Universal Access of Information Society (2020)
https://doi.org/10.1007/s10209-020-00766-z

LONG PAPER

A comparative study on user gestural inputs for navigation in NUI-based 3D virtual environments.

Author (s). Viviana Barneche-Naya, Luis A. Hernández-Ibañez

Abstract: This paper describes the results of a case study intended to compare three different user movement paradigms (metaphoric, symbolic and natural) designed to control the visit of virtual environments for a NUI-based museum installation. It also evaluates the effects of previous expertise with 3D video games in the results for users that took part in the study. The study evaluates the performance of each movement scheme for the navigation of the environment, the degree of intuitiveness perceived by the users, and the user experience (UX). The analysis is based on the data collected in an experiment with 28 participants sorted into two groups, separating users with less previous expertise in 3D videogames from those who considered themselves as frequent players. During the experiment, the participants completed two different tasks with every movement scheme in random order. During the course of the test, the system monitored and recorded the user movements in order to extract relevant data about time to complete the task, number of collisions and time spent in a collision condition. A post-task questionnaire was carried out immediately after completion of every task. At the end of the session, users also took a test questionnaire. In addition, the authors asked users for general comments and recommendations for improvement. The results show that the natural movement scheme stands out as the most adequate for the contemplation of the virtual environment and the most balanced at a general level for the three variables considered. The symbolic scheme proved the most efficient. The natural movement scheme and symbolic scheme appear to be the most appropriate to navigate such digital environments as those present in museum installations for any kind of user.

Keywords: User Experience · Natural User Interfaces · Virtual navigation · Virtual Museum.

Author
Viviana Barneche-Naya
videaLAB – Universidade da Coruña (Spain)
e-mail: viviana.barneche@udc.es

Author (✉)
Luis A. Hernández-Ibáñez
videaLAB – Universidade da Coruña (Spain)
e-mail: luis.hernandez@udc.es

1 Introduction

The emergence of the New Museology [1-4] entailed a radical redefinition of many concepts related to the classical idea of a museum. Museums have evolved from being just static exhibitions to provide their visitors with new and exciting learning experiences [5]. To attain the active user participation, museums employ different strategies, including the use of multimedia contents, computer simulations, immersive virtual reality environments, etcetera, leading to new museological forms such as virtual interpretation centres and virtual museums.

Undoubtedly, the application of interactivity, enriched contents and novel technologies are valuable tools to help museums to achieve their actual mission to preserve and disseminate knowledge.

Many museum installations use natural interaction. This is frequently seen in science exhibitions, but it can also be found in many others, including those related to historical heritage. Many of these installations currently use natural interaction, aiming to avoid complex interaction schemes, thus offering the user a more intuitive and comfortable operation, especially compared with handling traditional input devices. Natural interaction can take many forms [6]. Many are designed to present contents in environments where information is readily available, while others also include a ludic component to reward such effort.

Among those types of natural user interfaces, Nintendo Wii and Microsoft Kinect offer good popular examples of gesture and motion tracking technologies. The use of this kind of controller has transcended the field of video games to enrich many other disciplines where virtual environments are explored by natural interaction [7].

Many experts in the historical heritage field want to go beyond visual and virtual representation of the past. They want to depict and interpret ancient buildings past just the artistic and technical aspects to offer vivid experiences. In recent years, many interesting examples of the use virtual models have been developed for their use in museums [8, 9].

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature’s AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s10209-020-00766-z]

Technology can now describe archaeological hypotheses and put them in a context the public better understands.

Nowadays, there is a tendency to use game engines such as Unreal Engine and Unity to develop such installations. Game engines are very powerful tools for interactive, real-time visualization, making possible a continuous feedback between the user and the virtual environment. Game engines are sometimes used with a depth camera (i.e. Kinect) to display visual recreations of archaeological reconstructions. [10-15].

This kind of museum installation demands good walkthrough paradigms for exploring the space and contemplating the environment and the objects on display prior to enabling further interactions [16]. In this sense, we must highlight the research by Bowman, Koller & Hodges about the analysis and evaluation of travel techniques for use in immersive virtual environments [17].

The aforementioned examples of digital reconstructions make use of different movement paradigms to navigate the virtual environment, although they don't provide a further analysis regarding the intuitiveness and ease to use of the set of movements and body gesture chosen.

This paper is based on previous studies [18, 41] carried out by the authors. We initially tested and analysed six walkthrough paradigms for virtual environments using Kinect based natural interaction. All the schemes tested in that experiment showed pros and cons and yielded different outcomes in terms of ease of use, intrusiveness, and interpretation of users' intentions, comfort, and precision.

Among the aspects considered when designing NUI interfaces, the heterogeneity of the users is one of the most important. Museum visitors can be very diverse in age, education, and previous experience with digital interfaces and contents. They can also come from different cultures and speak different languages. If we consider their previous experience with such technologies, some studies suggest a positive impact of previous experience playing 3D video games in the ability of participants to perform those tasks related to navigation and interaction inside the virtual world [19,20]. In the case of a NUI-based installation, a lack of previous experience may result in a handicap for this user group if they tend to focus their attention in the control of the system instead of the enjoyment of the virtual visit.

This paper describes the results of a case study of the application of three different movement paradigms in a NUI museum installation and evaluates the effect of user's previous expertise playing 3D video games. The collected research data should contribute to the study of paradigms of navigation inside virtual architectural environments such as those used in the field of virtual archaeology.

2 Objectives

The objective of this work is to identify which of the three movement paradigms analysed (metaphorical, symbolic and natural) is the most suitable for museum visitors, both for expert and non-expert users with regard to their previous exposure to 3D video games.

The study compares the user's performance in navigation tasks, the perceived intuitiveness of the interface and the user experience (UX) of 28 users separated in two groups based on their previous expertise

The test-bed used consisted in a virtual interactive recreation of a *domus olearia* (a Roman country house and olive farm dedicated to oil production). This recreation rebuilds the existing archaeological remains of an ancient site near Seville (Spain) dated from the fifth century of our time. It also includes some interesting mosaics found on the excavation (Fig. 1).

The virtual reconstruction of the villa is also used to put in context and facilitate the interpretation of the mosaics, whose actual remains are displayed in the museum. The visitors can contemplate the fragments found in the dig along with their original appearance, and their placement in the floors of the different rooms of the house. The installation seeks to promote the experiential value of the virtual visit by means of the perception of the spaces and interaction with cultural multimedia content present in the virtual museum, creating an educational atmosphere.

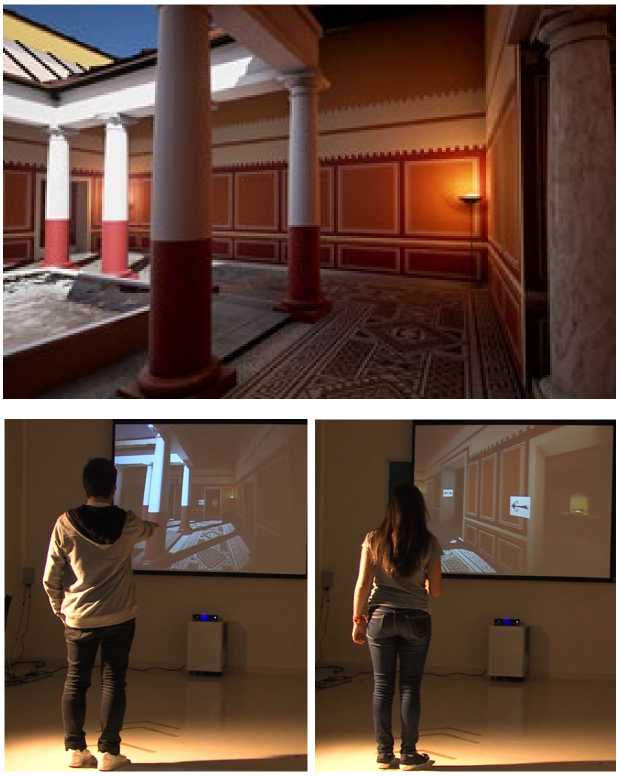


Fig.1 Virtual *Domus*: Natural interaction in the digital reconstruction

Visitors used natural interaction schemes to explore the digital reconstruction with a Kinect v2 Sensor designed for video game interaction, the device detects players' presence, body gestures and motion.

3 Methodology

This study adopted a mixed-method research approach, combining quantitative and qualitative methods to achieve complementary results [21]. The control interface of the installation uses the depth camera of the Kinect attached to the system to read and interpret different user actions depending on the interaction model chosen. The authors gathered quantitative data that includes the time spent in completing the tasks, collisions with the virtual walls and the Likert values for the users' responses to the questions regarding different aspects of the experience. Furthermore, the study obtained qualitative data, including videos, notes, "think aloud" feedback and open interviews with questions regarding aspects such as the ease to contemplate the villa, the mosaics and other objects that could be found during the walkthrough.
The study was carried out strictly following the guidelines of the Ethical Code for Research of the University of A Coruña (*Comité de Ética da Investigación e da Docencia*).

3.1 Movement paradigms

Several studies in the HCI literature explore the cognitive mechanisms that underlie in the intuitiveness of the interaction [22-25] Depending on the conceptual relations to the everyday gestures that they mirror, three different models, and their corresponding movement schemes can be established:

- Metaphoric: Actions that evoke the desired behaviour of the system by the existence of a correspondence and similarity. In a metaphoric interface, there is a mapping of concepts and operations between two domains. In this case, it's the virtual world and the reality so that an interaction suggested by the metaphor source domain corresponds to the execution of the application implementing the metaphor target domain. [26] (i.e. moving a hand, left and right in the air to browse a sequence of pictures in a projection screen).
- Symbolic: In symbolic natural interaction, their visual, aural, and maybe in the future, touch sensitive clones represent the objects. They are naturally manipulated, but they are still representations and not real things [27] (i.e. driving a virtual car by moving hands to control a virtual steering wheel).
- Natural: The user applies gestures for the exact same action in the real world. (i.e. the action of grabbing to handle a virtual object.)

Based on these models, three movement schemes were implemented corresponding with every interaction model. Table 1 illustrates those schemes according to the cognitive mechanism that they activate.

	Metaphoric	Symbolic	Natural
March	Point forward	Lean forward	Step forward
Turn	Point sideways	Twist upper body	Twist upper body
			

Table 1 Cognitive mechanism and movement paradigms.

- Metaphoric: Point forward/sideway with arm: This approach uses the movement of the user's arm to control both displacement and orientation. In order to explain the functioning of this paradigm to the users, they should imagine themselves holding the leash of a guard dog that should lead them in the direction that they expected to go. Users found this metaphor very clear and easy to embrace.

  The player controller analyses how much the user lifts his or her hand and measures the angle formed by the wrist and the elbow. Both in the horizontal plane (yaw) relative to the direction to the Kinect device, which is used for turning, and the inclination in the vertical plane (pitch) relative to the vertical axis, which is used to move forward and control the speed. An idle arm pointing to the floor means a zero angle in both directions.

  The pitch angle may be negative, thus allowing for backward displacement by pointing the arm just slightly backwards.
  By lifting his or her arm, the user increases or decreases the speed. The user can change his or her orientation at the same time by pointing sideways with the same arm. The user can turn and control the displacement speed simultaneously.

  While the user points his or her hand left or right for steering, the deviation from the initial idle angle determines the magnitude of the turn in radians per second up to a maximum angular speed. The arm's sideways movement must go past a threshold angle to start the turning mode.
- Symbolic: Lean forward/ Twist upper body: Users are told to move as if controlling a Segway-like vehicle. The user is instructed to lean forward slightly to begin the walk, now at the max speed, and straighten to stop. The system analyses the vector with origin in the base of the neck pointing forward and compares it to the horizontal plane. Again, a postural threshold value is considered, and the walk begins once the user leans past this threshold.

The user can change his or her orientation by twisting the upper body, following the natural rotation that occurs while changing the walking direction. The system measures the angle between the vector that connects both shoulders and the screen plane and applies an angular speed to the user's camera based on that angle. Again, there is a threshold angle and a maximum angular speed.

The magnitude of the effect of every gesture into the user's movement depended on several coefficients (gesture amplitude, threshold angles, minimum displacements, etc.). Those coefficients were determined in previous usability tests.

- Natural. Step forward/Twist upper body: In this third approach, users explore the virtual building by using movements similar to those used in a real walk. In this scheme, user may increase or decrease the walkthrough speed by stepping forward or backward and turn by twisting the upper body clockwise for a right turn and counter clockwise to steer to the left. The system considers the location of the starting point as the initial distance. It takes into account a threshold distance to and from this point as a safe area for the user to stand before moving to avoid unwanted displacements.

  When the user steps forward, thus exiting this safe threshold distance and approximating to the Kinect device the distance from the starting position is divided by the threshold distance to use it as an intuitive ratio between the safe zone and the area travelled. This helps to calibrate the system for different kind of users, establishing threshold distance close their approximate step size. One, two or three steps have meaningful influence in the speed of the displacement: one to exit the safe zone and begin moving, two to increase speed and three to reach maximum speed. To stop, the user has to step back to the start position. A step back from the start position initiates backward movement.

  The user can change his or her orientation by turning the upper side of the body. This movement combines with the previous one so the user can also turn and control the displacement speed simultaneously. The system measures the orientation of the vector corresponding to the neck joint of the user's skeleton, which points in a direction normal to the chest plane in the idle pose. The system evaluates the variation of the angle of this vector rotating around the z axis, namely the twist angle.

A previous test carried out with another set of participants determined the values required to configure the different parameters involved, such as thresholds, angular and linear speeds, etcetera. This test permitted to calibrate the system prior to the experiment for every movement paradigm.

3.2 Participants

The set of users involved in the experiment was selected to obtain a sample similar to the profile of the potential museum visitors, with a bias to the population segment of young and adult people with good educational level. [28].

The authors contacted and recruited participants from the university's students, faculty and other staff. Twenty-eight participants (14 male and 14 female), with ages ranging from 17 to 54, with an average of 25 years. Participation data per age segment were as follows: young, 17-24 years old (67.8%); adults, 25- 45 years old (25%); and adults, 45-plus years old (7.2%).

We created two groups segmented for self-reported expertise on video games. Fourteen people composed each group to ensure statistically significant evidence in the sample size [29]. A group integrated by users without or with little previous experience in 3D videogames (43% male, 57% female), and a second group composed by experienced users (57% male, 43% female). Each session lasted approximately 30 min for each individual.

3.3 Experiment design

Before the beginning of the test, the moderator explained the mechanics of the session to the participant and required the user to fill out a brief demographic questionnaire (age, gender, educational level, self-reported expertise in 3D videogames).

The experiment set consisted of a dimly lit room with a projection screen with a Kinect sensor underneath, and marks on the floor; one of them indicating the starting point of the experience, located 3.80 m in front of the screen. In this range, the depth and skeleton views from the sensing device cover the entire user's body.

Fig.2 Virtual *Domus*: Proposed path for task #1

Participants had to complete two different tasks for every paradigm tested. In both cases, the values of time spent, number of collisions with walls and objects, and number of frames in collision state were measured. The tasks that users had to carry out were as follows:

- Walkthrough to reach a specific place in the house: The purpose of this task was to gather metrics to analyse the performance of the navigation and the ease to use of the system. The user began in front of the main door of the *Domus*, then he or she had to cross the vestibule to the main atrium, surround it and exit the atrium through a door located in the side opposite to the entrance leading to the cubiculum containing the lararium (Fig.2).

  This task required the users to perform actions such as following corridors and crossing doorways. Such actions involve an accurate perception of distances between the viewpoint and the different objects present in the virtual environment in order to avoid collisions.

- Walk along a narrow and winding path: The second test measured the accuracy of every paradigm. This one required a more precise driving, since the user had to pass between two rows of objects and pass around another object to finish stopping in a given place. In this task, users had to perform more complex manoeuvers to avoid colliding with objects while applying a more precise control to carry out full turns.

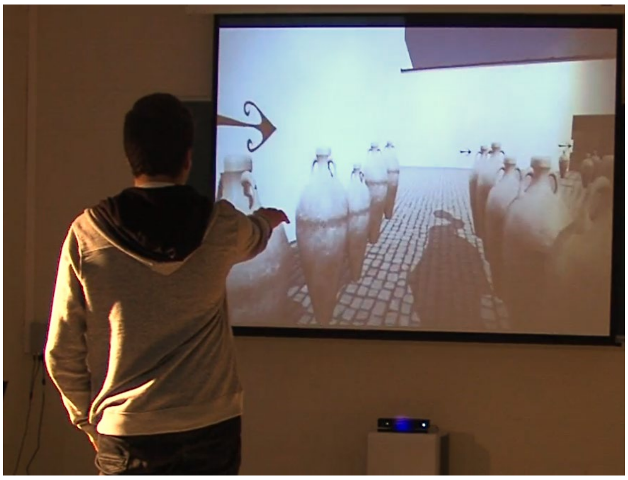


Fig.3 Virtual *Domus*: User performing task #2

The subjects tried the different movement schemes in random order to counter the effect of an increase of their navigation skills through repetition (Fig.3).

During the course of the test, the system monitored and recorded the user moves to extract relevant data about time to complete the task, number of collisions and time spent in a collision condition. A post-task questionnaire was carried out immediately after completion of every task.

At the end of the session, users also took a test questionnaire. Post-task questionnaires and test questionnaires, together with the record of the comments made freely by the participants, provided a good source for subjective data.

3.4 Metrics

3.4.1. User navigation performance

The experiment compared the performance in users' navigation for the three movement paradigms, evaluating the influence of their previous expertise with 3D video games.

Speed and skill were the variables that were taken into account. The first one was measured considering the time spent to complete task 1. In order to measure the navigation skill, the system counted the number of frames of the simulation that every user spent colliding (even tangentially) to walls or objects. Considering a frame rate of 30 fps, we can obtain the time in collision state and the percentage of the task performed in collision state in relation to the total task time.

Furthermore, since this study deals with a 3D interactive installation, it is important to include certain measurable characteristics of the quality of navigation [17] as a mean to solve certain tasks such as velocity, accuracy and spatial awareness.

3.4.2. Interface intuitiveness

One of the goals of natural interface design is to develop systems that interfere as little as possible with the user's experience. When the interface behaves as we expect it to [30], it responds to user's desires in a fluent, comfortable and confident way, allowing the user to focus attention on the experience instead of on the interface. We used two variables to measure the intuitiveness of every paradigm, namely *attention* and *reliability* (confidence on a trustworthy and consistent behaviour of the paradigm to follow the user's intentions).

3.4.3. User Experience

To measure the user experience in this test, we used the approach described by Tullis and Albert [31], where UX is the combination of all behaviours and attitudes people have while interacting with an interface. These include and go beyond traditional usability [32] and broader metrics dealing with users' attitudes and perceptions. In order to follow this approach, it is necessary to combine both objective and subjective measures to enable satisfaction analysis as a "subjective sum of the interactive experience" [33].

Beginning with the concept of usability [34], [35], we measured effectiveness by the completion rate, the number of collisions and the number of frames that the system registered where the user collided with walls and objects, which can be expressed as time in collision state.
Efficiency was measured by the mean time taken to complete each task in seconds.

We measured user satisfaction through the users' responses to questions related to several aspects of the experience, such as: ease of learning, physical fatigue, user comfort and degree of motivation and pleasure with the experience

[36]. Furthermore, the emotional factor influences the potential of learning new skills and acquiring new knowledge, which are key points of this kind of installation.

3.5 Data collection

Data were collected in several sessions. During those sessions, two moderators observed and interacted with the users as they were completing the tasks using the different movement paradigms. We videotaped the users while they were performing the different tasks and noted their spontaneous comments about their impressions related to the experience. During each task, the system registered the tasks' completion times and frames in collision state.

After each task and the end of the session, the participants rated the movement paradigms on a 10-point rating scale.

Table 2 Post-tasks and test questionnaire

	1	10
Q1- How difficult or easy did you find the task? (Post-Task#1)	Very difficult	Very easy
Q2- Did the system respond accurately to your actions(Post-Task#2)	Not accurate	Very accurate
Q3- How much of your attention was devoted to contemplate the environment and how much was given over to control the system?	No attention put on the environment	Most attention put on the environment
Q4- Is it easy to observe all the objects and mosaics?	Very difficult	Very easy
Q5- How well does the system interpret your intentions??	Very badly	Very well
Q6- Do you feel tired or relaxed after using the system?	Very tired	Very relaxed
Q7- Did you find the system comfortable?	Very uncomfortable	Very comfortable
Q8- Did you find the experience enjoyable?	Not enjoyable	Very enjoyable

- Post-Task#1: Based on Single Ease Question [37]: The user rated the difficulty of the activity they just completed.
- Post-Task #2: The user rated the accuracy of the system for the activity they just completed.

Subsequently, users completed a general test questionnaire that collects data about the perception of the experience as a whole. It includes five subjective measures:

- Reliability: how well the system interpreted user's intentions.
- Attention: the level of attention they put into the experience instead of on controlling the system.
- Physical fatigue: the level of fatigue
- User comfort: the level of comfort
- User pleasure: how exciting and enjoyable the experience was.

In addition, the authors asked users for general comments and recommendations for improvement.

Table 3 Metrics: variables and data collection

	Variable		Collection method
User Navigation Performance	Velocity		Task1 completion time.
	Navigation skill		Percentage of time in collision state for Task1
	Accuracy		Task2 completion time. Likert scale: Post-task2 question
	Spatial Awareness		Likert scale: test questionnaire
Intuitiveness	Attention		Likert scale: test questionnaire
	Reliability		Likert scale: test questionnaire
User Experience	Effectiveness		Tasks 1-2: completion rate, Tasks 2: Number of collisions and percentage of time in collision state
	Efficiency		Tasks 1-2: completion time.
	User Satisfaction	Ease of learning	Likert scale: post-task1 question
		Physical fatigue	Likert scale: test questionnaire
		User comfort	Likert scale: test questionnaire
		User pleasure	Likert scale: test questionnaire
		Comments Videos.	General observation. Notes.

4 Results

This study used the IBM SPSS 26 statistics software. In order to know if user's previous experience had a significant effect on every variable being analysed, the authors applied statistical inference with an contrast statistical significance level of α=0.05.

Particularly, and with regard to the use the results for the design of museum installations, this work tried to detect if any of the paradigms could be especially adequate for all users independently of their previous experience.

Since the sample is small, with a non-normal distribution, this study used a nonparametric Kruskal-Wallis test. Samples were analysed as follows:

1- Comparative analysis of the results obtained by group 1 (No experience/Casual Player) and group 2 (Frequent Player) performing every movement. (G01-PP/G02-PP, G01-LT/G02-LT, G01-ST/G02-ST)

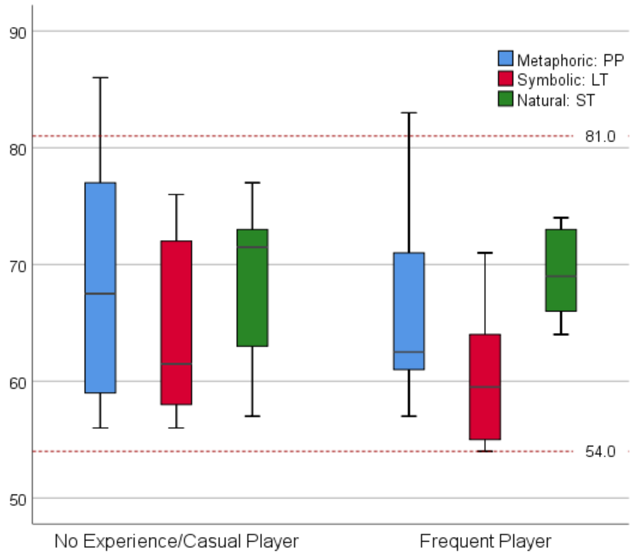
2- Comparative analysis of the results obtained by all users, by movement. (PP-LT/PP-ST/LT-ST)

3- Comparative analysis of the results obtained by every user group for every movement scheme. (G01-PP/G01-LT/G01-ST/G02-PP /G02-LT/G02-ST)

4.1 User navigation performance

4.1.1. Velocity: Mean values of times to complete task 1 were calculated for the three movement paradigms and for both user groups (Fig.4). From the analysis of the influence of previous expertise for every movement scheme on the time spent, we did not find statistically meaningful differences (α=0.05) between both user groups using the same paradigm (PP: p=0.352; LT: p=0.146, ST: p=0.817). The symbolic scheme (LT) obtained the best times for the task for both user groups.

Some differences appeared (α=0.05) in the time taken for task 1 if we compare the symbolic with the metaphorical schemes (p=0,04) and comparing it with the natural scheme (p=0,00059). No differences were round between the metaphorical and natural schemes. (p=0.285).



Gestures	User's previous skills	N	Mean	SD
Metaphoric: PP	No Experience/Casual Player	14	69,1	9,4
	Frequent Player	14	66,0	7,6
Symbolic: LT	No Experience/Casual Player	14	64,2	7,3
	Frequent Player	14	60,5	5,7
Natural: ST	No Experience/Casual Player	14	69,5	5,6
	Frequent Player	14	69,0	3,4

Fig.4 Task1 completion time by movement scheme and user group

The difference detected for all users between the symbolic (LT) and the rest of movement schemes is confirmed by the comparative analysis considering previous expertise. There are differences with a significance α=0.05 between symbolic (LT) and natural (ST), both for group 1 (p=0.026) and group 2 (p=0.001).

4.1.2. Navigation: The natural scheme (ST) was the one with smaller percentage of time in collision state (4.3%) for both groups.

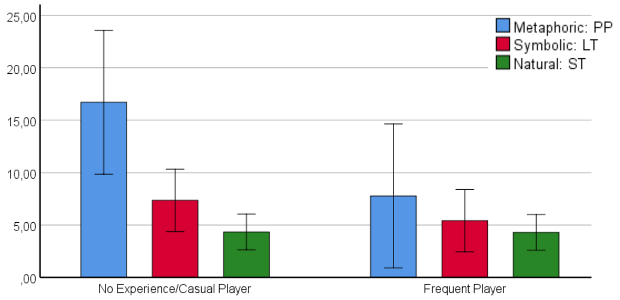


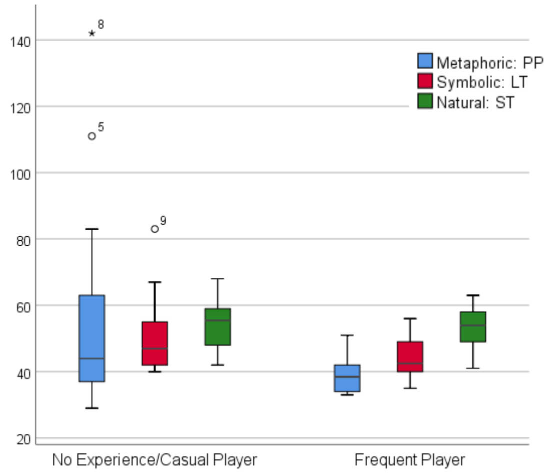
Fig.5 Percentage of task1 collision time by movement scheme and user group

The metaphoric scheme (PP) reached the highest percentage of time in collision state both for group 1 (16.7%) and group 2 (7.8%). In the analysis by movement type, no statistically significant differences were detected (α=0.05) in time spent in collision state between the two user groups (Kruskal-Wallis PP: p=0.183; LT: p=0.290, ST: p=0.505).

Comparing the three movement schemes, and considering all users, we found some significant differences between the natural and the metaphoric schemes (p=0.004), but no meaningful differences among the rest of movements.

Making the analysis by movement considering previous expertise, differences were found in group 1 between metaphoric and natural schemes (p=0.015), and no significant differences for group 2.

4.1.3. Accuracy: We calculated the average time to complete task 2 (Fig.6) for the three movement schemes and user groups. If we compare globally, considering all users, there are meaningful differences (α=0.05) between symbolic and natural schemes (p=0.033), but no difference for the rest of combinations.



Gestures	User's previous skills	N	Mean	SD
Metaphoric: PP	No Experience/Casual Player	14	57,6	32.8
	Frequent Player	14	39.6	6.2
Symbolic: LT	No Experience/Casual Player	14	51.1	11.9
	Frequent Player	14	44.1	6.4
Natural: ST	No Experience/Casual Player	14	54.0	7.7
	Frequent Player	14	52.8	6.3

Fig.6 Task2 completion time by movement scheme and user group

The symbolic scheme obtained the best time for task 2 by users in group 1, while metaphorical helped to obtain the best time for group 2. Nevertheless, considering the responses to the different schemes inside each group, hence taking into account previous user expertise, group 1 did not mark differences, but group 2 presented meaningful differences between natural and metaphoric schemes (p=0,0002) and natural and symbolic (p=0,017).

After completion of task 2, users were asked to fill out a questionnaire to rate the accuracy of the system in the activity they just completed.

Generally, the percentage of positive answers (score 6-10), defining all movement scheme as accurate was high. Comparing the score median for accuracy for all users, no significant differences among movements were found.

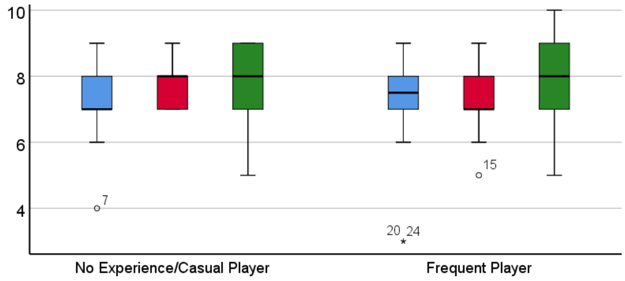


Fig.7 Accuracy by movement scheme and user group

The natural scheme (ST) was considered as the more precise, obtaining the best scores (9 y 10 points) for group 1 (71.5%) and group 2 (58.10%).

Analysing by movement type, the metaphoric scheme was the only one which obtained some low scores on accuracy (below 5) for both user groups. (7.1% for group 1 and 14.3% for group 2).

4.1.4. Spatial awareness: In order to evaluate this aspect, the authors analysed the subjective answers obtained from the open interview about the easiness to observe the rooms, objects and mosaics for each movement paradigm. Concurrently, (Fig. 8) we analysed the responses about the degree of conscious attention put into enjoy the museum experience (6-10 points) instead of controlling the system (≤ 5 points)

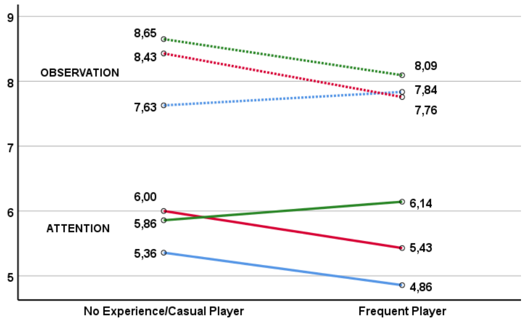


Fig.8 Observation and attention by movement scheme and user group

The ease of contemplating the objects and the scene received a good evaluation by both user groups in three paradigms. Contemplation of the mosaics, presented more difficulties for PP than for LT and ST.

The natural scheme (ST) was the one which obtained the best values for this variable.

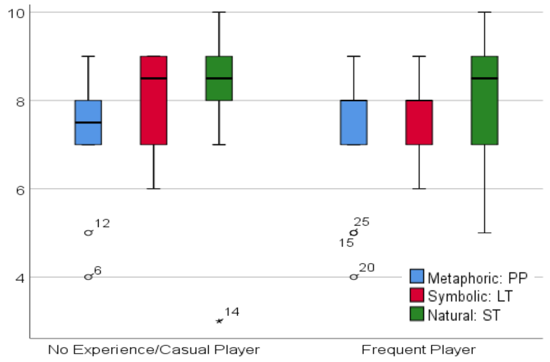
Comparing the ease of contemplation of the virtual environment among the three movements for all users, there are some differences with a level of significance α=0.05 between the natural and the metaphoric schemes (p=0.010) and natural with symbolic (p=0.011).

Regarding the level of attention devoted to contemplation of the virtual environment, a meaningful percentage of neutral values appears (5-6 points). There is a small difference, not statistically important, between group 1 and 2. The first considered the symbolic scheme (LT) as the one which allowed for a better concentration in the environment, while group 2 gave slightly better to the natural scheme (ST).

4.2 Intuitiveness:

4.2.1. Attention: Generally speaking, more than 30% of users considered that they put more or less the same degree of attention to observe the environment than to control the movement. (5-6 points: PP=32.1%, LT=42.9%, ST=35.7%). Considering the degree of previous expertise, the symbolic scheme (LT) was the one that best permitted the users to focus their attention in the museum experience, whereas the natural scheme (ST) obtained the best results for group 2. In both cases, the percentage of positive values (≥8 points) reached 42.9%

4.2.2. Reliability: In order to analyse the capacity of the system to adequately interpret the user's intentions, the medians of the responses were calculated and compared for the three movement paradigms by user group (Fig.9).



Gestures	User's previous skills	N	Median	SD
Metaphoric: PP	No Experience/Casual Player	14	7.50	1.31
	Frequent Player	14	8.00	1.59
Symbolic: LT	No Experience/Casual Player	14	8.50	1.02
	Frequent Player	14	8.00	0.86
Natural: ST	No Experience/Casual Player	14	8.50	1.68
	Frequent Player	14	8.50	1.59

Fig.9 Reliability by movement scheme and user group

The symbolic (LT) and natural (ST) schemes obtained the better scores for group 1, whereas the natural scheme (ST) got the best values for group 2. The Kruskal-Wallis test did not present significant differences considering previous expertise in any of the three movements.

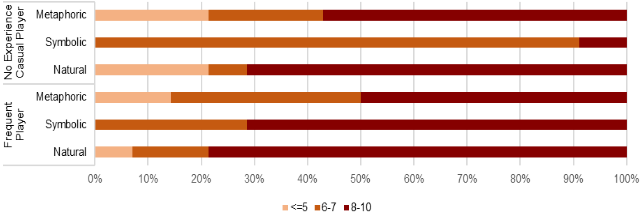


Fig.10 Reliability by movement scheme and user group

Completing the previous analysis, Figure 10 displays the degree of reliability on interpretation of users' intentions for both user groups and for every paradigm.
The natural scheme (ST) obtained the best scores (8 to 10 points) for 71.4% of users from group 1 and for 78.6% of those from groups 2.

4.3 User experience

4.3.1. Effectiveness: All the tasks were completed successfully. Data collected by the system during the execution of the tasks indicate that sometimes as the user advanced, he or she stuck laterally to the walls and continued this way, instead of returning to the centre of the path, or the user stuck frontally against obstacles for a while.

Collisions are a kind of unintended action a user makes while trying to do something on an interface even though the goal is correct. Norman used the term "slips" for these kinds of actions [38].

In our case, the analysis of the number of collisions in task 2 and time in collision state (Fig.11) helped to figure out how quickly the user learned to drive the system properly.

We used the data related to collision obtained in task 2 only. Since it is a collision test. It works better to check the influence in the interaction with the virtual environment.

The Kruskal-Wallis test shows meaningful differences in the percentages of time in collision state for the natural scheme (ST) related to previous expertise between group 1 and group 2 (p=0.024), without significant differences for the rest of the movements. The natural scheme (ST) obtained the best results, with smaller percentage of time in collision state for both user groups.

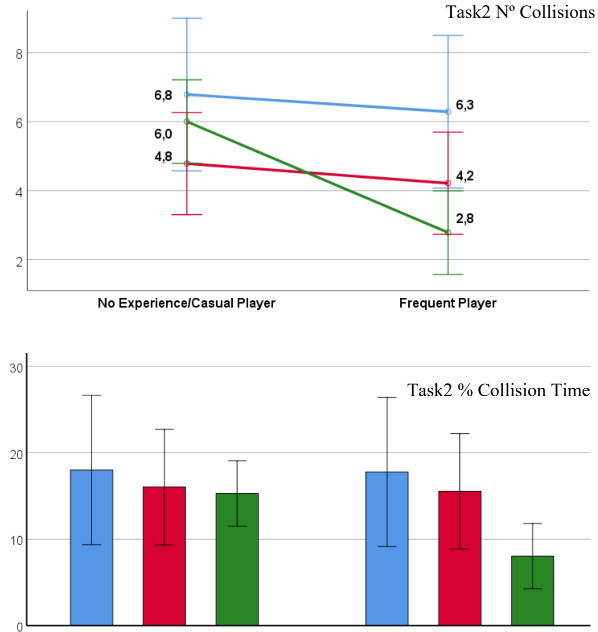


Fig.11 Number of collisions and percentage of time in collision state

Regarding the number of collisions detected, we found differences with a significance level of $\alpha=0.05$ between the metaphoric (PP) and the symbolic scheme (p=0.024) and with the natural (p=0.030). The symbolic scheme (LT) obtained the smaller number of collision for group 1 and the natural scheme (ST) got the best results for group 2.

The longest time in collision state and the highest number of collisions were observed on the metaphoric scheme (PP) for both groups. In any case, it is important to note that collisions did not affect the completion of the task. Therefore, this factor has only a relative effect in the performance of the system [39].

4.3.2. Efficiency: The symbolic movement (LT) obtained the best results for task 1 for both user groups.

Analysing the three movements for all users, the best average time to completion for task 1 was 54 s (Fig.4). This value establishes an ideal time for this task to obtain the average task time estimate for a small sample. In order to set up a benchmark, we considered as efficient times all times smaller than the ideal time multiplied by 1.5 [29]. A time shorter than 81 s should be considered efficient.

Regarding the benchmark, all the averages for task 1 fall inside the range of efficient time for the natural and symbolic schemes, while 89.3% of users reach this score for the metaphoric scheme.

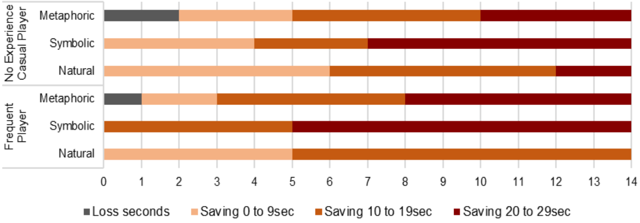


Fig.12 Time savings by movement scheme and user group.

Comparing by user group and type of movement (Fig.12), the symbolic scheme (LT) obtained the most efficient times, reaching time savings with respect to the benchmark up to more than 10 s for 71,4% of groups 1 users and for 100% of users of group 2. The metaphorical scheme (PP) presents a remarkable fact: although 10.7% of users needed more time than the benchmark value (81 s.), 78.6% of users of group 1 and 64.3% of users of group 2 obtained time savings higher than 10 s.

4.3.3. Overall Satisfaction: Upon finishing the test, the users answered a questionnaire related to their ease of learning, physical fatigue, comfort and user pleasure with the usage and performance of the corresponding movement scheme.

- *Ease of learning:* After task 1, we measured the degree of difficulty perceived by users during the performance of the task with a Single Ease Question (SEQ), from very difficult to very easy. Analysing the results, at a general level, the perception of easiness to learn for all movement paradigms is high for both user groups. (Fig.13).

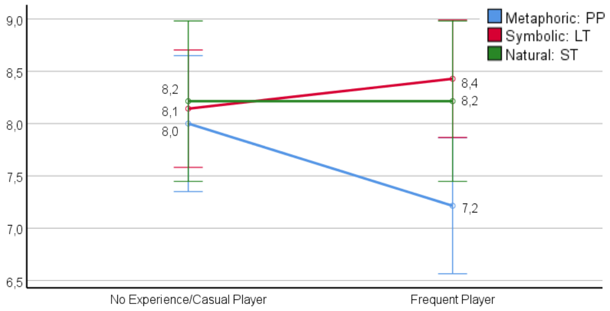


Fig.13 Single Ease Question by movement scheme and user group

Considering the whole set of users, LT and ST received the best evaluations. If we take into account the previous expertise, group 1 users considered the natural scheme (ST) as the easiest to learn. Users of group 2 considered the symbolic scheme (ST) to be the easiest.

Comparing the medians of the score obtained for ease of learning, we found differences with a level of significance α=0.05 between the metaphoric and symbolic schemes (p=0.046) for group 2. No meaningful differences were found for group 1.

The percentage of users describing the task 1 as easy ranges from 85.7% to 96.4%. If we consider values equal or higher than 8 as positive, and separating users by previous expertise, the symbolic scheme (LT) was considered the best for group 1, with 85.7% positive values. Group 2 gave the best values to the natural scheme (ST), with a 78.6% of positive responses.

- *Physical fatigue:* The question to evaluate this aspect asked the user to rate the level of fatigue, from very tired to very relaxed. The natural scheme (ST) was considered as the most relaxed for both user groups (Fig.14).

  Arranging the results by movement scheme and user type, ST stands out as less fatiguing for the no experience/casual player user (8.1, 95% CI 7.3,9.2) and frequent player (8.4, 95% CI 7.4, 9.4). The most tiring movement schemes, by group, are LT for the no experience user (7.6, 95% CI 6.6, 8.5) and PP for frequent player (6.7, 95% CI 5.6, 7.9).

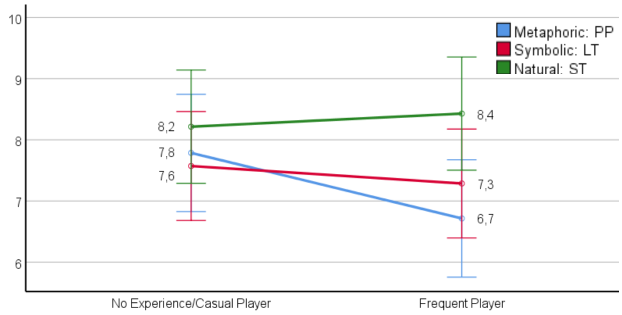


Fig.14 Fatigue by movement scheme and user group

Comparing the level of fatigue for the three movement schemes, and for all users, there are differences with a meaningful level of significance α=0.05 between natural (ST) and metaphoric (PP) (p=0.031) and between natural (ST) and symbolic (LT) (p=0.028), no meaningful differences were found between PP and LT.

The ST movement scheme obtained the highest percentage of positive answers (>= 8 points: being less tiring) for the no experience user (78.6%) and frequent player (64.3%), the LT scheme was the one which received less positive values both for group 1 (57.2%) and for group 2 (35.6%).

- *User comfort:* This question asks the user to rate the level of comfort, from very comfortless to very comfortable. The Kruskal-Wallis test does not display meaningful differences for the three movements taking into consideration users' previous expertise.

  Natural scheme (ST) appeared as the most comfortable (Fig.15) for no experience/casual players (8.0, 95% CI 7.2, 8.9) and frequent players (7.7, 95% CI 6.6, 8.8). The metaphoric scheme (PP) resulted to be the less comfortable both for group 1 (7.1, 95% CI 6.5,7.7) and group 2 (6.1, 95% CI 5.1,7.0).

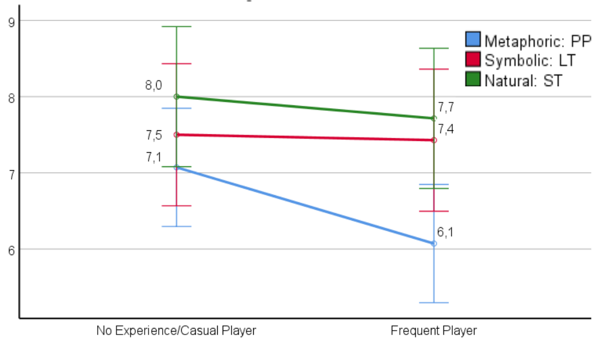


Fig.15 Comfort by movement scheme and user group

Comparing the three movement schemes, and considering all users, there are only significant differences between the natural (ST) and metaphoric (PP) (p=0.0004). Of all movement schemes, ST obtained the highest percentage of positive answers (>=8: very comfortable), for the no experience users (78.6%) and frequent players (58.1%). PP was the one which received fewer positive responses both for group 1 (35.7%) and group 2 (42.9%).

- *Pleasure:* This question asks the user to rate how exciting and enjoyable the experience was, from very unlikable and boring to very pleasing and exciting.

  Generally, all user groups perceived all movement schemes as pleasing and exciting, giving average values > 7.5/10. Arranging the results by groups, the most pleasing schemes were ST for no experience users (8.4, 95% CI 7.7, 9.0) and frequent players (8.1, 95% CI 7.2, 8.9).

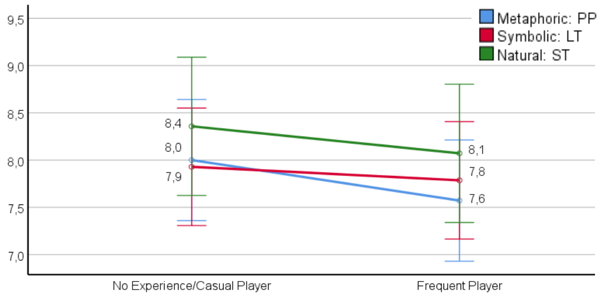


Fig.16 Pleasure by movement scheme and user group

The rest of the movement schemes obtained very close valuations, without significant differences among them. Finally, the percentage of positive answers with respect to user perception of pleasure is very high in ST (78.6%) for all users.

5 Discussion

From the data and analysis described in the previous section, Table 4 summarizes the performance of every movement scheme for every user group considering the variables that correspond to every aspect of the analysis.

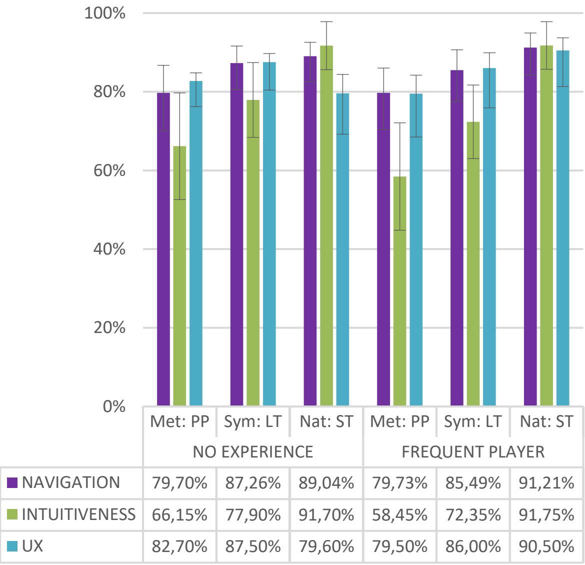
Table 4. Summary. Performance of every movement scheme for every variable and user group.

🟧 Poor 🟨 Medium 🟩 Good

Variables		Collection method	NO EXPERIENCE			FRECUENT PLAYER		
			PP	LT	ST	PP	LT	ST
Velocity		Task1time	🟨	🟩	🟧	🟨	🟩	🟧
Navigation skill		T1 collision time	🟧	🟨	🟩	🟧	🟨	🟩
Accuracy		Task2 time	🟧	🟩	🟨	🟩	🟨	🟧
		Q2 Question	🟧	🟩	🟨	🟧	🟨	🟩
		Q2 Best scores	🟧	🟨	🟩	🟨	🟧	🟩
Spatial Awareness		Q4 Question	🟧	🟨	🟩	🟧	🟨	🟩
		Q3 Question	🟧	🟩	🟨	🟧	🟨	🟩
Attention		Q3 Question	🟧	🟩	🟨	🟧	🟨	🟩
		Q3 Best scores	🟧	🟩	🟨	🟧	🟨	🟩
Reliability		Q5 Question	🟧	🟩	🟨	🟧	🟨	🟩
		Q5 Best scores	🟨	🟧	🟩	🟧	🟨	🟩
Effectiveness		Completion Rate	🟩	🟩	🟩	🟩	🟩	🟩
		T2 Nº collisions	🟧	🟩	🟨	🟧	🟩	🟨
		T2 collision time	🟧	🟨	🟩	🟧	🟨	🟩
Efficiency		Tasks 1 time	🟨	🟩	🟧	🟨	🟩	🟧
		Saving time	🟨	🟩	🟧	🟨	🟩	🟧
Satisfaction	Ease of learning	Q1 Question	🟧	🟨	🟩	🟧	🟨	🟩
		Q1 Best scores	🟧	🟩	🟨	🟧	🟨	🟩
	Physycal fatigue	Q6 Question	🟨	🟧	🟩	🟧	🟨	🟩
		Q6 Best scores	🟨	🟧	🟩	🟨	🟧	🟩
	User comfort	Q7 Question	🟧	🟨	🟩	🟧	🟨	🟩
		Q7 Best scores	🟧	🟨	🟩	🟧	🟨	🟩
	User pleasure	Q8 Question	🟨	🟧	🟩	🟧	🟨	🟩
		Q8 Best scores	🟧	🟨	🟩	🟨	🟧	🟩

The analysis of the previous results, supported with the users' comments and the notes taken by the authors during the experiment, provides clues to characterize the behaviour and performance of the movement schemes and their suitability for their use on a museum environment for virtual walkthroughs.

From the data and the analyses described in the previous paragraphs, and combining multiple usability metrics into a single metric for every dimension analysed, some characteristics for every movement scheme can be obtained (Fig.17).



	Met: PP	Sym: LT	Nat: ST	Met: PP	Sym: LT	Nat: ST
	NO EXPERIENCE			FREQUENT PLAYER		
NAVIGATION	79,70%	87,26%	89,04%	79,73%	85,49%	91,21%
INTUITIVENESS	66,15%	77,90%	91,70%	58,45%	72,35%	91,75%
UX	82,70%	87,50%	79,60%	79,50%	86,00%	90,50%

Fig.17 Single metric for every dimension by movement and user group.

- Metaphoric: Point forward/sideway with arm:

This scheme seems to be the less adequate to navigate in virtual environments for museum installations due to several reasons:

- It required the highest levels of attention devoted to control the system, making it the most intrusive.
- It obtained the highest percentage of time in collision state and the greater number of collisions for all users.
- It was considered the least adequate for the contemplation of the environment, especially with regard to the observation of the mosaics.
- All users considered it as the least comfortable.

During the test, some users expressed confusion about how to hold the arm in the air. The notes taken by the test supervisors also indicate that some users raising their arm completely in the air, unintentionally adopting a tiring pose, while others kept their arm flexed, making difficult for the system to identify the pose.

Considering the degree of previous expertise, this scheme was considered as the most tiring for frequent players.

Metrics give this scheme the worst percentage in intuitiveness and user navigation performance. Regarding UX metric, it obtained the lowest percentage for frequent players and the second worst for no experience/casual players.

- Symbolic: Lean forward/ Twist upper body:

Globally, this movement scheme obtained a high evaluation, standing out in time efficiency, and obtaining time savings of more than 10 s with respect to the benchmark.

Furthermore, this scheme was perceived for all users as the most reliable to interpret users' intentions. Users commented verbally during the test about the accuracy of this scheme to control acceleration, turn, and stop.

Considering previous expertise, this scheme permitted to obtain the smallest number of collisions for no experience/casual players.

Although this scheme obtained positive feedback regarding how much users can enjoy the environment, the individual values for this question in group 2 display a high deviation. Frequent players in group 1 considered this scheme the easiest to learn.

This movement scheme obtained the second best percentage in intuitiveness and user navigation performance. It also obtained the best percentage in UX metric for no experience/ casual players, and the second best for frequent players.

- Natural. Step forward / Twist upper body:

This scheme constitutes a good candidate for museum installations. Globally, it obtained very good evaluations for all users:
- This was the movement scheme that obtained smaller times spent in collision state.
- It was valued as the most comfortable and least tiring, while obtaining the highest scores in accuracy.
- It obtained the best scores for contemplation of the environment.
- Users perceived it as the most pleasing and exciting for both groups.
- It obtained the second place in efficiency (mean average time to finish tasks) and reliability, after Lean forward /Twist upper body.

Considering the degree of previous expertise, it was the easiest to learn for no experience/casual players. For frequent players, it was also the scheme that allowed the users to pay a better conscious attention to the scenario. Additionally, it yielded the least amount of collisions

This paradigm obtained the best values in intuitiveness and user navigation performance. It also was the best valued in UX metrics for frequent players and the second one for no experience/casual players.

6 Conclusions

Virtual museums can offer to their visitors much more than the mere visual representation of things. They may be seen as sources of new experiences, fostering a deeper interpretation and a more persistent memory that a simple exhibit. Natural user interfaces such as Kinect are extremely useful to achieve this goal, but it is important to find the movement scheme that best facilitates, in terms of HCI, visiting the virtual spaces.

Among the three schemes presented here, Natural Movement (Step forward /Twist upper body) stood out as the most balanced in all aspects with the best valuations in satisfaction, user navigation performance and intuitiveness. This movement scheme appears to be the most appropriate to navigate such digital environments for any kind of user, regardless of the previous expertise of the visitor in video games.

Symbolic Movement (Lean forward /Twist upper body) proved to be the most efficient scheme to interpret the user intentions. The degree of previous expertise did not mark significant differences in this result. This movement scheme seems to be very satisfactory for the generic profile of museum visitors.

Although this study evaluates the performance in navigation, intuitiveness and user experience, we should not forget that learning is one of the main goals of any museum installation. Users should be able to construct concepts through the observation and experience of the content provided. Therefore, it is necessary to find out how the UX metrics obtained relate to users' cognition, immersion and flow [40]. The results obtained may also be used to achieve a higher level of attention to the contents inside the virtual environment as a learning tool for visitors to the museum.

It is necessary to continue researching the optimization of these movement schemes and the acquisition of even more transparent movement interfaces. For instance, it is desirable to develop a schema that could permit a free combination of gestures from the three paradigms as a way to find out a pattern that results the most intuitive and natural for the user. This could constitute a future line of research.

The authors expect that the performance metrics and UX results presented here will be useful for designers of virtual environments to choose the natural interaction walkthrough scheme that may fit best with their needs based on the particular features of their installation.

Disclaimer

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Desvallées, A., Mairesse, F.: Key concepts of museology. International Council of Museums. Armand Colin. Paris (2010).
2. Ross, M.: Interpreting the new museology. Museum and society, 2(2), pp. 84-103 (2004).
3. Vergo, P.: New museology. Reaktion books. United Kingdom. (1997).
4. Hooper-Greenhill, E: Counting visitors or visitors who count? In The museum time machine, pp. 221-240. Routledge (2003).
5. Doering, Z. D.; Strangers, Guests, or Clients? Visitor Experiences in Museums. Curator, 42 (2), pp. 74–87 (1999).
6. Hernández Ibáñez, L. A., Mihura López, R., & Barneche Naya, V.: Nuevas tecnologías en los museos, de la contemplación a la experiencia. In ¿Cómo se cuelga un cuadro virtual? Las exposiciones en la era digital, pp.19-100. Gijón: Trea (2009).
7. Lun, R., & Zhao, W.: A survey of applications and human motion recognition with microsoft kinect. International Journal of Pattern Recognition and Artificial Intelligence, 29(05), 1555008 (2015).
8. Pletinck,D., Capurro, C., Ferdani, D., and Pescarin, S.: Deliverable Report. D 8.9, European Virtual Museum Observatory. http://www.vmust.net/library/documents/d89-eu-virtual-museumob-servatory-evmo-Report.html, pp. 6, 16-18. (2015). Last accessed 2019/11/21.

9. Ray, C., Pescarin, S., Pagano, A.: Deliverable Report. D7.1 Virtual Museum Quality Label. http://www.v-must.net/sites/default/files/D7.1c%20Quality_v2.pdf, pp. 32-39. (2015). Last accessed 2019/11/21.
10. Sheng, W., Ishikawa, K., Tanaka, H. T., Tsukamoto, A., & Tanaka, S.: Photorealistic VR Space Reproductions of Historical Kyoto Sites based on a Next-Generation 3D Game Engine. Journal of Advanced Simulation in Science and Engineering, 1(1), pp.188-204 (2015).
11. Lercari, N., Mortara, M., & Forte, M.: Unveiling California History through Serious Games: Fort Ross Virtual Warehouse. Games and Learning Alliance, pp.236-251. Springer International Publishing (2014).
12. Pietroni, E., & Adami, A.: Interacting with virtual reconstructions in museums: The Etruscanning Project. Journal on Computing and Cultural Heritage (JOCCH), 7(2), 9 (2014).
13. Richards-Rissetto, H., Robertsson, J., von Schwerin, J., Agugiaro, G., Remondino, F., & Girardi, G.: Geospatial Virtual Heritage: a gesture-based 3D GIS to engage the public with Ancient Maya Archaeology. Archaeology in the Digital Era, pp.118-130. (2014).
14. Cappelletto, E., Zanuttigh, P., & Cortelazzo, G. M.: 3D scanning of cultural heritage with consumer depth cameras. Multimedia Tools and Applications, pp.1-24. (2014).
15. Fanini, B., d'Annibale, E., Demetrescu, E., Ferdani, D., & Pagano, A.: Engaging and shared gesture-based interaction for museums the case study of K2R international expo in Rome. In 2015 Digital Heritage, Vol. 1, pp. 263-270. IEEE (2015).
16. Bowman, D. A., Johnson, D. B., & Hodges, L. F.: Testbed evaluation of virtual environment interaction techniques. Presence: Teleoperators & Virtual Environments, 10(1), pp. 75-95. (2001).
17. Bowman, D. A., Koller, D., & Hodges, L. F.: A methodology for the evaluation of travel techniques for immersive virtual environments. Virtual reality, 3(2), pp. 120-131. (1998).
18. Hernández-Ibáñez, L. A., Barneche-Naya, V., & Mihura-López, R.: A comparative study of walkthrough paradigms for virtual environments using kinect based natural interaction. In Virtual System & Multimedia (VSMM), 2016 22nd International Conference, pp. 1-7. IEEE. (2016).
19. Burigat, S., & Chittaro, L.: Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. International Journal of Human-Computer Studies, 65(11), pp.945-958 (2007).
20. Smith, S., & Du'Mont, S: Measuring the effect of gaming experience on virtual environment navigation tasks. 3D User Interfaces - 3DUI 2009, pp. 3-10. IEEE (2009).
21. Creswell, J. W.: Research design: Qualitative, quantitative, and mixed methods approaches. 3rd ed. Sage publications. (2009).
22. Macaranas, A., Antle, A. N., & Riecke, B. E.: What is intuitive interaction? Balancing users' performance and satisfaction with natural user interfaces. Interacting with Computers, 27(3), pp. 357-370 (2015).
23. Holland S.: Asymmetrical Multi-User Co-operative Whole Body Interaction in Abstract Domains. CHI'10. Whole Body Interaction Workshop, Atlanta, Georgia, USA. ACM Press. (2010)
24. Antle A.N. Corness G. Droumeva M.: What the body knows: exploring the benefits of embodied metaphors in hybrid physical environments. Interact. Comput. , 21, pp.66–75 (2009).
25. Hurtienne J. Weber K. Blessing L.: Prior Experience and Intuitive Use: Image Schemas in User-Centered Design. In Langdon P. Clarkson J.P. Robinson P. (eds), Designing Inclusive Futures. Springer, Berlin, pp. 107–116 (2008).
26. Celentano, A., Dubois, E.: Metaphors, analogies, symbols: in search of naturalness in tangible user interfaces. Proceedings of the 6th International Conference on Intelligent Human Computer Interaction. Paris, France. (2014).
27. Sperka, M.: Past and future of Human-Computer interaction. Proceedings of the International Conference on Current Issues of Science and Research in the Global World. Vienna Austria. (2014).
28. Ministerio de Cultura: Conociendo a todos los públicos. ¿Qué imágenes se asocian a los museos? Secretaría General Técnica de la Subdirección General de Publicaciones, Información y Documentación (Ministerio de Cultura). https://sede.educacion.gob.es/publiventa/descarga.action?f_codigo_agc=14315C. (2012). Last accessed 2019/12/01.
29. Sauro, J., & Lewis, J. R.: Quantifying the user experience: Practical statistics for user research. Morgan Kaufmann (2012).
30. Spool J.M.: What makes a design seem 'intuitive'? User Interface Eng. http://www.uie.com/articles/design_intuitive/ (2005)
31. Albert, W., & Tullis, T.: Measuring the user experience: collecting, analyzing, and presenting usability metrics. Newnes. (2013).
32. ISO 9241-11:2018: Ergonomics of human-system interaction-Part 11: Usability: Definitions and concepts.https://www.iso.org/obp/ui/#iso:std: iso: 9241:-11:ed-2:v1:en:en. Last accessed 2019/12/01.
33. Lindgaard, G., & Dudek, C.: What is this evasive beast we call user satisfaction? Interacting with Computers, 15(3), pp.429–452. (2003).
34. Shackel, B.: Usability–context, framework, design and evaluation. Artikkeli teoksessa Shackel, B. and Richardson, S. Human Factors for Informatics Usability. Cambridge University Press, Cambridge, 21, p38. (1991).
35. Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K.: Usability measurement and metrics: A consolidated model. Software Quality Journal, 14(2), pp.159-178. (2006).
36. Hassenzahl, M.: Hedonic, emotional, and experiential perspectives on product quality. In Encyclopaedia of Human Computer Interaction, pp.266-272. IGI Global. (2006).
37. Sauro, J., & Dumas, J. S.: Comparison of three one-question, post-task usability questionnaires. In Proceedings of the SIGCHI conference on human factors in computing systems, pp. 1599-1608. ACM. (2009).
38. Norman, D.: The design of everyday things: Revised and expanded edition. Basic books. (2013).
39. Rubin, J., & Chisnell, D.: Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons. (2008).
40. Csikszentmihalyi, M.: Flow. The psychology of optimal experience. New York: Harper & Row. (1990).
41. Barneche-Naya, V. & Hernandez-Ibáñez, L.A. UX Aspects of Kinect-based Movement Schemes inside Virtual Envoronments for Museum Installations. In Learning and Collaboration Technologies. Ubiquitous and Virtual Environments for Learning and Collaboration. pp 133-150. Springer (2019)