

表情の表現を含む手話アニメーションの試作

児玉哲彦¹ 安村通晃²

概要

手話は聴覚障害者にとって重要なコミュニケーション手段である。コンピュータグラフィックス技術の進展とともに、手話アニメーションによる情報提示の利用が注目されている。我々は、その中でも表現力の向上が望まれる、話者の表情の合成手法に着目し、感情表現を伴う実映像からのアニメーション制作手法を提案し、合成システムを実装した。表情の符号化手法として Facial Action Coding System (FACS) を用いている。実際に3編の手話アニメーションを試作し、WWWを通じて公開して文の理解度と親しみやすさの評価を行った。

Experimental Development of Animated Sign Language Including Facial Expression

Akihiko Kodama³, Michiaki Yasumura⁴

The use of sign languages is one of the important means of communication for people with hearing disabilities. Along with the progress of computer graphics technology, presenting information with animated sign language has been of keen interests among the researchers. However, current animations lacks facial expressions and more improvement in this area is desired. This paper proposes a new technique that creates animations from video-recordings of live performances including emotional expressions, in order to synthesize the speaker's facial expressions. We implemented an animation development system by using Facial Action Coding System (FACS) as a coding technique for facial expressions. We, then created three pieces of prototype animations using the system and evaluated the degree of understandability of the intended messages and the familiarity of the animations.

1 はじめに

1.1 背景

手話は聴覚障害者のコミュニケーション手段として広く用いられている言語である。手話への工学的なアプローチは手話工学と総称され、近

年目覚ましい発展を遂げている [1]。手話工学は、言語としての手話の調動の解明、電子的な記述法、認識、画像生成、辞書制作を扱い、インターフェースとしての手話の利用と、手話使用者のスムーズなコミュニケーションの実現を最終目標としている。

その中でも、手話画像生成は、CGにおける技術の蓄積を直接活用することが可能で、応用分野も辞書、学習コンテンツ、コミュニケーションツール等、幅広く考えられる。

¹慶應義塾大学政策・メディア研究科

²慶應義塾大学環境情報学部

³Graduate School of Media and Governance at Keio Univ.

⁴Faculty of Environmental Information at Keio Univ.

わかりやすく親しみやすい手話アニメーションを実現する上で課題となっているのが、表情の表出である。本研究では、より豊かな表情表出を伴う手話アニメーションを実現する手法を提案する。実際に制作した3編の例文アニメーションをWWWを通じて公開し、アンケート調査を行ってシステムの有用性を評価した。

1.2 先行研究

京都工芸繊維大学の黒川らを中心に、手話アニメーションの包括的な研究がなされてきた[2]。また、郵政省通信総合研究所ユニバーサル端末研究室は、手話認識生成技術を開発した[3][4]。実用化された手話アニメーション技術として、日立製作所のMimehandが挙げられる[5]。

日立と通信総合研究所の研究において共通するのは、モデル化を行わずに用意したパターンの組み合わせによって表情を表現しているため、表現できるパターンの数が限られる。筑波技術短期大学の学生を対象に、通信総合研究所のシステムについての簡易な調査では、表情の少なさから手話の魅力を表現しきれていないという意見が多かった[6]。しかし、顔の表出は非常に複雑であり、あらゆるパターンを用意するのは現実的ではない。多くの表情アニメーションの研究では、表情の認知と表情筋に関する基礎検討を基盤とした、表情の要素分解とモデル化を用いている[7]。

黒川らの研究においても、手話の理解のために必要な動作要素を抽出し用いているが、本研究ではより幅広い感情表現を伴う表情の実現のために、Ekman、Friesenらの開発したFacial Action Coding System[8]の利用を提案する。アニメーションモデルについては、本研究でも黒川らの手法[2]を概ね踏襲し、ポリゴンメッシュの骨格構造による変形によって動かした。

2 手話の調動

日本国内で用いられている手話は、日本語の語順と接続詞を伴う日本語対应手話と、独自の語順で主に非手指動作でアクセントやイントネーションを表現する日本手話とに分かれる。制作

したアニメーションは正確な日本手話ではなく、両者の混淆した文法である。

手話言語においては、副詞的情報(様態、程度、陳述)が手指単語と共起する非手指動作(顔の表情など)によって表現されることが知られている[9]。また、語の切れ目やアクセント、イントネーションなどのことばとしての情報(周辺言語情報)、さらには喜怒哀楽などの心理性やその人らしさのような個人性(非言語情報)を伝達する役割を果たしている[10]。

本研究では、感情表出のための顔部動作に加えて、瞬き/頷き/頭傾動作を加えた。また、後述するように顔部動作を強調した。

手話の発話は、多くの場合音声の口形を伴い理解の助けとしている。本研究においても顔面下部は口形で音素情報を伝え、感情表出は主に顔面上部の動作で表現している。

人間の表情における感情表出は、Ekmanらの研究[11]によって、人種や地域を問わず中立と幸福/悲しみ/嫌悪/怒り/恐怖/驚きの6つの基本感情に分類できることが知られている。また、他の多くの感情表出は、6感情の合成で表現できる。

本研究では手話においてもこの分類が有用であると仮定し、6感情のうち3感情の表出を伴うアニメーションを制作し、感情の認識率を測定した。

3 表情アニメーション合成システム

3.1 システムの構成

本システムによるアニメーション合成において必要となるのは、

1. モデル化、記述方式
2. 調動の分析
3. キャラクターモデリング
4. アニメーション生成

の四つのステップである。図1に示すように、まず手話演技をビデオ映像に撮影する。撮影した映像から、時間軸に沿った調動の変化を、複

数の記述方式で符号化する。その記述に基づき、動作の合成を行い、最終的にアニメーションを出力する。

また生成したアニメーションを提示する評価実験も行った。

制作工程

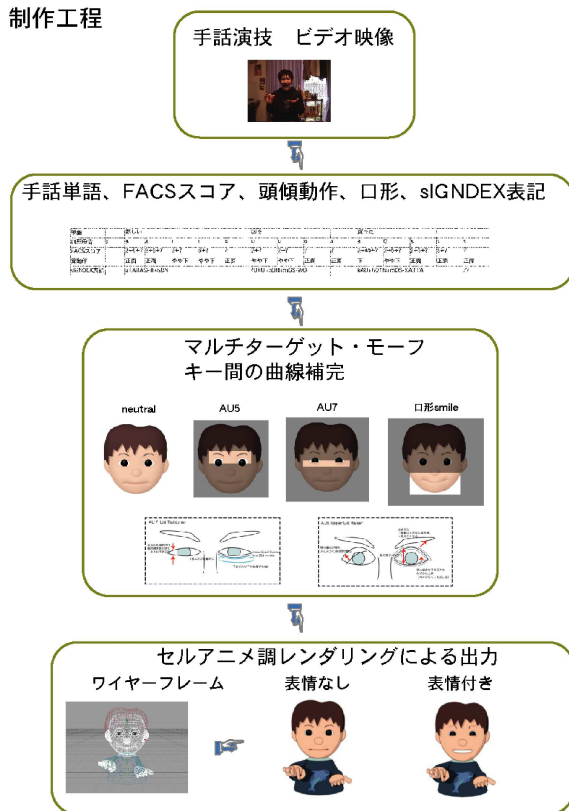


図 1: 制作工程

3.2 モデル化、記述

手話調動の記述には、複数の要素を時間軸状に記録する必要がある (図 2)。

第一に、sIGNDEX 表記法 [12] を用いる。これは、神田らが提唱した、手話を電子的に扱うための符号化手法である。単語表現と、頭傾や瞬きなどの非手指表現の双方を記述できる。

一方、sIGNDEX では不十分な表情表出を補うため、FACS を利用する。これは、表情を視覚的に認知しうる最小単位 (Action Unit) に分解し、その組み合わせで表情を表現する記述方式である。表情の研究において広く用いられており、また時間的な変化の記述にも適している。

手話における顔部の動作は、音素と対応した

口形の変化を含むため、単語の音素も記録した。



図 2: 記録した要素

3.3 調動の分析

手話使用者が、感情表出を伴う文章の発話の様態を演じる様子を、ビデオを用いて撮影し、調動の分析と記録を行った。この過程では、手話の表情の感情表現における特性を確認することを目的とした。

3.4 キャラクターモデル

感情を誇張して表現するため、デフォルメしたデザインのキャラクターを制作した。図 3 が画像の例である。特に、手と顔を拡大し、強調した。またセルシェーディング技術を用いて、CGの質感の不自然さを回避した。



図 3: キャラクター 「Teo」

3.5 アニメーション生成

開発環境としては、CGアニメーション制作ソフトウェアである Cinema 4D XL を利用した。同ソフトウェアにはオブジェクト指向スクリプティング言語 C.O.F.F.E.E. が搭載されており、表情メッシュの制御システムを開発した。

体幹動作は、ボーン変形によって手動で作成した。顔の動作は、まずキーとなるメッシュを合成する。FACS の動作単位 (AU) に基づく顔面上部動作と口形を、マルチターゲット・モーフィングした。図 4 に、AU1 (内側の眉上げ)、AU44 (薄目)、[a] の口形を合成して悲しみの表情を合成する様子を示している。さらに、合成されたキーメッシュ間を、曲線的な補完により通時的に滑らかに接続した。今回は 3 編の例文アニメーションを制作した (図 5)。

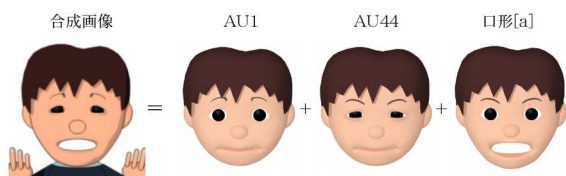


図 4: メッシュ合成



図 5: 作成画像例

4 評価実験

聴力に関して先天性失調／中途失調／健聴のそれぞれの手話使用者がおり、また聴覚障害者においても聴力の差、教育環境の違い等で、手話の習熟度や利用する手話は一様ではない。今回は、慶應義塾大学、筑波技術短期大学聴覚部、難聴児教育団体トライアングル、慶應義塾大学手話サークル MiMi、東京大学先端科学技術研究センター等より被験者を募り、なるべく多様な手話使用者を対象とした。調査には WWW を利

用し理解度と主観評価のアンケートをとった。

手法としては、表情動作を付与した映像と表情動作なしの映像の双方を被験者に提示し、アンケートに回答させた。被験者は映像を何回でも再生することができる。アンケート項目は表 1 の通りである。85 件の有効な回答があった。

項目	回答方式
文意	自由記述
感情	6 感情＋わからない
表情なしの親しみやすさ	5 段階
表情ありの親しみやすさ	5 段階
表情なしのわかりやすさ	5 段階
表情ありのわかりやすさ	5 段階
意見・感想	自由記述
聴覚障害の有無	選択
手話使用歴	選択
年齢層	選択

表 1: アンケート項目

5 実験結果

5.1 文意と感情の正答率

まず、文意と感情の正答率の結果を見ると、図 6 の通りである。

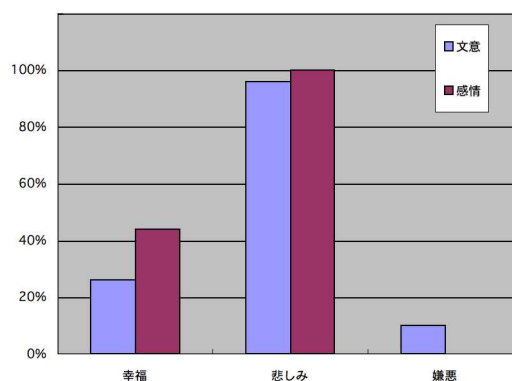


図 6: 文意と感情の正答率

一見して分かる通り、文章ごとのばらつきが大きくなっている。悲しみの例文が突出して正答率が高い。これは 2 語文であることから他の 2 文と比較して単語数が少なく、また単語も非常

に基本的な単語であるため、理解しやすかったためである。

一方幸福の例文は、感情は半数近くの被験者が正答しているものの文章の正答率が低い。有意な相関は得られなかったが、傾向として手話使用歴の長い高齢層の正答率が高く、また正答者ほとんどが聴覚に障害があった。感情の回答には年代や手話使用歴による偏りは全く見られなかった。感情の正答者の約1/3が文の正答者であり、手話文を読み取れるかが感情のわかりやすさにも反映されている。

嫌悪については文意の正答率は10%、感情は0%であった。文意に付いては、特殊な文脈における文章で、単語も一般的でなかったことから、文脈なしには理解しにくかったと考えられる。また、嫌悪という感情の例文として適切な文章ではなかったことが、感情の正答を困難にしていた。

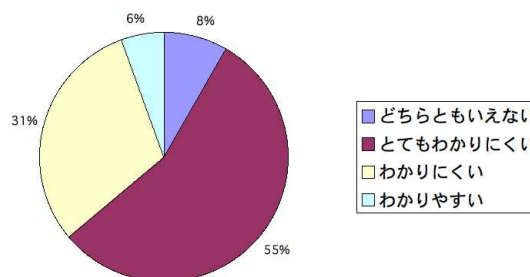
5.2 わかりやすさ、親しみやすさ

この項目は、主として表情の導入効果を主観的に調査するために導入した。理解度の差の大きかった幸福のアンケートの結果は図7、図8の通りであった。わかりやすさについて(図7)は、表情がない場合には「とてもわかりにくい」と「わかりにくい」が合計で86%を占めていたが、表情付きについては50%まで減少し、代わって6%だった「わかりやすい」が33%に増加している。

親しみやすさ(図8)は、さらに顕著で、表情なしの場合には親しみやすいという回答はなく、「とても親しみが持てない」が39%、「親しみが持てない」が33%であったのが、それぞれ11%、8%へと減り、代わって「親しみが持てる」が48%、「とても親しみが持てる」も8%となっている。

いずれの項目についても表情への評価が非常に高い。これは残りの2例文でも、ほぼ同様の結果となっている。

表情なしのわかりやすさ



表情付きのわかりやすさ

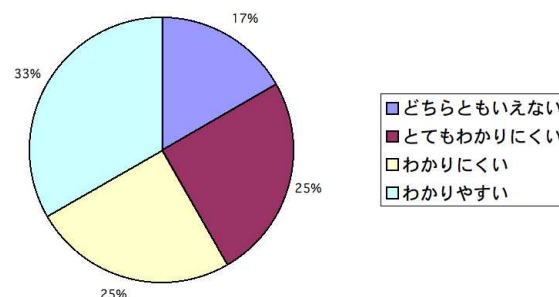


図7: 表情導入によるわかりやすさの比較

5.3 自由記述による回答

自由記述で多かったのは、口形が不自然で読み取れないという意見だった。他には、アニメーションの速度が速すぎて読み取りにくい、またセルシェーディングを用いたために手指が平面的で、形が読み取りにくいという意見も多かった。正しい手話文法にのっとっていないという指摘もあった。キャラクターは親しみやすいという意見が多かった。

5.4 システムの有用性

文意においてはやや厳しい結果ながら、正答率0となることはなかった。今後、正しい日本手話の文法、頭傾・瞬き以外の非手指動作、速度、口形、立体感等の面で改善を加えていくことで、理解しやすさを上げていくことができると考えられる。

一方、本研究の主眼であった感情の理解は、嫌

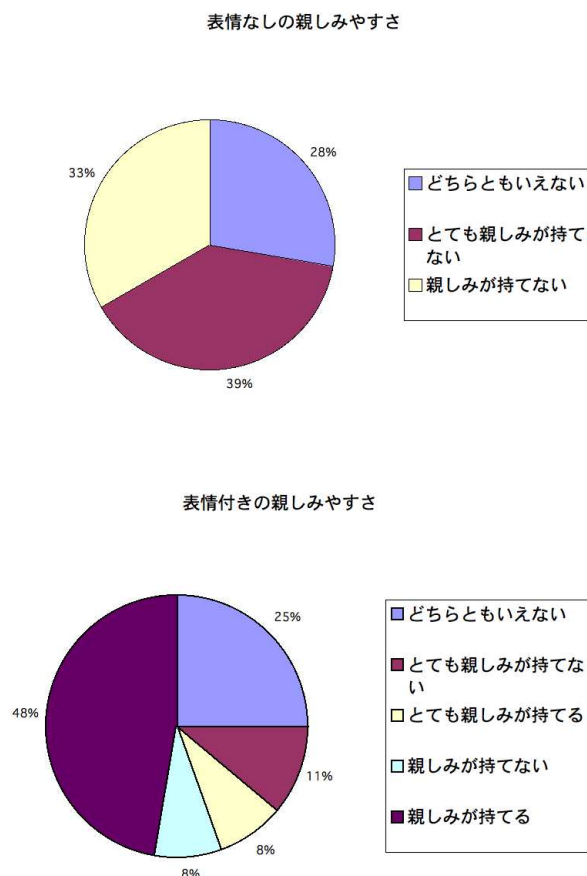


図 8: 表情導入による親しみやすさの比較

悪については制作した例文の問題などで 0% という結果に終わったが、幸福と悲しみは良好な結果を得た。表情による感情表出の効果は確かめられたといえる。

コンテンツ開発の効率化という観点からは、キーフレームごとにパターン化されたターゲットを制作する既存の手法の場合、この 3 例文に相当する表情アニメーションを作るには 40 のターゲットが必要となるが、本手法においては 14 のターゲットのタイミングの指定で自動的にアニメーションを合成する。このことから、関連研究と比べて手話の認識率ではまだ充分ではないが、より多様な表情を少ない労力で制作することは実現されたといえる。

6 おわりに

本研究では、親しみやすく、かつ感情表出を伴う手話アニメーションキャラクターについて

研究した。具体的には、手話アニメーションにおける非手指動作、特に表情の制作手法の提案を行った。

アンケート調査の結果、親しみやすさ／わかりやすさとも表情の導入効果が確認できた。特に親しみやすさは評価が高く、若い年代ほどその傾向が強かった。感情の認識率も一定の水準を達した。一方で、口形のわかりにくさ、手話文法の精度などから、文意の認識率には課題が残った。

アニメーションにおける表情の制作手法としては、既存のもの比べて少ない作業量を実現し、手法としての有効性が確かめられた。今後、文意の理解率を高めていくことで、学習コンテンツ等への応用が可能になると考えられる。

7 謝辞

本研究を行うにあたって、アンケート調査等に御協力いただいた筑波技術短期大学岡本明教授、大沼直紀教授、石原保志助教授、工学院大学の長嶋祐二助教授をはじめ、手話サークル MiMi、筑波技術短期大学を含む多くの方に感謝いたします。また手話の演技をお願いした石川尚貴君に感謝いたします。

参考文献

- [1] 加藤雄士, 神田和幸, 長嶋祐二, 市川熹, 黒川隆夫; 手話工学の現状と将来の研究課題; ヒューマンインターフェースシンポジウム論文誌'97, pp.37-44, 1997
- [2] 仙波和人, 緒方誠人, 黒川隆夫, 上田剛, 畑亮太, 中俊弥; 手話アニメーションモデルへの表情の導入; 第 13 回ヒューマン・インターフェース・シンポジウム論文集, pp.237-242, 1997
- [3] 郵政省通信総合研究所ユニバーサル端末研究室; <http://www2.crl.go.jp/jt/a131/indexj.html>
- [4] 猪木誠二; 郵政省通信総合研究所 H.11 年度調査報告書, pp.45, 1997

- [5] 日立製作所; Mimehand2;
<http://www.hitachi.co.jp/Prod/comp/app/shuwa/index.html>
- [6] 大沼直紀, 石原保志; 筑波技術短期大学内部資料, 2002
- [7] Demetri Terzopoulos, Barbara Mones-Hattal, Beth Hofer, Frederic Parke, Doug Sweetland, Keith Waters; Facial Animation: Past, Present and Future; Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp434-436, 1997
- [8] Paul Ekman, Wallace V.Friesen; Manual for the Facial Action Coding System, 1978
- [9] 市田泰弘, 江藤雄二; 日本手話学会第26回大会予稿集, pp16-17, 2000
- [10] 市川熹; 人と人をつなぐ声・手話・指点字, 岩波書店, 2001
- [11] Paul Ekman, Wallace V.Friesen, 工藤力; 表情分析入門, 誠信書房, 1987
- [12] 神田和幸, 市川熹, 長嶋祐二, 寺内美奈, 加藤雄士; 手話の文表記法 - Signdex V.2 -; ヒューマンインターフェースシンポジウム論文集'98, pp.761-766, 1998