

ダンスモーションにおける表現力付与システムの提案

岡田 成美[†] 福里 司[†] 森島 繁生[‡]

本研究では、光学式モーションキャプチャシステムを用いて基準となるダンスモーションデータと、異なる表現を行っているダンスモーションデータを取得し、それらを用いて、テンポを考慮しつつ表現力の乗ったダンスモーションを容易に作成する手法を提案する。

踊り手の意識が鑑賞者に伝わるかの主観評価実験を行い、その結果を基に、基準となるモーションに表現力を付加するフィルタを作成する。なお作成するフィルタは、テンポを一定に保ったまま、緩急と関節角度に注目して変換を行う。これにより得たフィルタを任意のダンスモーションに付加することで、表現力の豊かなモーションの作成を可能にした。

Conversion of a Motion Captured Dance into Expressive Motion

Narumi OKADA[†] Tsukasa FUKUSATO[†] and Shigeo MORISHIMA[‡]

We propose a method that transforms arbitrary dance motion into more expressive motion by filtering in accent and power. Original dance motion is divided into segments and converted to expressive dance to keep the original tempo. The expression conversion rule is extracted by analyzing motion capture data from training dance motions that include neutral and expressive motions labeled by subjective assessment.

1. はじめに

キャラクタアニメーションにおいて、身体に表れる表現力というものは非常に重要な要素である。例えば、単純に歩くという動作においても、その歩き方によって、その人が楽しい気持ちで歩いているのか、それとも悲しい気持ちで歩いているのかということを読み取ることが出来る。つまり、身体に現れる表現力とはその作品をより見応えがあるものにするにあたり、非常に重要な要素であると言える。特に感情や意志を伝達することに優れているダンス動作にとって、表現力は必要不可欠である。しかし、CGキャラクタのダンスモーションは見応えや表現力という点では不十分と言わざるを得ない。その主な原因として、ダンスの動き一つを取っても、腕の運びやタイミング・テンポなどと、表現の幅が広いことが挙げられる。近年CGキャラクタのモーションを手軽に生成するツールの普及に伴い、キャラクタを音楽に合わせて踊らせて楽しむユーザーの数が年々増加傾向にある。しかし、ダンス知識のないユーザーにとって、そういった表現力の幅を持ったダンスモーションの生成を行うことは難しく、多くの経験と時間が必要となる。以上のような背景から、幅広いダンスモーションを手軽に生成できるシステムに大きな需要がある。

モーション生成の従来手法としてキーフレームアニメーションが挙げられる。この手法では複雑な動きも

滑らかに表現することが可能だが、キーフレームを作成する手間がかかる。また、Kinect・モーションキャプチャシステムで逐一動作を取得する方法があるが、キャプチャ精度と時間的コストがトレードオフの関係にある。よっていずれの手法でも、ダンスモーション作成の際に、表現力まで考慮することは現実的ではない。

そこで本研究では、手軽に表現力を考慮する手法を提案する。手順としては、事前にモーションキャプチャシステムを用いて基準となるダンスモーションデータと、異なる表現を行っているダンスモーションデータを取得する。それらのダンスモーションについて、踊り手の意識が鑑賞者に伝わるかの主観評価実験を行う。その結果を基に表現力を定義し、基準となるモーションに表現力を付加するフィルタを作成する。これにより、任意のダンスモーションから、表現力の豊かなモーションの生成を実現する。

2. 関連研究

2.1 表現力に関する研究

本研究を行うにあたり、ダンスにおいて身体に現れる表現力がどのようなものであるかの定義が必要となる。松本らは運動とイメージとの関係の規範となる7つの特性として Natural, Happy, Lonely, Sharp, Flowing, Solemn, Dynamicで構成される7motivesを提唱した[1]。また、阪田らは、身体動作の物理的特徴量や、鑑賞者が主観的に評価した動きの特徴と、7motivesとの関係を提唱した[2]。これらは、ダンスにおける表現力の存

[†] ‡ 早稲田大学/JST CREST
Waseda University
E-mail: † narumi-pink@fuji.waseda.jp, ‡ shigeo@waseda.jp

在を実証するにあたり、有効な研究となっている。

一方、ダンスモーションの表現力に関する研究としては、Tsuruta らの研究[3]が挙げられる。速度・ダンス動作の大きさに注目し、それらのパラメータを変化させることにより、ダンスモーションの表現力を豊かにした。

これら 3 つの研究に共通する問題点は、振付のテンポ等も特徴として考えた、振付自身に現れる表現力についての解析がなされていることである。しかし、あくまでもダンスとは曲のテンポ内で表現されるため、テンポと振付が同じであっても、表現力を変化させてダンスを行うことは可能であると考えられる。

2.2 モーションの誇張に関する研究

キャラクタのモーションを誇張する研究は数多く存在する。今間らはモーションの加速度に着目し、タイミングと身体の形状を変形させることにより、誇張されたモーションの生成を行っている[4]。また、佐藤らはその時点前後での動きの変化によって、8 種類に大別した誇張フィルタの付与、ボーンの長さの変換を行うことにより誇張されたモーションの生成を行っている[5]。

これらの研究は共に「コンピュータシミュレーションによって生成された物体運動は、しばしば観客にとって不自然なものに映ることがある」という現象を前提にしている。つまり、既存の表現力を付与する研究は、いずれもアニメーターが経験的に決定する技法(タイミング、形状変化)を定式化するものである。一方、本稿で挙げるダンスモーションの表現力とは、ダンサーの身体表現に現れるもの(曲テンポ内で表現される動作の速度など)であり、従来の研究で挙げられる誇張表現とは異なる。

2.3 モーションの変換に関する研究

基準となるモーションを変換し、新しいモーションを生成する研究も数多く存在する。Hsu らは基準となるモーションとそれに対してスタイル(姿勢や歩き方など)が異なるモーションから、両者の姿勢、時間差分を抽出して未知のモーションを生成している[6, 7]。

これらの手法は基準となるモーションと、スタイルが付与されたモーションの二つを逐一必要とする。基準のモーションが歩行動作など単純であれば、正解とするモーションの取得は比較的容易である。しかし、ダンスモーションにおいては多数の表現力が存在する。また「表現力が載ったダンス」を踊ることは多くの時間と経験が必要となるため、曲や振付が異なる際に逐一ダンサーのモーションデータを取得することは現実的でない。

以上を踏まえ、本研究では、逐一表現力が付与されているデータを取得することなく、事前に解析したデータを基に、任意のダンスモーションに表現力を付与することを目標とする。そのために解析するダンスモーションは振付が同じであることを前提とし、表現力のみが異なるデータに対して解析を行う。データの取得には、モーションキャプチャを用い、それらのデータから表現力の要素を抽出し、フィルタを作成する。こうすることで、曲と同期されたダンスモーションデータを入力として使用した場合、出力として、曲と同期されたまま、表現力のみが変換されたダンスモーションを得ることが可能となる。

3. 表現力の異なるダンスモーションの取得

本研究において、データの取得には光学式モーションキャプチャを使用する。今回は同じ振付に現れる表現力についての解析を行いたいため、テンポは曲で同期を取り、ダンスの知識を持つ筆者の創作した振付に対し、表現力のみを変えて踊っているデータを取得した。また、個人差をなくすために、ダンサーは全て筆者自身とする。今回使用した表現力は 7motives[1]に基づき、表現の基準となる Natural の他 Happy, Flowing, Lonely, Sharp の 4 種類である。この 4 種類を選定した理由としては、視覚的に表現の違いが判断し易いこと、また振付による依存性が低いことが挙げられる。

4. 主観評価実験

取得したデータの表現力に関する主観評価実験を行った。実験の目的は 2 つあり、1 つはダンサーが意図した表現力を、鑑賞者が認識できるかどうかを検証し、取得したモーションデータの有意性を調べることである。もう 1 つは、鑑賞者が表現力を判断するにあたり、どの要素に着目しているかを調査することである。

実験方法は、被験者に Natural のダンスモーションを見てもらってから、Happy, Flowing, Lonely, Sharp の 4 つのダンスモーションの中から Happy, Flowing の表現力の判断するものとする。ただし被験者には、Happy と Flowing 以外の 2 つの表現力が何であるかは明示しなかった。Happy と Flowing のみを判断してもらった理由としては、7motives において、Happy と Lonely, Flowing と Sharp がそれぞれ対局に存在するものであり、Happy と Flowing の判断がつく場合は 4 つの表現力全てを正しく判断することができるためである。使用した動画は、モーションキャプチャで取得したデータをキャラクタに対してスキニングを施したものであり、問題数は 8 問とし、被験者は 17 人である。



(a)Happy (b)Flowing (c)Sharp (d) Lonely

図1 実験に用いた動画(例:鏡音リン)

左から順に Happy, Flowing, Sharp, Lonely の表現力で踊ったダンスモーションデータである。これらを同時に提示し、Happy, Flowing をそれぞれ選んでもらった

表1 鑑賞者の表現力の認識率

	Happy	Flowing
平均認識率[%]	88.2	73.5
判断基準	動きの大きさ 緩急 重心の運び	腕の動き ための少なさ 重心の運び

また、Happy と Flowing の判断を行う際に、なぜその判断に至ったのかの理由を回答してもらった。実験の参考画像を図1、実験の結果を表1に示す。

主観評価実験の結果より、踊り手の意識する表現力と、鑑賞者が感じる印象は、高い確率で一致することが分かった。よって、この結果を基に、表現力フィルタを作成するに当たり、考慮する要素の決定を行った。

フィルタの種類については、Happy, Flowing 共に認識率において高い数値を得たが、本研究では、表現力を定量的に解析するため、より認識率の高かった Happy フィルタの作成を行うことにした。Happy フィルタとはダンスモーションにおける Happy の表現力を抽出したものを指す。

Happy の表現力を判断するにあたり、鑑賞者は特に、動きの大きさ、ためやキレなどに現れる緩急、重心の動きの大きさについて着目していることが分かった。これを踏まえ、Happy フィルタにおいて変換する要素は、大きさと緩急の2つとした。

5. 表現力フィルタの作成

表現力フィルタは第4章で述べた通り、Happy フィルタを作成する。具体的には、モーションキャプチャで取得した Natural のデータと Happy のデータをそれぞれ比較することで Happy における表現力の抽出を目指す。また、生成した Happy フィルタを任意の Natural データに付与することで、Happy という表現力の載ったモーションの自動生成を実現する。また、フィルタ

作成の際に取得した Natural に対する Happy のデータをフィルタの正解値とする。

本章では Happy フィルタを構成する、緩急変換フィルタ・関節角度変換フィルタの具体的な作成方法について述べる。作成するにあたり用いたダンスモーションデータは、ステップを踏む動作である。

5.1 緩急変換フィルタ作成

5.1.1 セグメント分割

緩急変換フィルタを作成するにあたり、Happy の動作が Natural の動作と比べて、どのように緩急が付いているのかを調べる必要がある。そこで Natural と Happy のデータの時間的な同期を行うことにより、動作の意味をあわせ、ダンスモーションにおける緩急の付き方の解析をする。

時間的な同期は、同じテンポで同じ振付を踊っている部分のデータに対して行う。しかし長いデータに対し一括して同期を行った場合、振付として対応していない部分で同期が行われる可能性がある。そこで今回対象とするデータがダンスモーションデータであることから、テンポを取っているとされる部分で一連の動作を分割し、切り分けたセグメントごとに同期を行った。

セグメントに分割するにあたり、Laban[8]の提唱する、人間の動きの印象を表す要素に着目した。印象を表す要素として、体の部位の動かし方に基づく”Effort”と体の姿勢に基づく”Shape”の2つが存在する。また、Nakataらの行ったユーザスタディ[9]から Shape よりも Effortの方が動きの印象を左右することや、Effortの一種である Weight Effort と人間が感じる「動きの力強さ」との間に相関があることが判明している。そこで、本研究では Weight Effort を導入し、ダンスモーションの推移を定量的に扱う。Weight Effort の大局的な極小値が各動作の繋ぎ目、つまりデンポを取っている部分とされる[10]。Weight Effort は以下の式(1)で表される。

$$W(f) = \sum_{i=1}^N \gamma_i \sum_{j=\{x,y,z\}} |\theta_{ij}(f) - \theta_{ij}(f-1)| \quad (1)$$

ただし、 N は関節数、 F はフレーム番号、 i は各関節、 θ は各関節の角度、 γ は重み係数である。

ステップを踏む動作を使用したため、脚の運びが最もテンポを取っていると仮定し、重み係数は、右脚を1、それ以外を0とした。ただし、極小値を全て検出すると、ノイズを誤検出してしまいう事態が発生する。それを避けるために、経験的に、極小値と、極小値間に現れる極大値に閾値を設定した。極大値がその閾値に満たない場合は、Weight Effort として検出しないとい

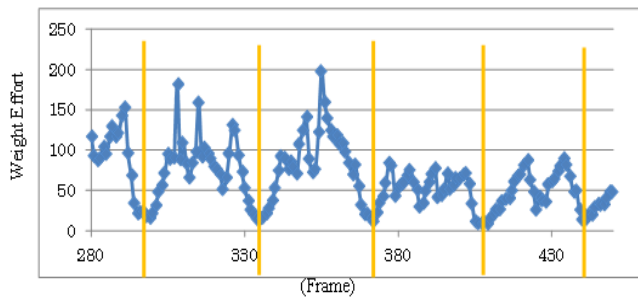


図2 Weight Effort によるセグメント分割

う制限を課した。図2に Weight Effort を用いてグラフを分割した例を示す。

5.1.2 時間的同期

それぞれのセグメントにおいて時間的な同期を行い、緩急の付き方を調べる。同期を行うに当たり、Dynamic Time Warping[11](以下 DTW)という手法を用いた。DTW とはデータ長さやピーク値をとるタイミングが一致するように2つの波形データの同期を取るアルゴリズムである。DTW を行うにあたり、2つのデータの姿勢の「見た目」の差異を定量化するような姿勢距離測度 d_p と呼ばれる指標を導入する。ステップを踏む動作であることを考慮し、右膝の角度を基に姿勢距離測度を算出した。姿勢距離測度 d_p を式(2)に示す。

$$d_p = \sum_i \omega_i \text{dist}(q_N, q_H) \quad (2)$$

ただし ω_i は加重係数、 q_N は Natural における右膝の角度、 q_H は Happy における右膝の角度、 $\text{dist}(q_N, q_H)$ は q_N と q_H 差分値を表す。

これを格納した姿勢距離行列を作成し、そのコスト最小経路を求めた。

DTW において、セグメント内で Natural と Happy で取る値の差が過度に大きいと上手く同期されない可能性がある。そのため、それぞれのデータで、最大値と最小値を基準に正規化を行った。さらに、DTW は、あくまでも最短経路の解法であり、フレームの連続性を考慮しない。そのため DTW の欠点として、動作の不連続性が挙げられる。これを回避するために、120(Frame/Second)のモーションキャプチャデータを1200(Frame/Second)へとアップサンプリングした。この際、アップサンプリング前後でデータの意味、見た目が変わらないこと、不自然なモーションにならないことが条件となる。今回はオリジナルのデータを基にスプライン補間を行い、10倍に再分割する手法を用いた。スプライン補間を選択した理由としては、オリジナルの点を必ず通り、点と点の間を1次導関数、2次導関数が連続となるように滑らかに繋ぐ補間方法であると

いうことが挙げられる。10倍にアップサンプリングしたモーションデータは数値的、視覚的にも滑らかなアップサンプリングを実現した。

上記の処理を行った場合でも DTW により動作不連続性が確認される場合がある。そのため、更に角速度による制限も施した。同期の際の対応するフレームとして同一フレームが30以上連続して検出された場合は、連続する前のフレームと、連続した30個目のフレーム間の角速度を算出した。角速度の閾値は、実測のデータの最小値と最大値を基に0.01(Degree/Frame)以上、0.35(Degree/Frame)以下とした。算出した角速度が閾値内の場合は、次のフレームを強制的に隣のフレームに更新するという処理、閾値外の場合は、その半分である、16フレーム目を強制的に隣のフレームに更新するという処理を行った。DTW 前後の右膝の角度の時間変化の結果を図3に示す。

以上の手順により、セグメントごとに、Natural が Happy に同期される際にどのような緩急が付加されるかのフィルタを得た。それらのフィルタの最初の部分に着目することによって、2種類に大別することが出来た。図4に得られたフィルタを示す。横軸が変換後のフレームであり、縦軸が変換前のフレームである。2種類のうち1つは、最初の部分の傾きが急になっている。つまり、セグメントの最初の部分がキレに似た動作になるような変換になっている。もう1つは、最初の部分の傾きが緩やかになっている。つまり、セグメントの最初の部分がゆったりとした、ために似た動作になるような変換になっている。これらの特徴を反映するフィルタを手動で作成した。作成した2種類のフィルタを図4に示す。ただし、最初の部分の特徴にのみ着目しており、変換前後でテンポを保つためにフレーム数を同じにするために、最初の部分で発生したフレーム数の誤差はセグメントの最後で調節するようなフィルタとした。

一意に関数フィッティングを行わなかった理由としては、セグメントの長さは Weight Effort の極小値の検出位置に依存し一定でない。そのため、一意に関数フィッティングを行うと、セグメントの長さによって、動作不連続に見える程度の緩急が付与されることが懸念される。よって動作不連続性を回避するために、一定値を圧縮、引き伸ばしを行うフィルタを緩急変換フィルタとした。

なお、任意のモーションに緩急変換フィルタを適用する際は、入力したモーションを Weight Effort の極小値で区切り、セグメントごとにフィルタを適用する。ダンスモーションにおける緩急は、セグメントの繋ぎ目において滑らかに変化するため、緩急変換フィルタ A, B を交互に選択した。

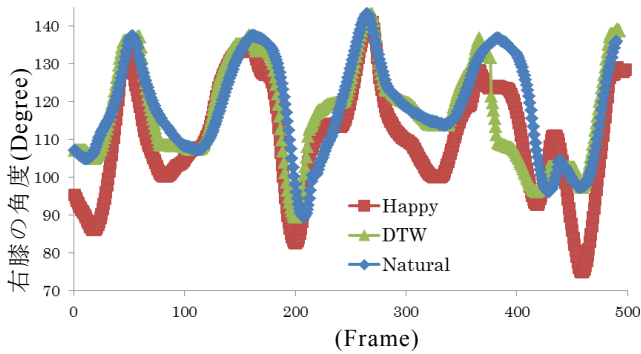


図3 DTW 前後における右膝の角度変化の様子

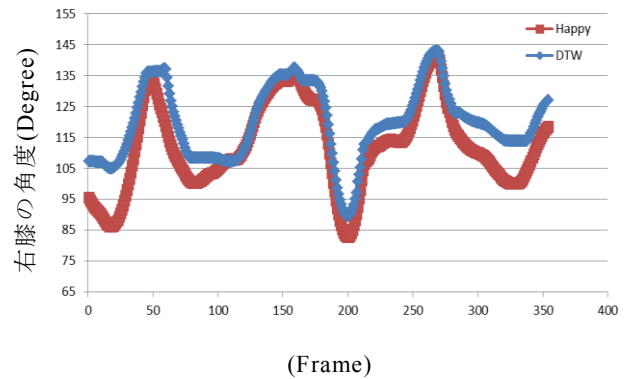
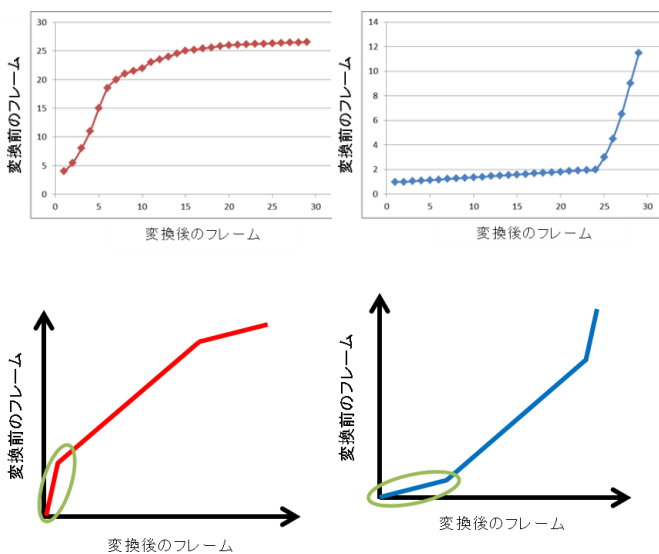


図5 角度変換前の右膝の角度の様子



フィルタ A

フィルタ B

図4 緩急変換フィルタ

5.2 関節角度変換フィルタの作成

動きの大きさを表現するにあたり、着目部位の関節角度を変換フィルタの作成を行った。関節角度の比較を行うために、この時点でオイラー角表示されている Happy のデータと、5.1 の処理により Happy に同期されている Natural のデータをワールド座標に変換した。ワールド座標を基に、内積の計算により膝の角度を算出し、比較を行った。この時点での右膝の角度の図を図5に示す。この図から分かるように、Natural と Happy において、極大値については差が見られなかった。これは、膝が伸びきっていると考えられる部分であり、人体の構造上このようなデータになったと考えられる。そこで、差が顕著に表れている極小値に着目した。極小値ごとに Happy つまり正解値と比較を行い、その平均比率を算出した。関節角度変換フィルタの内容としては、極小値を平均値倍、極大値を1倍として、そ

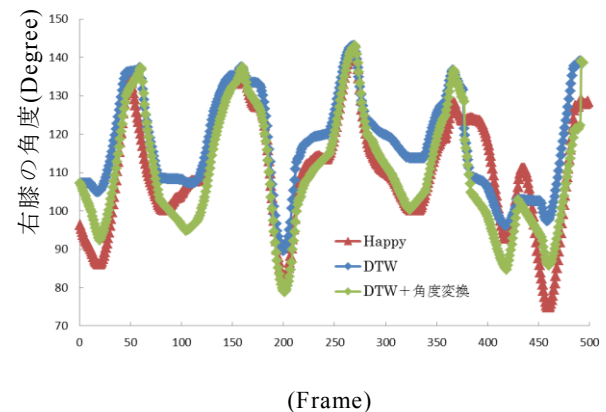


図6 角度変換前後での正解値との比較

の間の比率を極小値と極大値間のフレーム数に応じて線形的に算出し、補間を行うものとした。Happy フィルタにおいて、関節角度の極小値の平均倍率は0.88倍であった。Natural に角度変換フィルタを付与した結果と正解値との比較を図6に示す。変換前と比較を行うと、正解値に近づいていることが分かる。

また、膝の角度の変換を行うと、脚の甲の長さが伸び、更に足先が地面から浮いている不自然なモーションになってしまう。これを回避するために、まず、BVH ファイルに記述されているボーンの長さの情報を基に、甲の長さを一定に保つ処理を行った。更に、体が宙に浮くことがないように、つま先の座標に着目し、両足が地面から離れていると判断されたフレームでは、より低い位置にある方の足先を変換前の座標に合わせるという処理を行った。その処理に伴い、全身の座標も平行移動させることにより、全身のボーンの長さを一定に保った。これらの処理により、膝がより大きく曲がるように変換されたフレームでは、重心の上下の運動も表現することが可能となった。

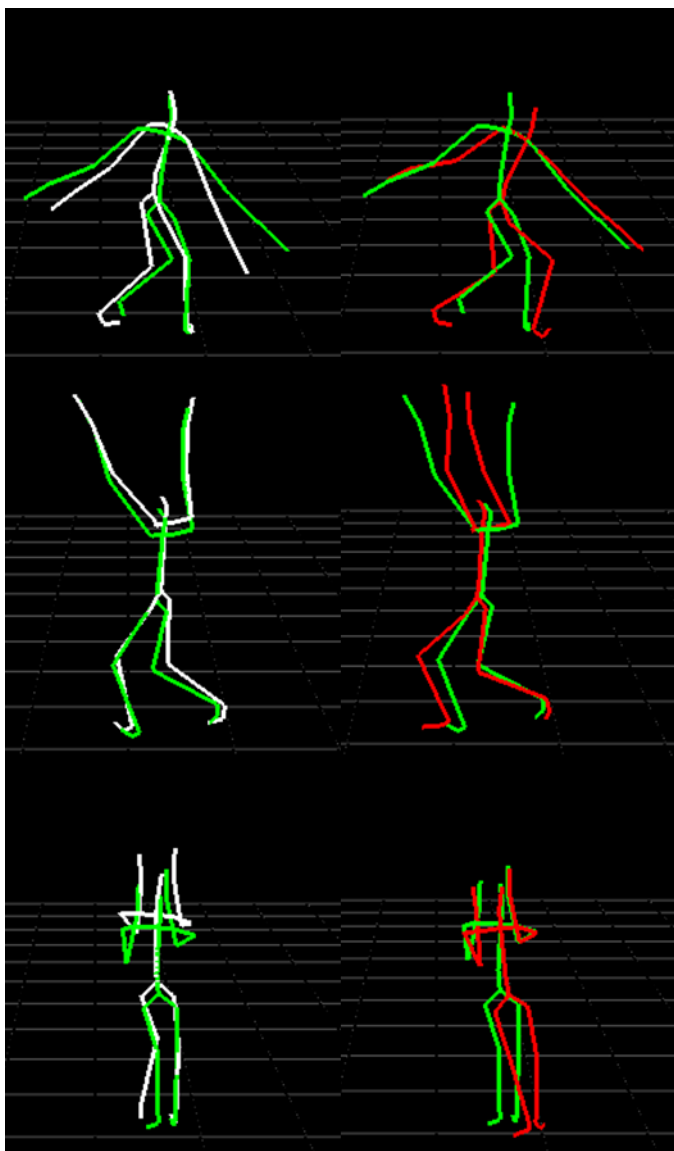


図7 変換後と、変換前・正解値との比較
 白：Natural 緑：Our Method 赤：Happy

6. フィルタ付与結果

6.1 正解値との比較

上記の手順で変換を行ったデータと、変換前、正解値(Happy)との比較を図7に示す。緩急フィルタをかけていることにより、脚のタイミングがHappyに合うようにずれていることが分かる。更に、膝の角度変換により現れた、重心の上下運動も、Happyの動作に似ていることが分かる。

6.2 任意動作付与結果

第4章にて作成したHappyフィルタを任意のダンスモーションデータに付与した。今回は、Perfume Global Siteで公開されているPerfumeのダンスモーションのBVHファイル[12]を使用した。このPerfumeのデータ

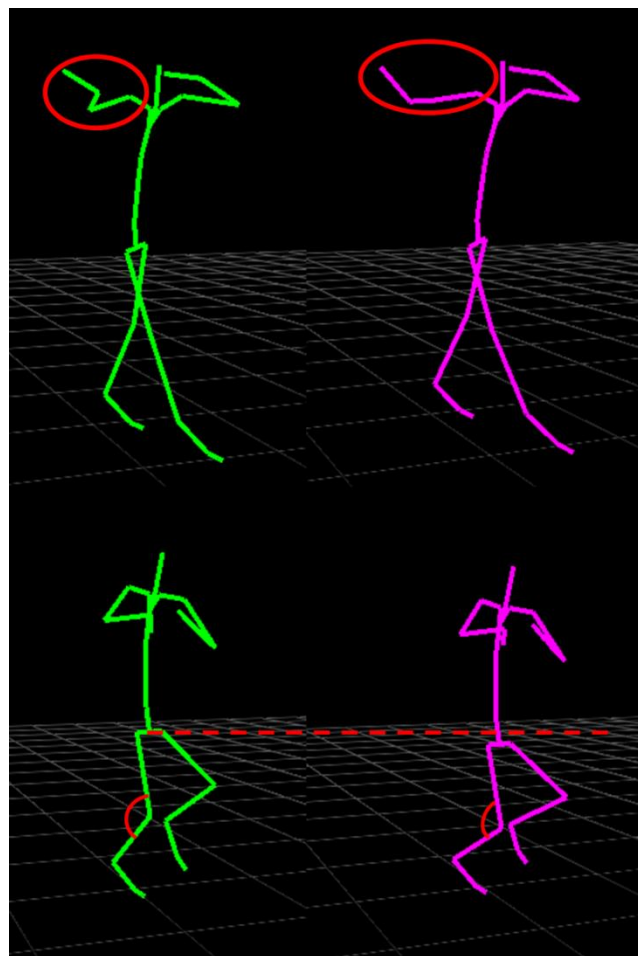


図8 フィルタ適用結果
 緑：Original 赤：Our Method

においても、右足のWeight Effortを算出し、セグメント分割を行った。セグメントごとに緩急変換フィルタを付与し、更に膝の角度について関節角度変換フィルタを付与することによってHappyという表現力を持つモーションを生成した。緩急変換フィルタについては第5章でフィルタAとフィルタBの2種類を作成し、繋ぎ目における緩急の滑らかさを踏まえ、フィルタAとフィルタBを交互に付与した。フィルタ付与前後の比較を図8に示す。図より、緩急変換フィルタによる、ため要素の表現や、角度変換フィルタによる膝の角度の変換、及び重心位置の振れ幅の大きさの表現がなされていることが分かる。これらにより、変換前のオリジナルのモーションと比べて、より躍動感のあるモーションが生成できていることが分かる。

7. まとめと今後の課題

本研究では、モーションキャプチャにより、テンポと振付が同じであり、表現力のみが Natural/Happy と異なるダンスモーションデータを基に、Happy の表現を抽出し、フィルタの作成を行った。作成したフィルタを Natural の動作に付与することにより、テンポを保ったまま、Happy な表現力が付与されたダンスモーションを簡易的に再現することが確認出来た。また、作成したフィルタを任意のダンスモーションに適用した結果、入力 of データに対して、より躍動感のあるダンスモーションを生成することを実現した。

本研究では、フィルタ作成、そして付与の際、着目部位を経験的に脚に着目した。しかし、ダンスモーションの振付により、脚よりも腕や体幹の動きに特徴がより表れる場合も考えられる。よって、実際に取得したモーションキャプチャデータを用いて、身体の部位ごとの Weight Effort 等を検出することにより、着目部位の自動決定、着目部位ごとのフィルタの作成を行う必要があると考えられる。

また、本手法は、関節角度変換フィルタ付与の際にオイラー角表示からワールド座標に変換を行っている。しかし、ワールド座標からオイラー角表示に一意に変換を行うことは困難である。その理由としては、ワールド座標への変換の際、ボーンの回転の情報が失われてしまうことが挙げられる。キャラクタに適用し、Happy などの表現力を視覚的に確認出来るかを検討する必要があるため、回転情報を保持した角度変換の方法を検討する必要がある。

更に、今回は Happy の表現力についての解析を行ったが、それ以外の表現力の付与も今後の課題である。

本研究の研究目的は、ダンスの表現力を定量的に解析し、任意動作に付与することである。つまり、あくまでもダンスと音楽の相関ダンスモーションに現れるテンポ情報しか考慮していない。今後は音楽の盛り上がりや雰囲気等を考慮し、それと連動してダンスモーションの表現力を変化させるシステムを目指したい。

関連研究

- [1] 松本千代栄, “舞踊研究・課題設定と課題会月学習 II : 運動の質と感情価”, 日本女子体育連盟紀要, p53-59, (1987).
- [2] 阪田真己子, 八村広三郎, “身体動作における感性情報の関係モデル構築”, 表現文化研究, 6, p194-202, (2007).
- [3] Seiya Tsuruta, Woong Choi, Kozaburo Hachimura, “Generation of Emotional Dance Motion for Virtual Dance Collaboration System”, Digital Humanities, UK, pp. 368-371, (2010).
- [4] 今間俊博, 近藤邦雄, 栗山仁, 古家嘉之, “モーションキャプチャデータを用いた加速度制御手法によるメンタルモーション生成”, journal="図学研究", (2005)-06-01.
- [5] 佐藤修一, 近藤邦雄, 佐藤尚, 島田静雄, 金子満, “アニメーション制作におけるキャラクタの動作強調手法 Motion Filter”, テレビジョン学会誌, 49(10), (1995).
- [6] Eugene Hsu, Kari Pulli, Jovan Popovic, “Style Translation for Human Motion”, ACM Transactions on Graphics (SIGGRAPH 2005), 24(3), pp. 1082-1089, (2005).
- [7] Eugene Hsu, Marco da Silva, Jovan Popovic, "Guided Time Warping for Motion Editing", ACM SIGGRAPH/Eurographics Symposium on Computer Animation 2007, pp. 45-52, (2007).
- [8] Rudolf Laban and Lisa Ullmann. “The mastery of movement”, July, (1971).
- [9] Nakata T, Mori T, and Sato T. “Entertainment and amusement robot technologies. analysis of impression of robot bodily expression”, J Rob Mechatron, Vol. 14, No. 1, pp. 27-36, (2002).
- [10] 白鳥貴亮, 中澤篤志, 池内克史, “音楽特徴を考慮した舞踊動作の自動生成”, 電子情報処理学会情報誌, p2242-2252, (2007).
- [11] Lucas Kover, Michael Gleicher, “Flexible Automatic Motion Blending with egistration Curves”, Symposium on Computer Animation, (2003).
- [12] <http://www.perfume-global.com/>