支援型検索への導入～”, 信学技報, 117, 329, CS2017-77, pp.101-106 (2017)
- 9) Z. Cao, et al.: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", CVPR (2017)
- 10) T. Simon, et al.: "Hand Keypoint Detection in Single Images using Multiview Bootstrapping", CVPR (2017)
- 11) S-E. Wei, et al.: "Convolutional pose machines", CVPR (2016)
- 12) 隅田圭祐, 曾我麻佐子: “VR技術を用いた練り供養の行列シミュレーションシステム”, 映情学技報, 40, 5, ME2016-17, pp.63-64 (2016)
- 13) 張子亮, 野地朱真, 須藤智: “ヴァーチャル茶室制作に関する試み～マルチディスプレイを用いた日本と中国の伝統的な茶室の提示～”, 映情学技報, 41, 4, ME2017-11, pp.39-42 (2017)
- 14) 佐々木久幸, 岡市直人, 渡邊隼人, 加納正規, 河北真宏, 三科智之: “直視型インテグラル立体表示の色モアレ低減技術～3台合成の検討～”, 映情学技報, 41, 26, ME2017-83, pp.1-4 (2017)
- 15) 岡市直人, 渡邊隼人, 佐々木久幸, 洗井淳, 河北真宏, 三科智之: “複数の直視型ディスプレイパネルを用いたインテグラル立体表示”, 映情学技報, 40, 26, ME2016-82, pp.1-4 (2016)
- 16) 三改木裕矢, 高橋翔, 小川貴弘, 秋山泰祐, 巖倉啓子, 長谷山美紀: “河川管理業務の堤防点検における技術者の視線データ取得の試み”, 映情学技報, 40, 6, ME2016-48, pp.169-172 (2016)
- 17) 高橋翔, 三改木裕矢, 小川貴弘, 長谷山美紀: “堤防点検における技術者の視線データと熟練度の分析に関する一考察”, 映情学技報, 40, 6, ME2016-50, pp.177-180 (2016)
- 18) 岡本健志, 松藤彰宏, 山本昇志: “視線検出を用いた系列位置効果の評価手法の開発”, 映情学技報, 41, 4, ME2017-1, pp.1-4 (2017)
- 19) 浦善晴, 加藤昇, 佐藤美恵: “シースルーヘッドマウントディスプレイを用いた情報探索に関する検討”, 映情学技報, 40, 5, ME2016-11, pp.41-44 (2016)
- 20) 蛭原大樹, 鈴木奏太, 佐藤美恵: “さまざまな形状の仮想物体の掴み動作に着目したARシステムに関する検討”, 映情学技報, 40, 5, ME2016-12, pp.45-48 (2016)
- 21) 岡田雅江, 樋口政和, 小室孝, 小川賀代: “携帯端末向けARタイピングインタフェースのためのオプティカルフロー時系列を用いた入力動作識別”, 映情学技報, 40, 18, ME2016-80, pp.51-55 (2016)
- 22) 大久保英彦, 高橋正樹, 加納正規, 池谷健佑, 三科智之: “被写体追跡可能なスポーツグラフィックスシステムの試作～複数の可動カメラを利用したボールの3次元リアルタイム追跡～”, 映情学技報, 41, 26, ME2017-85, pp.9-12 (2017)
- 23) 高橋翔, 長谷山美紀: “サッカー映像における試合内容の理解を促すデータの可視化”, 映情学誌, 70, 5, pp.722-724 (2016)
- 24) 中村友香, 橋本直己: “複数台プロジェクトを用いた舞台演出システムのための幾何補正手法”, 映情学技報, 41, 4, ME2017-3, pp.9-12 (2017)
- 25) 庭田直也, 橋本直己: “高速な特徴点検出を用いた動的な空間型ARの実現”, 映情学技報, 40, 5, ME2016-15, pp.93-96 (2016)
- 26) 宮崎大希, 橋本直己: “3次元非剛体物体への動的プロジェクトマッピング”, 映情学技報, 41, 4, ME2017-12, pp.43-46 (2017)
- 27) MPEG: "Omnidirectional Media Application Format", <https://mpeg.chiariglione.org/standards/mpeg-a/omnidirectional-media-application-format>
- 28) MPEG: "Future Video Coding", <https://mpeg.chiariglione.org/standards/exploration/future-video-coding>
- 29) O. Russakovsky et al.: "ImageNet Large Scale Visual Recognition Challenge", IJCV (2015)
- 30) K. He et al.: "Deep Residual Learning for Image Recognition", in CVPR (2016)
- 31) K. Simonyan et al.: "Very Deep Convolutional Networks for Large-Scale Image Recognition", in ICLR (2015)
- 32) C. Szegedy et al.: "Going Deeper with Convolutions", in CVPR (2015)
- 33) S. Zagoruyko et al.: "Wide Residual Networks", in BMVC (2016)
- 34) D. Han et al.: "Deep Pyramidal Residual Networks", in CVPR (2017)
- 35) K. He et al.: "Identity Mappings in Deep Residual Networks", in ECCV (2016)
- 36) S. Xie et al.: "Aggregated Residual Transformations for Deep Neural Networks", in CVPR (2017)
- 37) G. Huang et al.: "Densely Connected Convolutional Networks", in CVPR (2017)
- 38) K. Ishihara et al.: "Detection of Gastric Cancer Risk from X-ray Images via Patch-based Convolutional Neural Network", IEEE Int. Conf. Image Processing, pp.2055-2059 (2017)
- 39) E. Cambria, et al.: "Extreme learning machines [trends & controversies]", IEEE Intelligent Systems, 28, 6, pp.30-59 (2013)
- 40) K. Maeda et al.: "Automatic Martian dust storm detection via decision level fusion based on deep extreme learning machine", IEEE Int. Conf. Image Processing, pp.435-439 (2017)
- 41) R. Girshick et al.: "Region-based Convolutional Networks for Accurate Object Detection and Segmentation", TPAMI (2015)
- 42) R. Girshick: "Fast R-CNN", in ICCV (2015)
- 43) S. Ren et al.: "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", in NIPS (2015)
- 44) W. Liu et al.: "SSD: Single Shot MultiBox Detector", in ECCV (2016)
- 45) J. Redmon et al.: "You Only Look Once: Unified, Real-Time Object Detection", in CVPR (2016)
- 46) J. Redmon et al.: "YOLO9000: Better, Faster, Stronger", in CVPR (2017)
- 47) Y. Li et al.: "Fully Convolutional Instance-aware Semantic Segmentation", in CVPR (2017)
- 48) K. He et al.: "Mask R-CNN", in ICCV (2017)
- 49) H.O. Song et al.: "Deep Metric Learning via Lifted Structured Feature Embedding", in CVPR (2016)
- 50) K. Sohn: "Improved Deep Metric Learning with Multi-Class N-pair Loss Objective", in NIPS (2016)
- 51) H.O. Song et al.: "Deep Metric Learning via Facility Location", in CVPR (2017)
- 52) J. Krause et al.: "3D Object Representations for Fine-Grained Categorization", in ICCV Workshop (3dRR) (2013)
- 53) P. Welinder et al.: "Caltech-UCSD Birds 200. Technical Report CNS-TR-2010-001", California Institute of Technology (2010)
- 54) 堀口ほか: “深層特徴の比較評価～分類器学習vs.距離計量学習～”, 映情学技報, ME2017-59 (2017)
- 55) D.G. Lowe: "Distinctive Image Features from Scale-Invariant Keypoints", IJCV (2004)
- 56) J. Philbin et al.: "Object Retrieval with Large Vocabularies and Fast

- Spatial Matching", in CVPR (2007)
- 57) G. Tolias et al.: "Particular Object Retrieval with Integral Mal-Pooling of CNN Activations", in ICLR (2016)
- 58) A.S. Razavian et al.: "Visual Instance Retrieval with Deep Convolutional Networks", ITE Trans. On MTA, 4, 3 (2016)
- 59) A. Gordo et al.: "Deep Image Retrieval: Learning Global Representations for Image Search", in ECCV (2016)
- 60) F. Radenović et al.: "CNN Image Retrieval Learns from BoW: Unsupervised Fine-Tuning with Hard Examples", in ECCV (2016)
- 61) V. Balntas et al.: "HPatches: A Benchmark and Evaluation of Handcrafted and Learned Local Descriptors", in CVPR (2017)
- 62) 田良島ほか: "グラフ分割を用いた画像からの繰り返しパターン抽出", 映像学技報, ME2017-89 (2017)
- 63) A. Shrivastava et al.: "Learning from Simulated and Unsupervised Images through Adversarial Training", in CVPR (2017)
- 64) H.A. Alhaija et al.: "Augmented Reality Meets Computer Vision: Efficient Data Generation for Urban Driving Scenes", in BMVC (2017)
- 65) S.R. Richter et al.: "Playing for Data: Ground Truth from Computer Games", in ECCV (2016)
- 66) <http://jpn.nec.com/physicalsecurity/solution/watch.html>
- 67) <http://www.ntt.com/about-us/press-releases/news/article/2017/0712.html>
- 68) <http://www.ntt.com.co.jp/news/cf16113001.html>
- 69) <http://www.fujitsu.com/jp/about/resources/case-studies/cs-201710-kawasaki-geological-engineering.html>
- 70) M. Abadi et al.: "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems", in CoRR (2015)
- 71) S. Tokui et al.: "Chainer: A Next-Generation Open Source Framework for Deep Learning", in NIPS Workshop (2015)
- 72) <https://dl.sony.com/ja/>
- 73) <http://pytorch.org/about/>
- 74) <https://research.googleblog.com/2017/10/eager-execution-imperative-define-by.html>
- 75) N.P. Jouppi et al.: "In-Datcenter Performance Analysis of a Tensor Processing Unit™", in ICSA (2017)
- 76) <http://pr.fujitsu.com/jp/news/2016/11/29.html>
- 77) <https://www.apple.com/jp/newsroom/2017/09/the-future-is-here-iphone-x/>
- 78) <http://consumer.huawei.com/en/press/news/2017/ifa2017-kirin970/>
- 79) <https://www.movius.com/>
- 80) 前田圭介, 高橋翔, 小川貴弘, 長谷山美紀: "道路構造物の維持管理における高精度な変状分類に関する検討〜タグデータと撮影画像に基づく識別結果の統合〜", 映像学技報, 41, 5, pp.51-54 (2017)
- 81) K. Maeda et al.: "Distress Classification of Road Structures via Adaptive Bayesian Network Model Selection", ASCE Journal of Computing in Civil Engineering, 31, 5, pp.04017044-1-04017044-13 (2017)
- 82) 石原賢太, 高橋翔, 小川貴弘, 長谷山美紀: "畳み込みニューラルネットワークを用いた地下鉄トンネルにおける変状検出に関する検討", 映像学技報, 41, 5, pp.81-86 (2017)
- 83) 林鍾勳, 藤井浩光, 山下淳, 浅間一: "視覚センサおよび聴覚センサを用いたコンクリートひび割れ検出", 映像学技報, 41, 4, pp.85-88 (2017)
- 84) 小林大輝, 加藤邦人: "車載カメラによる道路標示劣化評価手法の検討", 映像学技報, 41, 29, pp.7-10 (2017)

- 85) 永井勇輝, 佐藤信夫, 川本英嗣, 今井寛, 島岡要, 矢野和男: "ウェアラブルセンサを用いた看護師の不規則な勤務形態と幸福感に関する研究", 映像学技報, 41, 28, pp.29-34 (2017)
- 86) 辻聡美, 佐藤信夫, 矢野和男: "ウェアラブルセンサを用いた身体リズムによる組織の状態推定技術", 映像学技報, 41, 28, pp.25-28 (2017)
- 87) 越前功: "現実とバーチャルの境界で情報を制御する〜メディアセキュリティ・プライバシーのこれから〜", 信学技報, 117, 201, EMM2017-50, pp.63-64 (2017)
- 88) 山本恭徳, 李雪霏, 島浦神吾, 姜錫, 坂本雄児: "監視カメラ画像の可逆的プライバシー保護技法", 映像学誌, 71, 3, pp.J110-J117 (2017)
- 89) 合田隆一郎, 高橋裕樹: "プライバシー保護のための風貌変換", 映像学技報, 41, 5, pp.115-120 (2017)
- 90) 渡辺友樹, 伊藤聡, 横井謙太郎: "画像認識による歩行者検出のための特徴量抽出技術の開発", 映像学誌, 71, 1, pp.J28-J34 (2017)
- 91) 長南孝弘, 河野隆二: "交差点における交通安全のための歩行者情報を用いた自動車速度制御", 映像学技報, 41, 5, pp.367-369 (2017)



長谷山美紀 1986年, 北海道大学工学部卒業。1988年, 同大学大学院工学研究科修士課程修了。1989年, 同大学電子科学研究所助手。1994年, 同大学大学院工学研究科助教授。1995年, ワシントン大学客員准教授。2006年, 北海道大学大学院情報科学研究科教授となり, 現在に至る。マルチメディア信号処理・次世代情報アクセスシステムの研究に従事。当学会副会長, ITE Transactions on Media Technology and Applications Editor-in-chief等を歴任。博士(工学)。正会員。



河村 圭 2004年, 早稲田大学理工学部電子・情報通信学科卒業。2005年, 同大学大学院国際情報通信研究科修士課程修了。2010年, 同大学大学院国際情報通信研究科博士課程修了。同年, KDDI(株)入社。現在, (株)KDDI総合研究所超臨場感通信グループ研究マネージャ。2017年, 当学会鈴木記念奨励賞受賞。主に, 動画像符号化方式の研究・開発および国際標準化に従事。博士(国際情報通信学)。



田良島周平 2009年, 東京大学工学部卒業。2011年, 同大学大学院新領域創成科学研究科修士課程修了。同年, NTT入社。現在, NTTメディアインテリジェンス研究所研究員, 東京大学大学院情報理工学系研究科博士課程在籍。画像認識に関する研究開発に従事。



新井 啓之 1989年, 東京理科大学理工学部物理学卒業。1991年, 北海道大学大学院理学研究科修士課程修了。同年, 日本電信電話(株)入社。図面認識技術, 画像認識技術の研究開発と実用化に従事。2017年より, 日本工業大学工学部情報工学科教授となり, 現在に至る。2001年〜2006年, 情報通信研究機構(NICT)ナチュラビジョンプロジェクト特別研究員。博士(情報科学)。正会員。