

## 審査の結果の要旨

氏名 谷口 行信

ビデオ技術の急速な発展と普及に伴って、日々、多くの映像データが作られるようになっており、今や、過去に作られ蓄積された膨大な映像データの中から再利用したいものを的確に探し出すことができないという深刻な問題が生じている。本論文は、この問題の深刻さに着目し、人の視覚の瞬間的な静止画把握能力と、コンピュータの動画処理能力を組み合わせる解決法を提案したもので、「映像要約一覧の自動生成法と応用に関する研究」と題し、6章から成る。

第1章「序論」では、蓄積された膨大な映像データから望みのものを検索する問題の深刻さとその背景を論じたあと、本論文のアプローチの基本的な考え方を述べている。それは、ビデオデータという時間変化を伴う情報を、紙面に静止画を配置した映像一覧という空間的な情報に変換することによって、静止画を瞬間的に把握できる人の視覚能力が生かせる環境を整えるというものである。次にこのアプローチと過去の関連研究との関係を論じ、最後に本論文の構成について述べている。

本論文の主要部分は2部から構成されている。第1部「映像一覧の自動生成法」には、第2章と第3章が当てられ、ビデオデータからリアルタイムで映像一覧を自動生成する方法が提案されている。

第2章「ショット切換え検出法」では、一台のカメラで連続的に撮影された映像区間をショットと名付け、これを時間的に変化する映像の最小単位とみなす。そして、一つのショットから次のショットへ変わるショット切換えの位置を検出する方法を提案している。基本的には、映像が急激に変化した時刻がショットの切換え時刻であるとみなすことが原則であるが、すべてのショット切換えをこの原則でとらえることはできない。たとえば、フラッシュがたかれた瞬間に映像は急激な変化を生じるが、これはショット切換えとみなすべきではなく、逆に、一つのシーンがゆっくりフェードアウトし、それと重なるようにもう一つのシーンがゆっくりとフェードインするところは、瞬間的な変化は大きくないにもかかわらずショットの切換えとみなさなければならない。この問題を解決するために本論文では、二つの静止画像の違いを表す非類似度という量を定義し、複数の時間スケールで求められた非類似度の時系列を解析することによってショット切換えを検出する方法を提案している。そして、ニュースと映画の映像に対して、この手法の有効性を実験的に確認している。

第3章は「PanoramaExcerpts: パノラマ混在型映像一覧」と題し、パン、チルト、ズームなどのゆっくりしたカメラ操作による映像の時間変化の全体を、静止した一つの合成画像によって表す方法を提案している。この方法では、時間的に隣り合う画像の間の変化を

拡大・縮小，水平方向への平行移動，垂直方向への平行移動の三つの動きに分解し，そのそれぞれをズーム，パン，チルトというカメラ操作に対応づける．そして，ゆっくりしたカメラ操作から生じる一連の映像データから一つの合成画像を構成し，それにカメラの動きを表す矢印などの補助的記号を書き加える．このようにして得られる合成画像をパノラマアイコンと名付けている．パノラマアイコンは，形も大きさもカメラの動きに応じてさまざまに変わるが，カメラ操作によって時間的にゆっくり変化するショットの内容全体を，一つの静止画として表現したものになる．これは，映像の時間変化を，映像の空間的広がりとして表す一種の「言語」の提案とみなすことができる．さらに，形が不定のパノラマアイコンを効果的に紙面に配置する方法も構成している．そして，第2章，第3章の技術を統合して，ビデオデータから実時間で映像一覧を自動作成するシステム **PanoramaExcerpts** を開発し，その有効性を実際のビデオデータで確認している．

第2部「映像一覧の応用」には，第4章と第5章が当てられ，第1部で提案した映像一覧の自動生成法を組み込んだ三つの実用的システムを構築して，その有用性を評価している．

第4章は「代表フレーム一覧を用いた映像アクセスインタフェース」と題し，ショット切換え検出法を用いて構成される代表フレーム一覧を2種類の映像アクセスインタフェースに応用している．

その第一は，**PaperVideo** と名付けられたシステムで，ショットの切換えを検出した後，各ショットをそれぞれ1枚の代表画像で表して並べることによって一覧をユーザに提供するものである．一方，その第二は，**TVRam** と名付けられたシステムで，テレビ放送された24時間分の映像データに対象を特化した映像アクセスインタフェースである．これによれば，たとえばテレビで放送される一つのチャンネルの一日分の映像を，CM部分を自動的に取り除いた上で1枚の一覧データにまとめたり，逆に，CM部分のみを集めて1枚の一覧データにまとめたりすることができる．さらに，この一覧は階層的な構造をもっており，注目したいところをユーザが指定すれば，その部分のみに対する詳しい映像一覧を表示することができる．

第5章「**SceneCabinet**: 映像インデクシングシステムの実装と評価」では，5種類の映像解析技術—ショット切換え検出，カメラワーク検出，テロップ検出，音楽検出，音声検出—を統合した映像インデクシングシステムを設計・実装している．このシステムでは，基本的には，各種のインデックスは自動編集されるが，その精度は百パーセントではないため，自動編集されたインデックスをユーザが効率的に修正したり，関連情報をユーザが簡単に編集するためのユーザインタフェースも備えている．このシステムの評価実験の結果，手作業の場合と比べておよそ半分の作業時間でインデックスが編集できることが確かめられた．

第6章「結論」では、以上の成果をまとめるとともに、今後に残された課題について述べている。

以上を要するに、本論文は、膨大で時間変化を伴うために検索が難しい映像データを、パノラマアイコンを含む空間的一覧の形でユーザに提供することによって、ユーザの検索作業を軽減するための新しい映像表現言語とそれを実現するインタフェースシステムを提案したもので、数理工学の発展に大きく貢献するものである。

よって本論文は、博士（工学）の学位請求論文として合格と認められる。