

振り付けの構成要素を考慮した ダンスモーションのセグメンテーション手法の提案

岡田 成美[†] 福里 司[†] 岩本 尚也[†] 森島 繁生[‡]

本研究では、CGのダンスモーションにおけるセグメンテーション手法を提案する。ダンスの振り付けには時間幅を持つ単位が存在し、それらを次々と繋げることで一つの振り付けが構成される。よって、ダンスモーションを再利用するにあたり、適切な箇所でのセグメンテーションは不可欠である。本研究では、ダンサーへの評価実験により、ダンスモーションのセグメンテーションについて重要である要素の検討を行い、セグメンテーションルールの構築を行う。具体的には、ダンス全体のテンポ、動作の左右対称性に加え、足の着地のタイミングに着目することにより、振り付けの構成要素を考慮したセグメンテーションを実現する。

1. はじめに

近年、動画共有サイトやテレビアニメ等でCGキャラクターを用いたコンテンツを頻繁に目にするようになった。中でも3DCGキャラクターがダンスをするコンテンツは動画共有サイトにおいて再生数が高くなっている。こういったCGキャラクターを用いたダンスコンテンツ制作手段として大きく2つの手法が挙げられる。

一つ目は、クリエイターがキーフレームにおけるキャラクターの姿勢を手動で逐次調整していく手法である。この手法ではクリエイター自身の表現したい姿勢や動きをキャラクターに直接反映させることが出来る点で優れている。しかしながら、CGキャラクターのダンスモーション制作には、曲の意味合いや流れを考慮する感覚的な技術が要求されるのに加え、一連の自然なダンスモーションを作成するために多くの手間と時間が必要となる。

二つ目は、モーションキャプチャシステムを用いて、ダンサーの動きを直接取得する手法である。この手法では、感覚的な要素を含めた滑らかなダンスモーションを取得し、キャラクターに適用することが可能であり、更に、姿勢が破綻することもない。しかしキャプチャには、高価な機器や撮影環境が必要である。また、新たなダンスモーションを生成するためには逐次キャプチャが必要となるため手間とコストがかかる。

このような背景から、既存のダンスモーションデータを再利用することで制作の自動化・効率化を図る手法が提案されてきている。一連のダンスモーションデータをセグメンテーションし、それらを繋ぎ合わせることで、簡易的かつ高品質なダンスモーションを生成することができる。

CGのモーションのセグメンテーションに関する従来研究としては、動作の周期運動や関節角度に着目し

た提案が多く、ダンスモーションのような非周期かつ、振り付けの遷移といった時系列での意味合いを持つモーションに特化したセグメンテーション手法は提案されていない。また、ダンスの振り付けには時間幅を持つ単位が存在する。振り付け師がダンスを考える際にも、ポーズをつなぎ合わせるのではなく、その単位動作を次々と繋げて、一連のダンスモーションの生成を行う。よって、ダンスモーションを高いクオリティで再利用するには振り付けを構成する単位動作（ステップ、ジャンプ等）に着目し、セグメンテーションする必要がある。

そこで本稿では、既存の一連のダンスモーションに対し、振り付けの構成要素を考慮したセグメンテーション手法を提案する。ダンスモーションのセグメンテーションルールの構築のため、現役ダンサーの協力により、一連のダンスモーションに対し適切とされるセグメンテーション分割を行ってもらい、その結果を基に、ダンスのセグメンテーションにおいて重要とされる要素に着目したルールを提示し、セグメンテーション結果に対する評価を行う。

2. 関連研究

CGコンテンツの増加に伴い、既存のモーションデータを再利用する研究がなされてきた。中でも既存のモーションを用いて、新しいモーションを自動生成する研究は多くなされている[1][2][3]。これらの手法では、複数のモーションデータ間の姿勢類似を計算し、類似度の高い姿勢同士を繋ぎ合わせのポイントとすることで、姿勢が破綻することのない自然な新しいモーションの生成を可能としている。

ダンスモーションにおいても自動生成を行う研究がなされている[4][5]。これらの研究では、入力音楽に対し、それに合うようなダンスモーションをデータベースから検索し、次々と繋ぎ合わせることで、新たな振り付けの提示を行っている。このように新しい振り付けを生成するにあたり、既存の一連のダンスモーションをセグメンテーションしデータベースとして保持

[†] 早稲田大学/JST
Waseda University/JST
E-mail:narumi-pink@fuji.waseda.jp
[‡] 早稲田大学理工学術院/JST
Waseda University/JST
E-mail:shigeo@waseda.jp

しておく必要がある。

既存のモーションにおけるセグメンテーションの研究としては、複数の動作が入ったモーションデータに対し、基本動作を抽出し、クラスタリングを行い、データの再利用に繋げるものが挙げられる[6][7][8]。これらの手法は歩行動作などの繰り返しの多いモーションに対して有効な手法であるが、複雑な手足の動きで構成されるダンスモーションに応用するには不向きである。

ダンスモーションのセグメンテーションにおいては、Laban[9]の提唱する、“人間の動きの印象”を表す要素に基づくセグメンテーション手法が多く使われる。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。Nakata はユーザスタディ[11]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を示した。この結果から、モーションデータの Weight Effort を算出し、セグメンテーションを行う手法が主流となっている。白鳥ら[10]はボーンの角速度の総和を Weight Effort と定義し、身体がほぼ静止している部分でテンポを取っているという仮定の基、その大局的な極小値でセグメンテーションを行っている。

また、白鳥らは適切に舞踊動作をセグメンテーションするために、手先や重心などの複数部位の速度に着目することによって、有識者によるセグメンテーション正解値に近い検出結果を得ている[12]。しかしこの研究においては、対象としている舞踊が日本舞踊であり、日本舞踊における「とめ」と言われるポイントを検出することは出来るが、一般的なダンスにおいては、止まっている部分が振り付けの切り替えし点であるとは限らない。よって振り付けにおける時間幅を持った単位を考慮できているとは言い難い。

時間幅を持った振り付けの単位を考慮するために、Masurelle らは足の着地の部分でセグメンテーションする手法を提案した[13]。この手法はサルサダンスなどの、ステップに特徴があるダンスモーションにおいては有効であるが一般的にコンテンツに使用されるようなダンスには適さない。

更に、ダンスの振り付けは音楽のリズムに合わせて作られるといった特徴に着目し、Fan らは、音楽のビートでダンスモーションのセグメンテーションする手法[14]を提案し、Lee らは音楽の変化する部分に合わせてセグメンテーションする手法[15]を提案した。しかし、同じ音の繰り返しであっても異なる振り付けを割り当てることも多々あるため、モーションの情報を全く使用せずに正しいセグメンテーションを行うのは難

しい。

いずれの手法でも、ダンスモーションの振り付けの単位を考慮し、セグメンテーションすることは実現されていない。ダンスモーションとは、ポーズの連続ではなく、時間幅を持ったセグメントが次々と繋ぎ合されることによって生成される。そのためセグメンテーションの際も、フレーム単位の情報を使用するのではなく、振り付けの単位に沿う必要がある。また、振り付けの単位を考慮することで、再利用の際に自然な流れのダンスモーションを生成することができる。

本研究では、ダンサーへの振り付けのセグメンテーションに関する評価実験を行い、その結果から振り付けの単位を考慮するにあたり必要な要素の考察を行い、振り付けセグメンテーションのルールを構築する。

3. 評価実験

3.1 目的

ダンスモーションを適切な部分でセグメンテーションする上で、セグメンテーションルールの有無及び、専門的な知識や感覚的な要因の検討を行うため、ダンサーへの評価実験を行った。

3.2 使用したデータ

実験データとして、音楽データと同期されたダンスモーションを使用した[17]。実験を行う際には、BVHファイルのボーンデータを3D キャラクタに対しスキニングを施したものを使用した。ただし BVH ファイルのボーンデータは、図1のように28点の関節を持つモデルである。

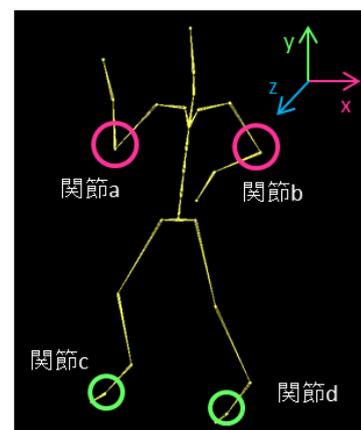


図1 使用したボーンモデル(関節数 28)

3.3 実験方法

音楽と同期してキャラクターがダンスを踊っている動画をダンサーに提示し、振り付けを区切るべきだと考えられる部分を回答してもらう。ただし以下2つの条件を設ける。

- 振り付けを考える際には音楽のテンポに合わせて8カウント(2小節分)を数えるという知見に基づき、音楽のテンポに合わせたカウントで分割をしよう。
- ダンサーの感覚と振り付けとしてどこで区切るのかで回答が異なることが予想される。そこで使用目的を説明した上で、振り付けとして区切るにあたり適切と思うカウントを回答してもらおう。

表1 ダンサーへの評価実験結果(全候補点:112個)

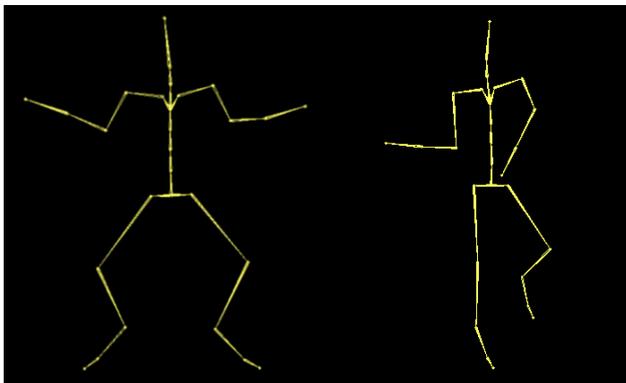
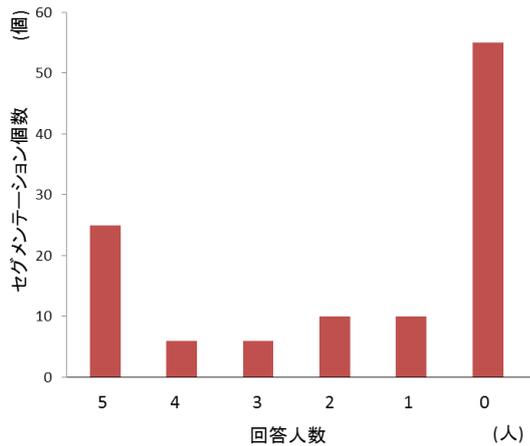


図2 左右対称の動作(左)と左右非対称の動作(右)

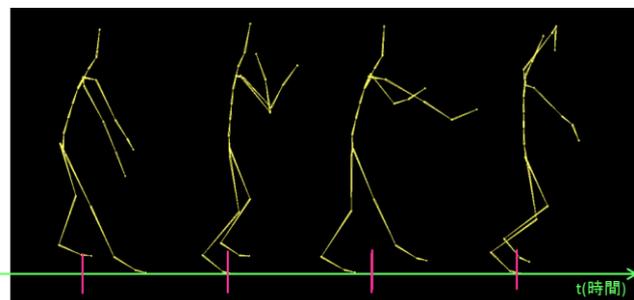


図3 足の着地のタイミング
同一セグメント内で見受けられた足の着地のタイミングが等間隔になっている例

3.4 評価実験結果

本実験は、振り付けについて知識を持つダンス歴9年以上の現役ダンサー5名に対して行った。実験結果を表1に示す。この結果から、112個の候補点のうち、全員が共通して回答するカウント、回答しないカウントが多数存在したため、振り付けに着目したセグメンテーションルールは存在するということが確認できた。回答人数が1~4人の箇所があることから個人差が出たことが分かるが、その理由としては使用したダンスモーションが複雑に作りこまれたものであったことと、個人の感覚の差が挙げられる。

これらの実験結果を踏まえ、ダンスモーションのセグメンテーションにおいてどのような要素が重要なかを考察する。

3.5 セグメンテーションにおける重要な要素の抽出

全員がセグメンテーションを行った箇所に着目すると、最も長いセグメントは音楽における8カウント分にあたり、更に8カウント目ごとに必ずセグメンテーション箇所があることがわかった。よって音楽のテンポの情報的重要であるということが言える。

また、振り付けに注目すると図2のように左右対称の動作を行っているセグメントと、左右で異なる動作を行っているセグメントが存在していることが分かった。よって左右対称の動作から左右非対称の動作(左右非対称の動作から左右対称の動作)に変化するタイミングでセグメンテーションが行われると考えられる。

更に、足の着地のタイミングに着目すると、図3のように同一セグメント内では足の動きが一定のリズムを刻んで着地を行っている傾向が見られた。よって、同一セグメントであることを判断するにあたり、足の着地の情報を考慮することが有効であると考えられる。

これらの考察を元に本稿では、音楽のテンポ・動作の対称性・足の着地に着目し、セグメンテーションされるべきカウントを自動で取得する方法を検証する。

4. 提案手法

評価実験によって決定した要素に基づき、セグメンテーションのルールの構築を行う。本研究では、音楽のテンポによるセグメンテーション、振り付けの左右対称性によるセグメンテーション候補点検出、足の着地のタイミングによる同一セグメント検出を提案する。

4.1 音楽のテンポによるセグメンテーション

評価実験から音楽における全ての8カウント目でセグメンテーションが行われるということが分かった。そこで、踊り出しのフレームを手動で決定したのち、

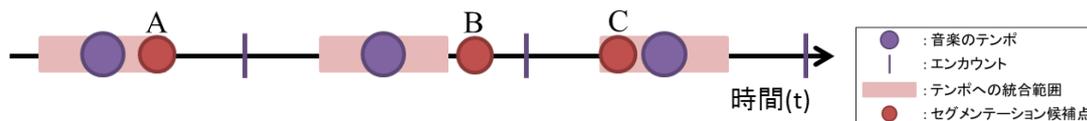


図4 候補点の絞り込み(例)

候補点がテンポへの統合範囲内に存在する場合、候補点の位置がテンポに修正される。この例では候補点 A, B, C のうち、A, C はテンポへの統合が行われる。候補点 B は左右対称性から求めたものであれば消去される。足の着地から求めたものであればこのフレーム数を保持する。

楽曲のテンポから 8 カウントにあたるモーションのフレーム番号を抜きだし、全てをセグメンテーション点とする。ただし今回はテンポが既知の楽曲を使用した。

4.2 左右対称性によるセグメンテーション候補点検出

左右対称の動作から左右非対称の動作(左右非対称の動作から左右対称の動作)への変化を検出する。その変化する瞬間のフレーム番号を求め、セグメンテーション候補点とする。

4.2.1 左右対称性の判断

身体の対称性を考慮するにあたり肘の関節(図 1 の関節 a, b)の角度に着目した。足は足踏みを行うなど、フレーム単位で左右対称な動作を行っていないことが多い。一方、肘は左右対称の動作が多いことに加え、ひねりを除けば関節の回転自由度が 1 であることから、左右対称性の判定が容易である。

判断方法としては、左右の肘の角度をフレーム毎に算出し、それらの差が閾値以下であれば左右対称、閾値以上であれば左右非対称と判断する。ただし左右対称(左右非対称)と判断されたフレーム数が α より少なければ誤差とみなし、左右非対称(左右対称)とみなす。以上の操作より求めた判断の切り替わりのフレーム番号をセグメンテーション候補点とする。今回は経験的に $\alpha=7$ とした。

4.2.2 候補点の絞り込み

左右対称性を考慮するだけでは候補点が絞り切れていないことが分かった。そのため、セグメンテーションのタイミングは音楽のテンポと同期しているという条件を基に候補点の絞り込みを行う。方法としては[16]を基に、音楽のテンポの前後一定時間内 β に候補点が存在するかを認識し、存在する場合はそのテンポのフレーム番号をセグメンテーション候補点とし、存在しない場合は候補から外す処理を行う。カウントとカウントの中間のエンカウントと呼ばれる部分が候補とされている場合に判断が可能であるよう、 β はテンポの $1/4$ の長さとする。図 4 に絞り込みの模式図を示す。

4.3 足の着地による同一セグメントの検出

評価実験から、同一セグメント内では足の動きが一定のリズムを刻んで着地を行っているという傾向が見て取れた。よって、ここでは足の着地のタイミングに着目し、タイミングが一定であった場合は同一セグメントであるという判断を行う。

4.3.1 足の着地の判断

足の着地のタイミングに着目すると、同一セグメント内では足の動きが一定のリズムを刻んで着地を行っているという傾向が見て取れた。よって、ここでは足の着地のタイミングに着目し、タイミングが一定であった場合は同一セグメントであるという判断を行う。

足の着地部分を検出するにあたり、足の関節は着地をしている間は速度が 0 に近くなる[16]という性質を用い、爪先(図 1 の関節 c, d)の速度を算出する。爪先を着地させたまま踵を浮かせるような動きも見られるため、そのような部分も着地と判断されるよう爪先の関節を使用した。単純に速度のみを使用すると着地と判断されるべき部分で速度が安定しないということが見て取れたため、以下のようなルールを設けた。

- ・着地の瞬間：速度が閾値以下になった瞬間
- ・離陸の瞬間(足が離れる瞬間)：加速度が閾値以上になった瞬間
- ・離陸と判断されていても、y 座標が閾値以上の値を持たなければ着地とみなす

ダンスの振付においては、足が着地する瞬間をテンポと同期させることが殆どであるため、着地の瞬間のフレーム番号を取り出す。

4.3.2 候補点の絞り込み

足の着地間のタイミングを取得するために、4.2.2 章と同様に、検出されたフレーム番号が音楽のテンポの前後 β に存在するかを認識し、存在する場合は音楽のテンポに統合する処理を行った。ただし、存在しない場合は取得したフレーム番号をそのまま用いる。その後、着地を判断されたフレーム番号全てにおいて、前後の間隔を算出する。前後の間隔が等しければ、その

カウントの部分は 4.2 章でセグメンテーション候補点と判断されていたとしても候補点から外し、同一セグメントと判断する。

表 2 テンポと左右対称性による結果

候補点検出手法	Recall[%]	Precision[%]
音楽のテンポ	100	60.0
左右対称性	40.0	40.0

表 3 本手法と既存手法の比較

セグメンテーション 手法	Recall[%]	Precision[%]
白鳥ら(2007)	35.3	72.0
本手法	55.6	80.0

5. 実験結果と既存手法との比較

5.1 結果

音楽のテンポ、左右対称性におけるセグメンテーション候補点検出結果を表 2 に示す。評価実験より、音楽の 8 カウントでは必ずセグメンテーションされるため、音楽のテンポを用いた結果 Recall は 100%であるが、Precision の値は十分ではないと言える。

足の着地のタイミングにより候補点を修正する過程では、同一セグメントの部分のみを検出することが出来た。足の着地の情報のみでは、セグメンテーションされるべき部分が誤って検出される箇所も存在したが、音楽のテンポ情報を優先することにより改善された。

5.2 既存手法との比較

本手法の有用性を検証するために、一連のダンスモーションをセグメンテーションし、精度比較を行った。ここで使用する正解値は、評価実験の結果、被験者の回答が全て一致した 25 点とする。比較手法として、白鳥らの Weight Effort の極小値を用いて [4]セグメンテーションした結果を用いる。ただし、セグメンテーション候補点は音楽のテンポと同期しているため、本手法と比較するにあたり、白鳥らの手法で検出されたフレーム番号を音楽のテンポに変換する。それぞれの手法により取得したセグメンテーション結果を表 3 に示す。

表 3 より、Recall・Precision 共に白鳥らの手法に比べ、本手法の方が高い値を得ることが出来る。ただし Recall は 55.6%となっており、誤検出が多いことが分かる。また、本手法では音楽のテンポの情報を利用し、8 カウントごとに必ずセグメンテーションを行うというルールを課したが、白鳥らの手法では音楽の情報を使用することなく、8 カウントの部分全てを検出する

ことができていることが分かった。

足の着地のタイミングにより、セグメンテーション候補点を修正する過程では、本手法においては 1 箇所しか修正されなかった。それに対し、白鳥らの手法に、本手法の着地のタイミングによる修正過程を加えてみた結果、8 箇所修正することができた。よって、ダンスモーションのセグメンテーションを行うにあたり、足の着地のタイミングに着目することは有効的であると考えられる。ただし、白鳥らの結果に着地のタイミングによる修正を加えた結果、Recall=41.9%となっており、本手法の有用性が確認できる。

また、本手法により誤検出された部分について検証を行った。その結果、正解値のカウントの一つ後が検出されている部分が 3 箇所、評価実験で 4 人がセグメンテーションしていた部分が 1 箇所存在した。この部分ではセグメンテーションが行われても、違和感はないが、セグメンテーションが 1 カウントずつ行われることは殆どないと考えられる。よって、連続したカウントが検出された場合にはどちらかを選択するといったことが必要だと感じる。また、その他 6 箇所の誤検出は評価実験で 0~2 人がセグメンテーションされるべきと答えた部分であり、個人の感覚的な違いが出た箇所であった。

6. まとめと今後の課題

本稿では、ダンサーへの評価実験による感覚的な知見を基に、CG のダンスモーションに対し、振り付けの単位を考慮したセグメンテーションを行う手法を提案した。音楽のテンポ、腕の左右対称性、足の着地のタイミングに着目することにより、既存のセグメンテーション手法よりも高い検出率を得た。

ただし、音楽のテンポが既知でないと、8 カウントごとにセグメンテーションを行うことができない。しかし Weight Effort の極小値を取得することにより 2 小節ごとのフレーム番号は取得出来ることが分かった。よって、Weight Effort を用いてモーションの情報からテンポの推定が出来れば、この問題を解決できると考えられる。また本手法は、サビの前等に 8 カウントの間に 4 カウントを刻むような楽曲も多々存在するが、そのような音楽に対するダンスには適用が出来ない。ダンスのジャンルや振り付けが大きく異なるダンスモーションにも適用していきたい。

最後に、インターネット上に存在するダンスモーションを適切にセグメンテーションし、それらを再利用することで新しいダンスモーションを自動生成するようなシステムを目指したい。

参考文献

- [1] Kovar, L., Gleicher, M., and Pighin, F. Motion graphs. *ACM Transaction on Graphics*, 21, 3, 473–482, 2002.
- [2] Arikan, O and Forsyth, D. Interactive Motion Generation from Examples. *ACM Transaction on Graphics*, 21, 3, 483–490, 2002.
- [3] Beaudoin, P., Coros, S., Panne, M., and Poilin, P. Motion-motif graphs. In: *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 117-126, 2008.
- [4] Shiratori, T., Nakazawa, A., and Ikeuchi, K. Dancing-to-Music Character Animation. *Computer Graphics Forum*, 25, No. 3, 449-458, 2006.
- [5] Kim, T.H., Park, S.I., and Shin, S.Y. Rhythmic Motion synthesis based on motion-beat analysis. *ACM Transaction on Graphics*, 22, 3, 392-401, 2003.
- [6] Wang, P., Lau, Rynson W.H., Pan, Z., Wang, J., and Song, H. An Eigen-based motion retrieval method for real-time animation. *Computer&Graphics*, 38, 255-267, 2014.
- [7] Vögele, A., Krüger, B. and Klein, R. Efficient unsupervised Temporal Segmentation of Human Motion. In: *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2014.
- [8] 大崎竜太, 嶋田光臣, 上原邦昭, 速度に基づく切り出しとクラスタリングによる基本動作の抽出. *人工知能誌*, 15, 5, 878-886, 2000.
- [9] Laban, R., and Ullmann, L. The Mastery of Movement. *Pre Textos*, 1971.
- [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*, 2242-2252, 2007
- [11] Nakata, T., Mori, T., and Sato, T., Analysis of Impression of Robot Bodily Expression. *The Journal of Robotics and Mechatronics*, 14, 1, 27–36, 2002.
- [12] 白鳥貴亮, 中澤篤志, 池内克史. 音楽のリズムを考慮した舞踊動作の解析. *電子情報処理学会情報誌*, 2242-2252, 2007.
- [13] Masurelle, A., Essid, S., and Richard, G. Multimodal Classification of Dance Movements using Body Joint Trajectories and Step Sounds. In *Proc. 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services (WIAMIS)*, 1-4, 2013.
- [14] Fan, R., Xu, S., and Geng, W. Example-Based Automatic Music-Driven Conventional Dance Motion Synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 18, 3, 501-515, 2012.
- [15] Lee, M., Lee, K., and Park, J. Music similarity-based approach to generating dance motion sequence. *Multimedia Tools and Applications*, 62, 3, 895-912, 2013.
- [16] Shiratori, T., Nakazawa, A., and Ikeuchi, K. Detecting Dance Motion Structure through Music Analysis. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004)*, 857-862, 2004.
- [17] <http://perfume-global.com/visualization.html>