

特徴を保持した統計的形状モデル

若木 良介^{1,a)} 延原 章平^{1,b)} 西野 恒^{1,c)}

概要: 本研究では、物体の幾何的特徴を保持した統計的形状モデルの生成手法を提案する。統計的形状モデルの生成には、データ間での頂点の対応付けが必要となる。通常、これは一様な（形状の意味的な対応を考慮しない）サンプリング等によって行われるため、その統計的形状モデルは特徴形状を精確に再現し得ない。そこで我々は新たに、幾何的特徴量の自動位置合わせによる意味的対応付けに基づいた統計的形状モデルを導出する。具体的には、まず3次元形状の2次元射影面上における法線画像の2次元位置合わせを行う。次に、幾何的特徴の対応付けの取れたサンプリングを行い、固有値分解による基底計算を行う。本研究では、車を例に取り、本手法の有用性を特徴的幾何形状の再現および補間精度を用いて実証した。その結果、我々の手法は一様サンプリングを用いた場合と比較して、有意に特徴的幾何形状を保つことが確認された。

1. はじめに

物体の3次元形状復元の手法の1つとして、統計的形状モデルを用いた復元がある。これは複数の学習用形状データを元に、形状を表現する空間（モデル）を生成し、その空間上の点で表現される形状によって復元を行う手法である。物体の3次元形状復元には、インバースレンダリング [16] をはじめとした様々な応用があり、それらが実用に供するには、高い精度で形状が復元できることが求められる。

統計的形状モデルの生成に際しては、学習用データの頂点数を揃える必要があるが、これは一般に、空間的に一様なサンプリングによって行われることが多い。こうしたサンプリングは、物体の特徴となる形状を考慮していない。したがって、得られる3次元統計的形状モデルは、物体の概形こそ表現できるものの、特徴となる形状は失われてしまうことが多く、精緻な物体形状復元の用途には向かない。これは、比較的丸みを帯びた形状である自然物に対してはまだしも、角の立った形状の多い人工物に対しては問題となる。

この問題に対処すべく、本研究では人工物として車を例に挙げ、特徴となる幾何形状をより正確に表現できる形状モデルを生成することを目的とする。具体的には、まず3次元空間で表現されるメッシュを2次元平面上に射影し、射影された2次元平面で、物体表面の法線を元に生成され

る特徴量画像の位置合わせを行う。次に、この位置合わせの結果を考慮したサンプリングを行う。これにより、同じインデックスで表される頂点が、全ての車で同じ形状的特徴上の意味を持つ。その後、次元圧縮処理を行うことで、エッジなどの幾何的特徴を保持した形状モデルを得ることができる。

2. 関連研究

形状の統一的な表現に関する既存研究としては、動物に関するもの [15] や、人間に関するもの [17]、車に関するもの [13] などがある。

動物の形状に関する研究では、パーツ分けされたテンプレートメッシュを用意し、学習用データの形状との一致度やパーツの変形に対して定義された損失関数を最小化するように、テンプレートの変形を施す。これにより、テンプレートメッシュと同様にパーツ分けされた学習用データが得られ、全学習用データ間での頂点の対応を取ることができる。人間に関する研究では、胸部やへそなどの特徴点を基準とした対応付けが考えられている。動物や人間の場合は、どちらもその主な特徴となる形状が必ず共通で存在する（四足である等）ため、正確に対応を取ることが可能である。しかし車のような人工物の場合は、特徴となる形状が必ずしも共通して存在しているとは限らない。例えば車ではフロントグリルの穴の数が車種によって異なりうる。

車に関する研究では、まず学習用データとなる各車に対して、角となる8頂点によって形成される6面体を生成する。その後、面の分割と新規頂点の移動を繰り返すことによって、同じ頂点数を持つ学習用データ集合を生成する。

¹ 京都大学大学院 情報学研究科
^{a)} wakaki@vision.ist.i.kyoto-u.ac.jp
^{b)} nob@i.kyoto-u.ac.jp
^{c)} kon@i.kyoto-u.ac.jp

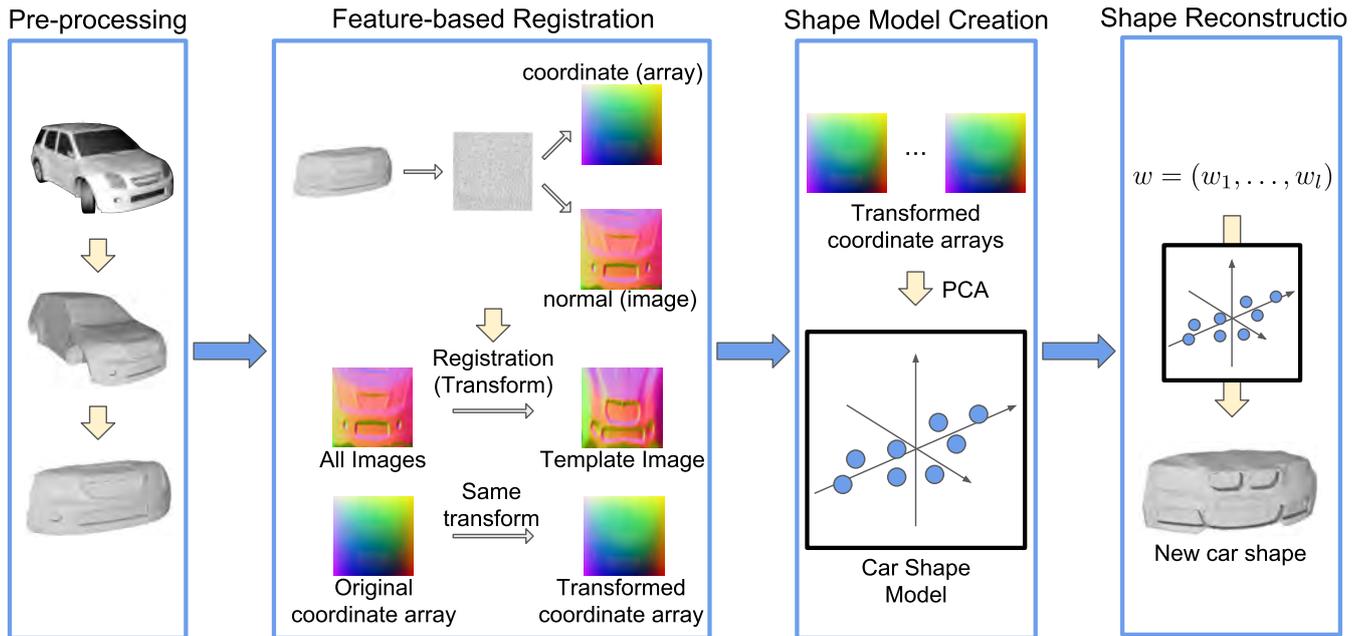


図 1 形状特徴を保持した統計的 3 次元形状モデルの生成フロー。

この手法は特徴となる共通形状を必要としない。しかし、空間的に一様なサンプリングに近いものであるため、同じインデックスで表される頂点でも、学習用データによってその形状的な意味上の場所は異なる。したがって、この手法によって生成される統計的の形状モデルは、物体の特徴となる幾何形状を保たない。

3. 提案手法

本研究で対象とするのは車の外装部分の形状であり、様々な車種での形状の変化を表現する統計的 3 次元形状モデルを生成することが目的となる。本研究における手法の全体図を図 1 に示す。

3.1 学習用データの準備

必要となる学習用データメッシュは、車の外装のみのメッシュである。しかし、データセットに含まれているメッシュは、内装や装飾がついていたり、穴が空いていたりするなどさまざまな問題がある。そのため、まずデータセット中のメッシュを、連続した 1 枚の面で構成される外装のみのメッシュに統一する。

また、得られる外装メッシュは車全体の形状を表すものであるが、本研究ではこれを 5 つの部分メッシュ（前部、右側部、左側部、後部、上部）に分割する。これによって、特徴となる形状が些末にならず十分に表現されるようになり、また後に生成する特徴量画像において、形状の特徴となるエッジがよく表現されるようになる。後にこれらはそれぞれ 2 次元平面に射影するため、部分メッシュがなるべく平面状になるよう分割することが好ましい。

具体的には、まず学習用データの車のメッシュ全てに対して、内装・装飾の除去、及び穴の補修を行う（図 2(b)）。これらは手動で行う。次に、各メッシュに対して複数の視点からの深度マップを取得し、マップ中の各点を 3 次元空間中に投影して点群に変換する（図 2(c)）。その後、Poisson Surface Reconstruction[6] を用いて外装のみの超高密度なメッシュを得る。Poisson Surface Reconstruction とは、点法線を持つ点群を入力として、頂点と面を持つメッシュを出力する手法である。一度点群に変換することにより、元のメッシュのトポロジによらず、すべてのメッシュが連続した 1 枚の面で構成される立体形状となる。この超高密度なメッシュを直接用いると、後に行う平面展開や、特徴量画像の生成において計算量が増えるため、メッシュの単純化 [9][10] を施し、約 100000 頂点のメッシュとして統一する（図 2(d)）。

次に、得られたメッシュに対して手動で特徴点を与え、その特徴点間を結ぶような、メッシュの辺をたどる最短経路によってメッシュを分割する（図 3）。最短経路の探索には Dijkstra 法 [3] を用いる。

3.2 平面展開図における特徴量位置合わせ

3.2.1 部分メッシュの平面展開

分割された部分メッシュを、頂点同士の接続関係を保ったまま平面に展開する。このとき、2 次元射影面上でのメッシュの各頂点について、元の 3 次元空間上での座標と法線を格納しておく。ここでは、2 次元射影面上で、3 次元座標と 3 次元法線の情報を持つ頂点が疎に分布している。この疎な頂点の間を補間することにより、座標情報と法線情



図 2 学習データの準備. (a):元データとなる車のメッシュ. (b):内装・装飾などを取り除いたメッシュ. (c):(b)のメッシュに対して各方向から深度マップを取得し、統合して得られる点群. (d):(c)で得られた点群に対して、Poisson Surface Reconstruction を施し、メッシュを再構成する. その後、約 100000 頂点となるようにメッシュの単純化を施す.

報を持った密な頂点集合を得ることができる.

部分メッシュの 2 次元射影面への変換は、Mean Value Coordinates を用いた平面展開 [4][5] を基本とした手法を用いて行う. ここでは、辺の長さが 1 の正方形に平面展開するものとし、平面を張る直交軸を (u, v) する.

メッシュに含まれる頂点は、部分メッシュ内での通し番号 (グローバルインデックス) か、部分頂点集合の中での通し番号 (ローカルインデックス) で表現する. ローカルインデックスを用いる場合は都度明示する. 初めに、メッシュの頂点 i, j を結ぶ辺を $e_{i,j}$ とし、全ての $e_{i,j}$ に対して重み $w_{i,j}$ を定義する

$$w_{i,j} = \frac{\tan \frac{\theta_{i,j,0}}{2} + \tan \frac{\theta_{i,j,1}}{2}}{|e_{i,j}|}. \quad (1)$$

ただし、 $\theta_{i,j,0}, \theta_{i,j,1}$ はそれぞれ、 $e_{i,j}$ が含まれる三角形 2 つの、 $e_{i,j}$ に対する対角である. この $w_{i,j}$ を用いて、まず部分メッシュの境界上の頂点の平面座標を決定する.

2 次元射影面における正方形の四隅に配置する頂点のインデックスを a, b, c, d とする. この頂点は、第 3.1 節で部分メッシュの切り出しに用いた特徴点から選択する. ここで、頂点 a, b の 2 次元座標をそれぞれ $(0, 0), (1, 0)$ とする. 頂点 a, b 間の切り出しの経路に含まれる頂点のインデックスを a 側から順に $a, 1, \dots, k, b$ としてローカルインデックス $1, \dots, k$ を定義すると、 $1, \dots, k$ の平面空間上での u 座標 u_1, \dots, u_k は、連立方程式

$$\begin{pmatrix} 1 & \frac{-w_{1,2}}{w_{a,1}+w_{1,2}} & 0 & 0 \\ \frac{-w_{1,2}}{w_{1,2}+w_{2,3}} & 1 & \frac{-w_{2,3}}{w_{1,2}+w_{2,3}} & 0 \\ 0 & \frac{-w_{2,3}}{w_{2,3}+w_{3,4}} & 1 & \dots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_k \end{pmatrix} = \begin{pmatrix} \frac{w_{a,1}}{w_{a,1}+w_{1,2}} \times 0 \\ 0 \\ 0 \\ \vdots \\ \frac{w_{k,b}}{w_{k-1,k}+w_{k,b}} \times 1 \end{pmatrix} \quad (2)$$

を解くことによって得られる. ただし、式 (2) の右辺は、

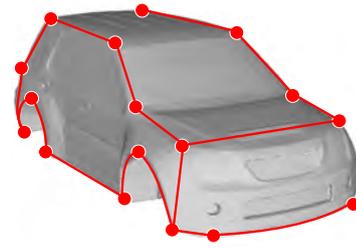


図 3 選択する特徴点と部分メッシュ分割の分割経路図. 車 1 台あたり 26 箇所の特徴点を手動で選択した. 分割は、Dijkstra 法によって求められる特徴点間の最短経路によって行う.

第 1 行が a の u 座標が 0, 第 k 行が b の u 座標が 1 であることに由来し、残りの成分はすべて 0 である. 残りの辺についても、同様の連立方程式を解くことによって頂点の座標を決定する.

その後境界上にないすべての頂点 i に対して、ローカルインデックス $1, \dots, N$ を与え、その頂点の座標 (u_i, v_i) を求める. (u_i, v_i) は連立方程式

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \quad (3)$$

$$\mathbf{u} = \begin{pmatrix} u_1 & v_1 \\ \vdots & \vdots \\ u_N & v_N \end{pmatrix}, \quad (4)$$

を解くことにより得られる. ここで、 \mathbf{A}, \mathbf{b} は次のようにして生成する. まず $\mathbf{A} = \mathbf{I}, \mathbf{b} = \mathbf{0}$ として初期化を行う (\mathbf{I} は単位行列). 次に、境界上にない全ての頂点 i について、隣接するすべての頂点となす辺の重みの総和を W_i , 隣接する頂点のインデックスを j とする. その全ての j に対して、

(i) j が境界上の点であるとき

$$j \text{ はグローバルインデックスとし,} \\ \mathbf{b}[i, 0] \leftarrow \mathbf{b}[i, 0] + \frac{w_{i,j}}{W_i} u_j, \quad (5)$$

$$\mathbf{b}[i, 1] \leftarrow \mathbf{b}[i, 1] + \frac{w_{i,j}}{W_i} v_j, \quad (6)$$

(u_j, v_j は前述の手順により既知)

(ii) j が境界上でない (内部の) 点であるとき

$$j \text{ はローカルインデックスとし,} \\ \mathbf{A}[i, j] \leftarrow \mathbf{A}[i, j] - \frac{w_{i,j}}{W_i}, \quad (7)$$

と値の更新を行う. ただし、 $[i, j]$ はその行列の i 行 j 列成分を表す. すべての頂点 i に対してこの処理を行い、求められた \mathbf{A}, \mathbf{b} に対して式 (3) を解くことで、頂点の射影面上での座標を決定する.

3.2.2 頂点情報の補間と特徴量画像の生成

平面に射影されたメッシュに対して、頂点と、座標・法線情報の補間は、補間したい点が含まれるメッシュ上の三角形に対して、重心座標系を用いて 3 頂点の値の重み付

き和を取ることで行う。まず平面に射影したメッシュの頂点のインデックスを i とし、その頂点をもつ 3 次元座標を $\mathbf{x}_i = (x_i, y_i, z_i)$ 、3 次元法線を $\mathbf{n}_i = (n_i^x, n_i^y, n_i^z)$ 、平面空間上での座標を $\mathbf{P}_i = (u_i, v_i)$ とする。なお、3 次元法線のノルムは 1 に正規化されているものとする。このとき、射影平面上の任意の点 \mathbf{P} の持つ 3 次元座標 \mathbf{x} 、3 次元法線 \mathbf{n} を

$$\mathbf{x} = w_0\mathbf{x}_i + w_1\mathbf{x}_j + w_2\mathbf{x}_k, \quad (8)$$

$$\mathbf{n} = w_0\mathbf{n}_i + w_1\mathbf{n}_j + w_2\mathbf{n}_k, \quad (9)$$

とする。ただし、 w_0, w_1, w_2 は、 $S(\mathbf{P}_i\mathbf{P}_j\mathbf{P}_k)$ を三角形 $\mathbf{P}_i\mathbf{P}_j\mathbf{P}_k$ の面積として、

$$w_0 = \frac{S(\mathbf{P}\mathbf{P}_j\mathbf{P}_k)}{S(\mathbf{P}_i\mathbf{P}_j\mathbf{P}_k)}, \quad (10)$$

$$w_1 = \frac{S(\mathbf{P}_i\mathbf{P}\mathbf{P}_k)}{S(\mathbf{P}_i\mathbf{P}_j\mathbf{P}_k)}, \quad (11)$$

$$w_2 = \frac{S(\mathbf{P}_i\mathbf{P}_j\mathbf{P})}{S(\mathbf{P}_i\mathbf{P}_j\mathbf{P}_k)}, \quad (12)$$

として求める。このとき、 $\mathbf{P}_i, \mathbf{P}_j, \mathbf{P}_k$ は、メッシュに含まれる面の 3 頂点であり、かつ平面空間においてその面の内部、もしくは辺上に \mathbf{P} を含むものとする。

上記の方法により、車の台数分の平面展開図が得られるが、異なる車の間では一般的に、同じ 2 次元射影面上の座標でも 3 次元空間上で指す形状的な意味上の位置は異なる。これは、生成される形状モデルが特徴形状を保たない原因となる。そこで、3 次元法線によって定義される特徴量画像を生成し、その位置合わせを行う。これにより、2 次元射影面上の任意の座標において、その 3 次元空間上の点がすべての車で形状的な意味が同一の場所を指すようにする。

特徴量画像の生成は以下のように行う。まず平面メッシュと同じ大きさの画像を考え、画素の射影平面上での座標を \mathbf{P} とする。式 (9) によって与えられる法線 \mathbf{n} を用いて、その画素の画素値 $\mathbf{C} = (b, g, r)$ を

$$\mathbf{C} = \frac{\mathbf{n} + 1}{2} \in [0, 1]^3 \quad (13)$$

とし、これをすべての画素に対して求める。法線のノルムは 1 に正規化されていることから、その各成分は $[-1, 1]$ の範囲にある。したがって、 \mathbf{C} の各成分は $[0, 1]$ の範囲にある。このようにして得られる画像は図 4 であり、3 次元形状における凹みの部分など、特徴となる形状をよく表現したものとなる。

3.2.3 特徴量画像の位置合わせ

得られた特徴量画像同士の位置合わせを行い、全頂点について形状的な意味上の対応付けを行う。位置合わせは、1 台の車の特徴量画像をテンプレートとし、他の特徴量画像全てに、テンプレートと一致するような変形を施すことで行う。

ここで、画像の四隅は第 3.1 節において手動で与えた特徴

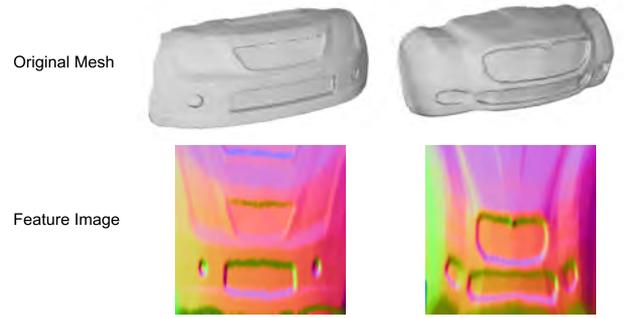


図 4 部分メッシュとその特徴量画像。特徴量画像は、メッシュの特徴となるエッジをよく表現している。

点であるため、意味上の対応付けは既に行われている。これに伴って、画像の各辺も部分メッシュ同士の境界部分と対応付けられており、この部分に関しては既に意味上の対応関係が取られている。内部の点に関しては、フロントグリルの凹みなどの部分同士で意味上の対応関係を考えることができる。これら内部の点の位置合わせのために、まず画像間の粗い対応点組を、Neural Best-Buddies (NBB)[7] を用いて取得する。NBB とは、画像認識ネットワークである VGG-19[12] を用いて、2 枚の画像中の対応点組を取得する手法である。対応点組が不足していたり、誤った対応点組が存在したりしている場合は、これらを手動で修正する。このとき、2 枚の画像組 I_s, I_t の四隅と各辺の各画素及び NBB を用いて得られる内部の対応点の座標を、 $(\boldsymbol{\mu}_i^s, \boldsymbol{\mu}_i^t) = ((x_i^s, y_i^s), (x_i^t, y_i^t))$ (以降「ランドマーク」と呼ぶ。 $i = 1, \dots, N$) とする。2 画像の位置合わせは、エネルギー関数

$$E = w_i E_{img} + w_\mu E_\mu + (w_d E_{div} + w_r E_{rot}) \quad (14)$$

を最小化するように、一方の画像に 3 次の B スプライン変形 [1] を施すことにより行う。ここで、 E_{img} は 2 画像間の二乗誤差、 E_μ はランドマーク座標の二乗誤差、 E_{div}, E_{rot} はそれぞれ変形の発散・回転に関する正則化項であり、 w_i, w_μ, w_d, w_r は各項の重みを決めるハイパーパラメータである。画像から画像への変形のベクトル場を $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ とすると、具体的には、

$$E_{img} = \sum_{x,y} \{I_t(x, y) - I_s(g(x, y))\}^2, \quad (15)$$

$$E_\mu = \frac{1}{N} \sum_{i=1}^N |\boldsymbol{\mu}_i^s - g(\boldsymbol{\mu}_i^t)|^2, \quad (16)$$

$$E_{div} = \int_{\mathbb{R}^2} |\nabla \text{div} g|^2 dx dy, \quad (17)$$

$$E_{rot} = \int_{\mathbb{R}^2} |\nabla \text{rot} g|^2 dx dy, \quad (18)$$

である [1] [11]。位置合わせを行う際の B スプライン変形の制御点の数は、初めは 2×2 とし、エネルギー関数が収束

した後に 8×8 として再度エネルギー関数の最小化を行う。

最後に、特徴量の位置合わせで得た変形のベクトル場 g を用いて、同じ変形を射影平面上の頂点集合に施す。これによって、平面上のすべての点に形状的な意味上の対応付けが行われる。

3.3 統計的 3 次元形状モデルの生成

一般に統計的モデルは、すべての学習用データが同じ形式である場合に生成することができる。今、第 3.2 節での処理によって、任意の 2 次元射影面上の点が、全ての車に対して同じ形状的意味上の場所を指している。したがって、この 2 次元射影面上で、適当なサンプリングを行うことで、幾何的特徴を考慮しながらも、全データの頂点数を揃えることができる。なお統計的モデルの生成方法として、本研究では主成分分析を用いる。

車の台数を M とし、各車 p に対して第 3.2 節で形状的な意味上の対応付けが行われた状態の頂点集合から、 125×125 点、計 15625 点の頂点をサンプリングし、その頂点集合を \mathcal{A}_p とする。各車のパーツは、その中心が車種によって様々であり統一されていないので、バウンディングボックスの中心が原点になるように中心化を行う。中心化する前の車の頂点集合 \mathcal{A}_p に含まれる頂点を $\mathbf{x}_{p,i} = (x_{p,i}, y_{p,i}, z_{p,i})$ 、中心化した後の車の頂点集合 \mathcal{A}_p^c に含まれる頂点を $\mathbf{x}_{p,i}^c = (x_{p,i}^c, y_{p,i}^c, z_{p,i}^c)$ とする。車 p のバウンディングボックスの中心を $\bar{\mathbf{x}}_p = (\bar{x}_p, \bar{y}_p, \bar{z}_p)$ とすると、

$$\bar{x}_p = \left(\frac{\min_i(x_{p,i}) + \max_i(x_{p,i})}{2} \right) \quad (19)$$

$$\bar{y}_p = \left(\frac{\min_i(y_{p,i}) + \max_i(y_{p,i})}{2} \right) \quad (20)$$

$$\bar{z}_p = \left(\frac{\min_i(z_{p,i}) + \max_i(z_{p,i})}{2} \right) \quad (21)$$

である。このとき、中心化した後の頂点の座標は、

$$\mathbf{x}_{p,i}^c = \mathbf{x}_{p,i} - \bar{\mathbf{x}}_p \quad (22)$$

として求められる。また、これを用いて車 p の頂点座標を 1 行のベクトル \mathbf{d}_p で表すとす。すなわち、

$$\mathbf{d}_p = (x_{p,1}^c, y_{p,1}^c, z_{p,1}^c, \dots, x_{p,N}^c, y_{p,N}^c, z_{p,N}^c) \quad (23)$$

とする。ただし、 $N = 15625$ である。これをすべての車に対して並べたものをデータ行列 \mathbf{D} ,

$$\mathbf{D} = \begin{pmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_M \end{pmatrix}, \quad (24)$$

とする。また、全車の平均形状を $\bar{\mathbf{d}}$,

$$\bar{\mathbf{d}} = \frac{1}{M} \sum_{i=1}^M \mathbf{d}_i, \quad (25)$$

とし、データ集合を中心化する

$$\mathbf{D}_c = \mathbf{D} - \begin{pmatrix} \bar{\mathbf{d}} \\ \vdots \\ \bar{\mathbf{d}} \end{pmatrix} = \begin{pmatrix} \mathbf{d}_1 - \bar{\mathbf{d}} \\ \vdots \\ \mathbf{d}_M - \bar{\mathbf{d}} \end{pmatrix}. \quad (26)$$

こうして得られた \mathbf{D}_c に対して、

$$\mathbf{E}\mathbf{\Lambda}\mathbf{E}^\top = \frac{1}{M} \mathbf{D}_c \mathbf{D}_c^\top \quad (27)$$

なる正方行列 $\mathbf{E} = (\mathbf{v}_1^\top, \dots, \mathbf{v}_M^\top) \in \mathbb{R}^{M \times M}$ と $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_M)$ なる対角行列を考える。このとき

$$\mathbf{u}_i = \frac{1}{\sqrt{M\lambda_i}} \mathbf{D}_c^\top \mathbf{v}_i \quad (\in \mathbb{R}^{3N}) \quad (28)$$

として求められる \mathbf{u}_i を、第 i 主成分ベクトルと言う。 $l \leq M$ なる l に対して、 $\{\mathbf{u}_1, \dots, \mathbf{u}_l\}$ によって張られる線形部分空間が、統計的形状モデルとなる。特徴量位置合わせの結果、この形状モデルはより車の幾何的形状に関する意味的特徴の変化に着目していると言える。

3.4 形状復元

第 3.3 節で得られた形状モデルはパーツごとの形状モデルであるが、応用にあたって生成したい 3 次元形状はパーツ単位の形状ではなく、全体形状である。全体形状の生成にあたっては、まず各パーツを各々の形状モデルを用いて生成する。次に、各パーツの 3 次元位置を調整し、隣り合うパーツの境界がなるべく一致するようにする。この状態では、各パーツ同士は接続されておらず、ただ複数の部分メッシュが 3 次元空間上に置かれているだけである。全体形状を 1 つのメッシュにするための手法としては、各パーツの頂点だけを抜き出して点群集合とし、第 3.1 節で用いた Poisson Surface Reconstruction を施すなどが考えられる。他にも、Shrink Wrapping[8] という手法を用いることも考えられる。

パーツ単位の復元は、そのパーツの形状モデルと、形状を表現する重みベクトル $\mathbf{w} = (w_1, \dots, w_l)$ を用いて行うことができる。具体的には、形状モデルの主成分ベクトル \mathbf{u}_i と、データ平均 $\bar{\mathbf{d}}$ を用いて、

$$\mathbf{d}^{new} = \mathbf{w} \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_l \end{pmatrix} + \bar{\mathbf{d}} \quad (29)$$

として、形状を得ることができる。

4. 評価実験

第 3 節で述べた手法を元に、実際に形状モデルの生成を行い、その評価を行う。第 4.1, 4.2, 4.3 節では、ShapeNet[2] の乗用車 30 台分のメッシュデータを用い、第 3.1 節で分割したパーツのうち、前面部分についての形状モデルについて行う。第 4.4 節では台数を増やし、42 台で行う。実施した実験は以下の 4 つである。

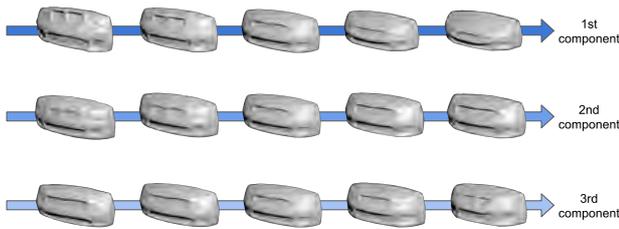


図 5 位置合わせを行ってから生成した形状モデルの各主成分表現. 各主成分の重みの変化による形状の動きは, 車の特徴となるエッジ部分の動きとして現れる.

主成分空間の可視化

実際に得られる形状モデルがどのような形状表現を実現しているかについて, 主成分ごとの形状変化を観察し, 定性的評価を行う.

2 台の車の中間表現

2 台の車の重みベクトルの加重平均をとり, その重みによって表現される形状を可視化することにより, 定性的評価を行う.

中間形状の意味的妥当性

2 台の車の中間形状が, 位置合わせ後の特徴空間において一定の特徴を示すか (意味の上で一貫性を持ち, 車らしい形状を保つか), 定量的評価を行う.

全体形状表現

2 台の車の中間形状に関して, 部分メッシュ単位ではなく全体メッシュの生成を行う.

4.1 主成分空間の可視化

求められた統計的形状モデルは, 第 3.3 節で述べたとおり, 主成分によって張られるデータ空間の線形部分空間である. 本節では, 各主成分方向に沿って変化する形状を実際に生成し, その特徴を観察する. また, 第 3.2.3 節の特徴量位置合わせを行わずに頂点のサンプリングを行って生成した形状モデルに対しても同様の実験を行い, 両モデルの比較を行う. 主成分分析によって得られた, 形状モデルを張るベクトル組を $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_M)$ とし (ただし, $M = 30$), 重みベクトル $\mathbf{w} = (w_1, \dots, w_M)$ を,

$$\begin{aligned} w_i &= x \ (\in [-a_i, a_i]), \\ w_k &= 0 \ (k \neq i), \end{aligned}$$

を満たす x により変化させる. ただし a_i は, 第 i 主成分の重み係数 w_i が学習用データ集合に対して $[w_i^{min}, w_i^{max}]$ の範囲で値を取るとき,

$$a_i = \min(|w_i^{min}|, |w_i^{max}|) \quad (30)$$

として求めるものとする. この \mathbf{w} を式 (29) に適用することで, 第 k 主成分の変化に対して得られる形状を観察する.

第 1, 第 2, 第 3 主成分の重みを変化させた場合の形状の変化を図 5, 図 6 に示す. いずれの主成分についても, 位

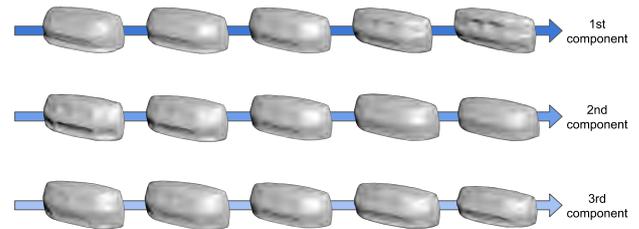


図 6 位置合わせを行わずに生成した形状モデルの各主成分表現. 各主成分の重みの変化による形状の動きは, エッジの消失と発現として現れる.

置合わせを行っていない場合は, 穴のふちのエッジが一度消失してから新たに別の部分のエッジとして現れる変形になっているが, 位置合わせを行った場合は, 穴のエッジを常に保ちながらその形状を変化させていることがわかる. これは, 位置合わせを行った場合は各頂点が意味の上で同じ場所に存在し続けるためである. より具体的には, 穴のふちにあった頂点は, 変形する過程の中でも常に穴のふちにあるためである.

4.2 2 台の車の中間表現

本実験では, 形状モデルを生成するに当たって用いた車のデータ 2 台分に対して, その中間形状の変化を観察することによって, 特徴量位置合わせの有用性を検証する. 2 台の車 A, B に対して, 形状モデル上でこれを表現する重みベクトルをそれぞれ $\mathbf{w}^A, \mathbf{w}^B$ とする. これらを用いて,

$$\mathbf{w}^{new} = (1 - a)\mathbf{w}^A + a\mathbf{w}^B \ (a \in [0, 1])$$

によって \mathbf{w}^{new} を求め, これを式 (29) に適用することで中間形状を得る.

用いた車と, これより得られた中間形状を図 7 に示す. 位置合わせを行っていない場合は, 頂点の意味上の位置が穴の外部, 穴のエッジ, 穴の内部と変化するため, 変形の中間過程では穴の外部からエッジになる過程の頂点と, エッジから穴の内部になる頂点が同時に存在する. その結果, 図 8 の赤枠に見られるような, 二重のエッジが現れることから, 形状的特徴の変化を適切に表現できていないことがわかる. 一方, 位置合わせを行った場合は, 変形のどの時点においても穴のエッジは 1 つであり, 車の形状として自然であることがわかる.

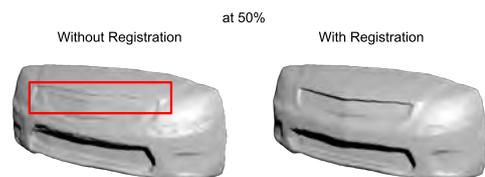


図 8 (A:B = 50:50) 点における中間形状. 赤枠の部分において, 二重のエッジ (エッジの重ね合わせ) が見られる.

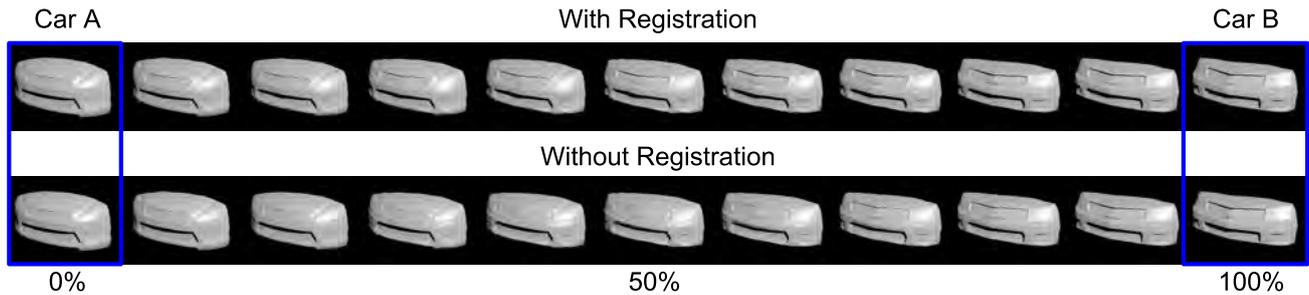


図 7 2 台の車の中間形状の推移. 位置合わせありのモデルから得られる中間形状は, エッジの移動として現れるのに対し, 位置合わせなしのモデルから得られる中間形状は, 2 つの車のエッジの重ね合わせとして現れる.

4.3 中間形状の意味的妥当性

第 4.2 節では, 2 台の車に対してその中間形状を観察することにより, 位置合わせありのモデルの方が, より車の幾何的形狀として適切な中間形状を表現できることを定性的に確認した. 本節では, その中間形状について特徴量画像を再び生成し, 2 台の車の特徴量位置合わせ時の中間変形を施す. その後, テンプレート特徴量画像との類似度を測ることで, 中間形状が車らしい特徴を保持しているかどうか評価する.

図 9 に実験の概要を示す. まず 2 台の車 A,B の中間形状データを両モデルより生成し, それを元に特徴量画像を生成する. 得られる特徴量画像は画像サイズが 125×125 px であるが, これをテンプレート特徴量画像と同じ大きさ (500×500 px) に拡大する. ここで, 位置合わせありのモデルから生成された形状の特徴量画像は, 既にテンプレートとの位置合わせが行われているが, 位置合わせなしのモデルから生成された形状の特徴量画像は, テンプレートとの位置合わせが行われていない. そこで, 後者のモデルから生成された特徴量画像に対しては, 車 A,B の位置合わせ時のベクトル場の平均を取り, その変形を施す. これによって, 特徴量の 2 次元的位置に関する条件を揃える. なお, 図 9 における R の文字は, テンプレート特徴量画像との位置合わせが行われている状態を示す.

学習用データ 30 台に対し, そのすべての組み合わせにおいて中間形状を生成し, 図 9 中の Similarity A と Similarity B を測る. 類似度の測定には, 周囲の画素値との相関によって評価を行う SSIM[14] を用いる. SSIM は $[0, 1]$ の値を取り, 値が大きいほど 2 画像の類似度が高いことを示す.

位置合わせありのモデルと, 位置合わせなしのモデルのそれぞれに対して, 類似度の平均と分散を表 1 に示す. 位置合わせありのモデルから生成された中間形状の特徴量画像の方が, テンプレート特徴量画像との平均類似度が高いことがわかる. すなわち, 位置合わせありのモデルから得られた中間形状の方が, より実在する車に近い形状特徴を持つことが示された.

4.4 全体形状表現

ここでは, 2 台の車の中間形状に関して, 全ての部分メッシュを統合し, 1 つの全体メッシュにする. 全ての部分メッシュが同じ 2 台の車の中間表現である場合は, 隣り合う部分メッシュ同士の境界は完全に一致する. そこで, それらを貼り合わせることで全体形状を生成することができる.

図 10 に, 位置合わせを行ったモデルから生成される中間表現の例を示す. いずれも実在しない架空の車であるが, そのほとんどが車らしい形状となっていることがわかる.

5. 結論

本研究では, 2 次元射影面上における特徴量位置合わせを用いることによる, 特徴的幾何形状を保った車の統計的 3 次元形状モデルの生成手法について述べ, その定性的・定量的評価を行った.

形状モデルの生成に際し, 複数のデータ間で特徴量位置合わせによって形状的な意味上の対応関係を取ることで, その形状モデルから得られる新たな形状は, その形状モデルのもととなった物体の特徴を保持することが確認できた. 特に, 学習に用いたデータの中間形状においては, エッジの場所など, 意味の上で中間と言うべき形状が得られ, 学習用データのカテゴリに属する物体として遜色ない形状が得られることが確認できた. また, 今回は車に対しての適用であったが, 同じカテゴリの物体に特徴となるような類似形状がある場合は, 本研究での提案手法によって, その特徴を保った形状モデルを生成できることが見込まれる.

今後解決すべき課題としては, NBB だけでは十分な精度が得られず, 手動での修正が必須であることが挙げられる. 本研究の展望としては, NBB に代わる, より高精度で手動での修正が必要ない対応点探索アルゴリズムの構築による, 位置合わせの完全自動化が目標となる.

謝辞

本研究の一部は JSPS 科研費 17K20143 の助成を受けたものです.

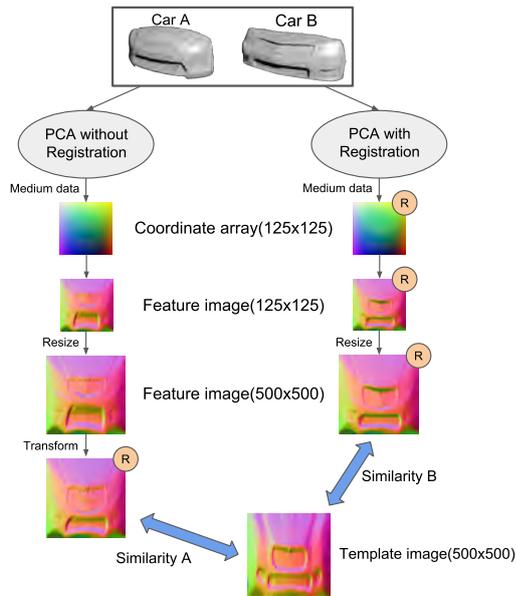


図 9 中間形状特徴量画像とテンプレート特徴量画像の類似度測定の概要。両モデルから生成されるメッシュ(頂点座標配列)の特徴量画像を再構成し、位置合わせなしのモデルから生成されるメッシュの特徴量画像に対しては、車 A,B それぞれの特徴量位置合わせに用いた変形のベクトル場の平均を取り、変形を施す。

参考文献

[1] Arganda-Carreras, I., Sorzano, C. O., Marabini, R., Carazo, J. M., Ortiz-de Solorzano, C. and Kybic, J.: Consistent and elastic registration of histological sections using vector-spline regularization, *International Workshop on Computer Vision Approaches to Medical Image Analysis*, Springer, pp. 85–95 (2006).

[2] Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L. and Yu, F.: ShapeNet: An Information-Rich 3D Model Repository, Technical Report arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago (2015).

[3] Dijkstra, E. W.: A note on two problems in connexion with graphs, *Numerische mathematik*, Vol. 1, No. 1, pp. 269–271 (1959).

[4] Floater, M. S.: Mean value coordinates, *Computer aided geometric design*, Vol. 20, No. 1, pp. 19–27 (2003).

[5] Hormann, K., Lévy, B. and Sheffer, A.: Mesh parameterization: Theory and practice (2007).

[6] Kazhdan, M., Bolitho, M. and Hoppe, H.: Poisson Surface Reconstruction, *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, SGP

表 1 各モデルにおける、中間形状の特徴量画像とテンプレート特徴量画像の類似度の平均と分散。

	位置合わせなし	位置合わせあり
平均類似度	0.76481	0.80672
分散	0.02328	0.02516

'06, Aire-la-Ville, Switzerland, Switzerland, pp. 61–70 (2006).

[7] Kfir, A., Jing, L., Mingyi, S., Dani, L., Baoquan, C. and Daniel, C.-O.: Neural Best-Buddies: Sparse Cross-Domain Correspondence, *ACM Transactions on Graphics (TOG)*, Vol. 37, No. 4, p. 69 (2018).

[8] Kobbelt, L. P., Vorsatz, J., Labsik, U. and Seidel, H.-P.: A shrink wrapping approach to remeshing polygonal surfaces, *Computer Graphics Forum*, Vol. 18, No. 3, Wiley Online Library, pp. 119–130 (1999).

[9] Lindstrom, P. and Turk, G.: Fast and memory efficient polygonal simplification, *Visualization'98. Proceedings, IEEE*, pp. 279–286 (1998).

[10] Lindstrom, P. and Turk, G.: Evaluation of memoryless simplification, *IEEE Transactions on Visualization & Computer Graphics*, No. 2, pp. 98–115 (1999).

[11] Sánchez Sorzano, C., Thévenaz, P. and Unser, M.: Elastic registration of biological images using vector-spline regularization, *IEEE Transactions on Biomedical Engineering*, Vol. 52, No. LIB-ARTICLE-2005-003, pp. 652–663 (2005).

[12] Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).

[13] Umetani, N.: Exploring generative 3D shapes using auto-encoder networks, *SIGGRAPH Asia 2017 Technical Briefs*, ACM, p. 24 (2017).

[14] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P.: Image quality assessment: from error visibility to structural similarity, *IEEE transactions on image processing*, Vol. 13, No. 4, pp. 600–612 (2004).

[15] Zuffi, S., Kanazawa, A., Jacobs, D. W. and Black, M. J.: 3D Menagerie: Modeling the 3D shape and pose of animals., *CVPR*, pp. 5524–5532 (2017).

[16] 佐藤洋一, 向川康博ほか: インバースレンダリング, 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM), Vol. 2004, No. 91 (2004-CVIM-145), pp. 65–76 (2004).

[17] 産業技術総合研究所: AIST/HQL 人体寸法・形状データベース 2003, <https://unit.aist.go.jp/hiri/dhrg/ja/dhdb/fbodyDB/model/index.html>.

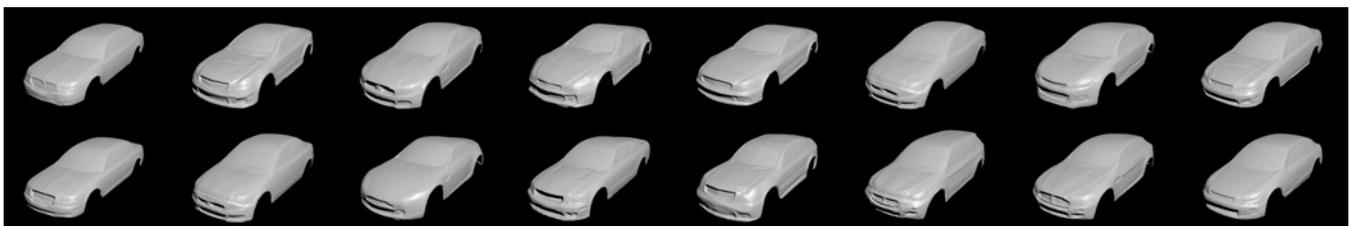


図 10 中間表現の全体形状。いずれも形状モデルから生成された架空の車である。