Clustering Social Touch Gestures for Human-Robot Interaction

Ramzi Abou Chahine[1(✉)], Steven Vasquez[2], Pooyan Fazli[3], and Hasti Seifi[3]

[1] University of East Anglia, Norwich, Norfolk NR4 7TJ, UK
r.abou-chahine@uea.ac.uk
[2] San Francisco State University, San Francisco, CA 94132, USA
svasquez7@sfsu.edu
[3] Arizona State University, Tempe, AZ 85281, USA
{pooyan,hasti.seifi}@asu.edu

Abstract. Social touch provides a rich non-verbal communication channel between humans and robots. Prior work has identified a set of touch gestures for human-robot interaction and described them with natural language labels (e.g., stroking, patting). Yet, no data exists on the semantic relationships between the touch gestures in users' minds. To endow robots with touch intelligence, we investigated how people perceive the similarities of social touch labels from the literature. In an online study, 45 participants grouped 36 social touch labels based on their perceived similarities and annotated their groupings with descriptive names. We derived quantitative similarities of the gestures from these groupings and analyzed the similarities using hierarchical clustering. The analysis resulted in 9 clusters of touch gestures formed around the social, emotional, and contact characteristics of the gestures. We discuss the implications of our results for designing and evaluating touch sensing and interactions with social robots.

Keywords: Social Touch · Touch Dictionary · Non-Verbal Communication · Crowdsourcing Study

1 Introduction

Social touch has been an active area of research for human-robot interactions (HRI) in the last decade. Social touch gestures refer to different ways that people use touch to communicate information or emotion and bond with other humans or robots [10]. For example, one may tap a robot's arm to get its attention or hug a robotic pet when stressed. A companion robot may stroke a user's hand to convey emotional support or guide the user's action by pushing their hand. Previous work has derived a set of social touch gestures and their definitions based on user interactions with robotic pets [28]. Others designed and evaluated touch interactions with humanoid robots [3,7]. The touch gestures from these studies have guided the development and evaluation of touch sensors for robots, helped examine user experience of robot-initiated touch, and informed the design of robot response to user touch.

A. Al. Ali et al. (Eds.): ICSR 2023, LNAI 14453, pp. 53–67, 2024.
https://doi.org/10.1007/978-981-99-8715-3_6

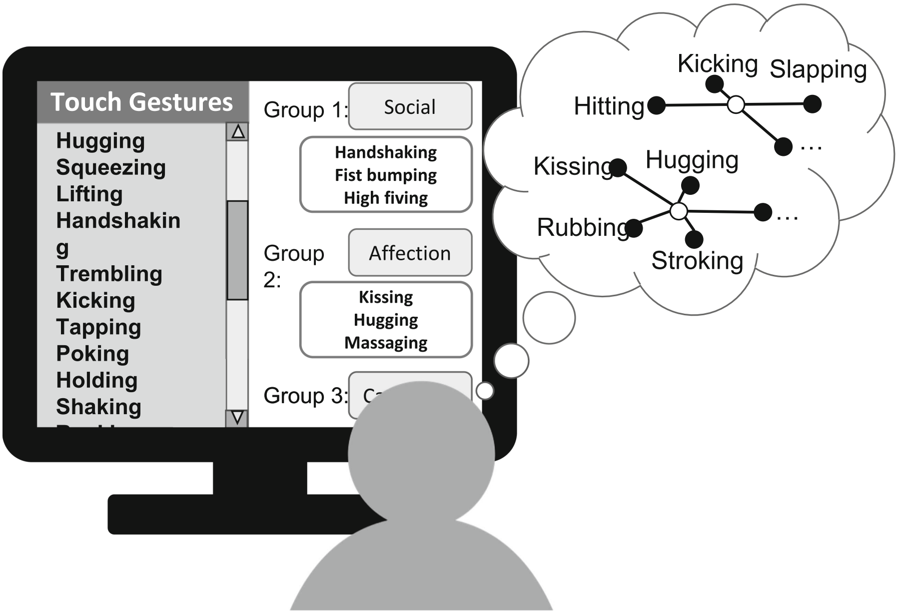

Fig. 1. In our study, users grouped social touch labels based on their perceived similarities (A). The resulting touch clusters can be used by robots to interpret and perform touch interactions with people (B).

Despite the abundance of interest in social touch communication, the semantic relationship(s) between various touch gestures remains unclear. Some gestures may be very similar or even identical in their contact characteristics (e.g., *tapping* vs. *patting*), while others may be similar considering the intended emotion or social context. People develop a mental structure for the semantics of touch gestures and their relationships. This mental structure shapes people's perception, interpretation, and use of touch [8]. Charting the relationship between social touch gestures can help HRI researchers select touch gestures for their studies (e.g., touch sensor evaluation) and develop robots that use touch in a socially intelligent manner. Yet, little data exists in the literature about how people perceive the relationships between social touch gestures.

As a first step toward addressing this gap, we asked how people perceive similarities of social touch labels (e.g., *stroking*, *hugging*). People can have unique styles in applying a touch gesture [12]. On the other hand, people often use natural language labels to refer to archetypal features of a touch gesture. The touch labels are also used in HRI studies to ask users to contact a robot (or a sensor) in a certain way [3,12] or to analyze user interactions with a robot [28]. The study of natural language labels for emotions has helped capture users' cognitive structure, leading to a circumplex model for affect [18]. Thus, as a first step, we investigated the semantic structure of social touch labels in the users' minds in this paper.

To chart the relationship between touch labels, we ran an online card sorting study with 56 users over Amazon Mechanical Turk (Fig. 1-A). The participant received the labels and definitions for 36 touch gestures from the literature, sorted them into 4, 8, and 12 groups successively based on their similarities, and provided descriptive names for each group. From this data, we identified 11 outliers by manually reviewing the data as well as analyzing the responses quantitatively. Then, we created a dissimilarity matrix for the 36 touch gestures with the data of the remaining 45 participants and applied agglomerative hierarchical clustering on the dissimilarity matrix. Furthermore, we analyzed the descriptive names that the participants had for their groupings using open codes (e.g., *social*, *aggressive*) and calculated the frequency of the codes for the gestures.

Based on the above analysis, we contribute 9 clusters for social touch gestures and the distribution of the top codes for each cluster. Using this data, we interpret the 9 clusters to capture the types of touch as follows: (1) social, (2) romantic affection, (3) caregiving affection, (4) hand contact, (5) aggression, (6) forceful press, (7) functional movement, (8) nervous contact, and (9) contact without movement (Fig. 1-B). Our results suggest that people primarily group touch gestures based on their social, emotional, and contact characteristics. These results provide the first data on cognitive structure(s) that people use to interpret and conceptualize social touch. We discuss how the results can help design and evaluate a robot to sense, interpret, and communicate via touch.

2 Related Work

2.1 Social Touch in HRI

The literature on social touch ranges from communication between humans to interactions between humans and robots. Hertenstein et al. studied how dyads use social touch gestures to communicate different emotions and found that people can decode the intended emotions with great accuracy when being touched [8]. Similar studies of human-human touch suggest that touchers can subtly but significantly vary contact attributes of their touch actions to communicate distinct messages [27]. HRI researchers have replicated Hertenstein et al.'s work to investigate how users and robots can use touch to communicate emotions. Some studies examined how humans communicate emotions to robots [11,28], while others examined whether a robot can communicate emotions to humans via touch [21,23].

Social touch gestures have also informed the development and evaluation of tactile skins for robots. Previous work in this area has proposed touch sensors with a novel working principle [5], sensors resembling the feel and structure of human skin [24], and low-cost do-it-yourself sensors for specific applications such as companion robots for children with autism [3]. To evaluate the sensor's efficacy, researchers select a set of social touch gestures and ask users to touch the sensor accordingly. Data from user contact with the sensor is then used to classify the gestures.

A variety of touch gestures are reported in the above studies. Yohanan and MacLean proposed a touch dictionary with labels and definitions for 30 gestures based on videos of user interactions with a furry lap-sized robot [28]. This dictionary has been widely used in social touch studies [3,11]. Others mentioned additional gestures for interactions with humanoid robots such as *fist bumping* [15–17], *handshaking* [2,17,26], or *kicking* [8,13,20]. To inform future work in this area, we collected common touch gestures from prior studies and examined how people conceptualize the relationship between these gestures.

2.2 Identifying Perceptual and Semantic Clusters

The psychophysics and interaction design literature has developed methods for estimating perceptual and semantic similarities of items through user studies. The pairwise rating method asks participants to rate the similarity of pairs of items in the set [1]. This method is effective for a small set of items (e.g., < 15) but it is prone to noise from local judgments and does not scale to large item sets [25]. The sorting methods, known as card sorting or cluster sorting, ask participants to group items into clusters based on their similarities. This process can be repeated with an increasing number of groups to obtain a fine-grained similarity matrix [22]. This method allows for collecting cognitive similarities of large item sets [18]. The similarity matrix is further analyzed using dimensionality reduction or clustering techniques [1,18]. Following this methodology, we used iterative cluster sorting and asked users to name their groups to obtain semantic clusters for social touch labels.

Natural language labels have been used to capture lay users' cognitive structure for sensory and emotional items. The circumplex model of affect by Russell [18] is based on a series of studies that use natural language labels for emotions. Also, studies of social touch often rely on user understanding of natural language labels for touch. In these studies, users receive labels for a set of social touch gestures (e.g., *tapping*, *stroking*) and are asked to touch the robot accordingly [3,11]. Similarly, studies on human-human and human-robot emotional communication sometimes provide a list of touch gesture labels for users to choose from, before applying the gestures [8,27]. The studies may also provide short definitions for each touch gesture e.g., from the touch dictionary by Yohanan and MacLean [28]. These studies rely on the users' knowledge of natural language labels for touch gestures. We follow a similar approach in our work to capture users' cognitive structure and similarities of social touch gestures.

3 Methods

To study how people perceive similarities of social touch gestures, we compiled a list of touch gestures from the literature, designed an online questionnaire for grouping the touch gestures, and ran a data collection study on MTurk.

Table 1. The 7 touch gestures that we added to the 29 gestures in Yohanan and MacLean's touch dictionary [28], resulting in 36 touch gestures for our online study.

Gesture Label	Gesture Definition
Finger Interlocking	Interlace fingers of one hand
Fist Bumping	Lightly tap clenched fists together
Handshaking	Shake clasped hands
High Fiving	Slap upraised hands against each other
Kicking	Strike forcibly with a foot
Picking Up	Take hold of and lift or move something
Squish	Press or beat into a pulp or a flat mass

3.1 Touch Gestures

We compiled 36 social touch gestures that are used for interacting with humans or robots. We focused our scope on gestures that are used in at least two publications in the social touch and HRI literature. Specifically, we included 29 touch gestures from the touch dictionary by Yohanan and MacLean [28]. Different subsets of these gestures are used in several other studies [11,12]. We removed *finger idly* from the touch dictionary as this gesture is not used in any other publication. We added seven other touch gestures that appeared in at least two publications including *finger interlocking* [8,9], *fist bumping* [15–17] *handshaking* [2,17,26], *high fiving* [6–8], *squishing* [4], *kicking* [8,13,20], and *picking up* [4,19,20].

We adapted the definitions provided in Yohanan and MacLean's touch dictionary by replacing the phrases related to their robotic pet (i.e., the Haptic Creature, or fur of Haptic Creature) with "something" in the definition. For example, we defined *lifting* as "raise something to a higher position or level." For the 7 actions that were not in the original touch dictionary, we created a definition with inspiration from sources such as the Britannica Encyclopedia. Table 1 shows the 7 newly added gestures and their definitions.

3.2 Questionnaire

We designed a Qualtrics survey to collect user demographics and data on the similarity of touch actions (Fig. 2). The first page of the survey asked users to enter their demographic information including their age, gender, and country where they grew up. The next three pages asked the users to divide the touch gestures into 4, 8, and 12 groups respectively. We call these 4 groupings, 8 groupings, and 12 groupings in this paper. Each page showed the list of touch gesture labels in a random order. The users could hover over a gesture's label to see its definition. The users were asked to group the touch gestures based on their likeness or similarity and provide a descriptive name for each group. Reasons for likeness were up to user interpretation. Having the users describe their groupings served multiple purposes. First, they helped us identify users' reasoning for the

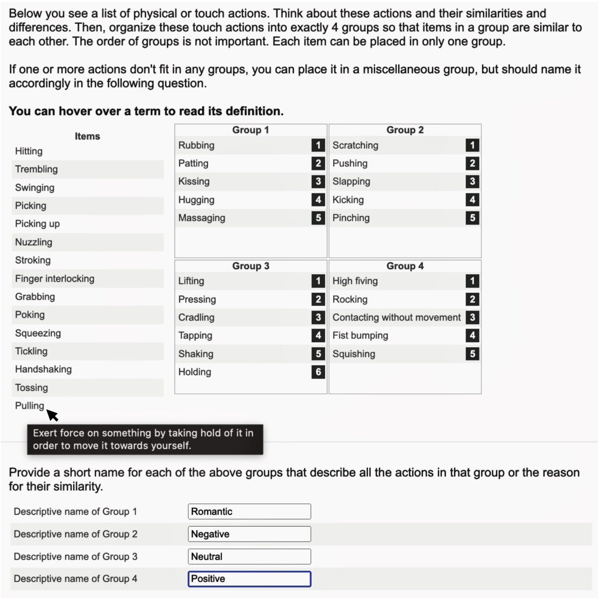

Fig. 2. A screenshot of the questionnaire for grouping the touch gestures in our study. The image shows touch gestures that are divided into four groups, the remaining list of gestures for grouping, and example descriptive names from one of the participants.

similarity of touch gestures. Second, the descriptive names served as an attention test and allowed us to detect those who did not do the task properly, e.g., if they organized the gestures into random groups.

We devised the above procedure based on common practices in studies of similarity perception and social touch gestures in the literature. First, the iterative cluster sorting method allowed us to collect users' holistic comparisons of the similarities of all 36 gestures. Second, following prior work on touch sensing and communication, the touch labels helped us abstract from a variety of styles that people use to apply the touch gestures (e.g., tapping one time or multiple times) to capture users' cognitive structure of the gestures.

4 Analysis and Results

We collected participant responses through MTurk. Eligible turkers were required to have at least 5,000 completed tasks with a minimum success rate of 97% and to speak English at the B2 level or higher. We analyzed their data in the following steps:

– **Identifying outliers.** We identified participants who did not follow the study instructions or appeared to group the touch gestures randomly (Sect. 4.1) and removed their data from the subsequent analysis. We also examined the effect of where participants grew up on their groupings.

- **Coding descriptive names for the groups.** To identify the themes behind the user groupings, we coded the descriptive group names from the participants. This step resulted in 25 codes (e.g., social, aggressive) to capture user logic for their groupings (Sect. 4.2).
- **Clustering touch gestures.** We calculated a dissimilarity matrix for the touch gestures based on the participants' groupings and identified semantic groups by applying hierarchical clustering on the dissimilarity matrix.
- **Interpreting the clusters.** Finally, we counted how many times a code from Step II was applied to the touch gestures in each cluster. The results helped us interpret and label each of the 9 social touch clusters (Sect. 4.3).

Below we detail these steps and their results.

4.1 Identifying Outliers

We marked and removed outliers who did not follow the study instructions or their groupings and descriptive names appeared random. One of the authors carefully examined all the responses from the 56 participants and marked potential outliers for further analysis. The author marked cases where no label was provided, the label was gibberish, or the description of group labels did not match with its gesture items. For example, if a participant grouped *kissing*, *nuzzling*, and *stroking* with *hitting* and labeled them as "fighting", we marked this as an unusual group. By the end of this step. 16 participants with several unusual groupings were marked as potential outliers.

Next, we calculated a similarity matrix where each cell showed similarity of the groupings provided by two participants (56×56 matrix). To obtain the best matching between groups from two different participants, we calculated the Jaccard Index values for all pairs of groups provided by them (e.g., 8 pairs for the 4 groupings) and averaged the highest Jaccard values as a measure of the similarity of the two participants.

We projected the participant similarities into two dimensions using a common dimensionality reduction technique known as non-metric Multidimensional Scaling (nMDS) and used clustering to assess outliers (Fig. 3). In addition, we conducted k-means clustering with a range of 2 to 10 clusters on the dissimilarity matrix. The value of the Gap Statistic suggested 3 as the optimal number of clusters (Fig. 3). Our analysis revealed that cluster 3 contained 11 out of the 16 participants that we had manually identified as potential outliers. Cluster 2 contained the remaining 5 potential outliers, as well as participants not considered to be outliers in our manual analysis. Thus, the two methods of manual and quantitative analysis of outliers largely overlapped and provided support that the cluster 3 participants either provided noisy data or judged similarities differently from the majority. Thus, we included the participants from clusters 1 and 2 ($n = 45$ participants) in further analysis.

The remaining 45 participants were from the United States (32), followed by India (7), Brazil (5), and Japan (1). They self-identified as man ($n = 29$), woman ($n = 16$), or nonbinary ($n = 0$). The mean age of the participants was

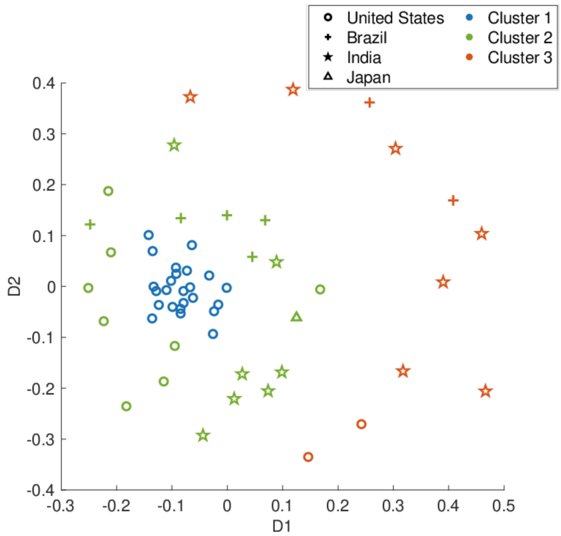

Fig. 3. MDS plot visualizing similarity of the 56 participants in grouping the gestures. Each mark represents one participant. The color and shape of the marks denote the clustering results and participant backgrounds, respectively. Participants in cluster 3 (red) were identified as potential outliers and were removed from further analysis. (Color figure online)

36.4 (±10.73) years and their ages ranged between 21–63 years. The participant background is denoted with the shape of the marks in Fig. 3. Participants who were not from the US are either in clusters 2 or 3. We analyzed this aspect further in our clustering results (Sect. 4.3).

4.2 Coding Descriptive Names for the Groups

To understand the reasoning behind group choices, we coded the descriptive names provided by the participants for each group. From 4 to 8 to 12 groupings, the codes became more complex as subgroups began to form. The process of identifying these codes was iterative. For example, when coding the descriptive names for 12 groupings, we used the codes identified from 8 groupings in the first iteration. If we found any new or more specific patterns, we added new codes and recorded the previous data accordingly. Upon completing the coding of all the groupings, we had a total of 25 codes. We found some descriptive names to be ambiguous and coded them as ‘vague’. We also found that some names did not match the social touches they were assigned to, we coded these descriptive names as ‘random’. In some cases, participants labeled a group as ‘other’ or ‘miscellaneous’. Thus, we also coded these groupings as ‘miscellaneous’. If a grouping contained only a single social touch, we coded it as ‘single action’. The remaining 21 codes included: aggressive, annoying, caregiving, direction, fingers, force, friendly, full-body, functional, grief, hands, holding onto, massage, nervous, playful, rapid, repetitive, romance, slow, social, and squeezing.

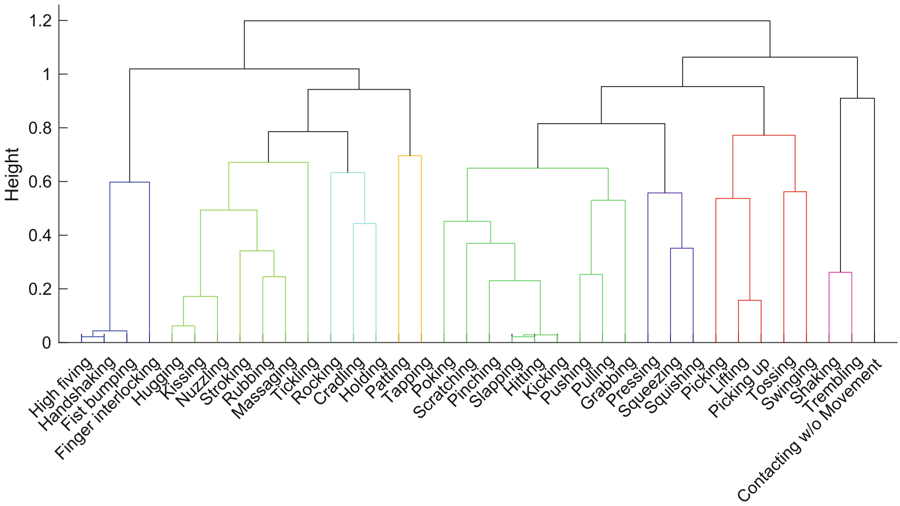

Fig. 4. Results demonstrating hierarchical clustering results for the social touch gestures. The Gap Statistic criterion suggested an optimal number of 9 clusters. The Cophenetic correlation coefficient is 0.85 suggesting strong correspondence with the dissimilarity matrix. Each color represents one cluster.

4.3 Clustering Touch Gestures

Using the grouping data of each participant, we created a similarity matrix of touch gestures following the same procedure described by Russell [18]. First, each pair of words was given a minimum similarity score of 1. If pairs of words were included in the same user-defined group, then their similarity score was increased by the number of groups being organized. For example, we increased the similarity score by 4 if a pair of words were in the same cluster for the 4 grouping mode for a participant. If a pair of words were included in the same group for 4, 8, and 12 groupings modes, then the words would have the maximum possible similarity of $1+4+8+12 = 25$. A single similarity matrix was calculated from the three grouping modes, and the matrix was subsequently normalized by dividing its entries by the maximum possible similarity value (i.e., 45 participants $\times$ 25 = 1125). We subtracted the normalized matrix from a matrix of ones to generate a dissimilarity matrix for all the gestures.

We applied clustering to the dissimilarity matrix and identified 9 clusters for the touch gestures. Specifically, we employed agglomerative hierarchical clustering using the unweighted pair group method with arithmetic mean (UPGMA) [14]. To determine the optimal number of clusters for hierarchical clustering, we utilized the Gap Statistic evaluation criterion with a range of 2 to 10 clusters.

Table 2. Our derived names and the top 5 codes with their percentages for the 9 clusters. All the single-item groups are coded as 'single action'.

C1	Social	social 50%	hands 14%	romance 14%	caregiving 4%	random 3%
C2	**Romantic Affection**	romance 33%	caregiving 11%	massage 6%	random 5%	vague 4%
C3	**Caregiving Affection**	caregiving 15%	romance 12%	vague 10%	functional 8%	random 6%
C4	**Hand Contact**	force 12%	hands 10%	social 9%	vague 9%	random 8%
C5	**Aggression**	aggressive 52%	functional 9%	random 5%	hands 5%	vague 5%
C6	**Forceful Press**	aggressive 18%	functional 12%	squeezing 12%	vague 9%	force 8%
C7	**Functional Movement**	functional 29%	vague 12%	aggressive 8%	random 8%	hands 5%
C8	**Nervous Contact**	nervous 30%	aggressive 14%	vague 11%	random 6%	force 5%
C9	**Contact w/o Movement**	single action 24%	miscellaneous 10%	social 10%	vague 7%	functional 6%

This analysis suggested 9 clusters (Fig. 4). The Cophenetic correlation coefficient was 0.85 for the 9 clusters, indicating a strong positive correspondence between the clusters and the original dissimilarity matrix. These clusters include:

- **Cluster 1:** *high-fiving*, *handshaking*, *fist bumping*, and *finger interlocking*
- **Cluster 2:** *hugging*, *kissing*, *nuzzling*, *stroking*, *rubbing*, *massaging*, and *tickling*
- **Cluster 3:** *rocking*, *cradling*, and *holding*
- **Cluster 4:** *patting* and *tapping*
- **Cluster 5:** *poking*, *scratching*, *pinching*, *slapping*, *hitting*, *kicking*, *pushing*, *pulling*, and *grabbing*
- **Cluster 6:** *pressing*, *squeezing*, and *squishing*
- **Cluster 7:** *picking*, *lifting up*, *picking up*, *tossing*, and *swinging*
- **Cluster 8:** *shaking* and *trembling*
- **Cluster 9:** *contacting without movement*

To test the effect of cultural background and English proficiency in our results, we repeated the above clustering analysis on data from 32 participants from the US. The analysis led to similar clusters with the exception that clusters 2 and 3 were merged into one cluster. Thus, we decided to continue with the above 9 clusters in our further analysis.

4.4 Interpreting the Clusters

We calculated the distribution of our codes for the descriptive names across these clusters to interpret the reason behind the groups. Table 2 presents the five frequent codes for the gestures in each cluster.

We named the clusters based on the distribution of their five top codes. For clusters 1 and 5, the majority of the codes ($\geq 50\%$) are 'social' and 'aggressive'. Thus, we call these clusters *Social* and *Aggression* respectively. Clusters 2, 7, 8, and 9 have one frequent code ($\geq 24\%$), followed by one or two codes with $\geq 10\%$ frequency. For cluster 2, the top code is 'romance' followed by 'caregiving', both of which reflect the affective nature of touch. Thus, we name this group as *Romantic Affection.* For cluster 7, the top code is 'functional', followed by 'vague'. This cluster includes a set of gestures that involve lifting and moving an object or person. Thus, we name it *Functional Movement.* Cluster 8 has a top code of 'nervous', followed by 'aggressive'. Thus, we call it *Nervous Contact.* Cluster 9 includes the single gesture of *contacting without movement.* This gesture was often put in a separate group by the participants and we coded it as 'single action'. Thus, we name this cluster as *Contact w/o Movement* to reflect its distinct nature in the participants' minds. Finally, clusters 3, 4, and 6 have a relatively flat code distribution. Cluster 3 has the same two top codes as cluster 2, representing affect, but in the reverse order. Thus, we name it *Caregiving Affection.* Cluster 4 has two codes of 'force' and 'hands' with more than 10% frequency. With two gestures of *patting* and *tapping*, we name this cluster as *Hand Contact.* The top codes ($\geq 10\%$) for cluster 6 are 'aggressive', 'functional', and 'squeezing'. Since the top labels indicate both the 'aggressive' and 'functional' aspects of the gestures in this cluster, we use a neutral label and call this cluster *Forceful Press.* Next, we discuss these clusters and their implications for HRI research.

5 Discussion

In this study, we present data on the user perception and description of touch gestures. Our findings indicate that users tend to assess the similarity of touch gestures based on their emotional and social connotations, in addition to the functional and contact characteristics. Specifically, cluster 1 includes touch gestures that are frequently annotated with 'social' names. Clusters 2 and 3 include gestures that are mainly coded with positive associations of 'romance' and 'caregiving'. Similarly, clusters 5 and 8 are coded with negative descriptors of 'aggressive' and 'nervous'. Finally, four clusters (i.e., 4, 6, 7, 9) seem to be mainly described based on the characteristics of the contact such as the body part (cluster 4), force (cluster 6), and whether the touch involved movement (cluster 7) or not (cluster 9). These clusters emerged without providing information on the context of interaction, suggesting that users have strong social, positive, negative, and functional associations with touch gestures even without context. Some clusters have a flat distribution of codes and show a notable mix of affective and functional interpretations (e.g., cluster 6 with *pressing*, *squeezing*, and *squishing*) suggesting that an individual's background or interaction context may notably shift their meaning. Interestingly *contacting without movement* was often regarded as different from the other gestures, which could be due to its neutral emotional content as well as the static nature of the touch.

These user-generated clusters are a step toward a framework for the analysis and understanding of social touch and can inform research on sensing, designing, and analyzing human-robot touch interactions. We anticipate the following use cases of the touch clusters for HRI:
(1) *Sensing touch from humans.* A desirable factor for robotic touch sensors is their ability to recognize a variety of gestures [11]. These clusters can aid researchers in selecting gestures that are different in their semantic and contact characteristics. For instance, the co-location of *stroking* and *rubbing* gestures in cluster 2 suggests that it might be appropriate to choose one of the two gestures. Relatedly, when evaluating the efficacy of a touch-sensing algorithm [3,5,11], HRI researchers can weigh misclassifications according to these semantic clusters. For example, misclassifying *stroking* with *slapping* should be penalized more than mistaking *stroking* with *rubbing* or *nuzzling.*
(2) *Interpreting and responding to touch from humans.* The proposed touch gesture clusters can aid robots in responding intelligently to human touch. These clusters can help robots identify the intention behind touch gestures. While the significance and purpose of social touch gestures may depend on the context, these clusters and their labels can help develop a probabilistic mental model for robots about a user's intent of a touch gesture. During an interaction episode, the robot can update these probabilities based on other contextual parameters and modes of communication such as the user's verbal utterances and body pose.
(3) *Touching people to communicate.* The semantic clusters can help design and evaluate robots that touch humans to communicate information or emotion [23]. Specifically, to evaluate the efficacy of a robot in using touch gestures, HRI researchers can determine the degree of dissimilarity between the intended touch gesture and the one identified by the human. Also, depending on the purpose of the interaction (e.g., social, emotional, or functional), the robot may use the clusters to select and use alternative gestures with similar connotations.
(4) *Analyzing human-robot touch interactions.* HRI researchers can use these clusters to code video recordings of touch interactions with a robot and aggregate touch interaction into higher-level themes. To support this, our work builds on the touch dictionary [28] by providing data on the relationship between touch gestures. Thus, these clusters provide an initial framework for the analysis of social touch interactions with robots.

6 Conclusion and Future Work

Our work is a first step toward charting the relationship of touch gestures for HRI. We anticipate that our results can pave the way for future work on designing and evaluating robots that use touch as a non-verbal communication channel.

We see several avenues for extending this work. First, the relationship between the user-generated clusters for touch gestures and signals produced by the gestures on different touch sensors is an open question. A good touch sensor should be able to create distinct signals for gestures that are in different clusters according to user perception. Also, robots should be able to create distinct sensations when touching users with gestures in different clusters.

Second, future work can examine the impact of presentation modality on the semantic relationship of touch gestures. In this paper, we presented text labels for social touch gestures, following the common procedure in user studies of touch sensing for social robots. This approach helped abstract different styles of applying the gestures and study the user's mental representations of archetypal touch gestures. Future studies can examine how people group the touch gestures using videos or by applying robot touch on the user's body and compare the results to the clusters we found in this work. These studies should capture a wide range of touch styles (e.g., contact, force) for each gesture to avoid biasing the results to a small sample.

Finally, the meaning of touch can vary based on contexts, cultures, and individuals. As a first step, we examined if any generalizable patterns could be found about the relationships between various touch gestures. Our study population primarily consisted of individuals that grew up in the United States. Participants from other cultures often fell into cluster 3 and around the borders of cluster 2. It is unclear whether this result is due to their lack of familiarity with touch labels or the difference in their cultural background. Future studies can examine how the clusters of social touch gestures differ across cultures by translating the text labels into different languages. A larger dataset can also allow future work to look into individual differences in perception of social touch.

References

1. Abou Chahine, R., Kwon, D., Lim, C., Park, G., Seifi, H.: Vibrotactile similarity perception in crowdsourced and lab studies. In: Seifi, H., et al. (eds.) Haptics: Science, Technology, Applications, EuroHaptics 2022. LNCS, vol. 13235, pp. 255–263. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-06249-0_29
2. Ammi, M., et al.: Haptic human-robot affective interaction in a handshaking social protocol. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 263–270 (2015)
3. Burns, R.B., Lee, H., Seifi, H., Faulkner, R., Kuchenbecker, K.J.: Endowing a NAO robot with practical social-touch perception. Front. Roboti. AI 86 (2022)
4. Burns, R.B., Seifi, H., Lee, H., Kuchenbecker, K.J.: Getting in touch with children with autism: specialist guidelines for a touch-perceiving robot. Paladyn J. Behav. Robot. **12**(1), 115–135 (2021)
5. Choi, H., et al.: Deep learning classification of touch gestures using distributed normal and shear force. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3659–3665 (2022)
6. Cramer, H.S., Kemper, N.A., Amin, A., Evers, V.: The effects of robot touch and proactive behaviour on perceptions of human-robot interactions. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 275–276 (2009)

7. Fitter, N.T., Kuchenbecker, K.J.: Analyzing human high-fives to create an effective high-fiving robot. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 156–157 (2014)
8. Hertenstein, M.J., Holmes, R., McCullough, M., Keltner, D.: The communication of emotion via touch. Emotion **9**(4), 566 (2009)
9. Hertenstein, M.J., Keltner, D.: Gender and the communication of emotion via touch. Sex Roles **64**, 70–80 (2011)
10. Huisman, G.: Social touch technology: a survey of haptic technology for social touch. IEEE Trans. Haptics **10**(3), 391–408 (2017)
11. Jung, M.M., Poel, M., Poppe, R., Heylen, D.K.: Automatic recognition of touch gestures in the corpus of social touch. J. Multim. User Interfaces **11**(1), 81–96 (2017)
12. Jung, M.M., Poppe, R., Poel, M., Heylen, D.K.: Touching the void-introducing cost: corpus of social touch. In: Proceedings of the International Conference on Multimodal Interaction (ICMI), pp. 120–127 (2014)
13. Li, B., et al.: Human robot activity classification based on accelerometer and gyroscope. In: Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 423–424 (2016)
14. Murtagh, F., Contreras, P.: Algorithms for hierarchical clustering: an overview, II. Wiley Interdiscip. Rev. Data Mining Knowl. Discov. **7**(6), e1219 (2017)
15. Pelikan, H.R., Broth, M., Keevallik, L.: Are you sad, cozmo? How humans make sense of a home robot's emotion displays. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 461–470 (2020)
16. Prasad, V., Koert, D., Stock-Homburg, R., Peters, J., Chalvatzaki, G.: Mild: multimodal interactive latent dynamics for learning human-robot interaction. In: Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 472–479 (2022)
17. Rognon, C., et al.: An online survey on the perception of mediated social touch interaction and device design. IEEE Trans. Haptics **15**(2), 372–381 (2022)
18. Russell, J.: A circumplex model of affect. J. Personal. Soc. Psychol. **39**, 1161–1178 (1980). https://doi.org/10.1037/h0077714
19. Salter, T., Dautenhahn, K., te Boekhorst, R.: Learning about natural human-robot interaction styles. Robot. Auton. Syst. **54**(2), 127–134 (2006)
20. Salter, T., Michaud, F., Létourneau, D., Lee, D., Werry, I.P.: Using proprioceptive sensors for categorizing human-robot interactions. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 105–112 (2007)
21. Seifi, H., Vasquez, S.A., Kim, H., Fazli, P.: First-hand impressions: charting and predicting user impressions of robot hands. ACM Trans. Hum. Robot Interact. (2023)
22. Ternes, D., MacLean, K.E.: Designing large sets of haptic icons with rhythm. In: Ferre, M. (ed.) EuroHaptics 2008. LNCS, vol. 5024, pp. 199–208. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-69057-3_24
23. Teyssier, M., Bailly, G., Pelachaud, C., Lecolinet, E.: Conveying emotions through device-initiated touch. IEEE Trans. Affect. Comput. (2020)
24. Teyssier, M., Parilusyan, B., Roudaut, A., Steimle, J.: Human-like artificial skin sensor for physical human-robot interaction. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 3626–3633. IEEE (2021)
25. Tsogo, L., Masson, M., Bardot, A.: Multidimensional scaling methods for many-object sets: a review. Multivar. Behav. Res. **35**(3), 307–319 (2000)

26. Wang, Z., Giannopoulos, E., Slater, M., Peer, A.: Handshake: realistic human-robot interaction in haptic enhanced virtual reality. Presence **20**(4), 371–392 (2011)
27. Xu, S., Xu, C., McIntyre, S., Olausson, H., Gerling, G.J.: Subtle contact nuances in the delivery of human-to-human touch distinguish emotional sentiment. IEEE Trans. Haptics **15**(1), 97–102 (2021)
28. Yohanan, S., MacLean, K.E.: The role of affective touch in human-robot interaction: human intent and expectations in touching the haptic creature. Int. J. Soc. Robot. **4**(2), 163–180 (2012)