

Leveraging Submovements for Prediction and Trajectory Planning for Human-Robot Handover

Kyle Lockwood
Northeastern University
Boston, USA

Tianjie Zhu
Northeastern University
Boston, USA

Garrit Strengé
Northeastern University
Boston, USA

Taskin Padir
Northeastern University
Boston, Massachusetts, USA

Yunus Bicer
Northeastern University
Boston, USA

Mariusz P. Furmanek
Northeastern University
Boston, USA

Tales Imbiriba
Northeastern University
Boston, USA

Deniz Erdogmus
Northeastern University
Boston, USA

Sadjad Asghari-Esfeden
Northeastern University
Boston, USA

Madhur Mangalam
Northeastern University
Boston, USA

Mathew Yarossi
Northeastern University
Boston, USA

Eugene Tunik
Northeastern University
Boston, USA

ABSTRACT

The effectiveness of human-robot interactions critically depends on the success of computational efforts to emulate human inference of intent, anticipation of action, and coordination of movement. To this end, we developed two models that leverage a well described feature of human movement: Gaussian-shaped submovements in velocity profiles, to act as robotic surrogates for human inference and trajectory planning in a handover task. We evaluated both models based on how early in a handover movement the inference model can obtain accurate estimates of handover location and timing, and how similar model trajectories are to human receiver trajectories. Initial results using one participant dyad demonstrate that our inference model can accurately predict location and handover timing, while the trajectory planner can use these predictions to provide a human-like trajectory plan for the robot. This approach delivers promising performance while remaining grounded in physiologically meaningful Gaussian-shaped velocity profiles of human motion.

CCS CONCEPTS

• **Computing Methodologies** → **Machine learning approaches**;
• **Human-centered computing** → *Interaction devices*; • **Computer systems organization** → *Robotics*.

KEYWORDS

human-robot interaction, object handover, Gaussian process regression, dynamic movement primitives, human inference, trajectory planning

ACM Reference Format:

Kyle Lockwood, Yunus Bicer, Sadjad Asghari-Esfeden, Tianjie Zhu, Mariusz P. Furmanek, Madhur Mangalam, Garrit Strengé, Tales Imbiriba, Mathew Yarossi, Taskin Padir, Deniz Erdogmus, and Eugene Tunik. 2022. Leveraging Submovements for Prediction and Trajectory Planning for Human-Robot Handover. In *The 15th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '22)*, June 29-July 1, 2022, Corfu, Greece. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3529190.3529220>

1 INTRODUCTION

The use of service robots, designed to perform daily tasks and support basic activities, has enormous potential to increase independence and improve the quality of life for individuals with health issues related to age and disability [7]. However, the effectiveness of these robots hinges on their ability to integrate seamlessly into our lives. Collaborative physical tasks in particular pose difficult modeling challenges, as robots must emulate the inference of intent, anticipation of action, and coordination of movement that comes effortlessly to humans [1]. Handover, or the exchange of objects between a human and their robot collaborator, represents an essential function played by service robots in a home or assistive living setting. Though human-human handover is a seemingly simple task, it involves a complex perception-action coupling to determine when and where the handover will happen and choose an appropriate trajectory to receive the object [10]. Despite recent advancements in sensing and control, human-robot handovers are far from approaching the fluidity and flexibility of human-human collaboration [2].

Here, we develop two models that leverage a well described feature of human movement: Gaussian-shaped submovements in human velocity profiles, to act as robotic surrogates for human

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

PETRA '22, June 29-July 1, 2022, Corfu, Greece

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9631-8/22/06...\$15.00

<https://doi.org/10.1145/3529190.3529220>

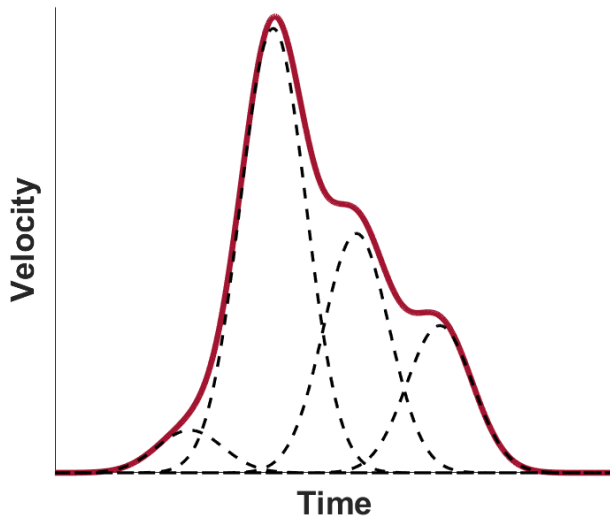


Figure 1: Velocity profile decomposed into its principle submovements. The complete velocity profile is depicted with a red solid line, while submovements (modeled as Gaussian curves) are depicted as dashed black lines.

inference and trajectory planning in a handover task. Continuous velocity profiles measured from a variety of human motions can be segmented into series of Gaussian-shaped components called submovements, see Fig. 1 [4]. Submovements have been utilized in autonomous robotics to generate complex motor behaviors that emulate human trajectories, and can adapt to the perturbations of a realistic environment [9]. This strategy has also been successfully utilized for robotic control of complex motor tasks including reaching movements, drawing 2D patterns, and tennis swings [3], hitting a baseball [8], and tasks for service and household robots [11].

The Gaussian shape of submovements inspired our approaches to modeling both human inference and trajectory planning. The model for human inference utilizes Gaussian process regression: a kernel based bayesian approach to regression. This approach allows us to leverage our knowledge about the functional form of human trajectories through the selection of a squared-exponential kernel. The shape of the kernel function serves to represent the Gaussian shape we observe in submovements. A similar approach using a hybrid Gaussian process and stochastic classification model showed promising results when applied to predicting human trajectories for human-robot interactions [5]. The model for trajectory planning uses a similar but reversed approach, whereby a sequence of submovements is calculated based on the results of the model for human inference, and combined to form a complete velocity profile. Submovements are calculated such that their individual contributions to the velocity profile account for updates to the predictions of handover location and timing that exceed an error threshold. This approach differs substantially from other approaches which train models on human example trajectories, and may fail when presented with handover cases which are not well characterized in the training data [14]. This paper demonstrates the efficacy of submovement inspired handover inference and trajectory planning models for robotic control by describing model performance in

reproducing human inference and trajectory planning extracted from experimental data collected during human-human handover.

2 METHODS

2.1 Experimental Setup and Data Collection

All protocols were conducted in conformance with the Declaration of Helsinki and were approved by the Institutional Review Board of Northeastern University. Two young healthy right handed individuals participated after providing institutionally approved written informed consent. Standing across from each other separated by a 60cm square table (adjusted to be roughly waist height) table, subjects performed 384 handovers in a single 180-min session, with the trials spread over 2 Roles (Giver, Receiver) $\times$ 2 Leads (Initiator, Follower) $\times$ 2 Objects (Small, Large) $\times$ 2 Initial position of the object (Right, Left of the Giver) $\times$ 4 Final position of the object (Right Shoulder (Top Right), Right Hip (Bottom right), Left Shoulder (Top Left), and Left Hip (Bottom Left) of the Receiver) $\times$ 6 Trials per combination of the other variables (5 standard trials, 1 perturbation trial). The duration of the inter-trial intervals was randomized and ranged from 1–3 s to prevent the giver from adopting a periodic pattern, which could be used by the receiver as temporal cue. So, each participant performed 192 trials as the Giver and 192 trials as the Receiver, of which he/she performed 96 trials as the Initiator and 96 trials as the Follower. Subjects wore earbuds, and separate recorded instructions to each subject were played prior to each trial describing the experimental condition. The start of each trial was indicated by an acoustical signal delivered to the initiator. The follower only received the instruction to grasp the object and was not given the start tone. All trials started with the object and the receivers hand placed on a copper tape-based electrical switch so that lifting the hand (or object) off the tape broke the circuit and provided a clear indicator of trial start.

A 21-IRE Motion Capture System (Qualisys AB, Sweden, 8 x Oqus 700+ and 13 x Miquis M3, 100 Hz) was used to collect full body position data. Separate 3D models were created for each participant using the modified Helen Hayes marker set with 55 markers per subject [6]. An additional 3 markers were placed on the object. IR LED markers were attached to the copper switch to synchronize motion capture data and the trial start indicator. Models were tested only on data collected with the large object where the initiator of the handover passed the object to the receiver (handover trials where the initiator reached to receive the object from the follower, as well as trials using the small object were not included in this analysis). All handover trials were segmented from movement onset of the initiator to receiver contact with the object. Trials were visually inspected for lapses in marker detection. All analyses and models were developed in MATLAB 2019b (Mathworks, Inc., Natick, MA).

2.2 Gaussian Process Regression

Gaussian process regression is a kernel-based Bayesian approach to regression. The model consists of a probability distribution over all possible functions that fit a set of training points, with a mean function that can be used for regression prediction. The covariance function (kernel) should reflect prior knowledge about the form of function being modeled. In the handover prediction model, we can leverage the representation of velocity profiles as linear

combinations of Gaussian submovements by choosing the widely-used squared-exponential kernel (reference equation). Here we outline the standard Gaussian process regression implemented in the model for human inference. [12]. Given a set of N input-output pairs $\{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^N$, $\mathbf{x}^{(i)} \in \mathcal{X} \subset \mathbb{R}^{n_x}$, $\mathbf{y}^{(i)} \in \mathbb{R}$ related according to an arbitrary model such as

$$\mathbf{y}^{(i)} = \psi(\mathbf{x}^{(i)}) + \eta^{(i)} \quad (1)$$

with $\eta \sim \mathcal{N}(0, \sigma_\eta^2)$, and $\psi \in \mathcal{H}$ considered to be a function of a reproducing kernel Hilbert space $\mathcal{H}$ defined over a compact set $\mathcal{X}$, GPs assume a Gaussian functional distribution as prior for the function $\psi|\mathbf{x}^{(i)} \sim \mathcal{N}(0, \kappa(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}))$, where κ is a kernel function such that $\kappa(\cdot, \cdot) \in \mathcal{H}$. For a set of input points $\mathbf{X} = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}]$ the prior distribution for ψ becomes $\psi|\mathbf{X} \sim \mathcal{N}(0, \mathbf{K})$, where $\mathbf{K} \in \mathbb{R}^{N \times N}$ is the Gram matrix with entries $[\mathbf{K}]_{ij} = \kappa(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$. For a given set of measurements $\mathbf{y} = [\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(N)}]^\top$ associated with the positions $\mathbf{X}$, the prior distribution becomes

$$\mathbf{y} \sim \mathcal{N}(0, \mathbf{K} + \sigma_\eta^2 \mathbf{I}). \quad (2)$$

The predictive distribution allows one to evaluate the function ψ at a new input value $\mathbf{x}$. Thus, we have $\psi|\mathbf{x} \sim \mathcal{N}(0, \kappa(\mathbf{x}, \mathbf{x}))$. Since $\mathbf{y}$ and ψ are jointly Gaussian their joint PDF is given by

$$\begin{bmatrix} \mathbf{y} \\ \psi \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{K} + \sigma_\eta^2 \mathbf{I} & \kappa(\mathbf{X}, \mathbf{x}) \\ \kappa(\mathbf{x}, \mathbf{X}) & \kappa(\mathbf{x}, \mathbf{x}) \end{bmatrix}\right). \quad (3)$$

Finally the predictive distribution can be obtained by conditioning ψ over the observation and its respective positions as

$$\psi|\mathbf{y}, \mathbf{X}, \mathbf{x} \sim \mathcal{N}(\mu(\mathbf{x}), s^2(\mathbf{x})) \quad (4)$$

with

$$\begin{aligned} \mu(\mathbf{x}) &= \kappa(\mathbf{x}, \mathbf{X}) \left(\mathbf{K} + \sigma_\eta^2 \mathbf{I} \right)^{-1} \mathbf{y} \\ &= \kappa(\mathbf{x}, \mathbf{X}) \boldsymbol{\alpha} \end{aligned} \quad (5)$$

$$\begin{aligned} s^2(\mathbf{x}) &= \kappa(\mathbf{x}, \mathbf{x}) - \kappa(\mathbf{x}, \mathbf{X}) \left(\mathbf{K} + \sigma_\eta^2 \mathbf{I} \right)^{-1} \kappa(\mathbf{X}, \mathbf{x}) \\ &= \kappa(\mathbf{x}, \mathbf{x}) - \boldsymbol{\beta}^\top \boldsymbol{\beta} \end{aligned} \quad (6)$$

where $\boldsymbol{\alpha} = \left(\mathbf{K} + \sigma_\eta^2 \mathbf{I} \right)^{-1} \mathbf{y} = (\mathbf{L}^\top)^{-1} \mathbf{L}^{-1} \mathbf{y}$, and $\mathbf{L}$ is the lower-triangular Cholesky decomposition of $\mathbf{K} + \sigma_\eta^2 \mathbf{I}$, and $\boldsymbol{\beta} = \mathbf{L}^{-1} \kappa(\mathbf{X}, \mathbf{x})$.

2.3 Human Inference Model

The Gaussian process regression model for human inference GP takes as input a one-second history of unidimensional velocity data $v_{k-L:k}$ (where L is one second's worth of samples), and outputs a mean velocity function $\hat{v}_{k:k+L}$ which is projected one second into the future for regression predictions.

$$\hat{v}_{k:k+L} = GP(v_{k-L:k}) \quad (7)$$

Input data are cropped at one second for each prediction to improve the algorithm's runtime, with the justification that an accurate prediction can still be obtained while omitting kinematics data from more than one second prior to the time when the prediction is made. Due to the squared-exponential covariance function, the predictions will invariably converge to zero, providing a time of handover, $\hat{t}_{h,k}$, based on the reasonable assumption that handover will occur when the giver velocity approaches zero. Using this point

of convergence (selected as the time after the current sample when velocity drops below a threshold of 0.005m/s), we can obtain a prediction of handover time.

$$\hat{t}_{h,k} = \arg(\hat{v}_{k:k+L} < 0.005) \quad (8)$$

By integrating the mean velocity function from the current time to the point of convergence ($\hat{v}_{k:k+L}$), we obtain a position offset in three dimensions from the initiator's current position obj_k^{xyz} to the locus of handover. This offset gives us our prediction for locus of handover $\hat{p}_k$.

$$\hat{p}_k = obj_k^{xyz} + \int_k^{k+L} \hat{v}_{k:k+L}(t) dt \quad (9)$$

2.4 Trajectory Planning Model

The trajectory planning model takes as input the current receiver hand position and the output from the inference model of the giver's intention, and returns a velocity profile for movement to the predicted locus of handover (arriving at the predicted timing of handover). As the inference model updates throughout the trial, the trajectory planning model waits for the error p_e between the previous predicted locus of handover $\hat{p}_{k-1}$ and the current predicted locus of handover $\hat{p}_k$ to exceed a distance threshold. This triggers the release of a Gaussian-shaped submovement whose amplitude A and standard deviation σ are calculated to account for the new offset in predicted locus of handover. The submovements are combined linearly to form a complete velocity profile. Several free parameters (defined below) were tuned via grid search over a range of possible values, and finding the combination of values that minimized root mean squared error between the model trajectory output and the human receiver's actual trajectory across all trials.

$$p_e = \hat{p}_k - \hat{p}_{k-1} \quad (10)$$

if p_e exceeds the error threshold of 1cm , and more than 50ms have elapsed since the last correction, a new submovement is triggered with the following amplitude:

$$s = 1/\sqrt{2\pi} \quad (11)$$

$$A = \frac{p_e}{(\sigma/s)} \quad (12)$$

σ will be calculated such that the correction will complete at the predicted timing of handover at time k , $\hat{t}_{h,k}$:

$$\sigma = \frac{\hat{t}_{h,k} - \mu}{4} \quad (13)$$

Reaction time r_t is a free parameter defined below. This parameter allows for an urgency factor that results in submovements with higher amplitude and smaller standard deviation when the target handover time approaches. The standard deviation for each submovement is calculated by taking the time difference between the current time and the predicted timing of handover, and dividing this value by 4. A minimum width for standard deviations is set to prevent non-physiological velocity profiles. The center of the submovement is defined as:

$$\mu = r_t + k \quad (14)$$

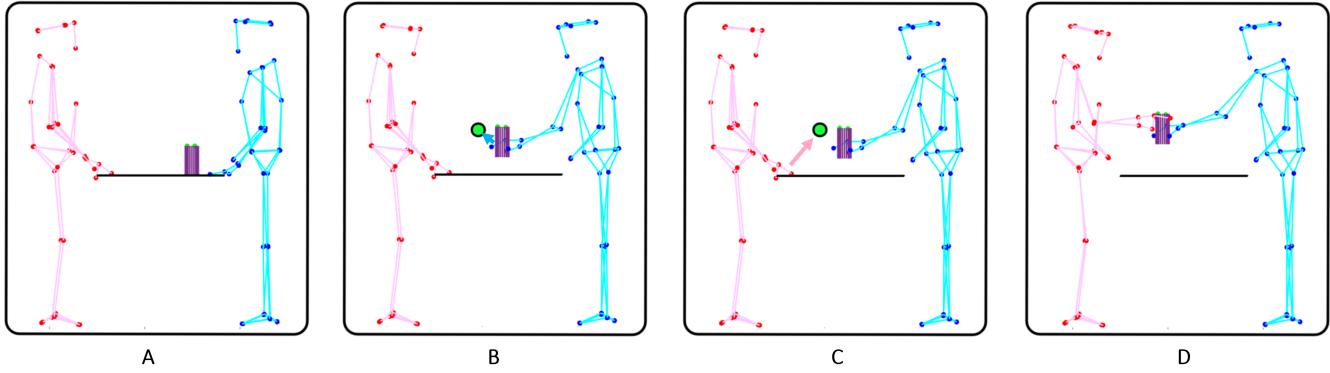


Figure 2: Handover experiment. In A both participants are in their starting positions (giver in blue, receiver in red). In B the giver begins to move towards the locus of handover. The human inference model uses motion capture data from the giver's index finger to output predictions for location and timing of handover (green dot). In C the trajectory planning model uses the output of the human inference model to generate and continuously update a trajectory plan that will bring the receiver's hand to the locus of handover at the predicted timing of handover. In D the receiver follows the trajectory plan to grasp the object, successfully completing the handover.

This prevents an instantaneous correction to a prediction error. Each submovement is defined using the probability density function for a Gaussian distribution:

$$f_x = \frac{1}{\sigma\sqrt{2\pi}} e^{(-0.5((x-\mu)/\sigma)^2)} \quad (15)$$

The complete velocity plan for the receiver is obtained by summing all individual submovements:

$$f_{x_{Total}} = \sum_{i=1}^n A_i \times f_{x_i} \quad (16)$$

- **Reaction time (283.3 ms):** the time between when the error threshold for prediction is exceeded, and the center of the submovement correcting for that error.
- **Correction latency (192.3 ms):** the delay following movement onset before the primary movement is overridden. Even if an error exceeding the error threshold is observed, a corrective submovement is not released until this delay is completed.
- **Distance threshold (1 cm):** the minimum distance between a current and previous prediction to trigger a corrective submovement.
- **Time threshold (50 ms):** the minimum time between successive submovement corrections.
- **Primary movement delay:** Primary movement delay μ_{prim} is calculated separately for X, Y, and Z components of the velocity. Three values: μ_{x_factor} , μ_{y_factor} , and μ_{z_factor} are multiplied by the standard deviation of the primary movement to determine this delay for each component.

$$\mu_{prim_x} = \mu_{x_factor} \times \sigma_{prim_x} \quad (17)$$

$$\mu_{prim_y} = \mu_{y_factor} \times \sigma_{prim_y} \quad (18)$$

$$\mu_{prim_z} = \mu_{z_factor} \times \sigma_{prim_z} \quad (19)$$

- **Minimum st_dev (300 ms):** minimum standard deviation for a submovement.

3 RESULTS

The performance of handover inference and trajectory planning models was assessed separately by comparing model outcomes to empirical data derived from human-human experiments.

3.1 Handover Location and Timing Prediction Results

Model outputs were analyzed in aggregate for location and timing prediction accuracy 1000ms, 800ms, 600ms, 400ms, 200ms and 0ms prior to receiver contact. The results for location prediction (see Fig. 4) show a marked reduction in mean location prediction error 600ms prior to handover, with the error consistently less than 3cm at this time point. Mean location errors further reduce to 1-2cm 300ms prior to handover. Similar results of low errors in location prediction 600ms prior to handover are shown for a single trial, see Fig. 3. Handover timing predictions showed initial improvement in prediction quality until around 200ms prior to handover before resulting in a mean overestimation of handover time. The prediction of timing suffers partially from small adjustments made by the giver immediately prior to receiver contact. These small movements can cause an overestimation of handover time due to their tendency to produce submovements with small amplitudes and large standard deviations. Given the model definition of handover time which searches for future convergence of the velocity prediction to zero, and the additive property of these submovements, final adjustments will often result in increasing overestimation of handover time.

3.2 Trajectory Planning Results

To assess the performance of the trajectory planning model, the model was run on the same data session used to test the human inference model. The location and timing prediction outputs from the human inference model were fed to the trajectory planning model, simulating a handover experiment. Using predictions of handover location and timing, the trajectory planning model returned a velocity profile for the receiver, and continuously revised the profile to accommodate for updates in these predictions. Root Mean Square Error (RMSE) was used to evaluate the similarity between the X, Y, and Z components of the final trajectory plan (integrated

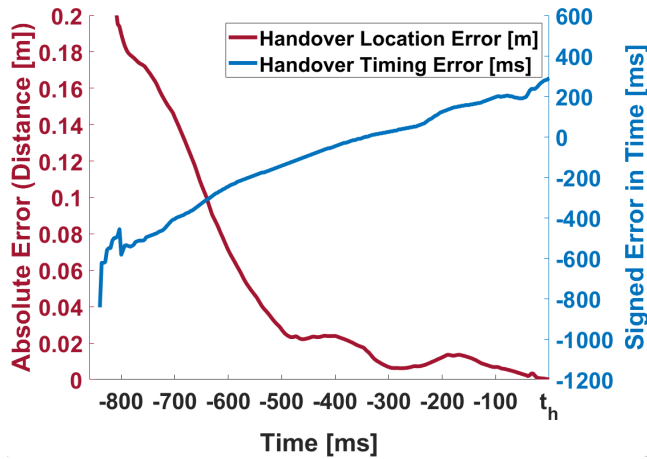


Figure 3: Visualization of the accuracy in model predictions of handover location (left axis) and timing (right axis). All traces are cropped from movement onset to receiver contact with the object. Handover location accuracy improves to within 2cm of the true locus of handover 350ms prior to handover. The handover timing prediction improves until 300ms prior to handover, before overestimating the true handover time. This late drop in accuracy is likely due to the small adjustments in the giver's trajectory immediately prior to receiver contact.

velocity), and the ground truth trajectory of the human receiver for that trial. A grid search was performed over a range of possible values to find those values that minimized RMSE across trials from all four handover locations. These values were as follows: Correction Latency (192.3ms), Reaction Time (283.3ms), μ_{x_factor} (3.09), μ_{y_factor} (3.42), and μ_{z_factor} (2.48) (the values for μ_{x_factor} , μ_{y_factor} , and μ_{z_factor} control the delay of the primary movement). Larger factors for a particular component (X, Y, or Z) resulted in a greater delay for that component of the primary movement.

The trajectory planning model was then run using optimal parameters. For comparison, variability of human receiver movements (per handover location and component position) were calculated as the RMSE between ground truth receiver trajectories and the mean receiver trajectory for each position. The RMSE between the model output and human receiver trajectories was comparable to that observed between those human receiver trajectories and the mean human receiver trajectory for all locations except bottom left, see Fig. 7. The bottom left location showed a larger RMSE between model and human trajectory than between human trajectory and mean human trajectory for X and Z components of the movement. These differences are also evident in the position traces for model outputs and receiver trajectories, see Fig. 5, where the model predicted earlier movement in the X and Z directions than what was observed in the corresponding human receiver trajectories. Representative traces for each location were plotted in three dimensional space for additional visualization of model output and human receiver trajectories, see Fig. 6.

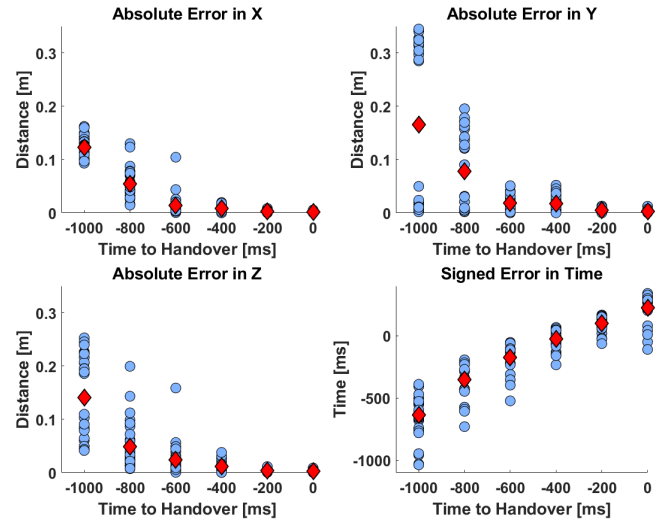


Figure 4: Prediction accuracy prior to handover across 32 handover trials. The mean location prediction improves to within 3cm of the target location as early as 600ms prior to handover. The variation in prediction quality across trials reduces significantly for predictions made later in the trial. Prediction of handover time improves until roughly 200ms prior to handover, after which the model overestimates handover time. This may be due to small adjustments made by the giver immediately prior to receiver contact which do not significantly change handover location, however, the submovements produced by small adjustments can lead to an overestimation of handover time.

4 DISCUSSION

The proposed models of human inference and trajectory planning showed promising similarity to human-to-human handover. The model of human inference could predict the location of handover within 2–3 cm, 600 ms prior to handover. The model also provided a timing prediction which, although overestimating the true handover time by an average of 200 ms, could still be used successfully in trajectory planning. The trajectory planning model produced smooth, sigmoidal position profiles characteristic of human motion for all conditions that were qualitatively similar to the human receiver. The RMSE between actual receiver trajectories and model trajectories fell within the variability observed for the human receiver, except for the bottom left handover location. The dependence of model performance on the task constraints of handover could be explored in future research. A potential improvement addressing the larger error observed for the bottom left handover location could be to delay the release of corrective submovements based on error in all three dimensions (X, Y, and Z). This condition could prevent the early release of corrections in the Y and Z dimensions, which appear to be contributing to these errors.

Our submovement based trajectory planner has potential advantages over existing trajectory planners designed to emulate a human-like reaching. During handover with a human giver, the object position is constantly changing, requiring frequent updating of the receiver trajectory. Models which require training on human

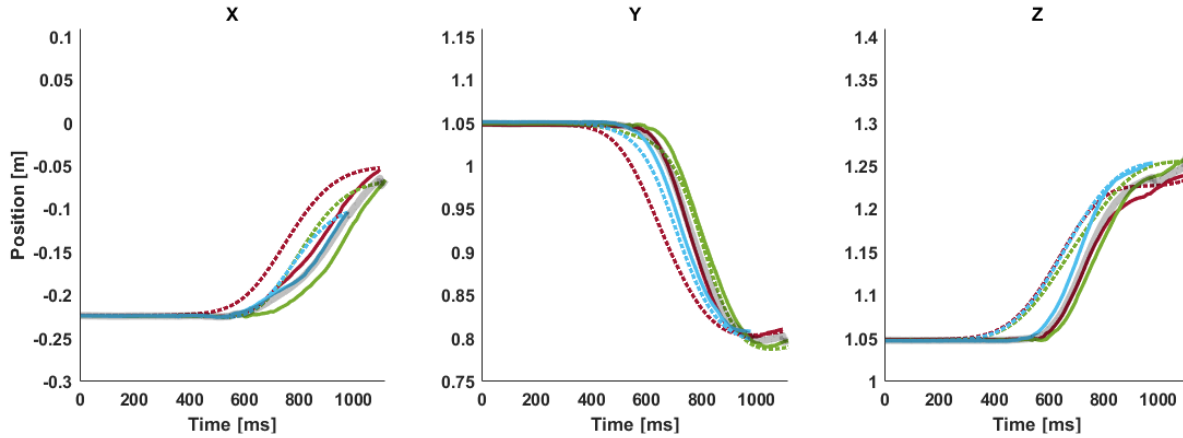


Figure 5: 2D visualization in time of the outputs of the receiver trajectory planning model. Each model output (dashed line) is paired by color with the true receiver trajectory for the same trial (solid line), and broken down into X,Y, and Z components. Three trials for the bottom left handover condition are depicted (red, green, and blue traces), along with the mean receiver trajectory (gray).

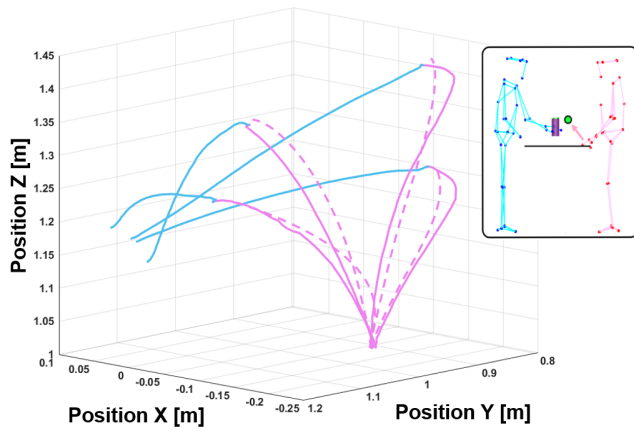


Figure 6: 3D visualization of the trajectory planning model output. One trial is selected for each of the four handover locations (top right, top left, bottom right, and bottom left). Motion capture traces of the object are depicted with solid blue lines. Ground truth receiver trajectories for those four trials are depicted with solid pink lines while model outputs are depicted as dashed pink lines.

examples may perform poorly or fail to maintain human-like motion when presented with giver movement patterns not included in the training set [10, 14]. Our submovement based approach both enables real-time adaptation to changes in object prediction and retains smoothness in the complete trajectory plan. Therefore, our additive Gaussian approach offers a generalizable model for trajectory planning to simulate the submovements we observe in humans as they adapt to changes in their internal predictions of handover location and timing.

In summary, the proposed models for human inference and trajectory planning offer promising performance for human-robot handover while remaining grounded in a physiologically meaningful feature of human motion: Gaussian submovements in velocity

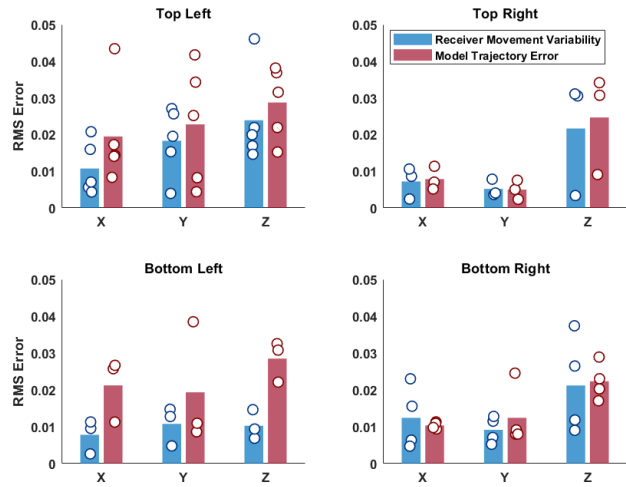


Figure 7: Results of root mean square error (RMSE) analysis on the trajectory planning model. RMSE was calculated between all individual receiver trajectories and the mean receiver trajectory (blue bars), as well as between each model trajectory and its corresponding receiver trajectory (red bars) across four handover target locations. Error between the model output and true receiver trajectories was comparable to the inter-trial variation observed across all receiver trajectories, with the exception of the X and Z components for the bottom left condition. Human receivers exhibited a delayed primary movement in the X and Z directions for this condition as compared to the trajectories predicted by the model.

profiles. Our study had several limitations. Data used to assess the model in this pilot study came from a single human dyad performing handover. We have plans to test performance of our models on a larger dataset consisting of multiple pairings of the same participant to explore the model generalizability to individual behavior.

The models described here utilize only index finger motion for inference and trajectory planning. Motion capture data used here was extracted from a larger dataset consisting of kinematic and gaze data collected on human-to-human object handovers using RGB-D cameras, 3D motion capture, inertial measurement units, force-sensitive finger sensors, and wearable eye tracking headsets. Data fusion from multiple sensors has the potential to increase model performance and robustness to intermittent loss of a single sensing modality as can be expected in real world environments.

Finally, in this pilot study, we did not compare our model output with those from previous models. Several directions for enhancements to inference and trajectory models can be pursued. One such direction is the use of a prior-based model in unison with the online model for human inference. This approach could provide highly accurate initial predictions of handover location and timing while retaining the ability to correct for unexpected changes in position of the object. We also plan to explore the effects of the location and timing error thresholds used to trigger the release of corrective submovements in the trajectory planning model and how these values correspond to humans' internal representation of prediction error [13]. These parameters likely have a profound effect on the shapes of the trajectories produced. We plan to incorporate these models, with improvements, into empirical testing of human-robot handover to investigate whether the proposed inference and trajectory planning models provide more fluid handover and greater user comfort when interacting with robots.

ACKNOWLEDGMENTS

This work was supported in part by NSF-CMMI-M3X-1935337 (DE, TP, ET, MY), NSF-CBET-1804540 (ET), NSF-CBET-1804550 (ET), NIH-2R01NS085122 (ET), and NIH-2R01HD058301 (ET).

REFERENCES

- [1] Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer, Kazuhiro Kosuge, and Oussama Khatib. 2018. Progress and prospects of the human-robot collaboration. *Autonomous Robots* 42, 5 (2018), 957–975.
- [2] Marco Costanzo, Giuseppe De Maria, and Ciro Natale. 2021. Handover Control for Human-Robot and Robot-Robot Collaboration. *Frontiers in Robotics and AI* 8 (2021), 132.
- [3] Auke Jan Ijspeert, Jun Nakanishi, and Stefan Schaal. 2002. Movement imitation with nonlinear dynamical systems in humanoid robots. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, Vol. 2. IEEE, 1398–1403.
- [4] Hermano Igo Krebs, Mindy L. Aisen, Bruce T. Volpe, and Neville Hogan. 1999. Quantization of continuous arm movements in humans with brain injury. *Proceedings of the National Academy of Sciences* 96, 8 (1999), 4645–4649. <https://doi.org/10.1073/pnas.96.8.4645> arXiv:<https://www.pnas.org/content/96/8/4645.full.pdf>
- [5] Muriel Lang, Satoshi Endo, Oliver Dunkley, and Sandra Hirche. 2017. Object handover prediction using gaussian processes clustered with trajectory classification. *arXiv preprint arXiv:1707.02745* (2017).
- [6] Alberto Leardini, Fabio Biagi, Andrea Merlo, Claudio Belvedere, and Maria Grazia Benedetti. 2011. Multi-segment trunk kinematics during locomotion and elementary exercises. *Clinical Biomechanics* 26, 6 (2011), 562–571.
- [7] Ester Martinez-Martin and Angel P del Pobil. 2018. Personal robot assistants for elderly care: an overview. *Personal assistants: Emerging computational technologies* (2018), 77–91.
- [8] Jan Peters and Stefan Schaal. 2008. Reinforcement learning of motor skills with policy gradients. *Neural networks* 21, 4 (2008), 682–697.
- [9] Stefan Schaal. 2006. *Dynamic Movement Primitives - A Framework for Motor Control in Humans and Humanoid Robotics*. Springer Tokyo, Tokyo, 261–280. https://doi.org/10.1007/4-431-31381-8_23
- [10] Kyle Strabala, Min Kyung Lee, Anca Dragan, Jodi Forlizzi, Siddhartha S Srinivasa, Maya Cakmak, and Vincenzo Micelli. 2013. Toward seamless human-robot handovers. *Journal of Human-Robot Interaction* 2, 1 (2013), 112–132.
- [11] Holger Urbanek, A Albu-Schaffer, and Patrick van der Smagt. 2004. Learning from demonstration: repetitive movements for autonomous service robotics. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, Vol. 4. IEEE, 3495–3500.
- [12] C. K. I. Williams and C. E. Rasmussen. 2006. *Gaussian processes for machine learning*. Vol. 2. MIT press Cambridge, MA.
- [13] Daniel M Wolpert, Jörn Diedrichsen, and J Randall Flanagan. 2011. Principles of sensorimotor learning. *Nature Reviews Neuroscience* 12, 12 (2011), 739–751.
- [14] Wei Yang, Chris Paxton, Maya Cakmak, and Dieter Fox. 2020. Human grasp classification for reactive human-to-robot handovers. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 11123–11130.