

# 仮想社交ダンスインストラクターにおける

## 細分化教示動作の自動抽出

### Automatic Motion Segmentation for Interactive Ballroom Dance Instruction

上亟 正樹 関 優樹 黄 宏軒 李 周浩 川越 恭二

Masaki Uejo, Yuki Seki, Hung-Hsuan Huang, Joo-Ho Lee, and Kyoji Kawagoe

立命館大学 情報理工学部

College of Information Science and Engineering, Ritsumeikan University

**Abstract:** The use of embodied conversational agents is expected in the learning of physical skills such as various sports. This paper describes about an ongoing project aiming to realize a virtual constructor for ballroom dance. In order to achieve natural and highly interactive instruction during the multi-modal interaction between the instructor and the learner, it is necessary to divide the learner's motion into meaningful segments as small as possible in real-time. In the case of ballroom dance, the smallest unit of dance steps is called count which should be synchronized with the beats of accompanied music. A method of automatic segmentation on learner's motion into counts based on trajectory similarity using AMSS is proposed in this paper.

## 1. はじめに

スポーツなど身体動作による技能の学習は、言葉だけでは伝授できず微妙な動きなどを習得するには練習を繰り返すしか他にない。それぞれの種目の技術と教え方に精通した教示者による一対一の指導であれば、個々の学習者の悪い癖を把握し丁寧かつピンポイントの指導が可能になる。また、一方的な説明ではなく対話的な指導が可能であれば学習者が随時インストラクター(教示者)の指導に割り込むことができ、自分がよく理解できなかった部分を小刻みに指導してもらうことができる。教示者も学習者の理解度合いに合わせてフィードバックの仕方を調整しながら教え方を調整することによって学習効率も上がることが期待できる。

しかし、教示者からの教示において以下の制約が存在する。最初は、時間的制約である。教示者は試合や出張が多いこともあり、学習者自身が指導を受けたい時間に時間を確保できない可能性がある。また、人気の教示者であればレッスンの予約を取る事が難しく受講する事さえ困難な場合がある。次に、場所的制約がある。教示者はレッスンを行う為の施

設に勤務しているが、受けたい教示者がいるにも関わらず、家から遠い場所にあり通学が難しいまたは不可能であるという場合がある。また、能力的制約がある。個人により上達の差があるために、全ての学習者が平等に同じ長さのレッスンを受講した場合でも、動作の習得に時間がかかる学習者がいるために動作の習得に差が開いてしまう。

動作の教示の中で表現型の運動は十分に上達した教示者から習う事が多い。今回取り上げる社交ダンスでも同じ事が言える。以上の制約を解消するために「擬人化会話エージェントを用いた社交ダンス教示システム」を提案する。本研究では、場所、時間の制約に関係なく自分の決めた場所で教示を受けられることを目指す。また、身体性を持つ擬人化会話エージェントを利用することで、複雑の動きを行う社交ダンスにおいて教示中にわからない部分などを言語による会話の中で理解を深められるようにする。動作の入力にはモーションキャプチャを利用し、入力された動作データを解析し会話エージェントによる教示に生かす。教示を行う際、社交ダンスの動作の区切りの点「カウント」の部分を利用しどのカウントの部分が間違っているのかを指示することで学

習者は理解しやすくなる。本稿では、エージェントの細かい粒度で教示を可能にするため、動きを「カウント」ごとに動作を細分化するアルゴリズムを提案、実装する。

## 2. 従来研究

従来研究では、Chua ら[1]は、HMD(Head Mounted Display)とモーションキャプチャによって太極拳の教示を行うシステムを構築している。学習者はHMDからバーチャル空間を見渡す。バーチャル空間には3D CGによって表現されたキャラクターがおり、学習者がキャラクターの動作を模倣することによって学習し、モーションキャプチャでその動作を入力する。このシステムではビデオのように手本動作に学習者の動作をリアルタイムで似せるのみでフィードバックや教示を受けることがない。また手本動作との比較はユークリッド距離によって行われるため、精度に欠ける。Akio Nakamura ら[2]は、前後に動くプロジェクタを設置し、動作を行っているキャラクターが3D CGで表現される。学習者の動作に同期して前後にプロジェクタが動き、見やすいようにしている。Chan ら[3]は、違ったフィードバックの手法で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た。3D CGによって動作を見せたあとにモーションキャプチャで動作を入力するが、結果を教示するときユーザの動作をスクリーンに表示し、間違いが起こっている体の部位をハイライトさせて分かりやすくしている。また教示動作と同じ動きの差分を点数化しユーザがどの部位の関節を直すべきか示している。[2][3]とも3D CGキャラクターを用いながら、ユーザがそのインターフェースから直感的に想像する対話的な教示が行われていない。以上に挙げたすべての研究において教示者と学習者の対話的な教示が行えておらず、[3]はシステムの使い方が自明ではなくユーザはダンスを学習する以前にシステムの使い方を学習させられる問題がある。

本研究では、モーションキャプチャを用いて動作の入力、解析を行い間違いが起こっている部分を発見し修正の方法を計算し、その結果について会話エージェントを用いることによって、学習者と教示者の会話による対話的なシステムの実現を目指している。

## 3. 予備実験

本研究を進めるにおいて、教示システムを想定した予備実験を行った。会話エージェントシステムでは、身体を接触するような教示が行えなくなるため、



図1 予備実験環境

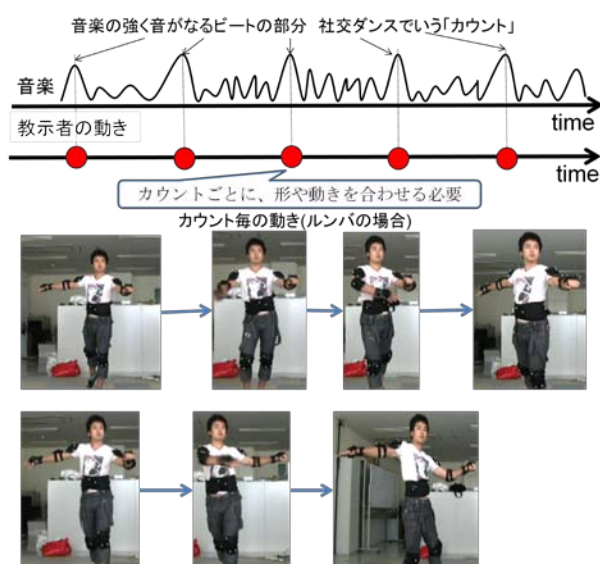


図2 カウントの定義

会話エージェントが人間に教えることを想定して、実際にどのように教示が変化するかを確認するために、人間同士による予備実験を行った。内容はラテン種目のルンバにおいて基礎的な動作ベーシックについての教示を行った。立命館大学競技ダンス部の上級生(社交ダンス学習歴6年間)がエージェント役を演じ、下級生(同1年間)が学習者役を演じた。二人は別々の部屋に分かれ、ビデオ通話でディスプレイ越しに教示を行った(図1)。実験を行った上で、教示者が学習者に対して動作の教示を行う際にどのように教示を行ったかを観察した。まず、社交ダンスには一つ一つの複数のルーティン(技)を組み合わせ一つのダンスになるが、ルーティンにそって決められた体の動きをしなければならず、体の形などが正しい動きかどうかに着目していた。形以外に一番共通して指摘されていた教示としては、カウントに関する発話が多かった。音楽には音が強くなる部分が「ビート」と呼ばれる。社交ダンスではその「ビー

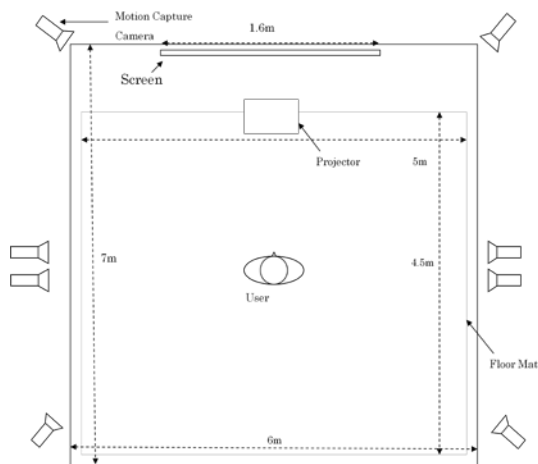


図3 システム構成図

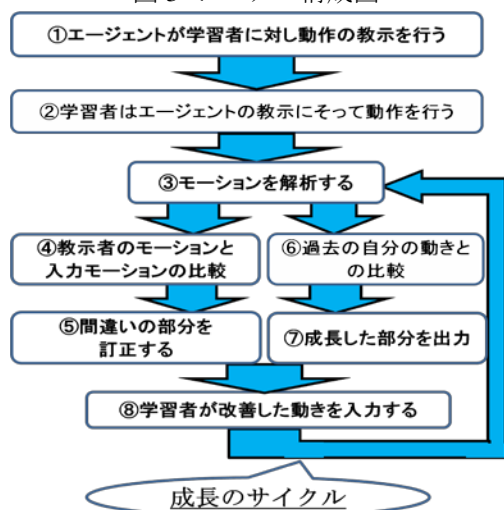


図4 システムフローチャート

ト」の部分で「カウント」と呼ぶ。社交ダンスでは「カウント」の部分に合わせて決められた体の部位を動かすことや形を作る必要がある(図2)。カウントの部分でルーティンごとに決められた形を作って止まっているか、またそのカウントに合わせたムーブメント(動作)が行えているかをチェックしていた。この結果からダンスの教示では、形の他にもカウントの概念を用いた教示の手法が必要であることが分かった。本稿では、動作の中からカウントと思われる部分を抽出することで学習者にとって有効な教示の手法を目指す。

#### 4. 提案システムと実験環境

本研究では、動作の会話エージェント、モーションキャプチャ(NaturalPoint OptitrackFLEX V100R2)、キャプチャスーツによってシステムを構成する(図4)。人間が十分動くことができる距離をとりキャプチャエリアを確保し、地面からの赤外線反射を防ぐためにマット(4.5m×5m)を敷いた。モーションキャプチャの設置は体全体をカバーするために上部に

6台、足先のキャプチャの撮りこぼしを防ぐために学習者からの距離の比較的短い両側のみ下部に2台を増設した。3章の実験結果を元に作成した教示システムの完成形のフローチャートを図4に示す。このシステムでは最初に会話エージェントが学習者に対し、教示動作のデモンストレーションを見せる(①)。動作の教示を受けた学習者は、エージェントの指示通りに、モーションキャプチャを用いて動作の入力を行う(②)。システムは入力されたデータを用いてモーションを解析する(③)。入力されたモーションは、2通りの処理がなされる。まず1つ目は、教示動作と入力動作を比較し(④)間違いの部分を検出し、適切な訂正手法を見つけエージェントがユーザに対して伝える(⑤)。もう一つは、教示動作とこれまでの学習者の入力動作との比較を行う(⑥)。これまでの動作と比較し、上達した点を伝えることによりどのような点が苦手だったのか、また成長しているのかを学習者に知らせ、学習者に気づかせモチベーションを上げる(⑦)。これらの情報から、学習者が動作の訂正箇所に対してデータの入力を行う(⑧)。③~⑧を繰り返し動作の成長を促す。

### 5. 動作解析について

#### 5.1 マーカ装着用のスーツの設計

人間には体格の大きさが違うので、同じ動きをしていても教示動作と入力動作を比較したときに同じ動きとして認識されない場合が考えられる。そのため、体格を考慮したうえで教示動作を学習者の体格に合わせて生成し比較に用いなければならない。またキャプチャ用のスーツが簡単に扱えることも望ましい。通常モーションキャプチャでは専用のスーツを着てキャプチャを行うことが多いが、専用のスーツは密着度が高いため、容易に装着できないことや、体型によってスーツのサイズを購入しなければならないなどの問題がある。そのため、簡易的に扱うことが出来るキャプチャスーツを開発する必要がある。上で述べた問題を解決するためにマーカ装着用のスーツは図5の左のように制作した。腰、肩、肘、手、膝、足の各関節部にマーカを付けたプロテクタを装着することにより、簡易的に脱着できるモーションキャプチャスーツを実現した。また、この配置により各関節間の肢体の長さが分かり、教示動作の体の大きさを入力動作の大きさにリサイズする。

#### 5.2 入力動作のカウント抽出について

##### 5.2.1 カウント抽出について

社交ダンスを指導する際には、3章で述べたよう

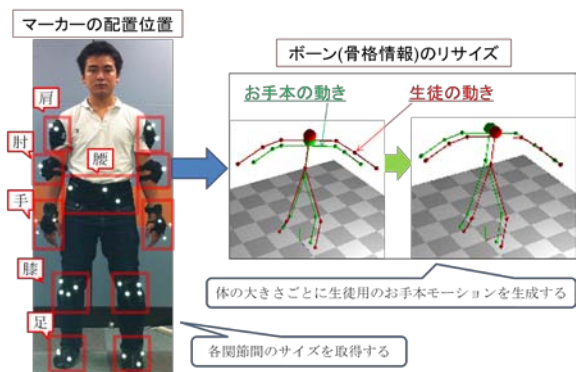


図5 マーカの配置

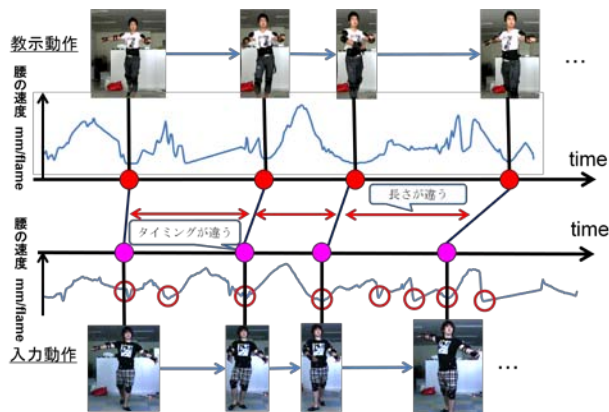


図6 動作中のカウントとカウント候補点

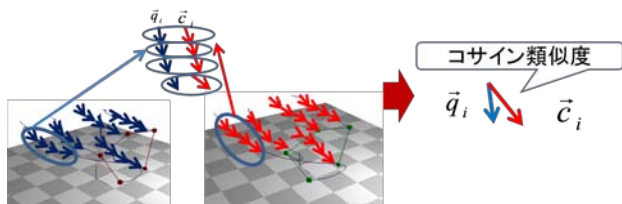


図7 AMSS法によるカウントの計算

にカウントごとの動作を確認し、直すべき部分を確認してから修正することが効率的な教示方法である。本システムでも、教示動作と入力動作のカウントの関係を調べ、入力動作を細分化しタイミングごとの詳しい比較が出来るようにする。しかし、図6のように、教示動作と学習者による入力動作では差異がある。また、初心者はカウントに合わせて踊る事が難しいために動作間の長さも異なる。これらの理由から、カウントの抽出は、動作の完全な一致ではなく、それぞれのカウントごとの動作の特徴を用いたアルゴリズムが必要になる。

カウントを抽出するために、社交ダンスのラテン種目において腰の速度が一番遅くなることから(図6)腰の速度でカウントを区切る位置の候補を抽出する。図6に教示動作と入力動作の腰の速度のグラフ

をプロットしてある。入力動作の速度が遅くなる部分をカウントの候補点とする(図7における入力動作の腰の速度グラフ、丸で囲まれた部分)。カウントの候補フレームに該当するすべての関節から教示動作との類似度を求める。教示動作と事前に処理した入力動作を比較する。 $q_i^n$ を教示動作 $c_j^n$ 入力動作の位置情報とする。 $n$ は関節の種類、 $i, j$ はフレーム数を示す。位置情報を連続する二つのフレーム間の差分を用いて三次元ベクトルを求める。

$$q_i^n = (X_i^n \square Y_i^n \square Z_i^n)$$

$$c_j^n = (X_j^n \square Y_j^n \square Z_j^n)$$

$$\vec{q}_i^n = q_i^n - q_{i-1}^n$$

$$\vec{c}_j^n = c_j^n - c_{j-1}^n$$

変換したベクトルを単位ベクトルに変換し、AMSS法(Angular Metrics for Shape Similarity, 図10)[4]により類似度を求める。AMSS法とは、単位ベクトル化したベクトル同士を下式の計算を用いて行う方法である。

$$sim(\vec{q}_i^n, \vec{c}_j^n) = \cos \theta = \frac{\vec{q}_i^n \circ \vec{c}_j^n}{|\vec{q}_i^n| |\vec{c}_j^n|}$$

$$sim(\vec{q}_i^n, \vec{c}_j^n) = 0 (if \theta > \frac{\pi}{2})$$

AMSS法で求められた数値はコサイン類似度で算出されるために、その値が1に近づくほど類似したベクトルであると言える。AMSS法に基づいて教示動作のカウント間のフレームのベクトルを学習者の動作のフレームのベクトルを用いてコサイン類似度を算出する。カウント候補点のフレーム前後25フレーム、50フレーム間(モーションキャプチャのフレームレートが100fpsであり、0.5秒間で大体のカウント間の動作が行われることから50フレームにした)でAMSS法を用いて類似度を求める。すべての関節(関節数は13個)において関節類似度を求め、類似度の平均(平均類似度)を求める。

$$Ave = \sum_{n=1}^{13} \left( \sum_{i,j} sim(\vec{q}_i^n, \vec{c}_j^n) / 50 \right) / 13$$

教示動作のカウントと入力動作のカウント候補点から平均類似度を求める。その他候補点も同じように平均類似度を求め、平均類似度が一番高い候補点を入力動作のカウント点として設定し、教示者動作のカウントと関係づける。カウント候補点が全てカウントであるとは限らない。

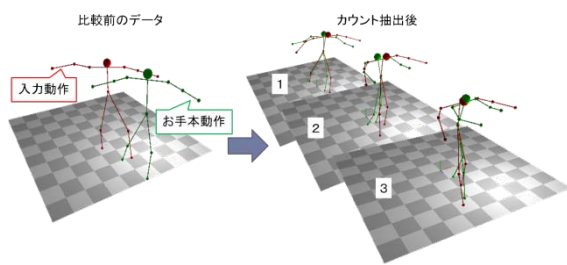


図8 カウントの抽出



図9 教示方法の検討について

## 5. 2. 2 実装

入力動作と教示動作を上述のアルゴリズムを用いて実装を行い、PC上で解析結果を表示させた(図8)。カウント抽出を行い、入力動作、教示動作のカウントごとの同期を行った。その結果計算結果を表示させると図8右のようになった。この結果を見ると教示動作と同じような姿勢を入力動作が行っている。

## 6. 教示方法の検討について

今後、カウントの抽出の結果を利用し、社交ダンスでも重要となる、基礎部分の教示方法に生かしたい。予備実験で指導者が数多く指摘した以下の点を中心に教示を行う予定である(図9)。

### ① 姿勢

背筋から腰骨にかけてのラインを軸という。軸が不必要に曲がっている場合や、膝が動いている時以外に曲がっている場合、姿勢が悪く見た目も悪い状態になってしまう。姿勢を保った状態とは、軸を不用意に曲げず、足の曲げ方が不用意に曲がっていないことである。カウントの部分では特に姿勢に気をつけなければならない。

### ② 形とタイミング

社交ダンスにおいて、動きとカウントのタイミングを合わせることは不可欠である。カウントに合わせることは、リズムに合わせるということでありカウント時に一定のポーズをとり、姿勢を保つ。またカウント間でも動く関節ごとに動くかす順番がある。カウントに合わせることによってきれいにする指標を作る。

### ③ 連続性(軌道の連続性)

ルンバベーシックにおいて動作が完全に停止することは無く、カウントに合わせる部分で制止に近い状態であっても、次の動作へ移行する為の動作を行う必要がある。すべての関節には、動き方や軌道に関係がある。カウント間の全ての関節の動作の動く順序を見ることにより、別々ではなく互いに関係あることを教示しなければならない。

## 6. まとめ

エージェントシステムを用いた社交ダンスの教示システムにおける、予備実験と入力動作から教示動作の情報に基づいてカウントを抽出した。また抽出したカウントをもとに、入力動作をどのような項目で評価するかの提案を行った。予備実験に関しては、プロフェッショナルの教示者を対象に実験を行いアマチュアとの教示の違いや学習効果についての差異を調査し、より学習効果の高い教示モデルを作る。カウントの精度向上をめざし、教示モデルに基づいた間違いを指摘するための解析について4章で述べたアルゴリズムの開発を行っていきたい。

## 参考文献

- [1] Philo Tan Chua, Rebecca Crivella, Bo Daly, Ning Hu, Russ Schaaf, David Ventura, Todd Camill, Jessica Hodgins, Randy Pausch: "Training for Physical Tasks in Virtual Environments: Tai Chi" Virtual Reality, 2003. Proceedings. IEEE, pp 87 - 94 2003.
- [2] Akio Nakamura, Sou Tabata, Tomoya Ueda, Shinichiro Kiyofuji, and Yoshinori Kuno : "Dance Training System with Active Vibro-Devices and a Mobile Image Display" Intelligent Robots and Systems, 2005. (IROS 2005) pp 3075 - 3080, 2005.
- [3] Jacky C. P. Chan, Howard Leung, Jeff. T. g, and Taku Komura: " A Virtual Reality Dance Training System Using Motion Capture", IEEE Transactions on Learning Technologies, PP Issue:99, 2010.
- [4] 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", 電子情報通信学会誌 D, Vol. J91-D, No.11, pp2579-2588, 2008.