Strategy for Folding Clothing on the Basis of Deformable Models

Yasuyo Kita(✉), Fumio Kanehiro, Toshio Ueshiba, and Nobuyuki Kita

National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan
{y.kita,f-kanehiro,t.ueshiba,n.kita}@aist.go.jp

Abstract. In this study, a strategy is given for automatically reshaping an item of clothing from an arbitrary shape into a fixed shape by using its deformable model. The strategy consists of three stages that correspond to the clothing state: unknown (before recognition), unknown to known (recognition), and known (after recognition). In the first stage, a clothing item that is initially placed in an arbitrary shape is picked up and observed after some recognition-aid actions. In the second stage, the clothing state is recognized by matching the deformable clothing model to the observed 3D data [1]. In the third stage, a proper sequence of grasps toward the goal state is selected according to the clothing state. As an instance of this strategy, a folding task was implemented in a humanoid robot. Experimental results using pullovers show that the folding task can be achieved with a small number of grasping steps.

Keywords: Robot vision · Clothes handling · Deformable model

1 Introduction

Domestic service robots and rehabilitation robots are expected to play an active role in graying societies, and thus it is becoming increasingly important for robots to autonomously handle daily necessities, such as clothes folding. Because of clothing's high deformability and complex self-occlusions, the techniques required for handling clothing are quite different from those required for handling rigid objects.

Recently, many studies on handling clothes have been conducted, such as those on skillful manipulation of a cloth [2][3] and on detection and classification of clothing [4],[5],[6]. Studies on automatic folding of clothing items are also becoming more common. Maitin-Shepard et al. [7] demonstrated that a robot can skillfully handle towels on the basis of corner detection by using multiple-view observation. However, extending this method to other types of clothing, which may have complicated shape variations and several types of characteristic points, seems difficult. Berg et al. [8] proposed a method of folding several types of clothing by assuming that the target clothing is neatly laid on a desk when the process begins. Although Miller et al. [9] eased the tidiness requirements of this

A. Campilho and M. Kamel (Eds.): ICIAR 2014, Part II, LNCS 8815, pp. 442–452, 2014.
DOI: 10.1007/978-3-319-11755-3_49

method, as a first step items still have to be flattened without self-occlusions. Flattening all parts of an item itself is not necessarily an easy task, as shown in some other studies (e.g., [5]) and requires extra procedures.

Because we exclude flattening in our strategy, it is desirable to recognize the clothing state while changing the grasping on items held in the air. Kaneko et al. [10] proposed a method that recognizes the clothing state by comparing the contour features (e.g., curvature and length ratio) of the observed data with models when clothes are held by two points in the air. The state of clothes held at two points, however, can vary tremendously because of the large number of two-point combinations. In addition, the detailed contour is difficult to extract robustly from real observations , which are highly sensitive to slight deformations of the clothing. Osawa et al. [11] and Cusumano-Towner et al. [12] proposed methods for estimating clothing state in a manner similar to the method proposed by Kaneko et al. [10] and showed a good recognition rate for several types of clothing. However, those methods require repeated grasping of the lowest point of a hanging item of clothing, which is time consuming and delays achieving the target state. In addition, reaching the lowest point with one hand while holding the item with the other hand might be impossible when the clothing is large. We[13][14] previously proposed a method that recognizes the state of a clothing item by using its deformable model and calculates the necessary information for grasping the item at any specific point, such as a shoulder. One advantage of our approach is that the overall configuration of the clothing item is considered through use of its deformable model. This is important both for handling the item without undesirable twisting of the item and for selecting the best action toward reaching the goal state.

In this study, we present a strategy for achieving a practical task: folding an item of clothing from an arbitrary shape into a fixed form on the basis of model-based recognition of the clothing state. In the next section, we introduce our complete strategy, which consists of the following three stages: actions to aid recognition, visual recognition of clothing state, and actions to complete a task on the basis of the clothing state. The first and second stages, which have been already presented elsewhere, are briefly explained in the same section. In Section 3, the key functions of the third stage and some methods to realize the functions are described. In Section 4, results of experiments with a humanoid robot are presented and discussed. Finally, in Section 5, current progress and future subjects are summarized.

2 Strategy for Folding Clothing

Fig. 1 shows the overall flow for reshaping an item of clothing from an arbitrary form into a fixed folded form. We treat “recognition of clothing state” as a core process of the strategy, because, once it is achieved, the desirable clothing shape can be positively led and traced. Hence, the total flow is divided into the following three stages:

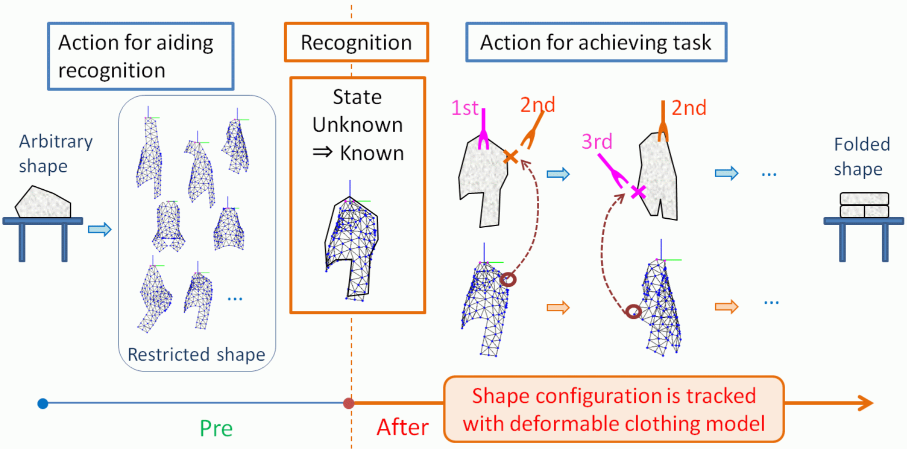
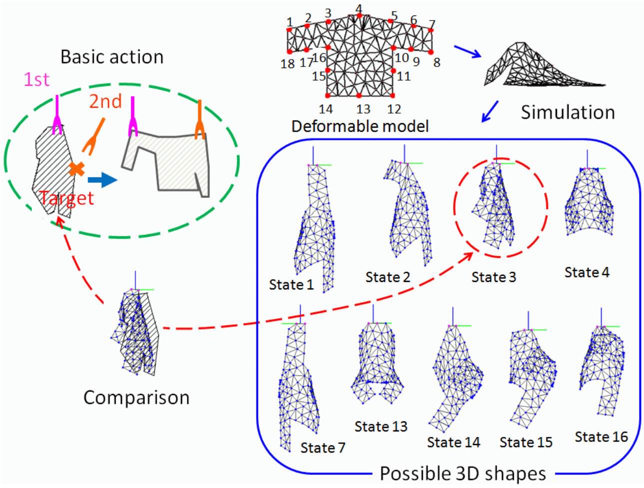


Fig. 1. Strategy based on model-driven analysis

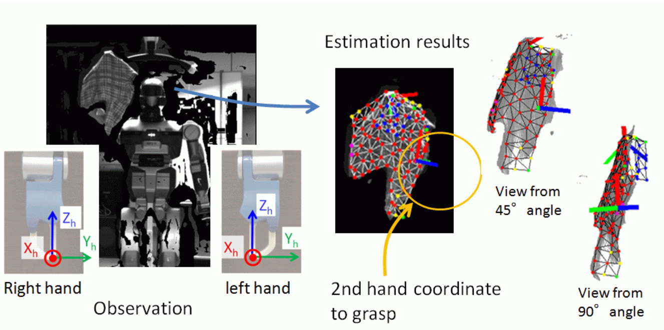
I. Action for aiding recognition
II. Model-driven visual recognition
III. Action for achieving task based on the recognition

At the first stage, the target item of clothing is brought into the shape which makes the following recognition process easier and more robust[14]. In our methods, two types of actions are used for this purpose: restrictions on shape configuration and acquisition of abundant information about the item. For restricting the shape configurations, we adopt the strategy of grasping the rim of the target item of clothing after picking it from the desk. Through this action, the possible shapes are limited to the states as shown in “State x” in Fig. 2(a). For acquiring abundant information about the clothing item, we use 3D observed data taken from different view directions in an integrated fashion [1][15].

At the second stage, we use the model-based state recognition methods [13]. Fig. 2 shows the core idea of these studies. By simulating the physical deformation of the target clothing based on an approximate knowledge about the target, such as the type (e.g., sweater, pants), approximate size, and softness, possible 3D shapes of the clothing being held in air are predicted as shown in Fig. 2(a). Here, the clothing states are classified into "State 1," "State 2," ... according to the position in which the article is held. After observation, to consider the variation in the shape of clothing that can arise even when it is held at the same point, one representative shape of each state is deformed to better fit the observed 3D data. The state that shows the best fit to the observed data is selected as the correct state. Once the clothing model is fit to actual observations, the position and normal direction of any part of the target article of clothing can be estimated by checking the 3D data corresponding to that part of the model. This is useful in determining the action plan for further handling of the part. Fig. 2(b) shows an example of the calculated action plan for grasping



(a)



(b)

Fig. 2. Model-driven strategy of clothes recognition for automatic handling: (a) basic action and model-driven recognition; (b) calculation of action to perform the basic action; to grasp a specific part, the hand should approach the part from outside along the blue line with its thumb open in the direction of the green line

a shoulder part. The red, green and blue lines superposed on the left and right hands show the definition of their hand coordinates respectively. These lines are used for illustrating the position and orientation of the hand in all the figures of this paper. For example, in Fig. 2(b), the red, green and blue lines illustrate the position and orientation of the left hand at the point when it should close its fingers to grasp the shoulder part; the hand should approach the part from outside along the blue line with its thumb open in the direction of the green line.

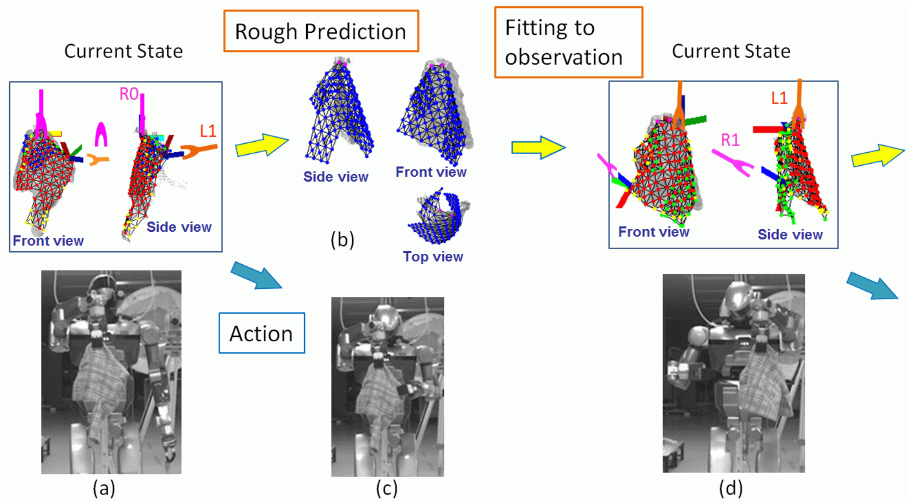


Fig. 3. Rough prediction and fitting to observations by deformable clothing model

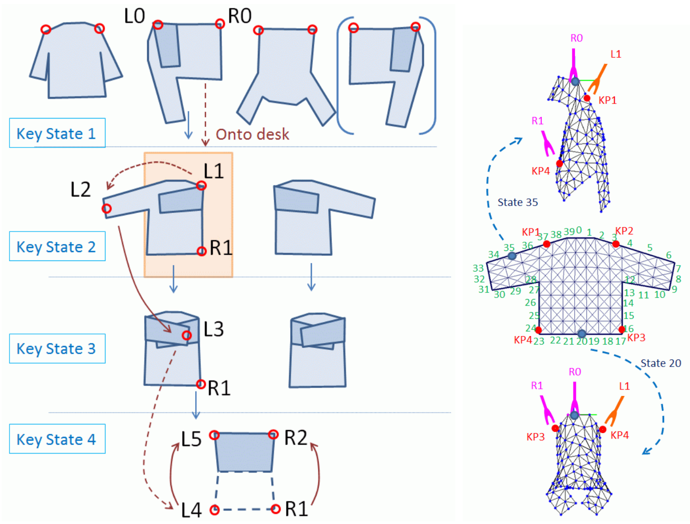
3 Actions for Achieving Folding Task

After the recognition process, a proper sequence of grasps to achieve the goal task is determined according to the resultant clothing state. Execution of the sequence while keeping the clothing state known is done by repeating the rough prediction of the clothing shape and fitting the shape to the observations at each grasp change. Fig. 3 illustrates this core process. The model in Fig. 3(a) shows an example of the clothing state obtained by the recognition process. The part to be grasped next is qualitatively determined from the state, such as grasping rim position 37 to move from state 25. The position and pose of the hand coordinates necessary to execute this action are calculated as indicated by the red, green, and blue lines superposed on Fig.3(a). During execution of this action (Fig. 3(c)), the 3D shape after this action is predicted, as shown by the model in Fig. 3(b). Then, the model is deformed and fitted to the observed data, as shown in Fig. 3(d). This gives a new "current state," and the system is then ready for the next grasping step.

In this study, we implemented the proposed strategy for the task of folding a pullover. Folding can be realized by transiting the target clothing item through some shape configurations. The sequence of key states for folding a pullover is shown in Fig. 4(a): Body opened, One sleeve folded, Two sleeves folded, and Body folded.

3.1 Opening the Body

Since gravity is useful for opening clothing items with few wrinkles, we set a subgoal of opening a pullover in air (Key State 1 in Fig. 4). By allowing three



(a) (b)

Fig. 4. Key states for folding a pullover: (a) folding action from S-B case; (b) automatic determination of opening action

types of opened states, S-S, B-B, and S-B (or B-S), the opening procedure can be realized by a small number of grasping steps from any initial state. Here S and B represent a shoulder and a corner of the bottom, respectively. The four red points in Fig. 4(b), $KP_i (i = 1, \ldots, 4)$, show the positions of parts S and B. The actions necessary to open the body of the clothing item can be systematically determined by selecting the pair KP_i according to the clothing state. In Fig. 4(b), two examples (states 20 and 35) are shown.

3.2 Folding on a Desk

For each of S-S, B-B, and S-B (B-S), the action sequence for folding the pullover can be determined in advance. Fig. 4(a) shows the sequence of the folding actions for case S-B.

During the folding process, two types of grasp steps are necessary. In Fig. 4(a), these are illustrated by the dotted lines and represented by the markers $L1 \rightarrow L2$ (grasping the tip of a sleeve) and $L3 \rightarrow L4$ (grasping a corner of the bottom). During the image processing for detecting the position and pose of the target part at each grasping step, the expected position of the target part is used to simplify the processing and obtain robust results. The details of the processes are as follows.

1. Detection of the tip of the hanging sleeve ($L1 \rightarrow L2$)

Before this process, the trunk of the item is placed at the edge of the desk so that the sleeve hangs down from the edge, as shown in Fig. 5(a). As a result,

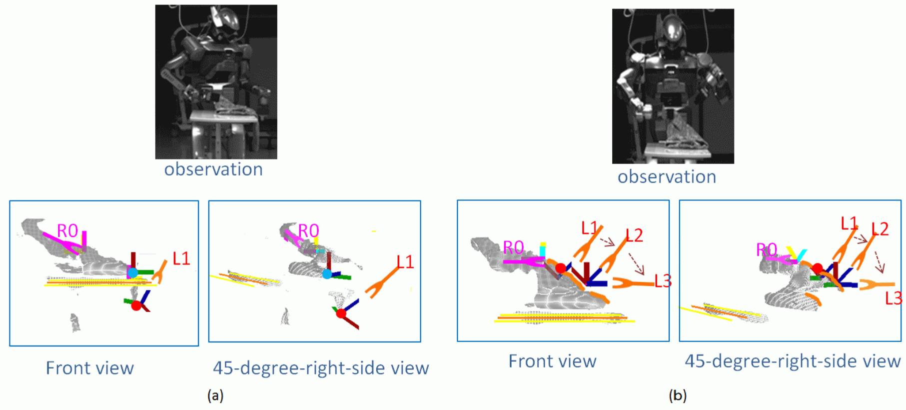


Fig. 5. Grasping actions during folding on a desk: (a) picking up a sleeve; (b) tracing a bottom line

the position of the tip of the hanging sleeve can be assumed to lie within a limited space and is robustly detected. Here, the extracted position of the sleeve tip and calculated position and pose of the left hand grasping the sleeve are superimposed on the front and side views of the observed 3D data and indicated by red points and red, green, and blue lines, respectively. Once the tip of the sleeve is grasped by a hand, the sleeve is folded by moving the hand in a circular trajectory whose center is on the edge of the desk (blue points in Fig. 5(a)).

2. Extraction of the 3D rim line ($L3 \rightarrow L4$)
Because the other corner of the bottom is held by the other hand (the holding hand) at this phase (e.g., R0 in Fig.5(b)), the corner can be grasped by sliding the grasping hand along the rim at the bottom. The rim can be detected by extracting the 3D boundary line of the region connected to the holding hand, as shown by the orange line in 5(b).

4 Experiments

We conducted experiments using two pullovers, Pullover A and B. Pullover B is softer than Pullover A. Eight experimental results are summarized in Table 1. In the experiments, we adopt the strategy of grasping the locally lowest rim of the hanging item first for restricting the shape configurations. After this first action, the pullovers were held by a tip of a sleeve (Nos. 1,4,5, and 7 of Table 1) or by a corner of the bottom (Nos. 2,3,6, and 8 of Table 1) . Fig. 7(a)–(d) shows one example (No. 5 of Table 1).

Then, the pullover is observed from multiple direction by the method proposed in [1]. During the process, the item is rotated around the vertical axis to be convex from the perspective of the camera automatically. In the cases where convex observation were not automatically obtained (Nos. 6 and 7), appropriate angles were manually input to continue the experiments.

Table 1. Experimental results for two pullovers (– means "no trial")

No.	Pullover	State	Recognition	Opening	Folding
1	A	8	○	○	○
2	A	17	×	×	–
3	A	23	○	○	–
4	A	8	○	○	△
5	A	32	○	○	○
6	B	17	○	○	–
7	B	8	○	○	△
8	B	17	○	○	–

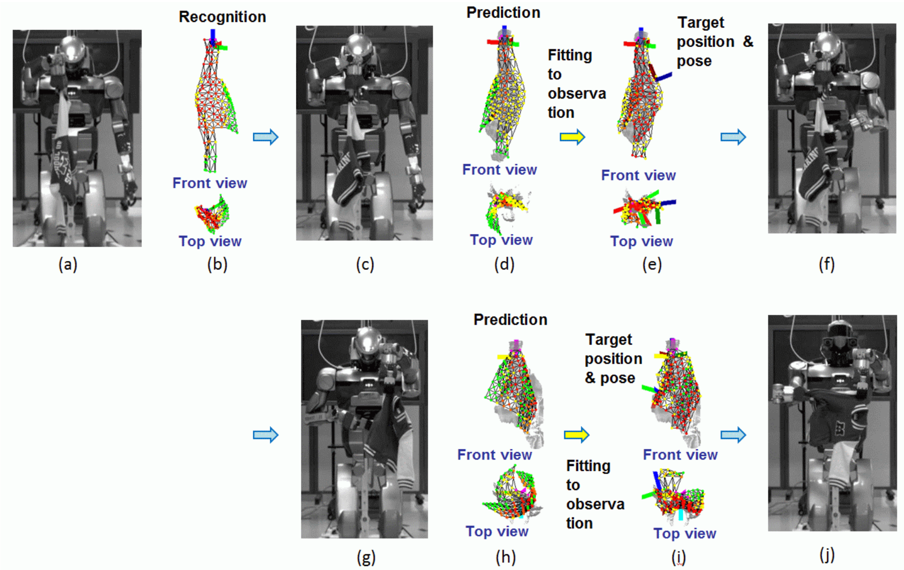


Fig. 6. Result of opening (Pullover B)

Once a pullover was observed from the side yielding a convex surface, the state of the pullover was successfully recognized and could be opened in most cases. Fig. 6 shows a case where Pullover B was successfully opened (No. 7 of Table 1). After the observation (Fig. 6(a)) was recognized as State 8, the model shape was fitted to the observed 3D data as shown in Fig. 6(b). Here, KP_2 and KP_3 were selected for the opening process. To grasp the shoulder part ($KP_2(Rim3)$) first, the item was rotated (Fig. 6(c)). Concurrently, the shape of the item after this rotation was predicted by rotating the current model shape, as shown by the model in Fig. 6(d). After fitting, this predicted shape was matched to the newly observed 3D data as shown in Fig. 6(e), where the position and

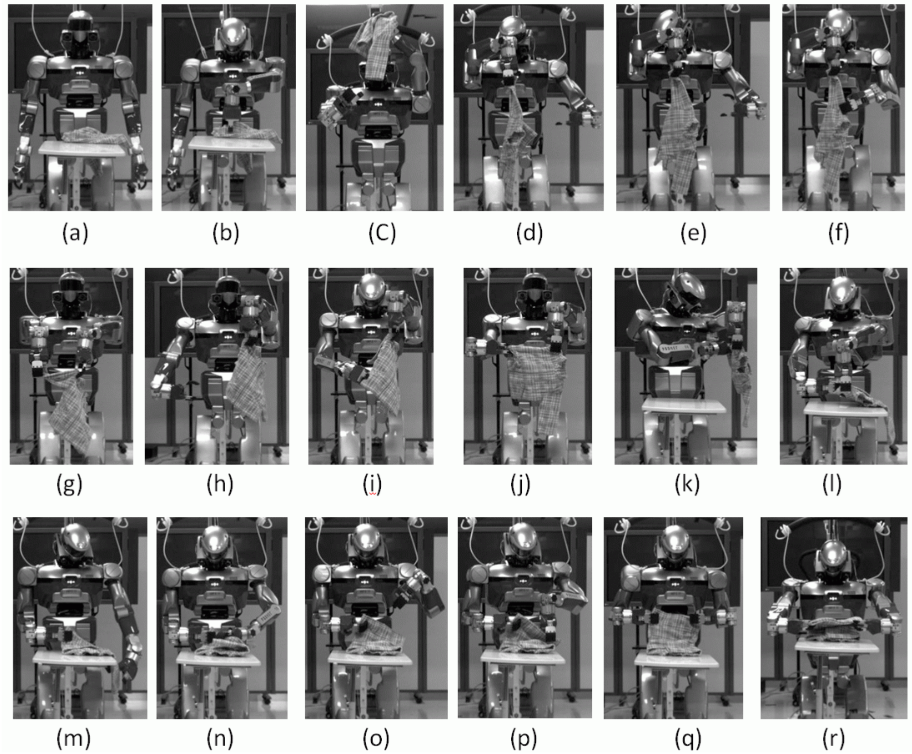


Fig. 7. Result of successful folding process (Pullover A)

pose of the hand to grasp the target part (rim position 3) was calculated as shown by the three axes illustrated as red, green, and blue lines superposed on Fig.6(e). After grasping the position according to these values (Fig. 6(f)) and releasing the sleeve from the right hand, a new observation was obtained (Fig. 6(g)). Here, again the process of rough prediction (Fig. 6(h)) and fitting (Fig. 6(i)) was repeated. As a result, $KP_3(Rim16)$ was detected and used to calculate the next motion for the right hand. Finally, the item was well opened in S-B mode, as shown in Fig. 6(j).

In the case of No. 2 in Table 1, although the correct state was selected, the clothing model failed to fit the surface. As a result, the estimation of the target part failed.

Fig. 7 shows an example of a successful folding process, from the first step to the last (No. 5). In the cases of No. 4 and No. 7, the folding procedure went well before the grasping of both corners of the bottom, at which time the robot failed in sliding its hand along the rim of the bottom.

5 Conclusions

We proposed a strategy for reshaping an item of clothing from an arbitrary shape into a fixed folded shape by using a deformable model to consider the overall configuration of the item. At each grasping step, the position and pose of the part to be grasped is determined on the basis of the corresponding part of the deformable clothing model, which is fit to the observed data. Through the experiments, we found that deforming the item to a desired shape is a key issue from a practical viewpoint. We showed the result with only one type of clothing (pullovers) in the experiments, but we believe that the proposed strategy is applicable to multiple types by changing clothing models. One of our future subjects is to enable handling of softer clothing items.

Acknowledgments. This work was supported by a Grant-in-Aid for Scientific Research, KAKENHI (22240019). We are thankful to Dr. K. Yokoi, Dr. Y. Kawai and Dr. K. Harada for their support in this research.

References

1. Kita, Y., Ueshiba, T., Kanehiro F., Kita, N.: Recognizing clothing states using 3D data observed from multiple directions. In: Proc. of International Conference on Humanoid Robots 2013, pp. 227–233 (2013)
2. Shibata, M., Ota, T., Endo, Y., Hirai, S.: Handling of Hemmed Fabrics by a Single-Armed Robot. In: Proc. of 2008 IEEE Conf. on Automation Science and Engineering, pp. 882–887 (2008)
3. Yamakawa, Y., Namiki, A., Ishikawa, M.: Motion Planning for Dynamic Folding of a Cloth with Two High-speed Robot Hands and Two High-speed Sliders. In: Proc. of IEEE Int'l Conf. on Robotics and Automation (ICRA 2011), pp. 5486–5491 (2011)
4. Yamazaki, K., Nagahama, K., Inaba, M.: Daily Clothes Observation from Visible Surfaces Based on Wrinkle and Cloth-Overlap Detection. In: Proc. the IAPR Conf. on Machine Vision Applications, pp. 275–278 (2011)
5. Willimon, B., Birchfield, S., Walker, I.: Classification of Clothing using Interactive Perception. In: Proc. of IEEE Int'l Conf. on Robotics and Automation (ICRA 2011), pp. 1862–1868 (2011)
6. Ramisa, A., Alenya, G., Moreno-Noguer, F., Torras, C.: Using depth and appearance features for informed robot grasping of highly wrinkled clothes. In: Proc. of International Conference on Robots and Systems, pp. 1703–1708 (2012)
7. Maitin-Shepard, J., Cusumano-Towner, M., Lei, J., Abbeel, P.: Cloth Grasp Point Detection based on Multiple-View Geometric Cues with Application to Robotic Towel Folding. In: Proc. of IEEE Int'l Conf. on Robotics and Automation (ICRA 2010) (2010)
8. van den Berg, J., Miller, S., Goldberg, K., Abbeel, P.: Gravity-based robotic cloth folding. In: Hsu, D., Isler, V., Latombe, J.-C., Lin, M.C. (eds.) Algorithmic Foundations of Robotics IX. STAR, vol. 68, pp. 409–424. Springer, Heidelberg (2010)
9. Miller, S., Fritz, M., Darrell, T., Abbeel, P.: Parameterized Shape Models for Clothing. In: Proc. of IEEE Int'l Conf. on Robotics and Automation (ICRA 2011), pp. 4861–4868 (2011)

10. Kaneko, M., Kakikura, M.: Planning 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)
11. Osawa, F., Seki, H., Kamiya, Y.: Unfolding of massive laundry and classification types by dual manipulator. JACIII **11**(5), 457–463 (2007)
12. Cusumano-Towner, M., Singh, A., Miller, S., O'Brien, J., Abbeel, P.: Bringing Clothing into Desired Configurations with Limited Perception. In: Proc. of IEEE Int'l Conf. on Robotics and Automation (ICRA 2011) (2011)
13. Kita, Y., Ueshiba, T., Neo, E.S., Kita, N.: A method for handling a specific part of clothes by dual arms. In: Proc. of International Conference on Robots and Systems, pp. 480–485 (2009)
14. Kita, Y., Kanehiro, F., Ueshiba, T., Kita, N.: Clothes handling based on recognition by strategic observation. In: Proc. of International Conference on Humanoid, pp. 53–58 (2011)
15. Ueshiba, T.: An Efficient Implementation Technique of Bidirectional Matching for Real-time Trinocular Stereo Vision. In: Proc. of 18th Int. Conf. on Pattern Recognition, pp. 1076–1079 (2006)