

審査の結果の要旨

氏 名 入江 豪

本論文は、「Video Affective Content Modeling for Analysis, Search, and Editing (情動的印象のモデル化に基づく映像の解析, 検索と編集)」と題し, 英文で書かれており, 7章よりなる。映像内容解析のゴールは, 人のように映像を理解することにある。そのためには, 映像のセマンティクスばかりでなく, 情動的要因の取り扱いが極めて重要である。例えば, 映像シーンが引き起こす感情をとらえることができれば, 人をよりひきつけるような要約を生成したり, 感情の好みも合わせて映像を検索したりすることが可能になる。本論文では, この映像内容の情動解析という新しい課題に取り組み, そのモデリング技術を中心として技術的な解決手法を提案, 評価し, 応用までも論じたものである。

第1章は, 「Introduction (序論)」であり, 研究の背景, 本論文の構成について述べている。情動的な情報処理に関しての取り組みを概観している。

第2章は, 「Degree-of-Edit: A Content-based Ranking Measure for Interest-Oriented User Generated Video Search (Degree-of-Edit: ユーザ生成映像の興味に基づいた検索のための内容に基づくランキング尺度)」と題し, ユーザ生成映像 (UGV) の新しいランキング尺度 Degree-of-Edit (DoE) について論じている。インターネット上の膨大な量に及ぶ UGV を検索するために通常は再生回数や投稿日時が用いられているが, それらの指標は映像内容を反映していない。ユーザがより興味をもつと推察される映像を, 映像と音響の内容を反映した編集の度合いに基づく新しい尺度としての DoE を提案した。70,000 以上の UGV を用いた実験を通して, ユーザスタディを行い, その有効性を示した。

第3章は, 「Attractive Face Keyframe Extraction based on Facial Parameters (顔のパラメータに基づく魅力的なキーフレームの抽出)」と題する。ホームビデオから印象に強いキーフレームを抽出する処理について論じており, 映像中の人の顔の数, 大きさ, 位置という構図的な要因や笑顔という情動的な要因を利用して, よいキーフレームのランキングを行った。ユーザスタディにより, ユーザの作成したランキングとの比較評価をし, 有効性を確認している。

第4章は, 「Joyful, Sad, and Exciting Video Segments Extraction based on Emotional Audio Event Detection (感情的なオーディオイベントの検出に基づく喜び, 悲しみ, 興奮のビデオセグメントの抽出)」と題する。Pleasure と Arousal が人の感情の大きな二つの要因であり, 前者に関連すると悲しみ, 後者に関連するの検出を行った。低次音響特徴量を用いて, 一般化状態空間モデルにより, 笑い声や大声といった感情的なオーディオイベントを検出した後, 感情の分類を行った。UGV に対しての評価を行い, 73.3%の精度を得た。

第5章は, 「Affective Scene Classification based on Affective Audio-Visual Words and Latent Topic Driving Modeling (情動的なオーディオビジュアルワードと潜在トピック駆動モ

デルに基づく情動的シーン分類)」と題し、映画シーンをユーザに起こる感情（8感情）に分類することを行った。まず、特徴抽出としては、感情に特化したオーディオビジュアルワードを提案した。分類に関しては、映像内容から駆動される潜在的トピック駆動モデルを導き、シーンごとに動的に変化する感情のモデル化を行った。24の映画から切り出した206の映像クリップを用い、16人の被験者に対する実験を通じて、その有効性を検証した。

第6章は、「Vid2Trailer: Automatic Generation of Emotionally Impactful Film Trailers (Vid2Trailer: 感情的にインパクトの大きな映画予告編の自動生成)」と題して、情動を考慮した映画の要約について論じている。ショットの切り替えに伴うインパクトを提案し、前章で論じた感情分類結果に合わせて、ロゴやテーマ音楽といったシンボリックな要素を合わせて用い、インパクトを最大にするようにシーンの接続を行うことで、要約の生成を行った。ユーザスタディによりその有効性の検証を行っている。

第7章は、「Conclusions(結論)」であり、本論文での成果をまとめるとともに、今後の課題について言及している。

以上これを要するに、本論文では、映像内容解析の新しい課題である情動解析に取り組んだものであり、映像シーンに対する動的な情動モデルによる感情分類を示し、さらに、感情を考慮してインパクトを最大化する要約の生成、映像内容への興味を考慮した検索のためのランキング、オーディオからの情動イベントの検出といった一連の課題について論じたものであり、電子情報学上貢献するところが少なくない。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。