

3次元ビデオ映像のためのデータ圧縮法

木村 雅之 和田 俊和 松山 隆司

京都大学大学院情報学研究科 知能情報学専攻

概要 物体の3次元形状情報を持つ映像メディアである「3次元ビデオ映像」はそのデータ量が膨大なため圧縮が必要となる。この際用いる方法が形状を表現するメッシュの構成要素を可能な限り削減するメッシュの簡略化である。従来の簡略化アルゴリズムでは元のメッシュに与えられた情報をいかに保存するかが重要であったが、3次元ビデオ映像の場合は生成されるテクスチャが撮影された画像をどれだけ忠実に再現するかも大きな問題となることから従来のアルゴリズムをそのまま用いることはできない。そこで我々は従来用いられている幾何形状の変化の評価に加えて生成されるテクスチャの精度も評価する新たな簡略化アルゴリズムを提案する。

Data Compression Method for 3D-Video Movie

Masayuki KIMURA Toshikazu WADA Takashi MATSUYAMA

Department of Intelligence Science and Technology, Graduate School of Informatics,
Kyoto University

abstract 3D-Video is an ultimate image medium recording a dynamic visual events in the world as is. 3D-Video contents consist of huge amount of data and it is necessary to compress them. In this paper data compression is conducted by surface patch compression in terms of "mesh simplification". In the ordinary simplification method, the most important topic is how to preserve information given to the original mesh. but these existing algorithms are not suitable for 3D-Video data compression because simplification should be conducted so that textures mapped from multiple images coincide with each other. So we propose the new mesh simplification method that evaluate not only shape distortion and also accuracy of mapped texture, and we show the result of simplification.

1 はじめに

映像情報処理技術の発達に伴い、実写映像を用いた映像合成技術が実用化されつつある。その中でも特に「実写映像からの3次元データの復元」という手法は、任意のパラメータで映像を生成することができ、なおかつ様々な映像効果を加えることができる優れた手法である。このようにして獲得される3次元データを元にして構成される新しい映像メディアを「3次元ビデオ映像」と呼ぶことにする。本稿ではこの3次元ビデオ映像のデータ圧縮について述べる。

3次元ビデオ映像の各フレームは前景である運動物体の形状およびテクスチャの情報と背景情報からなっており、前景の情報は以下の方法で獲得される。(図1参照)

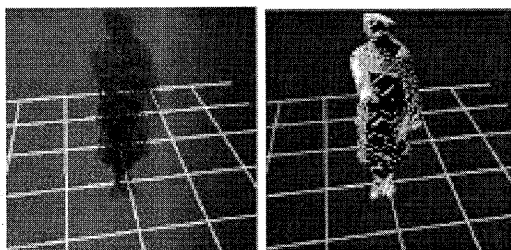
1. 多視点から同期撮影された画像から撮影対象のシルエットを抽出し、視体積交差法を用いて対象の3次元形状を復元する。[1]
2. 得られた形状データに対して離散マーチング・キューブ法 [2] を用いてメッシュ構造を獲得する。
3. 得られたメッシュの各面に対して撮影画像群から表面テクスチャ情報を作成する。

ところが、このようにして獲得されたデータを時間軸方向に並べるとデータ量が非常に膨大になるので圧縮が必要となる。データ圧縮のアプローチとしては

時間軸方向の冗長性を利用した圧縮

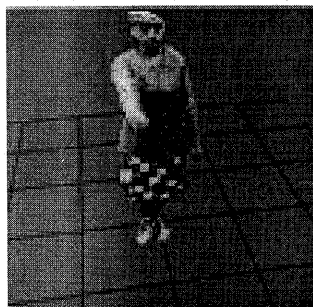
個々のフレームを構成するデータの圧縮

の2つが考えられるが、本研究では後者の個々のフレームを構成するデータの圧縮を取り扱う。そしてこの際に用いる方法がメッシュの構成要素を可能な限り少なくするメッシュの簡略化である。また、メッシュの簡略化にはデータ量の削減のほかにも副次的ではあるがメリットが期待できる。それはメッシュの質の向上である。現在3次元ビデオにおいては視体積交差法とマーチング・キューブ法が用いられているが、これらの結果得られるデータは精度が粗く、また冗長な部分が多い。メッシュの精度の粗さは生成されるテクスチャの質の低下の原因となり、冗長なデータは描画速度の原因となる。簡略



(a) volume data

(b) surface patch



(c) texture mapped mesh

図 1: 3次元ビデオフレーム

化によってこれらの問題が解決されることが期待できる。

以下本稿の構成を記す。第2章でメッシュの簡略化の方針についても少し説明する。第3章で簡略化の具体的な方法を述べ、第4章では実際に簡略化を行った際の結果を述べる。

2 メッシュの簡略化

メッシュの簡略化自体はすでにコンピュータ・グラフィックの分野で広く研究されており、様々なアルゴリズムが提案されている。この際に基本となっている考え方は「元のメッシュに与えられている幾何形状や面の法線、あるいは表面テクスチャなどの情報をいかに損なわないように簡略化を進めるか」というものである。しかし3次元ビデオの各フレームに対してこのような既存のメッシュ簡略化アルゴリズムを用いるのは適当でない。なぜなら3次元

ビデオフレームのテクスチャ情報は生成されたメッシュに撮影画像を投影することで獲得されるため、その映像的価値を高めるには「生成されるテクスチャが元の撮影画像をどれだけ忠実に反映しているか」を簡略化の際にあわせて評価する必要があるからである。言い換えるならば3次元ビデオ映像のデータ圧縮は元のメッシュに与えられた情報をできるだけ損なわず、なおかつ生成されるテクスチャが撮影画像を忠実に再現するようなメッシュを構成できるメッシュ簡略化処理によって実現される。

3 簡略化の方法

メッシュの簡略化の際にメッシュに加える基本的な操作は Edge Collapse と呼ばれる図2のような操作である。

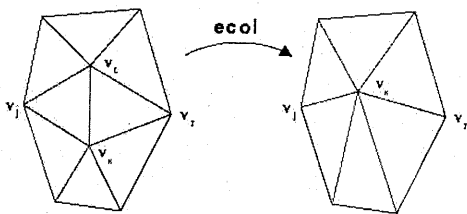


図 2: Edge Collapse

これはメッシュのある辺を縮退させ、ひとつの頂点にする操作である。メッシュのどの辺にこの操作を行い、新しくできる頂点の位置をどこにするかを決定するために、次の二つの評価基準を求める。

3.1 評価基準その1: 簡略化に伴う幾何形状の変化

これはメッシュのある辺に注目した際に、その辺の近傍面と簡略化後の頂点位置との距離の平方和を表す Quadric Error Metric という評価値を [3],[4] を用いている。この手法は計算が簡単である上にこれを最小化することで得られる頂点位置が元の形状をうまく表すという性質がある。以下この評価値の計算方法を述べる。

メッシュ中の任意の面 f について、 f を含む平面と \mathbf{R}^3 上の点 v との距離の平方を $Q^f(v)$ とする。

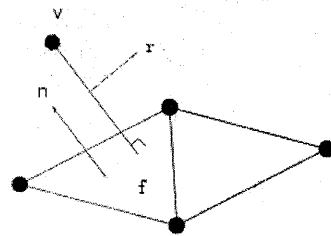


図 3: Quadric Error Metric

メッシュの各頂点 v について、これを要素に持つ面の $Q^f(v)$ をその面積で重みづけしてたし合わせたものを新たに $Q^v(v)$ として定義する。すなわち、

$$Q^v(v) = \sum_{f \ni v} \text{area}(f) \cdot Q^f(v) \quad (1)$$

である。さらに辺 $e = (v_s, v_e)$ について、 $Q^e(v) = Q^{v_s}(v) + Q^{v_e}(v)$ という値を定義する。この $Q^e(v)$ を最小化する v が、辺 e に対して Edge Collapse 操作を適用した際の新しい頂点 v_n の位置になる。簡略化の際にはメッシュのすべての辺に対して $Q^e(v)$ の最小値を求め、その値が小さいものほど優先的に Edge Collapse の対象になる。

3.2 評価基準その2: テクスチャ情報の利用

テクスチャの状態を判断する基準として各撮影画像から生成されるテクスチャ間の画素値のばらつきを利用するものとする。これが小さいほど生成されるテクスチャは元の撮影画像を忠実に反映している、すなわち良いテクスチャであると考えられることができる。実際の評価は注目している辺に対して Edge Collapse を適用し、変形した面に対して画素値の分散を求め、その総和を評価値とする。

3.3 簡略化アルゴリズム

以上のような評価値を用いて以下のように簡略化を進める。

1. メッシュの各辺に対し簡略化に伴う幾何形状の変化の大きさ $Q(e)$ を求める。

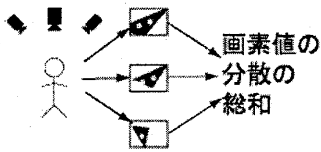


図 4: テクスチャの評価

2. $Q(e)$ の値が小さい順にいくつか (本実験では 12 に固定) の辺を選び, これらの辺についてテクスチャの評価を行い, 評価値 $\sigma^2(e)$ を求める.
3. 1.2. で得られた評価値をそれぞれ最小値が 0, 最大値が 100 になるように規格化し, 以下の式から総合評価値 $m(e)$ を求める.

$$m(e) = (1 - w) \cdot Q(e) + w \cdot \sigma^2(e) \quad (2)$$

w は 2 つの評価値に対する重みづけを行うパラメータで, $(0 \leq w \leq 1)$ を満たす.

4. $m(e)$ が最小となる辺に Edge Collapse 操作を適用し, メッシュを簡略化する
5. 所望の回数だけ 1. から 4. を繰り返す.

4 結果とその評価

4.1 評価方法

簡略化処理の結果を評価するための評価量として, 簡略化後のメッシュに対してマッピングされたテクスチャを撮影画像側に逆投影し, 得られた画像と対応する撮影画像の正規化相互相関を求める. これはマッピングされたテクスチャが元の撮影画像とどれだけ類似しているかを表し, これによってテクスチャマッピングの精度が数値的に評価可能となる. 正規化相互相関は以下のようにして求められる.

投影されたテクスチャの点 (x, y) における画素値を $I_a(x, y)$, 対応する撮影画像の同じ点の画素値を $I_b(x, y)$ とする. 正規化相互相関 c は,

$$c = \frac{\sum_x \sum_y I_a(x, y) I_b(x, y)}{|I_a| |I_b|} \quad (3)$$

ただし,

$$|I| = \left\{ \sum_x \sum_y \{I(x, y)\}^2 \right\}^{\frac{1}{2}} \quad (4)$$

である.

4.2 実験結果

実験に使用したフレームは面の数が 2524, 辺の数が 3786 のメッシュと, 7 台のカメラから撮影された画像によって生成されたテクスチャを持つ. 重みづけパラメータ w の値を

1. $w = 0$

2. $w = 0.5$

3. $w = 1$

としてそれぞれの方法で簡略化を行った. $w = 0$ の場合は従来の幾何形状のみを評価する手法と同一である. それぞれの方法で簡略化を行った際の正規化相互相関の平均値の推移をグラフおよび表にしたものが図 5 および表 1 である. また実際に簡略化を行った際の結果を図 6 に記す.

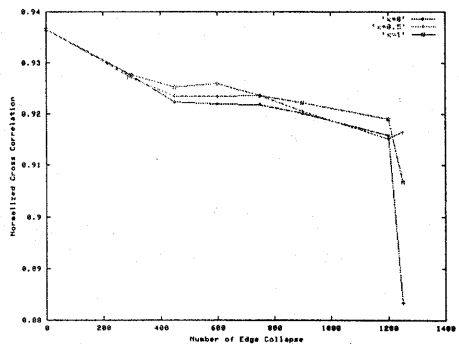
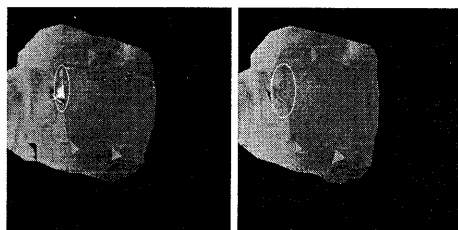


図 5: 正規化相互相関の推移

グラフを見てもわかるようにいずれの手法でも簡略化に伴い正規化相互相関は低下している. このことは簡略化がメッシュの質の向上にはあまり役に立っていないことを表していると考えられる. しか

簡略化の回数	w=0	w=0.5	w=1
0	0.936597	0.936597	0.936597
300	0.927519	0.927107	0.927666
450	0.922433	0.923505	0.925252
600	0.921947	0.923461	0.926022
750	0.921875	0.923661	0.923701
900	0.920195	0.920603	0.922312
1200	0.915954	0.915169	0.919092
1250	0.883344	0.916514	0.906683

表 1: 簡略化過程における正規化相互相関の推移



(a) 従来の手法 (w=0) (b) 新しい手法 (w=1)

図 6: 簡略化の例

し従来の幾何形状のみを評価した簡略化手法よりも本研究の手法のほうがその低下の度合いが小さい。これは本研究で用いた手法が従来の手法に比べて簡略化に伴うテクスチャ情報の劣化を抑えることができてきていることを示している。これにより「生成されるテクスチャを撮影画像により忠実にする」という目標が達成できたと考えることができる。

5 まとめ

本論文では3次元ビデオ映像のためのデータ圧縮と題して幾何形状の変化とテクスチャの精度をあわせて評価する簡略化アルゴリズムを提案し、生成されるテクスチャの質の向上にいくらかの効果があることがわかった。今後はアルゴリズムの高速化や、簡略化後のメッシュの形状決定の際にテクスチャの評価を反映させること、時間軸方向の冗長性を利用した圧縮、およびテクスチャデータの圧縮などに取

り組む予定である。

参考文献

- [1] Toshikazu Wada, Xiaojun Wu, Shogo To-lai, Takashi Matsuyama. Homography Based Parallel Volume Intersection: Toward Real-Time Volume Reconstruction Using Active Camera. *CAMP2000 Computer Architectures for Machine Perception Proceedings*(2000), 331-339
- [2] 剣持 雪子, 小谷 一孔, 井宮 淳. 点の連結性を考慮したマーチング・キューブ法 電気通信情報学会技術研究報告 (PRMU98-218)(1999), 197-204
- [3] Garland, M., and Heckbert, P. Surface simplification using quadric error metrics. *Computer Graphics (SIGGRAPH '97 Proceedings)*(1997), 209-216
- [4] Hoppe, H. New Quadric Metric for Simplifying Meshes with Appearance Attributes. *IEEE visualization '99 Proceedings*(1999), 59-66