動的計画法に基づくマッチングによる運動認識*

太田 順** 潮崎 成弘*** 新井 民夫**

Motion Recognition using DP Matching

Jun OTA, Shigehiro SHIOZAKI and Tamio ARAI

This paper presents recognition methodology of motion from 2 dimensional images. The following two problems are discussed: (1) how to define features for the recognition. (2) how to realize pattern matching. For (1), authors classify features into two: temporal ones and sequential one. Distances are defined with respect to two features. For (2), dynamic programming is introduced because of robustness of phase deflection, expansion and contraction. Recognizing experiments on the finger language showed the effectiveness of the proposed algorithm.

Key words: motion recognition, features, image understanding, dynamic programming, finger language

1. 序 論

近年、ロボティクスやマン・マシンインタフェースの分野において、運動認識に関する研究が多く行われている。

ここで対象とする認識には以下の二つのレベルがあると考えられる。(1) 運動の構造の認識、(2) 運動の意味の認識。運動の構造の認識とは、運動を実現した2次元または3次元構造の解析である。すなわち、特定の運動、もしくは特定の幾何的構造下での任意の動きを対象として、その運動を実現したパラメータを抽出することである。この分野の従来研究としては、手の形状の画像間でのマッチングを取り、運動パラメータを計測する研究¹⁾、人の手足をリンク機構でモデル化し、そのパラメータを測定する研究²⁾⁻⁵⁾が行われている。これらの研究はパラメータ同定の実時間性、正確さを目標に置き、トラッキングや運動の予測技術を伴い、移動ロボットの障害物回避等に使用されている。

一方の運動の意味の認識では、構造的解析を目的としないか、もしくは後段の異なる処理系に委ねることとして、あらかじめ用意した複数の運動から観測した運動に適合するものを識別することを行う。これは異なる運動を持つことを想定しているために、幾何的な構造等の一つの完結したモデルを持つことを避け、異なる運動シーンを出来るだけ同じ特徴量空間に投影することで、各モデル間の“距離”を設定することを目指す。路上の移動体の観測を行い、移動部の濃淡値における周期性の有無から人の足の動きの検出を行い、人とそれ以外の移動物との区別を行う研究⁶⁾がある。これはセンサによる監視の例と考えられる。またマン・マシンインタフェースとしての位置付けから、手

話の動画認識⁷⁾や人のジェスチャ認識^{8) 9)}、また基礎研究レベルでの円運動や正弦運動の識別¹⁰⁾等がある。文献10)は、カルマンフィルタを通して得られる状態変数としての速度の分散を入力としたニューラルネットワークにより、直線運動、正弦運動、円運動を識別したものである。これらは、認識という大きな枠組みの中において、画像による運動認識の分野を担っており、対象ごとに完結した結果に陥る危険性を克服すれば、ビジョンがロボットのシステム構築において制御部等の他の機能の“従”の役割のみを担うだけにとどまらない発展性が見い出せる。そこで、本論文では“運動の意味の認識”を行うことを目標とする。特に、以下の2つの問題を解決する必要がある。

- ・ 特徴量空間をどのように設定するか

物体の形状が表現する特徴と物体の運動が表現する特徴の両者が存在するため、その両者を考慮した特徴量空間を設定する必要がある。

- ・ どのようにマッチングするか

上記の特徴量空間の問題と関連するが、形状のマッチングと運動シーケンスのマッチングの両者を行う必要がある。また、一連の運動の中から意味のある情報を切り出す機能も必要である。

本論文では、上記の問題を解決する認識アルゴリズムを提案する。従来研究において文献8)と9)は、両者とも画像全体に映された人が腕を動かす等の運動をしているものを認識するために、特徴量としては画像全体のパターン、すなわち形状特徴のみを使っている。しかし、このモデルは画像上における位置のずれを考慮していないために、マッチングの失敗が起りやすい。すなわち、これらの研究は、マッチング時の特徴量の選択という観点から十分ではなかったため、本論ではこの改良を目指す。当然のことながら、それに伴いマッチング手法も改良する必要がある。本論文の検証対象として、指文字の運動認識を行う。まず、次章では上に示

* 原稿受付 平成8年3月13日

** 正会員 東京大学(東京都文京区本郷 7-3-1)

*** 三菱重工業(株)(高砂市荒井町新浜 2-6-47)

した問題を解決するための認識機構の概念設計を行う。

2. 運動認識アルゴリズムの概念設計

2.1 モデルの記述に必要な特徴量

前述したように本論文の認識対象では、認識すべき特徴量が多様かつ複雑であるため、ここでは、最初に特徴量の種類と構造を明らかにする。

運動認識における特徴量はその処理段階によって、ある時間での1枚の画像から得られる瞬間的な特徴と、それら瞬間的な特徴の時系列から得られる時間的変化の特徴の両者を認識する必要がある。以後、これらの特徴を順に瞬間特徴、時間特徴と呼ぶ。例えばエッジやコーナ、あるいは物体の重心の位置はある時間における1枚の画像から得られる瞬間特徴である。この瞬間特徴はある瞬間において得られた画像から得られる特徴で、2次元座標空間の各空間座標方向に対する処理によって得られる。一方の時間特徴は時系列から得られる特徴であり、時間軸方向の処理によって得られる特徴といえる。例えば、瞬間特徴におけるエッジは空間座標方向の微分値の絶対値であり、これに相当する時間特徴におけるエッジは速度の絶対値である。このような時間特徴は時系列を記述する上では不可欠なパラメータであり、運動のモデルを特定するためにはこの時間特徴を抽出する必要がある。

また一方で、それらの特徴は物体の形状を示すものと物体の画像空間中の位置姿勢といった状態を示すものがある。前者を形状特徴、後者を位相特徴と呼ぶ。なお、この両者は独立でない場合が多々存在する。すなわち、形状特徴に基づき位相特徴を算出する場合があることに注意されたい。上記の特徴を表1に示す。

Table 1 Types of character to be detected

	Instantaneous feature	Time feature
Geometrical feature	Shape	Transformation
Phase feature	Position and orientation	Motion

本論文では、上に示す複雑な特徴量を以下のようにマッチングする。

- ・ 瞬間特徴量はパターンマッチングによりマッチングする。
- ・ 時間特徴は離散化して時系列モデルのシーケンスとし、モデルと実画像のシーケンス同士をマッチングして運動認識を図る。

以下、それぞれについて説明する。

2.2 マッチングの方法

2.2.1 パターンマッチング

本論文では人の動き、特に手話などに見られる人の手指の動き等を認識対象とする。この対象においては、画像全体に映された形状が変形もしくは移動をするも

のではなく、部分的に変形、移動する場合も多く存在する。この場合、画像のサイズ全体を形状特徴として持つことはマッチングの精度の低下を引き起こすために適切ではない。そこで、本論文では、これらの特徴量から作られる2次元特徴量空間上の距離を以下のように設定する。

$$Q(t, u) = \alpha \cdot \text{diff}_{img} + \beta \cdot \text{diff}_{phase} \quad (1)$$

但し、 Q はパターン間距離、 diff_{phase} は位相の差、 diff_{img} は通常のテンプレートマッチングにより得られる形状の差である。従来の研究では、 diff_{img} のみを用いてマッチングを行っているのに対し、本研究では新たに diff_{phase} をマッチングに用いることで画像中の部分的な変形・運動にも適応できるようになっている。また、指文字が簡単なパターンから構成されることから、本論文では、手先の形状の変形については扱わない。この特徴量空間における距離を使い、時系列パターン同士のマッチングを行う。

2.2.2 シーケンスのマッチング

シーケンスのマッチングにおいて、時間特徴量のモデル化手法としては、

- ・ 物理モデル：動力学方程式に基づいて運動をモデル化する方法
- ・ 時系列モデル：パタンのシーケンスとして運動をモデル化する方法

の2種類が考えられる。

運動の意味の認識は、運動をモデルとしたモデルベース認識であり、運動の構造的解析を必要としていない。従って、用意するモデルは“直線運動”等の状態変数等の力学的に表現できる運動から“テニスのフォアストローク”等の動きに対して人が便宜上付けた運動の名称まで包含させることに問題はない。

また、運動の認識を行う必要性は1章で述べたように、以下に分類できる。

- ・ 運動の構造解析を続けて行うためのモデルの特定
 - ・ 特定の運動を行うものを検出する監視
 - ・ マン・マシンインタフェース
 - － バーチャルリアリティにおけるデータグローブに相当する非接触の入力法
 - － 手話認識

いずれの場合も識別すべき運動は運動シーンを一定時間以上観測した結果識別できるものであり、連続する画像から得られる瞬間的な運動の計測ではない、高次の認識を必要としている。従って、その運動モデルは数個のパラメータとして用意した物理量で記述できるようなモデルではない。これはモデルが力学的に構造が複雑、または力学的根拠がないために構造化が不可能であることが原因となっている。こういった運動を認

識する場合には、画像から直接得られる低次の特徴量を時間軸に沿って並べた特徴量時系列によってモデルを記述することが妥当である。

時系列モデルを使った運動認識としては連続DPを用いたもの⁸⁾とHMMを用いたもの⁹⁾がある。ここで、HMM⁹⁾は時系列の始末端が既知であること、すなわち切り出しが既になされていることを前提としており、この時系列端を調整しながらのマッチングにはHMMは非力であると考えられる。そこで本論文では、DPマッチングを採用する。

次章では、時系列同士のマッチングにおける、時間位相合わせを考慮したマッチングの手法として、動的計画法によるマッチングを導入する。

3. 動的計画法による認識

3.1 連続DP

時系列同士のマッチングでは、ある特徴量の時系列をモデルとして用意し、入力時系列から抽出した同じ特徴の時系列をマッチングさせる。この場合、一方の系列の始点か終点に他方の系列のある時点を一一致させ、残りの系列の最適な対応を求めた後に、両系列のマッチングの距離を求める。連続DP (Continuous Dynamic Programming) は離散単語音声のパターンマッチングの距離計算の手法の一つであるDPマッチング¹¹⁾を連続単語音声に拡張したものである。以下に連続DPによるワードスポッティング¹²⁾方式の概要を述べ、画像の時系列を使った場合の特徴量の設定を考える。

各時系列を構成する特徴パターン同士の距離を $Q(t, u)$ とする。ここで t は時間を、 u はモデルの系列を意味する。時系列間の距離は状態量 $P(t, u)$ として、以下のように定義される。

$$P(t, u) = \min \begin{cases} P(t-2, u-1) + 2Q(t-1, u) + Q(t, u) & \text{(a)} \\ P(t-1, u-1) + 2Q(t, u) & \text{(b)} \\ P(t-1, u-2) + 2Q(t, u-1) + Q(t, u) & \text{(c)} \end{cases} \quad (2)$$

$$0 \leq u \leq U \quad (3)$$

$$0 \leq t \leq T \quad (4)$$

これより得られた状態量 $P(t, U)$ は入力時系列において t を終端とした時のモデルとの最適マッチング距離を示し、これを積分経路の長さ $C(t, u)$ で割って正規化したものがDP値 $D(t)$ となる。この計算を終端時間0から T の間でそれぞれ行う。この値はモデルに最も一致する時系列ブロックの終点に当たる t_{min} で局所的最小値 $D(t_{min})$ を取る(図1)。

この特徴パターンを適当に選ぶことにより音声以外の時系列同士のマッチングさせることが可能になる。こ

Fig.1 Search region of continuous DP

こで距離値はスカラー量であるため、特徴パターンが多次元空間上に投影されている場合、その空間での特徴パターン間距離 $Q(t, u)$ を定義する必要がある。データグローブを使った手話認識の研究¹³⁾ではグローブから得られる関節角などの32次元空間での特徴パターンを形成し、ユークリッド距離を使ったマッチングを行っている。本論文では画像時系列を使った場合の時系列マッチングのために、多次元の特徴量による画像パターンのパターン距離を設定する。

3.2 動画認識のためのモデルと特徴量の設定

前章においてマッチングの距離関数として式(1)を示した。具体的には、時系列モデルを記述するための特徴量として、瞬間特徴として画像中の主要部のパターン形状、時間特徴としての位相の一つである瞬間速度方向を使う。モデルの画像時系列 $Model(v, w, u)$ 、および入力パターンの画像時系列 $Input(x, y, t)$ を基として、「画像パターン」そのものと「原画像上での瞬間移動方向」を特徴量として抽出、利用する。

$$Model(v, w, t) \rightarrow (M(v, w), a_m)(t) \quad (5)$$

$$Input(x, y, t) \rightarrow (I(x, y), a_i)(t) \quad (6)$$

但し、 $M(v, w)$ をモデルパターン、 $I(x, y)$ を入力パターン、 a_m, a_i を瞬間移動方向とする。

パターン同士のマッチング距離 $diff_{img}(t, u)$ を以下のように定義する。

$$diff_{img}(t, u) = \min_{x, y} \frac{\sum_{v, w} |I(x+v, y+w) - M(v, w)|}{\sum_{v, w} |I(x+v, y+w) + M(v, w)|} \quad (7)$$

移動方向については、全周を8方向に離散化したものを使用し(図2)、角度の差をその離散値の差の大きさにより求める。

$$\text{diff}_{\text{angle}}(t, u) = (0 \sim 4)/4 \quad (8)$$

Fig. 2 Quantized direction of velocity

両特徴量とも正規化されており，これらの線形和をパターン距離とする。

$$Q(t, u) = \text{diff}_{\text{img}} + \text{diff}_{\text{angle}} \quad (9)$$

ここでは，簡単のため，式(1)の α, β をそれぞれ1と設定した。

この特徴量空間は2次元であり，各特徴量における独立な距離の線形和によって距離を定義しているため，2次元画像におけるCity Block Distanceに相当するものと考えられる。

本章ではDPマッチングによる時系列マッチングを導入した。時系列を構成する各特徴量は式(9)で導入したものに従い，この特徴量からパターン距離を定義するために，瞬間運動特徴を離散値化し距離を定義した。

次章では，指文字を取り上げ，認識対象とする運動を明確にする。

4. 濃淡画像を使った実験

4.1 指文字

手話認識は，手指の形状認識を対象とした研究^{1) 14)}とその運動を含めた認識の研究^{7) 13)}がある。バーチャルリアリティのための入力装置への応用として手指の構造解析^{1) 14)}では物理モデルの構築を最大の目的としているために，その時系列を扱う段階ではない。一方，手話認識のための手指の動画認識では並進成分の検出⁷⁾や，移動方向の検出¹³⁾で，速度をDP等のマッチングそのものの特徴量として使ったものはない。

本論文では手指の形状と，位相の一部である瞬間速度の方向をそれぞれ特徴量ベクトルとし，それらが構成するベクトル場上の距離をパターン距離として連続DPによる時系列マッチングを行う。

対象とする運動は人差し指で指文字の“ん”を示す動きをするものである。この指文字は指文字の中でも運動(時間的変化)により表される文字である。この運動により表される指文字としては，濁音，拗音(撥音)，“の”，“ん”等がある。

“ん”の動きは人差し指で片仮名の“ン”をなぞる動きであるために，画像上の形状は見る角度によって色々

な形をとる。従って，ここでは候補となる手の形をそれぞれ用意し，単純なテンプレートマッチングによって生じる誤差をDPマッチングにより吸収することを期待する。

この実験では，手の形状とその速度方向をパターン距離としたDPマッチングの有効性を示すために，入力時系列の画像サイズと同じ大きさの画像列をモデルとしたものとの比較を行う。比較では特徴パターン距離を変えて，同じDPマッチング法により時系列マッチングを行う。

比較対象として，特徴量を以下の3つの場合で組み合わせた時系列モデルを使用する。

モデル構成1

- ・ 形状特徴として入力画像と同じサイズの画像を使用
 - ・ 位相特徴は未使用
- すなわち，

$$Q(t, u) = \text{diff}_{\text{img}} \quad (10)$$

モデルを入力画像と同じ大きさの画像として，画像同士の単純なマッチングによる距離をパターン距離とした場合である。用意するモデル画像はモデルの原画像そのもので，実験環境の背景を暗くする背面処理を行う。従って，画像上における手の位置のずれがパターン距離に大きく影響することが予想される。

モデル構成2

- ・ 形状特徴として手の形状を使用
 - ・ 位相特徴として画像上における位置を使用
- すなわち，

$$Q(t, u) = \text{diff}_{\text{img}} + \text{diff}_{\text{location}} \quad (11)$$

モデルを手の形に切り出した画像と，その画像上の位置とする。この場合もモデル構成1と同様に画像上における手の位置のずれがパターン距離に影響することが予想される。

モデル構成3(本論文で主張する特徴量)

- ・ 形状特徴として手の形状を使用
 - ・ 位相特徴として瞬間移動方向を使用
- すなわち，

$$Q(t, u) = \text{diff}_{\text{img}} + \text{diff}_{\text{angle}} \quad (12)$$

モデルを手の形に切り出した画像と，その移動方向とする。画像上の手の位置のずれに対してロバストに認識できることが期待できる。

4.2 モデル時系列の作成

原画像は256×256のサイズの256階調の濃淡白黒画像を使用する(図3)。背面処理を行い，手の部分を切り出してから認識処理を行う。時系列画像はサンプリングタイム1/10秒である。

モデル時系列は時間長8で，入力時系列とは異なる画像時系列から作ったモデル群より必要なものを選ぶ。

図 3 に対するモデル構成 1 のモデルを図 4 に示す。モデルの画像はあらかじめ背面を処理し手の部分のみを切り出したものであり、白黒の濃淡画像であるが、大きさはモデルにより異なる (図 5)。

モデル構成 2 とモデル構成 3 の例を以下に示す。モデル構成 2 では手の形状と位置のペアを、モデル構成 3 では手の形状と移動方向のペアを、それぞれ採用している。

モデル構成 2 の例 [, $x = 160, y = 60$], ...

モデル構成 3 の例 [, $\alpha = 7$], ...

モデル時系列を 3 つの構成法に従って構築し、比較を行う。具体的には、以下の実験を行う。

- (1) “ん”の動きを 1 度だけ試行した場合の認識実験
- (2) “ん”の動きを、画像での位置を変えて 2 度連続して試行した場合の認識実験
- (3) 連続した“ん”の動きの認識実験

4.3 単独の“ん”の運動の認識

入力時系列画像は全時間長 9 で 256×256 の大きさの白黒の濃淡画像である。図 6 に示すように、あらかじめ背面を処理しておく。

Fig.3 Original images

Fig.4 Images after masking (model No.1)

Fig.5 Detected patterns

この実験においては、入力時系列とモデル時系列とは比較的差が少ない (画像上での差が数十画素程度) 運動をしたものを用意している。

モデル構成 1 による DP マッチングの結果を図 7 に示す。画像上の位置のずれが小さいためにマッチングに成功している。

モデル構成 2 の場合もモデル 1 と同様に画像上の手の位置のずれがパターン距離に影響を及ぼすが、画像上の位置のずれが小さいためにマッチングに成功している (図 8)。

Fig.6 Input image sequence

Fig.7 Results of experiment 1 with model No.1

モデル構成 3 を用いた場合の DP マッチングの結果を図 9 に示す。画像の直接マッチングと同様の結果が得られる。

どのモデル構成の場合もマッチング可能であったが、 $D(t)$ の値の極小の程度を比較すると、モデル構成 3 の場合が最も谷が深く (他の値の 80% 程度、モデル構成 1 と 3 は 95% 以上)、安定的に認識していることがわかる。

4.4 画像上の位置を変えた場合の運動認識

次に画像上の手の位置が入力画像とモデル画像で異なる場合についての実験を行う。入力時系列画像は全時間長 37、前半 16 枚と後半 21 枚がそれぞれ同じ動きを連続で行う。但し、画像上における位置が異なる運動とした。あらかじめ背面を処理しておく (図 10)。モデル時系列は時間長 8 とする。

モデル構成 1 を用いた場合の DP マッチングの結果を図 11 に示す。位置のずれがマッチングに影響し、前半の運動は認識できるが後半はマッチングに失敗していることがわかる。

モデル構成 2 を用いた場合の DP マッチングの結果を図 12 に示す。モデル時系列は入力時系列中、前半の運動を行っている時の手の位置を採用したが、位置のずれがマッチングに影響し、前半の運動は認識できるが後半はマッチングに失敗していることがわかる。

モデル構成 3 を用いた場合の DP マッチングの結果を

Fig.8 Results of experiment 1 with model No.2

Fig.9 Results of experiment 1 with model No.3

Fig.10 Input image sequence

図 13 に示す。画像上の位置はテンプレートマッチングの段階で吸収され、その速度方向のみを時系列マッチングに使用しているため、両運動とも認識できている。DP マッチングのための特徴パターン間距離の設定が良好であることが確認できる。

これより、モデル構成 3 が他の 2 種類のモデルと比較して特徴形状の位置の誤差にロバストであることがわかった。

4.5 連続した運動の切り出し

次にモデル構成 3 を使い、連続して何度か“ん”の動きを試行したものに対して処理を行い、各運動の切り出しの実験を行う。

9 回連続して行ったもののマッチングと、10 回連続して行ったもののマッチングを比較する。但し、9 回のは前項で行ったものと同じ正面からの撮影で、10 回のは約 60° 側方から撮ったものである。どちらの場合もモデルとしては正面から撮影したものを用いている。それぞれの結果を図 14 に示す。側方から撮った

Fig.11 Results of experiment 2 with model No.1

Fig.12 Results of experiment 2 with model No.2

Fig.13 Results of experiment 2 with model No.3

ものは DP 値の変化が顕著ではなく、局所 minimal の個数が 10 個以上となり、誤認識を含んでいる (図 14(b))。

Fig.14 Results of DP matching

また、 $\pm 60^\circ$ 内においては、50 回程度の試行に対して、9 割以上の認識率が得られた。これより、 $\pm 60^\circ$ 程度の角度方向のずれに対しても本手法はロバストに認識できることが示された。

4.6 評価

特徴パターン間距離として式 (1) が妥当であるかを検討した。DP マッチングは時間軸方向の時間位相ずれや伸縮に強いマッチング法であるが、これは図 1 における探索方向に依存する。今回の実験では 2 フレーム以上の大きなずれが連続すると、誤認識の原因となった。

画像パターンとその画像上の位置を特徴パターン間距離とした場合も、画素 20 以上のずれが出ると誤認識しやすくなる。誤差に弱い原因には位置を特徴量とする問題以外にモデル系列長が短く、適当なマッチング結果が得られるものが少なかったことがある。

また、DP マッチングを使用する場合はその探索範囲により認識率が大きく変わる。なぜなら、ある程度の幅を持たせる探索範囲でなければ、時間位相ずれや伸縮によりパターンが不連続になった場合に、その部

分を飛び越えた経路の探索が出来なくなるためである。従って、これは入力時系列長とモデル時系列長との差に大きく依存することから、時間軸方向の伸縮に冗長性を持たせるにはこの部分の調整に関して考慮する必要がある。

5. 結 論

本論文では、手指などの人の動きを認識するために時系列マッチングの手法としてDPマッチングによる認識系の構築を行った。

- (1) 認識のための特徴量を整理し、瞬間特徴としてのパターンの形状、時間特徴としての運動の方向を使い、これらで構成される2次元特徴量空間での距離を定めた。
- (2) 時間位相のずれや伸縮に強い時系列マッチングの手法として、動的計画法によるマッチング(DPマッチング)を導入した。
- (3) 提案した手法を指文字認識に適用することで、有効性を示した。

参 考 文 献

- 1) 岩井 儀雄, 八木 康史, 谷内田 正彦: 単眼動画からの手の3次元運動および位置の推定, MIRU'94, 2 (1994) 207.
- 2) 加藤 克巳, 大西 昇, 杉江 昇: バイオロジカルモーションに対する運動立体視機構のモデル, 電子通信学会技術報告, MBE87-144 (1988).
- 3) M. Asada and S. Tsuji: Representation of Three-Dimensional Motion in Dynamic Scenes, Computer Vision, Graphics, and Image Processing, 21 (1983) 118.
- 4) J. O'Rourke and N. Badler: Model-Based Image Analysis of Human Motion Using Constraint Propagation, IEEE Trans. PAMI, PAMI-2, 6 (1980) 522.
- 5) M. Yamamoto and K. Koshikawa: Human Motion Analysis Based on a Robot Arm Model, IEEE Proc. ICCV CVPR'91, (1991) 664.
- 6) S. Yasutomi and H. Mori: A Method for Discriminating of Pedestrian based on Rhythm, IROS'94, (1994) 988.
- 7) 田村 進一, 川崎 伸吾: 手話動画認識システム, 情報処理学会研究資料, CV44 (1986) 1.
- 8) 高橋 勝彦, 関 進, 小島 浩, 岡 隆一: ジェスチャ動画のスポットティング認識, 電子情報通信学会論文集, J-77-D-II (1994) 1552.
- 9) 大和 淳司, 倉掛 正治, 伴野 明, 石井 健一郎: カテゴリーVQを用いたHMMによる動作認識法, 電子情報通信学会論文集, J77-D2, 7 (1994) 1311.
- 10) 梅田 和昇, 大内 茂樹, 新井 民夫: 運動の特徴量の抽出による時系列距離画像からの運動の学習・認識, 第11回日本ロボット学会学術講演会予稿集, 1 (1993) 347.
- 11) 迫江 博昭, 千葉 成美: 動的計画法を利用した音声の時間正規化に基づく連続単語認識, 日本音響学会誌, 27, 9 (1971) 483.
- 12) 岡 隆一: 連続DPを用いた連続音声認識, 音響学会研究資料, S78-20 (1978) 145.
- 13) 佐川 浩彦, 酒匂 裕, 太平 栄二, 崎山 朝子, 阿部 正博: 圧縮連続DP照合を用いた手話認識方式, 電子情報通信学会論文集, J77-D, 2 (1994) 753.
- 14) 亀田 能成, 美濃 導彦, 池田 克夫: シルエットを利用した手指の三次元形状推定法, MIRU'92, 2 (1992) 239.