

## 離散的な曲率と捩率を利用した身体形状の分類とジェスチャ認識

武佐 和宏<sup>†</sup> 渡辺 弥寿夫<sup>‡</sup>

<sup>†</sup> ‡ 金沢工業大学〒921-8501 石川県石川郡野々市町扇が丘 7-1

E-mail: <sup>†</sup> musa@venus.kanazawa-it.ac.jp, <sup>‡</sup> watanabe@infor.kanazawa-it.ac.jp

あらまし 身体形状が曲線によって構成されるとし、線分列のために定めた離散的な曲率と捩率を用いてその形状を分類する方法を提案する。曲率と捩率による表現はオイラー角による表現よりもパラメータが少なく、元曲線と一意に対応している点で優れている。ジェスチャ認識のための高速な分類器として決定木を用い、腕形状 100 クラスの分類実験を行うことで、決定木がジェスチャ認識においても充分高速に判別可能なことを確認した。分類結果として得られた身体形状パターン配列に対し、最適なローカルアラインメントを求めることにより、既知の運動パターンが現れる部分配列が求められることを確認した。

キーワード 曲率, 捩率, 決定木, ローカルアラインメント, ジェスチャ認識

## Shape Classification of Human Bodies and Gesture Recognition Using Discrete Curvature and Torsion

Kazuhiro MUSA<sup>†</sup> and Yasuo WATANABE<sup>‡</sup>

<sup>†</sup> ‡ Kanazawa Institute of Technology

7-1 Ohgigaoka, Nonoichimachi, Ishikawa-gun, Ishikawa, 921-8501 Japan

E-mail: <sup>†</sup> musa@venus.kanazawa-it.ac.jp, <sup>‡</sup> watanabe@infor.kanazawa-it.ac.jp

**Abstract** First of all, we define the discrete curvature and torsion for a series of line segments of a human body, which is taken as a space curve. These descriptions have a good property, because they have less parameters than Euler angles and are possible to reconstruct the original curve shape uniquely. And then, we propose a method for shape classification of human bodies and gesture recognition using discrete curvature and torsion. Some experiments shows that a decision tree works fast enough to recognize human gestures for one hundred classes of arm shapes. It is proved that an optimal local alignment for the classified pattern arrays is found as a part of arrays corresponding to the given motion patterns.

**Keyword** Curvature, Torsion, Decision Tree, Local Alignment, Gesture Recognition

### 1. はじめに

人の動作を認識することは様々な分野において現実的な要請があり、近年ではテレビゲームやセキュリティ分野、スポーツにおける動作の分析・指導などにも利用されている。

ジェスチャ認識の研究では、従来、画像から比較的容易に得られる人の重心座標や手などの速度情報を利用することが多かったが、本研究では、身体形状を空間曲線とみなして、これを規定する曲率と捩率を利用した身体形状の分類方法を提案する。ジェスチャ認識のための高速な分類器として決定木を用いることで身体形状の多クラス分類を行い、最終的には、決定木による分類結果として得られる時系列身体形状パターン配列に対し、最適なローカルアラインメントを求めることにより、指定配列区間中に現れる、既知の部分動作パターンを検出することを目的とする。

### 2. 動作データの取得

Anima Zoo社<sup>[1]</sup>の Gypsy Gyro を用いて動作をキャプチャした。モーションキャプチャにより得られる動作データは BVH(BioVision Hierarchy)形式である。BVH は HIERARCHY 部と MOTION 部からなる。HIERARCHY 部では腰部を根ノードとした木構造によって骨格形状が定義されている。MOTION 部には各フレーム・各関節におけるオイラー角変位が記述されている。BVH によって記述される各ノードの身体点の空間座標を復元して利用する。

### 3. 身体形状表現

BVH では腰部を根ノードとして手足先に向かう線分列がオイラー角によって記述されていたが、空間曲線の場合は、曲率と捩率のみから元曲線を一意に復元できることから、特徴量としては曲率と捩率の 2 つのパラメータがあれば充

分である。本研究では離散的な点列からなる線分列を扱うため、空間曲線に対する曲率と振率を離散的に定義し、人の腕や脚を線分列とみなしてそれを離散的曲率・振率から類別することを試みる。

### 3.1. 空間曲線に対する曲率・振率の定義

弧長  $s$  をパラメータとする滑らかな空間曲線  $c(s)$  の曲率  $\kappa$  および振率  $\tau$  は以下の式(3-1)および(3-2)より定義される[2]。

$$\kappa(s) = \left| e_1'(s) \right| \quad (3-1)$$

$$\tau(s) = e_2'(s) \cdot e_3(s) \quad (3-2)$$

ここではパラメータ  $s$  による微分で、接ベクトル  $e_1$ 、主法線ベクトル  $e_2$ 、従法線ベクトル  $e_3$  は以下の式(3-3)による。

$$e_1(s) = c'(s) \quad e_2(s) = \frac{e_1'}{|e_1'|} \quad e_3(s) = e_1(s) \times e_2(s) \quad (3-3)$$

×は外積を表す。

### 3.2. 線分列に対する離散的曲率・振率の定義

空間中の離散的な点列からなる曲線に対しては微分ができないため、直接に曲率・振率を計算することが出来ない。そのため、ここでは  $n$  点からなる曲線の点列を  $[P_0, \dots, P_{n-1}]$  とし、以下のようにして離散的に曲率と振率を定める。ここでは定義(3-1)のように定めた  $\alpha_i, B_i, N_i$  を、それぞれ  $i$  番目の接ベクトル、従法線ベクトル、主法線ベクトルとして扱う。

接ベクトル	
$\alpha_i = \frac{P_{i+1} - P_i}{ P_{i+1} - P_i } \quad (i = 0 \dots n-2)$	
従法線ベクトル	主法線ベクトル
$B_i = \alpha_i \times \alpha_{i+1}$	$N_i = B_i \times \alpha_{i+1}$
$(i = 0 \dots n-3)$	
定義 (3-1)	

これらから離散的曲率  $C_i$ ・振率  $T_i$  を定義(3-2)のようにして定める。

離散的曲率	
$C_i = \cos^{-1} \left( \frac{\alpha_i \cdot \alpha_{i+1}}{ \alpha_i   \alpha_{i+1} } \right) \quad (i = 0 \dots n-3)$	
離散的振率	
$T_i = \cos^{-1} \left( \frac{B_i \cdot B_{i+1}}{ B_i   B_{i+1} } \right)$	
if $(0 < \frac{N_i \cdot B_{i+1}}{ N_i   B_{i+1} }) \quad T_i = -T_i$	
$(i = 0 \dots n-4)$	
定義 (3-2)	

復元する際は  $\alpha_i$  の長さが必要になるため、以下のようにして予め線分長を保存しておく。

$\alpha_i$ の長さ $l_i$	
$l_i =  \alpha_i  \quad (i = 0 \dots n-2)$	
定義 (3-3)	

### 3.3. 離散的な曲率・振率値からの復元

得られた離散的な曲率・振率値を用いて、初期動標構と復元開始点から線分列を復元するためにはアルゴリズム(3-1)のような手順を繰り返す。

ベクトル $a$ を、単位ベクトル $b$ を軸として右ネジ方向を正方向にとり、 $\theta$ だけ回転したベクトルを返す関数を $\text{rot}(a, b, \theta)$ と定義する。	
for $(i = 0 \dots n-4)$ {	
$B_{i+1} = \text{rot}(B_i, \alpha_{i+1}, T_i)$	
$\alpha_{i+2} = \text{rot}(\alpha_{i+1}, B_{i+1}, C_i)$	
$P_{i+3} = P_{i+2} + \alpha_{i+2} \cdot l_{i+2}$	
}	
アルゴリズム (3-1)	

$C_i, T_i$  から線分列を復元するには初期条件と

しての平面が必要であり、動標構  $\alpha_1, B_0, (N_0)$  とその位置を規定する点の座標が必要であるため  $P_0, P_1, P_2$  が必要であり(ベクトル  $P_1 \cdot P_0$  および  $P_2 \cdot P_1$  は互いに独立なベクトルであるものとする), さらに元の線分列と空間中で完全に重なるように復元するためには  $l_i$  が必要となる。

#### 4. 決定木による身体形状の分類

離散的な曲率・振率値から元曲線を復元できることから、曲線形状を分類するためにこれらを用いることを考える。人の動作を認識するためにはなるべく短時間で各フレームにおける姿勢形状を分類できなければならないため、実行時に特別な計算を必要とせず、最悪の場合でも木の高さに応じた比較を行うだけで高速に類別可能な分類器である決定木を用いる。身体形状の分類結果は記号として配列に保存しておき、動的計画法によって既知の動作との最適ローカルアラインメントを求めするために用いる。

離散曲率・振率値を保存するためにデータベース(PostgreSQL8.0)を用いて以下のようなテーブルを各腕・脚について作成した。

表 4-1 テーブルの定義

列名	定義	
class	integer	NOT NULL
c1	real	NOT NULL
c2	real	NOT NULL
c3	real	NOT NULL
c4	real	NOT NULL
t1	real	NOT NULL
t2	real	NOT NULL
t3	real	NOT NULL

ここで、列名 class は以下の列 c1~t3 の所属する形状クラス番号を表す。c1, c2, c3, c4 は離散曲率値を格納し、t1, t2, t3 には離散振率信号値を格納するものとする。

#### 4.1. 木の評価

決定木の評価は識別率と木のサイズを調べることで行い、木の分割基準として用いた情報利得の閾値を変化させたときの両者の変化をみる。右腕形状 100 クラスに対して、各々 100 のデータをサンプルとして採り、各クラスが 50 サンプルもつようにデータを半分にして 2 つのテーブルそれぞれに格納した。

1 つのテーブルのデータを用いて木を作成し、残りのテーブルを用いて正しく分類できるかをテストした結果を以下表 4-2 に示す。さらに、閾値 0.01 としたとき得られたあいまいさの無い木に対して、事後枝刈りを行ったが、識別率は

66%程度であった。(実行環境: Windows Home Edition SP2/Pentium4 CPU3.2GHz/1GB RAM 以下同様)

表 4-2 閾値と作成された決定木の識別率

thres	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.01
correct	2858	3119	3153	3198	3214	3240	3265	3265
incorrect	2142	1881	1847	1802	1786	1760	1795	1795
correct rate	0.572	0.624	0.631	0.641	0.643	0.648	0.641	0.641

作成された木による分類を 100 回行い、その平均実行時間と分類に用いた葉ノードの高さの平均を表 4-3 に示す。

表 4-3 木による分類に要する時間

average proc time(us)	average leaf height
5.394	7.756

### 5. 最適ローカルアラインメントによる時系列動作の分節化と認識

#### 5.1. ローカルアラインメント

大域アラインメントを求める場合、通常は動作の開始点と終了点が既知であることが前提であるが、動作の開始点と終了点を同定するという作業は、それ自体が動作に対する知識を必要とするため、任意の動作に対してその区間を抜き出すことは容易ではない。

大域的に一致するようなパターンを見つけるのではなく、2 つの配列を与えたときに、部分的に最も一致するような区間が得られれば速度による動作の分節化を必要とせずに、マッチングが可能である。式(5-1)により得られる最適ローカルアラインメント[4]を求めることで、既知の動作パターン配列が部分的に最も一致するような動作区間を得ることができる。

$$\begin{aligned}
 D[i,0] &\leftarrow 0 \quad i = 0, \dots, m \\
 D[0,j] &\leftarrow 0 \quad j = 0, \dots, n \\
 D[i,j] &\leftarrow \max\{ \\
 &0, \\
 &D[i-1,j] + d, \\
 &D[i,j-1] + d, \\
 &D[i-1,j-1] + w(s_1[i], s_2[j]) \\
 &\} \\
 \end{aligned} \tag{5-1}$$

式(5-1)において  $d$  はギャップペナルティ、 $D[i]$  はアラインメントに対する最大スコアを格納す

るための2次元配列, $w(c1,c2)$ は文字  $c1$  と  $c2$  の対応に関するスコアを与える関数である。

### 5.2. スコア関数の定義

データベースには、離散的曲率・振率の情報が格納されているが、腕および脚の形状を記号で表すために、その記号間のスコアを離散的曲率・振率空間内の距離で置き換えることを考える。各クラスの平均ベクトルを計算し、全クラスの平均ベクトル間距離を計算して記号間スコアとするが、距離の近いクラス間に同程度のスコアを割り当てることで、決定木による誤分類の影響をある程度緩和することを期待する。ここで、離散的曲率の値域は $[0 - \pi]$ であるが、離散的振率の値域は $[-\pi - \pi]$ となっており環状になっている。そのため、離散的振率の値に関しては、距離が  $\pi$  を超える場合にそれを  $2\pi$  から差引くことで調整している。平均ベクトル間距離はSQLにより容易に得られる。

表 5-1 SQL 発行結果例

class integer	class integer	distance double precision
0	0	0.000
0	1	0.426
0	2	2.459
0	3	5.066
0	4	5.149
...	...	...
99	94	2.487
99	95	2.865
99	96	1.692
99	97	1.671
99	98	1.427
99	99	0.000

表 5-1 のように得られたクラス間距離は DP マッチング用スコアテーブルに保存して用いるが、最大スコアを与えるアラインメントを求めるためのスコアであり、距離が近いものほど高いスコアを与えるようにしたい。そのため、使用時には各スコアを  $5-distance$  と変換してもちいることとする。最適ローカルアラインメントを求めるための再帰式を、最小距離を求めるように修正するよりも、与えるスコア関数を修正するほうが容易である。

### 5.3. 認識対象動作の収集

読み込まれた BVH ファイルから復元された線分列より、離散曲率・振率値を計算することができる。認識したい動作を選び、選択区間の離散曲率・振率値をファイルとして出力するものと

する。似たような動作をしている複数の区間から選ぶと、複数のファイルが得られるが、DP マッチングによって2つの配列の最適ローカルアラインメントを求めることで、動作の認識を行うため、認識対象動作として使用する配列として、複数ある候補パターン配列の中から代表パターンとしてふさわしい配列をひとつ選択しなければならない。

代表パターンを選択するときの一般的な方法は信号の平均値を代表とすることである。区間幅がそろっていれば区間の始点から終点まで、各々対応するフレームの離散的曲率および振率の値の平均を求めることができるが、今回は選択される区間の幅が各々異なっており、単純に平均値を求めて代表とすることはできない。

そこで今回は、式(5-2)のようにして得られた  $c1$  を、決定木による類別により得られた同一動作パターン集合  $P$  から得られた代表パターンとする。DP(A,B)は動作パターン配列 AB間の最適グローバルアラインメントにおける最大スコアを求める関数である。

$$\operatorname{argmax}_{c_1 \in P} \left( \sum_{c_2 \in P, c_1 \neq c_2} DP(c_1, c_2) \right) \quad (5-2)$$

### 5.4. 実験

ラジオ体操第1の運動中に現れる、いくつかの動作パターンを収集し式(5-2)に基づいて代表となるパターンを選択した。逐次最適ローカルアラインメントを計算し、そのスコアの変化を調べるが、既知の動作パターンの検索対象範囲は過去 49 フレームから現在のフレームまでとし、決定木により得られた、配列長 50 の形状パターン配列との最適ローカルアラインメントを計算する。簡単のために右腕の動きについてのみ実験した。

そのとき得られたスコアと選んだ運動パターンを図 5-1~5-4 に示す。これらの運動パターンには、決定木によって分類された結果のクラス番号が配列の形で与えられている。使用する BVH ファイルには1秒当たり120フレームの情報が記述されているが、使用時には計算量の削減のために15FPSとして使用している(0~23592までのフレームがあるが8フレームずつ読み込み2950フレームのみ用いる)。

図中“・”で示す部分が実際にその運動が現れる部分である。該当パターンが出現した直後には高いスコアが得られるため、最適ローカルア

ラインメントを求めることで連続した運動の中から、指定動作が現れる区間を抽出することができる。

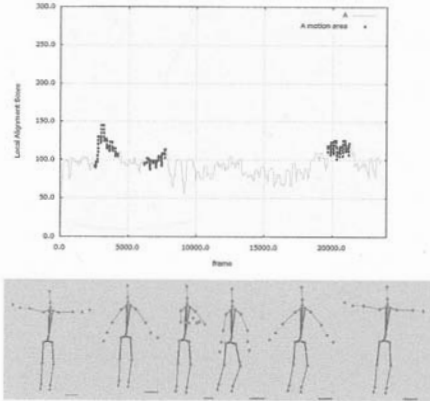


図 5-1 動作 A(配列長 29)とスコア

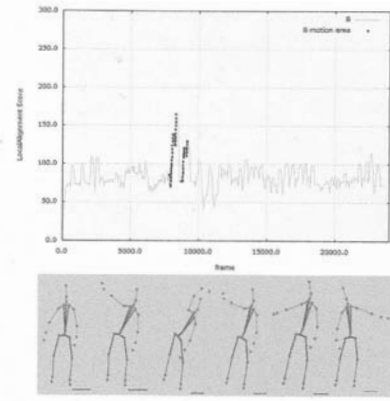


図 5-2 動作 B(配列長 33)とスコア

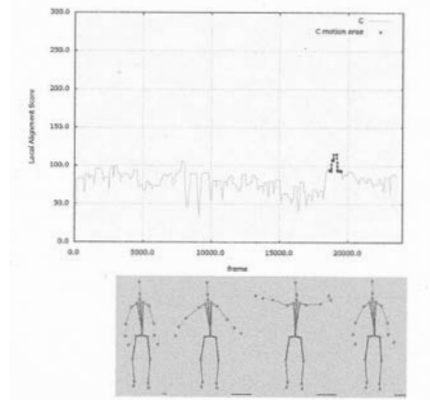


図 5-3 動作 C(配列長 23)とスコア

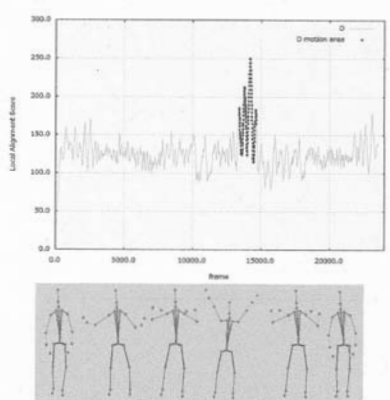


図 5-4 動作 D(配列長 54)とスコア

波形のピークと実際の運動が現れる区間にずれがあるのは、最適ローカルアラインメントを求めるための配列がたまるまで、しばらく待つ必要があるからであり、実際に運動が開始されていても、その運動が終了するまでは完全な運動パターンを配列として得ることができない。そのため、最適ローカルアラインメントのスコアのピークが現れるのは完全に運動が終了してからである。各運動パターンによって、配列長の違いがあるため、最適ローカルアラインメントを求めて得られるスコアは一様ではない。そのため、認識のための閾値を設定する際は運動パターン毎に事前に充分に調べておく必要がある。運動パターン A~D のスコアを見る限り、パターンの配列長が長いものほどピークが大きく現れ

るため、認識したかどうかの見分けが付きやすい。それだけゆとりを持たせて閾値を設定できるため誤識別も少なくできると期待できる。各運動の最適ローカルアラインメントのスコアのピークが現れるフレーム区間は重複しておらず、それぞれが形状の変化そのものに対して反応していることが窺える。

今回はギャップペナルティ  $d$  を  $-1$  として最適ローカルアラインメントを求めたが、スコアテーブル中の他のスコアに正であるものが多数含まれることに対し、 $d$  を  $-1$  としたことはパターンの伸長をほとんど許さないことになるとも考えられる。これでは、動作の緩急に柔軟に対応できない可能性があるため、 $d$  に対して適切な値をどのように設定すればよいか課題となる。

2950 フレームにわたって、最適ローカルアライメントを求めるために要した平均時間を動作パターンごとに表 5-2 に示す。

表 5-2 最適ローカルアライメントの平均処理時間

	A	B	C	D
average proc time(ms)	0.413	0.493	0.356	0.845

## 6. ソフトウェア

### 6.1. 決定木作成ソフトウェア myD-tree

決定木作成アルゴリズムを実装したソフトウェアについて説明しておく。データベースのテーブルに予め格納しておいた、腕および脚形状に関する、ラベルつき離散的曲率・振率値を用いて、領域内のあいまいさを測る最大利得基準に従って逐次離散的曲率・振率空間を2分割し、木を作成する。作成した木による空間分割結果を表示することができ、ここで作成した決定木を、実際に形状分類のために使用する。

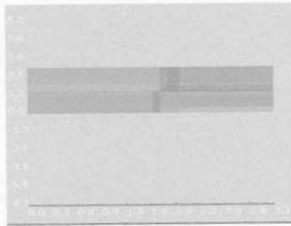


図 6-1 特徴空間の決定木による分割

### 6.2. 動作解析ソフトウェア Motion

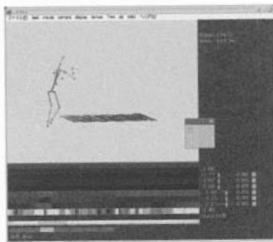


図 6-2 ソフトウェア

BVH ファイルを読み込み、復元した空間中の身体点座標から定めた、腕と脚に相当する線分列の時系列離散的曲率・振率値を表示するソフトウェアを開発した。決定木作成プログラムで作成した両腕・両脚用の形状分類のための決定木を読み込み、時系列で得られる離散的曲率・振率値を逐次分類する。それらの時系列分類結果パターン配列に対し、既知の運動パターン配列を問

い合わせ用の配列とする。ユーザが指定した検索対象区間から、既知の運動パターン配列が部分的に最も一致する区間を見つけるために、最適ローカルアライメントを計算し、そのスコアと部分区間を表示することができる。さらに、離散的曲率・振率値の加工と補間ができ、それらより復元された身体形状を表示することも可能である。

## 7. おわりに

身体形状に基づくジェスチャ認識として、人間の身体形状が線分列で構成されるとみなし、その形状を、離散的に定義した曲率と振率に基づいて分類する方法を提案した。離散的曲率・振率による表現はオイラー角表現に対して必要なパラメータが少なく、また形状に対して一意に対応しており、復元も容易であることが利点である。今回得られた決定木の認識率はおよそ66%程度しか得られておらず、まだ改善の余地があるが、決定木による分類はリアルタイムな処理に耐えるほど充分高速に行うことができた。また、実際に決定木により分類された動作パターン配列に対して最適ローカルアライメントを求めることで、指定した動作パターンが出現する区間を発見できることを確認した。しかし、最適ローカルアライメントを求めることはある程度のコストを伴う処理であり、これを用いて多数の運動パターンを単位フレーム時間内に処理することは現実的ではなく、パターン数を増やすことによる単純な拡張はできない。離散的な曲率・振率という身体形状に関する知識に基づいてそれを高速に分類するための決定木を作成したように、運動間の同時確立密度分布など、運動パターンそのものの知識に基づいて可能性の高い運動パターンを選択するようなアルゴリズムが必要となる。

## 文献

- [1] AnimaZoo, "http://www.animazoo.com/", 。
- [2] 小林昭七, "曲線と曲面の微分幾何学," 裳華房, 1995.
- [3] J.R.キンラン, 古川康一(監訳), "AI によるデータ解析," トッパン株式会社, 1995.
- [4] T.F.Smith, M.S.Waterman, "Identification of Common Molecular Subsequences," J.Molecular Biology, pp195-197, 147, 1981.
- [5] 武佐和宏, 渡辺弥寿夫, "曲率と振率を利用した3次元軌跡曲線の解析手法," 第6回情報科学技術フォーラム(FIT2007)講演論文集, H-30, pp69-70, 2007.