

# ビデオ画像からの形状復元のための三角網の最適化

中辻 敦忠\* 菅谷 保之† 金谷 健一†

\*NEC エンジニアリングインターネットターミナル事業部 †岡山大学大学院自然科学研究科

前報 [10] で提案したシーンの多面体表示のための三角網の最適化手法を, 2 画像だけでなく連続ビデオ画像にも適用可能な形に拡張する. また, テンプレートサイズを可変にして計算時間を効率化する方法を示す. そして, 実画像例を用いて従来手法と比較し, 提案手法が優れていることを実証する. その過程で従来手法の問題点を明らかにし, 提案手法の特性と対比する. 提案手法は入力画像間を比較するのみで, シーンの 3 次元情報は一切必要としない. また, 判定のための何らのしきい値も必要としない.

## Optimizing a Triangular Mesh for Shape Reconstruction from Video Images

Atsutada Nakatsuji\* Yasuyuki Sugaya† Kenichi Kanatani†

\*Internet Terminals Division, NEC Engineering, Ltd., Yokohama-shi, Kanagawa 224-0053 Japan

†Department of Computer Science, Okayama University, Okayama-shi, Okayama 700-8530 Japan

We extend our previously proposed method [10] for optimizing a triangular mesh for polyhedral representation of the scene from two images to a video stream. We also introduce a variable-size template to make the computation efficient. Using real images, we demonstrate that our method is superior to existing methods. At the same time, we reveal the problems inherent in existing methods and make clear the characteristics of our method. Our method is based only on comparison of input 2-D images, not requiring any 3-D information about the scene. Also, no thresholds to be adjusted for judgment are required.

### 1. まえがき

画像からシーンの 3 次元形状を復元するとき, 復元した形状をどう表現するかが大きな問題である. 代表的な方法は, 未校正カメラで撮影した画像上に特徴点を抽出し [7], それら特徴点間の対応を決定し [6, 15], その特徴点の 3 次元位置を計算し [1, 5], それを頂点とする三角網を定義して物体を多面体表示し, テクスチャマッピングを施して表示する方法である. これに必要な三角網は, 指定した画像上で特徴点を頂点とするドロネー三角形網を用いればバランスのとれたメッシュが自動的に生成できる.

しかし, 前報 [10] で指摘したように, シーン自体が多面体の場合は, ドロネー網によって定義される多面体の辺と物理的な辺が合致せず, 矛盾した 3 次元形状が復元されることがある. そこで前報 [10] では, 矛盾を検出するテンプレートを導入して三角網を物理的な形状に適合するように最適化する手法を提案した. 室内や建物などの人工的な環境では物体のほとんどが多面体であるから, この方法は実際の応用に非常に有益である.

しかし, 前報 [10] では 2 画像間の特徴点の対応に対する三角網の最適化のみを考えた. 本論文では連続ビデオ画像から特徴点を追跡して因子分解法 [4, 12]

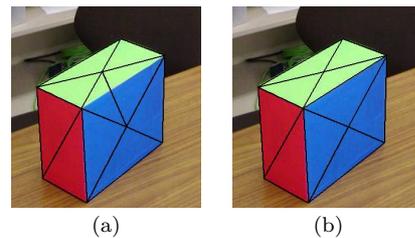


図 1: (a) 物理的な辺に矛盾する三角網. (b) 物理的な辺に適合する三角網.

によって 3 次元形状復元することを想定し, 前報 [10] の方法を多画像による方法に拡張する. そして, 実画像を用いて従来手法 [9, 11] と比較し, 提案手法が優れていることを実証するとともに, 従来手法の問題点を明らかにして, 提案手法の特性と対比する.

また, ビデオ画像ではフレーム数に比例して処理時間が増加するが, 本論文では前報 [10] では固定していたテンプレートサイズを可変にし, 正解率を落さずに計算が効率化できることを示す. さらに多面体シーンだけでなく曲面物体にも適用して, 形状表現に適した三角網が生成されることを示す.

### 2. 最適化の原理

図 1(a) のように物体頂点が特徴点として選ばれたとする. これから図 1(a) に示すドロネー網が得られるが, この多面体表現は物体形状を正しく表していない. 一方, 図 1(b) の三角網は物体形状を正しく表している. 本論文でいう三角網の“最適化”とは図 1(a)

\*223-0053 横浜市都筑区池辺町 4388, Tel: (045)939-2867

Fax: (045)939-2874, a-nakatsuji@pb.jp.nec.com

†700-8530 岡山市津島中 3-1-1, Tel/Fax: (086)254-8173, {sugaya,kanatani}@suri.it.okayama-u.ac.jp

のような物理的な辺と矛盾する三角網を図 1(b) のような物理的な辺に適合する三角網に自動的に変換することである。

従来からよく研究されているのは、(1) 多数の頂点による密な三角網を 3 次元形状の記述を損なわないように簡素化する手法、および (2) 少数の頂点からなる疎な三角網から出発して、3 次元データに適合するように辺や頂点を追加して記述を詳細化する方法である。前者では例えば Vogiatzis [13] らが真の形状は平面部分が多いという事前確率を用いるベイズ推定にアンニリングを適用し、後者では例えば Yu [14] らが物体の形状と光反射モデルを推定しながら 3 次元形状を詳細化する手順を述べている。

これらはセンサデータから 3 次元形状を復元する全過程における処理であるが、本論文では 3 次元構造や 3 次元位置データを用いずに、入力 2 次元画像を比較するのみで、与えられた三角網の接続関係を変更する最適化を考える。これは、これを一つの基本処理として独立させ、3 次元復元に限らず多くの応用の部分技法として確立させるためである。

このための基本演算は、誤った辺を、それを囲む四辺形のもう一つの対角線（以下「対角辺」と呼ぶ）と入れ換えることである。ただし、“誤った辺”とは図 1(a) にあるような、シーンの異なる平面上の 2 点を結ぶ辺のことである。このような対角辺との入れ替えを続ければ、物理的に適合する多面体表現が（あれば）得られる [9]。

問題はどの辺が誤っているかをどう検出するかである。これに対して Morris ら [9] と Perrier ら [11] は、3 次元データを用いずに（3 次元復元は行わずに）、2 次元画像のみから計算する方法を提案した。その原理は、シーンの平面部分を表す三角形パッチは他の画像の対応する三角形パッチにアフィン変換<sup>1</sup>で写像され、平面部分でなければ写像後に輝度差が現れるという事実である。

これを利用して、Morris ら [9] は全パッチを対応するパッチに写像したときの残差平方和を最小にするように辺の入れ換えを繰り返すグリーディ法を用い、Perrier ら [11] は残差平方和の代わりに正規化相関を用いた。各三角形パッチの輝度値から定数を引いてパッチ内の平均輝度値を 0 に正規化し、さらに定数で割ってパッチ内の輝度値の分散を 1 に正規化しておけば、その残差平方和の最小化と正規化相関の最大化は等価であるから（付録 A）、Morris ら [9]

<sup>1</sup> 実際のカメラでは、平面部分は射影変換で結ばれるが、小さく分割した三角形パッチ同士を考える限り、アフィン変換との差は無視できる。

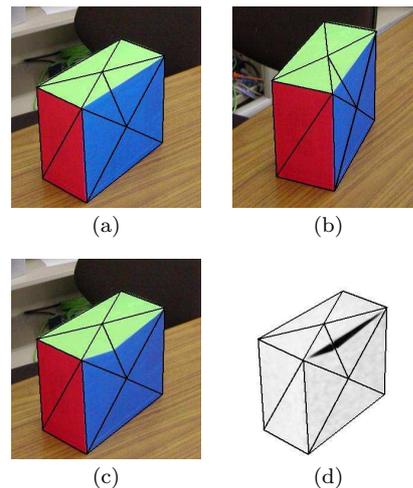


図 2: (a), (b) 入力画像と初期三角網。(c) 図 (b) を図 (a) にテクスチャマッピングしたもの。(d) 図 (a) と図 (c) との差画像。

の方法と Perrier ら [11] の方法は本質的には輝度値を正規化するかしらないかの違いでしかない<sup>2</sup>。

わが国でも飯田ら [2] が Morris ら [9] と同じような考え方を述べていた。栗原ら [8] は Morris ら [9] の方法で実画像実験を行い、2 画像間の照明変化を打ち消す必要性を指摘した。これは実質的に Perrier ら [11] に他ならない。

これに対して前報 [10] では幾何学的、光学的考察から輝度値の差が最も現れやすい領域（「不適合領域」と呼ぶ）を検出し、それを辺の入れ換えの尺度とした。そして、そのような不適合領域を最もよく検出するテンプレート（「不適合検出テンプレート」と呼ぶ）を設計し、実画像を用いてその有効性を実証した。まず次節で前報 [10] の方法の要点を述べ、その後これをビデオ画像へ拡張する方法を述べる。

### 3. 不適合検出テンプレート

前報 [10] の誤った辺の検出法の原理を図 2 に示す。図 2(a), (b) は多面体物体の画像上に初期三角網（図 2(a) の頂点に対するドロネー網）を表示したものである。図 2(c) は図 2(b) の各三角形パッチを図 2(a) の対応する三角形パッチ上にテクスチャマッピングしたものであり、図 2(d) は図 2(a) と図 2(c) の差画像<sup>3</sup>である。このように、輝度差は誤った辺を横切る細長い三角形領域（不適合領域）に顕著に現れることがわかる。そこでこれを検出するために次の不適合

<sup>2</sup> それ以外に計算の仕方や探索の仕方に若干の相違があるが、本質的ではない。

<sup>3</sup> 本論文ではカラー画像を考え、R, G, B 値の各々の差の二乗和の平方根を「輝度差」、それを画像として表示したものを「差画像」と呼ぶ。

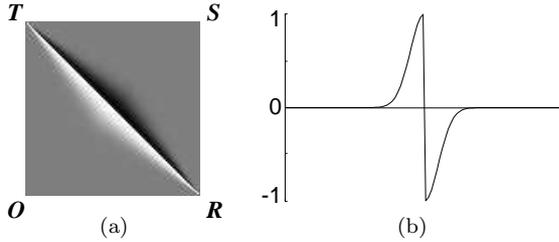


図 3: (a) 誤った辺を検出するテンプレートの濃淡表示 (白が正, 黒が負). (b) 対角線  $OS$  に沿った切り口の値.

検出テンプレートを定義する.

$$T(x, y) = \begin{cases} e^{-\frac{(x+y-l)^2}{2\alpha^2(x-y-l)^2}} & x+y < l, x \geq y \\ T(y, x) & x+y \leq l, x < y \\ -T(l-y, l-x) & x+y > l \end{cases} \quad (1)$$

図 3(a) はこれを濃淡表示したものである. 大きさは  $l \times l$  であり<sup>4</sup>, 実験では  $\alpha = 0.1$  とした.  $T(x, y)$  の等高線は  $R, T$  を端点として  $OS$  上に頂点を持つ折れ線となっている. そして,  $OS$  に関して対称であり,  $RT$  に関して反対称 (対称な値の符号を換えたもの) である. 図 3(b) はその対角線  $OS$  上の切り口の値のグラフであり, 平均  $l/\sqrt{2}$ , 標準偏差  $\alpha l/\sqrt{2}$  の正規分布の片側の符号を換えたものになっている.

反対称なテンプレートを用いるのは, 差の現れる不適合領域がどちら側に現れるかあらかじめわからないためである. 反対称にすれば, それ領域がどちらにあっても反対側の輝度差はほとんど 0 であるから, 相関の絶対値を計算することによって不適合が検出できる. さらに, テクスチャマッピングのずれのためにパッチ全体に分布するランダムな誤差パターンや照明変化による一様な輝度変化はこのような反対称なテンプレートによって正負が打ち消される.

ただし, 特徴点の検出の誤差によって三角形パッチの  $\triangle ORS, \triangle OTS$  上への写像にわずかなずれが生じることがあるので, 図 3 のテンプレートの対角線  $RT$  上の画素, および対角線  $OS$  の上下 2 画素とフレームの周囲の 2 画素の値は 0 とする.

前報 [10] ではこの不適合検出テンプレートによる出力の大きい辺から順に対角辺と入れ換えることによって三角網の最適化を行なった.

<sup>4</sup>前報 [10] では  $l = 100$  としたが, 本論文では後述のように可変にする.

## 4. ビデオ画像への拡張

### 4.1 最適化の手順

Perrier ら [11] は 2 画像の場合しか考慮していないが, Morris ら [9] は多画像の場合に各三角形パッチの全フレームに渡る平均と各パッチとの残差平方和を最小化している<sup>5</sup>. 本論文では, 与えられた辺  $AB$  の不適合の程度を測る前報 [10] の評価値  $w(AB)$  を次のようにビデオ画像に拡張して最適化を行う.

1. 辺  $AB$  の片側に一つの三角形しかない場合は境界辺であるから,  $w(AB) = -1$  を返す.
2. 辺  $AB$  の両側の三角形を  $\triangle ABP, \triangle ABQ$  とするとき, 四辺形  $APBQ$  が凸でないフレームが一つでも存在すれば,  $w(AB) = 0$  とする.
3. 値を 0 に初期化した  $l \times l$  正方形領域  $ORST$  を用意し, 以下を  $\kappa = 1, \dots, M$  に対して計算する.
  - (a) 第  $\kappa$  フレームの四辺形  $APBQ$  を正方形領域  $ORST$  に射影変換し, 輝度値を足す.
  - (b) 第  $\kappa$  フレームの  $\triangle ABP, \triangle ABQ$  をそれぞれ正方形領域  $ORST$  の  $\triangle OSR, \triangle OST$  にアフィン変換し, 輝度値をそこに書かれている値から引く.
4. 得られた正方形領域  $ORST$  の値と検出テンプレートとの相関の絶対値を  $w(AB)$  とする.

ステップ 3 はどのフレームも同等に扱う対称な操作であり, その意味を図 4 に示す. ステップ 3(a) は考えている辺  $AB$  を囲む四辺形の形がフレームごとに異なるので, 同じ形 (正方形) にそろえる操作である.

ステップ 3(b) のように辺  $AB$  両側の三角形をそれぞれアフィン変換すると, 辺  $AB$  が正しければ写像した結果はステップ 3(a) で得られる射影変換とほぼ同じになるが, 誤っていれば輝度差が辺  $AB$  を横切る不適合領域に現れる. これはシーンの見え方によって辺  $AB$  を横切る対角線のどちらかに現れるか一定しないが, どちら側かで符号が異なる. これを全フレームに対して合計すると, 図 4 の右のように, 辺  $AB$  を横切る対角線を境界として両側に符号の異なる不適合領域が現れる (一方のみに現れる場合もある). これと図 3(a) のテンプレートとの相関をとることによって不適合の程度を測定する.

ステップ 2 は, 四辺形  $APBQ$  がどのフレームかで凹であれば (付録 B), 入れ換えるべき対角辺が内部にないので, 入れ換えは行わないという意味である.

以下, 前報 [10] と同様に,  $w$  の値の大きい辺から入れ換えていく. そして, 入れ換えた辺の  $w$  を計算

<sup>5</sup>Morris ら [9] はそのパッチが見えないフレームを除去し, 見えるパッチの面積で重みづけしている.

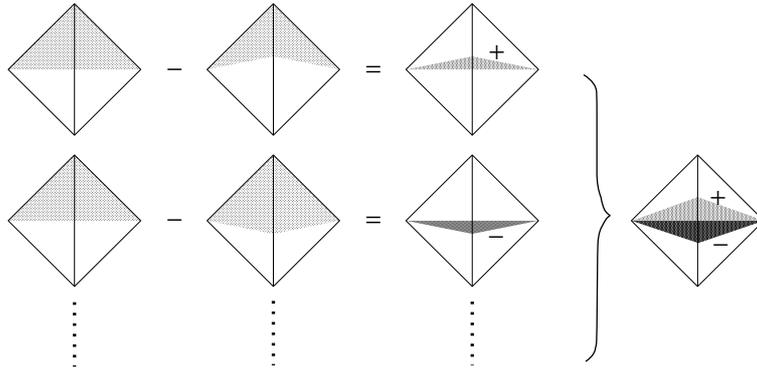


図 4: ビデオ画像からの不適合の検出の原理 .

し,それがより大きくなれば入れ換えを元に戻す.これをすべての辺に行い,前報 [10] と同様にもはや三角網が変化しなくなるまでこれを反復する.この過程で何らのしきい値を導入する必要はない.この操作を2画像 ( $M = 2$ ) に施すと,前報 [10] と同じ計算になる .

#### 4.2 裏返しの除去

初期三角網はどれかのフレームで,与えられた特徴点を頂点とするドロネー分割を行い,それを他のフレームに同型写像する.ただし,2画像の場合と同様にどの三角形パッチも全フレームで同じ“向き”であり,“裏返し”が生じてはならない.カメラを一方方向に移動した場合は,経験的には画像系列の中間のフレームを選んだほうが裏返しが生じにくい.

もしどれかのフレームかで裏返しになれば(付録 C),前報 [10] と同様にしてこれを除去する.すなわち,それが境界辺で生じるならその辺を除去する.内部で生じるときは,その三角形のどの辺かを対角辺と取り換えて裏返しが消されるものを選ぶ.どの辺を取り換えても解消されないものがあれば<sup>6</sup>,その3辺は以降では交換しない.

#### 4.3 可変テンプレート

前報 [10] では固定した  $100 \times 100$  画素の不適合検出テンプレートを用いたが,三角形パッチやフレームの数が増えると計算時間が極めて増大する.しかし,同時に各々のパッチは小さくなり,これを  $100 \times 100$  領域に拡大しても実質的なメリットはない.画像上のパッチとほぼ同じ大きさの領域に写像して比較すれば十分である.そこで本論文では, $l \times l$  がすべてのフレームのすべての三角形パッチの平均面積のほぼ2倍であるように  $l$  を定めた.これに伴って,パッチ境界の影響を除く範囲も幅  $0.02l$  画素とした.

<sup>6</sup>これは,すべてのフレームでシーンが隠れなしに見える状況では生じない.

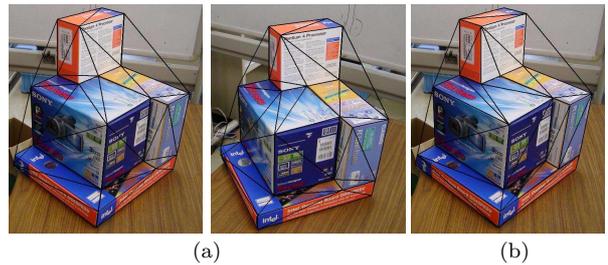


図 5: (a) 初期三角網 (辺数 58) . (b) 提案手法による最適化 (3.43 秒/3.89 秒) .



図 6: (a) 初期三角網 (辺数 31) . (b) 提案手法による最適化 (3.15 秒/2.89 秒) .

#### 4.4 偽の境界線除去

さらに前報 [10] では物体境界の凹部に発生する偽の境界線を除去する手法を述べた.これは各画像ごとに実行できるが,前報 [10] では2画像の両方で偽の境界線と判定された場合にその辺を除去するようにした.これは,偽の境界が残っても実際上の応用では大きな支障ではないのに対して,真の境界を誤って除去すると内部の形状が正しく表せず,したがって2画像で判定が分かれるあいまいな辺を除去するのは危険だからである.本論文でも同様に考え,各フレームでチェックし,全フレーム数の77%以上のフレームで偽の境界線と判定された場合のみその辺を除去するようにした.

### 5. 実画像実験

#### 5.1 手法の比較

Morris ら [9] の方法,Perrier ら [11] の方法,および提案手法を前報 [10] と同じ2画像例(特徴点は手

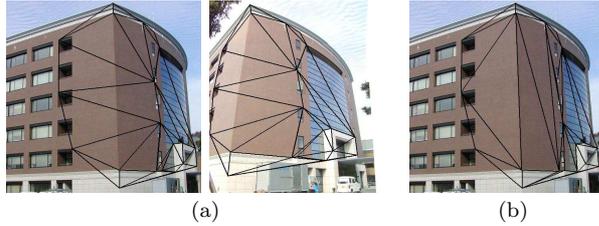


図 7: (a) 初期三角網 (辺数 47) . (b) 提案手法による最適化 (4.03 秒/5.26 秒) .

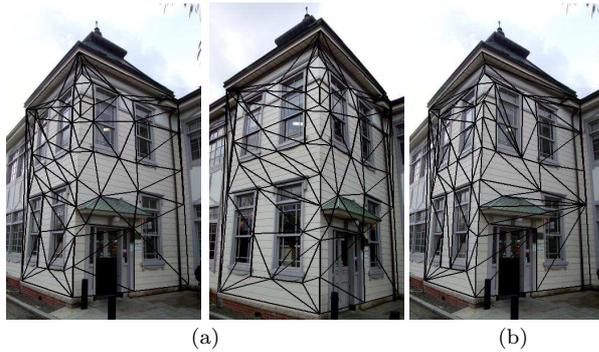


図 8: (a) 初期三角網 (辺数 157) . (b) 提案手法による最適化 (11.85 秒/50.40 秒) .

動で指定) に適用して性能比較を行った<sup>7</sup> .

図 5~9 の (a) は入力 2 画像と初期三角網, (b) は提案手法で最適化した三角網を第 1 画像上に重ねたものである (偽の境界線は残している) . 図説中に辺数と計算時間 (秒) を示す (前の数字が可変テンプレートを用いる場合, 後の数字は 100 × 100 固定テンプレートを用いる場合) . ただし, CPU には Pentium 4 3.2GHz, 主メモリ 2GB, OS には Linux を用いた .

この実験から, 可変テンプレートを用いても前報 [10] と同じ結果が得られることが確認される . そして, パッチ数が多いときに実行時間が著しく減少している . 一方, 図 6 のように, パッチ数が少なく可変テンプレートサイズが固定テンプレートより大きくなるとときには計算量は当然増加する .

表 1 は 3 手法による正解率 (%) を示す . 正解率とは境界辺を除くすべての辺に対する正しい (すなわち端点と同じ平面上にある) 辺の割合である . 異なる特徴点が同一平面上にあるかどうかは目視で判定した .

3 手法を比較すると, 全般的に提案手法が Morris ら [9] や Perrier ら [11] の方法より優れていることがわかる . 図 8 に対してパッチ全体を比較する Morris ら [9] の方法が提案手法を上回ったのは, 誤りを生じ

<sup>7</sup>Morris ら [9] は交換するとパッチ間の残差平方和の減少が最大になる辺から交換しているが, Perrier ら [11] は両側のパッチ間の正規化相関が最小の辺から交換している . ここでは輝度値の正規化以外は Morris ら [9] の方法に合わせて比較した .

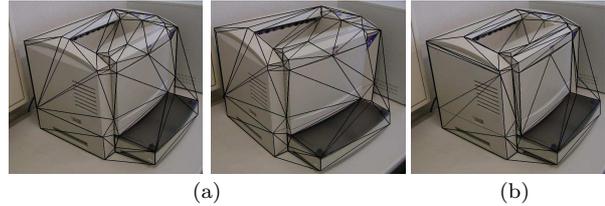


図 9: (a) 初期三角網 (辺数 114) . (b) 提案手法による最適化 (3.15 秒/3.34 秒) .

表 1: 最適化した辺の正解率 (%) .

	図 5	図 6	図 7	図 8	図 9
Morris ら	100	73.9	89.2	100	92.4
Perrier ら	95.7	69.6	91.9	98.7	85.7
提案手法	100	100	100	98.7	96.2

た三角形パッチが小さ過ぎて不適合領域が十分に抽出できなかったためと思われる .

一方, Perrier ら [11] の方法は照明変化を打ち消すために輝度値を正規化したものであるが, 2 画像間で日照変化の激しい図 7 以外は Morris ら [9] の方法に比べて正解率が悪化している . 輝度値を正規化するのは対応するパッチ間の類似性を増加させる目的であるが, 副作用として対応しないパッチ間の類似性も増加してしまう . それに対して, 提案手法は反対称テンプレートを用いているので, 照明変化に影響されない .

図 7 のように複数のパッチが物理的な辺を横切ると, 対角辺の入れ換えが一回では不十分で, 連続した入れ換えが必要である . このような場合の不適合領域を観察すると, 図 2(d) のような明瞭な三角形ではなく, やや狭まった領域に縮んでいる . しかし, 図 3(a) の検出テンプレートは依然として有効であり, 誤った辺が次々に修正されて, 最終的にすべての辺が正しい位置に置かれた .

## 5.2 パッチの類似度と正解率の関係

使用した画像例の中で, 提案手法に比べて Morris ら [9] の方法が劣った図 6 の例について, その原因を調査した . 図 10(a) は Morris ら [9] の方法の反復回数 (横軸) に対する対応パッチの輝度値の残差平方和 (左縦軸) と辺の正解率 (右縦軸) をプロットしたものである . Morris ら [9] の方法は残差平方和を最小化するものであり, 確かに残差平方和は単調に減少している . そして最初の反復で辺の正解率は増加している . しかし, 次の反復で残差平方和が減少しているにもかかわらず正解率が減少している .

図 10(b) は Perrier ら [11] の方法の反復回数 (横軸) に対して輝度値を正規化した残差 (左縦軸) と

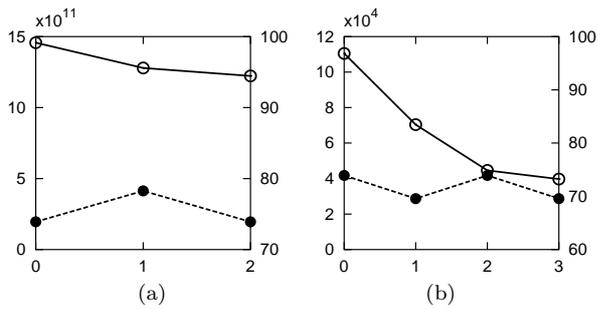


図 10: 横軸: 反復回数. 左縦軸: 対応パッチの類似度の評価 (—○—). 右縦軸 (%): 正解率 (---●---). (a) Morris らの方法. (b) Perrier らの方法.

正解率 (%) をプロットしたものである. Perrier ら [11] の方法はこの正規化残差平方和を最小化するものであり, 実際に単調に減少している. しかし, 正解率は大きく変動している.

以上より, 輝度値を正規化してもしなくても誤った三角網のほうが正しい三角網よりパッチ全体の残差が少ないという逆転現象が生じることがあり, Morris ら [9] や Perrier ら [11] の用いた対応パッチの類似度 (残差平方和および正規化相関) が辺の正しさの評価として適切ではないと結論される. それに対して, 提案手法はパッチ全体ではなく, 不適合が現れる狭い領域を選択して評価しているため, 100%の正解率が得られたと考えられる.

### 5.3 曲面物体への適用

提案手法はシーンが多面体であるという前提で, 不適合が生じやすい領域を検出しているが, これを曲面物体に適用するとどうなるかを調べた. 図 12 に 2 例を示す. いずれも左がドロネー網であり, 右が提案手法による最適化である. これは 2 画像を用いた例であり, 第 1 画像のみを表示している. 最初の例では金澤・金谷の方法<sup>8</sup>[6] で特徴点とその対応を自動的に抽出し, 後の例では車両上の特徴点を手動で指定した. いずれも提案手法によって物体形状に近い多面体表現に変換されている.

### 5.4 ビデオ画像への適用

2 画像のみ用いる場合と連続ビデオ画像を用いる場合の比較を行った. 図 11 の (a), (b), (c) はそれぞれ 30 フレームの画像列であり, その中から 5 フレームを抜き出したものである. 画像中には追跡した特徴点をマークしている. この特徴点は初期フレームに手動で指定し, 以降のフレーム上を Kanade-Tomasi-

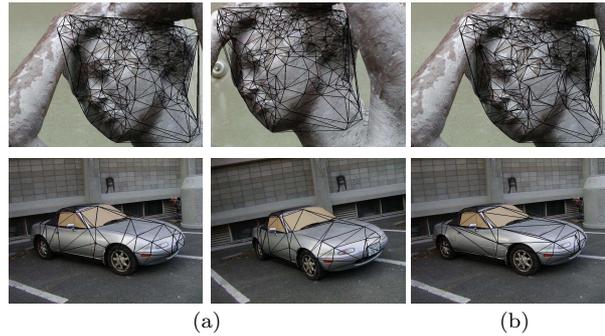


図 12: 曲面物体の三角網. (a) 初期三角網, (b) 提案手法による最適化.

Lucas の方法<sup>9</sup>によって追跡した. ただし, 追跡が途絶えたら手動で再追跡を開始した.

次に, 得られた特徴点を用いて, 系列の中央のフレームを用いてドロネー網を定義した. 5 節で述べたように, こうすると裏返しが生じにくいからである. 得られた初期三角網を 3 手法で最適化し, 正解率 (%) と実行時間 (秒) を表 2 に示す. また, 比較のために初期フレームと最終フレームの 2 画像のみを用いた最適化も行った.

Perrier ら [11] は 2 画像の場合しか考慮していないが, ここでは Morris ら [9] と同様にして多画像に拡張した. また, Morris ら [9] も Perrier ら [11] もパッチ同士を比較しているが, ここでは提案手法の技法を入れてテンプレート領域に射影変換して形を合わせてから比較した. また, 効率化のためにそのテンプレート領域も提案手法と同様に可変サイズとした.

表 2 から, 提案手法は全部のフレームを用いても前後の 2 フレームのみでも正解率に変化がないことがわかる. これ以外の多くのビデオ画像でも実験すると, 全フレームを用いるより 2 フレームのみほうが正解率が高い場合もあった. これは特徴点の追跡が不安定で, 真の位置からフレームごとに多少揺らいている場合である. これは対応するパッチの重ね合わせ (図 4) に誤差があるためと思われる. 一方, 特徴点の追跡が非常に安定している場合に全フレームのほうが正解率が上がる例もあった. しかし, 圧倒的に多くに場合は全フレームと 2 フレームに正解率の差がなかった.

それに対して Morris ら [9] や Perrier ら [11] の方法では, 提案手法に比べてかなり正解率が低下し, しかも全フレームを用いると正解率が悪化することが多い. これも追跡する特徴点の不安定が原因と思われる. 彼らの方法では対応するパッチ全体を重ねて

<sup>8</sup>下記のサイトにプログラムが公開されている.  
<http://www.img.tutkie.tut.ac.jp>

<sup>9</sup>下記のサイトにプログラムが公開されている.  
<http://vision.stanford.edu/~birch/klf/>

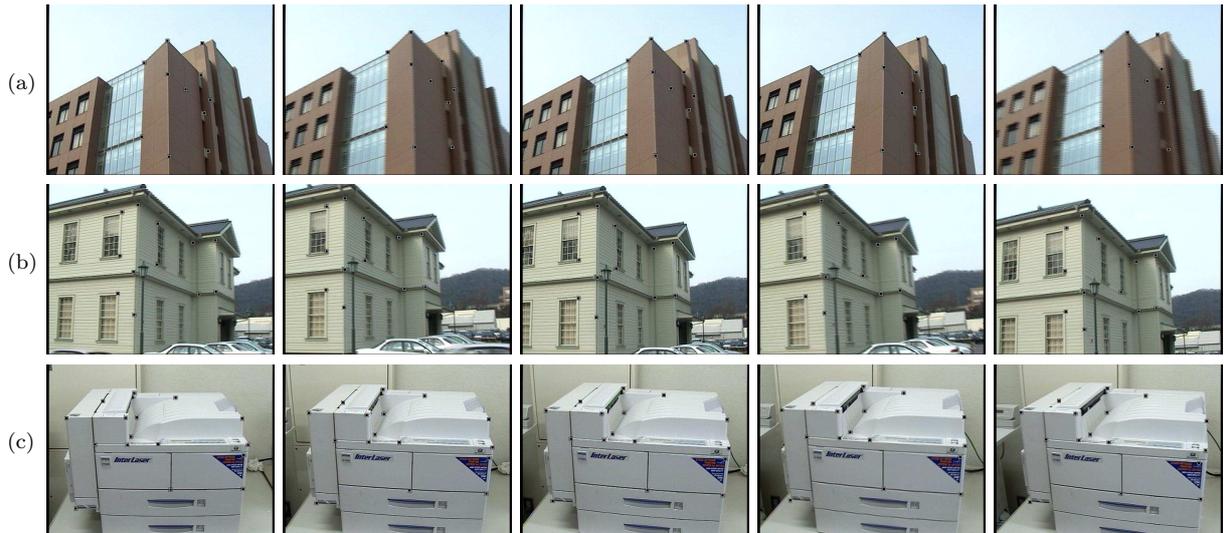


図 11: ビデオ画像上の特徴点の追跡 (5 フレームを抜き出したもの)。

表 2: 図 11 のビデオ画像の最適化の正解率 (%) と実行時間 (秒) . それぞれ (全フレームを使う場合の値) / (前後の 2 フレームのみを使う場合の値) .

(a)	正解率 (%)	実行時間 (秒)
Morris ら	100/81.3	12.8/0.8
Perrier ら	62.5/56.3	11.6/1.0
提案手法	100/100	25.2/1.2

(b)	正解率 (%)	実行時間 (秒)
Morris ら	81.0/81.0	12.0/0.9
Perrier ら	71.4/61.9	14.8/1.1
提案手法	100/100	31.0/1.9

(c)	正解率 (%)	実行時間 (秒)
Morris ら	77.4/77.4	24.6/1.8
Perrier ら	71.0/67.7	32.1/2.1
提案手法	80.6/90.3	60.9/5.5

比較しているので、重ね合わせのずれによって周囲のパッチと混じり合うためと思われる。それに対して提案手法はパッチ全体ではなく、不適合が現れる狭い不適合領域を選択して評価しているため、パッチの周辺の状態には影響されず、高い正解率を維持していると考えられる。

以上より、提案手法が Morris ら [9] や Perrier ら [11] の方法よりも優れているといえる。ただし、実行時間が多少余分にかかる。また、最適化には全フレームを用いる必要はなく、前後の 2 フレームのみで十分であることがわかる。図 11 以外にもさまざまなビデオ画像で実験を行ったが、どの場合もほぼ同様の結論が得られた。また偽に境界線の除去 [10] についても、同じことが観測され、多画像では対応のずれが蓄積するので、前後の 2 フレームのみ用いるほうがより正しい結果が得られた。

## 6. まとめ

本論文では、前報 [10] で提案した多面体表示のための三角網の最適化手法を、2 画像から連続ビデオ画像に適用可能な形に拡張した。また、プレートサイズを可変にして計算時間を効率化する方法を示した。そして、実画像例を用いて Morris ら [9] や Perrier ら [11] の方法と比較し、次の結論を得た。

1. 可変プレートを用いれば、正解率を落とさずに計算が効率化する。
2. 提案手法は Morris ら [9] や Perrier ら [11] の方法より高性能である。
  - Morris ら [9] や Perrier ら [11] が最小化したパッチ間の類似度は必ずしも辺の正しさ反映していない。
  - 照明変化を打ち消す輝度値の正規化 (正規化相関) は誤りの検出には有効ではない。
3. 最適化はビデオ系列の初期フレームと最終フレームのみを用いれば十分である。
  - Morris ら [9] や Perrier ら [11] の方法では中間フレームを用いると正解率が悪化する。
4. 提案手法は曲面物体に対しても、形状表現に適した三角網を生成するのに有効である。

提案手法は Morris ら [9] や Perrier ら [11] の方法と同様に、入力画像間を比較するのみで、シーンの 3 次元情報は一切必要としない。また、判定のための何らのしきい値も必要としない。

残る課題は頂点の追加・削除である。提案手法は与えられた特徴点に対して最適な三角網を生成する方法であるが、多面体物体のどれかの頂点が特徴点として選ばれなければ、三角網を最適化しても 3 次

元形状が正しく表現できない．また，多すぎる特徴点も処理効率を悪化させる．Perrierら [11] は極端に細長い三角形パッチを生成する特徴点を除去し，大きい三角形パッチの内部に特徴点を追加する手法を提案しているが，そのような処理も重要であろう．

謝辞：本研究の一部は文部科学省科学研究費基盤研究C (2) (No. 15500113) によった．有益なコメントを頂いた米国 Nothrop Grumman 社の Daniel Morris 博士，および実験に協力して頂いた熊平製作所の村田正和氏に感謝します．

## 参考文献

- [1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- [2] 飯田亮介, 出口光一郎, 平面を利用したステレオ画像からの任意視点画像の合成, 情報処理学会研究報告, 99-CVIM-114-10, pp. 73–80, January 1999.
- [3] 金谷健一, 「形状CADと図形の数学」, 共立出版, 1998.
- [4] 金谷健一, 菅谷保之, 因子分解法の完全レンビ, 電子情報通信学会技術報告, PRMU2003-118, pp. 19–24, October 2003.
- [5] 金谷健一, 三島等, 未校正カメラによる2画像からの3次元復元とその信頼性評価, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, 42-SIG 6 (2001-6) pp. 1–8.
- [6] 金澤靖, 金谷健一, 大域的な整合性を保証するロバストな画像の対応づけ, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, 44-SIG 17 (2003-12), pp. 70–77.
- [7] 金澤靖, 金谷健一, コンピュータビジョンのための画像の特徴点抽出, 電子情報通信学会誌, 87-12 (2004-12), 1043–1048.
- [8] 栗原祐介, 日吉久礎, 金谷健一, 太田直哉, 画像の一致度を考慮したサーフェスモデルの構築, 情報処理学会研究報告, 2001-CVIM-127-8, pp. 51–58, May 2001.
- [9] D.D. Morris and T. Kanade, Image-consistent surface triangulation, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, Hilton Head, SC, U.S.A., Vol.1, pp.332–338, June 2000.
- [10] 村田正和, 中辻敦忠, 菅谷保之, 金谷健一, 画像からの形状復元に適合した三角網の生成, 情報処理学会研究報告, 2004-CVIM-145-2, pp. 9–16, September 2004.
- [11] J. S. Perrier, G. Agin, and P. Cohen, Image-based view synthesis for enhanced perception in teleoperation, in J. G. Verly (Ed.), *Enhanced and Synthetic Vision 2000*, Proc. SPIE, Vol. 4023, June 2000.
- [12] C. Tomasi and T. Kanade, Shape and motion from image streams under orthography—A factorization method, *Int. J. Comput. Vision*, 9-2 (1992-10), 137–154.
- [13] G. Vogiatzis, P. Torr and R. Cipolla, Bayesian stochastic mesh optimization for 3D reconstruction, *Proc. British Machine Vision Conf.*, Norwich, U.K., September 2003, Vol. 2, pp. 711–718.
- [14] T. Yu, N. Xu and N. Ahuja, Shape and view independent reflectance map from multiple views, *Proc. 8th Euro. Conf. Comput. Vision*, Prague, Czech., May 2004, Vol. 4, pp. 602–615.
- [15] Z. Zhang, R. Deriche, O. Faugeras and Q.-T. Luong, A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Artif. Intell.*, 78 (1995), pp.87–119.

## 付録 A . 残差平方和と正規化相関

2 画像  $I_1(i, j)$ ,  $I_2(i, j)$  の領域  $\mathcal{R}$  内の輝度値の分布を比較する基本的な尺度は次の残差平方和である．

$$RSS = \sum_{(i,j) \in \mathcal{R}} (I_1(i, j) - I_2(i, j))^2 \quad (2)$$

これが小さいほど画像  $I_1(i, j)$ ,  $I_2(i, j)$  は類似している．一方，正規化相関は次のように定義される．

$$NC = \frac{\sum_{(i,j) \in \mathcal{R}} (I_1(i, j) - \bar{I}_1)(I_2(i, j) - \bar{I}_2)}{\sqrt{\sum_{(i,j) \in \mathcal{R}} (I_1(i, j) - \bar{I}_1)^2} \sqrt{\sum_{(i,j) \in \mathcal{R}} (I_2(i, j) - \bar{I}_2)^2}} \quad (3)$$

ただし， $\bar{I}_1$ ,  $\bar{I}_2$  はそれぞれ画像  $I_1(i, j)$ ,  $I_2(i, j)$  の領域  $\mathcal{R}$  中の平均輝度値である．定義より，画像  $\bar{I}_1$ ,  $\bar{I}_2$  に別々に定数を加えても定数倍しても NC は変化しない．また，シュワルツの不等式により  $0 \leq NC \leq 1$  である．これは照明変化の影響を打ち消す目的でよく用いられるが，反面，画像の定数差や定数倍が検出できない．

一方，画像  $\bar{I}_1$ ,  $\bar{I}_2$  に領域  $\mathcal{R}$  内で輝度値の平均が 0 になるように定数を加え，分散が 1 となるように定数倍して次のように正規化したとする．

$$\begin{aligned} \tilde{I}_1(i, j) &= \frac{I_1(i, j) - \bar{I}_1}{\sqrt{\sum_{(i,j) \in \mathcal{R}} (I_1(i, j) - \bar{I}_1)^2}} \\ \tilde{I}_2(i, j) &= \frac{I_2(i, j) - \bar{I}_2}{\sqrt{\sum_{(i,j) \in \mathcal{R}} (I_2(i, j) - \bar{I}_2)^2}} \end{aligned} \quad (4)$$

この正規化画像の残差平方和は次のようになる．

$$\begin{aligned} RSS &= \sum_{(i,j) \in \mathcal{R}} (\tilde{I}_1(i, j) - \tilde{I}_2(i, j))^2 = \sum_{(i,j) \in \mathcal{R}} \tilde{I}_1(i, j)^2 \\ &\quad - 2 \sum_{(i,j) \in \mathcal{R}} \tilde{I}_1(i, j)\tilde{I}_2(i, j) + \sum_{(i,j) \in \mathcal{R}} \tilde{I}_2(i, j)^2 \\ &= 2(1 - NC) \end{aligned} \quad (5)$$

ゆえに正規化相関 NC を最大化することと，正規化した画像の残差平方和を最小にすることは等価である．

## 付録 B . 凸四辺形の判定

注目する辺  $AB$  の両側に点  $P$ ,  $Q$  があるとき，それらの座標を  $A : (a_1, a_2)$ ,  $B : (b_1, b_2)$ ,  $P : (p_1, p_2)$ ,  $Q : (q_1, q_2)$  とすると，四辺形  $APBQ$  が凸である条件は次のように書ける [3] ．

$$\begin{vmatrix} p_1 - a_1 & q_1 - a_1 \\ p_2 - a_2 & q_2 - a_2 \end{vmatrix} \cdot \begin{vmatrix} p_1 - b_1 & q_1 - b_1 \\ p_2 - b_2 & q_2 - b_2 \end{vmatrix} < 0 \quad (6)$$

## 付録 C . 四辺形の向き判定

$\triangle ABC$  の符号を  $A$ ,  $B$ ,  $C$  が反時計回りに回転するとき正，時計回りのとき負，線分に退化した場合に 0 と定義する．これは  $A : (a_1, a_2)$ ,  $B : (b_1, b_2)$ ,  $C : (c_1, c_2)$  に対して次のように計算される [3] ．

$$\text{sgn} \left( \begin{vmatrix} b_1 - a_1 & c_1 - a_1 \\ b_2 - a_2 & c_2 - a_2 \end{vmatrix} \right) \quad (7)$$

ただし  $\text{sgn}()$  は符号関数であり，引数が正，負，0 のときそれぞれ 1, -1, 0 を返す．