

## 論文の内容の要旨

### 論文題目 AI-based Automatic Sound Motion Picture Generation with Screenplay GUI from the Perspective of Film Director

シナリオ GUI を利用した映画ディレクターの知識による AI ベースの映像自動生成に関する研究

氏 名 申 金紅

ブロードバンドネットワーク技術の進化により、圧縮されたデジタルビデオなどマルチメディアコンテンツを、WWW を介してやりとりし、個人の PC 上で表示できるようになった。しかし、どのような人でも複雑ではない方法で、個人のプレゼンテーションを作成し、配信することを可能にする環境を提供することが、依然として Web 上の通信におけるもっとも重要な問題の一つにある。

本論文は、デスクトップソフトウェアシステム EMM (Electronic Movie Maker) に用いるための、知的な動画生成のための方法論について述べている。EMM は、視覚エフェクトを利用した、3次元アニメーションや実写映像、合成映像などのデジタル映像の生成を自動化するよう設計されている。特に本論文では、EMM において、映像制作の決まりに従った、キャラクターアニメーションの自動同期化技術と、ビデオデータベース・ウェブビデオライブラリからの自動映像検索技術に、主眼が置かれている。これらの 2 つの技術は、マルチメディア制作とマルチメディア理解という、対極にある技術で、まだそれらが補完されていない技術である。

このアニメーション生成システムは、パーサを通してユーザの入力したシナリオを理解し、人の変わりになるバーチャルディレクターの指示のもと、しかるべき動画像に自動的に変換することができる。バーチャルディレクターは、ドメイン知識ベースに蓄えられた映画の経験に従い、シーンの設定と、適切なショットタイプおよびショットシーケンスの決定および、バーチャルカメラワークのプランニングを行い、知識ベースの手法により、ユーザの意図通りの演出を実現する。ビデオドキュメントの再利用は、低レベルな映像の特徴量と高レベルな概念の両方を用いた、セマンティックアノテーションに基づいた階層的な構造を持つ EMPVR (EMM Video Retrieval) によって実現される。EMPVR は、

以下の 3 つの機能から成り立っている。意味内容の自動抽出と、特定のデータ形式への意味内容の保存、メタデータに従ったセマンティックビデオコンテンツの検索である。ここで映像制作の知識とルールベースの推論技法は、CLIPS というエキスパートシステム言語にエンコードされており、ビデオデータは XML 形式でエンコードされ、MPEG-7 標準で記録されている。

EMM ソフトウェアのプロトタイプ的设计に関して、概念と原理、方法、実装についての説明の後、インタラクティブアプリケーションにおけるこれらの技術を用いた、シーンの記述能力に関する実験結果を示し、評価を行う。

映像自動生成システム EMM の開発と実装により、最先端の理論と技術水準の自然言語処理ならびに知的マルチメディアプレゼンテーション、言語視覚化、人工知能アプリケーション、映画理論の各分野の統合と改善に貢献した。最後に、この技術開発が、専門的なデジタル映像制作にどのような影響を与えるかを議論する。

本論文の第 1 章は、背景の紹介と EMM ソフトウェアシステムの目的について述べられている。一般的な人が自分のビジュアルコンテンツを制作し、簡単に配信できるような、簡単に学べて簡単に利用できるデスクトップソフトウェアツールを開発することは、デジタル映像制作の処理を簡便化したいというニーズに答えるための一つの提案である。まず私は、システムプラットフォームを作成することから研究を始めたが、試作を通して、人とコンピュータ間のインタフェース変換を実現する設計の理論的な基礎となる、ビジュアル・非ビジュアルコミュニケーションの基本概念に気がついた。続く第 2 章は、言語による自動映画生成システムに関する関連研究のアイデアと、この論文と関係するその他の最先端の技術について述べる。

第 3 章は、コンピュータアニメーションの技術と実写映像制作法、映像検索について分析し、バーチャル 3D ワールドで用いられる、映画理論の観点に基づいた新しい映画制作技法を提案し、自動デジタル動画作成実現のための、知識ベースに基づいた概念フレームワークについて述べる。既存の商用アニメ制作ソフト・ビデオ編集ソフトでは、出力されるアニメーションシーケンスのデザインは、基本的にすべて人の手を必要とする。本提案では、今までのようなアニメ生成システムと同様なタスクレベルのコマンドによる演技生成だけでなく、人工知能を利用した、動的なコンピュータグラフィックス生成のルールベース手法を取り入れた。

第 4 章は、シーンレイアウトおよび撮影方法を含む知識表現と、CLIPS 言語で記述されたキャラクターの演技の設計の仕方について説明している。2 種類のシナリオユーザインタフェースを、設計の基本理論に基づいて説明する。この視覚化という技術の方向は、社会トレンドも後押ししている。ビデオの再利用に関するパートでは、主に映画制作方法に基づいたビデオデータモデルを紹介する。そのデータモデルは、情報の抽出と検索のためのビデオの大量な特徴量を表現するのに十分なセマンティクスを持っている。まず、ビデオ検索サブシステムのアーキテクチャを示し、それから映画ディレクターの観点から AI アプローチを利用したアニメの自動生成をいかにして実現するか、詳細に説明する。

第 5 章では、シーンを作るための映画の根本原理をいかにして用いるかを詳細に説明するための、映画の経験に基づいたシステムの実装を示す。動的な絵の生成は、1) オブジェクト表現、2) オブジェクト動作、3) カメラワークショットとショットシーケンスに関わってくる。CLIPS は 3 つの異なるプログラミングパラダイムルールベース、オブジェクト指向、手続き型をサポートするための、広いさまざまな知識を扱う結束したツールを提供しているが、またデジタル動画の生成をプログラムするための、先述した要求も満たすことができる。この章では、ショットアンドショットシーケンスの生成によって、EMM の実現と機能を示す。

最後に、第 6 章では、以下の 3 つの節で成り立っている。結論と、このシステムへの貢献の要約、および今後の研究についての議論である。他のシステムとの比較を示した後、現在の研究の貢献について結論付ける。

プログラミングを介さない、映画知識ベース環境の EMM システムにより、専門家でない人でも簡単に自分でデジタル映画を作ることが可能になる。このシステムは、2 種類の補完しあうシナリオ入力フォームと、フィルムディレクターの観点を活用した、知識ベース手法による知的なアニメーション生成を用いている。今後の研究では、誇張表現や二次的な演出など伝統的なアニメーション技法を用いることと、キャラクターの演技を人間らしく向上することと、知識ベースを改善することと、人-コンピュータのインタフェースを改良していくことにより、アニメーションの効果を向上させることである。