

面の対応付けによる家具可動部の全周形状モデリング

3D Shape Modeling of Articulated Part of Furniture Based on Temporal Surface Correspondence

正 山崎 公俊 (信州大/東大) 宗玄 清宏 山本 貴史 (トヨタ自動車) 正 稲葉 雅幸 (東大)

*Kimitoshi YAMAZAKI (Shinshu Univ. / The Univ. of Tokyo), Kiyohiro Sogen,
Takashi Yamamoto (Toyota Motor Corp.) and Masayuki INABA (The Univ. of Tokyo)
kyamazaki@shinshu-u.ac.jp

This paper describes a modeling method that enables to generate 3D shape model with articulated link. Furniture such as refrigerator, shelf and cabinet are targeted in this research. The method is based on the tracking of 3D surfaces extracted by using depth images captured from a range camera. Both exterior and internal sides of an articulated part can be modeled as a 3D mesh model. The position of articulated link can also be estimated.

Key words: 3D Shape Modeling, Articulated Object.

1 はじめに

生活環境下で支援行動を行うロボットには、物品や家具等を操作する能力が求められる。操作を行うためには、形状や可動構造を定義したモデルをあらかじめ与えておく必要がある。しかしながら、計算機上でそのようなモデルを人手で逐一作成することは、手間のかかる作業である。

本稿では、冷蔵庫や引き出しといった可動構造を持つ家具のモデルを生成する方式について報告する。人が家具可動部を操作している最中に三次元距離画像センサで観測を行い、そこから得た距離画像列より面を抽出し、その時間方向の連続性を利用して、表面・裏面を一体とする形状復元を行う。また、可動軸の推定も行う。

可動構造を持つ物体のモデル化については、従来からいくつかの研究成果がある。Sturm ら [4] は、可動部の前板を発見し追跡を行うことで、マーカー等を付加することなく、家具のモデリングを実現した。Yamazaki ら [5] は、家具の見た目と操作方法を記憶するための表現として IM モデルを提案し、いくつかの家具について、人間による教示とロボットによる操作までを実現した。これらの研究は、引き出しの前板など可動部分にのみ着目しており、モデル化の範囲が部分的であった。

2 家具可動部モデリングの課題とアプローチ

生活環境下には扉や引き出し等の可動部が付いた家具が多く存在する。本研究では、収納物を単純に隠す目的としての扉だけでなく、たとえば冷蔵庫などのように内側にも収納機能を持つ扉といった、可動構造をもつ家具を対象とする。そのような家具のモデルを認識と操作の両方で用いるためには、可動構造と、可動部の全周形状の両方が必要である。なぜなら、外側から見える面はその家具を発見してアプローチするために必要であり、内側の面は、家具を操作した後に収納行動などを計画する

ために必要だからである。

本研究では、家具可動部のモデルを人手で逐一定義するのではなく、人が家具を操作する様子をセンシングする、もしくはロボット自身が家具を試行錯誤的に操作している最中にセンシングを行い、そこで得られたデータから自動でモデリングを行う方針を採る。

上記のような状況下でモデリングを行う場合、下記のような課題がある。

1. 単視点で計測を行う限り、可動部の前面と背面が同時に観測されることはないため、全周形状を得るには時系列データを適切につなぐ必要がある。
2. センサと家具の距離を小さく設定せざるを得ない場合があり、センサの観測範囲が狭いと、可動部全体を常に計測することが難しい。また、人の手などにより可動部の一部が隠される状況も存在する。

従来、時系列データをつなぐための方法として、画像特徴点を抽出・追跡し、三次元形状復元を行うものがある。しかしながら、冷蔵庫や棚等の家具は十分なテクスチャを持たない上、光を反射しやすい素材で構成されていることが多く、そのような方針が適用できない場合も多い。

これらの課題の元で、本研究では次のようなアプローチでモデリング手法を構築する。まず、センサとして三次元距離画像センサを利用する。得られる三次元点群から面情報を抽出し、それを時系列で追跡することによって、可動部全体の移動変化を推定する。その後、各時刻で得られている三次元点列を三次元空間上で統合し、一つの可動部形状を得る。

3 面の抽出と対応付け

3.1 面の抽出

三次元点群から面情報を得る手段として、Region growing 法を用いる。この手法は、ある基点を選んだ後、その

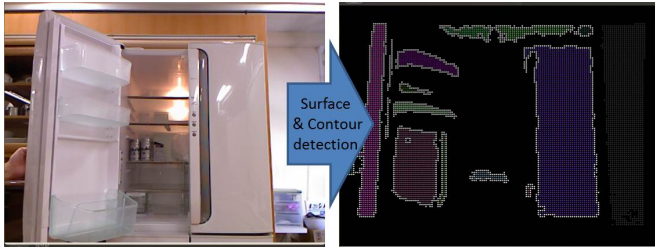


Fig. 1: A results of Surface and contour detection

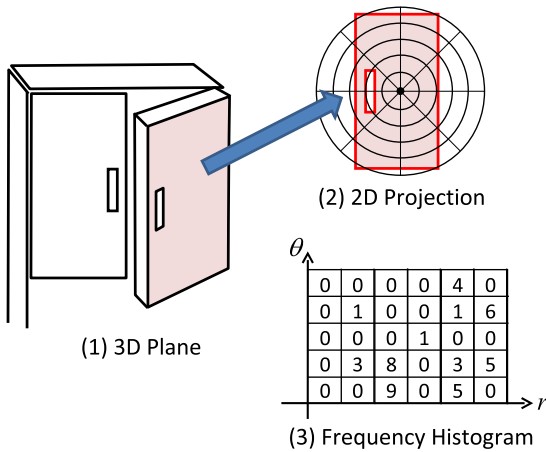


Fig. 2: Shape context

基点の近傍で基点と性質が近いものを同じクラスとして記録していき、同族領域を拡張していく手法である。

Region growing を適用するための前処理として、計測された三次元点のそれぞれについて法線成分を求める。この処理は、各点の近傍点数を参照して面パッチを生成し、その法線成分を計算するものである。すなわち、一つの画素には、奥行き値 d と法線成分 $\mathbf{n} = (n_x, n_y, n_z)$ の4成分が保持されることになる。

この結果に対して、似た法線成分を持ち、奥行き値が大きく変化しない画素を連結させていく。すなわち、以下を満たす場合について、領域拡張を行う。

$$\begin{aligned} |d(i, j) - d(i+n, j+m)| &< d_{\text{threshold}} \\ \cos^{-1}(\mathbf{n}(i, j) \cdot \mathbf{n}(i+n, j+m)) &> \theta_{\text{threshold}} \end{aligned} \quad (1)$$

ここで、 (i, j) は距離画像上の現在の座標値を示しており、 $(i+n, j+m)$ はその近傍の座標値を意味する。 \cdot は内積を表す。この方針によれば、平面検出法を適用して面要素を抽出すること [3] と比較し、ある程度の曲率を持った面に対しても適用できるという利点がある。Fig.1 に、冷蔵庫を開けた状態において面抽出を行った結果を示す。

3.2 面の対応付け

時刻の異なる距離画像からそれぞれ面群を抽出し、時間間での対応付けを行う。このためには、面と面との類

	time									
Serial number of surfaces	1	1	1	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	0
	0	1	1	1	1	0	0	0	0	0
	0	0	1	1	1	0	0	0	0	0
	0	0	0	0	1	1	1	1	1	1
	0	0	0	0	0	1	1	1	1	0
	0	0	0	0	0	0	1	1	1	1
	0	0	0	0	0	0	0	0	1	1
	0	0	0	0	0	0	0	0	1	1
	0	0	0	0	0	0	0	0	1	1

Fig. 3: Surface Tracking Matrix

似性を評価できる基準が必要である。

各面についての特徴量として、Shape context [1] を利用する。この特徴量の計算方法は次の通りである。まず、各面に所属する三次元点について主成分分析を行い、第一主軸、第二主軸から成る平面を求める。この平面に対して各三次元点を投影した後、面の輪郭部分に位置する点のみを抽出する。この処理は、距離画像中での近傍探索を行ったとき、近傍の数が一定以下であった画素のみを残すことで実現する。なお、Fig.1 右図における白点は、各面から抽出された輪郭線である。そして、この輪郭情報に対して Shape Context と同様の方法で頻度ヒストグラムを生成する。すなわち、面の中心から放射状に区切った平面領域において、輪郭点が存在する区切り (bin) には、点の数だけの投票を行い、この投票結果から、中心からの距離 r と水平線に対する bin の角度 θ を軸とする二次元ヒストグラムを生成する。なお、このヒストグラムの r 軸方向の bin 数は面の大きさに依存した値になる。この流れを Fig.2 に示す。

上記のように求めた時刻 t の各面 $s_t^j \in S_t = \{s_t^1, s_t^2, \dots, s_t^n\}$ が、次時刻 $t+1$ のセンサデータから得られた面群 $S_{t+1} = \{s_{t+1}^1, s_{t+1}^2, \dots, s_{t+1}^m\}$ に含まれているかどうかを調べる。これは、 s_t^i と類似した $s_{t+1}^j \in S_{t+1}$ を探索することに等しく、二次元ヒストグラム同士の比較を行うことで実現できる。この比較処理においては、対応づけを行う二つの面が連続する距離画像間で大きく動いていないか、ヒストグラムの r 方向の大きさが二つの面で大体同じであるか、等の条件によってフィルタリングを行うことにより、処理速度とロバスト性を向上できる。

4 可動部の全周形状モデリング

4.1 可動部に属する面の選出

面の対応付け処理を行うことで対応関係のリストが得られるが、このリストは時間方向における二つの面の関係を並べただけである。そこで、計測時刻 $t=0$ から $t=T$ までの間で、どの面がいつからいつまで観測できたかを記述した行列を新たに生成する。これを面追跡行列と称する。Fig.3 は面追跡行列の例である。横軸は時刻であり、縦軸が面のシリアル番号としている。表中の数字 1 はその面が観測されていたことを示している。この行列を参照すれば、どの面がどの期間に存在していたかを調べ、さ

らに同期間に存在していた面を探索することが容易になる．この行列とは別に，シリアル番号順にならべた面データ（中心座標，楕円体近似の結果，三次元点群）のリストを記録しておく必要がある．

可動部に属する面の選出は以下の手順で行う．まず，面追跡行列の中から面を一つ選び出し，それに相当する時系列面群を $S^{ref} = \{s_{st}^{ref}, s_{st+1}^{ref}, \dots, s_{et}^{ref}\}$ とする． st, et は s^{ref} が存在していた最初の時刻と最後の時刻である．面を選ぶときの基準は，観測期間が十分に長く，静止しておらず，ある程度の大きさを持っていることとする．たとえば，Fig.3 に指示したように，もっとも長期間で観測できていた面を選ぶ方法がある．そして， s^{ref} が可動部の一部であると想定し， s^{ref} と共に動いている面を探索する．

この探索処理は， s^{ref} が存在している期間と同様の期間もしくは一部期間に存在している面に対して行うものである．今，そのような面を s^c と記す．この s^c の属性が「可動部」か否かを定める処理は次のように行う． s^{ref} と s^c が同時に存在している期間から時刻を一つ選び， td とする．現在時刻 t における s_t^{ref} と，時刻 td における s_{td}^{ref} の間の座標変換を求める．これは，各面を楕円体近似したときの中心座標 \mathbf{x}^{ref} と楕円体の各軸方向として定義された方位 \mathbf{P}^{ref} について，時刻の違いによる姿勢変化を意味する回転行列 \mathbf{R} と平行移動ベクトル \mathbf{T} を求めることに等しい．そして，この \mathbf{R}, \mathbf{T} を用いて s_i^c の姿勢を変化させる．その結果の $s_{td}^{c'}$ と，実際の面 $s_{td}^{c'}$ が重なるのであれば， s^c は s^{ref} と同時に移動していると考えられるため， s^c と s^{ref} は同グループであるとする．

ただし，ある面 s^c が基準面 s^{ref} と同グループであるかどうかの判定を単一の s_i^c のみから判定することは，計測誤差等の影響により困難である．そこで， st, et の間で t を複数個サンプリングして，面群 S_i^c のそれぞれについて評価を行い，それらを足し合わせる．ここでの評価とは，各面が 1. 可動部，2. 背景，3. その他，のいずれに属する可能性があるかを数値化することを意味する．なお，3. は，可動部を操作している人の腕などが該当する．そして，この結果から s^c の属性を決定する．

$$\{\text{可動部, 背景, その他}\}^c = f(s_t^c, s_{td}^c, s_{td}^{c'}) \quad (2)$$

$f(\cdot)$ は，蓄積加算された3つの評価値 $C_{\text{可動部}}, C_{\text{背景}}, C_{\text{その他}}$ から，最大のものを選択してその種類を返す関数である．各評価値は以下の式で表現できる．

$$\begin{aligned} C_{\text{可動部}} &= \sum_t g_{\text{可動部}}(s_t^c, s_{td}^c, s_{td}^{c'}) \\ C_{\text{背景}} &= \sum_t g_{\text{背景}}(s_t^c, s_{td}^c, s_{td}^{c'}) \\ C_{\text{その他}} &= \sum_t g_{\text{その他}}(s_t^c, s_{td}^c, s_{td}^{c'}) \end{aligned} \quad (3)$$

ここで，面の位置の類似度関数 $h_{dist}(x, y)$ と面の方位の類似度関数 $h_{ang}(x, y)$ の出力を以下のように定義するとき，

$$D_{a-p/b-q} = \begin{cases} 1 & \text{if } h(s_p^a, s_q^b) < \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

可動部については以下ようになる．

$$g_{\text{可動部}}(s_t^c, s_{td}^c, s_{td}^{c'}) = w_1^{\text{可動部}} D_{c'-td/c-td}^{dist} + w_2^{\text{可動部}} D_{c'-td/c-td}^{ang} \quad (5)$$

背景については，上式の $c'-td/c-td$ を $c'-td/c-t$ とする． $g_{\text{その他}}$ については $g_{\text{可動部}}, g_{\text{背景}}$ のどちらからも0が得られたときに， D が定数を返すように設定する．

4.2 全周形状モデルの作成

前節に処理により，可動部に属する面が時系列データとして抽出できる．この結果から，可動部の裏表を合わせた一つの形状モデルを生成する．

まず，モデルを生成するための基準時刻 tb を定め，この時刻における面 s_{tb}^{ref} を得る．そして，距離画像列全体から，表面，裏面，側面等をそれぞれよく観測できているキーフレームを選択する．それぞれの距離画像上に存在する s_t^{ref} について， s_{tb}^{ref} との間での姿勢変化を求め， $\mathbf{R}_{t-tb}, \mathbf{T}_{t-tb}$ とする．そして，可動部と判定された面群 S_i^c について $\mathbf{R}_{t-tb}, \mathbf{T}_{t-tb}$ を用いた姿勢変換を行う．以上により，一つの空間上に可動部の面の集まり S が生成される．

ただし，面群 S は可動部形状をくまなく表せていない可能性がある．たとえば，形状の凹凸が激しく面検出が安定しなかった場所などは，三次元計測はできているものの，上記の処理では形状データが得られないことになる．そこで次の処理として，面群 S を包含するような直方体を求める．直方体の向きを可動部の傾きに合わせるため，主成分分析 (PCA) を行う．ただし，三次元点の座標を基にした PCA では，面の存在に偏りがあった場合に主軸が本来の可動部の向きから傾いて生成される恐れがある．そこで，NPCA [2] を利用する．本研究の想定では，Region growing を行った段階で，距離画像上の各点について法線成分を求めることができているため，各面に属する点に対応した法線を集めた下記行列を固有値分解することで，主軸を得る．

$$\mathbf{C} = \sum_{i=0}^p \mathbf{n}_i \mathbf{n}_i^t \quad (6)$$

ここで， p は主成分分析を行うために利用する法線の合計数であり，すべての面に属する点の数としても良いし，間引いても良い． \mathbf{n} は 3×1 の法線ベクトルである．ここで求めた主軸によって張られる空間上で，可動部に属する三次元点の最大座標・最少座標を求めれば，可動部を包括する直方体を得られる．

最後に，もとの時系列距離画像を用いて，この直方体に含まれる三次元点を求める．また，近傍の点同士をつないでメッシュを生成する．以上の手順により，密な三次元形状が得られる．

4.3 可動軸の推定

可動軸の推定は， S^{ref} の時系列変化を利用して行う．手順を下記に記す．

面の中心座標の収集

まず，時刻 st から et における各 s^{ref} から，面の中心座標のリスト $\mathbf{X} = \{\mathbf{x}_{st}, \mathbf{x}_{st+1}, \dots, \mathbf{x}_{et}\}$ を得る．

面の交線の算出

\mathbf{X} の差分ベクトル $\mathbf{V} = \{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{et-st-1}\}$ を求める．そして， \mathbf{v}_i に垂直で， $\mathbf{a}_i = \mathbf{x}_{st+i} + 0.5 * \mathbf{v}_i$ を通る平面を求め，平面群を得る．この平面群から二つの平面

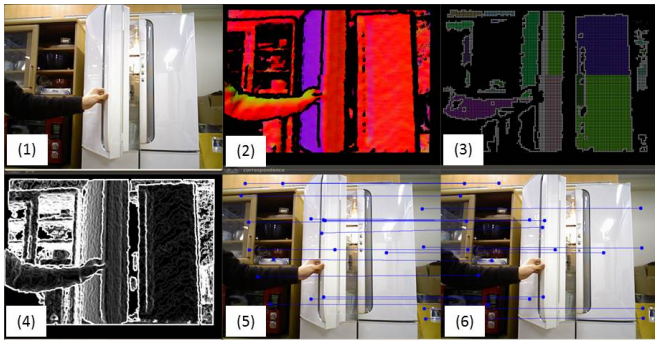


Fig. 4: Contour detection and tracking

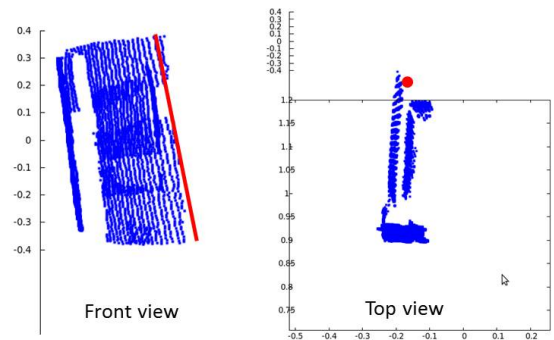


Fig. 6: Estimated rotation axis

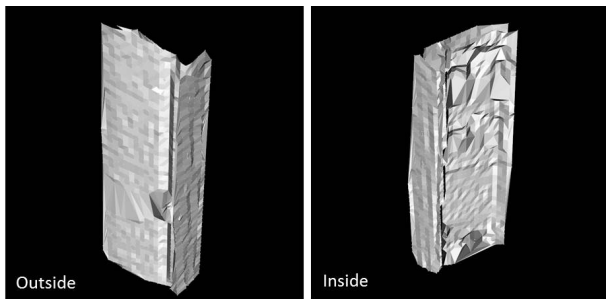


Fig. 5: A 3D model of the door of a frig

を取り出してその交線を求めることを繰り返し、直線群 $L = \{l_0, l_1, \dots, l_{et-st-1}\}$ を得る。

可動軸の推定

直線 l を順に選択し、 X との距離を計算する。もし、各点の軸からの距離がほぼ等しいのであれば、その直線は円筒の中心軸を近似したものとみなす。

より高精度の結果を得たいのであれば、上記処理の結果を初期値として、非線形最小化を行う方式を適用することができる。

5 形状モデリング実験

冷蔵庫を用いて提案手法の検証を行った。両開きの冷蔵庫を対象として、三次元距離画像センサを冷蔵庫から約 1.5 [m] ほど離れたところに設置した。人が扉を開ける様子をそのセンサで観測し、時系列の距離画像データを得た。ここから、法線成分計算 (Fig. 4, (2)), 面抽出と輪郭抽出 (Fig. 4, (3)), 面の対応付け (Fig. 4, (5), (6)) をそれぞれ行い、データを蓄積した。各フレームごとに

かかる処理は、2.67 GHz CPU で 120 から 150 [msec] 程度であった。

閉じた状態の扉を開けて内側が十分に観測できるまで、98 フレームの距離画像データを用いて、Fig. 5 に示すような三次元形状モデルを得た。また、Fig. 6 に赤線・赤点で示すように可動軸の位置と向きを得た。4 章に述べた処理はデータ取得後にバッチ処理で行い、三次元モデルと可動軸の推定に要する時間はおおむね 1 秒程度であった。

6 おわりに

本稿では、冷蔵庫のように開き戸型の可動構造を持つ家具のモデルを生成する方式について述べた。三次元距離画像センサから得た三次元点列より面を抽出し、その時間方向の連続性を利用して、表面・裏面を一体とする形状復元を行った。また、可動軸の推定も行った。

参考文献

- [1] S. Belongie, J. Malik, and J. Puzicha: "Shape Matching and Object Recognition Using Shape Contexts," IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 24 No. 24, pp. 509–521, 2002.
- [2] P. Papadakis, I. Pratikakis, S. Perantonis, T. Theoharis "Efficient 3D shape matching and retrieval using a concrete radicalized spherical projection representation," Pattern Recognition, vol. 40, no. 9, pp. 2437 – 2452, 2007.
- [3] J. Poppinga et al.: "Fast Plane Detection and Polygonalization in Noisy 3D Range Images," Int'l Conf. on Intelligent Robots and Systems, 2008.
- [4] J. Sturm, A. Jain, C. Stachniss, C. Kemp and W. Burgard: "Robustly Operating Articulated Objects based on Experience," in Proc. of Int'l Conf. on Intelligent Robots and Systems, 2010.
- [5] K. Yamazaki, T. Tsubouchi and M. Tomono: "Furniture Model Creation through direct teaching to a Mobile Robot," Journal of Robotics and Mechatronics, vol.20, no.2, pp.213-220, 2008.