

RESEARCH ARTICLE

Open Access



Stable haptic feedback generation for mid-air gesture interactions: a hidden Markov model-based motion synthesis approach

Dennis Babu^{1*} , Masashi Konyo¹, Hikaru Nagano¹, Ryunosuke Hamada^{1,2} and Satoshi Tadokoro¹

Abstract

Generation of stable and realistic haptic feedback during mid-air gesture interactions have recently garnered significant research interest. However, the limitations of the sensing technologies such as unstable tracking, range limitations, nonuniform sampling duration, self occlusions, and motion recognition faults significantly distort motion based haptic feedback to a large extent. In this paper, we propose and implement a hidden Markov model (HMM)-based motion synthesis method to generate stable concurrent and terminal vibrotactile feedback. The system tracks human gestures during interaction and recreates smooth, synchronized motion data from detected HMM states. Four gestures—tapping, three-fingered zooming, vertical dragging, and horizontal dragging—were used in the study to evaluate the performance of the motion synthesis methodology. The reference motion curves and corresponding primitive motion elements to be synthesized for each gesture were obtained from multiple subjects at different interaction speeds by using a stable motion tracking sensor. Both objective and subjective evaluations were conducted to evaluate the performance of the motion synthesis model in controlling both concurrent and terminal vibrotactile feedback. Objective evaluation shows that synthesized motion data had a high correlation for shape and end-timings with the reference motion data compared to measured and moving average filtered data. The mean R^2 values for synthesized motion data was always greater than 0.7 even under unstable tracking conditions. The experimental results of subjective evaluation from nine subjects showed significant improvement in perceived synchronization of vibrotactile feedback based on synthesized motion.

Keywords: Vibrotactile feedback, Motion synthesis, Stable haptic feedback, Hidden Markov model, Occlusion, Mid-air interaction, Unstable tracking

Introduction

The advent of affordable, small-sized motion tracking sensors such as Leap Motion¹ and Kinect², have made mid-air interactions viable in new application areas such as desktop computers, interactive tabletops, and inside cars. Portable virtual and augmented reality interfaces have also recently accelerated the use of mid-air interactions as a human–computer interaction technique.

Haptic feedback plays an important role in mid-air gesture interactions to give information about the physical presence of objects, which the users are interacting. Previous research has shown that delivering appropriate haptic cues to users during gesture input can help in gesture training [1] and task performance [2], as well as improve overall user experience [3].

*Correspondence: dennis@rm.is.tohoku.ac.jp

¹ Graduate School of Information Sciences, Tohoku University, Aramaki-aza Aoba, Sendai 980-0871, Japan

Full list of author information is available at the end of the article

¹ <https://www.leapmotion.com/>.

² <https://www.xbox.com/en-US/xbox-one/accessories/kinect/>.

One of the main limitations affecting gesture-based haptic feedback generation is noisy and volatile motion data during mid-air interactions. The occlusion of the tracked motion and range limitations of motion tracking sensors deteriorate the haptic feedback based on mid-air fingertip motion [4, 5]. Conventional filtering approaches may not be able to provide stable motion data generation because human gestures and interactions are highly arbitrary and do not have specific frequency distinctions with anomalies. In this paper, we intend to propose a new method for generating stable and realistic haptic feedback even during unstable tracking conditions.

Two primary requirements for concurrent vibrotactile rendering [6] in mid-air interactions are the stability and real-time motion data control over the vibrotactile signals. Continuous motion data such as fingertip position, velocity, and acceleration data have been used as the control elements for physically-based vibrotactile rendering models [7, 8]. Thus, in the context of real-time vibrotactile rendering, the mere recognition of gestures from erroneous motion patterns is insufficient. The recognized gestures may further be utilized to replicate stable and real-time motion data for controlling haptic feedback stimulation even “during” the motion.

In this paper, a method for synthetic motion element synthesis for haptic rendering using hidden Markov models (HMM) is proposed. The proposed method was inspired by the embodied motion pattern generation for robots as detailed in [9, 10]. In the aforementioned work, the authors generated self-motion elements from recognized motion patterns in robots to replicate human motion patterns in robots. Here we use mimesis theory to recreate stable, real-time motion patterns for different gestures. The unstable motion patterns are fed to an HMM-based gesture recognition algorithm which recognizes the hidden states corresponding to an identified gesture. Primitive motion elements associated with each state are synthesized to recreate the ideal motion paths associated with each gesture. An algorithm for adaptive modulation of primitive motion element compared with changes in the real-time execution speed by users is also proposed. An objective analysis of the comparative performance of the synthesized motion data with the stable motion data obtained from a reference sensor is conducted to estimate the viability of the proposed model. Further, a subjective evaluation of vibrotactile feedback based on the proposed model was conducted to confirm the performance of the proposed methodology.

The main contributions of this work are as follows.

- A method for real-time haptic feedback generation during mid-air interactions using motion synthesis even in the presence of unstable motion tracking is proposed.
- Estimation of the reference motion pattern of a gesture and definition of primitive motion elements to be synthesized in real-time.
- Validation of the proposed method for four gestures namely—tapping, three-fingered zooming, vertical dragging and horizontal dragging.

Initial work in this regard was implemented in [11] wherein the proposed method was verified for tasks involving zoom. In this work, we improve the method and confirm its suitability for multiple gestures involving more participants.

Related works

Significant research has been conducted in both motion recognition and mid-air haptic feedback generation techniques in recent past. However, to the best of our knowledge no research effort focused on the use of unstable motion data to generate stable haptic feedback in mid-air interactions have been conducted. Thus, we broadly classify the related works into two categories: Motion recognition in mid-air interactions and haptic feedback technologies for mid-air interactions.

Motion recognition

Multi-camera optical tracking systems and inertial measurement units (IMU) have been used in the past to track the position and orientation of the human body during mid-air interactions. Optical tracking systems with a constellation of cameras capable of tracking multiple markers attached to the human body [12] are used for accurate modeling and data acquisitions. The main limitations of such systems are their high cost and immobility of the test bed, thus making it an experimental reference test. Won et al. [13] presented a novel methodology based on one position sensor and an IMU to estimate position and orientation with the integration of filter tools. Though this method could obtain a relatively accurate position estimation, it needs extra assistance from several markers to get the 3D reference world coordinates of the tracking point.

Although depth cameras such as leap motion controllers (LMC) and Kinect have been used extensively in mid-air interfaces recently, it has many limitations. As the current study uses LMC as the tracking sensor, a detailed review of the limitations of the commercial depth cameras was conducted. Kinect, which has two infrared cameras for depth detection and one standard visual-spectrum camera for image recognition, can

obtain the depth information and color images of the operator which are used to calculate the position and the orientation. The LMC also uses similar technology that is similar to Kinect but offers precise tracking of fingers using a built-in hand model. Breakthroughs were also made in predicting a self-occluded hand, e.g., [14]. Until recently, this was a severe obstacle for optical tracking devices. Unstable motion patterns are still prevalent in the case of multiple hand interactions with incorrect initialization of the human postures. Moreover, as reported in [4] LMCs have uneven sampling frequency which causes discontinuous motion capture data along the time frame with increasing drift as capture duration increases. To complicate the matter further, it was found that the instability in motion capture is a function of distance and field of view from that of the LMC [4]. Another limitation of LMC is the unstable orientation information collection procedure. This leads to an incorrect position estimation as position tracking inherently depends on orientation information.

Many researchers have used advanced filtering technologies and learning based methods to recognize hand motions using a single depth camera. One of the straightforward approaches for haptic rendering from measured data is to apply a low pass [15] or bandpass filter [16] on the measured data for haptic feedback rendering. The proposed approach works well for a master–slave system of teleoperation and measured data based haptic feedback systems. However, the methodology may fail to have the desired effect in mid-air interactions as there are no definite frequency component differences between unstable and stable motion patterns. In [17] the authors used a Kalman filter to generate stable motion data from LMC to control robot motion. This methodology needs a precise mathematical model for each gesture, thereby requiring a heuristic approach for multiple gestures. Keskin et al. [18] proposed a discriminative method using a multi-layered random-forest to predict hand parts and thereby to fit a simple skeleton. The system runs at 30 Hz on consumer CPU hardware but failed under occlusion. Xu et al. [19] estimated the global orientation and location of the hand, and selected the correct posture by minimizing reconstruction error. The system ran at 12 Hz, and the lack of tracking occasionally led to noisy pose estimates occasionally.

Haptic feedback technologies

In the past, many haptic feedback methods with different form factors have been proposed in the past for mid-air virtual and augmented reality interactions. Previous works includes the finger-worn gloves with tactile feedback applied at the fingertip [20, 21] and the

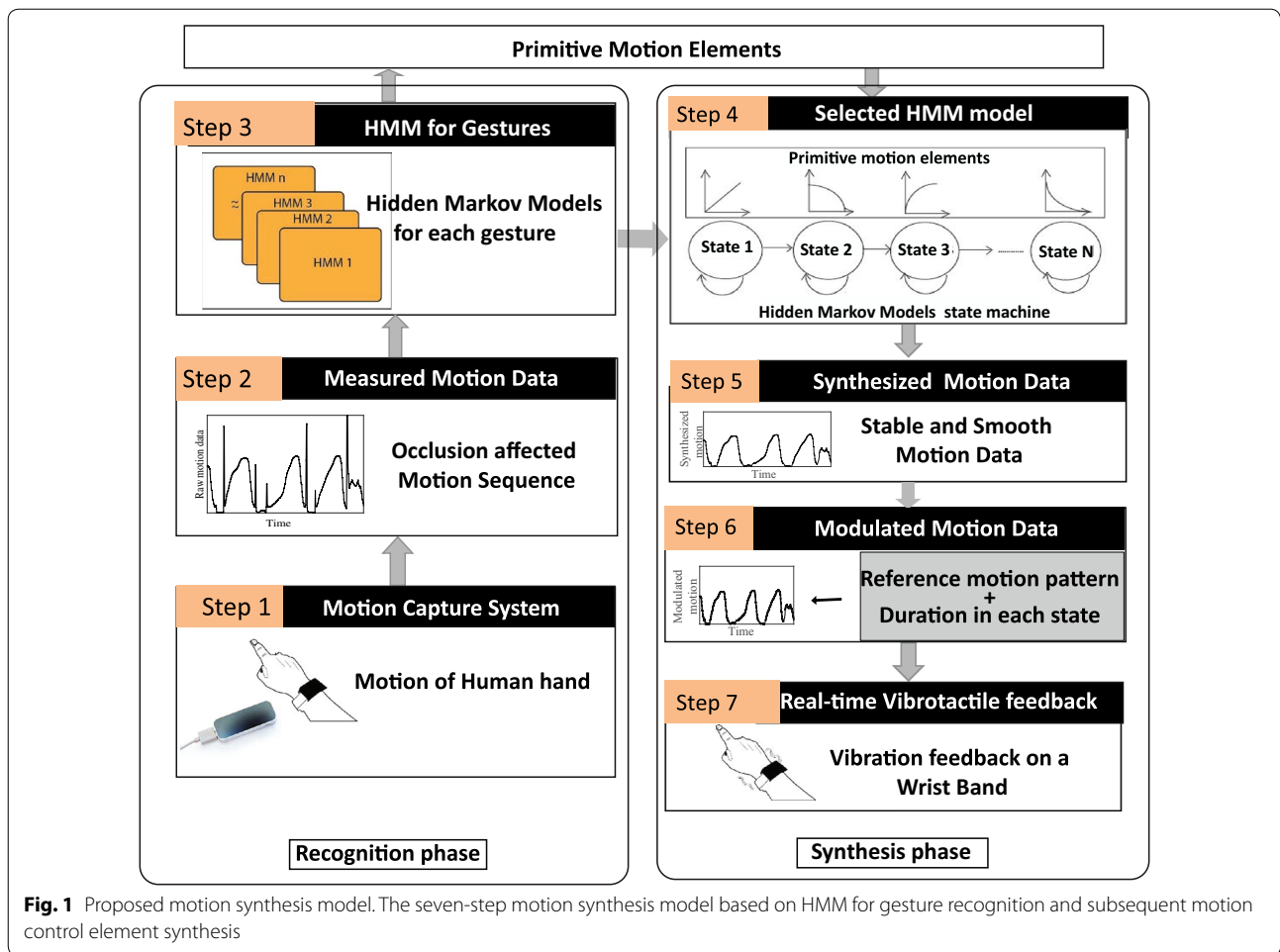
real-time auditory and vibrotactile feedback [20]. In [22], the authors presented a wireless haptic ring (HapRing) for spatial interaction. This device provided vibrotactile signals and vibration cues on a finger base using a haptic actuator. Novel techniques in the field, use ultrasonic transducers [23] and air vortex rings [24] to create focused haptic feedback in mid-air. Most recently, Lee et al. [25] reported the feasibility of using indirect laser radiations for mid-air tactile rendering.

While most of the above mid-air haptic feedback technologies uses depth motion sensors for motion sensing, the effects of specific research efforts were made to address unstable motion patterns on affecting haptic feedback have not been addressed. As is evident from the literature review, there is a gap to be addressed in linking the unstable motion patterns and real-time haptic rendering. Therefore, this work will discuss the proposed motion synthesis approach to address this critical research problem.

Motion synthesis model and primitive motion element synthesis using HMM

The proposed method uses a hidden Markov model to recognize the motion states and further synthesize corresponding motion profile, thus recreating actual motion pattern. Figure 1 shows the general outline of the proposed approach. Motion synthesis-based haptic feedback generation have two phases, namely motion recognition phase and motion synthesis phase. The first three steps involve motion recognition phase, where motion data is acquired from tracked finger joints during gesture execution. HMM with multiple hidden states are selected in real-time from these observed motion data. Each gesture corresponds to a unique HMM with a unique number of states. Further primitive motion elements corresponding to each state are estimated based on reference motion data obtained from a stable motion sensor. These primitive motion elements are stored in a lookup table and synthesized according to the recognized state as shown in step 4. With changes in execution speed of the gesture, the primitive motion elements are adaptively modulated by estimating the duration of HMM states. The synthesized motion data is then used as the control function to modulate the vibrotactile waveform as shown in steps 5 and 6. Finally, the vibrotactile feedback is finally fed to the users via a wrist band as shown in step 7.

We adopted discrete hidden Markov models (DHMMs) [26] to describe the relationship between the sequence of motion patterns and primitive symbols. DHMM is a stochastic process that generates time series data. Each gesture is modelled by a unique DHMM and the probability that a specific time series vector is generated by a DHMM can be calculated by recursive maximum



likelihood functions. Here finger position and velocity vectors obtained from depth camera sensor are treated as the time series data. For tapping and dragging gestures, the index finger distal phalange tip position was used as the motion data for HMMs. Both these gestures used a pointing gesture to start the gesture execution at the intended locations. For the zoom gesture, the radius from an imaginary circle connecting three finger joints namely, thumb distal phalange tip, index finger distal phalange tip and middle finger distal phalange tip positions were used as the motion data for HMM. These measured motion data vectors are normalised and discretised in real-time to obtain output symbols, $o_j(t)$ since DHMM generates discrete symbols at each time instant. An Expectation–Maximisation algorithm is then used to train unique DHMMs for each gestures, to obtain the following three parameters unique for each gestures.

- $A = a_{ij}$ is a states transition matrix. Here a_{ij} indicates a probability of transition from state q_i to state q_j .

- $B = b_{ij}$ is an output probability matrix. b_{ij} indicates a probability of output symbol o_j from state q_i .
- π is a vector of probability vector describing the distribution of initial states.

Since the assumed HMM is a left-right model we initialise π as a constant value for each gesture. The number of hidden states in HMM for all the gestures were selected as four to optimize the gesture recognition accuracy. For training the HMM in each gesture, motion data from multiple subjects with different interaction speeds and positions were used to ensure high recognition rate. The trained HMM had an accuracy of 94.4%, 94.6%, 95.1% and 81.778% for horizontal drag, vertical drag, zoom and tapping gestures respectively. The model was tested by multiple participants, executing the gesture with different speeds and tracking stability conditions.

Primitive motion elements

The HMM-based motion synthesis model uses smooth primitive motion patterns stored in the lookup tables to

generate motion elements for haptic feedback control. The primitive motion elements are defined by a polynomial curve fit for each state of the HMM models. Therefore, we define the proto symbol P 's as follows.

$$P_i = (A_i, B_i) \quad (1)$$

The Viterbi algorithm calculates an ideal path of $P(O|(A, B))$ over a given time frame T of output motion sequences, O and selects the most suitable HMM state sequence in real time. It further renders the corresponding primitive elements associated with each recognized state, after adjusting to the speed of gesture execution. Thus the determination of reference motion patterns for each gesture and the definition of primitive motion elements to be synthesized by the HMM states are detailed in the next section.

Determination of reference gesture motion patterns and primitive motion elements

Participants

The reference motion patterns are obtained from five subjects (three males and two females) with mean age of 22.88 and SD: 2.56. None of them had prior experience using a haptic interface and were naive towards the goal of the study.

Gestures and reference motion elements

We analyzed four distinct dynamic gestures for motion pattern synthesis using the proposed approach. The gestures were (1) horizontal drag, (2) vertical drag and (3) tapping using the index finger and (4) three-fingered zooming. Figure 2 shows the illustrative images of different gestures and the corresponding finger positions tracked. The above gestures were selected to analyze and validate the efficiency of motion synthesis to provide continuous and event-based haptic feedback triggering. Each subject repeated different gestures thirty times at different interaction speeds. The reference motion patterns to be synthesized from each gesture for controlling haptic feedback (shown in Fig. 2 pictorially) are detailed here.

For tapping and dragging gestures, the finger joint position of distal phalange tip was used. Both these gestures use a pointing gesture to start the gesture execution. The tapping distance (collision with the virtual surface) is defined as the reference motion element (shown by $tapp_D$ in Fig. 2a) to control the haptic feedback. In the vertical and horizontal dragging gestures, the distance of dragging tasks (shown by $hdrag_D$ and $vdrag_D$ respectively in Fig. 2c, d) is defined as the motion element to control haptic feedback.

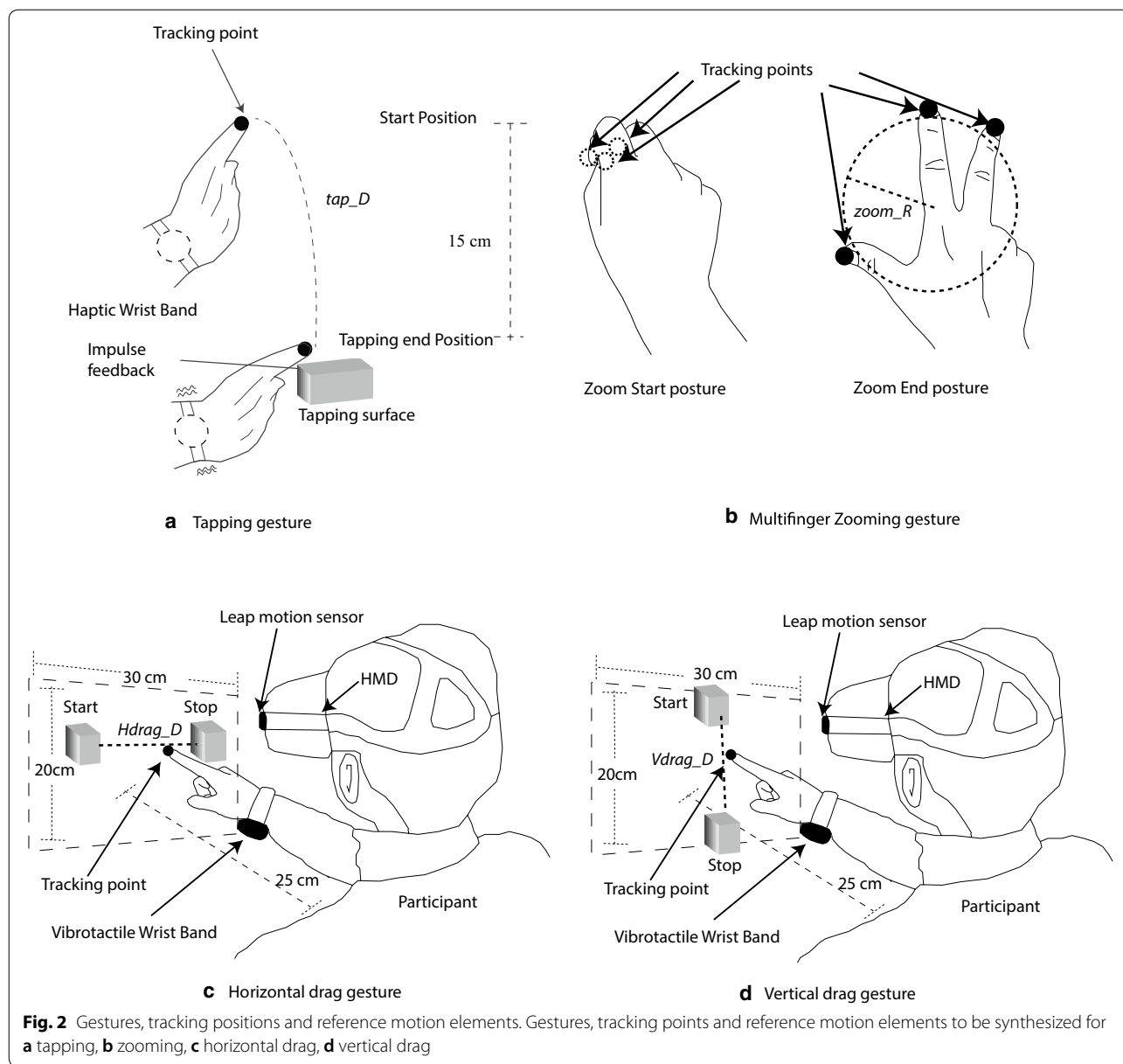
In the zoom gesture, the three finger joints namely, thumb distal phalange tip, index finger distal phalange tip and middle finger distal phalange tip positions are used for obtaining the radius of an imaginary circle connecting three joints (shown as $zoom_R$ in Fig. 2b). This virtual zoom radius is used as the motion element to control the haptic feedback. The z-axis data of each fingertip positions are neglected to obtain a planar circle approximation. Standard geometric equations shown in [27] we used in the calculation of zoom radius in real-time.

Experimental setup

We use a stable non-line-of-sight motion tracking system (Polhemus Liberty, USA) to obtain the reference gesture motion data. A fixed sampling interval of 60 Hz is used to obtain steady motion data. The subjects wore a head-mounted display (HMD) used in the case of VR/AR setups which displayed visual feedback of gesture execution. A unity application displayed the visual feedback while performing the task with actual fingertip position was rendered to the subjects to make the task easy. Care was taken to ensure the position of the subjects to be same every time to avoid improper tracking. Reference motion tracking is performed on a high definition PC with specifications of 32 GB RAM, 4 MHz, I7 processor and GE Force M470 graphics card.

Results: Reference motion elements and defining primitive motion elements

Motion patterns for each gesture obtained from multiple subjects are normalized in both time and amplitude to obtain a generic fitted curve. Figure 3a–d shows the normalized motion data obtained for horizontal drag, vertical drag, tapping, and zooming gestures respectively. The unnecessary motion patterns obtained during the experimentation are filtered out using a linear regression analysis where the R^2 value of the mean motion element is correlated with each motion element. Here the mean normalized curve of entire data set of motion curves for each gesture is calculated initially. Then the correlation between each curve in the data set and this mean curve is estimated using R^2 values. Here we treat the mean curve as the regressed curve and each curve in the data set as data points to be regressed or fitted to this mean curve. Thus, a perfect fitting will generate a correlation score of 1, while an imperfect fitting results in 0. All the motion patterns with $R^2 < T_h$ are treated as outliers and are excluded from the analysis. This is indicated by the unshaded region in Fig. 3. The threshold T_h is obtained by trial and error method for each gesture separately.



The gestures are then divided into multiple states where each state has a primitive motion element given by the mean of the motion curve at each state. These primitive elements are synthesized from polynomial curves with the best fit in each state. The standard deviations due to variations between subjects at each state are also included in the primitive elements as shown in Fig. 3. Each of these states is recognized by the HMM detailed in the previous section and the corresponding primitive motion element is rendered in real time.

Evaluation method

Participants

Nine volunteers (two female and seven male, mean age 22.88 and SD: 1.19) participated in the study. None of the volunteers reported any visual or tactile deficits. All of them were naive about the goal of the study and signed written consent forms approved by the University Ethics Committee.

Experimental setup

Figure 4a shows the experimental setup used to evaluate the proposed motion synthesis model. We used a Leap

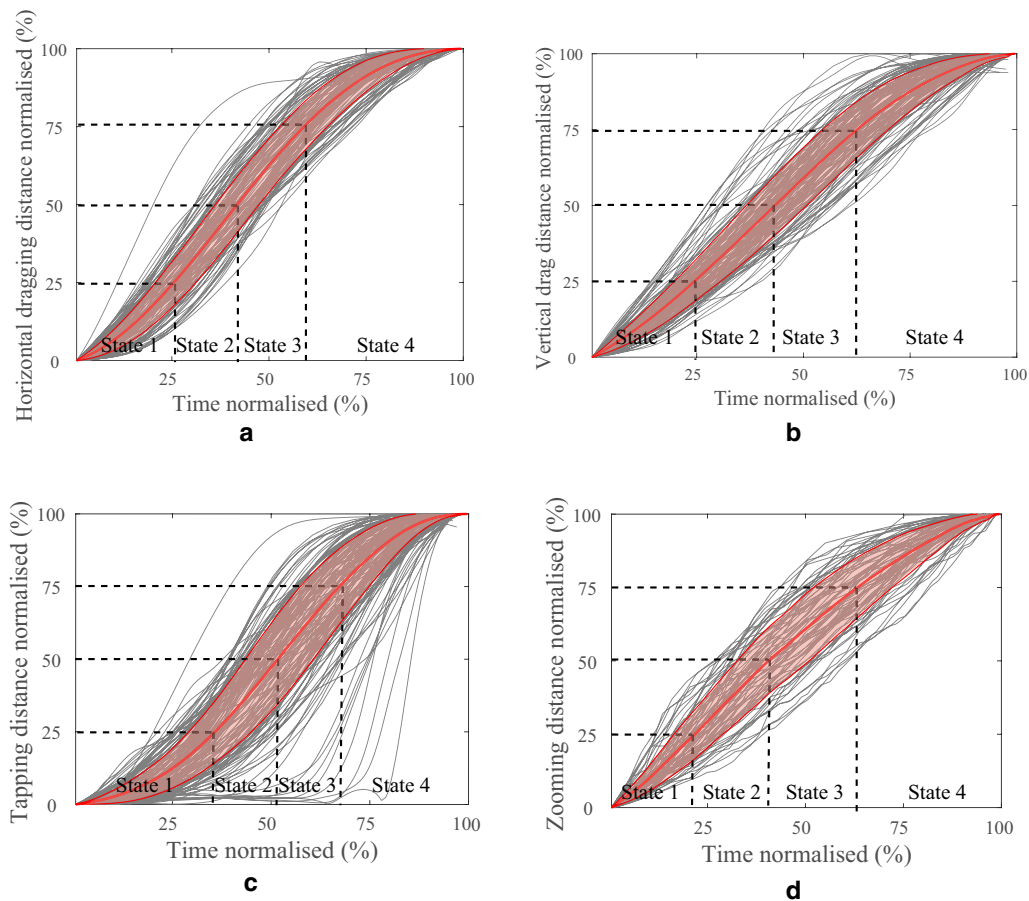


Fig. 3 Ideal motion pattern and primitive motion element definition. Definition of primitive motion element from reference motion pattern for different gestures. **a** Horizontal drag, **b** vertical drag, **c** tapping, **d** zooming. Here the reference motion pattern is obtained from different subjects and the outlier motion elements are removed based on R^2 value criteria to get the mean and standard deviation as shown by the shaded region in the figure. Finally, a curve fitting tool is used to define primitive motion elements, mathematically in each state

Motion sensor to track hand motion in real-time and render haptic feedback to the subjects during gesture interactions. We used a non-line-of-sight electromagnetic motion tracking system (Polhemus Liberty, USA) as the reference motion tracking system to ensure stable and accurate tracking during active gesture interaction. The system had a maximum update rate of 240 fps and delivers 6 DoF motion data with less than 4 ms latency. Stable motion data obtained from the reference motion tracking system was used to render visual feedback of fingertip positions to the subjects using a HMD (Oculus Rift, DK2). This ensured stable visual feedback irrespective of the gesture execution positions.

We use a voice coil actuator (VP2, Acouve, Japan) as the haptic feedback device. The voice actuator was placed in a 3D printed circular box of diameter 50 mm and height 15 mm and attached to a user's wrist as shown in Fig. 4a. Wrist-based haptic feedback was selected for

its commercial viability in current mid-air gesture interaction systems. We used pseudo-haptic vibrotactile representations for haptic rendering with a processing loop of 1 kHz. These are described in detail in the following section.

Task, experimental conditions, and vibrotactile feedback

We analyzed the same four gestures (horizontal drag, vertical drag, tapp and zoom) described earlier for evaluations. The 41-point moving average filtered motion data having a theoretical delay of 265 ms was also compared with the measured and synthesized motion data to assess its viability. The filtered motion data was used to compare the horizontal drag and zoom gestures as both the gestures had concurrent vibrotactile feedback.

For each of the four gestures, the subjects executed the gesture in two positions. One position has unstable tracking and the other with stable tracking using the

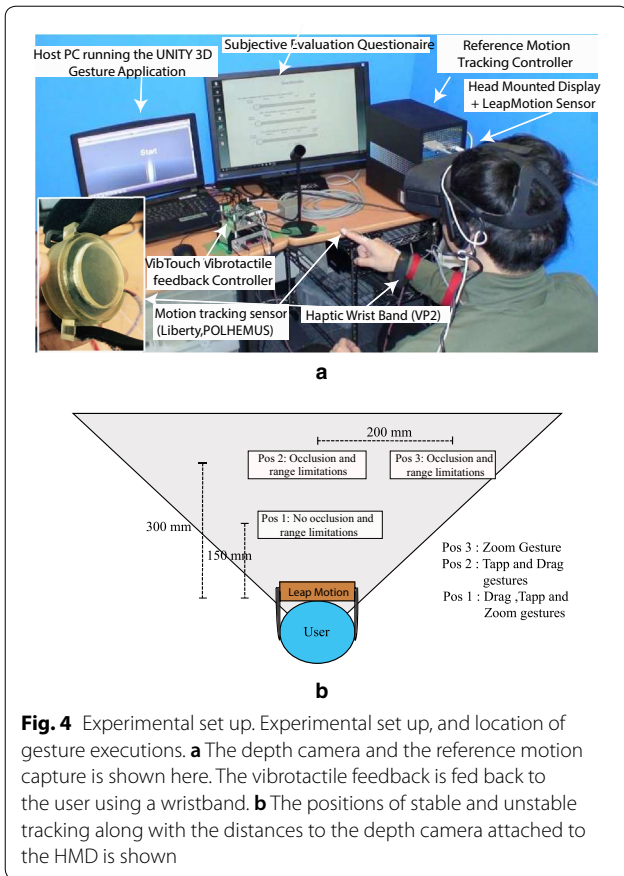


Fig. 4 Experimental set up. Experimental set up, and location of gesture executions. **a** The depth camera and the reference motion capture is shown here. The vibrotactile feedback is fed back to the user using a wristband. **b** The positions of stable and unstable tracking along with the distances to the depth camera attached to the HMD is shown

Leap Motion sensor as shown in Fig. 4b. The instability in tracking occurs due to two factors, self-occlusion of the tracked position and sensor range limitations.

Self-occlusion of the fingertip is caused by the pointing posture in tapping and dragging gestures. In both of these gestures, a pointing pose is used, in which the index finger is extended and all other fingers are closed are used. Such a posture coupled with the Leap motion camera being placed in the VR headset causes self-occlusion in which the camera cannot track the fingertip joint position but rather approximates the joint locations. In the zoom gesture, the tracking of three finger tip positions also causes self-occlusion since the tracked points are occluded by the user’s palm.

The Leap motion sensor has a maximum sensitivity at a range of 100–500 mm in a circular arc shape radius [5]. As the finger position which is being tracked goes beyond this level, the accuracy of tracking reduces at 250 mm from the centre of Leap motion [5]. The position of execution of different gestures are selected such that above-mentioned instabilities arises as shown in Fig. 4b. Both occlusion and range limitations are selectively avoided by

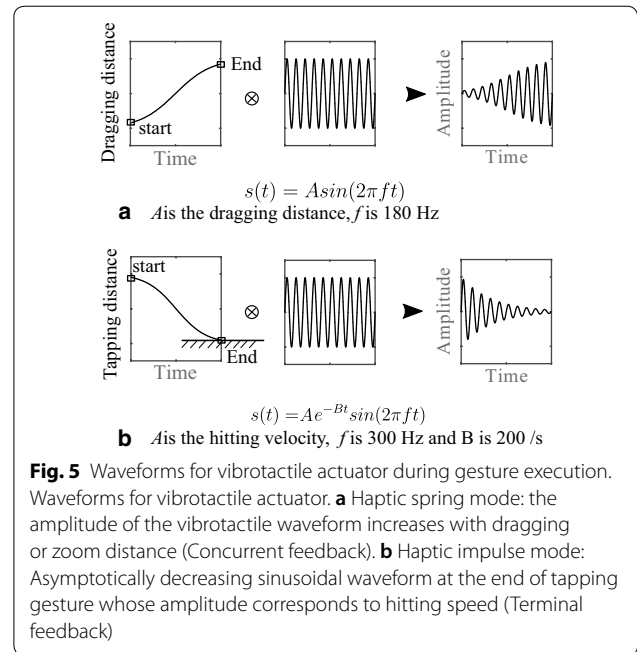


Fig. 5 Waveforms for vibrotactile actuator during gesture execution. Waveforms for vibrotactile actuator. **a** Haptic spring mode: the amplitude of the vibrotactile waveform increases with dragging or zoom distance (Concurrent feedback). **b** Haptic impulse mode: Asymptotically decreasing sinusoidal waveform at the end of tapping gesture whose amplitude corresponds to hitting speed (Terminal feedback)

opting a closer position of location, Pos1 of execution of the gesture as shown in Fig. 4b.

This was used to evaluate the performance of the motion synthesis method under stable and unstable motion tracking conditions. The participants maintained their positions during gesture execution using a visually rendered cylinder.

In the horizontal and vertical drag gestures (shown in Fig. 2c, d), the participant’s task was to drag a box of size 5 cm × 5 cm × 5 cm horizontally and vertically from a start position to a stop position separated by a distance of 20 cm. Two boxes of size 4 cm × 4 cm × 4 cm were rendered to indicate the start and stop position. The color of the boxes changed to blue and red to indicate the start and stop time respectively. While executing gestures, a continuous vibrotactile “spring-like” feedback [28] as shown in Fig. 5a was rendered to the subjects. Here, the amplitude of the sinusoidal waveform was modulated by the distance of the fingertip from the start position so that the participants could feel a virtual spring-like feeling as they drag the box. At the starting point of the dragging gesture, the transient vibration $s(t)$ was defined as

$$s(t) = M \sin(2\pi ft), \tag{2}$$

where the amplitude M is the normalized distance of the fingertip from the start position. The frequency f is set to 180 Hz, which corresponds to the maximum sensitivity of Pacinian corpuscles. On completion of the drag gesture, a vibrotactile inertial and viscous mode was

employed as the cue for task completion. The amplitude of the sinusoidal waveform depended on the velocity and acceleration of fingertip motion. A constant value for finger speed coupled with actual finger acceleration was used to provide feedback for a duration of 60 ms. This assignment was motivated by the findings reported in [8], where it was reported that during finger sliding, the acceleration and velocity of finger motion influenced the perception of two different physical parameters: mass and viscosity respectively.

During the tapping gesture (shown in Fig. 2a), an impulse feedback as shown in Fig. 5b was rendered when the finger contacted the interaction surface. The vertical tapping distance was set at 15 cm and the color of the surface changed to red when the finger makes contact with the surface. From the point of fingertip collision with surface, the transient vibration $s(t)$ was defined as

$$s(t) = Me^{-Nt} \sin(2\pi ft), \quad (3)$$

The amplitude M is velocity at the point of impact and the decay constant N was set to 200. This generated vibrations lasting less than 40 ms. The frequency f was set to 300 Hz to give the sensation of a hard surface. The above parameters were chosen to optimize the user perception of vibrotactile feedback based on the previous research [29].

During the zoom gesture (shown in Fig. 2b), the participant's task was to zoom the size of a sphere to three-times its original size. They could repeat the gesture as many times as required to complete the task. The size of the sphere was controlled by the virtual radius of the three fingers during the zoom gesture. A spring-like vibrotactile feedback was fed back to the subjects controlled by the zoom radius, similar to that of the drag gesture. However, no terminal feedback was provided when the gesture was completed. Here the amplitude of the sinusoidal waveform was modulated by the zoom radius so that the participants could feel a virtual spring-like feeling as they execute the gesture.

Procedure

At each gesture execution position, the participants repeated the task ten times in blocks of 5 trials each. Trials were performed for each experimental condition as described in previous section. The order of vibrotactile feedback (synthesized, measured, and filtered) was randomly changed to avoid any training effect on the subjects. Before the start of each gesture, a practice session was conducted. This allowed the positions of unstable and stable tracking of gestures to be recalibrated for each subject. Each gesture was completed in 30 min and the entire experiment, including rest time and the practice

session, required around 3 h. After each trial, the measured motion data from the depth camera sensor, filtered motion data, synthesized motion data, and the reference motion tracking data were recorded by the application. Only cases where the gesture was correctly recognized by the system was used in the analysis.

Objective analysis based on measured motion data

The following objective parameters were used to evaluate the effectiveness of the motion synthesis method.

- To evaluate the shape of the motion data compared to the reference motion, we use the R^2 values of the measured, filtered, and synthesized motion data and compared them with the reference motion curve. The R^2 values give the relative comparison of the shape of two motion curves on a scale of 0–1 after adjusting to the shift in time differences.
- To evaluate the time difference of endpoints of measured and synthesized motion data compared to that of the reference motion data, we define T_{diff} given by $(Endtime_{ref} - Endtime_{measured}$ and $Endtime_{ref} - Endtime_{syn})$.

Subjective analysis based on user experience

We also asked the participants to rate the VR mid-air interaction system in each condition after every five trials. The following three questions were answered by the participants on a 5-point Likert-type scale from 0 (strongly disagree) to 4 (strongly agree) .

- Synchronization judgement: The haptic feedback was synchronized with my hand motion.
- Smoothness judgement: The haptic feedback was very smooth.
- Task completion judgement: The haptic feedback helped me in the task completion.

Statistical data analysis

A one-sample Kolmogorov–Smirnov test of subjective and objective data across all subjects suggested a normal distribution. Thus, all the statistical analysis reported henceforth are conducted using two-way repeated measures ANOVA followed by post hoc tests using Bonferroni corrected two sample T-tests. The two-way repeated measures ANOVA estimated the main effects of tracking stability (stable and unstable) and the type of motion profiles used (measured, synthesized, and filtered) across all subjects.

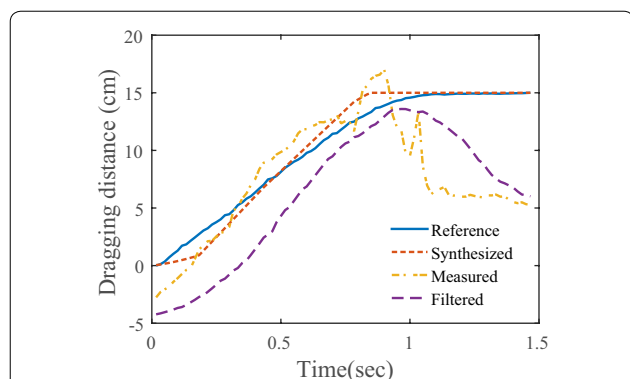


Fig. 6 Qualitative results. The comparative plot of different motion profiles during horizontal drag gesture execution under unstable tracking conditions. The synthesized motion data is smooth and best replicates the reference motion pattern. By comparison the measured motion data has high fluctuations making the vibrotactile feedback unrealistic and unpleasant

Evaluation results

Figure 6 shows the comparative plot of reference motion data along with measured, moving average filtered and synthesized motion curves during horizontal drag gesture for subject 8. From the figure, it is evident that while synthesized motion curve closely replicates the reference motion data, the measured motion data have instability at the tag end of the gesture. This instability occurs due to the poor tracking accuracy as the finger position moves away from the depth camera sensor. The moving average filtered data has a different shape and is significantly delayed compared to the reference motion data. This leads to a delayed vibrotactile feedback with diminished amplitude.

Objective evaluation results

The R^2 values and T_{diff} of the four different gestures—horizontal drag, vertical drag, tapping, and zooming—are shown in Fig. 7a–d respectively. Under poor tracking stability, we observed an improvement in R^2 values and T_{diff} for the synthesized motion data. Under high tracking stability, the performance of the measured and synthesized motion data had similar objective evaluation scores. Moreover, the filtered motion data was significantly delayed in all the gesture conditions.

The synthesized motion data from the horizontal drag gesture had high correlation in shape and timing with the reference motion data when the tracking was unstable. When the tracking was stable, however, there was no difference between measured and synthesized data. The motion profile used had a statistically significant effect on R^2 values given by ($F_{2,16}=4.67$, $p = 0.0251$) for the horizontal drag task as shown in Fig. 7a. The

interaction between the main effects were also significant ($F_{2,16} = 5.85$, $p = 0.0123$). The effect of motion profiles had significant effects on T_{diff} ($F_{2,16} = 91.025$, $p < 0.00001$), however tracking stability had no significant effect on the horizontal drag gesture. This shows that T_{diff} had similar trends regardless of tracking stability. The post hoc test showed significant differences in R^2 values between different motion profiles. Synthesized motion profiles had higher R^2 values compared to measured motion profiles ($p = 0.0002$) during unstable tracking. But there was no significant difference in R^2 values between synthesized and measured motion profiles when the tracking was stable. The post hoc test on T_{diff} for both tracking conditions shows that the filtered motion data is significantly delayed compared to the measured and synthesized data for the horizontal drag gesture ($p < 0.00005$).

For the vertical drag gesture, the end timings from the synthesized motion data had a significantly better correlation with the reference motion data. The general trend shows an increased performance of measured data under high tracking stability conditions and vice versa. There was no significant effect of tracking stability or motion profiles on the R^2 values, albeit R^2 values from the measured motion profiles were lower. Significant differences in T_{diff} were observed for both the tracking stability ($F_{2,16}=8.18$, $p = 0.0211$) and motion profiles ($F_{2,16}=8.71$, $p = 0.0184$). A comparison of the measured data to the synthesized motion data under unstable tracking conditions shows that the end time from the measured data was significantly delayed ($p = 0.0408$). Under stable tracking conditions, the end times of the measured motion profile were synchronized with the reference motion profiles while the synthesized motion data reached the end state much earlier than the actual gesture ($p < 0.00005$).

For the tapping gesture, there were no improvements in the shape and timing of the synthesized motion data compared to measured motion profile. While measured motion data had a higher correlation with the reference motion curve under stable tracking conditions, the performance was similar when the tracking was unstable. Both the tracking stability ($F_{2,16}=10.13$, $p = 0.0133$) and motion profiles ($F_{2,16}=21.6841$, $p = 0.0016$) had significant effects on the R^2 values. The R^2 values were higher for measured data in both tracking conditions and we observed significantly higher R^2 values under stable tracking conditions ($p = 0.0042$) as shown by the post hoc test. The post hoc analysis showed that the increase in R^2 values were not significant when tracking was unstable. The tracking stability had significant effects on T_{diff} ($F_{2,16}=12.83$, $p = 0.0072$) but there was no significant effect of type of motion profile (measured,

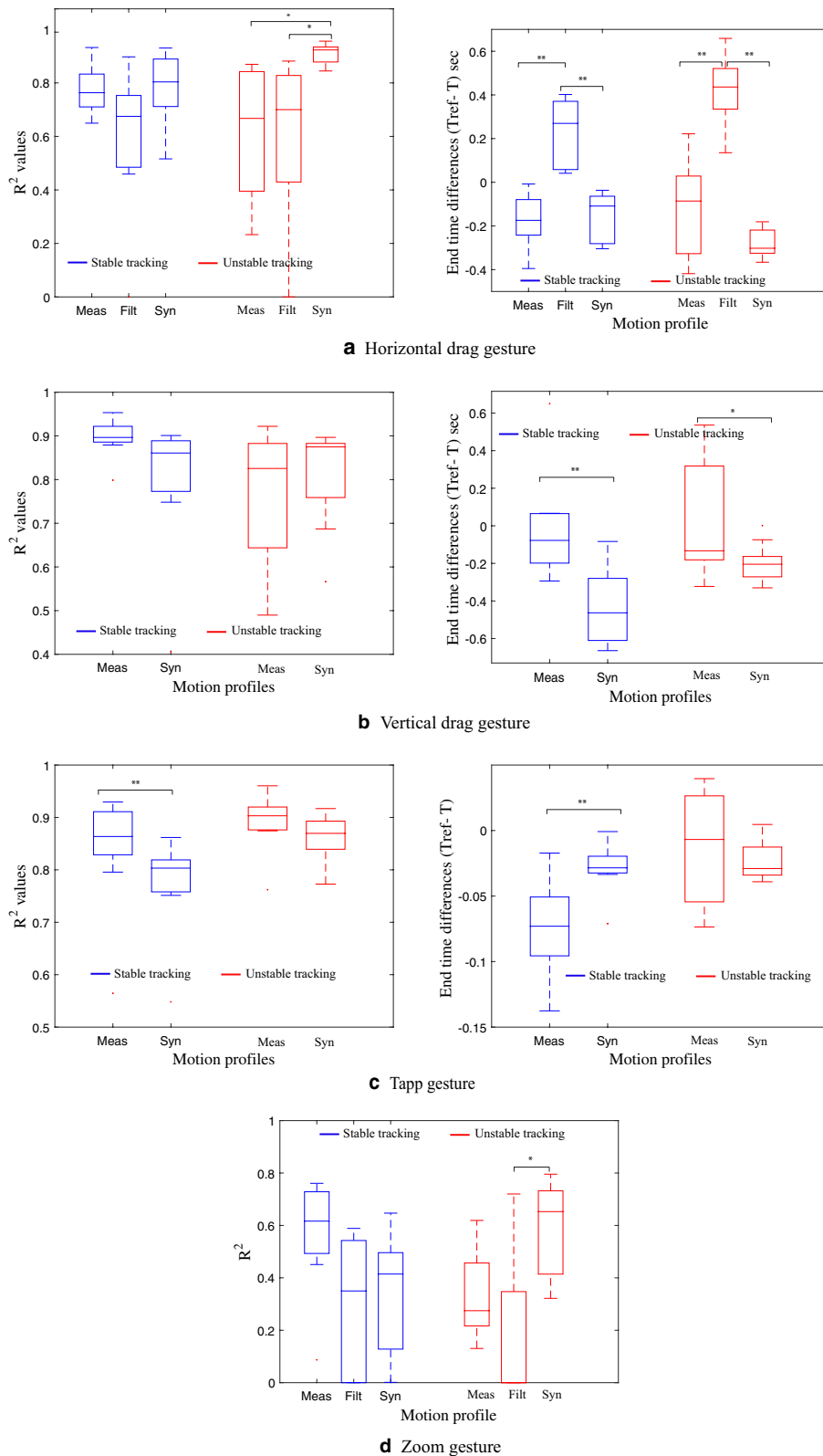


Fig. 7 Quantitative evaluation results. Quantitative evaluation results showing the shape and end timings of the motion profiles during gestures. **a** Horizontal drag, **b** vertical drag, **c** tapping, **d** zooming. Here the *Meas* corresponds to measured motion data, *Syn* synthesized motion and *filt* filtered motion data respectively (** $p < 0.001$ and * $p < 0.05$. Adjustment for multiple comparisons: Bonferroni)

synthesized or filtered). When tracking was stable, the end time difference between the reference motion curve and the measured data was significantly earlier compared to synthesized motion curve ($p = 0.003$) as shown by the post hoc analysis. With unstable tracking, however, there was no difference in the end times of the motion profiles for the tapping gesture.

In the zoom gesture, the synthesized motion data had a high correlation in shape with the reference motion data under unstable tracking conditions. When the tracking was stable, however, there was no difference between the measured and synthesized data. The end-timing was not analyzed for zoom gesture as no terminal vibrotactile feedback was provided at the end of gesture execution. The motion profiles had significant effects on R^2 values ($F_{2,16}=3.43$, $p = 0.057$) but tracking stability conditions had no significant effect. The interaction between the above main factors were statistically significant ($F_{2,16}=10.38$, $p = 0.0013$). The R^2 values were significantly higher for synthesized motion profiles compared to filtered motion data during unstable tracking ($p = 0.0004$).

Subjective evaluation results

The box plot of three subjective scores for different gestures is shown in Fig. 8. The general tendency was an improvement in subjective scores for synthesized motion data based on vibrotactile feedback, especially synchronization judgement, when the tracking stability was poor. When the tracking stability was high, the user ratings of measured and synthesized motion data based on vibrotactile feedback had similar scores. Moreover, the vibrotactile feedback based on filtered motion data had lower user ratings compared to others in all cases.

In all three subjective scores, the synthesized motion curve based on vibrotactile feedback for horizontal dragging gesture (Fig. 8a) had higher scores compared to measured and filtered motion curves based on feedback under unstable tracking conditions. However, the same was not true for stable tracking. The synchronization judgement rating scores were significantly higher for vibrotactile feedback based on synthesized motion data for tracking stability ($F_{2,16}=10.74$, $p = 0.0112$) and motion profiles ($F_{2,16}=4.62$, $p = 0.00261$). The two-way repeated Measures ANOVA gave statistically significant higher scores with tracking stability changes for the smoothness judgement score ($F_{2,16}=12.815$, $p = 0.0072$) and task completion judgement score ($F_{2,16}=18.89$, $p = 0.0025$). Subjective scores on the effect of motion profiles was not statistically significant in these cases. The post hoc analysis gave statistically significant differences only between synthesized and filtered motion data in the synchronization judgement score ($p = 0.002$).

The subjective evaluation scores of the vertical dragging gesture with changes in tracking stabilities and motion profiles are shown in Fig. 8b respectively. There were statistically significant higher scores for synchronization judgement with the two main factors. The two-way repeated measures ANOVA applied to changes in the stability motion profiles returned values of ($F_{2,16} = 18.185$, $p = 0.0027$) and ($F_{2,16}=11.636$, $p = 0.0092$) respectively. However, the post hoc test only gave statistically significant higher ratings in synchronization judgment scores for synthesized motion curves based on vibrotactile feedback under unstable tracking conditions ($p = 0.0037$). For other subjective scores, the increase in ratings for synthesized motion curves based on vibrotactile feedback was not statistically significant.

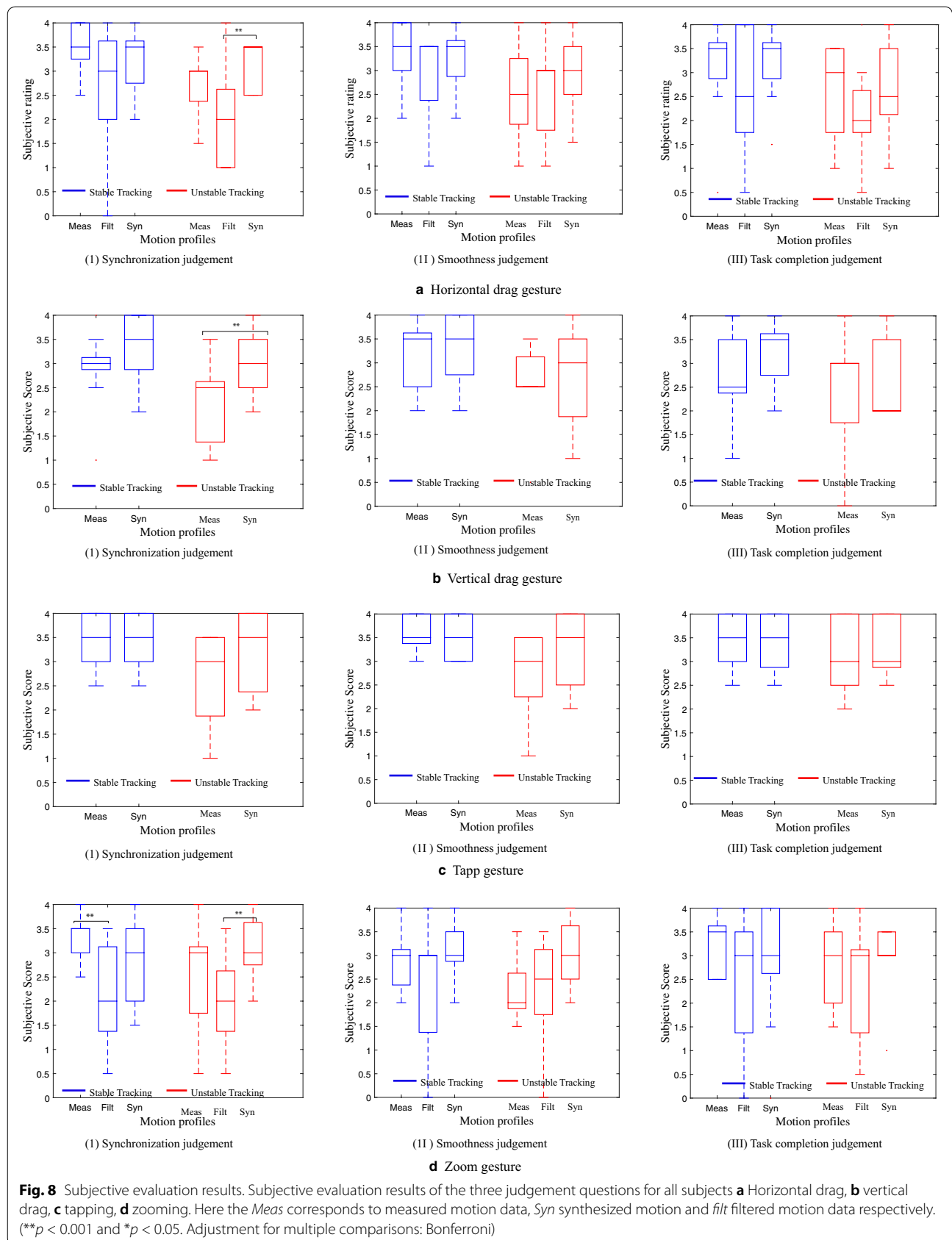
For tapping gesture (shown in Fig. 8c) there was statistically significant changes in subjective rating for all the three judgement criteria with changes in tracking stabilities given by ($F_{2,16}=11.93$, $p = 0.0086$), ($F_{2,16}=10.388$, $p = 0.0122$) and ($F_{2,16}=6.4$, $p = 0.0353$) respectively. However, the subsequent Bonferroni corrected paired T-test did not show any statistically significant increase in subjective rating scores when the tracking was unstable.

Figure 8d shows the subjective evaluation scores for the zoom gesture with changes in tracking stabilities and motion profiles. There were statistically significant higher scores for synchronization judgement with changes in motion profiles ($F_{2,16}=3.636$, $p = 0.0427$). There were also statistically significant interactions between motion profiles and changes in tracking stabilities on the subjective scores ($F_{2,16}=5.724$, $p = 0.0133$). Post hoc analysis showed that the individual scores were significantly higher for the synthesized curve compared to the filtered motion curve when the motion tracking was unstable ($p = 0.008$). But when the tracking was stable, the scores were similar for both measured and synthesized curves. For smoothness and task completion judgments ratings, the improvement in scores for synthesized motion curves was not statistically significant compared to measured and filtered motion data.

Discussions

Objective evaluation

In general, the motion curve shape metric increased for synthesized data when the tracking was unstable. This is underlined by the significant increases in R^2 values for horizontal drag and zoom gestures. The performance of different motion profiles was similar under stable tracking for all gestures. This shows that the proposed motion synthesis algorithm maintains the shape of the motion data during both stable and unstable tracking conditions. One can observe small R^2 values for filtered motion data. This shows that filtering of the motion data alters



the shape of the motion profiles. This change in shape of the motion data leads to diminished vibrotactile feedback based on the altered motion profile. The shape of the motion curve during gesture execution is significant for concurrent haptic feedback, where haptic feedback is fed continuously to the user from the start of the gesture until the endpoint.

Differences in endpoint timing are important for terminal haptic feedback, where haptic feedback is fed to the users at the end of gesture execution. When the tracking was stable, the endpoints of the measured motion profiles during drag gestures were close to that of the reference motion curves. However, when tracking was unstable, the endpoints of measured motion profiles were delayed compared to the actual motion profile. One can observe a significant delay in the endpoints of the filtered motion profiles compared to the measured and synthesized data for the drag gesture.

The synthesized curves had endpoints earlier to the actual endpoints by 200 ms and 30 ms for dragging and tapping gestures respectively. This may be caused by two factors: the minimum time required to recognize a gesture and the modulation of motion primitives stored in the look-up table. The gesture recognition system requires some time duration to recognize each gesture in realtime. Additionally, the modulation of the motion primitives in real-time is implemented by estimating the duration in each state and comparing with the ideal motion pattern recorded previously. This method led to overcompensation and/or under compensation and may lead to changes in execution timings even though the shape of the curve is maintained.

Unlike other gestures, the R^2 and T_{diff} values of the synthesized motion curves for the tapping gesture are not significantly improved. This can be explained by the duration of the tapping gesture as compared to dragging and zooming gestures. The average duration of the tapping gesture was 0.35 ± 0.15 s compared to 1.26 ± 0.25 and 0.88 ± 0.123 s for horizontal dragging and zoom gestures respectively. Thus, the tracking instabilities have a smaller effect on the shape of motion curves and timing differences. Moreover, when the speed of gesture execution is fast, controlling the instability using changes in the position of gesture execution becomes less efficient.

Subjective evaluation

Out of the three judgment questions, the synchronization judgment scores showed a significant increase for vibrotactile feedback based on synthesized motion curves when the tracking stability was poor. End timing differences were easier to perceive compared to the continuous increase in the amplitude of the vibrotactile waveform. This can be explained by previous studies which show

that terminal haptic feedback has prominent effects on user perception compared to concurrent haptic feedback [6]. Previous research by Jay et al. [30, 31] and Lee et al. [32] have also shown that the user interaction metrics such as task completion times and penetration depth will be significantly affected when the terminal haptic feedback is delayed by more 150 ms. The significantly lower user ratings of filtered motion data in the current study further reiterate the above findings.

Even though the task completion judgment scores increased for all subjects under unstable tracking conditions, the increase in ratings was not statistically significant when compared to synchronization judgment. This sheds light on the necessity of haptic feedback for task completion in mid-air gesture tasks. Although haptic feedback improves the pleasantness of virtual interactions, participants used visual feedback as the primary cue for estimating the end of a gesture rather than haptic feedback. The above argument is in line with the previous studies conducted by Jay and Hubold [31] for 1 DoF tapping and target acquisition tasks. They concluded that the absolute essentiality of haptic feedback is a primary factor in task execution improvements.

The low rating score improvement of the tapping gesture for the motion synthesized curve based on haptic feedback is consistent with the quantitative evaluation. The R^2 and T_{diff} values of tapping task had no significant changes with the change in motion synthesized curves. Thus, when the haptic feedback was based on these curves, the changes in subjective ratings was also detrimental.

Advantages of the proposed methodology

The proposed method offers a more general and stable approach to generate haptic feedback for mid-air haptic interfaces compared to the conventional approaches. In the event of instability in motion tracking, which frequently occurs in mid-air interactions, the proposed motion synthesis maintains the stability of motion profile and in turn the haptic feedback, thereby improving overall user satisfaction and task performance. Moreover, it has been shown by our previous studies [33] that addition of realistic tactile feedback changes the human motion profile as well, making it more streamlined and predictable. This increased stability of human motion pattern, in turn increases the recognition accuracy and thereby the stability of the tactile feedback as a result [33]. The conventional studies for haptic feedback based motion correction and training does not take into consideration this instability in motion tracking.

The motion synthesis method can be used to increase the frame rate of the motion data. The current commercially available depth cameras have an update rate

of 50–60 Hz with variable frame rates depending on the computational load. The motion synthesis method can be used to increase and stabilize the frame rate as motion data is synthesized by a separate thread once the gesture is recognized.

The proposed method could be extended for other feedback modalities of virtual interactions as well. In the mid-air interactions, controlling one's motion profile to be the ideal is hard especially when interactions involve motion based feedback. In the current study, we used the motion synthesis method to reproduce this ideal motion curves irrespective of the errors in the user's gesture execution. Although the synthesized motion data is used only for haptic feedback generation in the current study, the proposed method can be extended to other feedback modalities such as visual and audio feedback as well. If the visual feedback is also stabilized using the proposed method, user satisfaction of the virtual multimodal interactions can be significantly improved as human perception is dominated substantially by visual feedback when judging size, position or shape [34]. This may ensure that users do not make any mistakes during the intended gesture execution even in the presence of unstable motion tracking. Moreover, the target applications in the environment can be customised, thereby allowing the HMM to recognize the suitable gesture in real-time quickly. Thus a predefined environment with known applications and corresponding positions can greatly improve the recognition of the HMM and accuracy of user motion execution.

Multiplicity in definition of tactile feedback

The proposed method allows multiplicity in the definition of haptic feedback signals. For e.g. a velocity curve based haptic rendering for dragging gesture can be achieved by applying a derivative of the dragging distance stored in the look-up table instead of the distance vector itself. Similar approaches can be extended to other gestures such as tapping, zooming etc also where velocity and acceleration curves can be obtained in real-time from the reference motion curves by applying derivatives in real-time. Moreover, for applications requiring a specific velocity rendering curve, the recognition phase remains the same, but the reference motion curve and the corresponding primitive motion elements stored in the look-up table only changes. This allows for easy modulation of the proposed approach with the changes in the definition of the haptic feedback controlling motion elements for each gesture.

Scalability

The proposed method incorporates a scalable architecture for extending to other gesture applications. For

the addition of a new gesture, a new HMM should be trained corresponding to the gesture. Also, a reference motion curve corresponding to the new gesture has to be stored in the look-up table as well. For example, for a circular motion gesture, first a visual set up enabling the corresponding task is set up first. The definition of an adequate visual set up will avoid unwanted user motions, thereby reducing the occurrences of false positives during gesture recognition process. The reference motion curve to be synthesized (in this case will be a circle) is stored in the memory as a look up table. Next, an HMM model for circular motion gesture recognition having multiple underlying states is trained with different user hand or finger motion features such as position, velocity profiles of hand palm etc. With the change in the size of the circle, a simple scaling of the radius of the circle and corresponding arcs in look-up table will suffice. Moreover, in the proposed approach, the primitive motion elements stored in the look-up table are adaptively modulated according to the real-time execution changes by the participants such as speed of execution. This allows changing the reference motion pattern in real-time according to the changes in user gesture execution characteristics.

Implications of the current study

Only basic gestures have been analyzed in the current study. Thus the recognition rate of the HMM is more than 90% in all cases in the current study. However as the number of gestures increases, the recognition accuracy of HMM may get strained which can adversely affect the synthesized curve profiles.

Simple gestures were used in the current study because it enabled a direct mapping between user's performance/satisfaction with that of the haptic feedback generated by the proposed method for both stable and unstable tracking conditions. A comparative analysis of subjective judgement scores (synchronization, smoothness and task completion) and objective scores (shape, timing of the motion profile) was possible for simple gestures. In case of complex gestures, multiple external factors affected user performance, thereby making the experimental set up difficult to standardize. For example, consider the case of a virtual driving application, in which the haptic feedback is controlled by the rotation of the steering wheel by both hands of the user. The quantification of stability/instability conditions for motion curves during driving is difficult to achieve for all participants as each participant have different driving patterns. Any differences between the participants with regards to stability would render the quantitative discussions involving subjective and objective parameters irrelevant.

Furthermore, previous studies of the authors [33] have shown that the use of proper vibrotactile feedback for mid air gestures such virtual mid air writing and tapping tasks improved user performances. The importance of timing of haptic feedback in pointing tasks have been widely reported by other researchers as well [30–32]. Studies conducted by Jay et al. [30, 31] showed that task performance decreased when delay between haptic and visual feedback exceeded 150 ms for pointing tasks. These studies underlines the viability of the proposed method for not only more complex gestures but also for simple gesture interactions as well.

Conclusion

In this paper, we proposed a motion synthesis method for real-time, stable haptic feedback generation during mid-air interactions. The proposed method uses an HMM to recognize the gestures. Motion elements were synthesized based on recognized gestures to control the vibrotactile feedback. Four gestures (tapping, three-fingered zooming, vertical dragging, and horizontal dragging) were used in the study to evaluate the performance of the motion synthesis method.

The ideal motion curves and corresponding primitive motion elements to be synthesized for each gesture were obtained from multiple subjects in different conditions using a reference motion tracking sensor. An adaptive control algorithm was implemented to modulate the primitive motion elements based on the user's actual gesture execution speed. Separate HMM models were trained for each gesture and motion patterns were synthesized in real time in spite of changes in speed and tracking irregularities. The shape and timing of the synthesized, measured, and moving average filtered motion data were compared with the reference motion curve obtained from a stable sensor. Moreover, user satisfaction levels for concurrent and terminal vibrotactile feedback based on different motion data were compared by a subjective evaluation using a questionnaire.

Both objective and subjective evaluation results showed improvements with motion synthesis method. The objective evaluation results showed a significant increase in shape and end timing performance of the synthesized motion curves for different gestures in unstable tracking environments. The subjective evaluation results also supported the viability of motion synthesis based on haptic feedback when tracking stability was poor. When the executed gesture was fast, as is the case in tapping, the effect of tracking instability was minimal, and motion synthesis had no significant improvements in objective and subjective scores. The subjective evaluation results showed that participants could better perceive synchronization of vibrotactile

feedback with hand motion when synthesized motion data was used.

The proposed method ensures scalability for multiple gestures and sensing platforms making it a general approach. In the future, we intend to extend the motion synthesis to other feedback modalities also so that a more stable VR environment can be synthesized.

Authors' contributions

DB and MK developed the original concept and analytical formulation. HN guided DB in experimental design and analysis of the experimental design. RH guided DB in the hidden markov model implementation. ST encouraged DB to investigate [a specific aspect] and supervised the findings of this work. DB wrote the manuscript with inputs from MK. All authors discussed the results and contributed to the final manuscript. All authors read and approved the final manuscript.

Author details

¹ Graduate School of Information Sciences, Tohoku University, Aramaki-aza Aoba, Sendai 980-0871, Japan. ² NiChe, Tohoku University, Aramaki-aza Aoba, Sendai 980-0871, Japan.

Acknowledgements

The authors would like to thank all the members of Human Robot Informatics Lab, Tohoku University for helping conduct the experiments and for their fruitful feedbacks.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported in part by ImPACT (Tough Robotics Challenge).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 1 December 2017 Accepted: 30 January 2019

Published online: 13 February 2019

References

- Okamura AM, Dennerlein JT, Howe RD (1998) Vibration feedback models for virtual environments. In: Proceedings 1998 IEEE international conference on robotics and automation (Cat. No. 98CH36146), vol. 1, pp 674–679. <https://doi.org/10.1109/ROBOT.1998.677050>
- Adams RJ, Olowin AB, Hannaford B, Sands OS (2011) Tactile data entry for extravehicular activity. In: 2011 IEEE world haptics conference. pp 305–310. <https://doi.org/10.1109/WHC.2011.5945503>
- Petzold B, Zaeh MF, Faerber B, Deml B, Egermeier H, Schilp J, Clarke S (2004) A study on visual, auditory, and haptic feedback for assembly tasks. *Presence Teleoperat Virt Environ* 13(1):16–21. <https://doi.org/10.1162/105474604774048207>
- Guna J, Jakus G, Poganik M, Sodnik Toma J Saand (2014) An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors* 14(2):3702–3720
- Weichert F, Bachmann D, Rudak B, Fisseler D (2013) Analysis of the accuracy and robustness of the leap motion controller. *Sensors* 13(5):6380–6393. <https://doi.org/10.3390/s130506380>
- Ahmaniemi T et al (2012) Dynamic tactile feedback in human computer interaction. Ph.D. thesis. Aalto University, Aalto
- Konyo M, Yamada H, Okamoto S, Tadokoro S (2008) Alternative display of friction represented by tactile stimulation without tangential force. In: Haptics: perception, devices and scenarios, pp 619–629
- Okamoto S, Konyo M, Tadokoro S (2011) Vibrotactile stimuli applied to finger pads as biases for perceived inertial and viscous loads. *IEEE Trans Haptics* 4(4):307–315

9. Inamura T, Nakamura Y, Ezaki H, Toshima I (2001) Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In: IEEE international conference on robotics and automation proceedings 2001 ICRA, vol. 4, pp 4208–4213
10. Inamura T, Toshima I, Tanie H, Nakamura Y (2004) Embodied symbol emergence based on mimesis theory. *Int J Robot Res* 23(4–5):363–377
11. Babu D, Nagano H, Konyo M, Hamada R, Tadokoro S (2016) Stablehaptic feedback generation during mid air interactions using hiddenmarkov model based motion synthesis. In: Hasegawa S, Konyo S, Kyung M, Nojima K-U, Kajimoto H (eds) *Haptic interaction*. Springer, Singapore, pp 225–231
12. Kirk AG, O'Brien JF, Forsyth DA (2005) Skeletal parameter estimation from optical motion capture data. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol. 2, pp 782–7882. <https://doi.org/10.1109/CVPR.2005.326>
13. Won HPS, Melek WW, Golnaraghi F (2010) A kalman/particle filter-based position and orientation estimation method using a position sensor/inertial measurement unit hybrid system. *IEEE Trans Ind Electron* 57(5):1787–1798. <https://doi.org/10.1109/TIE.2009.2032431>
14. LeapMotion Orion Update. <https://developer.leapmotion.com/orion/>
15. Ye Y, Liu PX (2009) Improving haptic feedback fidelity in wave-variable-based teleoperation orientated to telemedical applications. *IEEE Trans Instrument Measur* 58(8):2847–2855
16. Yao H-Y, Hayward V, Ellis RE (2005) A tactile enhancement instrument for minimally invasive surgery. *Comput Aided Surg* 10(4):233–239
17. Du G, Zhang P, Liu X (2016) Markerless human–manipulator interface using leap motion with interval kalman filter and improved particle filter. *IEEE Trans Ind Inform* 12(2):694–704
18. Keskin C, Kırac F, Kara YE, Akarun L (2012) Hand pose estimation and hand shape classification using multi-layered randomized decision forests. In: *European conference on computer vision*. Springer, Berlin, pp 852–863
19. Xu C, Cheng L (2013) Efficient hand pose estimation from a single depth image. In: *Proceedings of the IEEE international conference on computer vision*, pp 3456–3462
20. Schönauer C, Mossel A, Zaiti I-A, Vatavu R-D (2015) Touch, movement and vibration: user perception of vibrotactile feedback for touch and mid-air gestures. In: *Human–computer interaction*. Springer, Berlin, pp 165–172
21. Martínez J, García A, Oliver M, Molina JP, González P (2016) Identifying virtual 3d geometric shapes with a vibrotactile glove. *IEEE Comput Graph Appl* 36(1):42–51
22. Nunez OJA, Lubos P, Steinicke F (2015) Hapring: a wearable haptic device for 3d interaction. In: *Mensch & computer*, pp 421–424
23. Subramanian S, Seah SA, Shinoda H, Hoggan E, Corenthy L (2016) Mid-air haptics and displays: systems for un-instrumented mid-air interactions. In: *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. CHI EA '16, ACM, New York, pp 3446–3452
24. Sodhi R, Glisson M, Poupyrev I (2013) Aireal: Tactile gaming experiences in free air. In: *ACM SIGGRAPH 2013 emerging technologies*. SIGGRAPH '13. ACM, New York, pp 2–121. <https://doi.org/10.1145/2503368.2503370>
25. Lee H, Kim J-S, Choi S, Jun J-H, Park J-R, Kim A-H, Oh H-B, Kim H-S, Chung S-C (2015) Mid-air tactile stimulation using laser-induced thermoelastic effects: the first study for indirect radiation. In: *World haptics conference (WHC)*, pp 374–380
26. Rabiner LR (1989) A tutorial on hidden markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286
27. Equation of Circle Passing Through Three Points. <https://www.qc.edu.hk/math/Advanced20Level/circle/>
28. Ito O, Kawabe H, Konyo M (2016) Input device, input system, electronic apparatus, and sense presentation method. US Patent 9,489,041
29. Okamura AM, Cutkosky MR, Dennerlein JT (2001) Reality-based models for vibration feedback in virtual environments. *IEEE/ASME Trans Mech* 6(3):245–252
30. Jay C, Hubbard R (2005) Delayed visual and haptic feedback in a reciprocal tapping task. In: *First joint eurohaptics conference and symposium on haptic interfaces for virtual environment and teleoperator systems*. World Haptics Conference, pp 655–656. <https://doi.org/10.1109/WHC.2005.29>
31. Jay C, Glencross M, Hubbard R (2007) Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. *ACM Trans Comput Hum Interaction* 14:2
32. Lee I, Choi S (2007) Discrimination of virtual environments under visual and haptic rendering delays. In: *2007 frontiers in the convergence of bioscience and information technologies*, pp 554–562
33. Babu D, Kim S, Nagano H, Konyo M, Tadokoro S (2016) Can haptic feedback improve gesture recognition in 3d handwriting systems? In: Kubota N, Kiguchi K, Liu H, Obo T (eds) *Intelligent robotics and applications*. Springer, Cham, pp 462–471
34. Rock I, Victor J (1964) Vision and touch: an experimentally created conflict between the two senses. *Science* 143(3606):594–596

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
