

学位論文及び審査結果の要旨

横浜国立大学

氏名	菅沼 雅徳
学位の種類	博士(工学)
学位記番号	環情博甲第1939号
学位授与年月日	平成30年9月15日
学位授与の根拠	学位規則(昭和28年4月1日文部省令第9号)第4条第1項及び横浜国立大学学位規則第5条第1項
研究科(学府)・専攻名	環境情報学府 情報メディア環境学専攻
学位論文題目	動画像における顕著性のあるイベント検出に関する研究 A Study on Salient Event Detection in Videos
論文審査委員	主査 横浜国立大学 教授 長尾 智晴 横浜国立大学 教授 田村 直良 横浜国立大学 教授 森 辰則 横浜国立大学 准教授 富井 尚志 横浜国立大学 講師 白川 真一

論文及び審査結果の要旨

近年ビッグデータの解析の必要性が急速に高まっているが、中でも監視カメラの増大などに伴って大量の動画像データが蓄積されるようになっており、ビッグデータとしての動画像をコンピュータで処理し、有効な情報資源として活用するニーズが高まっている。動画像を最大点有効に活用するためには、動画像の自動解析を行ない、意味がある部分を自動抽出する必要がある。しかしながら、それは対象とする動画像に依存するとともに、抽出すべき重要な部分や事例(正例)・抽出する必要がなくあまり重要ではない部分や事例(負例)が与えられている場合とそうではない場合があり、自動的に抽出することはこれまでは困難であった。これに対し本論文は、抽出すべき部分を他とは異なる特徴をもつ目立つ部分という意味から“顕著性のあるイベント”と定義し、それらを動画像から自動的に抽出する動画像処理・認識方式を開発している。論文中では、動画像中に頻繁に出現すると考えられる顕著性のないイベントを「正例」、それらに対して出現頻度は低い重要なイベントを「負例」とし、正例・負例の両方が与えられている場合、正例だけが与えられている場合、両方与えられていない場合のそれぞれについて提案手法の有効性を検証したものである。なお、負例だけが与えられている動画像は通常では考えられないため、本論文では対象としていない本論文は序論(第1章)と結論(第6章)を含めて6章から構成されている。

第2章「関連研究」では、始めに第3章で取り扱う覚醒下脳腫瘍摘出術、動画像を用いた手術工程の解析に関する先行研究について述べた後、正例データのみを用いたイベント検出に関する先行研究、正例および負例のラベルデータを用いないイベント検出に関する先行研究を紹介している。最後の例としては、GWR (Grow When Required)、刺激の選択性を用いた領域検出ネットワーク、適応的背景モデル、スパースコーディングに基づいたオンライン学習手法について述べている。

第3章「正例および負例のラベルデータを用いたイベント検出」では、動画像中の正例と負例の両方が与えられた場合の手法を提案している。処理対象としては、申請者が所属する研究室と東京女子医科大学 先端生命医科学研究所(FATS)との共同研究として実施されたプロジェクトにおいて、FATSが保存している脳腫瘍摘出術の記録動画を扱っている。この場合の負例(抽出対象)は、熟練の執刀医が行った電気刺激のタイミングと電極先端位置である。ここでは、正例・負例のもつ画像特徴量を用いる手法を提案するとともに、医用応用の観点から、検出漏れを限りなくゼロにするための措置として、複数の手法の統合方式を採用している。実験において、提案手法の精度を検証し、有効性を確認している。

第4章「正例のラベルデータのみを用いたイベント検出」では、正例のみが与えられている場合について検討している。対象としては、多数の歩行者が登場する監視映像からのイベント検出を扱っている。ここでは、正例として登場する多数の歩行者の中に、スケートボードや自転車などに乗った人物や、自動車などがたまに登場する。これらの人物や車両がこの場合の負例(抽出対象)である。ここでは、Convolutional Autoencoderによる復元誤差を

利用した手法を提案している。正例を学習する際、第1フレームと第3フレームからそれらの中間の第2フレームを生成する **Autoencoder** を作っている。これにより、正例の場合には正しく再現できる回路網を生成した後、動画像に適用することによって、正しく復元できないものとして正常とは少し異なる負例を検出しようとするものである。提案手法を世界的なベンチマーク問題として知られている動画像データセットである **UCSD Pedestrian dataset** に適用したところ、問題の独自性を利用した従来手法より高精度な認識率を達成したと述べている。

第5章「正例および負例のラベルデータを用いないイベント検出」では、正例および負例が与えられていない場合について考察している。対象としては、監視カメラ映像を対象にしている。一般に監視カメラは様々な場所に設置され、設置された場所や状況によって正常・異常は異なる。このため人があらかじめ正例・負例を与えたり、長時間に渡る動画像に対してそれらのラベルをつけることは不可能である。このため、単に動画像を入力するだけで、コンピュータが自ら正例・負例を判断して学習することが期待される。本章では、画像の各位置の情報量の変化を検出し、それが新たな情報に新規性があればそれを画像に反映させて **Saliency** として表示するとともに、刺激が継続されて頻繁に入ってくるものについては馴化して徐々に反応が低下するような独自のネットワークモデルを提案している。そして提案手法を、監視カメラが固定されている場合と、自動旋回する場合に適用している。実験では、世界的なベンチマーク問題だけでなく、学内で撮影された実動画像に対して適用し、従来手法と比較することによって提案手法の有効性を定量的に評価している。

第6章「結論」では、本論文の成果を総括している。

以上のように、本論文では、今後膨大に蓄積されていくと考えられるビッグデータの一種である動画像データに対して、正例・負例が与えられている場合、正例だけが与えられている場合、いずれも与えられていない場合に自動的に処理して関心のある、顕著性のあるイベントを選択的に抽出する動画像処理・認識手法を提案し、実験を通してそれらの有効性を検証したものであり、学術的新規性が高いだけでなく、監視カメラ映像処理など、実用上も利用価値が高い。本論文の成果は、査読付き論文誌の清論文2編、正論文相当のフルペーパー査読付き国際会議論文2編、および多数の口頭発表などを通して公表され、いずれも高く評価されている。

以上から、本論文は博士（工学）の学位論文として十分な価値を有すると審査員全員一致して認め、平成29年7月12日（水）13時00分から14時30分まで、総合研究棟1階S113室において博士論文公聴会を開催した。博士論文公聴会では充実した質疑応答がなされた。

同日14時30分から15時00分まで、総合研究棟1階S113室において審査委員全員出席のもと、菅沼雅徳君の博士論文最終試験を行った。始めに、同君より同君の博士論文「動画像における顕著性のあるイベント検出に関する研究（A Study on Salient Event Detection in Videos）」の概要と、博士論文に関連する学術論文の公表状況、大学院講義の単位取得状況などについて説明があった。これに引き続き質疑応答を行った。審査員からは博士論文に関する質問、一般専門知識に関する質問などがあり、学力及び博士号取得資格の確認を行なった。発表論文については、学術論文（正論文）が2編、査読付国際会議論文（正論文相当）が2編採録されていることから、博士号取得の資格を十分に満たしていることを確認した。外国語については、本人が国際会議において英語で発表していることから学力を確認した。また、履修単位が修了要件を満たしていることを確認した。以上により、同君は最終試験に合格であると、審査委員全員一致で判定した。

以上の論文審査委員会の結論に基づき、平成29年7月24日に開催された環境情報学府 情報メディア環境学専攻会議にて審議し、全員一致で本論文を博士（工学）の学位論文としての価値があるものとして環境情報学府教授会に付議することを決定した。その後、環境情報学府学務委員会での確認を経て、平成29年9月4日（月）に開催された環境情報学府教授会において審議を行い、無記名投票により、菅沼雅徳氏に博士（工学）の学位を授与することを決定した。