

動作特徴ベクトルの導入による SignWriting 表記の手話単語の類似度定義の改良と評価

岡田 紳太郎 松本 忠博 加藤 三保子
 岐阜大学大学院 自然科学技術研究科 豊橋技術科学大学

1 はじめに

Sutton の SignWriting (以下 SW) [1] は日常的に手話を読み書きするための表記法であり, 複数の国でろう児教育などへの利用が試みられている. 我々は SW による手話文書編集システム JSPad[2] を開発し, その付加機能として SW を手話単語の入力形式とする手話-日本語辞書機能を実装してきた.

SW では手の形, 位置, 動き, 表情などを表す図的な文字 (International SignWriting Alphabet, ISWA) を 2 次元的に配置して手話単語を視覚的に表現する (図 1). その表記は人にとって分かりやすい反面, 自由度が高く, 表記の揺れが大きいため, 辞書検索のための単語の識別は容易ではない. そこで, SW 形式の手話単語の識別のために, 単語を構成する ISWA 記号の類似性と位置に基づく単語間類似度を定義して利用した [3, 4]. しかし, 手の動作を表す記号 (図 1 右の矢印記号など) は記号選択や配置の自由度が特に大きく, 十分対応できていなかった.

本研究では, 動作記号が表現する運動を複数の特徴量に分解し, 動作記号を動作の特徴ベクトルとして表現することで単語間類似度の定義の改良を試みた. 評価のため, SW で記述された日本手話 (JSL) とアメリカ手話 (ASL) の単語による単語識別実験を行った結果, JSL 単語については従来に比べて識別精度を大きく向上させることができた.



図 1 手話単語の SW による表記例

表 1 SW 記号とその SSS の例

	カテゴリ	グループ	基底記号	変種	塗り	回転
	01	03	001	01	02	03
	02	03	001	02	01	08
	04	04	001	01	01	01

2 SW による単語表記と表記揺れ

SW では ISWA と呼ばれる図的な記号 (文字) を 2 次元的に配置して単語を表記する. ISWA は Unicode の文字集合に含まれており, TrueType フォントも公開されている*1. 表 1 に示すように, 各記号は文字コードのほか, 記号の種類や向きなどを表す Sign-Symbol-Sequence (SSS) と呼ばれる数字列によっても表される. 例えば, 表 1 の第 1 行の手型記号は 01-03-001-01-02-03 と表される. 手型記号の場合, 塗りは手のひらの向き, 回転は指先の向きを表す.

ISWA2010 の基本記号は 600 種類以上用意されており, 手形記号の場合, 指先の向きを 45 度単位で, 手のひらの向きを 90 度単位で指定できる. SW では手話の 3 次元的な動きを 2 次元平面に射影して表すが, 類似した記号のうちどれを選択するか, 記号の向きや位置をどのように近似するか, 省略可能な記号があるときに省略するか, 視点 (front/top view) はどうするかなど, 書き手が恣意的に選択できる要素が多い. また, 現実の手話にも表現の揺れが多く存在する. これらの要因により, SW による手話の表記には大きな揺れが生ずる.

3 従来の単語間類似度の定義

ここでは, 従来の研究 [3, 4] で定義された単語間類似度について概略を述べる.

*1 https://github.com/Slevinski/signwriting_2010_fonts
http://www.signbank.org/SignWriting_Fonts.html

SW 表現における手話単語 w は、単語を構成する ISWA 記号の集合 $\{s_1, s_2, \dots, s_n\}$ であり、各記号 s_i は記号を識別する SSS とその単語内での相対位置 pos の組 (sss, pos) である。SSS は記号のカテゴリ C, グループ G, 基本記号番号 B, 変種 V, 塗り F, 回転 R の 6 項組 (C, G, B, V, F, R) である。

定義 1 (記号間類似度) 2 つの記号 s_1 と s_2 の SSS を元にした類似度から、次のように定義される距離ペナルティ $DP(s_1, s_2)$ を差し引いた値を s_1 と s_2 の記号間類似度 $\text{sim}(s_1, s_2)$ とする。

$$DP(s_1, s_2) = \begin{cases} 0 & (d \leq \theta) \\ a \cdot (d - \theta) & (d > \theta) \end{cases}$$

ここで d は記号間のユークリッド距離、 a は重み、 θ は閾値を表す。

定義 2 (単語間類似度) 単語 w_1 と w_2 の単語間類似度 $\text{Sim}(w_1, w_2)$ を次のように定義する。2 つの単語に含まれる記号の個数の差が上限 k を超える場合、 $\text{Sim}(w_1, w_2) = 0$ とする。それ以外の場合は、単語 w_1 の記号 $\{s_{11}, \dots, s_{1m}\}$ と単語 w_2 の記号 $\{s_{21}, \dots, s_{2n}\}$ の記号間類似度の合計が最大になる組み合わせ (図 2 では対角線上) での記号間類似度の平均値を、単語内記号数の差により調整した値を $\text{Sim}(w_1, w_2)$ の値とする。

4 単語間類似度定義の改良

前節で述べた単語間類似度を用いて単語の識別実験を行ったところ、書き手による表記ゆれがうまく吸収できないケースが見られた。その原因の一つとして、手型記号に比べ、手の移動や接触などを表す動作記号

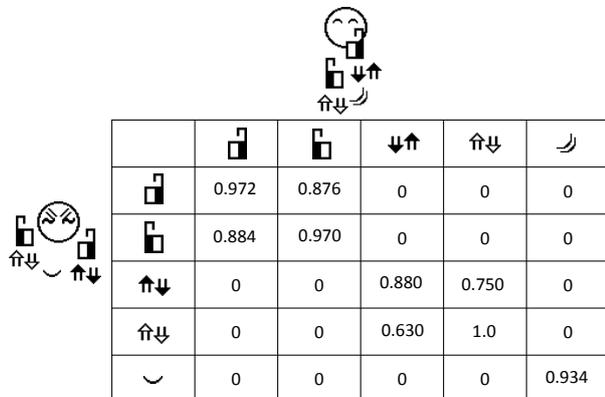


図 2 単語を構成する記号間の類似度

表 2 動作記号が表す動作の特徴量と重み

No.	動作特徴名	動作内容	重み
1	XposR	右方向への移動 (右手)	800
2	XnegR	左方向への移動 (〃)	800
3	YposR	上方への移動 (〃)	800
4	YnegR	下方への移動 (〃)	800
5	ZposR	前方への移動 (〃)	800
6	ZnegR	後方への移動 (〃)	800
7	XrepR	左右の反復移動 (〃)	50
8	YrepR	上下の反復移動 (〃)	50
9	ZrepR	前後の反復移動 (〃)	50
10	MultiDirR	複数方向への移動 (〃)	10
11	CircleR	円運動 (〃)	2000
12	CurveR	曲線運動 (〃)	50
13	CrossR	交差 (〃)	500
14	BendR	移動方向の途中変化 (〃)	500
15	CircleHitsWallR	壁に垂直な円運動 (〃)	500
16	CircleAxisR	回転軸 (〃)	500
17	CircleOtherR	その他の円運動 (〃)	500
18	XposL	右方向への移動 (左手)	800
19	XnegL	左方向への移動 (〃)	800
:	:	:	:
34	CircleOtherL	その他の円運動 (〃)	500
35	Contact	接触	10
36	Finger	指の動き	10
37	Dynamics	動作の緩急強弱など	5

は、書き手による表記のゆれが大きく、類似度を十分に表せていないことがあげられる。そこで本研究では、動作以外の記号については従来の方法で類似度を求め、動作記号については動きの向きや回数、軌跡などの情報を単語の特徴量としてベクトル化してベクトル間の類似度を記号間の類似度とする。

設定した動作特徴量とその重みを表 2 に示す。ISWA の動作記号には、左右の手の区別があるものと、ないものがある。動作特徴情報 1~17 は右手、18~34 は左手の動きに対するものである。35~37 の特徴情報には左右の手の区別はない。

定義 3 (単語の動作特徴ベクトルとその類似度) 単語 w を構成する記号 s_i の動作特徴ベクトルを v_i とする。このとき、単語 w の動作特徴ベクトル $V(w)$ を次のように定義する。

$$V(w) = \sum_i v_i \circ W$$

ここで、 W は表 2 の重みのベクトル、 \circ はベクトルの要素ごとの積を表す。なお、 s_i が動作記号でないとき、 $v_i = \mathbf{0}$ とする。動作特徴ベクトル間の類似度 Sim_{vec}

としてはコサイン類似度を用いる。

$$\text{Sim}_{\text{vec}}(w_1, w_2) = \cos(V(w_1), V(w_2))$$

定義 4 (単語間類似度) 動作記号については Sim_{vec} を、動作以外の記号については従来の記号の形 (SSS) と相対位置による類似度 Sim_{shp} を用いて、単語間類似度 Sim を次のように定義する。

$$\text{Sim}(w_1, w_2) = c \cdot \text{Sim}_{\text{vec}}(w_1, w_2) + (1-c) \cdot \text{Sim}_{\text{shp}}(w_1, w_2)$$

後述の評価実験では $c = 0.1$ とした。

そのほかの改良として、類似度算出の前に冗長な記号を削除する前処理の過程を加えた。これは、SW では手の移動の前後で手型が同じなら、移動後の手型の記述を省略するなど、(人が) その単語と認識できる範囲で記述を省くことができるが、省くかどうかの判断は書き手の主観に依存しており、それが表記の揺れの原因になっていたためである。

5 評価実験

5.1 JSL の単語による評価

新たに定義した単語間類似度の評価のため実験を行った。これまでに、累計 33 名の被験者に手話単語のイラスト 20 語、動画 10 語 (の一部) を見て JSPad を使用して SW で書き取ってもらい、計 30 種類 406 語の表記データを収集している。明らかな表記上の誤りを含むものには修正を加えて利用した。修正が困難なものは対象から除外した。それぞれを辞書検索し、辞書中の当該単語 (正解) が類似度順で何位以内に入っているかを調べた。辞書に登録されている単語数は約 2000 である。なお、前節で述べた単語間類似度定義のパラメータの調整などにもこれらのデータを使用しているため、評価実験はクロードテストとなる。

図 3 に被験者が入力した単語の例を、図 4 に単語の検索結果の表示例を示す。この例では被験者が入力した単語〈忙しい〉が、検索結果の第 1 位になっている。被験者の単語では、手の移動を表す動作記号 (矢印記号) が左右 2 つずつ、合計 4 つ使われているのに対し、辞書登録された〈忙しい〉では、上・下・上または下・上・下の連続動作を表す動作記号が左右 1 つずつ、合計 2 つ使われている。従来の単語類似度では記号数が異なるため、記号間の対応がうまく取れず、類似度を下げる要因となっていたが、動作記号を動作特徴ベクト



図 3 被験者が入力した手話単語の例 (〈忙しい〉)

表 3 日本手話の単語による実験結果

正解率 (1 位)		正解率 (5 位以内)	
改良前	改良後	改良前	改良後
54.8%	74.4%	72.0%	92.1%

ルとして表すことで、類似度の低下を防ぐができている。

被験者が選択した手型記号と辞書登録単語の手型記号では見た目が異なるが、これは視点の違いによるものである (被験者は上から見た手型、辞書単語は話者から見た手型)。このような視点による記号の見た目の違いについては、記号間類似度を算出する際に考慮するようにしている。

30 種類の単語に対する平均正解率を表 3 に示す。「正解率 (1 位)」はテストデータと辞書中の正解単語の類似度が最も高かったとき正解とした場合の正解率、「正解率 (5 位以内)」は類似度順で 5 位以内に入ったとき正解とした場合の正解率を表す。従来の単語間類似度を用いた場合に比べて精度は 19.6% (1 位) および 20.1% (5 位以内) 向上し、平均すると 9 割以上が 5 位以内に入った。

5.2 ASL の単語による評価

本研究では前節の実験で収集した日本手話の単語の表記サンプルを元にして、単語間類似度や記号間類似度の設定・調整を行った。オープンテストによる評価を行うため、SW コミュニティの Web サイトから ASL の単語辞書^{*2}と、SW で書かれた ASL 版の童話 “Goldilocks and the Three Bears”^{*3} (3びきのくま) を取得し、この童話に含まれる 155 個の単語を使用して評価実験を行った。

ASL の辞書登録語数は約 10,500 語であり、我々が構築した JSL 版の辞書の約 5 倍のサイズがある。ただし、ASL 版辞書は多くのユーザが辞書登録を行っており、同じ単語が重複して登録されている。実験では正解単語のうちテストデータとの類似度が最も高かった

^{*2} <http://signbank.org/signpuddle/>

^{*3} <http://www.signwriting.org/library/children/>



図 4 動作特徴ベクトルを用いた単語間類似度による単語検索結果の例

表 4 ASL 単語による評価実験結果（表情記号削除前）

正解率（1 位）	正解率（5 位以内）
48.3%	69.9%

表 5 ASL 単語による評価実験結果（表情記号削除後）

正解率（1 位）	正解率（5 位以内）
47.7%	72.2%

ものの順位を評価結果とした。155 単語の平均正解率を表 4 に示す。

JSL 単語による評価実験の結果に比べて、正解率は大幅に下がった。辞書登録語数の大きな差、JSL と ASL で使用される手型の違いのほか、顔の表情を表す記号の影響が考えられる。童話から単語を収集したためか、テストデータには単語の弁別には関係しない顔の表情を表す記号が多く含まれていた。そこで、顔の表情を表す記号を単語から取り除く前処理を加えて、再度評価を行ったところ、表 5 に示すように 1 位を正解とした場合の正解率は 0.6% 下がったが、5 位以内を正解とした正解率では 2.3% 向上する結果となった。

6 おわりに

本研究では、SW 表記の手話単語に含まれる記号のうち、動作記号をその動作の特徴ベクトルとして表すことで単語間類似度の定義を改良し、単語の識別精度向上を図った。評価実験の結果、JSL 単語を対象としたクローズドテストでは類似度 5 位以内を正解とした平

均正解率が従来の手法に比べて 20.1% 向上し、90% を超える結果が得られた。一方で、ASL 単語によるオープンテストでは 80% に届いておらず、さらなる改良が必要と考えている。本研究の手法とは別に、ニューラルネットワークなどを用いた機械学習による単語識別手法についても現在検討を始めている。

謝辞

本研究は JSPS 科研費 JP18K11430 の助成を受けたものです。

参考文献

- [1] V. Sutton, Lessons in SignWriting, The Deaf Action Committee for SignWriting, 2002.
- [2] T. Matsumoto, M. Kato and T. Ikeda, JSPad—a sign language writing tool using SignWriting, Proceedings of the 3rd International Universal Communication Symposium, pp.363-367, 2009.
- [3] 高瀬友宏, 松本忠博, 加藤三保子, SignWriting を利用した手話文書エディタ JSPad における手話-日本語辞書について, 言語処理学会第 18 回年次大会発表論文集, pp.1220-1223, 2012.
- [4] 杉山真也, 松本忠博, 加藤三保子, 手話文字 SignWriting の手話-日本語辞書への応用, 電子情報通信学会 2013 年総合大会講演論文集 A-19-13, p.274, 2013.