

## 審査の結果の要旨

論文提出者氏名 Nagul Coohajananone (クーハロッチャナノン ナググン)

本論文は「Shot-Based Adaptive Sampling Approach to Video Summarization (適応的なサンプリングアプローチによるショットに基づいた映像要約に関する研究)」と題し、英文でかかれており、9章よりなる。ビデオカメラといった映像機器やネットワークが広く行き渡るにつれ、個人が有する映像は飛躍的に増えつつある。膨大な映像データを管理するためには、その映像内容を適切に要約することが必要不可欠であり、本論文では、その要約表示のための画像処理手法についての研究をまとめている。本論文では、映像をショットの長さや動きといった特徴量を利用した適応的なサンプリング手法を提案し、重要と思われるシーンをより密にサンプリングすることによる要約表示の検討を行っている。要約結果に対する主観評価や既存手法との比較を通してその性能評価を行い、さらに、サポートベクタマシンを利用する適合性フィードバックを行うことにより、より個人の好みを反映させたキーフレームの取得を検討している。

第1章は、「Introduction(序論)」であり、映像インデキシングに関するマルチメディア処理についてふれるとともに、本論文で主たる特徴とするホームビデオの特徴について論じ、本論文の背景と目的、および論文の構成について述べている。なお、本提案は、ホームビデオ以外の一般の映像に対しても適用可能である。

第2章は、「Related works on video summarization(ビデオ要約に関する関連研究)」と題し、映像の要約に関する動向をまとめている。ハイライト、キーフレームによる要約など網羅的に現状の動向をまとめている。

第3章は、「Shot-based adaptive sampling approach to video summarization (ショットに基づく適応サンプリングによるホームビデオの要約)」と題し、本論文で構築する自動要約システムの全体にふれ、さらに、要約のための提案についてまとめている。自動的な要約のためには、高次の特徴にあまり入ることなく、低次の特徴を効果的に利用するのが望ましい。このため、本論文では、まず映像のショット分割を行い、各ショット毎にそのショットの重要性を表す特徴量と考えられるショット長とショット内の動き量をもとめ、その2つの特徴量の関数としてショットに対する代表的なフレーム(R-フレーム)の数を求める。最終的に、各ショットに対して、変動量を考慮した適応的なサンプリングを行うことで、必要な数のR-フレームを抽出する。要約映像は、このR-フレームをもとに切り出した映像クリップを接続することにより求めることができる。なお、所望の要約の長さに応じて、R-フレーム数、平滑化のパラメータを変化させ、自動的に対応することができる。この提案手法を、MPEG-7のテストシーケンスを含む複数の実映像に適用している。

第4章は、「Evaluation by users(ユーザ評価)」と題する。前章にて、論じた要約手法の主観評価を行っている。明瞭度、簡潔性、一致性の観点から異なる長さの要約映像を5段

階で 9 人の評価者によって評価した。実験に取り上げた 2 つの映像では、15%程度での要約において、3 つの指標のバランスがもっとも高くとれることが確認された。また、自動生成された要約に対してのコメントもまとめられている。

第 5 章は、「The improvement of shot-based adaptive sampling approach(適応的サンプリングアプローチの改善)」と題し、4 章で論じた要約手法に対する改善を論じている。ショットに対しての R-フレーム数の算出への改善を施した。その改善により、より多くのショットから R-フレームを抽出できるようになる。改良前の結果と比較して、映像全体をよりよく表現する要約の生成が可能となった。

第 6 章は、「Comparison between shot-based adaptive sampling and other conventional sampling algorithm(ショットに基づく適応的サンプリング手法と従来のサンプリング手法との比較)」と題する。本論文で提案するショットの特徴量に基づき R-フレームを抽出する手法と従来からの代表的なフレーム抽出法との比較を行った。従来法としては、等間隔で映像からフレームを取り出す均一サンプリング、及びフレーム間の動きに基づき R-フレームを抽出する R-シーケンス法を取り上げている。ショット単位で見ると、従来手法はいずれも、抽出されるフレームが著しく偏ってしまう。両者と比較して、提案手法は、偏りを十分に抑え、適度にショット毎に R-フレームの密度を変えることができたことが確認された。

第 7 章は、「User relevance feedback(ユーザによる適合性フィードバック)」と題する。抽出された R-フレームに対して、ユーザが評価を与えることで、よりユーザの好みに合わせた R-フレームの抽出を行える可能性がある。本研究では、抽出された R-フレームに対して、ポジティブな評価、ネガティブな評価を入力し、それをトレーニングデータとして学習したサポートベクターマシン (SVM) を用いた。SVMにより、さらに、原データから候補となるキーフレームを抽出することができる。このフィードバックを繰り返すことで、最終的な R-フレームを決定し、要約映像を得ることができる。

第 8 章は、「結論」であり本論文の成果をまとめている。

以上これを要するに、本論文では、膨大な映像データの要約手法として、ショット特徴量を反映した適応的サンプリング手法を提案し、その評価、改良について論じており、従来より示されているサンプリング手法に比べて優れたサンプリング手法であることを検証している。また、よりユーザの嗜好を反映するための適合性フィードバックについても論じている。本論文で論じた映像要約手法は、将来の映像コンテンツの利用に寄与することが期待され、電子情報工学上貢献するところが少なくない。よって本論文は博士 (工学) の学位論文として合格と認められる。