The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

# A method for handling a specific part of clothing by dual arms

Yasuyo Kita*, Toshio Ueshiba*, Ee Sian Neo** and Nobuyuki Kita**
*Information Technology Institute **Intelligent Systems Institute
National Institute of Advanced Industrial Science and Technology (AIST)
{y.kita, t.ueshiba, rio.neo, n.kita}@aist.go.jp

*Abstract*— **In this paper, we propose a strategy for a dual-arm robot to pick up a specific part of clothing with one hand while holding the item of clothing with its other hand. Due to the large deformability of clothing, the handling requirements differ from those required for the handling of rigid objects. In the case of holding a specific part of clothing, large deformation leads to a large variety of positions and orientations of the target part, requiring flexibility of both visual recognition and motion control. On the other hand, since the clothing can flexibly curve over the hand, a relatively large range of suitable actions is allowed for grasping the clothing. By considering these characteristics, the following three-stage strategy is proposed. First, the state of the clothing is recognized from visual observation of the clothing using a deformable-model [1]. Then, the theoretically optimal position and orientation of the hand for handling a specific part of the clothing is calculated based on the recognition results. Finally, the position and orientation of the hand is modified by considering the executable motion range of the dual arms. Preliminary experimental results using actual observations of a humanoid robot were used to validate the effectiveness of the proposed strategy.**

## I. INTRODUCTION

Since home and rehabilitation robots are expected to become increasingly important in an aged society, these robots must have the ability to handle objects, such as clothing, which are used in daily life. However, due to large deformability, the handling requirements of clothing are different from those for rigid objects. First, large deformation leads to a large variety of positions and orientations of the part of clothing that needs to be grasped for the next action, requiring considerable flexibility with respect to both visual recognition and motion control.

Compared to handling soft, string-type objects, such as ropes and electric cables [2][3], it is more difficult to recognize the state of clothing being handled due to complex states of self-occlusion. Here, the term “recognize the state” refers to the recognition of not only the geometrical shape, but also each part of the clothing. This state information is necessary if we want to pick up a specific part (in this study, the shoulder part of a sweater) which is a basic action for handling tasks such as folding clothing into a desired shape.

Although some studies have been conducted in the area of automated clothing handling [4][5][6][7], few have examined aspects related to the recognition of the clothing state. Kaneko et al. [5] proposed a method which recognizes the clothing state by comparing the contour features (e.g., curvature, length-ratio) of observed data with models for clothing that is held at two points. However, the detailed contour features are difficult to robustly extract from real observations and are very sensitive to any slight deformation of the clothing. Additionally, the learning processes required to distinguish model shape characteristics from actual observations are complicated.

In [6],[1], a model-driven method, which recognizes the state of clothing held at a point, has been proposed. The method predicts possible appearances using a deformable model of the clothing and selects the one that best fits the observed appearance. By using such a model-driven strategy, once the results are obtained, the 3D position and orientation of a specific part can be estimated by checking the position and orientation of the model segment that most closely corresponds to that part. Consequently, the most appropriate position and orientation of the second hand for handling a specific part of clothing can be obtained using this straightforward strategy.

However, another difficulty exists with regard to practical handling of clothing. Because of the considerable variety of clothing deformations, the resultant positions and orientations are varied, and so some can easily fall out of the executable motion range of actual robots. Even so, large deformability of clothing also enables a number of “suitable ways” of handling. For example, since clothing can bend over the hand, the orientation of the hand when it holds the clothing is not strictly limited.

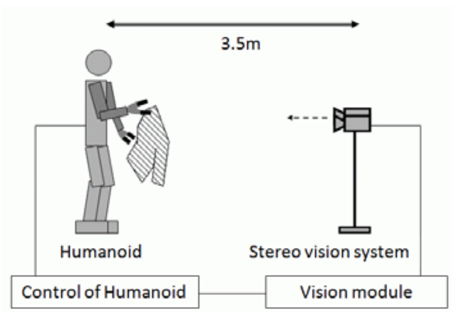
Based on these considerations, we propose a three-stage strategy for picking up a specific part of clothing: 1. clothing state estimation; 2. calculation of the theoretically optimal position and orientation of the hand for the next action; 3. modification of position and orientation so they can be executed. In this paper, we implement this strategy based on a deformable-model-driven recognition method [1] and evaluate the robustness and effectiveness of the strategy using a humanoid robot, as shown in Fig. 1.

In Sections 2 and 3, our system and clothing estimation method are briefly explained. In Section 4, methods for determining the position and orientation of the hand for grasping a specific part of clothing are proposed. In Section 5, experimental results using actual observations of a humanoid robot are discussed.

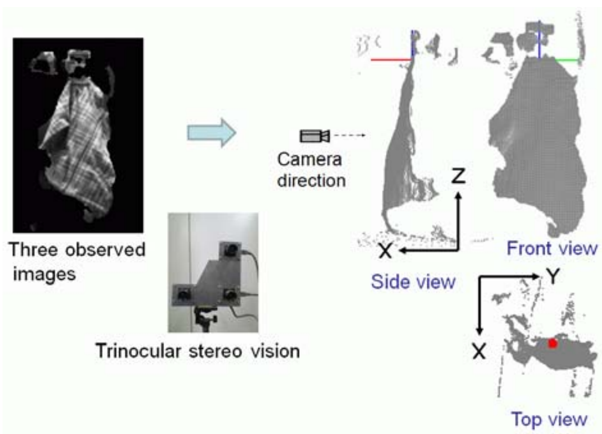
## II. OUR SYSTEM

Figures 1(a) and (b) show our system for grasping clothing. The system consists of a humanoid robot, HRP2 [8], and a trinocular stereovision system [9]. Although HRP2 has its

(a)




(b)




(c)

Fig. 1. System for handling clothing presented in this study: (a) calibration procedure; (b) relation between the humanoid and trinocular stereo camera system; (c) three-dimensional data obtained from the stereovision system (the view direction of the camera system is $-X$, with the red dot in the top view illustrating the holding position).

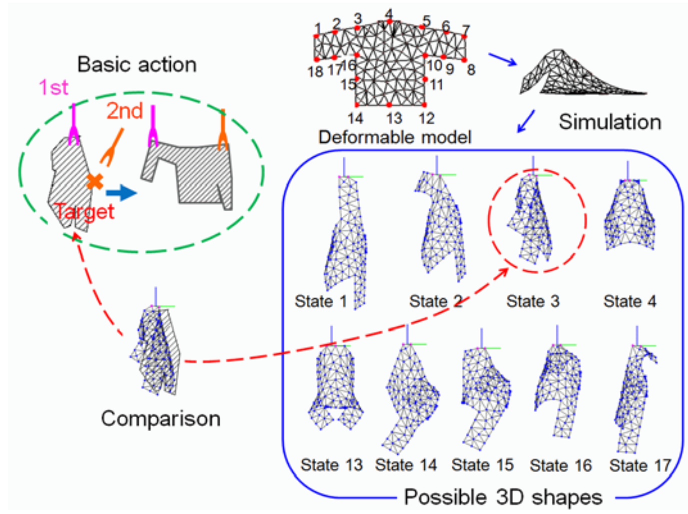



Fig. 2. Illustration of the basic action and the model-driven strategy to perform the basic action.

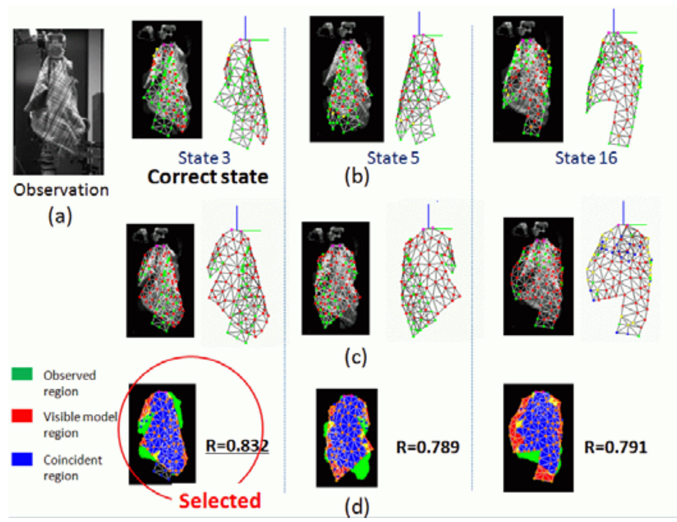
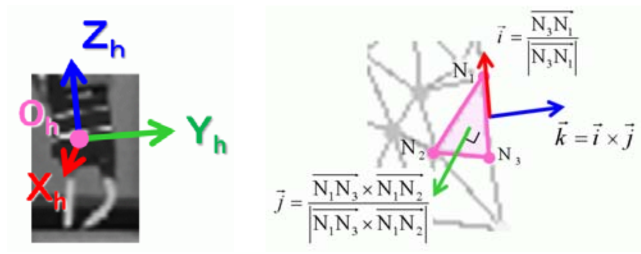



Fig. 3. Example of clothing recognition: (a) observed image; (b) predicted shapes obtained from physical simulation; (c) modified shapes after the adjustment from observed 3D data; (d) selection of state [1].

own vision system, it is currently unable to provide detailed 3D data of the entire clothing item. To compensate for this limited field of view, we used a stereovision system capable of recording dense 3D information with high accuracy, as shown in Fig. 1(c). Although this vision system is able to capture 3D information at a rate of 30 frames per second, we only used static 3D data taken at one time. The calibration of HRP2 and the stereovision system was performed using a marker, which required HRP2 to move around the marker as shown in Fig. 1(a). The accuracy of the 3D measurement obtained from 8∼10 point measurements is approximately $3 \sim 8$ mm. To extend the vertical field of view, three cameras were set on the stereo board after rotating them by 90 degrees; HRP2 and the vision module were connected via LAN, with communication between the module and HRP2 implemented using Common Object Request Broker Architecture (CORBA).

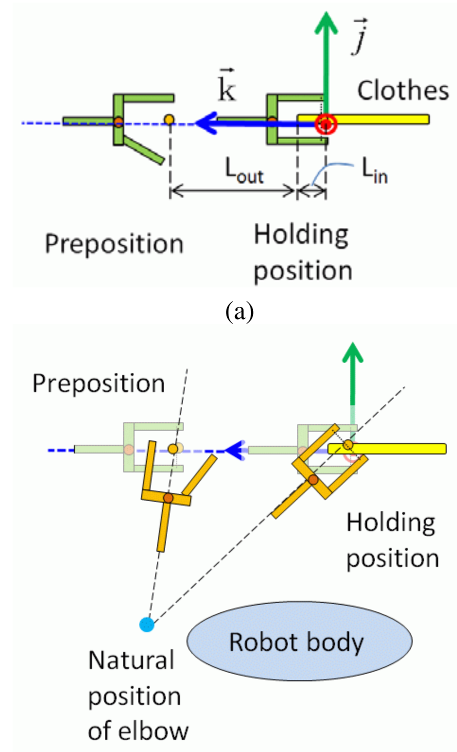
## III. Model-driven clothing state recognition

One important and basic action for handling clothing is grasping a specific part on the hem of an item of clothing that is held in the air, as shown in Fig 2. By iterating this action with two hands, clothing can be held at a variety of desired states. For carrying out this basic action, the 3D position and orientation of the target part are essential information. Even though dense 3D information, such as that shown in Fig. 1(c) is obtained using the system, it is almost impossible to recognize how the clothing is folded and where each part




(a) (b)

Fig. 4. Basic information for calculating hand coordinates: (a) hand coordinates of HRP2; (b) orientation of the model patch.




(b)

Fig. 5. Hand coordinates for grasping clothing: (a) best hand coordinates for the clothing; (b) one modification to make the model executable.

is in a data-driven way from the observed data. Therefore, we selected a model-driven strategy based on the assumption that simple knowledge about the target clothing, such as type (e.g., sweater, pants etc.) and approximate size and softness, is known in advance.

By simulating physical deformation of the target clothing based on this information, possible 3D shapes of the hanging clothing are obtained as shown in Fig. 2, At the current time, we have conducted this simulation using the cloth function of Maya 4.5 [10]. To simplify the problem, we assume that the front and back sides of the clothing are not separated and no thickness is given to the model. Based on this assumption, the model becomes a surface which deforms three-dimensionally. Here, we classify the clothing states according to the position at which the clothing is held, shown as "State 1," "State 2," and so on, in Fig. 2. We think this classification is natural since the grasping position is only one condition that explicitly determines the shape.

To carry out this approach despite the large variations in clothing shape, we use an estimation process consisting of two steps. First, a small number of representative 3D shapes are calculated using physical simulations of the hanging clothing. Then, after observing the clothing, each representative shape is deformed to better fit the observed 3D data. To select the correct state, the consistency between the adjusted shapes and the observed data is checked using overlap ratio $R$. Figure 3 shows one example of the recognition state. State 3 was correctly selected as the recognition result of the observation of Fig. 3(a). Although the number of experiments was small, the rate of successful recognition was approximately 80% [1].

Notice that once the clothing states are recognized with this model-driven strategy, the position and orientation of any specific part can be calculated from the 3D information of the model segment corresponding to the part.

## IV. Calculation of Hand Coordinates for Grasping Action

The hand coordinates of our humanoid are shown in Fig. 4(a). The origin, $O_h$, is set at the center of the rotation axis of the wrist. The $Y_h$-axis is set along the rotation axis of the wrist, and the $X_h$-axis is set perpendicular to the flat front surface of the hand. The $Z_h$-axis completes the right-handed coordinate system; the hand has two fingers, a straight thumb and a curved finger. Only the thumb opens by rotating around its base in the $-Y_h$ direction.

To grasp a specific part, the hand should approach the part from outside the clothing with its thumb open in the normal direction of the part. When the fingertips reach a position well suited for grasping, the thumb should close to pinch the part with the other finger.

### A. Calculation of best action for the clothing

The position and orientation of the triangular model patch corresponding to the target part can be calculated from the 3D position of the three vertices, $N_1, N_2$ and $N_3$, which enclose the patch shown in Fig. 4(b).

Specifically,
the best pinching direction is as follows:

$$\vec{j} = \frac{\overrightarrow{N_1N_3} \times \overrightarrow{N_1N_2}}{|\overrightarrow{N_1N_3} \times \overrightarrow{N_1N_2}|}$$

The best approach direction is this:

$$-\vec{k} = \vec{i} \times \vec{j}$$

where,

$$\vec{i} = \frac{\overrightarrow{N_3N_1}}{|\overrightarrow{N_3N_1}|}$$

Using this information, one of the best hand coordinates for holding the part is automatically obtained as shown in

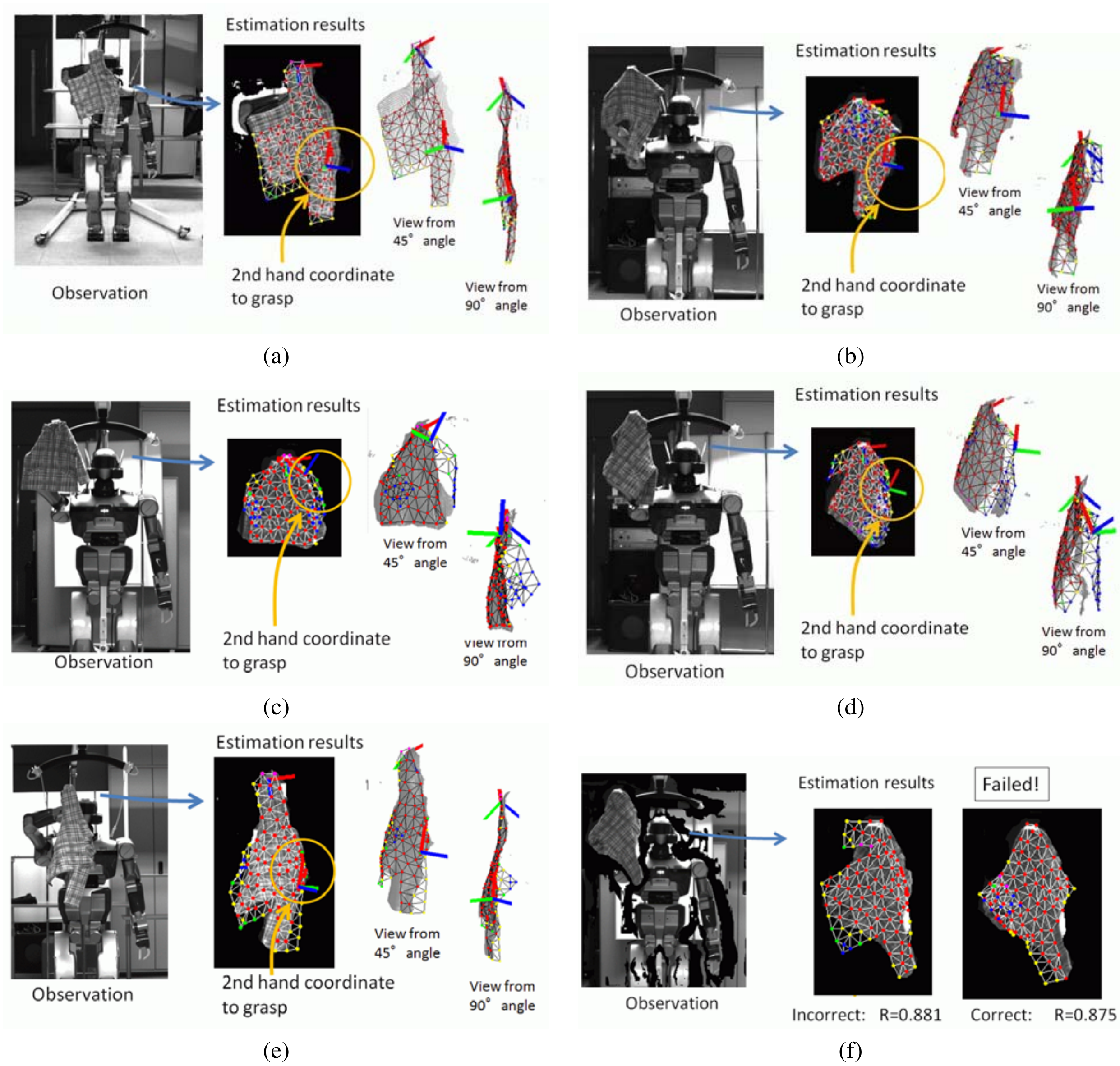



Fig. 6. Resulting hand coordinates calculated from visual information.

Fig. 5(a). The position for holding is set at the position where is $L_{in}$ inside from the middle point of $N_1, N_3$. The axes of $X_h, Y_h$ and $Z_h$ are set to $\vec{i}$, $\vec{j}$ and $\vec{k}$,respectively. To determine the approach direction, the position where is $(L_{in} + L_{out})$ away in the $\vec{k}$ direction from the holding position is used as the preposition.

## B. Modification to an executable action

Due to the limited executable motion of the dual arms, the hand coordinates calculated in the previous subsection often become unexecutable. For example, if the calculated hand coordinates appear as shown by the green hands in Fig. 5(b), then it is impossible to achieve the orientation at that position and the hand coordinates must be modified to executable ones. Fortunately, as stated previously, clothing deformability has relatively wide allowances for "the optimal way", as the hand can grip the clothing while bending it once the hand grips it between its fingers.

For any modification, a variety of approaches can be considered. If movement of the first hand (the hand that is holding the clothing) is permitted, then more action plans are possible. In the current study, we use a simple modification of the second hand, as shown in Fig, 5(b). Specifically, the $Z_h$-axis is set on the line connecting the position of the fixed natural elbow position and the target position. Note that the normal direction of the specific part is still used to determine the approach direction.

# V. Experiments

We conducted several experiments involving changing the shape of a sweater held by the first hand (right hand) of HRP2 and present the results of the six trials in Figure 6. The estimation results are shown as mesh models superposed

Table 1 Differences of hand coordinates obtained by applying the proposed method from those manually obtained.

| | Difference | | | |
|---|---|---|---|---|
| | X-axis | Y-Axis | Z-axis | Origin |
| | | (degree) | | (mm) |
| Fig. 6(a) | 10.2 | 8.8 | 12.8 | 53.5 |
| Fig. 6(b) | 5.8 | 13.8 | 13.6 | 26.3 |
| Fig. 6(e) | 12.2 | 13.2 | 12.4 | 14.5 |
| Average | 9.4 | 12.0 | 13.0 | 31.4 |

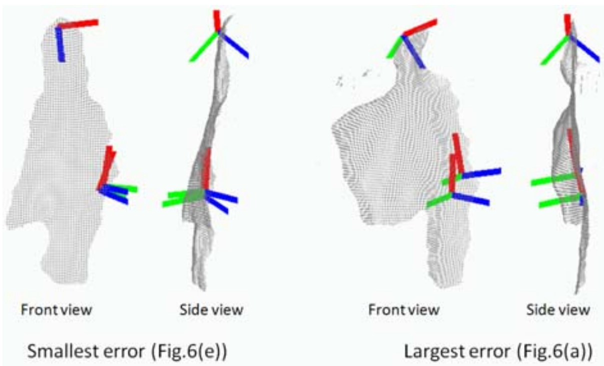



Fig. 7. Comparison of hand coordinates obtained by applying the proposed method to those manually obtained.

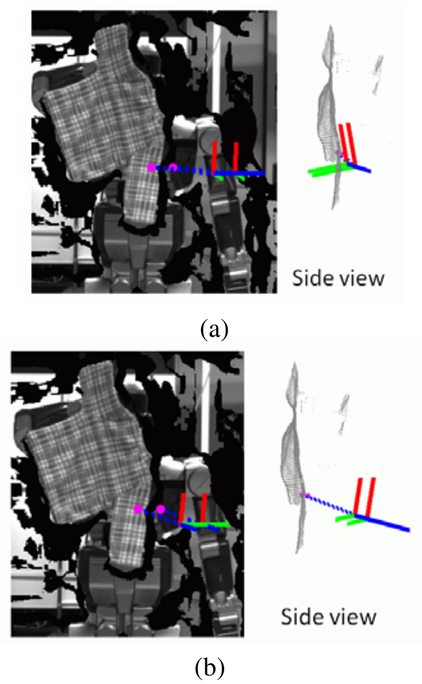


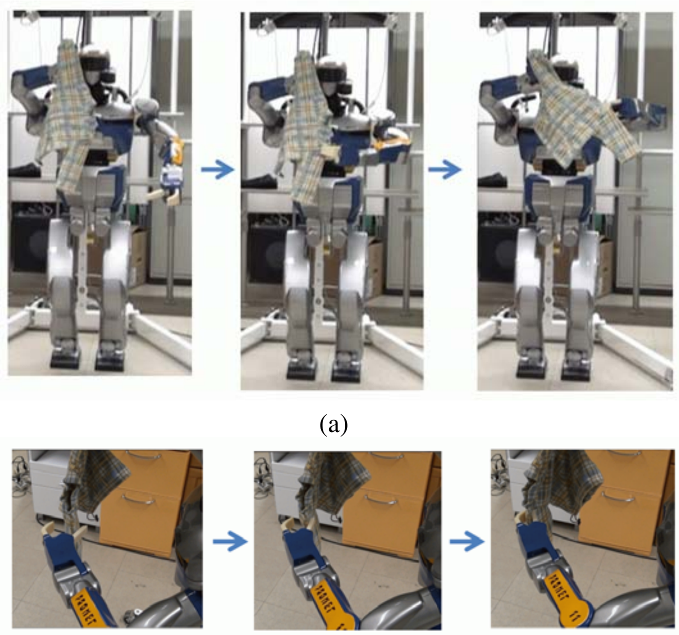

Fig. 8. Example of modification of hand coordinates: (a) best hand coordinates for the clothing; (b) coordinates modified to be executable.

on the observed 3D data. In five cases, (all cases except Fig. 6(f)), the state of the sweater was successfully recognized. For these five cases, the best position and orientation of the second hand coordinates for grasping the shoulder part were calculated by the proposed method: the resultant $X_h$-, $Y_h$- and $Z_h$-axes are illustrated by the red, green and blue lines, respectively. For purposes of illustration, the origin of the coordinates are set at the grasping position (2 cm inside the hem of the sweater) in the figures.

Notice that even in the case where the target part was occluded, as shown in Fig. 6(c) and (d), the deformed model shape was able to estimate the position and orientation of the part. This information can tell from which direction the sweater should be observed next to obtain accurate 3D information of the target part. Actually, in these two cases, the tangent of the shoulder patches from the image plane is 92 degrees and 76 degrees, respectively. Considering that the 3D data of the planes with a tangent of less than 45 degrees can be measured by our stereo system, the proper rotation angles around the vertical axis of the first hand (current holding hand), which are -47 degrees and -31 degrees, respectively, can be derived.

Using three cases where the shoulder is visible (Fig. 6(a),(b) and (e)), we compared the resultant coordinates to those obtained manually. The manually obtained coordinates were selected from three 3D observed points corresponding to the shoulder; the differences in the origins and axes are summarized in Table 1. The differences in the orientation was approximately 10 degrees. Considering the flexibility of the clothing when it is pinched between the fingers, the errors are within allowable ranges. The difference in the position of the origin was 14.5 $\sim$ 53.5 mm. Figure 7 shows the smallest and largest cases regarding to the errors. As can be seen in the figure, the deviations of the origin in both cases are along the clothing surface. Therefore, the resultant coordinates are still effective for handling action when accuracy of the position along the surface is not very important. The reason for the large deviation observed in Fig. 6(a) is because the estimated shape is vertically shifted from the actual shape. This happened because the model shape was obtained by deforming the shape of State 1 (Fig. 2), although the position held by the first hand was actually between State 1 and State 2. This can be corrected if adjustment processes, such as the “vertical translation” process proposed in [6], are added.

In four cases (except the case shown in Fig. 6(e)), the resultant hand coordinates were unexecutable by the second hand of HRP2. This finding indicates the importance of modifying the position and orientation of the hand by considering the executable motion range of the dual arms. Figure 8 shows a modification result for the case of Fig. 6(a), which was calculated off-line after the experiments. In the figures, in a different way from Fig. 6, the hand coordinates, with the origin at the wrist, are directly illustrated and connected to the fingertip position (violet circles) with dashed lines. As shown, the best action (Fig. 8(a)) requires unnatural hand orientation and is unexecutable. By changing the orientation of the hand coordinates, as explained in Section 4B, executable orientation of the second hand could be achieved, as shown in Fig. 8(b).



(a)

(b)

Fig. 9. Example of actual handling processes.

Figure 9 shows actual handling of the executable case (Fig. 6(e)). Figure 9(a) consists of photos of the three consecutive actions: approaching to grasp, grasping, and picking up the shoulder part. Figure 9(b) are photos taken from another angle close to the body of the robot during the consecutive actions of approaching to grasp and grasping of the shoulder part. Despite the inaccuracy of 0.8 cm for the observed 3D data, the clothing was successfully grasped (a video showing the handling process is provided as Supplementary Material).

## VI. Conclusions

We propose a method for automatically grasping a specific part of clothing based on the recognition results of the clothing state. By using the model-driven strategy, once the clothing state is estimated, the best position and orientation of the hand coordinates for grasping a specific part can be calculated from 3D information of the model segment corresponding to that part. Several experiments using actual observations of the clothing held by a humanoid robot showed promising results.

Through the experiments, we found the fact that the position and orientation of the hand which are calculated based on the clothing states often fell out of the executable motion range of the dual arms. In this paper, a simple method of modifying them to be executable was tested. For more robust and natural handling, more sophisticated actions such as moving the first hand, must be studied.

In addition to calculating the hand coordinates for grasping a specific part of clothing, the estimated clothing state can be used for a variety of other actions. One such example is shown in Section V, where we show how the appropriate rotation angle can be obtained from 3D data of the model segment corresponding to the occluded target part. Additional investigation of more complex actions derived from the recognized clothing state is one of our future tasks.


## Acknowledgment

This work was supported by a Grant-in-Aid for Scientific Research, KAKENHI (19300066). We are thankful to Dr. K. Yokoi and Dr. K. Sakaue for their support in this research.


## References


[1] Y. Kita, T. Ueshiba, E. S. Neo and N. Kita:"Clothes state recoginition using 3D observed data", In *Proc. of International Conference on Robotics and Automation*, pp. 1220–1225, 2009.

[2] 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.

[3] 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.

[4] E. Ono, H. Okabe, H. Ichijo and N. Aisaka:"Robot Hand with Sensor for Cloth Handling", In *Proc. 1990, Japan.U.S.A. Symp. on Flexible Automation*, pp. 1363–1366, 1990.

[5] 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.

[6] Y. Kita, F. Saito and N. Kita: "A deformable model driven visual method for handling clothes", In *Proc. of International Conference on Robotics and Automation*, pp.3889–3895, 2004.

[7] K.Yamazaki and M.Inaba: "A Cloth Detection Method Based on Wrinkle Features for Daily Assistive Robots", In *Proc. the IAPR Conf. on Machine Vision Applications*, pp. 366–369, 2009.

[8] E. S. Neo, K. Yokoi, S. Kajita and K. Tanie:: "Whole-Body Motion Generation Integrating Operator's Intention and Robot's Autonomy in Controlling Humanoid Robots", *IEEE Trans. on Robotics*, Vol. 23, No. 4, pp. 763–775, 2007.

[9] T. Ueshiba: "An Efficient Implementation Technique of Bidirectional Matching for Real-time Trinocular Stereo Vision", In *Proc. of 18th International Conference on Pattern Recognition*, pp.1076–1079, 2006.

[10] D. A. D. GOULD: *"Complete Maya Programming"*, Morgan Kaufmann Pub, 2004.