Active Clothing Material Perception using Tactile Sensing and Deep Learning

Wenzhen Yuan\textsuperscript{1}, Yuchen Mo\textsuperscript{1,2}, Shaoxiong Wang\textsuperscript{1}, and Edward H. Adelson\textsuperscript{1}

Abstract—Humans represent the objects in the same category using their properties, in order to well discriminate and understand them, and an intelligent robot should be able to do the same. In this work, we propose a robot system that can automatically perceive the object properties through touch. We work on the common object category of clothing. The robot moves under the guidance of an external Kinect sensor, and squeezes the clothes with a GelSight tactile sensor, then it recognizes the 11 properties of the clothing according to the tactile data. The target properties are the physical properties, like thickness, fuzziness, softness and endurance, and semantic properties like wearing season and preferred washing methods. We collect a dataset of 153 varied pieces of clothes, and make 6616 exploring iterations on them. To learn the useful information from the high-dimensional sensory output, we applied Convolutional Neural Networks (CNN) on the tactile data for recognizing the clothing properties, and on the Kinect depth images for selecting exploration locations. Experiments show that using the trained neural networks, the robot can automatically explore the unknown clothes and learn their properties. This work proposes a new architecture for active tactile perception system with vision-touch system, and has potential to enable robots to help humans with varied clothing related housework.

I. INTRODUCTION

A core requirement for an intelligent robot is to well understand the properties of objects in the natural environment. Among the common objects, clothing is an important part. Humans evaluate an article of clothes largely according to its material properties, such as thick or thin, fuzzy or smooth, stretchable or not, etc. The understanding of the clothes’ properties helps us to better manage, maintain and wash the clothes. If a robot is to assist humans in daily life, understanding those properties will enable it to better understand human life, and assist with daily housework such as laundry sorting, clothes maintenance and organizing, or choosing clothes.

For perceiving material properties, tactile sensing is an important method. Lederman and Klatzky \cite{1}, Tiest \cite{2} demonstrated that humans use different exploratory procedures to sense different properties of objects, such as roughness or compliance. Researchers have been trying to make a robot to learn the material properties through touch as well. As an example, Chu et al. \cite{3}, Xu et al. \cite{4} developed setups to perceive properties of general objects using tactile sensors and a set of pre-set procedures like squeezing and sliding. However, making the robots explore the more refined properties in the natural environment remains a big challenge, and discriminating the subtle difference between the objects in the same category, such as clothing, is even more difficult. The challenge mainly comes from two sides: how to obtain and understand adequate information from a tactile sensor, and how to generate an exploration procedure to obtain the information.

At the same time, roboticists have been interested in clothes related tasks, in both manipulation and recognition. The majority of the related works use only vision as sensory input that measures the clothes’ global shapes, and the recognition is mostly restricted to the rough classification of the clothing type. The perception of fine-grained properties of clothes, or the study based on common clothes with a wide variety, is still undeveloped.

In this paper, we design a robotic system that perceives the material properties of common clothes using automatic tactile exploration procedures. The hardware setup of the system is shown in Figure 1(a). We address the two challenges

\textsuperscript{1}Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT, Cambridge, MA 02139, USA
\textsuperscript{2}Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China
of tactile exploration of object properties: to collect and understand the high-resolution tactile data, and to generate exploration procedures for sensing. The tactile sensor being applied is a GelSight sensor [5, 6], which senses the high-resolution geometry and texture of the contact surface. A GelSight sensor uses a piece of soft elastomer as the contact medium, and an embedded camera to capture the deformation of the elastomer. The exploration procedure is squeezing a part of the clothes, mostly a wrinkle, and recording a set of tactile images with GelSight (see Figure 1(c)-(f)). Then we train a Convolutional Neural Network (CNN) for multi-label classification to recognize the clothing properties. For generating exploration procedures automatically, we use an external Kinect sensor to get the overall shapes of the clothes, especially the positions of the wrinkles, and train another CNN to pick up preferable points on the wrinkles. The robot will follow the Kinect detection for effective exploration. We also make the exploration closed-loop: if the tactile data is not good, which means the neural network cannot recognize the properties with high confidence, then the robot will re-explore the clothing on another location, until it gets good tactile data and confident results.

The 11 properties we studied are the physical properties, including thickness, fuzziness, smoothness, softness, etc., and the semantic properties that are more related to the application of the clothes, including wearing seasons, preferred washing methods, and textile type. To make the system robust to a wide range of common clothes, we collect a dataset of 153 pieces of clothes for the training, and the dataset covers different clothing types, materials and sizes. Experimental results show that the system can recognize the clothing properties for both seen and unseen items, as well as detecting effective locations to generate tactile exploration. The robot can use the trained networks to do closed-loop exploration on the clothes. To our knowledge, this is the first work on studying fine-grained clothing properties with robot tactile sensing. The methodologies of this work will enable robots to understand common clothes better, and assist humans on more housework such as washing laundry and clothing sorting.

II. RELATED WORK

A. Clothes classification

The robotic community has been interested in clothing related topics for years, especially for the home assistant robotic tasks. The major focus has been clothing manipulation and recognition/classification. Researches related to clothing manipulation are mostly on grasping, folding and unfolding. For the researches on clothing recognition or classification, most of them use vision as sensory input, and classify the clothes according to their rough types, such as pants, t-shirts, coats, etc. Willimon et al. [7], Li et al. [8], Gabas et al. [9] introduced methods for clothing classification by matching the global shape of the clothing, either the 2D shape from color images or 3D shape from RGBD sensors, to the clothes dataset. Sun et al. [10] proposed a method to recognize clothing type from stereo vision, where they applied more local features, such as the clothing’s wrinkle shapes and textures.

For multi-modal clothing perception, Kampouris et al. [11] proposed a robotic system to classify clothes’ general types and materials. They used an RGBD camera to capture the global shape, a photometric stereo sensor to sense surface texture, and a fingertip tactile sensor to measure dynamic force when rubbing the clothing. They showed that the multi-modal input, especially the texture perception from the photometric stereo sensors, largely improve the recognition precision of the clothing materials. However, recognizing more fine-grained clothing properties on a more common clothes dataset remains a challenge.

B. Tactile sensors and GelSight

Tactile sensing is an important sensory modality for robots. In the past decades, different kinds of tactile sensors have
been developed, as reviewed in [12, 13]. The majority of the existing tactile sensors measure force or contact distribution over an area. Tactile sensing has been used for object classification according to the shape inferred from force distribution (e.g. [14]), and for exploring material properties by combining with motion. Chu et al. [3] and Xu et al. [4] show two systems of using tactile sensors to explore multiple material properties. They used either a robot or a motion system to control the sensor to press or slide on the object surface, and thus classifying different properties or adjective descriptions from the tactile signals.

In this work, we apply a GelSight tactile sensor [5, 6] in the exploration of the clothes. The GelSight sensor is an optical-based tactile sensor that measures the surface geometry with very high spatial resolution (ranging from 1 micron to 20 microns for different designs). Moreover, the printed markers on the sensor surface enable it to measure the contact force or shear [15]. Li and Adelson [16] showed that the high resolution of GelSight makes it very effective to discriminate different material categories by surface texture. The sensor can also estimate the physical properties of the objects through contact. An example is [17], where the researchers showed that the GelSight video can be used to estimate objects’ hardness using the change of the contact shape under the increasing force. Yuan et al. [18] studied the GelSight’s performance on fabric perception, where they tried to discriminate different fabrics using an embedding vector from the vision or tactile modalities that described the overall properties of the fabrics.

C. Deep learning for texture recognition and tactile data

Texture provides significant information about material properties of objects. Early works on texture recognition mostly used hand-crafted features like Textons [19], Filter Banks [20] or Local Binary Patterns (LBP) [21].

In recent years, The Convolutional Neural Networks (CNN) have achieved many state-of-the-art performances on computer vision tasks, such as image classification, object detection and segmentation. It was successfully applied to texture recognition: Cimpoi et al. [22] proposed FV-CNN, which combined the CNN with Fisher Vectors (FV) to better extract localized features. The convolutional layers in FV-CNN are from the VGG model [23], pre-trained on ImageNet [24], and they served as filter banks; the FV was used to build the orderless representation. [25] proposed T-CNN for texture classification which used an energy layer after the convolutional layers, which is similar to traditional texture analysis techniques. It also showed that the pre-trained convolutional layers of Alexnet [26] significantly increased the classification accuracy.

The CNN models developed for computer vision also proved effective in processing tactile information: [17] and [18] used the CNNs on GelSight data images, with the tasks of hardness estimation and fabric property extraction, while the networks are pre-trained on images. The CNNs presented a good performance at extracting features from the complicated and highly nonlinear spatial signals.

III. DATA COLLECTION

The aim of this project is to develop a robotic system that can automatically perceive clothes and classify them according to material properties. We build the robot’s hardware system with a robot arm and a gripper with the tactile sensor GelSight. We use an external Kinect camera to guide the exploration, but vision is only used for planning the robot motion. The robot explores the clothes by gripping on the parts of the clothes, and during the process, the GelSight sensor will record a sequence of tactile images of the clothes’ shapes. We collect 2 kinds of data: the GelSight image sequences, and the gripping points on the Kinect depth images. The GelSight images help the robot to recognize the clothing properties, and the depth images and the exploration results help the robot to learn whether a gripping position is likely to produce good tactile data.

A. Clothing Dataset

We collect a dataset of 153 pieces of new and second-hand clothes. The clothes are all common clothes in everyday life, but widely distributed on types, materials and sizes. We aim to make the dataset cover all kinds of common clothes in an
TABLE I
CLOTHING PROPERTY LABELS

| Thickness (5) | Smoothness (5) |
|--------------|----------------|
| 0 - very thin (crepe dress) | 0 - very smooth (satin) |
| 1 - thin (T-shirt) | 1 - smooth (dress shirt) |
| 2 - thick (sweater) | 2 - normal (sweater) |
| 3 - very thick (wool coat) | 3 - not smooth (fleece) |
| 4 - extra thick (down coat) | 4 - rough (woven polo) |

| Fuzziness (4) | Season (4) |
|---------------|------------|
| 0 - not fuzzy (dress shirt) | 0 - all season (satin pajama) |
| 1 - a little fuzzy (dress shirt) | 1 - summer (crepe top) |
| 2 - a lot fuzzy (terry robe) | 2 - spring/fall (denim pants) |
| 3 - winter (cable sweater) | 3 - winter (cable sweater) |

| Textile type (20) | Washing method (6) |
|------------------|--------------------|
| cotton; satin; polyester; denim; gabardine; broad cloth; parka; leather; crepe; corduroy; velvet; flannel; fleece; hairy; wool; knit; net; suit; woven; other | machine wash warm; machine wash cold; machine wash cold with gentle cycles; machine wash cold, gentle cycles, no tumble dry; hand wash; dry clean |

| Labels with binary classes: Softness, stretchiness, endurance, wool surface, wind-resistance |
|---|

ordinary family. The dataset also includes a small amount of other fabric products, such as scarfs, handkerchiefs and towels. Some examples of the clothes are shown in Figure 3.

The property labels we choose are a set of common properties that humans use to describe clothes. We used the 11 labels, with either binary classes or multiple classes. The labels and examples of the classes are shown in Table I.

B. Robotic System Setup

The robotic hardware system is shown in Figure 1(a), and it consists of 4 components: a robot arm, a robot gripper, a GelSight tactile sensor, and a RGBD camera. We use the ROS system to control the entire system. The arm is a 6 DOF UR5 collaborative robot arm from Universal Robotics, with a reach radius of 850mm and payload of 5kg. We use MoveIt! library for the motion planning. The parallel robotic gripper is a WSG 50 gripper from Weiss Robotics, with a stroke of 110mm, and an approximate force reading from the current. We mount GelSight on the gripper as 1 finger, and the other finger is a 3D printed finger with a curved surface, which helps GelSight get in full contact with the clothes. The GelSight sensor we used in this project is the revised Fingertip GelSight sensor [27], as shown in Figure 1(b). The sensor has a soft and dome-shaped surface for sensing, and a sensing range of 18.6mm × 14.0mm, spatial resolution of 30 microns for geometry sensing. The elastomer on the sensor surface is about 5 Shore A scale, and the peak thickness is about 2.5mm. The sensor collects data at a frequency of 30Hz. The external RGBD camera we used is a Kinect 2 sensor, which has been calibrated and connects to ROS system via IAI Kinect2 [28] toolkit. It is mounted on a fixed supporting frame that is 106mm above the working table and a tilt angle of 23.5°, so that the sensor is able to capture the top view of the clothes.

Note that in the grasping procedure, due to the repetitive large shear force, the GelSight surface would wear off after a series of grasping, so that we have to change the sensing elastomer for multiple times. Since the elastomer sensors are manually made, they have slightly different marker patterns and shape parameters, which result in some overall differences in the tactile images.

C. Automatic Data Collection

The training data is automatically collected by the robot, and the flow chart of the process is shown in Figure 4.

Choosing gripping positions from Kinect images

The robot is most likely to collect good tactile data when it grips on the wrinkles on the clothes. The wrinkles are higher than the surrounding area, which will be captured by Kinect's...
depth images $D$. We firstly transfer the depth map into the world frame, thus obtain the depth map $D_W$ using

$$D_W = T_{K2W} \cdot K^{-1} \cdot D$$  \hspace{1cm} (1)

where $K$ is the camera matrix that expanded to 4×4 dimension, and $T_{K2W}$ is the 4×4 transformation matrix from the Kinect frame to world frame. We set the $x-y$ plane in the world frame as the table, so that the ‘depth’ value of $D_W$, which is represented as $z$, corresponds to the real height of the clothes on the table. An example of the transformed $D_W$ is shown in Figure 5(c). The edges of $D_W$, which could be easily picked by Laplacian operation, show the wrinkles on the clothes. We apply the pyramid method to down-sample the clothes. We randomly choose 1 point as the target gripping position.

Before gripping, the gripper should rotate to the angle perpendicular to the wrinkle. We calculate the planar direction of the wrinkle at point $(x, y)$ by

$$\text{Dir}(x, y) = \arctan\left(\frac{\partial D_W(x, y)}{\partial y} / \frac{\partial D_W(x, y)}{\partial x}\right)$$  \hspace{1cm} (2)

Gripping on the wrinkles Once the target point on the wrinkle is selected, the robot will move about the point, with the gripper in a perpendicular direction, and then descend to the position below the wrinkle to grip the clothing, with a low speed of 5mm/s. The gripper stops closure when the motor current reaches a threshold, which indicates a large impedance force. The GelSight records videos during the closure. Typically the GelSight records 10 to 25 frames for one gripping iteration.

After the gripping, we judge whether the contact is valid using GelSight images. If the GelSight image shows no contact with the clothing, we mark this tactile data invalid, and mark the gripping location as a failure case.

IV. CLOTHES CLASSIFICATION USING DEEP LEARNING

A. Networks for property perception

To perceive the properties of the clothes, we use a CNN for the multi-label classification of the GelSight images. The labels correspond to the clothing properties, and they are independently trained. We use two kinds of networks: one takes a single GelSight image as the input (Figure 6(a)), and the other one takes the video (Figure 6(b)). The CNNs for GelSight images are VGG19 [29], and the CNN for depth images is VGG16 [29]. The CNNs are originally designed for object recognition for general images, and pre-trained on the image dataset ImageNet [24].

For the network with a single input frame, we choose the GelSight image when the contact force is the maximum, and use a single CNN to classify the image. For recognizing the multiple properties, we train the same CNN with classification on multiple labels, which correspond to the clothing properties. The architecture is shown in Figure 6(a).

Additional to learning the properties from the single GelSight image, we also try to learn the properties from the video sequence. The GelSight video includes a set of images when the sensor squeezes the clothing with increasing forces, thus the frames record the surface shapes and textures under different forces. The videos are more informative than the static images. To train the video data, we use the structure connecting CNN and a long short-term memory units (LSTM) [30] with a hidden state of 2048 dimensions, as shown in Figure 6(b). We use the features from the second last layer fc6 from VGG16 as the input of LSTM.

For each video sequence, we select 9 frames, with a equal time stamp interval until the image of max contact. Since the gripper closes slowly and evenly when collecting the data, the gripper’s opening width between the frames is equal. As a result, some of the thick clothes would deform largely in the squeezing process, so that the selected sequence starts after the contact; while when gripping thin clothes, the maximum contact point is easily reached, and the selected sequence starts with several blank images.

B. Networks for gripping point selection

We train a CNN to learn whether the gripping location is likely to generate good tactile data. The network architecture is shown in Figure 6(c). The input data is a cropped version of $D_W$, the depth image in the world frame, and the output is a binary class on whether the image represents a potentially successful gripping. To indicate the gripping location in the depth image, we crop the depth image to make the gripping location the center of the new image, and the window size is 11cm×11cm.

C. Offline training of the neural networks

We divided the perception data from the 153 items of the clothes into 3 sets: the training set, the validation set, and the test set. The training set and validation set make of data from 123 items of clothes, and the testing set contains data from the rest 30 items. For the 123 items, we randomly choose data from 85% of collecting iterations as the training set, and 15% of the data as the validation set. The division of clothes for training and testing is manually done ahead of the network training, with the standard that the clothes in the test set should be a comprehensive representation of the entire dataset.

The output of the CNN of Kinect Depth image should be 0 or 1, depending on whether the proposed location in the image is a good location (1) or bad one (0). We make the ground truth as 0 under two circumstances: the gripper can not grip on the clothes; the gripper can not well contact the clothes, so that the tactile data is not explicit.

We train the model using stochastic gradient descent as the optimizer. The weights of the Depth CNN (Figure 6(c)) and single GelSight image(Figure 6(a)) is initialized with the weights pre-trained on ImageNet [24], and the weights of the CNN for the GelSight video (Figure 6(b)) are initialized with the weights of Figure 6(a). For the video network, we jointly train the CNN and LSTM, with 500 epochs, dropout rate at 0.5.
For training the network for GelSight images, we apply data augmentation to improve the performance of the network, by adding random values to the image intensity in the training. When training the video, we choose the input sequence slightly differently on the time stamp.

D. Online robot test with re-trials

We run the robot experiment online with the two networks: at the start of the exploration, the robot generates a set of potential exploration locations from the depth image, and use the depth CNN to select a best one. After collecting tactile data by gripping the clothing at the selected location, we use the tactile CNN to estimate the clothing properties. At the same time, the robot evaluates whether the collected tactile data is good, by analyzing the output probability of the tactile network. If the probability is low, it is likely the data is ambiguous and the CNN is not confident about the result. In this case, the robot will explore the clothing again, until the data is considered good. In the experiment (Section V-C), we choose the property of washing method and the probability threshold of 0.75.

V. EXPERIMENT

We conduct both offline and online experiments. For the offline experiments, We use the data that the robot collected with 6616 iterations, which includes 3762 valid GelSight videos, while the rest 2854 iterations did not generate good data because of inadequate gripping locations. The invalid data, is picked either automatically by GelSight or manually.

A. Property perception

In the experiment of property perception, we use 3762 GelSight videos from the 153 clothing items, and classify the tactile images according to the 11 property labels. The training set includes 2607 videos, the validation set includes 400 videos from the same clothes, and the test set includes 742 videos from novel clothes. We try the networks with either a single image as input, or multiple images from a video as input. The results are shown in Table II.

From the results, we can see that for both seen and novel clothes, the networks can predict the properties with a precision much better than chance. Specifically, the precision gap between the validation set and test set indicates the model overfits to the training set. We suppose the major reasons for the overfit are:

- The dataset size is limited. Although the dataset has a wide variety of clothing types, the number of the clothes in each refined category is small (2 to 5).
- We used 5 GelSight sensors in data collection, and they have some different optical properties, which result in some difference in the images.
- The CNNs are designed for visual images, which is not the optimum for the GelSight images.
- Some properties are not only related to the materials but also the purpose of the clothing. For example, satin is mostly used for summer clothes, but a satin pajama, which feels exactly the same, is worn for all seasons.

We also experiment with other CNN architectures for

| Property   | Chance | Validation set | Test set |
|------------|--------|----------------|----------|
|            |        | Image | Video | Image | Video |
| Thickness  | 0.2    | 0.89  | 0.90  | 0.67  | 0.69  |
| Smoothness | 0.2    | 0.92  | 0.93  | 0.76  | 0.77  |
| Fuzziness  | 0.25   | 0.96  | 0.96  | 0.76  | 0.76  |
| Softness   | 0.5    | 0.95  | 0.95  | 0.72  | 0.76  |
| Stretchness| 0.5    | 0.98  | 0.98  | 0.80  | 0.81  |
| Endurance  | 0.5    | 0.97  | 0.98  | 0.95  | 0.97  |
| Wool surfaced | 0.5   | 0.98  | 0.98  | 0.90  | 0.89  |
| Wind resist | 0.5   | 0.96  | 0.96  | 0.87  | 0.89  |
| Season     | 0.25   | 0.89  | 0.90  | 0.61  | 0.63  |
| Textile type | 0.05 | 0.85  | 0.89  | 0.44  | 0.48  |
| Wash method | 0.17  | 0.87  | 0.92  | 0.53  | 0.56  |
the multi-label classification task, including VGG16 and AlexNet, but the results are not satisfactory. VGG16 performs relatively better. We suppose for the given task of tactile image classification, AlexNet and VGG16 are not deep enough to extract all the useful features.

We believe given enough resource, that we can collect a much larger dataset with more clothing items and gripping iterations, the property perception with unseen clothes will significantly improve. Another possible improving direction is to explore network structures that are more suitable to the tactile images, instead of the ones developed for general images.

B. Exploring planning

We experiment on picking effective gripping locations from the Kinect depth image, using the Kinect images from the 6616 exploration iterations. The images are also divided into the training set, validation set (on the same clothes), and test set (on unseen clothes). On both the validation and test sets, the output of the neural network has a success rate of 0.73 (chance is 0.5). The result indicates the identification of the clothing item has limited influence on the result of gripping location selection. In the training process, the network quickly reaches the point of best performance and starts to overfit. For achieving better results for exploration planning, we plan to develop a more robust grasping system, and collect more data or use online training.

C. Online robotic test

In this experiment, the robot runs the exploration automatically using the depth CNN and tactile CNN. In the exploration, if the property estimation from the tactile CNN is not confident, which is most likely caused by the bad tactile data, the robot will re-do the exploration. The tactile CNN in this experiment is the CNN for single image input (Figure 6(a)).

We experiment on the test clothes (30 items), and each clothes is explored 5 times. The result is shown in Table III. Here we compared the result of ‘without re-trail’ which means the system would not judge the data quality, and ‘with re-trail’. Note that the ‘without re-trail’ results are worse than the results in Table II because the tactile data here is all the raw data generated by the robot, while Table II is only from good data. Another reason is that the gel sensor in this experiment is a different one, and not seen in the training set before, so that there is some slight difference in the lighting distribution. The results also showed that with the re-trails, the precision of property classification increases largely. On average, the robot makes 1.71 trials for each exploration, but 77.42% of the clothes are ‘easy’ for the robot, that it takes less than 2 grasps to get a confident result, and it turns out the property estimation is more precise. The rest clothes are more ‘confusing’, that the robot need to explore them for multiple times, but the properties are still not well recognized.

VI. Conclusion

The perception of object properties is an essential part to make an intelligent robot to understand and interact with the environment, and among the common objects, clothing is an important category. By better understanding clothing properties, the robots will be able to better understand the humans’ lives, and better assist humans in housework. In this paper, we introduce a robotic system that actively perceives the material properties of the common items of clothes. The system uses a GelSight tactile sensor as the sensory input, and recognize 11 properties of the clothes, which helps the robot to have a relatively comprehensive understanding of the clothing material. We use Convolutional Neural Networks (CNN) that are developed for computer vision to classify the GelSight tactile images, and the model can well recognize clothing properties of seen clothes, and effectively recognize the unseen clothes. At the same time, we use a Kinect sensor to guide the robot to explore the clothes in a natural environment, and use a method that combines a hand-crafted model and CNN to find the effective contact locations. To our knowledge, this is the first work on perceiving fine-grained properties of the comprehensive set of common clothes, and provides a new example of active tactile perception of object properties for robots. The system and methodologies in this paper will help robots to better assist humans in clothing related housework, as well as inspiring other works on active tactile sensing on object properties.

ACKNOWLEDGMENT

The work is financially supported by Toyota Research Institute. The authors sincerely thank John-Elmer Canfield and Christopher Cullity for the help on building the clothes dataset and labeling the clothing properties, thank Siyuan Dong for help building the GelSight sensor, and thank Bolei Zhou for helpful discussions.

REFERENCES

[1] S. J. Lederman and R. L. Klatzky, “Extracting object properties through haptic exploration,” Acta psychologica, vol. 84, no. 1, pp. 29–40, 1993.
[2] W. M. B. Tiest, “Tactual perception of material properties,” Vision research, vol. 50, no. 24, 2010.
[3] V. Chu, I. McMahon, L. Riano, C. G. McDonald, Q. He, J. M. Perez-Tejada, M. Arrigo, N. Fitter, J. C. Nappo,
T. Darrell, and K. Kuchenbecker, “Using robotic exploratory procedures to learn the meaning of haptic adjectives,” in *Robotics and Automation (ICRA)*, 2013 IEEE International Conference on. IEEE, 2013.

[4] D. Xu, G. E. Loeb, and J. A. Fishel, “Tactile identification of objects using bayesian exploration,” in *Robotics and Automation (ICRA)*, 2013 IEEE International Conference on. IEEE, 2013, pp. 3056–3061.

[5] M. K. Johnson and E. Adelson, “Retrographic sensing for the measurement of surface texture and shape,” in *Computer Vision and Pattern Recognition (CVPR), 2009 IEEE Conference on*. IEEE, 2009.

[6] M. K. Johnson, F. Cole, A. Raj, and E. H. Adelson, “Microgeometry capture using an elastomeric sensor,” in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4. ACM, 2011, p. 46.

[7] B. Willimon, S. Birchfield, and I. Walker, “Classification of clothing using interactive perception,” in *Robotics and Automation (ICRA)*, 2011 IEEE International Conference on. IEEE, 2011, pp. 1862–1868.

[8] Y. Li, C.-F. Chen, and P. K. Allen, “Recognition of deformable object category and pose,” in *Robotics and Automation (ICRA)*, 2014 IEEE International Conference on. IEEE, 2014, pp. 5558–5564.

[9] A. Gabas, E. Corona, G. Alenyà, and C. Torras, “Robot-aided cloth classification using depth information and cnns,” in *International Conference on Articulated Motion and Deformable Objects*. Springer, 2016.

[10] L. Sun, S. Rogers, G. Aragon-Camarasa, and J. P. Siebert, “Recognising the clothing categories from free-configuration using gaussian-process-based interactive perception,” in *Robotics and Automation (ICRA)*, 2016 IEEE International Conference on. IEEE, 2016.

[11] C. Kampouris, I. Mariolis, G. Peleka, E. Skartados, A. Kargakos, D. Triantafyllou, and S. Malassiotis, “Multi-sensorial and explorative recognition of garments and their material properties in unconstrained environment,” in *Robotics and Automation (ICRA)*, 2016 IEEE International Conference on. IEEE, 2016.

[12] R. S. Dahiy, G. Metta, M. Valle, and G. Sandini, “Tactile sensing from humans to humanoids,” *IEEE Transactions on Robotics*, vol. 26, no. 1, 2010.

[13] Z. Kappassov, J.-A. Corrales, and V. Perdereau, “Tactile sensing in dexterous robot hands,” *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[14] Z. Pezzementi, E. Plaku, C. Reyda, and G. D. Hager, “Tactile-object recognition from appearance information,” *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 473–487, 2011.

[15] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, “Measurement of shear and slip with a gelsight tactile sensor,” in *Robotics and Automation (ICRA)*, 2015 IEEE International Conference on. IEEE, 2015.

[16] R. Li and E. Adelson, “Sensing and recognizing surface textures using a gelsight sensor,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1241–1247.