

弾性メッシュモデルを用いた 多視点画像からの高精度3次元形状復元

延原章平[†] 和田俊和[†] 松山隆司[†]

多視点画像から物体の3次元形状を高精度に復元する手法を提案する。提案手法では、視体積交差法によってシルエットという geometric な情報を用いて対象の概形を安定に求めた後、弾性メッシュの変形によって photometric な情報を用いて形状の高精度化を行う。つまり前半では安定な復元が、後半では復元の高精度化がなされ、安定性と高精度化の両立を目指す。本論文では弾性メッシュモデルを、1) “photo-consistency” 制約、2) シルエット保持制約、3) 単純閉曲面制約、という3つの制約条件を満たしながら変形を行うように設計するとともに、実験によって物体形状の高精度化が可能であることを示す。

3D Shape From Multi-viewpoint Images Using Deformable Mesh Model

SHOHEI NOBUHARA,[†] TOSHIKAZU WADA[†] and TAKASHI MATSUYAMA[†]

This paper presents a two-stage method for 3D shape reconstruction from multi-viewpoint images. In the first stage, a rough 3D shape of an object is estimated by using volumetric intersection from geometric information of the object, i.e., silhouettes. In the following stage, the estimated shape is refined by using “deformable mesh model”. This model changes its shape so as to satisfy the following constraints: 1) “photo-consistency” constraint, 2) silhouette constraint, 3) smoothness constraint. The volumetric intersection estimates rough but stable shape using geometric information, and the deformable model refines the shape using photometric information. Hence, the method has two advantages of stability and accuracy. Some experimental results are also shown for demonstrating these advantages.

1. はじめに

本論文では、複数カメラを用いて対象の詳細な3次元形状を復元する手法について述べる。形状復元する対象は人物を、撮影環境として対象を取り囲むように配されたカメラ群を想定している。

これまで多視点2次元映像から立体形状を復元するアプローチとして、視体積交差法¹⁾²⁾³⁾⁴⁾とステレオ法が広く用いられてきた。視体積交差法はシルエット制約⁵⁾に基づく手法であり、シルエット制約とは2次元画像上のシルエットを3次元空間に逆投影して得られる錐体“視体積”の内側に対象が含まれるということである。この制約に基づいて、異なる視点の2次元シルエットから複数の視体積を求め、その交差領域をもって対象の形状とする(図1)。一方ステレオ法は、異なる視点から撮影された2次元画像上で物体表面上

の点の対応づけを行い、三角測量の要領で対象までの距離を求める手法である。

これらの手法を比較すると、視体積交差法は対応づけを行わないために誤対応が生じず、安定な形状復元が行える。さらに出力は距離画像のような2.5次元データではなく、特に張り合わせを行わなくても物体の全周囲形状が復元できるというメリットがある。しかし原理的に形状の visual hull (図1)を求めているため、対象の詳細な形状は復元することはできない。このことは、視体積交差法が対象のシルエット情報(輪郭情報)という画像の geometric な情報しか利用していないことに起因している。

一方ステレオ法の場合、画像の photometric な情報に基づいて対応付けを行い、対応付けの後は幾何計算によって奥行きを算出する。この場合正しい対応づけができれば物体表面上の窪みのある部分でもその3次元位置を計算することができるというメリットがある。しかし本論文のように対象を取り囲むようにカメラを配置する場合、限られたカメラ台数では視点間の

[†] 京都大学 大学院情報学専攻 知能情報学専攻
Department of Intelligent Science and Technology,
Graduate School of Informatics, Kyoto University

距離 (基線長) が長く、視線方向も大きく異なったものと成らざるを得ず、ステレオ法による形状復元は困難となる⁶⁾。

そこで本論文ではこれら 2 つの情報を共に用いる手法、すなわちまず画像の geometric な情報に基づいて対象の概形を安定に復元し、さらに画像の photometric な情報に基づいて詳細な対象の三次元形状を得る方法を提案する。

このようなアプローチとしては、これまでも視体積交差法と space carving⁷⁾ を組み合わせた手法⁸⁾などが提案されている。しかし対象表面のうちテクスチャがない部分では、形状変化に対する “photo-consistency”⁷⁾ の変化がないため、真の物体形状を求めることは困難である。このような状況に対処する一つのアプローチとして、photo-consistency によって形状を決定できる部分はこれを採用し、そうでない部分はその周囲の形状を元に補間するというアプローチが考えられる。このような復元を行うため、本論文では対象形状を voxel 集合としてではなく、弾性メッシュモデルと呼ぶ表面モデルによって表し、この弾性メッシュモデルの変形によって形状復元を行うこととする⁹⁾¹⁰⁾¹¹⁾。このモデルによって形状を表現することにより、頂点間に連結関係が保証され、photo-consistency による形状の推定が信頼できない部分については、周囲の信頼できる部分との間で力学的制約に基づく補間を行うことができる。

以上、本論文では対象を周囲から撮影するように配されたカメラ群を用いて、まず視体積交差法で対象の概形を求め、次にこれを “弾性メッシュモデル” に変換し、このメッシュを photo-consistency 制約、力学的制約、シルエット制約が成立するように変形させることによって詳細な形状を求める手法を提案する。以降、visual hull の取得とその特徴、弾性メッシュモデルの設計、評価実験の順に述べる。

2. Visual Hull

2.1 計算方法

visual hull の定義は先に述べた通りであるが、具体的には、対象が存在すると思われる空間を voxel に分割し、各 voxel について、“その voxel を撮像面に投影したとき、対象領域 (シルエット) 内に投影されるか否か” を各撮影視点で調べ、全ての視点でシルエット内であるものを残せばよい。こうして、視体積 (visual cone) の交差領域 (visual hull) が得られる。

2.2 特徴

先に述べたように、対象の真の形状は visual hull

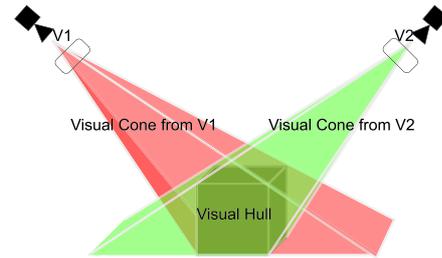


図 1 visual hull
Fig. 1 visual hull

カメラ v_1, v_2 それぞれから視体積 (visual cone) が生成され、その交差領域が visual hull となる。

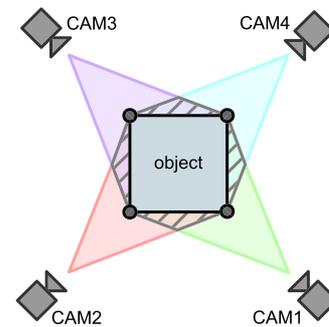


図 2 visual hull と対象の真の形状の接点
Fig. 2 visual hull & object

対象の真の形状と visual hull は必ず接している。そして接点は各カメラにおいてシルエットの輪郭として捉えられる (図中黒丸の点)。図 6 参照。

内に含まれる。このことは、visual hull を求めることによって、対象の存在する領域が限定できることを意味する。さらに visual hull の定義から明らかのように、対象の真の形状は各視点でのシルエットを保つように visual hull に内接している (図 2)。したがって、最終的に復元される形状においても、各視点から見たシルエットは保存されるべきである。

以上のことから、visual hull からさらに変形を進める際には、“シルエット保持” という制約を課すべきであるといえる。

3. 弾性メッシュモデル

1 節で述べたように、以下の方針で形状復元の高精度化を行う。

視体積交差法によって得られた対象の voxel モデルから対象の “表面モデル” を求め、それを photo-consistency が成立するように変形する。但し、変形によって表面同士が交差せず、しかもモデルを各カメラで撮影した場

合に、各画像上の対象シルエットと一致するように変形を行う。

このような変形を行うための対象の表面モデルを弾性メッシュモデルと呼ぶことにする。以降、この弾性メッシュモデルの詳細な設計法について述べる。

3.1 初期形状の獲得

本手法では、対象の表面形状を三角形メッシュで表現する。また初期形状である visual hull は voxel 集合として得られるため、本論文ではこれを離散マーチングキューブ法¹⁴⁾を用いてメッシュデータに変換する。この手法を用いる理由は、

- m 連結離散閉曲面 ($m = 18, 26$) が得られることが保証されている。
- 2 値の voxel 集合からの変換が一意である。

という特徴があるためである。

3.2 変形

弾性メッシュモデルの変形は、メッシュを構成する各頂点の移動によって実現する。このような変形を行う SNAKES¹²⁾ では全頂点で定義されるエネルギー関数が最適化されるように全頂点の移動を行うが、本研究の場合では、

- メッシュを構成する頂点・面数が膨大である。
- 後述する“シルエット保持力”を、エネルギーとして表現することが困難である。

などの理由から、各頂点に働く力を定義し、それに沿って各頂点に変形するという方式を採用する。

3.2.1 頂点に働く力

頂点 v に働く力 $F(v)$ を定義する。

visibility

頂点を v で表したとき、その3次元座標を q_v で表す。また、頂点 v を観測することができるカメラの集合を C_v で表す。例えば図3の場合、 $C_v = \{CAM_2, CAM_3\}$ となる。明らかに、各頂点毎にそれを観測することのできるカメラ集合 C_v は異なる。以降 C_v を求めることを、“ v の visibility を調べる”と表現する。

外力 $F_e(v)$

まず、弾性メッシュモデルを対象形状に近づけるための力として外力項 $F_e(v)$ を導入する。

$$F_e(v) \equiv -\nabla E_e(v). \quad (1)$$

ここで $E_e(v)$ は、その頂点における撮影画像間の相関値(図3)である。つまり、 $F_e(v)$ は、より撮影画像間の相関が高くなる方向へと作用する力である。このために、 $E_e(v)$ は、 C_v に含まれるカメラ c から、頂点 v に投影されるテクスチャ情報 $p_{v,c}$ と、その平均 \bar{p}_v

具体的には撮影された画素値など

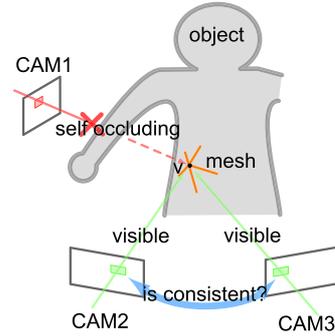


図3 外力項と visibility

Fig. 3 photometric consistency and visibility

頂点 v にとって、CAM₁ はセルフオクルージョンによって写っていないと判断され、CAM₂ と CAM₃ は写っていると判断される。このことから、photo-consistency の判定時には CAM₁ を除き、CAM₂・CAM₃ を用いる。

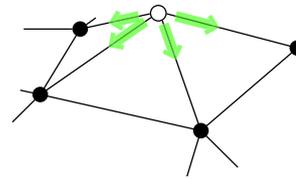


図4 内力

Fig. 4 internal force

頂点(白丸)には、隣接する頂点(黒丸)との間で張力(矢印)が作用する。

を用いて、

$$E_e(v) \equiv \frac{1}{N(C_v)} \sum_{c \in C_v} \|p_{v,c} - \bar{p}_v\|^2 \quad (2)$$

と表すことができる(ただし $N(C_v)$ は C_v に含まれる要素の数)。

内力 $F_i(v)$

上記 $F_e(v)$ のみで変形を行うと、メッシュの交差が起きる可能性がある。交差が起きると、先に述べた visibility の判定などで不具合が生じる。これを防ぐために、次式で定義される内力項 $F_i(v)$ を導入する。

$$F_i(v) \equiv \frac{\sum_j^n q_{v_j} - q_v}{n}, \quad (3)$$

但し v_j は v に隣接する頂点、 n を v に隣接する頂点の数、 q_v は v の位置ベクトルとする。これは頂点間の張力として働き(図4)、メッシュ形状の滑らかさを保つ働きを持つ。

シルエット保持力 $F_s(v)$

2.2 節で述べたように、弾性メッシュモデルは常に visual hull に外接し、撮影視点から見たときの輪郭を保つべきである。このような条件を成立させるため、シルエット保持力 $F_s(v)$ を導入する。 $F_s(v)$ は、カメラ c から見たときのシルエットを保持するために v に

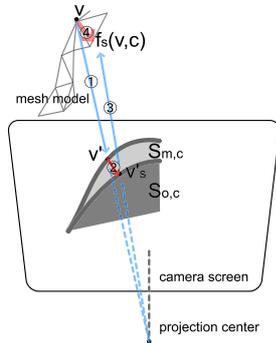


図 5 シルエット保持力

Fig. 5 silhouette preserving force

v が投影 (①) される先 v' を、撮影シルエットに近づける (②) ため、逆に移動先を推定し (③)、その方向への力 (④) を加える (図は本来の撮影シルエット $S_{o,c}$ よりも復元シルエット $S_{m,c}$ が外にある場合)。

加える力 $f_s(v, c)$ を、各カメラについて合成したものとして定義する。以降、撮影画像から得られる対象本来のシルエットを“撮影シルエット”、得られているメッシュを投影して得られるシルエットを“復元シルエット”と呼ぶ。 $f_s(v, c)$ は、カメラ c での復元シルエットを撮影シルエットに近づける力である。

本論文では $f_s(v, c)$ を、

- 1) v が撮影シルエットの領域外に投影されるとき、あるいは
- 2) v が撮影シルエットの領域内に投影され、且つ、復元シルエットの輪郭を成すとき

に

v の投影位置 (図 5 v') を、撮影シルエットの輪郭に一致させる最短移動ベクトル (同②)、④) を $f_s(v, c)$ とする。

と定義する。そして先に述べたように、 $F_s(v)$ を $f_s(v, c)$ の和として定義する。

$$F_s(v) \equiv \sum_{c \in C_v} f_s(v, c). \quad (4)$$

このシルエット保持力は、見方を変えると見かけ上の輪郭に対応する対象表面上の領域、すなわち“contour generator”⁸⁾ を推定していることに相当する (図 6)。

頂点に作用する力

以上から、係数 α, β, γ を用いて、頂点に作用する

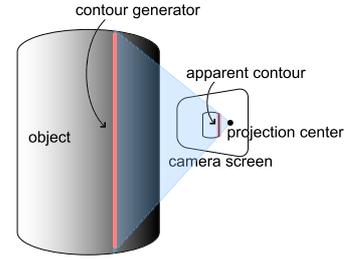


図 6 Contour Generator

Fig. 6 Contour Generator

撮影像上で観測される見かけ上の対象の輪郭 (apparent contour, シルエットの輪郭) には、対応する対象表面上の“稜線” (contour generator) が存在する。

力 $F(v)$ を

$$F(v) \equiv \alpha F_i(v) + \beta F_e(v) + \gamma F_s(v) \quad (5)$$

と定義する。外力 $F_e(v)$ とシルエット保持力 $F_s(v)$ が対象形状を復元するための力であるのに対して、内力 $F_i(v)$ は平滑化および補間を行う力としての働きを持つ。これらについては後の節で考察を行う。

3.2.2 パラメータの動的制御

以上、式 (5) に示したように、頂点に働く力 $F(v)$ の 3 つの項の設計について述べた。しかし、これだけでは互いの力が干渉し、高精度な形状復元が達成できない。そこで以下の 2 つの場合について考察し、式 (5) を修正する。

1) F_e と F_i

常に一定の内力が働き続けてしまうと、外力による頂点の移動が阻害される。これを防ぐために、

外力が強いとき、つまり対象の真の形状に近づこうとしているときには、それに任せて変形を進めるために内力を弱める

ということを実現するための外力・内力の重み $W_e(v)$ 、 $W_i(v)$ を以下のように定義する。

まず外力によって形状を決定するためには、1 節で述べたように対象表面にテクスチャがない (明度変化がない) 場合を考えると、式 (1) のような photo-consistency のみによって形状を決定するべきではない。単に photo-consistency が成り立つだけでなく、その頂点にテクスチャが存在し、明度変化があるならば、形状を決定してもよいはずである。したがって、

- テクスチャ情報 $p_{v,c}$ の変化量 $\|\nabla p_{v,c}\|$ が大きく、
- 且つその変化量の大きさが各カメラ間で互いに近いとき

に外力が強く働くべきである。このような重み付けを行うため、係数 $W_e(v)$ を、 $\|\nabla p_{v,c}\| \in C_v$ での標準偏差 $\sigma_{\nabla p_v}$ 、平均 $\overline{\nabla p_v}$ を用いて

$$W_e(v) \equiv \overline{\nabla p_v} - \sigma_{\nabla p_v} \quad (6)$$

なお、本論文では保持させるべきシルエットそのものの精度については議論していない。しかし明らかにシルエットの精度は論文で提案する手法全体に強い影響を及ぼす。この問題点と将来の解決案については、後の考察の項で取り上げる。

とする．

次に内力の重み $W_i(v)$ は，外力の強さに応じた大きさとするため，常に $K > \|F_e(v)\|$ となるような定数 K を用いて

$$W_i(v) \equiv K - \|F_e'(v)\| \quad (7)$$

とする．

2) F_s と F_i

上とは逆に，シルエット保持力 $F_s(v)$ についてはパラメータの調整を行わない．これは以下の理由による．

- シルエット保持力は外力とは異なり，常に真の contour generator を選択しているとはいえず，“contour generator 候補”を選択しているに過ぎない．
- 表面が滑らかな対象の場合，contour generator は，局所的に見れば対象表面上において連続した領域として存在するはずである．先に述べたシルエット保持力 F_s の定義ではこのような“連続性”に対する保証は入っていないが， F_s が内力 F_i と同時に働くならば，内力が推定中の contour generator 候補に対するローカルサポートとして働き，“局所的連続性”が確保できることが期待される (図 7) ．

以上，式 (5)，(6)，(7) から，頂点 v に働く力を $F(v) \equiv \alpha W_e(v)F_e(v) + \beta W_i(v)F_i(v) + \gamma F_s(v)$ (8) と定義する．

この式 (8) に含まれる $F_s(v)$ の定義 (式 (4)) から，シルエット保持力は頂点に選択的に働くことがわかる．モデルの変形をエネルギー関数最適化としてでなく，力による変形としたことにより，このように柔軟な変形を実現することができる．

3.2.3 変形計算

変形計算の目的は，先に定義した力の平衡する位置を求めることである．本手法ではこれを反復計算で求める．これは，

- モデルを構成する頂点数が膨大である．
- 各頂点に作用する力は，互いに影響を及ぼしあう．(ある頂点の移動による visibility の変化は，他の頂点の外力項を変化させる) ため，部分最適化問題へと分割できない．

という理由から，全頂点に働く力の平衡位置を一度に求めることが，計算上困難であることによる．そこで反復計算の各反復内において，頂点の移動を微小移動として互いの visibility の変化を無視することで計算を簡略化する．つまり各頂点に作用する力をそれぞれ独立に計算できることとする．

このような反復計算は，反復が進むことを時間が経

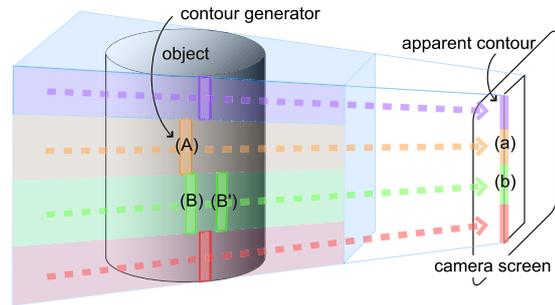


図 7 contour generator の候補

Fig. 7 candidates for the contour generator

apparent contour 上の領域 (b) に対応する contour generator は，(b) に投影される位置にある (B)，(B') などが候補となる．ここで (a) に対応する contour generator が (A) であるならば，より (A) に近い (B) が (A)・(B) 間の内力によるローカルサポートによって，contour generator となる．

過することを見なせば，物理現象のシミュレーションのように各頂点に働く力を微小時間ごとに求めながら，変形過程を求めていることと同じである．すなわち，

- 内力の設計は，力学解析における素材のモデル化に相当し，
- 外力・シルエット保持力の設計は，力学解析における外部からの圧力に相当する

と考えることができる．

初期形状

上のような反復計算で得られた解は，必ずしも大局的最適解であるとは限らず，初期形状によっては局所解となることは十分考えられる．しかしながら，本研究では変形の初期形状として visual hull を用いること，そして visual hull がかなり対象に近い 3 次元形状であることから，実用上問題がないものとする．

3.3 力の物理的意味

3 で設計した弾性メッシュモデルは，物理的には

- contour generator (図 6) を“梁”とする
 - シルエット保持力によって，contour generator の推定を行う．
- “梁”の上に極小曲面を張る膜を載せる (図 8(b))
 - 内力によって，表面を内挿する．
- 膜を支える点を追加してゆく (図 8(c))
 - 外力によって，対象の真の形状表面であるとされた部分は固定され，その間を再び内力によって内挿する (図 8(d)) ．

のような“膜 + 梁”モデルであると解釈できる．ただし，実際には反復計算の過程で“梁”，“面”，“支える点”は同時に復元が進められる．

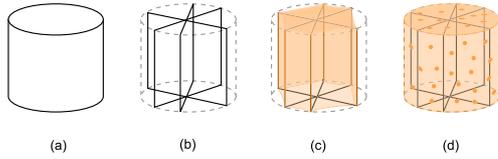


図 8 弾性メッシュモデルの意味
Fig. 8 physical model

左から (a) 真の形状, (b) “梁” としての contour generator, (c) “梁” のうえに内力 $F_i(v)$ による膜を載せたとき, (d) さらに外力 $F_e(v)$ で膜を支えるとき. 図 (b) の contour generator はシルエット保持力 $F_s(v)$ で復元される. 内力 $F_i(v)$ のみときは, “梁” の間を内挿する膜ができる (図 (c)). この膜を支える点 (図 (d) 内の水玉模様の点) を外力 $F_e(v)$ で保持することで, 対象形状に近づける.

4. 評価実験

設計した弾性メッシュモデルの特性を評価するため, シミュレーションと実際の画像を用いて実験を行う. この実験での実験環境は以下の通り.

- カメラ 9 台
- モデルを構成する面数は約 8000
- 1 回の反復計算に要する時間は約 1.5 秒 (Pentium 4 1.7GHz)

また, 弾性メッシュモデルのパラメータ (式 (5) の α, β, γ) は人手によって与えた固定値を用いる.

4.1 シミュレーションによるモデルの評価

シミュレーションによって, 真の形状が既知である物体に対して, 1) 設計した力の各項の効果と, 2) 本論文で設計した弾性メッシュモデルがどの程度適用可能であるかを示す.

まず, シミュレーションで用いる真の形状は超 2 次関数面 (super quadric ellipsoid) とする. これはパラメータ u, v を用いて,

$$\begin{aligned} x(u, v) &= a_1 \cos^n u \cos^e v \\ y(u, v) &= a_2 \cos^n u \sin^e v \\ z(u, v) &= a_3 \sin^n u \end{aligned} \quad (9)$$

$$-\frac{\pi}{2} \leq u \leq \frac{\pi}{2}$$

$$-\pi \leq v \leq \pi$$

で与えられる.

実験 1

これを図 9(b) のように配置し, 周囲の視点から仮想的に撮影し, 得られた撮影画像を用いて形状復元を行う. 対象表面全体にわたってランダムに生成された十分細かなテクスチャが存在する. また評価基準として, 真の形状との誤差を用いる. これは, 復元された形状 (メッシュモデル) を構成する各頂点から楕円面までの距離の和とする.

カメラと対象を図 9(b) のように配置し, 実験を行っ

た結果を図 9 に示す. 図 9 では, 3 つの項の組み合わせに応じて復元結果が異なっており, 真の形状 (図 9(c)) に対し, 全ての項を用いた場合 (図 9(i)) が最も復元精度が高いことが図 9(a) のグラフからわかる. この結果は表 1 のように解釈することができる. 特に図 9(i) から, 視体積交差法で得られた visual hull (9(d)) では復元できなかった部分 (図 9(c) 中の凹の部分) が, 提案手法では復元可能であることがわかる. なお 9(h) と 9(i) を比較すると, (h) に対してシルエット保持力 F_s を有効にした (i) の方が対象形状の輪郭がより正確に復元されていることが確認できるが, グラフ (図 9(a)) では他の場合に比べて差が小さい. これは, (h) と (i) の差となる部分 (この場合は対象形状の “端” の部分) が対象全体に占める割合が小さいためである.

実験 2

先に述べたように, テクスチャがない領域は周囲から補間を行えることを意図して弾性メッシュモデルを設計した. そこで, 前の実験では対象物体表面にはランダムに生成された十分細かなテクスチャを持たせていたが, 逆にテクスチャがない領域を持たせることで, テクスチャが復元に与える影響について考える. 前と同様に図 9(b) のように配置した状態で, 対象物体の表面テクスチャを変化させたときの結果を図 10 に示す. これらはすべて同じパラメータ (式 (5) の α, β, γ) のモデルで復元を行っている.

同図 (c) からわかるように, テクスチャがない領域が増えると, contour generator による “梁” の推定に失敗し, 大きく異なる形状に収束する. 一方同図 (b) のように, ある程度まではテクスチャの欠損が復元に影響を与えない.

このことは, 内力が contour generator 候補に対するローカルサポートとして働く (図 7) ことを考えると, 部分的にテクスチャが存在しなくとも, 内力によって “梁” の推定が可能であるためと考えることができ, “梁” の推定に十分なテクスチャの細かさは内力の大きさに依存するといえる.

実験 3

最後に式 (5) の (n, e) を変化させて作成した, 様々な形状に対する復元結果を図 11 に示す. カメラ配置および十分細かなテクスチャが存在することは実験 1 と同様である. これらはすべて同じパラメータ (式 (5) の α, β, γ) のモデルで復元を行っている.

この図の $(n, e) = (3.0, 1.0)$ の場合と $(n, e) = (5.0, 1.0)$ の場合を比較すると, 凹領域の復元に限界が生じ, $(n, e) = (5.0, 1.0)$ の場合には復元が不完全であ

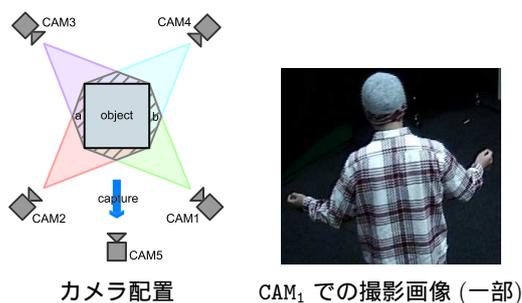


図 12 評価実験時のカメラ配置

Fig. 12 camera location

CAM₁ ~ CAM₄ を形状復元に使用し, CAM₅ は評価用として用いる. CAM₅ から見たとき, visual hull の段階では図内の領域 a および b が, 対象本来のシルエットの外側にはみ出して見えるはずである. なおこれはカメラ配置を上から見た図であり, 実際には各カメラは対象をやや見下ろす状態になっている.

ることがわかる. これは“梁”の間に張られた“膜”に働く 2 つの力, 内力と外力が, (5) の α, β によって決まるある曲率で平衡状態となり, 対象がその曲率以上の領域を持っていた場合であると考えることができる.

4.2 実画像による評価

実際の人物撮影像で実験を行い, 本手法で復元精度が改善できることを示す. ただしこの場合, 対象の真の形状は不明であるため, 形状の評価尺度として“復元に用いなかった視点から見たシルエットの一致度”を用いる. つまり図 12 の CAM₅ のように形状復元に用いなかった視点から見たとき, 復元されたシルエットが対象本来のシルエットに近づくかどうかを調べる.

カメラ配置を図 12 のように復元用 4 台・検証用 1 台のカメラを用いて実験を行った結果を図 13, 14 に示す. この結果から, 復元に用いなかった視点 (図 13) においても, シルエットが復元されていることがわかり, 本手法が実画像に対しても適用可能であることがわかる. また図 15 に, このときの最終的なレンダリング結果を示す. 股の下など, どの実視点からも見えにくい部分に注目すると, 本来の形状と異なっていることがわかる. これは, このような部分では内力による補間の影響が強く, “膜”が張られるためである.

5. 結 論

本論文では多視点映像からの 3 次元形状復元という問題に対し, 対象の輪郭形状という geometric な情報から安定に対象形状を絞り込んだ上で, photometric な情報を用いて高精度に 3 次元形状復元を行う手法を提案し, 実験・評価を行った.

提案した手法は, space carving にシルエット制約を導入したものと比較することができる. 提案手法が

モデルフィッティングによる復元であるのに対して, space carving は voxel 集合の削り込みによって形状の復元を進めるものであるが, いずれにしてもテクスチャがない部分では, 復元を行うことは本質的に困難である. この問題への自然なアプローチの 1 つとして“テクスチャがない部分では周囲の信頼できる復元結果を基に補間を行う”という手法が考えられるが, 本手法ではこれが“内力”として自然な形で取り入れられている. これは本手法が space carving とは異なり, 表面モデルで対象を表現していることの利点であるといえる.

提案したモデルが表現可能な形状は, 式 (5) のパラメータ α, β, γ で決まる. 実験では, 固定パラメータでいくつかの典型的幾何形状の復元を行うことで, 復元可能な曲率に限界があること, そしてこれが“膜 + 梁”のモデルで解釈できることを示した.

今後, 本論文のモデルフィッティングによる手法は, 異なる時刻の多視点画像へのフィッティングへと拡張が可能である. このことは, フィッティング先の時刻における対象形状が求まると同時に, 異なる時刻間での対象表面上の点の対応関係が得られることを意味している. この対応関係は, 対象の動作解析や記録した 3 次元形状のデータ圧縮のために必須の情報であることから, モデルフィッティングによる 3 次元形状復元の将来性が期待できる.

また提案手法で得られる 3 次元形状の精度は, 対象シルエットの切り出し精度に強く依存する. このためシルエットの輪郭推定を行う“シルエットモデル”を導入し, 提案した弾性メッシュモデルと各視点のシルエットモデルがそれぞれ同時に対象形状と対象輪郭を復元するなどしてこの制約を緩和することが今後の課題としてあげられる.

謝 辞

本研究は科学研究費補助金 基盤研究 A 13308017 および特定領域研究 13224051 の補助を受けて行った.

参 考 文 献

- 1) H. Baker. Three-dimensional modelling. In *Fifth International Joint Conference on Artificial Intelligence*, pages 649–655, 1977.
- 2) B.G. Baumgart. Geometric modeling for computer vision. Technical Report AIM-249, Artificial Intelligence Laboratory, Stanford University, October 1974.
- 3) W. N. Martin and J. K. Aggarwal. Volumetric description of objects from multiple views.

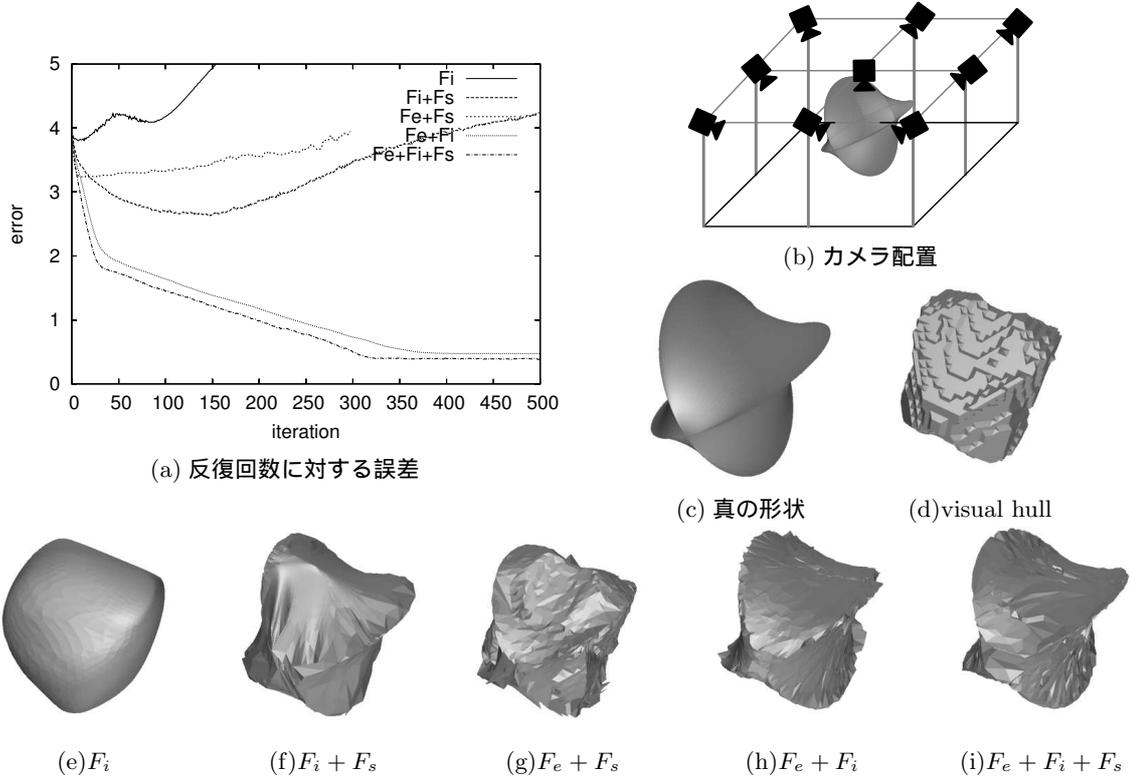


図 9 シミュレーション実験 1
Fig.9 results : simulation 1

(e) は縮小してつぶれるため、(g) はメッシュの交差によって変形計算が破綻するため、途中で処理を打ち切っている。また (f) では図 8 で述べた“梁”とその間に膜を張った状態になっている。しかし“梁”の推定に失敗した箇所付近では、大きく形状が崩れていることがわかる。
なお図 (b) のように、カメラは対象を見下ろすように配置されている。このため、どのカメラからも撮影されない部分 (c) - (i) の左下側) は変形・評価の対象外としている。
印刷の都合上、立体感を出すために表面テクスチャを消している。

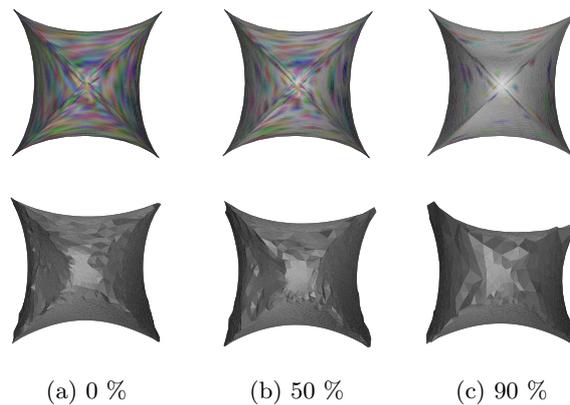


図 10 シミュレーション実験 2
Fig.10 results : simulation 2

テクスチャと復元結果の関係。上段が真の形状，下段が復元結果。
左から順番に 0%、50%、90% の面積がテクスチャのない領域 (灰色単色領域) である。これらはいずれも復元に用いていない仮想視点からの像である。

| F_e | F_i | F_s | |
|-------|-------|-------|---|
| - | - | | visual hull (初期形状) の段階でシルエットは保存されているため, 変形は起こらない. |
| - | | - | 縮小する (図 9(e)). |
| - | | | 復元に用いた視点ではシルエットが保存されるが, 用いなかった視点ではシルエットが崩れる (図 9(f)). |
| | - | - | 各頂点が初期位置近傍の局所的に photometric な評価が高い位置へと移動し, 変形が止まる. |
| | - | | 多くの頂点で (, - , -) の組み合わせと同じであり, F_s が作用した箇所では変形が進む. しかし次第に局所的交差の発生によって visibility の判定に支障をきたしはじめ, 変形計算が破綻する (図 9(g)). |
| | | - | 変形によって多くの頂点で真の形状に近づく. しかし全体として輪郭が崩れる (図 9(h)). |
| | | | (, - , -) の組み合わせに加え, 撮影視点でのシルエットを保存する (図 9(i)). |

表 1 使用する力と復元結果の関係

Table 1 reconstruction results against combinations of forces

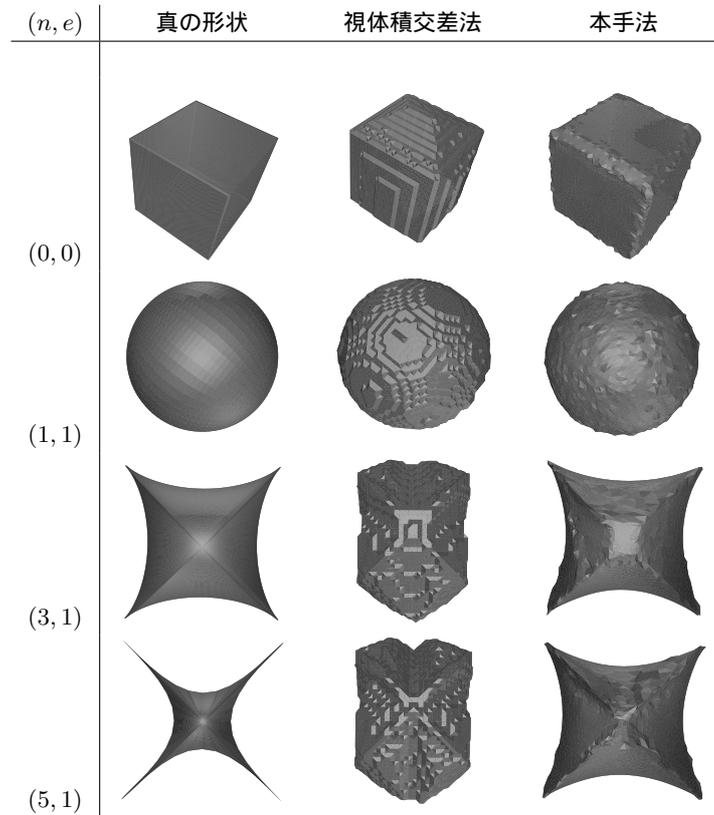


図 11 シミュレーション実験 3

Fig. 11 results : simulation 3

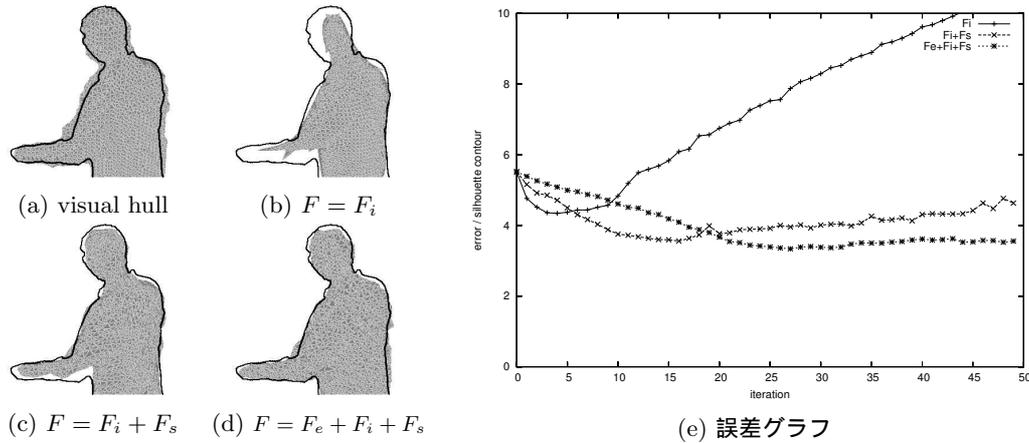
式 (9) の (n, e) を変化させたときの復元結果. これらはすべて同じパラメータ (式 (5) の α, β, γ) のモデルで復元を行った結果である.
印刷の都合上, 立体感を出すために表面テクスチャを消している.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 5(2):150–158, 1987.

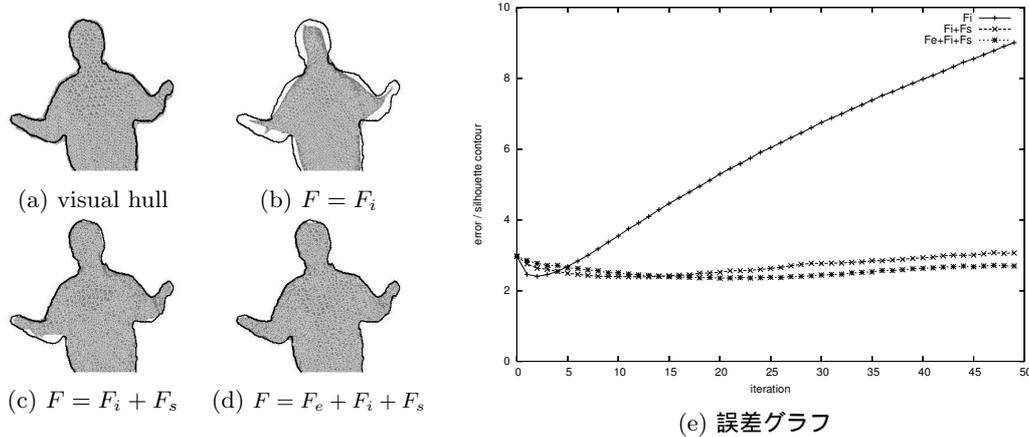
- 4) P. Srivasan, P. Liang, and S. Hackwood. Computational geometric methods in volumetric intersections for 3d reconstruction. *Pattern Recognition*, 23(8):843–857, 1990.
- 5) A. Laurentini, “How far 3d shapes can be understood from 2d silhouettes”, *IEEE Transactions on Pattern Analysis and Machine Intelli-*

gence, 17(2), pp. 188–195, 1995.

- 6) Philip Pritchett and Andrew Zisserman, “Wide Baseline Stereo Matching”, *International Conference on Computer Vision*, pp. 754–760, 1998.
- 7) K. N. Kutulakos and S. M. Seitz, “A theory of shape by space carving”, *IEEE International Conference on Computer Vision*, pp. 307–314, 1999.

図 13 実画像による実験 (CAM₅)Fig. 13 results : real image (CAM₅)

(a) ~ (d) 中の黒線が撮影シルエットの輪郭 . (e) の横軸は反復回数 , 縦軸は撮影シルエットと復元シルエットの誤差 .

図 14 実画像による実験 (CAM₁)Fig. 14 results : real image (CAM₁)

各図の意味は図 13 と同じ .

- 8) Geoffrey Cross and Andrew Zisserman, "Surface reconstruction from multiple views using apparent contours and surface texture", *Conference of Computer Vision and Computer Graphics*, NATO Science Series 3. High Technology, Vol. 84, Kluwer, Dordrecht, 2000, pp. 25-48.
- 9) W. Brent Seales and Oliver D. Faugeras, "Building Three-Dimensional Object Models from Image Sequences", *Computer Vision and Image Understanding* 61(3), pp. 308-324.
- 10) P. Fua and Y. G. Leclerc, "Object-Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading", *International Journal of Computer Vision*, 16, 1995, pp. 35-56
- 11) M. Hebert, K. Ikeuchi, and H. Delingette, "A Spherical Representation for Recognition of Free-Form Surfaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 7, July, 1995, pp. 681-690.
- 12) M. Kass and A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models", *Proc. of IEEE Conference on Computer Vision*, 8-11, pp. 259-268, London, England, 1987.
- 13) Tony Heap and David Hogg, "Towards 3D Hand Tracking using a Deformable Model", *Second International Conference on Automatic Face and Gesture Recognition (FG96)*, Vermont, New England, October 1996.
- 14) 剣持 雪子, 小谷 一孔, 井宮 淳, "点の連結性を考慮したマーキング・キューブ法", 信学技報 *TECHNICAL REPORT OF IEICE*, IE98-218 (1999-01), pp. 197-204.

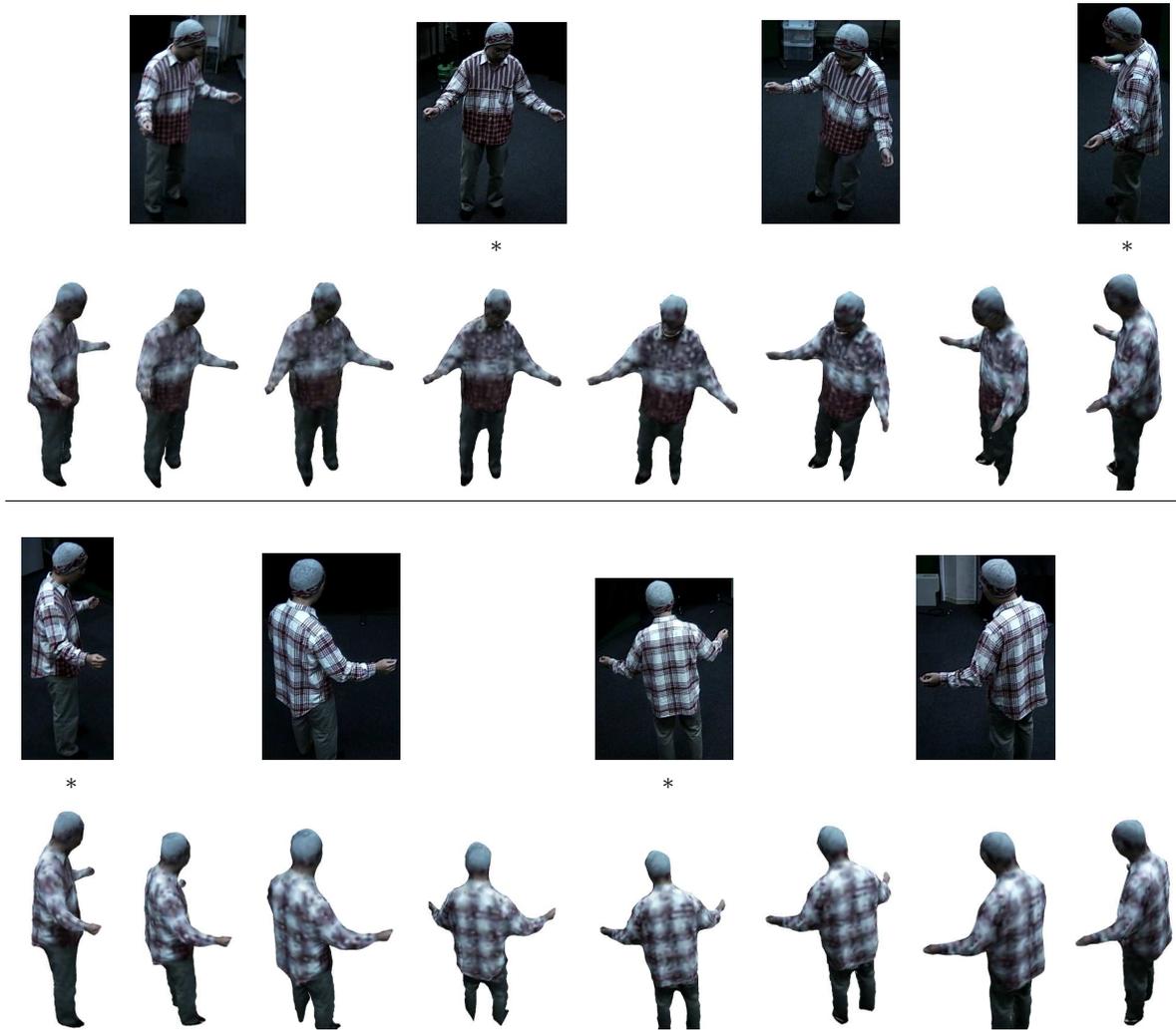


図 15 実画像と生成画像

Fig. 15 captured and rendered images

8 方向の実画像 (奇数行) と 16 方向の生成画像 (偶数行) . 生成画像のうち, 上段偶数列・下段奇数列の 8 枚は実画像と全く同じ視点から, 上段奇数列・下段偶数列の 8 枚は仮想視点からの像) . また, *がついているものが復元に用いた視点 .

15) 高井勇志, 松山隆司, “3D ビデオ映像の生成・編集・表示”, 映像情報メディア学会技術報告 VOL.25, NO.53, pp. 17-22, 2001.

(平成 0 年 0 月 0 日受付)

(平成 0 年 0 月 0 日採録)



延原 章平 (学生会員)

2000 年京都大学工学部電気電子工学科卒業 . 2002 年京都大学大学院情報学研究科博士前期課程修了 . 現在京都大学大学院情報学研究科博士後期課程在学中 . コンピュータビジョン, コンピュータグラフィックスに関する研究に従事 .



和田 俊和 (正会員)

1990 年東工大大学院博士課程修了。同年岡山大学工学部助手。1998 年京都大学大学院工学研究科助教授。2002 年より和歌山大学システム工学部教授。工学博士。画像理解, パターン認識の研究に従事。平成 7 年 David Marr 賞。平成 9 年情報処理学会山下記念研究賞。平成 11 年電子情報通信学会論文賞各受賞。



松山 隆司 (正会員)

1976 年京大大学院修士課程修了。京大助手, 東北大助教授, 岡山大教授を経て 1995 年より京大大学院電子通信工学専攻教授。現在同大学院情報学研究科知能情報学専攻教授。2002 年学術情報メディアセンター長, 京都大学評議員。工博。画像理解, 人工知能, 分散協調視覚の研究に従事。最近は「人間と共生する情報システム」の実現に興味を持っている。1980 年情報処理学会創立 20 周年記念論文賞, 1990 年人工知能学会論文賞, 1993 年情報処理学会論文賞, 1994 年電子情報通信学会論文賞, 1995 年第 5 回国際コンピュータビジョン会議 Marr Prize, 1996 年国際パターン認識連合 Fellow, 1999 年電子情報通信学会論文賞, 2000 年画像センシングシンポジウム優秀論文賞。情報処理学会前理事。