

1. 担当 PM

石黒 浩 PM

(大阪大学 大学院基礎工学研究科 システム創成専攻 教授 (特別教授))

(ATR 石黒浩特別研究室室長 (ATR フェロー))

2. 採択者氏名

クリエイター：古川 翔一 (早稲田大学)

3. 委託金支払額

2,304,000 円

4. テーマ名

音声に合う口の動きを実現した吹替え映像を容易に作成するソフトウェア

5. 関連 Web サイト

なし

6. テーマ概要

本プロジェクトでは「音声に合う口の動きを実現した吹替え映像を容易に作成するソフトウェア」というテーマの下で、音声のみから容易にかつ高速にキャラクターの口パクアニメーションを自動生成するシステムを開発した。本システムは現行のアニメ制作フローに沿うように Adobe After Effects のスクリプトとして開発されており、ユーザは使用する音声の性別を選択するのみで口パクアニメーションが生成可能となっている。本システムの開発に際して、キャラクターの口パクの遷移情報を計算するアルゴリズム及び Adobe After Effects 用のスクリプト、スタンドアロンに動く GUI を作成した。

7. 採択理由

外国語映画の日本語吹替において、リップシンクが不完全であることが映像メディアの質を下げていることは明らかで、その問題を解決しようとする実用性の高い提案であった。現在の画像処理技術であれば、十分にそれを実現できる可能性があり、是非とも実用レベルのシステムに仕上げてほしいと考えた。また、映画だけではなく、同時翻訳機能付きのテレビ会議システムなどへの応用も視野に入れると、さらに実用性は高くなると考えた。

8. 開発目標

本プロジェクトでは音声のみから容易に口パクアニメーションを生成するシステムの開発に取り組んだ。システムは以下の要件を満たすことで、アニメータの制作支援を実用的なレベルで実現することを目指した。

- 音声の性別を選択するだけで、高速かつ高精度に口パクアニメーションを生成すること
- アニメータが制作フローで一般的に利用するツールのプラグインとして開発すること
- ユーザがインタラクティブに編集できる機能を備えること

9. 進捗概要

本プロジェクトは開始当初、口の動きと音声とを同期させるアルゴリズムを幅広い領域（例えば Skype 等のビデオ通話）へ応用することを想定していた。しかし、プロジェクトを進めていく中で実用性、つまり実際にユーザに使ってもらえるシステムへと完成させることを重視することにし、その結果としてターゲットを 2D アニメーションの制作現場に絞り、口パクアニメーションの生成システムの開発に注力することとなった。

最終的に本プロジェクトで開発したシステム「Mouth Motion Generator」は、任意の人物の音声ファイル（wav 形式、サンプリング周波数 44100kHz 以上、モノラル形式推奨）から高速かつ高精度にそれに同期するアニメキャラクタの口の動きを計算することが可能である。本システムはアニメ制作で一般的に用いられる Adobe After Effects のスクリプトとして実装されており、現行のアニメ制作フローへの導入が容易となっている。図 1 に本システムの GUI を示す。本 GUI ではユーザは使用する音声は男性のものか女性のものかを選択し、アニメーションを出力したいコンポジション及び使用したいキーフレーム画像シーケンス（口の開閉度合ごとにユーザが用意した画像、例：開口、中口、閉口のイラスト）をドロップダウンリストより選択して実行をするだけで高速にアニメ

メーションを生成することが可能である. 本システムの使用画面を図 2 に示す.

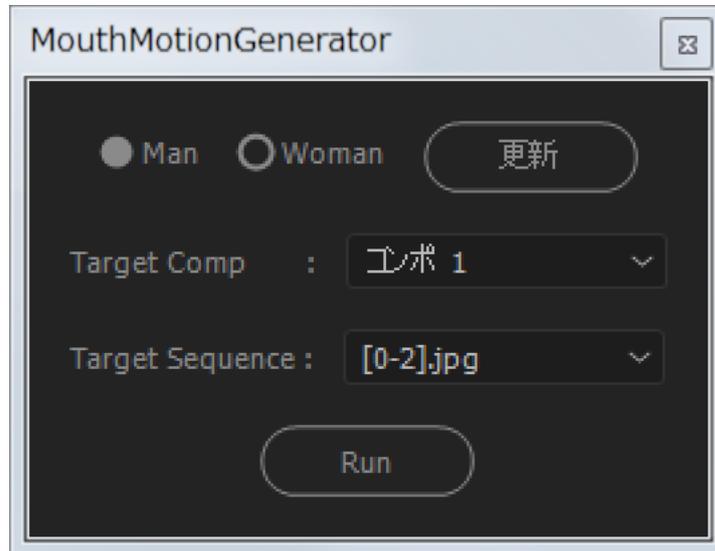


図 1 GUI

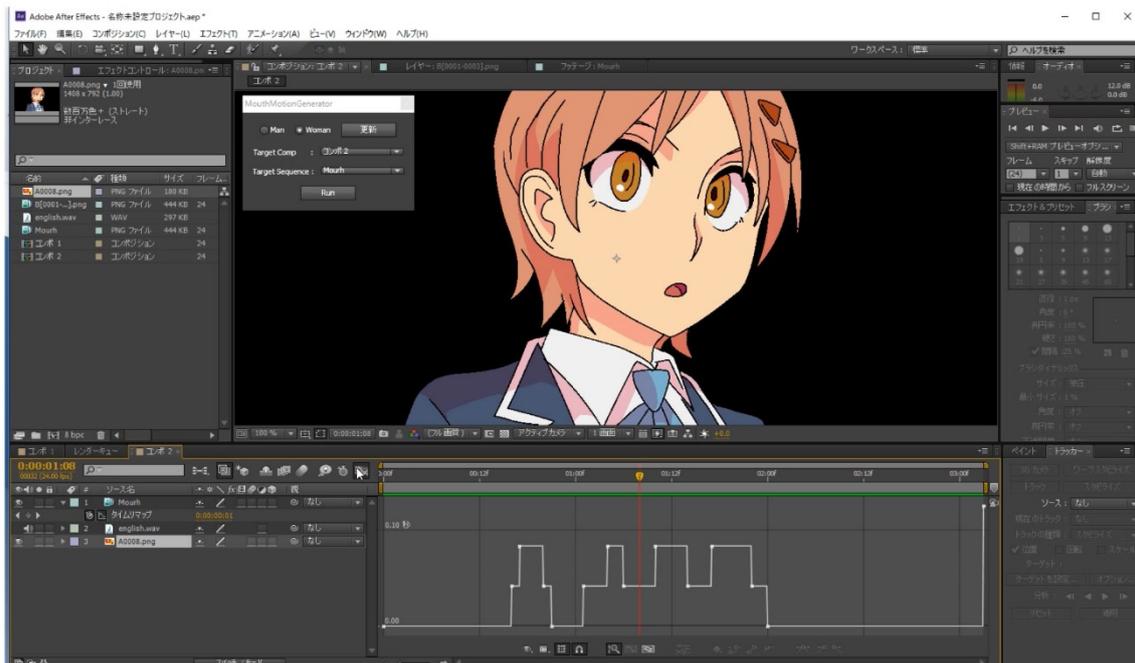


図 2 システム使用画面

本システムの特徴として以下の点が挙げられる.

- アニメらしさの実現
実写人物の口の動き情報を単に削減し, アニメキャラクタの口の動きに量

子化した場合には、不自然な口の動きをすることが確認される。これはアニメ作品におけるキャラクタの口の動きのように、限られた枚数の画像から如何に自然な動きを実現するかといった“アニメらしい”動き情報が存在するためである。本システムはこの点に初めて着目し、音声のみから数値的に“アニメらしい”口の動きを計算することが可能となっている。

- 多言語への適用可能性

本システムは言語に依存せずにアニメキャラクタの口の動きを推定することが可能である（例：日本語，英語，中国語など）。そのため，一次制作における制作支援はもちろんのこと，吹替え作品の制作といった二次制作支援も実現することが可能である。

- タイムリマップ形式の出力

本システムではキャラクタの口の動きを計算した後に，アニメータが容易にかつインタラクティブに編集を行えるように「タイムリマップ形式」で出力されるように実装した（図 3）。タイムリマップ形式とは口の動きの遷移を 0, 1, 2 の値に基づいてグラフ的に表示したものであり，口の動きの遷移がアニメータに理解しやすい形で可視化されている。また，タイムリマップ形式にはキーフレームとなる点が打たれており，これをマウス操作することでユーザは容易に口の動きを編集することが可能である。そのため，本システムによって自動で生成された口の動きの遷移に基づいて部分的に開閉値を強調したり，開口のまま発話シーンを終わらせたりするといった編集も容易であり，アニメータの制作支援を実現できる。

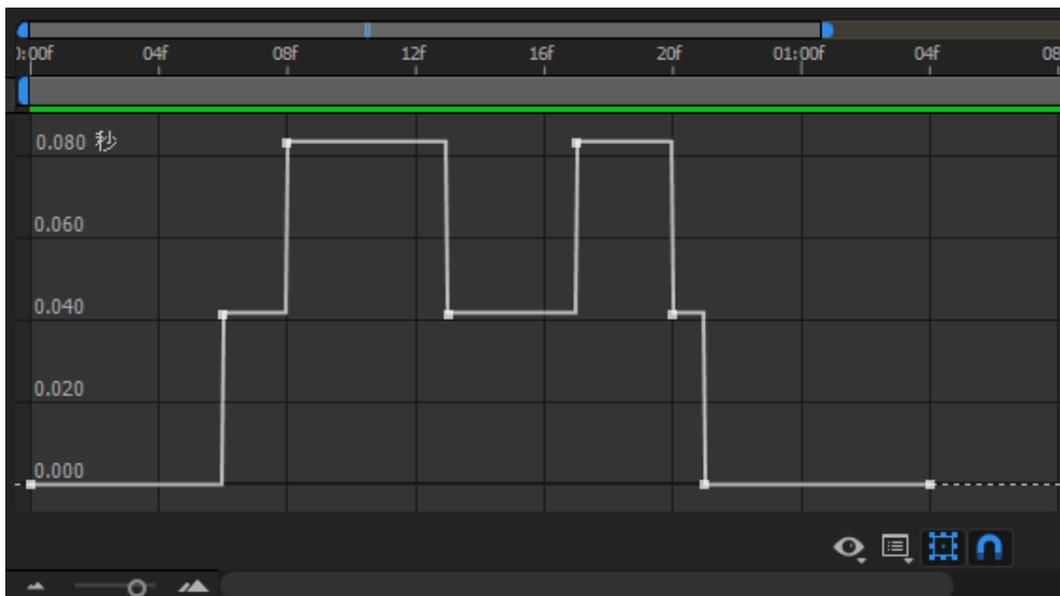


図 3 タイムリマップ形式

10. プロジェクト評価

アニメーションを対象にするか実写を対象にするか、プロジェクト開始当初は悩んだが、アニメーションを対象にして実用性が感じられるところまで実装を進めることができたことは評価に値する。将来実際のアニメ制作において利用される可能性は十分にある。

11. 今後の課題

本システムは取材やユーザテスト等を通して、多くのアニメータからの好評を得るに至ったが、本システムを用いたアニメ制作がスタンダードになるよう、更なる機能の追加と改善、普及活動が必要となる。