

論文

動画像からの動作感性情報の抽出

近藤 拓也^{†*} 山際 貴志[†] 山中 光司^{††} 山本 正信[†]

Detection of Skill in Skiing by Motion Analysis

Takuya KONDO^{†*}, Takashi YAMAGIWA[†], Kouji YAMANAKA^{††},
and Masanobu YAMAMOTO[†]

あらまし これまでの動画像の解析では、連続する画像間の対応付けに始まり、もっぱら対象の運動パラメータや3次元情報といった物理的・幾何学的情報の抽出が中心課題であった。これに対して、我々は人間の動作について「上手である」とか「美しい」とかの印象を定量的に把握することを目指している。本論文ではスキーにおける滑りを取り上げ、滑りの動画像を解析することにより、滑りの上手さを判定することを試みた。実際、動画像から得られた動作の運動パラメータから、動作の対称性や滑らかさなど動作の特徴を抽出した。これらの特徴を用いて滑りの上手さを判定することができた。この判定結果は、同じ動作を観察したときの人間の受ける印象と一致している。

キーワード 感性情報, スキー, ロボットモデル, 時空間こう配法, 動画像処理, 人間動作

1. まえがき

コンピュータはときとして人々に「冷たい」といった印象を与えてきた。それは、コンピュータが主に論理だけで組み立てられてきたからである。論理だけではなく、「美しい」とか「心地よい」とかを感じる感性をコンピュータに与えることができれば、人間に近い「暖かみ」のあるコンピュータを期待することができる。そのためには、感性の仕組みを明らかにしていく必要がある[1]。

人間が視覚を通して得る情報のうち、形や色はさまざまな印象を引き起こすことが知られている。形や色の感性情報については既に多くの研究が行われている[2]~[4],[7]。形や色以外に、動きについても人間はさまざまな印象を得ることができる。実際、「美しい立ち居振舞い」とか「きびきびした動作」などの言い表し方があるように、我々は人間の動作にさまざまな印象を抱いている。また、新体操やフィギュアスケートのような一部のスポーツでは、演技に対してその印象を点数化し競い合っている。本研究は人間の動作に

ついて、「上手である」とか「美しい」とかの印象を定量的に把握することを目指している。特に、本論文ではスキーにおける滑りを取り上げ、滑りの動画像を解析することにより、滑りの上手さを判定することを試みた。

これまでの動画像の解析では、連続する画像間の対応付けに始まり、もっぱら対象の運動パラメータや3次元情報といった物理的・幾何学的情報の抽出が中心課題であった。最近では、動作の種類を識別することが行われているが[6],[8]、本論文では更に、動作の上手・下手といった感性情報までを抽出しようとするものである。

2. 動作感性情報

スキーの滑りをビデオ映像で観察したとき、スキーについて多少の知識があれば、上手な滑り下手な滑りを見分けることができる。この上手・下手の判定は人間の感性によるもので、判定結果である上手な滑りや下手な滑りを動作の感性情報と呼ぶ。

一方、滑りを動画像から測定したとき、スキーヤーの動作を速度や角速度のような物理情報で表すことができる。更に、これらの物理量から動作についてある評価値が計算できているとする。2人のスキーヤーの動作について、両者の評価値を比較したとき、大きな評価値に対応する動作を上手な滑り、小さな評価値に

[†] 新潟大学工学部情報工学科, 新潟市
Niigata University, 8050, Ikarashi2-nochou, Niigata-shi, 950-21
Japan

^{††} テクノア, 岐阜市
Technoa, Gifu-shi, 500 Japan

* 現在, 富士ソフト株式会社

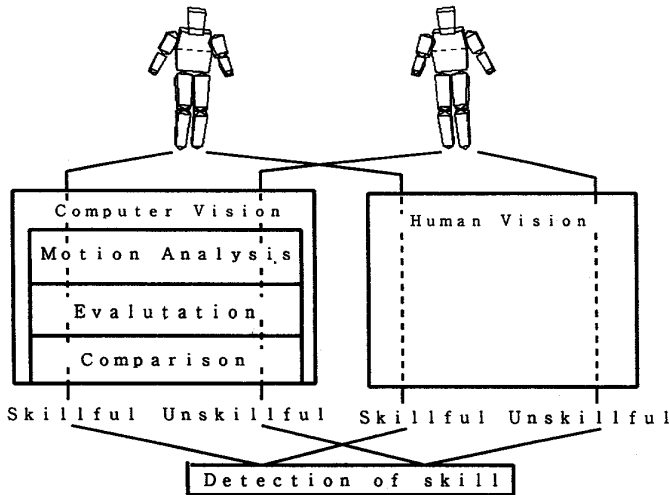


図1 感性情報のコンピュータによる抽出。
Fig.1 Detection of skill.

対応する動作を下手な滑りと定義する。この物理情報より得られた上手・下手の判定が、人間の感性から得られた判定と一致したとき、感性情報がコンピュータにより抽出できたことになる。図1はこの過程を示したものである。

次章から6.まで、スキー動作の動画像解析について述べる。7.で動作の評価関数の構成法を示し、8.で実際の動作について、コンピュータによる判定結果と人間による判定結果との比較を行う。

3. 動作解析の概要

動作の解析は図2に示すような手順により行った。まず、人間の体を多関節ロボットでモデル化する。このモデルは、CADモデリングシステムによりコンピュータ上に実寸で作成する。追跡開始フレームにおいて、このロボットモデルを画像中の実際の人間に手動でフィッティングする。そして時空間こう配法を用いて人間の動きをロボットモデルの運動パラメータとして求める。求めたパラメータによりモデルを移動させて画像を1フレームずつ進めていくことにより、逐次画像中の人間を追跡する。この操作を最終フレームまで繰り返すことにより人間の動きが運動パラメータの時系列として定量化される[9]。なおノイズ処理のため、得られたパラメータ列にカルマンフィルタをかけて平滑化を行った。

4. 時空間こう配法による運動パラメータの推定

3次元シーンに直交座標系 (x, y, z) を設定する。

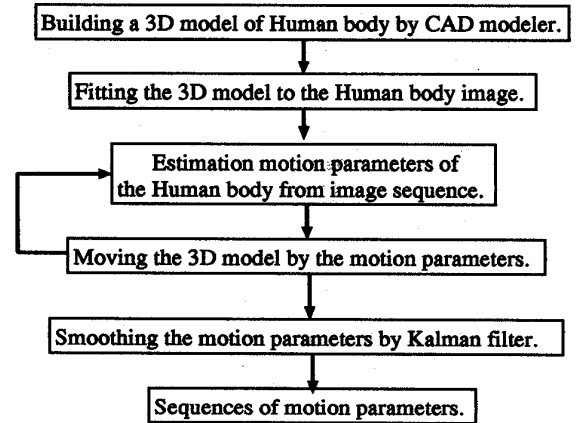


図2 解析手順
Fig.2 Steps for processing of human motion.

シーン中の点は座標系の原点を視点として、 z 軸に垂直な平面上に中心投影されるとする。投影面座標 (X, Y) は、 z 軸と投影面の交点を原点に、 X, Y 軸をそれぞれ x, y 軸と平行になるように定めておく。人体上の点 (x, y, z) は、画像上の点 (X, Y) へ透視変換式 $X = x/z, Y = y/z$ で中心投影される。画面上の速度ベクトル (V_X, V_Y) と人体上の3次元速度ベクトル (v_x, v_y, v_z) は

$$\begin{aligned} V_X &= \frac{1}{z}(v_x - v_z X) \\ V_Y &= \frac{1}{z}(v_y - v_z Y) \end{aligned} \quad (1)$$

で関係づけられる。また画面上の速度ベクトルは、次の時空間こう配法の方程式に拘束されることが知られている。

$$E_X V_X + E_Y V_Y + E_t = 0 \quad (2)$$

ここで (E_X, E_Y) は明るさの空間こう配で、 E_t は時間こう配である。この式に式(1)を代入して整理すると、3次元速度ベクトルの拘束式が得られる。

$$\frac{E_X}{z} v_x + \frac{E_Y}{z} v_y - \frac{1}{z}(X E_X + Y E_Y) v_z + E_t = 0 \quad (3)$$

一方、人体を多関節ロボットアームでモデル化したとき、人体上の位置 $\mathbf{p} = (x, y, z)^T$ はアームの関節角などをパラメータ $\mathbf{q} = (q_1, q_2, \dots, q_n)^T$ として次の関係式で表される[5]。

$$\mathbf{p} = \mathbf{f}(\mathbf{q}) \quad (4)$$

\mathbf{f} は一般に非線型であるが \mathbf{q} の時間変化率 $\dot{\mathbf{q}}$ に

対する速度ベクトル $\dot{p} = (v_x, v_y, v_z)^T$ はヤコビ行列 $J = \partial f / \partial q$ を使い、次の線形式で関係づけられる。

$$\dot{p} = J\dot{q} \tag{5}$$

\dot{q} は人間の動作を表す運動パラメータとみなせる。この式を式 (3) に代入すれば運動パラメータ \dot{q} の線形推定式が得られる。ロボットモデル上の点で得られた推定式を連立させて最小 2 乗法により運動パラメータの最適解を推定する。

5. CAD モデルのパラメータ設定

人体を多関節ロボットアームでモデル化したとき、その形状は図 3 に示すように、CAD モデラを用いて作られる。これは人体を構成しているパーツを胴体と頭、左右上腕、下腕、上脚、下脚の合計 9 種類と考え、これらが関節リンクで結ばれている。このモデルの運動パラメータ \dot{q} は、この各関節リンクの角速度およびモデルの親物体である胴体の並進速度と回転角速度である。今回は、表 1 の○印で示された合計 22 個を運動パラメータとして設定した。

追跡開始フレームでは、モデルを投影面上に投影し、その投影像が人体像と重なるようにモデルの位置・姿勢を手動で調整する [9]。モデルは 3 次元であるので、

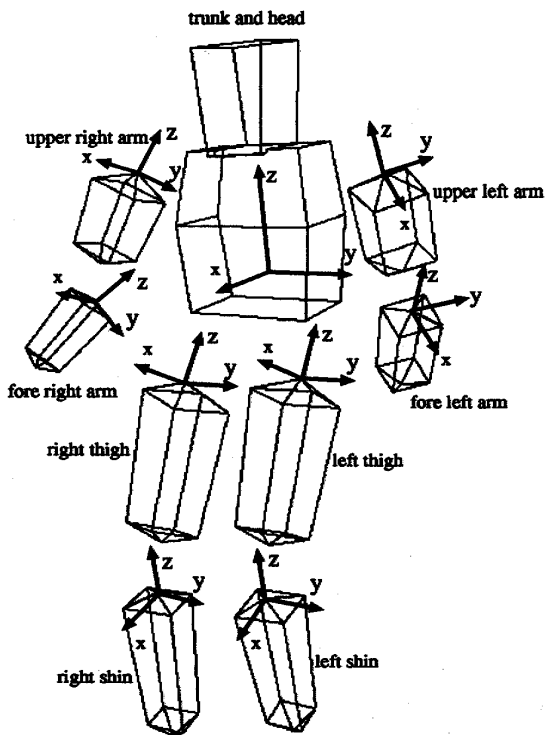


図 3 人体ロボットモデル
Fig.3 CAD modeling of parts of human body.

この手動調整は 3 次元的なフィッティングである。スキーヤーの正面から撮られた画像では、膝や腰の曲げ角度を知ることは難しい。そこで、本論文では画像内の演技者と同じ姿勢を試み 3 次元姿勢を推測することにした。この方法は正確ではないが、正しい値から大きくずれることはない。後節で動作の評価を行う際に、位置・姿勢の値そのものを用いることはないので、正確な位置・姿勢を求めることは必ずしも重要ではない。

6. スキー動作の解析

スキーの技術レベルの判定を行うために、インラインスケートを用いて動作を擬似的に再現した。これは、より等しい環境下でのスキーの動作を比較するためである。図 4 で示したようにカメラの光軸と平行な直線上に等間隔でマーカ（障害物）を置き、その間を滑走するシーンを撮影した。実演者には、(1) 下半身や胴体の動きの違いが顕著に現れるように、腕をなるべく動かさない、(2) 自然な状態での滑りを得るため、

表 1 解析に使用した運動パラメータ
Table 1 Motion parameters of human body.

パーツ	回転成分			並進成分		
	x	y	z	x	y	z
胴体	○	○	○	○	○	○
右上腕	○	○	—	—	—	—
右下腕	○	○	—	—	—	—
左上腕	○	○	—	—	—	—
左下腕	○	○	—	—	—	—
右上脚	○	○	○	—	—	—
右下脚	—	○	—	—	—	—
左上脚	○	○	○	—	—	—
左下脚	—	○	—	—	—	—

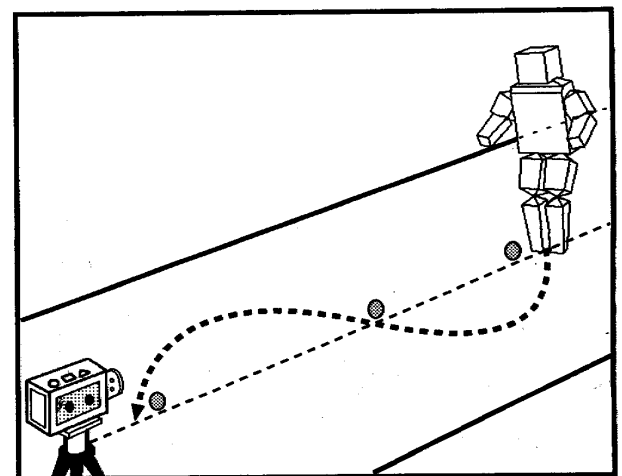


図 4 滑走するシーンの人の動きとカメラの配置
Fig.4 A view of the scene containing camera and human movement.

各自の好みの速度で滑走する、等の指示をした。

滑走を行ったのは、競技スキー部に属し経験年数の差から技量においてかなりの開きのある4人である。これら4人の滑りについて、観察者（複数）は自らの感性に基づき上手・下手の判定を下した。その結果、最も上手な者を上級者Aとし、以下上手な順に、上級者B、中級者、初級者とした。この序列を上手な順に右から並べると、次のようになる。

初級者、中級者、上級者B、上級者A (6)

上級者2人については技術レベルはほぼ同じであるが、上級者Aが上級者Bよりわずかに上手である。この序列は、観察者の主観に基づくものであるが、スキーについて多少の知識があれば、容易に判定が下せる。また、観察者間において共通性や観察者自身において

再現性のあるものであった。

滑走の追跡フレームは、80~100フレームで、ほぼ2周期弱のスラロームに相当する。図5から図8はそれぞれ初級者、中級者、上級者B、上級者Aの追跡結果を20フレームごとに抜き出して示したものである。人体のロボットモデルが、画像上の人間に白線で重ね合わされている。

7. スキー動作の判定基準

スキーの滑りが上手あるいは下手という印象は、観察者のあいまいな感性によって下されている。しかしこの印象も、例えば「両足が同じ動きをしている」とか「上半身が安定している」といったような基準（指標）に照らし合わせて上手か下手かを判定し、それら

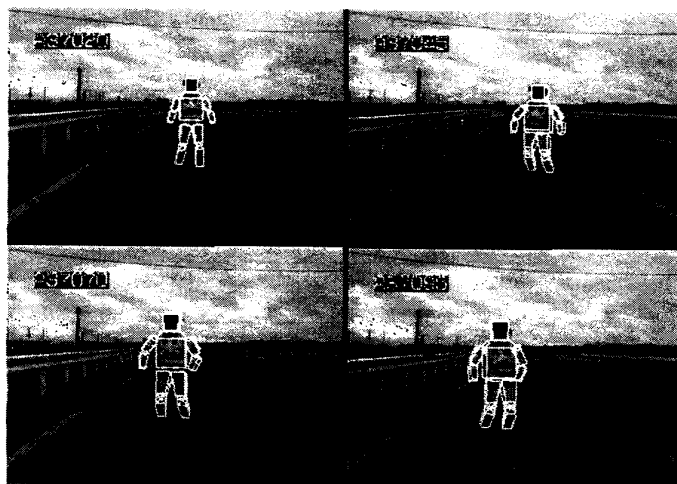


図5 解析結果：初級者
Fig.5 Tracking result: Beginner.

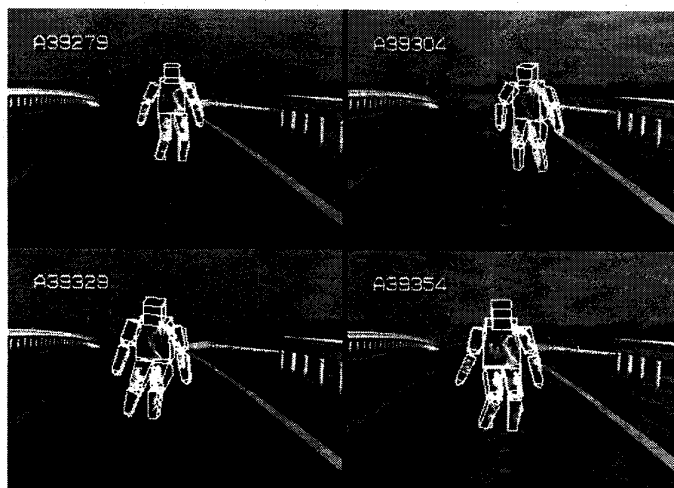


図7 解析結果：上級者B
Fig.7 Tracking result: Expert-B.

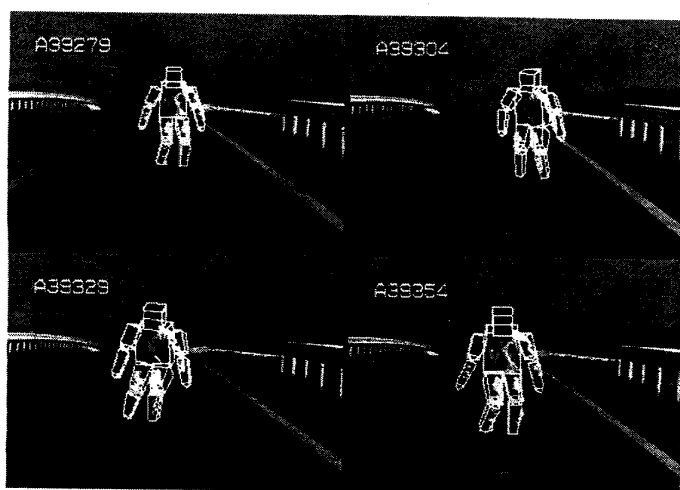


図6 解析結果：中級者
Fig.6 Tracking result: Intermediate.

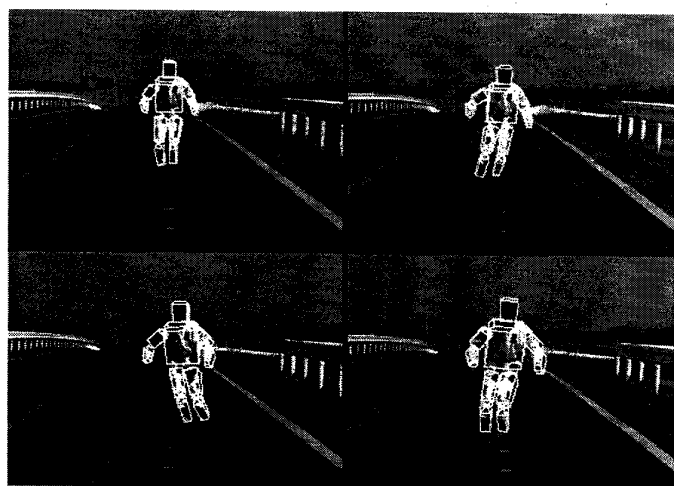


図8 解析結果：上級者A
Fig.8 Tracking result: Expert-A.

論文/動画像からの動作感性情報の抽出

表2 観察者の技術レベル別による判断基準 (新潟大学競技スキー部による)

Table 2 Bases for skill judgement in skiing.

判断 (指標)	観察者のレベル		
	初級	中級	上級
両足が同じ動き	○	○	○
上半身が安定		○	○
体軸の位置がセンター		○	○
スタンスが一定			○
重心移動がスムーズ			○
動作が途切れず, 滑らか			○
左右対称			○

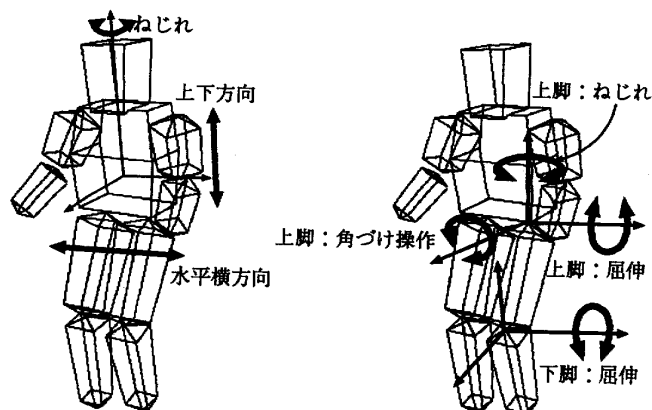


図9 検討に使用する運動

Fig.9 The motion parameters for skill judgement.

の判定についての情報を統合して総合的な判定を下していると思われる。これらの基準は、観察者のもっている専門的知識に依存する。実際、表2はスキーにおける滑りの上手さを判定するとき、観察者の技術レベルによって判定する基準を示したものである。○印が判定者により提示された基準である。初級の観察者は両足の同期のみで判定しているが、より上級者になり専門的知識が多くなっていくほど基準がより詳しく、多くなっている。このことから、人間のあいまいな感性によって下されているはずの判定も、専門的知識の量に比例して正確に、詳しく具体的に記述することができる。従って、これらの具体的な基準(指標)を運動パラメータといった物理情報で計算することができれば、感性をコンピュータにもたせることが可能となる。

8. 動作の評価

解析処理によって人間の動作が運動パラメータの時系列として抽出することができた。このパラメータ列を用いて動作の評価を行う。表2の判断基準のうち、「両足が同じ動き」および「動作が途切れず, 滑らか」という指標について評価を行い、その結果を統合して最終的な判定を下す。これらの評価を行うために、解析した運動パラメータの中から、図9に示したような動きに着目した。

8.1 両足のそろい具合

両足が同じ動きをしているかどうかを判定するために、図9の右図に示されるように、両上脚の足の付け根を中心とし、左右の方向に回転する運動(角付け操作)、上下方向の運動(屈伸運動)、ねじれ運動、そして下脚の屈伸運動について速度の時間変化を調べる。図10から図13までは、それぞれ初級者、中級者、上級者B、上級者Aの左右上脚の角付け操作における速度の時間変化を示したものである。実線が右足、点線が左足に対応している。この基準において重要なのは

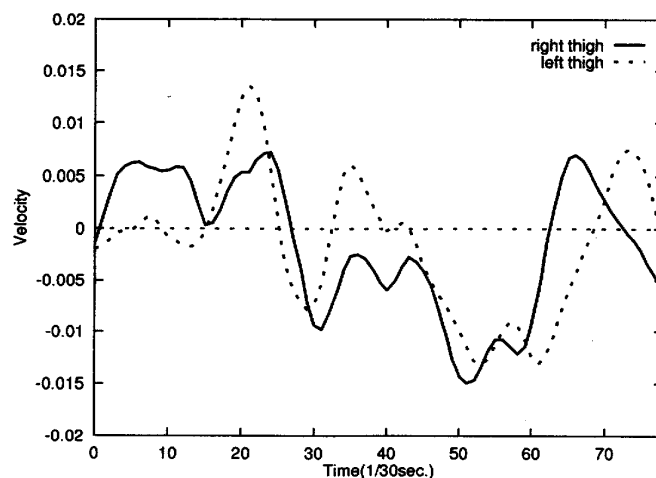


図10 角付け操作時の足の速度変化: 初級者
Fig.10 Edge handling: Beginner.

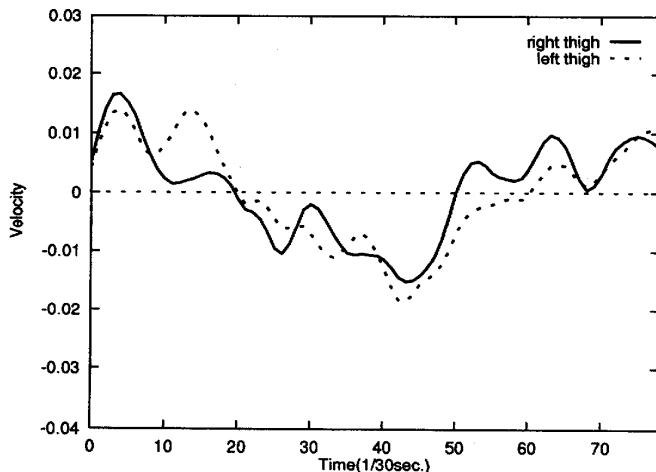


図11 角付け操作時の足の速度変化: 中級者
Fig.11 Edge handling: Intermediate.

変化の大きさではなく、左右の足が同じような動きの中で、スムーズな動作が行われているかという点である。この運動を評価するために、両足の動作が周期的

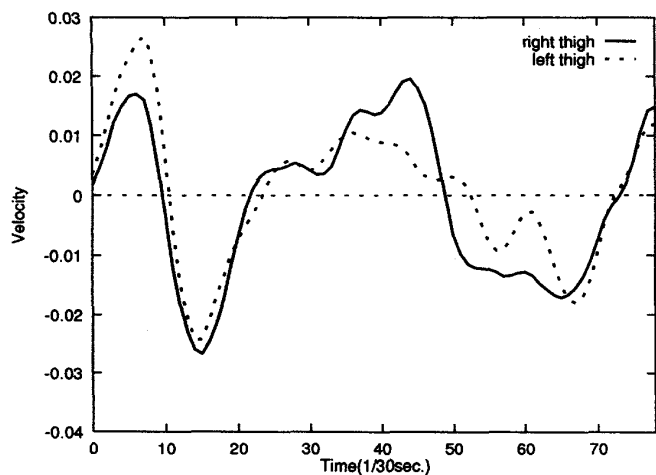


図 12 角付け操作時の足の速度変化：上級者 B
Fig.12 Edge handling: Expert-B.

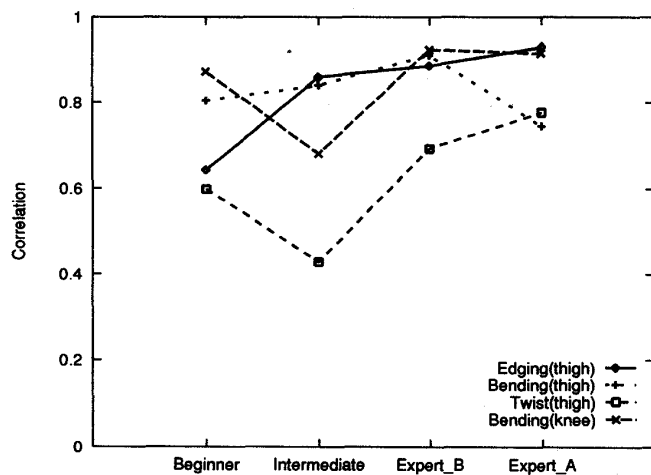


図 14 各種運動における両足のそろい具合の類似度
Fig.14 Similarity of movements of both legs.

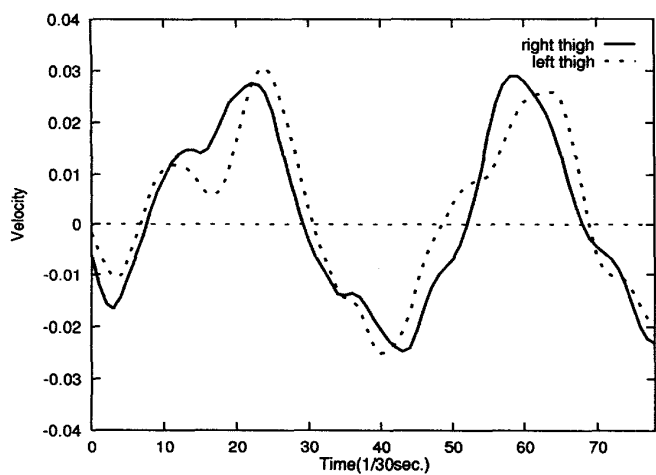


図 13 角付け操作時の足の速度変化：上級者 A
Fig.13 Edge handling: Expert-A.

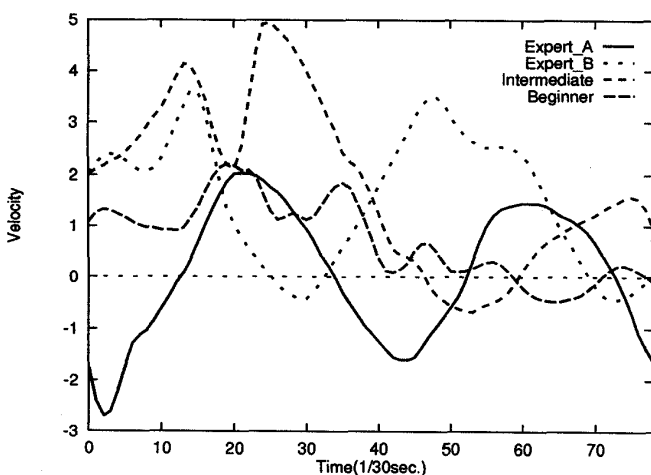


図 15 胴体の水平横方向の運動における速度の時間変化
Fig.15 Velocity of trunk motion.

に見てどれだけ結び付きをもっているか調べる。つまり、4人それぞれの左右の足について速度変化の類似度は、相関係数を使って式(7)により求めた。ここで、 r_i, l_i は i フレーム目での左右の足の角速度で、 \bar{r}, \bar{l} は角速度の平均値である。

$$\text{類似度} = \frac{\sum_{i=1}^N (r_i - \bar{r})(l_i - \bar{l})}{\sqrt{\sum_{i=1}^N (r_i - \bar{r})^2 \cdot \sum_{i=1}^N (l_i - \bar{l})^2}} \quad (7)$$

図 9 の右図に示される 4 種類の運動について、類似度を計算した結果が図 14 に描かれている。どの運動に関しても若干の違いがあるが、おおむね上級者 A、上級者 B の 2 人が中級者、初級者と比べて高い値を示している。これらの値を 4 人それぞれ合計して、平均を

計算した結果を図 18 に + 印で示す。

8.2 動作の滑らかさ

次に、滑らかな動作を評価する。これは、類似度を求めた際に使用した左右上脚の角付け操作、屈伸運動、ねじれ運動、下脚の屈伸運動に加えて、図 9 の左図に示される胴体の水平横方向の運動、上下方向の運動、ねじれ運動の合計 11 個の運動についての速度の時間変化を調べた。図 15 は 4 人の胴体の水平横方向運動についての速度の時間変化を示したものである。この中からスラローム 1 周期分の時系列データ $x_j (j = 0, \dots, n-1)$ を取り出し、それを式 (8)~(10) を用いてフーリエ変換し、振幅スペクトル S_k を求める。

$$a_k = \frac{2}{n} \sum_{j=0}^{n-1} x_j \cos \frac{2\pi k j}{n}, \quad \left(k = 0, \dots, \frac{n}{2}\right) \quad (8)$$

$$b_k = \frac{2}{n} \sum_{j=0}^{n-1} x_j \sin \frac{2\pi k j}{n}, \quad \left(k = 1, \dots, \frac{n}{2} - 1\right) \quad (9)$$

$$S_k = \frac{\sqrt{a_k^2 + b_k^2}}{2} \quad (10)$$

滑らかな動作は、高調波成分を含まないスラロームである。従って振幅スペクトルをある一定のしきい値で低周波成分と高調波成分に分け、低周波成分を主成分とし、主成分の割合が大きいほど動きが滑らかであると評価する。本論文ではしきい値を $k=3$ とした。主成分比は、次式で計算される。

$$\text{主成分比} = (S_1 + S_2 + S_3) / \sum_{k=1}^n S_k \quad (11)$$

図 16 は、4人それぞれの胴体の各運動について振幅スペクトルの主成分比を計算した結果である。また、左右脚部の各運動についても同様に求め、図 17 に示した。これらの運動に関して言えることは、相関値を求めたときと同様に一部の運動を除いて、上級者2人が中級者と初級者よりも高い値を示している。4人それぞれ主成分の値を足し合わせて計 11 個の評価値の平均を計算した結果を図 18 に □印で示す。

8.3 判定

前節で取り上げられた二つの指標を統合して、滑りの上手・下手の総合評価を行う。これには、左右の足のそろい具合を調べた際の結果である 4 個の類似度 i ($i = 1, \dots, 4$) と、動作の滑らかさを調べた際の 11 個の主成分比 i ($i = 1, \dots, 11$) について、それぞれ

の重み付き総和を実演者ごとの総合評価値とした。式 (12) は総合評価の計算式である。

総合評価値

$$= \sum_{i=1}^4 \alpha_i \times \text{類似度}_i + \sum_{i=1}^{11} \beta_i \times \text{主成分比}_i \quad (12)$$

ここで、 α_i, β_i はそれぞれ 類似度 $_i$, 主成分比 $_i$ にかかる重みである。

$\alpha_i = 1/4, \beta_i = 0$ の場合は「両足が同じ動き」という指標のみの評価となる。各演技者に対して、式 (12) で計算された総合評価値は図 18 に + 印で示されている。このときの評価値の大きさに基づく上手・下手の判定は、右から上手な順に次のとおりである。

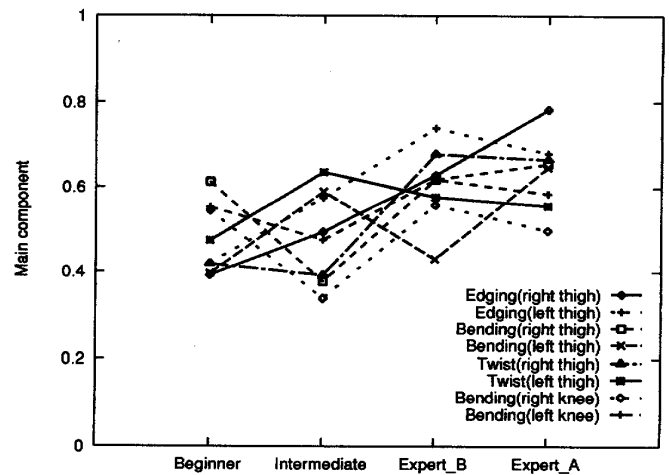


図 17 脚部の運動における速度変化の振幅スペクトルの主成分

Fig. 17 The ratio of main component in power spectrum.

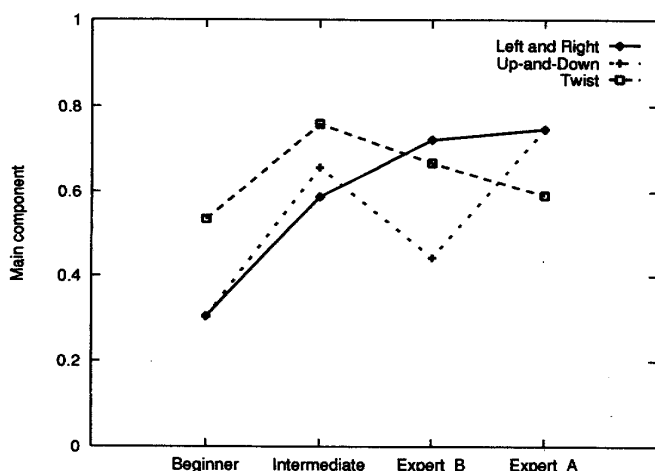


図 16 胴体の運動における速度変化の振幅スペクトルの主成分

Fig. 16 The ratio of main component in power spectrum.

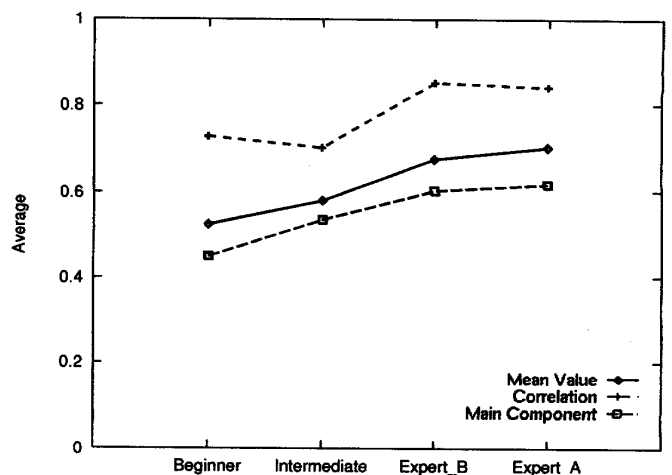


図 18 実演者ごとの動作の評価値

Fig. 18 Evaluated values of skill in skiing.

中級者, 初級者, 上級者 A, 上級者 B (13)

上位グループ (上級者 A, 上級者 B) と下位グループ (初級者, 中級者) との序列は, 感性による序列 (6) と一致している。しかし, 各グループ内での序列は感性による判定 (6) と異なっている。「両足が同じ動き」という指標のみでの判定は, 初級レベルの観察者に相当する (表 2)。初級レベルの観察者は, 上級者 A と B や中級者と初級者の区別がつかない可能性を示している。

$\alpha_i = 0, \beta_i = 1/11$ の場合は, 「動作が途切れず, 滑らか」という指標のみの評価となる。式 (12) で計算された演技者ごとの総合評価値が図 18 に □ 印で示されている。また, 「両足が同じ動き」と「動作が途切れず, 滑らか」という二つの指標を平等に評価する場合は, $\alpha_i = \beta_i = 1/15$ とする。このとき演技者ごとの総合評価値を図 18 に ● で示す。以上の二つの場合に, 総合評価により下された上手・下手の判定は,

初級者, 中級者, 上級者 B, 上級者 A (14)

となる。これは, 人間の感性により下された判定結果 (6) と一致する。

9. むすび

スキーの滑走を観察したとき, スキーについての知識が多少ともあれば, 滑走の上手・下手を判定することができる。この判定は人間の感性に基づくものであるが, 観察者はさまざまな判定基準に照らし合わせて総合判定を行っている。これらの基準の中のいくつかを, 動画像から測定した運動パラメータを使って評価した。評価値の高い順に上手であるとの判定を下したところ, 人間の感性に基づく判定結果とほぼ一致した。

今回は判定基準として, 「両足のそろい具合」と「動作の滑らかさ」という二つの指標を用いたが, この他にも表 2 に示したような指標を用いて判定を行えば, より正確な結果が得られるものと思われる。

総合評価では, それぞれの指標を構成する類似度と主成分比の単純平均を求めるだけに留まったが, 個々の評価値の重みを変えることにより, さまざまな視点における判定を下すことが可能である。例えば, 実際の観察者に重み付けを行わせることにより, 観察者の感性により近い判定結果を得ることも可能であろう。

謝辞 動画像データの撮影および表 2 の作成に御協力頂きました新潟大学学友会競技スキー部の皆さん

に感謝致します。本研究は一部, 科研費重点領域研究「感性情報処理」(No.06212208) の支援を受けた。

文 献

- [1] 井口征士, 猪田克美, 小林重順, 田辺新一, 長田典子, 中村敏枝, “感性情報処理,” オーム社, 1994.
- [2] 木村一郎, 武井 努, 黒江康明, “ニューラルネットによる色感性の実現,” 計測自動制御学会論文集, vol.32, no.2, pp.224-230, 1996.
- [3] 松山隆司, 山口 修, “代数的制約記述に基づく配色デザインシステム,” 情処学論, vol.34, no.5, pp.941-953, 1993.
- [4] 長田典子, 亀井光仁, 赤根正樹, 中嶋紘之, “感性計測技術に基づく真珠品質評価システムの開発,” 電学論 C, vol.112, no.2, pp.111-116, 1992.
- [5] R.P. Paul, “Robot manipulators,” Mathematics, Programming, and Control, The MIT press, 1981.
- [6] P. Polana and R.C. Nelson, “Detecting activities,” Journal of Visual Communication and Image Representation, vol.5, No2, pp.172-180, 1994.
- [7] 佐藤宏介, 井口征士, “イメージの画像解析に基づく絵画のデータベース化と自動作画,” 人工知能学会研究会, SIG-HICG-9303-1, 1994.
- [8] 高橋勝彦, 関 進, 小島 浩, 岡 隆一, “ジェスチャー動画像のスポッティング認識,” 信学論 (D-II), vol.J77-D-II, no.8, pp.1552-1561, 1994.
- [9] 山本正信, 川田 聡, 近藤拓也, 越川和忠, “ロボットモデルに基づく人間動作の 3 次元動画像追跡,” 信学論 (D-II), vol.J79-D-II, no.1, pp.71-83, 1996.

(平成 8 年 3 月 22 日受付, 6 月 19 日再受付)



近藤 拓也

平 6 新潟大・工・情報工学卒。平 8 同大学院情報工学専攻修士課程了。同年, 富士ソフト株式会社入社。在学中は動画像処理の研究に従事。



山際 貴志

平 7 新潟大・工・情報工学卒。現在, 同大学院情報計算機工学専攻博士前期課程在学中。感性情報処理に関する研究に従事。



山中 光司

平7新潟大・工・情報工学卒。同年テクノアに入社。在学中は競技スキー部において活躍。



山本 正信 (正員)

昭48九工大・工・制御卒。昭50東工大大学院修士課程了。同年、電総研入所。動画像処理、コンピュータビジョン等の研究に従事。平1~2カナダ国立研究協議会招聘研究員。昭62情処学会研究賞受賞。平4より新潟大工学部情報工学科教授。工博。情処理学会、IEEE各会員。