Instance Recognition of Clumped Clothing Using Image Features Focusing on Clothing Fabrics and Wrinkles

Kimitoshi YAMAZAKI*1,

***Abstract*— This paper describes about a method of instance recognition for clumped clothes. The method assumes to be used for building autonomous systems, with the purpose of recognizing day-to-day clothing thrown casually. A set of Gabor filters is applied to an input image, and then several image features that are invariant to translation and rotation are generated. To cope with a situation in clumped clothes, the feature description is applied after image over-segmentation and boundary edge detection. Experiments of state description and instance recognition using real clothes show the effectiveness of the proposed method.**

I. Introduction

In daily environments, people use various types of clothing. One of the effective contributions of robots will be to have the ability of doing laundry because people have an excess of tedious housework to accomplish. In this paper, an instance recognition method for clumped clothing is introduced. The method is useful for building autonomous systems, with the purpose of recognizing day-to-day clothing thrown casually. Fig. 1 shows the problem that we address. From one grayscale image that captures several articles of clothing, the method outputs guessed instances of them.

Compared to the approaches used for the recognition of solid objects, soft objects such as clothing pose significant challenges because of their variable shape and appearance. Although well-known image features [10] [13] [16] provide highly reliable results for the purpose of object detection, these features cannot be used for clothing because there is an assumption that the transformations are limited only to solid objects. In accordance with the fact, some robotics researchers who build robots to perform cloth manipulation have proposed the use of various types of image features. Ono et al. [19] targeted square-shaped cloth and proposed a description of the bending state based on its image contours. Kakikura et al. [9] proposed selecting target clothing by using color information and achieved success for an isolated task. Willimon et al. [33] also proposed clothing classification. In their isolation phase, graph-based segmentation algorithm was used for deciding target clothing.

In this previous work, type of the clothing was given in advance. Otherwise, knowledge that was useful for the target identification was simple such as material color. Thus, if there is a recognition module using more generic information, it is useful for developing a general-purpose autonomous system for applications such as daily assistance and industrial laundry.

*1 Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, Nagano, 380-8553, Japan. {kyamazaki}@shinshu-u.ac.jp

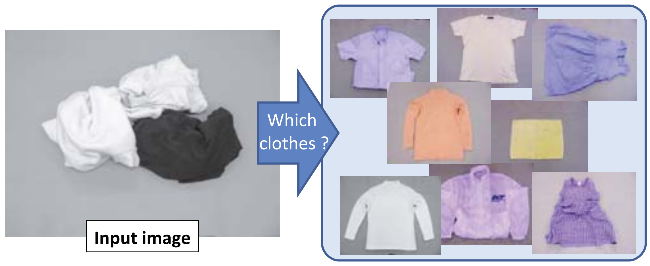

Fig. 1. Concept of our clothing instance recognition. One grayscale image is input, and proper fabric goods is selected from the list of clothing given in advance.

Other researches have adopted a policy of manipulating a target clothing for recognition. Osawa et al. [21] achieved this type of classification while handling clothing. 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. [11] utilized a 3D deformable model, and obtained a correspondence between the model and an input pointcloud that was captured by a trinocular stereo camera. Recently, Doumanoglou et al. [6] 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. [26] achieved to unfold an item of clothing based on shape estimation using the proposed polygonal model.

Manipulation-based recognition is suitable for clothing because active manipulation improves the accessibility of information needed for recognition. However, problems such as processing time and handling error remain for the recognition. That is, if there is a recognition module that enables a robot to identify each clothing without touching, it will be possible to plan an effective handling motion that responds to the clothing. However, to our best knowledge, such scheme have been achieved only using strongly-limited visual features: e.g. material color. In some cases color information is useful, however it is not always true in daily environment because there are many clothing having same color.

In this paper, we aim to identify clothing instance, without performing any handling and without using color information. A grayscale image capturing clumped clothing is our input data. To reserve generality to clothing, fabrics and wrinkles are used to describe distinctive representation. In our assumption, the knowledge of individual clothing is given in advance, and a main process of recognition is to estimate their existence and each of instance from the input image.

The paper is organized as follows: Section II explains how to extract the information of fabrics, wrinkles and cloth overlaps. Section III introduces superpixels-based clothing classification. Section IV shows experimental results, and Section V presents the conclusions of this paper.

II. Image Feature Description Focusing on Clothing Fabrics and Wrinkles

Our previous paper [29] describes a feature description suitable for clothing classification. Multi-scale and multi-orientation filtering are applied to an input image that captures a place of clothing, and the feature is extracted and analyzed. In this section, the feature description is explained.

A. Pre-processing for extracting clothing information

The leftmost figure in Fig.2 shows clothing that is placed on a floor. This image shows a shirt with long sleeves, which is made out of cotton. We can divide this clothing image into several parts: (i) cuffs or other specific parts, (ii) wrinkles and (iii) cloth overlaps. The information in (i) tells us what type of clothing it is, but the result depends on the placement of the clothing. In contrast, (ii) and (iii) are always observable, and could provide us with useful information. Moreover, the type of wrinkle depends on the clothing fabrics. From this reason, we extract and analyze them.

We focus on the fact that the contrast in the image region with respect to the clothing shows gradual changes in the frequency domain. In other words, some parts of the clothing can have stripe-shaped states due to soft material.

To analyze this property, a set of Gabor filters is applied to an input image. This approach is similar to a filter bank [8] [25] [31], the most frequent application of which is texture classification [14]. In our case, because we assume that wrinkles and cloth overlaps in images are derived from a combination of waves that have directionality and gradual frequency, Gabor filters are suitable to describe them. In the filtering, the parameters of the wave profile change, and then, helpful information is extracted from the convolution results. For example, high frequency coefficients often highlight contours and cloth overlaps, whereas low frequency coefficients persistently respond to wrinkles.

Fig.2 shows the concept of the filtering. Various kernel functions of Gabor filters are prepared, and the filtering results are used for generating a resulted image called "maximum magnitude image" for the remainder of this paper [1]. Each pixel in this image has a value that is related to a Gaussian variance, which provides a maximum magnitude during the scale space.

B. Maximum magnitude image

For generating a maximum magnitude image, the following procedure is needed. First, the multi-orientation filtering is performed. We obtain an image whose pixels are the sums of the result of the Gabor filtering with varying the orientation parameter θ. (In Fig.2, 'Sum of Gabor' indicates

[1]The original method generates two images as the filtering result. However, we only use one image in the proposed method of this paper.

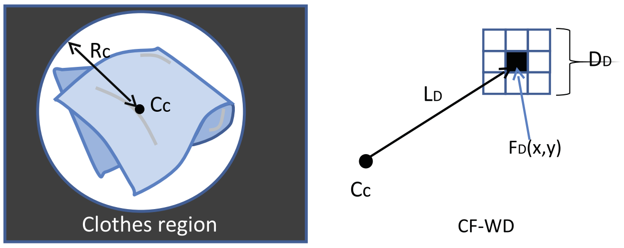

Fig. 3. Feature description

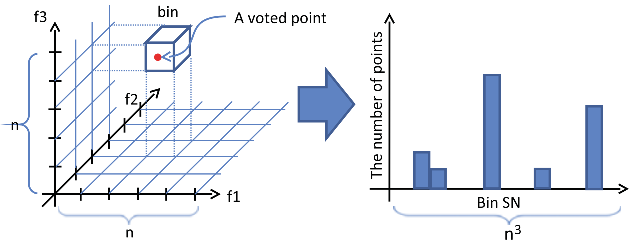

Fig. 4. Histogram calculation from parameter distribution

it.) We call the image 'temporal image' in the rest of this section.

By varying the frequency parameter σ, a set of temporal images is generated. The maximum magnitude image consists of pixels that indicate the maximum radiance in the temporal images. We define a pixel of the maximum magnitude image $I_{mag}(\mathbf{x})$, which can be written as the following:

$$I_{mag}(\mathbf{x}) = \arg \max_{\sigma} F(\mathbf{x}, \sigma), \tag{1}$$

where

$$F(\mathbf{x}, \sigma) = \int_{\theta} \int_{w} f(\mathbf{x}) g(\mathbf{x} + \mathbf{x}_0, \theta, \sigma) d\mathbf{x}_0 \mathbf{d}\theta. \tag{2}$$

g shows Gabor kernel. The window width w and the frequency are decided from the σ value automatically.

C. Feature description

First, the clothing region is extracted from the input image. In the present form, we apply the mean shift-based image segmentation [2], and remove the pixels that belong to the background. After that, as shown in the left panel in Fig.3, a circle that sufficiently includes the target clothing is defined. The radius R_C and its center coordinates C_C are used in the process for making features, as described next.

Clothing used in daily living is made of various materials such as cotton, polyester, acrylic. In addition, there are many types of clothing fabric, such as pile and shirring. Extracting these characteristics from the visible surface of clothing is a great help in classifying them. Moreover, when analyzing the organization of wrinkles, clothing fabric is dominant to the organization. A maximum magnitude image makes it possible to describe these elements distinctively.

A feature descriptor derived from clothing fabric and wrinkle density is generated from all of the pixels that form

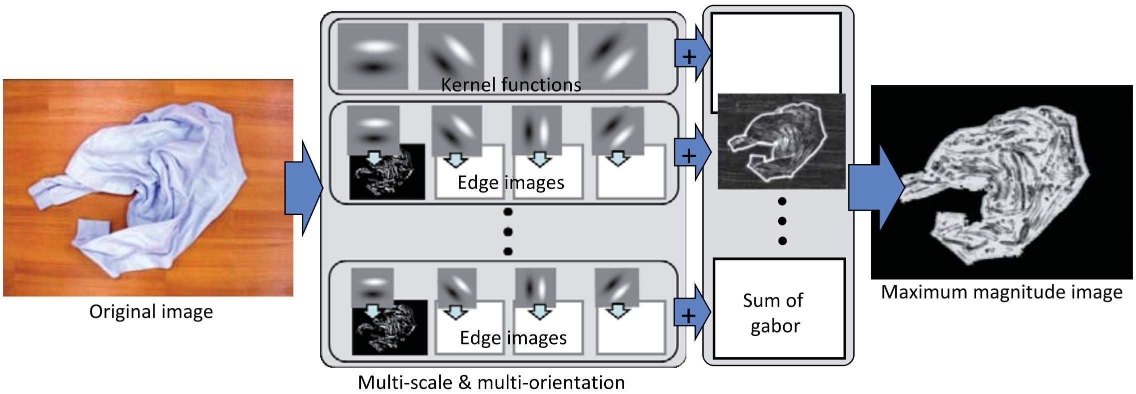

Fig. 2. Multi-scale & multi-orientation filtering. In the maximum magnitude image, the brighter pixel indicates a pixel which reacts in the larger scale parameters.

the clothing region in a maximum magnitude image. There are three criteria as follows (see the right panel of Fig.3):

1) The proportion of R_C to L_D, where L_D is the distance between C_C and an arbitrary pixel D that is in the clothing region with coordinates (x, y).
2) Pixel brightness $F_D(x, y)$.
3) The sum of the difference between D and its surrounding 8 neighbors, in other word, $D_D = \sum_{i,j \in W}(F_D(x, y) - F_D(x+i, y+j))$.

These results are parameterized according to the above three criteria, and then the results are projected into the 3D parameter space. After that, as shown in Fig.4, the space is regarded as divided into small voxels, and number of points in each voxel are counted. As a result, we can obtain a frequency histogram to describe the clothing in the image.

III. Extension of Feature Description for Clumped Clothing

The feature description method mentioned in Section II assumes a situation in which a clothing item is placed in front of a camera. However, it is often possible that several clothing items are deposited in one place. To cope with such situation, we propose a combination of the feature description with an over-segmentation method. The result is used to create feature descriptions locally, with classification performed in each local region.

A. Classification method based on local regions

The leftmost figure in Fig. 1 shows an example picture that we target. To create the feature description, image regions that capture individual clothing must be specified. That is, image segmentation is required in advance. However, it is difficult to obtain complete segmentation results because appropriate boundary edges between different clothes cannot always be obtained, and mistaken edges might be extracted from wrinkles or cloth-overlaps. For this reason, over-segmentation is first applied to an input image. An image is divided into superpixels, and the clothing class is then given at each superpixel.

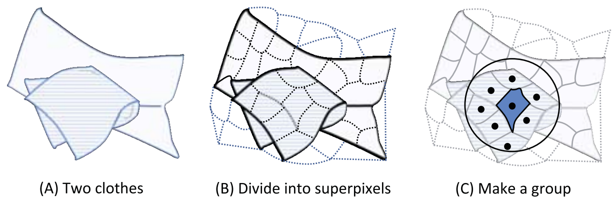

Fig. 5. Feature description

The following is an outline of the classification. Let s_i be the i th superpixel, and $\{s_1, s_2, \cdots, s_N\}$ be a set of superpixels. The probability that s_i is classified into class k is defined as follows:

$$p(k|s_i) = F(k, \bigcup_{s_j \subset G_d(s_i)} s_j), \tag{3}$$

where $F(k, R)$ is a function for calculating the likelihood that a local region R belongs to k. In the proposed method, the function is a discriminative function that outputs a classification result from trained feature descriptions. $G_d(s_i)$ is a set of superpixels neighboring s_i:

$$G_d(s_i) = \{s_j \big| \|x(s_i) - x(s_j)\| < d\}, \tag{4}$$

where $x(s_i)$ denote the center of gravity of s_i.

As shown in Eq. (4), one criterion for deciding whether s_j adjoins in s_i is the Euclidean distance from s_i. However, if a strong image edge exists between them, the adjacency is eliminated. This is because the edge has the potential to be a boundary edge that separates two cloth items. Fig. 5 shows an example. A black point depicted in (C) indicates the center of gravity of a superpixel. Only superpixels satisfying the following two criteria are regarded as belonging to the same group: (i) existing in a circle of radius d, (ii) no boundary edge between s_i and s_j. The method to obtain boundary edge candidates is presented in Section III-C.

After obtaining a group of superpixels, feature description is performed using the pixels that comprise the superpixels. R_C and C_C in Section II-C are substituted by d and $x(s_i)$, respectively. The result is used to estimate the clothing type

k, and then the estimated class is assigned to the centering superpixel s_i.

B. Superpixel calculation

As mentioned in the former section, pixels belonging to the background are initially removed using mean-shift-based image segmentation. Superpixels are calculated from the remaining image region. Next, the class of each superpixel is investigated. If coherent superpixels of the same class are obtained, the region is regarded as a clothing item.

One advantage of superpixels is that they consist of groups of pixels with keeping original strong edges. There are many methods for calculating superpixels, and the shapes of the generated superpixels depend on the method. The mean-shift-based [3] and watershed-based [32] methods enable us to obtain superpixels quickly. However, it is difficult to control the shapes and sizes of superpixels produced using the two methods. For the purpose of superpixel grouping shown in Fig. 5, inhomogeneous superpixels are not suitable. On the other hand, other methods such as normalized-cuts [34] and related methods [23] provide superpixels with nearly homogeneous sizes. For this reason, we apply TurboPixels [15] to obtain over-segmentation.

In TurboPixels, seed points are placed at regular interval, and are then extended on the bases of the following equation:

$$S_I(x, y) = [1 - \alpha\kappa(x, y)]\phi(x, y) - \beta[\mathbf{N}(x, y) \cdot \nabla\phi(x, y)], \tag{5}$$

where (x, y) denotes image coordinates, and κ denotes a curvature factor. $\mathbf{N}(x, y)$ is a normal at (x, y), and α and β are weight coefficients. The first and second terms at the right member are called the reaction-diffusion and doublet terms, respectively. ϕ is a term that decreases the extension speed, where the gradient value is large. The equation is as follows:

$$\phi = e^{-E(x,y)/\nu}, E(x, y) = \frac{||\nabla I||}{G_\sigma * ||\nabla I|| + \gamma}, \tag{6}$$

where ν and γ are constant values, and $G_\sigma *$ represents a convolution calculation that considers pixels in a circle of radius σ with a weighted mean. The extension process in itself is similar to the level set method [20]. However, there is a penalty to stopping the extension if a focusing point is on a boundary candidate edge.

C. A detection of boundary edge candidates

To obtain a boundary edge between two different clothes, cloth overlaps are detected by the following procedure. Edge detection process is first applied to an input image. This process can be substituted to use a certain temporary image (Sum of Gabor in Fig.2) which is filtered by a relatively high frequency σ. Thresholding and peak tracing based on edge magnitude are applied to the image, edge lines are obtained from pixels having large brightness difference. Other image regions with gradual brightness changes are ignored because they derive from wrinkles in many cases.

The obtained image edges have possibilities to represent cloth overlaps, however other causes such as clothing texture can be considered. To relieve the uncertainty, a verification for cloth overlaps is performed. The right enlarged panel in Fig.6 shows an example of cloth overlaps. Because there is some spatial gap between two part of cloth, we can find gradual brightness changes in downside of the cloth. This fact is used to select image edges derived from cloth overlaps.

Fig. 6. A example of cloth overlap. Gradual brightness changes are found in the downside part.

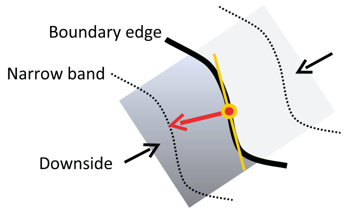

Fig. 7. A verification of cloth overlaps. Narrow band is defined original detected edge line, and the brightness gradation on downside part is investigated. Red arrow shows the direction to calculate the gradation.

A narrow band is set around pixels belonging to the edge lines that are detected at the beginning process, and a brightness value of pixels included in the band is then investigated. If pixels belonging to downside part of cloth overlap have gradually brightness changes as shown in Fig.7, the edge pixel is given high probability as cloth overlap.

In our implementation, which is described in the next section, we selected a temporary image that was filtered by $\sigma = 4.0$ Gaussian, and we set the band-width to 5 pixels.

IV. Proof with Experiments on Clothing Classification

A. Settings for the experiment of clothing classification

An image database that consisted of 15 clothing items was prepared, so that experiments could be performed on clothing instance recognition. The list of clothing is shown in Fig.8.

About 200 to 300 images the size of which was VGA (640×480 pixels) were captured from each piece of clothing by throwing it randomly, so that the total number of images in the database was over 2100. The distance between a camera and a cloth was about 900 [mm]. Fig.9 shows some images in the dataset. There were various state of clothing that were scrumpled or smoothed.

The database is used for generating feature library and basic evaluation. On the other hand, 20 images that capture some piece of clumped clothing were also prepared to investigate the effectiveness of the proposed method.

Fig. 8. The list of clothing used for generating image database

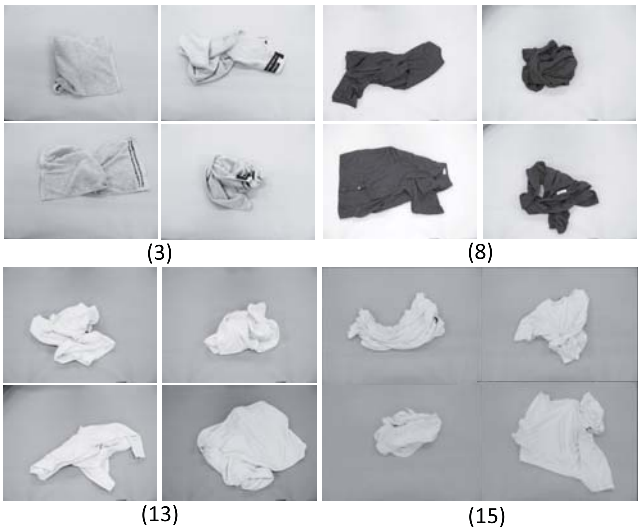

Fig. 9. Examples of image in database

B. Basic evaluation for classification performance

A discriminative function was generated from training data that were prepared from the image dataset introduced in above subsection. The procedure was as follows: (i) one grayscale image capturing an item of clothing was selected, (ii) superpixels were calculated, (iii) boundary candidate edges were calculated, and then (iv) feature descriptions were calculated using the method described in Section III. The feature descriptions are coordinated by each clothing item, and then a discriminative function is generated from them using a multi-class SVM.

In the classification phase, the procedure (i) to (iii) was also performed. After that, each superpixel arising from clothing region was input to the discriminative function. The result was represented by 15 floating values, which were the inverse of the distances from discriminative boundary in feature space. As it can be regarded as a representation of likelihoods for each clothing item, a class producing the highest value was finally given to the superpixel.

Classification performance was investigated using these images. In this experiment, an image that was not included in the learning data was input, and superpixels were in-

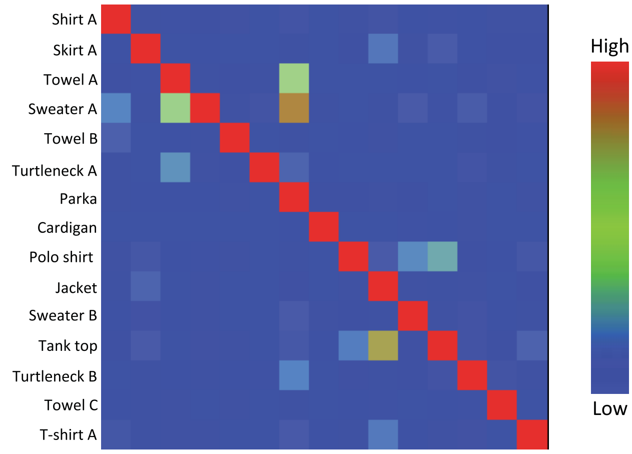

Fig. 10. Confusion matrix

TABLE I

CLASSIFICATION RESULTS

	Ave. (n=50)	Var. (n=50)	Ave. (n=200)	Var. (n=200)
(1)	0.687	0.0029	0.752	0.0068
(2)	0.694	0.0050	0.745	0.0036
(3)	0.689	0.0083	0.750	0.0068
(4)	0.727	0.0043	0.780	0.0041
(5)	0.667	0.0046	0.741	0.0042
(6)	0.711	0.0202	0.746	0.0244
(7)	0.700	0.0076	0.756	0.0057
(8)	0.832	0.0100	0.879	0.0066

dividually classified by the above-mentioned procedure. If all of superpixels are accurately classified, we should have a group of superpixels on which same class was labeled. Fig. 10 shows a confusion matrix of classification results by leave-one-out cross validation. Each cell shows above-mentioned inverse distances from discriminative boundary. Basically correct results were acquired; diagonal elements had high value.

Table I shows a part of classification results. It shows two patterns of experiments: in every clothing item, (i) 50 images were used to generate a discriminative function ($n = 50$), or (ii) 200 images were used ($n = 200$). Because we could obtain dozens of feature descriptions from one image, the number of feature vectors used in the learning phase was more than ten thousands in the case $n = 200$. In each clothing item, 50 images that were not included in learning data were used to test the classification performance. The ratio between the number of accurately classified superpixels and all superpixels was calculated. Table I shows the averages (Ave.) and the variances (Var.) of the ratio for clothes whose number (1) to (8). When we used 200 images for learning, classification accuracy was about 77%. This means that we can perform instance recognition by a simple majority voting scheme using this result.

In our past research, the significance of our feature description was proven by comparing with well-known methods of generic object recognition [28]. In this experiment,

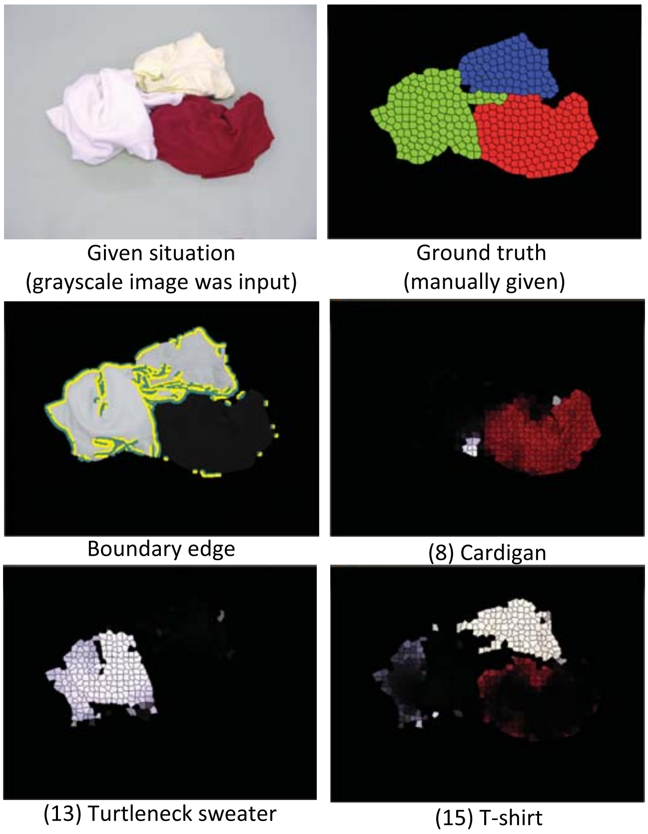

Fig. 11. Results of clumped clothing classification

we also confirmed that our feature description has higher discriminative power than other texture-based features: MR8, Leung-Marik, and Schmid [14].

C. Clumped clothing classification

The upper left panel in Fig. 11 shows an example of a classification of clumped clothing. It includes three clothing items from Fig.8: (8)cardigan, (13)turtleneck, and (15)T-shirt. As preparation, superpixels were calculated from an input image. An operator then manually divided the superpixels into the three clothing items as shown in the upper right panel. A discriminative function is generated by the same procedure mentioned in above subsection.

In the classification process, a grayscale image capturing clumped clothing was input, and then a background region was removed using mean-shift-based segmentation. Next, boundary candidate edges and superpixels were calculated, and then each superpixel was classified using the discriminative function. The center left panel in Fig. 11 shows boundary candidate edges that are a criterion for coordinating superpixel groups used for feature description.

The remaining three panels in Fig. 11 show the classification result. In this experiment, each superpixel was classified individually. If a superpixel was regarded as belonging to a specific class, it was painted with a brighter value in a panel named after that class. The brightness value was determined from the SVM classifier output. In this example, more than

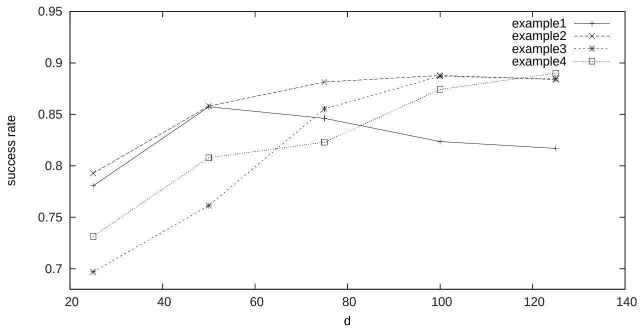

Fig. 12. d vs. success rate

92% of the superpixels were classified accurately. As further examples, we prepared 20 images capturing three or four items of clumped clothing, and the same classification experiments were performed on them. An average classification accuracy was 80.79%.

Parameter settings in this experiment were as follows: the seed point interval for calculating TurboPixels was set to 12 pixels, an empirically selected value. The distance threshold d presented in Eq.(4) was set to 100.

D. Performance differences according to distance threshold

Additional experiments were performed to investigate the performance with various distance threshold d because classification performance was influenced significantly by d.

Classification accuracy was investigated with the value d as $25, 50, 75, 100$ and 125. Four representative images were used in this experiment. In Fig.12, the horizontal axis shows the size of d and the vertical axis shows the accuracy rate (1.0 indicates 100 %). A large d value tends to provide higher accuracy performance, but there is almost no change between 100 and 125.

One issue is that superpixels close to the boundary have a special propensity to yield incorrect classification results, as can also be seen in the lower two panels in Fig. 11. This is because boundary detection is incomplete; thus, a superpixel group might be derived from two or more different items of clothing. However, perfect boundary detection is extremely difficult: hence, a reasonable practical approach is required to obtain a final classification result through comprehensive grouping of superpixels.

V. Conclusions

In this paper, we proposed a method of instance recognition of clumped clothes. A set of Gabor filters are applied to an input image with a range of frequencies and directions, and useful information such as fabrics and wrinkles are detected based on the maximum magnitudes. To cope with a situation in clumped clothes, the feature description is applied after image over-segmentation and boundary edge detection. Experiments using real images, we achieved a success rate greater than 80%.

Future work would include more feasible information about the state of the clothing, which would be added to the classification method. After that addition, we will attempt to develop a method that enables a daily assistive robot to handle daily laundry. Collaboration this research with clothing detection method [27] would have a important role for such application. As another extension, a combination with grasp point selection would make it possible to clothing manipulation such as folding laundry. Previous work by Ramisa et al. [22] and ours [29] would broaden our vision.

ACKNOWLEGEMENT

The author appropriates Professor Masayuki Inaba and Dr. Ryo Hanai, who discuss with me and give important advices. This work was supported by JST PRESTO program and JSPS KAKENHI Grant Numbers 26700024.

REFERENCES

[1] J. Bigü and J. M. du buf: "N-folded symmetries by complex moments in gabor space and their application to unsupervised texture segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-16(1):80–87, 1994.
[2] C. Christoudias, B. Geogescu and P. Meer: "Synergism on Low Level Vision," Int'l Conf. on Pattern and Recognition, pp. 150 – 155, 2002.
[3] D. Comanicui and P. Meer: "Mean shift: A robust approach towead feature space analysis," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp.603 – 619. (2002)
[4] G. Csurka, C. Bray, C. Dance, and L. Fan: "Visual categorization with bags of keypoints," in Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pp. 59 – 74, 2004.
[5] S. Cuén-Rochín, J. Andrade-Cetto and c. Torras: "Action Selection for Robotic Manipulation of Deformable Planar Objects," in Proc. of Frontier Science Conference Series for Young Researchers: Experimental Cognitive Robotics, pp. 1-6, 2008.
[6] A. Doumanoglou, A. Kargakos, T. Kim and S. Malassiotis: "Autonomous Active Recognition and Unfolding of Clothes using Random Decision Forests and Probabilistic Planning," in Proc of IEEE Int'l. Conf. on Robotics and Automation, 2014.
[7] M. Galun, E. Sharon, R. Basri and A. Brandt: "Texture Segmentation by Multiscale Aggregation of Filter Responses and Shape Elements," Proc. of IEEE Int'l. Conf. on Computer Vision, pp. 716-723, 2003.
[8] J. Geusebroek, A. Smeulders and J. Weijer: "Fast Anisotropic Gauss Filtering," IEEE Trans. on Image Processing, 12(8):938-943, 2003.
[9] K. Hamajima and M. Kakikura: "Planning Strategy for Unfolding Task of Clothes – Isolation of clothes from washed mass –," in Proc. of Int'l. Conf. on Robots and Systems, pp. 1237 – 1242, 2000
[10] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool: "SURF: Speeded Up Robust Features, " Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346–359, 2008
[11] Y. Kita, F. Saito and N. Kita: "A deformable model driven method for handling clothes," Proc. of Int. Conf. on Pattern Recognition, 2004.
[12] H. Kobori, Y. Kakiuchi, K. Okada and M. Inaba: "Recognition and Motion Primitives for Autonomous Clothes Unfolding for Humanoid Robot," in Intelligent Autonomous Systems 11, pp.57–66, 2010.
[13] K. Mikolajczyk and C. Schmid: "Scale and Affine Invariant Interest Point Detectors," Int'l Journal of Computer Vision vol. 60, No. 1, pp.63 – 86, 2004.
[14] T. Leung and J. Malik: "Representing and recognizing the visual appearance of materials using three-dimensional textons," International Journal of Computer Vision, 43(1):29-44, June 2001.
[15] A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson and K. Siddiqi: "TurboPixels: Fast Superpixels Using Geometric Flows," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 12, pp.2290 – 2297. (2009)
[16] D. G. Lowe: "Distinctive image features from scale-invariant keypoints," Int'l Journal of Computer Vision, vol. 60, No. 2, pp. 91-110, 2004.
[17] J. Maitin-Sphepard, M. Cusumano-Towner, J. Lei and P. Abbeel: "Cloth Grasp Point Detection based on Multiple-View Geometric Cues with Application to Robotic Towel Folding," Int'l. Conf. on Robotics and Automation, pp.2308 – 2315, 2010
[18] S. Miller, M. Fritz, T. Darrell and P. Abbeel: "Parametrized Shape Models for Clothing," In the proceedings of the International Conference on Robotics and Automation (ICRA), 2011.
[19] E. Ono, H. Okabe, H. Ichijo and N. Aisaka: "Robot Hand with Sensor for Cloth Handling," In Proc. 1990, Japan, U.S.A. Symp. on Flexible Automation, pp. 1363–1366, 1990.
[20] S. Osher and J. A. Sethian: "Fronts propageting with curvature dependent speed: Algorithm based on Hamilton-Jacobi formation," Journal of Computational Physics, Vol. 79, pp. 12 – 49. (1998)
[21] F. Osawa, H. Seki, and Y. Kamiya: "Unfolding of Massive Laundry and Classification Types by Dual Manipulator," Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.11 No.5, pp. 457– , 2007.
[22] A. Ramisa, G. Alenya, F. Moreno-Noguer and C. Torras: "Using Depth and Appearance Features for Informed Robot Grasping of Highly Wrinkled Clothes," in Proc. of IEEE Int ' l Conf. on Robotics and Automation, pp. 1703 - 1708, 2012.
[23] X. Ren and J. Malik: "Learning a classification model for segmentation," In Proc. 9th Int. Conf. Computer Vision, vol. 1, pp. 10-17. (2003)
[24] Y. Rubner and C. Tomasi: "Coalescing Texture Descriptors," ARPA Image Understanding Workshop, pp. 927 – 935, 1996.
[25] C. Schmid: "Constructing models for content-based image retrieval," In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 39–45, 2001.
[26] J. Stria, D. Prusa, V. Hlavac and L. Wagner: "Garment Perception and its Folding Using a Dual-arm Robot," in Proc. of International Conference on Intelligent Robots and Systems, 2014.
[27] K. Yamazaki and M. Inaba: "A Cloth Detection Method Based on Image Wrinkle Feature for a Daily Assistive Robots," IAPR Conf. on Machine Vision Applications, pp.366–369, 2009.
[28] K. Yamazaki and M. Inaba: "Clothing Classification Using Image Features Derived from Clothing Fabrics, Wrinkles and Cloth Overlaps," in Proc. of IEEE/RSJ Int'l Conf. on Robots and Systems, pp. 2710 – 2717, 2013.
[29] K. Yamazaki: "Grasping Point Selection on an Item of Crumpled Clothing Based on Relational Shape Description," in Proc. of IEEE Int'l Conf. on Intelligent Robots and Systems, 2014.
[30] Y. Zhao and G. Karypis: "Comparison of agglomerative and partitional document clustering algorithms," University of Minnesota - Computer Science and Engineering Technical Report, No. 02-014, 2002.
[31] M. Varma and A. Zisserman: "A Statistical Approach to Texture Classification from Single Images," International Journal of Computer Vision Volume 62, Numbers 1-2, 61-81, 2005.
[32] L. Vincent and P. Soille: "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 6, pp. 583 – 598. (1991)
[33] B. Willimon, S. Birchfleld, I. Walker: "Classification of clothing using interactive perception," in Proc. of IEEE Int'l Conf. on Robotics and Automation, pp. 1862 - 1868, 2011.
[34] S. Yu and J. Shi: "Multiclass spectral clustering," IEEE Int'l Conf on Computer Vision, Vol. 1, pp. 313 – 319. (2003)
[35] "Tutorial on Gabor Filters," http://mplab.ucsd.edu/tutorials/tutorials.html
[36] C. Chang and C. Lin: "LIVSVM – A Livrary for Support Vector Machines," http://www.csie.ntu.edu.tw/ cjlin/libsvm/.