

Artificial haptic recognition through human manipulation of objects

David Miralles*¹, Carlota Parés*, Guillem Garrofé*, Alberto Soto*, Albert Llauradó*

Gerard Serra*, Àlex Falcó*, Hans Op de Beeck†, Haemy Lee Masson†

*Grup de Recerca en Tecnologies Media, La Salle - Universitat Ramon Llull, Barcelona 08022, Catalonia, Spain

†Laboratory of Biological Psychology, KU Leuven, Leuven 3000, Belgium

Abstract

Object recognition has been extensively explored in the computer vision literature, and over the last few years the results in this field have sometimes even surpassed human performance. One of the main reasons for this success is the growing number of images available to generate training datasets for machine learning. In comparison to computer vision, haptic approaches to object recognition have received relatively little attention, probably due to the inadequacy of available sensors to gather the huge amount of data needed to train the modern machine learning algorithms. Consequently, the performance of machine haptic recognition of objects is still far from being comparable with humans. In this paper, we first present a new sensor system capable of capturing part of the information that humans produce during the haptic manipulation of objects and an artificial haptic intelligence that classifies shapes from the dataset created by the sensor system. Secondly, we compare the haptic object recognition performance between humans and a machine. The current study sheds new light upon the novel approach used to capture human haptic exploration and provides evidence that artificial haptic intelligence outperforms human haptic recognition abilities.

Keywords: haptics; shape processing; artificial haptic recognition;

Introduction

Robotic systems need to have capabilities similar to the human haptic system in order to perform complex grasping, manipulation, and object recognition tasks using dexterous hands. Currently, according to the type of data acquisition, the methods of haptic object recognition can be divided into two categories: based on the distributions of contact points (Allen & Roberts, 1989; Meier, Schopfer, Haschke, & Ritter, 2011; Navarro et al., 2012); and based on the pressure patterns in tactile arrays (Lin, Calandra, & Levine, 2019; Luo, Mou, Althoefer, & Liu, 2019). None of them gets accuracies or reaction times similar to humans. In fact, they do not even try to compare their algorithms results with humans.

On the other hand, in humans, tactile sense is the earliest developing sensory organ enabling an infant to actively explore the world. Previous investigations have demonstrated that humans are surprisingly good at judging shape of the

object using the haptic modality (Klatzky, Lederman, & Metzger, 1985). Our abilities to use the sense of touch for identifying and categorising the shapes are well supported by extensive structural and functional brain architecture implicated in haptic processing (Masson, Bulthé, De Beeck, & Wallraven, 2016; Masson, Kang, Petit, & Wallraven, 2018). In the current study, we present a novel method used to capture haptic shape exploration of humans with which an artificial agent can be trained. Afterwards, we compare the haptic shape recognition abilities between trained machine and humans.

Materials and Methods

Stimuli

Stimuli used in this experiment were modelled with the 3D extension of the Superformula (Gielis, 2003), a formula proposed by Johan Gielis in 2003 that is supposed to describe many complex shapes and curves that are found in nature.

$$\frac{1}{r} = \sqrt[n_1]{\left|\frac{1}{a} \cos\left(\frac{m\phi}{4}\right)\right|^{n_2} + \left|\frac{1}{b} \sin\left(\frac{m\phi}{4}\right)\right|^{n_3}} \quad (1)$$

3D parametric surfaces are obtained by multiplying two Superformulas, r_1 and r_2 . To generate a family of smoothly varying shapes as a set of stimuli, starting from a rounded cube, we have modified the values of exponent n_2 in latitude and longitude. The 2D images of resulting 9 stimuli are shown in Figure 1.

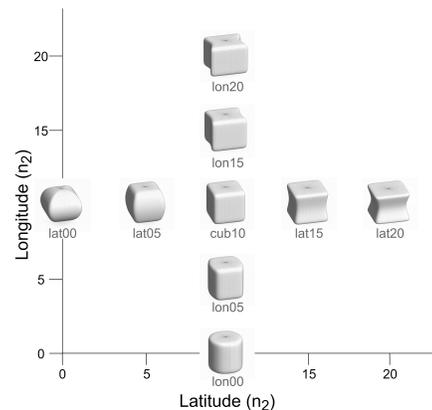


Figure 1: Images of 3D printed shapes.

The stimuli were printed out as tangible objects on a 3D printer (BQ, Witbox2, Spain). In order to capture the human

¹david.miralles@salle.url.edu

haptic object manipulation, we attached twenty-four copper pads equally distributed to the surface of each object (Figure 2). Each pad is connected to an electronic board placed inside each object, acting as a capacitive sensor.



Figure 2: Objects covered with copper pads connected to an electronic board placed inside. Each face is tagged with a letter.

Data Acquisition

The electronic board placed inside the object collects data from all the sensors (touched or not touched) and sends it to a computer via Bluetooth. The sampling frequency is 40Hz. For every sample, an array of 24 elements, each one being 1 (not touched) or 0 (touched), is received. Having sensors numbered from 0 to 23, we can observe, in this example (Figure 3), that sensors 2,5,6 and 7 are in the touched state while the rest remain untouched.

1 1 0 1 1 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1

Figure 3: An example of a sample obtained from 24 sensors attached to a stimulus.

To train our haptic system, we collected an independent dataset containing 20 minutes (15 for training and 5 for testing) of human haptic exploration per object. To obtain a complete exploration of the objects and to standardize haptic exploratory procedure over time, we instructed a participant to perform a task. The task is to match a letter displayed on the monitor in front of a participant with a letter on the top surface of the object for each trial. For this task, objects faces were tagged with a letter (see Figure 2).

During the task, a participant who sat in a comfortable chair had to explore the object with both hands looking for the displayed letter on the monitor at the object surface (Figure 4). Once a letter was found, a participant was instructed to push a button placed under his feet to display a new letter for the next trial. The same letter never appeared consecutively. Four series each lasting for five minutes have been recorded per object, giving a total of 36 (4 series, 9 objects) data logs with 12000 lines each.



Figure 4: Participant exploring an object.

Machine Learning Algorithm

To process the captured tactile signal, a naive Bayes classifier capable of differentiating the 9 stimuli was implemented. We have split the dataset into two halves: one for training the classifier and one for testing the accuracy of the machine. The dataset for training contains three series out of four recorded per stimulus, while the testing dataset contains one remaining series per stimulus.

From now on, every sample, a line of a log file containing an array of 24 binary digits, will be called state. Each state represents which copper pads of the object were touched in a moment in time (every 25 ms). From the 2^{24} possible states, humans, while exploring the geometry of each stimulus, generate distinctive states over time per stimulus. To prevent the machine from classifying objects based on how the sensors are placed, we installed sensors in the same order for all objects. This was possible because our set of stimuli shares similar geometry. Thus, if equivalent sensors were frequently touched in both stimuli, it means that these stimuli share similar states.

Bayes classifier determines, given a series of states, the probability of states belonging to a stimulus, i.e., $P(y | x_1, \dots, x_n)$ where x_i is a state configuration, n is the number of consecutive states ($n/40$ secs) and y is a label identifying each stimulus. The resulting prediction \hat{y} is:

$$\hat{y} = \arg \max_{y \in [1,9]} \left[P(y) \cdot \prod_{i=1}^n P(x_i | y) \right] \quad (2)$$

$P(x_i | y)$ is the probability of finding a x_i state given a stimulus y . We could obtain this probability from our dataset. The n -product is the naive condition, and $P(y)$ is the probability of every stimulus, in this case $1/9$. Since conditional probabilities for each stimulus, given a state, are very small, when multiplied together, they result in very small values, which can lead to floating point underflow. To fix this, we used the logarithm of the probabilities instead of the raw probabilities on its implementation. Moreover, some states that appeared on testing data could be new to the classifier. In this case, we added this new state with a probability of 0. However, not having a certain state in the training data for a certain stimulus does not guarantee that the testing data does not belong to that stimulus. Based on this logic, and to avoid the possibility of the probability becoming 0, we applied Laplace smoothing.

Human Haptic Experiment

To determine to which extent humans can classify objects based only on their haptic experiences, we conducted a haptic experiment. Adults (N=15 (male = 9), the range of age = 20-37 years old) with no prior diagnosis of neurological or perceptual deficits took part in this experiment. The experiment consists of two sessions. For both sessions, participants were blindfolded and explored each stimulus with both hands. The first session is the training session in which participants were instructed to explore each object for 10 seconds and to give a label/tag. Importantly, they were informed that the tags were going to be used to identify each stimulus in the testing session. When the training session started, to give a participant an idea of the range of objects shape used in the experiment, the first object given to a participant was always the cub10 and the second object given was always the lon20 that was perceived most differently to cub10 in our previous pilot study. The rest of the stimuli were presented in random order.

After a brief break, participants performed the haptic object recognition task during the testing session. The objects were given one by one in random order. In this session, participants were instructed to take their time and give a corresponding tag as correctly as possible. We use both accuracy and reaction time (i.e., the time taken to identify the object) as indices of human performance.

Results

Machine Haptic Recognition

To obtain the accuracy of the classifier over time, we split the testing dataset into multiple files with their sizes being proportional to the time of testing. Then, every subfile was classified using the presented algorithm. In the end, accuracy was computed as the number of times the classifier was able to classify correctly the subfile over the total number of subfiles tested. The results for each stimulus are shown in Figure 6.

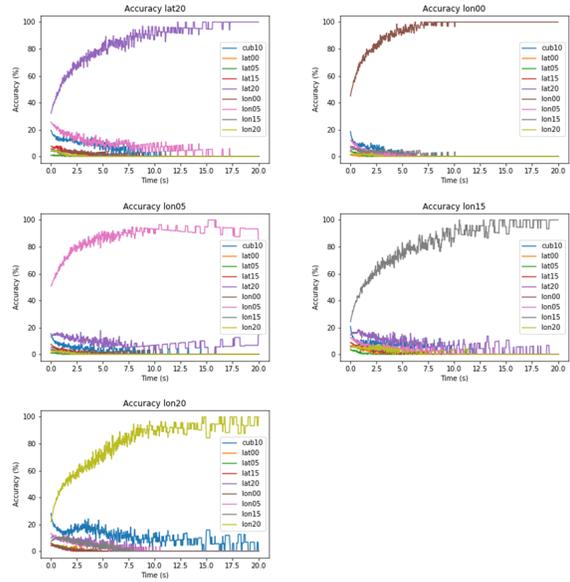
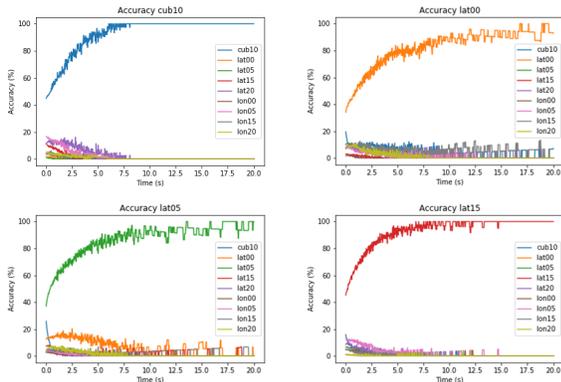


Figure 6: Naive Bayes classifier accuracy per stimulus versus time (secs).

The results demonstrate that classification accuracy is all above chance level across 9 stimuli (Figure 6). The performance of machine haptic recognition will be compared to human performance in the following section.

Human Haptic Recognition

When computing the average reaction time per participant, two participants with outlying reaction time were detected and removed from the following graphs and further analysis. The resulting average accuracy and reaction time across objects per participant are shown in the figures 7 and 8.

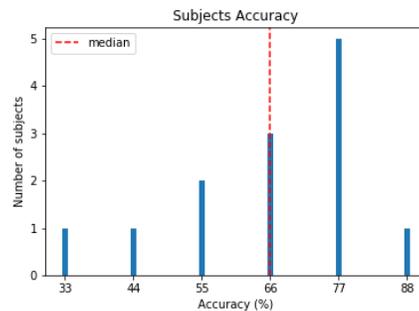


Figure 7: Average accuracy per participant.

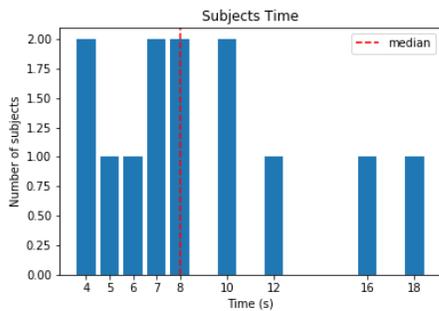


Figure 8: Average time per participant.

We observed that all participants were capable of identifying each object above the chance level (i.e., 11.1%). The group median accuracy is 66% and the reaction time is 8 seconds.

Comparison Between Machine and Human Haptic Recognition

Our human participants spent an average of 8 seconds to identify an object. For the machine trained with Bayes classifier, the performance at 8 seconds is 89.19% (Figure 6), which is higher than 66% (95% confidence interval=44-77%). Considering that the upper end of the confidence interval value of human accuracy (77%) is smaller than 89%, the machine seems to perform haptic object recognition task better than humans. In terms of accuracy, the machine was able to achieve the same accuracy (i.e., 66%) as humans in 1.75 seconds (8 seconds for humans; 95% confidence interval=6-12 seconds).

Conclusions

For the first time, a haptic recognition comparison between machine and human has been made. The results are clear and striking: the machine is capable of classifying novel 3D shapes much better than humans. It is important to consider that our algorithm uses just a small part of the data generated during human manipulation of shapes and the system does not know any geometrical relationship among sensors attached to different parts of the object. In this study, we first demonstrate that our novel methods can be used to investigate haptic recognition in both a machine and a human. It is worthwhile to mention that the results from human haptic experiments need to be analysed more profoundly. We have observed that human recognition performance across shapes does not follow a normal distribution, which could be due to two possible reasons. The first reason could be related to an issue of memory overload because of the high number of objects that are very similar to each other. Another reason could be due to a difference in the degree of difficulty in recognition across objects. For instance, a couple of objects that are familiar to humans, cub10 (cube) and lon00 (cylinder). They could be considered almost as outliers in terms of accuracy and reaction time compared with the other less familiar

shapes (Klatzky et al., 1985). In contrast, the performance of the proposed system did not show such bias.

Although, the proposed system and humans are able to classify objects according to their shapes, this does not mean that they do so under the same parameters or premises. As mentioned above, our system is trained by using only a small subset of the data collected during human shape manipulation. Moreover, humans and machines have different learning processes and decision making. In order to understand and compare the haptic internal representation of humans and machines, future study is required.

Acknowledgments

The authors are grateful to all the participants in the experiments. This work has been partially funded with the support of the European Social Fund and the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement of the Catalan Government for the pre-doctoral FI grant No. 2019 FI.BI 00171. David Miralles would like to thank Ronda de Dalt for the large amount of time we shared thinking about this work.

References

- Allen, P., & Roberts, K. (1989). Haptic object recognition using a multi-fingered dextrous hand. *Proceedings, 1989 International Conference on Robotics and Automation*.
- Gielis, J. (2003). A generic geometric transformation that unifies a wide range of natural and abstract shapes. *American Journal of Botany*, 90(3), 333-338.
- Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An expert system. *Perception & psychophysics*, 37(4), 299-302.
- Lin, J., Calandra, R., & Levine, S. (2019). Learning to identify object instances by touch: Tactile recognition via multimodal matching. *arXiv preprint arXiv:1903.03591*.
- Luo, S., Mou, W., Althoefer, K., & Liu, H. (2019). iclap: shape recognition by combining proprioception and touch sensing. *Autonomous Robots*, 43(4), 993-1004.
- Masson, H. L., Bulthé, J., De Beeck, H. P. O., & Wallraven, C. (2016). Visual and haptic shape processing in the human brain: unisensory processing, multisensory convergence, and top-down influences. *Cerebral Cortex*, 26(8), 3402-3412.
- Masson, H. L., Kang, H.-m., Petit, L., & Wallraven, C. (2018). Neuroanatomical correlates of haptic object processing: combined evidence from tractography and functional neuroimaging. *Brain Structure and Function*, 223(2), 619-633.
- Meier, M., Schopfer, M., Haschke, R., & Ritter, H. (2011). A probabilistic approach to tactile shape reconstruction. *IEEE Transactions on Robotics*, 27(3), 630-635.
- Navarro, S. E., Gorges, N., Worn, H., Schill, J., Asfour, T., & Dillmann, R. (2012). Haptic object recognition for multi-fingered robot hands. *2012 IEEE Haptics Symposium (HAPTICS)*.