

実時間画像解析に基づく映像の分節化と検索法に関する研究

学位論文内容の要旨

近年、ビデオ映像が、情報の表現および伝達手段として広く利用されるようになってきている。しかし、映像は、従来のテキスト等のメディアに比べて格段に大きなデータ量を持ったメディアであり、そのような膨大な映像データを作り出したり、必要な情報を見つけ出したり、あるいは利用することは多大な時間と労力を必要とする。

本研究は、コンピュータを用いて映像を実時間で画像解析し、映像を構成要素に分節化して蓄積・管理することにより、映像の構成や内容に関して利用者が指定した検索条件を満たす映像部分を高速に検索し提示することができる映像データベースの実現を目指して行なわれた一連の研究をまとめたものであり、特に、映像の分節化法と検索法の研究に焦点を当て、処理の実時間性、多種多様な映像に対応する柔軟性、そして、低コスト性を満たす手法の提案を行なっている。

第1章では、研究の背景、目的、位置付けと、論文の概要を述べている。第2章では、映像処理技術の研究動向を、制作支援技術、検索技術、視聴支援技術に分けて紹介している。

第3章では、申請者が開発した映像の自動カット分割と実時間ビデオモザイクの各手法について述べている。映像の時間的および空間的な構造を簡潔に表現するため、映像をカット(一台のカメラで撮影された連続した映像区間)単位に切り分けて、カットごとに1枚の代表画像を抽出して一覧表を作成し利用者に提示することを提案している。カットとカットの変わり目の自動検出手法として、隣接画像間の画像相関の変化に着目し、色ヒストグラムを隣接画像間の相異度を測る特徴量として用いる手法を提案している。デゾルブやワイプなどのエフェクトが用いられた映像に対しても、デゾルブ使用時には輝度が時間的に単調増加ないしは単調減少することを用い、ワイプ使用時には、新たに画面に現れる画像領域と、その出現前の同位置の画像領域との相関が低いことを利用し、これらのエフェクトに対処可能なカット点検出法を提案している。提案手法は、計算量が少ないことから一般的なパソコンで、3種類のカット変化の検出をリアルタイムで同時実行できる。また、カットの代表画像にカット内容を示す多くの情報を凝縮できるように、新しいビデオモザイク(パノラマ化)の手法を提案している。映像のコマ毎に水平・垂直両方向に輝度の投影分布を求め、得られた投影分布の相関に基づいて、カメラワークによる画像の移動量の推定を行なうことにより、パン・チルト・ズームの移動量の実時間推定を行い、連続画像を繋ぎ合わせてパノラマ画像を求めることができる。

第4章では、映像中の被写体およびカットシーケンスの高速検索の手法を提案している。被写体検索手法については、現れる場面によって同一被写体が大きさや形状を変えても検索できるように、被写体の特徴量として、色の組み合わせと大凡の位置関係とに着目した

手法を提案している。カットシーケンスの検索については、映像区間中に含まれるカット点に着目し、カット点の画像のみを時間順に組み合わせたパターンで映像区間全体の識別を行っている。映像を1コマ1コマ比較するのに比べて高速な照合ができ、単純に間引いて比較した場合のような性能の劣化も少なかった。本手法をワークステーションに実装し、ユーザが例示した映像区間と同じカット点画像の組み合わせを持つ映像区間を検索する実験を行ったところ、約5時間の映像に対して、123種の映像を1.9秒で同時に検索できることを報告している。

第5章では、映像中の特徴場面検出と実時間クラスタリングの手法を提案している。特徴場面検出は、映像の自動要約への適用を目標としている。本研究の手法では、映像中の演出技法に着目することで、複雑な意味内容の解析なしに、放送の受信と同時に実時間で特徴場面を選び出し、映像を再構成することができる。放送番組においては、重要な場面には、それを強調するように様々な演出がなされるが、そのような演出は、実時間で十分検出可能な画像や音声の変化にも反映されるということを利用して利用している。実時間クラスタリングは、ユーザの視聴した映像をコンピュータと一緒に視聴してすべて憶えておき、それと同様の映像区間が放送中に現れた場合に、自動的に録画を開始したり、ユーザに通知を行うことを可能にする技術である。次々と放送される新しい映像を入力し、その特徴を圧縮して記憶すると同時に、圧縮されたままの特徴量を用いて高速に照合を行いクラスタリングする新しい映像照合手法を提案している。標準的なパソコン上に実装した結果、特徴量を映像1秒あたり平均20バイト以下で記憶でき、数時間分の過去の映像から、実時間で同一映像区間を見つけられる。

第6章は、以上の技術が、サブリミナルの検出や個々のCFの放送実施の監視など、映像検査技術に応用できることを実際の応用例を上げて論じている。

第7章は上記の成果をまとめている。

学位論文審査の要旨

主 査 教 授 田 中 謙
副 査 教 授 原 口 誠
副 査 教 授 嘉 数 侑 昇
副 査 教 授 北 島 秀 夫

学位論文題名

実時間画像解析に基づく映像の分節化と検索法に関する研究

近年、ビデオ映像が、情報の表現および伝達手段として広く利用されるようになってきている。しかし、映像は、従来のテキスト等のメディアに比べて格段に大きなデータ量を持ったメディアであり、そのような膨大な映像データを作り出したり、必要な情報を見つけ出したり、あるいは利用することは多大な時間と労力を必要とする。

本研究は、コンピュータを用いて映像を実時間で画像解析し、映像を構成要素に分節化して蓄積・管理することにより、映像の構成や内容に関して利用者が指定した検索条件を満たす映像部分を高速に検索し提示することができる映像データベースの実現を目指して行なわれた一連の研究をまとめたものであり、特に、映像の分節化法と検索法の研究に焦点を当て、処理の実時間性、多種多様な映像に対応する柔軟性、そして、低コスト性を満たす手法の提案を行なっている。

第1章では、研究の背景、目的、位置付けと、論文の概要が述べられている。第2章では、映像処理技術の研究動向が、制作支援技術、検索技術、視聴支援技術に分けて紹介されている。

第3章では、申請者が開発した映像の自動カット分割と実時間ビデオモザイクの各手法について述べている。映像の時間的および空間的な構造を簡潔に表現するため、映像をカット（一台のカメラで撮影された連続した映像区間）単位に切り分けて、カットごとに1枚の代表画像を抽出して一覧表を作成し利用者に提示することを提案している。カットとカットの変わり目の自動検出手法として、隣接画像間の画像相関の変化に着目し、色ヒストグラムを隣接画像間の相異度を測る特徴量として用いる手法が提案されている。デゾルブやワイプなどのエフェクトが用いられた映像に対しても、デゾルブ使用時には輝度が時間的に単調増加ないしは単調減少することを用い、ワイプ使用時には、新たに画面に現れる画像領域と、その出現前の同位置の画像領域との相関が低いことを利用し、これらのエフェクトに対処可能なカット点検出法を提案している。提案手法は、計算量が少ないことから一般的なパソコンで、3種類のカット変化の検出をリアルタイムで同時実行できる。また、カットの代表画像にカット内容を示す多くの情報を凝縮できるように、新しいビデオモザイク（パノラマ化）の手法を提案している。映像のコマ毎に水平・垂直両方向に輝度の投影分布を求め、得られた投影分布の相関に基づいて、カメラワークによる画像の移動量の

推定を行なうことにより、パン・チルト・ズームの移動量の実時間推定を行い、連続画像を繋ぎ合わせてパノラマ画像を求めることができる。

第4章では、映像中の被写体およびカットシーケンスの高速検索の手法を提案している。被写体検索手法については、現れる場面によって同一被写体が大きさや形状を変えても検索できるように、被写体の特徴量として、色の組み合わせと大凡の位置関係とに着目した手法を提案している。カットシーケンスの検索については、映像区間中に含まれるカット点に着目し、カット点の画像のみを時間順に組み合わせたパターンで映像区間全体の識別を行っている。映像を1コマ1コマ比較するのに比べて高速な照合ができ、単純に間引いて比較した場合のような性能の劣化も少なかった。本手法をワークステーションに実装し、ユーザが例示した映像区間と同じカット点画像の組み合わせを持つ映像区間を検索する実験を行ったところ、約5時間の映像に対して、123種の映像を1.9秒で同時に検索できることが報告されている。

第5章では、映像中の特徴場面検出と実時間クラスタリングの手法が提案されている。特徴場面検出は、映像の自動要約への適用を目標としている。本研究の手法では、映像中の演出技法に着目することで、複雑な意味内容の解析なしに、放送の受信と同時に実時間で特徴場面を選び出し、映像を再構成することができる。放送番組においては、重要な場面には、それを強調するように様々な演出がなされるが、そのような演出は、実時間で十分検出可能な画像や音声の変化にも反映されるということを利用して、実時間クラスタリングは、ユーザの視聴した映像をコンピュータが一緒に視聴してすべて憶えておき、それと同様の映像区間が放送中に現れた場合に、自動的に録画を開始したり、ユーザに通知を行うことを可能にする技術である。次々と放送される新しい映像を入力し、その特徴を圧縮して記憶すると同時に、圧縮されたままの特徴量を用いて高速に照合を行いクラスタリングする新しい映像照合手法が提案されている。標準的なパソコン上に実装した結果、特徴量を映像1秒あたり平均20バイト以下で記憶でき、数時間分の過去の映像から、実時間で同一映像区間を見つけられることが報告されている。

第6章は、以上の技術が、サブリミナルの検出や個々のCFの放送実施の監視など、映像検査技術に応用できることを実際の応用例を上げて論じている。

第7章は上記の成果をまとめている。

これを要するに、著者は映像中のカットやオブジェクトといった対象を、映像再生に同期した実時間処理可能な前処理によって分節化し、これらの対象を高速に検索処理する分節化法と検索法を提案し、そのシステム技術を確立するとともに、実用性を実証し、電子情報工学に対して貢献するところ大なるものがある。

よって著者は北海道大学博士（工学）の学位を授与される資格あるものと認める。