Unfolding of a Rectangular Cloth Based on Action Selection Depending on Recognition Uncertainty

Hiroyuki YUBA*1, Solvi ARNOLD*1, Kimitoshi YAMAZAKI*1

Abstract—We propose a method for unfolding a rectangular cloth placed on a table in an arbitrary unarranged shape, using a dual arm robot. There are many situations where the manipulation of fabric products by dual arm robots takes time due to operation complexity. Also, observation of fabric products in unarranged shapes can be fraught with uncertainty, posing further difficulties for robotic manipulation. In this article, we addressed these problems for our specific task, implementing a "pinch and slide motion" to address the former issue, and an operation selection mechanism based on partially observable Markov decision process (POMDP) to address the latter. We used this approach to let a robot unfold a rectangular cloth, thereby experimentally verifying the effectiveness of our proposal.

I. INTRODUCTION

The need to handle cloth items is common in a wide variety of settings, from restaurants to hospitals. Manipulation of cloth items is a subject of active research, in the pursuit of automated handling of soft objects [1], [2], [3], [4]. The goal of the present research is fast and efficient manipulation of cloth items by a two-armed robot.

In this paper, we aim to let the robot unfold a rectangular cloth placed on a table. Numerous works have addressed the challenge of unfolding cloth. Cuén-Rochín et al. [5] proposed action selection for manipulating deformable planar objects. By using physical model, a real robot straightened a square-shaped cloth. Kita et al. [6] utilized a 3D deformable model, and obtained a correspondence between the model and an input pointcloud that was captured by a trinocular stereo camera. Abbeel et al. [7] achieved a cloth classification method by shifting an item of clothing from one arm to another. Their methods focused on clothing contours extracted from an image. Recently, Doumanoglu et al. [13] succeeded in recognizing clothing type and shape using a 3D range camera while unfolding. Their framework also provided a next grasping point. Stria et al. [8] achieved to unfold an item of clothing based on shape estimation using the proposed polygonal model.

Above mentioned studies employed a two-armed robot, and proposed a method to unfold a clothing item that was entirely held up. Meanwhile, a different method of unfolding is given in Figure 1. This method was successfully demonstrated in [9] [10]. The robot could hold a rectangular cloth and unfold it in mid-lift, using two-finger hands fitted with semi-spherical fingertips. Drawing from the way humans handle cloth, this method allows for quick and efficient manipulation. In the present work we adopt this method as one behavior primitive, and aim to realize an efficient unfolding operation for our two-armed robot. Below, following the terminology of [9], we will refer to the operation shown in Figure 1 as "pinch and slide unfolding".

In [9], the authors assumed a starting position where both hands already have a grip on the cloth. In this study, we aim to perform the unfold from a more general starting point, namely the situation where the cloth is placed on a table in an arbitrary and unarranged shape. Hence we must first determine suitable grasping points by means of image processing, and then let the robot grasp the cloth. As with other soft objects, accurate recognition of cloth is a hard problem for current day image processing techniques, and unavoidably introduces some degree of uncertainty. In the present paper, we introduce an action-selection mechanism that explicitly takes this uncertainty into account.

The paper is structured as follows: Section 2 explains our method for unfolding a rectangular cloth by a two-armed robot. Section 3 discusses the difficulties encountered when trying to let a robot unfold cloth items, as well as counter-measures to circumvent these difficulties. Section 4 presents our experiments and results in applying the counter-measures from the preceding section, as well as our evaluation of these results. Section 5 concludes the paper and lists directions for future work.

II. UNFOLDING OF A RECTANGULAR CLOTH BY TWO-ARMED ROBOT

A. Movement sequence for cloth unfolding

A rectangular cloth is placed on a table in an arbitrary unarranged shape. First we distinguish the following three cases as illustrated in Fig. 2.

- **State 1**: One corner without bends or wrinkles can be observed.
- **State 2**: Two corners without bends or wrinkles can be observed, in addition to a straight hem connecting them.
- **State 3**: No visible corner or hem.

We define basic steps for unfolding the cloth from these states.

---

*1 Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, Nagano, 380-8553, Japan. kyamazaki@shinshu-u.ac.jp
1) Classify the cloth’s state as one of the three states above.
2) If in step 1) state 1 is observed:
   a) detect hems,
   b) unfold the cloth by the “pinch and slide” motion [9].
3) If in step 1) state 2 is observed, proceed to unfold the cloth directly by grasping and lifting the two visible corners simultaneously.
4) If state 3 is observed, perform a cloth rearrangement manipulation.

Item 3) above is outside the scope of this study, because unfolding from state 2 is easily achieved by picking the cloth up by the two detected corners. Note that state 2 also covers situations where we observe three or more corners without bends or wrinkles. Hence the three states given above cover all possible cloth configurations. Hence we focus on items 1), 2)-a), 2)-b) and 4).

B. Cloth state estimation based on corner and hem detection

Step 1) above employs image-based pattern matching. A 3-D range image sensor is used for this process. An overhead view is acquired by mounting the sensor on the robot’s head, capturing both color and a depth images. These images are used to distinguish the cloth states (1-3) described in the previous section.

To detect corners and hems of an item of cloth, we apply image-based pattern matching to the color image. The procedure is as follows: template images reflecting a corner or hem of the cloth are gathered in advance. Fig. 3 shows examples. The detection process proceeds as follows: First we search the input color image for regions that match well on the template images. For regions showing a good match (meaning the similarity score exceeds a pre-defined threshold), we obtain the corresponding region from the depth image and compute the depth variance within it. Regions with depth variance over a pre-defined threshold are discarded, as these represent corners or hems in an overly bended state. The regions that remain constitute the final detection result.

In our primary experiment, we found that the success rate was over 90 % for hem detection, whereas it was about 70 % for corner detection. To improve the corner detection process, we added the following verification process.

First, a feature vector is calculated for every image region containing a corner of the cloth. Let $D$ be an $m$ by $n$ matrix that holds depth values in each cell. We represent its element by $D_{i,j}$ ($i=1, \ldots, m$, $j=1, \ldots, n$). A value $\Delta_{\text{max}}$, which is an element of a feature vector $\mathbf{V} = [\Delta_{\text{max}1}, \ldots, \Delta_{\text{max}n}]$, is calculated as follows:

$$\Delta_{\text{max}} = \max_j [D_{a,j} - D_{a,j+1}] \quad (a: \text{const})$$

These values are expected to be small when there are no wrinkles and back-folds on the surface of the cloth. We exploit this fact to make a classifier: We obtain several dozen image regions that show either a corner or other area of the cloth, and use this data to generate the relevant classifier using a Support Vector Machine (SVM) [14]. Using this classifier, we obtain a success rate of over 90 %. Fig. 4 shows some detection examples.

C. State recognition of the hem of the cloth

We assume that a robot identifies State 1 accurately and picks up one corner of the cloth. Repeatedly lifting the cloth by hand, we observed that the shape of a cloth held up by one corner can be classified into three possible states, as illustrated in Fig. 5 (here we assume that the corner is held by the right hand).

(a) A hem runs straight down,
(b) A hem runs diagonally down and left from the grasping point, or
(c) No hem can be observed from the robot’s viewpoint.

Fig. 5 shows images captured from the anterior view. We see a hem running approximately straight down. Classification of this situation as one of the above cases ((a) - (c)) proceeds in two stages, as detailed below.
(i) Preliminary discrimination using high-frequency edge detection: First, a Laplacian filter is applied to the input image. Fig. 8 shows an example. We can observe the following characteristics.

- In (a), the hem, located in the center of the region occupied by the cloth, acts as a boundary between areas of strongly differing brightness.
- In (b), high brightness pixels are seen at left-most side of the region occupied by the cloth.
- In (c), brightness values are high globally.

To distinguish these cases, a raster scan is applied with a search window of size \( n[\text{pixel}] \times n[\text{pixel}] \), and the brightness value average is calculated. If several regions with averaged brightness values over a pre-defined threshold \( T_a \) are found at consecutive rows, they are regarded as part of a straight hem, and hence we categorized this state as case (a). The equation is as follows:

\[
M_{i,j} - M_{i,j+1} > T_a,
\]

(2)

where \( M \) is a matrix whose element shows an average value. Here \( M_{i,j} \) indicates the element at the \( i \)th row and the \( j \)th column.

(ii) Refinement of classification result by means of low-frequency edge detection: Depending on the lighting conditions, the method described above may fail to adequately detect the visual edge produced by the hem, as illustrated in Fig. 10. Therefore, we refine the classification result obtained with the method above using a method for detecting gradual brightness differences. First, a two-dimensional Gabor filter is applied to the input image.

Gabor filter has several parameters; frequency, phase, kernel size, and direction. While the former three parameters are fixed, the direction range is changed over \( 0\,\text{[deg]} \) to \( 170\,\text{[deg]} \) in increments of \( 10 \). Next, each of the resulting images is divided into \( n[\text{pixel}] \times n[\text{pixel}] \) meshes. We let \( G_i(p;q) \) denote the area \((p;q)\) in the mesh cut of image \( G_i \). Next, we let \( N_i(p;q) \) denote the number of pixels in \( G_i(p;q) \) exceeding a given threshold value \( \tau \). By computing this value for every \( i \), we obtain the histogram \( H_{p,q} \) as follows:

\[
H_{p,q} = [N_0(p,q), \ldots, N_{17}(p,q)]
\]

(4)

As illustrated in Fig.10, the center hem line often aligns in parallel with the \( y \) axis. If we take this angle as 90 degrees, we expect to see high element values in \( H_{p,q} \) near \( N_8(p,q) \) and \( N_9(p,q) \). Based on this expectation, we again construct a classifier by means of an SVM. As illustrated in Fig.11, this classifier distinguishes, on basis of the mesh cut of \( H_{p,q} \), whether or not there is a hem edge present in a specific region underneath the hand grasping the cloth corner.

Finally, the results of processes (i) and (ii) are integrated. First, the classification result ((a) - (c)) of process (i) is obtained. Meanwhile process (ii), separately, determines whether case (a) applies or not. If multiple regions deemed to
likely contain part of the hem align vertically, we determine that a hem is indeed present. If a positive for case (a) is obtained by process (ii), then this result replaces the result of process (i). Otherwise, the result produced by process (i) is used as the final output.

D. Pinch and slide motion

We describe the “pinch and slide” motion performed after the cloth’s state is identified as state (a) or (b). Here we assume that a corner of the cloth has been picked up by the right hand. The robot, as shown in Fig. 6, grasps a part of hem by approaching from the front side for state (a) or from the right side for state (b).

Let \( O \) be a right-handed coordinate system configured in the center of the waist of the robot. \( x \) axis is the anterior direction and \( z \) axis is the upward direction. As shown in Fig. 7, initially the \( x \) and \( z \) coordinates of both fingertips are aligned [(1), (2)], and then the right hand is pulled in the \(-y\) direction in a straight path [(3)]. This is the “slide” part of the motion. Next, both hands are raised up, and the motion is finalized.

In the case that cloth’s state is identified as state (c), the above procedure is postponed. Instead, a motion to rearrange the cloth’s shape is chosen as the next motion. Section IV explains this case in detail.

E. The cloth rearrangement action

When state 3 is detected, it should be transformed into state 1 or state 2. First, a robot picks up the cloth with grasping it at two arbitrary points. This should hopefully produce a shape with at least one visible corner. Next, assuming a right-handed coordinate system, the following action is performed:

1. The bottom of the cloth is brought into contact with the table by moving both hands in the \(-\Delta z\) direction.
2. The lower part of the cloth is turned to face the robot by moving both hands quickly in the \(\Delta x\) direction.
3. The cloth is placed back on the table.

An example of this motion is shown in Fig. 12.

Here we explain why the second motion should be performed quickly. The quick motion makes it possible to selectivity move only the upper part of the cloth. The bottom part, in many cases a corner of a cloth, move only by a limited amount. This often results in presence of one graspable corner in front of the robot. Then, either of the basic actions 2) or 3) mentioned in Section II-A is applied to unfold the cloth.

III. ISSUES AND SOLUTION IN UNFOLDING ACTION PLANNING

In the previous section, we explained our method for unfolding cloth from various initial states of the cloth. In the present section, we discuss problems encountered with this approach. From our experiments, we learned that the cloth state classifier we introduced above identifies state 1 and 2 accuracies of 0.88 and 0.86, respectively. If we perform action selection on basis of these simple conditional rules alone, we are bound to see mistakes and failed actions. In order to improve on this point, our motion planning method should take the uncertainty in the classification process into account. To this end, we implement action selection as a Partially Observable Markov Decision Process (POMDP).

POMDPs have been used for action selection in robot control. For example, a study on interaction control for appropriate action determination of an agent can be found in [11], and a study on how to separate one item of cloth from a pile in [12]. A particularly impressive study is presented in [13], on action selection for unfolding cloth products with a dual arm robot. This study employed a 3-D range sensor and takes uncertainty into account. Following these studies, we adopt POMDP-based operation selection. Our POMDP consists of the following elements: States: \( S \), Actions: \( A \), Observations: \( O \), State transition probabilities: \( T \), Observation reliability: \( P \), Rewards: \( R \) Discount factor: \( \gamma \), Initial belief distribution: \( b_0(s) \). We first define \( S, A \), and \( O \):

- State \( S \) ... State of the cloth: \( s_1, s_2 \)
- Action \( A \) ... The next action: \( a_1, a_2, a_3 \)
- Observation \( O \) ... State of the cloth as determined by the classifier: \( o_1, o_2 \)

Here States \( s_1, s_2 \) refer to states 1 and 2 as explained in section 2, and observations \( a_1 \) and \( a_2 \) correspond to classifications of a state as \( s_1 \) or \( s_2 \), respectively, by the classifier. Actions \( a_1 \) and \( a_2 \) correspond to steps 2 through 4 of the unfolding sequence shown in section 2. In section 2 we classified the state of the cloth on the table into three cases. In preliminary experiments we found that state 3 was recognized near-flawlessly. Hence we decided to select motion \( a_3 \) directly whenever the classifier reports state 3, and considered only the remaining two states in constructing the POMDP system. The probability of transitioning to state \( S' \) when performing action \( A \) in state \( S \), i.e. transition probability \( T(s'|s,a) \), is defined as follows. In order to move to the unfolding operation and conclude the present task instance, for \( a_1 \) and \( a_2 \) we let

\[
T(s'|s,a) = b_0(s)(i = 1, 2)
\]

As for the initial belief distribution \( b_0(s) \), following the assumption that the initial state is 'placed on the table in arbitrary and unarranged shape', we let \( b_0(s_1) = b_0(s_2) = \ldots \)
0.5. The transition probabilities for action $a_3$ were determined experimentally. The results are shown in I. For the observation reliability $P$, we use the previously stated values. As for the reward $R$, we adopt the following policy: For a correct action, we give a positive reward. For an erroneous action we give a large negative reward. Lastly, for action $a_3$, rearranging the cloth, we give a small negative reward. For discount factor $\gamma$ we experimentally determined a value of 0.7 to be appropriate.

For the above settings, we computed optimal values for the parameters relating belief to action choice, using Point Based Value Iteration. The resulting optimal action selection policy is as follows:

- **Condition 1:** If $b(S_1) > 0.5912$ then perform action $a_1$
- **Condition 2:** If $b(S_1) < 0.3892$ then perform action $a_2$
- **Condition 3:** If $0.3892 \leq b(S_1) \leq 0.5912$ then perform action $a_3$

The process flow of our POMDP system is shown in Fig.13. The robot observes the cloth on the table and picks its next action on basis of conditions 1 3 from the policy above. In the case condition 3 applies, the robot selects action $a_3$ and rearranges the cloth, after which it receives a new observation $o'$ and updates its belief distribution $b(s)$ to $b'(s')$ using the equation below (6):

$$b'(s') = \frac{P(o'|s') \sum_s T(s'|s, a)b(s)}{\sum_{s'} P(o'|s') \sum_s T(s'|s, a)b(s)}$$

(6)

It then selects the next optimal action on basis of the updated belief distribution. In this article, we assume that the task is finished once either action $a_1$ or $a_2$ has been performed, so upon completion of either of these actions we re-initialize the belief distribution, and proceed to the next trial. In the case condition 1 applies, the robot proceeds to perform the ‘pinch and slide’ sequence explained before, and upon completion reinitializes the belief distribution $b_0(s)$. In case condition 2 applies we have two visible corners with a connecting edge between them, so the robot simply picks up the cloth by those two corners, and upon completion reinitializes the belief distribution $b_0(s)$. In addition to the above, if the classifier has identified the current state as neither $s_1, s_2$, then the robot follows the same process as when condition 3 applies.

**IV. EXPERIMENT**

**A. Experimental procedure**

Fig.14 shows our experimental setup. We used a HIRO robot (Kawada industries, Inc). This robot runs the QNX OS and is equipped with an ATOM N270 CPU (1.6 GHz). It has a total DOF of 15 (arm:6 × 2, neck:2, waist:1). Its hands each have 2 fingers with each 2 joints. The hands use RS301CR servos (Futaba corporation). We mounted a Microsoft Kinect on HIRO’s head to obtain RGB and depth images of the state of the cloth. We also equipped the robot’s torso with high resolution camera (VLG-22C, Baumer) to obtain detailed RGB images to guide the pinch and slide motion (see chapter I-D).

Each experimental trial was carried out as follows:
1) Observation of the state of cloth on a table.
2) Condition checking. If condition 3 applies, the robot repeats action $a_3$ until either condition 1 or 2 applies.
3) If condition 1 or 2 applies, the robot proceeds to unfold the cloth using action $a_1$ or $a_2$.
4) Assessment of the success or failure of this trial.

A trial is deemed successful if the robot holds a hem line of the cloth in mid-air without a wrinkle or fold. Fig.15 shows an example. We carried out the experiment using a cloth. (Size: 335mm by 340mm, Weight: 34.8g, Thickness: 1.74mm)

Performing the experiment requires decisions about arm postures assumed when performing action $a_2$ or $a_3$. However in this article, we do not consider these decision processes in detail. Hence, human operation was used to bring the robot’s arms to the relevant cloth corners in the case of action $a_2$ (after a robot observed the corner position), and to achieve proper cloth posture for holding the cloth in case of action

**TABLE I**

<table>
<thead>
<tr>
<th>State transition probabilities when rearranging the cloth</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T(s'</td>
</tr>
<tr>
<td>$s_1$</td>
</tr>
<tr>
<td>$s_1'$</td>
</tr>
<tr>
<td>$s_2'$</td>
</tr>
</tbody>
</table>

![Fig. 13. Process flow of POMDP system](image)

![Fig. 14. Experiment environment](image)
In this experiment, selection of pinch and slide motion when the cloth was in state (c) was considered an unfolding failure, and trials were aborted accordingly.

B. Result

We performed 40 trials of the experiment. The number of successful trials was 28 times. Causes of the observed failures were as follows: misrecognition of hem state of the partially lifted cloth, i.e. mistaking state (a) for (b) or vice versa (6 times), and slipping of the cloth from between the robot’s fingers during the pinch and slide motion (6 times).

C. Discussion

In 40 trials, we observed only one failure of our classifier to recognize the initial state of the cloth (mistaking $s_1$ for the state with no corners). In all other trials we observed correct judgements. State $s_2$, from which robot can unfold the cloth with relative ease, occurred 5 times. All occurrences were recognized correctly by our classifier, which contributed to the success rate. It can be said that in combination with the POMDP, the classifier contributes to more accurate action selection. In this experiment, we did not perform any quantitative evaluation of the POMDP. This evaluation is left as future work.

As noted, failures were caused by confusion of the hem states ((a) and (b)) and slippage of the cloth during the pinch and slide motion. One of the causes of the former was that the threshold $T_a$, $T_b$ of the Laplacian filter was set to a fixed value. This led to inconsistent results across the lighting conditions of different parts of the day, which likely has affected identification rates. Determination processing by means of Gabor filter and SVM also produced misclassifications. We used only a limited area in the image captured by the torso camera as target for image processing, and in some cases the hem line under the robot’s hand fell outside this target area. We aim to improve robustness to varying lighting conditions, and explore more specific classification schemes for distinguishing hem line shapes (along with suitable action-selection schemes). Slipping of the cloth from between the fingers during the pinch and slide motion also caused failures. Presently, after the robot grasps a corner of the cloth, it simply moves the hand to a fixed position, closes the fingers of the other hand around the hem and pulls the cloth along a straight line, without using any sensor feedback. In other words, there is presently no processing of tracking of the hem line or recognition of the cloth’s state while the robot is performing the slide motion. We believe that this cause of failures could be eliminated by means of tactile sensors and such, and aim to pursue this strategy in future work.

V. CONCLUSION

We proposed a procedure for unfolding a rectangular cloth placed on a table in an arbitrary unarranged shape. Given the pinch and slide motion as one of the choices of action, the robot efficiently unfolded the cloth. We implemented a system for accurate action selection using the POMDP framework. We experimentally verified the robot’s ability to unfold the cloth on basis of the observed state and the belief distribution, and we discussed the results.

Directions for future work are improvement of the success rate of recognition of the state of the hem line after lifting a corner, and improvement of the pinch and slide motion by means of sensor feedback, among others.

ACKNOWLEDGEMENT

This work was partly supported by JSPS KAKENHI Grant Number 26700024.

REFERENCES