

カメラ付き携帯電話を用いたゴルフスイング自動診断のための 手ブレ補正と人体シルエット抽出

○ 神山泰典*, 松本俊昭**, 井垣輔*, 白井良明**, 植田勝彦***

○ Yasunori KOYAMA*, Toshiaki MATSUMOTO **, Tasuku IGAKI*

Yoshiaki SHIRAI ** and Masahiko UEDA ***

*: 立命館大学大学院理工学研究科, koyama@i.ci.ritsumeit.ac.jp

** : 立命館大学情報理工学部 ***: 住友ゴム工業㈱

屋外環境下においてカメラ付き携帯電話で撮影したゴルフスイングの動画像からスイングを自動で診断するために必要な画像処理技術の開発を目指す。まず、人体シルエット抽出の際に問題となる手ブレの補正の手法を述べ、次に、診断に必要な人体シルエットを動画像から抽出する手法を述べる。

<キーワード>人体シルエット抽出, 手ブレ補正, ヒストグラム, 特徴点追

1. はじめに

本研究は、カメラ付き携帯電話で撮影したゴルフスイングの動画像(以下、スイング動画像)を用いてスイングを自動診断するために必要な人体シルエットの抽出を目的とする。住友ゴム工業㈱が開発した「ダンロップ スイング自動診断システム」¹⁾は、2台の固定 TV カメラで撮影した背景画像とスイング動画像から自動で人体シルエットなどの診断材料を検出し、スイングを診断するシステムである。また、このシステムを屋外環境下で撮影にも適応させるための研究²⁾も行われている。これらのシステムは、時間的・金銭的制約の多いインストラクター直接指導に比べ、短時間・低料金で利用可能であるが、設備が大掛かりであるため、運用環境が限られている。そのため、多くの人が持つカメラ付き携帯電話で撮影したスイング動画像を用いて自動診断できるシステムの開発が望まれている。カメラ付携帯電話で撮影したスイング動画像から診断に必要な人体シルエットを抽出するには2点の問題がある。1つ目は、動画像に手ブレが混じること、2つ目は、背景画像を撮影しないために背景差分法を用いることができないことである。本稿では、これらの問題を解決するために背景領域の特徴点に着目した手ブレ補正法、スイング動画像中の各画素の画素値の分布に着目したシルエット抽出法について述べる。

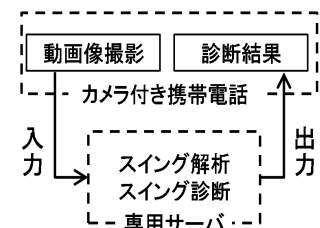
2. スイング自動診断システム

2.1 システムの構成

スイング動画像の撮影時の制約を図1に示す。スイング動画像は、アドレス姿勢のユーザの手元が図1の中央枠内、ボールが右下枠内に収まるようにカメラ付き携帯電話を用いてアドレス姿勢からフィニッシュ姿勢まで撮影する。



(図1: 撮影時の制約)



(図2: システムの構成)

システムの構成を図2に示す。ユーザは、撮影したスイング動画像を診断用サーバに送る。診断用サーバは、スイング動画像からスイングを解析、診断し、診断結果をユーザに送る。

2.2 診断処理の流れ

診断処理の流れについて簡単に説明する。まず、撮影されたスイング動画像から診断に必要な姿勢が映

る6枚の画像(以下, チェック画像(図3))をゴルフクラブや人体の推移を基に抽出する.



(a)アドレス (b)テイクバック (c)トップ



(d)ダウン (e)インパクト (f)フィニッシュ

(図3: チェックポイント画像)

各チェック画像における人体シルエット, 関節位置情報, クラブの傾きなどの診断材料となる情報を抽出する. インストラクターの知見に基づいて構成された診断アルゴリズムを用いて評価し, スイングの点数, 長所・短所, ユーザに見合った練習メニューを提示する.

3. 手ブレ補正

手で持ったカメラ付き携帯電話でスイング動画を撮影した場合, 手ブレが入る. アドレス姿勢とトップ姿勢の差分画像を図4(a), アドレス姿勢とインパクト姿勢の差分画像を図4(b), アドレス姿勢とフィニッシュ姿勢の差分画像を図4(c)に示す.



(a) (b) (c)

(図4: 補正前のアドレス姿勢と他姿勢の差分画像)

図4の背景領域の差分はスイング動画画像に手ブレが入っていることを示唆している. 手ブレは, シルエ

ット抽出の障害となる. そのため, アフィン変換を用いて背景領域の特徴点のズレを正すことで手ブレを補正しシルエット抽出に適したスイング動画画像を作成する. 動画画像はNフレームで構成されているとし以下に手法を示す.

- ① 第1フレームの背景領域から Harris オペレータを用いて t 個の特徴点座標 $p_1^1, p_s^1, \dots, p_t^1$ を検出する. 図5に抽出した特徴点を黒点で示す.



(図5: 背景領域の特徴点)

- ② フレーム f ($2 \leq f \leq N$)でフレーム $f-1$ の特徴点座標 $p_1^{f-1}, p_s^{f-1}, \dots, p_t^{f-1}$ との対応点座標 $p_1^f, p_s^f, \dots, p_t^f$ を検出する. ここで, 手ブレによって生じる背景の移動は微小であり, 局所領域内における背景の移動ベクトルは類似していることから, 対応点の検出には Lucas-Kanade 法を用いる. また, 特徴点座標 $p_1^{f-1}, p_s^{f-1}, \dots, p_t^{f-1}$ と対応点 $p_1^f, p_s^f, \dots, p_t^f$ の近傍でテンプレートマッチングを用いて対応点が正確に追跡できているか確認する.

- ③ 第1フレームの特徴点座標 $p_1^1, p_s^1, \dots, p_t^1$ とアフィン変換後の対応点座標 $p_1^f, p_s^f, \dots, p_t^f$ の差の和 D^f が最小となるようフレーム f ($2 \leq f \leq N$)を変換することで平行移動, スケール変換, 回転によるズレを補正する.

式(1)はアフィン変換後の対応点座標 $p_1^f, p_s^f, \dots, p_t^f$ を示し, 式(2)は第1フレームの特徴点座標 $p_1^1, p_s^1, \dots, p_t^1$ とアフィン変換後の対応点座標 $p_1^f, p_s^f, \dots, p_t^f$ の差の和 D^f を示す.

$$p_s^{f'} = A^f \times p_s^f + B^f$$

$$p = \begin{bmatrix} x \\ y \end{bmatrix}, A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B = \begin{bmatrix} e \\ f \end{bmatrix} \quad \dots (1)$$

$$D^f = \sum_{s=1}^{s=f} |p_s^{f'} - p_s^1| \quad \dots (2)$$

式(1)(2)を用いて、変換パラメータ A^f, B^f を求め、フレーム f の各画素を変換する。

この時、第 1 フレームの特徴点座標 $p_1^1, p_s^1, \dots, p_t^1$ とアフィン変換後の対応点座標 $p_1^f, p_s^f, \dots, p_t^f$ の差が大きいペアは誤対応点とし、誤対応点を省く。

手ブレ補正後のスイング動画のアドレス姿勢とトップ姿勢の差分画像を図 6-(a)、アドレス姿勢とインパクト姿勢の差分画像を図 6-(b)、アドレス姿勢とフィニッシュ姿勢の差分画像を図 6-(c)に示す。



(a) (b) (c)

(図 6 : 補正後のアドレス姿勢と他姿勢の差分画像)

図 6 の背景領域の差分は図 4 と比べて減少していることから手ブレが補正されたことが分かる。

4. 人体シルエット抽出

4.1 ヒストグラムを用いた抽出手法

スイングを診断するためには、人体シルエットが必要である。従来の診断システム^{1) 2)}は、背景画像を撮影していたが、本研究ではユーザの負担軽減のために背景画像を撮影しない。背景差分法を用いることなく人体シルエットを抽出するためにスイング動画画像中で背景のみ映る画素は、画素値の変化が少ないことに着目する。図 8 は、図 7 の各白点座標について動画画像中で変化する輝度値と色相値のヒストグラムである。図 8 が示すように画素の状態により画素値のヒストグラムの分布は異なる特徴がある。



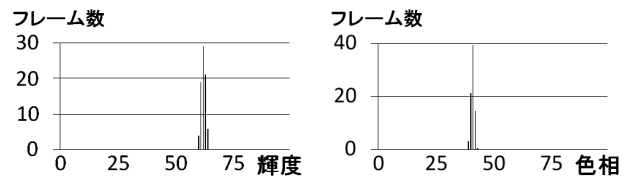
(a)アドレス (b)インパクト (c)フィニッシュ

右白点: 背景のみ映る画素

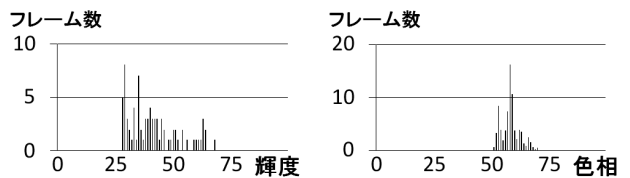
左白点: 人体のみ映る画素

中央白点: 混合して映る画素

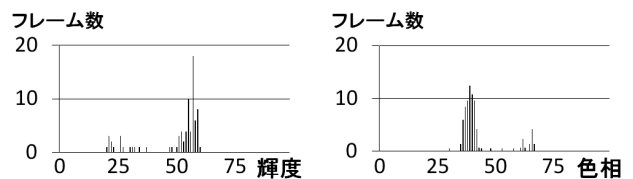
(図 7 : 画素の状態)



(a)背景のみ映る画素 (図 7-右白点)



(b)人体のみ映る画素 (図 7-左白点)



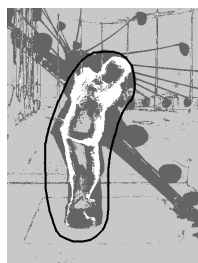
(c)背景と人体が混合して映る画素 (図 7-中央白点)

(図 8: 画素の状態による分布の違い)

背景のみ映る画素(以下、背景画素(図 7-右白点))は単峰な分布のヒストグラム(図 8-(a)), 人体のみ映る画素(以下、人体画素(図 7-左白点))は分散した分布のヒストグラム(図 8-(b))を持ち、背景と人体が混合して映る画素(以下、混合画素(図 7-中央白点))は背景画素と人体画素のヒストグラムの分布が合わさった分布のヒストグラム(図 8-(c))持つ。この特徴を用いて各画素で人体、背景の映るフレームを分離し、人体シルエットを抽出する。以下に、抽出の流れを示す。

- ① 各画素において動画画像中に変化する輝度と色相のヒストグラムを作成する。
- ② 各画素のヒストグラムが単峰なクラスを持つ時、そのクラスを背景クラス(背景が映るフレーム集合)、それ以外を人体クラス(人体が映るフレーム集合)とする。

るフレーム集合)と判別する。単峰なクラスを持たない画素とマスク画素は未判別とする。図9にクラスの判別結果を示す。



薄灰:背景クラスのみ持つ
灰:両方のクラスを持つ
白:単峰なクラスを持たない
黒枠内:マスク(人体が長い間映る予想される画素)
薄灰, 灰:判別済み画素
白, 黒枠内:未判別画素

(図9: クラスの判別結果)

図9に示す未判別画素について距離変換を行う。(図10)



白:判別済み画素
それ以外:未判別画素

判別済み画素との距離によって濃淡値を変化

(図10: 距離変換画像)

- ③ 未判別画素は図10に示す距離方向に下記の判別を繰り返す。

未判別画素のヒストグラムが判別済みの近傍画素のヒストグラムの背景クラスと似た形状のクラスを持つ時, そのクラスを背景クラス, それ以外を人体クラスと判別。

各画素のヒストグラムの判別結果に基づき画素の状態を判別した結果を図11に示す。



薄灰:背景画素
灰:混合画素
白:人体画素

(図11: 画素の状態判別結果)

- ④ 各フレームにおいて, 各画素のヒストグラムの人体クラスに属している画素を人体シルエットとして抽出する。(図12)



(a)アドレス (b)テイクバック (c)トップ



(a)ダウン (b)インパクト (c)フィニッシュ

(図12: 人体シルエット)

図12が示すように各姿勢で足元付近を除き大まかな人体シルエットが抽出できている。

4.2 足元付近のシルエット抽出法

逆光の撮影環境では, 足元と足元付近の背景の輝度値が下がり似た色(図13)となり, 4.1で述べたシルエット抽出法では, 正確な足元シルエットが抽出できない場合がある。(図14)



(a)アドレス (b)インパクト (c)フィニッシュ

(図13: チェックポイント画像(図3 拡大図))



(a)アドレス (b)インパクト (c)フィニッシュ

(図14: 足元シルエット(図12 拡大図))

これを解決するため, 下記の手法を用いて足元シルエットを抽出する。

- ① アドレス姿勢とトップ姿勢の差分画像を用いて求めたアドレス姿勢における手元とクラブヘッドの座標(図 15・白点)を基準に足元を含む初期の探索領域(図 15・枠内)を決める.



(図 15 : 初期の探索領域)

- ② 足元にマットが敷かれている場合, 初期の探索領域にマットの端が含まれないように探索領域を調節する. (図 16・太枠内)



(図 16 : 探索領域と背景サンプル)

- ③ 探索領域内で等間隔に設定した n 個の小領域内(図 16・白枠内)の平均値を背景サンプル値 $S_1, \dots, S_m, \dots, S_n$ とする. 式(3)を用いて, 探索領域内の画素 (x, y) の画素値 $V_{x,y}$ と各背景サンプル値 $S_1, \dots, S_m, \dots, S_n$ との差分値の和 $D(V_{x,y})$ を求め, 足元シルエットを抽出する.

$$D(V_{x,y}) = \sum_{m=1}^{m=n} (|V_{x,y} - S_m| \times W_m) \quad \dots (3)$$

重み W_m は, $D(S_t)$ が $1 \leq t \leq n$ で等しい値になるよう決定する.

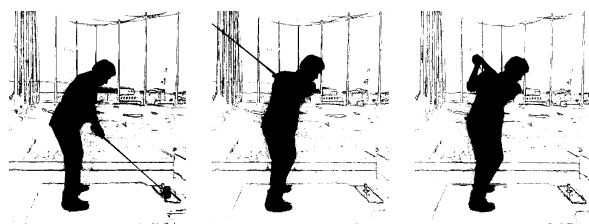
図 17 に足元シルエットの抽出結果を示す.



(a)アドレス (b)インパクト (c)フィニッシュ
(図 17 : 足元シルエット)

図 17 は図 14 と比べて足元のシルエットが正確に抽出されていることが分かる.

4.1 と 4.2 の手法を用いて手ブレ補正前のスイング動画像から抽出した人体シルエットを図 18 に, 手ブレ補正後のスイング動画像から抽出した人体シルエットを図 19 に示す. また, 図 20 に手ブレ補正前と補正後のシルエット誤抽出の比較を示す.



(a)アドレス (b)テイクバック (c)トップ



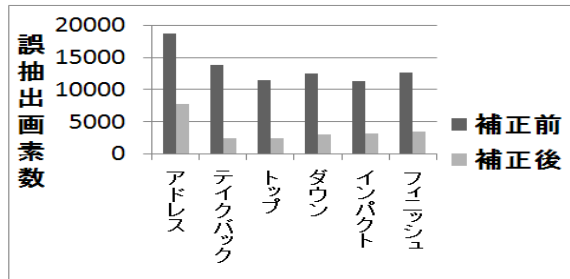
(a)ダウン (b)インパクト (c)フィニッシュ
(図 18 : 人体シルエット(手ブレ補正前))



(a)アドレス (b)テイクバック (c)トップ



(a)ダウン (b)インパクト (c)フィニッシュ
(図 19 : 人体シルエット(手ブレ補正後))



(図 20 : プレ補正前, 後のシルエット誤抽出の比較)

図 20 の示すように図 18 は図 17 の背景領域の誤差に比べ半減している。手ブレ補正をすることにより本稿で述べた人体シルエット抽出法の精度が向上しているといえる。

5. まとめ

ゴルフスイング診断に必要となる人体シルエットを背景画像を用いずスイング動画のみから抽出する手法を述べた。ヒストグラムを用いた人体シルエット抽出法で正確に抽出できなかった足元付近のシルエットについては、足元付近の背景情報との差分を用いることで解決した。本手法は、スイング動画に手ブレがある場合、正確な人体シルエット抽出が行えないが、背景領域の特徴点のズレに着目し手ブレを補正することでこの問題を改善した。

今後の課題として、背景色と似た色の服を着たユーザの人体シルエットを抽出した場合、人体シルエットが欠けてしまう問題がある。解決案として、近傍画素との関係性を考慮できないか検討中である。また、次のステップとして、首、脇、腰などの座標を取得し、それらを端点として人の外輪郭を滑らかに繋げないか考えている。

6. 参考文献

- [1] 植田勝彦, 白井良明, 島田伸敬, 大貫正秀 : “TVカメラからの映像を用いたゴルフスイング自動診断システム”, 動的画像処理実利用化ワークショップ 2006 講演論文集, pp.228-232 (2006)
- [2] 望月智則, 植田勝彦, 白井良明, 島田伸敬 : “屋外環境下におけるゴルフスイング自動診断のための画像特徴抽出”, VIEW2007(ビジョン技術の実用ワークショップ論文集, pp.53-58 (2007)