

指文字認識システムにおける効果的な辞書生成の検討

渋谷 美彩[†] 田中 翔平[‡] 菊池 優介[‡] 山崎 泰幸[‡] 福井 和広[‡]

概要

距離動画像を用いた指文字認識システムにおいて、手本となる正しい手形状を如何に学習するかが重要な課題である。本論文では、システム立ち上げの初期学習に対して、日々の追加学習により、識別性能がどのように向上するかを評価する。

1 まえがき

聴覚障がい者にとって手話は重要なコミュニケーション手段である。手話には見出し語は 4800 語しかなく、これに含まれない人名、地名、外来語、専門用語などは、日本語 50 音および濁音・半濁音に対応した指文字 (74 種類) を用いて表現される。このため手話学習者にとって指文字の習得は必須と言えるが、しかし指文字には手形状の表裏の違いなど初学者が間違いやすい手形状が多く、アニメーションなどを見るだけの一方向型の学習では自己習得が極めて難しい。指文字の自学習手段としては、指文字表や”ゆびもじ練習あいうえお”ソフト [1] のようなアニメーションを見ながら行う方法が一般的である。しかし、このような一方向の学習では指文字を言語として記憶することが困難である。

この課題に対して、筆者らは、学習者が提示した指文字と模範指文字との 3 次元形状の類似度を距離動画像を用いたパターン認識で測定し、その結果に基づいて、インタラクティブな学習システムを開発している。これにより提示した指文字が正しいか、どの指文字と混同され易いかを、学習者自身が直感的かつ大局的に把握でき、それに応じて指形状を適切に修正できる指文字認識システムを開発している [5, 2, 7, 6]。指文字認識システムの処理の流れを図 1 に示す。

このような認識システムにおいて、手本となる指文字 (特定の手形状) を如何に効率良くシステムに学習させるかが重要な課題となる。同じ手形状を呈示しても、距離センサーと手の相対位置関係により、取得できる 3 次元手形状 (距離情報) は大きく変動する。このようにセン

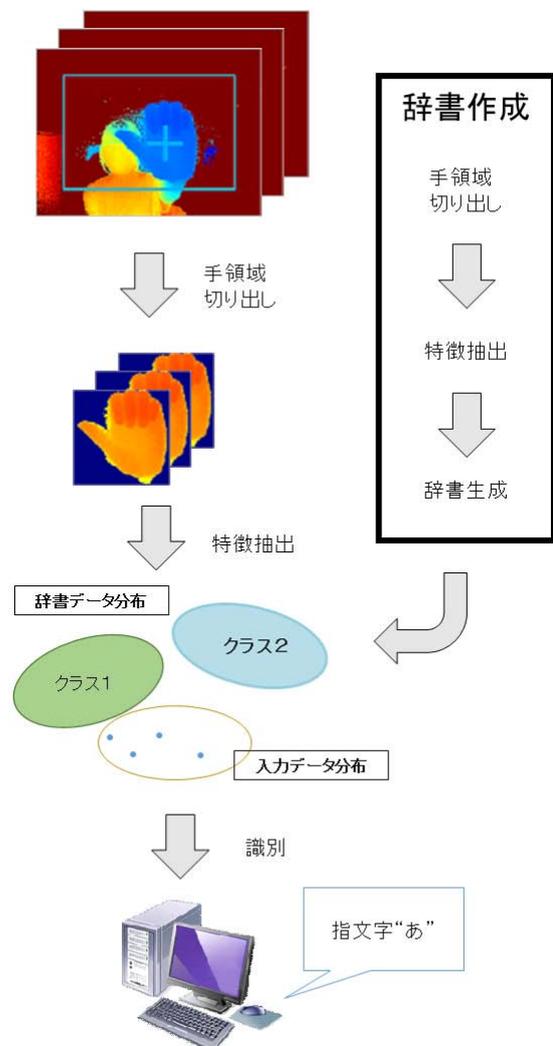


図 1: 指文字認識システムの処理の流れ

[†] 筑波大学 情報学群情報科学類

[‡] 筑波大学大学院 システム情報工学研究科

サーに対して完全に同じ手形状を呈示できない要因として様々な要因が考えられる。例えば、その日の体調や気分により手形状が微妙に変動したり、同じ日においても朝と夕の疲労度合いによっても変動する可能性がある。拘束器具などを用いて、強制的に同じ手形状を呈示させる事も可能だが、指文字認識システムのユーザビリティの観点から、拘束することは好ましくない。

以上を整理すると、上記変動に対応するためには、できるだけ多くのバリエーションを含んだ学習データが不可欠となる。それゆえその収集法やシステム学習法について検討が必要と考えられる。日々の変動や不確定な要因を考慮すると、一度に大量の学習データを収集してシステム学習するより、ある程度の期間において断続的に収集したデータを学習に利用する追加学習が有効と考えられる。

以下では、まず距離動画像を用いた手形認識システムの概要について述べる。次に 7 日間に渡り収集したデータを用いて、追加学習の有効性の検証実験を行い、その結果に対する考察を述べる。

2 距離動画像を用いた手形状認識システム

2.1 システム概要

データグローブ、カラーグローブ、3次元レンジファインダーなどの特殊なデバイスを用いる従来法 [11, 12, 13] では、事前に用意された 3 次元モデルと比較し、それに基づいて学習者に修正指示を出していた。これらの方法の問題点として、特殊デバイスを使用するために学習者の動きを拘束する点、学習者毎に煩雑なキャリブレーションが必要な点が挙げられる。カメラを用いた方法も提案されているが、1 枚の単視点画像に基づくために形状復元の安定性に欠け、微妙な形状差を検出できる精度ではない。さらに特殊デバイス、カメラを用いるいずれの方法でも、学習者へのフィードバックに改善の余地がある。正否の結果表示、あるいは混同された指文字の提示だけでは効果的な自学習は望めなかった。この問題に対して、筆者らは距離動画像から得られる複数の距離動画像セットを用いた識別システムを開発している。距離動画像セット同士を比較することで、安定で高精度な手形状を実現している [5, 2, 7, 6]。



図 2: 実験に用いた指文字の 5 形状

2.2 距離動画像列の KOMSM による識別

本指文字認識システムの識別処理にはカーネル直交相互部分空間法 (KOMSM)[3] が用いられている。KOMSM は線形の直交相互部分空間法 (OMSM)[4] の処理をカーネルトリックを用いて、高次元特徴空間上で行う手法である。OMSM は画像セット同士の類似度を比較する上で有効な方法である相互部分空間法 (MSM)[4] の拡張版である。MSM では、比較すべき 2 つの画像セットを部分空間で表し、2 つの部分空間の成す正準角を類似度とする。

3 追加学習の有効性検証

追加学習の有効性を検証するため 2 つの実験を行った。まず、(1) 追加学習による平均的な識別性能の推移の評価、次に (2) 各学習データの有効性の評価である。

3.1 平均的な識別性能の推移

まず実験に用いるデータと実験方法について述べる。評価対象の指文字には、41 形状の中で動きを伴わず、特に類似した形状を有する 5 形状 (図 2 に示す あ、え、お、さ、は) を選択した。データ収集は、各収集セット毎に 150 枚/形状の距離動画像を収集した。このデータ収集を 1 日に 3 回、7 日間に渡り実施し、合計で 21 データセット (set1~21) を収集した。例えば、set1~3 は初日に収集したデータセットである。収集の際には、手を微妙に動かしながら撮影を行うことで疑似的な多視点データを取得した。今回は追加学習の効果を検証することを第一目的として、被験者は一人というもっともシンプルなケースで評価を行った。入力部分空間は set19~21 (7 日目のデータ) を用いて生成した。辞書部分空間に関しては、まず set1 のみから辞書を生成し始め、逐次 1 セットずつ辞書に加えていきながら、辞書部分空間の精度を高めていった。ここで、辞書部分空間と入力部分空間の次元は 15 および 3 とした。

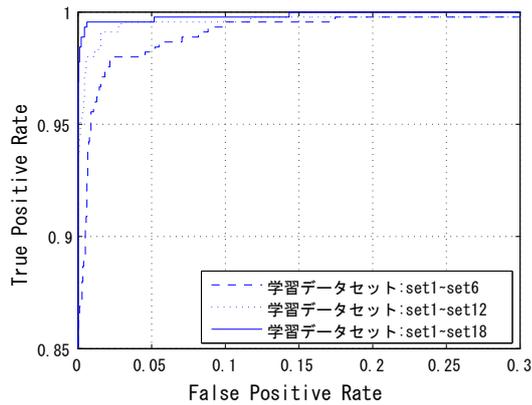


図3: 追加学習に対するROC特性の変化

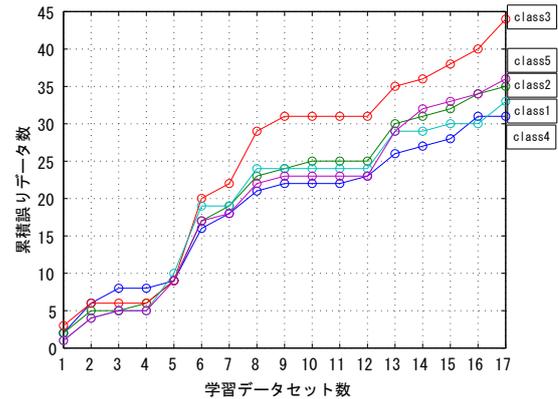


図5: 新規学習データ追加時の誤りデータ数の変化

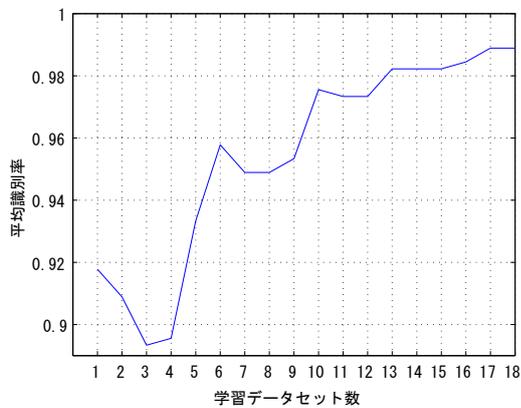


図4: 追加学習に対する平均識別率の変化

表1: 追加学習に対するERとEERの推移

	1	3	6	12	18
ER (%)	8.22	10.67	4.22	2.67	1.11
EER (%)	3.78	5.08	2.19	1.56	0.64

この実験から得られた性能推移の結果を図3と図4に示す。図3は追加学習毎のROC特性の変化を代表して3本表している。ROC曲線はクラス間の重なりが少ない程、グラフの左上に近づいていく曲線であるため、図3では追加学習を行う毎に識別性能が向上している様子が見て取れる。また、図4は最大類似度を用いて識別した全クラスの平均識別率の推移である。節目での具体的な値を表1に示す。追加学習の経過において一部識別率の低下が見られたが、全体的には向上している傾向が見られ、ROC特性と同様に識別率の改善も確認できる。以上により、追加学習を行うことで平均的な識別率は徐々に向上すると考えることができる。

3.2 各学習データの有効性

先の評価実験では、平均的な識別率の改善度合いを評価した。この実験では追加学習において、それまでに生成された辞書に対し、新規追加する学習データセットをテストデータとして識別、これに基づき追加する学習データの有効性を評価する。もし、新規追加する学習データセットの識別を誤るならば、それまでに生成した辞書での識別は不完全であることを意味し、新規追加する学習データが有効であると言える。この実験の目的は学習データセットの有効性を評価することで、今後の追加学習の効率を高めることにある。以下に具体的な実験方法を示す。

- (1) 時系列に取得した最初のデータセット (set1) により辞書生成。
- (2) 辞書生成に用いた次のデータセット (辞書に新規追加したい set2) をテストデータとして識別。
- (3) 各クラスごとに誤識別したデータ数を計数。
- (4) テストデータとして用いたセットを新規追加し辞書生成。
- (5) 以降、(2)~(4) を学習データとして用いた set18 まで繰り返す。

以上の方法で計数した累積誤りデータ数の変化を図5に示す。図5では、横軸が辞書生成に用いたデータセット数、縦軸が累積誤りデータ数を表している。図中において、set8~set12のようにグラフが平坦になっている箇所が存在する事がわかる。つまり、これらのセットには識別誤りデータが存在せず、全て正解だったということが言える。誤り識別データが無いということは、新規

追加しようとする学習データセットの分布がそれまでに作成した辞書に含まれている事を示唆し、例えば set9～set12 の学習データセットは辞書作成において冗長であると言える。これら冗長な学習データセットを辞書に追加しないことで、辞書のデータ量削減および、辞書データ数に計算時間が依存する KOMSM の高速化が期待できる。

また、図5において各クラスともに識別誤りデータ数の推移はそれほど変わらないことが見て取れる。しかしながら、class3 の識別誤りデータ数が他クラスに比べ多い傾向が読み取れる。このことから、クラス毎の辞書改善度合いにバラツキがあることがわかった。したがって、学習を強化したほうが良いクラスの選定が可能であり、それによっても追加学習の効率を高めることが可能である。

4 まとめ

距離動画像を用いた指文字認識システムにおいて、手本となる正しい手形状の学習方法について基礎評価を行った。その結果、一定期間の追加学習により識別性能が向上する事が確認できた。さらに累積誤りデータ数に基づいて各セットの追加学習における有効性を評価した。今後、これを用いて不要な学習データセットを追加対象から除く機構を実現したい。またクラス毎の辞書改善度合いにバラツキがある事がわかった。このような違いを考慮した追加学習法も検討したい。

謝辞

本研究は科研費(25282173)の助成を受けたものである。

参考文献

- [1] 原田泰, 生田目美紀, 豊田由美, 佐藤淳, "「指文字練習あいうえお」: 手話学習導入に焦点を絞った初心者のための指文字学習支援プログラム (Information Design)", デザイン学研究作品集 9(9), pp.12-15, (2004).
- [2] 高林大輔, 大川泰弘, 瀬戸山浩平, 田中陽土, 加藤伸子, 岡崎彰夫, 福井和広, "フィードバック機能を備えた指文字練習システムの試作", 信学技報, Vol. 112, No. 483, Human Information Processing, pp.79-84, (2013).
- [3] Y. Ohkawa, K. Fukui, "Hand Shape Recognition based on Kernel Orthogonal Mutual Subspace Method", Machine Vision and Applications (MVA2009), pp.122-125, (2009).
- [4] 福井和広, 山口修, "制約相互部分空間法と直交相互部分空間法の比較", 部分空間法研究会 Subspace2006, pp.63-71, (2006).
- [5] 田中陽土, 高林大輔, 瀬戸山浩平, 大川泰弘, 加藤伸子, 福井和広, 岡崎彰夫, "Kinect を用いた指文字練習システムにおける識別処理の検討", 第096回ヒューマンインタフェース学会研究会, pp.29-34, (2013).
- [6] 瀬戸山浩平, 高林大輔, 田中陽土, 大川泰弘, 加藤伸子, 福井和広, 岡崎彰夫, "Kinect を用いた指文字練習システムにおけるリアルタイムフィードバック表示の検討", 第096回ヒューマンインタフェース学会研究会, pp.35-40, (2013).
- [7] 福井和広, 大川泰弘, 加藤伸子, 岡崎彰夫, "複数視点からの画像を用いた手形状識別", 信学技報, vol. 112, no. 483, HIP2012-85, pp. 59-64, (2013).
- [8] 福井和広, 山口修, "経時変化を抑制する顔辞書更新法", 信学技報, PRMU99-25, pp.43-50. (1999).
- [9] K.Fukui, O.Yamaguchi "The Kernel Orthogonal Mutual Subspace Method and its Application to 3D Object Recognition", Asian Conference on Computer Vision, Vol. 2, pp.467-476, (2007).
- [10] S.Taylor, N.Cristianini, "Kernel Methods for Pattern Analysis", Cambridge University Press, (2004).
- [11] 島田貢明, "マルチメディアと画像認識を活用した初心者のための指文字学習ソフトの試作", 仁愛女子短期大学研究紀要, (2012)
- [12] 田畑慶人, 黒田知宏, 眞鍋佳嗣, 千原國宏, "手型認識を用いた指文字教育システム", 教育システム情報学会誌, Vol.18, No.2, pp.172-178, (2001).
- [13] 三宅太一, 若月大輔, 内藤一郎, "距離画像を用いた動きをともなう指文字認識に関する基礎的検討", 筑波技術大学テクレポ, Vol.20, No1, pp.7-13, (2012).