Usability Study of Static/Dynamic Gestures and Haptic Input as Interfaces to 3D Games

Farzin Farhadi-Niaki[1], Jesse Gerroir[2], Ali Arya[3], S. Ali Etemad[4], Robert Laganière[5],
Pierre Payeur[6], Robert Biddle[7]

[1] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada, ffarh101@uottawa.ca
[2,3] School of Information Technology, Carleton University, Ottawa, Canada
[4] Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada
[5,6] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada
[7] Department of Computer Science, Carleton University, Ottawa, Canada

Abstract— **In this paper, the quality of the interaction of users with a 3D game using different modalities is studied. Three different interaction methods with a 3D virtual environment are considered: a haptic 3D mouse, natural static gestures (postures), and natural dynamic (kinetics) gestures. Through a comprehensive user experiment we compared the pre-defined natural gestures to each other and also to a haptic interface which is designed for the same game. The experiments analyze precision (error), efficiency (time), ease-of-use, pleasantness, fatigue, naturalness, mobility, and overall satisfaction as evaluation criteria. We also used user-selected ranks of importance as weight values for evaluation criteria to measure the overall satisfaction. Finally, our user experiment presents a learning curve for each of the three input methods which along with the other findings can be a good source for further research in the field of natural multimodal Human-Computer Interaction.**

Keywords— Usability study, static/dynamic gestures, haptics, 3D game, human factors.

I. INTRODUCTION

Due to the substantial growth in computational capabilities, both in computer hardware and software, human computer interaction (HCI) systems have progressed, becoming more accurate, efficient, and practical.

In general, two major aspects of HCI systems need extensive study and analysis: functionality and usability. The reason for such categorization is that while the former is often bound by machine hardware, software, and complex algorithms, the latter depends largely on another complex process: the human experience.

Recently, different aspects of HCI have been subject to research and improvement. Modalities involved in such systems often play a determining role in the outcome and experience of employing the device. Multimodal interfaces (i.e. those with different input/output methods) are becoming more popular due to their variety, efficiency, and ability to adapt to user needs [1][2]. Active systems, capable of dynamically and intelligently adapting to different scenarios are also becoming more practical. Although there have been some advances in research on multimodal interfaces, numerous questions remain unanswered and until today, very few systems have integrated multiple modalities flawlessly and effectively [3]. Among these questions are:

1) What modalities are more suited to different tasks?

2) What is the right balance of input/output methods in multimodal systems?

The first question is the main focus of this paper where different human factors when playing a 3D computer game are measured, studied and analyzed. Following our previous studies on usability of natural interfaces [1], in this paper we have focused on two input methods that try to provide a more natural interaction: gesture-based input using a 3D camera and haptic force-feedback input using a 3D mouse. For gesture-based method we have considered static hand/finger gestures and dynamic arm gestures. By "static" we refer to those gestures that are defined by a single state. The term "dynamic" is used to refer to gestures that are identified by a certain movement. This provides three separate input options, all applied to a 3D computer game, a simple slingshot game developed using Microsoft XNA. Successive to implementation, the system is tested with multiple users which provide the feedback needed to analyze the usability of such systems with respect to factors such as precision, efficiency, ease-of-use, fun-to-use, fatigue, naturalness, mobility, and overall satisfaction.

The major contributions of this study are: a) choice of natural gestures, b) system design (UI and gesture recognition) and novel use of existing API's to implement gesture recognition and haptic force feedback methods, c) usability study for gesture-based and haptic 3D mouse inputs, d) overall satisfaction analyses, directly from the users feedback, and indirectly from average of multiplying the normalized weighted satisfaction criteria, and e) presenting a learning curve for each of the three modalities.

In the following sections, the complete process of construction of our proposed system is discussed. In Section 2 a review of some key literature and usability studies in the field of gestural and haptic HCI is carried out. Section 3 deals with methodology including the UI design, natural gestures selection, gesture and haptic recognition modules along with their algorithms designed to control the UI, and

eventually more detail of the experiment process and our evaluation method. In Section 4 the experimental results are analyzed and discussed. Finally in Section 5 some concluding remarks are presented.

II. RELATED WORK

A system, capable of recognizing human gestures and providing haptic feedback, is considered a major step towards a more natural and viable multimodal system. For example, in an augmented reality system to play table tennis, it would be ideal if the user could provide the system with controlling commands using either speech or gesture and feel force feedback when they hit the ball using the racquet. In such systems, human gestures play a critical role which needs to be preferably detected, classified, and interpreted through computer vision and pattern recognition means in order to avoid "non-natural" sensors.

A. Gestures

Humans, and many other living organisms, have been employing motions of limbs for expression. The broad definitions for gestures in biology and sociology allow researchers to describe the gestures proprietarily [4]. Kendon [5] classifies a gesture as: gesticulation (impulsive movements of hands/arms during speech), language-like gestures (replace words/phrases), pantomime-like gestures (depict objects/actions), emblems (common gestures, e.g., the "V" sign for victory), and sign languages (sets of gestures/postures defining linguistic communication systems, e.g., ASL, the American Sign Language). From gesticulation to sign languages, the association with speech decreases and social parameter increases.

When interacting with computers, gestures can be utilized through vision based techniques. Hand movements and poses such as pointing, grabbing, and moving, are extremely intuitive and content rich (both spatiotemporally and perceptually), and therefore, perfect means for interaction purposes [3]. For eventual gesture utilization in an HCI system, modeling (2D, 3D, or 4D), analysis (feature extraction through region of interest), and recognition (combined features to provide the scene's information) need to be designed precisely [6].

Although static arm/hand gestures, also called hand postures (shape-based recognition algorithms), have had the main focus in gesture recognition field, recently researchers are showing more interests in incorporating dynamic gestures (temporal-based classification systems) in their study, due to broader domain of hand's dynamic actions comparing to hand postures [7].

B. Vision-Based Modality

There has been significant study in the field of vision-based modalities. The hands and line of sight (LoS) combination, as the interaction method, can lessen the fatigue compared to a one hand pointing interaction [8].

Examples by Villaroman et al. [9] are presented to demonstrate how Kinect-assisted instruction can be utilized to accomplish certain learning results in Human Computer Interaction (HCI) courses. Moreover, the authors have confirmed that OpenNI, in addition to its accompanying libraries, are adequate and beneficial in enabling Kinect-assisted learning activities.

Based on a usability evaluation, Bhuiyan and Picking [10] recommend that a gesture user interface application, titled Open Gesture (available for standard tasks, such as making telephone calls and operating the television), can improve the lives of the elderly and the disabled users by creating more independence while some challenges still remain to be overcome.

An experimental study by Ebert et al. [11], on touch-free navigation through radiological images, measured the response period and the practicality of the system compared to the mouse/keyboard control. The image recreation time using gestures was 1.4 time longer than using mouse/keyboard. However it does remove the risk of infection, for both patients and staff.

In spite of all developing and improving facts in above mentioned works, it seems there is a significant lack of studies in terms of natural gestures selection. Designing a suitable user interface for the following usability studies is also crucial. Finally, we believe that there are more usability features which need to be studied than those in above mentioned works.

C. Haptic Modality

Haptics denotes the human's sense of touch for feeling or manipulating a virtual object. Haptics has been supported by a collaboration of various disciplines such as psychophysics, neuroscience, biomechanics, robot design, modeling and simulation, and software engineering. A wide range of applications have emerged and span many fields such as product design, education, video games, graphic arts, medical trainers, and rehabilitation [12].

Regarding the main forms of feedback, haptic devices can be classified into three groups [7]: ground referenced force feedback (e.g., Phantom device [13]), body referenced force feedback (e.g., CyberGrasp [14]), and tactile feedback (e.g., CyberTouch [14]).

Considering the continuous contact of user with a haptic device in most haptic systems, the user's perception of the virtual environment is hindered, and it is also impossible for the user to feel a new tactile sensation. The latter drawback hampers the integration of tactile and force feedbacks in one haptic device [15][16]. Moreover, most haptic devices have a very limited workspace (e.g., 13cm × 18cm × 25cm in the Phantom Premium 1.0A model [17][18]). Therefore, constant contact in such a limited space impairs to incorporate rich interaction elements in an extended virtual environment. To overcome these limitations and to generate an inclusive haptic experience, Ye [3] designed an augmented reality system that employs visual tracking to seamlessly merge force feedback with tactile feedback.

To evaluate what can be profoundly achieved in creating synthetic haptic experiences, technology development and quantitative investigations are indispensable. Having this combination, preferably in a multimodal system, would help to assign the right balance of input/output methods to different tasks in order to build an interactive system

between human and machine as much natural and effective as possible.

III.   METHODOLOGY

We initially defined a series of static (Fist and Palm) and dynamic (Circle and Push) gestures to be linked to the tasks in our test game. Then we designed proper algorithms to detect our predefined arm/finger gestures using the Kinect sensor and relatively novel existing API's (OpenNI [19], NITE [20], OpenCV) and to interact with our slingshot 3D game interface (using Microsoft XNA). Other than the vision-based modality, we incorporated a haptic modality to also interact with the same game interface (Figure 1).
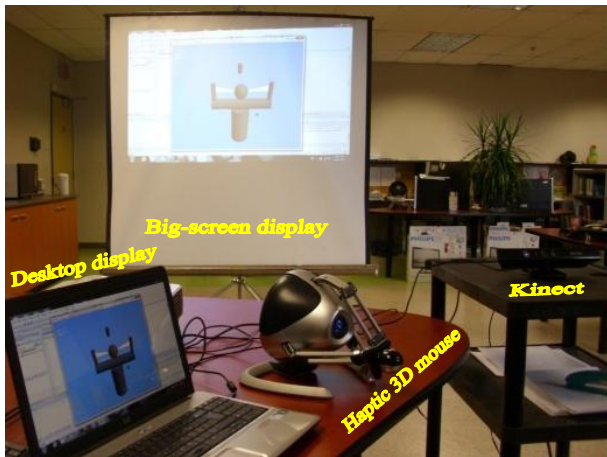
Figure 1.   The experiment devices (Falcon haptic 3D mouse, Microsoft Kinect camera, big-screen display using a high quality projector), and testing environment.

Using the static gestures, our prototype grabs the virtual ball when a fist is detected in the ball's space, then releases it when it recognizes a palm. This process is reproduced in dynamic gestures by detecting a circle drawing in the ball's space to grab it, and pushing to throw the ball towards a pin object for scoring purpose. Furthermore, using a 3D Falcon haptic device (the world's first consumer 3D touch device, which allows users to use their sense of touch in computing), the user grabs the ball (similarly to the vision-based procedure, when the 3D position of the pointer is in contact with the ball's space) by pressing the button (at this point the force feedback of the elastic band is felt), and throws the ball by releasing the button.

We then statistically compared the three input methods (static postures, kinetic/dynamic gestures, and haptic force feedback) considering the following human factors: precision (errors), efficiency (time), ease-of-use, fun-to-use, fatigue, naturalness, mobility, and overall satisfaction. As another contribution in this research we analyzed the overall satisfaction of users in two ways: i) directly from the users feedback, and ii) indirectly from average of **multiplying** *the normalized weighted satisfaction criteria* (easiness, pleasantness, fatigue, naturalness, and mobility), which have been ranked separately by users, **to** *the respected satisfaction feedback* rated from users per devices, in order

to provide a more significant and reliable rate for a practical overall satisfaction (more details in section IV.A.8.b).

Finally, our user experiment presents a learning curve for each of the three modalities by recording the time between any hit occurrences for the first 10 successful shots since the beginning of training session.

A.   User Interface

For the user interface we have designed a slingshot 3D game virtual environment (using Microsoft XNA). The design is kept as simple and minimalistic as possible, with neutral colors to reduce user error or bias (Figure 2). An option for changing the camera view is also included that users can use when they wish. The same virtual environment is controlled by the three modalities (static postures, dynamic gestures, and haptic controller) independently.
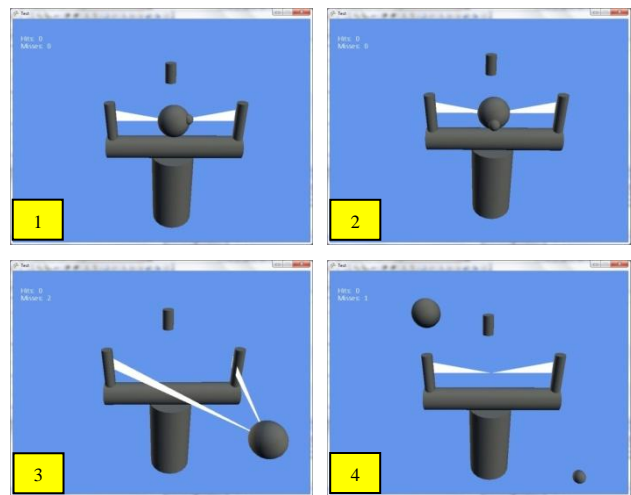

Figure 2.   Slingshot 3D game virtual environment.

B.   Vision-Based Module

Using the Kinect unit (as a commonly used vision-based input device) enables us to identify the depth of every single pixel in the frame by projecting a pattern of dots with the near infrared projection over the scene, and establishing the parallax shift of the dot pattern for each pixel in the detector. In addition, using OpenNI and NITE has helped us to secure our system with a higher stability and efficiency, and to develop a capable algorithm to recognize the arm and finger gestures.

Using the above explained method we can conserve the developing time (no need for making samples and efforts in training/testing sessions) and running time (CPU performance) for gesture recognition and user interaction compared to traditional learning-based method.

1)   Static Gestures

Considering the natural gestures to represent the tasks in interaction with our slingshot game interface, the selected static postures definition is presented in Table I.

TABLE I.   DESIGN FOR FINGER GESTURES SET.

Description	*Fisting*	*Palming*
Static gesture		
Action	*Grabbing the ball*	*Releasing the ball*

In order to recognize these gestures, as shown in Figure 3, we first needed to detect the fingertips through: i) depth thresholding, ii) contour extraction, and iii) assuming vertices of convex hull to be fingertips if their interior angles of hull corners (T_θ is the angle spanned by the predecessor edge and successor edge incident to vertex facing to the inside of contour polygon) are small enough.
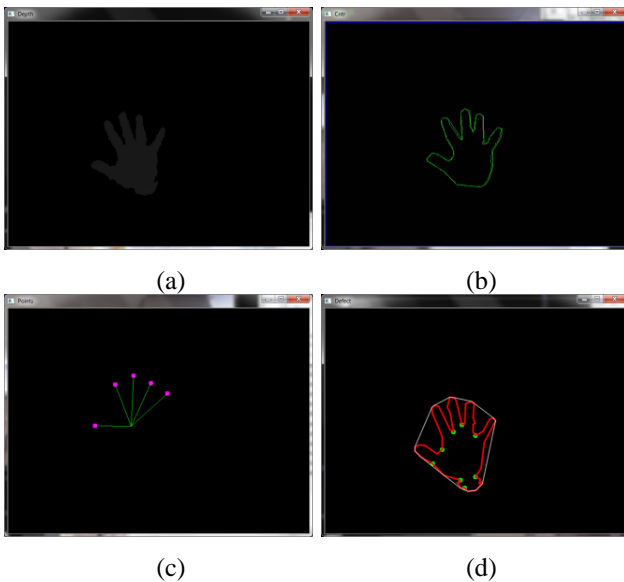

(a)   (b)

(c)   (d)

Figure 3.   Images in hand and finger detection processes, (a) depth thresholding, (b) contour extraction, (c) fingertips detection, and (d) convexity defects (depth points between two fingers).

$$T_\theta = \cos^{-1} \frac{(v_x \times v'_x) + (v_y \times v'_y)}{\|v\| \times \|v'\|} \qquad (1)$$

where $v_{x,y} = \big|$ successor of corner index (x, y) – corner index $(x, y) \big|$ and $v'_{x,y} = \big|$ predecessor of corner index (x, y) – corner index $(x, y) \big|$. Moreover, we cropped the depth map to remove the wrist out of the frame in order to improve accuracy.

Following the above process of fingers detection, the system will recognize fist (no fingers) and palm (at least four fingers), and interact with the user interface through Algorithm 1.

Algorithm 1. The algorithm controlling UI using gesture recognition.

```
1:   if (an initial Wave gesture happens) then
2:       the pointer appears (session starts)
3:       while the pointer follows the hand position (session updates)
4:           if (a grabbing gesture Fist/Circle is detected in the object area) then
5:               grab the object (save (x,y,z) based on distance deviation)
6:           if (a releasing gesture Palm/Push is detected in the object area) then
7:               release the object (transfer data)
```

2) Dynamic Gestures

Based on the available gesture recognition module in NITE, we have selected the following dynamic gestures definition as shown in Table II. During the controlling sessions of OpenNI and built-in gestures in NITE, the system will recognize circle and push, and interact with the user interface through the similar algorithm mentioned for finger gesture recognition (Algorithm 1), only replacing Fist and Palm gestures to Circle and Push.

TABLE II.   DESIGN FOR ARM GESTURES SET.

Description	*Circling*	*Pushing*
Dynamic gesture		
Action	*Grabbing the ball*	*Releasing the ball*

C. Haptic Module

We designed our haptic events controller traditionally, as shown in Table III. As the user moves the grip in three dimensions (right-left and, up-down like a mouse, but also forwards-backwards, unlike a mouse), the Falcon's software keeps track of where the grip is moved and creates forces that a user can feel. The default grip is a small spherical grip with 4 buttons on the top.

TABLE III.   DESIGN FOR HAPTIC CONTROLLER.

Description	*Click & Hold the button*	*Release the button*
Haptic 3D mouse		
Action	*Grabbing the ball*	*Releasing the ball*

D. User Experiments

Evaluation of the different interface modalities will be based on a number of criteria. These criteria are summarized in Table IV. Users have been asked to rate each criterion on a scale of 1 to 7. To ensure an unbiased sequence of used modalities during the user experiments (training and test sessions) we divided the participants into three groups (A, B, C) and permutated them such that each mode (M_1, M_2, M_3) had an equal share of sequence as first, second, and third used modalities (ABC, BCA, CAB).

1) Training Session

The experimental evaluation starts with a training session of about 15 minutes (five minutes on each device) until the participants feel comfortable to start the test session. The participants try the following four primitive tasks to get used to the application in order to run the test session precisely:

 a) *Move the pointer* (tool) around

 b) *Grab the ball* and point towards the pin (object)

 c) *Pull the ball* (elastic band)

 d) *Throw the ball* towards the pin (object)

Finally, the time between any hit occurrences for the first 10 successful shots is recorded since the beginning of training session, in order to study the learning curve.

TABLE IV.   EVALUATION CRITERIA AND THEIR REPLYING CONTEXTS.

Evaluation Criteria	Weighted Coefficients (ranked by participants)	Questions (answered by participants)	Measurements (measured by the observer)
Ease of Use	$\widetilde{E}$	How easy was it? (E)	Normalized $\widetilde{E}$ ($\widetilde{E}_N$)
Fatigue	$\widetilde{F}$	How non-fatiguing was it? (F)	Normalized $\widetilde{F}$ ($\widetilde{F}_N$)
Naturalness	$\widetilde{N}$	How natural was it? (N)	Normalized $\widetilde{N}$ ($\widetilde{N}_N$)
Pleasantness	$\widetilde{P}$	How pleasant was it? (P)	Normalized $\widetilde{P}$ ($\widetilde{P}_N$)
Mobility	$\widetilde{M}$	How flexible was it? (M)	Normalized $\widetilde{M}$ ($\widetilde{M}_N$)
Overall Satisfaction		How satisfied are you overall? (S)	Adjusted-Weighted-Overall Satisfaction ($\widetilde{S}$)
Efficiency			Required time for 5 shots
Effectiveness			Error rate (misses in 5 shots)
Learning Curve			Time rate for the first 10 hits

2) Test Session

The second part of the evaluation is a test session during which the user tries 5 shots using the haptic 3D mouse, static postures and kinetic gestures to complete a later satisfaction questionnaire. An experiment observer keeps a record of the total time (during the 5 shots) and the scores (number of hits/misses out of 5 shots) from the start to the end point.

IV. RESULTS AND DISCUSSIONS

This study has been conducted using 20 participants (11 males and 9 females). 17 participants were right-handed. They ranged in age from 15 to 45, with an average age of 29 years old. Ethics approval was secured for participant surveys. None of participants were familiar with the use of our three modalities before. The participants first read the experiment instructions and were given introductions to the tasks they were to complete during the trial.

The trial was divided into three phases:

1) Training phase: to get familiar with the applications
2) Test phase: the main process to observe the results
3) Satisfaction phase: to complete a questionnaire

A. Analyses

To analyze the effects of the different human factors being studied, one-way repeated measures analysis of variances (ANOVA) [21] is carried out in MATLAB, for the modality/input device variable (haptics vs. static postures vs. dynamic gestures). All analysis are concluded at p < 0.05 significance level and for 20 participants.

Notation: Through the following analyses of human factors, we calculate the mean and standard deviation for different variables in the forms of $M_{variable}^{human\ factor}$ (e.g., M_H^{Time} is the mean of time for haptic device) and $SD_{variable}^{human\ factor}$ (e.g., $SD_D^{Precision}$ and $SD_S^{Precision}$ are the standard deviations

of precision for dynamic and static gestures). Moreover, F(df,MS) is the test statistic (F-ratio) in which df and MS are the degree of freedom and mean square respectively for the variables (within subjects). The F-ratio is calculated using MS*variable*/MS*error* and P is the probability value.

1) Time

The analysis illustrates that for time (duration of test session), there is significant effect in the modality, $F(2,186.189) = 6.375$, $P = 0.0041$. Further repeated measures of ANOVA were used to make post hoc comparisons between each paired modalities. This reveals that the effect of dynamic gestures vs. static postures is significant on time, $F(1,370.88) = 15.76$, $P = 0.0008$, where static postures show to be significantly faster compared to dynamic gestures (Figure 4).
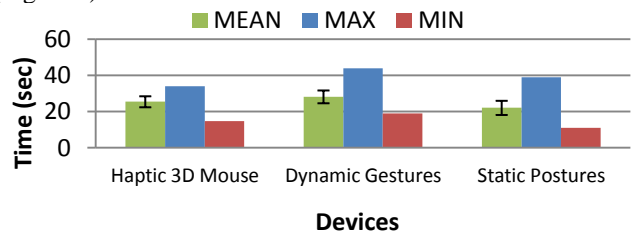

Figure 4.   Temporal statistical facts.

2) Precision

The analysis illustrates that the modality has significant effect on the precision (number of hits in five shots) when playing the game, $F(2,3.80) = 5.92$, $P = 0.0058$. Paired repeated measure ANOVA shows significant difference on scores between haptic 3D mouse vs. static postures, $F(1,4.90) = 7.11$, $P = 0.015$, and between dynamic gestures vs. static postures, $F(1,6.40) = 9.65$, $P = 0.0058$, where static postures are significantly more precise than the two other modalities (Figures 5 and 6).
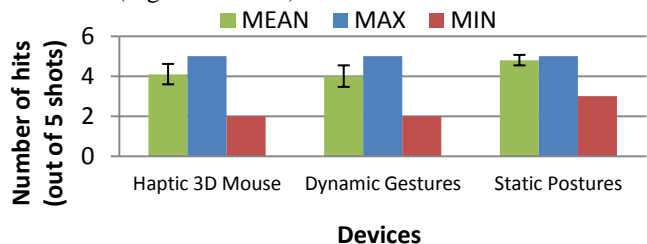
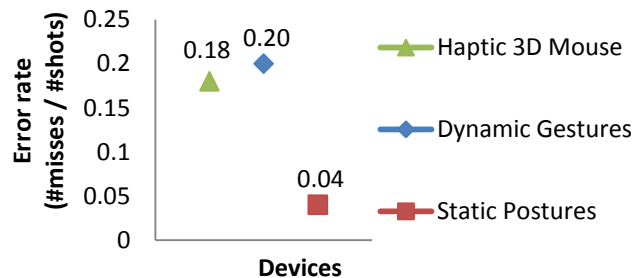

Figure 5.   Score-based statistical facts.

Figure 6.   Error rate per device.

3) Easiness

Analyzing the results indicates significant effect for the modality on easiness (how easy to interact with the UI) ratings, $F_{(2,5.55)} = 5.42$, $P = 0.0085$. Paired analyses of variances indicate that the haptic 3D mouse is significantly easier to be worked with compared to dynamic gestures, $F_{(1,3.60)} = 4.75$, $P = 0.0421$, while static postures are significantly easier than dynamic gestures, $F_{(1,11.02)} = 7.91$, $P = 0.0111$ (as shown in Figure 7).

4) Fatigue

The analysis on feedback for fatigue (how fatiguing to interact with the UI) reveals that the modality has significant effect $F_{(2,8.02)} = 6.93$, $P = 0.0027$. Paired analyses reveal that the static postures modality is significantly less fatiguing than the haptic 3D mouse, $F_{(1,9.02)} = 7.01$, $P = 0.0159$. Furthermore, it is revealed that static postures are significantly less fatiguing than dynamic gestures, $F_{(1,14.40)} = 15.54$, $P = 0.0009$ (as shown in Figure 7).

5) Naturalness

For this factor (how natural to interact with the UI), it is shown that the modality has significant effect, $F_{(2,9.62)} = 11.63$, $P = 0.0001$. Paired analyses indicate that static postures are significantly more natural than both the haptic 3D mouse $F_{(1,18.22)} = 18.95$, $P = 0.0003$, and the dynamics gestures, $F_{(1,9.02)} = 16.37$, $P = 0.0007$ (Figure 7).
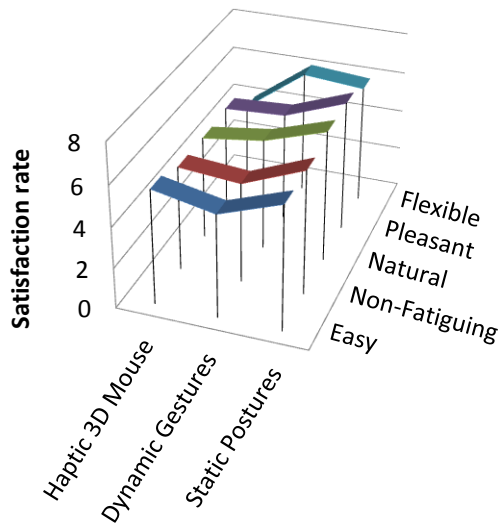

Figure 7. Direct-MEAN: rate of satisfaction criteria (direct criteria feedback).

6) Pleasantness

When analyzing the participants' feedback for pleasantness (how pleasant to interact with the UI), a similar trend to that of naturalness is observed. The effect of modality is significant, $F_{(2,9.32)} = 9.15$, $P = 0.0006$. Moreover, similar to naturalness, paired analyses indicate that static postures are significantly more pleasant than both the haptic 3D mouse, $F_{(1,15.62)} = 19.96$, $P = 0.0003$, and dynamics gestures, $F_{(1,12.10)} = 17.82$, $P = 0.0005$ (as shown in Figure 7).

7) Mobility

Analysis for participants feedback on mobility (the working space to interact with the UI), shows significant effect, $F_{(2,27.65)} = 20.72$, $P = 0.0000$. Paired analyses indicate that participants experience significantly less mobility when using the haptic 3D mouse compared to dynamic gestures, $F_{(1,46.22)} = 24.212$, $P = 0.0001$, and static postures, $F_{(1,36.10)} = 31.32$, $P = 0.0000$ (Figure 7).

8) Overall Satisfaction

a) Directly Gained (S)

In the overall satisfaction (how overall satisfactory to interact with the UI) rating obtained from participants, significant effect is observed, $F_{(2,5.12)} = 5.76$, $P = 0.0065$. Paired analyses indicate that static postures are significantly more satisfactory than both the haptic 3D mouse, $F_{(1,7.22)} = 9.62$, $P = 0.0059$, and the dynamic gestures, $F_{(1,8.10)} = 9.11$, $P = 0.0071$ (as shown in Figure 7).

b) Indirectly Gained ($\tilde{S}$)

As part of the questionnaire, the participants were asked to rank the five satisfaction criteria (easiness, non-fatigue, naturalness, pleasantness, and mobility) from 1 to 5 (the more important satisfaction factor gets the higher rank) to playing a 3D computer game (Figure 8). We name these new parameters as weighted satisfaction criteria or weighted coefficients ($\tilde{E}, \tilde{F}, \tilde{N}, \tilde{P}, \tilde{M}$). In order to refine the results and acquire an unbiased set of coefficients, we normalize the rank of satisfaction criteria per user through the following equation:

$$\tilde{X}_N = \frac{\tilde{X}}{\sum_{i=1}^{5} \tilde{X}(i)} \qquad (2)$$

where $\tilde{X}_N$ is the normalized rank of satisfaction criteria per user (Figure 9); and $\tilde{X}$ is the weighted coefficient ($\tilde{E}, \tilde{F}, \tilde{N}, \tilde{P}, \tilde{M}$), ranked from each user for satisfaction criteria.

We produce the weighted rate of satisfaction criteria (weighted criteria feedback) per device and per user (ω) through the following equation:

$$\omega = X \times \tilde{X}_N \qquad (3)$$

where X is the rate of satisfaction criteria (direct criteria feedback) per device and per user (E, F, N, P, M).

Finally, we infer a new practical rate ($\tilde{S}$) defined as Adjusted-Weighted-Overall Satisfaction (average of weighted criteria feedback) per device and per user through the following equation:

$$\tilde{S} = \frac{\sum_{i=1}^{5} \omega(i)}{5} \qquad (4)$$

The modality shows significant effect on the indirectly gained (computed) overall satisfaction, $F_{(2,0.29)} = 11.54$, $P = 0.0001$. Paired analyses reveal that similar to the overall satisfaction rates acquired directly from participants, static postures are significantly more satisfactory compared to the haptic 3D mouse, $F_{(1,0.54)} = 25.95$, $P = 0.0001$, and to the dynamic gestures, $F_{(1,0.30)} = 13.48$, $P = 0.0016$ (Figure 10).

c) Comparison

Figure 11 illustrates a comparison between the directly and indirectly gained "overall satisfaction" (S vs. $\tilde{S}$).
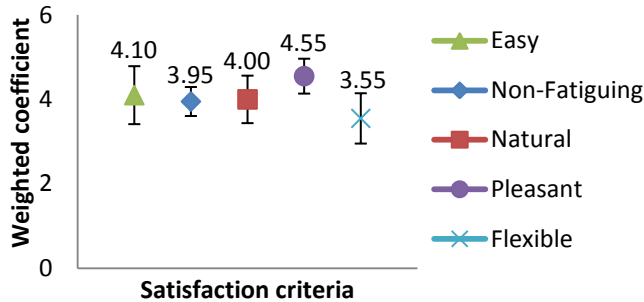
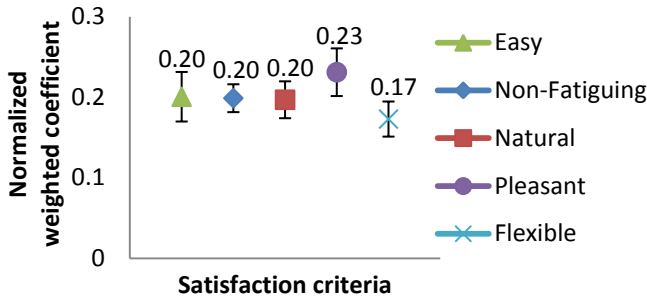Figure 8.   Rank of satisfaction criteria (weighted coefficient).

Figure 9.   Normalized rank of satisfaction criteria (normalized weighted coefficient).
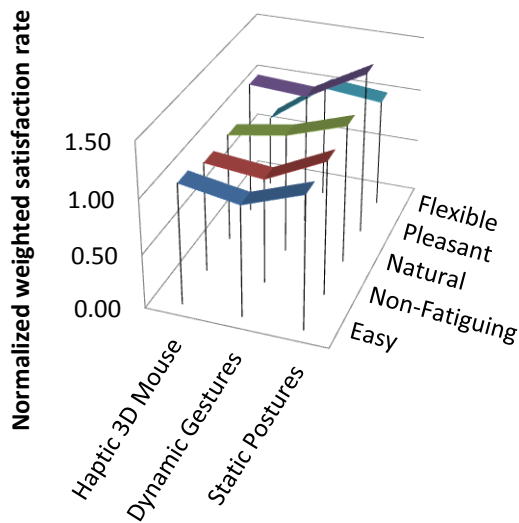

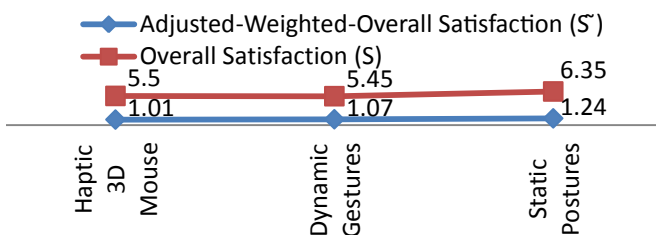Figure 10. Weighted-MEAN: weighted rate of satisfaction criteria (weighted criteria feedback).

Figure 11. Overall satisfaction (gained directly vs. indirectly).

9)   Four Primitive Tasks

Through the experiments, the participants were also asked to rate their satisfaction for the main tasks of moving the pointer, grabbing the ball, moving the ball, and throwing the ball for each device.

For moving the ball, the modality shows significant effect, $F(2,4.62) = 5.14$, $P = 0.0105$. Paired analyses indicate that it is significantly easier to move the pointer in static postures application compared to dynamic gestures application, $F(1,9.02) = 11.081$, $P = 0.0035$.

When grabbing the ball, the effect is significant also, $F(2,13.07) = 17.82$, $P = 0.0000$. Paired analyses show that all three modalities cause significant differences. In other words, it is significantly easier to grab the ball using haptic 3D mouse compared to dynamic gestures, $F(1,10.00) = 14.61$, $P = 0.0011$, and it is significantly easier to grab the ball using static postures compared to the haptic 3D mouse, $F(1,3.60) = 4.44$, $P = 0.0486$, and dynamic gestures, $F(1,25.60) = 36.30$, $P = 0.0000$.

For pulling the ball, the effect is again significant, $F(2,4.82) = 4.61$, $P = 0.0161$. It is significantly easier to pull the ball in static postures application than using haptic 3D mouse, $F(1,7.22) = 11.18$, $P = 0.0034$, and in dynamic gestures application, $F(1,7.22) = 6.16$, $P = 0.0226$.

Finally for throwing the ball, the effect is significant, $F(2,7.72) = 8.65$, $P = 0.0008$. It is significantly more difficult to throw the ball using dynamic gestures than using both haptic 3D mouse, $F(1,11.02) = 9.32$, $P = 0.0065$, and static postures, $F(1,12.10) = 10.50$, $P = 0.0043$.

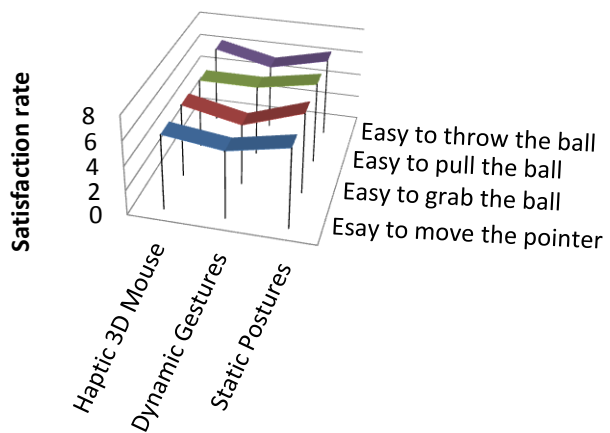Figure 12 illustrates the comparisons among the tasks and devices.

Figure 12. Satisfaction rate for primitive tasks.

10) Learning Curve

In order to study the learning curve, we have recorded the time between any hit occurrences for the first 10 successful shots since the test session starts (Figure 13). This data also presents an initial speed rate (hits/sec) during the beginning of the training session (Figure 14 shows the reversed speed).

For the average learning curves acquired for different modalities, there is significant effect, $F(2,6.38) = 4.36$, $P = 0.0286$. Paired analyses reveal that the mean learning time

per hit for the static posture is significantly less than that of the haptic 3D mouse modality, $F(1,12.72) = 5.89$, $P = 0.0381$.
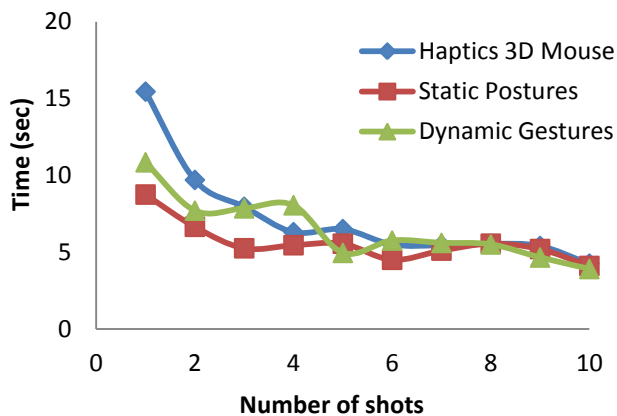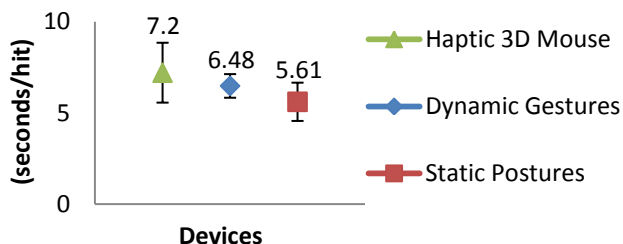

Figure 13. Learning Curves.

Figure 14. Reverse speed (average time (sec) per hit): MEAN/STD DEV per device.

For the average learning curves acquired for different modalities, there is significant effect, $F(2,6.38) = 4.36$, $P = 0.0286$. Paired analyses reveal that the mean learning time per hit for the static posture is significantly less than that of the haptic 3D mouse modality, $F(1,12.72) = 5.89$, $P = 0.0381$.

B. Extra Observations

Figures 15 and 16 present some more feedbacks from participants regarding their preferences in combination of devices to be used for this game application (the right balance of input/output methods), and their computer skills.
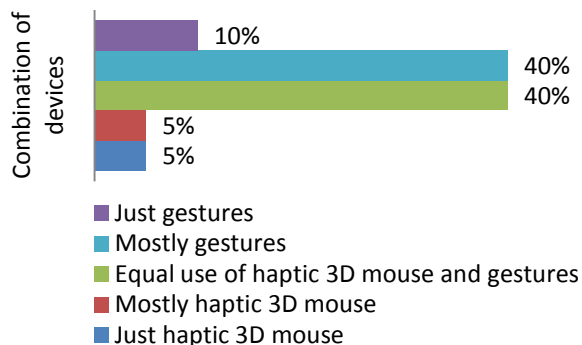

Figure 15. Users' preferences in combining the devices for this application.
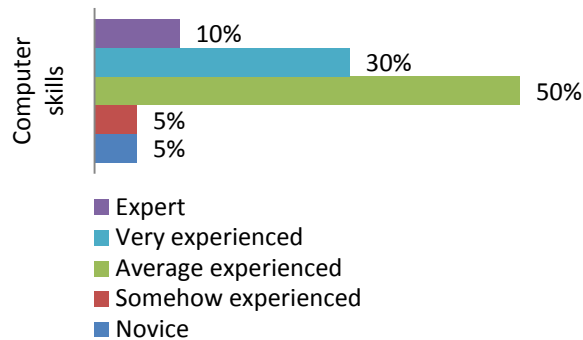

Figure 16. Participants' computer skills.

C. Discussions

Computer skills of participants widely ranged from novice to expert. As a result, it would be valid to conclude that the findings of this study are extensively applicable for practical purposes and not just tech-savvy users. Furthermore, the wide age span of the participants ensures that the findings are comprehensively general and age-independent.

According to the provided statistical analyses, we can summarize that static gestures are shown to be faster and easier than dynamic gestures while being more precise, less fatiguing, more natural, and more pleasant than both other modalities. It is faster and lighter to perform hand and arm static postures due to fewer movements involved in the process compared to dynamic gestures. As a result, they would be less fatiguing, more pleasant, and more precise.

Furthermore, continuous attachment to the haptic 3D mouse reduces the naturalness of experiences through the 3D virtual environment, while the ergonomic design and force feedback increase fatigue and error rate. The haptic 3D mouse is easier than dynamic gestures meanwhile having less mobility (space of interaction) than both other modalities. While the easiness factor was discussed earlier, the mobility can be argued in terms of the stationary nature and spatial boundaries of the haptic 3D mouse.

Overall, static postures are directly and indirectly more satisfactory than the other two modalities. Since most of the human factors showed to be superior for static postures compared to the other modalities, it is valid to expect higher rates for the direct overall satisfaction feedback as well. The weighted overall satisfaction is in complete correlation with the direct method. This indicates that intuitive feedback regarding satisfaction on a modality is a reliable means for design evaluation.

The experiments also reveal various conclusions regarding primitive tasks, namely moving the pointer, grabbing the ball, pulling the ball, and finally throwing the ball. It is shown that static postures application is more approved for moving the pointer compared to dynamic gestures application while maintaining superiority over the other two modalities when grabbing and pulling the ball. Finally, it is more difficult to grab and throw the ball with dynamic gestures compared to the other two modalities.

None of the users were familiar in applying any of these three modalities before this experiment. As a result, the obtained learning curves can be valid measures for further investigation of the modalities. The acquired learning curves indicate that the mean learning time per hit for the static posture is less than that of the haptic 3D mouse modality. This can be interpreted based on at least the two factors of naturalness and fatigue which influence the learning capabilities of participants.

According to Figure 15, direct feedback from participants demonstrates that most of them suggested "equal use of haptic 3D mouse and gestures" with "mostly gestures" as their preferred combination of modalities.

V. CONCLUSION

A 3D slingshot game was implemented using XNA, OpenNI, NITE, and OpenCV. Three modalities were defined using a haptic 3D mouse and Kinect. Two types of vision-based input methods were developed for Kinect as static and dynamic gestures. User experiments were conducted to study the different human factors associated with modalities. Precision (error) and efficiency (time) along with satisfaction criteria such as ease-of-use, fun-to-use, fatigue, naturalness, and mobility were rated in each modality and ranked independently. Static postures proved to be most efficient, precise, fun, and natural to use compared to the other modes. Furthermore, overall satisfaction was also acquired as a direct feedback from participants. Alternatively overall satisfaction was also computed by integrating the satisfaction criteria's ranks in their rates. The result of the computed overall satisfaction showed to be in complete conformity with the direct satisfaction ratings, yielding that intuitive feedback on satisfaction can be valuable means for design studies. Overall, static postures are directly and indirectly more satisfactory than the other two modalities. In terms of learning to utilize the modalities, static postures once more showed superiority against the others. Finally, it should be mentioned that even though static postures maintain superiority in many aspects over dynamic gestures and haptic 3D mouse, the latter two modalities cannot be completely ignored from being incorporated in HCI systems. This is because of the fact that dynamic gestures and haptic 3D mouse provide a vaster domain for possible gesture selection and real 4D tasks (e.g., precise following of a trajectory).

REFERENCES

[1] F. Farhadi-Niaki, R. Ghasem-Aghaei, and A. Arya, "Empirical study of a vision-based depth-sensitive human-computer interaction system," *10th Asia Pacific Conference on Computer Human Interaction*, Matsue, Japan, pp. 101-108, 2012.

[2] A. Jaimes and N. Sebe, "Multimodal human computer interaction: a survey," *Computer Vision and Image Understanding*, 108(1-2), pp. 116-134, 2007.

[3] G. Ye, "Applying vision to intelligent human-computer interaction," Ph.D dissertation, Department of Computer Science, The Johns Hopkins University, 2005.

[4] J. Nespoulous, P. Perron, and A. R. Lecours, *The biological foundations of gestures: motor and semiotic aspects*, Hillsdale, MJ: Lawrence Erlbaum Associates, 1986.

[5] A. Kendon, *Current issues in the study of gesture. In the biological foundation of gestures: motor and semiotic aspects*, Hillsdale: Lawrence Erlbaum Assoc., pp. 23-47, 1986.

[6] V. Pavlovic and T. Huang, "Visual interpretation of hand gestures for human-computer interaction: a review," *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):667, 1997.

[7] R. Vatavu, S. Pentiuc, and C. Chaillou, "On natural gestures for interacting in virtual environments," *Advances in electrical and computer engineering*, 24(5), 2005.

[8] B. Yoo, J. Han, C. Choi, H. Ryu, D. S. Park, and C. Y. Kim, "3D remote interface for smart displays," *The ACM conference on Human Factors in Computing Systems*, Vancouver, Canada, 2011.

[9] N. Villaroman, D. Rowe, and B. Swan, "Teaching natural user interaction using OpenNI and the Microsoft Kinect sensor," *The ACM Special Interest Group on Information Technology Education*, West Point, New York, USA, 2011.

[10] M. Bhuiyan and R. Picking, "A gesture controlled user interface for inclusive design and evaluative study of its usability," *Journal of Software Engineering and Applications*, 2011.

[11] L. C. Ebert, G. Hatch, G. Ampanozi, M. J. Thali, and S. Ross, "You can't touch this: touch-free navigation through radiological images," *Special Area Group of Educators*, 2011.

[12] M. A. Srinivasan, *What is haptics?*, [Online]. Available: http://www.sensable.com/documents/documents/what_is_haptics.pdf [Accessed: April 2012].

[13] Sensible Technologies, *The Phantom devices*, [Online]. Available: http://www.sensable.com/products/phantom_ghost/phantom-omni.asp [Accessed: April 2012].

[14] Immersion Corporation, *The Cyber Grasp*, Cyber Touch, Cyber Glove devices, [Online]. Available: http://www.immersion.com/3d/products/cyber_grasp.php [Accessed: April 2012].

[15] V. Hayward and M. Cruz-Hernandez, "Tactile display device using distributed lateral skin stretch," *In Proc. 8th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, vol. 2, pp. 1309–1314, 2000.

[16] C. Wagner, S. Lederman, and R. Howe, "A tactile shape display using RC servomotors," *In Proc. 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator*, pp. 354–355, 2002.

[17] M. Cavusoglu, D. Feygin, and F. Tendick, "A critical study of the mechanical and electrical properties of the Phantom haptic interface and improvements for high performance control," *Presence: Teleoperators and Virtual Environments*, 11(6):555–568, 2002.

[18] T. Massie and J. Salisbury, "The Phantom haptic interface: a device for probing virtual objects," *In Proc. Symposium on Haptic Interfaces for Virtual Environment and Teleoperator*, 1994.

[19] OpenNI Organization, *Documentation: Programmer's Guide*, OpenNI™, 2010. [Online]. Available: http://www.openni.org/Documentation/ProgrammerGuide.html [Accessed:Feb.2012].

[20] PrimeSense Inc., *Prime Sensor™ NITE 1.3 Controls Programmer's Guide*, PrimeSense Inc., 2010. [Online]. Available: http://www.andrebaltazar.files.wordpress.com/2011/02/nite-controls-1-3-programmers-guide.pdf [Accessed: Feb. 2012].

[21] S. W. Huck, *Reading statistics and research*, 3rd ed., New York: HarperCollins, 2000.