

論文内容要旨

論文題目 The Study on Semantic Video Content Generation System
(和訳：意味的な内容に基づいた映像コンテンツ生成システムに関する研究)

氏名 張 文利

近年、衛星を利用したBS及びCSデジタル放送、地上波デジタル放送が始まり、デジタル放送の普及に伴い、デジタル映像量が急増している。このような映像コンテンツを利用するため、ここ数年、ビデオ・オン・デマンド、バーチャル・リアリティ、インターネットを経由する映像ストリーミングなど、多くの映像応用アプリケーションが開発されている。しかし、効率的に映像処理し、効果的に管理する能力はまだ欠如である。このため、映像内容生成管理できる新しい技術とツールを備えたシステムが必要とされる。

本論文では、意味的な内容に基づいた映像コンテンツ生成システムに関する研究を行う。この分野の研究は主に二つがある。(1)メディア間協調の結果を自動、半自動あるいは手作業によって映像コンテンツに高次記述を付与システム (2)低次特徴量を利用して、映像コンテンツを自動検索するシステムがある。前者のシステムの場合、関連研究の多くは映像のすべての内容に対して、高次記述を付けようとしているが、結果として、記述作業は大変コストがかかることになる。我々の研究グループはメディア間協調結果を利用するシステムでは、ドラマ映像を対象にしたドラマ映像管理利用フレームワーク **Video-Stream Description Language for TV movie show (VSDL-TV)**を提案し、ドラマ映像とそれに付随する台本との対応づけによって、ドラマ映像コンテンツに高次記述し、**VSDL-TV**に基づいたシステムの提案と実装を行った。映像対象を限定することになり、一般映像を対象としない問題がある。一方、後者のシステムの場合、低次特徴量の組み合わせで自動的に検索を行うが、あくまで「イメージ検索」で、より高次の意味内容までは生成できない欠点がある。それに画像検索技術は限界があることで、なかなか満足できる検索結果が得られないのが現状である。

本論文では、上記の二種類の手法を統合し、一般映像を対象に映像内対象物認識を行う手法を開発。本手法のもとに映像をもつ低次特徴量から高次記述まで生成可能な映像コンテンツ生成システムを提案。

本システムは、まず映像ショットからキーフレームを検出し、画像セグメント手法を用いて、分割された各領域の一色、面積、フレームにおける位置等の低次特徴量を検出する。次に、領域ベースのモデルマッチング手法によって同一オブジェクトが存在する映像ショットが自動的に検出される。画像認識技術は限

界があること、また低次特徴量と高次記述の間のセマンティックギャップを埋めるには、三つの面で改善を行う。

その一：オントロジー理論を元に映像内にある対象物に関する"セマンティックオブジェクト"モデルデータベースを構築する。

その二：緩めたモデルを利用しリコー率の高い検索結果を検出する。

その三：代表的な対象物に対して、検索結果からコンテンツ提供者が対話的に該当オブジェクトが存在するショットを選択し、システム側はモデルにつけられた情報に従って自動的に選択されたオブジェクトに高次記述を付与する。従って高位記述と低位記述の組み合わせ利用で映像内一般的対象物にも対応可能。このようなデータベースシステムに基づいて、実装した応用演算を組み合わせることによって、より自由度の高い映像コンテンツが高精度で生成する環境を構築。

現段階では、80%の映像内対象物再現率を実証し、約 10 時間の映像データベースを構築し、評価を実施。効率的かつ実用性をもつシステムであると確認されている。また本システムに基づき、ドラマ登場人物人気投票システム、映像要約システム、映像空間検索利用システム、E コマースシステムなどさまざまなオブジェクトベースアプリケーションの開発し、システムの有効性をさらに確認されたとする。

具体的に論文の構成は以下のようになる。

第一章： 研究背景、目的について述べるほか、論文の構成の要旨についても説明する。

第二章： 高次記述をもつ映像コンテンツ生成システムについて述べる。

第三章： 低次記述をもつ映像コンテンツ生成システムについて述べる。

第四章： 低次記述に基づいたシステムの応用演算体系について述べる

第五章： 低次記述に基づいたシステム実装と評価について述べる

第六章： 提案したシステムに基づく応用アプリケーションについて述べる。

第七章： 全体のまとめと今後の予定について述べる。