

視覚障害者のエレベータ操作における筆記動作認識

Motion Recognition for Blind Person in elevator

望月 広一† 西村 広光‡ 富川 武彦‡
Hirokazu MOCHIDUKI Hiromitsu NISHIMURA Takehiko TOMIKAWA

1. はじめに

現在、キーボード、マウス、音声入力など多彩な入力インタフェースが存在するが、視覚障害者にとって利用しやすいものは非常に少ない。その中でも、既存のエレベータ操作に利用されているボタン押下型のインタフェースは、ボタンの位置がわからないときには、利用することができないなど多くの問題が存在する。

本研究では、このような状況でも利用可能な入力インタフェースについて検討を進め、筆記動作認識法を考案し、その中核である動作認識部を試作・検討し、有効性について検討した。

なお、このような認識系を構築するに当たっては、オンラインの時系列の動作認識で優れた能力が知られているHMM等を利用することや、複数の認識系を複合してより高い性能を発揮するような認識モデルにすることが考えられる。本研究では、複合認識系を利用したモデルでの認識を考えているが、その第一歩として、階層構造のニューラルネットワークを利用して、提案手法の認識性能の検証を行うことを目的とした。

2. 筆記動作認識

本論文で記す筆記動作認識とは、カメラを利用して、空間にエレベータのボタン操作に相当する文字を描画する動作を認識する手法である。本研究では図1のJIS規格にあわせたエレベータ環境を作成し、赤い手袋をつけて空間に英数字を描く動作をカメラで撮影し、認識を行うこととした。

具体的には、筆記動作にあわせて動画像を採取し、採取した動画像をフレームごとに切り出し、特徴抽出を行い、フレーム間の特徴の変化をニューラルネットワーク(以下NNと記す)の入力として、認識を行うこととした。

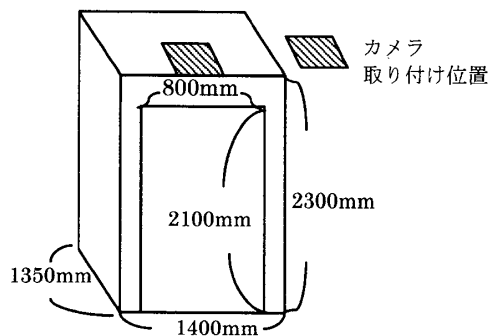


図1 擬似エレベータ

3. 処理の流れ

採取した動画をフレーム毎に切り出し、各フレームで手袋領域の重心を抽出し、フレーム間の手袋の移動をCode変換し、扱いが容易な3層の階層型NNを利用し認識システムを構築し、提案認識手法の有効性を検証した。

3.1 手袋領域の重心抽出

図2のように採取した動画像をフレーム毎に静止画に変換した後、画像から手袋領域の抽出を試みた。しかしRGB表色系では赤色手袋領域抽出が失敗することが多かった。そこでYIQ表色系を利用して赤色手袋領域抽出を行い、重心を求めることとした。

また、ノイズによる重心のずれを補正するためにラベリングを利用して、最大連結領域を求めてそれ以外を削除し重心を求めることとした。



図2 サンプル画像

3.2 Code列変換

連続する2フレーム間で先に求めた手袋領域の重心が移動する方向を図3に従って8方向にCode化し等間隔29フレームで長さ28のCode列に変換し、NNの入力とした。

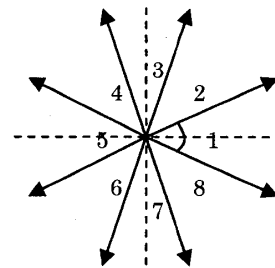


図3 8方向のコード化

† 神奈川工科大学大学院 工学研究科

‡ 神奈川工科大学 情報学部 情報メディア学科

4. 実験

第一段階として、マルとバツの2種類の筆記動作を認識、さらにエレベータ操作に必要な0から9及び、B, F, Rの筆記動作認識実験を行った。

4.1 2種類の筆記動作認識実験

マルとバツの筆記動作各160サンプルを学習データセットに、40サンプルを検査用データセットとし認識実験を行った。なお、認識に用いるNNの入力ユニット数、中間ユニット数、出力ユニット数はそれぞれ28, 12, 1とする。また学習効率を0.75, 学習回数を5000回とした。認識結果を表1に記す。

表1 2種類の認識結果

被験者数	1人	5人
認識率	97.00%	84.75%

被験者1人のデータでは約97%、被験者5人のデータでは約85%の認識率を得ることができた。被験者1人より被験者の5人の方が認識率は落ちるが、書き順が異なる場合があるために認識率が低下したと考えられる。

しかし、被験者1人、被験者5人の実験で共に認識率80%以上を得ることができ、筆記動作2種類の場合では提案手法の有効性の確認ができたといえる。

4.2 7種類の筆記動作認識実験

エレベータの設置が必要となる31メートルを超える建築物は、5階建て又は6階建ての建築物になることが多い。そこで6階建てのエレベータを想定し、1から6の数字と、地下を表すBを認識する実験を行った。被験者1人につき各筆記動作20サンプルを学習データセットとし各10サンプルを認識用データセットとする。なお、本実験では中間層のユニット数を14に変更した。認識結果を表2に記す。

表2 7種類の認識結果

被験者数	1人	2人	3人
認識率	61.43%	57.14%	49.52%

被験者が1人のときの認識率は約61%、2人の場合約57%、3人の場合では約50%となった。被験者の増加に伴い認識性能が低下した。これは、4.1節の実験と同じく、被験者による書き順の違いが影響していることが考えられるが、2種類の動作から7種類の動作とクラス数が増加したことで、約30%認識性能が低下したことは大きな問題である。

そこで、認識率をそれぞれの文字で調べたところ、ここまで認識率が落ちている原因が6の筆記動作の認識率だけが18.8%しかなく、他の筆記動作の半分以下の認識率となっていることが判った。この原因を検討したところ、次の2つのことが影響していると考えられる。

1つは、6の筆記動作が4の筆記動作と類似しているという点である。認識結果を見てみると6の動作60データ中16データが4の筆記動作と見なされている。このことから、4と6の筆記動作をより判別できるような特徴抽出法を今後検討する必要があると考えられる。

2つ目に認識結果集計時の優先順位である。今回の実験ではNNの出力が同じ場合に結果を1, 2, 3, 4, 5, 6, Bという優先順位で出力することとした。そのため、NNの出力が4と6の筆記動作で同一であった場合には、優先順位の高い4が認識結果となってしまったため、優先順位の低いクラスで認識結果が低下していると考えられる。このことは、学習データ数がまだ十分な数収集できていないことの影響もあると考えられるため、今後データ収集を進め、サンプル数を増加させて検証を進めていきたいと考える。

4.3 13種類の筆記動作認識実験

最後に、本提案システムの拡張性について検討するため、7階建て以上の建物のエレベータを想定し、0から9の数字及び地下を表すB, 地上の階を表すF, 屋上を表すRを認識する実験を行った。被験者1人につき各筆記動作20サンプルを学習データセットとし各10サンプルを認識用データセットとする。NNの各ユニット数は7種類の筆記動作認識実験を行ったときのものと同じである。認識結果を表3に記す。

表3 13種類の認識結果

被験者数	1人	2人	3人
認識率	42.30%	37.30%	31.28%

被験者が1人のときの認識率は約42%、2人の場合約37%、3人の場合では約31%となった。被験者数増加に伴う認識率低下は4.2節の7種類実験とはほぼ同じ傾向がみられた。

ここでも筆記動作の種類数を13種類にしたことによつて約20%落ちた。4.2節の実験と本実験から、現在の手法では認識するクラス数を増加させると大きく認識率が低下することがわかる。この点については、今後学習サンプル数を増加させてさらに検証を進め、より効果的な学習モデルを構築する方法について検討を進めていきたいと考える。

5. まとめ

4節で行った実験の結果より、被験者数、クラス数の増加した場合の認識率低下に関しては、今後学習サンプル数を増加させ、検証を進めていきたいと考える。

また、4と6や5と8のように類似する筆記動作に関して、より判別できるような方法を検討する必要がある。具体的には、新たな特徴抽出方法の検討、より効果的な学習モデルの構築が考えられる。

参考文献

- [1]佐藤 大志, 佐藤 雅之:ニューラルネットワークを利用したオフライン手書き文字認識システムの試作, pp4-5, 神奈川工科大学工学部電気電子工学科卒業論文, (2003)
- [2]Bahman, Kermanshahi, : 知能工学4 ニューラルネットワークの設計と応用, pp1-3, 株式会社昭晃堂, (1999)
- [3]Richard, O., Duda, Peter, E. Hart., David, G. Stork. 尾上守夫 訳: パターン識別, pp288-289, 株式会社新技術コミュニケーションズ