

手話認識性能向上のための複数の識別器の統合方法の検討

小澤 辰典[†] 坂本 一樹[‡] 西村 広光[†] 田中 博[†]

小林 大輔^{*} 岩本 典夫^{*} 加藤 秀司^{*}

神奈川県立大学 大学院 情報工学専攻[†] 神奈川県立大学 情報学部 情報工学科[‡]

株式会社 ケイ・シー・シー^{*}

1.はじめに

手話は、一般にろう者が用いるコミュニケーション方法の1つである。しかし多くの聴者は手話によるコミュニケーションを習得していないため、ろう者と聴者とのコミュニケーションに大きな隔たりが存在する。そこで著者らはスマートフォンに内蔵されているカメラとCPUを用いた手話の自動翻訳の実現を最終目標に取り組んでいる。

2.カラー手袋による特徴量の抽出

手話において各指や、両手の動き、手の表裏等の情報は重要な要素である。そこで著者らは専用に制作したカラー手袋の色情報を用いることで各要素の検出を行っている。カラー手袋は各指の先端にそれぞれ異なる色を5色、両手首に異なる色を2色、掌に1色、合計8色を彩色し、それらの各色領域の重心位置と面積から6種の特徴量を抽出している[1]。

3.認識手法

3.1.使用する識別器

著者らは手話認識手法として、識別器を複数使用している[2]。本稿では単独の識別器による判定処理を「識別」、複数の識別器を組み合わせた判定処理を「認識」とよぶ。各識別器は、判定の基準が異なり、これらを適切に組み合わせることで、判定性能の向上が期待できると考えられることから各識別器を組み合わせる方法を提案する。

使用する識別器として、Hidden Markov Model (HMM), Support Vector Machine (SVM), Discriminant Analysis (DA), Linear Classification

“Investigation of Integration Method of Multiple Classifiers for Improvement of Sign Language Recognition Performance”
Tatsunori Ozawa[†], Kazuki Sakamoto[‡], Hiromitsu Nishimura[†], Hiroshi Tanaka[†],
Daisuke Kobayashi^{*}, Michio Iwamoto^{*}, Shuji Kato^{*}
Course of Information and Computer Sciences,
Graduate School of Kanagawa Institute of Technology[†]
Department of Information and Computer Sciences,
Kanagawa Institute of Technology[‡]
KCC corporation^{*}

Model (LCM), k-Nearest Neighbor algorithm (k-NN), Decision Tree (DT) の6種を用いることとした。

3.2.各識別器の結果の統合方法

著者らはカラー手袋の色領域から抽出した特徴量を各識別器6種にそれぞれ入力し、各識別器の単独の性能である識別率を算出する。その識別率から各識別器の順位づけを降順に行う。次に各識別器から得られた識別結果を組み合わせることで判定性能の向上を試みる。本稿では各識別器の統合方法を2種提案する。2種の方法は、それぞれ「統合法I」、「統合法II」と表記する。

「統合法I」では6種の識別器の中から識別率の高い順序に識別器を用いた。この方法は識別率の高いもの同士を組み合わせることで、性能を向上させるという考え方である。

「統合法II」では6種の識別器の中から最も認識率の高い識別器を基準に、その識別器が誤識別した各単語を抽出し、その単語の正答率が高い識別器を選び組み合わせた。この方法は、最も識別率の高い識別器が苦手とする単語を、得意とする識別器を用いて、その弱点を克服し、性能を向上させるという考え方である。

ここで、「統合法I」、「統合法II」は共に識別器を追加するにあたり、組み合わせ後の認識率が低下する直前までに使用した識別器を用いる。

4.手話認識実験

4.1.データセット

認識実験に使用した単語とデータセットを表1に示す。この単語35種は、健康・病気に関する単語で、右手のみを使用した単語である。各単語を各手話者で、学習用に20サンプル、検討用、評価用に10サンプルずつ用意した。検討用データを用いて組み合わせに使用する識別器を選定し、その後、評価用データを用いて得られた識別器の結果を組み合わせ、認識性能を評価した。

表1 認識実験に使用するデータセット

	手話者	サンプル/手話者	合計/単語
学習用データ	3 (A, B, C)	20	60 (3 × 20)
検討用データ	2 (H, I)	10	20 (2 × 10)
評価用データ	3 (D, E, F)	10	30 (3 × 10)

4.2.統合する識別器の選定

学習用データと検討用データを用いて得られた各識別器の識別率を表2に示す。表2の結果から最も識別率の高い識別器はSVMであった。

「統合法I」, 「統合法II」の方法により検討データによる識別結果を組み合わせると、認識性能の変化を調査した結果を図3, 図4に示す。

「統合法I」はSVM, DA, DTの順序で、識別器を3種組み合わせると認識率が最も高く、4種以上組み合わせると認識率が低下した。「統合法II」は識別器SVM, DTの順序で、識別器を2種まで組み合わせると認識率が向上したが、3種目を組み合わせると認識率が低下した。以上の結果から「統合法I」で使用する識別器はSVM, DA, DT, 「統合法II」で使用する識別器はSVM, DTとした。

表2 各識別器の識別率

識別率	学習用データ A, B, C & 検討用データ H, I					
	識別器					
	HMM	SVM	DT	DA	LCM	KNN
	57.1%	78.1%	75.2%	76.0%	74.0%	75.1%

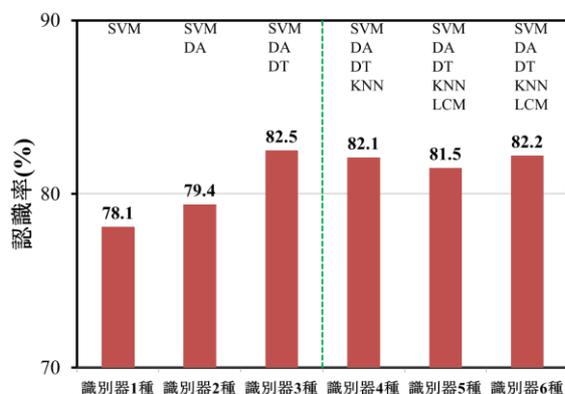


図1 「統合法I」による認識性能の変化

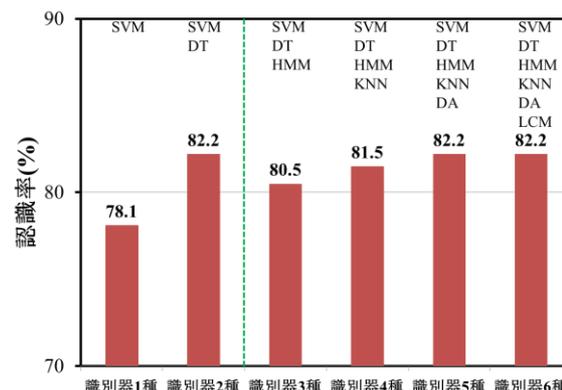


図2 「統合法II」による認識性能の変化

4.3.統合による認識性能

表3に各識別器の1位と3位まで累計した識別率および「統合法I」, 「統合法II」の1位と3位まで累計した認識率を示す。表3から「統合法I」, 「統合法II」の1位と3位累計の認識率は共に各識別器の識別率よりも高いことから、性能が向上することを確認した。そして「統合法I」, 「統合法II」の1位の認識率は76.4%と76.8%であり、また3位累計の認識率は、92.8%と92.3%であり、現状の対象の手話では有意な差はないと思われる。

表3 各識別器と統合方法による認識性能

認識率 (識別率)	学習用データ A, B, C & 評価用データ D, E, F							
	識別器							
	HMM	SVM	DT	DA	LCM	KNN	統合法I	統合法II
1位	53.6%	73.1%	70.1%	72.5%	70.8%	71.8%	76.4%	76.8%
3位累積	79.0%	90.8%	90.1%	90.4%	88.1%	91.5%	92.8%	92.3%

5.まとめ

カラー手袋の色領域から抽出した特徴量を各識別器6種にそれぞれ入力し、各識別器の識別率を算出した。そして各識別器から得られた結果を統合することで性能の向上を試みた。統合方法は2種類提案し、統合方法2種共に単独の識別器の性能よりも高いことを確認した。

参考文献

- [1]Okayasu, Y., Ozawa, T., Dahlan, M., Nishimura, M., Tanaka, H.: Performance enhancement by combining visual clues to identify sign language motions. In: IEEE Pacific Rim Conference, 4 pages. 2017
- [2]小澤辰典 他, ”特徴量と複数の識別器の統合による手話認識性能向上”, HCG シンポジウム 2017, HCG2017-A-8-3, 2017