

ファジィ・スプライン関数近似による手話テキスト作成の効率化について

渡辺 淳[†] 山田 晃 嗣[†] 中村 剛 士[†]
何 立 風^{††} 伊 藤 英 則[†]

手話学習のための教材の1つに、テキスト教材がある。これは、イラストや写真などの静止画像を利用して構成したものであり、最も一般的な手話教材の1つである。しかし、テキスト教材作成には、イラスト描写の専門的技術や写真構成技術が必要であり、人的にも経済的にも多大なコストを必要とする。本稿では、テキスト教材作成にかかるコストを軽減するために、ファジィ・スプライン関数近似による効率化を図り、実験において生成したいくつかの手話テキストを例示する。

A Supporting System to Make Sign-language Textbooks Using Fuzzy Spline Approximation

ATSUSHI WATANABE,[†] KOJI YAMADA,[†] TSUYOSHI NAKAMURA,[†]
LIFENG HE^{††} and HIDENORI ITOH[†]

You can common use illustrated textbooks as a kind of means for learning sign language. They are composed of static images, e.g., illustrations, photographs and so on, and are known as most general teaching materials. You need some professional skills or high cost to make the textbooks. In this paper, we propose the supporting system to make sign-language textbooks using fuzzy spline interpolations, and attempt to make the cost lower. The experimental results would show the efficiency of the system.

1. はじめに

近年、聴覚障害者の社会進出が目覚ましくなり、職場や公共の場所など様々な状況下で、手話はコミュニケーション手段として必要不可欠になりつつある。現在、日本の手話人口は20万人強に届くといわれ³⁾、健聴者が手話を学ぶ機会も多くなっている。また、地域住民が自主的に手話サークルを作るケースも増え、現在では、手話の地域サークルだけでも2000を超えるという報告もある。

手話を学習するためには、当然ながら学習教材が必要である。手話学習に使用される教材は一般に、ビデオなど動画像を使用したものと、テキストなど静止画像を利用したものに大別できる。これら2種類の教材

には、それぞれ学習における利点があり、手話学習用の教材として両者とも必要性が高いが、再生装置を必要とするビデオ教材に比べ、紙メディアで供給可能なテキスト教材は、学習する場所を選ばないという点で利便性が高いと考えられる。図1³⁾にその例を示すが、テキスト教材は、手話動作を描いたイラストや写真で構成され、一目で手話動作の流れを確認でき、反復学習に向けた教材であるといえる。しかしその一方、テキスト教材を作成するには、イラスト描写の専門的技術や写真構成技術が必要であり、人的にも経済的にも



図1 手話テキストの例³⁾

Fig. 1 An illustration of sign language.

[†]名古屋工業大学知能情報システム学科

Department of Artificial Intelligence and Computer Science, Nagoya Institute of Technology

^{††}愛知県立大学情報科学部地域情報科学科

Faculty of Information Science and Technology, Aichi Prefectural University

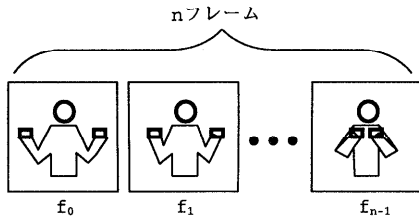


図2 手話単語動画のフレーム構成

Fig. 2 A frame sequence of a sign-language word.

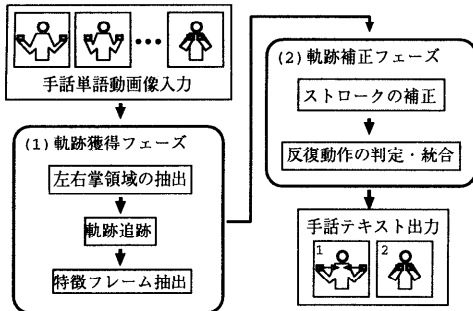


図3 システムの処理プロセス

Fig. 3 Processing flow of the system.

多大なコストがかかるといった短所がある。

本稿では、テキスト教材作成にかかるコストを軽減するために、ファジィ・スプライン関数近似による効率化を図る。このシステムでは、データ獲得が比較的容易な手話動作のビデオ映像を入力として、それを処理・加工し、手話テキストを生成する。

2. 手話学習用テキスト作成支援システム

本システムで用いる入力ビデオ動画像 \star は、 n 個の連続したフレーム列で構成される1つの手話単語とする。図2にその概念図を示すが、フレーム列の第 i 番目のフレームを $f_i (i = 0, 1, \dots, n - 1)$ として表す。また、システムの処理プロセスを図3に示す。処理プロセスは、(i) 手話動作の抽出を行う“軌跡獲得フェーズ”と、(ii) 手話テキストを生成する“軌跡補正フェーズ”の2つに大別することができる。

2.1 軌跡獲得フェーズ

軌跡獲得フェーズでは、両手の動作軌跡として掌領域の重心サンプリング点列 $p_r(i)$ (右掌領域)、 $p_l(i)$ (左掌領域) $\star\star$ をそれぞれ抽出し、特徴点および特徴フ

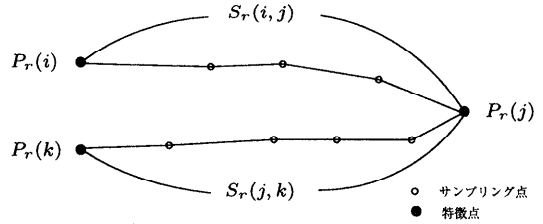


図4 動作軌跡の例

Fig. 4 An example of a hand moving locus.

レームを抽出する。なお、左右掌領域の重心サンプリング点は、各フレーム上にそれぞれただ1つしか存在しないのは自明である。

特徴点については、サンプリング点列のうち第0フレームおよび第 $n - 1$ フレーム上のサンプリング点と、軌跡が曲折した部分のサンプリング点を表すが、ここでは、Braultらの定義とその抽出アルゴリズム^{1), 2)}に従った。また、第 i フレーム上に存在する右掌領域の特徴点座標を $P_r(i)$ 、同様に、左掌領域の特徴点座標を $P_l(i)$ として表す。図4にサンプリング点および特徴点の例を示す。さらに、第0フレームから第 $n - 1$ フレームまでの特徴点総数を左右掌領域それぞれ σ_l 、 σ_r としたとき、以下の式(1)を満たす場合、右掌領域の特徴点座標 $P_r(i)$ が存在するフレームを特徴フレーム F_r とする。逆に、式(1)を満たさない場合は、 $P_l(i)$ が存在するフレームを特徴フレーム F_l とする。

$$\sigma_r \geq \sigma_l \tag{1}$$

動きをともなう手話単語については、そのほとんどが、左右いずれか片方の手のみが動作する場合と、左右の手がほぼ同期して、対称動作または同一動作をする場合の2種類に分けることができる^{3), 5)}。前者の場合、動作軌跡を形成する片方の掌領域のみに着目して、その特徴点から特徴フレームを抽出すればよいが、後者の場合、両手の動きが完全に一致して同期しているとは限らず、近傍フレームには、左右掌領域の特徴点がそれぞれ存在すると推測できるが、同一フレーム上に双方の特徴点が存在するとは限らない。したがって、ここでは、左右掌領域のうち、片方の特徴点に着目して特徴フレームを決定し、その手話動作を近似的に表現・特徴付けることとした。なお、特徴フレーム同定後は、特徴フレーム F_r 上に存在するサンプリング点を新たに特徴点 $P_r(i)$ および $P_l(i)$ として用いる。

\star 本稿では、手話工学研究会が製作し、研究用資料として公開している Signdex V.1 (SVHS) に記録されたカラー動画像データを入力動画像として用いた。なお、ここで用いた入力動画像においては、肌色領域は顔領域および手首から先の掌領域のみであり、背景領域は複雑ではない。

$\star\star$ 本研究では、第0フレームにおける各掌領域の重心サンプリング点についてはユーザが指定するものとした。第1フレーム以降については、システムがファジィ推論を用いて各 f_i 上の肌色領域を抽出し、各掌領域の重心サンプリング点を求め、自動追跡を実行する。

2.2 軌跡補正フェーズ

軌跡補正フェーズでは、軌跡獲得フェーズで得られた動作軌跡を補正し、手話テキストを生成・出力する。ここで、連続する2つの特徴フレーム F_i と F_{j^*} において、右掌領域について特徴点 $P_r(i)$ を始点、特徴点 $P_r(j)$ を終点としたとき、始点・終点間、すなわち、連続する特徴フレーム間に存在するサンプリング点列 $p_r(i), p_r(i+1), \dots, p_r(j)$ をストローク $S_r(i, j)$ として定義する。また、左掌領域についても同様に、ストローク $S_l(i, j)$ を定義する。ここで定義したストロークを最小単位として、次章で述べるように、ファジィ・スプラインを用いた軌跡の補正を実行する。

3. ファジィ・スプラインによる軌跡の補正

3.1 軌跡の補正

本システムでは、ストロークを手話動作の最小単位としてとらえ、軌跡の補正を実行する。まず、ストロークを構成するサンプリング点列を平面的な広がりを持った円錐底面として表現することによって「曖昧さ」を与え、これらすべてをファジィ制御点列としてファジィ・スプライン曲線（以下、FS 曲線）を求める。ここで、ストローク $S_r(i, j)$ のFS 曲線を $\tilde{S}_r(i, j)$ と表記する。 $S_l(i, j)$ のFS 曲線も同様に、 $\tilde{S}_l(i, j)$ と表す。

次に、サンプリング点列からいくつかを制御点として選択し、通常のBスプライン曲線により1本の補間曲線を求める。このとき求められたBスプライン曲線全体がFS 曲線内部に位置するように制御点を選択する。図5⁶⁾に示すように、FS 曲線は、円錐底面の集合として表現される。なお、本稿では、真野らの提案する手法⁴⁾を用いてFS 曲線を求め、ストローク単位で軌跡の補正を行う。

3.2 反復動作の判定・統合処理

手話の手振り動作には、反復動作を含む場合が多々見られる。一般の手話テキスト教材では、反復動作を明示的に呈示することによって簡明なテキスト表現を実現している（図1参照）。

ここでは、連続するストロークが反復動作であるか否かを判定し、反復動作である場合にはそれらを動作方向のみ異なる同一動作として、図1に示すような表示によって統合し、そうでない場合にはそれらを連結した表示を行う。反復動作の判定には、前章で求めたFS 曲線を用いる。図6に例示するように、2つの連

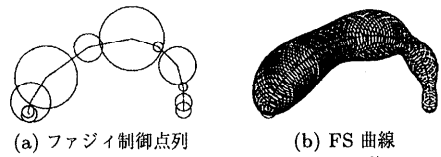


図5 ファジィ・スプライン補間の例⁶⁾
Fig. 5 An example of fuzzy spline interpolation.

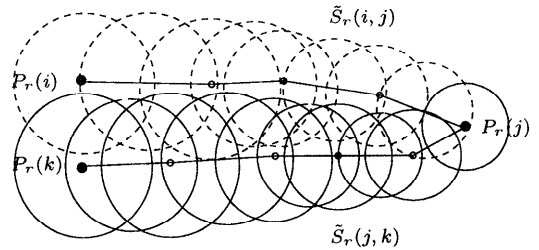


図6 連続するストロークのFS 曲線
Fig. 6 FS curve representation of successive strokes.

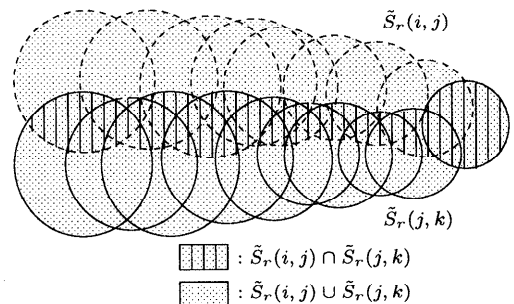


図7 FS 曲線の共有部分と和
Fig. 7 Union and intersection of FS curves.

続するストローク $S_r(i, j)$, $S_r(j, k)$ が存在するとき、このFS 曲線はそれぞれ、 $\tilde{S}_r(i, j)$, $\tilde{S}_r(j, k)$ と表される。このとき、 $\tilde{S}_r(i, j)$ の内包する領域と $\tilde{S}_r(j, k)$ の内包する領域の2つの領域の和および共通部分のそれぞれを式(2)、式(3)で表す（図7参照）。

$$\tilde{S}_r(i, j) \cup \tilde{S}_r(j, k) \quad (2)$$

$$\tilde{S}_r(i, j) \cap \tilde{S}_r(j, k) \quad (3)$$

$$\varepsilon_r = \frac{\tilde{S}_r(i, j) \cap \tilde{S}_r(j, k)}{\tilde{S}_r(i, j) \cup \tilde{S}_r(j, k)} \quad (4)$$

また、式(4)は、FS 曲線の和に対する共通部分の比率、すなわち共有率 ε_r を表す。 ε_r は、連続するFS 曲線の類似度を示しており、 ε_r が1.0に近いほど類似度が高いといえる。本システムでは、式(1)を満足し、かつ共有率 ε_r が閾値 (Θ) 以上のとき、連続するストロークを反復ストロークとして判定し、2つの連続するストローク $S_r(i, j)$, $S_r(j, k)$ のうち、 $S_r(j, k)$ を消去し、 $S_r(i, j)$ のみを手話テキスト上に表示すること

* $i < j$ の関係を有し、 F_i と F_j の間には特徴フレームが存在しないとき、この2つのフレームを連続する特徴フレームと呼ぶ。

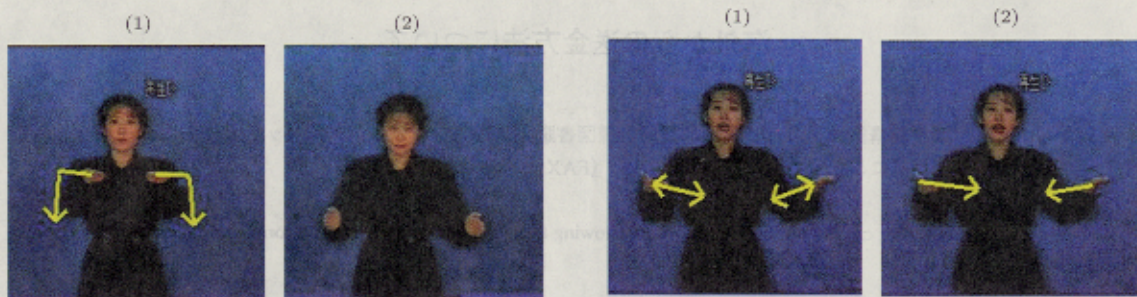


図8 出力例1(机)

Fig. 8 Output example 1 (desk).

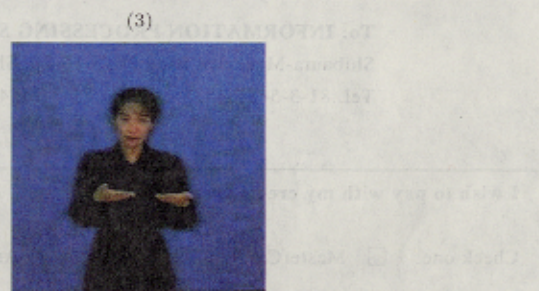


図9 出力例2(仕事)

Fig. 9 Output example 2 (job).

で統合を実現する。なお、本稿における実験では、 Θ を0.40とした。ただし、 ε_r によって反復ストロークの判定を行うのは式(1)に示す条件式を満足するときである。式(1)を満足しない場合については、左掌領域を対象として、共有率 ε_l を求め、同様の判定・統合処理を行う。

$$\varepsilon_r \geq \Theta (= 0.40) \quad (5)$$

4. 実験

本稿の提案手法によって出力した手話テキストの例を図8, 図9に示す。図8, 図9ともに、手話動作の終了状態を明示的に示すために、最終フレームを最後のコマに付加した。また、図9の第1コマと第2コマはさらなる反復・統合処理によって1コマに統合可能であるが、動作の反復回数が視覚的に分かるように、あえて3コマでこれを表現した。なお、図8, 図9ともに、動作順序は左から右の方向へ進む。

5. おわりに

本稿では、ファジィ・スプライン補間法を用いた手話学習用テキストの作成支援システムを提案した。本システムを用いることによって得られた手話テキストは、一般の手話テキスト教材に類似した簡潔かつ明瞭な表現が実現でき、従来の手話テキスト製作にかかる多大な人手とコストの削減が見込まれる。また、本システムのハードウェア構成は、民生用のビデオカメラとパソコンさえあれば実現でき、比較的安価に構成可能である。したがって、一般の手話教材では扱われないが、実生活において高い需要がある“ローカル手話”^{*}のテキスト教材を作成することも低予算で可能となる。

* 手話はその言語的性格からローカル性が高く、一部地域、局所的環境でのみ使用する方言的な手話(ローカル手話)が存在する。新語彙の発生速度が速いコンピュータなどの専門用語についても、定義・普及に時間がかかる共通手話の制定を待たずに、新語彙に対応した手話を独自に定義して使用する場合が少なくない。

今後の課題としては、まず、細やかな手・指の動きの抽出およびそのテキスト表現の実現、さらに、本システムによって得られたテキストの手話動作について、解説文を自動的に付加することなどがあげられる。

参考文献

- 1) Brault, J. and Plamondon, R: Segmenting handwritten signature at their perceptually important points, *IEEE trans. Pattern Analysis and Machine Intelligence*, Vol.15, No.9, pp.953-957 (1993).
- 2) Lin, C. and Chen, C.: A new approach to the verification of chinese signature with variant orientations and scales using relaxation and state-space search methods, *Pattern Recognition*, Vol.31, No.6, pp.665-674 (1998).
- 3) 篠田三郎, 全日本ろうあ連盟: 手話のハンドブック, 三省堂(1993).
- 4) 真野淳治, 江野脇宏, 中村剛士, 何立風, 伊藤英則: 毛筆書体における運筆状態規則を用いた平仮名毛筆文字生成について, *日本ファジィ学会誌*, Vol.11, No.1, pp.140-148 (1999).
- 5) 全日本ろうあ連盟: わたしたちの手話(1)~(4), (1966).
- 6) 佐賀聡人, 牧野宏美, 佐々木淳一: 手書き曲線モデルの一構成法: ファジィ・スプライン補間法, *電気情報通信学会論文誌*, Vol.J.77-D-II, No.8, pp.1610-1619 (1994).

(平成11年3月11日受付)

(平成11年6月3日採録)