

意味的共起関係を用いた手話認識方式

竹内 勝、 櫻井彰人

RWCP¹⁾新機能日立-1²⁾研究室

手話は、格標識が少ない、時制や様相を表す語が少ない、また、語順が自由であるといった特徴を有する、統語的な制約が弱い言語である。このため、手話認識を行なうには、意味に関する情報を統合することが必要となる。本稿は、意味情報を利用した手話認識方式に関する、最初の実験結果の報告である。この実験は手話認識における意味情報の利用の有効性を確認することを目的とする。実験の結果、意味共起関係を利用した手話認識は有望であるとの感触を得た。しかし、実用規模の意味共起高階グラフの手作業での作成は非常に困難であるとの知見も同時に得た。

Sign Language Recognition System Using Semantic Co-occurrence Relations

Masaru Takeuchi and Akito Sakurai

Novel Functions Hitachi Laboratory 1, RWCP

Effectiveness of using semantic information for recognizing Japanese sign language(JSL) was confirmed experimentally. JSL has few case markers, few words that represent voice, tense, and aspect, and relatively free word order in a sentence; that is, it is syntactically for less restrictive than oral languages. Thus, integration of semantic information is necessary for machine recognition of JSL. The semantic co-occurrence relation was used in the experiments as semantic clues. Necessity of full automatic acquisition of the semantic co-occurrence relations was realized.

1. はじめに

手話には、格マーカが存在せず、時制や様相を表す語が少ない。また、「主語」+「目的語」+「述語」が基本的な語順であるが、会話ではこの語順が頻繁に無視される。手話は、統語的な制約が弱い言語である¹⁾。このため、手話認識を行なうには、意味に関する情報を統合することが必要となる。

本稿は、意味情報を利用した手話認識方式に関する、最初の実験結果の報告であり、手話認識における意味情報の利用の有効性を確認することを目的とする。

2. 手話認識問題

手話認識システムは、データグローブからの入力される、手の位置・方向、指の関節角に関する3次元の波形データから手話単語を連続的に認識し、その結果をもとに、日本語文を生成し、ディスプレイ上

1) RWCP: Real World Computing Partnership (新情報処理開発機構)

2) 株式会社 日立製作所 基礎研究所内

に表示するシステムである⁽²⁾。

手話単語認識装置は単語束と云われる形式のデータを出力する。単語束は、切片とよぶ4項組、(距離, 単語ラベル, 始点, 終点)の集合である。ここで、距離は、提示された手話動作と、手話単語認識装置で用いる各単語の標準パターンとの差を表し、切片sに対して、 $D_s(s)$ で表す。始点、終点は認識装置により検出された手話動作の開始時刻と終了時刻である。例えば、手話単語認識装置は「触れると痛みます。」という手話を入力した場合、

$$\{ (1.249, "だけ", 28, 67), (1.364, "腫れる (膨れる)", 1, 43), (1.448, "触れる", 1, 33), (1.464, "腕", 4, 43), (1.659, "悪い", 37, 65), (1.673, "痛い (痛み)", 74, 116), \dots \} \quad (1)$$

を出力する。図1に、この単語束の概念を示す。「**...*」及び「++...+」の部分が切片である。

発話: 5 データ: 1 文: 触れると痛みます。

大きさ: 72 手話単語系列: (触れる だけ 痛い (痛み))

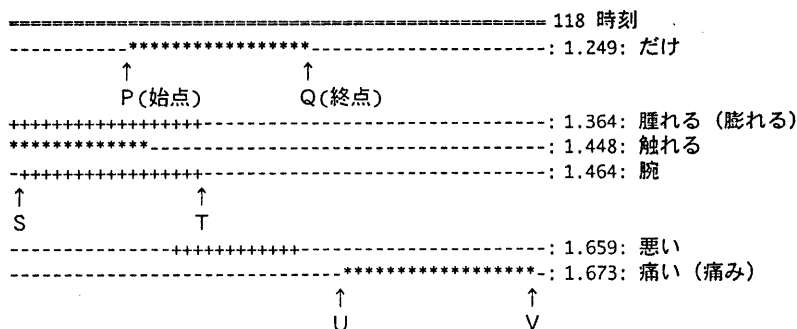


図1: 手話単語束の説明図

まず重複率と被覆率を定義する。

重複率 $0_c(s_1, s_2)$ とは、2つの切片の重なり具合を表す。図1における2つの切片、 $S_1=(1.249, "だけ", 28, 67)$ と $S_2=(1.464, "腕", 4, 43)$ の重複率は $0_c(s_1, s_2)=PT/SQ=0.238$ であり、重なりのない2つの切片、 $s_1=(1.249, "だけ", 28, 67)$ と $s_3=(1.673, "痛い (痛み)", 74, 116)$ の重複率は $0_c(s_1, s_3)=0$ である。

被覆率 $C_c(S)$ とは、系列を構成する複数の切片が全区間を覆う割合を表す。図1の「**...*」で示す3つの切片で構成される系列Sの被覆率は $C_c(S)=(SQ+UV)/118=0.890$ である。

我々は手話認識問題を、単語束から最適な単語系列を生成する最適化問題であると捉え、評価関数、

$$f(S) = \alpha \sum_{s_1 \neq s_2, s_1, s_2 \in S} 0_c(s_1, s_2) / \{ \#(S) * (\#(S) - 1) / 2 \} - \beta C_c(S) + \gamma \sum_{s \in S} D_s(s) / \#(S) - \delta h(S) \quad (2)$$

の単語束Lにおける最適化、

$$\min_{S \subset L} f(S) - \delta h(S) \quad (3)$$

を行なう。ここで、 $h(S)$ は、系列において連続する切片の単語の意味的共起頻度(主観確率)の平均値を与えるものとする。また、 α 、 β 、 γ 、 δ は定数であり、 $\#(S)$ は集合の濃度を表す。

3. 認識アルゴリズム

意味情報を利用することの有効性を検証するため、簡単な系列生成アルゴリズムを作成した。

このアルゴリズムは、重複率、被覆率に関する、(2)式の第1項、第2項を制約条件とした最適化問題、

$$\begin{aligned} \min \quad & \gamma d(S) - \delta h_2(S) - h_3(S), \quad d(S) = \sum_{s \in S} D_s(s) / \#(S) \\ \text{s.t.} \quad & \\ & \text{where } 0_c(s_1, s_2) < \alpha' \text{ for } s_1 \neq s_2, s_1, s_2 \in S \text{ and } C_c(S) > \beta' \end{aligned} \quad (4)$$

において、制約条件を満たす系列を生成するアルゴリズムである。ここで、 α' 、 β' 、 γ 、 δ 、 ε は定数である。

本稿では、意味情報としては単語の共起頻度を用いる。また、単語の共起頻度は重み付きの2階、3階の高階グラフで表す。 h_2 、 h_3 の添え字はグラフの階数を表す。

以下、アルゴリズムを示す。

0. 制約条件を満たす系列の集合 $A = \phi$ とし、単語束 L を各切片 s の距離によりソートする。
1. 系列を要素とする集合 $B = \phi$ とする。(Bをバッファとよぶ。)
2. L の先頭の要素、つまり、距離の最も小さい切片を s_0 とし、 $L = L \setminus \{s_0\}$ とする。
3. 重複率に関する制約条件に反する系列を生成する。つまり、 $0_c(s_0, s) \geq \alpha'$ for $s \in B \in B$ を満たす切片 s を全ての系列 b から取り除いた集合を B_0 とする。
4. B_0 全ての系列 b を $b \cup \{s_0\}$ に置き換え、これを B_1 とする。
5. 被覆率に関する制約条件を満たす系列を生成する。つまり、 $C_c(b) > \beta'$ for $b \in B_1$ を満たす b の集合を B_2 とする。 $A = A \cup B_2$, $B_1 = B_1 \setminus B_2$ とする。
6. $B = B \cup B_1$ とする。
7. 終了条件を満たすまで、2へ。
8. 制約条件を満たす系列の集合 A を出力する。

4. 病院での会話に関する実験

本実験では、病院での会話を想定し、表1に示す32例文を用いる。これらは57単語で作成されている。

4.1 意味共起高階グラフの作成

これらの単語の間の共起関係を表す高階グラフの作成方法としては、コーパスを用いて自動的に作成する方法が考えられる。日経コーパス92年、1年間の例文中の57単語の出現回数は、および、動詞を中心とした共起関係を抽出した結果、生起頻度、共起頻度が小さいため、新聞記事を用いて病院での会話に関する意味共起情報を抽出することが困難であることが判明した。

このため、手作業で2項辺、3項高階辺に、以下の3種類の操作で主観確率を割り振る作業を行なった。

1. 辺を提示し、明らかに共起しないと思われる辺を削除する。
2. 辺を提示し、主観的な共起確率を入力する。
3. 各辺を2組提示し、その優劣を数値として入力し、共起確率を修正する。

その結果、510個の2項辺、1140個の3項高階辺からなる意味共起関係を表す高階グラフを作成した。表2に高階グラフの一部を示す。左側から単語のインデックスによる辺の表示、主観評価値、評価回数を表し、「;;;」以降はコメントである。

4.2 評価実験

表1の32例文を手話通訳士に6回発話してもらい、その結果得られる単語束を用いて、重複率、被覆率に関して制約条件を満たす系列を生成した。このとき、表1に示す手話単語系列が出力された順位を記録した。なお、生成した系列の数が1000を越えた時点で、系列生成を打ち切った。

次に、各発話について、評価値 $d(S)$ 、 $h_2(S)$ 、 $h_3(S)$ を計算し、平均0、分散1に正規化を行い、
 $\gamma d(S)$ 、 $\gamma d(S) - \delta h_2(S)$ 、 $\gamma d(S) - \delta h_2(S) - \varepsilon h_3(S)$ (5)
を計算し、表1に示す正しい手話単語系列の順位を求めた。

日本語文	手話単語系列
痰が出ます。	痰/ある
熱が有って酷い風邪です。	熱/風邪(咳)/重い
下痢しています。	下痢/PT1(私)
風邪を引きました。	風邪(咳)/PT1(私)/変わる
触れると痛みます。	触れる/だけ/痛い(痛み)
痛みが足に走ります。	痛い(痛み)/足/感じ(感じる)
時々痛みます。	時々/痛い(痛み)
夜眠れません。	夜/眠れない
心臓が悪いのです。	心臓/悪い/PT1(私)
腹が痛みます。	お腹(腹)/痛い(痛み)
腹痛がします。	お腹(腹)/痛い(痛み)/ある
胃の調子が悪いのです。	胃/様子/悪い
胃の持病です。	胃/病気/持つ/PT1(私)
顔がむくんでいます。	顔/太る
軽い風邪を引きました。	軽い/風邪(咳)/変わる/PT1(私)
寒気がします。	寒い/ある/PT1(私)
喉が腫れて、喋りにくいのですが。	喉/腫れる(膨れる)/話す/そぐわない/けれども
喘息の発作が有ります。	喘息/発作/ある
喘息が私の持病です。	喘息/病気/持つ/PT1(私)
吐き氣がします。	吐く(吐気)/ある
仕事によく引き付けを起こします。	仕事/中/時々/ケイレン/起きる(起こる)
痛みはなくなりました。	痛い(痛み)/消す(消える)
下腹部が痛い。	お腹(腹)/下/それ/痛い(痛み)
虫に腕を刺されました。	虫/腕/刺す(刺さる)/PT1(私)
痒くてたまりません。	痒い/我慢/出来ない
口臭が有ります。	口/臭い/PT1(私)
辛いのです。	きびしい/PT1(私)
子供が引き付けを起こしました。	子供/ケイレン/起きる(起こる)
耳鳴りがします。	耳鳴り/PT1(私)
痛くて涙がでできます。	痛い(痛み)/涙/泣く
涙が出てきます。	涙/泣く
遠くの物がボーッとかすんで見えます。	視野/かすむ

表1：手話単語系列

((42 34) (621 10)) ;;; 621	■吐く(吐気)	■消す(消える)	■
((43 1) (856 10)) ;;; 856	■熱	■ある	■
((43 12) (496 10)) ;;; 496	■熱	■下	■
((43 14) (649 10)) ;;; 649	■熱	■我慢	■
((43 46 32) (525 5)) ;;; 525	■熱	■風邪(咳)	■重い
((44 14 33) (626 5)) ;;; 626	■発作	■我慢	■出来ない
((44 20 16) (550 5)) ;;; 550	■発作	■軽い	■感じ(感じる)
((45 10 50) (546 5)) ;;; 546	■病気	■悪い	■様子

表2：高階グラフの例

発話: 17 データ: 2 文: 喉が腫れて、喋りにくいのですが。

大きさ: 126 手話単語系列: (喉 腫れる (膨れる) 話す そぐわない けれども)

=====	268
-----*****-----	: 0.764: 腫れる (膨れる)
-----*****-----	: 0.821: そぐわない
-----*****-----	: 0.956: 話す
-----+++++++-----	: 0.992: 悪い
+++++++-----	: 1.045: 涙
-----*****-	: 1.067: けれども
-----+++++++-----	: 1.100: 熱
+++++++-----	: 1.134: 口
*****-----	: 1.174: 喉

図 2 : 意味共起関係高階グラフの効果

4.4 考察

図 2 に意味共起情報が有効であった典型例を 2 例示す。この例では「喉」と「腫れる」の共起によって、順位が上がっている。

また、アルゴリズムの不備は、被覆率のいき値小さくとったため、発話時間の短い切片が落ちたことに起因している。実験中は、バッファに残った系列に全ての単語束の切片を加えたとき、制約条件を満たす系列を 2 次出力とした。しかし、その後の検討で、アルゴリズムの修正、

5. 被覆率に関する制約条件を満たす系列を生成する。つまり、 $C_i(b) > \beta'$ for $b \in B_i$ を満たす b の集合を B_2 とする。 $A = A \cup B_2$ とする。また、 $C_i(b) > \beta'$ for $b \in B_i$ を満たす b の集合を B_3 とする。 $B_1 = B_1 \setminus B_3$ とする。ただし、 $\beta' < \beta$ 。

により解決できることが判明している。

本実験の場合、57 単語に関する高階グラフの作成には、これまででの 60 時間を費やしている。現在、122 例文での実験を検討中である。122 例文の場合、単語数は 148 となる。単語数が 3 倍に増加すると、辺の数は 9 倍に、3 次の高階辺の数は 27 倍に増加する。従って、代替となる高階グラフの作成手法の検討が必要である。

5. 結び

今回の実験により、意味共起関係を利用した手話認識は、認識方法として有望であるとの感触を得た。しかし、病院での会話を想定した実用規模の意味共起高階グラフの手作業での作成は非常に困難であり、代替手法の検討が必要であるとの結論を得た。

参考文献

- (1) 米川明彦: 手話言語の記述論的研究, 明治書院 (1984)
- (2) 佐川浩彦, 他: 圧縮連続 DP 照合を用いた手話認識方式, 信学論 J77-D-II, 4 (1994)