

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 1 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2010～2012

課題番号：22650016

研究課題名（和文） パッチマッチング法による実写 3 次元映像の圧縮と処理

研究課題名（英文） Compression of time-varying 3D meshes based on patch-based matching

研究代表者

山崎 俊彦（YAMASAKI TOSHIHIKO）

東京大学・大学院情報理工学系研究科・准教授

研究者番号：70376599

研究成果の概要（和文）：

実写 3 次元映像の圧縮と処理に関する研究を行った。実写 3 次元映像は各フレーム独立に生成されることが多く、時間的に連続するフレームにおいても頂点数や結線関係が保存されていない場合が多いため、圧縮が困難であった。そこで、研究代表者が考案したパッチマッチング法に基づき、実写 3 次元映像の頂点情報、結線情報、色情報を同時に圧縮可能な技術を研究した。具体的な手法について述べる。まず、測地線距離を用いて各フレームの 3 次元メッシュモデルをほぼ等面積になるようにパッチに分割し、それを圧縮処理の基本単位とした。各パッチの幾何情報、色情報に対してキルヒホッフ行列によるメッシュ周波数解析やベクトル量子化を施すことによりフレーム内圧縮を実現した。さらに、隣接フレーム間でもっとも類似するパッチを探索し、対応パッチからの残渣情報のみを符号化することでフレーム間圧縮を実現した。これにより、フレーム内圧縮で 1/8 程度、フレーム間圧縮で 1/12 程度の圧縮率を実現した。

研究成果の概要（英文）：

We have developed intra-frame and inter-frame coding algorithms for 3D mesh sequences generated by multiple cameras, which we call Time-Varying Meshes (TVMs). For this purpose, mesh segmentation into patches with patch alignment using principal component analysis (PCA) is proposed. The patches are used as minimum units to eliminate spatial and temporal correlation of the TVMs. For intra-frame coding of geometry data, spectral compression using the Kirchhoff matrix was employed and that of color texture was conducted using vector quantization (VQ). The inter-frame coding of geometry data and color texture was achieved by the combination of patch-based matching and simple scalar quantization. As a result, the TVM data were efficiently compressed to 1/8 and 1/12 by intra-frame and inter-frame coding, respectively.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010 年度	1,300,000	0	1,300,000
2011 年度	900,000	270,000	1,170,000
2012 年度	500,000	150,000	650,000
総計	2,700,000	420,000	3,120,000

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：3 次元映像、圧縮、パッチマッチング

1. 研究開始当初の背景

2009年当時、よりリアルな映像表現を求めて富士フィルムの3次元デジタルカメラ、パナソニックの3次元シアターシステム、ソニーの3次元液晶テレビなど、3次元映像を記録・表示するデバイスが次々と発表され、映像の3次元化が一般に浸透し始めた頃であった。しかし、表示されるコンテンツの多くはCG、もしくはステレオで撮影された実写映像であり、真の「実写」3次元映像は扱われていなかった。

我々は2004年からNHK技研・ATRらと共同で、多視点カメラを用いて動きのある実世界の物体を高精細に3次元モデル化する研究に取り組んできた(図1参照)。

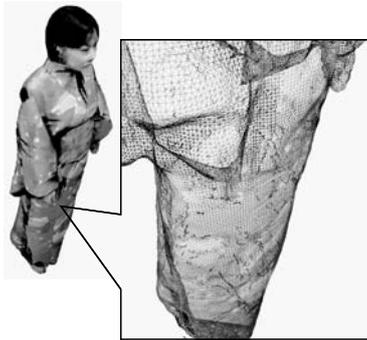


図1 実写3次元映像の例

実写3次元映像に関するこれまでの研究は、取得・生成に関するものが殆どであった。すなわち、取得・生成以降の処理技術、例えば実写3次元映像の圧縮に関しては、まだ不十分な状態にとどまっていた。実写3次元映像は3次元物体の表面情報(頂点座標、結線情報、色など)から構成されており、従来の2次元画像・映像とはデータ構造が全く異なる上にデータ量が膨大であるため、新しい圧縮アルゴリズムの開発が必要であった。

研究代表者らはこの様な状況を踏まえ、これまで拡張3次元ブロックマッチング法や階層的ランレングス法などによって、実写3次元映像の頂点情報を圧縮する手法を開発してきた。これらはIEEEの論文誌(S. Han, IEEE TCSVT, 2007)に掲載され、国内でデジタルコンテンツシンポジウム船井賞など各種の賞を受賞した。ただし、それまでの研究は頂点の幾何情報のみが圧縮可能であり、色情報や結線情報も圧縮できる技術が必要であった。

2. 研究の目的

本研究の目的は、パッチマッチング法という申請者の新たな着想に基づき、実写3次元

映像の頂点情報、結線情報、色情報を同時に圧縮可能な技術を確立することである。実写3次元映像は各フレーム独立に生成されることが多く、時間的に連続するフレームにおいても頂点数や結線関係が保存されていない場合が多いため、圧縮が困難であった。本手法は実写3次元映像に関連する全てのデータを圧縮しようとする、新たな試みである。さらに、2次元動画の圧縮アルゴリズムも他のアプリケーションに応用されているように、本研究で開発したパッチマッチング法を様々な分野に応用・展開することも目指す。これらの技術により、まだ萌芽的技術である実写3次元映像の実用化に資する。

3. 研究の方法

(1) パッチベース圧縮技術の基礎検討

実写3次元映像の圧縮の難しさは下記の2点にある。1つめは、被写体は服を着た人物など非剛体であることが多いため、各フレームが独立に生成される点である。時間的に隣接するフレーム同士であっても頂点数や結線関係は保存されないことが多い。そのため、隣接フレーム間には明示的な対応関係や動き情報がない。2つめは、3次元映像には2次元動画の「ピクセル」のような基本単位構造がなく、3次元モデルを構成する数万から数十万個の頂点が3次元空間の任意位置に分布している点である。そのため、処理のための基本ブロックを作りにくい。

測地線距離などを用いて3次元メッシュモデルの表面をある程度一定の面積で分割し、生成されるパッチを種々の処理の基本単位とする方式(パッチマッチ法)に基づいて上記の問題を解決する(2次元動画の「ブロックマッチング」に対応して、「パッチマッチング」と呼ぶ)。パッチを基本構造とすることで下記のような処理が可能となる。

- ・ パッチの空間冗長性を利用したフレーム内圧縮
- ・ 主成分分析によるパッチの正規化とパッチ同士の差分処理によるフレーム間圧縮

このパッチマッチング法について基礎的な実験を行い、方式の体系化を試みると同時に、パッチそのものの圧縮、およびパッチ間の残差情報の圧縮について下記の検討を行う。

パッチ自身の圧縮(フレーム内圧縮)

分割されたパッチ状のメッシュの頂点・色情報の圧縮について検討を行う。スペクトル分解、ウェーブレット解析、ブロックトランケーション、ベクトル量子化などを適用し、それらの圧縮効率や演算速度などの比較検討を行う。また、結線情報の圧縮手法について検討する。

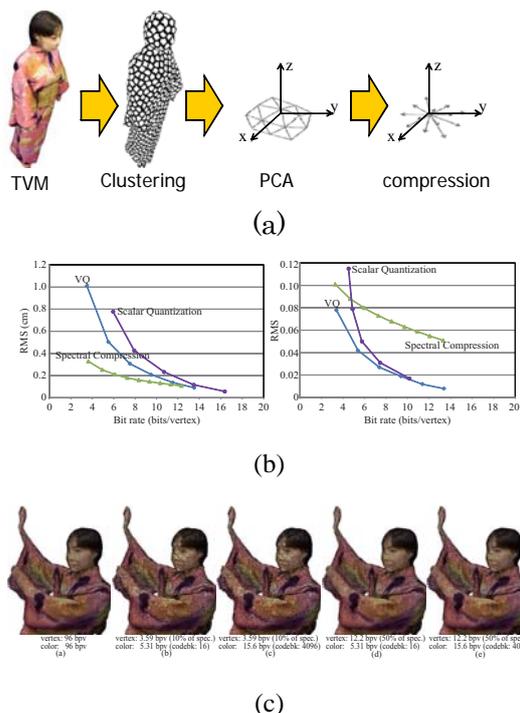


図 2. フレーム内圧縮の検討。(a)概念図、(c) 圧縮手法ごとの符号量と品質の関係、(d)レンダリング結果。

パッチ間の相関と圧縮（フレーム間圧縮）隣接フレーム間のパッチ同士の頂点や色情報を比較し、類似したパッチを探索する（パッチマッチング）と共に、効率よくフレーム間の差分を圧縮する方式について検討する。

(2) パッチベース圧縮技術の最適化 アルゴリズムの最適化

基礎検討の結果を基に、フレーム内圧縮、フレーム間圧縮それぞれについて最適な手法の選定・実装を行う。また、データ構造、量子化幅、キーフレームの頻度の最適化などを行い、圧縮・復号の総合的なアルゴリズムを完成させる。

マルチコア・クラウドコンピューティングによる並列化

実写 3次元映像はデータ量が膨大であるため、処理には大変な時間がかかる。そこで、マルチコア技術を利用した並列処理による高速化を試みる。また、クラウドコンピューティングを利用することによって、ネットワーク上の大量の仮想マシンを利用し、超高速な演算処理へ発展させる。

ソフトウェアの公開

開発した圧縮ソフトウェアに関して、ドク

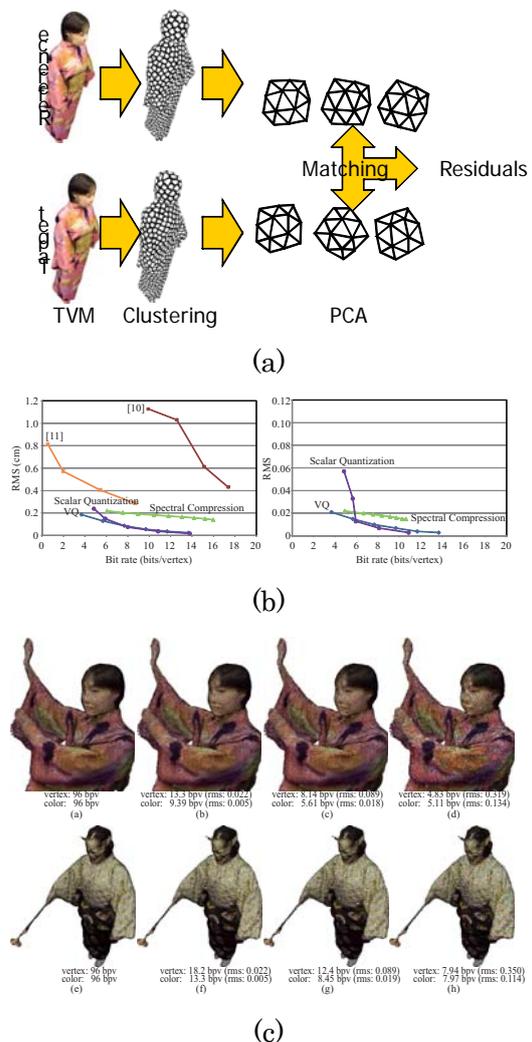


図 3. フレーム間圧縮の検討。(a)概念図、(b) 圧縮手法ごとの符号量と品質の関係、(c)レンダリング結果。

ュメントやサンプルコードを整備すると共に、成果をオープンソースとして公開する。

(3) パッチマッチング法の応用展開

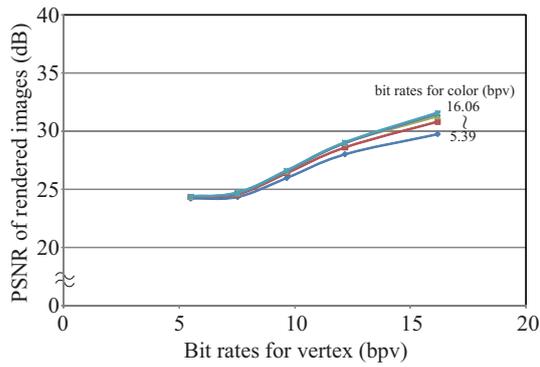
2次元動画の圧縮で利用されているブロックマッチング法は、単に映像を圧縮するためだけでなく、他の用途にも転用されている。同様に、パッチマッチング法の応用展開を試みる。具体的には、パッチごとに動きの方向や量を解析することにより、3次元モデルの動き解析や動き追跡を行う。これにより、例えば行動認識や運動解析などが可能となり、また、類似検索も可能となる。

4. 研究成果

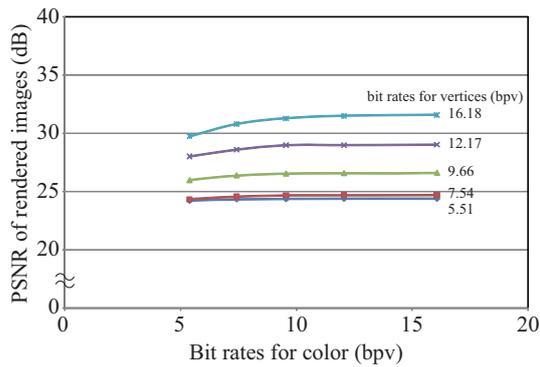
(1) パッチベース圧縮技術の基礎検討

パッチ自身の圧縮（フレーム内圧縮）

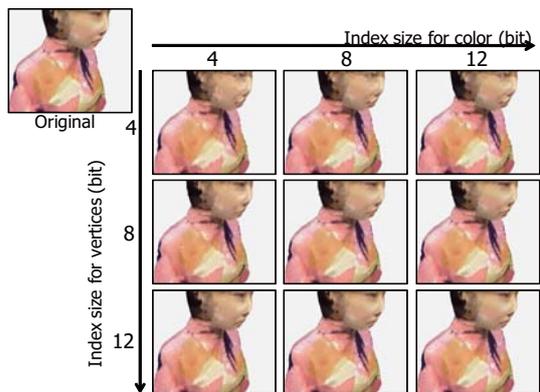
フレーム内圧縮、すなわちパッチ自身の圧縮について検討を行った。結果を図 2 に示す。



(a)



(b)

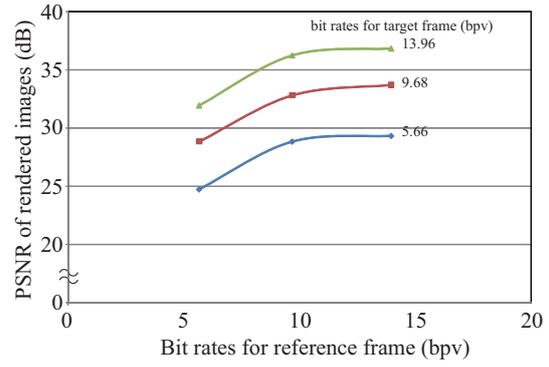


(c)

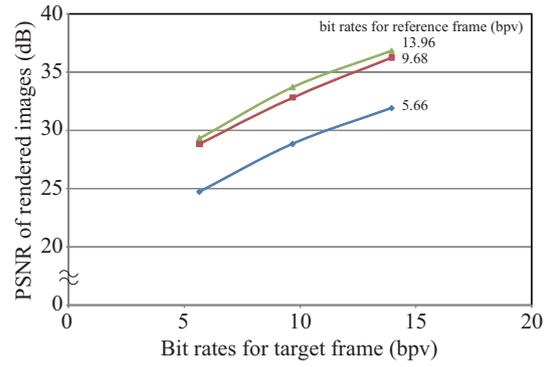
図4. フレーム内圧縮の最適化。(a)幾何情報圧縮のビットレートと品質 (b)色情報圧縮のビットレートと品質 (c)レンダリング結果

図 2(a)に示したとおり、与えられた 3 次元メッシュに対して測地線に基づいた領域のクラスタリングを行い、パッチを生成する。生成されたそれぞれのパッチに主成分分析 (PCA) をかけて、パッチの方向と座標を正規化する。正規化されたそれぞれのパッチに対して符号化処理を施すことによって圧縮を行う。

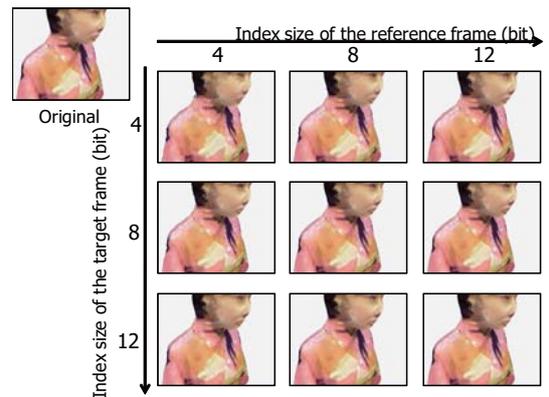
符号化手法としてはキルヒホッフ行列を用いたスペクトル圧縮方式、ベクトル量子化



(a)



(b)



(c)

図5. フレーム間圧縮の最適化。(a)幾何情報圧縮のビットレートと品質 (b)色情報圧縮のビットレートと品質 (c)レンダリング結果

方式、スカラー量子化方式について検討を行った。符号効率の検討結果を図 2(b)に示す。この結果、幾何情報の圧縮にはスペクトル圧縮方式が、色情報の圧縮にはベクトル量子化方式が適していることが明らかとなった。

実際に、着物を着た女性のシーケンスに対し、圧縮処理を行った(図 2(c))。様々なビットレートで圧縮実験を行ったところ、符号化効率 1/8 程度までは人間の目には劣化が見取れないことがわかった。

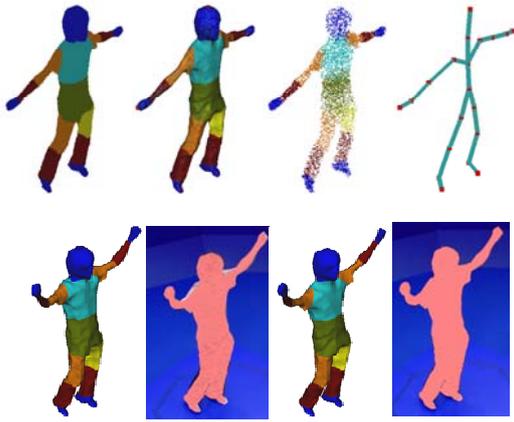


図 6. 応用例: 人物姿勢抽出に基づいた運動追跡手法

パッチ間の相関と圧縮（フレーム間圧縮）
 フレーム間圧縮、すなわち隣接するフレーム同士の圧縮について検討を行った。フレーム間圧縮は、基準フレームと対象フレームの差分だけを符号化すればよく、フレーム内圧縮に比べてさらに効率のよい圧縮が期待できる。

図 3(a)に示したとおり、基準フレームと圧縮対象フレームの両方からフレーム内圧縮と同じ方式でパッチを生成し、正規化する。正規化されたパッチ同士の類似度を計算し、基準フレームに含まれるパッチから圧縮対象パッチにもっともよく似ているものを探索し、差分情報のみを符号化する。

検討の結果、幾何情報の圧縮にはベクトル量子化およびスカラー量子化が、色情情報の圧縮にはベクトル量子化が適していることを明らかにした。実験結果を図 3(b)に示す。

着物を着た女性のシーケンスおよび鬼のシーケンスに対し、圧縮処理を行った(図 3(c))。様々なビットレートで圧縮実験を行ったところ、符号化効率 1/12 程度までは人間の目には劣化が見て取れないことがわかった。

(2) パッチベース圧縮技術の最適化 アルゴリズムの最適化

フレーム内圧縮における頂点情報と色情情報のビットレートの分配について。実写 3次元を構成する頂点・色・結線情報の 3種類の情報の内、結線情報の圧縮についてはこれまでに様々な手法が提案されている。それに対し、頂点と色情情報についてはあまり検討がなされていなかった。そこで、全体のビットレートが与えられたときに結線情報と色情情報のビットレートの最適な分配方法について検討した。その結果、最終的な映像品質は色情情報よりも頂点情報の復号精度に大きく

左右され、色情情報のビットレートは 9-14bpv 程度で十分であることを明らかにした。結果を図 4 に示す。

フレーム内圧縮とフレーム間圧縮のビットレートの配分については、フレーム間圧縮の復号品質はフレーム間の差分情報のほうが支配的で、フレーム内圧縮のレートはある程度(10bpv 程度)高くしておけば十分であることが明らかになった。また、フレーム間圧縮はフレーム内圧縮に比べて 2 倍程度符号化効率が高いことも明らかとなり、本研究で提案するパッチマッチング法の優位性が示された。結果を図 5 に示す。

マルチコア・クラウドコンピューティングによる並列化

パッチマッチング法は、それぞれのパッチをそれぞれ独立に処理できるため、並列処理に展開しやすい。OpenMP を使ったマルチコア処理、クラウドコンピューティングを利用した処理について検討を行った。特に、マルチコア処理については並列処理部分を実装して動作を確認した。

ソフトウェアの公開

ソフトウェア公開に向けてソースコードの整備を行った。今後、研究プロジェクトのサイト整備を行い、ソースコードを公開していく。

(3) パッチマッチング法の応用展開

高次元空間における人物姿勢抽出に基づいた運動追跡手法を提案した。まず同期された多視点動画像から得られた 3 次元ボクセルモデルのシーケンスのうち、第 1 フレームのテンプレートモデルをキューブマッチング法にて得る。次にテンプレートモデルを 15 の部位にセグメンテーションし、セグメントに応じて各ボクセルにラベルを与える。そして次フレームの人物姿勢を高次元運動追跡手法である annealed particle filter 法を用いて推定しテンプレートモデルを変形する。変形されたモデルは誤差を含むので、シルエット画像を用いて高精度化する。概念図を図 6 に示す。本手法によって、超点数や結線情報がほぼ保存される。そのため、さらに高効率な圧縮や、3 次元モデルの骨格抽出、3 次元モデルの動き記述・制御など様々な応用展開が可能となった。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計20件)

[1] T. Yamasaki and T. Chen, "Relative-Distance-Based Soft Voting for Human Attribute Analysis using Top-View Images," *ITE Transactions on Media Technology and Applications*, vol. 1, no. 2 p. 138-147, 2013.

[2] T. Yamasaki, T. Matsunami, and T. Chen, "Human attribute analysis using a top-view camera based on two-stage classification," *IEICE Transactions on Information and Systems*, Vol. E96-D, No.4, pp. 993-996, Apr. 2013.

[3] 山田健太郎, 山崎俊彦, 相澤清晴, "シルエット抽出とモデル生成の相互フィードバックによる Time-Varying Mesh の高精度化 (Time-Varying Mesh Generation Based on Iterative Feedback between Silhouette Extraction and Geometry Modeling)," *電子情報通信学会論文誌 D*, vol. J93-D, no, 8, pp. 1533-1543, Aug. 2010.

他、17件

[学会発表](計19件)

[1] W. Luo, T. Yamasaki, and K. Aizawa, "Marker-less Human Pose Estimation and Surface Reconstruction Using a Segmented Model," *Proceedings of IEEE International Conference on Image Processing*, pp. 901-904, Brussels, Belgium, 11-14 Sep. 2011.

[2] W. Luo, T. Yamasaki, and K. Aizawa, "3D pose estimation in high dimensional search spaces with local memorization," *Proceedings of 28th Picture Coding Symposium (PCS2010)*, pp. 354-357, Nagoya, Aichi, Dec. 7-10, 2010.

[3] T. Yamasaki and K. Aizawa, "Bit allocation of vertices and colors for patch-based coding in time-varying meshes," *Proceedings of 28th Picture Coding Symposium (PCS2010)*, pp. 162-165, Nagoya, Aichi, Dec. 7-10, 2010.

[4] S.R. Han, T. Yamasaki and K. Aizawa, "Automatic Preview Video Generation For Mesh Sequences," *2010 IEEE 17th International Conference on Image Processing (ICIP 2010)*, pp. 3433-3436, Hong Kong, China, Sep. 26-29, 2010.

[5] T. Yamasaki and K. Aizawa, "Patch-Based Compression for

Time-Varying Meshes," *2010 IEEE 17th International Conference on Image Processing (ICIP 2010)*, pp. 2945-2948, Hong Kong, China, Sep. 26-29, 2010.

[6] S. Nakagawa, T. Yamasaki, and K. Aizawa, "Deformation-Based Compression of Time-Varying Meshes for Displaying on Mobile Terminals," *3DTV Conference 2010 (3DTV-Con 2010)*, USB Proceedings, #60, Tampere, Finland June 7-9 2010.

他、13件

[図書](計1件)

[1] T. Yamasaki and K. Aizawa, "Refinement of Visual Hulls for Human Performance Capture," *Modeling and Simulation in Engineering*, InTech, editor: Catalin Alexandru, Chapter 7, pp. 159-174, ISBN 978-953-51-0012-6, 298 pages, 2012.

[産業財産権]

取得状況(計1件)

名称: 3次元動画画像の圧縮に用いる基本単位及び当該基本単位の生成方法、当該基本単位を用いた3次元動画画像の圧縮方法

発明者: 山崎俊彦、相澤清晴

権利者: 同上

種類: 特許

番号: 特開 2012-68881

取得年月日: 平成 24 年 4 月 5 日

国内外の別: 国内

[その他]

ホームページ等

<http://www.hal.t.u-tokyo.ac.jp/ja/index.html>

6. 研究組織

(1)研究代表者

山崎 俊彦 (YAMASAKI TOSHIHIKO)

東京大学・大学院情報理工学系研究科・准教授

研究者番号: 70376599

(2)研究分担者

なし

(3)連携研究者

相澤 清晴 (AIZAWA KIYOHARU)

東京大学・大学院情報学環・教授

研究者番号: 20192453