

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7146247号

(P7146247)

(45)発行日 令和4年10月4日(2022.10.4)

(24)登録日 令和4年9月26日(2022.9.26)

(51)国際特許分類

F I

G 0 6 T 7/20 (2017.01)

G 0 6 T 7/20 3 0 0 A

G 0 6 T 7/143(2017.01)

G 0 6 T 7/143

G 0 6 N 7/00 (2006.01)

G 0 6 N 7/00 1 5 0

請求項の数 15 (全17頁)

(21)出願番号 特願2018-164554(P2018-164554)
 (22)出願日 平成30年9月3日(2018.9.3)
 (65)公開番号 特開2020-38440(P2020-38440A)
 (43)公開日 令和2年3月12日(2020.3.12)
 審査請求日 令和3年8月17日(2021.8.17)

(73)特許権者 504137912
 国立大学法人 東京大学
 東京都文京区本郷七丁目3番1号
 (74)代理人 100103137
 弁理士 稲葉 滋
 (74)代理人 100145838
 弁理士 畑添 隆人
 (72)発明者 中村 仁彦
 東京都文京区本郷七丁目3番1号 国立
 大学法人東京大学内
 (72)発明者 高野 渉
 東京都文京区本郷七丁目3番1号 国立
 大学法人東京大学内
 (72)発明者 池上 洋介
 東京都文京区本郷七丁目3番1号 国立
 最終頁に続く

(54)【発明の名称】 動作認識方法及び装置

(57)【特許請求の範囲】

【請求項1】

コンピュータを用いて対象の動作を認識する動作認識方法であって、

対象の動作を規定する時系列データを複数のセグメントに分節し、
 各セグメントを、当該セグメントの時系列データを用いて基本動作のクラスに分類し、
 前記対象の動作を、連続する基本動作からなる上位記号の系列に変換するステップであ
 り、上位記号と基本動作列の組からスコアを算出する関数が用意されており、動的計画法
 を用いた最適化計算により、前記対象の動作を上位記号の系列に変換する、
 動作認識方法。

【請求項2】

前記動的計画法を用いた最適化計算は、
 以下の関数漸化式

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N) = \max_{l < m < N} [g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m) + f_{k-1}(\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_N)]$$

を用いるものであり、
 ここで、

Nは、対象の動作を分節して得られたセグメントの数、

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N)$$

10

20

は、基本動作の時系列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_N$$

から k 個の上位記号を抽出した時に得られる最大スコアであり、

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

は、

基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_m$$

10

を 1 つの上位記号として識別したときに得られるスコアである、

請求項 1 に記載の動作の認識方法。

【請求項 3】

上位記号に対応する 1 つあるいは複数の基本動作列及びスコアを備えたテーブルが用意されている、

請求項 1、2 いずれか 1 項に記載の動作の認識方法。

【請求項 4】

前記セグメントは、隠れマルコフモデル (HMM) を用いて、基本動作のクラスに分類される、

請求項 1 ~ 3 いずれか 1 項に記載の動作の認識方法。

【請求項 5】

前記上位記号は、運動競技における技である、請求項 1 ~ 4 いずれか 1 項に記載の動作の認識方法。

【請求項 6】

前記スコアには、技の難度が反映されている、請求項 5 に記載の動作の認識方法。

【請求項 7】

得られた上位記号をセグメントとして、前記対象の動作をさらに上位の動作を表す上位記号列に変換することを含む、請求項 1 ~ 4 いずれか 1 項に記載の動作の認識方法。

30

【請求項 8】

動作の認識装置は、記憶部と処理部とを備え、

前記記憶部には、複数の基本動作のクラスが格納されており、

前記処理部は、

対象の動作を規定する時系列データを複数のセグメントに分節し、

各セグメントを、当該セグメントの時系列データを用いて、基本動作のクラスに分類し、

前記対象の動作を、連続する基本動作からなる上位記号の系列に変換する、

ように構成されており、

前記記憶部には、さらに、上位記号と基本動作列の組からスコアを算出する関数が規定

40

されており、前記処理部は、前記スコアを用いて、前記対象の動作を、動的計画法を用いた最適化計算により、上位記号列に変換する、

ように構成されている、

動作の認識装置。

【請求項 9】

前記処理部は、

以下の関数漸化式

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N) = \max_{l < m < N} [g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m) + f_{k-1}(\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_N)]$$

50

を動的計画法によって最適化計算するものであり、
 ここで、
 N は、対象の動作を分節して得られたセグメントの数、
 $f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N)$

は、基本動作の時系列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_N$$

10

から k 個の上位記号を抽出した時に得られる最大スコアであり、

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

は、

基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_m$$

を 1 つの上位記号として識別したときに得られるスコアである、

20

請求項 8 に記載の動作の認識装置。

【請求項 10】

前記記憶部には、上位記号に対応する 1 つあるいは複数の基本動作列及びスコアを備えたテーブルが用意されている、

請求項 8、9 いずれか 1 項に記載の動作の認識装置。

【請求項 11】

前記処理部は、隠れマルコフモデル (HMM) を用いて、セグメントを基本動作のクラスに分類する、

請求項 8 ~ 10 いずれか 1 項に記載の動作の認識装置。

【請求項 12】

30

前記上位記号は、運動競技における技である、請求項 8 ~ 11 いずれか 1 項に記載の動作の認識装置。

【請求項 13】

前記スコアには、技の難度が反映されている、請求項 12 に記載の動作の認識装置。

【請求項 14】

前記処理部は、得られた上位記号をセグメントとして、前記対象の動作をさらに上位の動作を表す上位記号列に変換するように構成されている、請求項 8 ~ 11 いずれか 1 項に記載の動作の認識装置。

【請求項 15】

コンピュータを、請求項 8 ~ 14 に記載の記憶部、処理部として機能させるコンピュータプログラム。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、動作認識方法及び装置に関するものである。

【背景技術】

【0002】

人間の動作を分節化して認識しようとする研究は存在する (非特許文献 1、2)。運動学・動力学計算に代表されるロボティクスのアルゴリズムは統計的手法を利用した人間の全身運動の分節化・識別・予測技術へと展開されている。これら計算技術の応用先は、スポ

50

ーットレーニング、リハビリテーションなど多岐に渡る。

【 0 0 0 3 】

スポーツにおける動作認識について考えると、人間の動作のセンシング技術の向上、人工知能のための機械学習技術の発展、スポーツにおける動作や技能の複雑化、高度化及びそれに伴う人による判定の限界等の背景から、審判員によって採点が必要な競技における機械による自動採点技術の利用が課題となる。身体運動計測と高度情報処理技術は、このような課題に対して一つの解決法を与えられ（非特許文献3）。

【 0 0 0 4 】

特許文献1には、対象の演技を評価する運動認識装置が開示されており、当運動認識装置は、分節部と、識別部と、評価部とを備えており、対象の演技データを分節化して演技の技および難度を評価するものである。

10

【先行技術文献】

【特許文献】

【 0 0 0 5 】

【文献】WO2018/070414

【非特許文献】

【 0 0 0 6 】

【文献】Wataru Takano, Yoshihiko Nakamura, "Real-time Unsupervised Segmentation of human whole-body motion and its application to humanoid robot acquisition of motion symbols," Robotics and Autonomous Systems, Vol.75, Part B, pp.262-272, 2016

20

Wataru Takano, Hirotaka Imagawa, Yoshihiko Nakamura, "Spatio-temporal structure of human motion primitives and its application to motion prediction," Robotics and Autonomous Systems, Vol.75, PartB, pp.2288-2296, 2016

J. Shin and S. Ozawa. A study on motion analysis of an artistic gymnastics by using dynamic image processing - for a development of automatic scoring system of horizontal bar -. In Proceedings of IEEE International Conference on Systems, Man and Cybernetics, pages 1037-1042, 2008.

【発明の開示】

【発明が解決しようとする課題】

30

【 0 0 0 7 】

本発明は、対象の動作を分節化して得られた基本動作の連続を考慮して、動作の認識を最適化問題として扱うことによって、動作の発見と評価を自動で行うことを目的とするものである。

【課題を解決するための手段】

【 0 0 0 8 】

本発明は、人間の動作の識別とその説明を行うための技術に関するものであり、人間の動作を分節化する計算、

切り取られたセグメントごとに基本動作と比較して識別する計算、及び、

基本動作の一連からなる動作を発見して、その動作が前後関係から適切な説明かどうかを評価して決定する計算、からなり、

40

基本動作の系列から動作、体操などの運動競技では「技」、を発見する問題を、動的計画法を適用して最適化問題の解として計算する。

本発明は、一般の人間の動作の解釈を自動的に行うシステムや、体操、フィギュアスケートなどの運動の技を認識して採点する自動採点システムにも応用できる。

【 0 0 0 9 】

本発明が採用した動作認識方法は、

対象の動作を規定する時系列データを複数のセグメントに分節し、

各セグメントを、当該セグメントの時系列データを用いて基本動作のクラスに分類し、

前記対象の動作を、連続する基本動作からなる上位記号の系列に変換するステップであ

50

り、上位記号と基本動作列の組からスコアを算出する関数が用意されており、動的計画法を用いた最適化計算により、前記対象の動作を上位記号の系列に変換する、
ことからなる。

【 0 0 1 0 】

本発明が採用した動作認識装置は、動作の認識装置は、記憶部と処理部とを備え、
前記記憶部には、複数の基本動作のクラスが格納されており、
前記処理部は、
対象の動作を規定する時系列データを複数のセグメントに分節し、
各セグメントを、当該セグメントの時系列データを用いて、基本動作のクラスに分類し、
前記対象の動作を、連続する基本動作からなる上位記号の系列に変換する、
ように構成されており、
前記記憶部には、さらに、上位記号と基本動作列の組からスコアを算出する関数が規定
されており、
前記処理部は、前記スコアを用いて、前記対象の動作を、動的計画法を用いた最適化計
算により、上位記号列に変換する、
ように構成されている。

10

【 0 0 1 1 】

本発明は、また、コンピュータを、上記装置の記憶部、処理部として機能させるコンピ
ュータプログラム、あるいは、当該コンピュータプログラムを記憶したコンピュータ可読
媒体として提供される。

20

【 0 0 1 2 】

前記動的計画法を用いた最適化計算は、
以下の関数漸化式

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N) = \max_{l < m < N} [g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m) + f_{k-1}(\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_N)]$$

を用いるものであり、
ここで、

N は、対象の動作を分節して得られたセグメントの数、

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N)$$

30

は、時系列としての基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_N$$

から k 個の上位記号を抽出した時に得られる最大スコアであり、

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

40

は、

基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_m$$

を 1 つの上位記号として識別したときに得られるスコアである。

なお、上記漸化式と形式上異なる式であっても、当業者において等価であると考えられ
る式は、本発明の範囲に含まれる。

【 0 0 1 3 】

1 つの態様では、上位記号に対応する 1 つあるいは複数の基本動作列及びスコアを備え

50

たテーブルが用意されている。

1つの態様では、上位記号と基本動作列の組合せのスコアを確率値（尤度）として計算する関数を学習しておく。この関数として、上位記号と基本動作列を入力として、尤度を出力する統計モデルやニューラルネット等を用いることができる。動作認識装置の記憶部には、上記関数のパラメータが記憶されており、前記処理部は、前記パラメータを利用して上位記号と基本動作列の組合せのスコアを算出する。

【0014】

1つの態様では、前記セグメントは、隠れマルコフモデル（HMM）を用いて、基本動作のクラスに分類される。

【0015】

1つの態様では、前記上位記号は、運動競技における技である。

各技に対応するスコアを、実際の採点と整合させるように設定することで、対象の動作から技を同定すると同時に、自動採点に用いることができる技の難度に関連するスコア（例えば、体操競技におけるDスコア）を出力することができる。

【0016】

1つの態様では、得られた上位記号をセグメントとして、前記対象の動作をさらに上位の動作を表す上位記号列に変換してもよい。

【発明の効果】

【0017】

本発明は、対象の動作データを上位記号列に変換することによって、対象の動作を認識するものである。

【0018】

本実施形態は、体操競技の技は得点が高くなるように切り出された基本動作の系列であることに着目して技を識別するものである。

【図面の簡単な説明】

【0019】

【図1】動作の認識装置の全体図である。

【図2】動作の認識装置のハードウェア構成を示す概略図である。

【図3】動作の認識工程を示すフローチャートである。

【図4】動作における技の同定工程を示すフローチャートである。

【図5】セグメントを基本動作のクラスに分類する分類器の構成を示す図である。

【図6】対象の動作に対応する基本動作列を上位記号列に変換する工程を示す図である。

【図7】あん馬の動作における分節点の条件を説明する図である。

【図8】本実施形態に係る動作認識を示す概念図である。

【発明を実施するための形態】

【0020】

[A] 本実施形態に係る動作認識装置の概要

図1に示すように、本実施形態に係る動作認識装置は、処理部と記憶部を備えており、モーションキャプチャで取得された対象の動作データに基づいて、対象の動作の解釈を行う。動作認識装置は1つあるいは複数のコンピュータから構成される。図2に示すように、モーションキャプチャは、1つの態様では、対象の動作の動画データを取得するカメラと、動画データに基づいて対象の動作を表す時系列データを取得する1つあるいは複数のコンピュータから構成される。コンピュータは、入力部、処理部、記憶部（RAM、ROM）、出力部を備えている。1つの態様では、出力部はディスプレイを含んでもよい。

【0021】

対象の動作は、モーションキャプチャによって取得された当該対象の姿勢の時系列データによって規定される。対象の姿勢は、対象の身体上の複数の特徴点（典型的には関節）によって特定され、各フレームにおいて、複数の特徴点の3次元座標値を取得することで、複数の特徴点の3次元座標値の時系列データから対象の動作を規定する。本実施形態に用いられるモーションキャプチャの種類は限定されず、特徴点を特定する光学式マーカを用

10

20

30

40

50

いた光学式モーションキャプチャ、加速度センサやジャイロスコープ、地磁気センサなどのいわゆる慣性センサを対象の身体に装着して、対象のモーションデータを取得する方式、光学式マーカやセンサを装着しない、いわゆるマーカレスモーションキャプチャ等を例示することができる。対象の自然な動作を妨げないという観点からは、マーカレスモーションキャプチャが有利である。マーカレスモーションキャプチャとしては、カメラと深度センサを備えたシステム（Kinectに代表される）を用いたモーションキャプチャ、あるいは、深層学習を用い、1視点あるいは複数視点からのRGB画像を解析してモーションデータを取得するモーションキャプチャを例示することができる。

【0022】

認識動作装置の処理部は、入力された対象の姿勢の時系列データを複数のセグメントに分節する分節部と、各セグメントを、当該セグメントの時系列データを用いて基本動作のクラスに分類する分類部と、対象の動作を、動的計画法を用いた最適化計算により、連続する基本動作である上位記号の時系列に変換する動作認識部と、を備えている。

10

【0023】

セグメントは、対象の動作を規定する時系列データを分節することで得られる分節化された動作データであり、複数フレームの時系列データから構成される。全てのセグメントの系列が対象の動作を規定する時系列データである。セグメントは、いずれかの基本動作に対応するように動作データから切り出される。対象の動作の分節化によって得られたセグメントは、基本動作のいずれかのクラスに属するように分類され、いずれかの基本動作として識別される。全てのセグメントがいずれかのクラスに分類されることで、対象の動作を分節化したセグメント列は、基本動作列に変換される。

20

【0024】

基本動作は、複数フレームの時系列データであり、時系列に連続することで対象の動作を構成する動作単位ないしエレメントである。連続する複数の基本動作（すなわち、基本動作列）を、ある纏まった動作として認識すること考え、この纏まった上位の動作を、上位記号と呼ぶ。上位記号は、典型的には、知られている基本動作の系列である。どのような動作を基本動作、上位記号として扱うかについては、対象動作、動作認識の目的、分節化手法等によって異なり得る。例えば、対象の動作から技を同定する場合には、上位記号として「技」があり、この「技」が連続する基本動作ないしエレメントによって構成されることになる。

30

【0025】

連続する複数の基本運動からなる上位記号としては、技以外にも、例えば、「切る」「炒める」「盛る」の基本動作列を上位記号である「料理する」、「片付ける」「掃く」「拭く」の基本動作列を上記記号である「掃除する」、「部品Aをはめる」「ねじを回す」「部品Bを載せる」の基本動作列を上位記号である「（工場ラインなどでの）組み立てる」として言語的な抽象化に用いることができる。

【0026】

認識動作装置の記憶部には、対象の動作を規定する対象の姿勢の時系列データ、分節点の条件、学習モデルを備えた基本動作のクラス、上位記号に対応する基本動作列及びスコアを格納したテーブル、が格納されている。1つの態様では、基本動作のクラスは、隠れマルコフモデル（以下、「HMM」という）によって特徴付けられている。また、上位記号と基本動作列の組からスコアを算出する関数が用意されていればよく、上記テーブルに代えて、あるは、加えて、上位記号と基本動作列の組合せのスコアを確率値（尤度）として計算する関数を学習しておき、記憶部に、上記関数のパラメータを記憶させてもよい。さらに、動作認識部によって取得された上位記号の系列が、対応する基本動作列と共に格納される。記憶部には、対象の動作を撮影した動画データを格納してもよい。

40

【0027】

図3に示すように、本実施形態に係る動作認識方法は、対象の動作を規定する時系列データを受信すること、受信された時系列データを複数のセグメントに分節すること、セグメントを基本動作のクラスに分類して、対象の動作を基本動作の列に変換すること、前記対

50

象の動作を、動的計画法を用いた最適化計算により、連続する基本動作からなる上位記号の系列に変換すること、からなり、変換された上位記号の列に基づいて動作認識ないし行動理解を行う。

【0028】

[B]対象の動作の分節

対象の動作を複数のセグメントに分節する。各セグメントは、対象の姿勢を特定する複数フレームの時系列データである。1つの態様では、対象の特定の姿勢や対象と環境との位置関係等が、分節条件として予め決められており、フレーム t における対象の姿勢や対象と環境との位置関係が分節条件に合致した時に、フレーム t を分節点とする。この場合、例えば、フレーム t が新しいセグメントの先頭フレームとなる。

10

【0029】

例えば、あん馬を例にとると、あん馬の分節点については、審判員にどのように分節点を見つけているかをヒアリングし、それをヒントに以下のような条件を設定することができる(図7参照)。

- (a)両手があん馬に接地する、
- (b)上体が下向きである(体躯ベクトルの z 軸成分が負である)、
- (c)上体あん馬長手方向に対して平行もしくは垂直である、

の3つの条件を全て満たしている点を動作の分節点として検出する。

あん馬における分節点の詳細については、特許文献1を参照することができる。

対象があん馬を演技する時の対象の姿勢の時系列データ、あん馬の所定部位との接触条件、あん馬に対する向きから、上記条件を満たす姿勢のフレームを検出することで、分節点を決定する。これによって、あん馬を演技する対象の動作は、複数のセグメントに分節される。

20

【0030】

あん馬以外の他の体操競技、また他の採点競技でも、審判員からのヒアリングやルールブックを参照することで、分節点を設定することが可能である。したがって、例えば鉄棒や平行棒等の他の体操競技、フィギュアスケート等の他の採点競技においても、分節点を設定して、対象の動作を分節することができる。

【0031】

上述のように、あん馬の場合は人手で分節点の条件を設定しているが、動作認識としての技識別を、一般の行動理解へ拡張する場合、分節点の設定には、あん馬と同様に、行動に対して分節点を人(教師)が教えて、運動と分節点の関係を機械学習する方法(教師あり学習)を採り得る。あるいは、動作データから運動パターンを自動で抽出する教師なし学習を行う方法、体の一部の速度が零となるストップモーションを検出する方法、運動の分布が大きく変化する点を検出する方法、運動の予測が難しくなる点を検出する方法等も採り得る。

30

【0032】

対象の動作を分節するにあたり、対象自身の姿勢データに加えて、環境や物体との接触条件、位置関係等を付帯情報として分節化に用い得ることが当業者に理解される。人間の動作の多くは、環境や物体と関わる。例えば、料理をする動作であれば、対象の動作に加えて、対象と食材や調理器具等との接触状態、食材や調理器具等に対する対象の位置や向き等が重要な情報となり得る。掃除をする動作であれば、対象の動作に加えて、対象と掃除用具との接触状態、掃除用具に対する対象の位置や向き、環境に対する対象の位置や向き、掃除用具と環境の接触状態等が重要な情報になり得る。組立作業時の動作であれば、対象の動作に加えて、対象の動作に加えて、対象と工具や部品との接触状態、組立体に対する対象の位置や向き、部品と組立体との接触状態等が重要な情報になり得る。人間の運動において、動作自体だけでなく動作と環境や動作と物体の関係から、動作の分節の境目を発見して、動作を分節することができる。

40

【0033】

[C]セグメントの分類・認識

50

動作認識装置の記憶部には、多様な基本動作が分類されたクラスないしカテゴリが用意されている。1つの態様では、各クラスは、学習モデル（例えば、HMM）によって特徴づけられている。動作認識装置の処理部は、学習モデル（例えば、HMM）を用いた分類器を構成しており、分類器は、セグメントの情報（時系列データ）と学習モデル（例えば、HMM）を用いて、入力されたセグメントが、どのクラスに属するかを決定する。

【0034】

本実施形態では、動作データを分節して得られたセグメントをクラスタリングし、各クラスに含まれる動作データをHMMを用いて学習する。HMMは基本動作のクラスを表現し、認識に当たっては分節化された動作データであるセグメントは、当該動作データが生成される確率が最大となるHMMとして分類される。

10

【0035】

基本動作の学習においては、多数のセグメントの分類に教師あり分類を採用する。具体的には、各セグメント（姿勢の時系列データ）に、例えば体操では審判や専門家の判断によるラベルを貼り教師データとして使う。審査員による採点例のデータを教師データに含めることで、主観による審美的な評価の情報を取り入れることが可能となる。これによって専門家の視点や感性評価を扱うことが可能になるとともに、競技データの判定を実際の審査員による判定に近づけることができる。体操等の採点競技において、審判などが着目するキーフレームに着目して動作の分節点を検出することで、得られたセグメントは直感に合致した運動であり、基本動作ラベルを容易に付与することができる。基本動作ラベルを教師として動作データ（セグメント）をクラスタリングして、基本動作ごとに統計モデルを学習する。この統計モデルを用いて、運動（セグメント）を識別することによって統計モデルに紐付いた基本動作ラベルとして識別することができる。すなわち、同一の基本運動ラベルが貼られた分節データのみをクラスとし、各クラス毎に統計モデルを学習することで基本動作を獲得する。本実施形態では、基本動作の獲得は、分類された各クラスを教師データとして生成確率を最大にするように学習する隠れマルコフモデル（HMM）の計算を行う。

20

【0036】

基本動作の学習についてより具体的に説明する。複数の対象の複数の演技から動作データを取得する。分節化により得られた多数のセグメント（分節化された動作データ）を、学習用の基本動作データ（すなわち、学習データ）として用いる。複数の基本動作のラベルを用意し、人手によって、いずれかの基本動作のラベルを学習データに貼り付ける。多数の学習データを、動的時間伸縮法（Dynamic time warping、DTW）を用いて、階層的に木構造のクラスに分類し、同じラベルが付されたセグメントのみから構成されるクラスを見つけ、そのクラスを、各ラベルを代表するクラスとする。各クラスに含まれるセグメントをHMMに符号化して、各クラスを基本動作のクラスとして扱う。Baum-Welchアルゴリズムを用いて、HMMのパラメータを最適化することで、学習データが生成される確率が最大となるようなHMMが形成される。HMMによる時系列データの分類器は当業者に知られており、対象の動作を認識する場合に、対象の動作データの分節化により得られた各セグメントは、各セグメントを規定する動作データが生成される確率が最大となるHMMとして分類・認識される。

30

40

【0037】

身体運動データの扱いにおいては、身体差に応じてスケールリング行って、スケールリング後の変数を用いてもよい。運動の速度についても個人差を吸収するようにスケールリングを行ってもよい。具体的には、HMMの分類器を構成する処理、および観測データ（セグメント）がどの基本動作のクラスに属するかを決定する処理において、スケールリングを行う。標準体型の身長と対象の身長の比率を計算し、その比率に応じて対象の運動データを拡大・縮小する。セグメントが、決められた固定長のセグメントとなるように時間方向に伸縮するようなフレーム補間処理を施してもよい。スケールリングによって、競技選手の身体の大きさやそれによる運動速度の違いを補正することができる。

【0038】

50

動作データのセグメント（姿勢の時系列データ）を分類・認識する手法は、HMMを用いて分節化されたデータを用いる分類手法に限定されるものではなく、当業者に既知の他の分類・認識手法（典型的には教師データを用いた手法）を用いることができ、例えば、ニューラルネットワークを用いてセグメントを分類してもよい。

【0039】

[D]連続する基本動作からなる上位記号の認識

動作認識部は、基本動作の一連からなる動作（上位記号）を発見して、その上位記号が前後関係から適切な説明かどうかを評価して決定する。動作認識部は、基本動作の系列から上位記号、体操などの運動競技では「技」、を発見する問題を、動的計画法を適用して最適化問題の解として計算する。

10

【0040】

動作認識を、競技の技の同定を例として説明する。対象の動作データは複数のセグメントに分節され、各セグメントを、HMMを用いて基本動作のクラスに分類することによって、対象の動作データは、基本動作の時系列（ $1, 2, \dots, N$ ）へ変換される。技は複数の基本動作の連続であり、技として認識されることに適した基本動作列を抽出する。

【0041】

基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_N$$

20

からk個の技を抽出した時に得られる最大スコアを、

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N)$$

とおくと、

以下の関数漸化式

$$f_k(\lambda_l, \lambda_{l+1}, \dots, \lambda_N) = \max_{l < m < N} [g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m) + f_{k-1}(\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_N)]$$

を導くことができる。

30

【0042】

ここで、

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

は、基本動作列

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_m$$

を1つの技として識別したときに得られるスコアである。

40

【0043】

動作認識装置の記憶部には、基本動作列と技及びスコアの対応を規定したテーブルが格納されている。具体的には、人手によって連続する基本動作に対して技のラベルが貼り付けられており、連続する基本動作は、各基本動作を特徴づけるHMMの列によって規定される。技のID及び技ラベルに対して、基本動作列（HMMの列）がデータベースとして格納されている。

【0044】

1つの態様では、
基本動作列

50

$$\lambda_l, \lambda_{l+1}, \dots, \lambda_m$$

がデータベース中にあれば、

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

のスコアは 1、なければ 0 を取るものとする。そして、上記基本動作列がデータベース中にある時には、基本動作列に対応する技として動作を識別する。

すなわち、各基本動作列（HMM の列）はスコア 1 が割り当てられており、基本動作列がデータベース中にあれば、評価関数 g のスコアは 1、なければ 0 を取るものとする。そして、漸化式を用いた評価関数 f が最大となるように、基本動作列の纏まりを発見しながら、基本動作列をその纏まりに対応する技として識別する。

【0045】

基本動作列と技及びスコアの対応を規定したテーブルを用いる場合について説明したが、上位記号と基本動作列の組からスコアを算出する関数が用意されていればよい。例えば、上位記号と基本動作列の組合せのスコアを確率値（尤度）として計算する関数を学習しておき、記憶部に、上記関数のパラメータを記憶させてもよい。そして、処理部が前記パラメータを利用して上位記号と基本動作列の組合せのスコアを算出してもよい。この関数としては、上位記号と基本動作列を入力として、尤度を出力する統計モデルやニューラルネットワーク等を用いることができる。

【0046】

上記漸化式は、基本動作 1 からどこまでの連続する基本動作列を 1 つの技として識別するのが最適であるかを計算していることになり、この関数漸化式を、動的計画法を用いて解くことによって、最適な基本動作列の纏まりとその纏まりに相当する上位記号（上記例では技）を求めることができる。動作認識部は、人間の動作を説明するために、分節化後に認識された基本動作の系列から、上位記号（典型的には知られている基本動作の系列）を発見して動作を判別する。このとき、前後の関係などによって判別結果が変わり得るが、この判別問題を評価関数を最大化する最適化問題として扱う。なお、上位記号の種類や動作認識の目的等によっては、最大スコアの上位記号列だけでなく、高スコアの複数の上位記号列を出力してもよい。

【0047】

[E] 体操競技における技の認識

既述のように、本実施形態に係る動作認識装置及び方法は、体操競技における演技中の全身運動データからの技の認識ないし同定に適用することができる。演技は、基本動作の連続としての技から構成される。本実施形態は、演技中の全身運動データから基本動作を識別し、技として適した基本動作の連続を見つけることを動的計画問題に帰着させることを特徴とする。

【0048】

具体的には、体操競技における技の認識処理は、

演技者の動作データ（演技者の姿勢の時系列データ）をセグメント（基本動作として認識されるように切り出される）に分節すること、

分節化された動作であるセグメントを基本動作のクラスに分類すること、具体的には、例えば、各セグメントを、HMM を用いて基本動作のクラスに分類することによって、動作データを基本動作の時系列へ変換すること、

技として適した基本動作の連続を動的計画法を用いて抽出することで、技を自動認識すること、

からなる。

【0049】

図 4、図 8 を参照しつつ説明すると、対象の演技を撮影した動画から対象の動作データを

10

20

30

40

50

取得する。動作データは、審判員もしくはルールブックを参考に設定した分節点を用いて、複数のセグメントに分節され、動作データはセグメントの時系列として表される。

【 0 0 5 0 】

各セグメントは、例えばHMMによって、いずれかの基本動作（a,b,c,d,...は、基本動作ないしHMMのラベルである）として認識されて基本動作のクラスに分類される。各セグメントがいずれかの基本動作として認識されることで、対象の動作を、基本動作の時系列（クラスに対応するHMMの列bdcafafbd c）に変換する。

【 0 0 5 1 】

基本動作系列から技を識別し、対象の動作は、技の系列（BAAB...）に変換される。体操の場合では、「技」は基本動作が一連の順番を持つものとして定まっている。基本動作の系列を技として識別可能な組み合わせは一般には唯一ではない。体操では競技者の採点結果が最大になるように識別して技の識別を行い、それが採点され結果となる。採点結果が最大になるように動作を識別する方法を最適化問題として表し、さらに、基本動作の生成過程を動的システムとして表現することで、採点を動的計画法によって求める。

【 0 0 5 2 】

動作認識装置の記憶部には、基本動作列と技及びスコアの対応を規定したテーブルが格納されている。

上記漸化式において、関数

$$g(\lambda_l, \lambda_{l+1}, \dots, \lambda_m)$$

は、評価値すなわちスコアを決定するいわば評価関数である。評価関数を用いたスコア算出について、簡単な例を示して説明する。

【 0 0 5 3 】

表1の例では、各技についてID及びラベルが付されており、技ラベルAに対応して複数の基本動作列abc,abd,abeが割り当てられている。全ての基本動作列abc,abd,abeのスコアは1である。動作データから変換された基本動作列において、基本動作列abc,abd,abeのいずれかが認識された場合には、評価関数はスコア1を出力する。

【表1】

ID	技ラベル	基本動作列	スコア
1	A	abc	1
		abd	1
		abe	1

【 0 0 5 4 】

表2の例では、各技についてID及びラベルが付されており、技ラベルAに対応して複数の基本動作列abc,abd,abeが割り当てられている。基本動作列abc,abd,abeのスコアはそれぞれ1、0.9、0.8と異なっており、動作データから変換された基本動作列において、基本動作列abc,abd,abeが認識された場合には、評価関数は、それぞれ、スコア1、0.9、0.8を出力する。スコアの設定手法は限定されず、例えば、大量の演技の動作データのサンプルから経験的に設定してもよく、あるいは、何らかの形で算出した尤度をスコアとして用いてもよい。

【表 2】

ID	技ラベル	基本動作列	スコア
1	A	abc	1
		abd	0.9
		abe	0.8

10

【0055】

表3の例では、各技についてID及びラベルが付されており、技ラベルAに対応して複数の基本動作列abc,abd,abeが割り当てられている。全ての基本動作列abc,abd,abeのスコアは1であり、動作データから変換された基本動作列において、基本動作列abc,abd,abeのいずれかが認識された場合には、評価関数はスコア1を出力する。技ラベルBに対応して複数の基本動作列ghij,ghikが割り当てられている。全ての基本動作列のスコアは3であり、動作データから変換された基本動作列において、基本動作列ghij,ghikのいずれかが認識された場合には、評価関数はスコア3を出力する。技ラベルAのスコアと技ラベルBのスコアの差は、技の難度を反映したものである。また、ここでの評価関数は、技ラベルA、技ラベルBに対応する全ての基本動作列にスコア1を設定し、技ラベルBに別途難度に依存した加算ポイントを設定することと等価である。

20

【表 3】

ID	技ラベル	基本動作列	スコア
1	A	abc	1
		abd	1
		abe	1
2	B	ghij	3
		ghik	3

30

40

【0056】

上記説明では技ラベルについて簡略して説明したが、実際には、表4に例示するように、人手によって連続する基本動作に対して技のラベルが貼り付けられている。技のID及び技ラベルに対して、基本動作列（HMMの列）がデータベースとして格納されている（表4では省略）。

50

【表 4】

Maneuver index	Maneuver description
1	For transition
2	Scissor backward
3	Any circle or flair in side support
4	Circles with travel fwd. in side support
5	Circles in cross support frontways on end
6	On the end, Russian wend. with 720° or 900°
7	Travel bwd. In cross sup. From one horse part or pommel to another (1/3)
8	Reverse stöckli or direct stöckli A through hdst.
9	Scissor forward
10	Russian wendeswing with 180° or 270°
11	Circles in cross support on one pommel (from or to 1/4t. fwd.)
12	Direct stöckli B
13	From side sup. on end, circle or flair to hdst. and travel 3/3 with 450° or more turn

10

【0057】

本実施形態では、技と基本動作列の組からスコアを算出する関数が用意されており、この関数によって算出されたスコアを用いて動的計画法を用いた最適化計算が実行される。1つの態様では、前記関数は、技に対応する1つあるいは複数の基本動作列及びスコアを備えたテーブルを用いてスコアを算出する。1つの態様では、技と基本動作列の組合せのスコアを確率値（尤度）として計算する関数を学習しておき、上記関数のパラメータを利用して技と基本動作列の組合せのスコアを算出する。

20

【0058】

要約すると、本実施形態に係る体操競技における技の認識装置は、

- （１）全身運動の切れ目をみつける分節化、
- （２）切り取られたセグメントを基本運動と比較して判別する計算、及び、
- （３）連続した基本動作系列から技を識別する計算

を順に処理することによって運動競技における技を認識するシステムであり、体操演技の分節化、基本動作の識別、その時系列から技を認識する計算論の基礎を提供する。

30

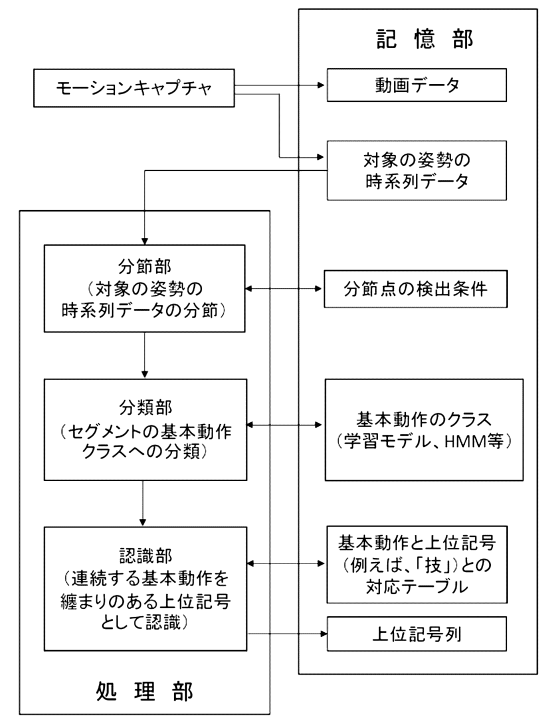
【0059】

体操競技における演技の得点はDスコアとEスコアの合計によって算出される。Dスコアは演技内容の難しさ、すなわち技の難度、を表したもので、Eスコアは演技実施の完成度を表す。本実施形態において、技と基本動作列の組からスコアを算出する関数を実際のDスコアの算出方法に整合させる（例えば、テーブルに格納されるスコアを、採点規則をまとめたルールブック等を参考にして実際のDスコアの算出方法に整合させる）ことで、体操競技の自動採点システムに適用することができる。したがって、本実施形態に係る動作認識装置及び方法は、演技の自動採点システム、客観的評価をフィードバックする運動指導や育成プログラムに利用可能である。

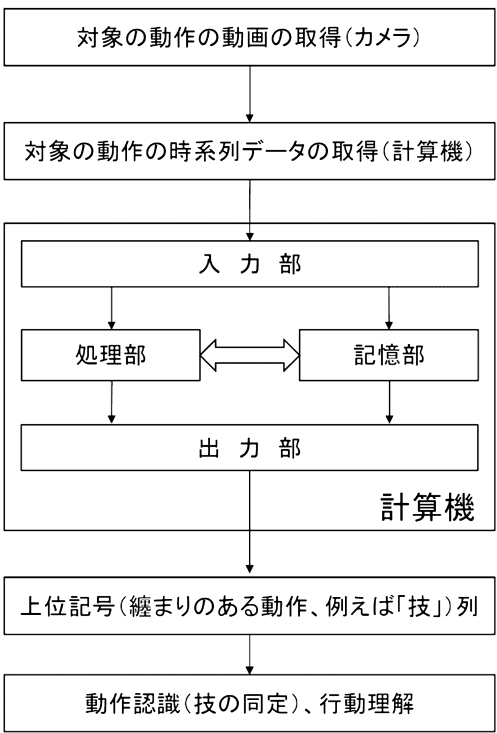
40

【 図 面 】

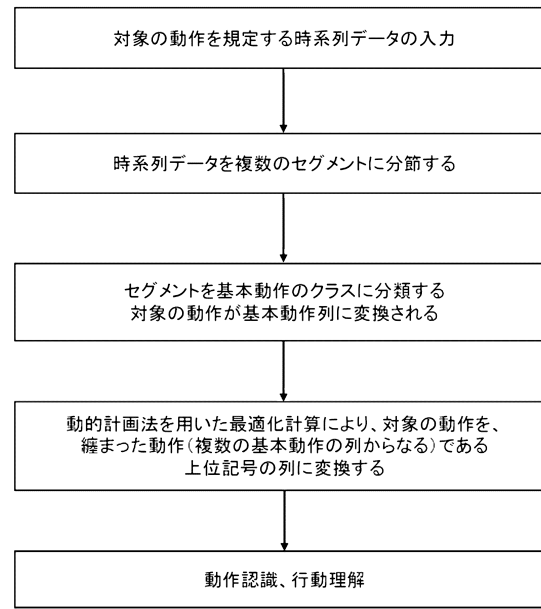
【 図 1 】



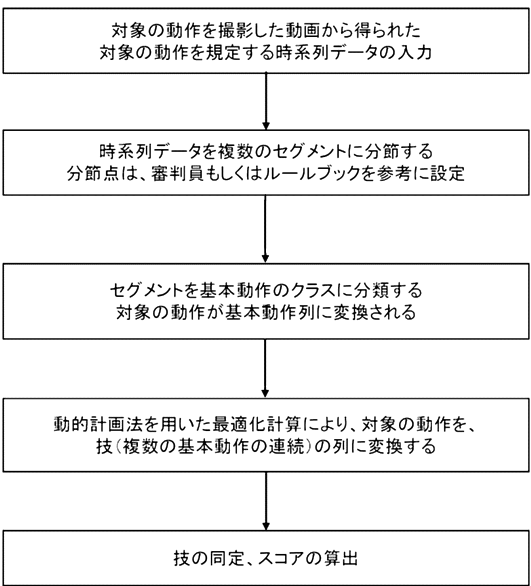
【 図 2 】



【 図 3 】



【 図 4 】



10

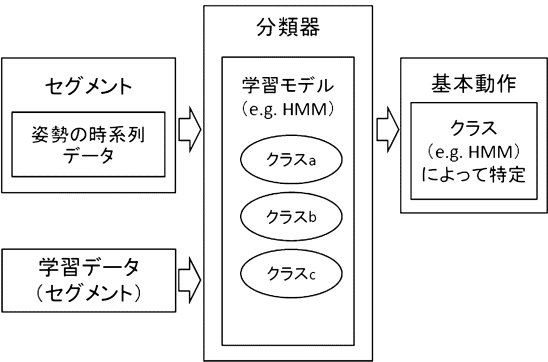
20

30

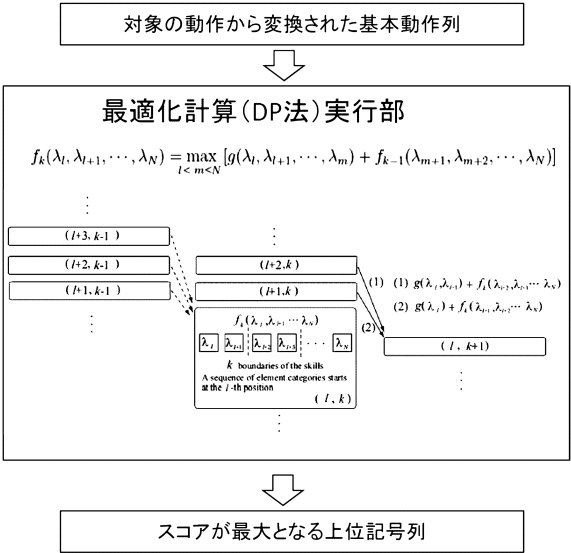
40

50

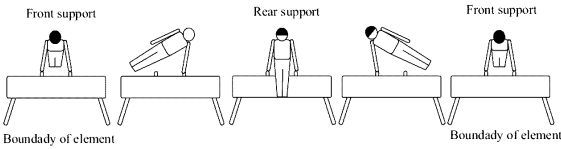
【 図 5 】



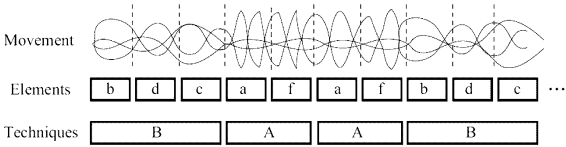
【 図 6 】



【 図 7 】



【 図 8 】



10

20

30

40

50

フロントページの続き

大学法人東京大学内

審査官 真木 健彦

(56)参考文献 国際公開第 2 0 0 7 / 1 3 8 8 8 5 (W O , A 1)

国際公開第 2 0 1 8 / 0 7 0 4 1 4 (W O , A 1)

特開 2 0 1 0 - 2 1 3 7 8 2 (J P , A)

(58)調査した分野 (Int.Cl. , D B 名)

G 0 6 T 7 / 2 0

G 0 6 T 7 / 1 4 3

G 0 6 N 7 / 0 0

G 0 6 N 9 9 / 0 0