

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 8 日現在

機関番号：10101

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25282048

研究課題名(和文) モバイルARアニメーションに基づくストーリーテリングシステムとその実践的評価

研究課題名(英文) A Storytelling System with Mobile AR Animation and Its Evaluations

研究代表者

杉本 雅則 (Sugimoto, Masanori)

北海道大学・情報科学研究科・教授

研究者番号：90280560

交付決定額(研究期間全体)：(直接経費) 12,800,000円

研究成果の概要(和文)：本研究の技術的な部分では、ストーリー中のキャラクタ表現のためのモーション取得が重要となる。そこで、モーションの各フレームを整数値を要素とする特徴ベクトルで表現することにより、Dynamic Time Warpingの計算を高速化し検索時間を短縮する手法を提案した。フレームにおける人型スケルトンの姿勢を表す特徴ベクトルに対して主成分分析を行い、各主成分の次元に対応する要素を2ビット量子化する。従来のDTWよりも30倍以上高速であり、既存手法とほぼ同程度の正解率と検索時間を示すことを確認した。提案システムを小学校でのストーリーテリング実験の評価手法ならびに計画を進めた。

研究成果の概要(英文)：For construction a storytelling system in mobile AR environments, we propose a rapid motion retrieval technique to represent motions of characters in stories.

Frames of motions are represented by feature vectors whose elements are integer values. The number of the feature vector dimension is reduced by using the Principal Component Analysis method and values of elements in the vector are quantized into two bits. Comparative experiments with existing methods have proved that our proposed technique can complete retrieval tasks more than 30 times faster than the traditional Dynamic Time Warping method and shown almost the same level of precision and rapid calculation time as existing methods. We also investigated evaluation methods of the proposed system to be used in elementary schools.

研究分野：情報工学

キーワード：コンピュータアニメーション ストーリーテリング

1 研究開始当初の背景

平成 23 年 4 月改訂（小学校）の新学習指導要領は、「生きる力」を育むという理念のもと、知識や技能の習得とともに思考力・判断力・表現力などの育成を重視”することを目標としている (http://www.mext.go.jp/a_menu/shotou/new-cs/index.htm)。ストーリーテリングは、児童の創造性、想像力、論理的思考力、表現力を高める教育手法の 1 つとして従来から採用されてきたが、上記の目標に合致した方法であることから今後ますます着目されると考えられる。コンピュータを用いたストーリーテリングは、幼児や児童を対象としてこれまでにさまざまな研究が行われており、主として ACM CHI, IDC, CSCL などの難関国際会議でその成果が発表されている。具体的な例として、マウスなどの入力デバイスでディスプレイ上に展開されるストーリーを作成、実行するシステム（たとえば [1][2]）、複数のユーザが互いに協調しつつ身体的なインタラクションを通してストーリーの表現を行うシステム（たとえば [3][4][5]）などが挙げられる。申請者らは、これまでにモバイルプロジェクトを用いて物理的なロボットを操作することによる協調的なストーリーテリングシステム（科研費挑戦的萌芽研究 2008-2010）、テーブルトップ環境での実ロボットを対象としたプログラミング（科研費基盤研究 B 2010-2012）とストーリーテリングへの応用を試みてきた。これらの実世界でのストーリーテリングは、主として小学生を対象に教育現場で実践的に評価された。その結果、没入感の高い拡張現実環境でロボットを用いることにより子どもたちの動機付けを高められること、ロボットを主人公とした第三者視点での論理的なストーリー構成能力を養えること、投影のための描画や実物体を活用した空間配置によるシーン構築を通して創造的かつ芸術的な表現力を涵養できる効果が示唆された。一方、実世界メディアとしてのロボットの表現力の限界が、児童や教員から指摘された。実世界において、ロボットの振る舞いや形状をアニメーションのキャラクターのように自在に変えつつ操作することは現時点では容易ではない。しかしながら、ストーリー表現において視覚的なリアリティをより高めるためには、ストーリーの展開や物理的な環境に応じてキャラクター自身が振る舞いを柔軟に変えられるようシステムが設計されることが望ましい。またその場合、ロボットが持つ実世界性という特徴を保持するキャラクターであることが、学習者の動機付けを高める効果に寄与すると期待される。そこで本申請では、申請者らがこれまでに進めてきたストーリーテリング支援 [6]、モバイルプロジェクトカメラシステムに

よる 3 次元形状取得技術 [7]、コンピュータアニメーション技術 [8]、ヒューマンロボットインタラクション技術 [9] を統合および発展させることで、よりリアリティの高いストーリーテリング環境を実現するとともに、その効果を教育現場で実証することを目指す。

2 研究の目的

本研究の目的は、モバイル AR アニメーション機能を持つストーリーテリングシステムの構築とその評価であった。これまでの科研費研究では、モバイル AR 環境でのストーリーテリングシステム（基盤研究 B、萌芽研究など）を構築している。技術的な課題の中心は、コンピュータアニメーションであった。ストーリー中のキャラクターアニメーションの表現の際、既存のモーションからの類似検索手法が用いられるが、高次元データの検索となり計算量が膨大となる。モバイル環境での実装に当たっては、計算量を軽減しつつ高い正解率を維持する必要がある。

本稿で提案する手法は、まず最初にモーション中の各フレームから抽出した実数値の特徴ベクトルに対し、主成分分析を適用することより次元削減を行う。次に、各主成分に対応する要素の値を 0, 1, 2, 3 の整数で 2 ビット量子化することで、各フレームを整数値の特徴ベクトルで表現する。フレーム間の距離は、各フレームに対応する特徴ベクトルを用いて生成される類似度行列より高速に求めることができる。モーション検索で用いる Dynamic Time Warping(以下 DTW) は、各フレーム間の距離の総和の最小値を求めるアルゴリズムである。よって、提案手法によりフレーム間の距離計算の高速化することで、モーション検索の高速化が可能となる。

提案手法の実装に当たり、各フレームに対応する実数値の特徴ベクトルに対して主成分分析を適用し、次元削減とモーション検索性能の関係を予備実験で調査した。提案手法の評価実験では、予備実験で最適と判断した次元数により各フレームを整数値の特徴ベクトルで表現した。既存手法との比較の結果、提案手法は従来の DTW の約 30 倍の高速化が可能であること、k-d tree を用いる Forebes らの手法 [10] とほぼ同等の正解率と計算時間でモーション検索が可能であることを、それぞれ確認した。

提案手法を実装したシステムを小学校での教材と組み合わせるに当たり、対象となる科目と単元構成、さらには提案システムの評価手法を検討する必要がある。申請者らは、実践協力校の教員との議論を踏まえつつ、これらの点を調査した。

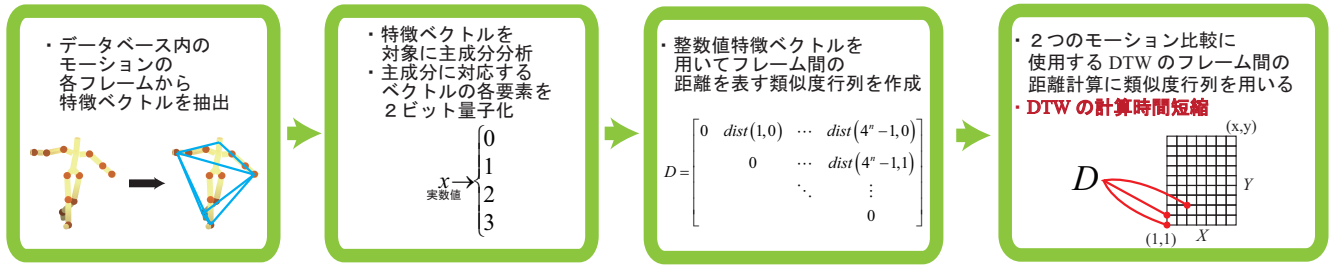


図 1: 提案手法の概要

3 研究の方法

提案手法では、モーション中の各フレームに対して整数値の特徴ベクトルを生成し、DTWの距離関数の計算を高速化することでモーション検索の実行時間を短縮する。提案手法の概要は図1に示す。モーションの各フレームについて、スケルトン表現のキャラクタの手足などの部位間の距離を基に17次元の特徴ベクトルを抽出する。その特徴ベクトルに対して主成分分析を行い、各主成分に対して設定された閾値を用いて特徴ベクトルの各要素を0, 1, 2, 3の4値で表す。2ビットで表された新たな特徴ベクトルを用いて、フレーム間の距離を表す類似度行列を作成する。DTWのフレーム間の距離計算に類似度行列を用いることで、モーション検索の計算を高速化する。

(1) 特徴量抽出

提案する特徴量抽出手法は、多次元の関節の角度情報(実数値のベクトル)が時系列順に並んだモーションデータファイルを、整数値のベクトルが時系列順に並んだデータに変換する。

モーションデータファイルには、関節の親子構造を含めた角度情報がフォワードキネマティクスで記述されている。インターネット上で公開されているCMUのデータベース[12]のモーションデータは、1フレームにつき32箇所の関節部位の3次元回転情報(計96次元)を持つ。高速なモーション比較をするためには次元削減を行いつつ、モーション間の類似度が正確に計測できる特徴ベクトルを抽出する必要がある。今回はMüller[11]らの手法を参考に、モーションの性質を表す特徴ベクトルとして身体の部位間の距離を用いた。

全モーションの各フレームの距離特徴量に対して、主成分分析を適用することにより特徴ベクトルの次元数を削減する。これにより、寄与率が上位 n 個の主成分からなる特徴ベクトルによって各フレームが表現される、主成分にかかる係数は以下の式により2ビットの整数に変換され、特徴ベクトルの要素の値として与えられる。 n を特徴ベクトルの次元数、 i を寄与率上位 i 番目の主成分 $i(1 \leq i \leq n)$ 、 a_{ji} をフレーム j の主成分 i に対応する係数、

μ_i を主成分 i の係数の平均、 σ_i をその標準偏差とする。フレーム j を表現する特徴ベクトルを $\mathbf{g}_j = (g_{j1}, \dots, g_{jn})$ とすると、主成分 i に対するフレーム j の要素 g_{ji} は式(1)のように計算される。

$$g_{ji} = \begin{cases} 0 & a_{ji} < \mu_i - \sigma_i \\ 1 & \mu_i - \sigma_i < a_{ji} < \mu_i \\ 2 & \mu_i < a_{ji} < \mu_i + \sigma_i \\ 3 & a_{ji} > \mu_i + \sigma_i \end{cases} \quad (1)$$

g_{ji} は2ビット量子化され、0, 1, 2, 3の4値のいずれかになる。以下の式(1)により特徴ベクトル \mathbf{g}_j を整数値特徴量 m_j にすることで、モーション中の全てのフレームは0から $4^n - 1$ までのいずれかの整数値に変換される。この整数値特徴量は、3.2節に示す類似度行列のインデックスとして用いられる。

$$m_j = \sum_{i=1}^n g_{ji} \times 4^{i-1} \quad (2)$$

(2) フレーム比較に用いる類似度行列の作成

前節で示した特徴ベクトルを用い、DTWの計算時間を短縮するために類似度行列を生成する。主成分 i に対応する固有ベクトルの寄与率を c_i 、フレーム j および k を表す特徴ベクトル $\mathbf{g}_j, \mathbf{g}_k$ の主成分 i に対応する要素をそれぞれ g_{ji}, g_{ki} とする。式(2)より $\mathbf{g}_j, \mathbf{g}_k$ は整数特徴量 m_j, m_k に対応する。式(3)(4)に示すように、 g_{ji} と g_{ki} の差の絶対値に c_i で重みづけした値を $weightedDiff$ で求め、その総和をフレーム j, k 間の距離 $dist$ と定義する。

$$dist(m_j, m_k) = \sum_{i=1}^n weightedDiff(g_{ji}, g_{ki}) \quad (3)$$

$$weightedDiff(g_{ji}, g_{ki}) = \begin{cases} 0 & |g_{ji} - g_{ki}| = 0 \\ c_i & |g_{ji} - g_{ki}| = 1 \\ c_i \times 4 & |g_{ji} - g_{ki}| = 2 \\ c_i \times 4^2 & |g_{ji} - g_{ki}| = 3 \end{cases} \quad (4)$$

各フレームは n 次元の特徴ベクトルで表現されているので、式 (2) に示した通り $2 \times n$ ビットの非負整数（すなわち 0 から $4^n - 1$ ）に対応付けられる。DTW の計算時間を短縮するために、各フレーム間の距離を式 (3) と式 (4) を用いて 4^n 次正方行列に格納する。この行列 D を類似度行列と呼ぶ。類似度行列 D の対角成分は同じ特徴ベクトルに対応するフレーム間距離となるため、値は 0 になる。類似度行列 D は対称行列となるため、式 (5) に示すように上三角行列または下三角行列で表現できる。各フレームの特徴ベクトルで計算される整数値特徴量 m を行と列に指定すれば、それらの距離が求まる。式 (5) で例えば、フレーム j, k 間の距離は、 m_j 行 m_k 列の要素である。モーション間の比較では、次節で示すように、対応付けられた全てのフレーム間の距離を加算するだけでよいので計算時間が短縮できる。

$$D = \begin{bmatrix} 0 & \text{dist}(1,0) & \text{dist}(2,0) & \cdots & \text{dist}(4^n-2,0) & \text{dist}(4^n-1,0) \\ & 0 & & & & \text{dist}(4^n-1,1) \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 0 & \text{dist}(4^n-1,4^n-2) \\ & & & & & 0 \end{bmatrix} \quad (5)$$

(3) 類似度行列を用いた DTW

上記の類似度行列を用いた DTW によるモーション間距離の計算は、以下のように実行される。2 つのモーションのフレーム数を x と y とする。

Step 1 $x \times y$ 行列を用意し、2 つのモーションの各フレーム間の距離を式 (5) の類似度行列から求めて格納する。

Step 2 動的計画法により行列の (1,1) から (x,y) までの距離の総和が最も低くなるパスを求める。

DTW では、上記の Step 1 の計算が占める割合が高い。提案手法では各フレームを既に整数値特徴ベクトルに変換しており、かつ特徴ベクトル間の距離を格納した類似度行列 D を作成済みである。DTW における各フレーム間の距離計算では図 2 のようにインデックスが用意されているため、Step 1 の計算時間が大幅に短縮できる。

4 研究成果

実験を通して得られた成果について述べる。モーション検索の正解率と計算時間を既存手法と比較する実験を行い、提案手法の有効性を確認した。なお、本手法は Windows 7 上で C++ で実装された。実験では、AeroStream RM5J-B41/S (Intel Core i5, 3.10 GHz, 4.00 GB memory) を用いた。

(1) 予備実験

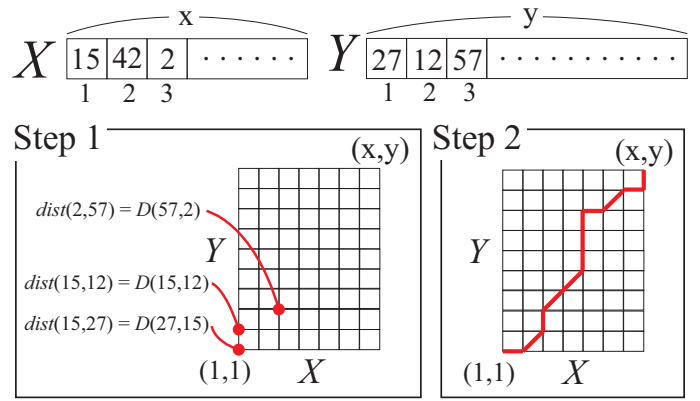


図 2: 提案手法における DTW

特徴ベクトルの最適な次元数を選ぶに当たり、次元数がモーション検索性能にどのような影響を与えるかを調査する予備実験を行った。

① 予備実験詳細

CMU のモーションデータベース [12] から複数のモーションの選択して独自のデータベースを作成する。5 種類の動作 (jump, walk, run, dance, kick) を選択し、各モーションにその動作のクラスをラベル付けしておく。このデータベースに対して 3 分の 1 を入力モーション、残りを参照モーションとする 3 分割の交差確認法でモーション検索を行った。入力モーションと同じクラスの参照モーションが結果として出力されれば正解とした。クラスの合致率と検索に要した時間を計測した。

予備実験では、次元数を 1 から 7 まで変化させた。用いたデータベース中のモーション数は 114 個 (全時間長 794 秒) である。

② 予備実験結果と考察

図 3 は、次元数と正解率の関係を示している。次元数 4 以上で正解率が 90% を超えた。提案手法の性質上、次元数が多いほど正解率はよくなるが、類似度行列の大きさも次元数の 4^2 倍で増大する。次元数の影響は検索の計算時間を示した図 4 にも現れている。次元数が 1 から 6 までは次元数の増大とともに計算時間がなだらかに増大しているが、7 のとき計算時間が大きく増えている。この原因は、類似度行列のサイズが大きくなる (4^7 次正方行列) 分割したためである。図 5 の累積寄与率は、次元数 2 以降は一定で増加している。また、累積寄与率が 70% 以上あれば、提案手法は 80% 以上の正解率を示すことを確認した。

以上の予備実験から、次元数 6 で高速かつ高正解率でモーション検索が出来ると判断し、以降の提案手法と既存手法との比較実験は次元数を 6 で行った。

(2) 比較実験

3 章で説明した提案手法と、既存の 2 種類のモーション検索手法を比較した。

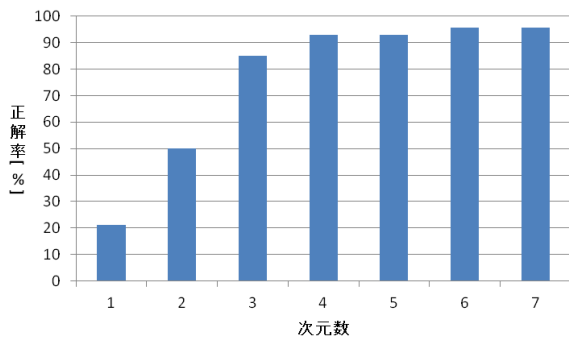


図 3: モーション検索のクラス正解率

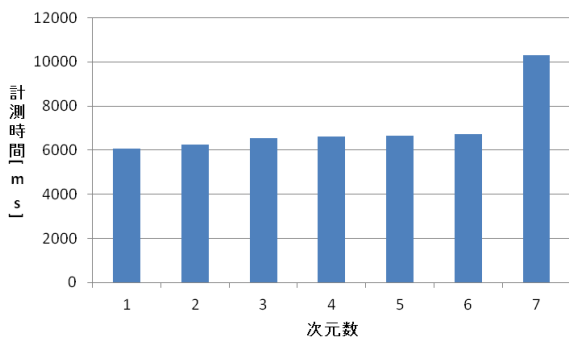


図 4: モーション検索の計算時間

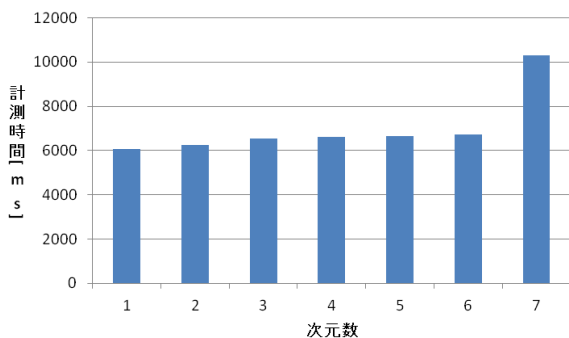


図 5: 射影された主成分の累積寄与率

① 比較実験詳細

既存手法の1つ目は、Müllerらの手法[11]を用いて実数値の17次元の特徴ベクトルを構成し、従来のDTWによりモーション検索を行う手法である。2つ目は、Forbes[10]らによる重み付き主成分分析とk-d treeを用いる手法である。

予備実験と同じモーションデータベースに対し3分割の交差確認法で実験を行い、クラスの合致率による正解率と検索時間を計測した。データベース内のモーションが114個(全体の長さが794秒)と162個(1278秒)の2ケースで実験した。

② 比較実験結果と考察

表1が114個、表2が162個のモーションを使用した実験の結果である。

表 1: 114 個 (794 秒) の結果

	提案手法	DTW	Forbes
時間 (秒)	7.768	240.579	6.869
正解率 (%)	94.737	97.368	95.614

表 2: 162 個 (1278 秒) の結果

	提案手法	DTW	Forbes
時間 (秒)	17.914	624.199	17.128
正解率 (%)	95.062	96.296	94.444

2つのケースで、従来のDTWよりも提案手法とForbesらの手法が約30倍計算時間が速いことが分かる。モーション数が少ない場合、提案手法とForbesらの手法との計算時間の差は大きいですが、モーション数が増えると提案手法が計算時間がやや小さくなる。提案手法では全モーションの各フレームを整数値にすることにより類似度行列の要素を抽出するので、距離計算時間への影響は小さい。一方、Forbesらの手法はk-d treeを用いて特徴量抽出・類似度行列作成を行うので、データ数が多いほどk-d treeが大きくなり、提案手法よりも検索時間への影響が大きい。

正解率に関しては提案手法、Forbesらの手法は主成分分析を用いて次元削減しているにも関わらず、従来のDTWをほぼ同程度の正解率を示すことが確認された。モーション数が多い場合に、Forbesらの手法よりも提案手法の正解率が若干高くなった。提案手法とForbesらの手法は全訓練モーションの各フレームをクラスタリングした後に、クラスごとの距離をあらかじめ類似度行列に格納する。そして、DTWのフレーム間の距離を計算する時に類似度行列を用いて計算時間を短縮する点で共通している。そのクラスタリングの手法が主成分分析と2ビット量子化(提案手法)なのか、重み付き主成分分析とkd木(Forbes)の違いである。データベースが大きいとより精度の高い主成分分析が行えるため、提案手法の方が正解率が高い。今後はデータセットの大きさやクラス数など多様な条件下で実験することで、2つの手法の詳細な性能を今後明らかにする必要がある。

(3) 結論と今後の課題

モバイルAR環境におけるストーリーリングシステムの構築とその評価に当たり、本研究ではその重要な要件となる高速モーション検索手法についての検討を中心に行った。提案手法は、モーションの特徴ベクトルを主成分分析し、2ビット量子化を通して整数値に置き換え類似度行列を作成することにより、モーション検索の正解率を維持しつつ高速化する。提案手法の有効性の確認のために、2種類の既存手法と比較実験を大小2つのデータセッ

トで行った。簡略かつ離散化された特徴ベクトルによるフレーム表現でありながら、提案手法は既存手法に比肩しうる性能を示すことが確認できた。今後はより多様なデータセットで実験を行い、提案手法の利点と限界、さらに実験結果に基づく改良を行う必要がある。また、提案手法に適した特徴量の抽出、特徴ベクトルの分布に基づく類似度行列の圧縮と計算時間の更なる短縮などが今後の課題として挙げられる。さらに、小学校での実践に向けての課題を明らかにし、評価実験を行う予定である。

参考文献

- [1] Howland, K. et al.: A Learner-Centered Design Approach to Developing a Visual Language for Interactive Storytelling, *Proc. of IDC2007*, pp. 45-52, 2007.
- [2] Raffle, H. et al.: Jabberstamp: Embedding Sound and Voice in Traditional Drawings, *Proc. of IDC2007*, pp.137-144, 2007.
- [3] Montemayor, J. et al.: Tools for Children to Create Physical Interactive StoryRooms, *ACM Computers in Entertainment*, Vol. 2, No.1, pp.1-24, 2004.
- [4] Stanton, D. et al.: Classroom Collaboration in the Design of Tangible Interfaces for Storytelling, *Proc. of CHI 2001*, pp. 482-489, 2001.
- [5] Ryokai, K., Cassel, J.: Computer Support for Children's Collaborative Fantasy Play and Storytelling, *Proc. of CSCL '99*, 1999.
- [6] Sugimoto, M.: A Mobile Mixed Reality Environment for Children's Storytelling using a Handheld Projector and a Robot, *IEEE Transactions on Learning Technologies*, No.4, Vol.3, pp.249-260, 2011.
- [7] Dao, V.N., Sugimoto, M.: A Dynamic Geometry Reconstruction Technique for Mobile Devices Using Adaptive Checkerboard Recognition and Epipolar Geometry, *IEICE Transactions on Information and Systems*, Vol. E94-D, No.2, pp.336-348, 2011.
- [8] Pantuwong, N., Sugimoto, M.: A Novel Template-Based Automatic Rigging Algorithm for Articulated-Character Animation, *Computer Animation and Virtual Worlds*, Vol. 23, Issue 2, pp.125-141, 2012.
- [9] Sugimoto, M., Fujita, T., Mi, H., Krzywinski, A.: RoboTable2: A Novel Programming Environment using Physical Robots on a Tabletop Platform, *Proc. of ACM ACE 2011*, Lisbon, Portugal, Article No.10, 2011.
- [10] Forbes, K., & Fiume, E.: An Efficient Search Algorithm for Motion Data using Weighted PCA. *Proc. of SCA 2005*, Los Angeles, CA, pp. 67-76, 2006.
- [11] Müller, M., & Röder, T.: Motion Templates for Automatic Classification and Retrieval of Motion Capture Data. *Proc. of SCA 2006*, Vienna, Austria, pp. 137-146, 2006.
- [12] C. G. Lab. CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu/>

5 主な論文発表等

[学会発表] (計 7 件)

- ① Pantuwong, N., Takahara, K., Sugimoto, M.: A Rapid Motion Retrieval Technique using Simple and Discrete Representation of Motion Data, *Proc. of IEEE ICITEE 2015*, Chiang Mai, Thailand, 2015 (2015 年 10 月 29 日).
- ② 高原, Pantuwong, 野中, 吉川, 杉本: 特徴ベクトルの簡略離散表現によるモーション検索高速化手法, インタラクシオン 2015, 東京, 2015 (2015 年 3 月 5 日).
- ③ Takeuchi, Y., Sugimoto, M.: Interacting with 3D Model on Tabletop and Mobile Paper Projection, *Proc. of Col-labTech 2014*, Santiago, Chile, pp.111-118, 2014 (2014 年 9 月 9 日).
- ④ Takahara, K., Pantuwong, N., Yoshikawa, T., Nonaka, H., Sugimoto, M.: An Iterative Motion Retrieval and Synthesis Technique Using a Multi-Touch Device, *Proc. of Motion in Games 2013*, Dublin, Ireland, 2013 (2013 年 11 月 8 日).
- ⑤ Nopparit, S., Pantuwong, N., Sugimoto, M.: A Kinetic Energy-based Feature for Unsupervised Motion Clustering, *Proc. of ICITEE 2013*, Yogyakarta, Indonesia, pp. 8-12, 2013 (2013 年 10 月 8 日).
- ⑥ Sugimoto, M.: Beyond Wow Effect: Making Fun Experience Learning and Learning Experience Fun, *Proc. of AECT-ICFER 2013*, Taichung, Taiwan, 2013 (招待講演, 2013 年 6 月 19 日).
- ⑦ Pantuwong, N., Sugimoto, M.: An Intuitive Computer Animation System for Enhancing Human Creativity, *Proc. of SSCI 2013*, Singapore, pp.71-78, 2013 (2013 年 4 月 18 日).

6 研究組織

(1) 研究代表者

杉本 雅則 (SUGIMOTO, Masanori)

北海道大学・大学院情報科学研究科・教授

研究者番号 : 90280560

(2) 研究分担者

稲垣 成哲 (INAGAKI, Shigenori)

神戸大学・大学院人間発達環境学研究科・教授

研究者番号 : 70176387

(3) 協力研究者

パントウオング ナタポン (Pantuwong, Natapon)

モンクット王工科大学ラート クラバン校・情報技術学科・講師