

Article

Evaluating Museum Virtual Tours: The Case Study of Italy

Katerina Kabassi ^{1,*}, Alessia Amelio ², Vasileios Komianos ³ and Konstantinos Oikonomou ⁴

¹ Department of Environment, Ionian University, 49100 Corfu, Greece

² Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica, University of Calabria, 87036 Arcavacata, Italy

³ Department of Audio and Visual Arts, Ionian University, 49100 Corfu, Greece

⁴ Department of Informatics, Ionian University, 49100 Corfu, Greece

* Correspondence: kkabassi@ionio.gr

Received: 19 October 2019; Accepted: 11 November 2019; Published: 14 November 2019



Abstract: Virtual tours in museums are an ideal solution for those that are not able to visit a museum or those who want to have a small taste of what is presented in the museum before their visit. However, these tours often encounter severe problems while users interact with them. In order to check the status of virtual tours of museums, we present the implementation of an evaluation experiment that uses a combination of two multi-criteria decision making theories, namely the analytic hierarchy process (AHP) and the fuzzy technique for order of preference by similarity to ideal solution (TOPSIS). AHP has been used for the estimation of the weights of the heuristics and fuzzy TOPSIS has been used for the evaluation of virtual tours of museums. This paper presents the exact steps that have to be followed in order to implement such an experiment and run an example experiment for virtual tours of Italian museums.

Keywords: virtual reality; museum tours; evaluation; multi-criteria decision making; fuzzy logic

1. Introduction

Virtual reality (VR) technology offers an ideal presentation medium for museums and other cultural heritage institutions [1]. These appealing technological systems have demonstrated their usefulness and value in science centers and traditional museums all over the world, thanks to the fact that visitors can view digitized artworks and explore reconstructed historical places by means of virtual museum-hosted installations [2]. However, the implementation of VR installation is complicated and expensive and, therefore, in most cases, the final environment may experience several usability problems.

Virtual museums' lack of usability is commonly attributed to non-comprehensible or too complicated navigation techniques being used [3]. Regarding navigation in virtual environments, users can face disorientation problems, loss of overview and difficulty relocating previously visited locations [4]. In this context, navigation assistance is considered to be an important factor of usability improvement and numerous approaches have been proposed in order to provide virtual museums in which virtual visitors have realistic navigation abilities and observation experiences of the presented exhibits.

Another factor that acts against the development of a VR environment is the lack of public experience in interacting with such environments. This problem is more obvious in cases where the users are elderly or novices in the usage of computers. In addition, the virtual museum and its functionality should be accessible to a wide range of users including people with diverse skills

and abilities including computer-users or non-users, children, elderly people or disabled people. For this purpose, VR environments should be as effective and friendly as possible in order to fulfill the communication goals of the museum. The provided navigation and interaction methods should be available for users with diverse abilities and different preferences.

As a result, specific usability evaluations have taken place for museum websites and virtual museum interfaces (e.g., [5,6]). Some developers and usability practitioners attempted to apply traditional usability engineering methods to the evaluation of VR systems [7,8]. However, these efforts revealed that most heuristics were not appropriate for these environments. Taking into account this observation, usability metrics (i.e., performance, satisfaction) for VR environments have been proposed by Bazargan et al. [9]. However, the proposed toolkit is rather limited.

A framework for usability evaluations of websites that include virtual environments and follow Nielsen's recommendations for expert evaluation [10] has been proposed by Sutcliffe and Gault [11]. Sylaiou et al. [12], on the other hand, conclude that five quality criteria are required for evaluating virtual museums in general: "Imageability", "Interactivity", "Navigability", "Virtual Spatiality" and "Narration". However, the main problem that occurs with heuristic evaluation experiments, in general, is that all heuristics are equally important in the evaluation process. Therefore, the research question that occurs is how the heuristics could be combined in the evaluation process of museum virtual tours.

The proposed solution was initially given by Kabassi et al. [13], who have used the fuzzy simple additive weighting (FSAW) theory [14] in combination with the framework for usability evaluation of the VR environments proposed by Sutcliffe and Gault [11]. However, FSAW does not have a formal way of calculating weights. Trying to improve this framework, we propose a combination of two multi-criteria decision making theories together with the heuristics proposed by Sutcliffe and Gault [11] for evaluating the VR environment. More specifically, the analytic hierarchy process (AHP) [15] has been used in combination with the fuzzy technique for order of preference by similarity to ideal solution (fuzzy TOPSIS) [16].

The evaluation of the virtual tours was done based on some usability heuristics proposed by Sutcliffe and Gault [11] and AHP provided a well-defined way of calculating the weights of these heuristics. The ability of AHP to make decisions by making a pairwise comparison of uncertain, qualitative and quantitative factors and also its ability to model expert opinion [17] were some of the reasons for its selection over other methods.

TOPSIS, on the other hand, aims at ordering evaluation items (which in our case are museum websites) through detecting distance between evaluated objects and optimal solutions [18]. In the particular evaluation experiment, fuzzy TOPSIS is used instead of TOPSIS to process the data, making estimations and ranking the alternatives. If the evaluated object is near the optimal solution and far away from the poor solution, it is the best.

These two theories have different reasonings but seem rather complementary. This combination is mainly reported in the evaluation of e-commerce websites and, more specifically, in the evaluation of websites of travel agencies [19] or group-buying [20]. Furthermore, fuzzy AHP has been combined with fuzzy TOPSIS for evaluating university websites [21] and e-government sites [22]. This combination has never been used before for the evaluation of virtual tours in museums.

The proposed approach has been applied for evaluating virtual tours in museums in Italy. The evaluation revealed that the framework can be easily applied to another set of museums and helped significantly in extracting conclusions about the state of virtual tours in museums in Italy. The paper is organized as follows: Section 2 presents the heuristics used for the evaluation of the virtual tours as well as the multi-criteria decision making theories that are used to combine these heuristics and evaluate the virtual tours. In Section 3, we present the implementation of the evaluation experiment, the estimation of the weights using AHP and how the heuristics of the virtual tours took their values from the evaluators. In Section 4, the implementation of fuzzy TOPSIS is described and the results are extracted. The results of the evaluation are discussed in Section 5 and the conclusions drawn by this work are analyzed in Section 6.

2. Materials and Methods

In this section, we present the materials and methods used for implementing the proposed model for the usability evaluation of virtual tours in museums. For this purpose, the heuristics proposed by Sutcliffe and Gault [11] that are specialized for the usability evaluation of VR environments have been used and are the following:

1. Natural engagement: how close the interaction is to the real world (Ne).
2. Compatibility with the user's task and the domain: how close the behavior of objects is to the real world and affordance for task action (C).
3. The natural expression of action: does the system allow the user to act naturally? (Na).
4. Close coordination of action and representation: quality of the response between user movement and Virtual Environment (Cr).
5. Realistic feedback: visibility of the effect of users' actions and conformity to the laws of physics (Rf).
6. Faithful viewpoints: the naturalness of change between viewpoints (Fv).
7. Navigation and orientation support: naturalness in orientation and navigation. Is it clear where they are and how they return? (No).
8. Clear entry and exit points: clearness of entry and exit points (Ce).
9. Consistent departures: consistency of departure actions (Cd).
10. Support for learning: promotion of learning (L).
11. Clear turn-taking: clearness of who has the initiative (Tt).
12. Sense of presence: the naturalness of the user's perception of engagement in the system and being in a 'real' world (P).

A simple analysis of these rules may provide a categorisation of the heuristics to four categories: VR experience (heuristics 1, 2 and 5), perception of presence (heuristics 3, 4, 11 and 12), navigation (heuristics 6, 7, 8 and 9) and the learning aspect (heuristic 10).

These heuristics are used as the criteria for the evaluation. Multi-criteria decision making theories are ideal for combining these criteria and evaluate different alternatives such as different virtual tours. For this purpose, AHP is used for calculating the weights of the criteria and then fuzzy TOPSIS is used for evaluating the different alternative virtual tours.

2.1. AHP

The steps of the implementation of AHP that are used for the estimation of the weights of the criteria, which have already been selected and are the heuristics presented above, are:

1. Form the set of evaluators: for the estimation of the weights of the heuristics, it is important to have the view of experts on the field. Therefore, the set of evaluators is only composed of human experts. The selection of expert-based evaluations has many advantages [5] and the correct choice of experts gives more reliable and valid results.
2. Setting up a pairwise comparison matrix of heuristics: in this step, a comparison matrix is formed so that the heuristics are pairwise compared.
3. Calculating the weights of criteria: after making pairwise comparisons, estimations are made that result in the final set of weights of the criteria.

As soon as, the weights have been calculated, the virtual tours are evaluated and the analysis of the results of the evaluation is made based on fuzzy TOPSIS.

2.2. Fuzzy TOPSIS

Fuzzy TOPSIS presupposes that the weights of the criteria have been estimated and it uses them to evaluate alternative virtual tours in museums. The main steps of fuzzy TOPSIS for the evaluation of alternative virtual tours are the following:

1. Forming a new set of evaluators: in this phase of the evaluation experiment, the set of evaluators was formed, following the taxonomy of types of users of cultural websites proposed by Sweetnam et al. [23]. This group may be the same as the one formed in step 1 of the application of AHP or may be different.
2. Assigning values to the criteria: in order to make this process easier for the user, especially for those that do not have experience in multi-criteria analysis, the users could use linguistic terms for characterizing the twelve heuristics presented above. The linguistic terms are presented in Table 1.
3. Linguistic terms are transformed into fuzzy numbers: each linguistic term is assigned to a fuzzy number, which is a vector $\tilde{a} = (a_1, a_2, a_3)$. The matches are presented in Table 1 [16].
4. Construction of the Multi-Criteria Decision Making (MCDM) matrix: a fuzzy multi-criteria group decision-making problem can be expressed in matrix format. Each element of the matrix is a fuzzy number. However, in order to aggregate all the values of the decision-makers in one single value, the geometric mean is used. The geometric mean of two fuzzy numbers $\tilde{a} = (a_1, a_2, a_3)$ and $\tilde{b} = (b_1, b_2, b_3)$ is calculated as follows:

$$\tilde{c} = (\sqrt{a_1 b_1}, \sqrt{a_2 b_2}, \sqrt{a_3 b_3}) \tag{1}$$

$$\tilde{D} = \begin{matrix} & & C_1 & C_2 & & C_n \\ \begin{matrix} A_1 \\ A_2 \\ \\ \\ A_m \end{matrix} & \begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} & & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & & \tilde{x}_{2n} \\ & & \tilde{x}_{ij} & \\ & & & \tilde{x}_{mn} \end{pmatrix} & , i = 1, 2, \dots, m; j = 1, 2, \dots, n, \tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij}) \end{matrix} \tag{2}$$

where i shows the alternative and j shows the criterion. Each $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ is a triangular fuzzy number.

5. The normalization of fuzzy numbers: to avoid the complicated normalization formula used in classical TOPSIS, the authors in [16] propose a linear scale transformation in order to transform the various criteria scales into a comparable scale. This particular normalization method aims at preserving the property that the ranges of normalized triangular fuzzy numbers belong to $[0,1]$. The normalization of a fuzzy number $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ is given by the formula:

$$\tilde{r}_{ij} = (\frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*}), \text{ where } c_j^* = \max_i c_{ij} \tag{3}$$

6. Calculating the weighted normalized fuzzy numbers of the MCDM matrix: considering the different importance of each criterion, which is imprinted in the weights of the criteria, the weighted normalized fuzzy numbers are calculated: $\tilde{u}_{ij} = \tilde{r}_{ij}(\cdot)\tilde{w}_j$ and these values are used to construct the weighted normalized fuzzy MCDM matrix $\tilde{V} = [\tilde{u}_{ij}]_{M \times N}, i = 1, 2, \dots, m; j = 1, 2, \dots, n$.
7. Determination of the fuzzy positive-ideal solution (FPIS) and the fuzzy negative-ideal solution (FNIS): the FPIS and the FNIS are calculated as follows.

$$\text{FPIS : } A^* = \{\tilde{u}_1^*, \tilde{u}_2^*, \dots, \tilde{u}_i^*, \dots, \tilde{u}_n^*\}, \tilde{u}_j^* = (1, 1, 1) \tag{4}$$

$$\text{FNIS : } A^- = \{\tilde{u}_1^-, \tilde{u}_2^-, \dots, \tilde{u}_i^-, \dots, \tilde{u}_n^-\}, \tilde{u}_j^- = (0, 0, 0) \tag{5}$$

8. Calculation of the distance of each alternative from FPIS and FNIS: the distances (d_i^* and d_i^-) of each weighted alternative (d_i^* and d_i^-) from FPIS and FNIS are calculated as follows.

$$d_i^* = \sum_{j=1}^n d_u(\tilde{u}_{ij}, \tilde{u}_j^*), i = 1, 2, \dots, m \tag{6}$$

$$d_i^- = \sum_{j=1}^n d_u(\tilde{u}_{ij}, \tilde{u}_j^-), i = 1, 2, \dots, m \tag{7}$$

where $d_u(\tilde{a}, \tilde{b})$ is the distance between two fuzzy numbers \tilde{a}, \tilde{b} . The distance between two fuzzy numbers $\tilde{a} = (a_1, a_2, a_3)$ and $\tilde{b} = (b_1, b_2, b_3)$ is calculated as:

$$d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3}[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]} \tag{8}$$

9. Calculation of the closeness coefficient of each alternative: the closeness coefficient of each alternative j is given by the formula $CC_i = \frac{d_i^-}{d_i^* + d_i^-}$, $0 \leq CC_i \leq 1$. According to the values of the closeness coefficient, the ranking order of all the alternatives is determined. The alternative that is closer to FPIS and further from FNIS as CC_i approaches 1. The values of the closeness coefficient of each alternative and the final ranking of the evaluated websites are presented in Table 2.

Table 1. Linguistic variables and fuzzy numbers.

Linguistic Term	Fuzzy Number
Very Poor	(0,0,1)
Poor	(0,1,3)
Fair	(3,5,7)
Good	(7,9,10)
Very Good	(9,10,10)

Table 2. d_i^*, d_i^- the closeness coefficient (CC) and the ranking of each alternative virtual museum.

	d^*	d^-	CC	Ranking
Maotorino	11.78306	0.288359	0.023888	15
Museivaticani	11.61684	0.457118	0.03786	5
Ebraico	11.479	0.583454	0.048369	3
dell’AraPacis	11.7397	0.328556	0.027225	14
Napoleonico	11.72471	0.344694	0.028559	13
Capitolini	11.70928	0.366162	0.030323	11
MercatidiTraiano	11.67833	0.399702	0.033093	7
Villatorlonia	11.70072	0.378629	0.031345	9
Spazzacamino	11.71159	0.363352	0.030091	12
Zoologia	11.87914	0.177064	0.014687	16
Geologia	11.70604	0.375118	0.03105	10
Motori	11.54909	0.52448	0.04344	4
Cavallogiocattolo	11.45735	0.594157	0.049301	2
Carezzonico	11.30168	0.735233	0.061082	1
Castellogrinzane	11.69436	0.37978	0.031454	8
Museodelletna	11.62822	0.43481	0.036045	6

3. Implementation of the Experiment

For the implementation of the experiment, we conducted a search on the websites of museums in Italy in order to locate the museums that offer a virtual tour. This search revealed that only a few

museums in Italy offer virtual tours on their websites (all these websites were retrieved in September 2019):

1. MAO Museum of Oriental Art (Maotorino)
<https://www.maotorino.it/en/node/36639>
2. Vatican Museums (Museivaticani)
<http://www.museivaticani.va/content/museivaticani/it/collezioni/musei/tour-virtuali-elenco.html>
3. Jewish Museum in Bologna (Ebraico)
<http://ucei.it/virtualtour/index.php?lang=it>
4. AraPacis Museum (dell'AraPacis)
http://www.arapacis.it/it/musei_digitali/tour_virtuali
5. Napoleonic Museum (Napoleonico)
http://www.museonapoleonico.it/it/musei_digitali/tour_virtuali
6. Capitoline Museums (Capitolini)
http://www.museicapitolini.org/it/musei_digitali/tour_virtuali
7. Imperial Forum Museum (Mercati di Traiano)
http://www.mercatiditraiano.it/it/musei_digitali/tour_virtuali
8. Villa Torlonia Museums (Villatorlonia)
http://www.museivillatorlonia.it/it/musei_digitali/tour_virtuali
9. Chimney Sweep Museum (Spazzacamino)
<http://santamariamaggiore.info/home/cultura/museo-dello-spazzacamino/>
10. Zoology Museum "P. Doderlein" (Zoologia)
<http://museozoologia.unipa.it>
11. Geological Museum Gemmellaro (Geologia)
<http://www.museogeologia.unipa.it>
12. Historical Museum of Engines and Mechanisms (Motori)
<http://www.museomotori.unipa.it/visita3d.php>
13. Toy Horse Museum (Cavallogiocattolo)
<http://www.museodelcavallogiocattolo.it/en/museum>
14. Ca'Rezzonico—Eighteenth-Century Venetian Museum (Carezzonico)
<https://carezzonico.visitmuve.it/en/il-museo/multimedialita/google-art-project/>
15. Langhe Museum (Castellogrinzane)
<http://www.castellogrinzane.com/en/Museo>
16. Etna Museum (Museodelletna)
<http://www.museodelletna.it/en/>

As soon as the set of alternative museum virtual tours is formed, AHP is applied to calculate the weights of the criteria used in the reasoning process of the decision-makers. For the application of AHP, first, we form the set of evaluators. The group of the evaluators involved one professional researcher in VR, one researcher in pattern recognition, one researcher in software engineering and one researcher in cultural heritage conservation.

The next step involved setting up a pairwise comparison matrix of heuristics. Each one of the four experts is asked to fill out the pairwise comparison matrix so that the heuristics are pairwise compared. As a result, four matrices are completed. In order to find the final pairwise comparison matrix of

heuristics, we calculated the geometric mean of the values of all corresponding cells belonging to the four matrices completed by the experts.

Finally, the weights of the criteria were calculated. In this step, the principal eigenvalue and the corresponding normalized right eigenvector of the comparison matrix gave the relative importance of the various criteria being compared. The elements of the normalized eigenvector were the weights of the heuristics. There are several methods for calculating the eigenvector. Multiplying together the entries in each row of the matrix and then taking the n -th root of that product approximates the correct answer. The n -th roots are summed and that sum is used to normalize the eigenvector elements to add to 1.00. In terms of simplicity, we had used the ‘Priority Estimation Tool’ (PriEst) [24], an open-source decision-making software that implements the AHP method, for making the AHP calculations. The weights of the criteria were estimated as follows: $w_{Ne} = 0.049$, $w_C = 0.038$, $w_{Na} = 0.042$, $w_{Cr} = 0.186$, $w_{Rf} = 0.055$, $w_{Fv} = 0.084$, $w_{No} = 0.163$, $w_{Ce} = 0.047$, $w_{Cd} = 0.015$, $w_L = 0.139$, $w_{Tt} = 0.054$, $w_P = 0.129$.

Then, fuzzy TOPSIS was used to implement the next phase of the experiment, which involved the evaluation of the museum virtual tours. At the beginning of the application of fuzzy TOPSIS, a set of evaluators had to be formed. We used the same set of four evaluators that were used in the application of AHP. Each evaluator visited each one of the virtual tours presented above. For each one of the museums selected to be evaluated, the users had to assign a linguistic term to each criterion. The linguistic terms are presented in Table 1 [16]. This resulted in four matrices with linguistic terms.

4. Results

The linguistic terms in the four matrices created by the evaluators were then replaced with fuzzy numbers. The fuzzy numbers that correspond to the linguistic terms are presented in Table 1. In order to aggregate all the values of the four matrices created by the decision-makers in one single matrix, the geometric mean was used. This process resulted in one single decision matrix:

$$\tilde{D} = \begin{matrix} & & C_1 & C_2 & & C_n \\ A_1 & & \tilde{x}_{11} & \tilde{x}_{12} & & \tilde{x}_{1n} \\ A_2 & & \tilde{x}_{21} & \tilde{x}_{22} & & \tilde{x}_{2n} \\ & & & & \tilde{x}_{ij} & \\ A_m & & \tilde{x}_{m1} & \tilde{x}_{m2} & & \tilde{x}_{mn} \end{matrix}, i = 1, 2, \dots, 16; j = 1, 2, \dots, 12, \tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$$

where i shows the alternative museum and j shows one of the 12 criteria. Each element of the matrix is a triangular fuzzy number. In our evaluation, $\tilde{x}_{12} = (0, 1, 3)$.

In order to transform the various criteria scales into a comparable scale, all values of the decision matrix should be normalized. For this reason, we used the linear scale transformation proposed by Chen [16]. This particular normalization method aims at preserving the property that the ranges of normalized values belong to using the formula: $\tilde{r}_{ij} = (\frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*})$, where $c_j^* = \max_i c_{ij}$. As a result, all triangular fuzzy numbers that exist in the decision matrix belong to $[0,1]$. In our evaluation, for example, $\tilde{r}_{12} = (0.00, 0.12, 0.36)$.

Taking into account that the criteria have different weights in the reasoning process of the decision-makers, we calculated the values of each criterion with the corresponding weight of the criterion. As a result, the weighted normalized fuzzy MCDM matrix $\tilde{V} = [\tilde{u}_{ij}]_{M \times N}, i = 1, 2, \dots, 16; j = 1, 2, \dots, 12$ is constructed. In this matrix, the element \tilde{u}_{12} was estimated to be $(0.0000, 0.0045, 0.0136)$. TOPSIS is based on the reasoning which estimates the Euclidean distances of each alternative from an ideal alternative and from the negative-ideal alternative. For reasons of simplicity, FPIS is considered as: $A^* = \{(1,1,1), (1,1,1), \dots, (1,1,1), \dots, (1,1,1)\}$ and FNIS as: $A^- = \{(0,0,0), (0,0,0), \dots, (0,0,0), \dots, (0,0,0)\}$. Considering these values, we have calculated the distances (d_i^* and d_i^-) of each weighted alternative museum $i = 1, 2, \dots, 16$ from FPIS and FNIS. In Table 2, one can see in the second column the distance of each museum from the fictitious ideal

museum and in the third column the distance of each alternative from the negative-ideal alternative. These values were further used to calculate the closeness coefficient of each alternative museum, which is presented in the fourth column of Table 2. The museum with the highest value of the closeness coefficient was considered to be the best. The final ranking of the evaluated websites is presented in the fifth column of Table 2.

The virtual tour of the museum Carezzonico was considered to be the best as it had the highest value of the closeness coefficient in comparison to all the others. The virtual tour of Cavallogiocattolo was ranked as second while the virtual tour of the Ebraico museum was ranked as third. The worst virtual tours were considered to be those of the Maotorino museum and the museum of Zoologia.

5. Analysis of the Results and Discussion

Searching for virtual tours of museums in Italy revealed that only a few museums that have a website also have a virtual tour. Also, searching on the webpage for the link from which the virtual tour can be accessed is sometimes a problem due to poor design of the webpage.

The reviewed virtual tours are implemented using worldwide web technologies, thus increasing the ability of users to access them using standard computing equipment and software. The virtual tours are based on 360-degree captured views and not on computed generated three-dimensional (3D) environments, thus providing a realistic view of the museum. A major drawback of virtual tours based on 360-degree views is the limited degree of freedom in navigating and the poor navigation experience. Additionally, most of the virtual tours do not provide interaction with the represented space or the hosted exhibits, and the information communication methods, when present, are mostly based on text descriptions of the exhibits.

For the evaluation of the museums' virtual tours and their classification, we used a combination of AHP with fuzzy TOPSIS. AHP has been used for the estimation of the criteria weights and then fuzzy TOPSIS was used to evaluate the different virtual tours of museums in Italy. As fuzzy TOPSIS was used, the evaluators could use linguistic terms for evaluating the different alternatives. This is one of the advantages of fuzzy TOPSIS and a reason that makes it easier for all evaluators to characterize the criteria. The linguistic variable is then transformed into triangular numbers and then all estimations take place.

Generally, the evaluators used the value 'Low (L)' much more than all the other values. Values such as 'Low (L)' were used in 32% of all rates, while values as 'Very High (VH)' were used very rarely (2% of all rates). The values 'Medium (M)' and 'High (H)' were almost equally used (28% and 25%, respectively). Generally, the 'Natural engagement—Ne' and the 'Compatibility with the user's task and domain—C' of the virtual tours were characterized as 'Low' (80% of all rates assigned to all evaluated virtual tours). Similarly, bad rates had the criterion 'Realistic feedback—Rf', which the evaluator characterized in all virtual tours as 'Very Low' or 'Low', and 'Natural expression of action—Na' has been characterized as 'Low' in 89% of cases. The most 'Medium' values of the criteria were assigned to 'Faithful viewpoints—Fv' and 'Clear entry and exit points—Ce'. The only criterion that was revealed to be good in almost all virtual tours as it received many 'High' values (85%) and some 'Very High' values was that of 'Clear turn-taking—XX'. However, the values cannot reveal how good a virtual tour is as all criteria do not have the same weight in the reasoning process of the decision-makers. For this purpose, a multi-criteria decision-making theory like TOPSIS seems very appropriate.

Generally, the interaction with the virtual environments was considered to be rather satisfactory. The main problems identified in general in the virtual tours of the museums in Italy were that very few of them were equipped with learning support, which could be very useful for understanding the context. Also, most of the virtual tours showed a middle–low naturalness of change between viewpoints. Further problems were identified related to: (i) the navigation functionality; (ii) the interactivity of the represented space and exhibits; (iii) the information communication methods; etc.

Due to these issues, the reviewed virtual tours received bad ratings for most of the evaluation criteria. The fact that the virtual tours are based on 360-degree views with limited degrees of freedom

in navigating is considered to be the reason why the evaluators state that the reviewed virtual tours 'Natural engagement—Ne', the 'Natural expression of action—Na' and the 'Realistic feedback—Rf' are 'Low' or 'Very Low'. The fact that the virtual tours do not provide interaction functionality is considered to be the reason that the 'Compatibility with the user's task and domain—C' is evaluated to be 'Low'.

6. Conclusions

Internet technologies have the tremendous potential of offering virtual visitors ubiquitous access via the worldwide web to a virtual museum environment [25]. However, these technologies often encounter several problems. Taking into account the usability problems that are often encountered in VR environments [9,11], the authors present a solution to the research question of combining heuristics for the evaluation of virtual tours of museums and present an evaluation experiment using a combination of two multi-criteria decision-making theories.

For the evaluation of the museums' virtual tours and their classification, we used multi-criteria decision-making theories. Indeed, the evaluation of a virtual tour is a multi-dimensional problem that requires the evaluation of several criteria. Therefore, we have used AHP in combination with fuzzy TOPSIS for the implementation of the evaluation experiment. These two theories seem rather complementary and have been used in several domains but never before for the evaluation of virtual tours.

AHP has the ability to model expert opinion and, therefore, was considered ideal for this evaluation experiment since the set of the decision-makers consisted only of experts. Therefore, AHP was used for the calculation of the weights of the criteria. The application of the theory assigned the highest weights to the three following criteria: 'Close coordination of action and representation—Cr', 'Navigation and orientation support—No' and 'Support for learning—L', while criteria such as 'Consistent departures—Cd', 'Compatibility with the user's task and domain—C' and 'Natural expression of action—Na' were found by the application of the AHP to be the least important criteria.

AHP was considered ideal for calculating the weights of the criteria but it was not selected to be used for the evaluation of the virtual tours because it is a time-consuming technique due to the mathematical calculations required. However, the main drawback of AHP is that the number of pairwise comparisons increases as the number of alternatives and criteria increases or changes. Since complexity rises with the increase in websites, the number of alternatives that can be compared is limited. This is one of the main reasons for our decision to combine AHP with another theory. The theory that was selected to implement the empirical method in the second phase of the evaluation experiment was fuzzy TOPSIS. This particular theory was selected because the complexity of TOPSIS's application does not increase at the same rate as AHP when the number of alternative websites increases. Therefore, this theory was considered to be more appropriate for the evaluation of virtual tours. A main drawback of TOPSIS is, however, that it does not provide a specified way for calculating the weights of criteria as AHP does.

Taking into account the advantages and disadvantages of AHP and TOPSIS, these two theories have different reasonings but seem rather complementary. The main advantage of fuzzy TOPSIS is that it uses linguistic variables and, therefore, it is easier for all evaluators to characterize the criteria. The linguistic variable is then transformed into triangular numbers and then all estimations take place.

The main problems identified in our evaluation of the virtual tours of the museums in Italy had to do with the implementation issues of the VR interaction, namely: (i) the navigation functionality; (ii) the interactivity of the represented space and exhibits, and (iii) the information communication methods. Most of the reviewed virtual tours do not provide additional information related to the exhibits, e.g., annotations, supplementary content, etc., which could be of use for learning purposes. Despite these minor problems that were identified, the general impression was positive. The interaction with the users was very good and friendly and the experts did not spot serious problems.

Author Contributions: Conceptualization, K.K. and A.A.; methodology, K.K.; investigation, A.A., K.K., V.K., K.O.; data curation, A.A., K.K.; writing—review and editing, K.K., A.A., V.K., K.O.; project administration, K.K.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Walczak, K.; Cellary, W.; White, M. Virtual museum exhibitions. *Computers* **2006**, *39*, 93–95. [[CrossRef](#)]
2. Barbieri, L.; Bruno, F.; Muzzupappa, M. Virtual museum systems evaluation through user studies. *J. Cult. Herit.* **2017**, *26*, 101–108. [[CrossRef](#)]
3. Latos, A.; Komianos, V.; Oikonomou, K. Interaction and Information Communication in Virtual Museums. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2018; Volume 364, p. 012038.
4. Komianos, V.; Kavvadia, E.; Oikonomou, K. Cultural heritage recommendations and user navigation in large scale virtual environments. *IJCISudies* **2015**, *4*, 151–172. [[CrossRef](#)]
5. Karoulis, A.; Sylaiou, S.; White, M. Usability Evaluation of a Virtual Museum Interface. *Informatica* **2006**, *17*, 363–380.
6. Othman, M.K.; Petrie, H.; Power, C. Measuring the usability of a smartphone delivered museum guide. The 9th International Conference on Cognitive Science. *Procedia Soc. Behav. Sci.* **2013**, *97*, 629–637. [[CrossRef](#)]
7. Hix, D.; Gabbard, J.L. Usability engineering of virtual environments. In *Handbook of Virtual Environments: Design, Implementation and Applications*; Stanney, K., Ed.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2002; pp. 681–699.
8. Stanney, K.M.; Mollaghasemi, M.; Reeves, L.; Breaux, R.; Graeber, D.A. Usability engineering of virtual environments (VEs): Identifying multiple criteria that drive effective VE system design. *Intell. J. Hum.-Comput. Stud.* **2003**, *58*, 447–481. [[CrossRef](#)]
9. Bazargan, K.; Falquet, G.; Metral, C. Testing the usability of information visualization techniques in interactive 3D virtual environments: A toolkit. In Proceedings of the IADIS International Conference Interfaces and Human Computer Interaction, Freiburg, Germany, 28–30 July 2010.
10. Nielsen, J. *Usability Engineering*; Academic Press: New York, NY, USA, 1993.
11. Sutcliffe, A.; Gault, B. Heuristic evaluation of virtual reality applications. *Interact. Comput.* **2004**, *16*, 831–849. [[CrossRef](#)]
12. Sylaiou, S.; Killintzis, V.; Paliokas, I.; Mania, K.; Patias, P. Usability evaluation of virtual museums' interfaces visualization technologies. In Proceedings of the International Conference on Virtual, Augmented and Mixed Reality, Crete, Greece, 22–27 June 2014; pp. 124–133.
13. Kabassi, K.; Maravelakis, E.; Konstantaras, A. Heuristics and Fuzzy Multi-criteria Decision Making for Evaluating Museum Virtual Tours. *Int. J. Incl. Mus.* **2018**, *11*, 1–21. [[CrossRef](#)]
14. Chou, S.-Y.; Chang, Y.-H.; Shen, C.-Y. A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *Eur. J. Oper. Res.* **2008**, *189*, 132–145. [[CrossRef](#)]
15. Saaty, T. *The Analytic Hierarchy Process*; McGraw-Hill: New York, NY, USA, 1980.
16. Chen, C.T. Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets Syst.* **2000**, *114*, 1–9. [[CrossRef](#)]
17. Mulubrhan, F.; Akmar Mokhtar, A.; Muhammad, M. Comparative Analysis between Fuzzy and Traditional Analytical Hierarchy Process. *MATEC Web Conf.* **2014**, *13*, 01006. [[CrossRef](#)]
18. Hwang, C.L.; Yoon, K. *Multiple Attribute Decision Making: Methods and Applications*; Springer: New York, NY, USA, 1981.
19. Soleymaninejad, M.; Shadifar, M.; Karimi, A. Evaluation of Two Major Online Travel Agencies of US Using TOPSIS Method. *Digit. Technol.* **2016**, *2*, 1–8.
20. Zhang, W. Group-Buying Websites Evaluation Model Based on AHP-TOPSIS under the Environment of Multi-Attribute Decision-Making. *Int. J. Multimed. Ubiquitous Eng.* **2015**, *10*, 31–40. [[CrossRef](#)]
21. Nagpal, R.; Mehrotra, D.; Kumar Bhatia, P.; Sharma, A. Rank University Websites Using Fuzzy AHP and Fuzzy TOPSIS Approach on Usability. *Int. J. Inf. Eng. Electron. Bus.* **2015**, *1*, 29–36. [[CrossRef](#)]

22. Büyüközkan, D.; Ruan, D. Evaluating government websites based on a fuzzy multiple criteria decision-making approach. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2007**, *15*, 321–343. [[CrossRef](#)]
23. Sweetnam, M.S.; Agosti, M.; Orio, N.; Ponchia, C.; Steiner, C.M.; Hillemann, E.-C.; Siochrú, M.; Lawless, S. User needs for enhanced engagement with cultural heritage collections. Proceedings of Second International Conference TPD, Paphos, Cyprus, 23–27 September 2012; pp. 64–75.
24. Sirah, S.; Mikhailov, L.; Keane, J.A. PriEsT: An interactive decision support tool to estimate priorities from pair-wise comparison judgments. *Int. Trans. Oper. Res.* **2015**, *22*, 203–382.
25. Sylaiou, S.; Liarokapis, F.; Kotsakis, K.; Patias, P. Virtual museums, a survey and some issues for consideration. *J. Cult. Herit.* **2009**, *10*, 520–528.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).