

適合フィードバックに基づく 好みを反映したダンス編集手法

柿塚 亮^{1,a)} 佃 洗撰² 深山 覚² 岩本 尚也¹ 後藤 真孝^{2,b)} 森島 繁生^{3,c)}

概要: 本稿では、ユーザの好みを反映したキャラクターダンスアニメーション制作システムを提案する。本システムで、ユーザは入力楽曲の任意の部分に対して、データベース中の好みの振付を割り当てる。ここで、ユーザは適合フィードバックに基づくダンス検索システムを通して、マウスクリックのみでデータベース中の好みの振付を探ることができる。そして、ユーザが振付を指定しなかった楽曲部分は、システムが自動的にダンスをつなぎ合わせ、楽曲に合った一連のダンスを生成する。本システムが実現されることで、ユーザが簡単に好みを反映したダンスを制作できることが期待される。

キーワード: 適合フィードバック, キャラクターアニメーション, モーションキャプチャ, ダンス

A choreographic authoring system reflecting a user's preference

RYO KAKITSUKA^{1,a)} KOSETSU TSUKUDA² SATORU FUKAYAMA² NAOYA IWAMOTO¹
MASATAKA GOTO^{2,b)} SHIGEO MORISHIMA^{3,c)}

Abstract: We propose a new system for constructing character dance animation by considering animator's preferences. First, a user of the proposed system assigns a preferred motion, obtained through a searching algorithm, to arbitrary part in the music. Then the proposed system automatically assigns motions to the other remaining parts of the music by using motions in a database. This system enables a user to create a new dance performance for character animation considering his/her preference.

Keywords: relevance feedback, character animation, motion capture, dance

1. はじめに

近年、3DCG キャラクタを音楽に合わせて踊らせる動画コンテンツが増加している。こういったコンテンツの制作手段として、キーフレームにおけるキャラクタの姿勢を逐次調整していく方法と、モーションキャプチャシステムを

用いて取得したダンサーの動きをキャラクタにリターゲットする方法の二通りがある。しかし、前者はダンスに関する知識や多くの手間と時間が必要であり、後者は高価な機器や撮影環境を用意しなければならないという問題があり、いずれも経験のないユーザにとっては敷居の高い制作過程となっている。このような背景から、ユーザが作成したモーションデータを web 上で公開し、他のユーザがそのデータを利用することで容易に動画コンテンツを作成するといったことが行われるようになった。日本で最も利用されている動画共有サービスのひとつであるニコニコ動画^{*1}では、あるオリジナル動画を元に新しい動画を作成する「N 次創作 [1]」が盛んである。その中には、web 上で公開されたモー

¹ 早稲田大学

Waseda University

² 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

³ 早稲田大学理工学術院総合研究所

Waseda Research Institute for Science and Engineering

a) kakitsuka.99821@ruri.waseda.jp

b) m.goto@aist.go.jp

c) shigeo@waseda.jp

^{*1} <http://www.nicovideo.jp>

ションデータを利用することで自分の好きなキャラクタをオリジナル動画の楽曲に合わせて躍らせる動画が多数存在する。しかし、特定の楽曲用に作成されたモーションデータをユーザが自分好みのモーションに修正するのは依然として困難である。また、既存のモーションデータを利用することで、入力楽曲にシンクロしたダンスモーションを自動生成することを目的とした研究も行われてきた [2][3][4][5]。これにより、専用のモーションデータが公開されていない楽曲に対しても、キャラクタのダンス動画を作成することが可能になったが、これらの研究の中ではユーザの好みを反映するという事は考慮されてこなかった。

そこで、本稿ではユーザの好みを反映したダンスアニメーション制作のために、大幅な振付の変更を容易に可能とするシステムの実現を目的とする。提案システムでは、ユーザが一連のダンスの中で編集したい部分をデータベース中の好みのモーションに簡単に置き換えることができる。その際、システムは候補となるダンスモーションを複数表示し、情報検索における適合フィードバックの枠組みを利用することで、ユーザが好みのダンスモーションを見つけるのを手助けする。そして、編集部分前後のダンスモーションについては置き換えたモーション間が自然につながるように自動生成される。これにより、ユーザの好みに合ったダンスを簡単に制作できるようになると期待される。

本研究の主な貢献は次の4点である。

- (1) ダンスアニメーションの制作において適合フィードバックの枠組みを導入した。
- (2) 適合フィードバックを用いてダンスモーションの検索を繰り返すことで、ユーザに揭示されるモーション間の平均非類似度が低くなる、つまりユーザの好みを反映したダンスモーションに探索範囲が絞られることを定量的に示した。
- (3) モーション間の連結性を考慮したダンスの自動生成手法を提案した。提案手法により、連結性を考慮しない場合に比べて、ユーザが指定した振付部分は固定であるという制約条件のもとで、自然なダンスモーションが生成されることを定性的に示した。
- (4) 視聴者の好みが反映可能なダンスアニメーション制作システムのためのインタフェースを提案し、システムを実装した。

本稿の構成は以下の通りである。2章では関連研究について述べる。3章ではまず提案システムの概要を示し、次に適合フィードバックを用いたダンスモーション検索手法およびモーションデータ間の連結性を考慮したダンスの自動生成手法を述べる。さらに、提案システムのインタフェースについて述べる。4章では提案手法の有効性を示すための評価実験について述べ、5章で本稿をまとめる。

2. 関連研究

入力楽曲にシンクロしたダンスを自動生成する研究の代表的なものとして、Shiratoriらの研究 [2] がある。Shiratoriらは、入力楽曲に対して盛り上がりやテンポが合うようにデータベース中のダンスモーションをつなぎ合わせていくことで、新たな振付の提示を行う手法を提案している。その他のダンス自動生成の既存研究では、音高と和音 [3]、メロディの概形 [4] などがダンスと対応付けて分析され用いられている。また、ガウシアンプロセス (GP) を用い、ダンスと音楽の対応関係に加えてダンス動作自体を学習することで、既存のダンスの断片を切り貼りするだけではなく、学習データにないダンス動作を自動生成する手法も提案されている [5]。さらに、生成される舞踊動作の盛り上がりや、ユーザが加速度センサによる入力によってリアルタイムでコントロールする手法もある [6]。しかし、こういった自動生成手法では生成された振付に対して編集ができないため、ユーザの好みを振付に反映できない。モーション編集に関する研究として、ソースモーションに対してタイムライン上の姿勢、動作のタイミングや運動の速度を変更することで出力となるモーションを得る手法がある [7]。こういった編集手法は、使用したいモーションが定まっておらず、それに対してより詳細な編集をする際には有用であるが、楽曲の中の同一部分においてベースとなる振付自体をほかのものに入れ替えて編集したい場合には不適である。そこで振付データベースから所望のモーションを選び入れ替える手法が考えられる。そのためには、データベース中から所望のモーションを検索する必要がある。しかし、多くのモーション検索手法 [8][9] では、クエリとして Kinect などモーションデータを入力する必要があるため、特にダンスモーション検索ではユーザへの負担が大きい。簡単にモーションを検索できる手法として、Choiら [10] はユーザがスケッチした棒人間を入力としてモーション検索を行う手法を提案した。しかし、ユーザのスケッチによるモーション表現には限界があり、複雑なモーションを多く含むダンスモーションの検索にはこの手法を用いることは難しい。このように、現状ではユーザの好みに合った振付を作ることが困難である。よって本研究では、簡単にユーザの好みに応じたダンス制作を可能とするモーション検索・編集手法の実現を目的とする。

3. 提案手法

本稿ではユーザの好みを反映したキャラクターダンスアニメーション制作システムとそれを実現するための手法を提案する。提案システムの概要を図1に示す。提案システムでは、まずユーザは入力楽曲の中でこだわりのある部分に対してデータベース中の好みのダンスモーションを割り

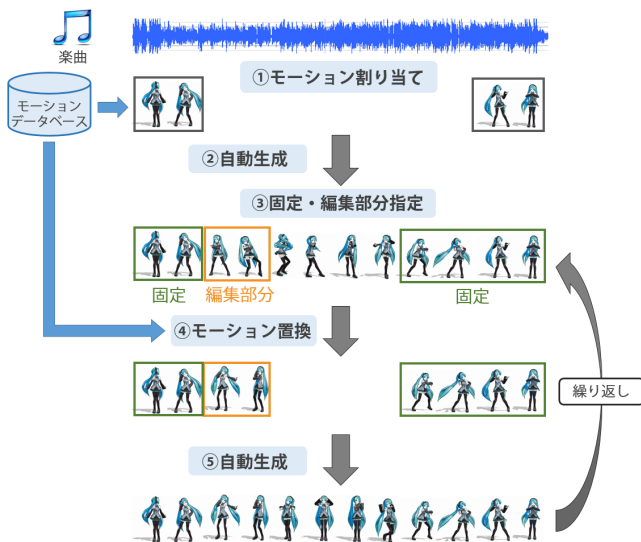


図 1. 提案システムの概要

当てる (図 1, ①). それ以外の部分についてはダンスモーションが自然につながるように自動生成される (図 1, ②). 次に, ユーザは生成された一連のダンスの中で気に入って固定しておきたい部分と編集したい部分を指定し (図 1, ③), 編集したい部分についてはデータベース中から好みのモーションを探して置き換える (図 1, ④). そして, ユーザが指定した固定・編集部分以外については再度自動生成を行う (図 1, ⑤). これを繰り返すことでよりユーザの好みを反映した一連のダンスモーションを制作していく.

図 1 に示した提案システムの中で, 解決しなければならない技術的課題は二点ある. 一つ目の課題は, ユーザがデータベースの中からどのように好みのダンスモーションを検索するのかということである. また二つ目の課題は, ユーザが選択したモーション間をどのように自動生成するのかである. 本研究で用いる手法の詳細はそれぞれ「3.1 モーション検索システムの構築」, 「3.2 ユーザ指定部分前後の振付生成」で述べる.

本研究で用いるデータベース中の各モーションデータは, インターネット上で公開されているダンスモーション [11] を 4 カウント毎に等分割して得られたものである.

3.1 モーション検索システムの構築

本節では, ユーザがどのように好みのモーションを探すかという課題に対する本研究の手法について述べる. ダンスの知識に乏しい多くのユーザにとっては, 実際に楽曲に合わせてダンスを見てみなければその振付の良し悪しや好みのダンスであるかどうかを判断することは難しい. そこで, 本研究では, ユーザが指定した楽曲中の任意の部分について候補の振付となるダンスモーションを複数提示し, 楽曲と合わせてダンスモーションをプレビューしながら, 探索的に好みの振付を検索できる手法を提案する. ユーザは

適合フィードバック [12] に基づくモーション再検索システムにより, 提示される候補モーションを変更することができる. 再検索の初期段階では, 好みの振付と出会う可能性を広げるため, 提示モーション間の非類似度が大きくなり, 再検索回数が増えるにしたがって非類似度が徐々に小さくなるようにした (以降, 提示モーション間の非類似度を「探索範囲」と呼ぶ).

本検索システムでは, Dou ら [13] のフレームワークを利用し, 編集中の楽曲部分と候補モーションデータの相性, ユーザが検索の過程でどのようなモーションに高評価をしたかを同時に考慮して, 次に提示する候補モーションを決定する. $n + 1$ 個目のモーションデータ m_{n+1} は

$$m_{n+1} = \operatorname{argmax}_{m \in U \setminus S_n} \{ \rho \cdot \operatorname{rel}(q, m) + (1 - \rho) \cdot \Phi(m, S_n, L_l) \} \quad (1)$$

と, $\operatorname{rel}(q, m)$ と $\Phi(m, S_n, L_l)$ の線形和を最大化するデータベース中のモーションデータ m で与えられる ($0 \leq \rho \leq 1$). 式 (1) における $\operatorname{rel}(q, m)$ は編集中の楽曲部分 q と候補モーションデータ m の関係を表しており, 編集中の楽曲部分のインテンシティと候補モーションのインテンシティがマッチしているほど大きい値となる. ここでは, 編集中の楽曲部分のインテンシティとして RMS を用い, 候補モーションデータのインテンシティとしては, 次式で表される全フレームにおける人体の各関節位置の速さの線形和を用いた.

$$W(f) = \sum_{i=1}^N \alpha_i \cdot |\dot{x}_i| \quad (2)$$

N は関節数, \dot{x}_i は各関節の速度, α_i は関節部位ごとに決められた重みである. そして, 各モーションデータのインテンシティはこの値の時間平均とした. 式 (1) における $\Phi(m, S_n, L_l)$ は提示される候補モーションの多様性を制御する項となっており, 次のように定式化した.

$$\begin{aligned} \Phi(m, S_n, L_l) = & \tau \cdot \min\{D(m, m_i) | m_i \in S_n\} \\ & + (1 - \tau) \cdot \max\{Sim(m, m_j) | m_j \in L_l\} \end{aligned} \quad (3)$$

S_n は編集中の楽曲部分に対する候補モーションとして既に提示された n 個のモーションデータの集合であり, L_l は編集中の楽曲部分に対する候補モーションとしてユーザが高評価を付けたモーションデータの集合である. また, $D(m, m_i)$, $Sim(m, m_i)$ はそれぞれモーションデータ m と m_i の非類似度, 類似度を表しており, 固有値解析に基づくモーション間の類似度尺度 [8] を用いて算出を行った. 式 (1) における U はデータベース中のモーションデータ全体の集合であり, m_{n+1} はデータベースから既に提示された n 個のモーションデータの集合 S_n を除いたデータの中から選ばれる. τ は, 再検索回数の増加に伴い減少させることで, 提示される候補モーションの多様性を徐々に収束させ

る役割を担う。 k 回目の検索における τ の値を次のように決めた。

$$\tau_k = \frac{19}{20} \tau_{k-1} \quad (4)$$

ただし、 $\tau_0 = 1$ である。

3.2 ユーザ指定部分前後の振付生成

ユーザが指定したモーションデータは固定であるという制約条件のもとで、その間の振付を、データベース中のモーションデータを用いてダンスが自然につながるように自動生成する。これにより、ユーザは楽曲の中でこだわりの強い部分に注力してダンス制作を進めることができる。

3.2.1 振付生成に用いるモーションセグメント選択

ユーザが指定したモーションデータの間の補間に用いるモーションセグメントを自動で選択する。まず、音楽のリズムとダンスモーションのリズムが一致するように全モーションセグメントを時間方向にリサイズする。次に、各モーションセグメント間のつながりやすさを表す連結評価関数 S の逆数を各セグメント間の遷移コスト $C (= 1/S)$ として、遷移コスト C の合計が最小となる経路をダイクストラ法 [14] を用いて求める (図 2)。この連結評価関数 S では、セグメント間のつながり目前後で姿勢に関する類似度 S_{pose} と動きに関する類似度 S_{move} を評価する。モーションセグメント A のフレーム i^A とモーションセグメント B のフレーム j^B との間における姿勢に関する類似度 S_{pose} と動きに関する類似度 S_{move} の算出方法について以下で述べる。

姿勢の類似度評価関数 S_{pose} は、人体モデルの各リンクの方向ベクトル \mathbf{v} の内積と各リンクに対する重み β_l を用いて算出する。ここで、各リンクの方向ベクトル \mathbf{v} は関数 h を通して同時座標系における単位ベクトルに変換する。

$$S_{pose}(i^A, j^B) = \sum_l^N \beta_l (h(\mathbf{v}_l(i^A)) \cdot h(\mathbf{v}_l(j^B))) \quad (5)$$

動きの類似度評価関数 S_{move} は、人体モデルの各リンクの速度により計算される。この速度は、関数 h を通して単位ベクトルに変換され、その内積を算出することでベクトルの大きさと向きを同時に評価できる。連結前のオリジナルの人体動作列の中でのリンクの方向ベクトルの速度 $\dot{\mathbf{v}}$ と、連結後の大まかなリンクの方向ベクトルの速度 $(\mathbf{v}_l(j^B) - \mathbf{v}_l(i^A))$ との内積で比較することにより、動きの類似度を算出する。

$$S_{move}(i^A, j^B) = \sum_l^N \beta_l \cdot g[h(\mathbf{v}_l(j^B) - \mathbf{v}_l(i^A)) \cdot h(\dot{\mathbf{v}}_l(i^A))] \cdot g[h(\mathbf{v}_l(j^B) - \mathbf{v}_l(i^A)) \cdot h(\dot{\mathbf{v}}_l(j^B))] \quad (6)$$

ただし、

$$g[x] = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

である。

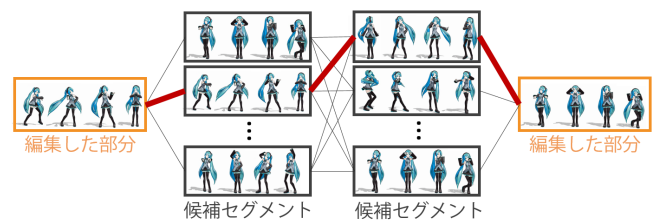


図 2. ダイクストラ法によるダンス自動生成：ダイクストラ法によるダンス自動生成。太線は遷移コストの合計が最小となる経路を表す。

最終的に連結性として、

$$S = S_{pose} + S_{move} \quad (8)$$

を算出する。

3.2.2 連結動作の生成

最後に、選ばれたセグメント列の各セグメント間の姿勢を三次関数補間で滑らかにつなげ、一連のダンスを生成する。補間の際、つなぎ目の前にあるセグメントの最終フレームと、つなぎ目の後にくるセグメントの開始フレームの間を補う処理をしてしまうと、補間をするたびに生成される動きセグメントの時間長が大きくなってしまいう問題が生じる。そこで、ここではつなぎ目の前にあるセグメントの最終 5 フレームとつなぎ目の後ろにあるセグメントの開始 5 フレームを補間で書き換えることとする。これにより、楽曲と動きの時間的なズレを生じることなく、各セグメントを滑らかにつなげることができる。

3.3 振付編集インターフェース

本手法により構築した振付編集インターフェースを図 3 に示す。1 小節ごとに区切られた図 3 下部の入力楽曲タイムラインの中で、モーションを割り当てたい小節をユーザが選択すると、図 3 左上部に複数の候補モーションが掲示される。これらの掲示される候補モーションは「3.1 モーション検索システムの構築」に記述したアルゴリズムによって選択される。ユーザは掲示された候補モーションの中から好みのモーションを選び「Set」ボタンを押すことで、選択中の楽曲部分にモーションを割り当てる。また、掲示されている候補モーションの中で気に入ったものを選択し「Search」ボタンを押すと、ユーザがどのようなモーションを気に入ったかという情報を利用してシステムが候補モーションを掲示しなおす。最後に「Generate」ボタンを押すことで、すでに割り当てられたモーションデータをもとに、入力楽曲に対する一連のダンスを自動生成できる。

4. 評価実験

4.1 候補モーション間非類似度の収束

ユーザが再検索を繰り返すことで、ダンスモーションの

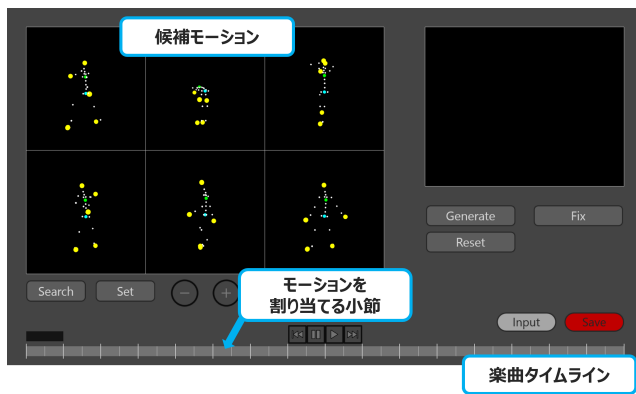


図 3. 振付編集インターフェース

探索範囲が絞られていくことを検証するために評価実験を行った。検証のために、再検索後に掲示される6つの候補モーションデータ間の平均非類似度を次式により計算した。

$$\frac{1}{15} \sum_{i=1}^5 \sum_{j=i+1}^6 D(m_i, m_j) \quad (9)$$

この値が小さくなるほど、掲示された候補モーション間の類似度は高くなり、探索範囲が絞られていることを表す。

ユーザによる再検索回数と掲示される候補モーション間の平均非類似度の関係を図4に示す。図4から、検索システム的设计意図通り、ユーザの再検索回数が増えるにしたがって掲示している候補モーション間の非類似度が徐々に小さくなり、段階的に探索範囲が絞られていることがわかる。再検索回数が3回を超えると、候補モーション間の非類似度は再び大きくなるが、これは掲示する候補モーションを決定する式(1)において、一度掲示したモーションが再び候補モーションとして掲示されないようにしているためである。

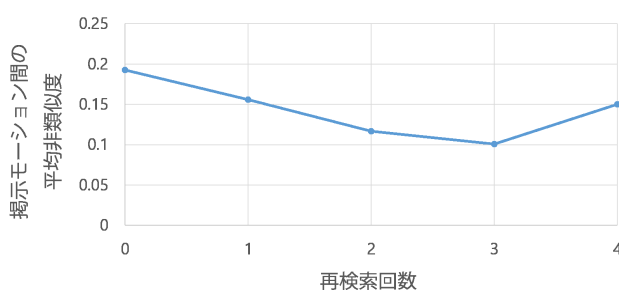


図 4. 再検索回数と探索範囲の関係

4.2 ユーザ指定部分前後の振付生成

連結性を考慮したダンス生成の有効性を検証するために、提案手法を用いて連結性の高いセグメントを4セグメントつなぎ合わせた振付生成結果と、連結性の低いセグメントを4セグメントつなぎ合わせた振付生成結果を比較した(図5)。連結性の低い生成結果では、セグメント間のつなぎ目

付近で意図せず急激な姿勢変化が起きてしまい、創作支援におけるこだわりが薄い部分の自動生成であることを考えると望ましくない。一方で、連結性の高いセグメント同

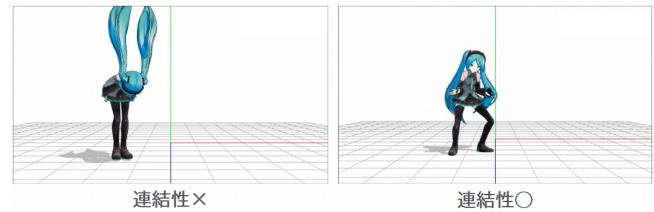


図 5. ダンス生成結果

士をつなぎ合わせた生成結果では自然なダンスモーションが生成された。このことから、連結性を考慮することの有効性を確認することができた。本手法では振付生成を行う際にセグメント間の連結性を考慮するため、自然なダンスが生成可能であるといえる。また、対話的に制作するために十分短い時間で、ユーザが指定したモーション間の振付を自動生成できることもわかった。

5. まとめと今後の課題

ユーザの好みを反映したキャラクターダンスアニメーション制作システムとそれを実現するための手法を提案し、手法の検証を行った。提案手法により構築したモーション検索システムでは、編集中の楽曲部分に対する候補モーションを複数掲示し、楽曲に合わせて候補モーションをプレビューしながら探索的に好みのモーションを検索できる。また、ユーザが選択したモーションセグメントの間のダンスの連結性を考慮してデータベース中のモーションデータで補間することで、一連の自然なダンスが生成可能となった。本手法により、簡単にユーザの好みに合う3Dダンスを制作できるようになると期待される。

今後は自動生成の中で楽曲と動きのインテンシティをマッチさせるなど楽曲特徴も考慮することで、自動生成される振付の質をより高いものにしていく。さらに、本手法で構築したインターフェースの改良を行い、ユーザ評価により有用性の検証を行う。

謝辞

本論文の図中に登場する3Dキャラクターは、ピアプロ・キャラクター・ライセンスに基づいてクリプトン・フューチャー・メディア株式会社のキャラクター「初音ミク」を使用した。またキャラクターの3DモデルはLat氏によって制作されたものである。最後に、本研究の一部はJST CREST「OngaCREST」プロジェクトの支援を受けた。

参考文献

- [1] 濱野智史 : CGM の現在と未来: 初音ミク, ニコニコ動画, ピアプロの切り拓いた世界: 5. ニコニコ動画はいかなる点で特異なのか: 「擬似同期」「N 次創作」「Fluxonomy (フラクソノミー)」, 情報処理, Vol. 53, No. 5, pp. 489–494 (2012).
- [2] Shiratori, T., Nakazawa, A. and Ikeuchi, K.: Dancing-to-Music Character Animation, *Computer Graphics Forum*, Vol. 25, No. 3, Wiley Online Library, pp. 449–458 (2006).
- [3] Ofii, F., Erzin, E., Yemez, Y. and Tekalp, A. M.: Learn2dance: Learning statistical music-to-dance mappings for choreography synthesis, *IEEE Transactions on Multimedia*, Vol. 14, No. 3, pp. 747–759 (2012).
- [4] Oore, S. and Akiyama, Y.: Learning to synthesize arm motion to music by example (2006).
- [5] Fukayama, S. and Goto, M.: Automated choreography synthesis using a Gaussian process leveraging consumer-generated dance motions, *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, ACM, p. 23 (2014).
- [6] 安永卓哉, 中澤篤志, 竹村治雄ほか : 加速度センサによるユーザコントロールを導入した音楽に合った舞踊動作の自動生成, 研究報告音声言語情報処理 (SLP), Vol. 2012, No. 25, pp. 1–6 (2012).
- [7] Mukai, T. and Kuriyama, S.: Pose-timeline for propagating motion edits, *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ACM, pp. 113–122 (2009).
- [8] Wang, P., Lau, R. W., Pan, Z., Wang, J. and Song, H.: An Eigen-based motion retrieval method for real-time animation, *Computers & Graphics*, Vol. 38, pp. 255–267 (2014).
- [9] Raptis, M., Kirovski, D. and Hoppe, H.: Real-time classification of dance gestures from skeleton animation, *Proceedings of the 2011 ACM SIGGRAPH/Eurographics symposium on computer animation*, ACM, pp. 147–156 (2011).
- [10] Choi, M. G., Yang, K., Igarashi, T., Mitani, J. and Lee, J.: Retrieval and visualization of human motion data via stick figures, *Computer Graphics Forum*, Vol. 31, No. 7, Wiley Online Library, pp. 2057–2065 (2012).
- [11] Perfume: Perfume GLOBAL SITE, <http://www.perfume-global.com/> (2012).
- [12] Rocchio, J. J.: Relevance feedback in information retrieval (1971).
- [13] Dou, Z., Hu, S., Chen, K., Song, R. and Wen, J.-R.: Multi-dimensional search result diversification, *Proceedings of the fourth ACM international conference on Web search and data mining*, ACM, pp. 475–484 (2011).
- [14] Dijkstra, E. W.: A note on two problems in connexion with graphs, *Numerische mathematik*, Vol. 1, No. 1, pp. 269–271 (1959).