

Leap Motion Controller を用いた 手話動作認識手法の実装と評価

小林実来[†] 武田敦志[†]

本稿では、日本語の手話の中でも五十音を表す指文字に着目し、Leap Motion Controller を用いることで手の形、及び動きを認識し、指文字の推定が可能であることを示す。まず、手の認識、及び指文字の推定方法を提案する。この推定方式では、Leap Motion Controller から得られる各座標から、それぞれの指のベクトルを算出し、これらのベクトルと基準ベクトルとのユークリッド距離から手話が示している文字を推定する。また、実験結果より、提案手法を用いた場合の認識率を示す。

Implementation and Evaluation of Sign Language Recognition by using Leap Motion Controller

Mirai Kobayashi[†] and Atsushi Takeda[†]

In this paper, we propose a sign language recognition system which determines characters of Japanese sign language. The proposed system uses Leap Motion Controller in order to obtain space coordinates of a right hand, and derives vectors of right fingers which are compared with reference vectors. Based on Euclidean distance between finger vectors and reference vectors, the proposed system estimates a character signed by the right hand. We also show recognition rate of the proposed system through experiment results in this paper.

1. はじめに

平成 25 年のデータ[1]によれば、日本における聴覚・言語障害を抱えている人はおよそ 34 万人である。これは、日本の当時の人口[2]から計算すれば約 2 千人に 1 人が障害を抱えて生活していることになる。彼らのコミュニケーション手段は手話であり、これを理解するためには英語や中国語等の他言語と同じように学習しなくてはならない。しかし、現代社会において手話を使うタイミングというものは少なく、そのために手話を理解している健常者は少ない。そこで、手話を使う人とのコミュニケーションを行う際に、コンピュータへ何らかの手段で手話の情報を送信し、手話を翻訳させることで、手話を理解していない人でも手話の意味が通じ、円滑なコミュニケーションが可能となる。

コンピュータへの入力装置として、赤外線を用いて手の情報を詳細に取得することのできる LeapMotionController[3] (以下 LMC) というものがある。過去に LMC を用いた手の形の認識研究としてアラビア語の手話認識[4]を行ったという例があり、こちらは高い認識精度で手の形を取得できた。しかし、アラビア語と日本語の手話の大きな違いとして、日本の手話には動作が必要なものが存在する。動作のないものならば 1 枚の座標データがあれば形の判定が可能だが、動きを認識するには複数の座標データを処理しなければならない。

よって本稿では、日本語の手話の中でも五十音を表す指文字に着目し、LMC を用いることで手の形、及び動きを認識し、指文字の推定が可能であることを示す。そのために手の認識、及び指文字の推定方法を提案し、実験を行い、認識率を調査する。なお、本システムのデータ取得対象は一人で行い、右手のみの認識とする。既存研究として、アルファベットを表す指文字を、ユークリッド距離とコサイン類似度を用いて判定し、認識精度を比較するというものがある[5]。本稿では既存研究からユークリッド距離を用いるとともに、認識率の向上のため指の本数での判定を行う。また、アルファベットを表す指文字にも動作が必要なものがないため、LMC から得られる座標をベクトルに変換し、認識を行う。

2. 指文字認識手法の提案

2.1 指文字認識手法の概要

本稿では、LMC を利用して日本語の指文字を推測する手法を提案する。提案手法では手の形を推定するために、LMC から得られる各座標間のユークリッド距離を利用する。始めに動かない状態での手の形を推測し、その形に応じて動きの認識が必要なのかを判断する。もし動きの必要でない手の形だった場合はその情報から、必要だった

[†] 東北学院大学教養学部情報科学科
Department of Information Science, Tohoku Gakuin University

場合には更に動きの測定の後に指文字を推測する．提案手法を用いることにより，動きのないものは素早く結果を，動きのあるものは手の形を推定してから動きを読み取るためより正確に結果を出すことができる．この章では，指文字の推定に用いる手法について述べる．

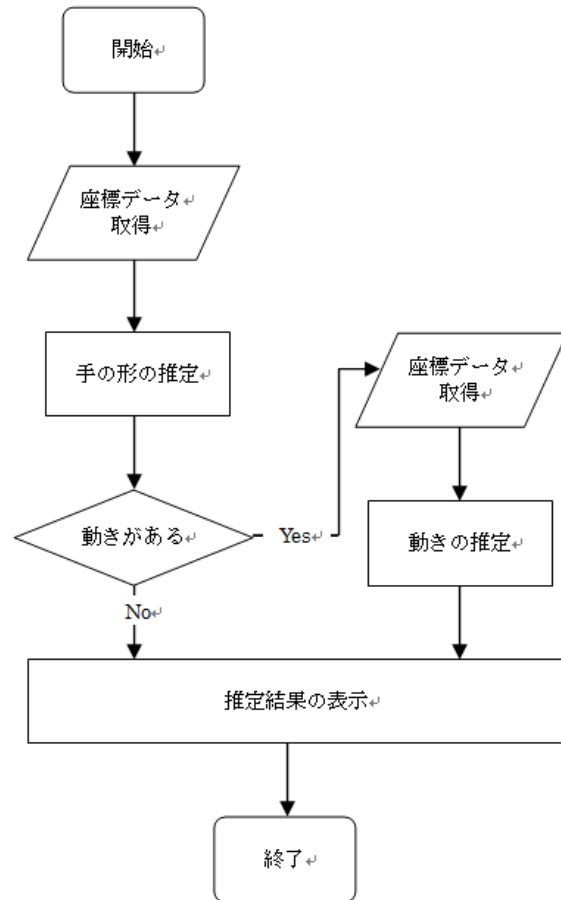


図 1 指文字認識フローチャート

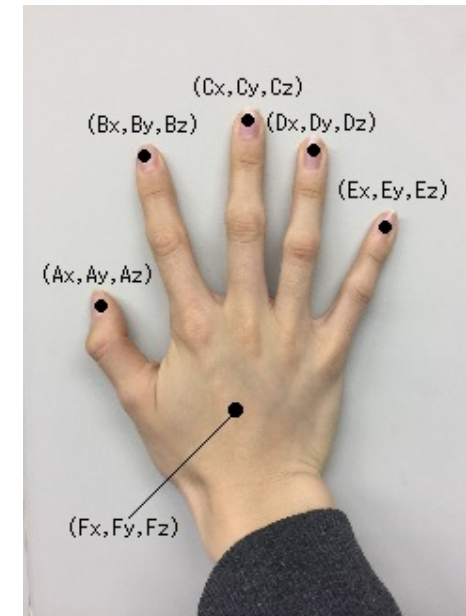


図 2 各点の座標

2.2 手の形の推定

提案手法では LMC より取得した情報のうち，各指先の座標と手のひらの中心座標，及び伸ばしている指の本数を用いる．取得した座標を用いて，手のひら中心から各指先までの距離を以下のように求める．

$$G_1 = \sqrt{(A_x - F_x)^2 + (A_y - F_y)^2 + (A_z - F_z)^2} \quad (1)$$

$$G_2 = \sqrt{(B_x - F_x)^2 + (B_y - F_y)^2 + (B_z - F_z)^2} \quad (2)$$

$$G_3 = \sqrt{(C_x - F_x)^2 + (C_y - F_y)^2 + (C_z - F_z)^2} \quad (3)$$

$$G_4 = \sqrt{(D_x - F_x)^2 + (D_y - F_y)^2 + (D_z - F_z)^2} \quad (4)$$

$$G_5 = \sqrt{(E_x - F_x)^2 + (E_y - F_y)^2 + (E_z - F_z)^2} \quad (5)$$

また、認識率を向上させるため、人差し指から中指、及び薬指までの距離を、以下の式から求める。

$$G_6 = \sqrt{(B_x - C_x)^2 + (B_y - C_y)^2 + (B_z - C_z)^2} \quad (6)$$

$$G_7 = \sqrt{(B_x - D_x)^2 + (B_y - D_y)^2 + (B_z - D_z)^2} \quad (7)$$

求められた7つの距離と、事前に準備しておいた各指文字の7つの距離から、ユークリッド距離を用いて判定値 G を求める。

$$G = \sqrt{(G_1 - g_1)^2 + (G_2 - g_2)^2 + \dots + (G_7 - g_7)^2} \quad (8)$$

文字ごとに算出された判定値を比較し、最も値が小さくなったものが定めた閾値以内であれば、その手の形であると判断できる。

2.3 手の動きの推定

日本語の指文字において、動作のある文字は5文字である。すべての指のデータを比較してしまうと、指先の一部が隠れて正しく認識できない。そこで、本稿では動作のある文字ごとの特徴を分析し、それらを用いて指文字の判定を行う。提案手法では伸ばしている指に着目し、その伸ばしている指の座標変化を測定に利用する。また、指を一本も伸ばしていないものに関しては、手のひらの座標変化を利用する。判定方法としては、複数時間分の座標データを取得し、それらを1つのベクトルに変換する。その後登録しておいた指文字のデータからユークリッド距離を算出し、判定値を求める。その判定値が定めた閾値以内であれば、同じ動きであると判断できる。

表 1 動作のある指文字の特徴

指文字	伸ばしている指	特徴となる座標
の	人差し指	(B_x, B_y, B_z)
ん	人差し指	(B_x, B_y, B_z)
を	なし	(F_x, F_y, F_z)
も	親指, 人差し指	$(A_x, A_y, A_z), (B_x, B_y, B_z)$
り	人差し指, 中指	$(B_x, B_y, B_z), (C_x, C_y, C_z)$

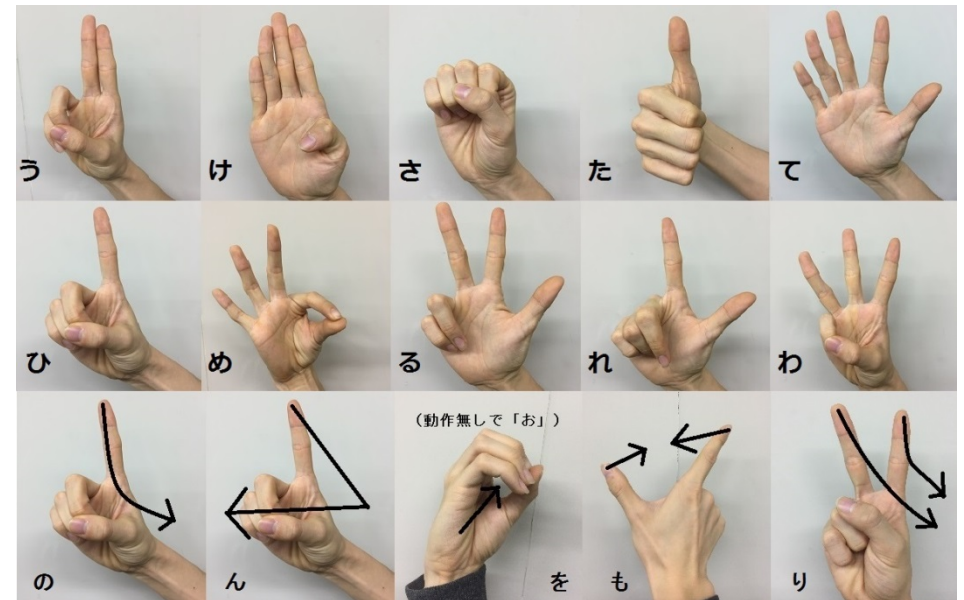


図 3 登録した指文字

3. 提案手法の実装と評価

3.1 手の形の推定手法の実装

本稿では、動作前の手の形を含む、動作のない手の形を計14件登録した。まず10F分の座標データを取得し、手のひらの座標でそれぞれ正規化を行う。その後各座標の平均値を算出し、それらを用いて G_1 から G_7 を求める。この作業を1文字につき10個ずつ行い、それぞれ判定値を算出した。求められた10個の判定値の平均から、最も値が小さかったものを、各文字のデータとして登録した。

判定する手からも同じように10F分の座標データを取得、正規化を行い、平均値から G_1 から G_7 を求める。その後、伸ばしている指の本数が一致している指文字データとのユークリッド距離を算出し、判定値を比較する。求められた判定値が最も小さいものが、定めた閾値よりも小さければ、その形であると判定する。閾値は、前述のデータ登録作業時の傾向から、100を上回らなければ間違った形ではないことがわかっている。ここで判定した指文字が動作のないものであればその文字の出力を、動作のあるものであれば動作の取得に移る。

3.2 手の動きの推定手法の実装

動きの推定に関しては、30F分のデータを取得する。取得し正規化した座標の中から、各指文字の特徴となる指の座標データを一つのベクトルに変換する。こうして算出したベクトルの成分を各指文字10個ずつ取得し、最も判定値が低いものをその指文字のデータとして利用する。形の推定から動きの取得に移行した際に、その指の形からどの特徴量を用いればいいのかを判断し、利用する指、もしくは手のひらの座標データ30個をベクトルに変換する。変換したベクトルと、形から判断される指文字のデータのユークリッド距離から判定値を算出し、最も小さく、且つ閾値以内のものであればその指文字であると判定する。なお、動きの推定に用いる閾値も、手の形の推定で用いた100とする。

3.3 手の形の認識率の評価

まず、固定された手の形を判定する実験を行った。本実験では、主に手を正面に向けるタイプの指文字を10個厳選し、それらのデータを準備した上で推定実験を行った。実験の内容としては、指文字ごとに50回の認識を行い、それぞれ正答率、及び失敗した場合の原因を調査した。結果として、認識率は50-96%となり、全体の認識率は約78%であった。誤認識の原因を解析したところ、伸ばしている指の本数を間違えたことによる認識ミスが多いことが分かった。伸ばしている指の本数は LMC

う	0.90					0.04						0.04	0.02
け	0.02	0.86					0.04					0.02	0.06
さ	0.02		0.74			0.18						0.04	0.02
た	0.08	0.02	0.20	0.50		0.06						0.08	0.06
て					0.96							0.04	
ひ	0.02		0.02	0.20		0.72						0.02	0.02
め		0.08					0.74					0.02	0.16
る								0.76					0.24
れ									0.76			0.08	0.10
わ		0.02		0.04		0.02	0.02			0.82		0.02	0.06
う	け	さ	た	て	ひ	め	る	れ	わ	認識切れ	Failed		

図 4 固定された手の形の認識率

(認識切れ：認識が途中で中断された Failed：判定値が閾値を超えた)

表 2 認識ミスのうち、本数を間違えたものの内訳

指文字	判定結果	伸ばした指の本数	誤判定した件数	指文字	判定結果	伸ばした指の本数	誤判定した件数	指文字	判定結果	伸ばした指の本数	誤判定した件数	
う	ひ	1	2	た	さ	0	10	る	failed	4	8	
	failed	3	1		う	2	4		failed	5	4	
け	認識切れ		2	て	failed	2	2	れ	failed	1	2	
	う	2	1		け	4	1		め	3	3	
	め	3	2	認識切れ		4	failed	3	2			
	failed	3	1	2	failed	3	2	認識切れ		4		
さ	failed	5	2	ひ	さ	0	1	わ	た	1	2	
	failed	1	1		う	2	1		ひ	1	1	
	認識切れ		1	failed	2	1	failed	2	1			
	ひ	1	9	め	3	1	け	4	1			
さ	う	2	1	け	4	4	failed	4	1			
	認識切れ		2	め	failed	4	7	認識切れ		2		
				認識切れ		1						

から直接取得しているため、LMC 側で指の本数を上手く数えられていないことになる。数えられなかった原因としては、LMC のデータ取得の性質上、指が隠れてしまった際に正確なデータが取れず、誤認識したと考えられる。もしも指の本数をより正確に数えられていたと仮定した場合に、本実験の結果は図 6 のようになる。こちらの場合だと全体の認識率は約 95%であった。このような結果から、指の本数を適切に計測できるようにすれば、より正確な認識が可能であることが分かった。

う	1.000												
け		1.000											
さ			0.974										0.026
た				0.862		0.103							0.034
て					1.000								
ひ				0.217		0.783							
め							0.974						0.026
る								1.000					
れ									0.974				0.026
わ										0.976			0.024
う	け	さ	た	て	ひ	め	る	れ	わ	認識切れ	Failed		

図 5 指の本数を加味した場合における固定された手の形の認識率

3.4 手の動きの認識率の評価

続いて、動きのある指文字を判定する実験を行った。本実験では動きのある指文字5種に加え、「の」「ん」と同じ手の形である「ひ」、「を」と同じ形である「お」を加えた計7つの文字の判定を行う。実験内容としては、指文字ごとに50回の動きの認識を行い、それぞれ正答率、および失敗した場合の原因を調査した。実験の結果、表 2 のような結果となった。手の形の認識失敗や動作中に手の認識が切れてしまった等の認識ミスが多く、平均として75%の認識率となった。動作中の認識切れに関しては、手の一部が隠れてしまう等の要因で上手く認識できない場合に、右手のはずなのに左手と誤認識してしまう、もしくは認識しているものが手ではないと判断してしまうことから起きている現象である。これらを除いた、純粋な動作のみの認識率は96%であった。このことから、指文字の動作は正確に推定可能である。

表 3 動作の推定結果

指文字	正答数	誤答数	認識ミス	指文字	正答数	誤答数	認識ミス
ひ	49	1	0	お	37	0	13
の	36	4	10	を	34	2	14
ん	33	4	10	も	34	0	16
り	38	0	12				

表 4 認識ミスの内訳

指文字	認識ミスの内容	認識ミスした件数
の	動作中に認識が切れた	10
ん	同上	13
り	同上	12
お	手の形を「さ」と認識した	13
を	同上	14
も	動作中に認識が切れた	3
	手の形を「り」と認識した	1
	手の形を認識できなかった	12

4. おわりに

本稿では、日本語の手話の中でも五十音を表す指文字に着目し、LMC を用いることで手の形、及び動きを認識し、指文字の推定が可能であることを示すため、手の認識、及び指文字の推定方式を提案し、実験を行い、認識率を調査した。調査の結果、手の形の推定は伸ばしている指の本数が正確に計測できた場合の認識率が 95%、動作の推定は認識切れや手の形の誤認識を除いた純粋な動作のみの認識率が 96%であった。このことから、LMC を用いることで日本語の指文字を認識することが可能であることを示した。今後、LMC の認識精度の向上により誤認識が減少することで、精度の高い認識システムが構築可能である。

参考文献

- 1) 内閣府, 障害者白書 平成 25 年版, 2013.
- 2) 総務省, 人口推移 (平成 25 年 10 月 1 日現在), 2013
- 3) Leap Motion, <https://www.leapmotion.com/>
- 4) M. Mohandes; S. Aliyu; M. Deriche, Arabic sign language recognition using the leap motion controller, 23rd International Symposium on Industrial Electronics (ISIE), pp.960-965, 2014.
- 5) Rajesh B. Mapari; Govind Kharat, Real time human pose recognition using leap motion sensor, International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), pp.323 - 328, 2015