A Cloth Detection Method based on Image Wrinkle Feature for Daily Assistive Robots

Kimitoshi Yamazaki IRT Research Initiative The Univ. of Tokyo Tokyo, Japan yamazaki@jsk.t.u-tokyo.ac.jp Masayuki Inaba Grad. School of Info. Science and Tech. The Univ. of Tokyo Tokyo, Japan inaba@jsk.t.u-tokyo.ac.jp

Abstract—This paper describes an image based clothes detection method. Our method enables daily assistive robots to find clothes which are readily placed on real environments. No prior knowledge related to shape model or color information is needed because wrinkles on a cloth are extracted as image feature. The feature is constructed through SVM learning by using training data which is the result of Gabor filtering. Clothes extraction based on the feature is combined with Graph Cut, image regions are divided into cloth regions and background. Effectiveness of the proposed method was illustrated through tidying clothes up by a real robot in living room.

I. INTRODUCTION

A lot of clothes exist in daily environment. One of the effective daily assistances by robots will be achieved if the robots can manipulate the clothes because human often has many tedious houseworks on it. We aim to develop recognition modules for a daily assistive robot which can tidy washed clothes up like a housekeeper.

In the viewpoint of robotics researches, vision application for robots have been developed [4][5][7][8][10]. However, these results were mainly focused on industrial robots which were emplaced on factories, so that manipulation skills were often researched under the condition of strictly constraints depending on environments or manipulation targets. On the other hand, daily assistive robots must do its tasks with moving around real environments. Therefore, finding clothes from the environment are also an important issue along with soft object manipulation and so on.

Although recognition methods can be applied if a shape model of a target object is given, it is difficult for us to give an accurate shape model in the case of soft objects. From this fact, we propose a novel feature description for detecting clothes which are placed on real environments.

Our method consists of a classifier based on image feature which reacts wrinkles of a cloth. Gabor filter is utilized to extract such features by generating 20-dimension feature vectors, and the vectors are applied to SVM (Support Vector Machine). This procedure enables to find wrinkles on the cloth from an image which is captured in daily environment. Moreover, graph cut is also used to extract regions having less wrinkle feature. In this process, extracted wrinkle regions are used as seed.

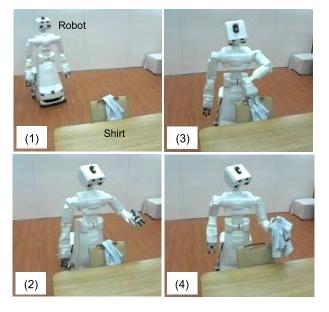


Fig. 1. A daily assistive robot tidying a cloth up

II. ISSUES AND APPROACH

A. The difference from previous works

The purpose of this research is to recognize clothes which are readily placed on daily environment as shown in Fig.1. The image shows a shirt placed on the back of a chair.

Some related works coping with clothes had been focused on soft object simulation[1]. For instance, modeling fabric structure and light refrection, trial fitting system was proposed[9]. Although these researches coped with a forward problem which could give a proper knowledge of the clothes, it is difficult for our case to adapt such approach

Image feature points are a good choice to recognize objects in real world. Several efficient methods such as SIFT had been proposed [6]. However, these features can only be used to textured solid object in many cases: a certain level of viewpoint change is permitted, but the change of 3D shape is not fundamentally assumed. This fact has no application in the case of soft objects.

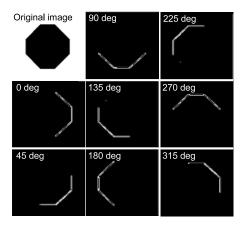


Fig. 2. Gabor filtering regarding 8 directions

B. Approach

Vision applications for present daily assistive robots should satisfy several conditions, for instance, the image processing works on relatively low resolution images, or less strictly knowledge is given in advance.

We aim to develop a method which points out wrinkles on a cloth. Wrinkles are often observed on every clothes, image features extracted from them can have wide recognition application. No solid shape model do not need in this approach.

By the way, combining the result of the feature with 3D position information obtained by a stereo camera enables the robot to pick up a target cloth as shown in Fig.1. After that, it will be assumed that the robot manipulates the cloth with recognizing its 3d shape. So this research has the role of first step to manipulate a soft object by daily assistive robots.

III. WRINKLE FEATURE EXTRACTION

A. Wrinkle Emphasis through Gabor Filtering

Because a cloth is a soft object, its 3D shape variously changes. Wrinkles on it also have various appearances. For feature extraction, we focus on the fact that image contrast of wrinkle region shows gradually changes on frequency domain. In other words, a cloth has strip-shaped states because of the soft body.

Through experiments of applying several image filters to captured cloth images, we selected gabor filter to emphasize wrinkle regions. This filter reacts to waves with particular frequency and direction.

One of the representations of 2D Gabor function is as follows:

$$F(\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma_x \sigma_y} e^a \sin(2\pi f_x (x - u_x) + 2\pi f_y (y - u_y)),$$
 (1)

where f_x and f_y indicate frequency domain. x and y are pixel coordinates, while u_x and u_y are the center of gaussian. The variable a is calculated as follows:

$$a = -\frac{1}{2} \left(\frac{(x - u_x)^2}{\sigma_x^2} + \frac{(y - u_y)^2}{\sigma_y^2} \right).$$
 (2)

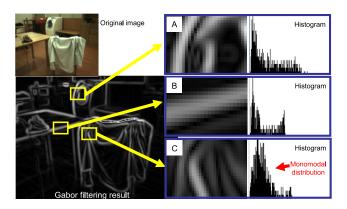


Fig. 3. Wrinkle feature characteristics

To design the filter to have strong reaction at wrinkles on a cloth, we take an approach to perform the filtering with changing f_x and f_y . The parameters are used to generate arbitrary direction as shown in Fig.2.

Left side of Fig.3 shows an example of gabor filtering. 320×240 sized image was used and the variance was set $\sigma=3$. In this case, 8 filters which were equally divided about its direction were defined, and all of the filtering results were summed. Strong reaction are steadily observed from wrinkles on the cloth.

B. Wrinkle feature leaning

Because wrinkles on a cloth can have various 3D shapes, its image features should have invariance to the shape. Although gabor filter emphasizes the image region considered as wrinkles, some other regions having similar frequency may also strongly react. Right side of Fig.3 shows examples in real environment. Because some edges extracted from background (A and B in Fig.3 shows the edges of a washer machine and a table) change its contrast gradually and indicate similar results as wrinkles on a cloth. However, the feature extracted these parts becomes a single edge in many cases. In other words, a true wrinkle has other wrinkles in its neighbor. From this fact, we take an approach to clip relatively large image region in each evaluation, and give good evaluation when the region has a number of wrinkle features.

Coping with various types of clothes, supervised learning from dozens of wrinkle features is performed by means of SVM (Support Vector Machine). Specifically, image regions of wrinkles are extracted from several images by manual, gabor filtering is adopted to the regions, and intensity histograms are calculated from the results of the filtering. The histograms are configured by 20 bins into which are equally divided.

Learning result is generated from these 20-dimension training data by means of following equation:

$$L(\mathbf{w}, h, \alpha) = \frac{1}{2} |\mathbf{w}|^2 - \sum_{i=1}^n \alpha_i \{ t_i(\mathbf{w}_i^t \mathbf{x}_i - h) \}, \quad (3)$$

where $L(\cdot)$ is lagrange function, \mathbf{w} and h are parameters of discriminant function. \mathbf{x}_i indicates ith training data.



Fig. 4. Cloth segmentation by using graph cut

Greatest descent method is utilized to calculate \mathbf{w} and h. These parameters are applied to following equation, cloth regions are distinguished.

$$y = \mathbf{w}^t \mathbf{x}_j - h,\tag{4}$$

where \mathbf{x}_j indicates the feature vector of jth extracted region.

Through this distinction, pixels are divided to 3 series of regions as follows: (1) highly reliable cloth regions which indicate good evaluation, (2) regions which indicate poor evaluation, (3) regions which are hard to decide whether background or not. By using graph cut described in next section, regions of (3) are segmented from positive or negative seeds which belong (1) or (2) regions.

In our implementation, each pixels are evaluated through gabor filtering by masking 15×15 pixels window.

C. Region expansion by means of graph cut

Using parameters \mathbf{w} and h described above, wrinkle extraction is achieved. However, because this approach enables to extract only a part of wrinkle region, other cloth parts will be discarded. To acquire more accurate cloth region, we utilize graph cut[2] by using the result of wrinkle extraction.

Graph cut is one of popular methods to segment image region. Fig.4 shows an example. Left image shows an original image. Center image shows interactively selected regions by manual. A white curved line indicates a part of foreground region, and black one is a part of background region. As shown in the right side image, only a part of cloth is extracted.

The segmentation by graph cut is performed to minimize following cost function:

$$E(X) = \sum_{v \in V} g_v(X_v) + \sum_{(u,v) \in E} h_{uv}(X_u, X_v), \quad (5)$$

where V indicates pixels. u and v means pixel coordinates, and X is a label. E indicates the group of edges which divide image region.

Data term g_v is represented as follows:

$$g_{v(p)}(l) = -\log \theta(I(p), l) \tag{6}$$

where I(p) means the color intensity of pixel p in image I. $\theta(c,l)$ takes values of 0 or 1. The 0 represents foreground and the 1 represents background.

Smoothing term h_v are set as follows:

$$h_{v(p)v(q)}(l, l') = \begin{cases} 0 & (l = l') \\ \frac{\lambda \exp(kI(p) - I(q))^2}{dist(p, q)} & (l \neq l'), \end{cases}$$
(7)

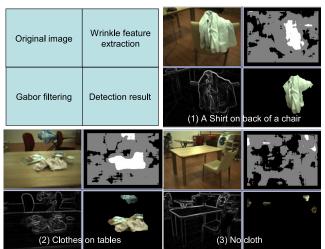


Fig. 5. Experiments of cloth extraction

where λ and k are positive constant. Obeying data term $g(\cdot)$ and smoothing term $h(\cdot)$, image pixels are segmented to minimize the cost of equation (4).

Because many of clothes have simple textures on its surface, and such clothes will provide pixels almost same color elements. This approach effectively works in the combination of wrinkle feature extraction and graph cut.

IV. EXPERIMENTS

A. Wrinkle feature leaning

Using dozens of images captured in real environment, learning of wrinkle feature was performed. "Original image" in Fig.5 shows examples of training image.

40 positive training data were cut out from wrinkle regions, and 40 negative data were also extracted from other than cloth regions. Histograms were calculated about each region after gabor filtering.

For constructing a good classifier, we particularly selected negative regions from background where strong reaction by gabor filtering were observed. This effort had benefit to reduce recognition error.

B. Cloth extraction

Fig.5 shows three kinds of experimental result. Upper right example shows a shirt on back of a chair. Firm result was acquired from wrinkles. After that, graph cut also played a role for enlarging accurate cloth region. If no cloth existed in the environment as shown in lower right example, almost no cloth candidates was provided by our classifier. On the other hand, lower left images show an example of miss-detection. In our experience, this approach still had several issues when highly textured objects were observed near cameras. Because complex edges were extracted from such images, these edges were misunderstood as wrinkles.

Several countermeasures can be considered, for instance, other type of image features such as SIFT can be utilized to recognize solid objects having highly texture. After leaving solid objects by the recognition result, clothes

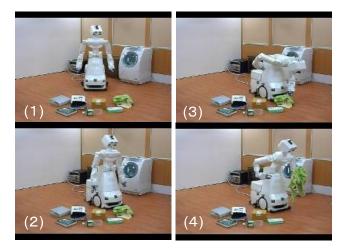


Fig. 6. Picking a cloth up by a robot

can be extracted from remained image region. Because daily assistive robots should recognize and manipulate many series of object in real environment, a number of recognizers are needed in either case.

C. Apply to a daily assistive task by a robot

A task of tidying clothes up was tried by using a real robot. Fig.6 shows a daily assistive robot in our use. The upper body was very similar to human, and a stereo camera was equipped on the head. Clothes were readily placed on floor and back of a chair. The purpose of this task was to find and pick up clothes, and to put them into a washing machine.

In this task, proposed method was applied to two situations, the one was to find a cloth on the floor or on the chair, another was to check whether the picking had succeeded or not. In the finding phase, the stereo camera was also utilized to detect 3d position of the cloth. The robot reached out its hand to the measured position. In the checking phase, the robot captured an image with turning its gaze on hand which grasped the cloth, and observed wrinkle features from the image. If sufficient features could not find, the robot retried the picking task.

Fig.6 shows an experiment of picking up a cloth on the floor. Several objects, for instance a paper sack or a cardboard box, were placed on the floor, the robot selected a cloth and picked it up. In this experiments, the results of wrinkle extraction and stereo reconstruction were merged, the robot detected its grasping target by selecting the result which had maximum 3d volume. Wrinkle features could be strongly extracted from the real cloth as shown Fig.7.

V. CONCLUSION

We proposed a cloth detection method from an image. No prior knowledge related to shape model or color information is needed because wrinkles on a cloth are extracted as image feature. We arranged the method to a task of a daily assistive robot which works in real environment, and showed the effectiveness of our method.

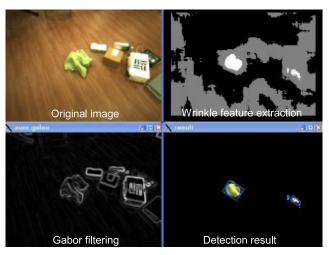


Fig. 7. Detected cloth on floor

Future works, more reliable cloth extraction should be achieved by analysing wrinkle features. Moreover, by merging other types of image features, we would like to construct robust object recognition system for daily assistive robots.

ACKNOWLEDGEMENT

This research is supported by Special Coordination Funds for Promoting Science and Technology, "IRT Foundation to Support Man and Aging Society".

REFERENCES

- D. Baraff and A. Witkin: "Large Steps in Cloth Simulation," Computer Graphics, SIGGRAPH Conf. Proc., pp.43–54, 1998.
- [2] Y. Boykov and V. Kolmogorov: "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," IEEE Trans. Pattern Analysis and Machine Intelligence, 26(9), pp.1124–1137, 2004.
- [3] "Tutorial on Gabor Filters," http://mplab.ucsd.edu/tutorials/tutorials.html
- [4] M. Inaba, H. Inoue: "Rope Handling by a Robot with Visual Feedback," Advanced Robotics, Vol. 2, No. 1, pp.39-54, 1987.
- [5] Y. Kita, F. Saito and N. Kita: "A deformable model driven method for handling clothes," Proc. of Int. Conf. on Pattern Recognition, 2004.
- [6] D. G. Lowe: "Distinctive image features from scale-invariant keypoints," Int'l Journal of Computer Vision, 60, 2, pp. 91-110, 2004.
- [7] E. Ono, H. Okabe, H. Ichijo and N. Aisaka: "Robot Hand with Sensor for Cloth Handling," Proc. 1990, Japan-U.S.A. Symp. on Flexible Automation, pp.1363-1366, 1990.
- [8] J. K. Parker, R. Dubey, F. W. Paul, R. J. Becker: "Robotic Fabric Handling for Automation Garment Manufacturing," Journal of Engineering for Industry, Vol.105, pp. 21-26, 1983.
 [9] K.Tagawa, T. Kakimi, T. Ojika and R. Kimima: "A Method for
- [9] K.Tagawa, T. Kakimi, T. Ojika and R. Kimima: "A Method for Dynamix Cloth Simulation and Application in the Virtual Fashion System," Proc. of VSMM, 1998.
- [10] J. Wolter and E. Kroll: "Toward Assembly Sequence Planning with Flexible Parts," Proc. of IEEE Int. Conf. on Robotics and Automation, pp.1517–1524, 1996.