

# コンピュータグラフィクスを使ったアナウンサーシステム

## The Announcer System Using Computer Graphics

○蘇 宇航\*, 大久保 重範\*\*, 及川 一美\*\*\*, 高橋 達也†

○ Yuhang Su\*, Shigenori Okubo\*\*,  
Kazumi Oikawa\*\*\*, Tatsuya Takahashi†

\*山形大学

\*Yamagata University

キーワード： モデリング (modelling) モディファイヤ (modifier) リップシンク (lip sync)

連絡先： 〒 992-8510 山形県米沢市城南 4-3-16 山形大学 工学部 機械システム工学科 大久保研究室  
蘇 宇航, Tel: (0238)26-3245, Fax: (0238)26-3245, E-mail: tr154@dip.yz.yamagata-u.ac.jp

### 1. 緒論

近年のコンピュータグラフィクス (CG) 技術の発達により、さまざまなデータを視覚的に表現することが可能になりつつある。特に、人間のように複雑な形状を持つ三次元物体や、その動きを、CGを用いて表現する技術の発展は目覚ましい。また、音声処理や、画像認識、状況判断、画像やモデルの合成などの分野で大きく発展している。

いわゆる、アナウンサーシステムと呼ばれるものは、CGを使った音声と画像のシンクロシステムで、「会話する人の顔」を表現することである。このシステムはマスメディア、コマーシャル、アニメーションの製作、教材テキストなどのさまざまな分野で期待が高まっている。

本研究はコンピュータ上で視覚的に表現され、3D Studio MAX で作ったキャラクターが発声の基本表現を行う。そして英語の発音の種類、主に口の形によって分類し、3Dオブジェクトを使って、

各種類の発音を表現し、発音過程での口の形の変化する様子をアニメーションする。特に、「視覚言語とリップシンクの統合手法」の提案に関して述べる。

### 2. 3次元キャラクターの口の動きの表現

リップシンクを表現するために、まず、以下の基本モデルを採用して、3次元キャラクターの口の動きの表現を製作する。

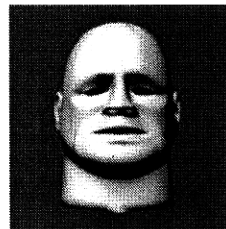


Fig. 1 正面図

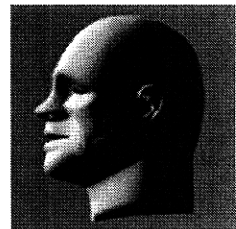


Fig. 2 側面図

## 2.1 基本モデルの作製

頭部は、モデルの中で最も手間のかかる部分であり、キャラクターの個性を表現する時にも最も重要な部分でもある。3次元キャラクターの製作は大変な手間がかかり、製作手順も大変複雑で、今回時間を短縮するために、主な製作手順を以下に示す。

- 1) 頭の輪郭の作製
- 2) 目、鼻、口、耳の作製
- 3) 眼球と歯の作製
- 4) 帽子の作製
- 5) モデリングの完成
- 6) マテリアルの作製
- 7) レンダリング

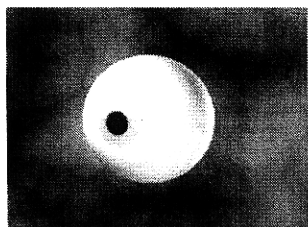


Fig. 3 眼球の作製

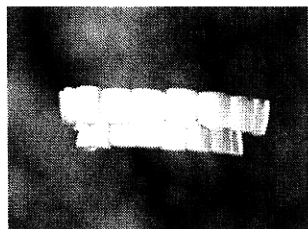


Fig. 4 歯の作製



Fig. 5 帽子の作製

作製したそれぞれの部分を合成し、配置する。その中で、歯をモデルの一部にしたり、独立して動かすことができる。そして、キャラクターのモデリングを完成させる。マテリアルの作製には、写真を使用することにより、かなり自然に近い画像の表現が可能になる。マテリアルを割り当てたワイヤフレームはレンダリングによって、作製した画像を画像ファイルとして保存することができる。最後に作製した画像をレンダリングしたものを以下に示す。



Fig. 6 正面図



Fig. 7 側面図

## 2.2 基本モデルにおける口のシェイプ形成の効率化

発音に対応する口の動きの表現を作製するには、モデルにおける口のシェイプ形成の効率化が必要であり、口のシェイプ形成の効率化によって、頂点を選択する操作が容易にできる。このプロセスでは、口の両端を形成する四つの頂点を選択し、3D Studio Maxの「修正」という機能を使用して、四つの頂点を「名前付き選択」にするという作製によって、口のシェイプ形成の効率化ができる。

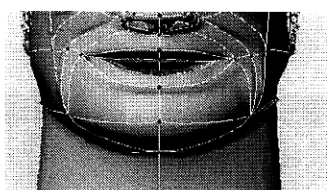


Fig. 8 口のシェイプ

## 2.3 口の位置ターゲットの作製

口の位置ターゲットを作製するときは、唇、歯、および顎の動きを含める。作成した頭部をコピーし、モデルのシェイプを修正して、口の位置ターゲットおよび表情のターゲットを作成します。モデルのシェイプを調整するには、頂点の接線ハンドルを使用する。このプロセスでは、話す動作を表現に大きく影響する唇、顎などの曲線、厚さ、及び位置を調整することができる。また、キャラクターの喜び、悲しみ、驚き、怒りなどの表情を表現することができる。

口のシェイプの修正手順：効率化した口の両端の外側の頂点を選択し、主軸はZ軸を確定して、上唇の曲線をコントロールする接線ハンドルを上に移動させる。口の反対側でも同じ操作を実行する。歯は、モデルの一部にすることも、別々にアニメートすることもできる。

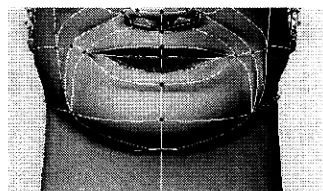


Fig. 9 修正前（微笑）

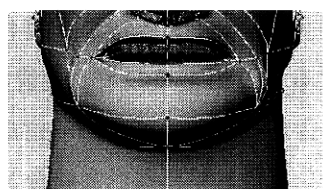


Fig. 10 修正後（怒り）

## 3. リップシング及び顔の表情

リップシンクおよび顔の表情の作製は、基本モデルにより、作成したキャラクターの頭部を使用する。元の頭部をコピーして修正し、リップシンクおよび表情のターゲットを作製する。修正手段は、キャラクターまたは基本モデルの頭部を選択してモディファイヤを適用する。

### 3.1 視覚言語

現実の人間が話す言語とは、

- 1) 音声
- 2) 口唇と舌の動き
- 3) 会話動作
- 4) 感情表現

が組み合わされたものである。

- 1) は聴覚により認知されるものである
- 2) はいわゆるリップシンク、発音記号と唇、舌の形を同期させる。
- 3) は頭の動き、「うなずき」など会話する時の動きである。
- 4) は、眉、頬などの顔面の筋肉の動きによる喜び、悲しみ、驚き、怒りなどの感情表現であり、

瞬きの速度や頻度を表現することもできる。

2)、3)、4) は視覚によるものであり、「視覚言語」と呼ぶことができる。本研究は視覚情報により英語の発音の音素を分類する。

### 3.2 視覚言語とリップシンクの統合

英語には26個のアルファベットがあるが、一つのアルファベットでも異なる単語の中で異なる発音をすることがあるので、発音は26個だけではない。音素は言語の最小の単位なので、今回のリップシンクは英語の音素を対象として研究している。英語には、50個の発音記号があるが、発音記号は母音(26個)と子音(24個)と分けられる。

母音には発音の長さによる短母音、長母音、二重母音に分けられる

Table 1 母音表[長さによる]

短母音	[æ] [ʌ] [ɑ] [ə] [ɚ] [ɪ] [e] [u]
長母音	[ɑ:] [ɑ:r] [ɚ:r] [ɔ:] [ɔ:r] [i:] [u:]
二重母音	[ai] [au] [ei] [ou] [ɔi] [iə] [iɚ] [ɛə] [uə] [aiə] [auə]

発音の音節による単母音、二重母音、三重母音に分けられる

Table 2 母音表[音節による]

単母音	[æ] [ʌ] [ɑ] [ə] [ɚ] [ɪ] [e] [u] [ɑ:] [ɔ:] [ɔ:r] [ɚ:r] [i:] [u:]
二重母音	[ai] [au] [ei] [ou] [ɔi] [iə] [ɔ:r] [ɑ:r]
三重母音	[iɚ] [ɛə] [uə] [aiə] [auə]

Table 1 と Table 2 によって、二重母音と三重母音に対応する口のシェイプは単母音から作られるので、本研究は母音の中の13個の短母音と長母音を含む単母音のリップシンクの特徴、24個の子音のリップシンクの特徴を提案する。母音と子音の発音特徴と視覚言語の対応表を作製した。以下に示す。

下に示す。

Table 3 母音の発音特徴・口形符号の対応表

発音	顎	歯	口の形	口形符号
[æ]	下げる	見えない	口を大きく開ける	[æ]
[ɑ] [ɑ:]	静止	見えない	口を大きく開ける	[ɑ]
[ɚ] [ɚ:r]	静止	離れて見える	口を少し開ける	[ɚ]
[ʌ] [ə]	静止	少し見える	口をあまり開けず	[ʌ]
[ɔ:]	下げる	少し見える	口を縦にあける	[ɔ:]
[i] [i:] [e]	静止	離れて見える	唇を左右に引く	[i]
[u] [u:]	下げる	少し見える	口を窄めて、唇を少し突き出す	[u]

Table 3 によって、作製した母音のリップシンクのターゲットを以下に示す。

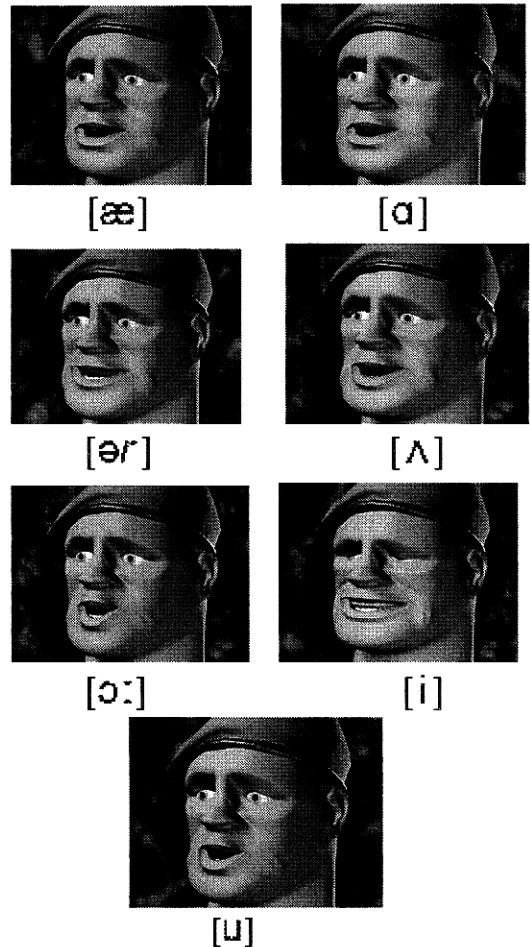


Fig. 11 母音リップシンクのターゲット

Table 4 子音の発音特徴・口形符号の対応表

音	発	顎	歯	口の形	口形符号
[p] [b] [m]	静止		見えない	唇を開める	[b]
[d] [v]	顎を少し水平にする	上の歯が見える		下唇を前歯で軽くおさえ、唇を左右に引く	[d]
[w]	下げる	見えない		唇を丸くして狭めて前に突き出す	[w]
[ɹ]	下げる	離れて見える		口を開け、唇を丸める	[ɹ]
[ʃ] [ç] [ʒ] [ʝ]	少し水平	歯を噛む。見える		口を開け、唇を少し丸める	[ʃ]
[s] [z] [j]	下げる。	歯を噛む。見える		口を少し開け、唇を左右に引く	[s]
[l] [θ] [ð]	下げる	離れて見える		口を自然に開けたまま	[l]
[t] [d] [k] [g] [ŋ] [h]	下げる。	歯は噛んで見える		唇を下げて、口を少し開ける。	[t]

Table 4 によって、作製した子音のリップシンクのターゲットを以下に示す。

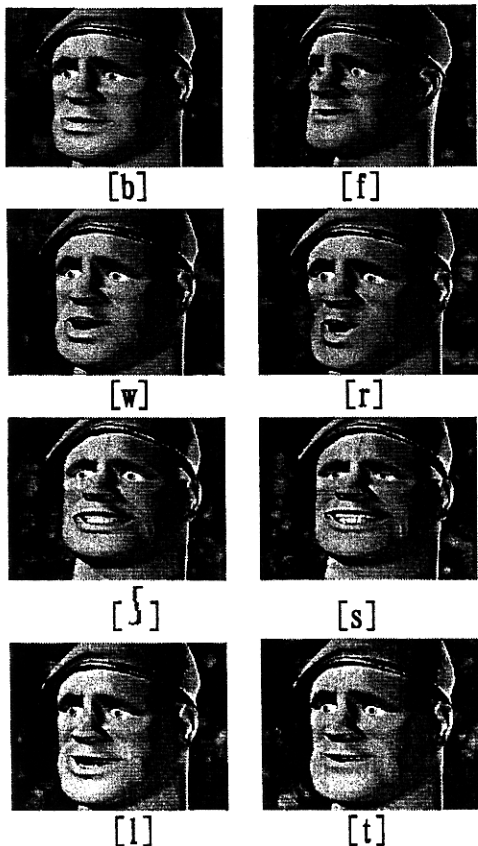


Fig. 12 子音リップシンクのターゲット

### 3.3 顔の表情のターゲット

キャラクターに必要なだけのたくさんの表情ターゲットを作製する。喜び、悲しみ、驚き、怒りなどの表情は、すべて独自のターゲットを持つ。作製のプロセスには、顎、頬、及び眉の各頂点を目標の位置になるまで修正する。キャラクターに歯がある場合、歯および基本の頭をコピーして新しいターゲットを作成する。時間を節約するため、必要に応じて顔の表情を作ったフレームを以下に示す。

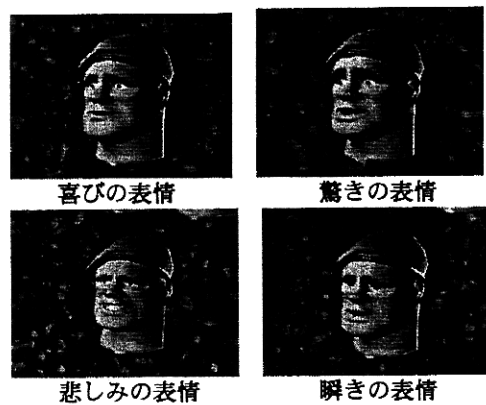


Fig. 13 主な表情のターゲット

### 3.4 リップシンク及び顔のアニメーション

ここまで、基本のモデルを作製し、視覚言語により英語の音素を分類し、さまざまな発音記号の特徴に対応したキャラクターのリップシンク及び表情のターゲットを作製した。それぞれのリップシンクおよび表情のターゲットをモディファイヤのチャンネルに割り当てている。作製したさまざまな発音記号に対応したリップシンク及び表情のターゲットをフレームとして扱って、英語の単語の発音によって、アニメーションができる。

英語単語の Good morning を実例として、キャラクターを使ってリップシンクのアニメーションをプレビューする。

good morning

gú:d mór:niŋ

発音記号は以上のように、[g]、[u]、[d]など九つの発音記号から構成され、九つに対応のターゲットをアニメートして、この中に、瞬きの表情のターゲットも含んで、Good morning のリップシンク及び顔のアニメーションができる。

同じ方法で、さまざまな発音に対応したリップシンク及び顔のアニメーションができる。

#### 4. 結論と課題

基本モデルを採用し、3次元キャラクターを製作した。単にキャラクターによるのアニメーションではなく、様々な発音記号の特徴に対応した口の動作を再現し、実際の発音に対応させ、キャラクターのリップシンクを忠実に再現することにより、より現実味のある発声動作を視覚的に表現することができた。更に視覚言語により英語の音素を分類し、短母音と長母音を含む単母音や、子音の様々なリップシンクの特徴を提案し、母音と子音の発音特徴と視覚言語の対応表を作成した。それにより、リップシンクの作製を効率よく行うことができる。

今後、視覚的な領域の作製が終わり次第、聴覚的な領域の研究に取り組み、サウンドシステムも本研究に取り入れていきたい。そして映像とサウンドをシンクロさせた、より現実味のある総合的なアウンサーシステムの製作を行っていきたいと考えている。

#### 参考文献

- 1) Duane Loose (著) トップスタジオ (訳) :  
3d studio max 3 Workshop, エムディエヌコーポレーション社 (2000)

- 2) Sung-Wook Su(著) 山本健介 (監訳) :  
3D Studio MAX 完全ガイド, オーム社 (1998)
- 3) 坂本 勉他: 現代言語学入門1, 岩波書店 (1999)
- 4) 鷺見 由里: 英語の発音が正しくなる本,  
ナツメ社 (2000)
- 5) 五十嵐 康男: CD BOOK 英語正しい発音ができる,  
ベレ出版 (2000)