

論理判定型 DP マッチングによる類似区間検出

森 明慧[†] 内田 誠一[†] 倉爪 亮[†] 谷口倫一郎[†]
長谷川 勉[†] 迫江 博昭[†]

Detection of Similar Sub-Sequence by Logical DP Matching

Akihiro MORI[†], Seiichi UCHIDA[†], Ryo KURAZUME[†], Rin-ichiro TANIGUCHI[†],
Tsutomu HASEGAWA[†], and Hiroaki SAKOE[†]

あらまし 本論文では、論理判定型 DP マッチングによる類似区間検出手法について提案する。論理判定型 DP マッチングとは、サポートと呼ばれる論理関数を基準として用いて二つのパターン間の非線形マッチングを行うアルゴリズムである。本手法の特徴は、パターン間に複数存在する類似区間の始端及び終端をマッチングの過程で最適に決定していく点にある。また、本手法の有効性を評価するための一応用として、ジェスチャの基本動作抽出についても検討する。実験の結果、本手法の基本的な性能を示すことができた。

キーワード 動的計画法, DP マッチング, 論理判定, ジェスチャ

1. ま え が き

本論文では、時系列パターンの類似区間の検出について検討する。類似区間とは、2 パターン中に存在する似通った区間を指す。例えば、音声データにおいて、二つの時系列パターンそれぞれに同じ単語が含まれている場合、この同じ単語の部分が類似区間となる。また、動画データにおいて、一つの時系列パターン中に何度も同じシーンが登場する部分も類似区間となる。類似区間の検出は、例えばある特定のパターンが他のパターン中に何回存在するかを調べる、あるいは頻繁に登場する部分パターンを抽出する、といった目的で用いられる。

類似区間の検出法は、クラスタリング等によるシンボル化が必要な方法 [1] ~ [3] と特徴量レベルで直接検出が可能な方法 [4] ~ [6] に大別できる。前者については、Bobick ら [1] によってジェスチャを特徴空間上の軌道として考え、同一パターンにおいて主曲線、及びその確率的な変動量を計算し、これをもとにパターンを分割・クラスタリングすることで類似区間を検出する手法が提案されている。また、Zhao ら [2] はジェス

チャの各フレームにおける特徴量をクラスタリングした後に、得られたクラスタに対して MDL (Minimum Description Length) 原理を適用することで類似区間を得ている。向井ら [3] は複数のジェスチャからシンボル化を介して動的に一つのネットワークを構築することでジェスチャ間の類似動作を抽出する手法を提案している。ただし、これらの手法は類似区間検出の際に、その始端及び終端がシンボルの境界に限定されてしまうという問題がある。

本論文では、特徴量レベルで直接類似区間を検出する手法として論理判定型 DP (Dynamic Programming) マッチング法を提案する。本手法では、サポートと呼ばれる論理関数を基準として用いて二つのパターン間の非線形マッチングを行う。サポートとは、各パターンにおける特徴間の類似度を論理値として定義したもので、このサポートが真の値をとり続ける区間が類似区間であると考えることができる。論理判定型 DP マッチングの最大の特長は、パターン中に存在する複数の類似区間を同時に検出できることである。特に、サポートの利用により類似区間の始端及び終端をマッチングの過程で最適に決定できる点は、ユークリッド距離を用いた従来の DP マッチングでは実現困難な特長である。なお、本手法も従来法と同様に、パターンの非線形伸縮の吸収が可能である。本手法と同じく特徴量レベルの類似区間検出法である文献 [4] ~ [6] との

[†]九州大学システム情報科学研究所, 福岡市
Faculty of Information Science and Electrical Engineering,
Kyushu University, 744 Motooka, Nishi-ku, Fukuoka-shi,
819-0395 Japan

比較については後に詳述する。

以下、本論文では本手法の詳細について述べる。また、本手法の一応用として、論理判定型 DP マッチングによるジェスチャの基本動作、すなわちジェスチャを構成する要素単位の抽出についても検討する。加えて本手法の有効性を評価すべく実際にジェスチャの類似区間検出実験、及び類似区間検出によって得られた基本動作の評価実験を行った結果について述べる。

2. 論理判定型 DP による類似区間検出

本章では、提案する論理判定型 DP を用いた類似区間の検出手法について、その詳細を述べる。また、検出精度を評価するために行った実験とその結果について述べる。

2.1 アルゴリズム

本節では、論理判定型 DP のアルゴリズムの詳細について述べる。以下では、類似区間の検出対象となる二つの時系列パターンを $A = a_1, \dots, a_i, \dots, a_I$ 及び $B = b_1, \dots, b_j, \dots, b_J$ で表す。また、特徴量 a_i と b_j の距離を

$$d(i, j) = (a_i - b_j)^2 \quad (1)$$

で定義する。

2.1.1 サポート

各時系列パターンを特徴空間上の軌道として表すと、パターン間の類似区間は図 1 のように 2 軌道の近接区間として考えることができる。この近接/非近接を定義するために、サポートと呼ぶ論理関数を次式で定義する。

$$s(i, j) = \begin{cases} 1 & \text{if } d(i, j) < \theta_1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

ここで θ_1 は定数である。すなわち、サポートとは a_i と b_j の距離がしきい値 θ_1 より小さければ 1 をとり、 θ_1 よりも大きければ 0 をとる論理関数である。サポートの値が 1 となるフレーム間対応を図 1 中に破線として示す。

論理判定型 DP マッチングは、類似区間としてサポート値 1 となるフレーム間対応が続く区間を検出するためのアルゴリズムである。今、図 2 のような $i-j$ 平面を考え、各ノード (i, j) においてそのサポート $s(i, j)$ が計算されているとすると、この類似区間検出問題は、サポート値 1 のノードを結ぶ経路の検出問題に帰着する。この経路は A, B の類似部分パターン

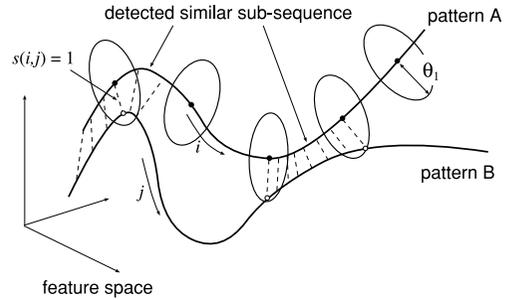


図 1 類似区間の検出の基本的な考え方。破線はサポートの値が 1 となるフレーム間対応を表す。

Fig. 1 Basic idea for similar sub-sequence detection.

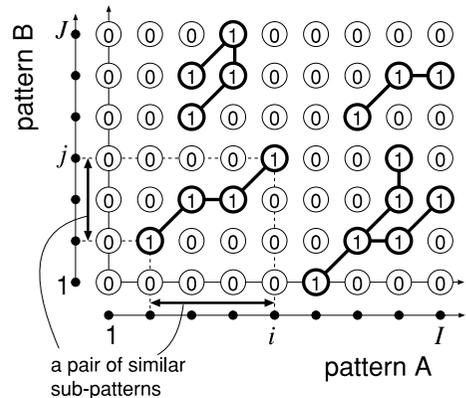


図 2 類似区間の経路表現。各ノード i, j の数字 ($\in \{0, 1\}$) はサポート $s(i, j)$ の値を表す

Fig. 2 Representation of similar sub-sequence by a matching path.

のマッチングを意味する。また、この経路は、類似区間の発生状況に応じて $i-j$ 平面中に複数存在する可能性があり、更にそれらが分岐・統合することもある。

2.1.2 漸化式計算と始端・終端の決定

本手法においては経路検出のために、以下の DP 漸化式を $i = 1$ から I まですべての j において計算する。

$$g(i, j) = s(i, j) \wedge \left(\bigvee_{k=1}^3 g_k \right) \quad (3)$$

ここで、表記の都合上、

$$\begin{aligned} g_1 &= g(i-1, j-1) \\ g_2 &= s(i, j-1) \wedge g(i-1, j-2) \\ g_3 &= s(i-1, j) \wedge g(i-2, j-1) \end{aligned}$$

としている。上式中の $g(i, j)$ は論理値 ($\in \{0, 1\}$) をとる。また g_1, g_2, g_3 も論理値であり、それぞれ図 3 中

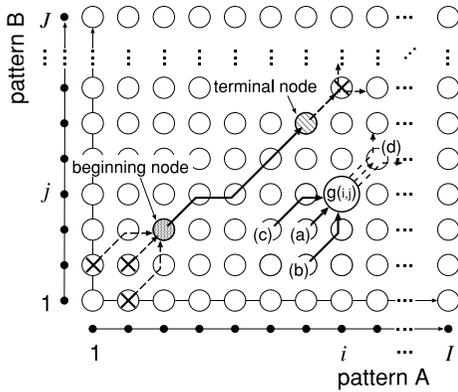


図 3 漸化式計算と始端・終端決定
Fig. 3 Calculation of DP recurrence and detection of beginning/terminal node.

の (a), (b), (c) を経由してノード (i, j) に到達し得る経路が存在するか否かを表している。したがって漸化式 (3) は、サポート $s(i, j)$ が 1 ならば、同図の (a), (b), (c) のいずれかを經由する経路にノード (i, j) をつなげ、 (i, j) に到達する経路を生成する処理に相当する。逆に (i, j) のサポートが 0、若しくは (a), (b), (c) を經由する経路がない ($g_1 = g_2 = g_3 = 0$) 場合、 $g(i, j)$ は 0 になり、 (i, j) に到達する経路は生成されない^(注1)。

類似区間を検出するためには、経路の始端・終端も決める必要がある。そこで、本手法では式 (3) の計算に同期して、以下のように始端・終端の検出を行う。まず次の論理関数が 1 となる場合、ノード (i, j) を始端とみなし、強制的に $g(i, j) = 1$ とする。

$$s(i, j) \wedge \neg \left(\bigvee_{k=1}^3 g_k \right) \quad (4)$$

これは、ノード (i, j) はサポート 1 であるが、その直前には類似区間が存在しない場合、すなわち図 3 の (a), (b), (c) いずれの経路においてもサポート 0 となる場合に相当する (図 3 中の beginning node)。なお、周辺ノード $(i, 1), (1, j)$ については、 g_k が存在しないため、 $s(i, j)$ の値が 1 ならば始端すなわち $g(i, j) = 1$ とおく。

次に、終端について考える。ノード (i, j) に対してノード $(i+1, j+1)$ は、今後 (a), (b), (c) のどの経路が選ばれても通過するノードである (図 3 (d))。このため、このノードがサポート 0 の場合はノード (i, j) で経路が途切れることになる。そこで、類似区間の終

端を次の論理関数が 1 となるノードとして定義する。

$$g(i, j) \wedge \neg s(i+1, j+1) \quad (5)$$

以上のように、本手法は式 (3) を基本として (4), (5) の処理を付加的に加えた DP マッチングアルゴリズムといえる。計算量は通常の DP マッチングと同じ $O(IJ)$ である。

2.1.3 バックトラックによる類似区間の確定

式 (5) によって定義された各終端からバックトラック処理を行うことで、2 パターン間の複数の類似区間を同時に確定できる。すなわち、式 (5) で定義した終端からスタートし、ノードを逆にたどることによってサポート値 1 をもつ (i, j) の列として経路 (部分マッチング) を求めることができる。

図 2 にも示したように、経路は分岐・統合することがある。すなわち、一部が重複したマッチング経路が複数存在する場合がある。こうしたすべての経路を類似区間とする考え方もあるが、以下ではいずれか一つだけを代表として選出し、ある始端・終端にかかわる経路は一つだけとする。選出の基準には様々なものが考えられるが、以下の実験ではその経路上の $d(i, j)$ の累積値を用いた。

こうして得られたマッチング経路のうち、経路長がしきい値 θ_2 を超えるものを最終的に類似区間として検出する。このようなしきい値処理を行うのは、ごく短い区間におけるマッチングを雑音とみなして除去するためである。

2.2 類似区間検出実験

本手法の性能を評価するために、ジェスチャを対象としての類似区間検出実験を行った。

2.2.1 実験諸元

実験に使用するデータとして、本実験では 18 種類のジェスチャを想定し (図 4 にその一部のジェスチャについて、特徴的なフレームを示す)、それぞれ成人男性 1 人によって得られた各 6 パターン、計 108 パターンのジェスチャパターンを用意した。これら 108 パターンの平均フレーム長は約 89 フレームであった。各フレームにおける特徴量は、顔の位置を基準とし

(注1): この定式化では、 $\frac{1}{2}$ 未満及び 2 倍より大きな速度の変動を補償することができない。しかしながら、一般に音声認識やジェスチャ認識の分野ではこの範囲に限定された DP マッチングで一定の成果を挙げており、本手法においても特に問題にはならないと思われる。実際、後述の実験ではこの問題によって類似区間が分断されてしまうといった例は見られなかった。また、速度の変動がこの範囲外となる場合であっても、式 (3) で定義される局所経路の形状を変更することで対処可能である。

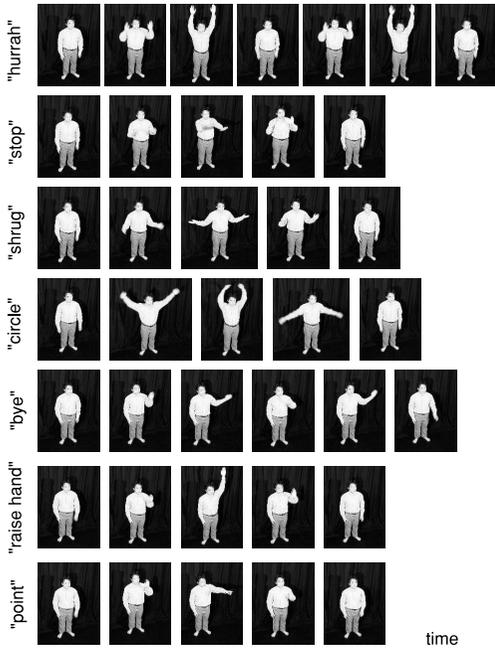


図 4 実験で想定したジェスチャの一部
Fig. 4 Snapshot of assumed gestures.

た右手先及び左手先の三次元位置からなる六次元特徴ベクトルである．この位置特徴は，(i) まずユーザの前方におかれた 2 台の IEEE1394 カメラ (Sony 製 DFW-X700, 15 フレーム/秒) により距離画像をステレオ計測し，(ii) 次に肌色検出により両手と顔部分を同定することで自動取得したものである．

本実験ではしきい値をそれぞれ $\theta_1 = 200.0$, $\theta_2 = 8$ に設定した．これらの値は実験的に定められたものである．なお $\theta_1 = 200.0$ とは，2 姿勢の間で両手先の位置の違いがおよそ 0.2 m 以下であればサポートを 1 とすることを意味する．また， θ_2 については，500 ms 以上の類似区間を抽出することを意味する．このように， θ_1 , θ_2 は物理量と対応づけながら設定できる．

2.2.2 類似区間の検出結果

実験結果の一部を図 5 に示す．図 5(a) は，ジェスチャ“left-hand bye”の同じパターン間 (すなわち $A = B$) に対して本手法を適用した結果である．同じパターンであるから，対角線上に経路が求まるのは当然であるが，それ以外にも二つのマッチング経路，すなわち類似区間が検出されているのが分かる．この類似区間は，ジェスチャ“left-handbye”において手を左右に振る繰返し動作であり，すなわち本手法により 1

回のジェスチャ中の 2 度の繰返しを同時に正しく検出できていることが分かる．

図 5(b) は，同じカテゴリ“left-hand bye”ではあるがデータとしては異なる 2 パターン間の類似区間を検出したものである．非線形伸縮があっても図 5(a) の場合と同様，繰返し部分が正しく検出されていることが分かる．

図 5(c) は，二つの異なるジェスチャ“right-hand bye”と“raise right-hand”の類似区間を検出した結果を示している．これらのジェスチャは，ジェスチャの冒頭部分に「右腕を肩の高さまであげる」，終了部分に「右腕を肩の高さから下げる」という二つの共通動作をもつ．実際に本手法を適用した結果，この共通動作を同時に類似区間として検出することに成功している．

図 5(d) は，ジェスチャ“hurrah” (万歳) と“circle”の類似区間を検出した結果を示している．これらのジェスチャはともに両手を頭上に上げる動作からなっているが，頭上に至る両手先の運動軌跡が異なるため，共通した動作が存在しない．実験では，誤検出することなく正しい結果を得ることができた．

図 5(e) は，異なるジェスチャ“hurrah”と“stop”に対する本手法の適用結果であり，失敗例である．実験の結果，冒頭の共通動作部を類似区間として検出することができなかった (図 5(e) 中の矢印で示した部分)．その理由としては，検出されたマッチング経路がしきい値 θ_2 より短くなってしまい，雑音として処理されてしまったためと考えられる．一歩進めて経路が短くなった原因まで考えると，サポートのしきい値 θ_1 が厳しすぎ，共通部末端の大きな変動に対応できなかった可能性も考えられる．

2.3 従来法との比較

本手法は特徴量レベルで直接類似区間の検出が可能な手法の一つである．1. で述べたように，こうした手法は従来よりいくつか提案されている．以下，本節では従来法と本手法との原理的な違いについて述べる．

2.3.1 線形フィッティング [4] との関係

本項では，DP 平面上での線形フィッティングによってパターン間の類似区間を検出する手法 [4] と本手法とを比較する (文献 [4] 中では DP 平面を時系列相関行列と呼ぶ)．この線形フィッティングによる手法では，本手法のサポートと同様の方式により DP 平面上において，類似フレームを見つける．次にその類似フレームがなす連結成分ごとに (すなわち図 5 中央列の

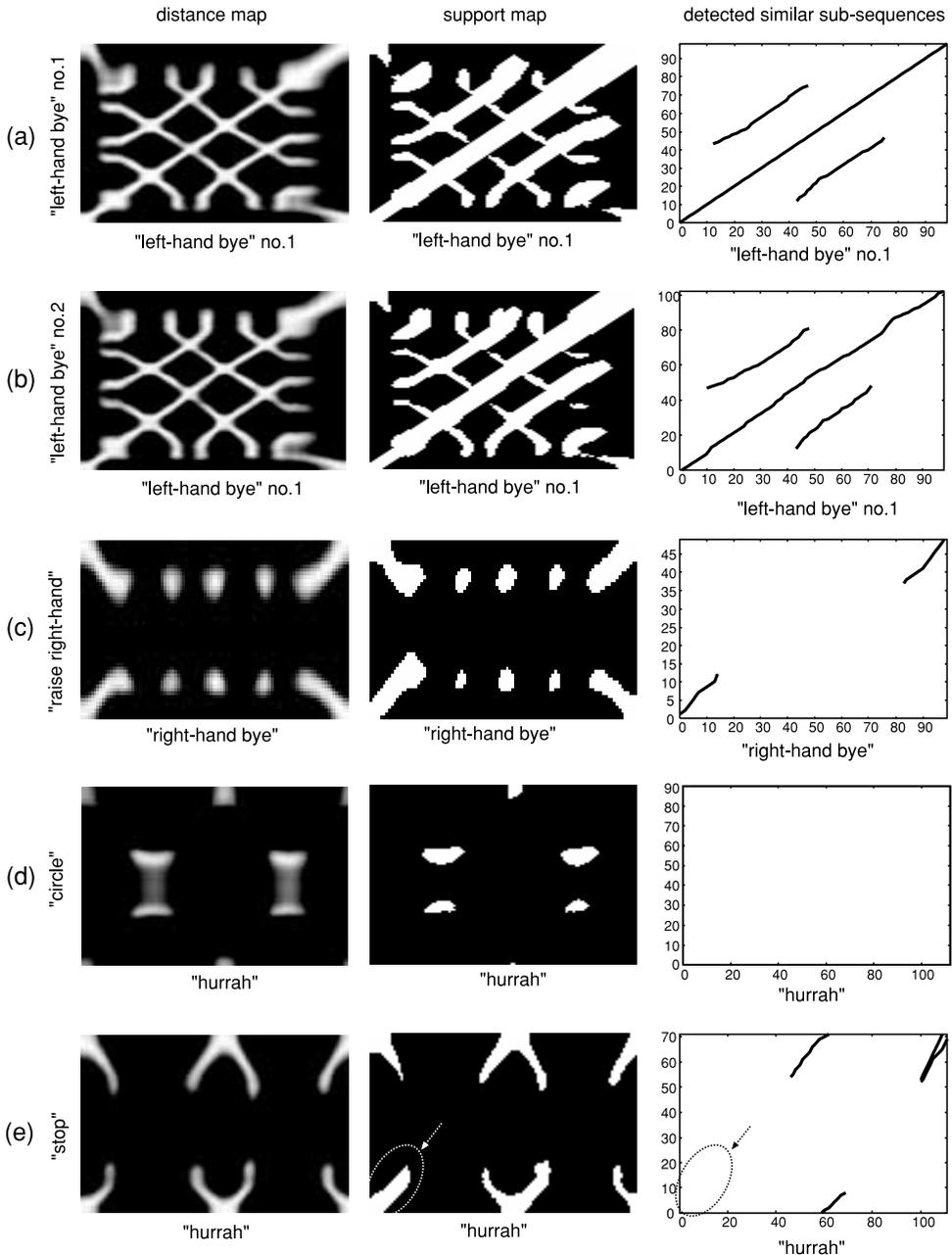


図 5 本手法によるジェスチャパターン間の類似区間検出結果．左から順に，2 パターン間の距離画像（明るいほど距離が小さい），サポート画像（白い部分がサポートを満たすノード），類似区間を表すマッチング経路である．

Fig. 5 Experimental result of similar sub-sequence detection.

サポート画像の各白領域に)，最小二乗法による直線フィッティングを行い，類似区間を定める．類似区間の始端及び終端は，連結成分のうち時間が最小・最大

となるノードから求めることができる．

しかし，線形フィッティングによる手法には問題点が存在する．上記の方法では図 5 (a) や図 5 (b) のよう

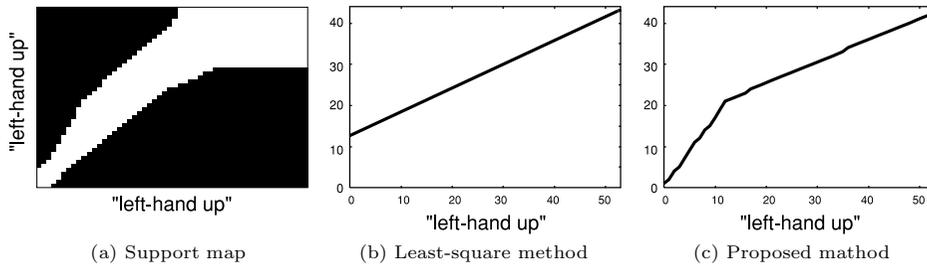


図 6 非線形伸縮への対応実験の結果 . (a) サポート画像 . (b) 線形フィッティングによる類似区間検出結果 . (c) 本手法による類似区間検出結果

Fig. 6 Comparison of experimental result between conventional linear-fitting algorithm and proposed algorithm.

に類似フレームが DP 平面上で複雑に連結している場合、正しい結果を得るのは困難である . また、図 5 (e) のように繰返し動作などによって、時間が逆進するような対応区間が連結成分に含まれる場合にも、検出に失敗してしまう . 更に、類似動作間に速度の非線形変動が存在する場合には、正しいフレーム間対応を得ることができない . これら線形フィッティングを用いる場合の問題点は、DP アルゴリズムを応用した本手法において原理的に改善されている .

実際に同じ「左手を頭上まで上げる」という動作について類似区間検出の比較実験を行った結果を図 6 に示す . 図 6 (a) は、横軸方向は一定の速度のパターン、縦軸方向は動作の冒頭部は高速だが次第に減速していくというパターンについてのサポート画像である . これらのパターンでは動作速度が非線形に変動しているため、上記の線形フィッティングによるマッチングでは類似区間を正しく検出できない . 実験の結果、図 6 (b) に示すように縦軸パターンの一部を類似区間として検出できなかった . 一方、本手法ではパターン間の非線形変動を DP マッチングのアルゴリズムによって吸収することで、図 6 (c) に示すとおりパターン全体を類似区間として検出することができた .

2.3.2 RIFCDP [5], [6] との関係

DP を用いた類似区間検出手法として、過去に伊東らによって Reference Interval-Free 連続 DP (RIFCDP) が提案されている [5] . 本項では、この RIFCDP、及びその発展型である重み減衰型 RIFCDP [6] について本手法と比較する .

RIFCDP とは、連続 DP の拡張にあたり、標準パターンの任意の区間と入力パターンの任意の区間のマッチングを行う手法である . 具体的には、入力パターンの任意のフレーム i と標準パターンの任意のフ

レーム j において、その対応ノード (i, j) に至る最適経路上の任意のノード $(i-t, j-\tau)$ を始端とした累積距離を t, τ を変化させながら各々計算する . この累積距離を、各始端からノード (i, j) までの整合度として用いることで類似区間を探索する .

RIFCDP はその有用性から様々な応用がなされているが、重要な問題を抱えている . すなわち、計算量とメモリ量が非常に大きいという問題である . また、最適性原理を満たしていないという問題も指摘されている [6] . これに対し、本手法は通常の DP と同等の計算量及びメモリ量であるため、RIFCDP に比べ非常に少ない計算量とメモリ量しか必要としない . また、本手法で検出された類似区間は最適性原理を満たす .

重み減衰型 RIFCDP [6] は、RIFCDP のもつ問題を改善した手法である . 具体的には、各フレームにおける局所距離の影響を過去にさかのぼるに従って指数関数的に減少させたものを累積距離として計算し、これを類似区間の整合度として使用することで計算量及びメモリ量を大幅に軽減する . 結果的に、その計算量及びメモリ量は本手法や通常の DP と同等の $O(IJ)$ となる .

ただし、この手法は区間検出問題の近似解法であり、また、実際筆者らも述べているように、類似区間が数フレーム遅れて検出されるという問題をもつ . 更に、徐々に重みを減衰させることから、短い類似区間と非常に長い類似区間とを同等の精度で同時に検出することはできない . これに対し、本手法では局所距離に対してサポートを適用するため、類似区間整合度に過去の影響を受ける重み減衰型 RIFCDP のように、類似区間の検出が遅れることはない . また、短い類似区間と非常に長い類似区間を同等の精度で同時に検出することが可能である .

3. 基本動作抽出への応用

本章では、前章で提案した類似区間の検出手法を基本動作の抽出に応用する方法について述べる。更に、本手法の評価のために行った実験について述べる。

3.1 基本動作

基本動作とはジェスチャを構成する単位要素である。基本動作はこれまでも様々に定義されてきた。大崎ら [7] は速度変化の大きな時点を基準としてジェスチャを分割し、得られたセグメントを DP 距離に基づいてクラスタリングすることで基本動作を得ている。澤田ら [8] も同様に加速度を用いて手話を基本動作に分解する方法を提案している。これらの手法は、個々のジェスチャについてその物理的な特性から基本動作を求めるものである。また、Sanger [9] はジェスチャ時系列の最適基底として基本動作を定義している。Fodら [10] は以上のアプローチを混合した方法を提案している。

これらに対し、本論文では論理判定型 DP を用いて得られた複数のジェスチャに共通する部分ジェスチャを基本動作として抽出することを考える。その際、類似区間として検出されなかった部分、すなわち複数のジェスチャに共有されずにある特定のジェスチャのみに現れる部分ジェスチャも併せて基本動作と考える。これにより任意のジェスチャは基本動作の列に分解される。この手法は、従来法で用いられているジェスチャの速度変化のような物理的な特性によらないため、不必要な分割を避けることができる。また、想定しているジェスチャ集合に応じて、適切な基本動作が得られるという特徴をもつ。このため、マッチングベースの認識において、標準パターンが部分的に重複するといった冗長性を取り除くことができる。

3.2 基本動作の抽出アルゴリズム

基本動作の抽出においては、学習用ジェスチャパターンから任意の 2 パターンを取り出し、論理判定型 DP マッチングを行い類似区間を検出する。これをすべての 2 パターンの組について行う。こうして検出された類似区間は冗長である。なぜならば、複数のジェスチャに共通した区間は何度も重複して検出されるためである。そこで、この重複した基本動作を統合するために、後述する実験においては次の処理を行っている。すなわち、得られた基本動作の候補間で通常の DP マッチングを行い、それぞれマッチングコストを求め、このマッチングコストがしきい値 θ_3 よりも小

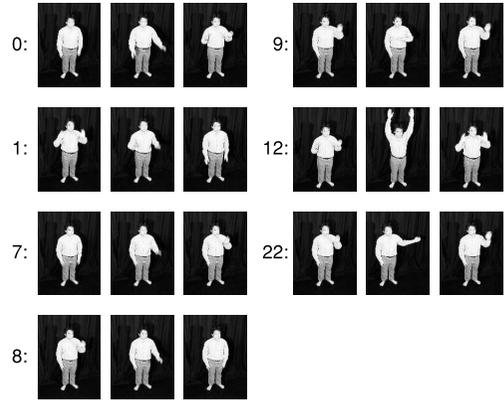


図 7 得られた基本動作の例（括弧内の数字は基本動作番号を表す。0：両手を肩まで上げる。1：両手を肩から下げる。7：左手を肩まで上げる。8：左手を肩から下げる。9：左手を内側に振る。12：両手を上まで上げる。22：左手を外側に振る）

Fig. 7 Snapshot of extracted motion primitives.

さな場合、それらの部分ジェスチャを重複した基本動作とみなし、これらの部分ジェスチャのうち一つを代表して基本動作とする。

3.3 基本動作抽出実験

2.2 の実験結果をもとに、上記の手法を用いて実際に基本動作の抽出を行った。18 種のジェスチャ各 6 パターンのうち各 1 パターンだけを学習パターンとし、それらすべての学習パターンについて総当たりで上記の手法を適用した。その結果、重複を含めて総計 142 の類似区間を検出した。これらの類似区間に、各ジェスチャに固有の姿勢列を加えた基本動作の候補から、3.2 で説明した手法によって 26 の基本動作を得ることができた。得られた基本動作の一部について特長的なフレームを図 7 に示す。

このようにして得られた基本動作を用いて、各ジェスチャにつき残る 5 パターンを分割表現した。分割表現の際には、文字列認識や連続音声認識の分野で多用されている、DP に基づく Recognition-based segmentation 法 [11] により基本動作を割り当てた。

図 8 (a) はジェスチャ“hurrah”を得られた基本動作で分割表現したものである。“hurrah”は図 4 に示すような両手の上下運動を 2 度繰り返すジェスチャであり、基本動作で分割した結果「0：両手を肩まで上げる」、「12：両手を上まで上げる」、「1：両手を肩から下げる」という三つの基本動作の繰返しで表現された。また、図 8 (b) はジェスチャ“left-hand bye”を基本動作で分割表現したものである。こちらは、「7：左手を

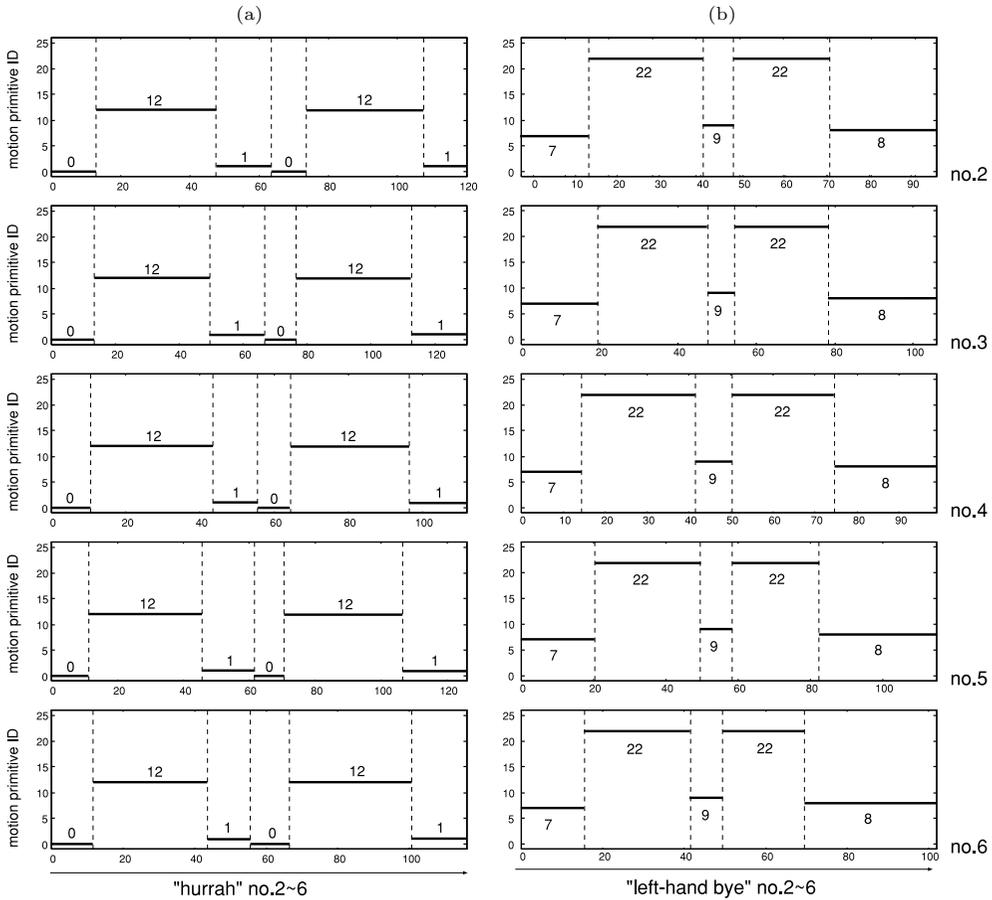


図 8 得られた基本動作でジェスチャを分割表現した結果 . (a) ジェスチャ“hurrah” についての結果 . (b) ジェスチャ“left-hand bye” についての結果 . グラフ中の数字は基本動作の番号を表す . また右端の番号はジェスチャパターンの番号を表す (no.1 は学習用データとして使用)

Fig. 8 Segmentation of gesture pattern by extracted motion primitives.

肩まで上げる」,「22 : 左手を外側に振る」,「9 : 左手を内側に振る」,「22 : 左手を外側に振る」,「8 : 左手を肩から下げる」という基本動作の列で分割表現された . このように , どちらのジェスチャについても 5 パターンすべてで同じ基本動作の列として表されるという結果が得られた .

その他の 16 ジェスチャについても , 11 ジェスチャについては 5 パターンすべてで同じ基本動作の列で表現できた . この結果から , 抽出された基本動作の多くは実際のジェスチャの共通部分及び非共通部分を適切に表現できているといえる . 残る 5 ジェスチャには複数の分割表現が含まれていた . 複数の分割表現が含まれていた原因としては , 重複動作の統合に失敗してい

ることが挙げられる . 具体的には , 「両手を肩の位置から下げる」という基本動作が複数存在していた . 重複動作の統合については , 後述するサポートの拡張を導入するなど検討の余地があるといえる .

4. む す び

本論文では , 論理判定型 DP を用いてパターン間に存在する複数の類似区間を同時に検出する手法について提案した . また , 本手法を基本動作抽出へ応用する方法についても提案した . 実験の結果 , 本手法の基本的な性能を示すことができた .

今後の課題としては , 以下のものが挙げられる .

- サポートの拡張

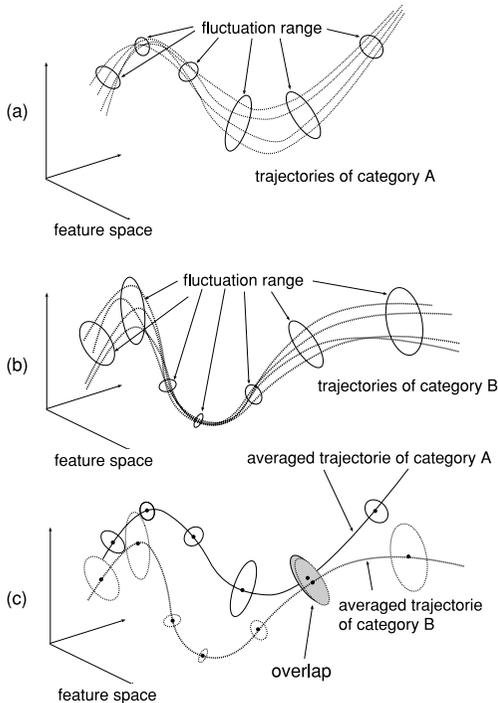


図 9 サポート定義の拡張。(a) 及び (b) は各パターンの変動と、それによって各フレームにおいて定義される変動幅を表す。(c) 各パターンの変動量を考慮したサポート定義の一例。カテゴリ A の変動幅内にカテゴリ B の平均軌道が含まれ、かつ B の変動幅内に A の平均軌道が含まれる場合に $s(i, j) = 1$ とする。

Fig. 9 Non-uniform support for coping with stochastic characteristics of gestures.

本手法では、式 (2) で定義した比較的単純なサポートを用いているが、これを拡張したサポートの定義を考えることができる。例えば、同一のカテゴリに属する複数のパターンから各フレームにおける特徴量の分布を調べ、この共分散から変動幅を定義する (図 9)。この変動幅を用いてサポートを定義することで、それぞれのパターンの各フレームにおける変動のしやすさを考慮に入れた類似区間検出が可能になると考えられる。

• ジェスチャ以外のパターンへの応用の検討

本論文では、本手法の具体的な応用例としてジェスチャの基本動作抽出について検討してきた。今後は、本手法を音声や動画画像などジェスチャ以外のパターンにも応用していきたいと考えている。

• ジェスチャネットワークの自動構築への応用

文献 [12] では、ジェスチャネットワークと呼ばれるグ

ラフ構造により複数のジェスチャの関係を表現し、それをジェスチャの早期認識及び動作の予測に利用している。このジェスチャネットワークの構築に本手法で得られた基本動作が利用できるものと考えている。

以上に加え、3. で抽出された基本動作について更に吟味が必要であると思われる。3. の方法で得られた基本動作は、3.1 で述べた様に従来の手法では得られない特長をもっている。本論文では、共通区間検出法の提案が最大の目的であるため、その一応用である基本動作についての比較実験は、今後行う予定である。

謝辞 本研究の一部は総務省戦略的情報通信研究開発推進制度 (SCOPE) の支援を受けた。

文 献

- [1] A.F. Bobick and A.D. Wilson, "A state-based approach to the representation and recognition of gesture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.19, no.12, pp.1325-1337, 1997.
- [2] T. Zhao, T. Wang, and H.-Y. Shum, "Learning a highly structured motion model for 3D human tracking," *Proc. 5th Asian Conf. Computer Vision*, vol.1 of 2, pp.144-149, 2002.
- [3] 向井理朗, 西村拓一, 遠藤 隆, 岡 隆一, "ジェスチャ動画画像の自己組織化ネットワークによるモデル化と要素動作の自動抽出," *信学技報*, PRMU97-128, Oct. 1997.
- [4] 山根 亮, 戸高千智, 川嶋幸治, 尺長 健, "動作データの時系列相関行列による舞踊動作解析," *信学論 (D-II)*, vol.J88-D-II, no.8, pp.1652-1661, Aug. 2005.
- [5] 伊東慶明, 木山次郎, 小島 浩, 関 進, 岡 隆一, "時系列標準パターンの任意区間によるスポッティングのための Reference Interval-free 連続 DP," *信学論 (D-II)*, vol.J79-D-II, no.9, pp.1474-1483, Sept. 1996.
- [6] 西村拓一, 古川 清, 向井理朗, 岡 隆一, "時系列標準パターン検索のための重み減衰型 Reference Interval-Free 連続 DP について," *信学論 (D-II)*, vol.J81-D-II, no.3, pp.472-482, March 1998.
- [7] 大崎竜太, 嶋田光臣, 上原邦昭, "速度に基づく切り出しとクラスタリングによる基本動作の抽出," *人工知能誌*, vol.15, no.5, pp.878-886, 2000.
- [8] 澤田秀之, 橋本周司, 松嶋俊明, "運動特徴と形状特徴に基づいたジェスチャー認識と手話認識への応用," *情処学論*, vol.39, no.5, pp.1325-1333, 1998.
- [9] T.D. Sanger, "Optimal movement primitives," *Advances in Neural Information Processing Systems*, vol.7, pp.1023-1030, 1995.
- [10] A. Fod, M.J. Mataric, and O.C. Jenkins, "Automated deviation of primitives for movement classification," *Autonomous Robots*, vol.12, no.1, pp.39-54, 2002.
- [11] 迫江博昭, 藤井宏美, 吉田和永, 巨理誠夫, "フレーム同期化, ビームサーチ, ベクトル量子化の統合による DP マッチングの高速化," *信学論 (D-II)*, vol.J81-D-II, no.6, pp.1251-1258, June 1988.
- [12] 森 明慧, 内田誠一, 倉爪 亮, 谷口倫一郎, 長谷川勉,

迫江博昭, “実体を伴うプロアクティブインタフェースのためのジェスチャの早期認識・予測に関する検討,” 日本ロボット学会誌, vol.24, no.8, pp.954-963, Nov. 2006.
(平成 18 年 10 月 10 日受付, 19 年 2 月 6 日再受付)



森 明慧 (学生員)

2005 九大・工・電気情報卒。2007 同大学院修士課程(知能システム学)了。現在(株)東芝勤務。在学中,ジェスチャ認識,画像処理に関する研究に従事。2006 本会パターン認識メディア理解研究会アルゴリズムコンテスト最優秀賞。日本ロボット

学会, ACM 各会員。



内田 誠一 (正員)

1990 九大・工・電子卒。1992 同大学院修士課程(情報)了。セコム(株)勤務を経て,現在,同大システム情報科学研究院知能システム学部門准教授。博士(工学)。画像パターン・時系列パターンの解析・認識に関する研究に従事。2003 本会 PRMU

研究奨励賞, 2006 MIRU 長尾賞各受賞。IEEE, 情報処理学会, 映像情報メディア学会各会員。



倉爪 亮

1991 東京工業大学大学院機械物理工学専攻修士課程了。同年(株)富士通研究所入社, 1995 同学機械宇宙学科助手, 2000

スタンフォード大客員研究員, 同年東京大学生産技術研究所博士研究員, 2002 九州大学システム情報科学研究院助教授, 2007 より同研究院教授。現在に至る。群ロボット, 歩行機械, レーザ計測, 医用ロボットの研究に従事。博士(工学)。日本ロボット学会会員。



谷口倫一郎 (正員)

1978 九大・工・情報工卒。1980 同大学院修士課程了。同年同大助手。1996 同大学院システム情報科学研究科(現研究院)教授。工博。画像処理, コンピュータビジョン, 並列処理等の研究に従事。1987 年度

本会篠原記念学術奨励賞, 1994 年度情報処理学会坂井記念特別賞, 2000 年度映像情報メディア学会丹羽高柳論文賞等を受賞。



長谷川 勉 (正員)

1973 東工大・工・電子物理卒。同年電子技術総合研究所勤務。1992 より九州大学工学部情報工学科教授。現在, 同大学院システム情報科学研究院教授。知能ロボットの研究に従事。工博。日本ロボット学会, 計測自動制御学会, 電気学会, 日本機械学

会等各会員。



迫江 博昭 (正員:フェロー)

1966 九大・制御卒。1968 九大大学院修士課程(電子)了。同年日本電気(株)入社。1989 より九大システム情報科学研究院知能システム学部門教授, 現在同大名誉教授。工博。音声認識及び画像パターン

解析の研究に従事。IEEE ASSP Senior Award, 本会業績賞, 米沢ファウンダーズメダル受賞記念特別賞及び論文賞(社)日本発明協会・内閣総理大臣発明賞, 九工大嘉村記念賞を各受賞。IEEE, 日本音響学会, 情報処理学会, 映像情報メディア学会各会員。