A deformable model driven method for handling clothes

Yasuyo Kita Fuminori Saito Nobuyuki Kita
Intelligent Systems Institute,
National Institute of Advanced Industrial Science and Technology (AIST)
AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Japan 305-8568
{y.kita, fumi.saito, n.kita}@aist.go.jp

Abstract

A model-driven method for handling clothes by two manipulators based on observation with stereo cameras is proposed. The task considered in this paper is to hold up a specific part of clothes (e.g. one shoulder of a pullover) by the second manipulator, when the clothes is held in the air by the first manipulator. First, the method calculates possible 3D shapes of the hanging clothes by simulating the clothes deformation. The 3D shape whose appearance gives the best fit with the observed appearance is selected as estimation of the current state. Then, based on the estimated shape, the 3D position and normal direction of the part where the second manipulator should hold are calculated. The experiments using actual two manipulators have shown the good potential of the proposed method.

1. Introduction

The handling of soft objects is attracting increasing attention in the robotics field[1][2][3][4][5]. Kaneko et. al[4] proposed a method which recognizes the clothes state by comparing the contour features (e. g. curvature, length-ratio) of an observed appearance with ones of model appearances under the situation that the clothes is hanging at the two points. However, the contour features are difficult to robustly extract from real observations and are very sensitive to a slight deformation of clothes. In [5], we proposed a method which recognizes the state of clothes hanging at a point in a model driven way. The method predicts the possible appearances by the simulation using a deformable model of the clothes and selects one which fits the observed appearance the best. The results of the preliminary experiments using a pullover held by a human hand were encouraging. Although the method succeeded in indicating the position of the point where the second manipulator should hold next on observed images, the information is not enough for actual handling. For actual handling, how to approach to the holding position and in which direction the manipulator gripper should close should be determined three-dimensionally.

In this paper, we propose a method for extracting the three-dimensional information required for manipulators' actions based on the estimation of the clothes state. The estimation processes take the same strategy as proposed in [5] and is improved to take consideration of appearance variations caused by the difference in the clothes softness. We show the model-driven strategy also has advantages in the stage of acquisition of the 3D information necessary for manipulator actions.

2. Prediction of possible 3D shapes

We suppose a system consisting of two manipulators and two cameras which are calibrated relative to the manipulators. In this paper, we consider a task to hold up a specific part of clothes (e.g. one shoulder of a pullover) by the second manipulator, under the situation that the clothes is held in the air at any point close to its hem by one manipulator. We assume that we know style of the target clothes (e.g. pullover, trousers) and its approximate sizes in advance.

Fig. 1a shows a common pullover model. the model consists of 20 nodes which are connected to each other by springs as illustrated with the lines. An individual deformable model for the target clothes is automatically built on its approximate sizes, like the width and length of the trunk and the length of the sleeves for a pullover. Possible deformed shapes when the pullover is held in the air at a point close to the hem is predicted through the simulation as shown in Fig. 1b [5]. Fig. 1c shows examples of the predicted 3D shapes, from State 1 to State 20, named after the number of the node held. State 3, 6, 9, 12, 17, 18, 19 and 20 are left-right symmetrical shapes of State 1, 4, 7, 10, 13, 14, 15 and 16, respectively. Here, by assuming the pullover is always observed from the direction perpendicular to the gripper plane, the viewing direction is fixed so that it coincides with the normal of the plane around the holding point.

Actually, the possible 3D shapes largely change depending on the softness of the clothes. Fig. 2 shows examples of shape change of clothes when its softness changes, that were artificially obtained by reinforcing the clothes with packing tapes attached to its back. Fig. 2a corresponds to the case when the clothes is thick and hard. It hardly folds and is close to a plane. Most of all other variations by drapes can be thought as distortion of the plane. When the clothes becomes softer, the number of drapes increase more as shown

0-7695-2128-2/04 $20.00 (C) 2004 IEEE

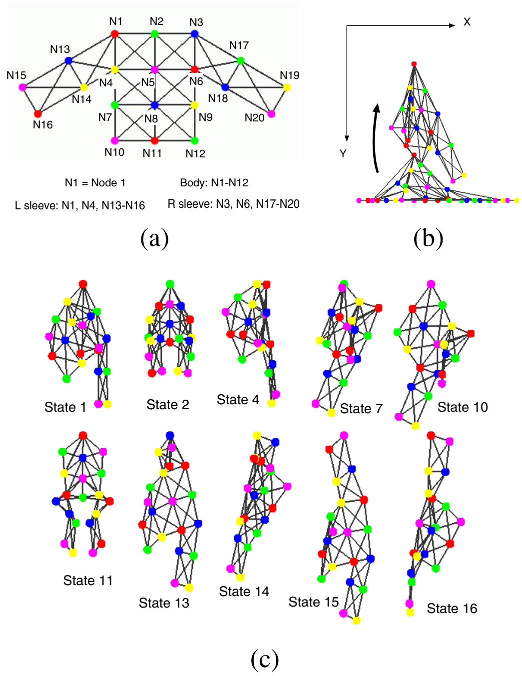


Figure 1. Model shapes: (a) common pullover model; (b) simulation processes; (c) possible 3D shapes when the pullover is held at a point

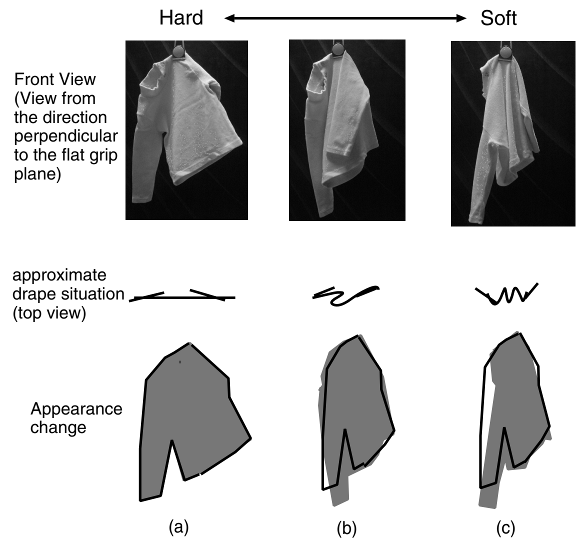


Figure 2. Appearance variation depending on softness of clothes

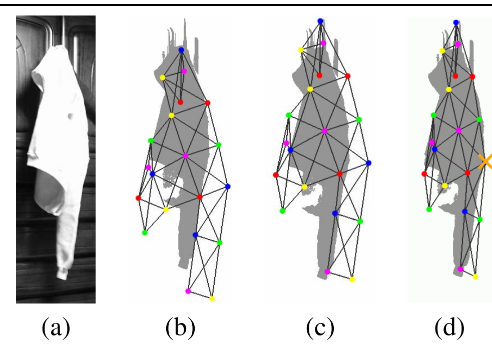


Figure 3. model adjustment: (a) observed image; (b) original position (gray region shows observed clothes region); (c) after vertical translation; (d) after width normalization.

in Fig. 2b and c. At the same time, the appearances from the front direction become narrower, however, they do not drastically change. For example, the observed appearances, illustrated with the gray regions in Fig. 2, are close to a horizontally shrunk shape of the appearance of the hardest case, illustrated with the black contours.

By taking these points into the account, we simulate clothes deformation assuming relatively hard clothes, which gives a characteristic view like Fig.2a, and use the deformed shapes as possible representative shapes. At the stage of estimating the state of the observed clothes, the model appearances corresponding to the representative shapes are used for all the softness by adjusting them based on the observed width.

3. Estimation of clothes states

At the current, we assume the clothes regions in observed images can be extracted in a bottom-up way. The gray region in Fig. 3b shows an example in the case of the observed image in Fig. 3a.

First, each possible 3D shape as shown in Fig. 1c is virtually set at the current holding position which is given from the manipulator module. Then, the appearances in the left and right images of the shape are calculated based on each camera projection matrix. In the example of Fig. 3b, the appearance at State 13 is overlapped on the observed clothes region. The following processes are applied to the left and right images respectively using these predicted appearances as model appearances. Only the process ii), which is largely different from ones proposed in [5] is described in detail.

i) By checking the appearance height, model appearances which are clearly different from the observed appearance are excluded from the candidates.

ii) Each model appearance left in the candidates is adjusted to the observed region as follows:

1. Vertical translation
   The observed clothes may be held at a point far from any 20 nodes, which are used as holding positions to predict representative model shapes. To remove the misalignment caused by this difference, model appearances having longer vertical length than the observed

one are vertically translated up on the image by the difference in their lengths. Fig. 3c shows the model appearance after this translation.

2. Width normalization
   Based on the considerations in Section 2, each model appearance is horizontally shrunk or extended so as to have the same width as the observed region. Fig. 3d shows a result after this normalization.

At the current, these adjustments on a model appearance are not feedbacked to the position and shape of the original 3D shape producing the appearance.

iii) After these adjustments, the overlap ratio, R, which is the sum of the ratio of overlapped area to model appearance area and the ratio of overlapped area to observed area, is calculated. The model state which has the highest value of R is selected as the estimation result.

iv) The image coordinates of the target point, which is the point to hold next, are detected by searching the edge point closest to the model node corresponding to the point. The orange cross in Fig. 3d shows the detected position as a shoulder in this example. During this process, consistency of the contour straightness around the target point between model appearance and observed clothes region, is checked. If inconsistency is detected, the model appearance is rejected and then, the model appearance which has next highest overlap ratio is tried.

4. Calculation of grip coordinates for grasping

If the clothes states are consistently estimated in the left and right images, next, the grip coordinates for grasping the indicated point are calculated according to the estimated model shape. Fig. 4a shows the gripper coordinates of our system. To grasp some part of the clothes, the manipulator should approach along the minus Z direction to the part with the Y axis set to coincide with the normal of the part. When the origin of the coordinates reached to the position to grasp, the gripper is closed. The grip coordinates satisfying these conditions are calculated as follows:

I. Calculation of the 3D coordinates of the indicated point
Based on the image coordinates of the left image indicated in process iv), the edge point in the right image is searched along the epipolar line. Then the 3D coordinates of the corresponding point are calculated based on the stereo principle.

II. Decision of model shape from symmetrical candidates
If the camera projections can be assumed to be weak-perspective, there are two possible 3D shapes for one appearance, which are symmetrical with respect to the vertical plane perpendicular to the view direction and goes through the current holding point. Fig. 4b is a top view illustrating the two symmetrical model shapes: the blue line

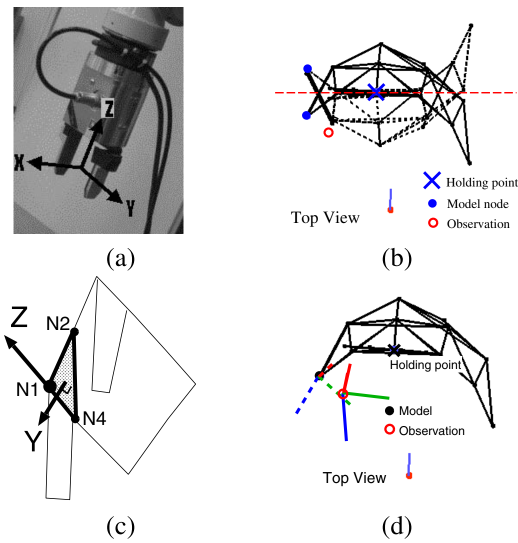


Figure 4. Determination of grip coordinates: (a) gripper coordinates of our manipulator; (b) decision of model shape from symmetrical candidates; (c) calculation of the coordinates based on model shape; (d) correction based on the observation.

sticking out from the red dot at the lower part of the figure illustrates the view direction; the cross shows the current holding position. After obtaining the actual 3D position of the target point, the model shape which has closer model node to the observation is selected. In this example, the red circle shows the obtained 3D position corresponding to the model node shown by blue dots, and the shape shown by the solid lines in Fig. 4b is selected.

III. Calculation of grip coordinates based on model shape
In order to determine the direction of the grip coordinates, the posture of the target part should be known. However, the posture is not easy to obtain from the observation. Therefore, instead of acquisition from the observation, we approximate it with the posture of the target part in the estimated model shape. Fig. 4c shows an example when the point to hold is one shoulder of a pullover. The normal direction of the adjacent triangle patch, that is the triangle surrounded by Node 1, 2 and 4 in this case, is used for the direction of the Y axis. The direction of the Z axis, which is approaching direction, is set to coincide with the line connecting Node 1 and 4. The X axis is determined to complete the right-handed coordinate system. The red, green and blue dashed lines in 4d respectively show the X, Y and Z directions obtained in this example.

IV. Correction of grip coordinates based on observation
If any good clues indicating the actual posture of the tar-

Table 1	Estimation results		
	Success	failure	total
Pullover	22	3	25
Trousers	11	3	14

get part is found in the observed images, the coordinates are corrected based on the information. Suppose that the shoulder line is three-dimensionally obtained from the observed image as shown with the red solid line in Fig. 4d. Then, the estimated grip coordinates are rotated to fit the observation as shown with the red, green and blue solid lines in Fig. 4d.

After the direction of the grip coordinates is fixed, its origin is set a few centimeters inside along the Z direction from the 3D position of the target point.

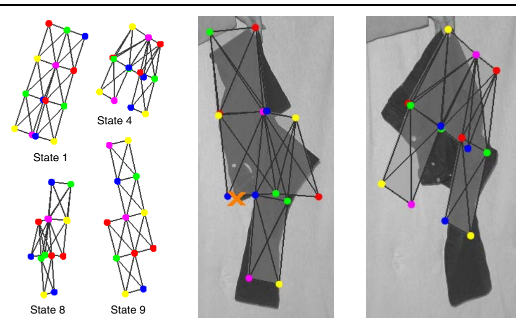
5. Experimental results

5.1. Experiments on estimation

First, for the purpose of examining the ratio of correct estimation, we have conducted experiments using single images of a pullover and trousers which were held in the air at a point by a human hand. The estimation results are summed up in Table 1. Regarding the pullover, we used the same input images as in [5]. The improvement of the estimation processes raised the success rate than before. As a common trousers model, we use a deformable model consisting of 14 nodes. Fig. 5a shows examples of the predicted shapes. Fig. 5b and Fig. 5c show examples of success and failure respectively.

5.2. Experiments using manipulators

We also conducted experiments using an actual system consisting of stereo cameras and two manipulators (Mitsubishi PA-10) as shown in Fig.6.



(a) (b) (c)

Figure 5. Estimation results on trousers: (a) examples of model appearances; (b) examples of success ; (c) example of failure.

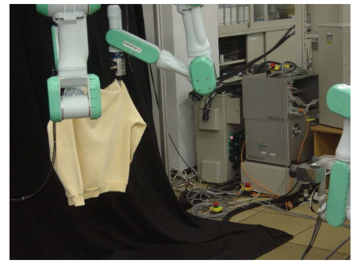


Figure 6. System consisting of two manipulators and stereo cameras

Fig. 7 shows experimental results using a toddler pullover which is made of rather soft material. Fig. 7a,b show observed left and right images. Fig. 7c,d show the estimation results of the both images, which were State 17. In this example, the model appearances were largely shrunk horizontally to fit the observed appearances as shown in Fig. 7c, since the actual 3D shape was folded more than the model shape. Because of the large folding, the 3D position corresponding to the shoulder node was far from the position of the model shape as shown with the red circle in Fig. 7e. In this case, since no clues on the normal of the target part was obtained from the observed images, the direction of the grip coordinates determined from the model shape was used as it was without any correction (red, green and blue solid lines in Fig. 7e,f). From manual measurement, the actual normal direction of the shoulder part was different from that of the estimated model shape by about 20 degrees. As a result, the approaching direction deviated from the best direction by about 20 degrees. Despite of that, the second manipulator successfully grasped the shoulder part since the clothes change shape along the gripper responding to its contact. Fig. 7h shows the newly observed image after the action with the predicted appearance, which was obtained through the simulation shown in Fig. 7g. The method can judge if the current estimation of the state is correct or not by comparing the expected appearance with the newly observed appearance [5].

At the current, we have done limited number of similar experiments using two kinds of pullover. The results are summed up in Table 2. As shown in the table, once the clothes state was correctly estimated, grip coordinates effective for handling tasks were obtained. We are now studying a method to select the correct estimation when the estimation results conflict between left and right images.

6. Conclusions

We proposed a deformable model driven method for determining the 3D grip coordinates required for holding a specific part of a hanging clothes by a manipulator. In the

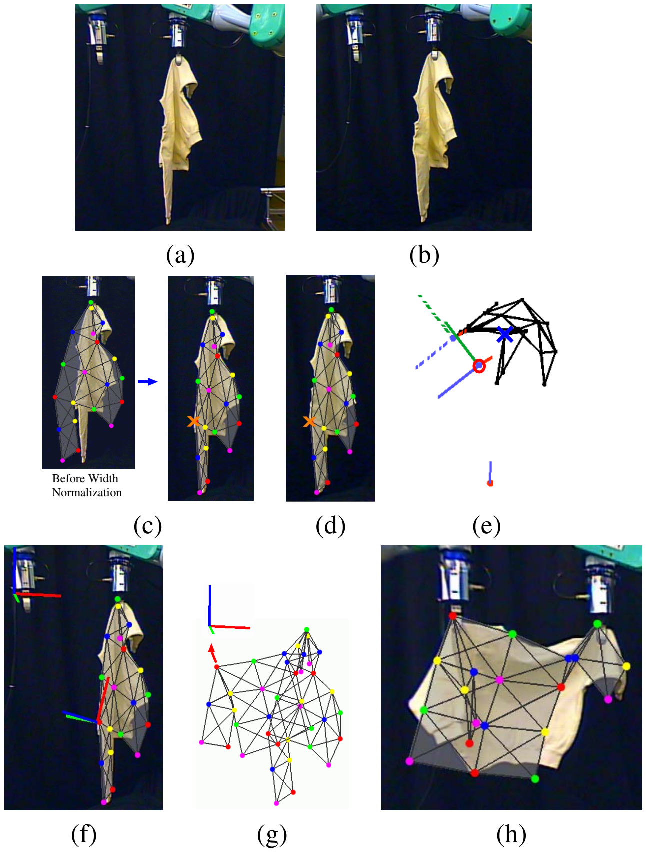


Figure 7. Handling experiment using relatively soft pullover: (a) observed left image; (b) observed right image; (c) estimation result(left); (d) estimation result(right); (e) resultant grip coordinates in a top view; (f) grasp and goal grip coordinates in the left image; (g) simulation on the action; (h) newly observed image with the predicted appearance superimposed.

task of handling clothes, accuracy in the grip coordinates are not so much required comparing to the case of handling rigid objects since clothes change shape along the gripper when it contacts the clothes. Considering this point, the proposed method determines the grip coordinates based on roughly estimated clothes shape. Preliminary experiments using actual manipulators showed the resultant coordinates are realistically effective for the handling task.

One of our future works is to improve the simulation of clothes deformation, which should reinforce the robustness of the estimation. However, we do not intend to use elaborate clothes model like the models studied in the field of computer graphics[6], because for our purpose, a deformable model which requires less prior knowledge of the clothes in question is desirable. Rather than that, key point

Table 2 Results using manipulators

	No.	Estimation	Holding up
Adult pullover	1	×	-
	2	○	○
	3	○	-(The manipulator could not realize the resultant grip coordinates although the coordinates look proper)
Toddler pullover	4	△(failed in left images)	-
	5	△(failed in left images)	-
	6	○	○
	7	△(failed in left images)	-
	8	○	○

is how to modify 3D predicted shape so as to make it close to actual shape based on the observed images. One way we are considering now is a feedback of the change in model appearances by the adjustment processes described in Section 3 to the original 3D predicted shape producing the appearances.

Other aspects we will consider include more collaboration with manipulations to simplify the recognition processes. For example, "shaking by a manipulator" should be effective to remove unstable deformations.

Acknowledgments

We are thankful to Dr. S. Hirai, Dr. K. Tanie, Dr. K. Sakaue, and Dr. T. Suehiro for their support to this research.

References

[1] M. Inaba and H. Inoue: "Hand eye coordination in rope handling", *JRSJ (Journal of Robotics Society of Japan)*, Vol. 3, No. 6, pp. 32–41, 1985.

[2] H. Nakagaki, K. Kitagaki, T. Ogasawara and H. Tsukune: "Study of Deformation and Insertion Tasks of a Flexible Wire", In *Proc. of IEEE International Conference on Robotics and Automation*, vol.3, pp. 2397–2402, 1997.

[3] S. Hirai, T. Tsuboi, and T. Wada: "Robust Grasping Manipulation of Deformable Objects", In *Proc. IEEE Symp. on Assembly and Task Planning*, pp.411–416, 2001.

[4] M. Kaneko and M. Kakikura: "Plannning strategy for putting away laundry –Isolating and unfolding task –", In *Proc. of the 4th IEEE International Symposium on Assembly and Task Planning*, pp. 429–434, 2001.

[5] Y. Kita and N. Kita: "A model-driven method of estimating the state of clothes for manipulating it", In *Proc. of 6th Workshop on Applications of Computer Vision*, pp.63–69, 2002.

[6] D. H. House and D. E. Breen: *"Cloth modeling and animation"*, A. K. Peters, Ltd., 2000.