MagicHand: In-Hand Perception of Object Characteristics for Dexterous Manipulation

Hui Li, Yimesker Yihun, and Hongsheng He

Wichita State University, Wichita KS 67260, USA
hongsheng.he@wichita.edu

Abstract. An important challenge in dexterous grasping and manipulation is to perceive the characteristics of an object such as fragility, rigidity, texture, mass and density etc. In this paper, a novel way is proposed to find these important characteristics that help in deciding grasping strategies. We collected Near-infrared (NIR) spectra of objects, classified the spectra to perceive their materials and then looked up the characteristics of the perceived material in a material-to-characteristics table. NIR spectra of six materials including ceramic, stainless steel, wood, cardboard, plastic and glass were collected using SCiO sensor. A Multi-Layer Perceptron (MLP) Neural Networks was implemented to classify the spectra. Also a material-to-characteristics table was established to map the perceived material to their characteristics. The experiment results achieve 99.96% accuracy on material recognition. In addition, a grasping experiment was performed, a robotic hand was trying to grasp two objects which shared similar shapes but made of different materials. The results showed that the robotic hand was able to improve grasping strategies based on characteristics perceived by our algorithm.

Keywords: Object characteristics identification · dexterous grasping · NIR spectrum · Neural Network.

1 Introduction

Dexterous grasping is a required ability for many tasks that are expected to be performed by robots, ranging from the assembly of cars in an automobile factory to manipulation of cooking utensils in a kitchen. A successful dexterous grasping system is essential to apply sufficient contact forces to the object and maintain grasp stability which requires the fingertips are placed on the relevant point of an object. To achieve this, two sub-problems are addressed: grasp planning and grasp execution. Considerable progress has been made in development of algorithms for efficient grasp planning [2, 3, 6]. Algorithms have been implemented to generate grasps in structured or unstructured environments for known, partially known or unknown objects. Despite these achievements, proposing grasp planning algorithms based on characteristics of an object remains a challenge.

Nowadays most object recognition methods are based on images or videos of an object [5]. An object is first recognized, the shape of that object is determined and then the grasping strategy is selected based on the shape of the object. However, there are two problems with this approach. Firstly, from pictures or videos of an object, useful properties which would affect the grasping such as mass, density, hardness and fragility can be hardly perceived. Secondly, when working in natural environments, there could be situations where the lighting conditions are too dim and dark to get a good picture or video for the algorithms to identify the object.

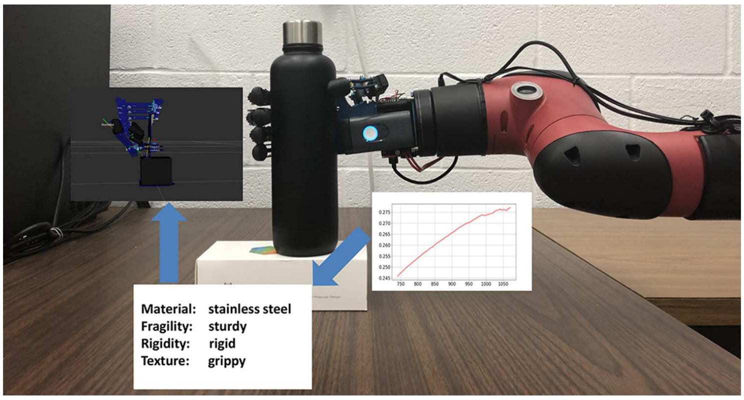

Fig. 1. Deciding grasping strategies based on characteristics of object: SCiO sensor takes NIR spectrum from object, then the material of the object are perceived based on the NIR spectrum and three characteristics are mapped to the perceived material. At last grasping strategies is decided based on characteristics of the object.

Near-infrared spectroscopy (NIR or NIRS) is a low-cost, simple, fast and nondestructive technique to analyze the spectrum of the material on molecular level. Recently, near infrared (NIR) spectroscopy, due to its fast analysis, good precision and accuracy for multi-parameters, is increasingly becoming one of the most efficient analytical tools [4, 8, 1]. In this paper, we propose a novel way to use NIR spectrum to recognize the properties of objects with specific intent of using the information for dexterous grasping. NIR can work functionally under dark environments and the characteristics of NIR allow to determime properties directly from an object. A multiplayer perceptron (MLP) neural network was implemented to fit the NIR data. A MLP is a class of feed forward artificial neural network which consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP neural network is very efficient for tabular dataset and classification prediction problems which fits the problem in this paper very well.

This technique could be helpful in many social aspects. For example, in manufacturing industry, one of the biggest issues in manufacturing and automation is safty. Knowing the characteristics of an object before hand, a robot hand is able to perfrom different grasp strategies on different objects. Much lower grasp force will apply to a human hand than a metal part. In this way many safty issues can be solved.

2 Characteristics Recognition

In this section, we describe the complete workflow of the proposed system step by step. In 2.1, a spectral pre-processing method is described. A new Multilayer Perceptron (MLP) neutral network is implemented to classify the data by materials of the objects in 2.2. In 2.3, a material-to-characteristics table was established to map characteristics to perceived material.

2.1 Data Pre-processing

The goal of pre-processing is to remove extraneous signals such as background offsets and noise while enhancing spectral features. However, in this work, all the data that used for training and validating the algorithm are collected by hand with a light proof shade (Fig. 2a) in an ideal environment. The background offsets and noise could be neglected. In this case, the data set is standardized to fit the MLP neutral network. This step is to constrain all the data in the same scale to improve learning ability of an algorithm.

a) Experiment data acquisition *b) Practical data acquisition*

Fig. 2. Spectrum acquisition of experiment spectrum (dataset A) and practical spectrum (dataset B)

2.2 Materials Classification

Multilayer Perceptron (MLP) neutral network is a classical and common type of artificial neutral network. This network is flexible, fast and especially efficient for tabular datasets and classification prediction problems. All these features fit the problem in this paper very well.

In this paper, a six layers MLP neutral network was developed specifically for the NIR datasets. Rectified Linear Units (ReLU) is used as activation function in the first five layers to train the network. In the first layer, instead of 331 neurons (number of features in data set), 2^{10} neurons was used so that more parameters can be generated to improve the accuracy. In the next four layers, for each layer, the number of neurons were decreased by half (2^{10},2^9, 2^8, 2^5) to further calibrate the model. In the last layer (the output layer) a Softmax function and six neurons are used to classify the data.

2.3 Characteristics Mapping

In this part, in order to map the characteristics of a material, a material-to-characteristics table was established base on Rao's object dataset. Three characteristics including fragility, rigidity and texture were extracted and assigned to each corresponding materials. After the material of an object is perceived, Three characteristics (fragility, rigidity and texture) values are mapped to that material. These information will then feed into Rao's grasping algorithm to improve the grasping strategies.

Table 1. Characteristics of material mapping List

Materials	Fragility	Rigidity	Texture
Wood	sturdy	rigid	rough
Glass	medium	rigid	smooth
Plastic	medium	soft	smooth
Ceramic	sturdy	rigid	slippery
Cardboard	medium	rigid	rough
Stainless Steel	sturdy	rigid	grippy

3 Experiment Evaluation

In this section, we investigated the accuracy of the material detection and properties mapping using the proposed method. The experiment uses two datasets to evaluate the developed algorithm in real world environment. The purpose of this experiment is to evaluate our algorithm in real world environment.

3.1 Experiment Setup

In this experiment, a SCiO sensor (Fig.3a), an AR10 robotic hand (Fig.3b) and a Sawyer robotic arm (Fig.3c) are assembled together as shown in Fig.3d. An object holder was attached to a fixed point on the work desk. A DELL OPTIPLEX 9010 workstation was used to control the whole system. On the software side, Robot Operating System (ROS), Sawyer SDK, AR10 control package, python and SCiO Lab were used in this experiment.

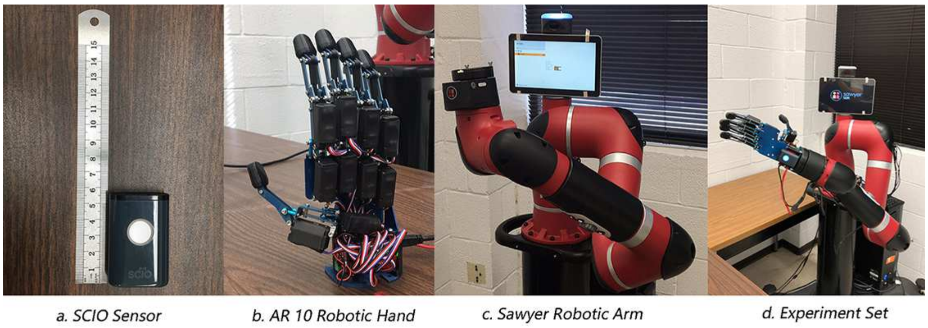

Fig. 3. Equipment Set: a) SCiO sensor: a small, hand-held near-infrared spectrometer which can detect signature on molecular level. b) AR10 Robotic Hand: a lightweight, reliable, powerful, servo-actuated robotic hand with 10 degree of freedom (DOF). c) Sawyer Robotic Arm: a high performance robot arm with 7 degree of freedom (DOF). d) Experiment Set: the AR10 hand is attached to Sawyer Arm and the SCiO sensor is fixed to wrist of the arm.

3.2 Datasets

In this paper, two datasets were collected. Dataset A is collected manually in an ideal environment, this dataset is used to train the MLP neutral network. Dataset B is collected by robotic hand in a practical work environment. This dataset is used to evaluate the performance of the algorithm in real world environments. For both datasets, NIR spectra of six different materials including ceramic (label 0), stainless steel (label 1), wood (label 2), cardboard (label 3), plastic (label 4) and glass (label 5) were collected. Each spectrum includes 331 features which are reflectance intensities of the surface of the object from wavelength 740 nm to 1071 nm. Fig. 4 and Fig. 5 show the spectrum of the two datasets.

Sample Fields In dataset A, we collected ten objects each for ceramic, wood, plastic and cardboard, nine objects for stainless steel and five objects for glass. For each object, ten samples were selected for spectrum collection. For samples of large size such as plates, stainless steel pans and cardboard boxes, around 30 scans were collected on each sample. Fewer scans were taken on small samples such as spoons, forks and plastic cups. In total, 15936 spectra were collected and most of samples were objects of daily use.

In dataset B, two objects for each material are selected. For each object two samples were chosen and five scans were taken from each sample. In total, there were twelve objects, twelve samples and 60 spectra in dataset B.

Sample Acquisition The data collecting system for dataset A includes a SCiO sensor, a light proof shade and a calibration device. All data in dataset A is collected by hand in a lab environment (room temperature 75 Fahrenheit degrees). While acquisition, the sample was placed steadily on a table in the lab.

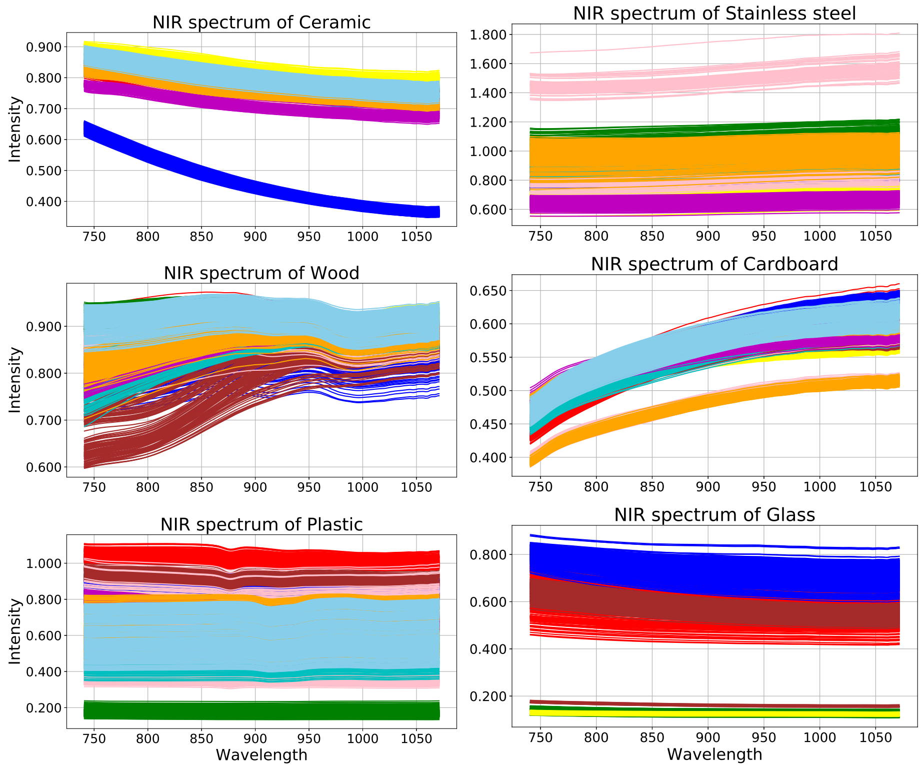

Fig. 4. NIR spectrum of dataset A

The shade was attached to the sensor. While scanning, the sensor was kept as close to the surface of the sample as possible (Fig. 2a). The acquisition time was around two seconds for each scan. Each sample was scanned about 30 times in different locations. The senor was re-calibrated after scanning each sample.

Dataset B was acquired using the same data collection approach except that the SCiO sensor was attached to the wrist of Sawyer arm (Fig. 3d). An object holder was placed on a fixed point and the sample was placed on the object holder. The spectrum data of that sample was collected using the robot hand (Fig. 2b). Five different scans were taken from each sample and the sensor was calibrated each time it finished scanning a sample.

3.3 Results and Discussion

Evaluating the Classifier on Ideal Dataset The algorithm was trained and evaluated using the spectra in an ideal dataset (dataset A) which was collected manually. First, we extract one object from each material in the dataset. This part of the data (1803 scans) was separated and used as the validation set to evaluate the final performance of the classifier. The rest of data in dataset A

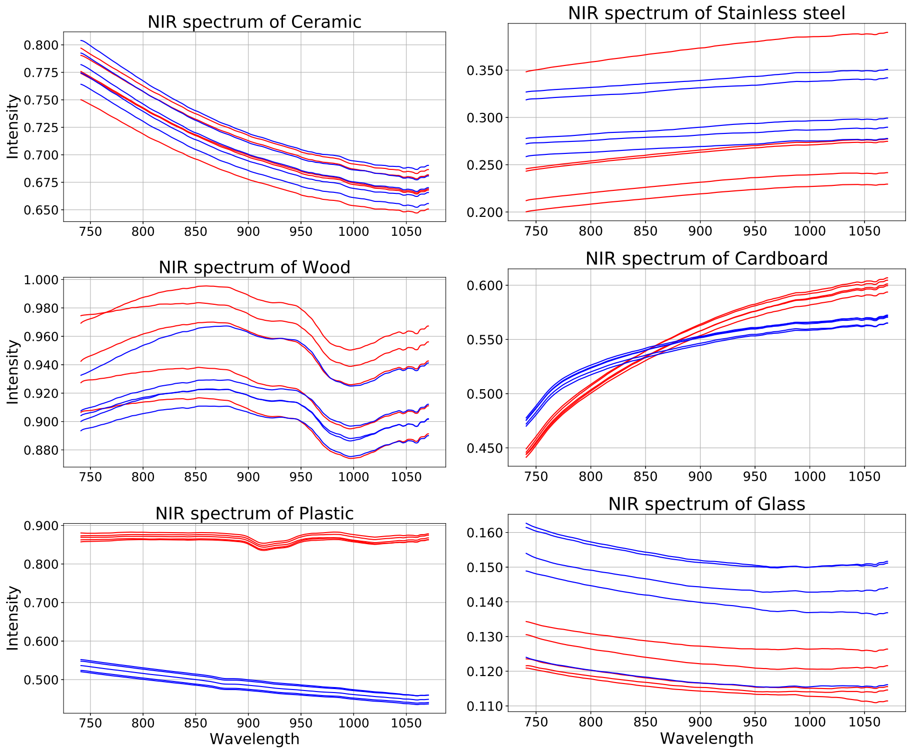

Fig. 5. NIR spectrum of dataset B

was divided into training and testing set. 70% of the data was selected randomly to train the network and the rest 30% of the data was used as the testing set. We tuned the hyper-parameters as follows: 64 batch size, 50 epochs and ADAM optimizer with 0.0001 learning rate. The algorithm achieved overall 99.96% accuracy.

To evaluate the performance of the network, the pre-extracted validation data (1803 scans), which was completely new for the algorithm, was fed into the network. A total 1609 correct predictions were achieved by the algorithm (Table 2). The algorithm made perfect predictions (100% accuracy) on ceramic, stainless steel, wood and cardboard and a good predictions on plastic (84% accuracy). However, predictions on glass (51.3% accuracy) were not satisfactory.

As shown in Fig. 4, the wave shape of ceramic, stainless steel, wood and cardboard are obviously different from each other. In the mean time, for those four materials, the wave shape of different objects that made from the same material are very similar. In this case, for these four materials, it is obvious that the algorithm can make correct predictions easily. Since plastic is a broad definition, different plastic products may have different chemical compositions. Hence, as seen in Fig. 4, the plastic objects have different wave shapes. In this

Table 2. Predicted results for dataset A

Material	Label	Correct classification	Accuracy
Ceramic	0	303	100%
Stainless steel	1	300	100%
Wood	2	300	100%
Cardboard	3	300	100%
Plastic	4	252	84%
Glass	5	154	51.3%
Total result		1609	89.2%

case, when classifying a new plastic object, the wave shape of that object could be out of the algorithm's knowledge base. This leads to a lower performance on plastics recognition. Being a transparent material, the spectrum of a glass object is affected by many external factors such as the background, the shape of the object or the thickness of the glass. These makes the spectrum spread over a wider range which leads to mis-classification.

Table 3. Predicted results for dataset B

Material	Label	Correct classification	Accuracy
Ceramic	0	10	100%
Stainless steel	1	10	100%
Wood	2	10	100%
Cardboard	3	10	100%
Plastic	4	10	100%
Glass	5	10	100%
Total result		60	100%

Evaluating the Classifier on Practical Dataset The proposed algorithm was evaluated on a practical dataset (dataset B). Using the data collected by robotic hand, the prediction result (Table 3) achieved a 100% (60/60) accuracy. There are two reasons that may lead to this perfect result. Firstly, the light proof shade greatly reduced external noises which is the largest difference between a practical and an ideal environment. Secondly, there were only sixty testing samples in this experiment, As the number of testing samples increase, the accuracy will be more practical. However, the current result is adequate to show that this algorithm is capable of working in a practical environment.

Characteristics Mapping In this part, a grasping experiment was performed on two objects, one plastic tube and one stainless steel tube. These two tubes have similar shapes but different characteristics (Fig. 6). From the characteristics we can infer that the two tubes may require different grasping methods. According to Rao's work [7], Precision-Prismatic Grasp (Fig. 7 a) should be used to

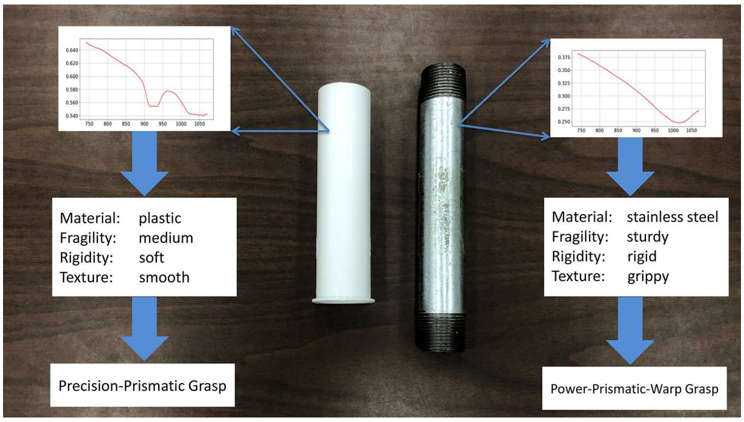

Fig. 6. Different grasping methods based on different characteristics

pick up the plastic tube and Power-Prismatic-Warp Grasp (Fig. 7 b) should be applied to stainless steel tube. In the experiment, the spectra of the tubes were collected by the robot hand and fed into our system to perceive their materials. Then characteristics were mapped to each material. Based on the characteristics, the robotic hand applied Precision-Prismatic Grasp on plastic tube and Power-Prismatic-Warp Grasp on stainless steel tube which matched our expectations.

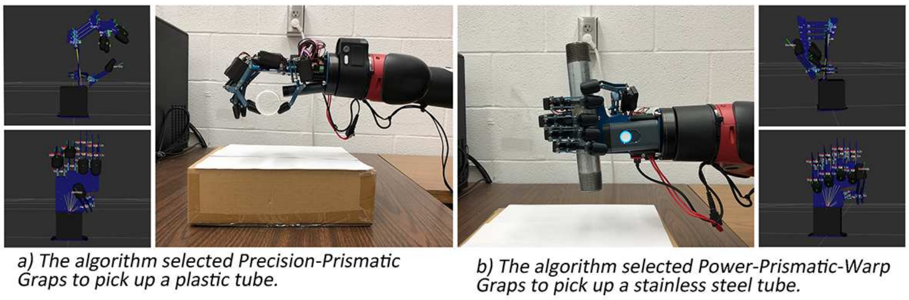

Fig. 7. The algorithm selected Precision-Prismatic Grasp to pick up a plastic tube

4 Conclusions

This paper has proposed a novel way to use NIR spectrum to perceive object characteristics for dexterous manipulation. The experiment results show that our algorithm works well on both ideal and practical NIR datasets on material recognition and characteristics mapping. The proposed algorithm shows great performances on material recognition and characteristics mapping in both ideal and practical environments. Our method will greatly aid grasping decisions for dexterous grasping by robots.

References

1. Balabin, R.M., Safieva, R.Z., Lomakina, E.I.: Comparison of linear and nonlinear calibration models based on near infrared (nir) spectroscopy data for gasoline properties prediction. Chemometrics and intelligent laboratory systems **88**(2), 183–188 (2007)
2. Brost, R.C.: Automatic grasp planning in the presence of uncertainty. The International Journal of Robotics Research **7**(1), 3–17 (1988)
3. Goldfeder, C., Allen, P.K., Lackner, C., Pelossof, R.: Grasp planning via decomposition trees. In: Robotics and Automation, 2007 IEEE International Conference on. pp. 4679–4684. IEEE (2007)
4. Li, J., Huang, W., Zhao, C., Zhang, B.: A comparative study for the quantitative determination of soluble solids content, ph and firmness of pears by vis/nir spectroscopy. Journal of Food Engineering **116**(2), 324–332 (2013)
5. Lowe, D.G.: Object recognition from local scale-invariant features. In: Computer vision, 1999. The proceedings of the seventh IEEE international conference on. vol. 2, pp. 1150–1157. Ieee (1999)
6. Miller, A.T., Knoop, S., Christensen, H.I., Allen, P.K.: Automatic grasp planning using shape primitives. In: Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on. vol. 2, pp. 1824–1829. IEEE (2003)
7. Rao, B.: Learning robotic grasping strategy based on natural-language object descriptions (2018)
8. Wilson, B.K., Kaur, H., Allan, E.L., Lozama, A., Bell, D.: A new handheld device for the detection of falsified medicines: demonstration on falsified artemisinin-based therapies from the field. The American journal of tropical medicine and hygiene **96**(5), 1117–1123 (2017)