

# SignWriting 表記の手話単語検索のための単語のベクトル表現

松本 忠博  
岐阜大学 工学部  
tad@gifu-u.ac.jp

東 香織  
岐阜大学 工学部

石橋 陽典  
岐阜大学大学院  
自然科学技術研究科

加藤 三保子  
豊橋技術科学大学 総合教育院

## 1 はじめに

手話には標準的な文字表現がないことから、主に手話の言語学的、工学的研究を目的として表記法がいくつか提案されてきた。Sutton の SignWriting[1] (以下, SW と略す) は主要な表記法の中で唯一日常生活の中で手話を読み書きする書記体系として提案され、かつ、その文字が Unicode 文字集合にも含まれている<sup>1)</sup>。これまでに複数の国で、ろう児教育の場などでの利用が試みられてきた。SW を使った手話文書エディタも複数作られており、我々も JSPad[2] と呼ぶシステムを開発し、公開している。

手話単語の構成要素(音素)は、手の形、手の位置、手の動き、および、顔の表情などの非手指要素と言われている。SW ではそれらを表す図的な記号を2次元的に配置して単語を表記する(図1)。この図的な表現は人にとって直観的に分かりやすい反面、自由度が高く、書き手による表記の揺れが大きい。そのため、機械的な単語の同定が単純には行えず、文書処理の基本操作である単語の検索も簡単ではない。そこで JSPad では、単語を構成する記号間の類似度の定義と、それをもとにした単語間類似度の算出アルゴリズムを考案し、SW を入力形式とする手話-日本語辞書機能を JSPad に実装した[3]。さらにその改良として、とくに表記の揺れが大きい手の動きの表記について、動きの特徴量を要素とする動作特徴ベクトルを定義して、単語識別精度の向上を図った[4]。

本研究では SW 表記の手話単語を辞書検索することを主な目的として、動作だけでなく、手の形や位置などの要素についてもベクトル化する方法を検

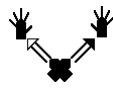


単語			
意味	晴れ, 明るい, 陽気	満員, 充滿, 満杯	見る, 見つめる, 注目する

図1 SignWriting による単語の表記例

討している。単語の辞書検索を、手書き文字認識のように多クラス分類問題として解くには、書き取りデータのサンプル数が少ないこと、そして、辞書に単語が追加された際の対応や、他の国・地域の手話単語辞書への適用がしやすいことなどから、ここでは単語間の類似度を用いて、検索対象の単語との類似度が高い辞書内の単語をユーザに提示する方式をとる。

## 2 SW による手話表記と表記揺れ

SW では図2に示すような図的な記号(International SignWriting Alphabet, ISWA<sup>2)</sup>)を2次元的に配置して手話単語を記述する。国際発音記号のように様々な国の手話に対応できるように、ISWA2010 では7つのカテゴリ(手形、動作、強弱、頭・顔、体など)、30のグループ、652種類の基本記号(base symbol)が用意されている。さらに、手形記号と動作記号には2つの修飾要素 fill(手形記号では手のひらの向きを表す6種類)と rotation(手形記号では指先の向きと左右の手の違いを表す16種類)を指定することができるため、1つの基本記号に対して最大で  $6 \times 16 = 96$  通りの記号が存在することになる。

音声言語では音素(子音・母音)は逐次的に表出され、文字も1次元的に並べられる。手話言語の音素(手の形、向き、位置、動き、顔の表情など)は

1) Sutton SignWriting (Unicode block): <https://unicode.org/charts/PDF/U1D800.pdf>

2) ISWA 2010: <https://movementwriting.org/symbolbank/index.html#ISWA2010>

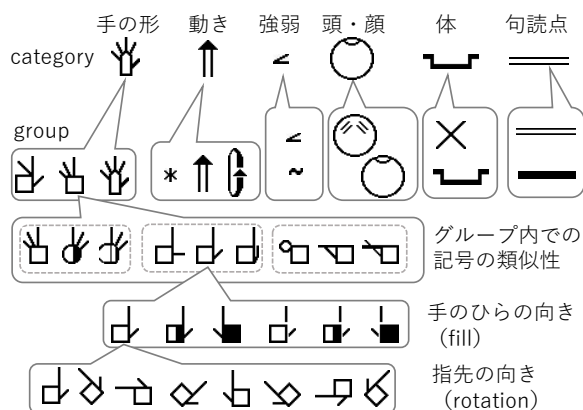








図2 SignWriting で使用される文字 ISWA

3次元空間上で同時的に表出され、SWではこれを2次元平面上に射影した形で表す。その際、類似した記号のうち（単語の弁別に影響しない範囲で）どれを選択するか（例：と）、指先や手のひらの向き（角度）がどの記号とも一致しない場合にどの角度で近似的に表すか、省略可能な記号があるときにその記号を省略するかしないか、視点をどう選ぶか（例：話者視点と、上からの視点）など、SWでは書き手が恣意的に選択できる要素が多い。同じ動きを表す複数の方法が存在する場合もある（例えば、右手を2度上に動かす動作は、記号で表すことも、記号を2つ並べて表すこともできる）。また、内部表現では、記号の位置はピクセル単位で指定できるため、ひとりの書き手が同じ単語を書いても、一般に記号の配置はまったく同じにはならない。さらに、実際の手話にも話者によって手の形や動きが微妙に異なる表現の揺れがあり、それが表記揺れにつながることもある。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ため、文書内の単語の検索のような基本的な操作も音声言語の文字と同じようにはできない。

### 3 SW形式の手話単語の検索

#### 3.1 関連研究

SWで書かれた手話単語の検索手法として、Costaら[5]は、2つの単語の照合時に記号の位置のずれを許容する単語類似関係（sign similarity relation）を定義し、それをを用いた照合手続きを提案している。また、Aertsら[6]は、検索対象となる単語の手の形と、その手が接触する体のゾーン、及び、接触の仕方をユーザに指定させることで、単語を検索する方法を提案している。いずれも記号の類似性について

は考慮しておらず、記号の違いは許容されない。



SWのWebサイト<sup>3)</sup>では、SWによるオンライン手話辞書 SignPuddle や、SWを使って手話文書を作成するソフトウェアなどがいくつか公開されている。これらのソフトウェアにも辞書内の単語を検索する機能が用意されているが、手話からその訳（英語等）を検索する方法は、単語を構成している記号を指定するものであり、記号間の類似性や記号の位置情報は考慮していない。

#### 3.2 JSPad における従来手法

JSPadでは、単語を構成する記号や配置が完全に一致しなくても、ユーザが記述した単語と類似した単語を辞書内から見つけ出し、候補をユーザに提示する機能を実装してきた。ここでは単語間の類似度を求めるための2種類の従来手法の概略を述べる。

##### 手法1

まず、比較対象となる2つの単語を構成する記号間の対応関係を推定する。そのために、すべての記号の対応関係のうち、対応する記号間の類似度の平均が最も大きくなるものを探索し、その記号間類似度の平均値を単語間の類似度とする。計算量を抑えるため、類似度0の記号の組み合わせが見つかった場合は、その対応関係は候補から除外する。また、記号数の差が一定以上の場合、単語間の類似度を0とする。

記号間の類似度は、記号が表す手などの形と、単語内での記号の相対位置をもとに算出する。ISWAの各記号は、SSS（sign symbol sequence）と呼ばれる6つの値の組によって識別される。これは、各記号のカテゴリ、グループ、基本記号番号、バリエーション、塗り（fill）、回転（rotation）を表すもので、例えば、手形記号は01-09-016-01-03-01、動作記号は02-03-001-02-01-01と表される。形の類似性は、このSSSをもとに算出する。また、同じグループに属す記号でも、記号が表す手の形が類似したものとそうでないものが混在する（図2）ため、独自にサブグループを設けて類似度の調整に用いている。

位置情報については、対応する記号の単語内の相対位置間のユークリッド距離により類似度を減じることで、類似度に反映させる。

3) <https://www.signwriting.org/>

## 手法 2

手法 1 の評価実験の結果、ISWA 記号のうち、とくに手の動きを表す記号の表記揺れが大きいことが分かった。そこで手の動きについては、単語に含まれる動作の成分（右手の垂直運動、左手の円運動、反復の有無、接触の有無など）を特徴量とする 39 次元の動作特徴ベクトルで表す。そして、2 つの単語の動作特徴ベクトルのコサイン類似度と、手法 1 による動作以外の要素から得られる単語間類似度の加重平均を単語間の類似度とした [4]。

## 4 手話単語のベクトル化

本研究では手の動き以外の、手形、頭・顔、体（肩、腕、腰など）の記号についてもベクトル表現を定め、それらをもとに単語ベクトルを生成する。ベクトル化することで、記号の対応関係を推定するための探索を行う必要がなくなる。

### 4.1 記号とその位置のベクトル化

比較対象となる 2 つの単語を構成する記号間の類似度は、記号そのものの類似度と単語内での記号の相対的な位置によって決まる。まず、記号の類似性により記号のベクトル表現を取得し、位置情報をそのベクトルにエンコードする。

#### 記号のベクトル化

手法 1 で取得した記号間類似度の算出結果を訓練データとして用いて、記号が表す手の形などの類似性に基づく記号ベクトルを取得する。学習のためのニューラルネットワークのモデルを図 3 に示す。入力は比較対象となる 2 つの記号の ID であり、図中の 2 つの埋め込み層の重みは共有されている。2 つの埋め込み層の出力ベクトルのコサイン類似度が、手法 1 で得られた結果に近づくように学習を行う。学習の結果得られた埋め込み層の重みを記号ベクトルとする。

すべての記号の組み合わせを学習対象にすると訓練データのサイズが大きくなり、訓練時間も長くなるため、学習はカテゴリごとに行った。カテゴリごとの記号数（種類）と埋め込みベクトルの次元数を表 1 に示す。

記号ベクトルのサイズはカテゴリによらず 185 次元の固定長とする。埋め込みベクトルデータを格納する位置をカテゴリごとに設定し、空き領域には 0 を埋める。カテゴリの異なる記号間のコサイン類似

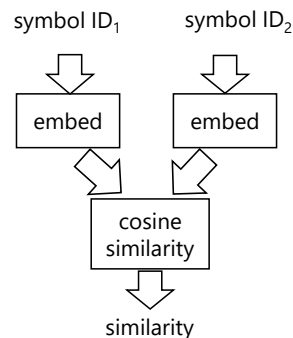


図 3 記号間類似度の学習モデル。

表 1 カテゴリごとの記号数（種類）と次元数

カテゴリ	手形	頭・顔	体	動作+強弱	計
記号数	24464	1059	762	11157	37442
次元数	100	26	20	39	185

度は 0 になる。

#### 位置情報のエンコード

手法 1 では、同一の記号でも単語内での相対位置の違いにより類似度が変化する。これをベクトルで表現するために、記号の位置情報をベクトルにエンコードする。単語内での相対位置をもとに正規分布の平均  $\mu$  を設定し、カテゴリごとの重みをかけて記号ベクトルに加える。位置情報には  $x$  成分と  $y$  成分があるため、例えば手形記号（100 次元）では、 $x$  成分を表す正規分布を先頭の 50 次元に、 $y$  成分を表す正規分布を残りの 50 次元に加える。

### 4.2 単語のベクトル化

単語に含まれる記号のベクトルをカテゴリごとに重みをつけて統合し、固定長の単語ベクトルを取得する。動作記号（と強弱記号）に対するベクトルについては、手法 2 の動作特徴ベクトルを流用する。統合の仕方としては、記号ベクトルの和をとる（sum pooling）、あるいは、各要素の最大値をとる（max pooling）、平均をとる（average pooling）などが考えられる。次節の実験では、単語に含まれる記号のベクトルの、カテゴリごとの平均を単語ベクトルとした。したがって、単語ベクトルのサイズは、記号ベクトルと同じ 185 次元である。

## 5 評価実験

単語ベクトルによる単語間類似度評価のために辞書検索実験を行った。

これまでに、累計 33 名の被験者に日本手話の単語のイラスト 20 語、動画 10 語（の一部）を見て 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。明かな表記上の誤りには修正を加え、修正が困難なものは除外した。表記サンプルのそれぞれを辞書検索し、辞書中の目的の単語（正解）が類似度順で何位になるか調べた。「1 位」、「5 位以内」、「10 位以内」を正解とした場合の正解率のマイクロ平均と、406 語の検索にかかった時間を表 2 に示す。

表中のベクトル欄「なし」は動作特徴ベクトルも単語ベクトルも用いない従来の手法（手法 1）、「動作」は動作記号のみベクトル化する手法（手法 2）である。

「記号」は手法 2 において、記号の形状に基づく類似度算出部分に記号ベクトルを使用した方法であり、記号の位置情報はベクトルにエンコードしていない。

「単語 1」は記号ベクトルに位置情報をエンコードして、（動作を除く）単語ベクトルを生成し、その単語ベクトルによる類似度と、動作特徴ベクトルによる類似度の加重平均（9:1）を単語類似度としたものである。

「単語 2」は動作特長ベクトルを単語ベクトルに組み入れて 1 つのベクトルにして類似度を求めたものである。

ベクトル間の類似度にはコサイン類似度を使用した。辞書の登録語数は 2048 である。

実験の結果、今回の単語ベクトル表現では、実行時間を短縮することはできたものの、正解率はいずれの場合も下がる結果となった。実行時間が短くなった理由としては、記号の対応関係の探索が不要になったことと、辞書単語のベクトルを予め求めて辞書に格納しておけるようになったことがあげられる。

なお、記号ベクトルを用いない従来手法において、記号間類似度や単語間類似度を定める際、どのような表記揺れがあるかを調査するためにこれらの表記サンプルを参照しているため、この実験はクローズドテストとも言える。

## 6 おわりに

SW で表記された単語を辞書や文書中から検索することを目的として現在検討している手話単語のベクトル表現について述べた。手の動きの要素については動作特徴量をもとにベクトル化し、手の形や頭・顔、体（肩、腕、腰）などを表す記号について

表 2 日本手話単語による評価実験の結果

ベクトル	正解率 (%)			時間 (秒)
	1 位	5 位以内	10 位以内	
なし	58.37	78.08	80.48	4.7
動作	<b>74.14</b>	<b>91.87</b>	<b>93.84</b>	5.0
記号	68.72	88.67	<b>93.84</b>	18.7
単語 1	53.69	79.96	84.98	3.2
単語 2	53.20	74.63	79.80	<b>3.0</b>

は、従来手法で得られた記号間類似度をもとに埋め込みベクトルを取得した。記号の位置情報は、正規分布を用いて記号ベクトルにエンコードしている。実験の結果、検索にかかる時間は短くなったが、今回の単語のベクトル化方法では正解率は低下した。

サンプルごとの結果の詳細な分析とそれに基づく手法の改善は今後の課題である。オープンテストとなる別のデータセットによる評価実験についても今後行う予定である。

## 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] Valerie Sutton. *Textbook and Workbook, 3rd ed.* The deaf action committee for SignWriting, 2002. <https://www.signwriting.org/lessons/books/>.
- [2] Tadahiro Matsumoto, Mihoko Kato, and Takashi Ikeda. JSPad—a sign language writing tool using SignWriting. In *Proceedings of the 3rd International Universal Communication Symposium (IUCS2009)*, 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] 岡田紳太郎, 松本忠博, 加藤三保子. 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. 言語処理学会第 25 回年次大会発表論文集, pp. 1141–1144, 2019.
- [5] Antônio Carlos da Rocha Costa, Graçaliz Pereira Dimuro, and Juliano Baldez de Freitas. A sign matching technique to support searches in sign language texts. In *Workshop on the Representation and Processing of Sign Languages, LREC2004, Lisbon*, pp. 32–36, 2004.
- [6] Steven Aerts, Bart Braem, Katrien Van Mulders, and Kristof De Weerd. Searching SignWriting signs. In *Workshop on the Representation and Processing of Sign Languages, LREC 2004, Lisbon*, pp. 79–81, 2004.