

## 手指動作特徴系列を用いた手話単語の類似検索システム

安達 久博

宇都宮大学工学部情報工学科

〒 321-8585 宇都宮市石井町 2753 番地

adachi@galileo.infor.utsunomiya-u.ac.jp

手話は聴覚障害者と健聴者との円滑なコミュニケーションを確立するための重要なコミュニケーション・パスの一つとして、位置づけることができる。このため、聴覚障害者の社会参加の割合が年々増加するに従い、手話を学習する健聴者の数も増加する傾向にある。しかし、他言語学習の核となる対訳辞書における提供形態は、従来の冊子体形式とその構造を踏襲した機械可読化形式であるという現状である。すなわち、電子化媒体の持つ潜在的な特徴である構造化や検索手段を十分反映した電子化辞書は現在、提供されていない。そこで、本稿では、最近特に注目され普及しつつある WWW 環境における手話単語の類似検索システムの構築方法を提案し、Java 言語ならびに CGI 技術に基づくシステムの一実現例を示す。本検索システムの特徴は、(1) 手指動作表現の指示手段 (GUI) として、Java アプレットにより実現された動作指定パレット上で、マウสดラッグ操作による直接ポイントを可能とした。(2) マウสดラッグの軌跡 (座標) 情報に基づき、動作特徴系列を自動生成し、システム内の手指動作記述文との類似性により、類似動作を持つ日本語対応ラベルの手話単語を検索する手段を提供する点にある。この類似動作単語の検索機構は、動作特徴の一部を置換することで別の意味または、類似の意味を持つ手話単語を構成する、手話単語の造語法の特徴をも効率良く修得できる利点となる。

### A Retrieval System for Japanese Signs Using Manual Motion Sequences

Hisahiro Adachi

Dept. of Information Science, Utsunomiya University

2753 Ishiimachi, Utsunomiya, 321-8585 Japan

Sign language is an important path to communicate with impediment people in hearing. Therefore, the number of hearing people learning sign language are on the increase. However, there is no signed electronic dictionary system that reflected the potential characteristics of electronic media. This paper describes a method of constructing a system to retrieve signs effectively and shows an implemented system over the World Wide Web. The system has the remarkable advantages of retrieving the several signs with similar motions simultaneously.

## 1 はじめに

手話は、聴覚障害者と健聴者との重要なコミュニケーション・パスの一つである。近年、この手話を学習する健聴者が増加傾向にある。また、最近ではインターネット上でも手話辞書を公開し、検索できるサイトも増えつつある。さらに、冊子体形式の辞書を電子化媒体 (CD-ROM) として市販されている電子辞書も幾つか提供されている。しかし、これらの辞書構造の多くは、源辞書である冊子体形式の index 構造をそのまま保持する形で電子化されているため、現状の検索キーは、基本的に日本語の単語見出しであり、日本語見出しからの単方向の検索手段が提供されている現状である。しかし、手話が顔の表情、視線の向き、手指動作を伴う身体表現により表出、受容される視覚言語である特徴から、従来の紙による媒体の制約から、イラストや静止画像でしか提供できなかったものが、動画像として提供される点は手話の学習者、特に、初心者にとって、より情報量は多く、かつ視覚的にその動作特徴を捉えることが可能となった。

一方、手話の電子化辞書に関する工学的な研究は、これまでに幾つかの検討と提案がなされている。これらの研究を大別すると、(1) 手話単語の表記法に関する研究、(2) 手話単語の検索に関する研究 (3) 手話単語の分類に関する研究の 3 つに分類することができる。特に、(2) の手話単語の検索に関する研究は、上で述べたように、日本語単語見出しからの検索手段が従来の音声言語の要素技術で構築可能なため、手指動作特徴に基づく手話側からの検索方法の研究が主である。これらの検索方式としては、(1) の表記法を利用する方式と自然言語文により記述された手指動作記述文を利用する方式が提案されている。(3) の分類についても同様に大別できる [1][2][3]。前者の方式は、表記法が手話単語の構造を精密に記述することを目的として研究されているため、すなわち、その記述形式から手話の動作表現が再現可能なように、精緻なコード体系に基づいており、検索条件の指定が複雑化し、初心者には利用しづらいとの指摘がある [4]。しかし、検索条件の段階的な指定により、候補を最終的に一つに絞りこめる利点がある。一方、後者の方式は、検索条件の指定を自然言語で記述する容易さはあるが、自然言語の持つ冗長さ、同一の動作に対

して、別の言語表現が可能のため、入力記述文とシステム内の手指動作記述文との表現の差が大きくなると検索効率が低下するという問題がある。また、最終的に検索候補を一つに絞り込まず、類似動作を持つ単語を検索することを目的としている。

本稿では、後者の問題点である入力記述文の正規化に焦点をあて、システム内の手指動作記述文との部分照合により、類似動作の手話単語を検索するシステムの実現例について報告する。入力記述文の正規化は、形態素解析等の自然言語処理技術を利用するアプローチも考えられるが、本システムでは、マウスのドラッグ軌跡 (座標計算) により、動作特徴系列を抽出し、この特徴系列に基づき入力記述文を生成するアプローチを検討した。

## 2 手話単語の類似検索の重要性

### 2.1 手話単語の特徴

手話言語は、視覚言語の側面を持つため、動作の類似性に意味の類似性を反映させて、新しい手話単語を造語する傾向がある。すなわち、手の形などの手指動作特徴の一部を別の要素で置き換えることで類似の概念に対応する別の手話単語を構成することができる。例えば、図 1 は、「甘い」と「辛い」の単語対を示す。この単語対は、両者の動作特徴の差が手の形だけが異なる手話単語の最小対を構成する例である。同時に、両者に共通の概念ラベルとして、「味覚概念」を持つことが分かる。この場合、概念の類似性を規定している動作特徴が、口の前で手指動作を行なうという、動作位置の共通性に起因している可能性が高いが、動作の類似性は、検索トリガとして利用できると考える。

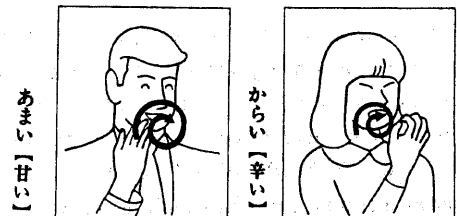


図 1: 動作の類似性が意味の類似性を反映している例

一方、図 2 は、手の形と掌の向きが異なるが、手の動作位置と動作が同じ類似の動作特徴を持つ単

語対「なめらか」と「横浜」を示す。この単語対には意味の類似性は無いが、語源として、どちらも「西洋カミソリ」に由来している。(髭剃り後は肌がなめらかになり、西洋剃刀が最初に陸揚げされた港が横浜であった。)



図 2: 動作の類似性が意味の類似性を反映しない例

このように、手話単語の動作の類似性は意味の類似性を反映している場合が少なくない。これは、英語が 26 文字の組み合わせによる配列順序の類似性が意味の類似性を表現する場合や別の意味を表現する場合 (book, look) があるのと同じである。また、日本語でも漢字の接続による派生語の生成 (考察, 考慮, 思考など) と同様である。手話を学習する場合、この動作特徴の類似性と差に関する情報は、単語の認識や表出の際の弁別要素を明確にするだけでなく、類似の動作単語の意味との関係を調べる際に重要な検索キーとなる。特に、手の形、手の位置、手の動きの 3 つの特徴素の中で、最も情報量が多く、その特徴がほぼ正確に記憶に残りやすいと考えられるのが、手の動きである。このため、動作特徴の類似性をキーとする検索戦略は、目的とする手話単語のみを検索する絞り込みの処理の負荷を減らし、関連する単語群を同時に提供することに意味がある。

### 3 手話単語の類似検索システム

#### 3.1 検索システムの概要

本検索システムの構成は、図 3 に示すように、WWW 上での動作環境を想定し、(1) クライアントからの動作特徴に基づく検索要求をサーバに転送する動作特徴認識生成部と、(2) 検索条件を満たす手話単語を単語データベースから検索し、対応する html 文書を動的に作成し、クライアントに転送する単語検索エンジン部からなる。ここで、(1) は Java 言語で記述された動作軌跡を認識する認識

アプレット、(2) は C 言語で記述された CGI により実現されている。

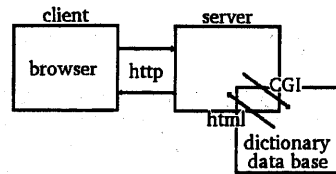


図 3: 検索システムの構成図

また、図 4 に、(1) のクライアント側の検索画面を示す。上部の左側が、動作軌跡の指定を行なう、認識アプレットであり、右側は、動作軌跡を表示する軌跡表示アプレットである。また、下部の左側に検索された単語リストを表示し、各単語をクリックすると、対応する手話単語のアニメーションが表示される画面構成となる。

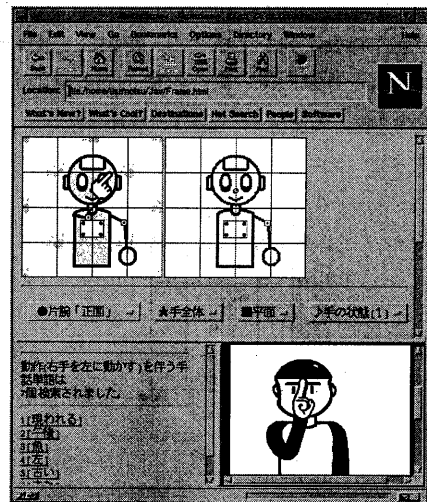


図 4: クライアントの検索画面例

#### 3.2 特徴系列の生成

認識アプレットでは、マウスの座標位置により、図 5 に示すように、9 桁のコード化された数字列の特徴系列を生成する。ここで、 $move(i)$  は、動作軌跡 (4 桁)、 $position(i)$  は、動作位置 (2 桁)、 $hand(i)$  は、

手の形状 (2桁), *change\_back* は, 背景画像 (1桁) を示す. なお, 初期値は (000|00|00|0) である.

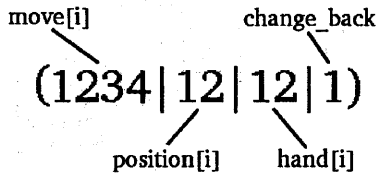


図 5: 動作特徴系列の構造

### 3.3 動作軌跡の認識方法

図 6 に示すように,  $200 \times 200$  ピクセルの認識パレットを  $4 \times 4$  の領域に分割し, マウスのドラッグ軌跡が, 別の領域に移動した時点で, 表 1 に従い, up(4), down(2), left(3), right(1) の 4 方向を決定し, 4 桁の動作特徴系列を生成する. 表 2 に, 正面背景画像と横向き背景画像の動作特徴系列と対応する動作を示す.

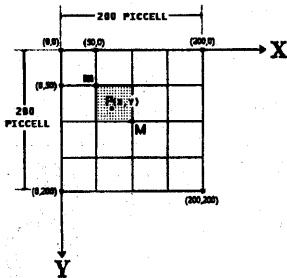


図 6: 認識パレットの領域分割

ここで, マウスドラッグ状態の位置座標を  $P(x, y)$  とすると, この領域の  $m(x_{min}, y_{min})$  と  $M(x_{max}, y_{min})$  は, それぞれ,  $x_{min} = x - (x \text{ mod } 50)$ ,  $y_{min} = y - (y \text{ mod } 50)$ ,  $x_{max} = x_{min} + 50$ ,  $y_{max} = y_{min} + 50$  となる. また, 新しい位置座標を  $P(X, Y)$  とすると, 移動軌跡は, 表 1 に示すように決定される.

表 1: 移動軌跡の決定

マウスの軌跡	方向	コード
$X > x_{max}$	Right	1
$X < x_{min}$	Left	2
$Y > y_{max}$	Down	3
$Y < y_{min}$	Up	4

表 2: 軌跡に基づく動作特徴系列と動作の関係

軌跡動作特徴系列	動作	
	正面	横
1234, 2341, 3412, 4123	回転	回転
1000	左	前方
3000	右	後方
1300, 1313, 3100, 3131	左右	前後
4000	上	上
2000	下	下
2400, 2424, 4200, 4242	上下	上下

### 3.4 検索条件 (背景画像) の指定

表 2 に示したように, 認識パレットの背景画像の違いにより, 同一の特徴系列に対応する動作が異なる場合がある. 従来の検索方法では, 手話単語の構成要素として, (1) 手の形, (2) 手の位置, (3) 手の動きの 3 種類の動作特徴素を検索条件として指定する方式が一般的であった. これは, (1) と (2) は, 比較的容易にメニューあるいはアイコンを利用して指定できるため, これらの情報により分類した検索辞書を用いて, 候補を絞り込むためと考える. しかし, これらの検索条件の指定を間違えると, (3) の情報にマッチする候補が存在しなくなる可能性がある. そのため, 本検索システムでは, 手話単語の動作特徴として, 手の動きを最優先とし, 動作位置に関しては, 無効とする戦略を採用した. 位置情報を有効に利用する場合は, 身体部位を指示する場合のみとした. また, 手の形に関しても, 顔の部分に指示する場合以外は無効とした. 背景画像は, 表 3 に示すように, 大きく片手手話と両手手話に分類し, 片手手話は, 正面, 横向き, 上部に, 両手手話は, 正面左右同時動作, 正面左右交互動作, からなる. 図 7, 図 8, 図 9 と 図 10 にそれぞれの背景画像を示す.

表 3: 背景画像と特徴コードの対応

手	身体, 動作	コード
片手	正面	0
	横	1
両手	正面左右対称	2
	正面点对称	3

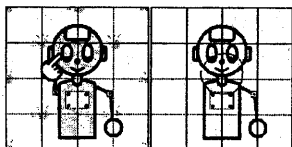


図 7: 片手・正面

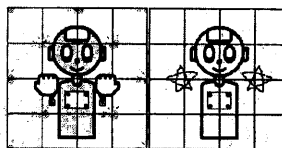


図 9: 両手・左右対称

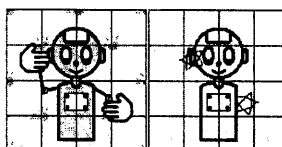


図 10: 両手・点对称

### 3.5 動作・指示位置の認識

手指動作が体のどの部分を指示しているかを、マウス座標の位置情報により、以下のように分類した。

- (1) 図 11 に示すように、顔の各部位を指示で規定する手話の場合は「額」「右目」「左目」「右耳」「左耳」「鼻」「口」「のど」の 8 領域に分類し、表 4 に示すコード番号を、
- (2) 図 12 に示すように、遷移動作を伴う手話の場合は「顔」「顔の右側」「顔の左側」「体」「体の右側」「体の左側」の 6 領域に分割し、表 4 に示すコード番号を対応付ける。

なお、本検索システムでは、以下の理由により、(2) のコード番号は検索条件に採用しない方針とした。

1. 本稿で対象言語データとして利用した手話辞典の 650 単語の手指動作記述文 [5] は、原則として、手話イラストを捕捉する目的で記述されており、身体の中の領域(部位)で、その手

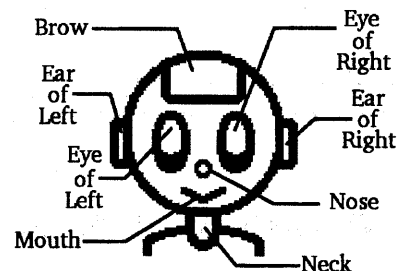


図 11: 顔の領域分割

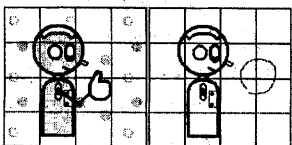


図 8: 片手・横向き

### 3.6 手指動作特徴系列文の生成

手指動作特徴系列文は、前節で述べた系列により以下の順序により生成される。

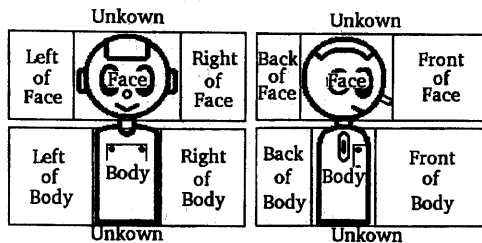


図 12: 身体の領域分割

表 4: 顔, 身体の位置情報と特徴コード

顔 領域	身体		特徴 コード
	正面	横	
口	顔	顔	1
鼻	胸	胸	2
右目	顔の右	顔の前方	3
左目	顔の左	顔の後方	4
右耳	身体の右	身体前方	5
左耳	身体の左	身体後方	6
額, 頭			7
首			8

- (1) 背景画像のコードにより,“(右手, 両手)|を”に置換される,
- (2) 動作軌跡のコードにより,“(右に, 左に, 左右に, 上に, 下に, 上下に, 前方に, 後方に, 前後に, 回転)”に置換される,
- (3) 両手の場合は, “交互に”, “同時に” が (1) と (2) により置換された, 文字列の間に挿入される.

### 3.7 動作特徴系列文と動作記述文の照合

与えられた動作特徴系列文とデータベース内の動作記述文とを分離記号間の文字列(トークン)を順序配列に従い照合操作を行なう。例えば, 動作特徴系列文として, 右手|を|回転が与えられた場合,

右手の掌を口の前にあて二度ほど回転させる

が検索される。なお, “右に” と “下に” の文字列照合の場合は, 直前の文字が, それぞれ, “左”, “上” でないかのチェック操作を行なう。

### 3.8 検索操作手順

図 4 に示した検索画面を例に検索操作手順を以下に述べる。

- (1) 最初に, 背景人物画像の選択を行なう。すなわち, 身体の前後に腕を動かす場合は横向きの背景画像に, 両手手話を対象とする場合は, 両手の背景画像というように, 検索条件を指定し, この指定条件にマッチする検索結果を表示する。この例の場合は, 片手手話で, かつ, 正面背景画像を指定している。
- (2) 次に, 動作の軌跡をマウスドラッグにより指示し, 右方向への動作と認識し, 動作特徴系列の文字列を引数として, 検索サーバの CGI を起動する。
- (3) CGI では, 引数として渡された動作特徴系列から, 右手を左に動かすという動作特徴系列文を生成し, システム内の手指動作記述文との部分照合により, 該当する手話単語群を検索し, html 文書を動的に作成し, クライアントに転送, 表示する。ここで, システムの手指動作記述文は, 手話動作主体の側から見た左右方向であるのに対して, 軌跡認識部の左右方向は逆の関係となる。
- (4) 検索結果画面の単語をクリックすると対応する手話アニメーションが提示される。

## 4 検討

本検索システムの検索効率とユーザインタフェースに関する評価は, 現在進行中であるが, 現状で明らかとなった問題点について, 述べる。本システムの検索要求(検索 CGI を起動する)のタイミングは, ある領域に位置するマウス座標が別の領域に進入した時点で, 検索処理が始まり, 検索結果が表示される, 動的類似検索の枠組を採用している。すなわち, 検索条件の指定(特徴系列の部分的な代入)と同時に検索が実行され, その時点の特徴系列の状態に基づく, 結果が表示される方式と定義され, 検索条件の変化に応じてリアルタイムな検索が可能となる。例えば, 回転動作の中間状態の “右手|を|上に”, “右手|を|右に”, “右手|を|下に”,

“右手 | を | 左に” に対するそれぞれの検索結果を表示する。この結果として、CGI からの応答を待つタイミングにより、マウスの動きが遅くなり、動作軌跡の入力操作が影響を受けるという問題がある。一方、検索効率の点でも、例えば、回転動作を持つ手話表現に、水平方向に回転する軌跡は、指定できないことが、明らかとなった。このため、背景画像として、頭上方向からの mode を追加するか、2次元平面上で疑似的に3次元空間の軌跡の表示と動作認識のアルゴリズムを検討する必要がある。また、複雑な動作を伴う手話表現に対して、自然言語では比較的容易に記述できるが、マウストラッグの認識処理部からの特徴系列と、この特徴系列から生成する特徴系列文の生成方法を検討する必要がある。なお、検索効率を向上させるためには、本稿では、検索条件に利用していない動作位置の情報の利用方法を検討する必要があると考える。このように、現時点で解決しなければならない問題はあがあるが、動作特徴素に基づき、分類した検索辞書を必要としない本提案手法は、単語数が大規模になる検索辞書システムの編集管理を容易にする可能性があると考えられる。

## 5 おわりに

本稿では、手話単語の動作特徴を検索キーとする手話単語の類似検索システムの一実現例を示した。検索方式の特徴は、マウストラッグ操作により得られる動作特徴系列から動作記述文を生成し、この記述文とシステム内の手指動作記述文との部分照合により、検索を行なう点にある。このため、事前に、動作特徴に基づき分類された検索辞書データベースを用意する必要がない。また、手話の学習を支援する検索システムの機能として、類似の動作特徴を持つ手話単語の類似検索の重要性について述べた。今後は、検索システムの検索効率に関する評価と従来方式との比較評価を行なう必要がある。また、ユーザインタフェースの評価データを収集するため、試験的な公開を現在検討中である。

## 謝 辞

本研究を進めるにあたり、日頃ご指導を頂く宇都宮大学工学部情報工学科鎌田一雄教授、ならびに貴重なご意見を頂く熊谷毅助教に深謝する。なお、本研究の一部は平成9年度文部省科学研究費補助金(重点領域研究(2):課題番号09204202)、平成9年度厚生科学研究費補助金(感覚器障害及び免疫アレルギー等研究事業(感覚器障害研究分野)),ならびに平成8年度(財)電気通信普及財団研究助成金、平成9年度(財)実吉奨学会研究助成金による。

## 参考文献

- [1] H. Adachi and K. Kamata, "A Classification Method for Japanese Signs Using Manual Motion Descriptions," *Proc. of COLING'94*, vol. II, pp. 961-967, Kyoto, August 1994.
- [2] H. Adachi and K. Kamata, "A Similar Retrieval Method from Japanese Signs Using Manual Motion Descriptions," *Proc. Int. Conf. on Computer Processing of Oriental Languages*, Hong Kong, April 1997.
- [3] 安達久博, 他, "日本語・手話電子化辞書の構成法について", 情報処理学会, NLP96-3, pp. 17-24, 1993.
- [4] 内藤一郎, 加藤雄士, "効率的な検索方法の確立を目的とした日本手話の形態的特徴の基礎的検討", 電子情報通信学会論文誌 A, Vol. J79-A, No. 2, pp.337-345,1996.
- [5] 丸山浩路編, イラスト手話辞典. KK-ダイナミックセラーズ, 1989.