

RESEARCH ARTICLE

Open Access



A method of picking up a folded fabric product by a single-armed robot

Yusuke Moriya¹, Daisuke Tanaka¹, Kimitoshi Yamazaki^{1*} and Keisuke Takeshita²

Abstract

This paper describes a method to pick up a folded cloth product by a single-armed robot. We focus on a problem on picking up a folded cloth, and organize tasks to attack it. Then, we propose methods of grasp position estimation composed of two stages: detection of the thickest folded hem and pose estimation of the cloth product. In addition, we attempt to search for appropriate grasping postures, and show that there are regions where the success rate of grasp was high. In experiments using an actual robot, we achieved a picking task with 92% success rate.

Keywords: A folded cloth, Deformable object recognition and manipulation, A single-armed robot

Introduction

One desired ability for autonomous robots engaging daily assistance is “pick-and-place an object” on a designated place. Among them, object grasping is difficult and important issue. A conventional approach to robotic grasping in daily assistance assumes to cope with rigid objects, and employs geometrical models, and then cooperates model-based recognition with motion planning [1–3]. In this procedure, how to grasp the object is one issue. In many cases, it is assumed that the object is a rigid body, and the point-to-point contact between the robot finger and the object is determined. However, in daily environment, we can find essential tasks that non-rigid objects are needed to manipulate. For instance people use various types of clothing in the course of their daily lives. If robots have an ability to handle a folded cloth, e.g. handing over a towel and putting a shirt in a chest, it might be one of the effective contributions of autonomous robots, especially for handicapped people [4]. When doing grasping of the folded fabric product, it is desirable to grasp a proper position of the fabric product so as not to destroy the original folded shape.

The purpose of this study is to develop a method to pick up a folded cloth item by a single-armed robot. Cloth products are often folded in a rectangular shape

when they are going to go to shelves and dressers. This is a common matter with various types of cloth products. Therefore, we will proceed with the study assuming the situation that the cloth product folded in a rectangle is placed on the horizontal plane. The cloth grasp assumed in this paper is required to be reversible deformation. That is, it is unacceptable that if the original folded shape of the cloth is collapsed when the gripped cloth is placed on a designated place.

The contributions of this paper are as follows:

- We focus on a problem on picking up a folded cloth, and organize tasks to attack it.
- We propose a method to determine the grasping position from a folded cloth product placed on a table. The proposed method consists of two stages: detection of the thickest folded hem and pose estimation of the cloth product.
- To obtain robust grasping, we attempted to search for appropriate grasping postures. As a result, we were able to find regions where the success rate of grasp was high.

The paper is organized as follows: “[Related work](#)” section shows related work, and “[Issues and approach](#)” section explains issues and our approach. “[Grasping position detection](#)” and “[Grasping motion determination](#)” sections explain the proposed method. “[Experiments](#)” section

*Correspondence: kyamazaki@shinshu-u.ac.jp

¹ Faculty of Engineering, Nagano, Nagano, Japan

Full list of author information is available at the end of the article

shows experimental results, and “**Conclusion**” section presents the conclusions of this paper.

Related work

In many previous studies on automatic operation of cloth products, there are a phase to make a suspended state. Osawa et al. [5] showed that the type of cloth product can be determined by repeating the following procedure: a robot holds a cloth product by hanging it with one hand, and then grasps the lower end portion by another hand, and finally hangs the product by the hand. This idea was later referred to by many researchers and contributed to implementation of several cloth product operations such as type discrimination and folding. Willimon et al. [6] introduced the task of picking one gripping point for suspending a single cloth product placed on a table casually. Kita et al. [7] proposed a method of matching the model with a 3D point cloud measured using a trinocular stereo camera using a deformable shape model for the hanging state. Abbeel et al. [8] succeeded in identifying the type of cloth product by a robot observing the contour and the position of the lower end point while operating the cloth product.

There are also studies that aimed at more efficient operation, sophisticated selection of gripping points and introduction of operation methods other than picking and moving. Doumanoglou et al. [9] succeeded in recognizing clothing type and shape using a 3D range camera while unfolding. Their framework also provided a next grasping point. Li et al. [10] proposed a framework for recognizing the categories and the poses of a deformable object. They used RGB-D data, and matched it with garment shape registered in database. Yuba et al. [11] proposed a method for unfolding cloth products placed casually in a few steps by introducing “pinch and slide” proposed by Shibata et al. [12].

In these studies, a robot manipulated a cloth product that was placed in a casual way or was suspended. Of course, they are difficult tasks due to being complex

shape state. However, it is clearly different from the approach we are assuming about grasping cloth products. In the abovementioned studies, they actively changed the shape of the cloth, to obtain information or to transform to the desired shape. On the other hand, the task assumed in this study is to grasp the folded cloth product without collapsing the shape as shown in the right side of Fig. 1. If we cannot select the folded hem properly, we must grasp by clipping multiple cloths together. In this case, it was often occurred in the our preliminary examination that because of difficult task of inserting fingers under the cloth, it was not possible to grasp or the shape of the cloth collapsed even if gripped. Based on the above, we have selected study topics from selecting the parts to be gripped, proposing and demonstrating the solution method.

Issues and approach

Successful grasp definition and issues

A single-armed robot exists in front of a folded cloth product. A parallel jaw gripper, which is a simple and popular equipment for robot manipulators, is attached as the end-effector. A 3D range image sensor is installed to observe robot’s workspace. The purpose is to pick up the cloth product from the table.

First, we define successful grasping state. When a cloth product is folded in a rectangle shape, if we grasp the thickest folded hem that was made when we folded at the end, we are often grasped without collapsing the shape. We can set such fact in various types of clothe products: towel, T-shirt, pants and so on. Therefore, we will proceed with the premise of such a way of folding. Let us assume that the grasping position assumed in this paper is on the middle of the thickest folded hem depicted as a red point in center picture of Fig. 1. If the robot can grasp that part and lifts it without breaking the shape of the cloth, it will be successful. However, if the shape of the cloth is irreversibly deformed after picking up, e.g. when the shape of the cloth collapsed because the place to

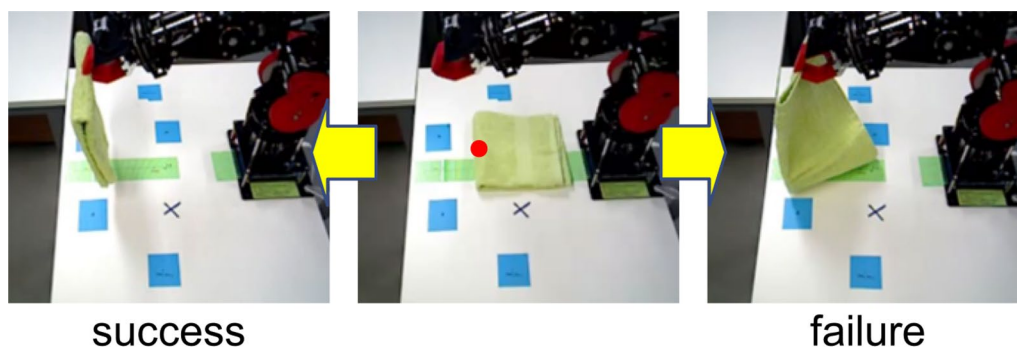


Fig. 1 Picking up a folded cloth product: success and failure cases

grasp was not properly grasped, it becomes failure case. Meanwhile, a robot grasps another points on the cloth, it is also failure.

This problem setting is pretty simple but includes outstanding issues as follows:

- How to detect a grasped position from a folded cloth product: since the shape of the cloth has a certain regularity, it is relatively easy to detect a hem portion as a border. However, it is necessary to verify whether the detected border is a suitable site for lifting without collapsing the shape of the cloth. That is, it is necessary to recognize the state of folding of each hem.
- How to make grasping motion sequence of the robot: cloth products are flexible material, so the success rate of grasping changes depending on how the hand is brought close to and how to grasp. Therefore, consideration should be given not only to the pose at the time of grasping but also how to bring the end-effector closer to the grasping position.

The next subsection introduces our approach to solving them.

Approach for acquiring method of grasping cloth products

The left flowchart in Fig. 2 shows a basic procedure for grasping a folded cloth product. First, the cloth placed on a table is measured by a 3D range image sensor, and a pair of color image and depth image are obtained. Using these images, a grasp position is determined, and then a grasp motion of a robot arm is determined. Finally, the result is performed by the real robot.

For the second and third block of the flowchart, two types of pre-experimented dataset was used, respectively. The first is information on the grasping point, which saves pairs of an instructed grasping point and a depth

image. The other is information for bringing a hand closer to the grasping point. It is composed of a pair of a grasping posture and via posture of the end-effector.

These pre-experimented data are collected in advance: that is, picking up a folded cloth product is performed with an instructed grasping position, and sensor data therebetween is recorded. In the remaining of this paper, we call one data unit (a pair of \mathbf{P} and \mathbf{R}) “task experience data,” and call a dataset consisting of all of the data “task experience dataset,” and a dataset collecting only successful case “successfully experience dataset.”

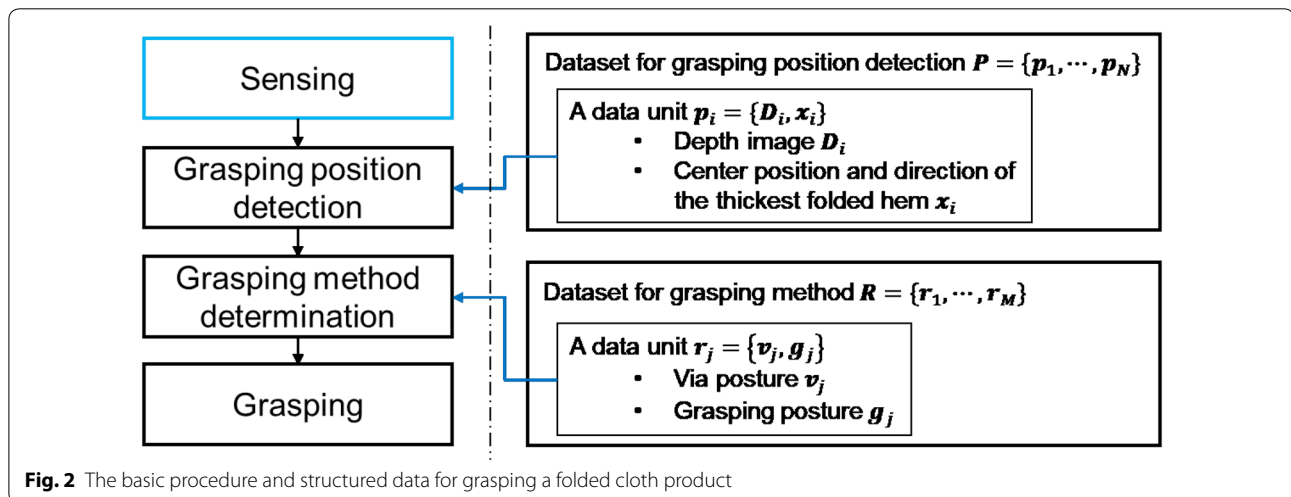
In order to solve the issues mentioned in the previous subsection, the following processing is performed by using experience data. In the following two sections, each of them will be explained in detail.

Grasping position detection

In order to determine a grasping position, it is necessary to recognize how a cloth product is placed and then to find the position to grasp. Recognition of the placed situation is accomplished by performing shape-based registration processing between a task experience data and the current sensor data. On the other hand, grasping point determination is accomplished by detecting visually recognizable borders and counting overlapping of cloth that can be observed there. However, since ambiguity remains, by observing the relationship of the number of the overlapping on a neighboring border, the determination accuracy is improved.

Grasping motion determination

We solve the problem of finding an appropriate posture transition from via posture to grasping posture of the end-effector. Via posture means the preparatory posture of the end-effector just before reaching the grasping posture. In order to obtain an appropriate combination of



these two postures, we take an approach to selecting the posture parameters from advance experiments that perform actual grasping a folded cloth product with various posture parameters.

Grasping position detection

Extraction of the area where a cloth product exists

A color image and a depth image are captured for a cloth product placed on a table. A three-dimensional point cloud is generated from the depth image, and a plane equation of the table top is calculated by plane detection. Here, by estimating the plane parameters to which RANSAC [13] is applied, a plane coincident with the table top is detected without being affected by the existence of the cloth product.

After that, only the three-dimensional points on the sensor side than that plane are selected, and they are projected on a two-dimensional plane. This two-dimensional plane virtually constitutes an image obtained by observing the top plate of the table from vertically above. As a result, we obtain a projected image of the point cloud belonging to the cloth product when looking from directly above.

Orientation estimation of cloth products

In order to find a border including a point to be grasped, the thickest folded hem is detected. For this purpose, a color image taken from obliquely above the cloth product is used. When comparing the thickest folded hem with other hems, a clear difference appears depending on whether there is a gap due to overlapping of cloths. Therefore, edge detection is applied to the obtained color image. Then, a processing focusing on the fact that the number of edges appearing depends on the type of hem is performed.

First, the Canny operator [14] is applied to the color image, and as shown in the top right panel in Fig. 3, gaps between the cloth parts are obtained as edges. On the other hand, as shown in the lower left panel, the contour of the cloth product is obtained. Then the folded cloth product is approximated to a quadrangle as shown in the lower right panel. From this shape, the edge positioned on the camera side is selected, and processing of state estimation of folded parts is performed.

Let $\mathbf{u}_c = (u_c, v_c)$ be image coordinates that belong to quadrangle line segments shown in Fig. 3 (4). L , the number of edge, is calculated as follows.

$$L = \frac{1}{l} \sum_{c=1}^l f_{u_c}(v_c), \quad (1)$$

where l is the number of pixels in the horizontal direction (u direction) in the area where a hem exists. $f_{u_c}(v_c)$ is the

number of edges detected when operating in the vertical direction in the column of the coordinate u_c . That is

$$f_{u_c}(v_c) = \frac{1}{2} \sum_{v=v_{min}+1}^{v_{max}} \{1 - \delta_{I(v)I(v-1)}\}, \quad (2)$$

where I represents the resulted image of canny edge detection, and $I(v)$ is the pixel values on the coordinates $(u_c, v_c + v)$. v_{min} is a negative integer whereas v_{max} is positive, and δ is Kronecker's delta. In Eq. (2), a column of pixels passing through the hem is selected in order. A pixel in the column is compared with another pixel immediately before that, 0 is added if it is the same value, and 1 is added if it is different value. That is, the number of times of crossing the white line is calculated. This process is applied throughout one border, after which the average is calculated by Eq. (1). Fig. 4 visually shows this process. By looking at the number of edges obtained, it is possible to judge whether the border of interest is the thickest folded hem or not. Also, by looking at the average of the number of edges in adjacent hems, it is possible to specify the orientation of the cloth product. These are described in "Experiment" section.

Fine adjustment of position and orientation

Due to the above-mentioned processing, the rough 2-dimensional position of the thickest folded hem is known. Next, in order to accurately obtain the grasping position, additional processing is performed. In this study, as we assume that cloth is folded in a rectangular shape, geometrical fitting of rectangular shape might be one convenient way. However, the depth data from the surface of cloth is affected by the location of the cloth or the existence of wrinkle, and some data might be missing. When geometrical shape fitting is performed on such data, we empirically confirmed that errors were remained particularly in the angular direction. Therefore, pose adjustment by particle filter [15] was adopted. The procedure is as follows.

First, one learning data whose placed direction was similar to the current cloth product is identified and used as reference data. In this identification, each learning data and input data are converted into an image on a view-point looked down from vertically above. Next, a process of collating the reference data with the shape of the input data is performed. If the matching degree between the two data is high, it is assumed that the grasping position recorded in the reference data is mapped on the input data, and the grasping position can be determined.

The reason for preparing dozens of learning data for the cloth is as follows. As the orientation of the cloth changes, the depth data for the cloth will also change.

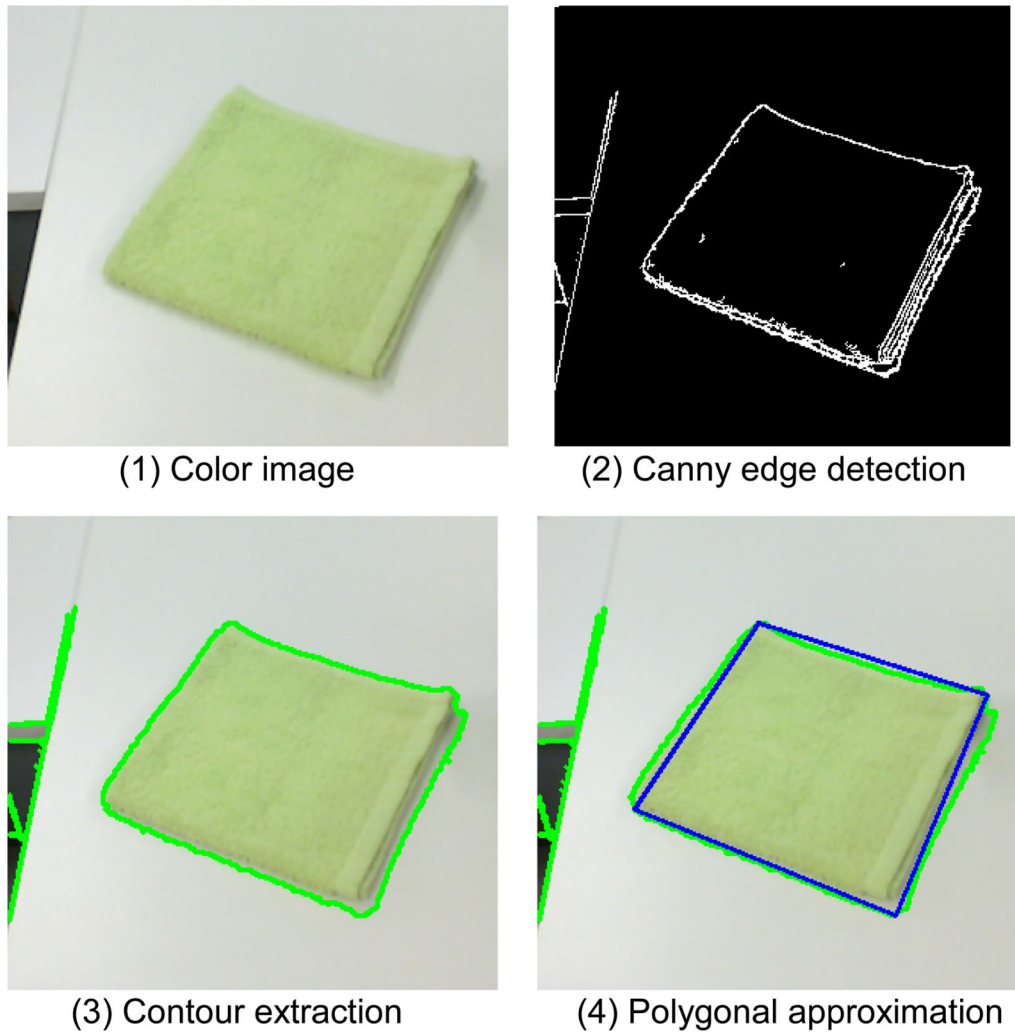


Fig. 3 The procedure of contour extraction

In particular, since the measurement result around hem changes, it directly affects the error of the gripping position. Therefore, we added the selection process to pick up the data that resembled the current placement.

An issue in the pose adjustment procedure is that the shape and inclination of the cloth product in the reference data is not completely the same as the input data. Therefore, in order to overlap the input data well, the reference data is aligned by means of a particle filter. Originally, there are six variables in the posture alignment. However, as described above, if the transformation for directly above viewpoint has been added, the posture variables can be thought of as a total of three degrees of freedom; two parallel movement parameters (x , y) and a rotation parameter θ on the plane.

In the particle filter, a posture of a target object \mathbf{x}_t is estimated from measurements \mathbf{z}_t by external sensors according to the following two equations:

$$\begin{aligned} p(\mathbf{x}_t|Z_{t-1}) &= \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|Z_{t-1})d\mathbf{x}_{t-1}, \\ p(\mathbf{x}_t|Z_t) &\propto p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|Z_{t-1}), \end{aligned} \quad (3)$$

where \mathbf{z}_t indicates a sensor measurement, that is, the perspected transformed image in our case. Z_t is a group of $\mathbf{z}_i^t (i = 1, \dots, n)$ at time t . In Eq. (3), it is a depth value obtained for each coordinate (u , v) on a depth image. The former equation is a prior probability which is calculated before image processing at time t , and the latter is a posterior probability which includes the estimation result. In our approach, a likelihood $p(\mathbf{z}_t|\mathbf{x}_t)$ is calculated by

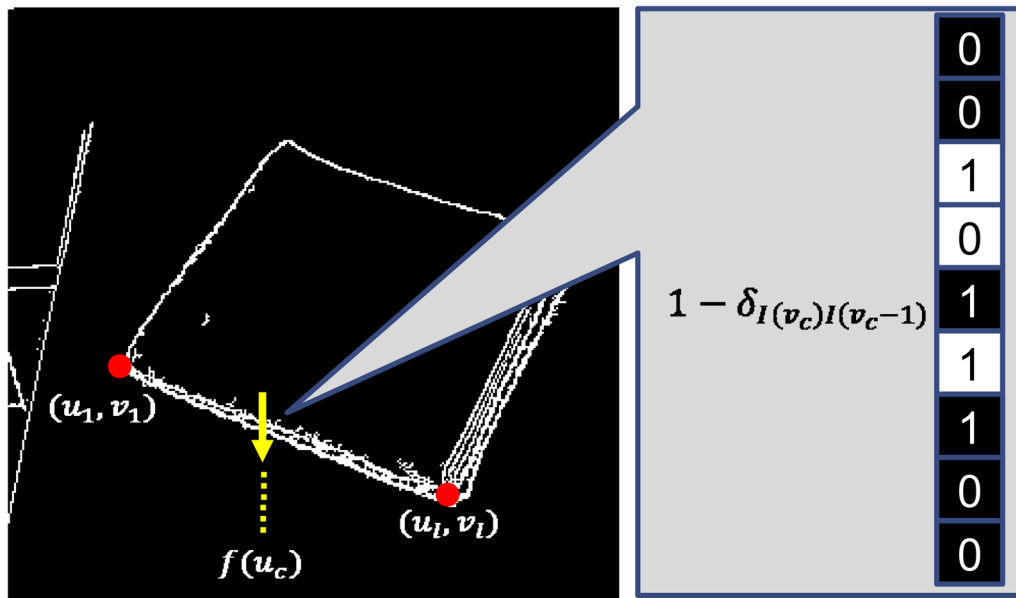


Fig. 4 Counting the number of edges from the result of canny edge detection

comparing 3D points derived from a cloth product. The evaluation equation is as follows:

$$p(\mathbf{z}_t | \mathbf{x}_t) = \sum_d \frac{1}{\{d_{ref}(u', v') - d_{input}(u, v)\}^2 + C}, \quad (4)$$

where (u, v) and (u', v') are image coordinates. (u', v') are a result of transformation using posture parameters (x, y, θ) . $d_*(u, v)$ is the depth value on (u, v) . *ref* indicates a training data for comparizon, and *input* indicates the input data. The subscript *d* of the symbol indicates all the depth values deemed to belong to the fabric product in the depth image. *C* is a constant.

In this equation, the difference from the input data is taken for all three dimensional points after posture conversion of the reference data. The more the many points overlap with small differences, the better the evaluation is obtained.

Grasping motion determination

Concept of determining grasping motion

Since cloth products are a flexible material, deformation of cloth might occur by touching it. That is, even in a grasping operation, it is possible to assume operations such as sliding a finger under the cloth or letting a part of the cloth between fingertips. Thereby there is a possibility that the success rate of grasping can be improved. From the above, it is desirable to take into consideration not only the hand posture of grasping but also posture change in sequential order up to grasping.

Therefore, we take a policy of looking for suitable posture sequence in advance. Instead of manually giving

a grasping posture in a descending manner, we take an approach to repeating trial and error according to various grasping methods. However, there is a big problem with this approach: As the number of target postures increases, the dimension of the parameter space to be searched becomes larger, so that it is not realistic to obtain an appropriate solution. Therefore, we decided to find an appropriate grasping method by limiting the end-effector postures to be searched to two kinds; a via posture and a grasping posture.

A proper grasping motion seems to have a part depending on the shape of end effector. Therefore, in the course of trying grasping, we try to clarify two aspects: elements generally common with two fingered hands and elements dependent on end-effector. If we grasp the former, it can be expected to find appropriate grasping motions through realistic number of trials even when using different end-effectors.

Search for an posture pair

A posture of an end-effector is represented by six parameters $(x, y, z, \phi, \theta, \psi)$. Therefore, it is necessary to consider a combination of a total of 12 variables in order to decide an appropriate grasping motion. However, even if it is simplified so far, the search space is still high-dimensional.

We select grasping postures according to the policy as follows. First, the grasping position (x_g, y_g, z_g) , which is fixed by the method in “[Fine adjustment of position and orientation](#)” section, is set as the center part of the thickest folded hem. Then nine posture variables are defined

as $(x_v, y_v, z_v, \phi_v, \theta_v, \psi_v)$ for a via posture and $(\phi_g, \theta_g, \psi_g)$ for a grasping posture. Next, they are randomly changed within a pre-defined spatial range to grasp the cloth product. Both the combination of variables at the time of success and the combination of variables at the time of failure are recorded, respectively. From the results, we identify the area where successful grasps are concentrated in the posture parameter space, and specify the posture parameters with high importance for stable grasping. Then, by selecting appropriate ranges of values for the posture parameters, set of parameters which are the center of the ranges are set to the via posture/grasping posture.

By this procedure, the number of posture variables to be noticed can be reduced. It is considered that posture variables defined by such procedure are also effective for hands having similar mechanical structure but another fingertip shape. Therefore, when using another two fingered hand, it is sufficient to search for two hand postures in the reduced low dimensional search space.

Experiments

Experimental settings

NEKONOTE 6 DOF for Academic manufactured by RT CORPORATION was used as an experimental robot. As a three-dimensional range image sensor, Xtion PRO LIVE manufactured by ASUS was used, and a web camera (BSW32KM04WH) made by Buffalo Co., Ltd. was also used. A color image and a depth image of the size of 640×480 [pixel] can be acquired from the three-dimensional range image sensor, and a color image of the same size can be acquired from the camera. As shown in Fig. 5, the manipulator was fixed on a

table, and the 3D range image sensor was installed at the point of view where a cloth product and the manipulator can be seen from above. Also, the camera was installed in a position where hems on the near side of the cloth product was easy to see. For cloth products, a rectangular cloth towel, which is 340×340 (mm) size, 35.7 (g) weight, and 1.23 (mm) thickness, was used. In doing grabbing task, we folded this towel in four and put it on the table.

Orientation estimation of cloth products

When the folded cloth product is shot with a camera, the number of hems that can be observed from the camera is one or two. As shown in Fig. 6, the types of observable hem can be classified as follows.

1. The thickest folded hem
2. There is one gap between overlapped cloths.
3. There are two or more gaps between overlapped cloths.

When these images were obtained, it was examined whether the position of the hem to be grasped can be specified by using the number of edges detected from the border.

Figure 7 shows the relationship between the type of hems, the position and the number of edges when only one hem is visible from the camera. Each numerical value is after rounding off. In the table, “position of the hem” shows the rough position of the hem when viewed from the sensor. From this table, when the average number of edges detected from the hem on the near side is about two, there is a high possibility that the hem is the thickest folded hem. This is because the average number of edges will be greater than three if it is another hem. Likewise, even if a hem exists in the front or the lower right corner and the average number of edges is about three, it can also be identified as the thickest folded hem.

On the other hand, Fig. 8 shows the relationship between the type and position of the hem when two edges are visible on the near side. In this case, the position of the visible hem is divided on the right side or the left side from the sensor. Naturally, it does not happen when the thickest folded hems appear on both sides. The same is true for cases that hems only with one gap appears on both sides. From these results, it was found that when the thickest folded hem is visible on the left side, the average number of edges becomes one or two, and when it is visible on the right side, the average number of edges becomes three.

These trends basically depend on the number of gaps due to overlapping of cloths. However, although

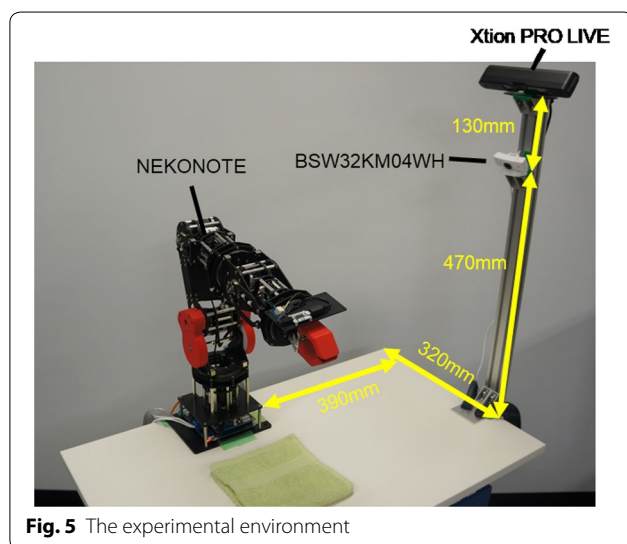


Fig. 5 The experimental environment

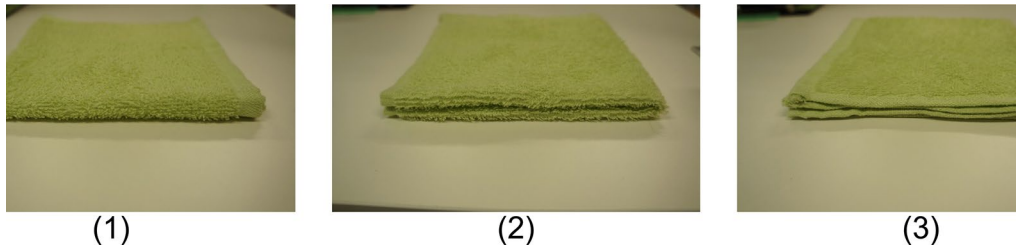


Fig. 6 Three types of hem

		Position of the hem		
		left	center	right
Type of the hem	(1)	2	2 or 3	2 or 3
	(2)	3 or 4	4 or 5	4
	(3)	6	6	4

Fig. 7 The relationship between the type of hems, the position and the number of edges when only one hem is visible from the camera. The number (1)–(3) means three types of hem explained in “Orientation estimation of cloth products” section

		Type of the hem on right side					
		(1)		(2)		(3)	
		Left	Right	Left	Right	Left	Right
Type of the hem on left side	(1)	-	-	1 or 2	4 or 5	2	4 or 5
	(2)	3 or 4	3	-	-	3 or 4	4
	(3)	3 or 4	3	3 or 4	4 or 5	3 or 4	4 or 5

Fig. 8 The relationship between the type and position of the hem when two edges are visible on the near side

the ratio of the number to edges of this experiment has some degree of invariance, the number itself is due to cloth hardness, lighting conditions, etc. It is necessary to clarify these experimentally. Another important point is that if only one side is visible and the hem type is (2) or (3) shown in Fig. 6, the thickest folded hem cannot be specified. Also, when two sides are visible, ambiguity remains between left (3)–right (2) and left (3)–right (3). In such a case, additional method such as actively moving the cloth and observing it again are necessary. However, in cases other than the above, it is possible to specify even if the position of the thickest border is not visible if we use the relationship in Fig. 8.

Figures 9 and 10 show examples of recognition results of hem on the basis of the above. Fig. 9 shows a case where one hem is visible on the near side, and Fig. 10 shows two cases. The right side is the output image of the recognition result, green line is the result of contour extraction, blue line is the result of specifying the hem,

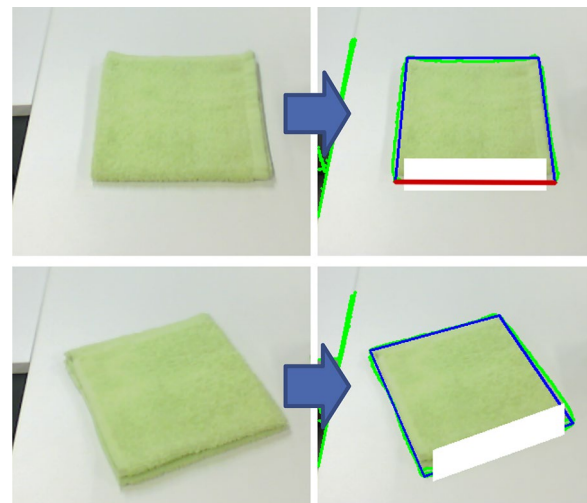


Fig. 9 Results of the thickest hem detection in a case where one hem is visible on the near side

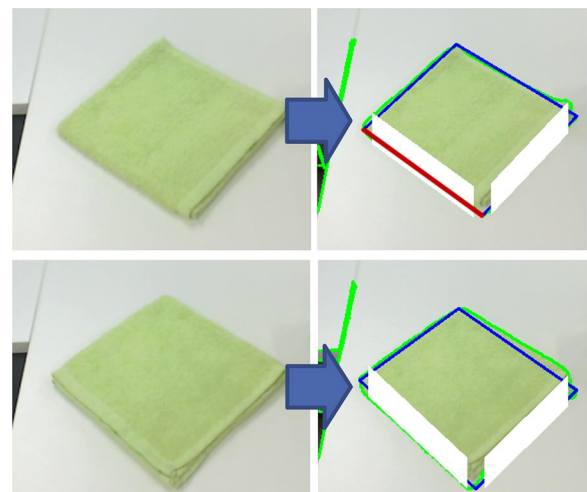


Fig. 10 Results of the thickest hem detection in a case where two hems are visible on the near side

the area painted white is an area used to calculate the number of edges. Results that a hem with red line was recognized as the thickest folded hem.

Fine adjustment of reference data with input data

By the processing mentioned in the previous section, the position of the thickest folded hem is specified. Next, a process of determining the grasping position is performed using particle filter. In this section, the procedure is explained. First, 30 pieces of learning data for grasping position determination were prepared. In collecting this data, a folded cloth product was randomly placed in a region where length 300 mm \times width 500 mm in front of the manipulator. With respect to the direction of the cloth, it was also randomly placed in the range of $-90^\circ < \theta < 90^\circ$, assuming $\theta = 0^\circ$ when the direction of the thickest folded hem is perpendicular to the axis in the front direction of the robot. The number of particles of the particle filter used in the alignment process was set to 250. The standard deviation of particles on prediction process was empirically set to $(x, y, \theta) = (10 \text{ mm}, 10 \text{ mm}, 5^\circ)$.

Examples of alignment using particle filter are shown in the Fig. 11. The red part is input data, the green part is learning data, and the part where the two data overlap is represented by yellow. The orange points represent grasping position candidates. Points without filling are the original grasping position linked to the learning data and another point with filling is the grasping position with respect to the input data newly obtained by the alignment process. As can be seen, the original grasping position was moved near the midpoint of the edge. This result shows that it was possible to determine an appropriate grasping position.

In the positioning process according to the Eq. (3), if processing was performed at all the existing points (10,000–15,000), it takes a long processing time. Therefore, we decided to thin out the points to be compared. We sampled points every n pixels while doing raster scanning, and examined the accuracy of alignment in

each sampling. As a result of reducing the sampling to 1/20, the accuracy of alignment was almost the same as before the thinning. On the other hand, the processing time was greatly reduced from about 11 seconds to about 0.6 seconds.

Parameter search for via posture and grasping posture

In this sub-section, we report experiments that determine the appropriate via posture and grasping posture through actual grasping trials. First, as shown in Fig. 5, a cloth product was placed in a predetermined position in front of a robot. The position of the thickest folded hem was made to be the farthest from the installation position of the robot. That is, in the case where the x axis is forward, the y axis is on the horizontal plane perpendicular to the x axis, and the z axis is upward, the orientation of the hem was parallel to the y axis.

Let (d_x, d_y, d_z) be the via position of the end-effector as seen from the coordinate system of the grasping position, and let $(\alpha_v, \beta_v, \gamma_v)$ be roll-pitch-yaw angles of the via posture and $(\alpha_g, \beta_g, \gamma_g)$ be roll-pitch-yaw angles of the grasping posture, respectively. $\alpha_g = \beta_g = \gamma_g = 0^\circ$ when grasping position is grasped from just above cloth product and the direction of fingertips are parallel to the thickest folded hem. The ranges are limited as follows: $-\pi/4 \leq \{\alpha_v, \beta_v, \alpha_g, \beta_g\} \leq \pi/4$, and $-30 \leq \{d_x, d_y, d_z\} \leq 30 \text{ mm}$ for the position of via posture. On the other hand, as for γ , it was clarified by prior examination that the success rate of grasping drops greatly unless the value is set to a value close to 0. Therefore, $\gamma = 0$ was fixed.

Within the above range, posture parameters of the end-effector were randomly selected according to a uniform distribution, and 100 grasping trials were performed. The method of determining whether or not grasping was

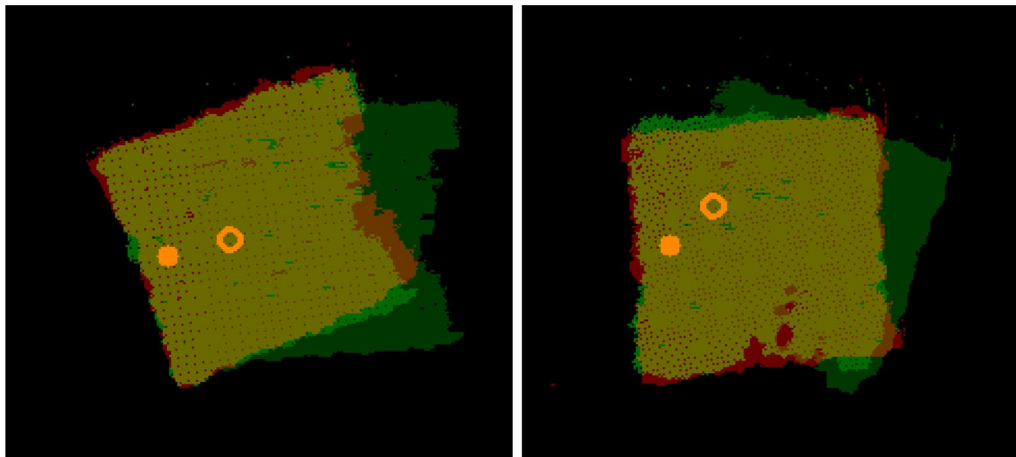


Fig. 11 Examples of fine adjustment for grasp point detection

successful was as described in “[Successful grasp definition and issues](#)” section. That is, if the robot grasped the thickest folded hem and lifted it without collapsing the shape of the cloth, the worker visually recognized and judged it to be successful. Otherwise, it was judged as a failure. The result was that the number of times of grasping succeeded was 50 times and that of failures was 50 times. The purpose of this experiment was to find the range of via/grasping posture that was easy to succeed.

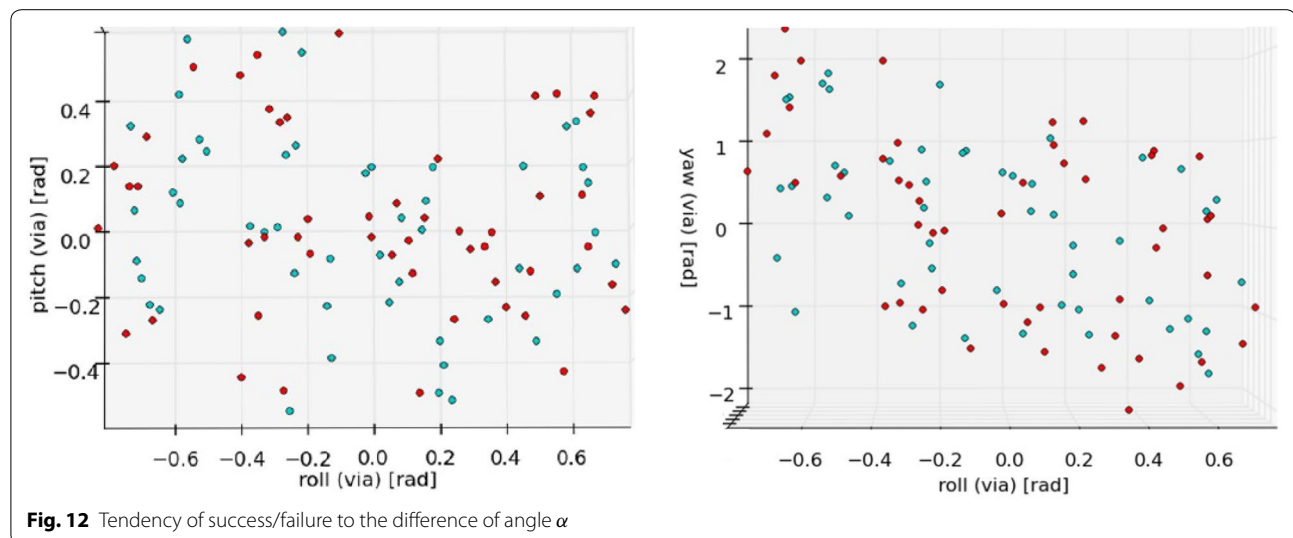
Figure 12 shows two graphs plotting success/failure with the wrist roll angle α as the horizontal axis. The blue dot indicates that the grasping was successful, and the red dot indicates that it failed. From this result, since there is no noticeable trend in the value of α , we decided to always set $\alpha = 0$ in “[Experiment](#)” of the next subsection. On the other hand, Figs. 13 and 14 shows the result of plotting four posture parameters: two positional parameters (d_x, d_z) and pitch angles (β_v, β_g) that were considered to have a large influence on successful grasping. Figure 13 plots the samples when $\beta_v < \beta_g$. A blue color means a success, a red color means a failure sample, a circle mark indicates a via posture, and a triangle mark connected by a line indicates a grasping posture shifted from the via posture. From these graphs, it turns out that successful grasps are concentrated when via posture was started from the area surrounded by green square. This means a movement that puts the fingertip between the cloth product and the desk. On the other hand, Fig. 14 plots samples for $\beta_v > \beta_g$. There were many successes in the square part of the figure. This was a grasping method in which the cloth product was pressed down with the fingertip and then the other fingertip was hooked on a hem. From the above results, it is appropriate to shift

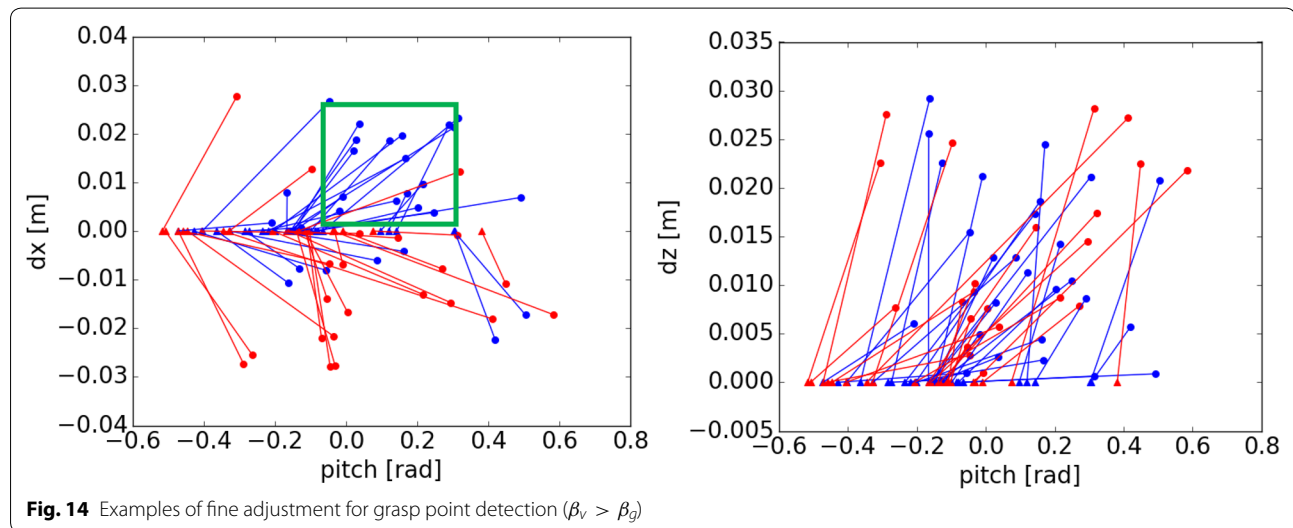
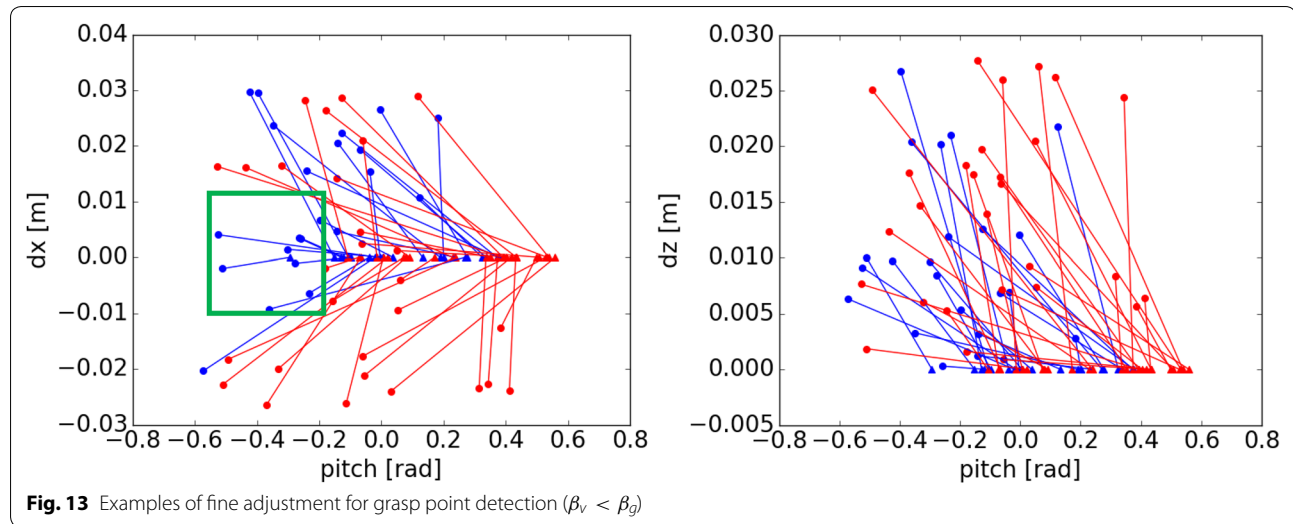
the posture so that the cloth product is pressed down by a fingertip through the back side as viewed from the robot ($\beta_v > \beta_g$ and $dx > 0$).

Experiment with integrated system

Based on the experiments introduced in “[Orientation estimation of cloth products](#)”, “[Fine adjustment of reference data with input data](#)”, “[Parameter search for via posture and grasping posture](#)” sections, a robot system that performs from detection of cloth products to grasping had conducted. With the same placement method as the learning data collection described in “[Fine adjustment of reference data with input data](#)” section, a folded fabric products was randomly placed in front of the robot, and it was investigated whether grasping can be done consistently. It included the detection of the thickest holded hem, determination of the grasping position, and determination of via/grasping posture.

As described in “[Parameter search for via posture and grasping posture](#)” section, end-effector pose where grasping is successful with a high success rate had already been investigated. For a proof experiment explained here, the average value of pose parameters of via/grasping postures (d_x, d_z, β_v and β_g) in the light blue area shown in Fig. 13 were used. That is, the relative via/grasping posture with respect to the cloth product was determined from the average value, and grasping operation was performed based on the inverse kinematics calculation according to the posture estimation result of the cloth product. The cloth product was in a quadrant state as shown in Fig. 1. The result was 46 successes and 4 failures out of 50 trials. All of the cause of the failure was that inverse kinematics of the robot arm could not be solved.





Conclusion

In this paper, we described a method to pick up a folded cloth product by a single-armed robot. We focused on a problem on picking up a folded cloth, and organize tasks to attack it. Then, we proposed methods of grasp position estimation composed of two stages: detection of the thickest folded hem and pose estimation of the cloth product. In addition, we attempted to search for appropriate grasping postures, and found that there are regions where the success rate of grasp was high. In experiments using a real robot, we achieved a picking task with 92% success rate.

As future work, we apply the proposed methods to other types of folded cloth products. It is also needed to perform the same experiment with another single arm

robot. Furthermore, it is desired to improve the proposed method so tha robots grasp even if there are multiple overlapped cloth products.

Authors' contributions

YM and DT implemented the proposed method and carried out actual experiments. KY and KT proposed the method and wrote the paper. All authors read and approved the final manuscript.

Author details

¹ Faculty of Engineering, Nagano, Nagano, Japan. ² Toyota Motor Corp., Toyota, Aichi, Japan.

Acknowledgements

None.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Available.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

No funding.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 19 July 2017 Accepted: 18 December 2017

Published online: 04 January 2018

References

1. Kitahama K, Tsukada K, Galpin F, Matsubara T, Hirano Y (2006) Vision-based scene representation for 3D interaction of service robots. In: Proceedings of IEEE/RSJ international conference on intelligent robots and systems, pp 4756–4761
2. Kuehnle J, Verl A, Xue Z, Ruehl S, Zoellner J, Dillmann R, Grundmann T, Eidenberger R, Zoellner R (2009) 6d object localization and obstacle detection for collision-free manipulation with a mobile service robot. In: Proceedings of international conference on advanced robotics
3. Lee S, Moradib H, Jangc D, Jangd H, Kime E, Lef PM, Seog J, Hanh J (2008) Toward human-like real-time manipulation: from perception to motion planning. *Adv Robot* 22(9):983–1005
4. Hashimoto K, Saito F, Yamamoto T, Ikeda K (2013) A field study of the human support robot in the home environment. In: Proceedings of IEEE workshop on advanced robotics and its social impacts, pp 143–150
5. Osawa F, Seki H, Kamiya Y (2007) Unfolding of massive laundry and classification types by dual manipulator. *J Adv Comput Intell Intell Inf* 11(5):457–463
6. Willimon B, Birchfield S, Walker I (2011) Model for unfolding laundry using interactive perception. In: Proceedings of IEEE international conference on intelligent robots and systems, pp 4871–4876
7. Kita Y, Saito F, Kita N (2004) A deformable model driven method for handling clothes. In: Proceedings of international conference on pattern recognition, vol 4, pp 3889–3895
8. Maitin-Sp Shepard J, et al. (2010) Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. In: 2010 IEEE international conference on robotics and automation (ICRA), pp 2308–2315
9. Doumanoglou A, Kargakos A, Kim T, Malassiotis S (2014) Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning. In: Proceedings of international conference on robotics and automation, pp 987–993
10. Li Y, Chen CF, Allen PK (2014) Recognition of deformable object category and pose. In: Proceedings of IEEE international conference on robotics and automation, pp 5558–5564
11. Yuba H, Arnold S, Yamazaki K (2015) Unfolding of a rectangular cloth based on action selection depending on recognition uncertainty. In: Proceedings of IEEE/SICE international symposium on system integration, pp 623–628
12. Shibata M, Ota T, Endo Y, Hirai S (2008) Handling of hemmed fabrics by a single-armed robot. In: Proceedings of IEEE international conference on automation science and engineering, pp 882–887
13. Fischer MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 24(6):381–395
14. Canny J (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 8:679–714
15. Isard M, Blake A (1998) Condensation—conditional density propagation of visual tracking. *Int J Comput Vision* 29(1):5–28

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)