Data in Brief 22 (2019) 109–117

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

A multi-sensor dataset of human-human handover

Alessandro Carfì *, Francesco Foglino, Barbara Bruno,
Fulvio Mastrogiovanni

Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Via Opera Pia 13, 16145 Genoa, Italy

ARTICLE INFO

Article history:
Received 19 October 2018
Received in revised form
13 November 2018
Accepted 16 November 2018
Available online 27 November 2018

ABSTRACT

The article describes a multi-sensor dataset of human-human handovers composed of over 1000 recordings collected from 18 volunteers. The recordings refer to 76 test configurations, which consider different volunteer's starting positions and roles, objects to pass and motion strategies. In all experiments, we acquire 6-axis inertial data from two smartwatches, the 15-joint skeleton model of one volunteer with an RGB-D camera and the upper-body model of both persons using a total of 20 motion capture markers. The recordings are annotated with videos and questionnaires about the perceived characteristics of the handover.

Specifications table

Subject area	*Engineering*
More specific subject area	*Robotics, Human-Robot Interaction*
Type of data	Tables and Videos
How data was acquired	*Motion Capture (Optitrack Flex 13); RGB-D camera (Microsoft Kinect v2.0); Inertial Measurement Unit (LG G Watch R); Video camera; questionnaire.*

---

* Corresponding author.
E-mail address: alessandro.carfi@dibris.unige.it (A. Carfì).

https://doi.org/10.1016/j.dib.2018.11.110

Data format	*Raw and processed.*
Experimental factors	*Depth data processed online to extract human skeleton all the other data are raw.*
Experimental features	*Different test configurations which vary in terms of class of the experiment (double-blind or single-blind), role of the volunteer (giver or receiver), starting position of the two persons (with approach or without approach), object to pass (10 with different size, weight and stiffness) and in single-blind case strategy adopted by the experimenter (6 different strategies).*
Data source location	*EMARO Lab, Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, Italy (44.402241, 8.960811)*
Data accessibility	https://data.mendeley.com/datasets/cyr9gmhspy/1 DOI:10.17632/cyr9gmhspy.1
Related research article	*N/A*

Value of the data

- The multi-sensor dataset allows for the study of human-human handover.
- The multi-sensor dataset can be used to develop techniques for human-robot interaction (HRI) allowing robots to understand that an handover is taking place.
- The multi-sensor dataset can inspire new methods for more natural human-robot handovers.
- The multi-sensor dataset can also be used to compare the efficiency/accuracy of different sensors in perceiving humans.

1. Data

In this article, we present a multi-sensor dataset of human-human handovers (see Fig. 3). Each data file contains: i) the trajectories followed by giver and receiver as extracted by an Optitrack Flex 13 Motion Capture system using 10 markers on each person (see Fig. 1); ii) the 15-element skeleton of only one person as extracted by a Microsoft Kinect sensor (see Fig. 2); iii) the right wrist linear acceleration and angular velocity of both giver and receiver as measured by a LG G Watch R smartwatch equipped with an inertial sensor. Therefore, each data files contains a 47-dimensional description of the handover. All sensors are synchronised and sampled at a rate of approx. 7 Hz; each sample is timestamped to provide the exact timing reference. Each data file is associated with a video recorded at 8 fps; data files related to double-blind experiments (see Table 1) are annotated with the volunteers' ranking of the objects (see Fig. 4 and Table 2) with respect to their manipulability; data files related to single-blind experiments (see Table 3) are annotated with the volunteers' answers to 6 questions related to the naturalness, practicality, comfort, safety, speed and timing of the handover.

2. Experimental design, materials, and methods

2.1. Equipment

In our experiment the sensors, described in the following Sections, communicate with a laptop (CPU Dual-Core processor Intel i7 4500 1.80 GHz, 6 GB RAM, running Ubuntu 14.04 LTS with ROS Indigo) through a Wi-Fi network, providing each sensory sample as a ROS message. On the laptop, sensory data are grouped and timestamped with the timestamp of the arrival of the sample from the slowest sensor. The whole system runs at an average frequency of 6.6 Hz.

2.1.1. Motion capture

We used an Optitrack Flex 13 Motion Capture system (1.3 MP, 56° Horizontal FOV, 46° Vertical FOV, 28 LEDs, 8.33 ms latency), covering an area of 5 m × 5 m × 4 m with 8 cameras, to track 10 markers placed on the head (3 markers, placed on the forehead, above the left and above the right ear), left shoulder, right shoulder, tailbone, bellybutton, right elbow, right wrist and right hand of each person. The exact placement of the markers on a person is shown in Fig. 1. The markers placed on the torso (denoted with yellow circles in Fig. 1) allow for accurately tracking the person's trajectory and estimating the handover pose (which characterize the mutual influence of giver and receiver [1]), while the three markers (blue circles) on the head allow for tracking the head direction (which is related to

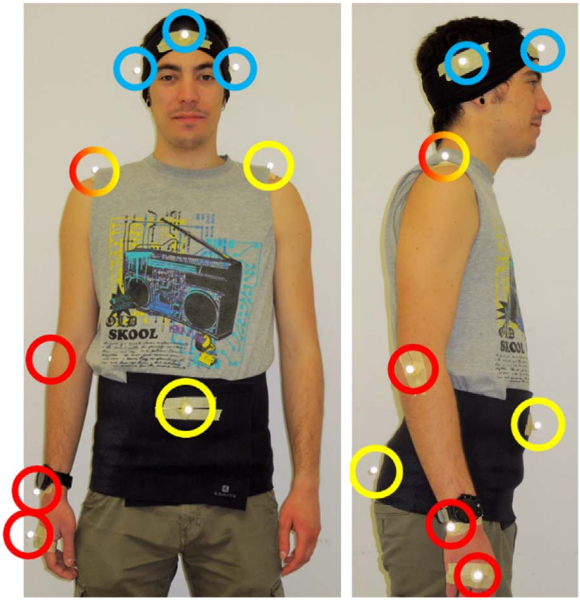

Fig. 1. Placement of the Motion Capture markers on a volunteer: blue circles denote the markers used to track the head direction, red circles denote the markers used to track the arm movement and yellow circles denote the markers used to track the approach trajectory and handover pose. The watch worn by the volunteer at the right wrist is the smartwatch used to collect inertial data.

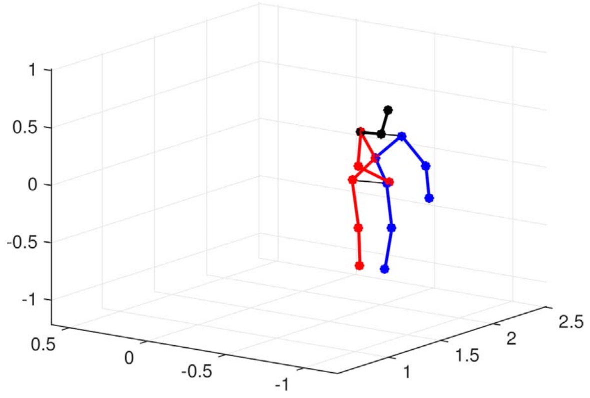

Fig. 2. The skeleton model of a volunteer acting as giver: circles denote the 15 skeleton joints and a red filling identifies the right side of the body.

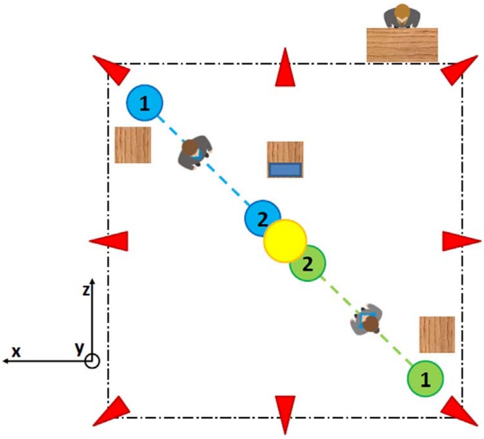

Fig. 3. Sketch of the test area. Red triangles correspond to the Motion Capture cameras, while the blue rectangle identifies the Kinect. The yellow circle denotes the area in which the handover takes place. The blue and green circles denote the starting positions (1 = with approach, 2 = without approach) of the people involved in the handover; in single-blind experiments, the experimenter uses the blue positions and the volunteer uses the green one. In Tables 1 and 3 the role is always specified with respect to the person starting from the green positions. The origin of the reference frame of the Motion Capture system is at the centre of the yellow circle. The reference frame in figure, describes the orientation of the Motion Capture reference frame whose origin is at the centre of the yellow circle.

Table 1
Double-blind test configurations.

ID	Role	Approach	Object	ID	Role	Approach	Object
1	giver	no	1	17	receiver	no	1
2	giver	no	2	18	receiver	no	2
3	giver	no	3	19	receiver	no	3
4	giver	no	4	20	receiver	no	4
5	giver	no	5	21	receiver	no	5
6	giver	no	6	22	receiver	no	6
7	giver	no	7	23	receiver	no	7
8	giver	no	8	24	receiver	no	8
9	giver	yes	1	25	receiver	yes	1
10	giver	yes	2	26	receiver	yes	2
11	giver	yes	3	27	receiver	yes	3
12	giver	yes	4	28	receiver	yes	4
13	giver	yes	5	29	receiver	yes	5
14	giver	yes	6	30	receiver	yes	6
15	giver	yes	7	31	receiver	yes	7
16	giver	yes	8	32	receiver	yes	8

gaze [2], used to exchange information by giver and receiver [3]). The four markers placed on the right arm and hand (red circles) contribute to the analysis of the giving/receiving movement.

2.1.2. Skeleton model

We used a Microsoft Kinect v2.0 to build a skeleton model composed of 3D information of 15 joints (left foot, left knee, left hip, left hand, left elbow, left shoulder, right foot, right knee, right hip, right hand, right elbow, right shoulder, torso, neck, head). The Kinect senses one person only in each experiment: since the volunteers alternated in the giver/receiver role, the dataset is equally split among recordings reporting the skeleton model of the giver and recordings reporting the skeleton model of the receiver. As an example, Fig. 2 shows the skeleton model of a volunteer acting as the giver, in the final moments of the transferring phase.

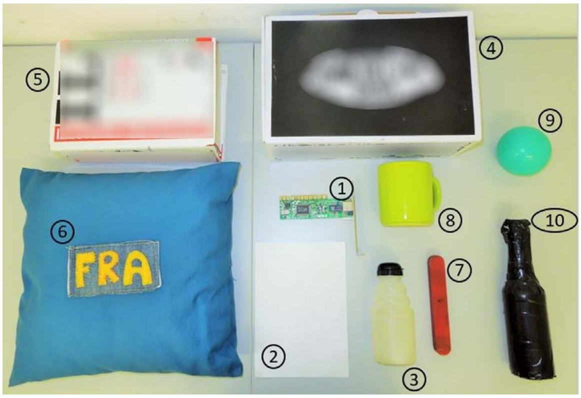

Fig. 4. The objects used in the experiments. We used the objects marked with numbers from 1 to 8 (included) in the double-blind set of experiments, and objects marked with numbers 8–10 in the single-blind set of experiments.

Table 2
Characteristics of the objects.

#	Name	Size	Weight	Stiffness
1	network card	small	light	hard
2	paper sheet	small	light	soft
3	full water bottle	small	heavy	hard
4	shoes empty box	big	light	soft
5	dictionary	big	heavy	hard
6	pillow	big	light	soft
7	stick	small	light	hard
8	mug	small	light	fragile
9	ball	small	light	soft
10	black bottle	small	–	hard

2.1.3. Inertial sensors

We used two LG G Watch R smartwatches (CPU Quad-Core Qualcomm Snapdragon 400 1:2 GHz, 512 MB RAM, 4 GB memory, 6-axis inertial sensor, 16 bits per axis precision), to acquire 6-axis inertial data of the right wrist of each person. The placement of the smartwatch on the arm is shown in Fig. 1.

2.2. Methods

We devised a number of test configurations, which vary in terms of class of the experiment, either involving two volunteers (double-blind class) or one volunteer and one experimenter (single-blind class), role and starting position of the two persons, object to pass and strategy adopted by the experimenter, when directly involved in the handover.

2.2.1. Double-blind experiments

Experiments belonging to the double-blind class follow the empirico-inductive approach and envision the experimenter to be a pure observer, while two volunteers directly perform the handover. Table 1 lists the 32 test configurations tested with this approach: the first column reports the unique ID of each test configuration and the second column specifies the class of the experiment; the last three columns define, respectively, the adopted role, starting position and object.

As anticipated above, a person involved in the handover can either have the role of giver or receiver. All the volunteers repeated each test configuration twice, once in the role of the giver and once in the role of the receiver, to reduce possible biases on the dataset. Fig. 3 shows a sketch of the

Table 3
Single-blind test configurations.

ID	Strategy	Role	Approach	Object
33	normal	giver	no	9
34	normal	receiver	no	9
35	normal	giver	yes	9
36	normal	receiver	yes	9
37	normal	giver	no	8
38	normal	receiver	no	8
39	normal	giver	yes	8
40	normal	receiver	yes	8
41	quick	giver	no	9
42	quick	receiver	no	9
43	quick	giver	yes	9
44	quick	receiver	yes	9
45	quick	giver	no	8
46	quick	receiver	no	8
47	quick	giver	yes	8
48	quick	receiver	yes	8
49	delay	giver	no	9
50	delay	receiver	no	9
51	delay	giver	yes	9
52	delay	receiver	yes	9
53	delay	giver	no	8
54	delay	receiver	no	8
55	delay	giver	yes	8
56	delay	receiver	yes	8
57	holding	giver	no	9
58	holding	receiver	no	9
59	holding	giver	yes	9
60	holding	receiver	yes	9
61	holding	giver	no	8
62	holding	receiver	no	8
63	holding	giver	yes	8
64	holding	receiver	yes	8
65	wrong pose	giver	no	9
66	wrong pose	receiver	no	9
67	wrong pose	giver	yes	9
68	wrong pose	receiver	yes	9
69	wrong pose	giver	no	8
70	wrong pose	receiver	no	8
71	wrong pose	giver	yes	8
72	wrong pose	receiver	yes	8
73	deceptive object	giver	no	10
74	deceptive object	receiver	no	10
75	deceptive object	giver	yes	10
76	deceptive object	receiver	yes	10

test environment, denoting with green and blue circles the starting positions of the two volunteers. The roles reported in Table 1 refer to the volunteer starting from the green positions.

The green and blue circles marked with number 1 in Fig. 3 denote the starting positions requiring the volunteers to walk towards each other for handing over the object (with approach, denoted with "yes" in Table 1), while the green and blue circles marked with number 2 allow the volunteers to immediately perform the handover (without approach, denoted with "no" in Table 1).

Fig. 4 shows the 10 objects used in the experiments, together with their identification number, while Table 2 details their characteristics and the rationale for their choice. With respect to the size, some objects are "small" (i.e., they can be comfortably grasped with one hand) and others are "big" (i.e., they can be barely grasped with one hand). With respect to the weight, we consider "light" objects (i.e., weighting less than 500 g) and "heavy" objects (more than 500 g). Finally, we consider an object as "soft" if it does not make noise nor damage when impacting the floor surface, "hard" if it makes noise in an impact, usually not damaging itself, “fragile” if it makes noise in an impact, usually damaging itself without affecting the floor surface.

The network interface card, albeit known to some, is an object that none of the volunteers manipulates in their daily activities. As Table 1 reports, in the experiments following the double-blind approach, we used the objects numbered from 1 to 8.

2.2.2. Single-blind experiments

Experiments belonging to the single-blind class follow the hypothetico-deductive approach and envision the direct involvement in the handover of the experimenter, who alternates with a volunteer in the roles of giver and receiver. The rationale is to provide data to understand the strategies applied by a human for giving and receiving an object by comparing his/her behaviour during a normal handover and when facing an unexpected movement from the counterpart (experimenter strategy). Fig. 3 shows a sketch of the test environment, denoting with green circles the starting positions of the volunteer and with blue circles the starting positions of the experimenter. The roles reported in Table 3 refer to the volunteer.

We identified 6 strategies to test for each role:

Volunteer giver – experimenter receiver

- *Normal*: the experimenter moves as he would normally do to receive an object.
- *Quick*: the experimenter moves his arm quicker than he would normally do to receive an object.
- *Delayed start:* once the volunteer initiates the transferring gesture, the experimenter keeps his arm still for about 2 s before reaching towards the volunteer to receive the object.
- *Holding:* the experimenter holds the object in place for about 2 s once both persons have touched it, i.e., he does not back away with the object after grasping it.
- *Wrong position:* once the volunteer initiates the transfer gesture, the experimenter moves his arm to reach a different position with respect to where the volunteer is aiming at. Based on observations of natural handovers, and due to the fact that all our volunteers are right-handed, the experimenter aimed at the volunteer's left shoulder.
- *Deceptive objects:* the object used for the handover is one out of four identical black bottles, which only differ in weight; the receiver discovers the weight of the chosen object at the transferring moment.

Volunteer receiver – experimenter giver

- *Normal*: the experimenter moves as he would normally do to give an object.
- *Quick*: the experimenter moves his arm quicker than he would normally do to give an object.
- *Delayed start:* once the volunteer initiates the transferring gesture, the experimenter keeps his arm still for about 2 s before reaching towards the volunteer to give the object.
- *Holding:* the experimenter holds the object in place for about 2 s once both persons have touched it, i.e., he does not release the object once the volunteer has firmly grasped it, as he would do.
- *Wrong position:* once the volunteer initiates the transfer gesture, the experimenter moves his arm to reach a different position with respect to where the volunteer is aiming at. Based on observations of natural handovers, and due to the fact that all our volunteers are right-handed, the experimenter aimed at the volunteer's left shoulder.
- *Deceptive objects:* the object used for the handover is one out of four identical black bottles, which only differ in weight; the receiver discovers the weight of the chosen object at the transferring moment.

The starting positions reported in Table 3 are defined as for the double-blind experiments, and the objects used for the handover (marked with numbers 8, 9 and 10) are shown in Fig. 4 and described in Table 2. In the “deceptive object” configurations, the giver had to pick one of four identical black bottles placed on a table at his/her right (see Fig. 3): the bottles were filled with different materials and had weights ranging between 200 g and 780 g.

In the single-blind experiments, each volunteer performs in a random order the 44 test configurations listed in Table 3. In the course of the session, the experimenter asks the volunteer to rate each configuration in terms of naturalness, practicality, comfort, safety, speed and timing, on a 4-items Likert scale. For the evaluation of timing, 1 means "very long" and 4 means "very short", while in all other cases 1 means "very little" and 4 means "very much".

Please notice that different volunteers interpreted the grades in different ways: a part of the volunteers assumed the grades to grow from the extreme lack of the given characteristic (grade 1), to an excessive presence of it (grade 4), while others simply assumed grade 1 to be the worst and grade 4 to be the best.

2.3. Environment

Fig. 3 is a sketch of the test environment, which is a $4 \times 4\,\mathrm{m}^2$ square. The red triangles mark the position of the Motion Capture cameras. We calibrated the Motion Capture system to maximise the accuracy along the considered diagonal (error $< 7.5 \times 10^{-3}$ m); its reference frame is oriented as shown in the Figure and centred around the yellow circle, which denotes the area where the handover takes place. The blue rectangle marks the position of the Kinect, which is placed on a table, to maximise the visibility of the green positions and which tracks the skeleton with respect to its centre. The position of the Kinect in the reference frame of the Motion Capture area has been recorded as $K = (-0.159;\ 0.996;\ 1.881)$ m. The smartwatch, worn as shown in Fig. 1, provides the inertial data with respect to its reference frame, which is a right-hand frame with origin at the centre of the case, the x-axis pointing towards the pivot and the z-axis pointing out of the screen.

The blue and green circles denote the starting positions (1 = with approach, 2 = without approach) of the people involved in the handover; in the single-blind experiments, the experimenter uses the blue positions and the volunteer uses the green positions. In Tables 1 and 3 the role is always specified with respect to the person starting from the green positions.

To ensure that the handover takes place in full view of the sensing equipment, i.e., inside the yellow circle, we drew the same circle on the floor. The circle has a diameter of 1.2 m, which corresponds to the boundary of the social space, i.e., the distance within which human-human cooperation actions usually occur, as noted in [4].

Lastly, we placed two tables at the right of the starting positions marked with number 1, to be used for the retrieval and laying of the objects.

To allow for an easier matching between different recordings, we asked the volunteers to stand in a specific pose, shown in Fig. 1, at the beginning (i.e., at the starting position) and the end of each test.

2.4. Dataset inspection

The dataset is composed of 1087 recordings, collected from 18 Italian right-handed volunteers performing the 76 test configurations. The volunteers are 12 men and 6 women, with ages ranging between 20 and 60. Details about the volunteers, including their familiarity and preferences on the chosen objects and their intimacy with the volunteer they were paired with, are provided together with the dataset.

The dataset is distributed on three folders: *sensor_data*, *questionnaires* and *videos*. The folders contain one subfolder for each volunteer, named *volunteer_[n]* (for the single-blind experiments) or *volunteer_[n]_volunteer_[m]* (for the double-blind experiments). The files in these subfolders are named *volunteer_[n]_[c]*, where *c* denotes the configuration ID. The files in the questionnaires folder contain the questionnaires (.xlsx format) filled by the volunteers during the single-blind experiments, while the files in the videos folder contain the video recordings (.m4v format) of all experiments.

Each file in the *sensor_data* folder contains the timestamped sensors readings of one experiment (.xlsx format), organised as:

$$row = \{t; p_{mc}; p_{sk}; \omega_{green}; a_{green}; \omega_{blue}; a_{blue}\}$$

where t corresponds to the timestamp of the sensors acquisition, measured in s, p_{mc} identifies a vector of 20 × 3 MoCap marker coordinates, measured in m, p_{sk} identifies a vector of 15 × 3 skeleton joint coordinates, measured in m, ω corresponds to the triaxial angular velocity measured in rad/s and a corresponds to the tri-axial acceleration measured in m/s^2. The green and blue labels identify, respectively, the smartwatch worn by the person starting from the green positions and the smartwatch worn by the person starting from the blue positions, shown in Fig. 3.

We provide MATLAB scripts for playing a video recording (*play.m*), for loading one or a number of sensor files (*loader.m*) and for loading a questionnaire file (*Qloader.m*). We also provide a MATLAB script for plotting the inertial (*inertiaplotter.m*), skeleton (*kinectplotter.m*) and markers (*mocaplotter.m*) data contained in one or more sensor files. Detailed information on the usage of the functions can be found in the HTML folder, or directly accessed within the MATLAB environment with the command “help *function name*”.

Acknowledgements

The research leading to these results has received funding from the European Commission ERASMUS+ Funding Scheme, under Grant agreement n. 2014-2616 (EMARO+).

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.11.110.

References

[1] A. Koene, S. Endo, A. Remazeilles, M. Prada, A. Wing, Experimental testing of the coglaboration prototype system for fluent human-robot object handover interactions, in: Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2014), Edinburgh, UK, 2015.

[2] M. Imai, T. Kanda, T. Ono, H. Ishiguro, K. Mase, Robot mediated round table analysis of the effect of robot’s gaze, in: Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2002), Berlin, Germany, 2002.

[3] K. Strabala, M. Lee, A. Dragan, J. Forlizzi, S. Srinivasa, M. Cakmak, V. Micelli, Towards seamless human-robot handovers, J. Hum.-Robot Interact. 2 (1) (2013) 112–132.

[4] E.T. Hall, The Hidden Dimension, Garden City, N.Y., Doubleday, 1996.