Designing and Evaluating Interactive Tools for a Robot Hand Collection

Ramzi Abou Chahine
r.abou-chahine@uea.ac.uk
University of East Anglia
Norwich, UK

Sid Padmanabhuni
spadmanabhuni@sfsu.edu
San Francisco State University
San Francisco, California, USA

Pooyan Fazli
pooyan@asu.edu
Arizona State University
Tempe, Arizona, USA

Hasti Seifi
hasti.seifi@asu.edu
Arizona State University
Tempe, Arizona, USA

ABSTRACT

Recent robot collections provide various interactive tools for users to explore and analyze their datasets. Yet, the literature lacks data on how users interact with these collections and which tools can best support their goals. This late-breaking report presents preliminary data on the utility of four interactive tools for accessing a collection of robot hands. The tools include a *gallery* and *similarity* comparison for browsing and filtering existing hands, a *prediction tool* for estimating user impression of hands (e.g., humanlikeness), and a *recommendation* tool suggesting design features (e.g., number of fingers) for achieving a target user impression rating. Data from a user study with 9 novice robotics researchers suggest the users found the tools useful for various tasks and especially appreciated the *gallery* and *recommendation* functionalities for understanding the complex relationships of the data. We discuss the results and outline future steps for developing interface design guidelines for robot collections.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in interaction design**; **Empirical studies in HCI**;

KEYWORDS

robot collections, robot hands, interface design, user experience, user study

ACM Reference Format:
Ramzi Abou Chahine, Sid Padmanabhuni, Pooyan Fazli, and Hasti Seifi. 2023. Designing and Evaluating Interactive Tools for a Robot Hand Collection. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23 Companion), March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3568294.3580099

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden

ACM ISBN 978-1-4503-9970-8/23/03…$15.00
https://doi.org/10.1145/3568294.3580099

1 INTRODUCTION

Hundreds of robots and robot parts (e.g., hands) have been developed over the last decades. Existing designs have a wide range of features (e.g., materials, colors), technical specifications (e.g., degree of freedom), and target applications (e.g., household, factory). Moreover, the robot designs can invoke different user impressions (e.g., humanlike, safe). Understanding the trade-offs of existing robot technologies is important for robotics researchers and practitioners when selecting an existing design or building a new one.

Several robot collections have been developed to showcase the variety of robot designs. For example, IEEE has a large online catalog of various robots, drones, and self-driving cars [2]. Phillips et al. developed the Anthropomorphic RoBOT (ABOT) database including 251 robots with their user ratings of humanlikeness and design features such as the existence of facial, locomotion, surface, and manipulation parts [1, 10]. Reeves et al. compiled images of 342 social robots, asked users to code the existence of 21 features (e.g., has a face, bipedal), and collected user ratings of the warmth and competence of the robots [3, 11]. Kalegina et al. created a dataset of 157 robot faces and coded 76 features for each face to study how variations in facial features may impact user ratings [8]. Similarly, Seifi et al. recently compiled a dataset of 73 robot hands with 15 design features (e.g., number of fingers) and 17 user ratings (e.g., humanlike, safe) and developed regression models to predict user ratings from the hand design features [4, 13, 14].

These online collections provide various interfaces for the users. For example, IEEE Robots provides an interactive gallery of robots that can be sorted by the robot's name, size, release date, type (e.g., humanoids, drones), and country. Users can see technical specifications and videos of a robot's icon in the gallery. Besides a gallery, the ABOT database allows users to filter the collection with sliders indicating the robot scores on humanlikeness or its design features. The ABOT interface also provides a 3D scatterplot of the robot scores and a tool for predicting the humanlikeness user rating for robots. Similarly, the RobotHands database provides three interactive tools: (1) a *gallery tool* listing all the hands, (2) a *similarity tool* with scatter plots of the hands, and (3) a *prediction tool* to estimate user ratings based on the hand design features (e.g., number of fingers). These interfaces provide different means of accessing their underlying datasets. Users may explore these collections to get a sense of the variety of robots designed by academia and industry, analyze trends in the field, or obtain information on a single robot or robot part. The interactive tools (e.g., *gallery*, *similarity*) are inspired by similar interfaces in other domains (e.g., information visualization [9], virtual reality [6]) or envisioned by robotics researchers (e.g., *prediction*) [1, 4]. Yet, the variety of interfaces to robot collections suggests a lack of data on how roboticists use these collections and what interface features they value the most.

We are interested in developing guidelines for the design of interactive tools for robot databases. As a step toward this goal, this late-breaking report focuses on the opinions and interactions of 9 novice robotics researchers with the RobotHands database [4]. We replicated the three existing interfaces to RobotHands (*gallery*, *similarity*, *prediction*) and added a *recommendation tool*. The *recommendation tool* allows users to specify their target user ratings for a hand and receive suggestions on which hand features to change. As such, the *recommendation tool* provides the inverse functionality of the *prediction tool*. The tools can assist users in selecting and purchasing an appropriate robot hand, support the design of new robot hands, and inform customization of an existing robotic hand for a target user impression (e.g., humanlikeness). In a user study, 9 robotics students completed a set of tasks with the tools, explored the tools freely, and rated and commented on the tools.

Our preliminary results suggest that participants found the tools useful for supporting different tasks (mean rating >4 on 1–5 rating scales). The *gallery* received the highest rating, followed by the *recommendation*, *prediction*, and *similarity* tools. The *recommendation tool* received more interest than the *prediction* and *similarity* tools, but such an interface is missing from all existing robot collections. The participants wanted data about the performance of the robot hands alongside user ratings and design features, pointing to another gap in existing robot collections. We discuss these early results and outline our ongoing work on developing guidelines for user interface design for robot databases. The tools are available online at https://robothands.org/.

2 ROBOTHANDS DATASET AND INTERACTIVE TOOLS

After presenting an overview of the RobotHands dataset, we summarize the three existing online tools for the dataset (*gallery*, *similarity*, *prediction*) and detail the fourth tool that we designed for the dataset, the *recommendation tool*. We obtained the code for the RobotHands website from the original authors [4, 14] and extended it to include the *recommendation tool* for our user study.

2.1 RobotHands Dataset

The dataset includes 73 robotic hands with their images, design features, and user ratings. The hands were selected as a representative subset of 371 robot hands for existing robots. The authors identified a set of 15 design features (e.g., number of fingers, color scheme) that can be discerned by a layperson rather than the technical specifications of the hands. These features refer to the visual appearance of the hand ($n = 11$ features), the types of materials used in the hand ($n = 2$), its grasping functionality ($n = 1$), and whether the hand is a commercial product or not ($n = 1$). The authors manually coded the features for all the hands in the dataset. The dataset also includes 17 user ratings for each hand on a 0–100 Likert scale. The ratings capture user impressions of the hand or a robot with this hand (e.g., humanlike, nice) as well as the user emotions (e.g., excited) and comfort in interacting with the hand (e.g., comfortable touching the hand). Finally, the dataset includes 17 regression models that can predict the 17 user ratings from the design features.

2.2 Interface Tools

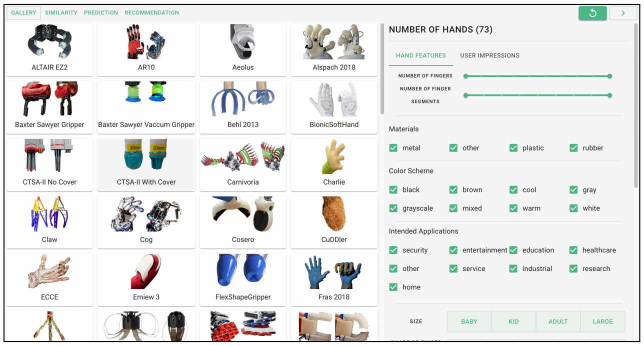
(1) Gallery Tool - This tool features a collection of cards with the names and images of the robot hands. Users can scroll through the collection of the hands in the dataset. A left click on a card opens a pop-up view of all the design features and user ratings for the corresponding robot hand. On the right side, the tool has a panel of filters including the range of design features and user ratings. The user can select a subset of design features or user ratings to narrow down the list of hands in the gallery. The *gallery tool* is commonly used for visualizing various technology datasets [1, 2, 6, 9].

(2) Similarity Tool - This tool allows the users to explore the perceptual space of robot hands using two scatter plots. Seifi et al.conducted a principal component analysis (PCA) on the 17 user ratings for the hands, resulting in three dimensions: *Comfortableness*, *Interestingness*, and *Industrialness* [14]. The values obtained for each hand on these PCA dimensions determine the hand's position on the scatter plots. Hovering over a data point on the plots displays information about the corresponding robot hand. Left-clicking on a point anchors the image of the hand on the top left allowing the user to compare it with other hands. The right panel is identical to *gallery* and allows filtering the hands on the scatter plots.

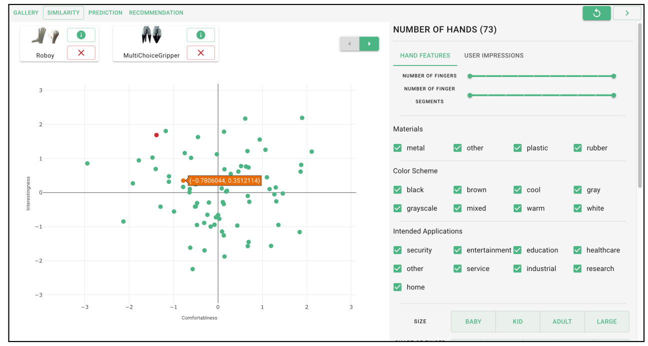
(3) Prediction Tool - This tool allows the user to specify a set of design features for a hand and get estimates of the 17 user ratings for the hand. The user can employ the tool to assess how people may perceive robot hands that are not included in the dataset. We used the linear regression models presented by Seifi et al. for the estimation of user ratings [13, 14]. These 17 models were trained on a dataset of 73 hands and used 5 design features as predictors [13]. No interaction effects among the design features were considered in the regression models. The user can specify the hand features with a set of sliders, radio buttons, buttons, and checkboxes on the left side and see the predicted user ratings in the right panel.

(4) Recommendation Tool - We created this tool to provide the reverse functionality of the *prediction tool*. We anticipated that robot designers might want to modify the features of a robot hand to achieve a target user rating (e.g., make it more humanlike). The *prediction tool* requires the designer to guess relevant design features (e.g., number of fingers) or do trial and error with various feature combinations to test if those features can obtain the target user ratings. The *recommendation tool* is meant to simplify this task. The designer can select a robot hand as a starting point, then specify their target user ratings (e.g., humanlikeness > 70) for the given hand and receive suggestions on the design features that should be modified to achieve user ratings close to their desired targets.

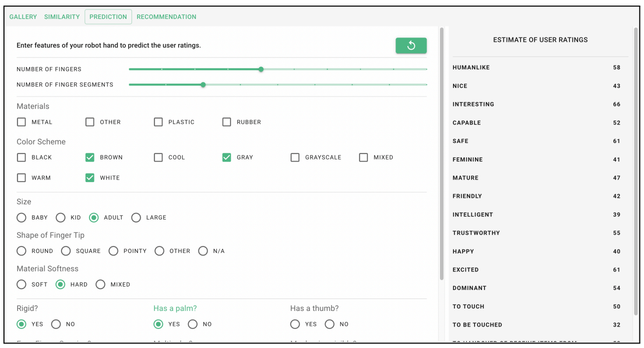
The tool is composed of four vertical panels. The left side panel has image cards for the 73 robot hands. One hand must be selected from this panel as a starting point. The right side panel includes 17 sliders corresponding to the 17 user ratings. The sliders allow the designer to define their target user ratings for the selected hand. After moving a slider, the tool computes the relevant feature change



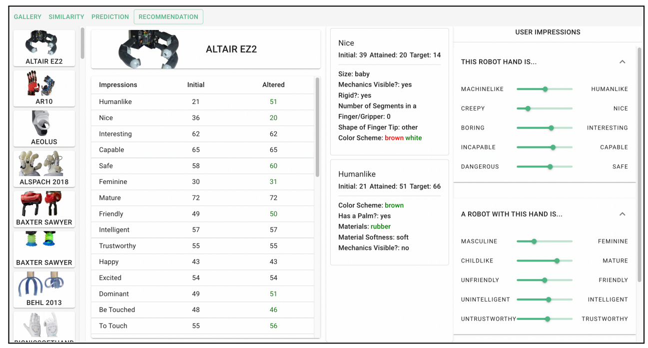
(a) Gallery tool presents a list of 73 robot hands with their images and names on the left and the filters for design features and user ratings on the right side.



(b) Similarity tool includes scatter plots of the robot hands on three dimensions of Comfortableness, Interestingness, and Industrialness.



(c) Prediction tool allows users to enter a set of design features for a robot hand on the left and see estimates of 17 user ratings for the hand on the right.



(d) Recommendation tool allows users to select a robot hand on the left column, adjust the sliders on the right column to indicate a target user rating, and receive suggested modifications to the hand features in the middle.

Figure 1: The four tools for the RobotHands database. We used the first three from [4, 14] and developed a *recommendation tool*.

suggestions for the hand and presents them in the middle columns. The middle left column shows the initial hand features and ratings as well as the modified features and new ratings. The middle right column provides a log of the user interactions with the sliders and the corresponding feature change suggestions. With each slider modification, a new entry is added to the log column.

To compute the feature change suggestions, we model the problem as a knapsack problem [12] and solve it with a greedy approach using the linear regression models from the *prediction tool*. For instance, assume the user specifies a target user rating of humanlikeness = 70 for a hand with a user rating of 50. The algorithm first determines the amount of change in the user rating, which is 20 in this case. Then the algorithm uses the coefficients of the design features for the corresponding linear regression model with a greedy strategy to determine whether to remove, add, or replace features to attain the closest user rating value to the specified target. After a design feature is selected by the algorithm, the humanlikeness user rating is recomputed and the algorithm iterates to identify the next feature change. If the target user rating cannot be reached given the hand features and regression models, the tool provides the closest user rating to the target that is attainable. After the relevant design features are obtained with this greedy algorithm, the tool recomputes all 17 user ratings with the newly-suggested design features and presents them to the user in the middle columns.

3 USER STUDY

We conducted a user study to evaluate the utility of the four interactive tools and their strengths and weaknesses in supporting user exploration of the dataset. To obtain insights into user interactions and opinions, we opted for a qualitative user study using a relaxed think-aloud protocol [7]. We recruited 9 students (2 female, 7 male) with backgrounds in robotics or a relevant field via the university mailing lists. The participants were between 22 to 30 years old (mean = 25.1±2.52). They were originally from India (n = 6 participants), United States (n = 2), and Iran (n = 1) and had 1–7 years of robotics experience (mean = 2.7±1.86) . Each study session took about one hour over the Zoom video conferencing tool. The sessions were video recorded. The participants received $15 USD as compensation.

After obtaining informed consent, the participants filled a pre-questionnaire about their background and a recent project with robots. Next, the experimenter briefly explained the dataset and each of the four tools. The participants were asked to complete a hands-on task immediately after they were introduced to each tool. For example, the task for the *prediction tool* was: "Find the humanlikness of a rubber hand with four fingers, including a thumb, brown color, and a palm." These tasks were meant to familiarize the participant with the functionality of the tool and provide hands-on training. After the completion of the introductory tasks, the participants could explore the tools freely for 10–15 minutes before filling out a post-questionnaire. The participant rated the usefulness of each tool on a scale of 1 (not useful at all) to 5 (very useful) and provided comments on the tool's strengths and future improvements.

3.1 Results

We analyzed user ratings and comments for the tools and cross-referenced them with video recordings of their interactions as needed. Due to the small number of participants, we do not analyze the ratings with statistical techniques (e.g., ANOVA). Instead, we provide summary statistics and user comments for each tool.

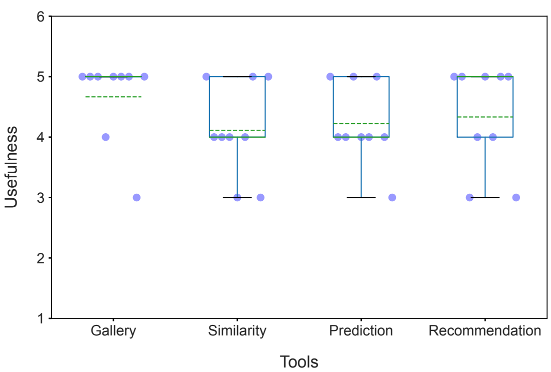


Figure 2: Usefulness ratings for the four tools on a scale of 1 (not useful at all) to 5 (very useful).

Figure 2 shows the user ratings for the four tools. All the tools received ratings between 3 and 5, suggesting that the participants found them useful for various tasks. The *gallery tool* received the highest ratings with 7 out of 9 participants giving the tool a rating of 5 and a mean score of 4.67±0.71. The *recommendation tool* had the next highest ratings with 5 out of 9 participants evaluating the usefulness with a score of 5 and a mean rating of 4.33±0.87. The *prediction* and *similarity* tools had the lowest support with 3 out of 9 participants giving a usefulness rating of 5 for each tool and a mean rating of 4.22±0.67 and 4.11±0.78 for the tools respectively.

The participants found the *gallery tool* easy to navigate (P6, P8) and helpful for getting a sense of various robot hands (P1, P2). They described the pop-up modal with the hand's information as practical and beneficial (P2, P9). Filtering of the hands using the side panel worked well (P4, P5, P9) and facilitated fine exploration and searching of the database (P2, P3). They found the *recommendation tool* easy to use (P8) and effective for narrowing down the features of a robot hand when the user does not have specific design features in mind (P1). The tool helped P3 to understand which design features influence user ratings, and P5 noted that the tool could convey copious amounts of information clearly. For the *prediction tool*, the ability to select features and view the estimated user ratings was interesting (P1, P4). The participants liked the idea of efficiently comparing hands (P1, P4, P6, P9) and exploring the dataset with scatter plots (P1, P8) in the *similarity tool*.

The participants had several suggestions for improving the tools. For the *gallery tool*, the suggestions focused on improving the color scheme and aesthetics and adding tooltips to the sliders to describe the user ratings (P1, P3, P4). For the *recommendation tool*, the participants proposed limiting the range of the user rating sliders to the maximum attainable ratings for the selected hand (P4) or providing an explanation for why a target user rating was unattainable for the hand (P2). For the *prediction tool*, P9 found the estimated user ratings unexpected at times. We conjecture that the tool should be more transparent about the underlying regression models. Other comments focused on changing the placement (P3, P4) or the color scheme (P1) for the predicted user ratings. For the *similarity tool*, the participants wanted to interact with and customize the axes of the similarity space (P4). They asked for further support for comparing the states of the perceptual space before and after applying a filter (P6). Also, the participants wanted to compare all the design features and user ratings side-by-side for two or more hands (P5). Similarly, P2 and P9 wanted to see all the design features of a hand when hovering over its mark on the scatter plots. Searching for a robot hand with its name was mentioned for various tools (P1, P3, P7, P9). Finally, P9 noted that while the *prediction* and *recommendation* tools offer information on design features and user ratings, they do not provide any insights into the functional performance of the robot hands. The participant expressed interest in understanding the trade-offs between attaining the desired user perception of robot hands and their performance (e.g., in picking up objects).

4 DISCUSSION AND CONCLUSION

The study results provided insights into the utility of the tools and future developments to improve their functionality. In particular, the positive reception of the *recommendation tool* compared to the *prediction tool* is interesting. Existing robot collections do not employ an interface similar to the *recommendation tool*, but our preliminary results suggest that the tool can effectively support the user workflows and goals and help them form an understanding of the relationship between the design features and user ratings. Also, we note that existing robot collections either provide technical specifications of robots (e.g., IEEE Robots [2]) or list design features and user ratings (e.g., RobotHands [4], ABOT [1]). No dataset exists that allows robot designers and researchers to explore both technical performance measures and user ratings to understand their trade-offs when selecting or designing a robotic technology.

We are extending this work in several ways. First, we are collecting a larger database and improving the underlying algorithms for the tools. A large set of robot hands can increase the accuracy of the linear regression models for the *prediction* and *recommendation* tools. Relatedly, we aim to use a non-greedy approach for the *recommendation tool* to optimize the feature suggestions further. Moreover, we aim to include technical specifications for the robot hands and study what interactive tools can support users in understanding the trade-offs among the performance measures, design features, and user impression ratings. We also hope to build on research in human-AI interaction to minimize user surprise and improve their understanding of the outputs of the *prediction* and *recommendation* tools [5]. Finally, our goal is to collect longitudinal data on the utility of these tools through anonymous logging in a user study and through an interactive website.

REFERENCES

[1] 2014 — 2023. ABOT Database. http://abotdatabase.info/. [Online; accessed 15-January-2023].

[2] 2014 — 2023. IEEE Robot Database. https://robots.ieee.org/. [Online; accessed 15-January-2023].

[3] 2014 — 2023. Stanford Social Robot Collection. https://osf.io/hz7p3. [Online; accessed 15-January-2023].

[4] 2022 — 2023. RobotHands Database. http://robothands.org/. [Online; accessed 15-January-2023].

[5] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 1–13.

[6] Massimiliano Di Luca, Hasti Seifi, Simon Egan, and Mar Gonzalez-Franco. 2021. Locomotion Vault: The Extra Mile in Analyzing VR Locomotion Techniques. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 1–10.

[7] Morten Hertzum, Kristin D. Hansen, and Hans H. K. Andersen. 2009. Scrutinising Usability Evaluation: Does Thinking Aloud Affect Behaviour and Mental Workload? *Behaviour & Information Technology* 28, 2 (2009), 165–181.

[8] Alisa Kalegina, Grace Schroeder, Aidan Allchin, Keara Berlin, and Maya Cakmak. 2018. Characterizing the Design Space of Rendered Robot Faces. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 96–104.

[9] Johannes Kehrer and Helwig Hauser. 2013. Visualization and Visual Analysis of Multifaceted Scientific Data: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 19, 3 (2013), 495–513.

[10] Elizabeth Phillips, Xuan Zhao, Daniel Ullman, and Bertram F. Malle. 2018. What is Human-like? Decomposing Robots' Human-like Appearance Using the Anthropomorphic RoBOT (ABOT) Database. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 105–113.

[11] Byron Reeves, Jeff Hancock, and Xun "Sunny" Liu. 2020. Social Robots are Like Real People: First Impressions, Attributes, and Stereotyping of Social Robots. *Technology, Mind, and Behavior* 1, 1 (2020).

[12] Harvey M. Salkin and Cornelis A. De Kluyver. 1975. The Knapsack Problem: A Survey. *Naval Research Logistics Quarterly* 22, 1 (1975), 127–144.

[13] Hasti Seifi, Steven A. Vasquez, Hyunyoung Kim, and Pooyan Fazli. 2022. Charting Visual Impression of Robot Hands. *arXiv preprint arXiv:2211.09397* (2022).

[14] Hasti Seifi, Steven A. Vasquez, Hyunyoung Kim, and Pooyan Fazli. 2023. First-Hand Impressions: Charting and Predicting User Impressions of Robot Hands. *ACM Transactions on Human-Robot Interaction* (2023).