

頭部ジェスチャの度合認識に関する研究

木戸 隆太郎 呉 海元 塩山 忠義

京都工芸繊維大学大学院 工芸科学研究科
〒606-8585 京都市 左京区 松ヶ崎御所海道町

E-mail: {wuh,y,shioyama}@ipc.kit.ac.jp

ビデオカメラで撮影した頭部ジェスチャの動画を時系列特徴ベクトルに変換する。この時系列特徴ベクトルを用いた頭部ジェスチャの識別と度合の認識方法を提案する。同じジェスチャでも、その動作の速さ・大きさは様々である。その度合は、時系列特徴ベクトルでジェスチャを表す要素パターンの時間の範囲・値の範囲に反映される。同種類の要素パターンならば時間の範囲・値の範囲が異なっても柔軟に、且つノイズが入っても安定にジェスチャを識別すると同時にその度合を認識できるように CDP(Continuous Dynamic Programming) を改良する。また、その時の要素パターンの時間の範囲と値の範囲を度合に関するパラメータとして出力できるようにする。そして、改良に伴って増加した計算量を主成分分析を利用して減らす試みをする。

キーワード：頭部ジェスチャ、三次元姿勢推定、度合認識、CDP の改良、主成分分析

Degree Recognition of Head Gestures

Ryutaro KIDO, Haiyuan WU, Tadayoshi SHIOYAMA

Department of Mechanical and System Engineering

Kyoto Institute of Technology

Matsugasaki, Sakyo-ku, Kyoto, 606-8585, Japan

E-mail:{wuh,y,shioyama}@ipc.kit.ac.jp

This paper presents an approach for degree recognition of human head gestures from color image sequence. The 3D pose of the head relative to the camera is estimated with the skin color region and the hair color region. The reference patterns of gestures are represented by the sequences of rotation angles in X, Y, Z axis, and skin area. The human head gestures are recognized by using CDP(Continuous Dynamic Programming) algorithm for spotting recognition to compare input image sequence with the reference patterns. Then CDP is improved to recognize the patterns flexibly and robustly, and it can output the parameter concerned in the degree of the expansion and contraction. Moreover, the computational burden is reduced by using principal component analysis.

key word : head gestures, 3D pose estimation, degree recognition,
improving CDP, principal component analysis

1 はじめに

色情報を基に頭部の3次元姿勢(頭部座標系の X, Y, Z 軸回りにおける回転角 $\theta_x, \theta_y, \theta_z$)、ならびに肌色面積(a_s)を推定し、得られたベクトル $(\theta_x, \theta_y, \theta_z, a_s)$ を特徴ベクトルとする。この時系列の変化を頭部ジェスチャの認識に使用する[1]。

ジェスチャは動作の違いでその表す意味に違いが出てくる。同じ意味を表すジェスチャでは、その動きの速さ・大きさにより度合の一部を表現している。ここで、特徴ベクトルの要素の時系列パターンを要素パターンと定義する。あるジェスチャを表す要素パターンにおいては、動作の速さがその始点から終点までのフレーム数(以下、時間範囲と呼ぶ)に、動作の大きさがその値の範囲(以下、値範囲と呼ぶ)に反映されていると考えて、要素パターンの時間範囲と値範囲の伸縮度合について調べていけばジェスチャの度合が理解できる(図1)。

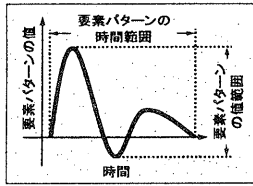


図1: 要素パターンの時間範囲と値範囲

時々刻々とジェスチャの識別と度合を認識するためには、入力画像のフレームに同期して、認識しようとするジェスチャの要素パターンに対して相似のパターンや類似のパターンを柔軟に取得して、その要素パターンの伸縮度合が分からなければならない。また、ノイズが入っても安定に認識できる必要がある。

そこで、本稿では頭部ジェスチャをCDP(Continuous Dynamic Programming)を用いてスポットィグ認識させる手法[2]を用いた。そして、上記の目的に適合するようCDPを改良し、その有効性を調べる。また、改良に伴って増加した計算量を主成分分析を利用して減らす試みをする。

2 参照パターンの作成

本稿では、頭部ジェスチャを表現する時系列特徴ベクトル I を次のように定義する。

$$I \equiv \{i(t) \mid t \in N\} \quad (N: \text{自然数の集合}) \quad (1)$$

但し、 $i(t)$ は次式のように定義する。

$$\begin{aligned} i(t) &\equiv \{i_k(t) \mid k = 1, \dots, 4\} \\ &\equiv \{\theta_x(t), \theta_y(t), \theta_z(t), a_s(t)\} \end{aligned} \quad (2)$$

後述のCDPマッチングに用いる頭部ジェスチャを表現するモデルとして参照パターンを作成する。参照パターンは、ある一つのジェスチャを表す4要素からなる特徴ベクトルの時系列として定義する。

画像系列に特徴ベクトルの抽出処理[1]を適用し、ジェスチャが表現された特徴ベクトル時系列を選び出す。これを入力したジェスチャに対する参照パターンとする。始点の時刻を1、終点の時刻を T とすれば参照パターンは、

$$P \equiv \{p(\tau) \mid \tau = 1, \dots, T\} \quad (3)$$

で表される。但し、 $p(\tau)$ は次式のように定義する。

$$\begin{aligned} p(\tau) &\equiv \{p_k(\tau) \mid k = 1, \dots, 4\} \\ &\equiv \{\theta_x(\tau), \theta_y(\tau), \theta_z(\tau), a_s(\tau)\} \end{aligned} \quad (4)$$

また、参照パターンの1要素に注目したときには、区別して参照要素パターンと呼ぶ。

3 従来型CDPのパターンマッチング

頭部ジェスチャを認識する手法としてCDPを用いる。以下に、従来型CDPによる参照要素パターンとテスト要素パターンのマッチング法を示す。ここで、テスト要素パターンとは、時系列特徴ベクトルの1要素でCDPの探索域内にあるものを指すことにする。

テスト要素パターン $i_k(\gamma)$ と参照要素パターン $p_k(\tau)$ の距離 $d_k(\gamma, \tau)$ を

$$d_k(\gamma, \tau) = |i_k(\gamma) - p_k(\tau)| \quad (5)$$

で定義する。CDPでは、 (γ, τ) における累積距離 $M_k(\gamma, \tau)$ を次のように定義[3]し、初期条件を式(6)とする。

$$M_k(-1, \tau) = M_k(0, \tau) = \infty \quad (\tau = 1, \dots, T) \quad (6)$$

$\gamma \geq 1$ について、漸化式を

$$\begin{aligned} (\tau = 1) \\ M_k(\gamma, 1) &= WB \cdot d_k(\gamma, 1) \end{aligned} \quad (7)$$

$$\begin{aligned} (\tau = 2) \\ M_k(\gamma, 2) &= \min \begin{cases} M_k(\gamma-2, 1) + WA_2 \cdot d_k(\gamma-1, 2) & (a) \\ +WA_1 \cdot d_k(\gamma, 2) & (b) \\ M_k(\gamma-1, 1) + WB \cdot d_k(\gamma, 2) & (c) \\ M_k(\gamma, 1) + WC_1 \cdot d_k(\gamma, 2) & (c) \end{cases} \end{aligned} \quad (8)$$

$$\begin{aligned} (3 \leq \tau \leq T) \\ M_k(\gamma, \tau) &= \min \begin{cases} M_k(\gamma-2, \tau-1) + WA_2 \cdot d_k(\gamma-1, \tau) & (d) \\ +WA_1 \cdot d_k(\gamma, \tau) & (e) \\ M_k(\gamma-1, \tau-1) + WB \cdot d_k(\gamma, \tau) & (e) \\ M_k(\gamma-1, \tau-2) + WC_2 \cdot d_k(\gamma, \tau-1) & (f) \\ +WC_1 \cdot d_k(\gamma, \tau) & (f) \end{cases} \end{aligned} \quad (9)$$

で与える。但し、上式の $WA_1, WA_2, WB, WC_1, WC_2$ は傾斜制限の重み係数で、それぞれ $WA_1 = 2, WA_2 = 1,$

$WB = 3, WC_1 = 3, WC_2 = 3$ である。上記の漸化式を表す傾斜制限を従来型 CDP の探索域の概念図である図 2 の中に示す。また、同図中の最適変形パターンとは参照要素パターンが CDP の探索域でテスト要素パターンに最も近い形に変形されたパターンで、このパターンとテスト要素パターンが照合されることになる。

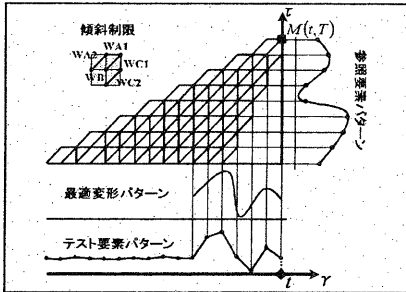


図 2: 従来型 CDP の探索域

ここで、時刻 $\gamma = t$ に動作が終了したと仮定する。すなわち、 $\tau = T$ とおくと、 $M(t, T)$ は参照要素パターンの区間 $[1, T]$ についてテスト要素パターンに最適に整合したときの累積距離を示す。この値を重み $WB \cdot T$ で正規化したもの

$$A_k(t) = \frac{1}{WB \cdot T} M_k(t, T) \quad (10)$$

がその時刻でのテスト要素パターンに対する CDP の出力値である。

N 種類の参照要素パターンがあれば、各時刻 t において N 個の CDP の出力値が生成される。各々の CDP の出力値は、各々の参照要素パターンの終点とみなせる値が入力された時点で最小になる。よって、ある時刻 t で、ある CDP の出力値が閾値より小さくなれば、その参照要素パターンに対応する要素パターンを認識したと判断する。

4 改良型 CDP のパターンマッチング

4.1 テスト要素パターンの値範囲の違いに対応できる CDP

従来型 CDP は参照要素パターンに対して時間軸方向の伸縮を全体的に持っているも局所的に持っているも、目的のテスト要素パターンを認識できる [4] が、要素パターンの値範囲に多少の違いがあると、CDP の出力値が閾値より落ちず認識できない。そこで、参照要素パターンに 13 段階の倍率を掛け、各倍率のときの CDP の出力値を計算し、その中の最小値を最終的な CDP の出力値とする。この方法でパターンの値範囲の違いを吸

収できるようにした。倍率の数値はテスト要素パターンの値範囲の $1/2$ 倍から 2 倍まで対応できる 13 段階 $[0.4, 0.5, \dots, 0.9, 1.0, 1.2, \dots, 2.0, 2.2]$ である。このテスト要素パターンの値範囲の違いに対応できる CDP を CDP-A と呼ぶことにする。

4.2 テスト要素パターンの値範囲と時間範囲の伸縮比を出力する仕組み

テスト要素パターンの値範囲の違いには参照要素パターンに 13 段階の倍率を掛けることで対応するが、このとき選ばれた倍率をテスト要素パターンの値範囲の伸縮比として出力する。テスト要素パターンの時間範囲の伸縮比は、CDP の探索域で最適経路が時間軸方向、つまり γ 軸方向にどれだけ進んだかをカウントし、参照要素パターンのフレーム数で割った値を出力する。カウントは CDP 漸化式と別途に設けた下式 (11), (12), (13) の漸化式で行う。

$$C_k(\gamma, 1) = 0 \quad \text{if 式 (7)} \quad (11)$$

$$C_k(\gamma, 2) = \begin{cases} C_k(\gamma-2, 1) + 2 & \text{if 式 (8.a)} \\ C_k(\gamma-1, 1) + 1 & \text{if 式 (8.b)} \\ C_k(\gamma, 1) + 0 & \text{if 式 (8.c)} \end{cases} \quad (12)$$

$$C_k(\gamma, \tau) = \begin{cases} C_k(\gamma-2, \tau-1) + 2 & \text{if 式 (9.d)} \\ C_k(\gamma-1, \tau-1) + 1 & \text{if 式 (9.e)} \\ C_k(\gamma-1, \tau-2) + 1 & \text{if 式 (9.f)} \end{cases} \quad (13)$$

最終的な参照要素パターンに対するテスト要素パターンの時間範囲の伸縮比は下式 (14) で出力する。

$$B_k(t) \equiv \frac{C_k(t, T)}{T-1} \quad (14)$$

以上のことを用いてテスト要素パターンの時間範囲と値範囲の伸縮比を時系列で出力する。その挙動の一例を図 3 に示す。同図中の最上部のグラフはテスト要素パターンを表し、値範囲の度合いが小さいものが①, ②, ③、中ごろのものが④, ⑤, ⑥、大きいものが⑦, ⑧, ⑨である。また、時間範囲の度合いが小さいものが①, ④, ⑦、中ごろのものが②, ⑤, ⑧、大きいものが③, ⑥, ⑨である。二番目のグラフは CDP の出力である。そして、三番目、四番目のグラフがそれぞれ時間範囲、値範囲の伸縮比を時系列で出力したときの挙動である。この伸縮比を表す三番目、四番目のグラフは二番目のグラフの CDP の出力値がある閾値より落ち、尚且つそのときの最下点である垂直線の時点で意味をなす。例えば、時間範囲が順に大きくなる組①, ②, ③を選ぶと、三番目のグラフでも垂直線の時点の出力は①, ②, ③の順で大きくなっていくのが分かる。また、値範囲が順に大きくなる組②, ⑤, ⑧を選ぶと、四番目のグラフでも垂直線の時点

の出力は”②, ⑤, ⑧”の順で大きくなっていく。この出力を利用してパターンの度合を認識する。

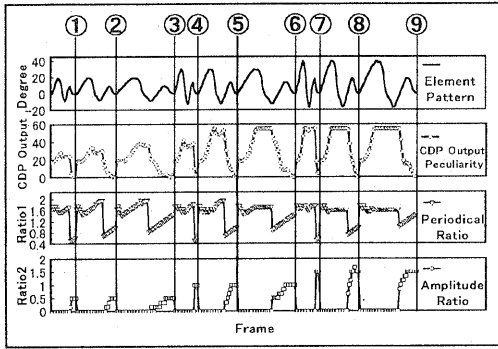


図 3: 伸縮比のパラメータの挙動

4.3 短時間のテスト要素パターンに対応できる CDP

短時間のテスト要素パターンは要素パターン全体に対し比較点が少なく、参照要素パターンから誤差の大きい比較点で無理に照合するため、認識すべき要素パターンでも CDP の出力値が閾値より落ちないことがある。そこで、この場合のパターンに限り二つの隣り合う比較点間にその平均値を補間して、要素パターン全体に対する比較点を多くしてから CDP の出力値を求める。この方法で短時間のパターンでも閾値を低い値に抑えることができ識別の誤りを防ぐことができる。また、この場合のパターンに限り比較点を増やしているので計算量の増加を抑えることができている。この CDP を CDP-S と呼ぶことにする。

CDP-S については、図 4 のように $\tau - T \geq \gamma - t$ と $\tau - T < \gamma - t$ とで場合分けをする。前者の場合は前述の従来型 CDP の距離と漸化式であり、後者の場合は下記の距離と漸化式になる。ここから、CDP-S はその探索域の概念図である図 5 中の (a), (b) のような複数の傾斜制限を持つ。 $i_k(\gamma)$ と $p_k(\tau)$ の距離 $d_k(\gamma, \tau)$ を $t - T \leq \gamma \leq t$ において次式のように定義する。

$$d_k(\gamma, \tau) \equiv \begin{cases} |i_k(\gamma) - p_k(\tau)| & (\gamma \in Z) \\ \left| \frac{i_k(\gamma - \frac{1}{2}) + i_k(\gamma + \frac{1}{2})}{2} - p_k(\tau) \right| & (\gamma \notin Z) \end{cases} \quad (15)$$

$(\gamma = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}, \dots), \quad (Z: \text{整数の集合})$

$\gamma \geq 1$ について、漸化式を次式のように与える。

$$\begin{aligned} (\tau = 1) \\ M_k(\gamma, 1) &= WB \cdot d_k(\gamma, 1) \\ (\tau = 2) \end{aligned} \quad (16)$$

$$M_k(\gamma, 2) = \min \begin{cases} M_k(\gamma - 1, 1) + WA_2 \cdot d_k(\gamma - \frac{1}{2}, 2) \\ + WA_1 \cdot d_k(\gamma, 2) \\ M_k(\gamma - \frac{1}{2}, 1) + WB \cdot d_k(\gamma, 2) \\ M_k(\gamma, 1) + WC_1 \cdot d_k(\gamma, 2) \end{cases} \quad (17)$$

$(3 \leq \tau \leq T)$

$$M_k(\gamma, \tau) = \min \begin{cases} M_k(\gamma - 1, \tau - 1) + WA_2 \cdot d_k(\gamma - \frac{1}{2}, \tau) \\ + WA_1 \cdot d_k(\gamma, \tau) \\ M_k(\gamma - \frac{1}{2}, \tau - 1) + WB \cdot d_k(\gamma, \tau) \\ M_k(\gamma - \frac{1}{2}, \tau - 2) + WC_2 \cdot d_k(\gamma, \tau - 1) \\ + WC_1 \cdot d_k(\gamma, \tau) \end{cases} \quad (18)$$

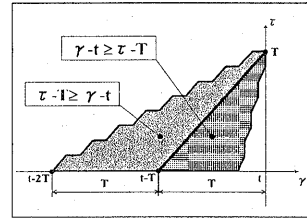


図 4: CDP-S の式の場合分け

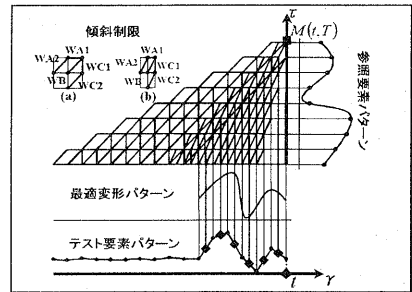
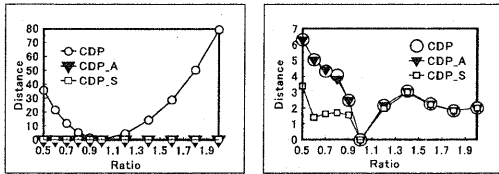


図 5: 改良型 CDP の探索域

図 6(a), (b) は従来型 CDP (同図中の CDP) と CDP-A、CDP-S との性能比較を示したグラフである。但し、CDP-S には CDP-A の機能も持たせてある。図 6(a) は基準となる参照要素パターンの値範囲に対するテスト要素パターンの値範囲の比を横軸に採り、比の異なるテスト要素パターンに対する各 CDP の出力値を縦軸に採ったグラフである。このグラフから CDP は参照要素パターンの値範囲からずれていくと、つまり、比 1 からずれていくと出力値が急激に大きくなっていき、値範囲の違いに対応できていないことが分かる。一方、CDP-A、CDP-S は出力値が小さいままで、値範囲の違いに対応できていることが分かる。図 6(b) は基準となる参照要素パターンの時間範囲に対するテスト要素パターンの時間範囲の比を横軸に採り、縦軸は (a) の場合と同じグラフである。こ

のグラフから CDP、CDP-A は、時間範囲が比 1 より大きいとき、出力値はそれほど大きくなりませんが、比 1 より小さいときは出力値が急激に大きくなっていき、短時間のパターンに対応できていないことが分かる。一方、CDP-S は出力値が他の二つの CDP に対して約半分に抑えることができおり、短時間のパターンに対応できていることが分かる。CDP-S は従来型 CDP の機能を保ちながら短時間のパターンを安定に認識できると言える。



(a) 値範囲の違い (b) 時間範囲の違い

図 6: 各 CDP の性能比較

4.4 ノイズに対応できる CDP

頭部ジェスチャの動画画像から得られる実際のテスト要素パターンには CDP で誤認識を招くようなノイズも含まれる。このノイズに対応するため、図 7 に示す復元性を持つ進路制限を実現する CDP の傾斜制限を考える。ここで、復元性とは図 8 に示すように、参照要素パターンの時系列データを終点から始点へと見てやったときに、値軸方向に上がり続けたり、下がり続けたりすることのないよう、常に参照要素パターンの値に向かう傾向を持つ性質のことを言う。

この傾斜制限は、同図中にあるような従来型 CDP の傾斜制限を三層構造にしたような形になっている。中央の層 (同図中、Middle Layer) は通常どおりに参照要素パターンが登録されている CDP の傾斜制限である。その左右の層 (同図中、Upper Layer, Lower Layer) は参照要素パターンの時系列の値がノイズの標準偏差だけ大きい時系列パターン (以下、上参照要素パターンと呼ぶ) と小さい時系列パターン (以下、下参照要素パターンと呼ぶ) がそれぞれ登録されている。累積距離は中央の層と、左右の層とはその取り方が異なる。中央の層では左右の層からも始点の乗り入れが許される。左右の層では自身以外の層からしか始点を取ることができない。

以上のことを次に具体的に式で表現する。テスト要素パターンと参照要素パターン、上参照要素パターン、下参照要素パターンとのそれぞれの距離 $d_k(\gamma, \tau; n \mid n \in \{M, U, L\})$ を次式のように定義する。

$$\begin{cases} d_k(\gamma, \tau; M) \equiv |i_k(\gamma) - p_k(\tau)| \\ d_k(\gamma, \tau; U) \equiv |i_k(\gamma) - \{p_k(\tau) + S_k\}| \\ d_k(\gamma, \tau; L) \equiv |i_k(\gamma) - \{p_k(\tau) - S_k\}| \end{cases} \quad (19)$$

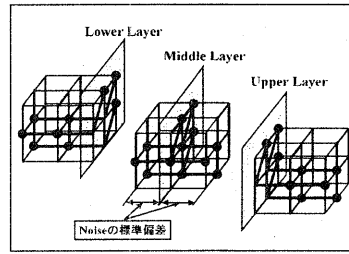


図 7: CDP-N の傾斜制限

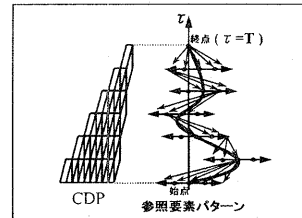


図 8: 復元性を持つ進路制限

ここで、 S_k は k 要素でのノイズの標準偏差である。この CDP では、 $(\gamma, \tau; n \mid n \in \{M, U, L\})$ における累積距離 $M_k(\gamma, \tau; n \mid n \in \{M, U, L\})$ を次のように定義する。初期条件を式 (20) とする。

$$M_k(-1, \tau; n) = M_k(0, \tau; n) = \infty \quad (1 \leq \tau \leq T) \quad (20) \quad (n \in \{M, U, L\})$$

$\gamma \geq 1$ について、 $n = M$ の場合の漸化式は

$$\begin{aligned} (\tau = 1) \\ M_k(\gamma, 1; M) &= WB \cdot d_k(\gamma, 1; M) \end{aligned} \quad (21)$$

$$\begin{aligned} (\tau = 2) \\ M_k(\gamma, 2; M) &= \min \begin{cases} \min_n \{M_k(\gamma-2, 1; n)\} + WA_2 \cdot d_k(\gamma-1, 2; M) \\ +WA_1 \cdot d_k(\gamma, 2; M) \\ \min_n \{M_k(\gamma-1, 1; n)\} + WB \cdot d_k(\gamma, 2; M) \\ \min_n \{M_k(\gamma, 1; n)\} + WC_1 \cdot d_k(\gamma, 2; M), \end{cases} \end{aligned} \quad (22) \quad (n \in \{M, U, L\})$$

$$\begin{aligned} (3 \leq \tau \leq T) \\ M_k(\gamma, \tau; M) &= \min \begin{cases} \min_n \{M_k(\gamma-2, \tau-1; n)\} + WA_2 \cdot d_k(\gamma-1, \tau; M) \\ +WA_1 \cdot d_k(\gamma, \tau; M) \\ \min_n \{M_k(\gamma-1, \tau-1; n)\} + WB \cdot d_k(\gamma, \tau; M) \\ \min_n \{M_k(\gamma-1, \tau-2; n)\} + WC_2 \cdot d_k(\gamma, \tau-1; M) \\ +WC_1 \cdot d_k(\gamma, \tau; M), \end{cases} \end{aligned} \quad (23) \quad (n \in \{M, U, L\})$$

また、 $n = U$ の場合の漸化式は

$$M_k(\gamma, 1; U) = WB \cdot d_k(\gamma, 1; U) \quad (24)$$

$$M_k(\gamma, 2; U) = \min \begin{cases} \min_n \{M_k(\gamma-2, 1; n) + WA_2 \cdot d_k(\gamma-1, 2; U) \\ + WA_1 \cdot d_k(\gamma, 2; U)\} \\ \min_n \{M_k(\gamma-1, 1; n) + WB \cdot d_k(\gamma, 2; U)\} \\ \min_n \{M_k(\gamma, 1; n) + WC_1 \cdot d_k(\gamma, 2; U)\}, \end{cases} \quad (25)$$

$(n \in \{M, L\})$

$$(3 \leq \tau \leq T)$$

$$M_k(\gamma, \tau; U) = \min \begin{cases} \min_n \{M_k(\gamma-2, \tau-1; n) + WA_2 \cdot d_k(\gamma-1, \tau; U) \\ + WA_1 \cdot d_k(\gamma, \tau; U)\} \\ \min_n \{M_k(\gamma-1, \tau-1; n) + WB \cdot d_k(\gamma, \tau; U)\} \\ \min_n \{M_k(\gamma-1, \tau-2; n) + WC_2 \cdot d_k(\gamma, \tau-1; U) \\ + WC_1 \cdot d_k(\gamma, \tau; U)\}, \end{cases} \quad (26)$$

$(n \in \{M, L\})$

そして、 $n = L$ の場合の漸化式は

$$M_k(\gamma, 1; L) = WB \cdot d_k(\gamma, 1; L) \quad (27)$$

$$M_k(\gamma, 2; L) = \min \begin{cases} \min_n \{M_k(\gamma-2, 1; n) + WA_2 \cdot d_k(\gamma-1, 2; L) \\ + WA_1 \cdot d_k(\gamma, 2; L)\} \\ \min_n \{M_k(\gamma-1, 1; n) + WB \cdot d_k(\gamma, 2; L)\} \\ \min_n \{M_k(\gamma, 1; n) + WC_1 \cdot d_k(\gamma, 2; L)\}, \end{cases} \quad (28)$$

$(n \in \{M, U\})$

$$(3 \leq \tau \leq T)$$

$$M_k(\gamma, \tau; L) = \min \begin{cases} \min_n \{M_k(\gamma-2, \tau-1; n) + WA_2 \cdot d_k(\gamma-1, \tau; L) \\ + WA_1 \cdot d_k(\gamma, \tau; L)\} \\ \min_n \{M_k(\gamma-1, \tau-1; n) + WB \cdot d_k(\gamma, \tau; L)\} \\ \min_n \{M_k(\gamma-1, \tau-2; n) + WC_2 \cdot d_k(\gamma, \tau-1; L) \\ + WC_1 \cdot d_k(\gamma, \tau; L)\}, \end{cases} \quad (29)$$

$(n \in \{M, U\})$

で与える。

この CDP を CDP-N と呼ぶことにし、CDP-N の探索域の概念図を図 9 に示す。

4.5 改良型 CDP

前述の全ての機能を採り入れると、CDP の認識の反応が敏感になり過ぎ、CDP の出力値が下がりがちになる (この CDP を CDP-N と呼ぶ)。そこで、CDP-A における倍率の最小値を 0.8 とし、認識の反応が少し鈍くなるよう調節する。この CDP を CDP-I と呼ぶことにする。図 10 と図 11 はパターンにノイズが入っているときの CDP-I と従来型 CDP、CDP-A、CDP-S との性能比較を示したグラフである。それぞれのグラフの設定は図 10 が図 6(a) と、図 11 が図 6(b) と同じである。テスト要素パターンに対する出力値は CDP-I が図 10、図 11 の

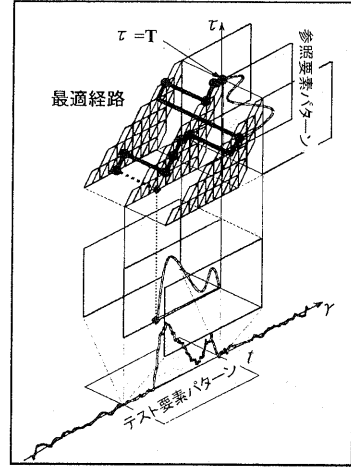


図 9: CDP-N の探索域

どちらのグラフを見ても他の CDP より、ほぼ低い値を示しており、ノイズに対応できていると言える。

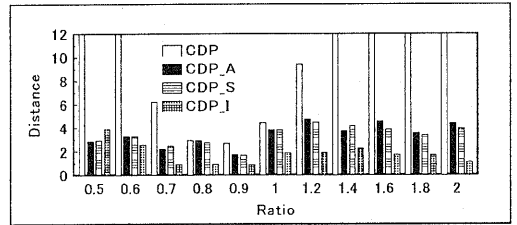


図 10: 値範囲の違いについての各 CDP の性能比較

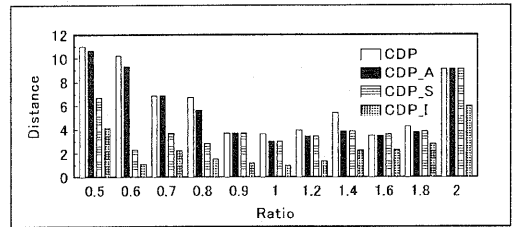


図 11: 時間範囲の違いについての各 CDP の性能比較

図 12 はパターンに実際のノイズを加えたときにおける CDP-I と CDP-A、CDP-N の出力値の挙動である。CDP-N は最も反応が敏感でパターンが入っていない状態のときでも出力値が低く、閾値の設定が困難になる。一方、CDP-I は目的のパターンが入ったときだけ出力値が鋭く下がり、閾値の設定が容易になる。また、同図中にはフレームが数 100 から 140 あたりのノイズだけのと

きと、フレーム数が150あたりから180までのパターンやノイズが何も無い状態であるときのCDPの出力の挙動も含めて示してある。

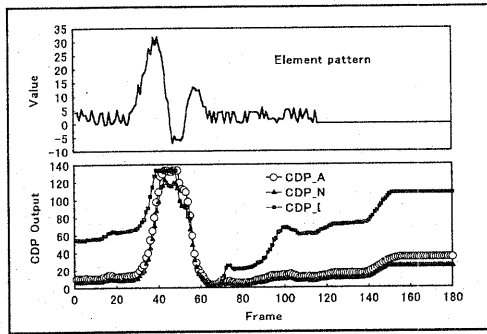


図 12: ノイズがあるときの各 CDP の出力値の挙動

5 主成分分析による計算量の縮小

改良型 CDP は、改良する前に比べて数十倍の計算量が必要で、現在使用しているコンピュータでリアルタイム処理を行うには計算量が多い。このことから、評価する特徴ベクトルの要素数を主成分分析を使用して減らす。また、度合の認識の観点からも少数の要素である特徴ベクトルの方が、評価が容易になる。

具体的手順としては、参照パターンの4要素を主成分分析して得た主成分ベクトルを、参照パターンとテストパターンのそれぞれの4要素に係数として掛ける。そして、4要素の一次結合である新要素のうちパターンの比較評価に必要な2要素だけを選び出す。これにより4要素の特徴ベクトルを2要素の特徴ベクトルに次元圧縮する。ここで、参照パターンの新要素によるパターンを新参照パターン、その要素を新参照要素パターン、テストパターンの新要素によるパターンを新テストパターン、その要素を新テスト要素パターンと呼ぶことにする。主成分分析をするときは要素パターンの値範囲に関する情報を損なわないために変量の正規化は行わない。表1、表2はそれぞれ首を横に振る単軸動作であるジェスチャ”No”と首を回す複数軸動作であるジェスチャ”Relax”の参照パターンを主成分分析した結果である。

ジェスチャが入力されて参照パターンが対象とする取得すべきテストパターンに遭遇すると、新テスト要素パターンに新参照要素パターンと同じ固有のパターンが現れる。このパターンどうしをマッチングする。

主成分分析は寄与率の大きい順に主成分を選んでいき、データに内在する主要な変動(複数の変量が互いに従属して生ずる変動で動きの大きいもの)を見出していく[5]。

表 1: ”No”の場合の主成分分析

主成分	第1主成分	第2主成分	第3主成分	第4主成分
寄与率	0.994	0.00429	0.00115	0.000451
変量	各主成分ベクトル			
$\theta_x(t)$	-0.0159	0.719	-0.653	0.238
$\theta_y(t)$	-0.998	0.0251	0.0416	-0.0286
$\theta_z(t)$	-0.0472	-0.669	-0.741	-0.0182
$a_x(t)$	-0.0264	-0.188	0.148	0.971

表 2: ”Relax”の場合の主成分分析

主成分	第1主成分	第2主成分	第3主成分	第4主成分
寄与率	0.577	0.376	0.0307	0.0157
変量	各主成分ベクトル			
$\theta_x(t)$	-0.829	-0.323	-0.322	-0.323
$\theta_y(t)$	-0.247	0.839	0.225	-0.430
$\theta_z(t)$	0.313	-0.405	0.374	-0.773
$a_x(t)$	0.392	0.169	-0.840	-0.336

逆に寄与率の小さい順に主成分を選らんでいくとデータに内在する不変動を見出すことができる。つまり、寄与率の大きい主成分と小さい主成分を特徴ベクトルの新要素に選ぶと主要な変動要素と不変動要素に分けることができる。

新要素の選び方は主成分の寄与率の状態により2種類あり、その内訳を説明していく。寄与率の一つの主成分に偏るとき、2要素として第1主成分と第4主成分を選ぶ。このとき不変動要素は特に意味のある情報であり、不変動要素の導入で主要な変動要素にしか注目していない場合に違いが区別できなかったテストパターンを区別できるようになる。表1からジェスチャ”No”の参照パターンがこれにあたり、その実例を図13の中の①に挙げる。

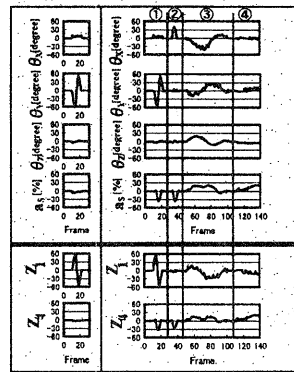


図 13: ”No”の場合の新テスト要素パターン

同図中の左上部はジェスチャ”No”の参照パターンである。同図中で右上部の①のパターンはジェスチャ”No”の参照パターンと要素 a_x を除いて同じである。しかし、パターンとしては別のものとして捉えなければならない。左下部はジェスチャ”No”の新参照パターンであるが、そ

の第1主成分である変動要素 z_1 だけに注目すると、①の
 パターンの z_1 と変わらず区別がつかない。しかし、ジェ
 スチャ”No”の新参照パターンでは不変動であるべき第4
 主成分の要素 z_4 が、①のテストパターンの新要素 z_4 で
 は変動しているため、不変動要素を注目することで、こ
 の新参照パターンと新テストパターンの区別ができる。
 ジェスチャ”No”と異なるジェスチャである②、③、④の
 新テストパターンもジェスチャ”No”の新参照パターンと
 は異なるパターンになるのでジェスチャ”No”と区別が
 つく。表1から第1主成分である要素 z_1 と第4主成分で
 ある要素 z_4 は次式(30)のようになる。

$$\begin{cases} z_1(t) = -0.0159 \cdot \theta_x(t) - 0.998 \cdot \theta_y(t) \\ \quad - 0.0472 \cdot \theta_z(t) - 0.0264 \cdot a_s(t) \\ z_4(t) = 0.238 \cdot \theta_x(t) - 0.0286 \cdot \theta_y(t) \\ \quad - 0.0182 \cdot \theta_z(t) + 0.971 \cdot a_s(t) \end{cases} \quad (30)$$

寄与率が偏らず分散する場合、寄与率の大きい上位二
 つの主成分、つまり第1主成分と第2主成分を取り、特
 徴ベクトルの新要素とする。表2からジェスチャ”Relax”
 の参照パターンがこれにあたり、その実例を図14に挙
 げる。

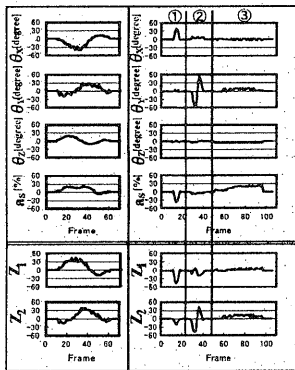


図14: ”Relax” の場合の新テスト要素パターン

同図中の左下部はジェスチャ”Relax”の新参照パター
 ンであるが、寄与率の大きい第1主成分と第2主成分を
 新要素に選んだため、2種類の変動が新要素 z_1 と z_2 に
 現れていることが分かる。そして、ジェスチャ”Relax”と
 異なるジェスチャである①、②、③の新テストパターン
 にジェスチャ”Relax”の新参照パターンと同様なパター
 ンは現れないのでジェスチャ”Relax”と区別がつく。表2
 から第1主成分である要素 z_1 と第2主成分である要素
 z_2 は次式(31)のようになる。

$$\begin{cases} z_1(t) = -0.829 \cdot \theta_x(t) - 0.247 \cdot \theta_y(t) \\ \quad + 0.313 \cdot \theta_z(t) + 0.392 \cdot a_s(t) \\ z_2(t) = -0.323 \cdot \theta_x(t) + 0.839 \cdot \theta_y(t) \\ \quad - 0.405 \cdot \theta_z(t) + 0.169 \cdot a_s(t) \end{cases} \quad (31)$$

以上の事実から特徴ベクトルの要素を4個から2個へ
 減らすことで計算量を半分にできる。

6 おわりに

様々な傾斜制限[6]を用いれば、パターンのより大きな
 時間軸方向の伸縮に対してCDPが対応できることが分
 かった(CDP-S)。また、ノイズに対してもCDPに新たな
 軸方向を設けることで対応できることが分かった(CDP-
 N)。(図1参照)

改良したCDP、特にCDP-Aは現在使用しているコン
 ピュータでリアルタイム処理を行うには計算量が非常に
 多い。CDP-Aで使われる13段階の倍率が全て必要かど
 うか、実験による再検討が必要であると考えられる。

今後CDPの更なる改良や計算量の縮少を進めていき、
 ジェスチャ入力精度の向上と実用性を高めていく予定
 である。

参考文献

- [1] 呉海元, 小林弘知, 陳謙, 塩山忠義, 島田哲夫: 色彩動
 画像からの頭部ジェスチャ認識システム, 情報処理学
 会論文誌, Vol.40, No.2, pp.577-584(1999).
- [2] 高橋勝彦, 関進, 小島浩, 岡隆一: ジェスチャー動画
 のスポット認識, 電子情報通信学会論文誌 D-
 II, Vol.J77-D-II, No.8, pp.1552-1561, (1994).
- [3] T.Takahashi and R.Oka: Spotting Recognition of
 Human Gestures from Time-Varying Images, Proc.
 2th Int. Conf. On Automatic Face- and Gesture-
 Recognition (ICAFGR), pp.318-322(1996).
- [4] 浦 昭二, 小野勝章, 川嶋弘尚, 芳賀敏郎 編: コン
 ピューティングと数学, 迫江博昭, 音声認識におけ
 る動的計画法への応用, pp.131-142, 共立出版(1983).
- [5] 河口至商, 多変量解析入門 I, 森北出版(1973).
- [6] H.Sakoe and S.Chiba: Dynamic Programming Al-
 gorithm Optimization for Spoken Word Recogni-
 tion, IEEE Trans. on Acoustics, Speech, and Signal
 Processeing, Vol.Assp-26, No.1, pp.43-49(1978).