

Kinect を用いた簡易 3D アニメーション制作システムの提案

古川 真行¹ 赤木 康宏¹ 河合 由起子² 川崎 洋¹

概要：映画等のような高品質な 3 次元アニメーション制作に用いられる機器は、高価であるばかりではなく利用に専門的な技能を要し、導入コストが高い。一方、教育や観光産業等の分野では、小規模な団体および個人がアニメーションを制作したいという需要が増加しており、簡単かつ安価なコンテンツ制作・編集システムが求められている。そこで、本研究では Kinect を用いた 3 次元アニメーション制作システムを提案する。本システムを実現する方法として、各モーション間の関連性を自動的に構造化する手法、および、モーション間のつなぎ目を滑らかに表示するための 3 次元形状補間手法を新たに提案する。また、制作した 3 次元アニメーションをタブレットデバイスを用いてインタラクティブに操作・閲覧することのできる既存のシステムと統合することで、制作から閲覧までを行なえる一貫したシステムを実現する。

3D Animation Creation System based on Auto-structurization Method

FURUKAWA MASAYUKI¹ AKAGI YASUHIRO¹ KAWAI YUKIKO² KAWASAKI HIROSHI¹

Abstract: Contents creation systems for create 3D character animations in movies and games require not only expensive but also those require professional skills. On the other hand, there are increasing demands for contents creation systems that are low-cost and easy to handle for the general users of small-scale groups and persons in the area of education, advertisement and tourism, especially to create human motion animations. To answer the demands, we propose a contents creation system which enable to synthesize multiple 3D motions of a human captured by Kinect. To realize the system, we propose an interpolation method which corrects gaps between 3D shapes and skeleton models of the input 3D motions to create transition motions between the inputs. Moreover, we also extend an existing interactive animation viewer to be able to show the proposed 3D animation contents on tablet devices.

1. はじめに

3 次元キャラクタを用いたアニメーション表現は映画やゲームのようなエンターテインメント分野ばかりではなく、航空機のインストラクション、運転者への安全教育、博物館のガイダンス、および電子マニュアル等の教育・案内用のコンテンツとしても幅広く利用されている。前者の映画制作等を目的とする場合では、膨大なコストをかけてでも、品質の高い映像が制作できることが第一目標であるが、後者のような教育・案内用コンテンツでは分かりやすい表現の映像を低コストに作ることが重視されている。さらに、

教育・案内用コンテンツでは、教育効果を高める目的で、受け手の要求や状況に応じて内容を柔軟に変更する必要もあり、コンテンツにカスタマイズ性やインタラクティブ性が求められるという特徴もある。

しかし、一般的な 3 次元キャラクタアニメーションの制作では、(1) 3 次元キャラクタモデルを制作する、および (2) モーションキャプチャシステム等を用いてキャラクタに動作を与える、という 2 つの工程が必要であり、どちらも専門的な知識および機材が必要となるので、3 次元アニメーションの制作には多くの金銭的、時間的コストが必要となる。よって、教育効果を高めるために対象者に合わせてカスタマイズした 3 次元アニメーションを提供したいという需要はあるものの、小規模な団体や個人が既存のシステムを用いてこのようなコンテンツを制作することは困難である。

そこで本研究では、人間の身体動作を対象とする 3 次元

¹ 鹿児島大学大学院理工学研究科, 鹿児島市
Graduate School of Science and Engineering, Kagoshima University

² 京都産業大学大学院先端情報学研究科, 京都市
Graduate School of Division of Frontier Informatics, Kyoto Sangyo University

さらに、提案システムにより生成した3次元アニメーションコンテンツを、先行研究において著者らが提案した、タブレットを用いた3次元アニメーション表示システムと組み合わせることで、インタラクティブ性のある3次元アニメーションを利用できる環境を実装する。以上により、前述した教育や観光、小規模小売店等でのアニメーション制作需要に適したアニメーション制作環境を実現する。

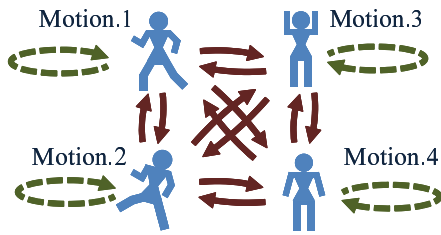
人間や自然物の動きには、例えば歩行動作や樹木の揺れなどのように、ある運動を一定の周期で繰り返しつつ、他の周期性のある運動へと滑らかに遷移するという特徴がある。この性質を利用して、ビデオなどから自然物の規則性を見出し、新たな映像の再構成に利用する研究がある。3次元キャラクタアニメーションを制作するために、人間の動作の周期性および遷移性に着目した研究として、Motion Graph と呼ばれる手法がある [1]。Motion Graph は、モーションキャプチャ等により取得した人間の骨格の動きを解析することで、入力モーション中に含まれるループを検出したり、複数のモーション間に含まれる共通の動作を見出すことで、それらの動作間の滑らかな遷移動作を生成するための手法である。Motion Graph を用いることで、アニメーションの編集者は個別に取得したモーションデータから、任意の組み合わせのアニメーションを制作することもできる。Motion Graph を応用した研究としては、いくつかのキーフレームを与えるだけで、既存のモーションデータから未知の動作を生成する研究 [2], [3] がある。また、Motion Graph の生成を行なうための基礎技術として、モーションデータ内で何度も繰り返される動作を発見し、基本構成要素（プリミティブ）として抽出する研究 [4], [5] なども提案されている。

(1) 計測 (2) 構造化および編集 (3) 3次元アニメーションのインタラクティブな操作

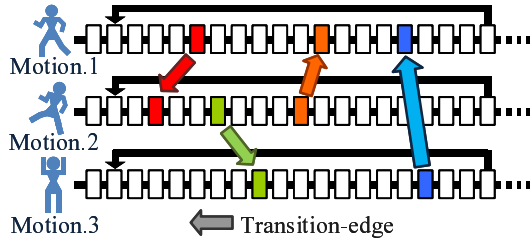
図 1 システムおよびモーションデータの構造.

以上のように，アニメーションを制作する際に，キャラクターの動作を任意回繰り返したり，複数の動作を任意に組み合わせたりするための手法は数多く提案されている．本研究においては，3次元キャラクターをアニメーションの中心としているので，モーションキャプチャにより得られた動作データを基に，Motion Graph を構築することで，ユーザが複数の動作を任意に組み合わせることのできるシステムを実現する．その際，提案システムでは一般的な Motion Graph で用いられる 3次元キャラクターモデルと骨格モデルの組み合わせではなく，Kinect により取得した 3次元点群と推定された骨格モデルを用いるので，3次元形状と骨格データ間に誤差が生じるという問題がある．本研究では，既存の Motion Graph の考えに基づきながら，簡易なアニメーション制作システムを実現するために，この 3次元形状と骨格データ間の誤差を解消する手法を提案する．

提案手法の概要を図1に示す。まず，Kinect で人間の動きを計測することで入力データを取得する（図1(a)-(1)）。Kinect を用いる理由としては，安価な3次元形状計測装置であるとともに，人体の骨格を認識し，モーションデータとして取得する機能を有しているので，今回想定する用途に適しているからである。この時取得されるデータは，人間の3次元形状および骨格の時系列データ（skeletal motion）である（図1(b)）。次に，複数の skeletal motion を解析することで，コンテンツ制作者による3次元アニメーションの編集を容易にするための核となる，Motion Graph を構築する（図1(a)-(2)）。その後，制作者は本システムを用い



(a) 3次元アニメーションの Motion Graph 構造. 矢印実線が Transition-edge, 矢印破線が Self-loop を表す.



(b) 複数の 3次元アニメーションをつなぐ Transition-edge.

図 2 状態遷移を伴うアニメーションの構成例

て複数の skeletal motion を任意に組み合わせることで、3次元アニメーションコンテンツを制作する．最後に、制作したアニメーションをネットワーク等を通じて閲覧用のシステムをもつエンドユーザに配信することで、3次元アニメーションをインタラクティブに操作しながら閲覧することができる (図 1(a)-(3))．以下に、Motion Graph の構造とインタラクティブな閲覧システムの概要を記述する．

3.1 Motion Graph による人体動作の表現

複数の skeletal motion から Motion Graph を生成するにあたり、1つの skeletal motion をノードと呼び、そのノード内にある繰り返し部分を Self-loop と呼ぶ．また、ノード間の遷移 (動作の切り替え) をエッジと呼び、特に具体的な skeletal motion 間の接続位置を Transition-edge と呼ぶ．図 2 は、skeletal motion の各フレームと、Self-loop および Transition-edge の関係性を表しており、ノードおよびエッジは複数フレームからなる時系列な骨格の集合で表される．提案システムにおける Motion Graph の構造化手法の詳細は 4.1 節および 4.2 節に記述する．

また、Kinect を用いた計測ではマーカーを用いずに、3次元形状から骨格を推定する方法をとっているため、形状と骨格の間に推定誤差が発生する．これが、Self-loop および Transition-edge を表示する際に形状のずれとして見えてしまうので、skeletal motion と 3次元点群を用いた補間手法を提案する．本補間手法の詳細は 4.4 節で述べる．

3.2 インタラクティブ 3次元アニメーション閲覧システム

著者らは先行研究として、専門の設備を持たない一般ユーザの 3次元アニメーション制作を支援するため、Motion Graph のような状態遷移グラフをもった 3次元アニメーションをタブレット端末上でインタラクティブに再生する

手法を提案している [9], [10]．しかし、この従来の手法では、手動で状態遷移グラフを編集する必要があった．そこで、提案手法のもつ状態遷移グラフの自動的な構造化手法を、この閲覧システムと統合することで、3次元アニメーションの制作から閲覧までを統一的行なうことのできる環境を実装する．

4. Motion Graph に基づいた 3次元アニメーション制作システム

本章では、Kinect により計測した複数の skeletal motion から Motion Graph を生成する方法と、Motion Graph 内で定義される Self-loop および Transition-edge を実際に再生する際に必要となる、繋ぎ目部分に相当するモーショ (本論文では、これを transition motion と呼ぶ) を補間する手法について述べる．まず、4.1 節において、Motion Graph を生成する際に必要となる、skeletal motion 内に含まれる短時間の骨格の動き同士の類似度を比較するための、動作間距離を定義する．

次に、4.2 節において、動作間距離の算出法に基づき、Motion Graph を構築する手法について説明する．Motion Graph を構築する際は、この動作間距離に基づき 2つの skeletal motion 間の距離が最も小さい箇所に Transition-edge を生成する．こうして生成した Motion Graph に従い、Self-loop および Transition-edge の transition motion をそのまま表示するだけでは、類似度が高いフレーム間で生成されたものであっても、わずかな形状のずれが原因で、不自然なアニメーションとなる．この問題を解決するため、skeletal motion 間の transition motion を補間することで、より自然な動きを提示する手法を提案する (4.3 節)．さらに、提案手法では Kinect から同時に取得した 3次元点群にも同様にずれの問題が生じるので、3次元点群に適した補間処理を併せて提案する (4.4 節)．

4.1 skeletal motion 間の距離の算出

本節では、skeletal motion から類似する動作を発見するための動作間距離の計算方法について述べる．Kruger らによって、人の姿勢の類似性を比較するための手法が考察されている [11]．この研究では、人の身体の特徴点を $\mathcal{F}_E^{n \times d}$ と表現しており、 n は特徴点の数、 d は特徴点の次元数を表している．本研究では、Kinect による推定で得られた人 20 箇所の 3次元関節位置を $\mathcal{F}_E^{20 \times 3}$ の特徴点として表現し、式 1 により姿勢間の距離を求める．

$$DistPose(i, j) = \sum_{k=1}^{20} \|\mathcal{F}_{E,j}^k - \mathcal{F}_{E,i}^k\| \quad (1)$$

ここで、 i と j は 2つのモーションの時刻 (フレーム番号) を表している．ただし、式 1 を繰り返し動作の検出に

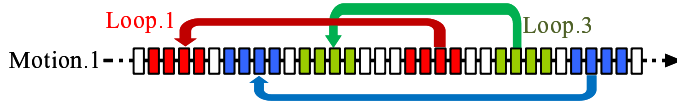


図 3 Self-loop の検出手法.

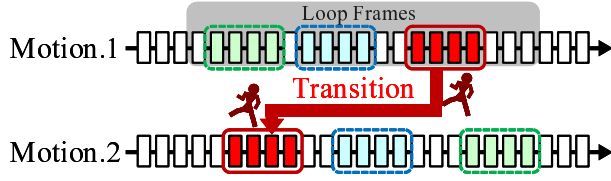


図 4 skeletal motion 間の Transition-edge の検出手法.

使用した場合、歩く動作のような振り子動作では、半周期ずれた位置に運動方向が異なるが同じ姿勢があり、間違っただけの遷移を生成する可能性がある。そこで、次のようにして複数フレームの姿勢間の距離（動作間距離）を算出する。

$$DistMotion(p, q) = \sum_{i=1-N}^N \frac{DistPose(p+i, q+i)}{2N} \quad (2)$$

ここで、 p と q は 2 つのフレーム番号、 N はウィンドウ幅を表している。式 2 を用いることで、動きの方向性を考慮した動作間距離を求めることができる。

4.2 Self-loop および Transition-edge の検出

図 3, 4 は Motion Graph の Self-loop と Transition-edge の検出方法を示したものである。式 2 を用いて、skeletal motion 内の時刻 (p, q) を基準とした、全ての 2 つのフレーム群の組み合わせの間の距離を算出する。Self-loop の検出の場合は、より短い Self-loop がより良い候補と考え、次の重み係数を掛けて距離を再計算する。

$$W(i, j) = \left| \frac{i - j}{FrameSize} \right|^2 \quad (3)$$

ここで、 $FrameSize$ は入力された skeletal motion のフレーム総数を表している。最終的に、フレーム (p, q) を基準として検出した最小の $W(p, q) * DistMotion(p, q)$ を Self-loop が生成されるフレーム組とする（図 3）。

Self-loop を検出した後に、全ての skeletal motion の組み合わせにおいて動作間距離を算出し、各組で距離が最小のフレーム間に Transition-edge を生成する（図 4）。この際、類似する動作がほとんど含まれていない入力モーションの組がある可能性を考慮し、動作間距離が閾値より大きい場合は、Transition-edge は生成されないものとする。

4.3 transition motion 中の skeletal motion の補間

前節で述べたように、Self-loop および Transition-edge で接続される骨格の組は完全に同じではないため、単純に連続的に表示した場合、不自然な動きとなる場合がある。そこで、transition motion に用いられる遷移前および遷移

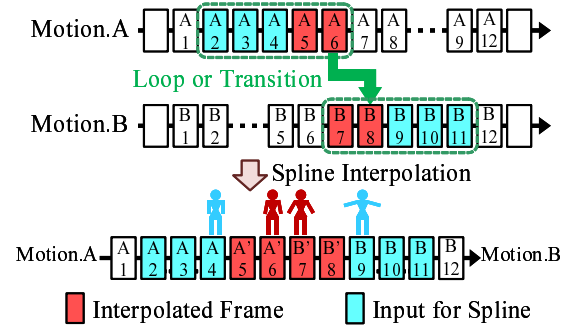


図 5 transition motion の補間手法.

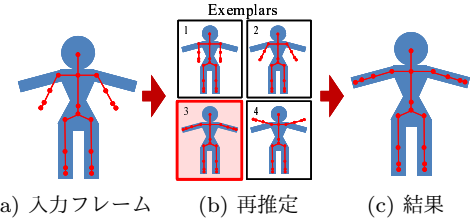


図 6 推定が低精度な骨格の exemplar based の再推定.

後の n フレーム分の skeletal motion に対して補間処理を行なう。補間処理では、Self-loop および Transition-edge の繋ぎ目のフレームの組 $(p, q + 1)$ を中心に前後 5 フレーム分の骨格の 3 次元位置にスプライン補間を適用する。本論文では、図 5 で示すように、 $n = 2$ として定義する。したがって、図 5 において、 $(A2, A3, A4, B9, B10, B11)$ がスプライン補間の入力として与えられ、本来の skeletal motion $(A5, A6, B7, B8)$ は、補間された skeletal motion $(A5', A6', B7', B8')$ に置き換えられる。

4.4 transition motion 中の 3 次元点群の補間

Kinect を用いることで、時系列に連続した 3 次元点群と、この点群から推定される時系列に連続した skeletal motion を同時に得ることができる。しかし、skeletal motion が不正確に推定されることにより、3 次元点群と骨格の間に誤差が生じる（図 6(a)）。この誤差を含んだフレームが Self-loop や Transition-edge になった場合、生成されるアニメーションに形状のずれが生じ、違和感を生じさせてしまう。そこで、この問題を解決するため、次のような手法を提案する。

- (1) skeletal motion の推定精度が低いフレームを検出。
- (2) 精度の十分な skeletal motion を Exemplar としての骨格再推定を行なう。
- (3) (2) を適用したフレームで、閾値より推定精度が低いものを Self-loop および Transition-edge の候補から除外する。

以上の手順により、骨格の推定精度が低いフレームを除外し、モーション間に滑らかでない遷移が生成されることを防ぐ。手順 (1) では、の推定精度の低いフレームを検出するために、点群内の各点と、それに最も近い位置の関節との距離を平均した値を算出し、閾値以上の距離をもつ

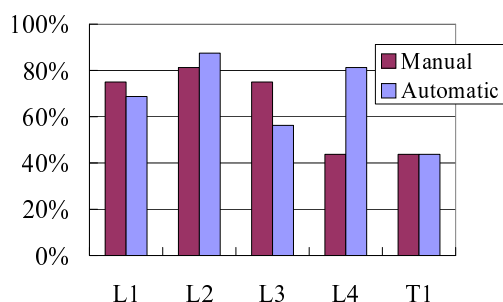


図 7 手動および自動で構築された Motion Graph のアニメーションの自然さの評価.

フレームを推定精度が低いフレームとする. この閾値は, Kinect を用いた予備実験により, 全体の 10% が修正候補となるような値として $0.1m$ と定めた. 次に, 手順 (2) では, 推定精度が低いと判定されたフレームに対して骨格の再推定を行なう. この手順では, 骨格が多くの反復動作を行なっていることに着目し (図 6 (b)), 全ての骨格を Exemplar とし, 骨格と修正対象の 3 次元点群との距離が最小となる骨格を探索して, 置き換える処理を行なう (図 6 (c)). 手順 (2) を適用したフレームの骨格と点群間の距離が $0.1m$ を超えていた場合, そのフレームを Self-loop および Transition-edge 生成の候補から除外する. Kinect の場合, 全体の 50% のフレームが候補として有効であることが実験により判明した. 骨格の再推定を行なった後, Self-loop および Transition-edge の検出を行い, skeletal motion が補間された繋ぎ目のフレームの組 $(p, q + 1)$ の前後 n フレームに対して, Sumner らの手法を用いて, 骨格を基準とした 3 次元点群の変形を行なう [12].

5. 実験および評価

本章では, 体操を指導するコンテンツ作成を想定した 3 次元アニメーション制作を行い, 提案手法の評価を行なう. Kinect を用いて計測した運動を以下に示す.

- 腕を上には伸ばす運動 (Motion.1)
- 腕を広げながら膝を曲げ伸ばす運動 (Motion.2)
- 腕を左右に回す運動 (Motion.3)
- 深呼吸する運動 (Motion.4)

次に, Self-loop および Transition-edge の補間手法の有効性の評価を行なうために, 補間手法を適用していないものと, 適用しているものを第三者に提示することで主観評価を行なう (5.2 節). 最後に, 直感的かつ容易な 3 次元アニメーション操作が可能なシステムを実装する (5.3 節).

5.1 Motion Graph の自動構築精度の評価

まず, 提案手法によって自動生成した Motion Graph に基づくアニメーションと, 手動で制作した Motion Graph に基づくものの品質および制作時間の違いを比較する. 比較例として, Self-loop (L1,L2,L3,L4) と Transition-edge

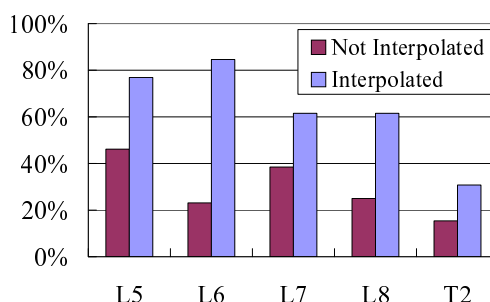
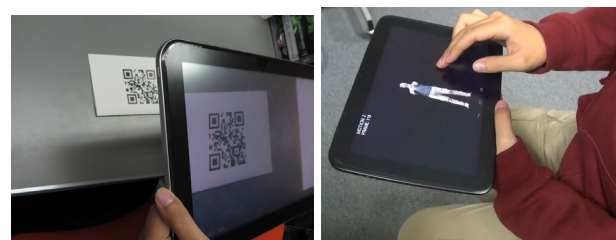


図 8 補間適用後および未適用のアニメーションの自然さの評価.



(a) 閲覧システムでの QR コードの認識.

(b) 3 次元アニメーションの操作.

図 9 タブレット端末での操作.

(T1) の transition motion の 5 種の動作について, 手動で作成した場合と自動で作成した場合の合計 10 種類のムービーを用意し, 手動か自動のどちらで作成されたかを伏せて 16 名の評価者に視聴させ, それぞれの動作を自然に感じたかどうかのアンケートを実施した. 結果は図 7 で示すように, L4 の例を除いて, 提案手法が手動によるものとほぼ同等の評価であった. これは, 自動で生成した Motion Graph が手動で制作した場合と同等に利用可能であることを示している. 自動生成により評価値が大きく向上した L4 に関しては, 制作者の主観による Motion Graph の設計が, 評価者の主観に沿ったものではなかったことが原因として考えられる. また, 手動および自動による Motion Graph の構築にかかる時間を比較したところ, 手動の場合は 19 分であり, 自動の場合は 8 分であった. 提案手法では補間処理の際に入力の 3 次元形状を全て再評価するという処理を行っているため, この処理の高速化が今後の課題である.

5.2 transition motion の補間手法の評価

次に, transition motion 補間手法の有効性についての評価を行なう. 4 種類の Self-loop (L5,L6,L7,L8) と 1 種類の Transition-edge (T2) について, それぞれ補間手法を適用したアニメーションと適用していないアニメーションの合計 10 種類を作成し, 13 名の評価者に補間手法を適用しているかどうかの情報を与えずにそれぞれを自然と感じたかどうかのアンケートを実施して評価する. 結果は図 8 で示すように, 全てのアニメーションにおいて補間手法適用後の方が評価値が高くなり, 補間処理により自然な transition motion が生成できているといえる.

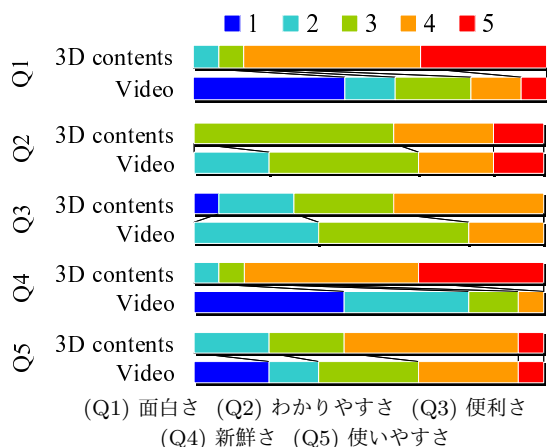


図 10 2次元ビデオおよび3次元アニメーションの評価値。1から順に不満、やや不満、普通、満足、大いに満足。

5.3 タブレット端末での3次元アニメーション閲覧システム

最後に、インタラクティブな3次元アニメーション閲覧システムについての評価を行なう。図9に実装した閲覧システムを用いてQRコードを読み取り、ネットワークを通じて入手した3次元アニメーションをタブレット端末上で操作している様子を示す。実験では、14名の被験者に3次元アニメーションを利用してもらい、同じコンテンツを扱った従来の2次元ビデオと比較した後に、図10に示す5つの評価項目による5段階評価のアンケート調査を行なう。図10に示す結果から、閲覧者は従来の2次元ビデオよりも3次元アニメーションに面白みを感じたことが示されている。また、新鮮さや使いやすさについても2次元ビデオより高水準であったといえる。さらに、わかりやすさ、便利さについても十分な水準の結果が得られた。したがって、インタラクティブな3次元アニメーション閲覧システムは、閲覧者にとって、従来の2次元ビデオよりも新鮮で魅力的なコンテンツであるといえる。

6. 結論

本論文では、Motion Graphの自動構造化によるインタラクティブな3次元アニメーション制作システムおよび3次元点群のモーションの効果的な補間手法を提案した。Motion Graphの自動構造化では、skeletal motion間の動作類似度の算出法を提案することで、手作業による制作と同等のMotion Graphを生成することに成功した。また、3次元点群に適したフレーム間の形状補間手法を提案することで、アニメーションの切り替え時に生じる動きのずれによる違和感を解消することができた。最後に、教育用コンテンツを試作することで、本研究で提案する3次元アニメーションは、従来のビデオコンテンツよりユーザを惹きつける効果があることが確認できた。今後の課題としては、入力モーションデータから全く未知の動作を生成する手法の構築や、補間処理の高速化がある。

謝辞 本研究の一部は、内閣府最先端研究(LR030)、文科省科研費(25870570)および総務省SCOPE(ICTイノベーション創出型研究開発)の助成を受けて実施された。

参考文献

- [1] Kovar, L., Gleicher, M. and Pighin, F.: Motion Graphs, *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '02*, New York, NY, USA, ACM, pp. 473–482 (2002).
- [2] Levine, S., Lee, Y., Koltun, V. and Popović, Z.: Space-time Planning with Parameterized Locomotion Controllers, *ACM Trans. Graph.*, Vol. 30, No. 3, pp. 23:1–23:11 (2011).
- [3] Zhao, L., Normoyle, A., Khanna, S. and Safonova, A.: Automatic Construction of a Minimum Size Motion Graph, *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '09*, New York, NY, USA, ACM, pp. 27–35 (2009).
- [4] Nakaoka, S., Nakazawa, A. and Ikeuchi, K.: An Efficient Method for Composing Whole Body Motions of a Humanoid Robot, *Proc. Tenth International Conference on Virtual Systems and Multimedia(VSMM2004)*, pp. 1142–1151 (2004).
- [5] Shiratori, T., Nakazawa, A. and Ikeuchi, K.: Detecting dance motion structure through music analysis, *Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, FGR' 04*, Washington, DC, USA, IEEE Computer Society, pp. 857–862 (online), available from <http://dl.acm.org/citation.cfm?id=1949767.1949917> (2004).
- [6] Schödl, A., Szeliski, R., Salesin, D. H. and Essa, I.: Video textures, *Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00*, pp. 489–498 (2000).
- [7] Couture, V., Langer, M. S. and Roy, S.: Panoramic stereo video textures, *Proceedings of the 2011 International Conference on Computer Vision, ICCV '11*, IEEE Computer Society, pp. 1251–1258 (2011).
- [8] Liao, Z., Joshi, N. and Hoppe, H.: Automated video looping with progressive dynamism, *ACM Trans. Graph.*, Vol. 32, No. 4, pp. 77:1–77:10 (2013).
- [9] Furukawa, M., Akagi, Y., Fukumoto, S., Kawasaki, H. and Kawai, Y.: Seamless Interaction Between Real Object and Animated 3D Model, *CHI 2013 Workshop on Human Computer Interaction for Third Places*, 13, No. 6, pp. 1–6 (2013).
- [10] Akagi, Y., Furukawa, M., Fukumoto, S. and Kawai, H. K. Y.: Interactive 3D animation system based on touch interface and efficient creation tools, *The 2013 IEEE International Conference on Multimedia and Expo (ICME) 2013 in "Multimedia for Humanity" theme track*, 318, pp. 1–7 (2013).
- [11] Krüger, B., Tautges, J., Weber, A. and Zinke, A.: Fast Local and Global Similarity Searches in Large Motion Capture Databases, *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '10*, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association, pp. 1–10 (2010).
- [12] Sumner, R. W., Schmid, J. and Pauly, M.: Embedded deformation for shape manipulation, *ACM SIGGRAPH 2007 papers, SIGGRAPH '07*, ACM, (online), DOI: 10.1145/1275808.1276478 (2007).