

特徴学習を用いた空手組手の動作識別 Classification of karate motions using feature learning

佐藤恵一[†] 栗山繁[‡]
Keiichi Sato Shigeru Kuriyama

1. はじめに

本報告では、空手の訓練に有効なスコアブックの自動作成を目的とした、組手動作の技単位での抽出と識別手法について提案する。本研究で扱うスコアブックは、一般に知られている野球のスコアブックと同様に、プレーの内容を表す特徴的な情報から試合中の技の短所と長所を知る手がかりとするために、統計的手法により弱点や勝利への戦略を発見し、訓練の目的と方向性の決定を支援する。しかしながら、空手等の高速かつ連続的に変化する動作に対してスコアブックを作成する手法は未だに提案されていない。

スポーツのスコアブック作成またはそれを用いたサービスに関連する研究として、サッカーを対象とした手法が提案されている。例えば、ビデオ映像の画像解析からゴール、ライン、選手、ボール位置を検出し、ゴールキック等の特定映像を抽出する手法[1]や、ボールタッチの回数でプレー種別を簡易的に自動識別する手法[2, 3]、および自動では識別できないプレーを GUI ツールにより手動で効率良く識別して作成したスコアブックからアニメーションを生成する手法[4]等が提案されている。

サッカー以外では、歓声や野球における打撃音や捕球音など音により映像データをプレーごとに自動分割し、手動で注釈を加えて映像版スコアブックを作成する手法[5, 6]が提案されている。しかしながら、[1-4]の手法の様な、画像処理による身体的位置検出だけでは空手の技単位での動作分割はできない。また、空手の場合は特徴的な音が技とは関連していないので、[5, 6]の手法の様に音を用いた技単位での動作分割は困難であり、その識別も不可能である。

空手の技を訓練する研究として、訓練者の手足に加速度センサーを装着して実時間計測したデータを音でフィードバックする手法[7]が提案されており、初級者の訓練に対する効果を実証されている。しかしながら、この研究では組手の動作を解析の対象とはしていない。

ロボットの人動作の模倣やアニメーションの合成等を目的とした従来の動作識別法は、動作を技などの意味の単位で分割するものは少ない。また、識別も半自動のものが多く、その精度に関しても詳細には言及されていない。

2. 動作の分割

本研究では、空手組手の動作解析のために光学式のモーションキャプチャ装置で3次元計測した組手の動作データを用いる。人の動作データから技の基本単位を抽出するには、データを最小単位の動作に分割する必要がある。

動作を識別する先行研究をその分割手法別に分類すると、以下の4種に分類できる。

(1) 関節部位の回転角に注目して動作分割し、クラスタ分析やHMM (Hidden Markov Model) で識別する手法[8, 9]。

(2) 音声分析技術の Chunk の概念を適用して動作分割し、HMM で識別する手法[10]。

(3) プレーヤーが技を繰り出す際、足が床に加える力のピークを利用して動作分割し、SVM (Support Vector Machine) で識別する手法[11]。

(4) 手足の運動速度が0になる点を利用して動作分割し、部位の軌跡情報からDP (Dynamic Programming) で識別する手法[12, 13]、または音声認識技術DTW (Dynamic Time Warping) を用いてDPで識別する手法[14]。

本研究では、空手の技単位で動作を抽出し、各種の特徴量の構成について動作識別を試み、その識別精度について詳細に調査する。ゆえに、空手の技に顕著な特性を考慮して、動作の分割方法として「手足の速度が0になる点を利

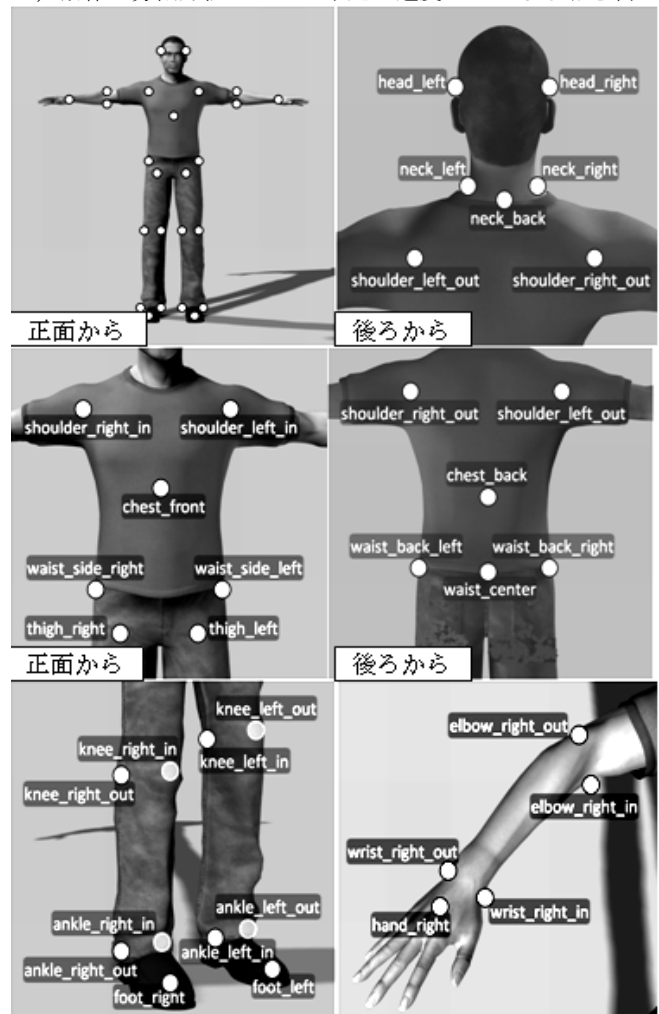


図1 人体に設置したマーカ

[†] 函館工業高等専門学校情報工学科

Department of Computer Engineering, Hakodate National College of Technology

[‡] 豊橋技術科学大学情報工学系

Department of Information and Computer Science Toyohashi University of Technology

用して動作分割する」方法を空手の技単位で動作分割する方法として拡張する。

従来の研究で扱ってきた特徴量は一人のプレーヤーから抽出したものに限定されているので、相手の攻撃に対する反応動作を識別することができない。反応動作以外にも、牽制動作は通常の突き、蹴り技と同じ動作であることが多く識別ができない。ゆえに我々は、二人のプレーヤーの動作から抽出した新たな特徴量を SVM で識別する手法を提案する。

2.1 空手組手の測定

今回の実験では、12 台の赤外線カメラを搭載した VICON 社製の光学式モーションキャプチャ装置を用い、上級者(三段)と初級者の伝統系空手の組手を 16 ラウンド(ただし、各ラウンドの所要時間は約 40 秒)を毎秒 120 フレームの時間解像度で測定した。各プレーヤーには図 1 に示すように 40 個の計測マーカーを身体やサポーターに直接貼付け[16]、その座標位置(x, y, z)から関節に相当する位置を推定した 21 個の節点(以降ノードと呼ぶ)を構成し、図 2 に示す仮想的な骨格構造を構築する。

組手測定において頻繁にマーカーを見失う事象が発生したため、測定後に計測データを手動で補完処理する際のガイドマーカーとして用いるための、ノードの位置計算に使用しないマーカーもプレーヤーの身体に貼付けた。また、肘(elbow_left, elbow_right)と膝(knee_left, knee_right)は、本来 2 つの測定マーカーからノード位置を推定した方が正確である。しかし、今回の組手は試合に近い動作を得るために実戦に近い形式で測定したため、肘と膝の計測マーカーの脱落損傷が頻繁に発生したので、肘と膝の外側に貼付けたマーカーの座標位置をそのままノードの位置とした。

2.2 技単位での動作抽出

スコアブック自動作成のためには、計測した動作データから意味のある技の単位で抽出しなければならない。空手組手における動作では関節部位の動きが高速であり、瞬時に静止するという特徴がある。ゆえに 1 章で述べたように、関節部位速度の極値検出に基づいた動作分割方法を採用する。

今回の実験では、右手(hand_right),左手(hand_left),右足(foot_right),左足(foot_left)の計測マーカーに対する 3 次元座標の時系列データから各々の速度ベクトルを計算する。しかしながら、座標データには計測雑音等の影響から微小な変動が含まれるので、生のデータを用いた場合には極めて小さな動作片まで検出してしまふ。そこで以下の式により各フレームでの速度ベクトルを平滑化する。

$$\begin{aligned} \vec{V}_{frm} &= (x_{frm}, y_{frm}, z_{frm}) \\ &= \frac{\sum_{i=frm-rg_a}^{frm+rg_a-1} \vec{GP}_i - \vec{GP}_{i+1}}{2 \times rg_a} \end{aligned} \quad (1)$$

\vec{V}_{frm} : フレーム毎の速度ベクトル
 frm : フレーム番号($rg_a < frm < \text{最終フレーム番号} - rg_a - 1$)
 rg_a : 速度の平均を求める範囲(フレーム数)
 \vec{GP} : グローバル座標系の位置

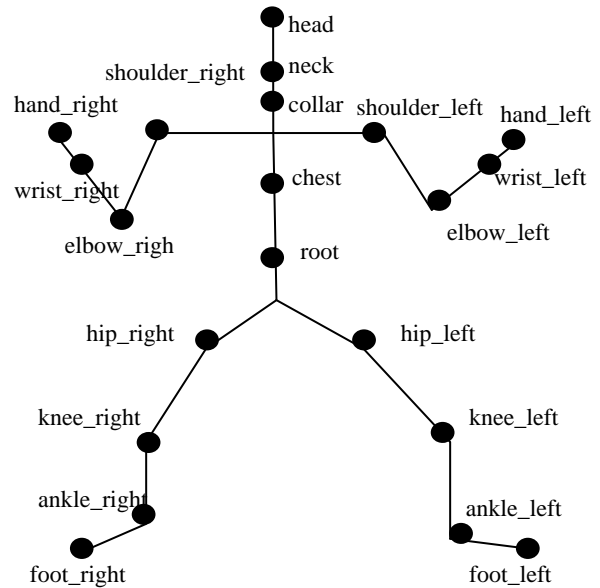


図 2 推測して得られた骨格構造

(1)式では、速度を求めるフレーム番号 frm を基準にして rg_a フレーム数前後の範囲でグローバル座標系における位置の移動平均を求めている。ただし、今回の実験では $rg_a=10$ と設定した。(1)式で求めたフレーム毎の速度ベクトルに対して以下の式により速度ベクトルのノルム(以後、単に速度と呼ぶ)を求め、この値から速度グラフを作成した。

$$V_{frm} = \sqrt{x_{frm}^2 + y_{frm}^2 + z_{frm}^2} \quad (2)$$

次に、 $rg_b < frm < \text{最終フレーム番号} - rg_b - 1$ の frm の最大値を以下の(a)から(e)の手順で求める。ここでは、最初に、測定した動作データ全ての最大値を求め、次にそれらの最大値間の最小値を求める。

(a)最大値を判定するフレーム番号 frm を基準に rg_b フレーム数前後の範囲で速度変化量の総和を(3)式で求める。今回の実験では $rg_b=20$ とした。

$$SV_{frm} = \sum_{i=frm-rg_b}^{frm+rg_b-1} V_i - V_{i+1} \quad (3)$$

frm : 最大値を判定するフレーム番号, V_i : フレーム番号 i の部位速度,
 rg_b : 速度変化量総和範囲

(b) $SV_{frm} < thresh_a \times 2$ (4)
 真なら frm は最大値でない。

(c)(4)式が偽なら(5)式の判断をする.

$$SV_{frm} > thresh_b \times 2 \quad (5)$$

(d)(5)式の判断が真なら $rg_b = rg_b/2$
偽なら何もしない.

(e)フレーム番号 $frm-rg_b < frm < frm+rg_b-1$ の範囲で速度値の大小関係を調べる. rg_b の値を変更した場合はもとに戻す.

最大と最小を求める上記手順について補足すると, (b)の判断は, 速度グラフの「微小な山」を無視するためのものであり, 実験では $thresh_a = 0.1$ とした. また, (c), (d)の処理は, 速度グラフで大きな山の直近に出現する「比較的小さな山」を無視するためのものである. これは大きく動作した技の衝撃等で生じるものである. 一方, この処理は同時に, 軸足を回転したときの動作のような全体的に速度変化が小さいプレイヤーの小さな動作は検出できる. このように(c),(d)で検出される動作は空手組手において意味ある動作分割に特化した処理である. ただし, 実験では経験的に $thresh_b=1.0$ に設定した.

上級者左手の速度グラフの極値を検出した結果を図3と図4に示す. グラフ上でマークされている点が検出された極値である. 図3のA-B間にほぼ同じ大きさのピークが2つあるが, これは「左突き」の動作である. 空手の技は多くの場合, このように単発の場合では大きさのほぼ同じ2つのピークになる事が実験により確認された.

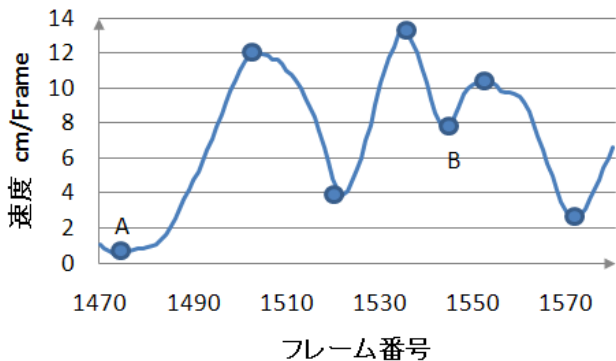


図3 上級者左手速度グラフ(a)

図4は複雑なグラフとなった例であり, A-B間に大きさの異なった2つのピークが見られる. これはプレイヤーが「左回し蹴り」を行っており, 腰の回転の逆方向に上半身をひねり, 平泳ぎのように左手を後方に大きくかく動作をしていることに起因しており, 上級者に多く見られる動作である. 次にB-D間に2つのピークがあるが, これは「回し蹴り」の大きな動作をした直後に相手が反撃しようとしたため, 素早く防御の構えとした動作を示している. このように複雑な動作となったために, C-D間で検出されなかった「比較的小さな山」と「微小な山」の2つのピークが存在する. 大きな変化量となったC点直後の「比較的小さな山」は, 極大値検出手順(c),(d)により検出されなかった. 一方, D点直前の「微小な山」は, 手順(b)により検出されなかった. ゆえに, 意味のある技単位での動作分割を目的とした本アルゴリズムが正しく機能していると考えられる.

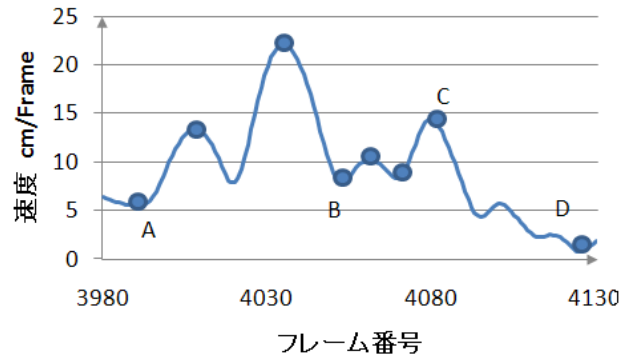


図4 上級者左手速度グラフ(b)

3. 技の識別

3.1 特徴量の抽出対象範囲

今回は手技として「突き, 受け, 牽制, 構え」, また足技としては「蹴り, 受け, 移動」の全7つの動作を識別した. 空手の基本技である「蹴り, 突き, 四肢による受け」の単位動作は2つの大きな極値を持つので, 2つのピークの区間内で特徴量を計算する.

図5に, 下肢関節の速度グラフから下肢動作の特徴量を抽出した例を示す. 図5は, 2歩「摺り足で前進」した(図5A区間)後, 「蹴り」を行い(図5C区間), 次に2歩「摺り足で前進」した(図5D区間)例を表している.

ここで, 移動の動作は, 組手でしばしば見られる「膝蹴りするようにして前進」, 「足を前に放り出して前蹴りするように前進」, 「進行方向とは反対方向へ足を振ってから前進」などは2つのピークが現れるが「摺り足による前進」は, 1つのピークしか現れない. 空手組手における移動動作のほとんどは「摺り足による前進」であるため, 図5に示すように今回は2回の「移動」で特徴量を計算する.

単位動作をすべて1つのピークに分解する方法も考えられるが, 「蹴り」, 「突き」の識別精度は2つのピークとそのピーク間の最小値の特徴量が大きく影響すると考えられるため, 「1個分のピークだけ進めて, 2つのピークごとに特徴量を計算する」(図5A区間からB区間)を繰り返す手順にした. 但し, 「蹴り」または「受け」と判定された場合は2個分のピーク進めて次の2つのピークの特徴量を計算する(図5C区間からD区間). これにより1つのピークしか現れない「移動」と2つのピークが現れる「移動」や「蹴り」, 「受け」の双方に対応した. 具体的な手順を下記の(a)~(d)に示す.

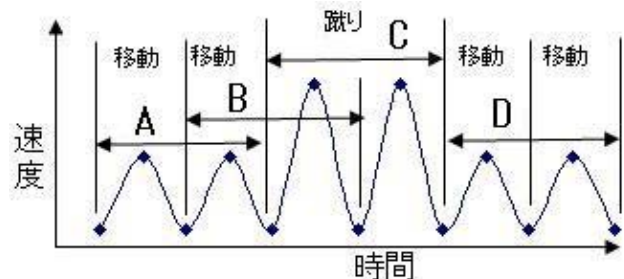


図5 下肢の特徴量抽出

- (a) 2つのピークに対して動作識別する(A区間).
- (b) (a)の識別結果が「移動」なら, 次の開始点を1個分

のピークだけ進めて次の2つのピークに対して特徴量を計算する(B区間).

- (c) 識別結果が「蹴り」または「受け」の場合(C区間), 次の開始点を2つのピーク進めて2つのピークに対して特徴量を計算する(D区間).
- (d) (a)から(c)を繰り返す.

さらに, 腕部に関する動作の「突き」, 「受け」, 「牽制」と「構え」についても同様に特徴量を算出する.

以上のように, 空手のすべての技を部位速度グラフの2つのピークで取り扱う.

3.2 特徴量の構成

被験者2名に実際に組手を実施してもらい, 計測したデータを組手主体者と対戦相手に分別した後に動作を解析した. ここで, 2つピークから計算する特徴量は, 組手主体者の動作識別に関するノードの速度, および主体者と対戦相手の全ノードの位置を, 図6に示す様に極値のフレームと隣接するフレームの時間的な中間値を抜き出した3, 5, および9フレーム分のベクトルデータで構成する.

$$\begin{aligned}
 & \text{主体者の部位速度ベクトル}(V_x, V_y, V_z) \\
 & + \{ \text{主体者の21ノード位置}(P_x, P_y, P_z) \\
 & \quad + \text{対戦相手の21ノード位置}(OP_x, OP_y, OP_z) \} \\
 & \times \text{抽出フレーム数 } 3, 5 \text{ または } 9 \tag{6}
 \end{aligned}$$

ここで, 次元数は, $3 + \{ (21 \times 3) + (21 \times 3) \} \times (3, 5, 9)$ となることに留意されたい.

(6)式の「主体者の部位速度ベクトル」は, 図7で示すようにグローバル座標系で算出する. 「主体者の21ノード位置」は, 主体者の腰(rootノード)を原点とするローカル座標系から算出する. また, 「対戦相手の21ノード位置」も主体者のローカル座標系により, 主体者に対する相対座標として算出する.

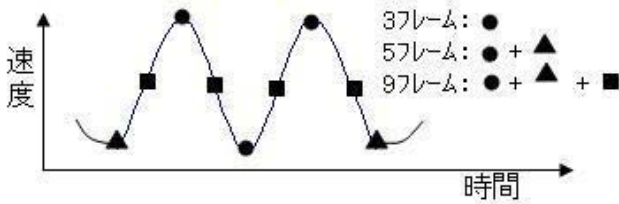


図6 姿勢データの標本化フレーム

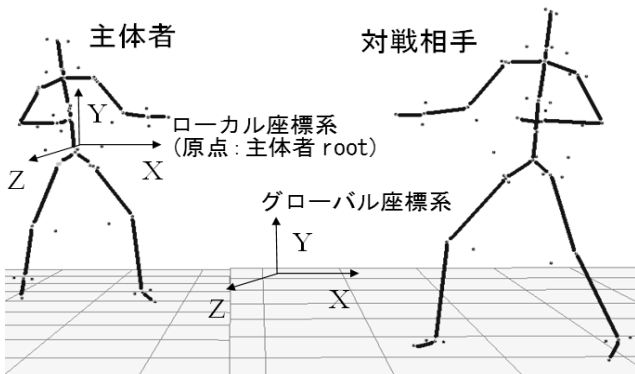


図7 特徴量算出で用いられる座標系

4. 動作識別結果

計測した動きをCG仮想人間を用いて可視化した映像とノードの速度をグラフ化した画像を観察し, 「蹴り」や「突き」等の注釈を動作ごとに設定して, SVMに学習させる教師データを作成した. ただし, SVMの機能は公開ライブラリである libsvm[17]を利用して実装した. 今回の計測で得られた各種動作の総数を表1に示す. ここで, 値が0のものは今回の組手で被験者がその動作を行わなかったことを意味する.

表1 実験で使用した動作数

手技	突き		構え		受け		牽制	
	左	右	左	右	左	右	左	右
上級者	26	37	1730	2143	11	0	14	0
初級者	51	28	1789	1789	23	4	1	0
合計	77	65	3519	3932	34	4	15	0
足技	蹴り		移動		受け			
	左	右	左	右	左	右		
上級者	11	15	1214	1148	0	0		
初級者	22	14	923	889	5	0		
合計	33	29	2137	2037	5	0		

組手の各ラウンドに対して, leave-one-out 交差検定法により動作を識別した. すなわち, 種類の識別に用いるラウンドの識別対象者のデータを除外したデータを教師データとしてSVMに学習させた. 例えば, 第1ラウンド上級者左手の識別を行う場合, 第1ラウンド初級者の左手の教師データと第2から第16ラウンド両者左手の教師データの, 全16ラウンド \times 2人 $-1=31$ 人分の教師データをSVMに学習させた. 以上の方法で組手の主体と相手の四肢動作をその種別毎に識別し, 再現率, 適合率, および再現率と適合率の調和平均であるf値を求めた. 上級者の動作を識別した結果を図8(a)から図9(b)に示す. ただし, 特徴量の構成フレーム数が3, 5, 9フレームの場合をF3, F5, F9とそれぞれ表記する.

SVMへ学習させる教師データ数が多い「移動」と「構え」は, ほぼ正確に識別された. また, F3の「突き」と「蹴り」は, 教師データ数が少ないにも関わらず精度良く識別された. さらに, F3からF9へと特徴量を構成するフレーム数が増えるとf値が減少する傾向が確認された. 図8(a)「牽制」のF5, 図9(a)「蹴り」のF5とF9は値が0となっており, その他でもF3よりF5, F9の値が低い傾向が確認できた.

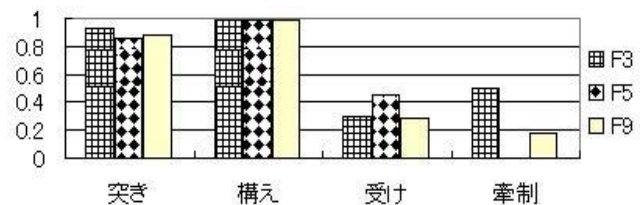


図8(a) 上級者の左手のf値

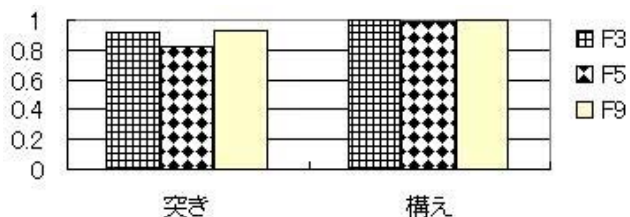


図 8(b) 上級者の右手の f 値

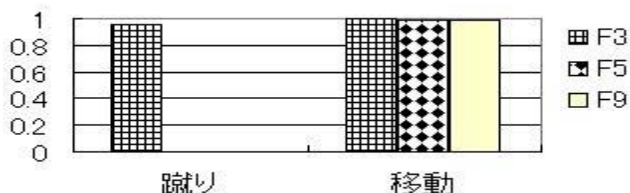


図 9(a) 上級者の左足の f 値

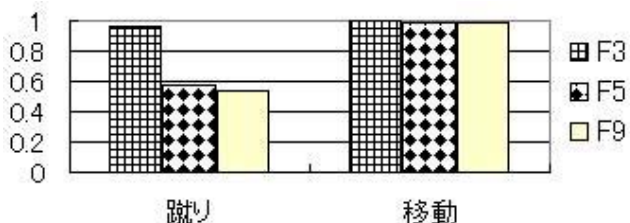


図 9(b) 上級者の右足の f 値

5. 考察

一般的には、特徴量が多いほど判断材料となる情報が増えて識別精度が向上するが、今回は F3 の場合、つまりノード速度の極値を取るフレームでの全ノード位置のみの場合に良好な識別性能が得られた。これは、極値以外のフレームでのノード位置が無意味な雑音成分として影響したと考えられる。ゆえに、極値における姿勢のみが当該動作の特徴的な情報を表すと考えられる。また、F9 が F5 よりやや良好な識別性能が得られたが、その理由として、F9 は F5 に加えて隣接する極値での姿勢を含んでいる事が考えられる。

「受け」の動作に関しては、教師データでは「構え」であるが SVM では「受け」と判定する事が誤検出を招き、精度の低下に影響したものと考えられる。そこでこの動作を再調査したところ、受け動作をしていたにも関わらず、相手が攻撃をしようとしただけで実際には攻撃していなかったことが判明した。教師データでは、相手が攻撃していない場合には受け動作でない事を正解としている。一方、空手では「受けの直後に攻撃を返す」を基本としており、上級者ほどこの動作を多く繰り返すので、SVM に対する教師データの設定が誤りであったと考えられる。さらに、「牽制」でもこのような事例が多く見られた。この教師データの曖昧さも精度を低下させる原因になったと考えられる。したがって、識別精度を向上させるには、「受け」動作のサンプル数を増やし、教師データを十分に吟味する事が重要である。

本スコアブックでは情報を集約して整理できるレイアウトとして、図 10 に示すユーザインタフェースを構築した。組手の様子は、動作識別された技が 3 DCG アニメーションを用いて画面中央に表示される。また、速度グラフの 2 つのピークを利用して、技の開始時、打撃の瞬間、および技の終了時における主体者と対戦相手の立ち位置（足跡）と姿勢を表示している。

技の姿勢、バランス、および立ち位置は重要であると考えられるが、スコアブックが扱う情報は、どのようなものが適当かは専門家の意見を今後調査する必要がある。

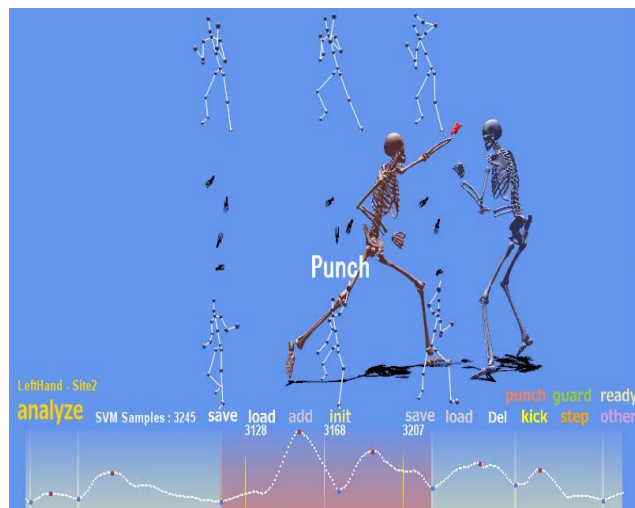


図 10 スコアブックの GUI による実行画面

6. おわりに

左右手足の速度から検出した極値を利用して空手の組み手動作を意味のある単位で分割する手法を提案した。

空手組手の技を部位速度グラフの 2 つのピーク単位で扱う手法を提案し、移動、蹴り、構え、突きの技が高い識別精度で抽出されたので、その有効性を基本技について確認することができた。

これまでの先行研究では、2 プレーヤーによる運動解析がほとんどなく、また相手の反応動作を識別するための情報の特徴量で扱う方法が存在しなかった。本研究では、対戦相手の姿勢データを特徴量として扱い、動作識別実験を行った結果、良好な結果が得られなかったものの、今後の改善方法が発見できた。

従来の動作識別手法は、一般に動作データにおける全フレームでの特徴量を用いている。しかし、本研究が対象とする空手など打撃系格闘技においては、部位速度の極値における姿勢のみの特徴量で突きや蹴り等の基本技の動作に関して、SVM を用いた識別が可能であることが確認できた。また、極値をとるフレーム以外の姿勢を特徴量に含むと f 値が減少するので、極値における姿勢のみが組手動作の種類を判別するのに有用な情報となる事が示唆された。

謝辞

本研究を行う際、多大なる助言や開発したシステムの基本となったライブラリを提供して下さった当時、豊橋技

術科学大学教員であった向井智彦先生, また, 豊橋技術科学大学の当時学生であった佐藤央君, 高橋星矢君, プログラム開発や測定を手伝ってくれました. 空手専門家として助言や組手を行っていただいた国際空手道連盟極真会豊橋本部道場指導員 金光宣明さん, 豊橋技術科学大学空手部小林利彰君, 以上の皆様に, この場をかりて心から感謝申し上げます.

参考文献

- [1] 丸尾二郎,岩井儀雄,谷内田正彦,越後富夫,飯作俊一,“サッカー映像からの特定映像イベントの抽出”,電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎 99(183),pp.31-38(1999).
- [2] 中川靖士,羽田久一,今井正和,砂原秀樹,“サッカー映像の自動ゲーム分析”,情報処理学会研究報告マルチメディア通信と分散処理研究会報告 2002(12),pp.193-198(2002).
- [3] 中川靖士,“サッカー映像の自動ゲーム分析方法の提案と評価”, UNISYS TECHNOLOGY REVIEW 第 76 号,(2003).
- [4] 橋本賢一,大城英裕,大崎保彦,山崎一城,井上誠喜,藤田米春,“サッカー競技を対象としたデジタルスコアブックのための画像処理支援”,電子情報通信学会技術研究報告 PRMU パターン認識メディア理解 103(737),pp.43-48(2004).
- [5] 三上弾,南田幸紀,森本正志,“野球映像とスコアブックの統合による効率的なメタデータ付与”,信学技報 PRMU2003-262,pp.37-42,(2004).
- [6] 佐野雅規,住吉英樹,柴田正啓,“映像番スコアブックの相撲への適用”,2001 年映情学年次大 17-8,pp.250-251(2001).
- [7] MasamiTakahata,KensukeShiraki,YutakaSakane,YoichiTakebayashi,“Sound Feedback for Powerful Karate Training”, NIME '04 Proceedings of the 2004 conference on New interfaces for musical expression,(2004).
- [8] ChunMeiLu,NicolaJ.Ferrier,“Repetitive Motion Analysis: Segmentation and Event Classification”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,VOL.26,NO.2,pp.258-263(2004).
- [9] TetsunariInamura,IwakiToshima,YoshihikoNakamura,“Acquiring Motion Elements for Bidirectional Computation of Motion Recognition and Generation”, Springer Tracts in Advanced Robotics,Volume.5,(2003).
- [10] Doo Young Kwon,MarkusGross,“Combining Body Sensors and Visual Sensors for Motion Training”, the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology,(2005).
- [11] TaesooKwon,Young-SangCho, SangIlPark,Sung Yong Shin,“Two-Character Motion Analysis and Synthesis”, IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS,VOL.14,NO.3(2008).
- [12] AtsushiNakazawa,ShinichiroNakaoka,KatsushiIkeuchi,KazuhitoYokoi,“Imitating Human Dance Motions through Motion Structure Analysis”, Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference,vol.3,pp.2539 - 2544 (2002).
- [13] Winfried Ilg, Johannes Mezger,MartinGiese,“Estimation of Skill Levels in Sports based on Hierarchical Spatio-Temporal Correspondences”,LectureNotesinComputer Science,Vol.2781,(2003).
- [14] 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”,人工知能学会誌,15 巻 5 号(2000).
- [15] OkanArikan,David A. Forsyth,James F. O'Brien,“ Motion Synthesis from Annotations”, SIGGRAPH '03 ACM SIGGRAPH 2003 Papers, Volume 22 Issue 3(2003).
- [16] 向井智彦,栗山繁,“テニス動作モーションキャプチャデータベース”,<http://www.val.ics.tut.ac.jp/project/tennis/>.
- [17] Chang,Lin,“Libsvm:Introduction and benchmarks.Tech.rep.”, Department of Computer Science and Information,Engineering, National Taiwan University (2000).