

# ニューラルネットワークを利用した筆記動作認識

望月 広一<sup>1</sup>・西村 広光<sup>2</sup>・富川武彦<sup>2</sup>

1 大学院工学研究科情報工学専攻(s045829@cce.kanagawa-it.ac.jp)

2 情報メディア学科({nisimura,tomikawa}@ic.kanagawa-it.ac.jp)

## Writing Motion Recognition Using Neural Networks

Hirokazu Mochiduki<sup>1</sup>, Hiromitsu NISHIMURA<sup>2</sup>, Takehiko TOMIKAWA<sup>2</sup>

### Abstract

There have been introduced some human interfaces in the past for partially sighted or blind person regardless of requiring its performance with reasonable price. We proposed a guide system to help handicapped persons to control elevator as a human aid.

In our system, it recognizes a motion of hand writing as a gesture in the elevator based on neural networks. All we need to build up this system are a personal computer and a low priced web-camera. This paper describes an application of writing gesture for handicapped person using simple, but not problem oriented algorithms in a sense.

Our experiments result in the recognition rate of roughly 90% in a limited 2 class classification and 40% in a limited 13 class classification.

We have learned that our system doesn't give an appropriate response of recognition rate with increasing the number of classes, which must be improved in our future experiments.

Key Words : human interfaces, writing motion, character recognition,

### 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、これらは障害を持たないユーザにとっての利便性を追求したものであり、それらが同じように障害者にとって利便性が高いとはいえない。特に視覚障害をもつ人が生活を行う上で入力を機器に求められた際に、意図した操作が困難であることが多い。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、乗りなれていないエレベータを利用した場合、ボタン付近に点字があっても、ボタン自体の場所が分からないため利用できないという問題がある。また、音声認識を利用して、対話型のインタフェースを構築することも考えられる。しかし、通常の会話がある中で、どのようにしてエレベータ操作の命令を区別するのか、複数同時発音時の音声認識が非常に困難であることなどの理由により、実用的なインタフェース構築は難しいと考えた。

そこで本研究では、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、雑音に強く導入コストが抑えられる方式として、安価なWebカメラを利用して、人間の動作を認識することでエレベータ操作を実現することを検討する。さらに、人間の動作を認識するに当たり、手話は後天的障害者にとっては利用しにくいと考えた。そこで、通常の文字筆記動作を認識する手法について検討を進め、中核となる認識部を試作し、本提案手法の有効性について検討する。

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にする。

## 2. エレベータ操作インタフェースの検討

エレベータ操作のインタフェースを構築する場合、様々な手法が考えられる。現在多く採用されているボタン押下型インタフェースのほか、音声入力によるインタフェース、エレベータ利用者の動作をカメラに入力しそれを認識させるインタフェースがある。

### 2.1 エレベータ操作における限定条件

本研究で対象としたエレベータ操作について、インタフェース構築における限定条件について検討を進める。

現在利用されている視覚障害者の利用を考慮したエレベータ操作のインタフェースとして、点字を利用したもの、音声認識を利用したものが存在する。

点字を利用したものは、既存のエレベータ操作パネルに点字シールを貼り付けたもので、安価に導入できるものであるが、エレベータ操作パネルの位置が規格化されていないこともあり、どこに点字があるのか分からないため、視覚障害者にとって操作しやすいものと言えない。

次に、音声認識を利用したものは、視覚を必用としないため視覚障害者にとって操作はしやすいと考えられる。しかし、通常会話がある不特定多数の中でエレベータ操作の命令だけを区別することは困難であり、実装コストも決して安価ではない。

そこで本研究で開発を行うにあたり、次のような利用を想定して利用可能なインタフェースの構築を検討した。

1. 雑音による影響を受けないこと
2. 複数人が同時に操作することが可能であること
3. 既存エレベータに低コストで設置できること
4. 点字という特殊なスキルを利用しないこと

この条件下でのインタフェース実現を検討したところ、低解像度のカメラ画像を利用して、通常の文字を空間に書く動作を認識することで実現することを検討した。

なお、以下論文ではこのカメラ画像を利用した空間に図形を描画する動作を認識することを筆記動作認識と記述する。

### 2.2 筆記動作を利用したエレベータ操作

カメラを利用して動作を認識する際に、どのような動作を認識するかが問題となる。本研究では、視覚障害者が容易に利用可能なエレベータ操作を行うという条件の下で、カメラをエレベータ天井に取り付け、カメラが天井にあることを音声案内し、そこから動作を入力、認識させるというインタフェースを考案した。

カメラの取り付け位置は、エレベータ横部に取り付けた場合、視覚障害者の向きや位置によってカメラの場所が変わってしまい複数のカメラを設置し制御する必要がある。そこで、カメラをエレベータ天井に取り付けることとした。これによって、音声案内で天井にカメラがあることを伝え、視覚障害者の向きに関わらず天井に向かって筆記動作を行うこととする。さらに、カメラを天井に1つ設置すれば、複数人の操作時に陰になり操作が見えなくなることを避けることができ、全体を見渡せるためカメラを複数設置する必要がない。

また、カメラを横部に設置した場合、混雑した状況下では身振りをすることが困難であると考え、カメラを天井に設置したことによって身振りも天井に向かって行うことから、多少の混雑では問題ないと考えた。加えて、天井につけたカメラに向かって筆記動作を行うため、カメラとの距離が近く、小さな筆記動作でも比較的大きな画像として取得することができ、利用面・システム開発面からも有利であると考えた。

### 2.3 データ採取

動画画像から数点の特徴を抽出し筆記動作の認識を行う。本研究では動画画像の特徴として軌跡の方向ベクトルを入力値として利用する。

本実験で使用する入力装置の仕様を以下に記す。表1の仕様のWebカメラを図1の擬似エレベータ入り口上部2300mmの位置に被験者を見下ろすように設置する。

表1 Logitech Quickcam Web Webcam の主な仕様

インタフェース	USB
フレームレート	最大 30 フレーム/秒
有効最大画素数	640×480

擬似エレベータの寸法は、幅1400mm、高さ2300mm、奥行き1350mmとする。この寸法はJIS標準規格11人乗りのエレベータと同じものである。<sup>1)</sup>

被験者は、擬似エレベータの中に赤い手袋をつけて入り、カメラに向かって指定の筆記動作を行う。この筆記動作を動画として保存し、フレームごとに352×288の静止画に変換する。特徴を抽出後に認識を行う。なお、このときの静止画1枚あたりの大きさは約300KBである。

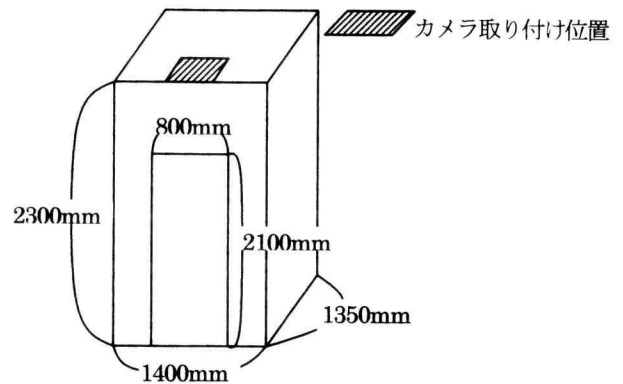


図1 擬似エレベータ

## 3. 処理の流れ

本研究では、筆記動作をカメラで撮影して、動画画像ファイルを生成し、フレーム毎に静止画に変換する。その静止画より特徴を抽出しニューラルネットワークを使用して認識を行う。

筆記動作認識システムの流れを図2に記す。



図 4 サンプル画像

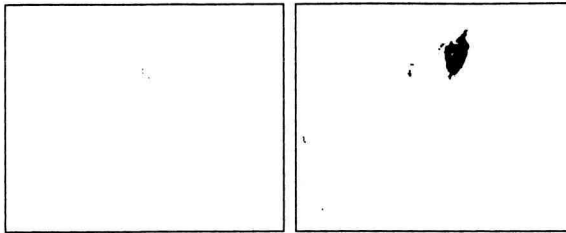


図 5 RGB 表色系 (左) と YIQ 表色系 (右) による手袋抽出

図 5 から分かるように、RGB 表色系を用いた左の画像より YIQ 表色系を用いた右の画像の方が正確に手袋の領域を抽出することが分かった。

### 3. 3 ラベリングによるノイズ除去

抽出した手袋領域より重心を求める。重心を  $g$ 、全画素数を  $n$  とすると以下の式で表される。

$$g(x, y) = \left( \frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i \right) \quad (2)$$

しかし手袋領域以外にもノイズが発生してしまい図 6 のように重心のずれが発生する場合がある。

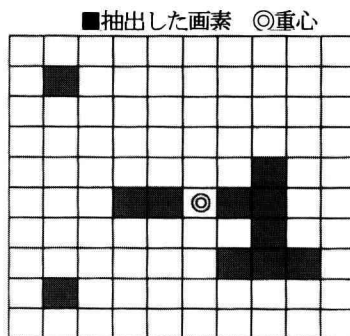


図 6 ノイズがある場合の重心位置

ここでラベリングを用いて抽出した画素のラベル付けを行うと次のように番号付けすることができる。

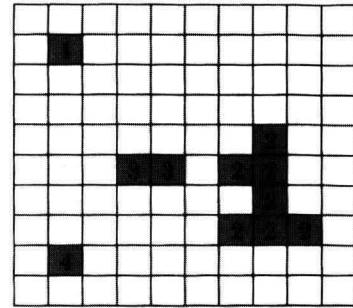


図 7 ラベル付け

ラベル付けした後、最大面積の連結領域以外を削除する。ここでは最大連結領域は 2 とラベル付けされた連結領域のため、それ以外をノイズとして削除する。削除した後に重心を求めることと図 8 のようにノイズを除去した後の重心の位置を求めることができる。

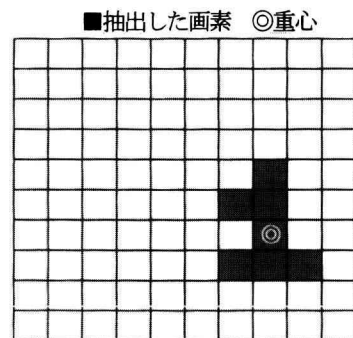


図 8 ノイズを除去した場合の重心位置

しかし、被験者が赤い服を着ている場合に、着ている服の方が最大連結領域として判断されてしまう。その結果、手袋領域がノイズとして削除され、着ている服の重心を求めてしまう。

また、最大連結領域だけを利用しているため、抽出した領域が 2 つの大きな連結領域として求められた場合には、正しく重心を求めることができないが、本研究でわずかな位置ずれは認識に大きく影響を与えないように特徴抽出を行っているため、今回の実験の範囲では結果に影響しなかった。

### 3. 4 Code 化

静止画から手袋領域の重心の位置である  $x$  座標と  $y$  座標を抽出後、次の静止画の重心の位置がどの位置にあるかを図 9 の Code 表に当てはめて移動方向ベクトルの方向を Code 化する。この Code 表は表 3 に示す移動方向ベクトルを方向別に数値化するものである。

方向ベクトルのコード化の際には、移動方向単位ベクトル化手法である freeman-code を利用することが多いが、本研究で利用したものはそれを 22.5 度傾けたものである。これは、筆記動作には水平、垂直方向移動することが多いため、水平・垂直方向に手を動かした際の微小な上下左右の揺れがコードに反映されないように考慮したものである。

この移動方向ベクトル化により、1 つの筆記動作が同時間で完了した場合に、移動方向変化をコード変化で表し、移動速度変化を連続コード列の長さ変化で表すことができる。

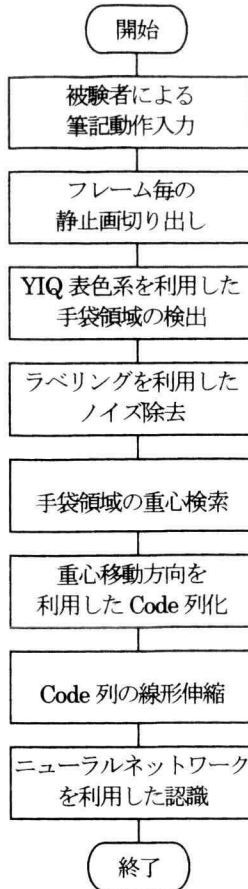


図 2 主な処理の流れ図

フレーム毎の切り出しには、TMPGEnc<sup>2)</sup>というフリーソフトを使用した。現段階で各動作のフレーム切り出しは、目視により手動で行っている。図 3 に「○」の記号を描画した動画を静止画に切り出した画像を示す。図 3 は、切り出したフレームをフレームごとに並べたものである。図 3 から分かるように、本研究で対象としている画像は、低品質なものを想定し、動作を表す手の領域がぼけている場合が多いものである。

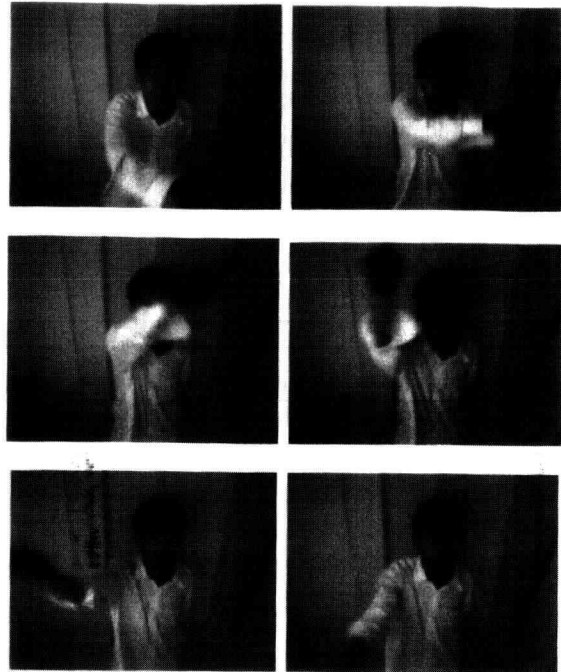


図 3 1 動作の静止画

### 3. 1 静止画の切り出し

利用した表 1 のカメラは、最大秒間 30 フレームを記録することができる。本研究では、秒間 30 フレームで筆記動作の動画を記録し、フレーム毎に静止画に変換し、特徴を抽出することとした。

筆記する文字によって筆記時間が異なるため、静止画の枚数も変化する。表 2 に本研究で採取した代表的な筆記動作について、それぞれ何フレームであるかを示す。表 2 から分かるように、1 の筆記動作は約 31 フレームであるのに対し、0 の筆記動作は約 46 フレームと非常に差異が大きい。この筆記動作ごとの動作フレーム数の差異は 3. 5 節で述べる Code 列の線形伸縮で補正して利用することとした。

表 2 各静止画の平均枚数

動作の種類	静止画の枚数
0 の動作	46
1 の動作	31
2 の動作	39
3 の動作	32
4 の動作	34
5 の動作	39
6 の動作	32
7 の動作	30
8 の動作	32
9 の動作	32
B の動作	40
F の動作	33
R の動作	35

### 3. 2 YIQ 表色系を利用した手袋領域の抽出

筆記動作の軌跡の方向ベクトルを求めるために、まず各静止画より被験者の手袋の重心を求める。そのために手袋領域を抽出しなければならない。そこで RGB 表色系を用いて R 要素、G 要素、B 要素の抽出範囲の閾値を手動で設定し図 4 の静止画より手袋領域の抽出を試みた。しかし、RGB 表色系のままではどの静止画も図 5 左のように手袋領域はほとんど抽出されず失敗することが多かった。

そこで、手袋領域抽出の際に RGB 表色系を用いるのではなく、YIQ 表色系を用いて手袋領域の抽出を行った。この YIQ 表色系は色と明るさを分離した表色系であり、Y は輝度、I はシアン、Q はマゼンダを表している。RGB 表色系から YIQ 表色系への変換式は以下の通りとなっている。<sup>3)</sup>

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

この YIQ 表色系を用いて、RGB 表色系で行ったときと同様に数枚の静止画から  $10 \leq Y \leq 50$ ,  $10 \leq I \leq 70$ ,  $0 \leq Q \leq 30$  の閾値処理を行い、2 値画像として手袋領域を検出する。抽出した画像を図 5 右に記す。

表 3 図 9 の Code 対応表

連続する重心の移動角度	対応する Code
337.5 度以上 22.5 度未満	1
22.5 度以上 67.5 度未満	2
67.5 度以上 112.5 度未満	3
112.5 度以上 157.5 度未満	4
157.5 度以上 202.5 度未満	5
202.5 度以上 247.5 度未満	6
247.5 度以上 292.5 度未満	7
292.5 度以上 337.5 度未満	8

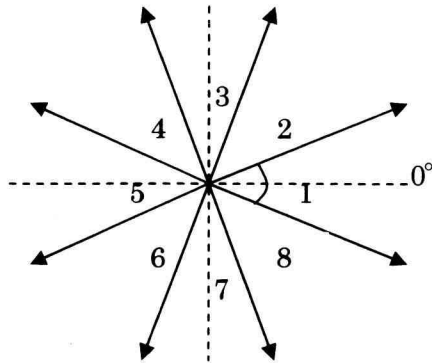


図 9 8 方向の Code 表

具体的に 3 枚の連続する静止画の重心が図 10 のように求められたとする。すなわち、重心が左上、左下、右下と移動したとき、図 11 のように重心移動を方向ベクトルで表すことができる。1 から 2 の移動方向ベクトルは下向きであるので、図 9 の Code 表を利用してコードが 7 と求められる。同様に 2 から 3 への移動方向ベクトルは 1 と求められる。

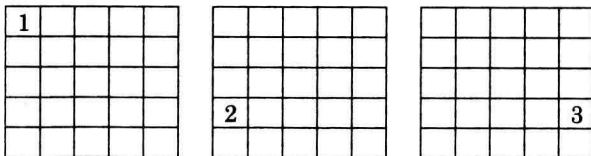


図 10 重心を抽出した連続する画像

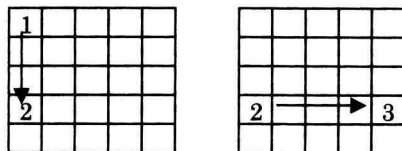


図 11 3 枚の重心の移動方向

このとき、図 12 のように 4 枚目の静止画の重心の位置が 3 枚目の重心の位置と同じだった場合、移動方向が無い場合 Code 化できないが、静止を表すコードを 0 として表すこととした。

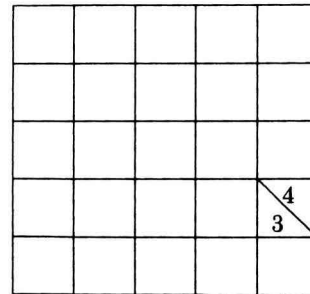


図 12 3 枚目と 4 枚目の重心の位置

### 3. 5 Code 列の線形伸縮

3.4 節で求めた Code 列をニューラルネットワークの入力値とするが、3.1 節で説明したように、各筆記動作をすべて Code 列化した場合に、この Code 列の長さは筆記動作ごとに大きく異なる。そこで Code 列の線形伸縮を行い、コード列の長さを 28 に揃えることとした。

線形伸縮には、図 13 に示す長さ  $m$  の Code 列を図 14 に示す長さ  $n$  のコード列に変換するとき、コード列 B の  $j$  番目のコードを(3)式を用いて求める。

$$j = i \left( \frac{n}{m} \right) \quad (3)$$

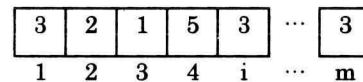


図 13 長さ  $m$  の Code 列 A

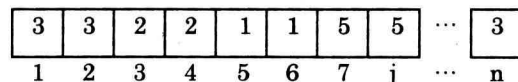


図 14 長さ  $n$  の Code 列 B

## 4. ニューラルネットワーク

求めた Code 列から筆記動作を認識する際に、本研究では階層型のニューラルネットワークを使用した。

### 4. 1 階層型ニューラルネットワークの原理

階層型ニューラルネットワークは、入力層、中間層、出力層により構成されているニューラルネットワークである。本研究ではこの階層型ニューラルネットワークのうち、入力層 1、中間層 1、出力層 1 の単純な 3 層型のニューラルネットワークを利用することとした。

階層型ニューラルネットワークでは、入力層で受け取った信号を中間層、出力層へと伝達していく前向き演算と、前向き演算で出力された信号を元に、出力層、中間層の荷重を修正していく後ろ向き演算の 2 段階の過程により、入力信号と教師信号から最適な結合荷重を求めることで学習を行う。認識は学習したニューラルネットワークの入力層に道信号を与え、その出力値を判定することで行う。

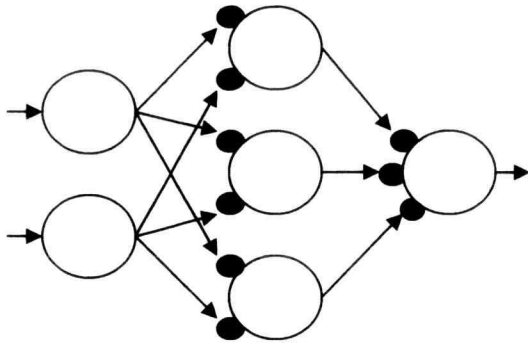


図 15 階層型ニューラルネットワークの例

4. 2 単一ニューラルネットワークによる認識モデル

4.1節で解説したニューラルネットワークは、通常2クラスの認識を行うモデルであるため、複数クラスの認識を行う場合には、認識モデルの構築方法を工夫する必要がある。

最も基本的な複数クラスの認識モデル構築法は、単一のニューラルネットワークの出力ユニットの数をクラス数だけ用意する方法である。すなわち、3種類の動作を認識させる場合、データが動作1のときには出力ユニットの値を1.0, 0.0, 0.0とし動作2のときには出力ユニットの値を0.0, 1.0, 0.0とする。動作3のときもこれにしたがって出力ユニットの値を決定する。

認識手法は、出力ユニットの値でどの動作であるかを判別する。具体的には、事前に学習させておいた単一ニューラルネットワークに未知動作を入力したときの出力値が図16(a)の場合、出力ユニット1の出力値が1番大きいいため動作1と判定する。

また、図16(b)の場合、出力ユニット3の値が1番高いので動作3と判定する。

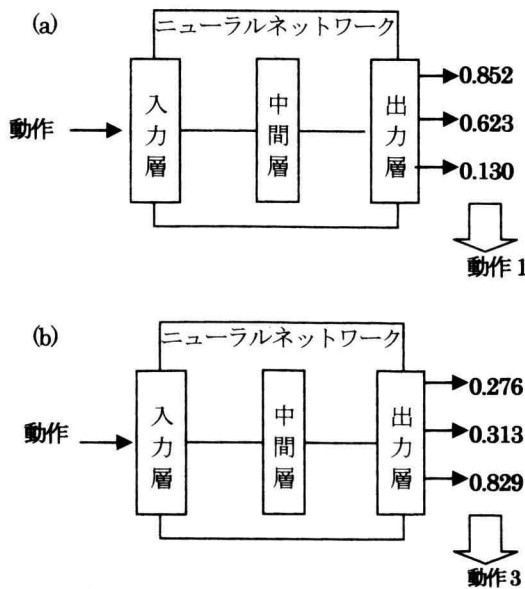


図 16 単一ニューラルネットワーク認識モデル

4. 3 複数ニューラルネットワークによる認識モデル

我々の研究室では、文字の認識において複数クラスの認識にニューラルネットワークを利用する際に、複数クラスを同時に認識する4.2節のように1つのニューラルネットワークを利用して認識モデルを構築するよりも、2クラスずつ識別する認識対象のクラス数だけのニューラルネットワークを利用して認識モデルを構築するほうが、高い認識性能を実現できることを確認している。<sup>5)</sup>

ニューラルネットワークを筆記動作の種類数分だけ構築し認識を行う。複数のニューラルネットワークによる動作認識モデルを図17に記す。

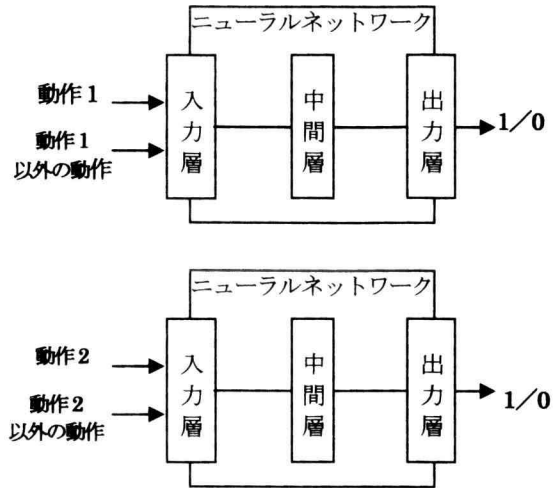


図 17 複数ニューラルネットワークによる動作認識モデル

4. 4 複数 NN を利用した筆記動作の学習

複数の筆記動作を学習させる場合、真のデータと偽のデータを交互に読み込ませて学習させる。これは出力ユニットが一方に偏り、データが正確に学習されないのを防ぐためである。

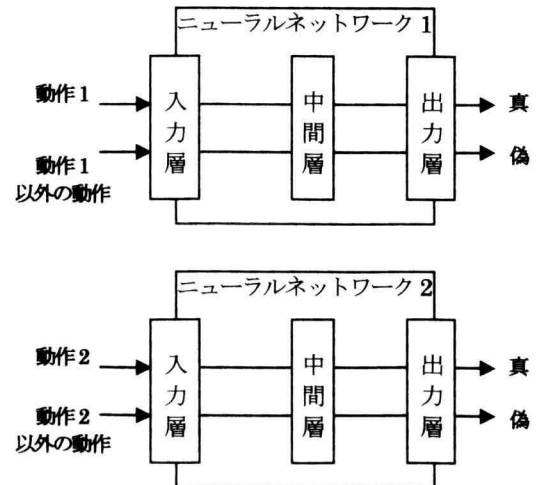


図 18 複数動作におけるニューラルネットワークの学習方法

4. 5 複数 NN を利用した筆記動作の認識

複数のニューラルネットワークで出力ユニット数が1の場合、出力ユニットの値が1番高いものを認識結果とする。

図 19 の 3 つのニューラルネットワークで動作[1-3]として正しいものかを判定させる場合、3 つのニューラルネットワークの出力値が最も大きいものを認識結果とする。図 19 では、0.981 を出力したニューラルネットワーク 1 に該当する、動作 1 を認識結果とする。

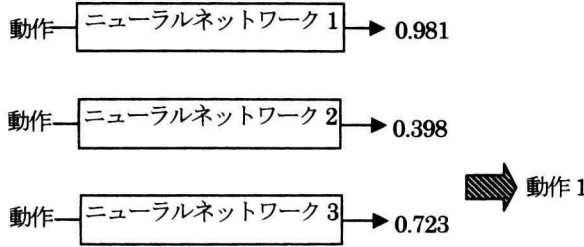


図 19 動作 1 の認識例

5. 実験

本研究ではマルとバツの 2 種類の筆記動作、1 から 6 及び B の 7 種類の筆記動作、また 0 から 9 及び B, F, R の 13 種類の筆記動作を認識する実験を行った。

なお、この 3 種類の実験を行う前にニューラルネットワークの入力層のユニット数、中間層のユニット数、出力層のユニット数、学習回数を決める必要がある。そこで予備実験を行った結果、入力層のユニット数を 28、中間層のユニット数を 12 及び 14、出力層のユニット数を 1、学習回数を 5000 回とした。

試作システムにおいて、ニューラルネットワーク認識系の処理時間を調べたところ、表 4 の環境下で認識時間 0.13ms と非常に高速であり、実時間処理に耐えるものであることを確認した。

表 4 実験環境

CPU	Pentium4 2.40GHz
メモリ	512MB
OS	WindowsXP Professional
開発言語	Visual Studio 6.0

5. 1 2 種類の筆記動作認識実験

提案手法である筆記動作認識の有効性の確認を行うため、マル「○」とバツ「×」の 2 種類の筆記動作の認識実験及び 1 動作あたりの認識時間の計測を行った。

実験方法はマルとバツの筆記動作合計 320 サンプルを学習用データ、80 データをテスト用データとし、入力ユニット数、中間ユニット数、出力ユニット数、学習係数をそれぞれ 28, 12, 1, 0.75 とした。結果を表 5 に記す。

表 5 2 種類の筆記動作認識結果

被験者数	1 人	5 人
認識率	97.0%	84.8%

被験者 1 人のときのデータでは 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 人の場合より被験者 5 人の方が認識率は低下するが、これは被験者によって書き順が異なる場合があるために誤認識したと考えられる。

しかし、被験者 1 人、被験者 5 人の実験で共に認識率を 80% 以上得ることができ、筆記動作 2 種類の場合では提案認識手法の有効性が確認できた。

5. 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<sup>9)</sup>。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の合計 140 サンプルを学習データセットとし、70 サンプルを認識用データセットとする。なお、今回の実験では中間ユニット数を 14 とした。認識実験の結果を表 6 に記す。

表 6 7 種類の筆記動作認識結果

被験者数	1 人	2 人	3 人
認識率	61.4%	57.1%	49.5%

被験者が 1 人のときの認識率は約 61%、2 人の場合では約 57%、3 人の場合では約 50% となった。こちらも被験者の増加に伴い認識性能が低下した。これは 5. 1 節の 2 種類動作実験と同様、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。

しかし、2 種類の動作から 7 種類の動作とクラス数が増加した際に、約 30% 認識性能が低下している。そこで、それぞれの文字ごとに認識率を調べたところ、6 の筆記動作で 18.8% と他の筆記動作の半分以下の認識率となっていることが分かった。この原因を検討したところ次の原因が考えられる。

6 の筆記動作が 4 の筆記動作と大変類似しているという点がある。図 20 に 4 の筆記動作と 6 の筆記動作の重心移動を記した。これを図 9 の Code 表に当てはめると、4 の筆記動作の Code は 671147 になり、6 の筆記動作の Code は 678146 となる。本研究の対象とする筆記動作認識においては、オンライン文字の認識に利用するペンアップ・ペンドアウンの違いを検出することができない。そのため重心移動だけで捉えた際には、文字形状が似ていなくても非常に類似したコードの識別になることがあり、これが認識性能低下に繋がっていると考えられる。

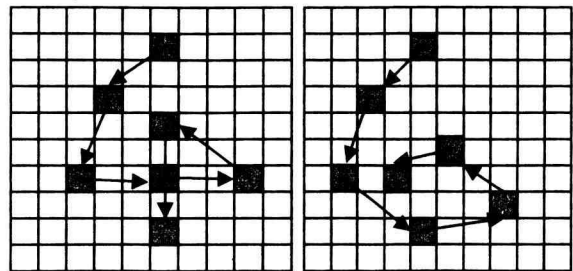


図 20 4 の筆記動作(左)と 6 の筆記動作(右)の重心移動

### 5.3 13種類の筆記動作認識結果

本提案システムの拡張性について検討するため、7階建て以上の建築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の合計260サンプルを学習データセットとし、130サンプルを認識用データセットとする。ニューラルネットワークの各ユニット数は5.2節で行った実験と同じものである。認識結果を表7に記す。

表7 13種類の筆記動作認識結果

被験者数	1人	2人	3人
認識率	42.3%	37.3%	31.3%

被験者が1人のときの認識率は約42%、2人の場合では約37%、3人の場合では約31%となった。被験者数の増加に伴う認識率低下は5.2節の7種類の動作実験結果とほぼ同じ傾向であった。

ここでも7種類から13種類へとクラス数を増加したことによって認識性能が約20%低下した。認識性能低下の原因として5.2節の4の筆記動作と6の筆記動作が類似している以外に、5の筆記動作と8の筆記動作も類似していることが考えられる。

図21に5の筆記動作と8の筆記動作の重心移動を記す。図9のCode表に当てはめると、5の筆記動作のCodeは7886431になり、8の筆記動作のCodeは6886424となる。この5と8の筆記動作が重心移動に関して類似していることが認識性能低下の原因として考えられる。

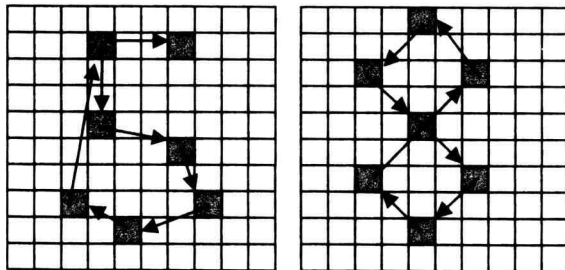


図21 5の筆記動作(左)と8の筆記動作(右)の重心移動

以上の実験結果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している。現在の実験においてサンプル数が十分とはいえないため、今後学習サンプル数を増加させてさらに検証を進める必要がある。

## 6. むすび

本研究は、エレベータ操作における視覚障害者も容易に利用できるインタフェース実現のため、カメラを利用した空間への文字筆記動作を認識するシステムを試作した。本章では、実験結果のまとめと、視覚障害者の利用を考慮したエレベータ操作インタフェースについての考察を述べる。

### 6.1 提案手法の評価実験の考察

今回認識を行った場合に必ず0から9及びB、F、Rのいずれかを認識結果として出力していたが、出力値が全て低い値であるとき、認識結果を出力しないようにし、拒否率が発生するが誤認識率を低下させるほうが有効であると考えられる。

実験の結果より、提案認識手法は実時間処理に耐えうる高速な処理時間での認識率最大97%であった。しかし、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の筆記動作と6の筆記動作や5の筆記動作と8の筆記動作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。

具体的には筆記動作だけではなく、筆記されたものを絵としオフライン文字として認識し、結果を統合して認識率を上げる方法があげられる。

### 6.2 視覚障害者利用を考慮したインタフェースの考察

本研究では、視覚障害者の利用を考慮した新たなエレベータ操作インタフェース構築方法として、カメラを利用して文字を空間に描く動作を認識するという手法を考案し、その基幹部分である認識部について試作・検討した。認識部の性能向上の課題があるものの、本手法をエレベータ操作のインタフェースとして取り扱うことについての考察を述べる。

本研究では、データ採取を実際のエレベータと同じ大きさの模型を使って行った。本研究では、あらかじめ被験者にカメラ位置を説明して、カメラに向かって操作させたが、実際の操作では、音声ガイドと協調したインタフェースを構築する必要がある。

残された課題として、複数人がエレベータ内にいる状態での身振り制限の問題がある。身振りを天井に向かって行っているため小さな身振りでも大きな画像データとして取得可能になっているものの、身長差が大きい人間では動作が隠れてしまう場合や、動作を行うこと自体に制限がある場合が予想される。今後混雑したエレベータ内での認識実験を進め、検討を進める必要がある。

また、本論文はエレベータ操作が手の動作認識で可能かの検討を進めたが、今後操作者以外の動きや、操作以外の動作を認識しないようにする手法の検討を進め、実用化に向けたシステム開発を進めていく必要がある。

以上より、本提案手法のみでの実用的なインタフェース構築は現段階では難しいが、今後、音声ガイドの導入や混雑した状況での認識率の向上を提案、検討することで優れたインタフェース構築につなげることができると考える。

## 謝辞

検証実験のためにデータの収集にご協力して戴いた富川研究室学生諸氏に心からの感謝の意を表します。

## 参考文献

- 1) 日本規格協会：“JISハンドブック9 建築II”，pp1006-1007，財団法人日本規格協会，(2004)
- 2) “TMPGEnc Net”，<http://www.tmpgenc.net/>.
- 3) “色のお話”，<http://anipeg.yks.ne.jp/color.html>.
- 4) 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
- 5) 佐藤大志，佐藤雅行，“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”，神奈川工科大学工学部電気電子工学科卒業論文(2003)
- 6) 日本建築士会連合会：“問答式建築法規の実務”，pp1573，新日本法規出版株式会社，(1985)