

修士論文

2017 年度

音源に同期する運指および表情に注目した 吹奏アニメーションの自動生成

堀井 絵里  
(学籍番号：81621728)

指導教員 教授 藤代 一成

2018 年 3 月

慶應義塾大学大学院理工学研究科  
開放環境科学専攻

## 論文要旨

音楽演奏を題材としたアニメーションは多く存在する。そこでは、実際に演奏する演奏者の身体の動きに近い演奏シーンが生成されている。しかし、楽器を演奏するキャラクタの運指に注目すると、音楽と指の動きが完全に同期されていないことがあり、強い違和感を感じる。また、完全に同期されているアニメーションを作成するには、多くの労力が必要となる。そこで本研究では、管楽器を演奏するキャラクタの吹奏アニメーションを、音楽に容易に同期させるため、音源から自動的に運指のアニメーションを生成することを目指す。より具体的には、ウインドシンセサイザを用いて作成した MIDI 音源を解析することにより、運指が音源に同期した、自然な吹奏アニメーションを実現する。

近年、コンピュータの処理速度が向上したことで、高品質な画像の高速な描画が可能となった。この影響により、ゲームや映像をユーザが完成版に近い品質で編集できる技術が発達している。このような CG 技術の台頭や IT の発達で、最新のコンピュータ上に仮想の社会や人間をつくる技術が研究者や CG 関係者の注目を集めている。しかし人間のパーツの多くは目や肌、髪の毛のようにやわらかい、小さい、半透明といった特徴をもつパーツで構成され、コンピュータには簡単に描画できないものが多い。特に頭髪は、人間にとって重要なパーツでありながら、その複雑さが原因で高速な描画が難しい物体として知られている。また頭髪状物体が落とす影も同様に、一般的な描画手法を使うとちらつきやありえない濃度の影が生じるため、計算が難しい描画対象として知られている。

頭髪は半透明で、かつ細長い円筒の集合であり、このような幾何形状をもつ物体は頭髪の形状、位置や視点、照明に変化がある場合、影を描画する際にちらつきが生じる。このちらつきのうち、頭髪が半透明で遮蔽判定が難しいことによるちらつきは、影を描く場所に光が届くまでに通過する頭髪密度の積算値を使い影を描くことで解決できる。一方、髪が細すぎてコンピュータが髪を認識できないことによるちらつきは、投影後にちらつきが生じている部分と周囲の色を平均化して抑制されてきた。しかしこの平均化処理は、結果の品質が投影後の単位面積あたりの画素数に依存してしまう。

そこで本論文では、頭髪の密度を積算する既存手法を拡張して、大きさが可変で不透明度分布をもつ正方形小板で頭髪の三次元形状を近似することにより、結果の品質が画素数に依存しないアンチエイリアシング手法を提案する。これに加えて既存研究では積極的に行われてこなかった、出力画像の定量的な品質評価をするために新たな評価方法を作成した。この評価方法を使用して既存手法と提案手法の出力画像の品質を比較した結果、提案手法は計算速度を大きく減少させることなく効果的にちらつきを抑制できることを確認した。

### キーワード

アニメーション, 吹奏楽, 金管楽器, 自動生成.

## Thesis Abstract

### An Object-Space Anti-Aliasing Method for High-Quality Shadowing of Hair-Shaped Objects

Recent improvement in CPU and GPU performance made it possible to render high-quality computer graphics in real time. Accordingly, technologies to edit video games or movies with nearly-final quality previews have been developed. With support of such rapid growth of CG technology as well as other IT developments, it has begun attracting much attention from related researchers and practitioners to construct realistic and dynamic virtual societies and humans on a modern computer. However, making virtual humans is a very difficult task because they are made up of many soft, small or transparent parts such as skin, **hairs** or eyes. Especially, human hairs are well-known as intricate objects which require much executing time to render and cause **aliasing**, even though they are essential parts of humans. Shadows cast from hairs are also known as difficult objects to render because they generate aliasing and too far dense shadows with commonly-used shadowing methods.

Hairs are **semi-transparent** and can be presented as mass of thin cylinders, but this style of geometric modeling causes aliasing to rendered shadows when users update a shape or a position of hair, camera angle and/or the position of lights. As for a type of aliasing caused by difficulty to detect occlusion due to the transparency of hairs, even previous shadowing methods can decrease it by calculating the accumulated **amount of density values of all hairs** that light goes through. On the other hand, aliasing caused by the problem that a computer cannot recognize too-thin hairs could be alleviated by image-space post-processing such as neighboring pixel averaging, though the quality depends strongly on the density of projected pixels.

Therefore, this thesis proposes an object-space method which realizes resolution-independent **anti-aliasing**. The method is an extension of the previous shadowing method that calculates the accumulated amount of density values of hairs, and **approximates a three-dimensional shape of hair** as a cloud of variable-size square splats with concentric opacity distribution. In this study, a novel evaluation method was also proposed to compare the extent of antialiasing in a quantitative manner. Comparison based on the novel evaluation method revealed that the proposed method outperforms the existing image-space anti-aliasing with minimum loss of computational speed.

#### Keywords

Hair; real-time rendering; anti-aliasing; object space; semi-transparent objects.

# 目 次

<b>1</b>	<b>序論</b>	<b>1</b>
1.1	研究背景 . . . . .	2
1.2	研究目的 . . . . .	2
1.3	本論文の構成 . . . . .	2
<b>2</b>	<b>関連研究</b>	<b>3</b>
2.1	音からアニメーションを自動生成する研究 . . . . .	4
2.2	アニメーションから音を生成する研究 . . . . .	5
2.3	音とアニメーションを同期させる研究 . . . . .	5
2.4	本研究との関連 . . . . .	5
<b>3</b>	<b>提案手法</b>	<b>6</b>
<b>4</b>	<b>結果と分析</b>	<b>7</b>
<b>5</b>	<b>結論</b>	<b>8</b>
	謝辞	10
	公開文献	11
	参考文献	12

## 図 目 次

2.1 落下アニメーションの自動生成 . . . . .	4
2.2 口元アニメーションの自動生成 . . . . .	4
2.3 ピアノを弾く手元のアニメーションの自動生成 . . . . .	5

## 表 目 次

---

## 第 1 章

### 序論

---

本章では、本研究の研究背景と研究目的について説明する。1.1 節で研究背景、1.2 節で研究目的、1.3 節で本論文の構成について述べる。

## 1.1 研究背景

音楽演奏を題材としたアニメーションは多く存在する。テレビで放映されたアニメーションの例を挙げると、『のだめカンタービレ』、『けいおん!』、『響け! ユーフォニアム』が相当する。これらのアニメーションは、セル画であったり 3DCG であったり、製作方法がさまざまであるが、いずれも実際に演奏する演奏者の身体の動きに近い演奏シーンが生成されている。楽器の輝きや形状なども忠実に再現されており、評判が高い。

一方、楽器を演奏するキャラクタの運指や身体の動きに注目すると、音楽と動きが完全に同期されていないことがあり、違和感を感じる。特に速いフレーズや、複雑なリズムを演奏するシーンで、このようなアーティファクトが起きやすい。また、表情からも不自然さを感じることもある。しかし、1 つ 1 つの音と動きや表情を合わせるには、多くの労力が必要となる。

## 1.2 研究目的

1.1 節で述べた課題を解消するため、本研究では、管楽器を演奏するキャラクタの吹奏アニメーションを、音源から自動的に生成することを目指す。より具体的には、実際の演奏アニメーション制作フローに沿わせるため、楽曲は楽器（ウインドシンセサイザ）を用いて、MIDI 音源として生成する。そして、作成した楽曲を解析することにより、吹奏の情報を得る。得られた情報をキャラクタの身体の動きや表情、そして管楽器に適用することにより、音源に同期した自然な吹奏アニメーションを実現する。

## 1.3 本論文の構成

本論文は次章以降、以下のように構成される。

次章では関連研究を説明する。第 3 章では本手法の仕組みを説明する。そして、第 4 章では自動生成結果を述べ、この結果に対する評価をまとめる。最後に第 5 章で結論を述べ、今後の課題に言及する。

なお、本研究の成果は Visual Computing/グラフィクスと CAD 合同シンポジウム 2017 にてポスター発表 [ii] を行った。そして、映像表現・芸術科学フォーラム 2018 (2018 年 3 月)、Cyberworlds (2018 年 8 月) において発表を行う予定である。



---

## 第 2 章

### 関連研究

---

本章では 2.1 節で音からアニメーションを自動生成する研究を紹介し、?? 節でアニメーションから音を自動生成する研究を紹介する。2.3 節では音とアニメーションを同期させる研究を紹介し、最後に 2.4 節にて本研究の新規性を述べる。

## 2.1 音からアニメーションを自動生成する研究

音とアニメーションを同期させることを目的として、音からアニメーションを自動生成する研究は多く存在している。

物が落下するアニメーションを生成するには、物の落下音と物が落下するタイミングを 1 つ 1 つ合わせる必要がある。Langlois ら [2] は、それらを正確に合わせるために、物が落下する音が収録されている音源から、物が落下するアニメーション 2.1 を自動生成する手法を提案した。



図 2.1 落下アニメーションの自動生成

キャラクタの口の動きと音声を合わせる、リップシンクも難しい課題となっている。キャラクタの口の動きに合わせて後から音声を録音するアフレコは比較的容易であるが、音声から、その言葉を発している口元のアニメーションの生成は、言葉と口の形、そして発するタイミングと口の動きを合わせる必要があり、ひじょうに難易度が高い。そこで、Edwards ら [3] は、台詞が収録されている音源および台詞が記載されているテキストファイルを入力すると、その台詞を話している口元のアニメーション 2.2 を自動生成する手法を提案した。本手法では、顎と唇のみを制御することにより、自然な口元のアニメーションを自動生成している。更に、後から表情を編集することが可能となっているため、最終的には自然な表情のアニメーションが完成する。

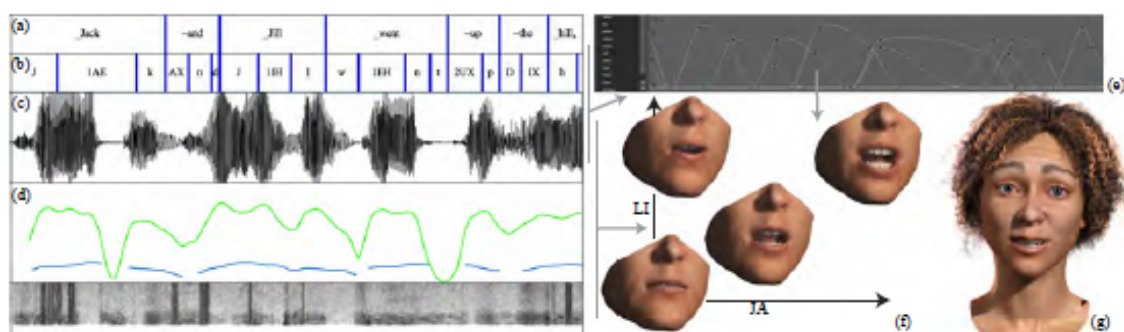


図 2.2 口元アニメーションの自動生成

本論文の目的と同じように、音源から楽器を演奏するアニメーションを自動生成する研究も多く行われている。Zhu ら [4] は、MIDI 音源からピアノを弾く指元のアニメーション 2.3 を自動生成する手法を提案した。ピアノのような楽器は、指と音が 1 体 1 で対応していないため、フレーズによ

り指使いを変える必要がある。ある指が他の指の上や下を通るような指使いの時は、指同士が衝突しないように考慮する必要もある。彼らの手法では、これらの課題を解消することが可能である。



図 2.3 ピアノを弾く手元のアニメーションの自動生成

## 2.2 アニメーションから音を生成する研究

音とアニメーションを同期させることを目的として、アニメーションから音を自動生成する研究も存在する。この研究はサウンドレンダリングという名称で一般的によばれている。CG-ARTS の教育リポート、『音を描き出す夢の実現』

## 2.3 音とアニメーションを同期させる研究

Lee ら [1] は、MIDI 音源およびキャラクタのモーションデータを入力とし、双方を最小限編集することにより、MIDI 音源と同期したキャラクタのダンスアニメーションを生成する手法を提案した。Lee らの研究は、本研究と同様、音と完全に同期したアニメーションを生成することが目的である。しかし本研究とは異なり、入力したモーションデータを編集することにより、アニメーションを生成している。

## 2.4 本研究との関連

---

## 第 3 章

### 提案手法

---

---

## 第 4 章

### 結果と分析

---

---

## 第 5 章

### 結論

---

本章では本論文の結論を述べ，今後の課題に言及する．

## 謝辞

本論文の執筆にあたり、ひじょうに多くのアドバイスを頂いたばかりでなく、私からの様々な相談もお聞きいただいた慶應義塾大学理工学部情報工学科の藤代 一成教授に深く感謝いたします。私の至らなから様々なご負担やご迷惑をおかけしたにも関わらず、辛抱強くご指導いただきありがとうございました。今後いかなる場で働く際にも、先生のご指導を思い出しながら仕事に励んでいきたいと思えます。

本研究に取り組むにあたり、実装面において多くのアドバイスをいただいただけでなく、コンピュータグラフィクス分野における様々な課題に出会う機会をいただいた、現在は株式会社 DeNA に所属しておられる松岡 徹氏、株式会社デジタル・フロンティア開発部の福田 啓氏、斎藤 弘氏、高山 耕平氏並びに開発部の皆様に深く感謝いたします。

研究を進めるにあたり時には楽しく話し合い、時には真面目に議論を行い、共に助け合った研究室の皆様に深い感謝をいたします。皆様の存在が無ければ、私の研究はここまで進めることができませんでした。

研究をするにあたり、科研費基盤研究 (B)25280037 及び慶應義塾大学博士課程教育リーディングプログラム オールラウンド型「超成熟社会発展のサイエンス」に多大なるご支援をいただいたことを感謝しております。

最後に、研究を進めるにあたりお世話になりました、すべての方に感謝を申し上げたいと思います。



## 公開文献

- [i] 池田 泰成, 藤代 一成, 松岡 徹, 「頭髪状物体を対象とするアンチエイリアシング処理されたりアルタイムシャドウイング」, 第 76 回 情報処理学会全国大会 講演論文集, 東京電機大学 東京千住キャンパス, 2014 年 3 月.
- [ii] 池田 泰成, 藤代 一成, 松岡 徹, 「頭髪状物体の描画に対応するアンチエイリアシング処理されたりアルタイムシャドウイング」, Visual Computing/グラフィクスと CAD 合同シンポジウム 2014 DVD 予稿集, No. 48, 早稲田大学, 2014 年 6 月.
- [iii] Yasunari Ikeda, Issei Fujishiro, Toru Matsuoka, “An Object Space Approach to Shadowing for Hair-shaped Objects,” *ACM SIGGRAPH 2014 Posters*, Vancouver, Canada, August 2014. doi:10.1145/2614217.2633392
- [iv] 池田 泰成, 松岡 徹, 藤代 一成, 「頭髪状物体の影のオブジェクトスペースアンチエイリアシング」, 電子情報通信学会 和文論文誌 D 分冊, Vol. J99-D, No. 3, 未決定, 2015 年 12 月 (早期公開). doi:10.14923/transinfj.2015PDP0018
- [v] 金子 徳秀, 村岡 遼太, 池田 泰成, 「研究室リレー訪問」, 芸術科学会誌 DiVA, Vol. 34, pp. 17-21, 2013 年 11 月.

## 参考文献

- [1] H. C. Lee and I. K. Lee, “Automatic Synchronization of Background Music and Motion in Computer Animation,” *Computer Graphics Forum*, Vol. 24, No. 3, pp. 353-361, September 2005.
- [2] T. R. Langlois and D. L. James, “Inverse-Foley Animation: Synchronizing Rigid-body Motions to Sound,” *ACM Transactions on Graphics*, Vol. 33, No. 4, Article No. 44, pp. 44:1-44:11, July 2014.
- [3] P. Edwards, C. Landreth, E. Fiume, and K. Singh, “JALI An Animator-Centric Viseme Model for Expressive Lip Synchronization,” *ACM Transactions on Graphics*, Vol. 35, No. 4, pp. 127:1-127:11, July 2016.
- [4] Y. Zhu, A. S. Ramakrishnan, B. Hamann, and M. Neff, “A System for Automatic Animation of Piano Performances,” in *Proceedings of Computer Animation and Virtual Worlds*, Vol. 24, No. 5, pp. 445-457, September 2012.
- [5] <http://takabosoft.com/domino> (最終アクセス日 : 2018/02/05).
- [6] <http://tf3dm.com/> (最終アクセス日 : 2018/02/05).
- [7] <https://jp.3dexport.com/> (最終アクセス日 : 2018/02/05).
- [8] CG-ARTS, “音を描き出す夢の実現,” 最終アクセス : 2018 年 2 月 5 日. [Online] [https://www.cgarts.or.jp/report/rep\\_kr/rep0901.html](https://www.cgarts.or.jp/report/rep_kr/rep0901.html)