

人体モデリングソフトウェアによる 手話のコンピュータグラフィックスアニメーション制作 Computer Graphics Animation of Sign Language using Human Modeling Software

平山 亮†

Makoto J. HIRAYAMA†

† 金沢工業大学

† Kanazawa Institute of Technology

E-mail: † mako@infor.kanazawa-it.ac.jp

1. はじめに

聴覚障害者への情報保障及びコミュニケーション支援に関しては、いくつかの手段が考えられるが、手話を日常的に使っている聴覚障害者へは、手話による情報保障が望まれる。

手話を提示するには、手話を第1言語とする人あるいは手話通訳者によるのが、理解度及び自然性の面から好ましいが、そのような人材は多くなく、手話通訳者の養成は必要であるが容易ではない。そこで、手話を提示するのに、あらかじめ録画したものを再生する方法、リアルタイムに生成して提示する方法が研究されてきている。実際の人間を撮影した映像をベースとするもののほか、コンピュータグラフィックス (CG) による人物モデルを作成してアニメーションをすることも研究されてきている。CGによる手話アニメーションシステムはこれまでにいくつか提案されているが、手話通信や手話通訳を想定して実時間表示を重視したシステム、情報サービス端末等での手話表示を想定して、辞書を内蔵し日本語を入力することで手話CGアニメーション作成が簡易に作成できる補助ツールなどが研究開発されてきている[1]～[13]。

これまでに、手話によるコミュニケーションシステムの構築を想定した手話認識及び手話合成の技術開発をいくつか行ってきたが[14]～[25]、本報告では、誰にでも比較的簡便かつ高品位な手話のコンピュータグラフィックスアニメーションをできる市販の人体モデリングソフトウェアによる手話アニメーション作成方法を紹介し、制作したアニメーションを紹介する。

本報告では、実時間アニメーションにはこだわらず、CGによる指文字及び手話単語・短文を部品として作成しておき、自由に組み合わせて、高品位なCGアニメーションを比較的簡便に作ることをできるようにすることを中心に行うことを目的とする。CG映像作品

制作などへの対応も可能なように、汎用のCGソフトウェアとの連携も想定している。市販の人体アニメーションソフトウェアを用い、指文字及び手話単語・短文のCG部品ライブラリを作成するという方法で行う。

2. 人体アニメーションソフトウェア

使用した市販ソフトウェアは、Smith Micro社のPoserである。執筆時点で、Poser 9及びPoser Pro 2012が最新である。両者は機能・操作ともにほぼ同じであるが、Poser Pro 2012の方はモデルを外部のコンピュータグラフィックスソフトとのデータのインポート/エクスポートの機能が充実している。Poser Pro 2012英語版を使用しているが、日本語版も近日中にe-frontier社から発売される予定である。本報告で説明する方法は、Poser 9、または、Poser 7以降であれば、ほぼ同様に使用することができる。実際、本報告で紹介するCGモデルのいくつかは古いバージョンで作られた。



図1. ソフトウェアの操作画面の例
(Poser 紹介ホームページより)

このソフトウェアでは、数種類の人体モデルがあらかじめ用意されていて、人体の骨格構造に応じた構造

がモデル化されている。CG ソフトウェアでいうところのいわゆるボーンとジョイント及び表面のメッシュの変形モデルがあらかじめ作られているため、ゼロから人体 CG モデルを作成していくのに比べ非常に手数が少なく済む。また、別の人体モデルを追加購入することもでき、男性、女性、人種、体系、顔、リアルな人間モデルか漫画モデルかなど、様々に違ったモデルを登場させることができる。人体モデルの基本構造はモデルごとに異なる場合があるが、数種類ある基本モデルと同じ基本構造をもつモデルどうしであれば、一つのモデルで作った動作を別のモデルにも適用することができる。

人体モデルに動作をさせる場合、例えば、指を曲げるという動作は、動かそうとする指をドラッグして引っ張るだけで曲げることができる。マウスのドラッグ操作だけだと厳密な操作はできないので、数値入力による姿勢指定も可能である。インバースキネマティクスも考慮されており、たとえば手先を目的の位置にドラッグすると、それに応じて、手首、肘、肩も動いてくれるため、関連するすべての関節角を一つ一つ決めていくといった作業もしなくて良い。

アニメーションを作るときには、タイムライン上に設定したキーフレームに姿勢を設定することによって、キーフレーム間の中間のフレームは線形または非線形に補間をする機能が備わっているの、コマ撮りアニメーションのように一コマずつちよつとずつ変えて制作していく作業が必要ない。

また、作った姿勢をライブラリとして登録しておくことができるため、同じ姿勢を何度も作らなくとも、登録してあるライブラリを呼び出すだけで、姿勢を設定できる。

3. 人体モデルの基本構造

基本の人体モデルの構造として、“G2” (Generation 2) と呼ばれる構造の、Poser 7J のサイモンという名の基本キャラクターを使用した。以前は人体モデルごとに基本構造が違って互換性がなかったのが、この G2 構造に準拠した人体モデルならば、ある人体モデルで設定した姿勢・手形状を、別の人体モデルにも適用可能であるという利点があり、制作する CG アニメーションにバリエーションをもたせることができる。CG 人体モデルの基本構造は、人体全体を部位に分け、腰部を親とした階層構造になっている。G2 構造に準拠した人体モデルの階層構造を次に示す。

```
--- 腰部
---- 腰
----- 腹部
----- 胸
----- 首
----- 頭部
----- 左目
----- 右目
----- 左肩
----- 左腕
----- 左前腕
----- 左手
----- 左親指 1
----- 左親指 2
----- 左親指 3
----- 左人さし指 1
----- 左人さし指 2
----- 左人さし指 3
----- 左中指 1
----- 左中指 2
----- 左中指 3
----- 左薬指 1
----- 左薬指 2
----- 左薬指 3
----- 左小指 1
----- 左小指 2
----- 左小指 3
----- 右肩
----- 右腕
----- 左前腕
----- 右手
----- 右親指 1
----- 右親指 2
----- 右親指 3
----- 右人さし指 1
----- 右人さし指 2
----- 右人さし指 3
----- 右中指 1
----- 右中指 2
----- 右中指 3
----- 右薬指 1
----- 右薬指 2
----- 右薬指 3
----- 右小指 1
----- 右小指 2
----- 右小指 3
<--- 下半身省略 --->
```

例えば、上腕－前腕－手という階層構造があり、ジョイントとボーン（関節と骨）が設定されている。姿勢データは、人体各部の基準姿勢からの、ジョイントの屈伸及びメッシュの変形をパラメータ数値であらわす。基本的な屈伸は、関節角をパラメータ化した数値で表されるから、変形を伴わない単純な関節角運動への変換は容易であり、したがって、MPEG-4 の Body Parameter (BP) [26] や H-Anim [27]パラメータへ変換して、通信向けのアプリケーションで使用することもできる。

頭部に関しては、細かい階層構造がなく、頭部、左目、右目という部位に分かれているだけである。ジョイントとボーンによる関節の屈伸による姿勢変更は、首と頭部の連結部での、前後、左右、ねじりだけであり、その他の器官の運動は、全て、メッシュの変形で表す。頭部に関しては、より細かい構造モデルが必要な場合はあるが、基本の人体モデルの構造変更は行わず、メッシュの変形により表情生成などを行った。

この構造は読み込み・書き出しが可能であり、頂点の追加・削除など構造が変わってしまう変更を行わない限りは、外部のCGモデリングソフトウェアにより外観は自由に変更できる。外観を変更するためのCGモデリングソフトウェアとしては市販ソフトウェアの3ds maxを使用している。

4. 指文字及び手話CG部品作成[18]

CGアニメーションのキーフレームに配置して使える部品として、基本となる手形状の姿勢を作り、ライブラリに登録する。ライブラリに登録することで、あとからその姿勢を呼び出して適用できる。まずは、指文字及び基本の手話CGを作成する。

4.1 指文字及び手話CG部品作成

部品には、静止した姿勢の部品と、動きのある姿勢の部品とがある。例えば、指文字“あ”は静止した部品として1フレームだけを登録するが、指文字“り”は開始時の指形状、終了時の指形状をキーフレームとした複数フレームのアニメーションを部品として登録する。

図2に指文字「あ」「い」「う」「え」「お」の部品データを示す。



図2. 指文字部品の作成
(あ, い, う, え, お)

4.2 部位の分割

姿勢、手型、顔（唇・舌を含む）の3つは、それぞれ別々に、作成し、登録する。

一般に手話における手指動作の音韻構造は、手の形状、位置、動きから成り立っているといわれる([28],[29]などを参照)。手の形状は、手型と向きとに分ける場合もある。向きを考慮しない場合(指文字“う”と指文字“と”は同じ)、手型は48種類([28])作成すればよく、これらをライブラリの“手”カテゴリとして作成した。

“手”カテゴリは、前腕で制御される手首の向きは含まず、手の向きは、姿勢で区別する。手の向きは、前腕のねじりで変える、手のひらが表向き、裏向き、内向きという種類と、上腕の上下で変える、手の先端(指先)が上向き、下向き、内向き、外向きがある。手の向きだけでなく、腕全体の姿勢・動きも腕の姿勢はパラメータで表す。又、うなずきなど、首の前後左右についても姿勢に含まれる。

顔は、全て頭部という一つの器官として作成する。骨格筋ではないから、ボーンとジョイントによる関節運動はなく、全て、メッシュの変形により、表情を作る。非手指動作として、口形、眉上げ、まばたきなどを作成する。

4.3 非手指動作について

手話においては、手指動作だけでなく、非手指動作

が文法的に重要な意味を示す[30]. 非手指動作については口形及び眉毛, 目, 首の動きを中心に, 顔出の非手指動作に関してはCG部品作成を行った. 口形として, 発話時の口の動きのアニメーションの研究[31][32]で得られた知見を参考に, “あ” “い” “う” “え” “お”, “ば” “び” “ぶ” “ぺ” “ぽ” を作成した. 眉毛は, “眉上げ” “眉下げ”, 目は, “まばたき” “首振り”, “首振り”, “上向き” を作成した. これらに加えて, 微笑み, 怒りといった感情表情も複合的に組み合わせることができる. しかし, 非手指動作に関しては, まとまった文法書がなく, わからない部分も多いので, 入門書[29]に記述してあるレベルのものを取り入れた.

図3は非手指動作の一例であり, 「非手指動作なし」「うなづき」「まばたき」「まゆあげ」を示す.



図3. 非手指動作 (“非手指動作なし” “うなづき” “まばたき” “まゆあげ”)

4.3 手話単語

図4は単語の例で, “石川” の地名手話単語である.



図4. 手話単語の例 (“石川”) (左上~右下)

5. 手話アニメーション

4. で作成した手話単語を組み合わせることでCGアニメーションを行う. アニメーションの際のフレームレートは, 15コマ/秒程度が実用的である.

図5は, 「こんにちは」 (あいさつ) の手話短文のアニメーションの一部である.



図5. 手話短文の例 (“こんにちは” (あいさつ))

6. キャラクターモデルの変更

状況や要求に合わせて, 様々な人物キャラクターを使うことができるので, 様々なコンテンツに合わせてそれにふさわしいキャラクターを使うことができる.

図6はMikiというキャラクターモデルを使って作成した例であり, 地震などの災害の速報などを行うため, BML を使って携帯電話等のワンセグテレビへ手話を配信する実験をしたときに作成したものである [19][20].

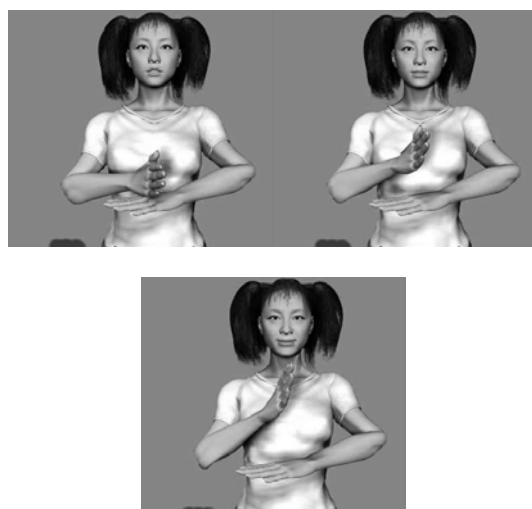


図6. 手話短文 (“ありがとう”)

また、図7はウェブのコンテンツにちょっとした手話のイラストを載せる必要があったときに作成した、にゃあみいというアニメ風キャラクタによるものである。これも販売されているキャラクタであり、ゼロからキャラクタ作成をするのに比べれば格段に開発時間が少なく済む。これらを用いて作成したコンテンツを公開することも著作権上問題ない。ただし、市販のキャラクタによって、体型や顔などの意匠を変えることができなかつたり、作品公開に制限がある場合など、著作権上の制限があるものがあり、キャラクタによって違うので、コンテンツを作って公開したり、それを販売したりする場合には、使用条件によく注意する必要がある。



図7. アニメ風キャラクタの例 ("I Love You.")

7. おわりに

市販の人体モデリングソフトウェアを使って、指文字及び手話の基本単語・短文のCG部品を作成し、それらの部品を呼び出して、連結することにより、手話のCGアニメーションを制作した。

本報告で紹介した方法により、比較的簡便に手話アニメーションができる。このような方法を使って、アクセシビリティに配慮したウェブページコンテンツ作成をはじめ、さまざまな場面での案内や説明のコンテンツを作成していくとよい。

参考文献

- [1] 比留間伸行, 清水俊宏, 梅田修一, 加藤直人, 宮崎太郎, 井上誠喜, 金子浩之, 長嶋祐二, CGによる手話アニメーションの自動生成システム, 画像電子学会誌, 41(4), 406-410, 2012.
- [2] 猪木誠二, 渡辺錬士, 呂山, "手話アニメーション作成・編集ツール," 電子情報通信学会論文誌, D-I, J84-D-1(6), 987-995, 2001.
- [3] 由井蘭隆也, 原康祐, 中山茂, "Web対応分散協調型手話アニメーション作成支援システムの開発と評価," 電気学論文誌, C, 123(10), 1745-1752, 2003.

- [4] 黒川隆夫, "日本語手話翻訳における手話アニメーション表示," 可視化情報学会誌, Suppl. 24(1), 273-276, 2004.
- [5] 河野純大, 黒川隆夫, "日本語手話翻訳システムの開発," 知能と情報: 日本知能情報ファジィ学会誌, 16(6), pp.485-491, 2004.
- [6] 浜上和也, 和泉千絵, 森本一成 他, "画像表示条件と手話アニメーション読取りの関係," ヒューマンインタフェース学会研究報告集, 9(1), pp. 1~6, 2007.
- [7] 才原崇文, 川村, 昌平, 森本一成, "3次元手話アニメーションの開発とわかりやすさの評価," ヒューマンインタフェース学会研究報告集, 9(1), pp. 7-12, 2007.
- [8] 出淵亮一朗, 末松亜斗夢, 八木良一, 長嶋祐二, "手話アニメーション合成エディターの開発," 信学技報, 106(57), pp. 53-58, 2006.
- [9] 出淵亮一朗, 八木良一, 姫野泰典, 長嶋祐二, "手話単語アニメーションライブラリと手話単語合成エディターの開発," 信学技報, 107(61), pp. 69-74, 2007.
- [10] 木村勉, 中田修司, 中村有希, 神田和幸, 原大介, "音素記述による手話アニメーション作成支援システムの開発," 信学技報, 105(684), pp. 71-76, 2006.
- [11] 大高崇, 西田昌史, 堀内靖雄, 市川薫, "単語の時間構造を考慮した手話CGアニメーションの生成," 信学技報, 105(507), pp. 37-42, 2006.
- [12] 木下健, 西田昌史, 堀内靖雄, 市川薫, "手話における手動作のモデル化に基づくCGアニメーション."
- [13] 竹内晃一, 鈴木雄介, 宮本一郎, 三樹弘之, "実写画像ベースの手話映像合成技術を利用した情報提供システム," 信学技報, 105(67), pp.1-6, 2005.
- [14] 平山亮, 辻原和昌: "手話アニメーションシステムの作成", 電子情報通信学会ヒューマン情報処理研究会, 2005.
- [15] 舟川政博, 平山亮: "指文字画像からの手指形状特徴量抽出方法の検討" FIT (情報科学フォーラム) 2006, 87-88.
- [16] 辻原和昌, 野志恭兵, 平山亮, 舟川政博: "人物コンピュータグラフィックスによる手話のアニメーション" FIT (情報科学フォーラム) 2006, 561-562.
- [17] 舟川政博, 平山亮: "指文字の静止画像識別システム—手領域の面積・重心・外接長方形による識別法", 電気関係学会北陸支部連合大会, F-71, 2006.
- [18] 平山亮: "指文字及び手話のCGアニメーション", 電子情報通信学会技術研究報告, HIP2007-122, 121-126, 2007.
- [19] 辻原和昌, 平山亮, 水野舜: "BMLを用いた手話による地震情報の伝達システムの検討", ヒューマンインタフェース学会研究報告集, 9(5), pp. 13-18, 2007.
- [20] Makoto J. Hirayama, Kazuaki Tsujihara, Shun Mizuno, "An Earthquake Information System for Hearing-Impaired via Digital Terrestrial Broadcast," 5th International Conference on Information Technology and Applications (ICITA 2008), pp. 21-26, 2008.
- [21] Makoto J. Hirayama and Masahiro Funakawa, "A parameter extraction method for recognizing

- Japanese finger alphabet," The 15th Japan-Korea Joint Workshop on Frontiers of Computer Vision (FCV2009), pp. 146-151, 2009.
- [22] 平山 亮, 舟川政博, 手袋型計測装置を用いた実時間指文字認識, 情報処理学会第 71 回全国大会, pp. 4-9-10, 2009.
- [23] Makoto J. Hirayama, Sign Language Animation using Human Modeling Software, NICOGRAPH International 2009, Kanazawa, Japan, P-04, 2009.
- [24] Makoto J. Hirayama, Qinglian Guo, "Sign Language Image Generation by Concatenation of Computer Graphics Animation," 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV2010), Hiroshima, Japan, pp.251-256, 2010.
- [25] 平山亮, 舟川政博, "ニューラルネットによる静止画像からの指文字認識," 情報処理学会創立 50 周年記念(第 72 回)全国大会, 東京, 1D-1, 2-13-14, 2010.
- [26] ISO/IEC, Face and body object decoding tables and definitions, ISO/IEC 14496-2:1999 Annex C, 1999.
- [27] Humanoid Animation Working Group, Specification for a Standard Humanoid, Version 1.1, <http://www.h-anim.org>.
- [28] 神田和幸・藤野信行, 基礎からの手話学, 福村出版, 1996.
- [29] 神田和幸, 手話学講義, 福村出版, 1994.
- [30] 木村晴美, 市田泰弘, はじめての手話, 日本文芸社, 1995.
- [31] 平山亮, "音声発話時の口唇周辺高速動画画像データベース," FIT(情報科学フォーラム)2003, F-025, pp. 257-258, 2003.
- [32] 平山亮, "日本語音声発話口形素の CG モデル," 日本音響学会 2004 年秋季研究発表会講演論文集, pp.389-390, 2004.