

変形予測に基づくモデル駆動型衣類状態推定 — 距離画像利用の検討 —

松川 徹[†] 喜多 泰代[‡]

[†] 筑波大学大学院システム情報工学研究科 〒305-0006 茨城県つくば市天王台 1-1-1

[‡] 産業技術総合研究所 〒305-8568 茨城県つくば市梅園 1-1-1 つくば中央第2

E-mail: [†] t.matsukawa@aist.go.jp, [‡] y.kita@aist.go.jp

あらまし 対象衣類の変形を予測しながら変形予測と観測画像とを照合することで、衣類の状態をモデル駆動型に推定する手法の研究を進めている。これまでに変形予測と観測画像との2次元画像照合を基本とする手法を用いてアプローチの有望性は示せた[3]が、実用のための頑健性が不十分であり、詳細な3次元情報を入力に用いることにより、より高度な状態推定手法の検討を開始している。この目的のために、3眼ステレオビジョン[4]とCGソフトウェア Maya[5]を用いたシステムを構築中である。本研究報告では、距離画像を用いた状態推定法に関する予備実験を行った結果を報告する。

キーワード 物体認識, 柔軟物体, 距離画像, 衣類シミュレーション, ロボットビジョン

A Deformable Model Driven Method For Handling Clothes — A Consideration Of Using Range Image —

Tetsu MATSUKAWA[†] Yasuyo KITA[‡]

[†] Graduate School of Systems and Information Engineering, University of Tsukuba,

Tennoudai 1-1-1 Tsukuba, 305-8573 Japan

[‡] National Institute of Advanced Industrial Science and Technology(AIST)

AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, 305-8568 Japan

E-mail: [†] t.matsukawa@aist.go.jp, [‡] y.kita@aist.go.jp

Abstract For the purpose of handling clothes by a robot, we are studying a deformable model driven method which estimates cloth state by matching the predicted shape of the target clothes with its observed shape. Though our previous work have shown good potential of this strategy[3], it did not have enough generality and robustness for practical applications. To improve this point, we start considering 3D matching between 3D predicted shape and range data obtained by a trinocular stereo vision. In this paper, we report our system using a real-time trinocular stereo vision [4] and a cloth simulation software Maya[5] and preliminary experimental results using this system.

Keyword Object Recognition, Deformable Object, Range Image, Cloth Simulation, Robot Vision

1. はじめに

家庭内では、衣服や紙等の変形を生じる物体が溢れており、これらの物体をロボットが的確に扱うためには柔軟物体の形状を認識できる視覚機能が必要となる。

最近では、家庭用ロボットでの利用を想定した衣類のハンドリングに関する研究がいくつか報告されている。金子らは対象輪郭の幾何学的特徴量に基づいた展開動作計画を立てている[1]が、日常環境において物体輪郭を厳密に抽出することは困難であると思われる。大澤らは、洗濯・乾燥後の数種類の衣類を対象として、ロボットによる展開と分類を同時に行い、作業台上に

乗せ、道具を利用して折りたたむシステムを開発している[2]が、展開開始状態として数回の持ち替え操作により少数の形状に収束させた状態を用いるため、余分な持ち替え動作が必要となっている。さらに、これらの手法では個別の衣類に対し人手により作業計画を立てる必要がある、ロボットが日常環境で様々な衣類を扱うことを考えると、より汎用性の高い状態推定方法が望まれる。

一方我々は、衣類のような自在に変形する物体を観測データのみからボトムアップに認識することは困難であると考え、対象に対する大まかな知識をもとにシ

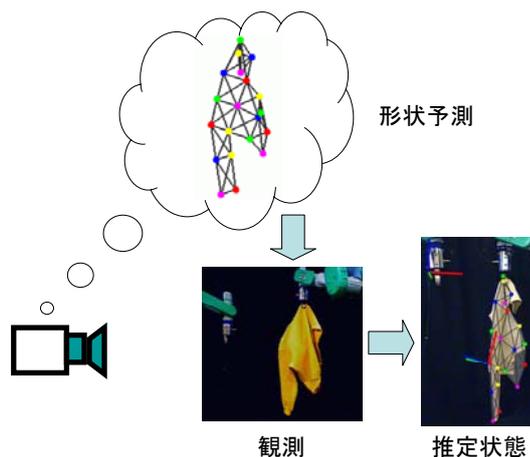


図1 モデル駆動型状態推定の概念図

ーンに対する予測を行い、予測と観測画像の情報を照合することで状態を推定するモデル駆動型のアプローチ(図1)を提案している。これまでに予測モデルとの2次元画像照合を基本とする手法で状態推定を行いこのアプローチの有望性を示した[3]。ステレオカメラをハンドの保持位置の算出等に部分的に使用したが、対象の詳細な3次元情報は用いていなかった。

最近では、複数台のカメラによりシーンの密な3次元情報である距離画像を実時間で生成する技術が確立されて来ている[4]。距離画像を入力に導入することで、3次元空間の一定範囲内の物体を抽出することで背景からの衣類領域抽出を容易に行え、詳細な3次元情報を有効に用いることで状態推定に頑健性を持たせることが出来ると考えられる。

本研究報告では、距離画像を用いた状態推定法を検討するために行った予備実験を報告する。

2. 課題設定

本報告では、双腕マニピュレータとステレオカメラからなるハンドアイシステムにより衣類を折り畳むタスクにおけるサブゴールとして、衣類の両肩をハンドで保持し展開した状態にすることを課題としている。一方のハンドで保持した状態の衣類の特定の部位をもう一方のハンドで掴み持ち上げるという動作を基本として、この基本動作の繰り返しによりこの課題を実現する。本報告では基本動作の内、ある1点で保持された未知の衣類の状態を代表形状と照合することにより推定する手法を検討する。この状態を推定すれば、[3]で提案した方法により第2ハンドの保持目標とする座標情報を算出し、ハンドリングが行える。

3. システム構成と対象衣類

入力装置に用いるリアルタイム3眼ステレオビジョンシステム[4]は、IEEE1394インターフェースを備え複数カメラの自動同期機能を持つ Point Grey Research Inc. の Dragonfly 3 台 (基準/右/上) のカメラを入力

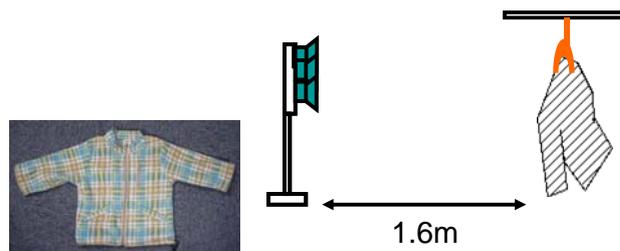


図2 対象衣類 (左) と実験環境 (右)

とし、マルチベースラインステレオと双方向視差探索を組み合わせたブロックマッチングに基づく320×240程度の距離画像を汎用PCワークステーション(Intel XEON 3.8GHz ×2)上に1秒間に30枚生成可能である。現時点では静止した状況の衣類の状態推定を対象とし、その内の1フレームのみを用いる。

衣類形状の予測には、衣服の変形シミュレーションが行えるCG(ComputerGraphics)ソフトウェアの1つであるAlias|Wavefront(現Autodesk)のMaya4.5[8]のCloth機能を用いる。

今回の実験では、入力カメラの画角の関係上、子供用トレーナーを対象としている。本報告では、マニピュレータを用いる前の予備実験として、洗濯ばさみにより保持し吊るし上げた状態で実験を行う(図2)。

4. 状態推定法

4.1 起こりえる形状の予測

対象としたトレーナーの形状の情報に基づき図3aに示す前身頃と後ろ身頃の離れることのないことを仮定した平面のモデルをMaya Clothで生成する。このモデルは厳密な衣類の情報を与えなくとも使えるように同一種類の衣類全てに対応できる汎用モデルを目指している。本報告では、厳密ではないおおまかな形状や柔らかさ等、ある程度の衣類の情報はわかっている場合を想定し、モデルの形状、ノード点数、曲げに対する衣類の抵抗、伸縮のパラメータを手動で設定している。具体的に衣類の袖部の幅を10cm、長さを24cm、身頃部の幅を30cm、高さを35cmとして各種サイズの三角形が不規則に配置されたメッシュモデルを作成する。1メートルのクロス内に含まれる三角形の数を表すBase Resolutionを25に設定する。衣類の質量、厚さ、摩擦係数、空気抵抗等、多くのパラメータはデフォルトの状態にしておき、柔らかさに関するパラメータのみを実際の衣類に合わせ、次のように調節する。モデルのサーフェイスを(U,V)空間でパラメータ表現[5]した時のU/V各方向への曲げに対する抵抗であるU/V Bend Resistanceを30にメッシュの3三角形が0度から180度へ折り返される時の非線形性を表すU/V Bend Rateを20にU/V各方向へのストレッチ動作に対する抵抗を表すU/V Stretch Resistanceを50に各々

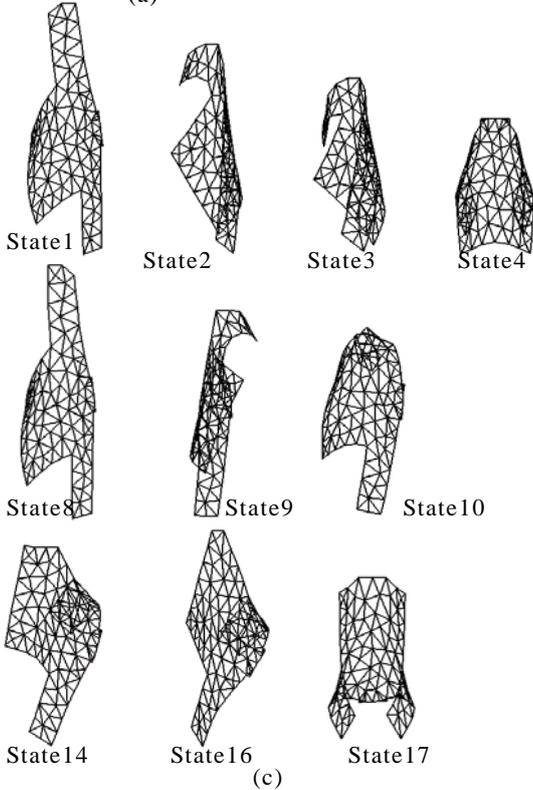
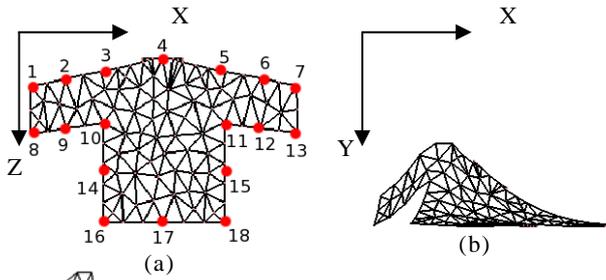


図3 モデル形状 (a)共通モデル; (b) 予測過程; (c) 予測されたモデル

設定する。

このモデルを用いて、トレーナーが図3(a)の赤い点で示す縁からほぼ等間隔に選んだ全18個の代表点の内、ある1つの代表点保持され吊るされた状況をシミュレーションする。我々が想定している保持は小面積を保有する平らなグリップによるものであり、画像が常にグリップ面から垂直に撮影される状況を想定している。ここで、Y軸を重力方向に一致させた左手系3次元座標系を基準座標系とする。まず、対象モデルをXZ平面に水平に図3aの状態に配置する。小面積を有するグリップでの保持の効果を代表点に近い縁の2点のノードを固定することで再現し、他のノードに対し重力を平衡状態に達するまで働かせる。図3bはその過程を表している。これを全18個の代表点に対し行い、収束した18種類の予測形状モデル(以下モデル)に縁の近い代表点で掴み吊るされた全ての状態を代表させる(図3c)。予測形状モデルの3次元座標とハンドの位置をmayaのAPIを利用し作成したMEL(Maya

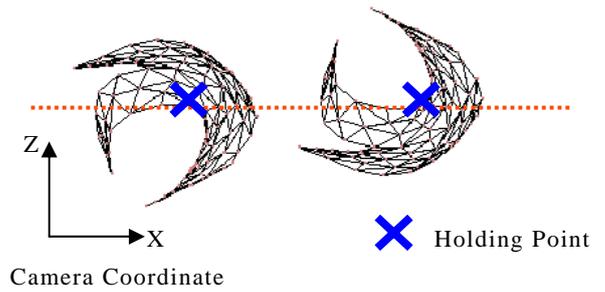


図4 保持位置を通るx軸に平行な線に対象な候補

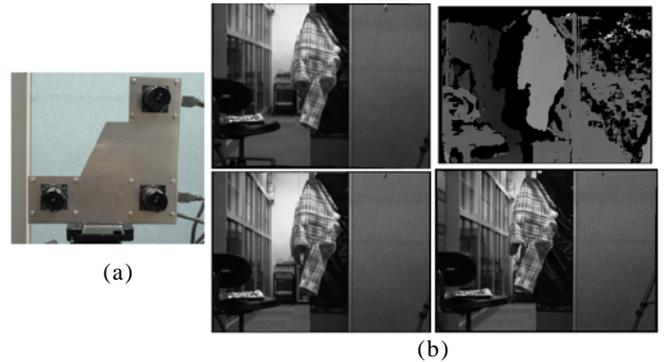


図5(a) 3眼カメラ (b) 濃淡画像(基準画像(左下), 右画像(右下), 上画像(左上)), 距離画像(右上)。

Embedded Language) コマンドにより、予測形状データベースに保存する。

4.2 観測画像と予測との照合

4.2.1 カメラ座標系

カメラ座標系は3台のカメラの内、基準カメラを原点に取り観測方向にz軸の正、重力方向をy軸の正、観測方向の右側をx軸の正とした直交座標系とする。距離画像には欠損の生じていない画素(u,v)に対し、対応するカメラ座標系での3次元座標値(x(u,v), y(u,v), z(u,v))が格納されている。以下の座標の記述はカメラ座標系での記述である。

4.2.2 衣類領域抽出

ハンドリングを行う際、ハンドの近傍の空間に対象物体以外が存在しない状態であることは自然な設定であると考えられる。この仮定により、保持点以下一定範囲内の物体を抽出することで衣類の色やテクスチャによらず衣類領域を抽出できる。保持ハンドの3次元位置情報(hand.x, hand.y, hand.z)は既知であるので、 $hand.y \leq y(u,v)$ でありかつ $|hand.z - z(u,v)| \leq 閾値 A$ を満たす画素(u,v)平面上の連続領域を衣類領域として入力距離画像より抽出する。衣類の広がりを見越して閾値Aを300mmとした。

4.2.3 全長に基づく候補選定

距離画像より衣類領域を抽出したら、処理時間短縮のため全長情報に基づいて候補を選定する。衣類領域

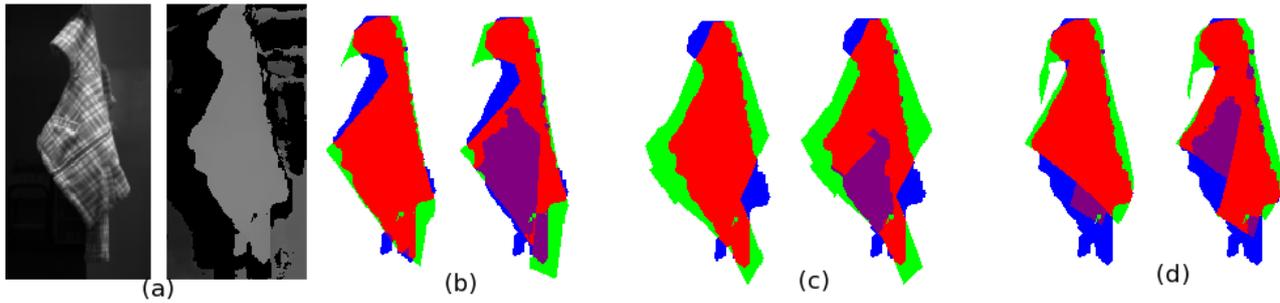


図 6 照合結果例 1 (a)観測画像(State2 に相当, 濃淡画像(左), 距離画像(右)), (b)(c)(d) モデル距離画像との重なり (b) State2(ROR=1.73,0.86),(c)State18(ROR=1.57,1.15),(d)State3(ROR=1.52,1.29).

表 1 照合結果 (代表点保持)

入力データ代表点.	1	2	3	4	8	9	10	14	16	17	合計
成功数	5/5	3/5	0/5	5/5	4/5	2/5	5/5	5/5	5/5	5/5	39/50

の最下点を (lowest.x, lowest.y, lowest.z) と表し, 全長を $|hand.y - lowest.y|$ と定義する。この全長のある一定の誤差の範囲内に正しい状態が含まれている。そこで, $|データの全長 - モデルの全長| / データの全長 \leq 閾値 B$ であるモデルを候補として残す。正しい状態のモデルが確実に残るように閾値 B を 0.2 とした。

4.2.3 保持位置を通る x 軸に平行な線に対象な候補

代表点を保持した場合に対し, 現実には様々な状態が存在するものを 1 つの予測形状に代表させてモデルを作成したが, この予測に対しハンド保持部分を通る x 軸に平行な線に対象な形状 (以下, 前後対象形と表記) があり得る (図 4)。ハンドリングを行う場合, 第 2 ハンドが掴むべき部位の座標やハンドの進入方向はこの 2 種類で大きく異なってくる [3]。そこで, 本研究報告ではこの 2 種類の形状を区別して考える。

4.2.4 モデル距離画像の生成

観測距離画像とモデルを照合するためにモデルを観測距離画像と同一形式のものにしておく。ハンドの保持位置 (hand.x, hand.y, hand.z) にモデル上の保持位置に対応する部分を合わせて配置し, グリップ面から垂直な方向から観測したモデルの見え方距離画像をカメラキャリブレーションにより既知である基準カメラの射影行列を用いて OpenGL のデブスバッファ等を利用し, 生成する。

4.2.5 距離画像照合

各画像 (u,v) において観測距離画像上での衣類領域の z 座標 (Dataz(u,v)) とモデル距離画像上の z 座標 (Modelz(u,v)) の双方の値が存在し, $|Dataz(u,v) - Modelz(u,v)| \leq 閾値 C$ を満たす画素を重なり点とする。重なり点の点数を数え, 次式で定義する距離オーバーラップ比 ROR (Range Overlap Ratio) を評価値とする。

$$ROR = \frac{\text{重なり点数}}{\text{観測距離画像点数} + \text{重なり点数}} / \frac{\text{重なり点数}}{\text{モデル距離画像点数}}$$

ROR (0~2.0) は 2 次元画像の重なり距離による制限を加えたものであり, 数値が大きくなるに従い高い重なり度合いを示している。閾値 C の設定によって照合性能に違いが生じるが, 実験による経験値により閾値 C を 100mm と設定した。

5 の実験では, 18 種類のモデルからその前後対象形を考慮した 36 種類のモデルの内, 4.2.2 の処理で残ったモデルに対し, モデル距離画像の生成, 観測距離画像との照合を行い, ROR が最大のモデルを照合結果として選定する。

5. 実験

観測データの再現性を考察するために一度衣類を振った後に, 4.1 で示した代表点の内, 左右対称な場所を除く 10 個の代表点 (1,2,3,4,8,9,10,14,16,17) に近い位置で対象衣類を保持した状態の距離画像を各々 5 枚撮影し, 合計 50 枚の距離画像を入力とした。図 5b に図 5a の 3 眼ステレオに入力する濃淡画像, 及び生成される距離画像の例を示す。距離画像は, z が小さくなるに従い白く表示し, 欠損画素は黒で表示している。

今回の実験では, ハンドの 3 次元保持位置 (hand.x, hand.y, hand.z) は手動で与えた。全ての画像で, 複雑な背景下, テクスチャの衣類を保持したシーンであるが, 衣類の領域抽出が行えた。入力装置の特性により若干の膨張領域や欠損領域が見られたが本研究目的としては十分な精度であると考えられる。全長に基づく候補選定により照合する 36 個のモデルを平均 19.5 個に減らすことが出来た。代表点 4 を保持したデータにおいては 2 個にまで減らすことが出来ていた。

図 6 に結果の一例 (代表点 2 で保持) を示す。図 6a は基準画像の濃淡画像 (左) と距離画像 (右) を示し,

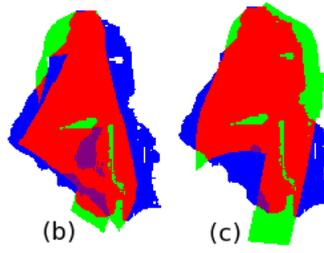
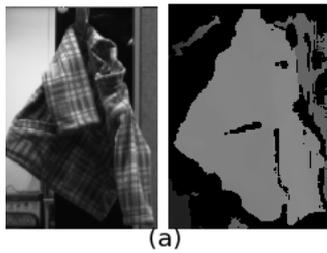


図 7 State3 の例 (a)観測画像 (濃淡画像 (左), 距離画像 (右)), (b)(c)モデル距離画像との重なり:
(b)State3(ROR=1.40), (c)State10(ROR=1.53).

図 8 State9 の例

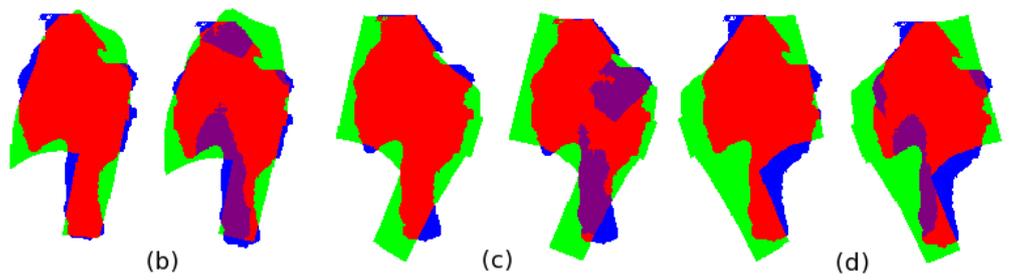
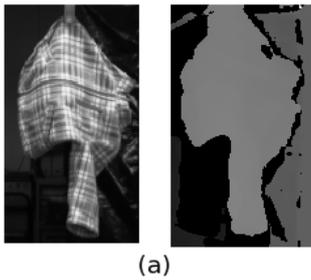


図 9 照合結果例 2 (a)観測画像 (State10-14 の中間, 濃淡画像 (左), 距離画像 (右)), (b)(c)(d) モデル距離画像との重なり:(b) State10(ROR=1.74,1.22), (c)State14(ROR=1.67,1.45), (d)State15(ROR=1.54,1.26).

表 2 照合結果 (代表点の中間地点保持)

入力データ代表点	1-2 間	2-3 間	3-4 間	1-8 間	8-9 間	9-10 間	10-14 間	14-16 間	16-17 間	合計
成功数	3/5	3/5	4/5	5/5	0/5	0/5	5/5	3/5	3/5	26/45

図 6bcd は距離画像より抽出した衣類領域とモデル距離画像との重なりを ROR が大きい順に前後対象形と共に示しており、赤で示す部分が重なり領域、青がデータのみ存在する領域、緑がモデルのみ存在する領域、紫がモデルとデータの双方が存在するが、双方の z の絶対差が閾値 C よりも大きいため重なりとしない領域を示す。実際の保持代表点(State2)のモデルとの ROR が最も大きくなっており (図 6b), 同一保持位置(State2 内)においても前後対象形を正しく区別することが出来た。

表 1 に代表点に近い位置で保持した場合における全照合結果を示す。50 枚中 38 枚で正しい状態が選定された。誤選定は、代表点 2,3,9 を保持した場合に集中していた。代表点 2,3 を保持した場合においては肩の折れの度合いが予測と大きく異なり、幅が変わっていることが原因であった。この例 (State3) を図 7 に示す。望ましいモデル図 7b(State3)では観測データより幅が狭くなり、異なるモデルである図 7c のモデル (State10)との重なりより小さくなっている。このような場合には、2次元画像に対してではあるが文献[3]で

行っていたような幅正規化等の処理を考えることよりある程度この違いを吸収できると考えられる。

代表点 9 を保持した場合では袖の下側の中心を保持していることにより、もう一方の袖の位置関係によって回転が引き起こされていることによりモデルが観測画像と大きく異なる形状をしていることが誤選定の原因であった (図 8)。今回の実験で代表点 9 に関しては 1 通りの状態ではなく、この回転により 2 通りの状態が現れる確率が高いことが確認された。そのため、代表点 9 に限ってはこの保持位置に対し 1 通りの予測では不十分であり、この回転を考慮した 2 通りの予測を行う必要があると考えられる。

次に、一度衣類を振った後各代表点の中間地点 (1-2,2-3,3-4,1-8,8-9,9-10,10-14,14-16,16-17 間) で保持した入力画像を各 5 枚、合計 45 枚用いて、同様に実験を行った。全ての画像で、衣類の領域抽出が行えた。全長に基づく候補選定により照合する 36 個のモデルを平均 23.4 個に減らすことが出来た。

図 9 に結果の一例(代表点 10-14 の中間を保持)を示す。図 9a は観測画像を示し、図 9bcd は距離画像より

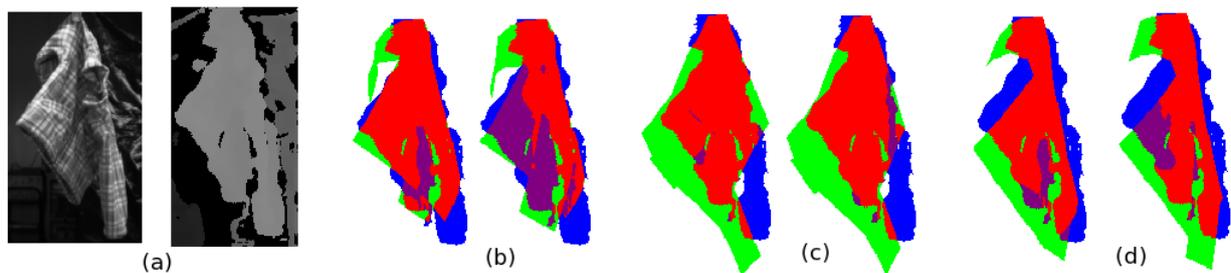


図 10 照合結果例 3 (a)観測画像 (Stae3-2 間, 濃淡画像 (左), 距離画像 (右)), (b)(c)(d) モデル距離画像との重なり: (b) State3(ROR=1.50,0.88) (c)State18(ROR=1.43,1.33), (d)State2(ROR=1.26,1.16).

抽出した衣類領域とモデル距離画像との重なりを ROR が大きい順に前後対象形と共に示す。実際の保持位置と近い代表点の 10, 14 のモデル両方で ROR が大きくなっている。

一方, 2つの中間地点を保持した場合の一方の代表点のみに形状が近くなる場合が見られた。その例を図 10 に示す。図 10a は代表点 2-3 間を保持した衣類の観測データである。図 10b のように代表点 3 を保持したモデルとの ROR は大きい, 2 番目に大きいモデルは図 10c で示す代表点 18 を保持した場合を代表するモデルとなっており, 図 10d に示す代表点 3 を保持したモデルは 6 番目に ROR が大きいモデルとなっていた。これは, ハンドの保持位置で観測距離画像とモデル距離画像を合わせているために代表点から離れた場所を保持したデータに対しては重ねるべき領域が垂直方向にずれているためと考えられる (図 10d 参照)。このように代表点以外を保持した場合に対応するためには, 文献[3]で行っていたような重ね合わせ位置を垂直方向に修正する等の対応策が必要と考えられる。

表 2 に代表点の中間地点を保持した場合の全照合結果を示す。ここでは, 最も ROR が大きいモデルが観測データで中間を保持している 2つの代表点の内の 1つであれば成功としている。45 枚中 26 枚で正しい状態が選定された。誤選定の状態は, 代表点の中間地点 (8-9,9-10 間) を保持した場合に集中していた。この要因は, 図 8 で示したように 9 の代表点には 2 種類の状態が考えられることにより 9 の代表形状とは大きく異なった観測データが存在することと, 図 10d のように重ねるべき領域が垂直方向にずれていることであることが確認されている。

7. まとめ

[4]で提案した衣類の状態をモデル駆動型に推定する手法に, 実用のための頑健性を持たせるため, 入力に詳細な 3次元情報を用いることを検討中である。本報告では距離画像入力化に関する検討実験を行い次の結果を得た。

(1)衣類近傍に他の物体が存在しないことを仮定して,

背景の複雑さや衣類の色やテクスチャによらず 3次元空間内の一定範囲を抽出することで衣類領域抽出が行えた。

(2)距離画像照合法として距離オーバーラップ比を定義し, 望ましい予測モデルが観測画像と形状の近い場合においては, これを用いて各代表保持位置のモデルから望ましいモデル及び前後対象形の選定が行えた。

今後の課題としては, 観測データが予測モデルの形状と大きく異なる場合に対応することや代表点から離れた位置で保持した場合において生じるモデルとデータの重ね合わせ位置のずれを補正していくことがあげられる。

また, 実際にロボットハンドを用いて実験を行い, 本状態推定法の有効性検証を行っていく。

謝辞

ステレオビジョンシステムを提供頂いた植芝俊夫主任研究員をはじめ, 産総研インタラクティブビジョングループの人々に感謝する。

文 献

- [1] M.Kaneko and M.Kakikura: Planning strategy for putting away laundry - Isolating and unfolding task -, Proc. Of the 4th IEEE International Symposium on Assembly and task Planning, pp.429-434, 2001.
- [2] F.Osawa, H.Seki, Y.Kamiya, "Clothes Folding Task by Tool-Using Robot", Journal of Robotics and Mechatronics, Vol.18, No.5, pp618-625, 2006.
- [3] Y.Kita, F.Saito, and N.Kita : "A deformable model driven visual method for handling clothes", Proceeding of the 2004 IEEE International Conference on Robotics & Automation, pp.3889-3895, 2004.
- [4] T.Ueshiba: "An Efficient Implementation Technique of Bidirectional Matching for Real-time Trinocular Stereo Vision", Proc. 18th Int'l Conf. on Pattern Recognition, vol.1, pp.1076-1079, 2006.
- [5] David A.D.GOULD, 中村達也訳: COMPLETE MAYA PROGRAMMING(日本語版), ボーンデジタル, 2004.