

# 手話アニメーション生成システムの開発と評価

電子部 上蘭 剛, 呂 山\*, 猪木誠二\*, 坂戸博之\*\*

## Development and Experiment of Synthesizing System for Japanese Sign Language Animation

Tsuyoshi UEZONO, Shan LU, Seiji IGI and Hiroyuki SAKATO

郵政省通信総合研究所では、聴覚障害者と健聴者とのコミュニケーションを支援することを目的として、手話アニメーション生成システムを開発している。このシステムでは、身体の各部位の動きに注目した手話単語記述方式を基に、コンピュータグラフィックス(CG)アニメーションにより手話単語動作を生成している。このシステムについて、口動作をはじめとする諸機能を追加し、かつ改良した上で、聴覚障害者と手話通訳者、計3名を対象として18手話単語について再度評価実験を行ったところ、全体で91.5%(改善前:63.8%)の認識率を得ることができた。

### 1. 緒言

これまで、人と機械、人と人のインターフェイスは、健聴者だけが利用することを想定していたため、聴覚障害者を持つ人々は健聴者と同等のサービスや情報を享受できないことがあった。ここで、手話を日常のコミュニケーション手段として利用している聴覚障害者に対するサービスを考えた場合、情報は手話で提供するのが適切であると考えられる。そのため、最近ではテレビ電話を通じて手話通訳者を呼び出す事ができる交番や、手話通訳のできる社員を配置する百貨店、銀行なども登場してきている。しかし、手話に関心が集まっているとはいえ、一般社会に広まっていない現状を考えると、上記のようなサービスを聴覚障害者のニーズに応じて提供できる状況ではない。こうした問題の解決策として手話通訳の機械化、すなわち手話の認識と生成の研究が盛んに行われている<sup>1)~4)</sup>。

郵政省通信総合研究所では、聴覚障害者と健聴者のコ

ミュニケーションを支援することを目的として、日本語からコンピュータグラフィックス(CG)により手話動作を生成するシステムを構築している。本報告は上記研究所において研究を行った際の、システムの改善、機能追加と、改善後の評価結果について報告するものである。

### 2. 手話単語記述方式

本システムで使用している手話単語記述方式は、手話単語動作における手や指の動作について位置、運動、手形、両手の関係をパラメータとして記述するものである。本方式を適用した記述例を表1に示す。特殊な記述例として手の開始、終了位置があるが、これは手話動作を行う領域を図1のように105分割し、対応するであろう場所のコードを記述している。生成された手話CGアニメーションの1コマを図2に示す。

表1 手話単語記述例

	パラメータ項目	記述例			パラメータ項目	記述例	
1	単語	結婚	速い	16	左手終了位置	13, 1	—
2	右手開始位置	16, 1	16, 1	17	左ひじ終了位置	B	—
3	右ひじ開始位置	B	B	18	左運動の軌跡	直線	—
4	右手終了位置	左手小指	15, 1	19	左運動の方向	—	—
5	右ひじ終了位置	B	—	20	左手首運動	—	—
6	右運動の軌跡	直線	直線	21	左指運動	—	—
7	右運動の方向	—	—	22	左手開始の形	い	—
8	右手首運動	—	—	23	左手の方向	上	—
9	右指運動	—	変形	24	左手甲の方向	前左	—
10	右手開始の形	た	も	25	左手終了の形	—	—
11	右手の方向	右	右後ろ	26	両手の関係	—	—
12	右手甲の方向	後やや右	前上	27	右手指定点(*)	親指	—
13	右手終了の形	—	む	28	左手指定点(*)	—	—
14	左手開始位置	15, 1	—	29	フレーム数(*)	20	8
15	左ひじ開始位置	B	—				

(\*)改善による追加パラメータ

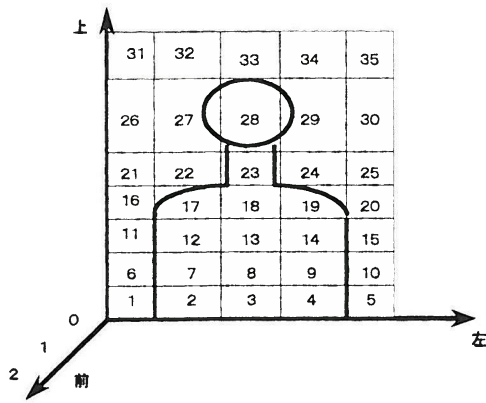


図1 開始, 終了位置コード表

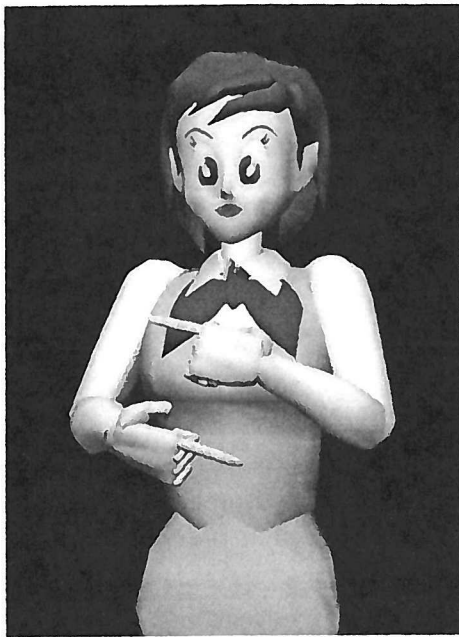


図2 手話CGアニメーション

### 3. 現状の問題点

改善前の問題点の代表的なものを下記に示す。

- 1) 特定の場所を指で指す場合に、正しく指されていない。
- 2) 体や口が動かないのは不自然である。また、表情にも変化がない。
- 3) 細かい位置指定ができないため、両手で微妙な形、動きを表現できない。

### 4. 改善・機能追加

以下の改善と機能追加を施した。下線の引いてあるものは主要な項目であるため、詳しい説明を後述する。

- 1) Windowsアプリケーションへの移植
- 2) 詳細部位指定方法の改善
- 3) 口型の追加
- 4) 表示速度定義機能の追加
- 5) 動作開始時の手首角度設定方法
- 6) パラメータの追加, 変更

#### 4. 1 詳細部位指定方法の改善

手話では、特定の指が特定の場所を指すことで意味を表現する場合が非常に多い。例えば“唇”を指して“赤色”を、“歯”を指して“白色”を表現したりする。このような場合、唇や歯といった詳細部位を正確に指すことが非常に重要になってくる。現在、手の位置は図1に示したとおり105分割された領域でしか指定することができない。詳細部位を指定しようとする場合、現システムでは単語ごとの微調整が必要となり、登録単語数を増やす際の大きな障害の1つになっている。そこで以下のような手法を用いて実現した。

準備として、現在の手話単語データに、どの指で指定するか（以下 指定点、例：人差指で唇を指す場合は“人差指”）というパラメータを左手右手それぞれ1個ずつ追加する。また、指定される点（以下 被指定点、上記例の場合“唇”）の3次元絶対座標をあらかじめ把握しておき、被指定点を示すラベル（上記例の場合“唇”という単語）に対応付けておく。

- ① 最初に、手のある決められた位置（指定点、被指定点によらず共通）に設定する。
- ② 指定点と被指定点の座標から、指定点が被指定点まで移動できるだけの方向と距離を求める。その距離の1/2（この値は指定点と手首の位置関係、つまり単語により適切な値が変化する。現在は指を伸ばした状態での適正值を採用している）の移動量を、唯一制御可能な手位置に反映させる。すると2点は一致しないまでもある程度近づく。
- ③ 手位置が新たに設定されるので、再度指定点を計算する。
- ④ ②と③を複数回繰り返すことで、指定点は被指定点に限りなく近づいていく。結果として指定点は被指定点を正確に指すことができるようになる。
- ⑤ ①～④を両手それぞれの開始、終了位置設定時に実行する。

つまり、現システムでは指定点を直接制御せず、手位置を介して制御する仕組みになっている。それは、手の角度、方向、形により変化する指定点を、直接制御するのは適切でないと考えるからである。また、手位置を変更すると腕の姿勢も変わり、指定点と手首の位置関係が変化してしまう。従って、一度に2点を完全に一致させることは、ほぼ不可能である。そのため、上記のような繰り返しを行うことで、2点間の距離を限りなく0に収束させるアルゴリズムを採用した。

#### 4. 2 口型の追加

手話は文字どおり手による言葉（言語）であるが、その他に重要な要素として表情、口の動きがあげられる。聴覚障害者が手話を見ている時の視線移動の様子を図3に示す

1)。ほとんど顔を、中でも口の近辺を凝視していることがわかる。そこで、口の動きについて若干の機能を盛り込んでみた。手法としては、かなに対応する口型の絵を作成し、顔のポリゴンにマッピングするようにした。

日本語は、口語において41音の かな を中心に濁音、半濁音、促音、長音を交えて構成されている。そのすべてについて口型を作るのは時間もかかるため、母音（‘a’、‘i’、‘u’、‘e’、‘o’）と”ん”（‘n’）についてのみ作成した。各口型のイメージを図4に示す。

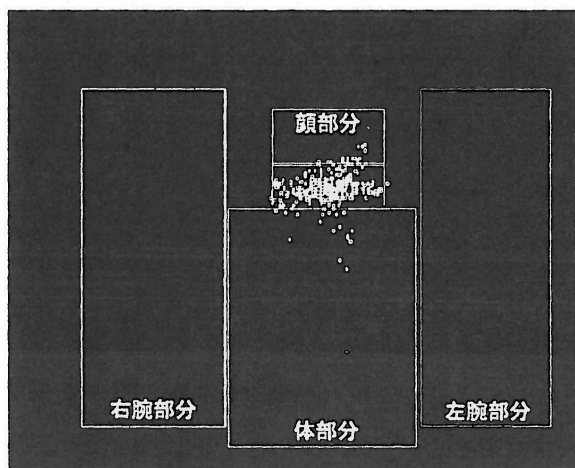


図3 聴覚障害者の手話認識時視線移動

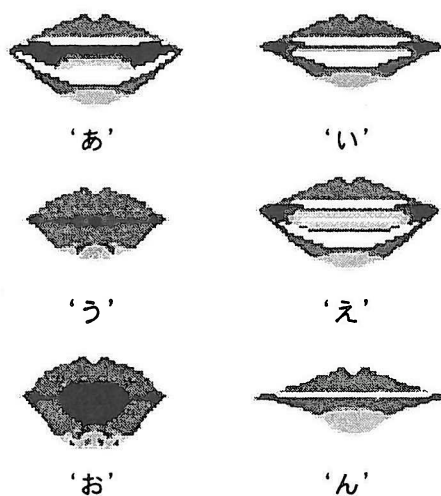


図4 各口型のイメージ

## 5. 実験方法

提示する手話単語は、郵便局での窓口業務で使われると思われる118単語を用意し、3度繰り返し見せた後、その動作が表現している手話単語を特定していただいた。被験者は日常生活で手話を使っている聴覚障害者2名と手話通訳者1名で、一人を除いて手話CGアニメーションを見るのは初めてである。

## 6. 実験結果

今回の評価はシステムとして2度目の評価になる。改善

前の評価において抽出された様々な問題点を改善することができたため、被験者数は少ないものの認識率は91.5%と改善前の63.8%から飛躍的に向上した。正答とした基準は、イラスト手話辞典<sup>5)</sup>、日本語・手話辞典<sup>6)</sup>を参考にして、回答が提示した手話単語名と異なった場合でも、動きが同一と思われるものは正答とした。正答率を表2に示す。また、改善前の評価実験の正答率を表3に示す。

表2 改善後正答率(単位:%)

被験者	簡単に特定	多少迷い特定	多少無理して特定
A	66.1	89.0	92.4
B	72.0	78.0	92.4
C	70.3	85.6	89.8
全体	69.5	84.2	91.5

表3 改善前正答率(単位:%)

被験者	簡単に特定	多少迷い特定	多少無理して特定
D	46.6	58.5	69.5
E	48.3	61.0	64.4
F	52.5	60.2	62.7
G	51.7	62.7	66.1
H	28.8	31.4	46.6
I	49.2	61.0	69.5
J	50.0	55.9	67.8
全体	46.7	55.8	63.8

## 7. 結 言

今回の改善において一番効果が現れたと思われる点はやはり口型を付けたことであった。被験者の意見として「口が動いていると安心する」「理解しようとする意欲がわく」「類似の単語も口の動きで判別できた」などがあつた。もちろん口型だけで単語の判別は不可能であるから、手話動作が良くなっている上での話である事は言うまでもない。

誤認識の原因の一つとして、手話が全国で完全に統一されていないという事実がある。表現方法が複数あつたり、方言のような地方独特の表現があつたり、省略した形での表現をしたり、人それぞれが異なった手話をする。そのため自分が日頃使用していない表現方法を見ても、理解することが難しいようである。

今後は単語だけでなく文章へも対応し、また表情なども追加していく事で、システムの完成度が更に向上すると考える。



## 参 考 文 献

- 1) 呂, 猪木: "ユーザの位置視線情報の検出とアクティブキャラクターの動作制御", 第13回ヒューマン・インタフェース・シンポジウム論文集(1997), p.281-286
- 2) 崎山, 大平, 佐川, 大木: "リアルタイム手話アニメーションの合成方法", 電子情報通信学会論文誌(1996), J-79-D-II, No.2, p.182-190
- 3) 河野, 仙波, 黒川: "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・インタフェース・シンポジウム論文集(1997), p.237-242
- 4) 長嶋, 亀井, 杉山: "形態情報駆動型手話アニメーションシステム", 電子情報通信学会技術研究報告(1996), E-T96-86, p.73-78
- 5) 丸山浩路編: "イラスト手話辞典", KKダイナミックシリーズ(1984)
- 6) 日本手話研究所編: "日本語・手話辞典", 日本手話研究所(1997)