Touching the Sphere: Leveraging Joint-Centered Kinespheres for Spatial User Interaction

Paul Lubos, Gerd Bruder, Oscar Ariza, Frank Steinicke
Human-Computer Interaction, Department of Informatics, Universität Hamburg
{lubos,bruder,ariza,steinicke}@informatik.uni-hamburg.de

ABSTRACT
Designing spatial user interfaces for virtual reality (VR) applications that are intuitive, comfortable and easy to use while at the same time providing high task performance is a challenging task. This challenge is even harder to solve since perception and action in immersive virtual environments differ significantly from the real world, causing natural user interfaces to elicit a dissociation of perceptual and motor space as well as levels of discomfort and fatigue unknown in the real world. In this paper, we present and evaluate the novel method to leverage joint-centered kinespheres for interactive spatial applications. We introduce kinespheres within arm’s reach that envelope the reachable space for each joint such as shoulder, elbow or wrist, thus defining 3D interactive volumes with the boundaries given by 2D manifolds. We present a Fitts’ Law experiment in which we evaluated the spatial touch performance on the inside and on the boundary of the main joint-centered kinespheres. Moreover, we present a confirmatory experiment in which we compared joint-centered interaction with traditional spatial head-centered menus. Finally, we discuss the advantages and limitations of placing interactive graphical elements relative to joint positions and, in particular, on the boundaries of kinespheres.

Categories and Subject Descriptors
H.5.2 [Information Interfaces and Presentation]: User Interfaces–Input Devices and Strategies, Evaluation / Methodology; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism–Virtual Reality

Keywords
Spatial user interfaces; 3D touch interaction; kinespheres; virtual environments; head-mounted displays

1. INTRODUCTION
The recent developments in the field of virtual reality (VR) and augmented reality (AR) head-mounted displays (HMDs) such as the Oculus Rift, HTC VIVE, Meta or Hololens have received much public interest and praise for their design, tracking and visual quality. However, they also received much critique for their spatial interaction and, in particular, for how they handled interaction with 3D graphical user interfaces (GUIs). By combining HMDs with depth sensors such as the Leap Motion it has become possible to leverage direct hand input for spatial interaction, as also seen in the Meta and Hololens. User interfaces based on direct hand interaction, however, received less user acceptance than expected, such that Oculus and HTC even abandoned it in favor of hand-held controllers. While portions of the challenge lie on the hardware side with still limited tracking ranges and accuracy, it has become necessary to revisit spatial menu designs and to develop methods to improve the user acceptance of free-hand spatial user interfaces by considering the human factors of 3D mid-air interaction [33]. In this paper, we revisit mid-air interaction based on the simple observation that for the human jointed arm structure the reachable space from each joint is naturally limited. We refer to such an interactive volume of space around the center of a user’s kinematic joint as a joint-centered kinesphere. While traditional head-centered menus are by design placed within arm’s reach in front of the user, it is seldomly the case that they can be reached without requiring movements of the shoulder joint and as a result inducing high levels of fatigue over time [10]. Moreover, most traditional menus are not matched to each user’s arm length but rather use population means, which means that for some users the menus are placed closer to their maximal reachable arm distance than for others. Requirements for the visual placement and size of spatial GUI elements for good visibility and high interaction performance have been extensively researched [5, 29], but their optimal location in the kinematic chain has received less attention.

To evaluate the optimal placement of menus in the scope of joint-centered kinespheres, we conducted an experiment in which we collected quantitative and qualitative data on usability and selection performance between interaction targets located at different distances on such kinespheres with a Fitts’ Law task. The results support the notion that the efficiency of selections in spatial menus within arm’s reach can be improved by placing interactive elements on joint-centered kinespheres at certain distances, in particular at the maximal reachable distance. We confirm the practical usability of joint-centered menus in an experiment by comparing them against traditional head-centered GUIs. We discuss the advantages and limitations of this method focusing on differences in the length of the kinematic chain for kinespheres centered around the shoulder, elbow or wrist joint.

In this paper we provide the following contributions:

- we evaluate the placement of interactive elements in joint-centered kinespheres with a view on usability and touch performance, and
• we present a joint-centered user interface and compare user acceptance against head-centered menus.

The remainder of this paper is structured as follows: Section 2 presents an overview of related work. Section 3 introduces the concept of joint-centered kinespheres. In Section 4 we describe the experiment that we conducted to evaluate the placement of interactive elements for 3D touch interaction in joint-centered kinespheres. Section 5 describes the confirmatory study that we performed comparing joint-centered and head-centered menus. Section 6 concludes the paper and discusses vistas for future work.

2. RELATED WORK

3D user interfaces and spatial interaction has been in the focus of many research groups over the last decades. According to Mine et al. [35], direct 3D interaction leads to significantly higher performance than manipulation of objects at a distance from the user’s hand. Most results from similar studies agree that optimal performance may be achieved when visual and motor spaces are superimposed or coupled closely within arm’s reach [12, 25, 43]. In contrast, interaction techniques offering the possibility to interact with distant objects by decoupling the virtual hand from the user’s real hand pose include the Go-Go technique [37] and HOMER [6], which work by nonlinear scaling of hand positions within arm’s reach, approaching infinity in the virtual environment (VE) when approaching the maximal reachable distance in the real world.

However, it is still an open research question, how the position of virtual target objects located within arm’s reach may affect interaction performance [29, 30]. Direct 3D interaction in HMD environments is subject to perceptual limitations, e.g., the accommodation-convergence mismatch, ghosting or double vision, which can result in strong misperception effects [8, 9, 11]. Depending on the location of virtual objects, users may be unable to discriminate interrelations between objects or they may perceive distances to objects to be smaller or larger than they are geometrically [28, 30]. Consistently, evaluations of 3D selection performance found that users made most errors along the view direction due to incorrectly judged distances of virtual objects [29]. Such distortions do not appear in the real world and are likely caused by limitations of current technology to correctly reproduce depth cues from the real world in a perfect way [44]. Internal representations of the 3D space are influenced and updated by both visual as well as motor input, which may affect interaction performance [41, 45].

Various interaction techniques have been proposed to improve upon traditional 2D menus anchored on an arbitrary plane in 3D space. For instance, Li et al. [26] propose using the orientation of a mobile device held with outstretched arms for menu control. Shoemaker et al. [40] propose using a user’s personal space to display menus at a comfortably reachable position on large displays. Ens et al. [14] propose a personal cockpit with menu items placed on a sphere centered around the user’s point of view. Also, Gerber and Bechmann [16] center their Spin Menu around a user’s wrist.

Selection Performance

3D selections by touching or grasping objects with a user’s hands can be split up into two phases, the ballistic phase and the correction phase [27]. The ballistic phase consists of focusing on the target object and bringing the hand in the proximity of the goal by using proprioceptive motor control. After that, visual feedback is used in the correction phase in order to incrementally reduce the distance from the hand to the goal.

For such selections, Fitts’ Law predicts the movement time

\[ MT \in \mathbb{R}^+ \]

for a given target distance \( D \in \mathbb{R}^+ \) and target size \( W \in \mathbb{R}^+ \) [13, 15, 31, 32]. They are brought together in a log term, which describes the difficulty of the task overall with

\[ MT = a + b \cdot \log_2(D/W + 1) \]

The values \( a \in \mathbb{R}^+ \) and \( b \in \mathbb{R}^+ \) can be empirically derived for different setups. The index of difficulty (ID) is given by the log term and indicates overall task difficulty; smaller or farther target results in increased difficulty, whereas larger or/and closer targets provide a decreased ID. The formula has been extended in order to get effective measures. The error rate is adjusted to 4% by resizing targets to their effective width \( W_e \in \mathbb{R}^+ \). This adjustment is supported by an international standard [21]. By calculating the average of the measured movement distances, \( D_e \in \mathbb{R}^+ \) can be determined. With that, the effective throughput can be computed as a useful combination of speed and accuracy:

\[ TP = \log_2\left(\frac{D_e}{W_e} + 1\right)/MT. \]

Wingrave and Bowman, as well as Murata and Iwase showed that Fitts’ Law also holds when pointing in VEs [46, 36]. They observed that \( D \) was related to the amplitude of the movement, and \( W \) to the visual size of the target. Poupyrev et al. [38] defined the size of an object \( W \) according to the vertical and horizontal angles that the object occupies in the user’s field of view (FOV). Similarly, Kopper et al. [24] propose a modification of Fitts’ Law as a model for human performance in distal pointing tasks. Their model is based on angular amplitude and angular width as they argue that, contrary to classic 2D Fitts’ Law tasks, the objects are floating in 3D space and the sizes and distances depend on the user’s position, which can be solved by using angular measurements. Ha and Woo [19] adopted Fitts’ Law for tangible AR environments with virtual hand metaphors by using the model established by Grossman et al. [17], which however was based on 3D objects arranged on a 2D plane.

3. JOINT-CENTERED INTERACTION

In this section we describe the motivation and rationale of joint-centered interaction, as well as the single components.

3.1 Kinespheres

Figure 1 illustrates the kinespheres at the maximal reachable distance for three joints typically involved in 3D mid-air interaction. According to the biomechanics and human factors literature, when assuming that a joint (e.g., wrist, elbow or shoulder) is rested and the remaining joints are stretched, the user’s fingertip will be approximately on a partial sphere centered on the joint [34, 39]. Finger positions close to a kinesphere are usually regarded as easily reachable and comfortable, whereas it is harder to reach the mid region of a kinesphere closer to the center joint. This fact has been utilized in previous research using joint angles and hyperplanes [18, 22].

3.2 Kinematic Chain

We refer to the term kinematic chain as a series of links connected by joints, or more specifically, a human’s arm from the shoulder joint to a fingertip. Previous studies have shown that interaction higher up the kinematic chain is less accurate and and causes higher energy expenditure if the corresponding muscle groups are used, and more accurate and less fatiguing further down [3, 39]. Optimally, a physical support at a joint as low as possible in the kinematic chain allows users to relax muscle groups higher in the kinematic chain, reducing exertion and increasing comfort. In contrast, traditional head-centered spatial menus are placed within arm’s reach in front of the user’s eyes, requiring movements of the entire kinematic chain.
implemented joint-centered 2D GUIs around the following joints:

circumscribed above. With respect to the kinematic chain we have
limitations by the maximal reachable distance and the biomechan-
ics described above. With respect to the kinematic chain we have
is limited by the maximal reachable distance and the biomechan-
ics described above. With respect to the kinematic chain we have
Thus, though placing menus slightly inwards from the boundary
chains are regarded as less comfortable than slightly relaxed joints.

For joint-centered menus we align the touch-enabled 2D GUI el-
elements or widgets at the maximal reachable distance, e.g.,
while ensuring visibility and low muscle exertion. Arranging 2D
GUI elements or widgets at the maximal reachable distance, e.g.,
from the shoulder, elbow or wrist, might thus provide advantages
in performance based on Fitts’ Law in 2D during the correction phase, using visual perception only to
correct for lateral offsets from the target.

2D GUI elements may be placed anywhere on the accessible part
of a kinesphere, i.e., on the opposite side from the user’s head,
while ensuring visibility and low muscle exertion. Arranging 2D
GUI elements or widgets at the maximal reachable distance, e.g.,
from the shoulder, elbow or wrist, might thus provide advantages
in performance based on Fitts’ Law over placing them anywhere
else within arm’s reach. Moreover, in contrast to pointing tech-
niques, this approach has the advantage of maintaining the super-
imposed perceptual and motor spaces as well as direct interaction
with the graphical elements, since they are located within arm’s
reach. However, we have to consider that fully-stretched kinematic
chains are regarded as less comfortable than slightly relaxed joints.
Thus, though placing menus slightly inwards from the boundary
may lose the advantages described above, it may lead to higher
user acceptance.

### 3.3 Boundary

The boundary of a joint-centered kinesphere defines a 2D man-
ifold in 3D space that, conceptually, reduces the interaction com-
plexity of touch gestures in terms of the degrees of freedom (DOF)
of arm movements from three to two since the finger’s distance
from the joint is clamped to its maximum and does not have to be
controlled consciously. Since large overshooting and undershoot-
ing are not possible as long as the corresponding kinematic chain
is stretched, selection performance can be described by Fitts’ Law
in 2D during the correction phase, using visual perception only to
correct for lateral offsets from the target.

1. Indication of target object by touching,
2. Confirmation of selection, e.g., with a touch or flip gesture
   with the dominant or non-dominant hand, voice command or
   timed response, and
3. Feedback generated by the system.

The joint-centered kinespheres described in this paper contribute
mainly to the first component. Interaction designers may combine
the joint-centered approach to indicate target objects with any of
the confirmation and feedback methods presented in the literature
as desired.

Overall, this describes the conceptual part of joint-centered ap-
proaches. We implemented the ideas using the Unity engine and a
PPT active IR optical tracking system. More information about the
hardware and the implementation is described in Sections 4 and 5.

### 4. EXPERIMENT

We conducted a Fitts’ Law experiment to evaluate selection per-
formance of GUI elements within arm’s reach focusing on the re-
lationships between joint, distance, performance, comfort and user
acceptance.

#### 4.1 Participants

20 participants (5 female and 15 male, ages 20 − 36, \(M = 26.1\))
took part in the experiment. The participants were members or
students of the local department of computer science, who ob-
tained class credit for their participation. All of our participants
had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the experi-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.

Participants had normal or corrected-to-normal vision. Seven participants wore
glasses and one participant wore contact lenses during the exper-
iment. None of our participants reported a disorder of equilib-
rium or a motor disorder such as an impaired hand-eye coordina-
tion. One of our participants reported a color blindness; no other
vision disorders have been reported.
and for the wrist joint ($M = 0.195\, \text{m}, \ SD = 0.027\, \text{m}$). We confirmed each participant’s ability to perceive binocular depth before the experiment via stereograms. 17 participants had prior experience with 3D stereoscopic display (cinema, games etc.). The total time per participant, including pre-questionnaires, instructions, experiment, breaks, post-questionnaires, and debriefing, was 1 hour. Participants wore the HMD for approximately 45 – 50 minutes. They were encouraged to take regular breaks between trials in order to rest their arm.

4.2 Material

As illustrated in Figure 2, participants were instructed to sit in an upright position facing towards the cameras of the optical tracking system. The experiment was conducted with the user wearing an Oculus Rift (Developer Kit 2) HMD with an attached active infrared (IR) target. Additionally, we attached an IR target to the index finger of the participant’s dominant hand (see Figure 2). The targets were tracked by an optical WorldViz Precision Position Tracking (PPT X4) system with sub-millimeter precision, which was chosen to offer maximal precision for the Fitts’ Law task. The visual stimulus displayed during the experiment showed a 3D scene, which was rendered with the Unity3D engine with an Intel computer with a Core i7 3.4 GHz CPU and an Nvidia GeForce GTX780TI. The Oculus Rift DK2 offers a nominal diagonal FOV of approximately 100 degrees at a resolution of 1920 × 1080 pixels (960 × 1080 for each eye).

4.3 Methods

The targets in the experiment were represented by spheres. All target spheres for one trial were always visible and colored white, except for the current target. The current target sphere was colored red when the marker was outside, and green when the marker was inside to give the participants a visual cue about selection performance. The spheres were highlighted in the order specified by the ISO 9241-9 standard for Fitts’ Law evaluations [21]. As illustrated in the inset of Figure 2, each trial consisted of an arrangement of 7 target spheres, forming a circle with each sphere at the same distance to the participant’s joint.

In the experiment we used a within-subjects repeated measures 3 (joints) × 4 (distances) × 4 (index of difficulties) × 3 (repetitions) full-factorial design. As described before, we considered the shoulder, elbow and wrist joints. As Figure 3 illustrates, we considered three distances (with $j$ being the distance between the joint and the finger in a comfortably stretched pose: $0.618 \times j \equiv$ short, $0.854 \times j \equiv$ medium, $1.0 \times j \equiv$ long) in the experiment, which were chosen using the knowledge that the distances between a human’s joints can be approximated with the golden ratio [34]. The fourth distance condition denotes the boundary technique. When touching an object at the maximal reachable distance it might be advantageous in practical scenarios to include a tolerance range to account for slight overshooting or undershooting in depth (cf. Section 2), hence in this technique we considered a range stretching from $[0.854, 1] \times j$. We evaluated the ecologically viable IDs 2, 2.5, 3 and 3.5. All conditions were tested three times and their order was fully randomized. We included 38 training trial (one for each condition), which were excluded from the analysis.

Before the experiment, all participants filled out an informed consent form and received instructions on how to perform the task. Furthermore, they filled out the simulator sickness questionnaire (SSQ) [23] immediately before and after the experiment, the Slater-Usoh-Sted (SUS) presence questionnaire [42], and a demographic questionnaire. We further observed the behavior of the participants during the experiment, and debriefed the participants after the experiment.

4.4 Results

We analyzed the results with repeated-measure ANOVAs and multiple pairwise comparisons with Bonferroni’s correction at the 5% significance level. We confirmed the assumptions of the ANOVA for the experiment data. A Shapiro-Wilk test did not indicate that the assumption of normality had been violated. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly’s test indicated that the assumption of sphericity had been violated.

4.4.1 Selection Performance

Figure 4(a) shows the mean time elapsed until a participant selected a target object for each condition in the experiment. Table 1 shows the results of the statistical analysis.

Figure 4(b) shows the mean error distances, i.e., the Euclidean distance between the participant’s fingertip and the center of the target sphere when a participant selected the target object for each condition in the experiment. Table 2 shows the results of the statistical analysis.

Figure 4(c) show the mean error rate between the times when the participant’s fingertip was within or outside of the target sphere when a participant selected the target object for each condition in the experiment. Table 3 shows the results of the statistical analysis.

Figure 4(d) shows the mean effective throughput for each condition in the experiment. Table 4 shows the results of the statistical analysis.

4.4.2 Questionnaires

Asked to rate their level of comfort during the joint conditions
Joint | M | SD | F | p | η²
--- | --- | --- | --- | --- | ---
wrist | 1093ms | 202ms | (2.38) | < .001 | .26
elbow | 1171ms | 203ms | 6.56 | .18
shoulder | 1146ms | 195ms | .79 |

Post-hoc | p
wrist–elbow | < .001 |
wrist–shoulder | .18 |
elbow–shoulder | .79 |

distance | M | SD | F | p | η²
--- | --- | --- | --- | --- | ---
short | 1214ms | 193ms | (3.57) | < .001 | .67
medium | 1155ms | 172ms | 38.30 | < .001 | .95
long | 1146ms | 236ms | .59 | .59 |
boundary | 1030ms | 186ms | .95 |

Post-hoc | p
medium–long | < .05 |
others | 5.99 |

Table 1: Statistical analysis for the movement time

Figure 4: Results for Fitts' Law trials with the joint on the horizontal axis and pooled results on the vertical axis for (a) movement times, (b) error distances, (c) error rate and (d) effective throughput. The error bars show the standard error. Red bars show the results for the short distance, yellow for medium distance, green for long distance, and blue for the boundary technique.

Figure 5: The results showing the subjective comfort ratings (x), and the joint and distance condition (y).

with 5-point Likert scales (0 = very low, 4 = very high) (see Figure 5).

We further asked the participants for each joint condition
which distance they preferred with the alternatives boundary, long, medium and short (see Figure 6). The responses show that most participants preferred the boundary technique and medium distance, followed by the long distance, but not the short distance.

We measured a mean SSQ score of 15.15 ($SD = 17.77$) before and 26.18 ($SD = 29.67$) after the experiment. The results indicate a typical increase in simulator sickness symptoms with an HMD over the time of the experiment. The mean SUS score for the sense of feeling present in the VE was 4.51 ($SD = .97$), which indicates a high sense of presence [42].

We further collected informal comments during the debriefing after the experiment. Notably, one participant explained the preference of the wrist condition by, “My arm was too heavy for keeping it up in the air with the shoulder-based interface.” Another participant stated, “I liked the wrist condition; similar to a mouse.”

### 4.5 Discussion

Overall, the results show that the interaction performance highly depended on the placement of targets within arm’s reach. In the following, we discuss the findings focused on the kinematic chain and kinespheres and summarize the main ones as practical guidelines.

#### 4.5.1 Kinematic Chain

We found the significantly highest effective throughput for the wrist joint condition, whereas we found lower effective throughput for the elbow and shoulder conditions, between which the results showed no significant difference. There are multiple possible explanations of this effect. First, from a biomechanical point of view a longer kinematic chain leads to more weight and inertia while requiring more muscle exertion and might thus have caused longer movement times [34, 39]. Moreover, the presence of a physical support by means of an elbow or a wrist cushion might have reduced fatigue and hand tremors by shortening the kinematic chain [18, 30]. Additionally, we have to consider that the wrist condition was similar to using a mouse and might have benefited from this similarity.

#### 4.5.2 Placement in Kinespheres

As expected, the practically-inspired boundary technique with the short tolerance range near the maximal reachable distance showed the overall significantly lowest errors, shortest movement times, and consequently highest effective throughput. This result might be explained by the reduction from 3 DOF to 2 DOF due to the stretched-out kinematic chains and short tolerance range, which additionally eliminated errors along the depth axis over the similar long distance condition.

While we expected an overall high performance for the long distance condition second only to the boundary technique due to the matching stretched-out kinematic chains, our results instead show performance that is very similar to the medium distance condition. Hence, we further analyzed the distribution of errors in the joint-centered kinespheres. With stretched-out limbs we observed mean selection errors of $-2.97mm$ ($SD = 1.43mm$) for the shoulder condition, $-2.97mm$ ($SD = 1.43mm$) for the elbow condition, and $-2.83mm$ ($SD = 1.40mm$) for the wrist condition as measured from the center of the kinesphere to the target. The negative values might be explained by the participants calibrating their hand position with a truly outstretched arm, while they later drew...
their arm back a few millimeters into a more comfortable pose. Without this systematic undershoot in depth it appears that the long distance condition would have, indeed, resulted in higher performance. However, most participants indicated the boundary technique or the medium distance condition as their preference while estimating the long distance (without boundary) and the short distance as least comfortable.

4.5.3 Guidelines

We found the highest performance for the wrist joint using the boundary technique, which we recommend for practitioners in the field of spatial user interfaces. However, overall performance and user acceptance also showed that placing menus at the medium tested distance is an ecologically viable alternative. We do not recommend placing menus near the long or short tested distances. We found similar performance for the elbow and shoulder conditions, but user preference based on estimated comfort clearly shows that the shoulder joint should be ruled out in practical applications.

5. CONFIRMATORY STUDY

In this section we present a confirmatory study in which we compared a traditional head-centered menu with a joint-centered menu that we designed based on the results of the first experiment. The menus were compared in a practical interior design application. While direct interaction is a great choice for many applications in spatial user interfaces, interior design is merely one of them. The concept of joint-centered user interfaces is applicable for any applications with context changes and tasks which require menus and which can benefit from direct interaction. These could be applications involving planning, construction, design, entertainment or smart environments. In this case, the example of interior design was picked because it offers many opportunities to offer the user a selection of multiple items.

5.1 Participants

20 participants (2 female, 18 male, ages 19–44, M = 27.6) took part in the experiment. The participants were students or experts from computer science or human-computer interaction, and one journalist. Most participants had normal or corrected-to-normal vision (10 no correction, 7 glasses, 3 lenses, 1 with strong eye dominance, 1 with dyschromatopsia (red-green color weakness) and strong eye dominance and astigmatism). 14 of the participants took part in a study with an HMD before. All participants were right-handed. 18 participants reported at least some experience with computer games. The total time per participant was approximately 30 minutes.

5.2 Material

The experiment setup is illustrated in Figure 7. Participants were seated in an MWE Lab Emperor Chair 1510 and wore an Oculus Rift CV HMD on their head, IR markers on their wrist and elbow and a haptic ring input device on their index finger (similar to [2]). The IR markers and the ring were tracked in 3 DOF using an optical WorldViz PPT X4 tracking system with sub-millimeter precision. Additionally, the ring includes an inertial measurement unit, which allows for 6 DOF tracking of the user’s index finger. The living room shown in the virtual scene was rendered in Unity3D on an Intel computer with a Core i7 4 GHz CPU, 16 GB RAM and an Nvidia GeForce GTX 980. The Oculus Rift CV offers a resolution of 1200×1080 per eye and approximately 94H×93V degrees FOV for both eyes. The in-house implemented interior design application is illustrated in Figure 8.

5.3 Methods

Half of the participants started with the joint-centered user interface (JCUI) and afterwards completed the second part of the experiment with the head-centered user interface (HCUI), and vice versa. The tasks for the participants were, first, to familiarize themselves with the spatial user interface, toggling the light by pointing at it and pressing the button on the input device, then moving a table by pointing at it, pressing a button and then pointing to the target location and pressing the button to confirm the position. Afterwards, the participants were instructed to change the wall, ceiling and floor materials by pointing at them, respectively, and then touching

<table>
<thead>
<tr>
<th>joint</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>p</th>
<th>ηp</th>
</tr>
</thead>
<tbody>
<tr>
<td>wrist</td>
<td>1.21bps</td>
<td>.38bps</td>
<td>1.4 (26.51)</td>
<td>&lt;.001</td>
<td>.57</td>
</tr>
<tr>
<td>elbow</td>
<td>1.89bps</td>
<td>.29bps</td>
<td>25.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>shoulder</td>
<td>1.94bps</td>
<td>.26bps</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>distance</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>p</th>
<th>ηp</th>
</tr>
</thead>
<tbody>
<tr>
<td>short</td>
<td>1.85bps</td>
<td>.21bps</td>
<td>1.5 (28.57)</td>
<td>&lt;.001</td>
<td>.39</td>
</tr>
<tr>
<td>medium</td>
<td>1.95bps</td>
<td>.25bps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>long</td>
<td>1.94bps</td>
<td>.24bps</td>
<td>12.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>boundary</td>
<td>2.12bps</td>
<td>.39bps</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>p</th>
<th>ηp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.05bps</td>
<td>.32bps</td>
<td>1.98 (37.61)</td>
<td>&lt;.001</td>
<td>.4</td>
</tr>
<tr>
<td>2.5</td>
<td>1.95bps</td>
<td>.24bps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.98bps</td>
<td>.29bps</td>
<td>12.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>1.87bps</td>
<td>.28bps</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Statistical analysis for the throughput

Figure 7: Illustration of the confirmatory study setup.

Figure 8: The virtual room for the confirmatory study.
one of the appearing menu buttons of their choice. The appearing menus were either positioned centered around the head, in the head-centered condition, or one of the three joints, in the joint-centered condition, depending on whether the participants were resting their wrist, elbow, or nothing (shoulder) as shown in Figure 9. Next, the participants had to move the couch and then change its color. For this they had to point at the couch, press a button, touch a context button (move or color), and then either move it or color it as described above. We positioned the menus and determined input based on the boundary technique described in Section 4.3.

After the experiment, the participants had to complete different questionnaires comparing the conditions. As a measure of usability, we used the simple usability scale [7]. Additionally the participants had to complete a NASA TLX [20] and a Borg15 Scale [4] to measure the task difficulty and physical exertion. Afterwards, the participants were asked to provide subjective feedback and answer questions regarding the demographics in a questionnaire.

### 5.4 Results

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>M</th>
<th>SD</th>
<th>t(19)</th>
<th>p</th>
<th>dCohen</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUS JCUI</td>
<td>73.625</td>
<td>18.308</td>
<td>.373</td>
<td>.713</td>
<td>.373</td>
</tr>
<tr>
<td>SUS HCUI</td>
<td>71.75</td>
<td>15.822</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLX JCUI</td>
<td>33.16</td>
<td>15.740</td>
<td>-1.538</td>
<td>.14</td>
<td>-.344</td>
</tr>
<tr>
<td>TLX HCUI</td>
<td>40.91</td>
<td>18.854</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borg15 JCUI</td>
<td>9.4</td>
<td>1.522</td>
<td>-2.463</td>
<td>&lt;.05</td>
<td>-.551</td>
</tr>
<tr>
<td>Borg15 HCUI</td>
<td>9.4</td>
<td>2.458</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Statistical analysis for the confirmatory study

The results were analyzed with a paired-samples t-test for each of the three main questionnaires and are shown in Figure 10 and Table 5.

Additionally, we analyzed the subjective feedback to the questions whether participants liked the joint-centered approach (15 positive), whether they liked head-centered interfaces (12 positive), whether joint-centered interfaces could be applied in other domains also (16 positive), which technique was less tiring (15 joint-centered), and which technique was more fun (11 joint-centered, 5 indifferent).

### 5.5 Discussion

Overall, the feedback was very positive concerning the joint-centered menu, and the results from the Borg15 Scale show that it is estimated as significantly less tiring than head-centered menus. Our results did not show significant differences for the SUS and NASA TLX usability questionnaires.

#### 5.5.1 Improvements

We received qualitative feedback and suggestions for improvement, which we plan to incorporate in the next iteration of the design. For instance, the tracking of the different joints with IR markers occasionally caused an unintended switch of the interface to another joint, which we aim to resolve using improved skeletal sensors. Moreover, our participants remarked that the level of fatigue during the few minutes with the interface in the experiment was still tolerable using either head-centered or joint-centered menus, but they estimated that their preference would be clearly shifted towards joint-centered interfaces during longer use. Also, as two participants commented, the joint-centered menu is positioned around the user’s arm, which is not necessarily always within the user’s FOV, but due to the matching perceptual and motor space it might be possible to interact even without vision, potentially supported by (vibro-)tactile feedback (cf. [2]).

#### 5.5.2 Implications

Our results imply multiple advantages of presenting 2D GUIs on joint-centered kinespheres instead of traditional head-centered GUIs:

- Physical support, if available, can be used to rest joints without the need to lift the arm up for GUI interaction.
- Users can move the GUI closer to their eyes if need be to increase accuracy and precision during interaction [29, 30].
- Matching joint-centered perceptual and motor spaces support motor training and may even be used for GUI selections without vision for experienced users.
- Joint-centered menus located at the maximal reachable distance ensure that false positives of unintended selections, when moving the hands casually within arm’s reach, are highly unlikely.
- Using the boundary technique that combines the positioning of menus on kinespheres with a small tolerance area can significantly improve interaction performance.

### 6. CONCLUSION

In this paper we presented and evaluated joint-centered kinespheres for efficient and comfortable spatial interaction in virtual and augmented reality. We performed a Fitts’ Law experiment, which showed that the interaction performance largely depends on the placement of menus along the kinematic chain and within the kinespheres centered around the shoulder, elbow or wrist joint. In particular, the effective throughput was highest for the wrist joint and menus that were placed on the boundary of the kinesphere with a small tolerance region for touch input. Based on the results, we
implemented a joint-centered user interface and compared it to a traditional head-centered user interface in an in-house developed interior planning application. We discussed our observations and presented practical guidelines for how to leverage joint-centered kinespheres for the design of menus for spatial user interfaces.

There are several possible paths to extend and to adapt the method of joint-centered spatial user interfaces. For instance, kinespheres could be applied at finger level, but also at the legs, trunk or back to incorporate further modalities. Moreover, placing interactive objects on joint-centered kinespheres may require a new class of optimized widgets or layered menu structures. Furthermore, since we observed both high efficiency and high comfort in the wrist boundary condition in our experiment, following up on this path may lead towards spatial user interfaces for productive long-time use. While joint-centered user interfaces provide an efficient solution for long term direct interaction, future work could offer a comparison with indirect interaction methods to place the joint-centered user interfaces in the wide field of spatial user interfaces.

Acknowledgments
This work was partly supported by grants from the Deutsche Forschungsgemeinschaft.

7. REFERENCES


