

# A Multimodal Data Set of Human Handovers with Design Implications for Human-Robot Handovers

Parag Khanna<sup>1</sup>, Mårten Björkman<sup>1</sup> and Christian Smith<sup>1</sup>

**Abstract**—Handovers are basic yet sophisticated motor tasks performed seamlessly by humans. They are among the most common activities in our daily lives and social environments. This makes mastering the art of handovers critical for a social and collaborative robot. In this work, we present an experimental study that involved human-human handovers by 13 pairs, i.e., 26 participants. We record and explore multiple features of handovers amongst humans aimed at inspiring handovers amongst humans and robots. With this work, we further create and publish a novel data set of 8672 handovers, bringing together human motion and the forces involved. We further analyze the effect of object weight and the role of visual sensory input in human-human handovers, as well as possible design implications for robots. As a proof of concept, the data set was used for creating a human-inspired data-driven strategy for robotic grip release in handovers, which was demonstrated to result in better robot to human handovers.

## I. INTRODUCTION

Robots have been used to increase productivity and carry out risky jobs in industrial settings with success. Such industrial robots are technically advanced and suited for work in isolation on particular tasks. However, significant recent advancements in robotic intelligence and technology have made it possible to foresee robots cohabiting with humans to help and collaborate. We envisage social robots having a shared environment with humans performing a variety of activities in the future. These activities include, but are not limited to, assisting at home [1], collaborating in factories [2], helping humans with restricted mobility and in old age homes [3]. Indeed, effective human robot collaboration (HRC) will also allow humans and robots to complement each other’s abilities. In a real world environment filled with uncertainties, the decision making skills of a human counterpart can largely benefit a collaborative robot. This has motivated many robotic platforms aimed at physical HRC [4]. While doing collaborative work, robots must be able to engage with people in a safe, natural, and convenient manner. Thus, it is crucial that we create robots that can interact with people physically. Handover is one such physically interactive task which frequents our daily social lives. To achieve effective human robot collaboration, fluent and natural robot human handovers are essential. As a result, a social robot needs to master the skill of handover.

According to many studies [5], [6], [7], humans prefer robot behaviour that mimics humans. Humans also make the ideal subjects for studying handovers due to their proficiency and adaptation skills. In this paper, we present an

<sup>1</sup> Division of Robotics, Perception and Learning (RPL), EECS, KTH Royal Institute of Technology, Sweden [paragk@kth.se](mailto:paragk@kth.se), [celle@kth.se](mailto:celle@kth.se), [ccs@kth.se](mailto:ccs@kth.se)

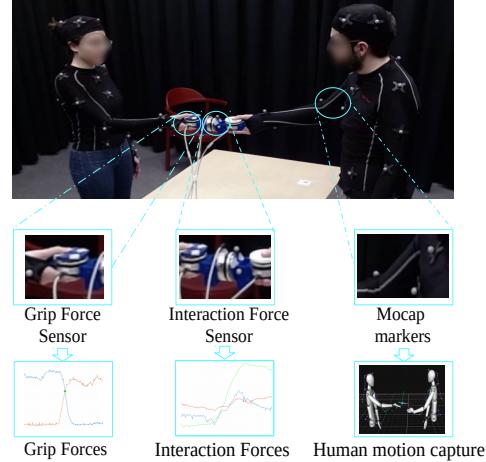


Fig. 1. A snippet of the human study exhibiting various aspects of a handover and the recorded signals

experimental study of human handovers, which is used to create an open source data set of handovers and to further analyze various aspects of human handovers as shown in Fig. 1. We also focused on studying the effect of increased weight of transfer in handovers. We analyze handovers and what happens when humans rely on just haptic feedback, comparing the cases with no visual information to those with visual and haptic feedback. The aim of this work is to learn from human handovers to inspire natural, safe, and efficient handovers between humans and robots. The main contribution of this work is a novel multimodal data set of 8672 human-human handovers with information from multiple sensors including grip forces, interaction force and human motion tracking. We also provide some qualitative observations including the effects of increased weight of transfer and visual sensory impairment on various aspects of handovers aimed to enhance handover capabilities for robots.

## II. BACKGROUND AND RELATED WORK

A task of handover is comprised of multiple phases that humans collaborate to perform seamlessly. Verbal or nonverbal communication is used to initiate a coordinated spatio-temporal movement of a giver and a taker. In case of a general handover, the giver determines a suitable location to transfer the object in the inter-personal space shared with the taker and starts moving the object towards that location [8] [9]. The taker, too, prejudices the handover location based on the giver’s motion with the object and starts approaching the location with the preferred hand. As handover progresses, both giver and taker adjust their arm speed based on visual

sensory information. The coordinated movement finally ends with a light impact on the object as the taker starts interacting with the object, trying to form a grip on it. During this interaction, the giver decreases the grip force on the object while the taker's grip force increases. Both giver and taker also share the responsibility of sharing the object's weight and maintaining the object's pose so that it does not fall in the process. The handover ends when the giver releases the object completely after determining a stable grasp on the object by the taker. Thus, a robot should be competent across multiple phases of a handover for fluent handovers.

Furthermore, there are more complex handover scenarios to which humans swiftly adapt. Humans are not bothered by different-weighted objects from their daily lives, for example, handing over a knife and a hammer with fluency [9], [7], [10]. There are also common scenarios when the weight of the object to be transferred turns out to be different than expected. One such scenario is the handover of a supposedly empty box, which turned out to be full. Both human giver and taker are capable of quickly adapting to changes caused due to these uncertainties regarding the weight of the transferred object. In some scenarios, a human giver does not look actively at the object, or does not look at all, while giving the object. In such a case, the giver relies on only the haptic feedback to perform the handover, i.e., sensing the forces in interaction. In these scenarios too, both giver and taker are able to suitably adapt and perform a safe and efficient handover. A social robot is likely to face these scenarios, owing to the dynamic social environment. It is expected to adapt and perform according to the given situation in order to accommodate safe handovers, while avoiding failures.

The interest in human inspired handovers for robots has led to multiple studies on handovers amongst humans [9]. Several human studies have been done to study human motion to model the movement of a robotic arm close to a human in handovers. Human hand motion was captured in human handovers by an electromagnetic tracker and markers placed on the hand in [11], and was further used to evaluate robot reaching profiles for handovers. Human arm motions were recorded via a camera in [12] to analyse and propose a joint motion model for a robotic giver for human like handovers. To understand verbal and non-verbal human cues, and joint coordination involved in human handovers, a human study was done in [7] in a kitchen environment with multiple color and depth cameras for analysis.

To study forces in human handovers, [5] used a baton with grip force sensors and a load force sensor for handovers in a human study. The baton was vertically transferred in handovers and a relationship between grip forces and load forces was measured, which was used to develop a controller for robotic grip forces given load forces. In [13], the effect of reaching velocities and sensory manipulations was seen on grip forces in human human handovers. In one of the studied cases, the giver's vision was manipulated by closing the eyes during the handover. It was observed that sensory manipulation had a strong effect on temporal aspects of the handover. The effect of object weight on handover location

and duration was investigated by a human study in [14]. It was shown that only the handover duration was impacted by object weight. In [8], a relation between the taker arm reaching speed and the giver's grip forces was found in a human study. The effect of restricting visual information on giver was also studied to conclude that coordinated arm-movement is impacted by lack of visual feedback.

Some publicly available data sets do exist for human handover tasks. The [15] data set includes human motion tracking via different sensors for participants in human handovers passing multiple objects. They provide 3D upper skeleton tracking via a Motion-Capture (MoCap) system and a depth camera, as well as inertial data of the hands from smartwatches. In [10], a human study was done to inspire proper object orientations in robot to human handovers. The corresponding data set includes MoCap tracking data of the upper skeleton and object pose in handovers. Human grasps were studied and labelled in [16], collecting over 5202 grasps. Their study showed that precision grasp is the most preferred one. Also most givers in human handovers preferred a precision grasp irrespective of object. In [17], an image based data set focused on hand poses and grasps via visual analysis of handovers is described, where the hand poses are tracked by markers placed on each finger.

We observe that the task of handover has multiple aspects and modalities spanning across spatio-temporal and force domains. Many of these aspects have been researched, but independently from each other. Publicly available data sets also lack information about grip and interaction forces. In this work, we present a comprehensive study that integrates multiple aspects of handovers and a multimodal data set containing handover forces with human motion tracking. In a multimodal interaction like handover, it is important to understand not only the effect of each modality, but also how different modalities are interrelated. Moreover, some of these modalities, like human-grip force, might not be available to a robot in handovers. However, a multimodal dataset makes it possible to train the robot on multiple modalities and to learn mapping from one to another. The robot can then use only those modalities that are accessible during online execution.

### III. HUMAN HANDOVERS STUDY

Humans practice handovers countless times over several years to perfect them. This motivated us to design an experimental study focused at human handovers. We aimed to get a unified perspective of the giver's and taker's motion, grip forces applied by them, and the interaction forces involved in human handovers. Our goal in this study is two-fold:

- 1) Recording and analysing *natural* human handovers.
- 2) Create a multi-sensor data set of human handovers.

#### A. Participants

For the study we recruited a total of 26 participants which included 22 male and 4 female participants between the ages of 14 and 28 (age  $23.23 \pm 2.55$  years). Only one male participant was left handed while others were right handed. This led to 13 participant pairs which comprised

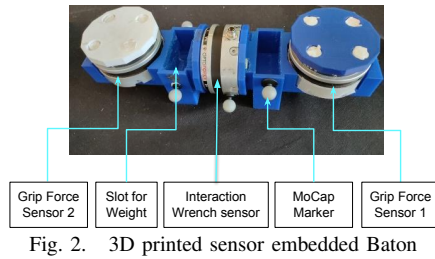


Fig. 2. 3D printed sensor embedded Baton

of one female-female pair, two female-male pairs and ten male-male pairs. One of these male-male pairs was a right handed-left handed pair. The participants had little to no prior acquaintance with each other. They were recruited by advertisement amongst the university students and were rewarded with a 100 SEK gift-voucher.

### B. Experimental Design

1) *Baton*: The experiment involved passing back and forth a sensor embedded baton which was made by 3D printing. The baton houses three 6D Force/Torque (F/T) sensors from Onrobot [18] as shown in Fig. 2, where each F/T sensor measures a 6D wrench, comprising 3 forces and 3 torques. The F/T sensor in the center of baton is used to measure the interaction wrench in the handover. The two sensors on the sides are used to measure the grip forces by the giver and the receiver. Additionally, 5 Motion Capture (MoCap) markers were placed asymmetrically on the baton to track its movement. The baton has a total weight of 0.8 kg and measures 0.28x0.085x0.075 m. It also has two slots to add external lead-weights, increasing the baton weight when needed. The top of the two sensors for grip force were color coded with blue and white parts. For simplicity, participant 1 was told to always pick up and receive the baton from the blue side while participant 2 would do so from white side.

2) *Experimental Setup*: To record the movements of participants, the experiment was conducted in a MoCap room by Optitrack [19] and participants were required to wear upper body MoCap suits. The participants performed handovers standing across a table as shown in Fig. 3(a). The experimental session started with participant 1, acting as giver, picking up the baton from position A and handing over the baton mid-air and horizontally to participant 2,

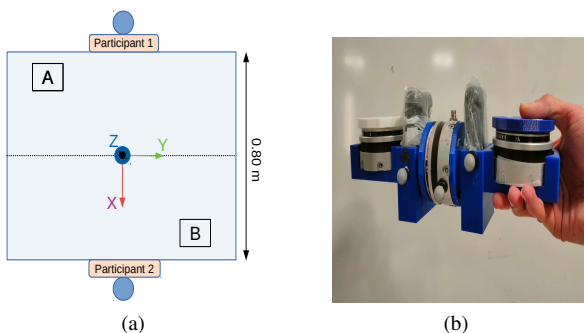


Fig. 3. (a) Table (0.8 m wide) setup for human-human handovers (b) Precision grasp on baton with lead weights added

the receiver, who would then place the baton at position B. Further, participant 2 acting as giver, would pick up the baton and hand it over to participant 1, who places it back at position A, and the process would be repeated. To preserve the naturalness of handovers, they were not given any instruction regarding the speed and location of handover. To reduce ergonomic load, positions A and B were marked close to the dominant hand of participants. The layout in Fig. 3(a) corresponds to two right handed participants.

3) *Experimental Settings*: The handovers occurred in three different settings, and the participants were required to do two sets of 9 minutes for each of the three settings:

- Setting 1 (S1) - **Normal**: In this setting, the baton was used for handover without any external weight. Neither giver nor taker was visually impaired by closing the eyes. The corresponding sets were named Set 1 and 2.
- Setting 2 (S2) - **Heavier Baton**: In this setting, two lead weights of 0.5 kg were added to the slots of the baton, leading to a total weight of 1.8 kg (Fig. 3(b)). This enables us to analyze the effects of a significant increase in weight on different aspects of handovers, particularly the interaction forces. These sets are called Set 3 and 4.
- Setting 3 (S3) - **Giver vision impairment**: In this setting, the giver was asked to close the eyes while handing over the baton without added weights. The giver would pick the baton normally, with eyes closed only during the giving phase as the receiver took the baton. The giver would open their eyes once the baton had been completely released. This scenario is similar to a robotic giver lacking vision input during handover and must rely solely on interaction forces. This setting allows us to analyze the effects of sensory impairment and the sets are called Set 5 and 6.

### C. Procedure

1) *Introduction and Instructions*: Before each experimental session with a participant pair, they were given a brief outline of what they were required to do in the experiment. They were given verbal and written explanations of all steps, signed an informed consent form regarding the experiment. They were also assigned as Participants 1 and 2. To ensure the naturalness of the handovers, they informed about the aim of the study only at the end of the experiment. They were suggested to get acquainted with each other and choose some mutual topics of interest to talk about while doing the handovers. This was an important step to reduce monotonicity in the experiment, which could lead to unnatural behavior.

The participants were given sanitized MoCap suits of their size to wear. After the MoCap markers were properly placed on their suits, they were allowed to enter the MoCap room and stand across the table. They were instructed to use precision grasp for holding the baton, which was then demonstrated as in Fig. 3(b), and to do some practice handovers. A precision grasp has been chosen because it involves only two opposing grasp surfaces, as opposed to multiple oblique contact surfaces involved in other type of grasps, making it easier and more accurate to measure forces using a single sensor. This grasp type has also been used in

similar studies for grip forces [5], [8], [20] and it is shown in [21], [16] that precision grips make up about 50% of all grips used in tool-intensive activities at homes and industry.

2) *Experimentation*: For all pairs, the experimentation started with Set 1 of handovers in normal setting. There was no restriction on how many handovers the participants need to do in the 9 minute allotted time for each set. The participants would then have a break of 2 minutes where they were allowed to sit in the chairs provided. The further order of sets were randomized so that no sets of similar strategies come after each other. After completing three sets of the experiment, they were given a longer break of minimum 5 minutes and the experiment continued when they were ready again. They were also allowed to take any breaks in between.

#### D. Data Collection

For each set of experimentation by a participant pair, multiple sensory data were recorded by reading through Robot Operating System (ROS) [22] as common interface. The data from F/T sensors is read using publicly available libraries for ROS at 333Hz. This data contains 6D wrench for the interaction and grip force sensors. The tracking data from MoCap has the pose tracking of baton and upper body skeletal representation of the two participants which contains pose tracking of 13 segments: Hip, Ab, Chest, Neck, Head, Left-shoulder, Left-upper arm, Left-forearm, Left hand, Right-shoulder, Right-upper arm, Right-forearm, Right hand. Each segment frame of the skeleton is tracked with respect to the world frame, which lies on the floor beneath the center of the table. This data is broadcasted by MOTIVE software from Optitrack [19] at a frequency of 120Hz and read by a customized ROSnode. Furthermore, this ROSnode [23] reads the original data from F/T sensors and republishes synchronized F/T and MoCap tracking data with common timestamps at 120Hz. For each body-segment, the tracking data is published in form of ROS-pose messages which contain 3D position and 4D quaternion rotation information. Finally, the recorded data for each experimental set involved the synchronized data from F/T sensors, the baton pose and skeletal representation for participant 1 and 2 at 120 Hz.

### IV. DATA SET

The continuous recordings of handovers for each pair was post-processed to separate individual handovers.

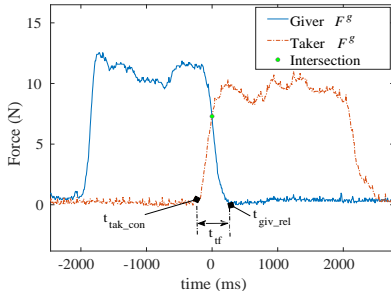


Fig. 4. Grip force ( $F^g$ ) variation in a particular handover.

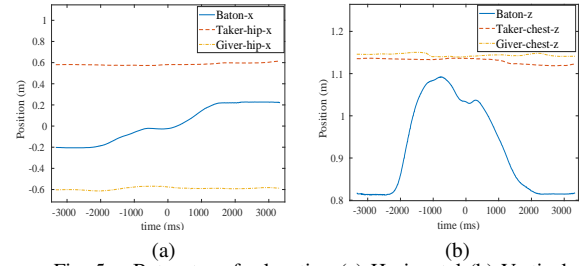


Fig. 5. Baton transfer location (a) Horizontal (b) Vertical

#### A. Data set Creation

In every handover, the giver lets go of the object as it is transferred to the taker, who takes hold of it. As a result, for each handover, the giver's grip force decreases to zero, whereas the taker's grip force increases from zero, resulting in an intersection of grip forces. In force-space, a handover is shown in Fig. 4 with the intersection point of grip-forces ( $F^g$ ) of giver and taker. The existence of this intersection point for each handover was leveraged to separate every handover in the recorded data. As can be seen in the Fig. 4, the handovers are centered about the intersection point at  $t = 0$  ms. The separated handovers are saved for a duration of 3.333 seconds (400 timestamps at 120 Hz) before and after the intersection point. For all handovers, this duration was sufficient to capture the baton being picked up from the table by the giver, the handover to the taker and baton being placed again on the table. The larger duration was used to verify that the baton was indeed transferred from the giver to taker as can be seen in Fig. 5. It also enables analysis of multiple aspects of human motion before and after the baton transfer. For this data set, we saved the recorded data with common timestamps i.e. the republished data. The data set is publicly available at: <https://github.com/paragkhanna1/dataset>. The data set contains a total of **8672** handovers, distributed among the 3 experimental settings as follows:

- S1 - **2999** normal handovers.
- S2 - **2764** handovers with heavy baton.
- S3 - **2909** handovers with giver vision impairment.

#### B. Features

Table I summarizes all information saved for an individual handover. It lists the signal names with which they are available in the data set. As stated earlier, each signal is

Signal No.	Signal Name	Signal Components
1	Wrench_interaction	Force (x,y,z) Torque(x,y,z)
2	Wrench_giver*	"—"
3	Wrench_taker*	"—"
4	baton_pose	Position(x,y,z) Rotation( $q_0, q_1, q_2, q_3$ )
5-17	giver-Skeleton	13 bodies-pose
18-30	taker-Skeleton	13 bodies-pose

\* Grip force is given by  $-Force(z)$

TABLE II  
TRANSFER TIME

Pair No.	$t_{tf}$ (Mean $\pm$ SD, ms)			t-test significance	
	S1	S2	S3	S1-S2	S1-S3
1	<b>414.6 <math>\pm</math> 127.0</b>	481.4 $\pm$ 114.5	540.3 $\pm$ 160.8	✓	✓
2	545.5 $\pm$ 185.2	579.4 $\pm$ 163.3	601.3 $\pm$ 198.1	X	X
3	<b>398.5 <math>\pm</math> 136.4</b>	484.7 $\pm$ 154.8	583.0 $\pm$ 198.5	✓	✓
4	<b>471.9 <math>\pm</math> 184.9</b>	632.9 $\pm$ 181.5	560.7 $\pm$ 183.0	✓	✓
5	<b>502.2 <math>\pm</math> 169.4</b>	610.9 $\pm$ 148.4	542.5 $\pm$ 166.6	✓	✓
6	<b>475.5 <math>\pm</math> 197.7</b>	513.6 $\pm$ 174.7	660.9 $\pm$ 181.9	✓	✓
7	<b>577.1 <math>\pm</math> 168.0</b>	622.6 $\pm$ 180.9	693.8 $\pm$ 176.2	✓	✓
8	<b>474.3 <math>\pm</math> 170.1</b>	654.1 $\pm$ 166.1	649.1 $\pm$ 178.4	✓	✓
9	414.9 $\pm$ 165.5	431.4 $\pm$ 127.8	412.2 $\pm$ 127.9	X	X
10	<b>439.5 <math>\pm</math> 157.7</b>	487.1 $\pm$ 151.0	512.1 $\pm$ 148.6	✓	✓
11	<b>490.3 <math>\pm</math> 181.7</b>	514.2 $\pm$ 148.3	561.4 $\pm$ 189.3	X	✓
12	489.1 $\pm$ 207.5	527.3 $\pm$ 161.2	490.2 $\pm$ 195.5	X	X
13	<b>581.8 <math>\pm</math> 173.4</b>	739.9 $\pm$ 214.4	714.3 $\pm$ 235.4	✓	✓

$t_{tf}$  values in **bold** show significance in one way ANOVA analyses,  $p < 0.05$

centered around the grip-force intersection point and has a total duration of 6.675 seconds. The metadata includes the height, arm-lengths, age, and handedness of the two subjects.

## V. DATA ANALYSIS

### A. Transfer time - $t_{tf}$

Considering the haptic interaction in a handover, the time taken to transfer the object has been defined as the time from initial contact by the taker to final release by the giver [5]. As shown in the Fig. 4, the taker's contact ( $t_{tak.con}$ ) was determined when the taker- $F^g$  rose above a threshold before the intersection point at  $t = 0$  ms. Similarly, the giver's final release ( $t_{giv.rel}$ ) is determined when the giver- $F^g$  falls below the threshold. This threshold was set to 0.4 N upon investigating the sensor noise in the recorded data. The transfer time is then given by

$$t_{tf} = t_{giv.rel} - t_{tak.con}$$

Variation in the mean transfer time for each pair and among the three settings is shown in Fig. 6. It can be distinctly seen that both a rise in weight and a giver sensory impairment result in an increase in mean  $t_{tf}$ . For further analysis, we needed to determine the significance of the effect on  $t_{tf}$  under different settings. The difference between the settings is found significant with  $p < 0.0001$  for a one-way ANOVA analysis across the entire trial population with pair-number as a random effect. For each participant-pair, we also performed t-test between the recorded  $t_{tf}$  for pair of settings (S1-S2, S1-S3) and one way ANOVA analysis among the three settings. A significance level of  $\alpha = 0.05$  was considered

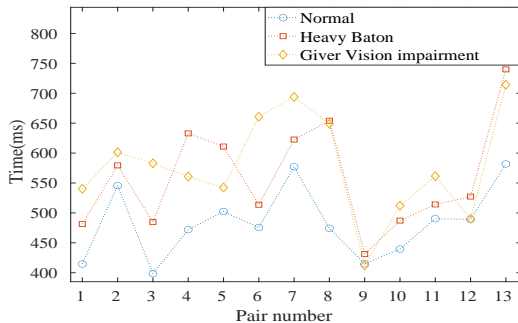


Fig. 6. Mean transfer time

TABLE III

Pair No.	$t_{gr}$ (Mean $\pm$ SD, ms)			t-test significance	
	S1	S2	S3	S1-S2	S1-S3
1	<b>400.1 <math>\pm</math> 114.3</b>	475.5 $\pm$ 105.7	509.6 $\pm$ 153.0	✓	✓
2	522.5 $\pm$ 163.1	570.7 $\pm$ 152.6	576.6 $\pm$ 187.5	✓	X
3	<b>388.6 <math>\pm</math> 123.6</b>	469.3 $\pm$ 140.9	555.2 $\pm$ 196.3	✓	✓
4	<b>437.5 <math>\pm</math> 160.1</b>	618.7 $\pm$ 169.0	514.2 $\pm$ 162.8	✓	✓
5	<b>473.4 <math>\pm</math> 154.6</b>	591.7 $\pm$ 133.9	485.4 $\pm$ 147.2	✓	X
6	<b>458.6 <math>\pm</math> 178.5</b>	505.1 $\pm$ 167.6	632.3 $\pm$ 170.7	✓	✓
7	<b>559.5 <math>\pm</math> 155.4</b>	609.6 $\pm$ 169.6	660.4 $\pm$ 167.2	✓	✓
8	<b>454.1 <math>\pm</math> 153.6</b>	633.5 $\pm$ 150.8	619.4 $\pm$ 163.8	✓	✓
9	397.5 $\pm$ 145.9	427.2 $\pm$ 121.8	392.5 $\pm$ 121.2	X	X
10	<b>420.5 <math>\pm</math> 137.1</b>	480.7 $\pm$ 145.2	476.4 $\pm$ 133.9	✓	✓
11	<b>449.3 <math>\pm</math> 155.5</b>	498.8 $\pm$ 141.3	520.1 $\pm$ 175.6	✓	✓
12	455.4 $\pm$ 175.8	514.5 $\pm$ 151.3	454.2 $\pm$ 173.7	✓	X
13	<b>553.6 <math>\pm</math> 153.3</b>	702.9 $\pm$ 192.8	646.3 $\pm$ 213.0	✓	✓

$t_{gr}$  values in **bold** show significance in one way ANOVA analyses,  $p < 0.05$

and the analysis is summarized in Table II which also shows the mean  $t_{tf}$  and standard deviation (SD). As per t-test analysis, an increase in weight caused a significant increase in  $t_{tf}$  for 69.2% participants. With giver vision impairment, a significant increase in  $t_{tf}$  was seen for 76.9% participants.

### B. Grip release time - $t_{gr}$

The grip release time is the total time a giver takes to release the object in handover, which we believe to be of particular importance for the design of a robotic giver based on human data. This time was calculated by observing the total time in which the giver's grip force ( $F^g$ ) decreased to zero. The giver grip release was considered to start when a first decrease in giver- $F^g$  was observed after the taker's contact, and it ended when the giver forces were reduced below the minimal threshold (at  $t_{giv.rel}$ ).

Fig. 7 shows the variation in mean  $t_{gr}$  for different pairs across different experimental settings and Table III summarizes the significance analysis. We observe that for most pairs,  $t_{gr}$  increases with heavier baton and giver vision impairment. The results of a one-way ANOVA analysis with pair number as random effect ( $p < 0.001$ ) show that the difference is found significant across all participants. As per t-test analysis, an increase in weight caused a significant increase in  $t_{gr}$  for about 92.3% participants. With giver vision impairment, a significant increase in  $t_{gr}$  was seen for approximately 69.2% participants.

### C. Interaction Forces - Pull force

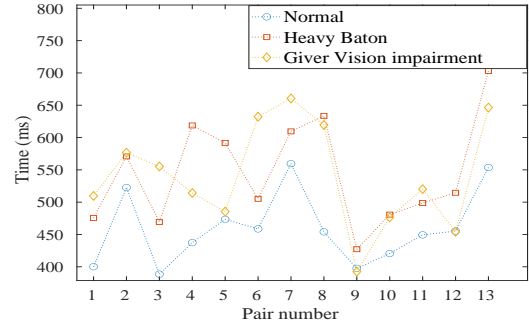


Fig. 7. Mean grip release time

In human-robot handovers, other than vision input, the robot can rely on interaction forces to detect, adapt, and

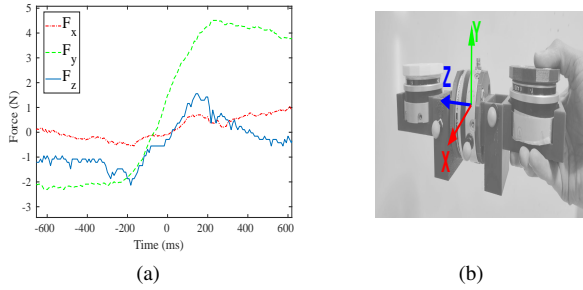


Fig. 8. Interaction Force: (a) Plot (b) Interaction sensor

perform the give or take in handover. For the input of interaction forces, robots are often equipped with wrist sensors or can also use end-effector torque estimation. Hence, we believe the analysis of interaction forces to be important. Multiple robot to human handovers experimentation studies used pulling force to command grip release [8], [9].

Considering our experimental case of horizontal transfers, the pulling force is in the direction of baton transfer. As a result, when a giver hands over the baton, the pull force in our baton is given by change in the  $F_z$  component of forces measured by interaction sensor (Fig. 8). We analyzed the maximum pull force seen in the handovers by considering the maximum change observed in  $F_z$  during the transfer time. Fig. 9 shows the average maximum-pull force normalized to object weight across different settings. Across most participant pairs, we found this ratio to be higher when giver vision was impaired as compared to normal handovers. This is supported by significance in one-way ANOVA analysis with pair number as random effect ( $p < 0.05$ ). However, a lower ratio is seen for heavier baton transfer. This led to another analysis inspecting the absolute value of maximum pulling force, whose average variation is shown in Fig. 10. We observe that the average pull force is higher for heavy baton transfer than the normal baton transfer ( $p < 0.05$ ).

#### D. Interaction Forces - Load-sharing

When humans carry an object, they feel the load (weight) of the object. This load reduces as the object is being handed over to the other person, as load-sharing between giver and taker occurs. The handover ends in complete transfer of load from giver to taker. In case of horizontal handovers, the felt load corresponds to the vertical component of interaction

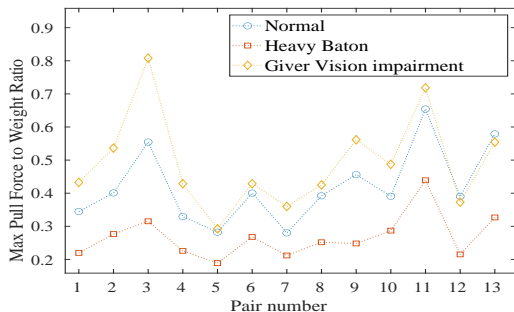


Fig. 9. Mean of maximum pull force normalized to weight

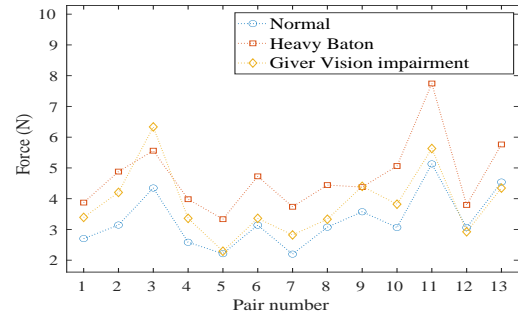


Fig. 10. Mean of maximum pull force

forces ( $F_y$ ) in Fig. 8. Load-sharing is considered by a robotic giver in [20], [24], where the object was considered to be in a *sharing state* when the measured vertical load ( $F_y$ ) reduced to a fraction of the initial load detected ( $F_y^i$ ) due to load-sharing by human taker. For robot to human handovers in [20], this fraction set to 0.5. However, the grip release was commanded by a pull force threshold detection.

An interesting analysis is to study correspondence between load-share shift/transfer to grip-force intersection. In the baton transfer, the load-share shift occurs when measured  $F_y$  changes its sign, i.e. in our case changes from negative to positive, that is at time

$$t_{ld\_shift} = \min(t : F_y(t) > 0, t \in [t_{tak.con}, t_{giv.rel}])$$

The dominant grip force changes at the intersection for grip forces at  $t = 0$  ms for each handover in our data set. Hence, it is sufficient to observe absolute  $t_{ld\_shift}$  values to compare the load-share shift to the intersection point of grip forces. A plot of mean  $t_{ld\_shift}$  observed is given in Fig. 11. A positive mean is observed across all settings and it is found that the values are significantly above zero ( $p < 0.05$  in t-test analysis) for all pairs. Thus, the load-share shift occurs after the grip forces' intersection, across all experimental settings. In other words, we validate that the dominant grip force changes before humans actually feel the load-share shift.

#### E. Height of Transfer

In [14], the analysis of a human-human study showed that transfer height does not depend on the object weight in handovers. In our analysis, for all handovers, we consider the baton transfer height as the measured Z coordinate of

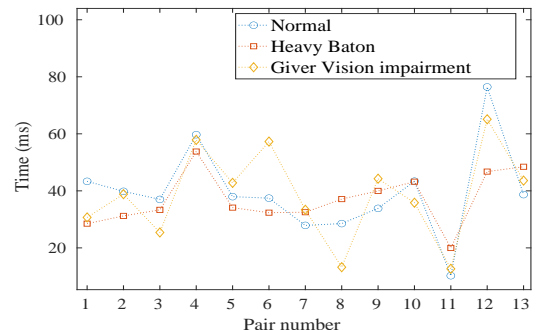


Fig. 11. Mean time of load-share shift

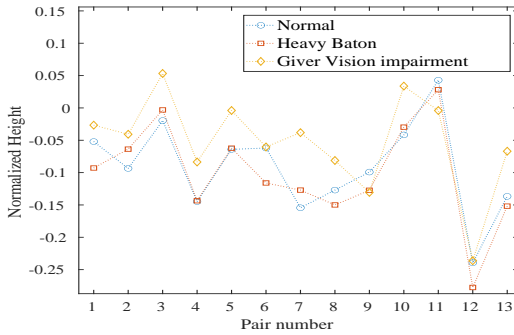


Fig. 12. Mean normalized height of baton transfer

baton at the intersection of grip forces, i.e. at  $t = 0$  ms. For proper analysis across different pairs, this was normalized to the average chest height of giver and taker. The mean variation in the normalized height across different settings is shown by Fig. 12. A significant difference is seen in the height of transfer when comparing the three via the results of one-way ANOVA analysis with pair number as random effect ( $p < 0.001$ ). Posthoc analysis show that the height of transfer is higher for the Setting 3. We observe a higher mean height of transfer with 69.23% participant-pairs in case with giver vision impairment. We also corroborate that the weight of the object had no significant effect on the height.

## VI. DESIGN IMPLICATIONS

Based on our analysis, we propose some design implications for robots in human-robot handover scenarios.

1) *Transfer Time*: Our overall analysis shows that the  $t_{tf}$  is close to 500 ms which agrees well with the reported average  $t_{tf}$  for human handovers in literature [9]. Our study further suggests that in human-robot handovers, a  $t_{tf}$  of 500 ms shall be targeted. As a significant increase in weight caused increase in  $t_{tf}$ , a robot should expect and plan for a longer  $t_{tf}$  for heavier objects or if the weight increases in the interaction. In cases when a giver is not looking, a longer  $t_{tf}$  suggests that both giver and taker adapt for sensory restrictions of the giver. The robot should also adapt for a longer  $t_{tf}$  if the giver is not actively looking at the handover.

2) *Grip Release time*: A higher grip release time with a heavier object for most participants is an interesting finding, as this indicates humans adapt their grip release time according to the object weight. In case of a similar, but heavy object, this might imply that more caution is exercised by humans to prevent handover failure or object fall. Thereby, in a human-robot interaction scenario, a human inspired robotic giver shall adapt its  $t_{gr}$  for heavier objects. If a robot, during the interaction, detects that the current object is significantly heavier than other objects, it should increase its  $t_{gr}$  for this handover. This will be especially important for robot givers relying on pre-planned grip force modulation [8]. Without vision input, the giver relies only on haptic feedback for grip release, which leads to a larger resistance by the giver before releasing, indicated by a larger  $t_{gr}$ . This leads to a significantly larger pull force by the taker. Thus, a robotic taker should expect more resistance in releasing the object

by a human giver who is not actively looking during the handover. A robotic giver which increases its  $t_{gr}$  should also expect a corresponding increase in pulling force as well.

3) *Interaction Forces*: With the heavier baton, the increase in pull forces corresponds to increased caution displayed by both giver and taker to enable successful transfer of the heavier baton. Overall, we can say that a heavier object would lead to higher pulling forces in the handover, but the increase in pulling forces is not directly proportional to the object weight. So, humans tend to pull more while taking a heavier object but will not generally pull or expect to pull above a certain force. A robotic giver or taker should also expect increased resistance and pulling force with a high increase in object weight.

It is important to note that a robot giver only has the input of interaction forces and not the grip forces of a human taker. Our analysis with  $F_y$  implies that it is safe for the robot to assume the taker has applied sufficient grip force when load-share shift occurs. If grip release is not yet triggered based on any other strategy, the robot can start decreasing the grip forces at the load-share shift. Such a strategy can ensure more safety in robot to human handovers, especially in cases of uncertainty like a sudden increase in the transfer object's weight that was not previously known to the human taker. However, this robotic grip-release will be commanded later than an ideal human giver's grip-release. Thus, the robot should adopt a higher grip-release time ( $t_{gr}$ ) to preserve the naturalness by aiming for a total  $t_{tf}$  of 500ms.

4) *Transfer Height*: As per our analysis, transfer height increases for a human giver when they rely on just haptic feedback and either do not or can not use the visual input. Thus, a robot taker should adapt by planning for higher transfer height when a human giver is not looking while giving the object.

## VII. LEARNING GRIP RELEASE

To perform a quantitative assessment of the data set, a human-inspired data-driven strategy was proposed to command robotic grip release in robot to human handovers. The creation of this strategy and further evaluation via human-robot experimentation have been discussed in detail in [25]. During a handover, it is necessary for a robotic giver to determine when to start the grip release as the human taker forms its grip on the object. The external wrench due to human interaction can be measured by a F/T sensor placed at the robot wrist. The idea is to investigate the interaction wrench ( $W_{int}$ ) in human handovers and then predict the start of giver grip release based on  $W_{int}$ .

1) *Data-driven strategy*: For each recorded handover,  $W_{int}$  comprises of a time series of 6 F/T components which vary significantly in the handover, as seen for forces in Fig. 8. As seen in Fig. 4, the giver's grip forces starts to decline before the intersection point ( $t = 0$  ms). The learning task is to predict the grip release start from the observed  $W_{int}$  before  $t = 0$  ms. Thereby, we sampled a time series of  $W_{int}$  ending at a time-step  $t_e$  from each saved handover:

$$X(t_e) = \{W_{int}^t, t \in [t_e - 100, t_e]\}. \quad (1)$$

This 100 step time series corresponds to 833 ms at 120Hz for our recorded data, with  $t_e$  varying between 0 to -1803 ms. Considering the average  $t_{tf}$  of 500 ms [5], [9], the labels ( $l$ ) to the samples were assigned based on  $t_e$ , given by Equation 2. If the sample's  $t_e$  lies after  $t = -250$  ms, i.e. in the interval  $(-250, 0]$  ms, the input time series of  $W_{int}$  corresponds to a grip release start ( $l = 1$ ). Otherwise, the sample does not represent a grip release start ( $l = 0$ ).

$$l(X(t_e)) = \begin{cases} 1, & \text{if } t_e > -250ms \\ 0, & \text{if } t_e \leq -250ms \end{cases} \quad (2)$$

The labeled samples were then used to train a long short-term memory (LSTM) based classifier. The trained LSTM is then used for robot to human handovers where  $W_{int}$  is measured. After the robot reaches the fixed final pose for handover, the time series of  $W_{int}$  is continuously fed to the classifier to command the grip-release. The LSTM classifier assigns the labels 1 and 0 implying whether the input  $W_{int}$  corresponded to a grip release start in human handovers.

2) *Robot to Human handovers experimentation*: The data-driven strategy was compared against loadshare and pull force based grip release strategies (Section 5.3, 5.4) in robot to human handovers. An experimental study was designed involving 20 participants taking objects from Baxter robot [25]. The robot giver measured  $W_{int}$  via a wrist F/T sensor and relied only on this sensory information to detect the handover, commanding its grip release. It was shown that the human-inspired strategy led to faster detection of handovers with the training object, reducing the  $t_{tf}$ . This data driven strategy was also ranked most natural by 95% participants and preferred over other strategies with high significance.

## VIII. CONCLUSION AND FUTURE WORK

A human study involving 26 participants in 13 pairs was done to explore multi-sensory information in human handovers. In this study, a proper experimental setup was devised to study the effect of a high increase in transfer weight and the effect of a lack of visual sensory feedback with a human giver. The data from study was processed to segregate individual handovers, leading to a multi-sensory large data set of human handovers. Further, we present a detailed analysis of the effect of the experimental settings on various aspects of human handovers. We also discuss possible design implications for a robot aiming for handovers between humans and robots. As a quantitative evaluation of the dataset, we propose and evaluate a data driven strategy for grip release of a robotic giver in robot to human handover.

In the future, we intend to investigate how we can make better use of our human handover dataset in a quantitative setting. We plan to further research on applying our work to produce more realistic human-robot handovers and what human-inspired metrics may be applied to assess their quality. It will also be interesting to study how human perceptions of a natural handover change when a robot is involved.

## ACKNOWLEDGMENT

This work was supported by Digital Futures at KTH.

## REFERENCES

- [1] A. Bilyea, N. Seth, S. Nesathurai, and H. Abdullah, "Robotic assistants in personal care: A scoping review," *Medical Engineering & Physics*, vol. 49, pp. 1–6, 2017.
- [2] S. Robla-Gómez, V. M. Becerra, J. R. Llata, E. González-Sarabia, C. Torre-Ferrero, and J. Pérez-Oria, "Working together: A review on safe human-robot collaboration in industrial environments," *IEEE Access*, vol. 5, pp. 26 754–26 773, 2017.
- [3] G. Bardaro, A. Antonini, and E. Motta, "Robots for elderly care in the home: A landscape analysis and co-design toolkit," *Int. J. Soc. Robot.*, vol. 14, 04 2022.
- [4] U. E. Ogenyi, J. Liu, C. Yang, Z. Ju, and H. Liu, "Physical human-robot collaboration: Robotic systems, learning methods, collaborative strategies, sensors, and actuators," *IEEE Transactions on Cybernetics*, vol. 51, no. 4, pp. 1888–1901, 2021.
- [5] W. P. Chan, C. A. Parker, H. M. V. der Loos, and E. A. Croft, "A human-inspired object handover controller," *The International Journal of Robotics Research*, vol. 32, no. 8, pp. 971–983, 2013.
- [6] S. Parastegari, B. Abbasi, E. Noohi, and M. Zefran, "Modeling human reaching phase in human-human object handover with application in robot-human handover," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 3597–3602.
- [7] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Toward seamless human-robot handovers," *J. Hum.-Robot Interact.*, vol. 2, no. 1, p. 112–132, 2013.
- [8] M. Controzzi, H. Singh, F. Cini, and T. Cecchini, "Humans adjust their grip force when passing an object according to the observed speed of the partner's reaching out movement," *Exp. Brain Res.*, vol. 236, 2018.
- [9] V. Ortenzi, A. Cosgun, T. Pardi, W. P. Chan, E. Croft, and D. Kulić, "Object handovers: A review for robotics," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1855–1873, 2021.
- [10] W. Chan, M. Pan, E. Croft, and M. Inaba, "An affordance and distance minimization based method for computing object orientations for robot human handovers," *Int. J. Soc. Robot.*, vol. 12, 01 2020.
- [11] M. Huber, R. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *RO-MAN - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, 2008, pp. 107–112.
- [12] R. Rasch, S. Wachsmuth, and M. König, "A joint motion model for human-like robot-human handover," *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, pp. 180–187, 2018.
- [13] F. R. Döhring, H. Müller, and M. Joch, "Grip-force modulation in human-to-human object handovers: effects of sensory and kinematic manipulations," *Scientific Reports*, vol. 10, no. 1, p. 22381, Dec 2020.
- [14] C. Hansen, P. Arambel, K. B. Mansour, V. Perdereau, and F. Marin, "Human-human handover tasks and how distance and object mass matter," *Perceptual and Motor Skills*, vol. 124, pp. 182–199, 2017.
- [15] A. Carfi, F. Fogliano, B. Bruno, and F. Mastrogiovanni, "A multi-sensor dataset of human-human handover," *Data in Brief*, vol. 22, pp. 109–117, 2019.
- [16] F. Cini, V. Ortenzi, P. Corke, and M. Controzzi, "On the choice of grasp type and location when handing over an object," *Science Robotics*, vol. 4, p. eaau9757, 02 2019.
- [17] R. Ye, W. Xu, Z. Xue, T. Tang, Y. Wang, and C. Lu, "H2o: A benchmark for visual human-human object handover analysis," in *IEEE Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 15 742–15 751.
- [18] Onrobots, "Onrobots force-torque sensors, webpage," <https://onrobot.com/en/products/hex-6-axis-force-torque-sensor>, [Accessed June 23, 2022].
- [19] OptiTrack, "Motion capture systems - optitrack, webpage," <https://optitrack.com/>, [Accessed Mar 23, 2023].
- [20] M. Costanzo, G. De Maria, and C. Natale, "Handover control for human-robot and robot-robot collaboration," *Frontiers Robot. AI*, vol. 8, p. 132, 2021.
- [21] J. Z. Zheng, S. De La Rosa, and A. M. Dollar, "An investigation of grasp type and frequency in daily household and machine shop tasks," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2011, pp. 4169–4175.
- [22] ROS, "Ros-webpage," <http://wiki.ros.org/>, [Accessed Mar 23, 2023].
- [23] P. Khanna, "Handover dataset," <https://github.com/paragkhanna1/dataset>.
- [24] J. R. Medina, F. Duvallat, M. Karnam, and A. Billard, "A human-inspired controller for fluid human-robot handovers," in *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, 2016, pp. 324–331.
- [25] P. Khanna, M. Björkman, and C. Smith, "Human inspired grip-release technique for robot-human handovers," in *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, 2022, pp. 694–701.