

人を観る

八木 康史[†] 中村 裕一^{††}

1. はじめに

コンピュータビジョンとイメージメディア研究会では、2001年7月と9月の2回の研究会にまたがって、オーガナイズドセッション「人を観る」を開催し、この分野の現状と将来像に関する講演と討論を行った。その1つのまとめとして、本特集号「人を観る」が企画された。

その構成は、展望論文・総合論文と一般論文からなり、前者では主に全身を対象とした研究に焦点を当て、その分野で活躍されている方々に執筆を依頼した。人体のようにはっきりした内部構造を持ちながらも外見がさまざまに変化する対象の計測、追跡、認識について概観することができる。また「人を観る」研究が扱う対象は人物の全身像から顔の造作まで幅広く、また処理内容も、検出から追跡、さらには3次元動作の解析や個人認識と多岐に渡る。そのため、一般論文では、全身に関する研究に限らず、人間を観る研究全般を公募の対象とし、優れた研究論文を集めた。

2. 「人を観る」特集号の論文

「人を観る」研究は、指、腕、頭部などの3次元形状や運動の解析ならびに全身動作の認識に関する基礎研究から、バーチャルリアリティ、サーベイランスやヒューマンインタフェースでの利用を想定した応用研究まで、さまざまな視点からさかんに研究が行われている。以下、展望論文、総合論文、研究論文に対し、位置づけを考えつつ内容について概観する。

山本正信氏による展望論文「コンピュータビジョン技術による次世代モーションキャプチャ」では、人体幾何モデルをもとにした動画からの身体動作の推定という問題に対する現状の技術と今後の展望を示してもらった。筆者らの用いている多関節の人体モデルに対する順運動学または逆運動学に基づく位置・姿勢推定方法を中心に、そのために必要となる特徴抽出から3次元運動推定までの一連の処理までが分かりやすく解説されている。

続く美濃導彦氏による総合論文「3次元人体モデル中心処理」ならびに岩井儀雄、谷内田正彦両氏による総合論文「動画からの人物動作計測と認識」では、ともに身体動作の復元を目的とした技術が解説されている。美濃氏は、まず3次元の人体モデルを用意し、画像と一致するようにこれを变形させていく方法を提案している。入力データとしては、2次元画像データまたは2.5次元データが用いられ、たとえば正面と側面の2つの方向からのシルエット画像と、人体モデルから計算したシルエット画像間で照合を行うことにより姿勢を求めている。この手法の特徴は、標準人体モデル、または、統計量を用いた個人の推定人体モデルを利用する点であり、画像とモデルの厳密な対応付けがなくとも、また部分的な観測データでも、姿勢の推定が可能となっている。一方、岩井氏の場合は、動画による身体動作計測・認識手法で、対象となる身体の形状モデルを变形させる点では、上記アプローチと共通点を持つが、形状モデルを用いる場合の問題点である計算コストを削減するために、手先などの身体上の特徴点を動画中で追跡することで、運動復元を行っている点が大きく異なる。また動き情報により次フレームでの姿勢を推測することで、解の信頼性を高めている点も特徴である。

次の、白井良明、三浦純両氏による「複雑背景における人物の追跡」は、人物の追跡に関する人を観る研究を紹介している。人物の追跡はサーベイランスやヒューマンインタフェースなどの分野で最も重要な技術の1つである。そこで、両氏にさまざまな条件下での複数人物の追跡方法について、一連の研究をまとめてもらった。

最後に、ヒューマンインタフェースのための身体動作の計測・認識に関する研究を、久野義徳氏による総合論文「ポインティングデバイスとしての身体動作」、および岡隆一氏、西村拓一氏、矢部博明氏による総合論文「ジェスチャー動作の動画からのスポッティング認識について」としてまとめていただいた。人間の意思を伝達するための手段としては、指示操作のように正確な位置・姿勢が要求されるものもあれば、ジェスチャーのように動作の意味が重要なものまである。久野氏の論文では、手先の位置・姿勢や顔の向きを画像処

[†] 大阪大学大学院基礎工学研究科

^{††} 筑波大学機能工学系

理により推定することで、ポインティングに利用する方法が述べられている。一方、岡氏らの論文では、動画像で観測される人物のジェスチャを実時間で認識する手法と、ジェスチャ認識と音声認識を組み合わせたマルチモーダルな認識の手法から、CG や音声合成でユーザと対話するシステムの構築など、著者らが RWC プロジェクトで手がけてきた研究について述べられている。

一般投稿に関しては、11 件の投稿に対し、厳正な査読の結果、5 件が採録された。1 件が追跡に関する基礎的な研究で、残りの 4 件が応用を意識した研究である。

採録された論文の概要は、以下のとおりである。

- (1) 「確信度付き仮説群の相互作用に基づく複数対象追跡」
複数の人物が行き来する空間では、人物のさまざまな動きやオクルージョンなどに対して頑健な人物追跡手法が必要となる。本論文では、condensation による確率論的な枠組みに基づき、各対象(人物)ごとに複数の仮説を生成し、画像情報の時間的・空間的連続性と一貫性を利用することで、複数対象を同時追跡する手法を提案している。
- (2) 「3D 似顔絵フィギュア製作の実用化の試み」
本論文は、人物頭部全周形状データから顔特徴点を自動的に抽出し、似顔絵を生成、さらに顔フィギュアを生成する手法である。従来 2 次元であった似顔絵システムを 3 次元の顔フィギュア生成に拡張した。その結果が、似顔絵作家の作品と類似したデフォルメであったことは大変興味深く、今後の実用化が期待される。
- (3) 「全方位画像列からの個人識別」
全方位視覚センサは、広範囲を同時に観測できることからサーベイランスなどでの利用が期待されている。この論文では、全方位動画像で抽出された顔の系列から個人同定を行う手法を提案しており、全方位視覚によるサーベイランス

の新たな展開を示すものである。

- (4) 「動作者の視点からのカメラ画像によるモーションキャプチャとロボット作業教示への応用」
ヒューマノイドロボットに関する研究がさかんになり、家庭内での活躍が期待されている。家庭内での利用のためには、ロボットに対して簡単に作業の教示ができなければならないが、この論文では、そのために、実演者の手先と作業環境内に取り付けたマーカを同時観測することで、実演者の指先の動きと頭部の動きを計測、認識する手法が提案されている。
- (5) 「複数教示動作の時系列上での統合に基づく人間作業のモデル化手法」
ロボットに対して提示された複数回の同一動作をデータグロブとステレオカメラで計測し、その結果から作業の本質を表す動作を抽出することで、ロボットに人間の動作を模倣させる方法を提案している。上記のロボット作業教示同様、ロボットと人間のインタラクションのために今後ますます重要となる分野である。

3. おわりに

以上、本特集号には、展望論文、総合論文、研究論文合わせて 11 編の論文が掲載されている。これらの論文は、コンピュータビジョンの技術を用いた身体動作の計測や認識処理の現状をよく表したものになっていると同時に、将来の技術を展望するための良い指針を与えてくれるものとなっている。今後、この分野の研究を行う方々の参考になれば幸いである。

末筆ながら、本特集の編集にあたり、貴重な研究成果を投稿いただいた著者、厳しいスケジュールの中、短期間での査読をお引き受けいただいた査読委員、多忙の中編集にご協力いただいた編集委員、ならびに「人を観る」オーガナイズドセッションの運営でご支援いただいた CVIM 研究会運営委員の方々、そして情報処理学会の渡辺様、CVIM 事務局の斉木様に心から深く感謝いたします。