

令和 3 年 5 月 20 日現在

機関番号：32663

研究種目：若手研究

研究期間：2018～2020

課題番号：18K18173

研究課題名（和文）音源を入力とした最適なアニメーションの自動生成手法の開発

研究課題名（英文）Development of an automatic generation method for optimal animation using sound sources as input

研究代表者

石川 知一（Ishikawa, Tomokazu）

東洋大学・情報連携学部・准教授

研究者番号：30635545

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究は、音源に適したCGのアニメーションを出力することを目的とした。特定のアニメーション制作時に、現在アナログで行っている作業をデジタル上で再現するツールを実装し、実際に利用してもらうことで必要なパラメータを確認する実験を行った。また、別のシーンではキーフレームで指定した通りに中割を作成する際、初期パラメータの最適化によって実現できるのかを確認した。映像から得られる特徴量と、BGMから得られる音楽の特徴量とこの印象語を関連付けられるかと、人間が最適だと感じる映像とBGMの組み合わせについてアンケート調査を行った。このデータを利用して映像に適したBGMの検索システムを提案した。

研究成果の学術的意義や社会的意義

デジタルアニメーションではアーティストのユーザビリティの観点からもいくつかのパラメータによって再現することが求められている。本研究では限定的なシーンではあるが、そのパラメータを特定することと、その最適化によってアニメーションを作成することが有効であることを確認した。また、本研究ではアンケート調査により人が音から受ける印象を感性語によって動画・音楽特徴量と関連付けることで、CG分野の制作工程に寄与するだけでなく、VRの分野においてもクロスモーダル観点から人が耳から受ける刺激と映像のマッチングについて貢献できると考えている。

研究成果の概要（英文）：The purpose of this study was to output CG animations suitable for sound sources.

We implemented a tool to digitally reproduce the work that is currently done in analog when creating a specific animation, and conducted an experiment to confirm the necessary parameters by having the user actually use our tool. In another scene, we confirmed whether it was possible to create a middle segment as specified by a key frame by optimizing the initial parameters.

A questionnaire survey was conducted to find out whether it is possible to associate the features obtained from the video with the features of the music obtained from the background music and these impression words, and to find out the combination of video and background music that people find optimal. Using this data, we proposed a system to search for suitable background music for video.

研究分野：コンピュータ・グラフィックス

キーワード：コンピュータグラフィックス 最適化 シミュレーション アニメーション

## 1. 研究開始当初の背景

申請者はこれまでに CG による映像制作の研究だけでなく、音・文章・静止画像・動画のようなメディアを相互に変換する研究を行ってきた。メディアはその 1 つ 1 つが研究分野であり、それぞれ研究とともに成熟してきている。コンシューマゲームやバーチャル・リアリティなどのデジタルコンテンツは、メディアを統合して制作されているためマルチメディアコンテンツと呼ばれる。また、日本のビデオゲームやアニメにおいては、映像作品・小説・音楽作品の多岐に渡って商品展開することをクロスメディア展開と呼んでおり、メディアを有機的に結びつけることは非常に有用であると言える。メディアの間には指向性を持った変換関係があり、これを本研究ではメディアの相互変換と呼称する。

申請者は雨音から降雨強度を判定して CG 映像を制作する手法について提案した。流体が映っている画像に Navier-Stokes 方程式を適用して流体の動画を得るための実験を行っている。また、アニメ業界におけるフィルムコミック制作時のクリエイターの負担を軽減する研究として、アニメ中の各フレームの特徴量を計算し、マンガ形式に自動でレイアウトを整える映像要約手法の研究開発を行った。

現在考えられているメディアの相互変換は主に不可逆変換で、音声・画像処理の方法を用いるか、物理現象の再現によって行われることが多い。しかしながら、逆方向の変換は解が複数存在し得るため、妥当な結果を得ることは容易ではない。例えば、文字から画像を検索する方法がこれに該当する。逆方向の変換について、共通した解決策は、ビッグデータを用いた学習によって、いくつかの制約を設けた最適化問題を解いていることにある。提案法でも音源データを収集してパラメータ決定のための回帰問題として解決することを考える。

## 2. 研究の目的

本研究は、音源を入力として、その音源に適した CG のアニメーションを出力することを目的とする。申請期間内に解決する課題は、次の 3 つである。提案法は情報が既知の音源を複数用意して、回帰モデルにかけたり、高度な機械学習をしたりすることによって、アニメーション生成のためのパラメータを確定させる。

**課題 1 . シーン生成のための数理モデルを構築し、音源からパラメータを決定すること**

**課題 2 . 映像制作ソフトのプラグインを開発し、制作効率が向上するか検証すること**

**課題 3 . 音と映像が適しているか人間が受ける感覚を統計的に検証すること**

申請段階で検討していた研究手順では、音波と各種パラメータがわかっているサンプルデータの収集を行い、それらの音波解析から決定すべきパラメータに対する大まかな傾向を分析するか、機械学習にかけられるかのアプローチになる。また、音波から特定したパラメータを入力として CG アニメーションを作成し、音から映像を生成することで作業効率が上がるかどうかや、人の感覚として音と生成された映像が適しているかどうかの評価実験を行うところまでを研究範囲とする。

また、音と映像のマッチングは、近年急速に普及している VR の分野においても臨場感や没入感の追求という意味で価値ある研究対象であると言える。本研究は人が音から受ける印象を視覚化することも評価の一部として取り入れることで、CG 分野の制作工程に寄与するだけでなく、VR の分野においてもクロスモーダル観点から人が耳から受ける刺激と映像のマッチングについて調査することによって貢献する。

## 3. 研究の方法

まず、CG アニメーション制作時に、最適化すべきパラメータについて実装して検証した。具体的なシーンとして、アニメで使われる髪の毛のアニメーションと、3DCG でモデリングを行うオーロラのアニメーションについて実装した。さらに、音楽と映像のそれぞれについて人間が受ける感覚をアンケート調査し、どの映像にどの音楽が適しているのかを解明しようとした。

### (1) 髪の毛のセルルックアニメーション

アニメ作品中に登場する髪や布が揺れるシーンを作成するために、必要なパラメータが未知であったため、実際にアニメーターが手書きで行う方法をデジタル上で再現できるようなユーザーインターフェースの作成から開始した。アニメの中で髪の毛の束が動く様子は先端が 8 の字を繰り返し描くように動かし、その動きが先端から根本まで伝搬するように作成されることがわかった。この動きを再現するために必要だと考えられる複数のパラメータ（手書きの 8 の字の軌跡とそのサンプリング点、波の伝搬速度と位相）を実装し、実装したシステムをユーザーに使用してもらうことで必要なパラメータを絞り込んだ。

## (2) オーロラの 3DCG アニメーション

オーロラのシミュレーションは電磁流体のシミュレーションを行う手法が一般的であるが、ユーザーが所望する形状への変形を物理則に従って行うことは難しかった。この研究では、オーロラが撮影されたときの 3D 形状を適当な仮定をもとに推定し、キーフレームとすることで、キーフレームとして与えた時間にその形状を満足するようなシミュレーションができるよう、初期値のパラメータを最適化することを考えた。この最適化するパラメータはオーロラを形成する沿磁束線電流の値の分布で、この分布を有限フーリエ変換した係数を遺伝的アルゴリズムで最適化することを実験した。

## (3) 音楽と映像の人間が受ける感覚調査とマッチング

映像と背景音楽（以下、BGM）が人間に与える影響をそれぞれ印象語としてラベリングされている研究は既に存在している。映像から得られる数値的特徴量と、BGM から得られる音楽の数値的特徴量とこの印象語を関連付けられるかと、人間が最適だと感じる映像と BGM の組み合わせについてアンケート調査を行った。実際に 17 名の協力者に、映像 50 本、BGM50 本を視聴してもらい、それぞれについて 16 個の印象語対に 7 段階で評価値をつけてもらった。動画の特徴量としてカラーヒストグラムやオプティカルフローの値、音楽特徴量としてメル周波数ケプストラム係数を利用し、これらと印象語対のベクトルを回帰分析によって関数化した。これらの結果から、映像を入力として最適な音楽を探す検索システムを構築し、評価実験を行った。

## 4. 研究成果

アニメで使用される髪の毛のなびきアニメーションについて作成したシステムを図 1 に示す。ユーザーが 8 の字をスケッチすると、動作して欲しい毛束の主軸方向に各位置がどのように動くかを自動計算する。また、微調整を行うことができるような GUI を用意し、アニメーション制作時にどのようなパラメータ調整が必要かを取得できるようにした。

アニメーション制作時に適切にキーフレーム間の中割画像をシミュレーションで作成するための初期値を最適化問題として解くことを考えた。オーロラのシミュレーションについて適用した結果を図 2 に示す。探索範囲を限定することで、この問題について良好な結果を得ることができた。初期値の最適化手法を用いて、音源を入力とした課題についても適用することを検証していきたい。

音楽と映像によって人間が受ける感覚についてアンケート調査を行い、BGM の検索システムの構築を行った。検索精度の評価を行い、ある程度の結果を得ることができたため、今後国際会議で発表する予定である。

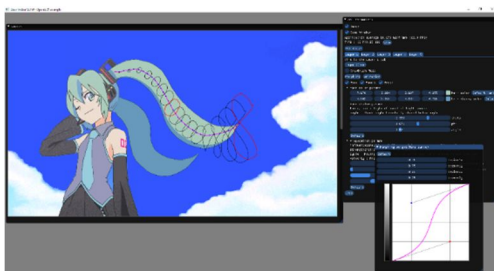


図 1: なびきアニメーション制作の操作画面と UI

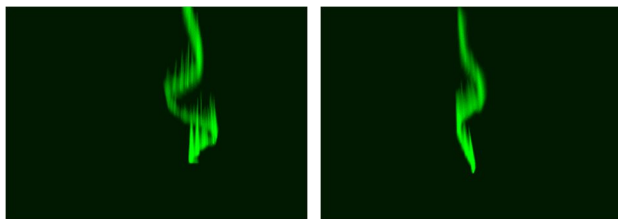


図 2: オーロラのビジュアルシミュレーション(指定したキーフレームの間の運動が適切に行われるように初期値のパラメータを求めた)

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 Tomokazu ISHIKAWA, Yusuke KAMEDA, Masanori KAKIMOTO, Ichiro MATSUDA, Susumu ITOH	4. 巻 6(2)
2. 論文標題 Rust Simulation Based on Chemical Reaction Processes	5. 発行年 2018年
3. 雑誌名 IIEEJ Transactions on Image Electronics and Visual Computing	6. 最初と最後の頁 82-88
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計14件（うち招待講演 0件／うち国際学会 6件）

1. 発表者名 Tomokazu Ishikawa, Ryota Nakazato, Ichiro Matsuda
2. 発表標題 Procedural animation of aurora and its optimization for keyframe animation
3. 学会等名 2019 International Symposium on Computer Graphics, Multimedia and Image Processing（国際学会）
4. 発表年 2019年

1. 発表者名 Naoaki Kataoka, Tomokazu Ishikawa, Ichiro Matsuda
2. 発表標題 A Method to Create Fluttering Hair Animations That Can Reproduce Animator's Techniques
3. 学会等名 SIGGRAPH Asia 2019 Posters（国際学会）
4. 発表年 2019年

1. 発表者名 石川知一
2. 発表標題 Energy-Brushes: Interactive Tools for Illustrating Stylized Elemental Dynamicsの実装
3. 学会等名 情報処理学会 コンピュータグラフィックスとビジュアル情報学研究会(IPSJ CGVI)研究会
4. 発表年 2018年

1. 発表者名 片岡直映, 石川知一, 亀田裕介, 松田一朗, 伊東晋
2. 発表標題 アニメータ技法に基づく髪の毛のデジタル再現
3. 学会等名 Visual Computing 2018
4. 発表年 2018年

1. 発表者名 櫻井惇, 石川知一, 亀田裕介, 松田一朗, 伊東晋
2. 発表標題 SVMによるアニメキャラクターの顔に対する作画崩壊の判別
3. 学会等名 Visual Computing 2018
4. 発表年 2018年

1. 発表者名 Naoaki Kataoka, Tomokazu Ishikawa, Yusuke Kameda, Ichiro Matsuda, Susumu Itoh
2. 発表標題 Digital Reproduction of Hair Waving Based on Animator Technique
3. 学会等名 SIGGRAPH Asia 2018 Posters (国際学会)
4. 発表年 2018年

1. 発表者名 Ichiro Matsuda, Tomokazu Ishikawa, Yusuke Kameda, Susumu Itoh
2. 発表標題 A Lossless Image Coding Method Based on Probability Model Optimization
3. 学会等名 Proceedings of the 26th European Signal Processing Conference (EUSIPCO-2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Jun Sakurai, Tomokazu Ishikawa, Yusuke Kameda, Ichiro Matsuda, Susumu Itoh
2. 発表標題 Discrimination of drawing collapse for animated characters by SVM
3. 学会等名 Proc. SPIE, The 22nd International Workshop on Advanced Image Technology (IWAIT 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Naoki Kataoka, Tomokazu Ishikawa, Yusuke Kameda, Ichiro Matsuda, Susumu Itoh
2. 発表標題 Procedural animation of waving cartoon hair by emulating animator's techniques
3. 学会等名 Proc. SPIE, The 22nd International Workshop on Advanced Image Technology (IWAIT 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 手塚建貴, 石川知一
2. 発表標題 HoloLens2を用いた料理時における情報提示方法の検討
3. 学会等名 映像表現・芸術科学フォーラム2021
4. 発表年 2021年

1. 発表者名 倉田大樹, 石川知一
2. 発表標題 超解像処理を用いたレンダリング高速化についての一検討
3. 学会等名 映像表現・芸術科学フォーラム2021
4. 発表年 2021年

1. 発表者名 杉本一就, 石川知一
2. 発表標題 動画を入力としたBGM検索システムについての研究
3. 学会等名 映像表現・芸術科学フォーラム2021
4. 発表年 2021年

1. 発表者名 石井 建, 石川知一
2. 発表標題 VR技術を用いたTRPGのゲーム性についての研究
3. 学会等名 映像表現・芸術科学フォーラム2021
4. 発表年 2021年

1. 発表者名 一重元希, 石川知一
2. 発表標題 お絵描きプロジェクションマッピングについての研究
3. 学会等名 映像表現・芸術科学フォーラム2021
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------