Modeling Cumulative Arm Fatigue on Large Multi-touch Displays

by

Zhe Liu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Public Health and Health Systems

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisor(s): James Wallace
Associate Professor, School of Public Health and Health System
University of Waterloo

Daniel Vogel
Associate Professor, Cheriton School of Computer Science
University of Waterloo

Internal Member: Philip Bigelow
Associate Professor, School of Public Health and Health System
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis includes first-authored peer-reviewed material that has appeared in conference proceedings published by the Association for Computing Machinery (ACM).

The materials from which I have adapted content are the following:

- Liu Zhe, Daniel Vogel, and James R. Wallace. "Applying the cumulative fatigue model to interaction on large, multi-touch displays." *Proceedings of the 7th ACM International Symposium on Pervasive Displays.* ACM, 2018.

Abstract

Large multi-touch displays have long been studied in the lab, and are beginning to see widespread deployment in public spaces. Although they are technologically feasible, research has found that large multi-touch displays are not always used, and that fatigue is commonly identified as a significant barrier. Fatigue, often called the 'gorilla arm' effect, prevents people from using large displays for extended periods of time. One solution to this problem is to design large-scale interfaces that can minimize actual fatigue in practice. A first step towards building such an interface is to quantify fatigue, and more importantly, to quantify it easily. While there have been methods developed to estimate arm fatigue in mid-air interaction, there remains little understanding of fatigue on touch-based interfaces. To address this gap, we propose that existing models for mid-air interaction may be effective for measuring fatigue on large multi-touch displays. We evaluated the accuracy of Jang et al.'s mid-air Cumulative Fatigue model for touch interaction tasks on a large display. We found that their model underestimates subjective fatigue for multi-touch interaction, but can provide accurate estimates of subjective fatigue after fine-tuning of model parameters. We discuss the implications of this finding, and the need to further develop tools to evaluate fatigue on large, multi-touch displays.

Acknowledgements

I would first like to thank my supervisors, Prof. Wallace, and Prof. Vogel. Their offices were always open whenever I had any questions about my research. They consistently led me in the right direction.

I would also like to acknowledge Prof. Bigelow, from the School of Public Health and Health System at the University of Waterloo, as the third reader of this thesis. I am gratefully for his very valuable comments on this thesis.

In addition, I would like to thank the Natural Sciences and Research Council of Canada for funding this research. I would also like to thank Sujin Jang and Wolfgang Steurzlinger for providing access to and supporting the use of their existing software, in-depth descriptions of their experimental protocol, and discussions about their CF model.

Finally, I must express my gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without them. Thank you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Overview

Large interactive displays, usually known as wall display, have been envisioned by the Human-Computer Interaction (HCI) community as a platform for ubiquitous interaction for long (e.g., [3, 4, 5]). The natural and intuitive interaction offered by multi-touch displays, combined with their attractive and inviting large-format, is well-suited to engage non-technical users. They serve as an ideal place for multiple users to gather around and share and explore data collaboratively. For these reasons, large multi-touch displays have been installed in schools [6], offices [7, 8], and public spaces [7, 9, 10].

However, as technologies have advanced and large displays become more capable, researchers have found barriers to interacting with large multi-touch displays. These include privacy concerns, feelings of social embarrassment[11, 12, 9], and more importantly, the fatigue and its consequences on user experience. Working on multi-touch displays for extended periods of time leads to arm fatigue, a phenomenon that occurs so frequently that it is referred to by the research community as the 'gorilla arm' effect [13]. Wang and Ren note that arm fatigue is the main drawback of multi-touch interaction, and that the larger the display, the more fatiguing interaction is on the user [14]. Given the prominence of the problem surrounding arm fatigue and the push towards bigger displays, there is a need to understand how large-scale interfaces can be designed to minimize fatigue.

Arm fatigue, as a specific type of muscle fatigue, not only affects comfort perception of users, but also their completion of operations and even their health condition, which may lead to serious consequences. Several studies identified a positive correlation between muscle fatigue and muscle injury and damage [15, 16]. Grujicic et al. [17] particularly researched the musculoskeletal fatigue during long-distance driving, indicating the safety issue caused by muscle fatigue. These results all lead to a demand for understanding arm fatigue and minimizing it.

Existing studies showed the negative effect of fatigue caused by large-scale interfaces. One specific workplace study conducted by Leme and Maia [18] aimed to evaluate the fatigue of teachers in the modern classroom setting. Modern classrooms are normally equipped with state-of-the-art resources, including projectors, interactive whiteboards, and computers. Teachers usually need to work with these devices for extended periods of time throughout the school day. Teachers who participated in this study expressed that they become fatigued due to their body posture and the utilized resources. The fatigue experienced can be intense for specific cases, including both visual fatigue and muscle fatigue in the neck and shoulder. More precisely, 29 % of the participants reported tiredness and visual fatigue, while over 14 % complained about the pain in the neck and shoulder muscles, thighs and legs. Such facts reinforce the importance of solving the arm fatigue problem before large displays become applied to a wider range of workplaces.

To minimize arm fatigue, one must first be able to measure it reliably. As Jang et al. [1] noted in their review of the literature, more than 300 research papers within the HCI community cite fatigue as an issue, and yet we still lack means of accurately measuring it. Traditionally, measuring arm fatigue is done either objectively with intrusive, expensive, and specialized test rigs [19, 20, 21], or subjectively using a fatigue self-report evaluation such as the Borg Scale [22]. Self-reporting is shown to be accurate for most purposes [21], but can be invasive since it requires frequent interruptions during the measured activity. As a middle ground, researchers have developed reliable fatigue models to measure mid-air arm fatigue using non-invasive whole-body tracking with a single consumer-level depth camera. The Consumed Endurance (CE) model from Hincapié-Ramos et al. is based on shoulder torque [23], and the Cumulative Fatigue (CF) model from Jang et al. includes the elbow as well. The CE model has already been used to justify and evaluate new mid-air gestures [24]. However, neither the CE nor CF models were explicitly designed for measuring fatigue on large multi-touch displays.

We expected that, models designed for mid-air interaction may also perform well for measuring fatigue on large multi-touch displays. Although large multi-touch interfaces contain a two-dimensional (2D) surface for the users to interact with, while mid-air interface allows movement in three-dimensional (3D) space [25], the basic operations for both interfaces are similar. The arm motions of both mid-air and large-display multi-touch interactions are similar. For both mid-air and large-display multi-touch interactions, users primarily operate with their dominant hand for precise actions [26], with the non-dominant hand to support more complicated operations using two-hand gestures. Furthermore, during interaction, users lift their arms and move their hands in space. Since mid-air fatigue models rely on bio-mechanical models and shoulder torque, we hypothesize that they can be applied towards large-display multi-touch interaction as well. Specifically, we predict that, the CF model proposed by Jang et al. can estimate cumulative arm fatigue for large multi-touch displays with little or no adjustments.

To validate this hypothesis, we carried out an experiment using a similar protocol to Jang et al., and evaluate the CF model's performance in estimating arm fatigue for large multi-touch displays. Our work extends the study protocol from Jang et al. to multi-touch tapping and dragging tasks over short and long distances. Using this protocol, we gather Borg CR10 Scale ratings and whole-body tracking data in a 24-participant experiment to quantitatively. From the results, we found the CF model underestimates subjective fatigue for multi-touch interaction, but can provide accurate estimates of subjective fatigue after fine-tuning of model parameters with a higher upper bound.

Our contributions include: (1) empirical validation of the CF model for estimating subjective arm fatigue in multi-touch interaction on large displays; (2) demonstrating that the CF model underestimates fatigue for multi-touch tasks, but shows high accuracy during rest periods; and (3) that though fine-tuning of model parameters, the CF model can accurately predict arm fatigue for touch interaction. Based on our investigation, we discuss opportunities to further refine measures of fatigue through improvements to the CF model.

Chapter 2

Literature Review

While the mouse and keyboard have evolved over decades of use and are highly efficient input devices, large multi-touch displays provide a relatively new and evolving means of interaction. Recent technological developments make it possible for multi-touch displays to be larger and less expensive, which promotes their real-world deployment, ranging from indoor to outdoor settings, from office and work contexts to public spaces and communities. As the HCI community has explored multi-touch interaction, numerous benefits have been identified. For example, it is considered 'intuitive' and 'natural' [27], and thus is accessible to a wide range of users without any barriers to learning. Touch input on large displays also allows a group of users to collaborate seamlessly around shared content and has been shown to support communication and workspace awareness [28, 29]. A variety of information can be presented simultaneously for collaborators and audiences to review, share opinions, modify, and leave comments if needed. These features explain the surge in usage of multi-touch displays in a wide variety of settings for different purposes.

Multi-touch input, however, is not always a perfectly satisfying user experience, and several issues need to be resolved before the a larger-scale adoption of large multi-touch displays can occur. Challenges that currently exist for public displays include attracting participation from potential users, dissolving users' social embarrassment, which refers to discomfort when witnessed by others, and protecting users' interaction and information privacy [30, 31, 32]. For more formal environments, such as office and work contexts, where users need to operate with a large multi-touch display for extended periods of time, an additional common limitation is arm fatigue.

Arm fatigue remains a challenge when operating large, multi-touch displays [33, 34, 35, 36]. In practice, fatigue is frequently cited as a drawback of multi-touch interaction, particularly over extended periods of use on large and vertical displays [37, 14, 13]. We know from previous research that fatigue is universal and applies to both horizontal (i.e., tabletop) and vertical displays [38].

Many research projects were conducted to address fatigue on large displays. For instance, researchers have explored how different display configurations, such as their height and angle, can influence force exertion over extended time during multi-touch interaction [39, 38]. Similarly, means of offloading interaction to nearby devices like smartphones [40, 41] and smartwatches [42], or even to the users' body [43, 44] have been developed. Users may interact with the elements duplicated on the smartphones or smartwatches [40, 42] or projected to their hands or arms as a portable touchpad [43] to control a connected large display. However, in practice, all of these solutions can be problematic in some ways. Adjusting the angle of the display to minimize arm fatigue is not ideal in collaborative scenarios because settings that decrease fatigue of one user may increase fatigue of another. Moreover, vertical displays are well-suited to sharing information within a group, and offloading interactions from a shared display to nearby devices, will hide that user's interactions from others, thus negatively impacting group awareness [45]. Thus, there remains a need to better understand and design for fatigue on conventional large displays.

Finally, while many researchers have designed interaction techniques to address fatigue, such as, At-your-side gestures like Gunslinger [46, 47], and smart devices like E-Pad [48], few have directly compared their proposals with alternative designs proposed by other researchers. We argue that the lack of such studies arises from the shortage of accurate, inexpensive, and easily implementable fatigue evaluation methods.

Here we review prior work related to objective and subjective evaluation of muscle fatigue, models to estimate cumulative fatigue, fundamental arm motion detection technology required by the models.

2.1 Evaluating Arm Fatigue

Existing measures for evaluating fatigue can be classified as either subjective or objective. Subjective measures are most likely to be familiar to the HCI community. For example, the NASA Task Load Index (NASA-TLX) in Figure A.1 [2] is frequently used in HCI, and its 'physical demand' dimension partly encapsulates fatigue. The Borg CR10 Rating of Perceived Exertion (RPE) [22] has also been used, particularly to understand fatigue during mid-air pointing (e.g., [1]), and has been empirically shown to strongly correlate with objective measures [49, 50, 51]. The Borg CR10 Scale has been shown to be best suited when there is an overriding sensation arising from a specific area of the body, for example, muscle fatigue. The CR10 Scale is designed as linear and interval with 0.5-length step, which makes it easy for people to understand in a short time.

Subjective measures have been used in the past because they are lightweight and easily included in most evaluations. They require participants to rate their experience on a single scale (or, for NASA-TLX, six scales) and are quickly and easily collected at the end of a trial. However, they are also limited in significant ways. First, they are subjective, and interpretation of a rating on the scale may vary significantly among users. Second, they cannot, by design, provide real-time feedback on users' fatigue since they require users to stop what they are doing and give their rating. In this situation, participants' ratings may be different from the true rating when they are not interrupted.

Objective measures of fatigue are typically collected by calculating an individual's maximal strength and the proportion of that strength consumed during an activity [1]. Such quantification always requires measuring changes in physiological markers, such as muscle activation [52], blood pressure [53], or heart rate [54]. While highly accurate and capable of providing fatigue measurements in real-time, physiological sensors can be expensive and are often impractical for HCI research [19, 20], since the measurements can be complicated or require a specific environment.

Overall, subjective measures are generally straightforward to collect, whereas objective measures can be gathered in real-time, are very accurate and easy to replicate. In search of measures that combine the benefits of both subjective and objective measures, the research community began to explore how mathematical models can be used to evaluate fatigue.

2.2 Modelling Arm Fatigue

While we are unaware of work that has specifically modeled fatigue for multi-touch displays, a number of models exist for mid-air interaction. For example, a popular approach for modeling mid-air fatigue is to model Endurance Time (ET), with several models including Rohmert's curve [55], the original Three Compartment Muscle (TCM) model [56, 57], and the TCM's more recent version with developments by Ma et al. [58, 59]. These models show potential in predicting fatigue for either static or dynamic load conditions, but lack experimental validation, particularly for settings involving human interaction, which makes them not that suitable for our study. In this work, we consider two recent models that have been developed and validated with HCI contexts in mind: The Consumed Endurance (CE) model [23], and the Cumulative Fatigue (CF) model.

Consumed Endurance (CE) [23] is a novel metric derived from the bio-mechanical structure of the upper arm, specifically modeling the torque on a user's shoulder joint during mid-air gestures. In this model, two major equations are used:

$$E(T_{shoulder}) = \frac{a}{(T_{shoulder}/T_{max} \times 100 - 15)^b} - c$$

$$CE(T, TotalTime) = \frac{TotalTime}{E(T_{shoulder})} \times 100$$

where

$T_{shoulder}$	=	the current shoulder torque
T_{max}	=	the maximum shoulder torque
$TotalTime$	=	the total time consumed by the current task
$E(T_{shoulder})$	=	the estimated endurance time
a, b, c	=	parameters calculated from the data

A specific benefit of this model is that it is easily calculated using readily available data from body tracking hardware, such as a Microsoft Kinect. In the validation of their approach, Hincapié-Ramos et al. showed a strong correlation between the model's prediction and subjective fatigue (i.e., Borg CR10 scale). However, the CE model is limited to static load conditions by design and does not take into account periods of rest, such as those between mid-air gestures when the users have their arms at their side. Another limitation is its assumption that users never exceed an exertion level lower than 15%. This assumption leads to inaccurate estimations of fatigue for interactions with lower exertion levels. These limitations were specifically addressed by Jang et al. and solved in developing their CF model.

Jang et al. extended CE to develop a Cumulative Fatigue (CF) model for mid-air interaction. Their model is the first to quantify cumulative arm fatigue in recovery-involved tasks, such as those where users may rest between gestures. Like the CE model, the CF model also relies on camera-based skeleton tracking via a device like Microsoft Kinect. However, the CF model includes the users' elbow joint and an additional degree of freedom for their shoulder joint. Jang et al. validated their model using a continuous mid-air pointing task that included rest periods of varying lengths. Their evaluation found a strong match between the CF model's estimation and subjective measures of arm fatigue using the Borg CR10 Scale. They also note that their model performed well for a range of exertion levels, outperforming the CE metric (including low exertion level under 15 %).

2.3 CF Model Quantifying Method

The CF model is built upon biomechanical arm model and the concept of three-compartment muscle (TCM) model.

From previous research, we know that the fatigue can be accurately modeled as a joint specific phenomenon [23, 1]. When it comes to arm movement, a person's shoulder fatigues faster than their elbow or wrist. Therefore, we assumed that arm fatigue is mostly attributable to shoulder-joint fatigue. To measure the real-time shoulder joint torque, we decide to build a biomechanical arm model, modeling the upper limb as rigid bodies (links) connected in series by joints. The biomechanical arm model is shown in Figure 2.1. With this biomechanical model and accurate measurements of each link's length and weight, we are able to calculate shoulder torque at any moment, which will be used in the TCM model.

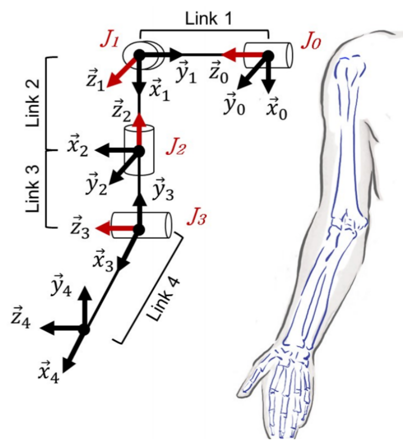


Figure 2.1: Biomechanic model of the upper limb [1]

In the TCM model, every motor unit or muscle involved in a task is in one of the three possible states: Active, Fatigue or Rest. Active motor units represent those units receiving neural activation and are contributing to the task. Fatigue units include fatigued motor units without activation. Rest motor units are inactive motor units not required for the task. The relations between these three states are shown in Figure 2.2. Here M_A, M_F, M_R, are the proportion of motor units that are currently in active, fatigued and rest state, respectively. Each of these proportions is expressed as a percentage of the maximum voluntary contraction (%MVC). For example, MA = 100% indicates that all motor units are recruited for an MVC task. A sub-maximal task implies MA $\leq$ 100%.

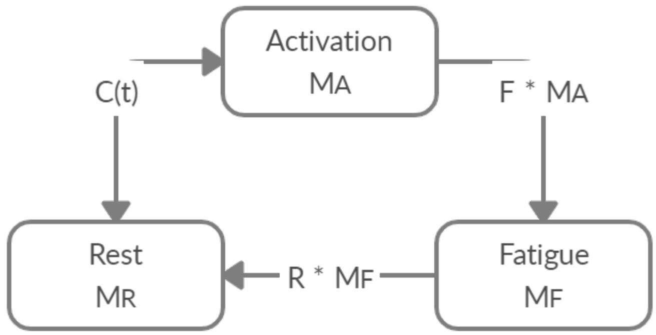


Figure 2.2: The Three Compartment Muscle (TCM) model

The transfer functions between the three motor unit states are defined as below.

$$\frac{dM_R}{dt} = -C(t) + R \times M_F$$
$$\frac{dM_A}{dt} = C(t) - F \times M_A$$
$$\frac{dM_F}{dt} = F \times M_A - R \times M_F$$

Here. F and R are the model parameters for FATIGUE RATE and REST RATE. F is the rate at which active motor units are fatiguing, and R is the rate at which fatigued motor units recover and enter the rest state. These two parameters will be identified in the model fitting stage. $C(t)$ is motor unit activation function, which is defined as:

$$C(t) = \begin{cases} L_D \times (T_L - M_A) & \text{if } M_A < T_L, M_R > T_L - M_A \\ L_D \times M_R & \text{if } M_A < T_L, M_R \leq T_L - M_A \\ L_R \times T_L - M_A & \text{if } M_A \geq T_L \end{cases}$$

where T_L is the target load defined as a torque ratio $[T_{current}/T_{max}]100(\%)$, L_D is the muscle force development factor, and L_R is the relaxation factor. The last two parameters are set to 10 based on the sensitivity analysis by Frey-Law et al. [37].

To identify the optimial value for F and R, the pattern search method was used in the model fitting:

$$\begin{aligned} \underset{\mathbf{F},\ \mathbf{R}}{maximize} \quad & \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\phi(M_F(i)) - B(i)\right]^2} \\ \text{subject to} \quad & F \in [F_{lb}, F_{ub}] \\ & R \in [R_{lb}, R_{ub}] \end{aligned}$$

where n is the number of fitting data, $M_F(i)$ is the fatigue level estimation, $B(i)$ is the Borg CR10 scale rating, and the upper and lower bounds of the parameters are defined as $F_{lb}, F_{ub}, R_{lb}, R_{ub}$. We assume a linear relationship between the Borg CR10 scale and %MVC (proportion of current and maximum torque) based on a review of their relationship [60]. $\phi(x)$ is a linear function mapping the fatigue estimation $M_F(i)$ to the Borg CR10 scale. We define this linear mapping as:

$$\phi(x) = 0.0875 \times x$$

Although the CE metric and the CF model have both been validated for their performance in estimating arm fatigue for mid-air interaction, there are no well-validated models for arm fatigue estimation of interaction with large, multi-touch displays. There are a number of similarities between mid-air and large-display interactions that lead us to believe that the CF model could be applicable. In both cases, users must hold and move their arms in front of their body, while exerting force from their shoulders. Yet, a significant difference is that, in multi-touch settings, users are exerting a force on the screen, which can add friction, provide support, and influence how users perceive fatigue. Therefore, we can hypothesize that, when users interact with large multi-touch displays, their arm fatigue may also be predicted with a CF-like model, while minor adjustments may be required to achieve a more accurate estimation.

2.4 Arm Motion Detection

Arm motion can be analyzed and reproduced with data about joints including the shoulder, the elbow, and the wrist. To collect data on the involved joints, body-tracking sensors need to be integrated into the system. One popular body-tracking sensor is the Microsoft Kinect, which is used by both CE and CF models.

A Microsoft Kinect [61] tracks body motion at a frame rate of 30 Hz; in other words, for each second, information on the arm motion of the user is captured 30 times. Equipped

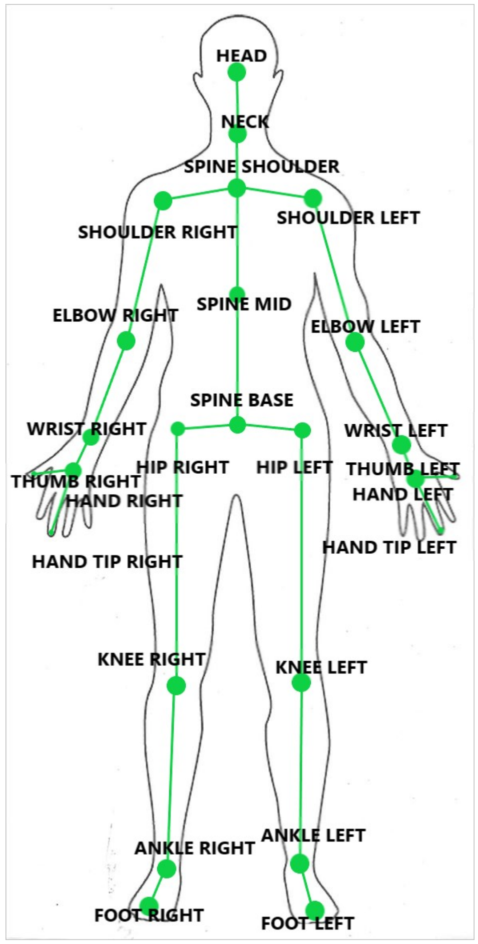


Figure 2.3: Kinect v2 25-joint skeleton

with an RGB color camera and an IR depth camera, the Kinect can track up to six skeletons at one time, each of which having 25 joints, as shown in Figure 2.3 [62]. Each joint has 11 properties: color (x, y); depth (x, y) ; camera (x, y, z) ; and orientation (x, y, z, w). The meaning of each property is explained in the Table 2.1. With this data point, a user's motion can be continuously tracked, stored, and later, reproduced and analyzed for specific attributes, including continuous arm fatigue conditions.

Joint property	**Meaning**
color (x, y)	Coordinates of the joint on the image from the color camera
depth (x, y)	Coordinates of the joint on the image from the depth camera
camera (x, y, z)	The 3D points of the joints in space from Kinects infrared sensor
orientation (x, y, z)	Representing yaw, pitch and roll
orientation (w)	w serves as a rotation represented as a quaternion

Table 2.1: Kinect 11 Joint Properties and the Meaning

Chapter 3

Experiment

With the experiment described below, we evaluate the suitability of Jang et al.'s [1] CF model for estimating cumulative shoulder fatigue during mid-air interaction. To do so, we replicated the 'virtual hand' mid-air pointing task that they used to validate their model, but moved interactions to a large, multi-touch display. This design closely mirrors their 'Condition C', which asked participants to interact with a display very near to their arm, but via mid-air interactions. We expected that by replicating this design, but instead having participants use touch interaction, we would show the adequacy of their model in estimating fatigue for large, multi-touch interfaces.

3.1 Participants

We recruited 24 right-handed volunteers (7 female) from a local university campus. During the recruitment, a self-reported screening measure was applied to ensure that no subject had a musculoskeletal disorder or neurological disease. Participants were asked about their general health condition and previous medical history about the musculoskeletal disorder or neurological disease before any further experiment. After that, in the foremost questionnaire (shown in section A.1), we collected information from participants, including their age, dominant hand, weight, height, experience with touchscreens and their daily exercise level. This daily exercise level was explained to participants within the same document to ensure they reported it correctly. In addition, we took measurements of their shoulder height, upper arm length, lower arm length, and hand length. According to the data collected, their ages ranged from 21 to 39 years ($\mu = 25.7$ yrs); heights ranged from 152 to 194 cm ($\mu = 171.8$ cm); weights from 47 to 95 kg (M = 61.3 kg); upper arm lengths from 29 to 37 cm (μ = 32.7 cm); lower arm lengths from 21 to 29 cm (M = 25.3 cm); and hand lengths ranged from 16 to 21 cm (μ = 18.5 cm). Participants' experience with touch screens ranged from 0.5 to 6 years (M= 1.8 yrs).

3.2 Apparatus

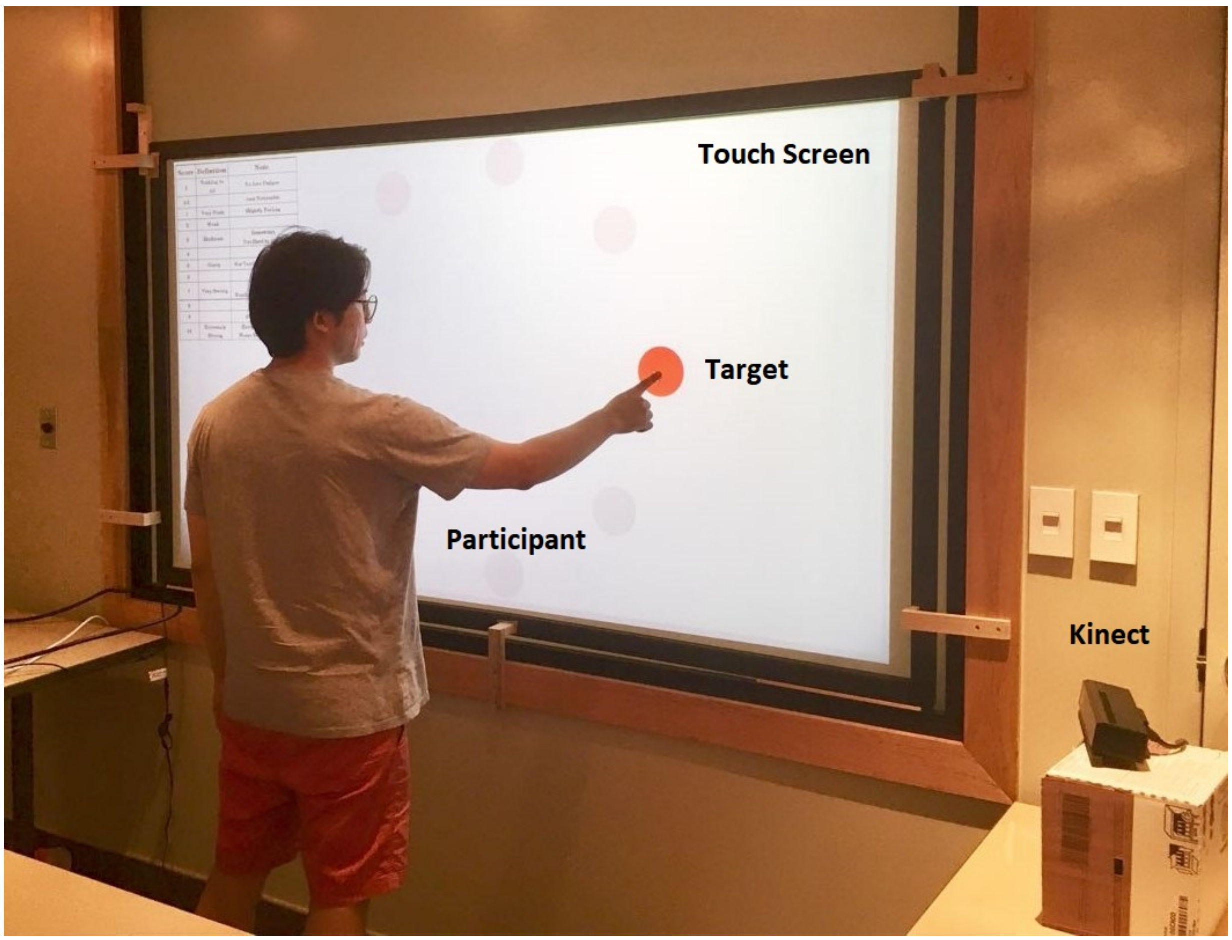


Figure 3.1: Participant performing tasks in the experiment

Participant completed the ISO 9241-9 standard [63] pointing task on the large, multi-touch display. Skeletal data was collected by a Microsoft Kinect V2, located to the participant's right.

We used a Microsoft Kinect version 2 sensor and the corresponding software development kit (SDK) to track arm movements. Data was sampled at 30 frames/second using a desktop with a Core i7 2.90GHz CPU. The computer was connected to a short-throw projector to project the task interface from back. The physical size of the projected screen on the wall was 2×1.5 m, and the Kinect camera was located 1 meter on the right of the screen and 1 meter above the floor to track participants' right arm movements. Participants stood about 0.5 - 1 m in front of the screen, at a distance where they felt comfortable working on the touch screen (Figure 3.1). The Kinect camera was placed at a slight angle facing the participant instead of strickly vertical to the wall to ensure that it captured the complete skeleton of the participant.

To support touch, the display was equipped with an 84" PQ Labs G4 IR frame. To minimize friction added through touching the screen, we piloted the study to find the optional touch screen calibration, and provided talcum powder to participants as lubricant. The setup was identical for all experiment blocks. The experimental software was written in JavaScript and HTML, running on Google Chrome browser v60. Actions were taken to minimize latency during experimental trials.

3.3 Experimental Tasks & Design

To replicate Jang et al.'s evaluation of their CF model, we compared the model's predicted arm fatigue with the participant's subjective ratings. As in their study, we collected fatigue ratings while participants worked through a series of increasingly fatiguing interactions with a nearby display. Our experimental design included 2 Task $\times$ 2 Distance as within-subjects factors, with 2 Rest Patterns as a between-subjects factor.

The study Tasks are based on the ISO 9241-9 standard [63] pointing task, with participants either Tapping or Dragging targets. This ISO 9241 standard is a multi-part standard proposed by the International Organization for Standardization (ISO) covering ergonomics of human-computer interaction. The ISO 9241-9 is widely applied in multiple scenarios of task design for human-computer interaction evaluation and is demonstrated to cover most fundamental physical characteristics of computer equipment. For both Tasks, 9 targets arranged in a circle centered on the middle of the display were shown. These tasks were chosen as representative of common types of multi-touch interaction, since interaction with large displays often involves tapping on icons or buttons, and moving content from one place to another via dragging.

Importantly, in this study, we aim to measure perceived fatigue only caused by muscular effort, and not by the task difficulty, so target sizes are intentionally selected to minimize

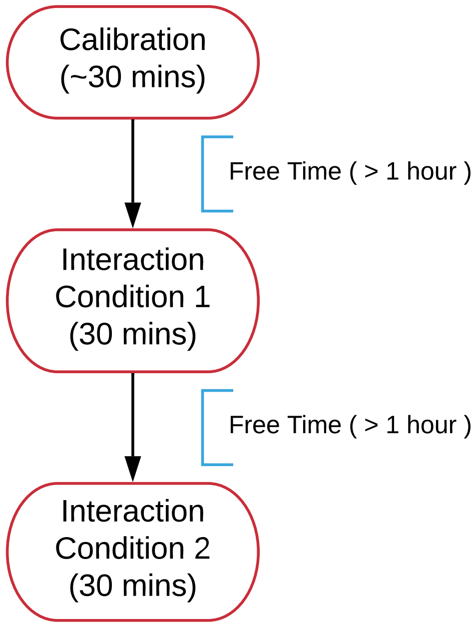


Figure 3.2: 3 major experiment sessions

Participants completed three sessions. The first session was used to calibrate the CF model for each user's body and maximum torque measurements. The second and third sessions were used to collect TAPPING and DRAGGING data, and were counterbalanced for TASK.

task difficulty. This potential effect was identified by Jang et al. during pilot testing, and so we decided to replicate their target widths: all targets were calibrated to a width of 10 cm. To study the accuracy of CF model for large screens, we included two DISTANCES: SHORT had a distance of 30 cm between targets and a Fitts's ID of 2.58, while LONG had a distance of 96 cm between targets and a Fitts's ID of 4.26. Here Fitts's ID is the index of difficulty defined by Fitts et al. for quantifying the difficulty of a target selection task [64]. The metric of Fitts's ID is:

$$ID = log_2(\frac{2D}{W})$$

where D is the distance to the center of the target, and W is the tolerance or width of the target.

Also, since Jang et al. reported that their models performance was robust to changes in the duration of rest, we included two REST PATTERNS as a between-subject factor: CONSTANT and DYNAMIC. Participants assigned to the CONSTANT rest pattern took

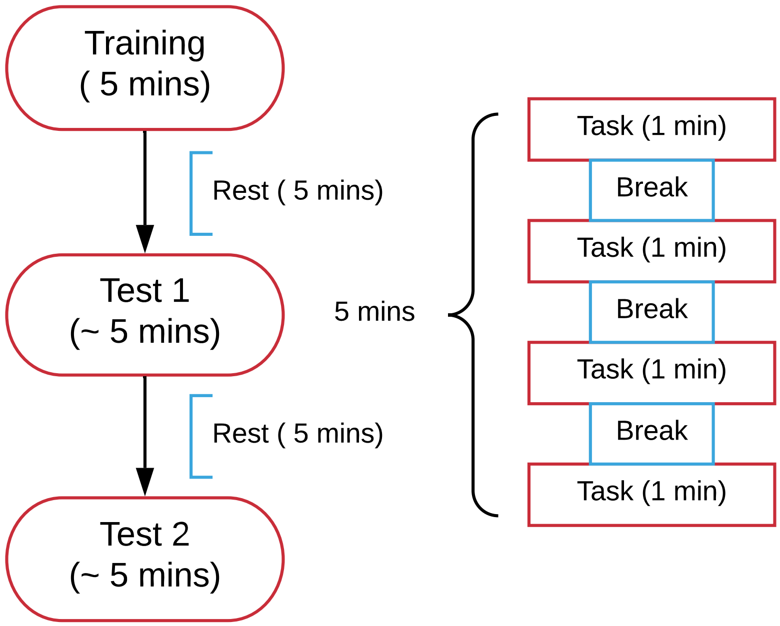


Figure 3.3: Detailed structure within each test session

The first sub-session was provided for participants to become familiar with the tasks and self-reporting experience. The second and third sub-sessions were used to collect LONG DISTANCE or SHORT DISTANCE data, and were counterbalanced for DISTANCE. Within each test sub-session, participants alternated between TASK blocks and BREAK blocks.

a 15-second break between each two 1-minute task block. Participants assigned to the DYNAMIC rest pattern received 5 seconds, then 10 seconds, then 20 seconds, then 15 seconds between the 1-minute blocks, replicating the rest pattern used by Jang et al. [1] .

3.4 Procedure

Data collection for each participant spanned three sessions: one for calibration, and one for each of TAPPING and DRAGGING interactions. To minimize the accumulation of arm fatigue during the study, each session was scheduled to last for less than 30 minutes, and at least one hour rest time between each session was mandated. Within each session, we also scheduled regular break times (i.e., REST PATTERN). The general procedure can be explain through Figure 3.2 and Figure 3.3.

During the calibration session, we replicated the protocol reported in Jang et al.'s validation experiment to calibrate the CF model for each participant. We measured participants' gender, age, height, weight, upper and lower arm lengths, and hand length. Participants were then shown an introductory video about the Borg Scale, trained in its use, and provided an opportunity to practice using it with the investigator.

Next, participants performed a calibration task where they held a 2.5-pound weight on their right wrist and held their arm horizontally for as long as they could. Throughout this period, participants were asked to practice the use of the Borg scale by regularly reporting their fatigue levels. By requiring participants to hold the weight for as long as possible, we were also able to calibrate experienced fatigue with descriptors on the Borg scale.

We then calculated each participant's maximum shoulder torque using the following equation from Jang et al.:

$$T_{max} = \frac{-b \times T_{avg}}{log(\frac{ET}{a})} \times 100$$

where ET is the measured endurance time, and a and b are constants provided by Jang et al. [1].

In this equation, we can easily see that T_{avg}, the average torque exerted, and ET, the endurance time, together contribute to the T_{max}, the maximum shoulder torque. Considering one's maximum shoulder torque as constant within a certain period of time, T_{avg} and ET are correlated. In the calibration task, the participant did not change their posture, therefore the average torque is only related to the weight they are holding. The heavier the weight is, the larger the average torque will be and the shorter time participant can endure.

In the pilot study, we tried several levels of weights for the calibration, including bare hand, 1 lb, 2.5 lbs, and 5 lbs. We found that when the weight was too light (bare hand or 1 lb), arm fatigue did not increase significantly enough for participants to notice, and the calibration task may last for too long. However, for 5-lb weight, we noticed that it was too heavy for some participants to even hold it up, and the arm fatigue increased so fast that participants skipped some Borg Scale levels, which was not what we intended. We regarded the calibration task as an extra chance for participants to get familiar with the Borg10 Scale self-reporting. As a result, we decided to use 2.5 lbs as the weight for formal calibration.

After at least 1 hour of rest time, participants came back for the second and third sessions. During the second and third sessions, participants completed the Tapping and Dragging tasks separately. In both sessions, participants were first instructed on their task and the structure of the session (including RESTING PATTERN), and allowed to practice the task using the Borg scale, and given the chance to ask questions. This training sub-session lasted for around 5 minutes. After the training was completed, participants were given 5 minutes to rest, before completing 5 minutes of trials for each DISTANCE condition, with a second 5 minute rest period in the middle, matching Jang et al.'s procedure. The order for DISTANCE was counterbalanced between participants. During each set of TAPPING or DRAGGING trials, participants were expected to continuously perform the task with the exception of pre-scheduled breaks according to their assigned RESTING PATTERN.

Finally, at the end of their third session, participants completed a short interview about their experience and were given an opportunity to comment on what they felt was most fatiguing about the large display interactions.

3.5 Data Collection and Analysis

Throughout each trial, the skeletal data of participants were recorded using a Microsoft Kinect V2 and used to calculate joint torques using the CF model. To replicate the experiment as closely as possible, we also applied a moving-average filter (15^{th} order) to smooth joint-torque trajectories. Computer logs also captured both start time and end time for each block, all touch positions, and all task events (click, miss, moving distance, responding time, etc.). Participants' subjective ratings for arm fatigue were recorded by the investigator verbally every 20 seconds. Participants' interview responses were recorded via field notes.

Before analysis, we first reviewed the data for correctness, and cleaned the incomplete, incorrect, inaccurate or irrelevant parts if there was any. We went through the Kinect captured videos with the skeleton tracked to confirm that all joints were tracked and labeled correctly. In addition, we checked the participants' subjective ratings for the correctness, all remaining within the range of [0, 10].

We used computer logs with blocks and touch information to calculate the trial number and error rate, which helped with the general understanding about participants' performance. We fed the body tracking data, together with the corresponding body measurements, to the CF model to calculate the shoulder torque and predict the real-time fatigue prediction. These predictions were then compared with participants' subjective ratings, which were collected during the experiment, to evaluate the performance of the CF model for large mulit-touch display interaction.

To investigate differences between participants' perceived fatigue and that predicted by the CF model, we calculated Root-Mean-Square-Error (RMSE) values every 20 seconds for each session. These values were calculated based on the difference between the CF model and participants' subjective ratings recorded during each trial. We then used a repeated-measures analysis of variance (RM-ANOVA) with TASK, DISTANCE, and REST CONDITION as factors to examine the RMSE and participants' subjective reports data for trends. After a visual inspection of the RMSE data, we performed a second RM-ANOVA with data summarized by BLOCK to investigate differences in RMSE over time.

All collected data met the assumptions of independence, equality, and normality, for the ANOVA tests. For all tests, $\alpha = .05$.

3.6 Results

On average, our participants completed 369 drags and 619 taps throughout their sessions, with an overall error rate of 4.6%. Our RM-ANOVA revealed no differences in RMSE for TASK ($F_{1,22} = .077, p = .784$), DISTANCE ($F_{1,22} = .441, p = .513$), or REST PATTERN ($F_{1,22} = .330, p = .572$). And throughout the entire study, we found an average RMSE for the CF model of 2.399 (Figure 3.4). Similarly, our analysis of participants' subjective ratings revealed no differences between TASK ($F_{1,22} = 1.365, p = .255$), DISTANCE ($F_{1,22} = .543, p = .469$), or REST PATTERN ($F_{1,22} = .186, p = .671$). Participants' average reported fatigue across the entire study was 2.72.

We also examined RMSE over time, and found a main effect for BLOCK ($F_{1,22} = 14.002, p = .001$), where RMSE rose from an average of .854 in Block 1 to an average of 2.28 in Block 14 (Figure 3.4). No interaction effects were found for BLOCK × TASK ($F_{1,22} = .164, p = .597$) or BLOCK × DISTANCE ($F_{1,22} = .337, p = .631$).

Further visual inspection of the graph suggested that the error in RMSE appeared significantly greater during periods of activity than those in which participants were allowed to rest. To examine these differences, we performed an additional RM-ANOVA comparing RMSE for periods of ACTIVITY and REST with REST PATTERN as a between-subjects variable. For this analysis, participants' blocks 3, 7, and 11 for the CONSTANT REST PATTERN, and blocks 3, 6, 10, and 14 for the DYNAMIC REST PATTERN were classified as REST.

We found a significant difference in RMSE ($F_{1,22} = 28.034, p \approx .000$), where RMSE was higher in blocks without a rest period ($\mu = 1.588, SE = .179$) than in those with a rest

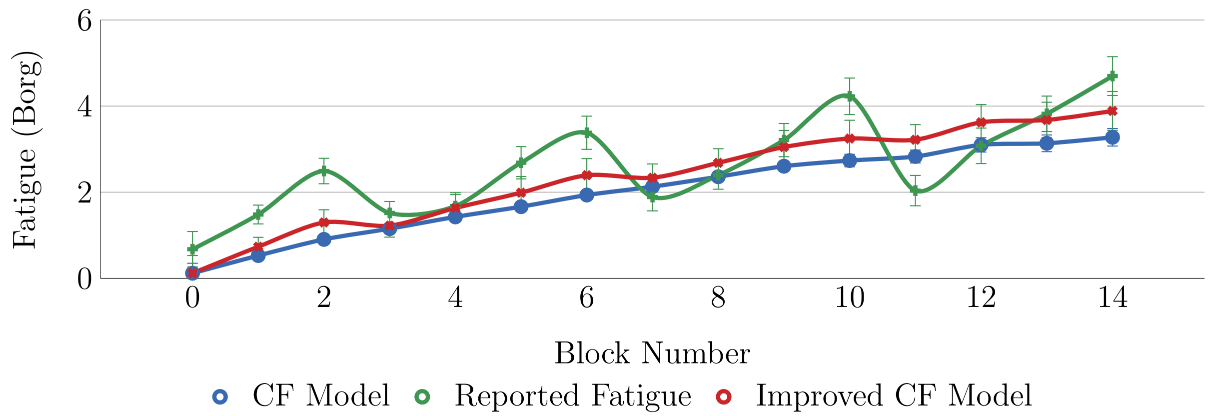


Figure 3.4: Overall CF model results

Study-wide averages for CF model predicted fatigue, participants' reported fatigue, and RMSE, by experimental block. Error bars show standard error.

period ($\mu = .698, SE = .090$). Our analysis found no interaction effect between BLOCK and REST PATTERN or BLOCK and ACTIVITY ($F_{1,22} = 1.160, p = .293$).

All of the fatigue predictions were calculated by the CF model with the suggested parameters from Jang et al., including the F and R. However, after a visual inspection of Figure 3.4, we could came to the conclusion that the model does not perform aggressively enough when fatigue is increasing or decreasing. We therefore adjusted the upper bound of both F and R to generate an improved model, where model parameters F and R lied outside of Jang et al.'s suggested range for mid-air interaction. Our improved CF model, with tuned F and R, showed improved RMSE when compared to Jang et al.'s mid-air model.

Table 3.1: Summary of participant data

ID	Gender	Age (Yr)	Height (cm)	Weight (Kg)	Upper Arm (cm)	Lower Arm (cm)	Hand (cm)	Endurance Time (min)	Average Fatigue Rating (Borg Scale)
1	M	25	168	55	30	25	17	3:43	1.72
2	M	25	176	67	31	28	18	3:29	2.41
3	M	28	181	79	33	27	21	5:43	1.52
4	M	23	185	95	37	28	20	5:30	1.24
5	M	27	173	65	33	27	19	3:11	1.68
6	M	27	194	75	35	27	21	6:06	1.89
7	M	30	172	78	32	26	18.5	4:55	1.83
8	M	30	170	72	32	24	19	3:31	2.38
9	F	21	157	47	29	25	17	2:01	2.26
10	F	29	160	50	30	21	17	1:44	3.2
11	M	22	165	68	33	24	16	4:38	1.64
12	F	27	157	65	29	23	17	2:14	2.69
13	F	39	162	55	31	22	16	2:55	1.82
14	M	21	178	57	37	27	19	8:42	0.84
15	M	25	183	77	37	27	21	3:43	2.55
16	M	33	176	63	35	29	19	5:17	1.26
17	M	25	165	59	29	24	17	3:06	2.73
18	F	21	152	57	30	23	18	3:18	1.92
19	M	22	179	65	34	26	20	3:55	1.93
20	F	21	175	76.5	32	25	18	3:09	2.09
21	F	22	165	55	29	24	18	2:34	1.51
22	M	23	183	60	31	24	19	3:56	1.96
23	M	28	170	65	31	25	19	4:48	2.00
24	M	22	178	62	33	26	19	2:05	2.85
Average		25.67	171.83	65.31	32.21	25.29	18.48	3:55	2.4

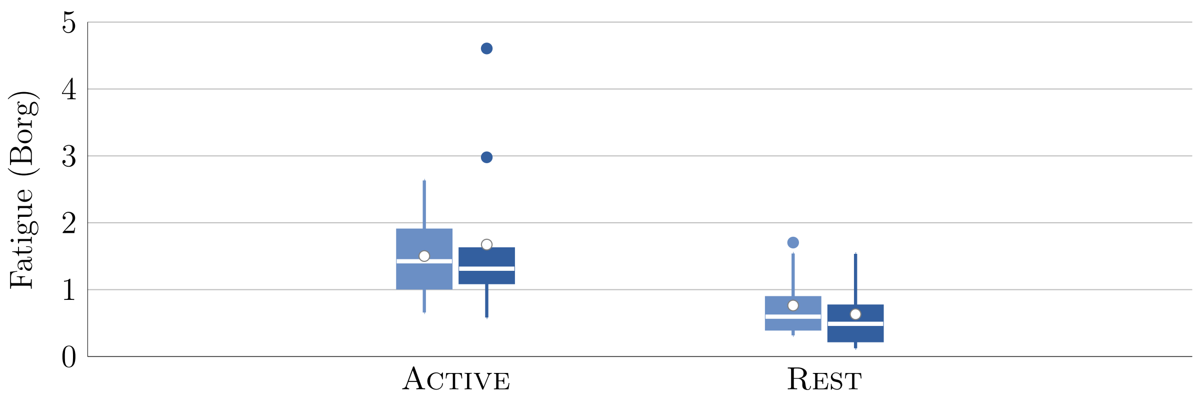


Figure 3.5: Box plots of RMSE between Active and Rest Blocks

Box plots of RMSE for Blocks containing no rest period (Active) and those containing a rest period (Rest), with Constant (left) and Dynamic (right) Rest Patterns. Our RM-ANOVA found a significant difference between Active and Rest periods, where the CF model was more accurate for resting periods. Median marked by white bar, mean marked by white dot.

Chapter 4

Discussion

Our research was motivated by the need for more accurate and easily deployable measures of fatigue for multi-touch interaction on large displays. We hoped that by leveraging Jang et al.'s [1] CF model, we would validate such a tool. Our initial analysis of RMSE found an average of 2.4 across all conditions. In their initial validation of the CF model for mid-air interaction, Jang et al. report RMSEs ranging from .79 to 1.38 with an overall RMSE of 1.33. These differences between reported and predicted fatigue are significantly smaller than what we observed for multi-touch interaction. Based on these measurements, we must conclude that their CF model does not directly translate to multi-touch interaction.

We performed an RM-ANOVA to understand whether any of our independent variables (i.e. Task, Distance, or Rest Pattern) influenced the error rate of the model, but found no differences. Based on this analysis, RMSE appears to be equally high across all of our independent variables. However, visual inspection of our data pointed to time (i.e. Block) as an influential factor of RMSE, and in particular, we found that fluctuations in the CF model's estimates corresponded to blocks in which participants were allowed to rest. This analysis revealed that RMSE increased throughout our study and that estimates for blocks with a rest period were significantly more accurate than those without.

Based on these results, we calibrated the model to more accurately predict fatigue up close to the display. The CF model depends on **F**atigue and **R**ecovery rate parameters, and they provide upper and lower bounds for F and R based on mid-air interaction. In a preliminary analysis of our data, we found that increasing F to the upper bound improved RMSE by 8.7%, and further optimization beyond that bound improved RMSE by 10.2% (RMSE 1.8) as shown in Figure 3.4. Notably, the optimal value is outside their suggested bound, indicating that interaction on large multi-touch displays may fatigue users differently than mid-air interaction, and that modeling such interactions requires the model to be calibrated for them.

One specific concern for the data analysis is that using the Borg Scale for reported fatigue may cause autocorrelation, since participants will be more likely to try to keep these scores consistent and less likely to rate lower than their previous rating as they keep performing the tasks. This symptom is expected, because in the TCM model, participants' muscle state won't directly transfer from fatigue to active, and only transfer from fatigue to rest when resting. In this way, the participants' ratings are supposed to keep increasing during task blocks. Besides, we provided the calibration session as a training session for the Borg scale rating and confirmed that every participant understood the Borg scale completely before proceeding to the next session. Another piece of evidence suggesting that autocorrelation does not bias the conclusion is that the starting rating of each session from every participant is 0. This guarantees that participants understood the Borg Scale and were able to give a consistent rating, and that participants always started the test sessions in a fully recovered condition.

We also performed a one-way ANOVA with participants' demographic data and body measurements to understand whether any of these independent variables (i.e. GENDER, AGE, or ENDURANCE TIME) influenced the general fatigue rating. We found no differences for GENDER($F_{1,22} = 1.508, p = .232$), AGE($F_{1,22} = 0.025, p = .877$), HEIGHT($F_{1,22} = 2.096, p = .162$), or ARM LENGTH RATION($F_{1,22} = 0.031, p = .863$). However, ENDURANCE TIME($F_{1,22} = 26.59, p \approx .000$) significantly contributed to the average fatigue ratings. Participants with longer endurance time reported lower fatigue ratings, which is expected since longer endurance time results in larger maximum torque.

We now discuss some of the qualitative findings from our study that may further elucidate differences between the CF model's predictions and the fatigue experienced by participants.

4.1 Subjective Fatigue Ratings

A reasonable question is whether participants' subjective reports showed differences for any of our independent variables. These reports serve as 'ground truth' for our analysis of RMSE, and a number of our participants in the follow-up interview mentioned that DRAG was more fatiguing than TAP, and that LONG was more fatiguing than SHORT. However, our RM-ANOVA of participants' subjective ratings revealed no statistically significant results for TASK, DISTANCE, or REST PATTERN.

The lack of a difference between these experimental controls is perhaps counter-intuitive, but can be explained by the task design — participants performed the target selection task continuously for about 5 minutes regardless of what condition they were in, as opposed to performing a fixed number of trials in each block. Nevertheless, these results support the lack of differences found in RMSE across our experimental controls.

It is worth noting that Jang et al. report in their mid-air validation that some participants were more sensitive in their reports of fatigue than others. We also found that participants varied in their subjective assessments of fatigue. For example, Participant 10 reported fatigue much higher than others, with an average of 3.2, whereas Participant 14 reported significantly lower ($\mu = .084$). While most of our participants reported similar fatigue ratings throughout the task, the variance may point to some of the difficulty that the CF model had in accurately predicting fatigue. Average fatigue ratings are presented by-participant in Table 3.1.

4.2 Participant Interview Feedback

We also examined participant feedback during their post-study interviews to look for trends in what they felt was most fatiguing about the task. 15 participants suggested that DRAGGING or LONG distances were more fatiguing than TAPPING or SHORT distances respectively, while another 5 participants claimed that SHORT distances were more fatiguing. Through looking back to their experiment, we found that participants who kept a constant speed to complete the tasks reported LONG distances to be more fatiguing because they had to move a longer distance during the experiment. However, some participants tried to accomplish a task as fast as they could, and clicked much more targets in the SHORT distances than in the LONG distances. In this case, the total arm movement in the SHORT distance was higher than the LONG distances, despite a single arm movement in the LONG distances being farther. However, the results of our statistical analysis of participant fatigue ratings do not support these suggestions as we did not find a difference between TASK, DISTANCE, or REST PATTERN. This indicates that in the future experiment, the number of tasks completed should be controlled.

13 participants mentioned that moving up was more fatiguing than down, and 4 participants mentioned that target selections on their left side were more difficult due to our restriction that participants only use their right arm. Based on our current study design for participants to report their fatigue levels every 20 seconds, we cannot compare specific sub-activities within each task (i.e., up/down or left/right drags). However, comparison to Jang et al.'s validation provides some insight into these differences.

The range of motion required for on-screen touch interactions is different than for mid-air pointing. When standing near a display, one has to reach 'up' and 'away', rather than pointing 'towards' a display. Thus, the arm movements that we observed in our study are more varied than those Jang et al. used to calibrate their model, and the movements that were described by participants as being particularly fatiguing were not studied either. We expect, however, that given additional data collection, the CF model can be improved further to better predict fatigue for these settings.

The height of the interaction zone also influenced participants' interpretation of fatigue. According to several participants whose heights were lower than 160 cm, the pain in the neck, shoulder and arm muscles increased significantly when the target was out of their reachable range. They would have to stand on their toes and bend to the side a little to ensure their hands were high enough to reach the target. This posture was awkward and more fatiguing than their comfortable operating posture. This statement can lead to a future requirement for an extended model that may consider extra joints including the neck, waist and even leg to better understand the general muscle fatigue of large multi-touch display interaction.

Chapter 5

Limitations and Future Work

From our evaluation of Jang et al. 's [1] CF model for mid-air interaction, we may have an incomplete picture of how fatigue occurs when working with large displays. There remain several limitations of this work including arm fatigue evoked from force pressure, and extended fatigue estimation model supporting more complicated and more flexible interaction. These limitations provide several directions for future works. In the end, we discuss the current model's contributions and its future contribution with potential applications.

5.1 Limitations

The CF model may be applicable to touch interaction on a large display and predict arm fatigue for large multi-touch display interaction. However, its accuracy can be an issue that is worth conducting deeper research into. The forces that are present in the scenario involving the touchscreen is more complicated than during mid-air gestures which may result in the inability of the CF model to accurately predict arm fatigue for large multi-touch display by either underestimating or overestimating its effect. Thus, we may conclude that our result from the validation experiment can place a lower bound on the accuracy of the model, and future work that addresses limitations is likely to make it even more accurate in practice.

5.2 Future Work

There are several improvements that we would like to address as future work.

First, force-sensitive displays have been explored for some time (e.g., [65, 66, 67]) and are commercially available in large display formats. A limitation of the CF model is that it relies on Kinect data for the shoulder and elbow joints only, and thus cannot take into account the force exerted by a user *into* the display. Predictive models of fatigue may benefit from incorporating the force data captured from these displays in the future.

Second, interaction with large displays increasingly involves more than just the arm of the user. Trends towards larger displays mean that users cannot stand in one place and reach all of the content they need to, they may need to walk around to navigate content [68], reach up or bend down to access new parts of the display, use bi-manual interaction techniques, or even kick it [69]. Accounting for the fatigue induced by any of these activities requires a more complete skeletal model that encompasses more than a single arm. For example, an upper-body model may be used for two-hand interaction and a whole-body model may allow participants to walk when interacting.

Finally, we evaluated the CF model's accuracy only for tapping and dragging; it is likely to be even less accurate for more complex interactions. For example, Olwal et al., [70] explored 'rubbing' techniques for the multi-touch selection and zooming and noted during their evaluation that these gestures were especially fatiguing to use. Other techniques often use arm and wrist rotation as input, which is also unlikely to be fully recognized by the existing model.

5.3 Contributions

We expect this validated CF model would provide a better understanding of the arm fatigue on large multi-touch displays and also become a useful tool for evaluating arm fatigue for various large multi-touch display settings and interactions. To make it clear, we would like to take an interaction designer who aims for a low-fatigue design for large multi-touch displays as an example.

Before the CF model was validated, the designer could only design based on the empirical knowledge and iterate upon the design according to potential users' feedback from the trials. Such qualitative information may give the designer a general impression of whether a design is more fatiguing or not. Nevertheless, the designer would not be able to identify the parts that cause more fatigue to make significant improvements.

Now that the CF model is validated for the large multi-touch displays, the designer can easily evaluate several designs' performance with more accurate results. With the system set up, the designer can simply invite potential users to try with the designs, and monitor their fatigue estimation at the same time. In this way, the estimation of fatigue can directly provide evidence on whether a certain interaction design is more fatiguing than another one, and the designer can selectively improve the design through iterations.

5.4 Potential Application: Fatigue-Aware Interface

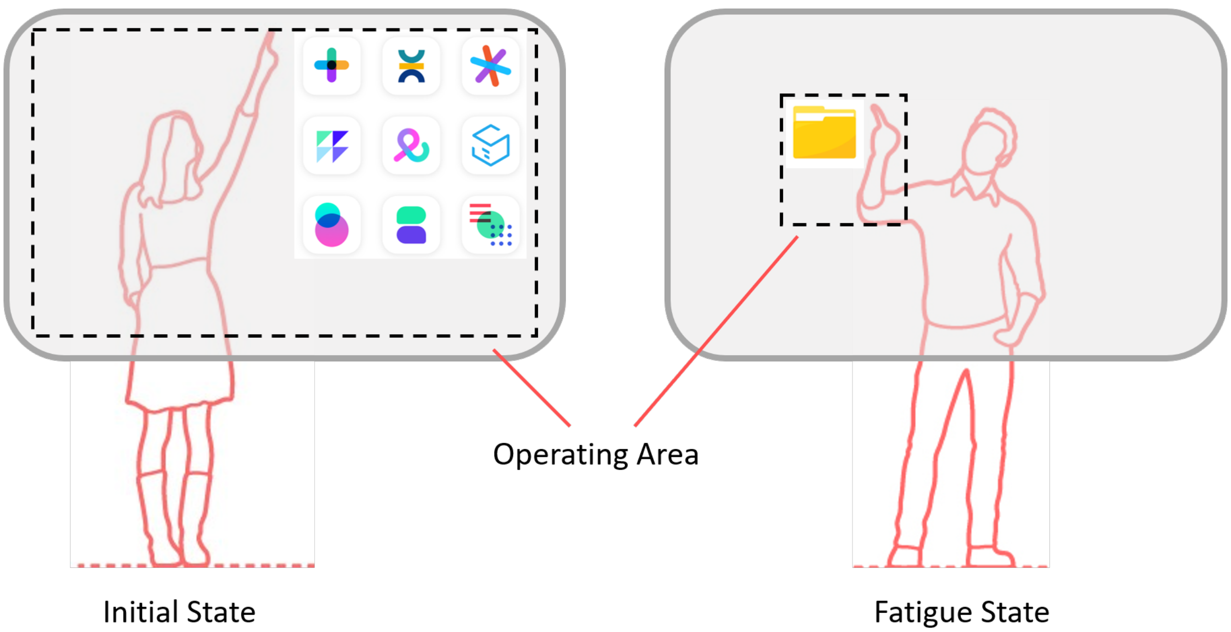


Figure 5.1: Fatigue-Aware Interface

We believe this CF model can be applied to multiple scenarios in the future. One potential application for this is the Fatigue-Aware Interface. A fatigue-aware interface is one type of adaptive user interfaces which can automatically self-adjust its layout and elements according to the user's fatigue level.

An adaptive user interface (also called AUI) refers to a user interface that changes its layout and elements based on the context and requirements of each user [71]. An AUI can be initially adapted according to the user's personalized settings, or it may automatically change later based on an analysis of the user's behaviours. Studies about AUI involve devices such as mobile phones [72] and personal computers, and are conducted in environments including smart settings [73] and outdoor navigation scenarios [74]. Their methodologies include machine-learning-based user modeling [75] and error-based self-improving.

By leveraging the AUI approach, we could propose the design of a fatigue-aware interface, which adapts with the users' fatigue conditions. Researchers have proved that a fatigue-related adaptive driving system can provide drivers with sufficient rest when needed, thereby significantly reducing the adverse effects of driver fatigue and workload [76]. We expect a fatigue-aware interface can similarly improve user experience with large multi-touch displays, through providing users with sufficient rest and thus reducing fatigue.

For example, initially, when the user is not experiencing fatigue, all elements on the interface are aligned by category, so that the user can access any of them within one touch. When the interface detects that the user's arm fatigue has reached a relatively high level and requires rest, it can automatically pack up the elements and organize them into a hierarchy. Through this adjustment, the operation area will shrink into a smaller size so that users can reach without large movements. Such layout may require longer operation procedures but all selections provide users with some time to recover from arm fatigue. Eventually, when the system identifies that the user's fatigue levels have decreased enough, the interface will unpack all of the elements back again. This kind of interface is theoretically effective to help reduce arm fatigue, and we hope the CF model will contribute to its implementation.

Chapter 6

Conclusion

Fatigue has long been a problem for multi-touch displays but designers have had few tools at their disposal to measure and understand it. To address this gap, we investigated the accuracy of Jang et al.'s mid-air Cumulative Fatigue (CF) model for multi-touch interactions on a large display. We found that their mid-air model underestimates fatigue, particularly during periods of activity and that it became more inaccurate the longer it was used. Based on these results, we calibrated the model for touch interactions, achieving error rates comparable to those for mid-air interaction. Based on our evaluation, we discuss how participants' subjective ratings and interview feedback can further inform the improvement of the CF model for multi-touch interaction, and a continued need for lightweight tools to assess fatigue on large, multi-touch displays.

References

[1] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani, "Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, (New York, NY, USA), pp. 3328–3339, ACM, 2017.

[2] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," *Advances in psychology*, vol. 52, pp. 139–183, 1988.

[3] S. Elrod, R. Bruce, R. Gold, D. Goldberg, F. Halasz, W. Janssen, D. Lee, K. McCall, E. Pedersen, K. Pier, J. Tang, and B. Welch, "Liveboard: A large interactive display supporting group meetings, presentations, and remote collaboration," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, (New York, NY, USA), pp. 599–607, ACM, 1992.

[4] N. Matsushita and J. Rekimoto, "Holowall: Designing a finger, hand, body, and object sensitive wall," in *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, (New York, NY, USA), pp. 209–210, ACM, 1997.

[5] M. Weiser, R. Gold, and J. S. Brown, "The origins of ubiquitous computing research at parc in the late 1980s," *IBM Syst. J.*, vol. 38, pp. 693–696, Dec. 1999.

[6] C. Ardito, P. Buono, M. F. Costabile, and G. Desolda, "Interaction with large displays: A survey," *ACM Computing Surveys (CSUR)*, vol. 47, no. 3, p. 46, 2015.

[7] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko, "It's mine, don't touch!: Interactions at a large multi-touch display in a city centre," in *Proceedings of the SIGCHI Conference on*

Human Factors in Computing Systems, CHI '08, (New York, NY, USA), pp. 1285–1294, ACM, 2008.

[8] A. Bellucci, A. Malizia, P. Diaz, and I. Aedo, "Don't touch me: Multi-user annotations on a map in large display environments," in *Proceedings of the International Conference on Advanced Visual Interfaces*, AVI '10, (New York, NY, USA), pp. 391–392, ACM, 2010.

[9] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Linden, M. Jurmu, S. Hosio, F. Kruger, and D. Zanni, "Multipurpose interactive public displays in the wild: Three years later," *Computer*, vol. 45, no. 5, pp. 42–49, 2012.

[10] T. Ojala, H. Kukka, T. Lindén, T. Heikkinen, M. Jurmu, S. Hosio, and F. Kruger, "Ubi-hotspot 1.0: Large-scale long-term deployment of interactive public displays in a city center," in *Internet and Web Applications and Services (ICIW), 2010 Fifth International Conference on*, pp. 285–294, IEEE, 2010.

[11] H. Brignull and Y. Rogers, "Enticing people to interact with large public displays in public spaces," in *Proceedings of INTERACT*, vol. 3, pp. 17–24, 2003.

[12] J. Müller, F. Alt, D. Michelis, and A. Schmidt, "Requirements and design space for interactive public displays," in *Proceedings of the 18th ACM international conference on Multimedia*, pp. 1285–1294, ACM, 2010.

[13] M. Bachynskyi, G. Palmas, A. Oulasvirta, J. Steimle, and T. Weinkauf, "Performance and ergonomics of touch surfaces: A comparative study using biomechanical simulation," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, (New York, NY, USA), pp. 1817–1826, ACM, 2015.

[14] F. Wang and X. Ren, "Empirical evaluation for finger input properties in multi-touch interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1063–1072, ACM, 2009.

[15] H.-J. Appell, J. Soares, and J. Duarte, "Exercise, muscle damage and fatigue," *Sports Medicine*, vol. 13, no. 2, pp. 108–115, 1992.

[16] S. A. Dugan and W. R. Frontera, "Muscle fatigue and muscle injury," *Physical Medicine and Rehabilitation Clinics*, vol. 11, no. 2, pp. 385–403, 2000.

[17] M. Grujicic, B. Pandurangan, X. Xie, A. Gramopadhye, D. Wagner, and M. Ozen, "Musculoskeletal computational analysis of the influence of car-seat design/adjustments on long-distance driving fatigue," *International Journal of Industrial Ergonomics*, vol. 40, no. 3, pp. 345–355, 2010.

[18] A. Leme and I. Maia, "Evaluation of fatigue at work in teachers using modern resources in the classroom," *Procedia Manufacturing*, vol. 3, pp. 4852–4859, 2015.

[19] S. K. Hunter, J. Duchateau, and R. M. Enoka, "Muscle fatigue and the mechanisms of task failure," *Exercise and sport sciences reviews*, vol. 32, no. 2, pp. 44–49, 2004.

[20] S. E. Mathiassen and E. Åhsberg, "Prediction of shoulder flexion endurance from personal factors," *International Journal of Industrial Ergonomics*, vol. 24, no. 3, pp. 315–329, 1999.

[21] T. Öberg, L. Sandsjö, and R. Kadefors, "Subjective and objective evaluation of shoulder muscle fatigue," *Ergonomics*, vol. 37, no. 8, pp. 1323–1333, 1994.

[22] G. Borg, *Borg's perceived exertion and pain scales.* Human kinetics, 1998.

[23] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani, "Consumed endurance: A metric to quantify arm fatigue of mid-air interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, (New York, NY, USA), pp. 1063–1072, ACM, 2014.

[24] J. Ruiz and D. Vogel, "Soft-constraints to reduce legacy and performance bias to elicit whole-body gestures with low arm fatigue," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, (New York, NY, USA), pp. 3347–3350, ACM, 2015.

[25] G. Bruder, F. Steinicke, and W. Sturzlinger, "To touch or not to touch?: Comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop surfaces," in *Proceedings of the 1st Symposium on Spatial User Interaction*, SUI '13, (New York, NY, USA), pp. 9–16, ACM, 2013.

[26] P. Kabbash, I. S. MacKenzie, and W. Buxton, "Human performance using computer input devices in the preferred and non-preferred hands," in *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, pp. 474–481, ACM, 1993.

[27] D. Wigdor and D. Wixon, *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier, 2011.

[28] S. D. Scott, M. S. T. Carpendale, and K. M. Inkpen, "Territoriality in collaborative tabletop workspaces," in *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, (New York, NY, USA), pp. 294–303, ACM, 2004.

[29] A. Tang, M. Tory, B. Po, P. Neumann, and S. Carpendale, "Collaborative coupling over tabletop displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, (New York, NY, USA), pp. 1181–1190, ACM, 2006.

[30] G. Mojgan, P. Marvin, C. Wong, J. R. Wallace, and S. D. Scott, "Increasing passersby engagement with public large interactive displays: A study of proxemics and conation," *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, 2018.

[31] V. Gentile, M. Khamis, S. Sorce, and F. Alt, "They are looking at me!: understanding how audience presence impacts on public display users," in *Proceedings of the 6th ACM International Symposium on Pervasive Displays*, p. 11, ACM, 2017.

[32] S. Deterding, A. Lucero, J. Holopainen, C. Min, A. Cheok, A. Waern, and S. Walz, "Embarrassing interactions," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2365–2368, ACM, 2015.

[33] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell, "A survey of design issues in spatial input," in *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology*, UIST '94, (New York, NY, USA), pp. 213–222, ACM, 1994.

[34] A. Kendon, *Gesture: Visible action as utterance*. Cambridge University Press, 2004.

[35] D. Vogel and R. Balakrishnan, "Distant freehand pointing and clicking on very large, high resolution displays," in *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, UIST '05, (New York, NY, USA), pp. 33–42, ACM, 2005.

[36] W. Yee, "Potential limitations of multi-touch gesture vocabulary: Differentiation, adoption, fatigue," *Human-Computer Interaction. Novel Interaction Methods and Techniques*, pp. 291–300, 2009.

[37] L. A. Frey Law and K. G. Avin, "Endurance time is joint-specific: a modelling and meta-analysis investigation," *Ergonomics*, vol. 53, no. 1, pp. 109–129, 2010.

[38] S. Al-Megren, A. Kharrufa, J. Hook, A. Holden, S. Sutton, and P. Olivier, "Comparing fatigue when using large horizontal and vertical multi-touch interaction displays," in *Human-Computer Interaction*, pp. 156–164, Springer, 2015.

[39] B. Ahlström, S. Lenman, and T. Marmolin, "Overcoming touchscreen user fatigue by workplace design," in *Posters and Short Talks of the 1992 SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, (New York, NY, USA), pp. 101–102, ACM, 1992.

[40] S. Boring, M. Jurmu, and A. Butz, "Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays," in *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OZCHI '09, (New York, NY, USA), pp. 161–168, ACM, 2009.

[41] K. Pietroszek, J. R. Wallace, and E. Lank, "Tiltcasting: 3d interaction on large displays using a mobile device," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, (New York, NY, USA), pp. 57–62, ACM, 2015.

[42] K. Katsuragawa, K. Pietroszek, J. R. Wallace, and E. Lank, "Watchpoint: Freehand pointing with a smartwatch in a ubiquitous display environment," in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '16, (New York, NY, USA), pp. 128–135, ACM, 2016.

[43] C. Harrison, S. Ramamurthy, and S. E. Hudson, "On-body interaction: Armed and dangerous," in *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, TEI '12, (New York, NY, USA), pp. 69–76, ACM, 2012.

[44] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay, "Mid-air pan-and-zoom on wall-sized displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, (New York, NY, USA), pp. 177–186, ACM, 2011.

[45] J. R. Wallace, S. D. Scott, T. Stutz, T. Enns, and K. Inkpen, "Investigating teamwork and taskwork in single- and multi-display groupware systems," *Personal and Ubiquitous Computing*, vol. 13, p. 569, Jun 2009.

[46] M. Liu, M. Nancel, and D. Vogel, "Gunslinger: Subtle arms-down mid-air interaction," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 63–71, ACM, 2015.

[47] S. Siddhpuria, K. Katsuragawa, J. R. Wallace, and E. Lank, "Exploring at-your-side gestural interaction for ubiquitous environments," in *Proceedings of the 2017 Conference on Designing Interactive Systems*, pp. 1111–1122, ACM, 2017.

[48] C. Arslan, Y. Rekik, and L. Grisoni, "E-pad: Large display pointing in a continuous interaction space around a mobile device," in *Proceedings of the 2019 on Designing Interactive Systems Conference*, pp. 1101–1108, ACM, 2019.

[49] Å. Dedering, G. Németh, and K. Harms-Ringdahl, "Correlation between electromyographic spectral changes and subjective assessment of lumbar muscle fatigue in subjects without pain from the lower back," *Clinical Biomechanics*, vol. 14, no. 2, pp. 103–111, 1999.

[50] D. L. LaChapelle and M. Finlayson, "An evaluation of subjective and objective measures of fatigue in patients with brain injury and healthy controls," *Brain injury*, vol. 12, no. 8, pp. 649–659, 1998.

[51] A. Troiano, F. Naddeo, E. Sosso, G. Camarota, R. Merletti, and L. Mesin, "Assessment of force and fatigue in isometric contractions of the upper trapezius muscle by surface emg signal and perceived exertion scale," *Gait & posture*, vol. 28, no. 2, pp. 179–186, 2008.

[52] M. Cifrek, V. Medved, S. Tonković, and S. Ostojić, "Surface emg based muscle fatigue evaluation in biomechanics," *Clinical Biomechanics*, vol. 24, no. 4, pp. 327–340, 2009.

[53] G. Sjøgaard, G. Savard, and C. Juel, "Muscle blood flow during isometric activity and its relation to muscle fatigue," *European journal of applied physiology and occupational physiology*, vol. 57, no. 3, pp. 327–335, 1988.

[54] S. C. Segerstrom and L. S. Nes, "Heart rate variability reflects self-regulatory strength, effort, and fatigue," *Psychological science*, vol. 18, no. 3, pp. 275–281, 2007.

[55] W. Rohmert, "Ermittlung von erholungspausen für statische arbeit des menschen," *European Journal of Applied Physiology and Occupational Physiology*, vol. 18, no. 2, pp. 123–164, 1960.

[56] J. Z. Liu, R. W. Brown, and G. H. Yue, "A dynamical model of muscle activation, fatigue, and recovery," *Biophysical journal*, vol. 82, no. 5, pp. 2344–2359, 2002.

[57] T. Xia and L. A. F. Law, "A theoretical approach for modeling peripheral muscle fatigue and recovery," *Journal of biomechanics*, vol. 41, no. 14, pp. 3046–3052, 2008.

[58] L. Ma, D. Chablat, F. Bennis, and W. Zhang, "A new simple dynamic muscle fatigue model and its validation," *International Journal of Industrial Ergonomics*, vol. 39, no. 1, pp. 211–220, 2009.

[59] R. Ma, D. Chablat, F. Bennis, and L. Ma, "Human muscle fatigue model in dynamic motions," in *Latest Advances in Robot Kinematics*, pp. 349–356, Springer, 2012.

[60] S. Morishita, S. Yamauchi, C. Fujisawa, and K. Domen, "Rating of perceived exertion for quantification of the intensity of resistance exercise," *Int J Phys Med Rehabil*, vol. 1, no. 9, pp. 1–4, 2013.

[61] Wikipedia contributors, "Kinect — Wikipedia, the free encyclopedia," 2019. [Online; accessed 23-May-2019].

[62] L. Jamhoury, "Understanding kinect v2 joints and coordinate system." Accessed: 2019-07-26.

[63] ISO, "Iso 9241-9 ergonomic requirements for office work with visual display terminals (vdts)-part 9: Requirements for non-keyboard input devices (fdis-final draft international standard)," *International Organization for Standardization*, 2000.

[64] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement.," *Journal of experimental psychology*, vol. 47, no. 6, p. 381, 1954.

[65] A. M. Grau, C. Hendee, J.-R. Rizzo, and K. Perlin, "Mechanical force redistribution: Enabling seamless, large-format, high-accuracy surface interaction," in *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, (New York, NY, USA), pp. 4137–4146, ACM, 2014.

[66] J. Y. Han, "Low-cost multi-touch sensing through frustrated total internal reflection," in *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, UIST '05, (New York, NY, USA), pp. 115–118, ACM, 2005.

[67] M. Hancock, T. ten Cate, and S. Carpendale, "Sticky tools: Full 6dof force-based interaction for multi-touch tables," in *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, (New York, NY, USA), pp. 133–140, ACM, 2009.

[68] R. Ball, C. North, and D. A. Bowman, "Move to improve: Promoting physical navigation to increase user performance with large displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, (New York, NY, USA), pp. 191–200, ACM, 2007.

[69] R. Jota, P. Lopes, D. Wigdor, and J. Jorge, "Let's kick it: How to stop wasting the bottom third of your large screen display," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, (New York, NY, USA), pp. 1411–1414, ACM, 2014.

[70] A. Olwal, S. Feiner, and S. Heyman, "Rubbing and tapping for precise and rapid selection on touch-screen displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, (New York, NY, USA), pp. 295–304, ACM, 2008.

[71] M. Schneider-Hufschmidt, U. Malinowski, and T. Kuhme, *Adaptive user interfaces: Principles and practice.* Elsevier Science Inc., 1993.

[72] N. Mitrovic and E. Mena, "Adaptive user interface for mobile devices," in *DSV-IS*, 2002.

[73] M. Kern, F. Trollmann, M. Blumendorf, and S. Albayrak, "Adaptive user interface assistance in smart environments," in *Proceedings of the Workshop on Meaning and Matching (AISB2010), De Montfort University Leicester, SSAISB*, Citeseer, 2010.

[74] B. VAN TONDER, *Adaptive user interfaces for mobile map-based visualisation.* PhD thesis, Masters thesis, Computer Science and Information Systems, Nelson Mandela , 2008.

[75] P. Langley, "Machine learning for adaptive user interfaces," in *KI*, 1997.

[76] P. A. Hancock and W. B. Verwey, "Fatigue, workload and adaptive driver systems.," *Accident; analysis and prevention*, vol. 29 4, pp. 495–506, 1997.

[77] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures," in

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11, pp. 143–146, 2011.

[78] F. Alt, S. Schneegaß, A. Schmidt, J. Müller, and N. Memarovic, "How to evaluate public displays," in *Proceedings of the 2012 International Symposium on Pervasive Displays*, p. 17, ACM, 2012.

[79] K. Swaminathan and S. Sato, "Interaction design for large displays," *interactions*, vol. 4, no. 1, pp. 15–24, 1997.

[80] U. Hinrichs, H. Schmidt, and S. Carpendale, "Emdialog: Bringing information visualization into the museum," *IEEE transactions on visualization and computer graphics*, vol. 14, no. 6, pp. 1181–1188, 2008.

[81] X. Bi, Y. Shi, and X. Chen, "Upen: a smart pen-liked device for facilitating interaction on large displays," in *Horizontal Interactive Human-Computer Systems, 2006. TableTop 2006. First IEEE International Workshop on*, pp. 7–pp, IEEE, 2006.

[82] A. M. Piper, E. O'Brien, M. R. Morris, and T. Winograd, "Sides: a cooperative tabletop computer game for social skills development," in *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pp. 1–10, ACM, 2006.

[83] E. M. Huang and E. D. Mynatt, "Semi-public displays for small, co-located groups," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 49–56, ACM, 2003.

[84] C. Ardito, R. Lanzilotti, M. F. Costabile, and G. Desolda, "Integrating traditional learning and games on large displays: an experimental study," *Journal of Educational Technology & Society*, vol. 16, no. 1, p. 44, 2013.

[85] L. Straker, C. Pollock, and J. Mangharam, "The effect of shoulder posture on performance, discomfort and muscle fatigue whilst working on a visual display unit," *International Journal of Industrial Ergonomics*, vol. 20, no. 1, pp. 1–10, 1997.

[86] D. G. Allen, G. D. Lamb, and H. Westerblad, "Skeletal muscle fatigue: cellular mechanisms," *Physiological reviews*, vol. 88, no. 1, pp. 287–332, 2008.

[87] Y. Rogers and S. Lindley, "Collaborating around vertical and horizontal large interactive displays: which way is best?," *Interacting with Computers*, vol. 16, no. 6, pp. 1133–1152, 2004.

[88] S. Elrod, R. Bruce, R. Gold, D. Goldberg, F. Halasz, W. Janssen, D. Lee, K. McCall, E. Pedersen, K. Pier, *et al.*, "Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 599–607, ACM, 1992.

[89] J. A. Paradiso, C. K. Leo, N. Checka, and K. Hsiao, "Passive acoustic sensing for tracking knocks atop large interactive displays," in *Sensors, 2002. Proceedings of IEEE*, vol. 1, pp. 521–527, IEEE, 2002.

[90] G. D. Morrison, "A camera-based input device for large interactive displays," *IEEE Computer Graphics and Applications*, vol. 25, no. 4, pp. 52–57, 2005.

[91] P.-A. Albinsson and S. Zhai, "High precision touch screen interaction," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 105–112, ACM, 2003.

[92] M. Bachynskyi, G. Palmas, A. Oulasvirta, and T. Weinkauf, "Informing the design of novel input methods with muscle coactivation clustering," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 21, no. 6, p. 30, 2015.

[93] A. Jude, G. M. Poor, and D. Guinness, "Personal space: User defined gesture space for gui interaction," in *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, (New York, NY, USA), pp. 1615–1620, ACM, 2014.

[94] K. Pietroszek, A. Kuzminykh, J. R. Wallace, and E. Lank, "Smartcasting: A discount 3d interaction technique for public displays," in *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, OzCHI '14, (New York, NY, USA), pp. 119–128, ACM, 2014.

[95] R. Wimmer, F. Schulz, F. Hennecke, S. Boring, and H. Hußmann, "Curve: Blending horizontal and vertical interactive surfaces," in *Adjunct Proceedings of the 4th IEEE Workshop on Tabletops and Interactive Surfaces (IEEE Tabletop 2009)*, 2009.

[96] M. Collomb, M. Hascoët, P. Baudisch, and B. Lee, "Improving drag-and-drop on wall-size displays," in *Proceedings of Graphics Interface 2005*, GI '05, (School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada), pp. 25–32, Canadian Human-Computer Communications Society, 2005.

[97] J. Schöning, F. Daiber, A. Krüger, and M. Rohs, "Using hands and feet to navigate and manipulate spatial data," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, (New York, NY, USA), pp. 4663–4668, ACM, 2009.

[98] S. Reeves, C. Varakamin, and C. Henry, "The relationship between arm-span measurement and height with special reference to gender and ethnicity.," *european Journal of clinical nutrition*, vol. 50, no. 6, pp. 398–400, 1996.

[99] S. Goel and R. Tashakkori, "Correlation between body measurements of different genders and races," in *Collaborative mathematics and statistics research*, pp. 7–17, Springer, 2015.

[100] L. Findlater and J. McGrenere, "Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces," in *CHI*, 2008.

APPENDICES

Appendix A

Materials for Experiment

A.1 Information Form for Validation Experiment

Subject No. :	________________
Gender :	________________
Dominant Hand :	________________
Weight :	________________
Height :	________________
Shoulder Height :	________________
Upper Arm Length :	________________
Lower Arm Length :	________________
Hand Length :	________________
Experiment with touchscreen (years) :	________________
Daily Exercise Level (1 - 5):	________________

Lifestyle	Example	PAL
Extremely inactive	Cerebral Palsy patient	≤ 1.40
Sedentary	Office worker getting little or no exercise	1.40-1.69
Moderately active	Construction worker or person running one hour daily	1.70-1.99
Vigorously active	Agricultural worker (non mechanized) or person swimming two hours daily	2.00-2.40
Extremely active	Competitive cyclist	≥ 2.40

Table A.1: 5 Lifestyle based on PAL level

Note: Daily exercise level is for participants to self-report how hard and how frequent they exercise daily. This scale is borrowed from Physical Activity Level (PAL).

$$PAL = \frac{TEE_{24h}}{BMR}$$

Here TEE_{24h} is defined as a person's total energy expenditure (TEE) in a 24-hour period, while BMR is the person's basal metabolic rate. Usually PAL is divided into five levels as different lifestyles

A.2 Borg10 Scale for Subjective Fatigue Report

Score	Definition	Note
0	Nothing At All	No arm fatigue
0.5	Very, Very Weak	Just noticeable
1	Very Weak	As taking a short walk
2	Weak	Light
3	Moderate	Somewhat but Not Hard to Go on
4	Somewhat Heavy	
5	Heavy	Tiring, Not Terribly Hard to Go on
6		
7	Very Strong	Strenuous. Really Push Hard to Go on
8		
9		
10	Extremely Strong	Extremely strenuous. Worst ever experienced

Table A.2: The Borg10 Scale for Subjective Fatigue

A.3 Interview Question List

1. What is your general feeling?

2. Comparing the two tasks, *Tapping* and *Draging*, which one is more fatiguing?

3. Comparing the two conditions, *Short* and *Long*, which one is more fatiguing?

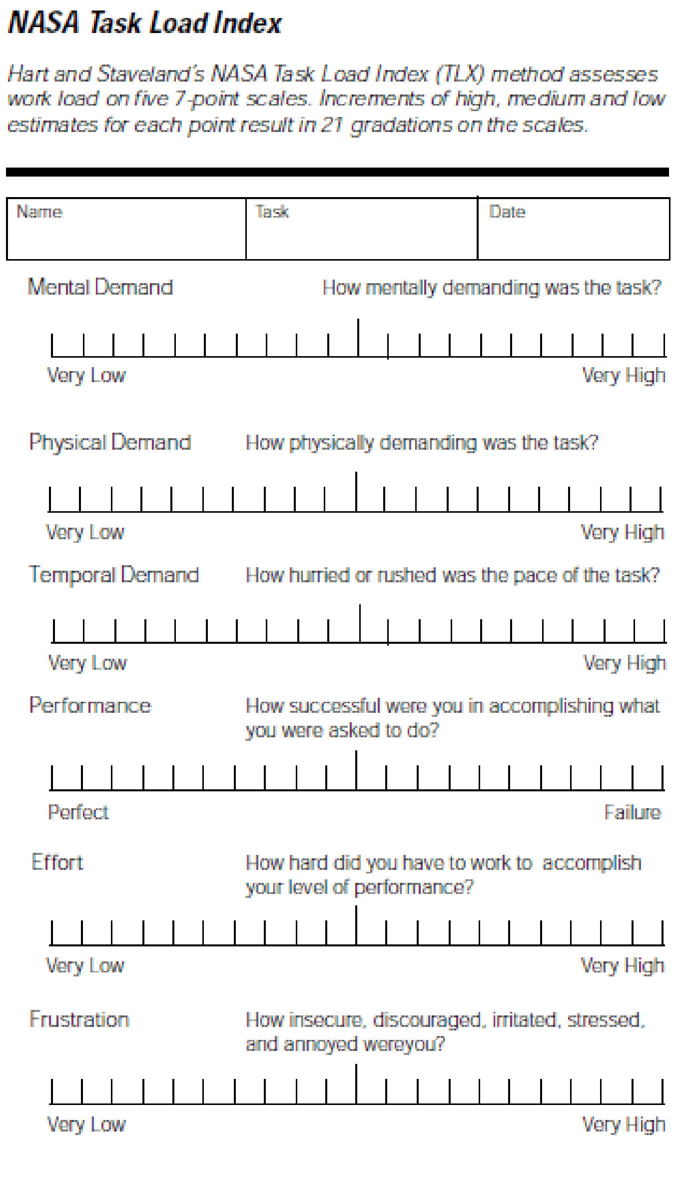
4. Are there any specific conditions that you noticed to be more fatiguing? What is it?

5. Can you totally understand the Borg10 Scale?

6. Do you have any preference to any interaction gestures?

7. To make this task less fatiguing, do you have any suggestions? What is that?

A.4 NASA Task Load Index



NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure A.1: NASA Task Load Index [2]

A.5 Required Resources

This research project aims to understand the arm fatigue effect of the large multi-touch display interactions, which requires studying and analyzing people's behaviour when interacting with large displays. Equipment and funding are required.

Table A.3 lists all the required equipment for the experiment.

Materials	Amount
Large multi-touch display	$\times 1$
Microsoft Kinect version 2	$\times 1$
Microsoft Kinect adapter	$\times 1$
Laptop with Windows system	$\times 1$
Small sandbag (2.5 lb)	$\times 1$
Talcum powder	$\times 1$
Kitchen towel	$\times 1$

Table A.3: Required Equipment for Experiments

Funding is required for the experiment. Participants recruited to participate in the experiment should be remunerated around $10 - $15 per hour. We recruited 24 participants for the validation experiment, and the experiment was generally 1.5 hours per person. Therefore, a budget for around $550 is required. In addition, we have to apply for the specific ethical approval to recruit participants.