

# Skin-off : 2次元平面への展開による3次元ビデオ映像の表現と圧縮

## Skin-off: Representation and Compression for 3D Video by Unfolding onto 2D Planes

曾良 洋介<sup>†</sup>  
Yosuke Katsura

波部 斉<sup>‡</sup>  
Hitoshi Habe

Martin Boehme<sup>†</sup>  
Martin Boehme

松山 隆司<sup>†</sup>  
Takashi Matsuyama

### 1. 平面による3次元ビデオ映像の表現と圧縮

複数視点のカメラを用いることで3次元世界の様々な対象の幾何情報(姿・形), 光学情報(色・模様)の時間変化をそのまま記録できる。それらは実写に基づく映像であるため写実性が高く, エンターテインメントや教育用コンテンツ, 伝統舞踊の記録保存といった多様な分野での利用が期待され, 幅広い研究が行われている。ここでは, 膨大なデータ量になる幾何・光学情報を効率的に表現・圧縮する方法の確立が一つの大きな課題になる。

3次元情報を獲得するためのカメラ配置は, a) 直線上にカメラを設置し平行に対象を観測するもの, b) 観測対象の周囲にカメラを設置し, 対象に向かって観測するもの, c) ある点から周囲を向くようにカメラを設置し, 全方位の映像を観測するもの, などに分けることができる。その中で, b), c) によって観測された情報は, 3次元空間中の物体表面がもつ幾何情報と, その表面にマッピングされた映像情報から構成されると考えられる。我々はこのようなメディアを3次元ビデオ映像と呼び, 本稿ではその表現・圧縮法を提案する (b) のカメラ配置から3次元ビデオ映像が得られる様子を図1に示す [5, 7]。

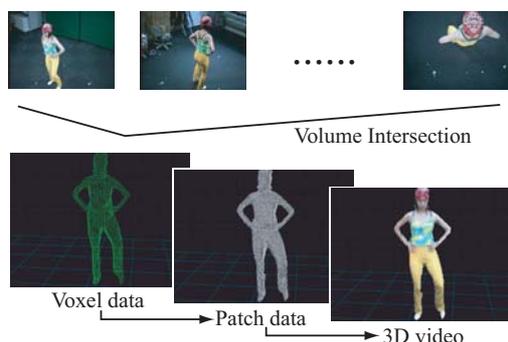


図 1: 3次元ビデオ映像

対象物体は一般に様々な形状をし, 時間経過とともにその形状が変化する。本手法では, それらの情報を効率よく表現・圧縮するために, 1) 任意形状の物体を切り開くことで2次元平面へ展開し, 2) 幾何・光学情報を2次元矩形画像上で記述し, 3) 通常の2次元映像に用いる映像圧縮法を適用する (図2)。この中の1)の操作は直観的には物体の「皮をむく (Skin-off)」操作に相当する。

CG分野を中心として, 3次元メッシュで表現された物体形状を圧縮する研究が数多く行われている [1, 6]。その中には同様に2次元平面への展開を行うものがあるが [2], いずれもCGによるアニメーションを想定しているため, 実写による3次元ビデオ映像にそのまま応用することは困難である。具体的な問題としては, 1) 各時刻で独立に求めた3次元メッシュの幾何情報の効率的な圧縮, 2) 誤差を含む幾何・光学情報からの写実的な映像の

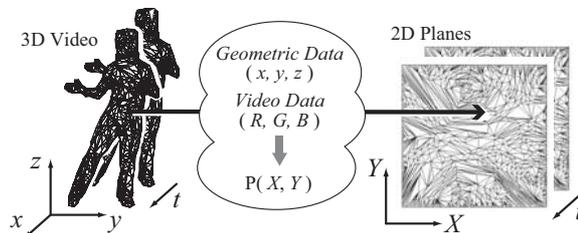


図 2: 3次元ビデオの2次元平面への展開

生成, という点が挙げられる。本稿ではこれらの問題に対する解決策を述べ, 本手法の有効性を明らかにする。

### 2. 3次元ビデオ映像の幾何情報の表現と圧縮

2次元矩形画像上で幾何情報を表現するために, 矩形画像の一つの画素を3次元メッシュの一つの頂点に対応させ, RGBの画素値で頂点の座標  $(x, y, z)$  を表現する。さらに, 矩形画像上での画素の隣接関係が3次元空間中の頂点の隣接関係をそのまま表現するようにする。これによって, 矩形画像 (Geometry Images と呼ぶ) で3次元メッシュの幾何情報をすべて表現でき, 既存の2次元映像用のコーデックを使用して圧縮できる。この手法はCGアニメーションに対して提案された [2] ものであるが, 前述のとおり実映像から得られた3次元ビデオに対しては問題が発生する。以下に手法の詳細を述べながら問題点とその解決策を示す。

#### 2.1 形状と運動の同時復元

2次元映像圧縮法を Geometry Images へ適用したときに高い圧縮率を得るには, Geometry Images の空間的・時間的な相関が高くなる必要がある。Geometry Images 中で隣接する画素は, 対応する3次元空間中の頂点も隣接していることから, 空間的な相関が高くなることは保証される。一方, 時間的な相関に関しては, 実写による3次元ビデオでは一般に個々の時刻毎に3次元メッシュが求まり, 時刻間での頂点同士の対応が取れない。これによって異なる時刻の Geometry Images の画素値 (頂点の座標) の相関が低くなり圧縮効率が低下する。

これに対し我々は, 時刻毎に3次元メッシュを得るのではなく, 時刻間でのメッシュの運動と各時刻でのメッシュ形状を同時に得るアルゴリズムを採用する [7]。ここでは, 複数視点映像から視体積交差法によって得られた3次元メッシュに対して, 視点間の光学的整合性, 形状の滑らかさ, 過去の運動履歴などによる制約条件を用いて, 精度の高いメッシュ形状とその運動を同時に得ることができる。

これによって, 実写による3次元ビデオ映像についても, 時刻間のメッシュの対応付けを正確に得ることができ, Geometry Images を用いた効率のよい表現・圧縮が可能となる。

<sup>†</sup> 京都大学情報学研究所. Grad.Sch.of Informatics, Kyoto Univ.

<sup>‡</sup> 京都大学工学研究所. Grad. Sch. of Engineering, Kyoto Univ.

## 2.2 3次元メッシュの矩形平面上への展開

こうして得られた3次元メッシュの矩形平面への展開は、1)3次元メッシュへ切れ目を加え、2)メッシュ頂点の矩形平面上へのマッピングの最適化を行い、3)矩形平面上で再サンプリングを行う、という手順で行われる。1)、2)を行う際には、3次元空間中の頂点間の距離の比が矩形平面上でも保たれるように最適化を行う。2)までの処理の結果、矩形平面上の任意の位置に頂点が配置されるため、3)では適当な間隔でサンプリングを行い、離散的な位置  $(X, Y)$  に対して画素値  $(r, g, b) = (x, y, z)$  が定義される Geometry Images を得る。

時系列データについては、ある時刻のデータについて1)、2)で切れ目と矩形平面へのマッピングを求め、他の時刻でも同じ切れ目とマッピングを適用して各時刻の Geometry Images を求める [3]。

## 3. 3次元ビデオ映像の光学情報の表現と圧縮

CGアニメーションにおいては、仮想物体表面の反射特性を与え、光環境を設定することによって任意視点映像を生成できるが、実写による3次元ビデオ映像では物体表面の反射特性の計測は一般に困難である。また、その幾何情報には若干の形状測定誤差が含まれる。このような状況で写実的な映像を生成するためには、陽に反射特性を計測せず、多視点映像から直接任意視点での映像を生成する方法が考えられる [7, 5]。ここでは、物体表面の光学情報(色)の表現法として以下の二つの手法が考えられる。

### 3.1 面ベースの表現

CGで通常行われているように、3次元メッシュの面上のテクスチャを画像として保持し、そこからテクスチャマッピングによって任意視点映像を生成する「面ベース法」が考えられる。CGアニメーションのように正確な幾何形状が得られている場合には、高解像度のテクスチャがマッピングされるので、写実性の高い映像を生成できる。しかし、先述のとおり、計測された幾何形状には誤差が含まれるので、面パッチの境界でテクスチャの不整合が生じ、生成された映像にちらつきが発生する。

### 3.2 頂点ベースの表現

もう一つの手法は、3次元メッシュの各頂点に色情報を持たせるものである。生成する映像の視点位置と撮影した各カメラ位置との関係から適切なカメラを選択して頂点の色情報を求め、各面のテクスチャは隣接する頂点の色情報を補間して生成する [5]。この場合には、隣接する面パッチが頂点を共有しているため、面パッチ間でのテクスチャの不整合は起こらず、滑らかなテクスチャを生成することが可能となる。

図3にそれぞれの表現手法で映像を生成した例を示す。先に述べたとおり (b) の頂点ベースの表現手法が滑らかなテクスチャを生成できていることがわかる。また、頂点ベースの表現手法では2.2節で述べた Geometry Images の画素値を色情報に置き換えるだけで光学情報を表現できるので、統一的な扱いが可能となる。

以上の理由により、実写による3次元ビデオの光学情報の表現・圧縮の手法として、頂点ベースによる方法が適当であると考えられる。



(a) Patch-Based Method



(b) Vertex-Based Method

図3: レンダリングの結果

## 4. 実験

ここまで提案した“Skin-off”による3次元ビデオの表現・圧縮方法の有効性を検証する。本稿ではその中でも幾何情報に着目して定量評価を行った。入力データとして、2.1節で述べた手法によって形状と運動を同時復元した3次元メッシュデータを8フレーム用いた。各フレームの頂点数は13661、ポリゴン数は27318である。本実験では Geometry Images に対して適用する圧縮手法としてウェーブレット変換を採用している。

まず最初の実験として本手法で必要になるデータ量を求めた。Geometry Images の圧縮では、各フレームを個別に圧縮するのではなく、フレーム間の差分値を圧縮する方法が考えられる。この方法は各頂点の動き情報を圧縮することに相当するので、一様な動きの場合には高い圧縮効率が期待される。

表1にフレーム間差分を行わない場合と行う場合のデータ量を示す。ここでは、最終的に復号して得られる形状がほぼ同じになるように Geometry Images の大きさ<sup>§</sup>とウェーブレット変換の圧縮率を設定している。また、表中のデータ量は1秒あたりの値に換算している。図4は圧縮の結果得られた3次元メッシュの例を示している。(a)がオリジナル、(b)がフレーム間差分なし、(c)がフレーム間差分ありのデータである。この結果をみると、非常に小さなデータ量で幾何情報を表現でき、その場合でも人の目には区別ができないほどの幾何形状が生成できることがわかる。なお、(c)のメッシュにはわずかな皺が見られるが、これは、非可逆な圧縮を行ったために画素値(頂点座標)が変化したものと考えられる。

表1: 幾何情報の表現に必要なデータ量

オリジナル	フレーム間差分無	フレーム間差分有
30780kbps	1920kbps	450kbps

次に、Geometry Images を用いることによる幾何形状の歪みを定量的に評価する。2.2節で述べたとおり、Geometry Images への変換は、その過程で矩形画像上の格子点での再サンプリングを行うため(ウェーブレット変換などの非可逆圧縮を行わなくても)非可逆な変換になる。このため生じる歪みを定量的に評価するため

<sup>§</sup>今回の実験では65x65とした。頂点数4225に相当する。

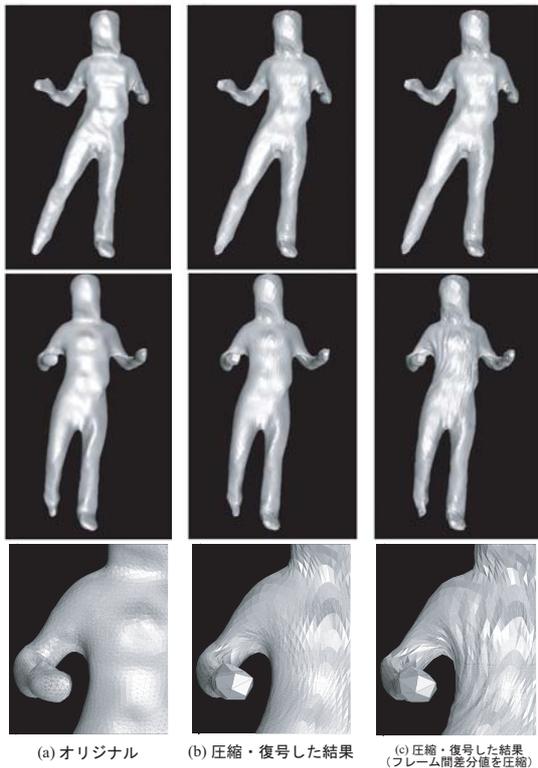


図 4: 3次元メッシュの例

に、頂点数で正規化したメッシュ間の距離  $D$  を  $D = \sum_{v \in M} d(v, M') / N_v$  と定義した。ここで、 $M, M'$  はそれぞれオリジナル、Geometry Image 変換後のメッシュであり、 $N_v$  は頂点  $v (\in M)$  の個数、 $d(v, M')$  は頂点  $v$  とメッシュ  $M'$  間の最短距離を表すものとする。

図 5 は Geometry Images のサイズとメッシュの歪みとの関係を示している。図中で横軸は時間経過、縦軸は  $D$  を示し、Geometry Images のサイズを 4 通りに変えて、それぞれの結果を示している。

この結果をみると Geometry Images が小さくなるほどメッシュの歪みが大きくなる、という基本的性質が確認できる。この中の 65x65 のデータは図 4 に示したデータに相当しており<sup>1)</sup>、この程度の歪みは人の目を見た場合はほとんど区別がつかないことが分かる。

今回の実験は幾何情報の表現・圧縮についてのものであったが、3. 節で述べたように頂点ベースの表現を用いれば幾何情報と光学情報を統一したデータ表現で扱うことができるため、提案手法は 3 次元ビデオ映像のすべての情報を効率的に表現・圧縮できるといえる。

## 5. まとめと今後の課題

本稿では“Skin-off”の考え方に基づいた、3次元ビデオ映像の表現および圧縮の手法を提案した。3次元ビデオ映像は、CGアニメーションとは異なり、1)各時刻で独立に求めた3次元メッシュの幾何情報の効率的な圧縮、2)誤差を含む幾何・光学情報からの写実的な映像の生成、という問題点がある。このような問題点に対し、

<sup>1)</sup>図 4(b) は 65x65 の Geometry Images を使い、ウェーブレット変換を行っているため図 5 のデータより歪みが大きくなっている。

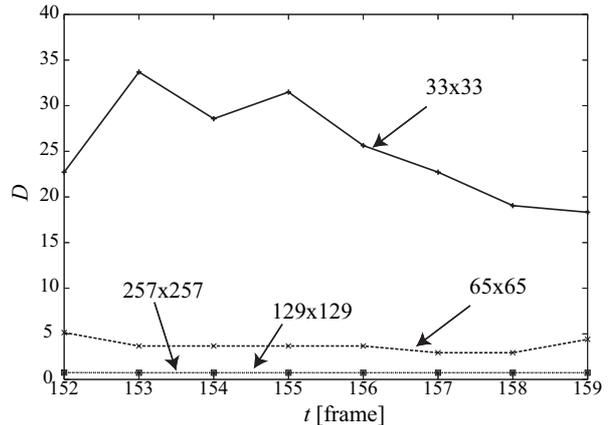


図 5: Geometry Images のサイズと歪みの関係

形状と運動の同時復元を用い、頂点ベースの色表現手法を導入することが有効であることを示した。

幾何情報の圧縮に関する実験の結果によると、低いビットレートでも十分観賞に耐えうる品質の幾何情報を伝送できることが示された。これは、携帯端末などでの 3 次元ビデオコンテンツの利用に際して、本手法が有効であることを示している。また、本手法は 2 次元矩形画像で 3 次元ビデオ映像を表現するため、圧縮処理に既存の映像コーデックを応用することができ、システムを容易に構築可能であると期待される。

今後の課題としては、2次元平面への展開アルゴリズムへの光学（テクスチャ）情報の導入、面ベースのテクスチャ情報表現の際のちらつき低減、さらには反射特性など視点独立な情報の抽出によるより写実性の高い映像生成の実現などが挙げられる。

なお、図 1 では分散配置されたカメラから視線がある対象に集中するような場合を挙げているが、逆に物体の中心にカメラを置いて周囲を見渡すような場合（全方位ビデオ）にも、Skin-off に基づく表現・圧縮法を適用できる。我々は、多面体スクリーンの展開によって全方位ビデオの表現・圧縮を行う手法を提案している [4]。

## 謝辞

本研究の一部は科学研究費補助金特定領域研究 13224051 の補助を受けて行った。

## 参考文献

- [1] G.Taubin et al.: Geometric Compression Through Topological Surgery, ACM Trans. on Graphics, Vol.17, No.2, pp.84–115, 1998.
- [2] X.Gu et al.: Geometry Images, ACM SIGGRAPH 2002, pp.355–361, 2002.
- [3] H.M.Briceno et al.: Geometry Videos: A New Representation for 3D Animations, ACM SIGGRAPH 2003, pp.136–146, 2003.
- [4] 波部他: 多面体モデルを用いた全方位ビデオの圧縮: 2003 年画像符号化シンポジウム, pp.57–58, 2003.
- [5] 松山他: 3 次元ビデオ映像の撮影・編集・表示, 日本 VR 学会論文誌, Vol.7, No.4, pp.521–532, 2002.
- [6] 上田他: 2 次元格子構造化とウェーブレット変換を用いたポリゴンメッシュデータの圧縮, 2003 年画像符号化シンポジウム, pp.93–94, 2003.
- [7] T.Matsuyama et al.: Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation, and High Fidelity Visualization for 3D Video, CVIU, 2004(in printing).