

指導者が注目する身体部位のトラッキングに基づく社交ダンス動作 指標抽出システム

The system to extract subjective criteria for ballroom dance through tracking important body parts on which a coach focuses

伊豆蔵 拓也 *¹ 大本 義正 *¹ 西田 豊明 *¹
Takuya Izukura Yoshimasa Ohmoto Toyoaki Nishida

*¹京都大学 情報学研究科
Graduate School of Informatics, Kyoto University

Most of arts are evaluated by the subjective criteria which are defined by the experts. When beginners learn an art, they have to understand what is important for the art. In this study, we focused on teaching a ballroom dance as an example and propose a method to extract the subjective and implicit evaluation criteria. Firstly, we conducted an experiment to capture real teaching scenes. After that, we analyzed and acquired the evaluation criteria which included important features for evaluation of dancing motions. We evaluated the criteria by using discriminant analysis. As a result, we confirmed that the method could accurately evaluate about 70% of beginner's dance steps. Furthermore, we also confirmed that one of the extracted criteria could accurately evaluate different ballroom dance steps.

1. はじめに

伝統工芸や芸術的な動作を初学者が学習する際、複雑な運動や行為を自力で学習することは難しく、教示者に頼って学習を行うシーンが存在する。これは、芸術的動作の技能の向上を図る際には、人間が定義した芸術的動作の指標が技能評価に影響を与えるものの、初学者のように分野に馴染みの薄い人からすると、動作の何が重要であるか、その指標を獲得することから学習を始める必要があるためである。

運動系動作においては、自身が身体を動かすときに感じる動作の不感状態（上手く動けていないと認識している状態）に対して、教示者による指導を受けることで、効率的に課題を認識して動作技能を向上させることができるといわれている[永山 10]。しかし、フィギュアスケートや社交ダンスといった芸術的動作の学習では、このように動作の不感状態を認識するためには、力学的な指標に加えて、人間が定義した芸術的動作の評価指標を自分が持っている必要がある。

例えば、社交ダンスは音楽に合わせて芸術的な身体表現を行う動作の集合と捉えることができ、様々な要素が混在しているが漠然と美しさは理解できる。しかし、社交ダンスを競技レベルに発展させた競技ダンスに対する採点項目は抽象的な記述にとどまっており、具体的にどう踊ることが正解なのかは定義されておらず、各々の判断に任せられている。従って初学者が名状しがたい具体的な評価指標を初学者が獲得しようとする際には、熟練者から口伝されたり表現されたものを「美しいものである」と自身の中で理解するしかないが、具体的にどうすれば良いのか、どう踊ることが正解かが始めは分からない。

本研究では、芸術的な運動の一例として社交ダンスを対象として、教師の指導によって評価指標と動作を同時に学習することに着目し、動作指導を通して評価指標を抽出する手法を開発することを目的とする。ダンスの評価指標は、ダンスの動作の修得過程時に他人が踊るさまを実際に身体の動きをもって観察し、自分の中で動きのコツを発見し、その踊りに解釈を付けていく過程を通して、ダンスに対する自分の動きの動感、つまり評価指標を独自に形成して自分固有のダンスを確立していく

からであるとされており[宮川 05]、本研究の目的に合致する。

本研究の具体的な目標は、社交ダンスの動作の様々な評価指標を抽出して、蓄積することができるシステムを作成する事である。今回は、「教示者毎に評価指標の内容は異なること」を前提として、特定の教示者の評価指標を抽出する。そして学習者の社交ダンス動作を、獲得した評価指標に基づいてシステムが自動的に評価することで、対象となる教示者が持つ評価指標に照らしあわせて学習者の動作を評価する事を目指す。また、本研究で提示した手法が将来的にどれくらいの範囲の運動に一般化することができるのかを、評価実験の結果やシステムの定性的な特徴の観点から実験的に考察する。

2. 先行研究

社交ダンスに限らず、運動系動作の模倣や学習を補助するための様々な研究が行われている。

山根ら[山根 05]は舞踊動作データの分析を行うために、同一舞踊に対する動作データの差異を検出し、姿勢の類似度を測定し、姿勢を定量的に比較する手法を提案して、踊りの「振り」の差異を検出している。

Chanら[Chan 11]は、モーションキャプチャを利用したダンスの指導システムを構築している。学習者の動作をモーションキャプチャによって捉え、事前に見せたお手本となるモデルの動作との違いを、身体の部位をハイライト表示させることで提示している。加えて教示動作と同一の踊りを行った箇所については動きの差分を点数化して、差異が大きい箇所を学習者の修正すべきポイントとして提示している。

Hung[Huang 12]らは社交ダンスの教示を目的としたシステムを提案している。この研究では社交ダンスの重要な観点としてカウントと呼ばれる音楽の拍に着目している。動作データを個々のカウント間の遷移毎に切り出して、学習者とお手本の動作データの遷移ベクトルの差異が大きい部分を教示する。

従来研究の問題点として、学習者とお手本データの全体を見て、その中でも動作差異が大きい箇所だけに着目して学習者に模倣学習を行わせる点に焦点をあてており、実際に運動系動作の評価を向上させるものなのかを明確に示していないことが挙げられる。しかし、現実の教示場面を観察すると、ある身

体動作の差異が他の身体動作の差異よりも小さくとも、教示者は採点項目に対する技能の向上を果たすために、各々の評価指標を持って教示を行うことがよくある。このような場面では、従来研究のように差分が大きい点のみ注目すると指標としての有用性を評価することができない。

本研究では動作の模倣のために画一的に全身を見て差異を検出することではなく、動作の評価向上のために必要な評価指標を、教示者の視点を通して、動作時系列の中から抽出することで、システムが社交ダンスの採点項目に対して評価指標を獲得することを目標とする。

3. 社交ダンス動作評価システムの概要

本研究では、社交ダンス動作に対する教示内容に表れている評価指標を抽出するために、教示シーンの教示者が注目する身体部位情報と、教示者動作データに着目する。具体的には教示時の言語情報や身体部位ポインティング動作から教示者が注目する身体部位を推定し、その重要身体部位に対する学習者の指導前、指導後、そして教示者の教示時動作データのパラメータを比較することで、学習者が改善できた際に教示者が重要であると判断した身体部位の重要パラメータを、教示内容の評価指標を提示する「評価指標パラメータ」として抽出し社交ダンス動作の判別に用いる。つまり重要身体部位に関して、学習後データを学習前データと比較し、教示者の動作データとの差分が小さくなっているパラメータを抽出し、そして動作の良し悪しを判断する判別式を導出する。また、教示者によって認められている指導後の動作は評価指標の観点から良い動作データであり、逆に教示者に良くないと言われている指導前の動作は悪い動作データであると仮定できる。従って、各パラメータ毎に正解データである教示者データとの類似度を計算し、類似度が指導前後で変化しているパラメータを評価指標パラメータとして抽出する。

上記を実装したシステムのフローを図1に示す。下記ではシステムのフローを追いながら各処理の概要を説明する。

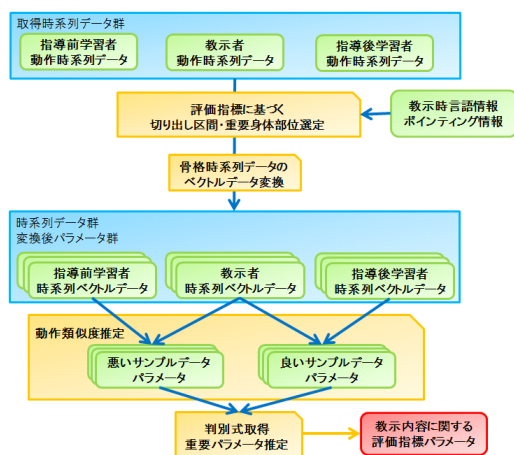


図1: 提案システムでのデータフロー

始めにシステムに入力するデータを、非接触で取得する。そのために、Microsoft Kinectを複数台設置して、人間の動作を記録、及び読み込みを行う、矢野[矢野12]のシステムを拡張して使用した。取得するデータは社交ダンス動作データである「動作時系列データ」と、評価指標を表すパラメータである「重要身体部位情報」とその「タイミング情報」である。

次に、それぞれの重要身体部位・タイミングに基づいて動作時系列データから適切な部分を抽出し、その時系列データの特徴を多面的に捉えるために、ベクトルデータに変換する。これらの「重要身体部位」「タイミング」を表すデータは教示内容に従って手動で取得する。

そして、動作時系列データの中でも、教示内容に関係ある身体部位に関するパラメータについて、(1)提示された「重要身体部位」のフレーム間の時系列遷移ベクトル、(2)提示された「重要身体部位」間の単位時間毎の相対ベクトル、(3)上記ベクトルの1次、2次微分ベクトル、としてエンコードする。これらの時系列ベクトルデータを、中村ら[中村08]のAMSS法を利用することで2つの可変ベクトル間の類似度を求める。そして、良いデータ集合(指導後の動作)と悪いデータ集合(教示前の動作)の類似度の違いから両者を分ける判別式を求め、その判別式から評価指標パラメータを抽出する。

以上で、教示シーンからの切り出し区間・重要身体部位の選定によって評価指標を抽出し、その評価指標に従って正解データと動作データを様々なパラメータから比較することで動作の良し悪しを判断する判別式を求め、評価指標パラメータを推定することができる。

4. 評価実験

本評価実験の目的は、初学者に与えられる教示内容に対してシステムが正しく動作を評価できるかを確認すること、そして取得の際に扱った評価指標に関する「重要身体部位」や「ベクトルパラメータ」といったパラメータに関する考察を行うことである。本実験では一連の教示シーンから評価指標を3パターン抽出し、システムが導き出す判別結果と実際の評価から正解率を計算し、システムの評価及び考察を行う。

4.1 教示タスク

教示を行うにあたって学習者に実践してもらった踊りのステップは、ルンバの踊りの基本ステップとして位置づけられているものであった。このステップの男性用のステップを1人で踊ってもらった中で教示を行った。ステップの教示は複数回にわたって行われ、学習者が十分にうまくなった事を、教示者が判断できるまで繰り返された。

4.2 実験参加者

実験の参加者は本学で募集した学生であり、合計12人(男性10人、女性2人)の初学者を対象に教示を行った。1番目の評価指標学習には8人、2個目の評価指標学習には8人、3番目の評価指標学習には6人の教示タスクを実行した。評価指標毎にデータ取得を行ったため、複数の評価指標の教示を受けた初学者もいる。

4.3 対象となる評価指標

本実験では、(1)ある瞬間の身体の姿勢の正しさ、(2)ある瞬間の腕の振りの正しさ、(3)ある区間における身体動作の正しさ、の3点の教示内容に関して実際に初学者を対象に教示者に教示を行わせ、そのデータを用いて評価指標を抽出した。それぞれのタスクからデータの欠損を除いて表1にある数のサンプルデータを取得した。なお、設定した評価指標に対する重要身体部位も表1に記述する。評価指標1、2はある瞬間の教示に対する指標であり、評価指標3は4拍分のカウントの教示に対する指標と定めた。

4.4 実験結果と考察

まず、評価指標に対するそれぞれの実験結果の分析を行い、評価指標毎に得られる結果の差異からシステムの有効性を考

表 1: 各教示タスクの重要身体部位, および各動作として取得したサンプル数.

重要身体部位	学習前	学習後	教示者
1 「頭部, 首, 左踵」	76	85	15
2 「右肩, 左肩, 左肘」	99	76	15
3 「右肩, 左肩」	73	79	21

察する. そして従来研究では実現できない本研究の貢献を考察し, 最後に獲得した評価指標を新しいステップに適用した場合の正当性を評価することで, システムの汎用性を考察する.

本実験では, 実際の社交ダンスの一般的な教示シーンから, 学習者データとしては「指導前学習者データ」・「指導後学習者データ」を獲得でき, 教示者データとしては「正解データ」を獲得することができる. 評価実験では「指導前学習者データ」を不正解データ, 「指導後学習者データ」を正解データとラベル付けすることで, 各実験参加者一人ずつをテストデータとしてクロスバリデーションを行うことでシステムを評価する.

4.4.1 判別率の分析

実験から得られた判別率を表 2 に示す. この表はは良いデータと悪いデータのサンプル数と, それらをクロスバリデーションを用いて正答率を導いたものである.

表 2: クロスバリデーションによる正答率

評価指標1, Head, Neck, Left Foot									
被験者	A	B	C	D	E	F	G	H	合計
Bad Sample	9	8	11	10	9	10	10	9	76
Good Sample	8	12	9	15	13	10	10	8	85
									161
クロスバリデーション後正解数									
Bad Sample	6	8	10	10	9	4	10	4	61
Good Sample	6	9	6	11	9	9	10	1	61
									122
									判別率 122 / 161 = 0.757763975

評価指標2, Right Shoulder, Left Shoulder, Left Elbow									
被験者	A	B	C	D	E	F	G	H	合計
Bad Sample	7	8	11	23	14	15	12	9	99
Good Sample	7	10	10	13	10	10	10	6	76
									175
クロスバリデーション後正解数									
Bad Sample	3	4	6	19	10	8	7	4	61
Good Sample	4	8	6	9	4	8	2	4	45
									106
									判別率 106 / 175 = 0.605714286

評価指標3, Left Shoulder, Right Shoulder									
被験者	A	B	I	J	K	L			合計
Bad Sample	13	12	12	12	12	12			73
Good Sample	15	12	11	17	12	12			79
									152
クロスバリデーション後正解数									
Bad Sample	13	12	12	12	11	12			72
Good Sample	14	12	11	15	12	12			76
									148
									判別率 148 / 152 = 0.973684211

評価指標 1, 3 の判別率は高く, 抽出された評価指標が妥当であったと考えられる. しかし, 評価指標 2 の判別率は他の 2 つと比較すると判別率が悪い結果となった.

それぞれの評価指標に対応する判別式の各ベクトル時系列データ (判別式の説明変数) に掛かる重みを分析したところ, 評価指標 1, 3 については学習データによって取捨選択の結果が若干異なる部分もあるが, 重みが大きくなる説明変数が全体を通して似通っていることがわかり, システムが有効なパラメータを判断することができていることが確認された. 一方, 評価指標 2 に関しては, 全体を通して重要視されるパラメータが分散しており, 説明変数の取捨選択自体がほとんど行われて

表 3: 全ての身体部位の遷移ベクトルを説明変数として与えた場合の判別結果. 右は本手法の結果.

評価指標1, 全骨格の時系列遷移ベクトル										本来の結果
被験者	A	B	C	D	E	F	G	H	合計	合計
Bad Sample	9	8	11	10	9	10	10	9	76	76
Good Sample	8	12	9	15	13	10	10	8	85	85
									161	161
クロスバリデーション後正解数										
Bad Sample	7	6	4	10	5	6	9	0	47	61
Good Sample	6	6	8	0	12	8	9	8	57	61
									104	122
									判別率 104 / 161 = 0.645962733	0.75776398

いなかったことが確認された.

評価指標 3 に関して詳細に分析したところ, 一定区間のデータを取得していることもあって, 身体の動作間のベクトルデータだけでなく, 「身体座標の時系列遷移データ」や「速度ベクトルデータ」といった, 時間軸によって値が変化するベクトルデータを表す説明変数が重要であることが分かった. この結果より, 長い区間の動作データを取得する場合でも, 高い判別率を持つ判別式を取得できることが実験から確かめられた.

評価指標 2 がうまくいかなかった理由を確認するために, 取得データを詳細に分析した. その結果, 重要身体部位と見なされるべき左手首のパラメータが取得できていなかったことが確認された. また, 「大きく見える」「なめらかに見える」などの評価において, 身体の大きさによって影響があることを考慮しておらず, 体格によって要求される評価指標が異なっていた可能性ある. この点については, 評価指標が同じであっても, 学習者の「悪い箇所」ごとに教示者が分類し, それぞれの観点から評価指標を抽出する必要があると考える.

4.4.2 従来研究との比較

ここでは本研究のように教示者の教示内容を参照することで, 判別にかかる重要身体部位を抽出することの有効性を検討する. そのために, 本研究の手法に使用する説明変数の決定に教示者の教示内容を参照せずに, 単純に全身体部位 (データ取得に失敗した腕を除く) の時系列遷移データを入力し, 良いデータと悪いデータの判別を行った結果を表 3 に示す.

表 3 の結果より, 均一的に多くのパラメータを入力するだけでは判別率が悪くなるのが分かる. この原因としては, パラメータ数が多くサンプルデータに過学習を起こしていることが考えられる. また, 今回注目した教示箇所は「ポイントを押さえていれば美しく見える」という芸術的動作の典型的な場面であったため, 個人ごとに変化があるパラメータが異なっていたことが確認された. これらの結果より, 重要身体部位を抽出することが有効である事を示唆した.

4.5 指標の一般性の検討

獲得した評価指標を, 他のダンスステップで適応させる事で, 他のダンスステップの学習を効率的に行うことができるかどうかを確かめるため, 計算した判別関数を, 同じ評価指標を持つ他のダンスステップに適用した. 評価指標 1 は, ダンスステップの最初の姿勢に対して, 一般的に適用することができると考えられたので, ここまでに抽出した評価指標 1 の判別式を以って判別が上手くいくのかを考察する.

ある学習者の動作の良し悪しを判断させる際には, 判断対象の人物が学習データに入っていない判別式を使用した. 評価実験のとくと同様に, 動作自体の良し悪しは指導前の動作を悪いサンプルデータ, 指導後の動作を良いサンプルデータとして判別を行った. 表 4 にその結果を示す.

この結果より 7 割程度の判別率を獲得することができたこと

表 4: 評価指標 1 で獲得した判別式を異なるステップに適用した場合の判別率。

評価指標 1 をダンスステップ開始時の姿勢に適用した場合									
被験者	A	B	C	D	E	F	G	H	合計
Bad Sample	7	8	6	5	8	6	3	5	48
Good Sample	6	9	6	5	6	5	5	5	47
									95
クロスバリデーション後正解数									
Bad Sample	7	8	4	2	8	1	0	2	32
Good Sample	1	7	6	0	6	4	5	5	34
									66
				判別率	66	/	95	=	0.694736842

が確認されたが、評価指標を違うステップに適用すると判別の精度が落ちてしまうことも分かる。判別率が下がる原因としては、踊るステップが異なることで、同じ評価指標で判断するにしても身体の様子が異なってくるのが理由だと考えられる。従って、より多くのサンプルデータを集め、さらにその中でも幾つかのパターン分けを教示者が行うことで、頑健な評価指標をシステムが抽出できるようになるのではないかと考えている。しかし、評価指標の適用範囲を教示者から獲得すれば、画一的に取得するだけでは得られない具体的な判断尺度を一定の水準で獲得することができたといえる。

5. 議論

本研究の一連のシステムによって抽出した評価指標毎の判別式の評価を行い、ダンス動作判別においての成果を得ることを確認した。本システムの有効性としては、教示者の教示が発生した区間においては、身体動作のキャプチャリングがうまくいった場合において、比較的少ないサンプルデータから教示内容に存在する評価指標をシステムが抽出できる点である。また、その抽出した評価指標に対応する判別式の重みが大きい説明変数を参照することで、システムが重要であると判断した重要パラメータを獲得することができ、学習者は自分の踊りがどのような評価指標から判断されて良いのか悪いのかを知り、どのような重要パラメータに着目して正解データに近づく努力をすべきだと分かることも利点であるといえる。従来の研究では画一的に差分が大きな部分の提示しか行わない研究が多く、こうした教示者の視線を獲得した評価指標の提示システムは、従来研究に比べて教示者のいないシーンでの学習の有効な補助になると考えている。また、説明変数として入力されるパラメータが多くなった場合には、不要であると判断されるパラメータの切り捨てをシステムが上手に行ってくれる点も確認でき、判別式に対する一定の頑健さを確かめることもできた。そして、ある評価指標に対する判別式を新しい動作に適用した場合においてもある程度有効な判別率をもって動作評価を行えることを確認した。

一方で課題となる部分も多く存在している。最も大きな課題は、同じ評価指標だが学習者の動作が異なるタイプに分かれている場合、あるいは、異なる評価指標だが同じ重要身体部位を設定する場合には、動作の質が本質的に異なっているにも関わらずシステムが同じ評価指標であるとして判断してしまう点である。本研究で提案するシステムは一人の教示者の教示内容から評価指標を獲得するという特徴を持っているので、上記の問題を回避する手段として、その一人の教示者が手動で別のラベル付を行うといった処理を追加する方法が考えられる。しかし、手動で教示者にラベル付けを行うことは教示者の負担でしかならないので、同じ動作中に含有される異なるタイプを

システムが自動で判断し、別々にラベル付けを行う処理を追加することを今後の課題としたい。

6. 結論

本研究では社交ダンスを例にとり、社交ダンスの採点項目を向上させるための評価指標の抽出を行うシステムを構築した。その際に教示シーンに発せられる教示内容やポインティング動作から教示内容に関係のある重要身体部位を推定し、その部位に関して学習者の指導前動作データ、指導後動作データ、正解データとなる教示者動作データから判別式を導出した。そして、評価指標毎に導出した判別式から動作の良し悪しの判断を行う評価実験を行い、システムによる動作評価が可能である事を確認した。

今回は教示者が提示したステップの区間に於いて有効な判別式を求める手法を提案したが、今後は多くの評価指標に対する判別式を集めていくことで、教示が提示されなかった部分に評価指標を拡大していくことを今後の課題とする。

また、獲得した指標と神田ら [神田 12] の手法等を組み合わせることで、本研究で獲得した評価指標に対応する課題内容を蓄積することで、動作評価からその後の課題提示までを実現する教示システムを提案して行くことを長期的な展望としたい。

参考文献

- [Chan 11] Chan, J., Leung, H., Tang, J., and Komura, T.: A Virtual Reality Dance Training System Using Motion Capture Technology, *Learning Technologies, IEEE Transactions on*, Vol. 4, No. 2, pp. 187–195 (2011)
- [Huang 12] Huang, H.-H., Uejo, M., Seki, Y., Lee, J.-H., and Kawagoe, K.: Realizing Real-time Feedbacks on Learners' Practice for a Virtual Ballroom Dance Instructor (2012)
- [永山 10] 永山 貴洋, 北村 勝朗: 動作のコツ習得過程における身体知の働きの質的分析: 高等学校女子バスケットボール選手を対象として, *教育情報学研究*, Vol. 9, pp. 33–44 (2010)
- [宮川 05] 宮川 則子, 渡邊 伸: ダンスの初心者指導について (教科教育), *信州大学教育学部紀要*, Vol. 116, pp. 103–110 (2005)
- [山根 05] 山根 亮, 戸高 千智, 川嶋 幸治, 尺長 健: 動作データの時系列相関行列による舞踊動作解析 (顔・身体動作認識, 特集: 画像の認識・理解論文), *電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理*, Vol. 88, No. 8, pp. 1652–1661 (2005)
- [神田 12] 神田 賢一, 山形 佳祐, 大本 義正, 西田 豊明: 練者の知識構造の構築と可視化による動作習得支援システムの構築 (2012)
- [中村 08] 中村 哲也, 瀧 敬士, 野宮 浩揮, 上原 邦昭: AMSS: 時系列データの効率的な類似度測定手法 (データマイニング), *電子情報通信学会論文誌. D, 情報・システム*, Vol. 91, No. 11, pp. 2579–2588 (2008)
- [矢野 12] 矢野 正治, 大本 義正, 西田 豊明: RGB・深度センサを用いた複数人数会話の三次元記録環境の構築, *人工知能学会全国大会 (第 26 回) 論文集* (2012)