

Cross-object user interfaces for personalised virtual reality museum tours

Xiangdong Li, Associate Professor at Zhejiang University, together with researchers from Stockholm University are using Cross-object User Interfaces (COUIs) that include a deep learning-based model to predict user's visual attention. The COUIs predict the users' personal interests by learning through each interaction and delivering the right content to each visitor in the context of their previous interactions. They provide each visitor with a virtual experience that has been tailored to suit their preferences and requirements.

Dragon Images/Shutterstock.com



Museums have traditionally been places for displaying artwork and exhibiting artefacts. However, museums are now becoming more than physical places to visit as they evolve to align with the digital age. Visitors have conventionally gathered information from object surfaces such as signs and posters, but museums are also opening their digital doors as they integrate digital media and virtual reality into their collections. Virtual visitors can access the work and narratives surrounding the exhibits. In the virtual reality museum, a multitude of videos exist within virtual objects accessed via user interfaces.

THE VIRTUAL EXPERIENCE

Designers and researchers are striving to provide visitors with a richer interactive experience combining the physical exhibition and artefacts with virtual objects. To improve the experience, museums want to provide real-time background information, accurate remote access to exhibits and immersive interactions, such as 360-degree views.

They also need to ensure smooth transitions between the physical objects and the virtual realities.

There is a variety of user interfaces available, but for museums to make an informed choice, they need to know what sort of experience their visitors can expect as the interfaces perform various video interactions. This is an area that still

requires further investigation, as current research into virtual reality museums has been mainly driven by the conventional desktop metaphor.

Xiangdong Li, Associate Professor at Zhejiang University, together with researchers from the ZJU-Alibaba IDEA lab and Stockholm University, are on a quest to fill this knowledge gap. They have carried out a study comparing participants' learning experience using cross-object user interfaces, conventional card-style user interfaces and the plain virtual reality user interfaces within the virtual reality museum and developed a deep learning model that could predict users' visual attention from history interaction patterns.

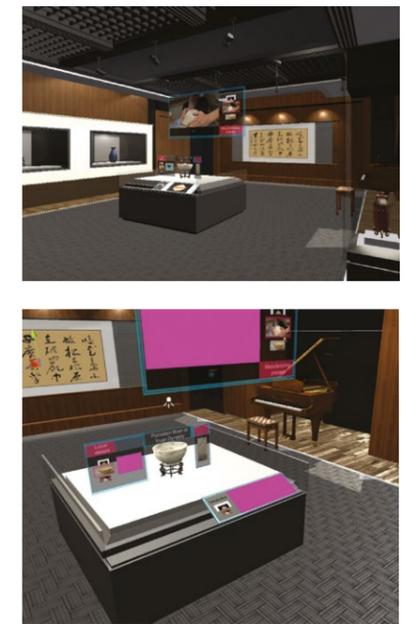
CROSS-OBJECT USER INTERFACES

Cross-object user interfaces (COUIs) are particularly suited to real-time adaption on virtual reality environments. Li explains: "Put simply, COUIs act as an instance of spatially distributed user interfaces that anchor to different virtual objects and can appear and disappear in run time. It differs from conventional graphic user interfaces on monitors, as it adopts the analogy of the information displayed on different objects' surfaces in the physical world."

In this study, the researchers have used COUIs that include a deep learning-based model that predicts users' visual attention from past eye movement patterns and



The COUIs in the virtual reality museum (the fully attached, semi-attached, and fully detached user interfaces for exhibits).



uses convolutional neural networks to suggest the next objects to be viewed. The COUI's delivery mechanisms determine when, where and what should be displayed next as it aims to predict the users' personal interests by learning through each interaction and delivering the right content to each visitor in the context of their previous interactions. Essentially, they provide visitors with a virtual experience that has been tailored to suit their preferences and requirements.

VIRTUAL REALITY MUSEUM

Museums offer visitors a rich learning experience through interaction with artefacts and multimedia. Together with lights, audio and video they can use interactive techniques, including webcams, multitouch and Bluetooth on tablets and smart phones, to recognise exhibits, determine the visitor's location and present information. Recent advances in head-mounted displays enable an enhanced sense of immersion with finely reproduced artworks and the spatial relationships between them and the exhibition environment being reproduced in virtual reality with spatially distributed user interfaces.

The researchers carried out a review of the literature and found a lack of understanding of COUIs for video

interaction in the virtual reality museum. There was specifically a lack of knowledge around how the COUIs affect the perceived system usability and how the COUIs influenced the users' learning experience through video interaction. These led the research team to propose two hypotheses:

- H1: In the virtual reality museum, COUIs for video seeking and watching have advantages pertaining to the perceived usability over the other two interfaces.
- H2: COUIs offer a better learning experience than the other two interfaces.

Museums are also opening their digital doors as they integrate digital media and virtual reality into their collections.

METHODOLOGY

The researchers recruited 45 students who were randomly assigned to three groups:

- Group A: using cross-object user interfaces
- Group B: using conventional card-style user interfaces
- Group C: using plain virtual reality museum user interfaces

Thirty-nine of the participants were new to head-mounted, display-based virtual reality interaction, while the remaining six had a little experience of using a VR headset. The participants' experience of

museums was recorded using a pre-study questionnaire. The researchers showed the participant a video explanation of their intended interactions with the virtual reality museum. This was followed by a demonstration of the virtual reality museum with the Oculus DK2 headset and video interaction equipment.

All participants had to perform a walkthrough of the virtual reality museum following a particular path, look at the exhibition item and learn related information. They were then asked to complete questionnaires and semi-formal interviews, providing feedback on the

museum scene and related interactions. The positions of the user interfaces were established from the participants' eye movements and the distance between the participants and the object.

ANALYSIS

Established system usability scales were used to measure how the different user interfaces' stimuli influenced the perceived usability of the COUIs. The researchers used a collection of metrics to measure the users' learning experience during study tasks. These were made up of the effectiveness of learning, the level of understanding, the level of interest, content appreciation and learning memory.

RESULTS

No strong advantages of the COULs over the other interfaces were apparent with respect to both the perceived usability and learning experience. All three interfaces had similar levels of effectiveness and ease of use. The COULs, however, had significantly lower satisfaction scores than the conventional card-style user interfaces, but this could be a result of the study tasks. The COULs task durations were shorter than the other two interface types which show greater efficiency with shorter eye fixation durations and higher saccade frequencies. (Saccade is a rapid jerk-like movement of the eyeball which redirects the visual axis to a new location.)

Given these findings, hypothesis 1 cannot be accepted as the COULs did not show significant advantages in perceived usability. Hypothesis 2 is also rejected as the COULs did not show a significantly different learning experience when compared with the other two interfaces.

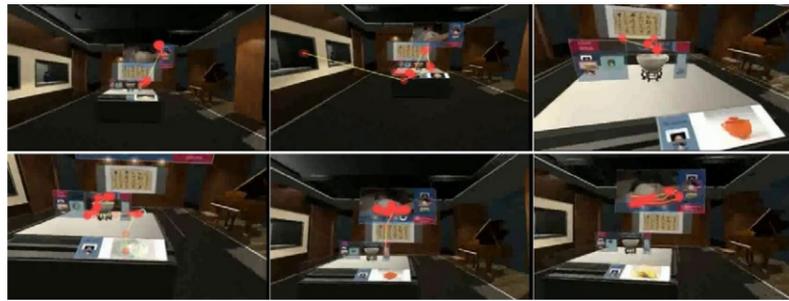
DEVELOPMENTS

Real-time adaptation is a significant issue that personalised human-computer interaction has to develop. Virtual reality interactions mean that while users operate tools in the physical world, they simultaneously view virtual objects in the virtual world, which adds to the complexities involved in personalising interactions in virtual environments.

To resolve this issue, Xiangdong Li and his team propose the COULs for personalised virtual reality touring. They again chose the virtual reality museum as the main study scenario as it incorporates many exhibits in a variety of forms, most participants are familiar with the museum and the tasks are not competitive.

EXPERIMENT

The researchers constructed a virtual reality exhibition of ancient vases and paintings in a room. A user's visual attention is made up of many factors such as the eye gaze position, speed, direction and virtual distances together with the individual's personality and preferences. The researchers concentrated



Above: COULs predict the users' next possible visual interests according to past eye movements.



Above: COULs adapt to prediction results and adjust virtual reality museum scenes and exhibits.

on eliciting correlations between a user's visual attention and the object that they will view next. They used a deep learning algorithm, a convolutional neural network (CNN), to process eye movements and predict the next virtual object that the user would choose to view in virtual 3D spaces.

This time a cognitive walkthrough was completed by five experts from the university's digital media department using the same Oculus DK2 headsets as before.

The participants were given the following information:

'You are about to experience a virtual reality exhibition room. During your tour around the room, you will be presented with personalised user interfaces at runtime based on personal interests captured from your eye movements.'

Then the researchers defined the tasks. The experts would enter the virtual reality environment, tour the virtual space, look at one or more objects, interact with the COULs that appear and navigate the virtual space. The experts were then asked to complete an evaluation questionnaire

and report on any factors that they felt might hinder the desired performance.

DISCUSSION

After iterative training, the model achieved 86.3% accuracy in predicting real-time eye movements. The predictive model considered the user-object distance in virtual 3D space and predicted which virtual object, or part of object, the user was about to view. This is valuable in helping researchers and designers understand individuals' cognitive states in 3D space and support personalising interactions.

The research team aim to personalise the user interface rather than the contents. The delivery mechanisms incorporate spatially distributed user interfaces that are driven by a real-time visual attention prediction model and combine a number of factors to determine the time, form and content of the COULs display.

Xiangdong Li stresses that COULs could be employed to personalise interaction in virtual reality and other environments including augmented reality and mixed reality. He also highlights that these studies have identified future research challenges such as constructing adaptive prediction models for personalised interaction in virtual environments.

Real-time adaptation is a significant issue that personalised human-computer interaction has to develop.



Behind the Research

Professor Xiangdong Li

E: axli@zju.edu.cn T: +86 0571 87952010 W: www.zju.edu.cn

Research Objectives

The flexibility of distributing user interfaces in any positions at any sizes in the virtual reality gives designers great challenges. To investigate how users perceive and interact with the user interfaces that are spatially integrated across virtual objects, we developed three representatives of cross-object user interfaces. Then, we evaluated how the cross-object user interfaces supported virtual reality museum navigation by comparing it with conventional user interfaces and card-style user interfaces. Given the understanding of cross-object user interfaces in the museum environments, we aimed to understand how the users manipulated their personal visual attentions during the navigation tasks and how we could predict the users' next possible visual interest according to the past activities. To fulfil the goal, we adopted deep learning methods to prototype a prediction model that was trained with a large set of eye-tracking data. The prediction result is promising and generalising for other virtual reality scenes.

Detail

Assoc. Prof. Xiangdong Li
109, Zetonglou
Yuquan Campus
Zhejiang University
38 Zheda Road
Hangzhou
China P.R., 310027



Bio

Xiangdong Li is an Associate Professor in design at the College of computer science and technology, Zhejiang University. His research interests mainly include cross-device computing and intelligent user interfaces, with a focus to leverage the interaction between humans and machines in physical, virtual, and hybrid spaces.

Funding

The project is supported by National Key R&D Program of China (2016YFB1001304), National Natural Science Foundation of China (61802341).

Collaborators

Professor Lingyun Sun and Professor Weidong Geng from Zhejiang University built the virtual reality museum environment and developed the deep learning model of visual interest prediction, respectively. Prof Preben Hansen from Stockholm University provided valuable feedback on the user study.

References

Sun, L., Zhou, Y., Hansen, P., Geng, W & Li, X. (2018). Cross-objects user interfaces for video interaction in virtual reality museum context. *Multimedia Tools and Applications*, 77, 29013–29041. Available at: <https://doi.org/10.1007/s11042-018-6091-5> [Accessed 01 March 2019].

Sun, L., Zhou, Y., Geng, W & Li, X. (2018). Towards Personalised Virtual Reality Touring Through Cross-objects User Interfaces. *Personalised HCI 2018*, book chapter, in publication (by April 2019).

Personal Response

What initially prompted your research into using Cross-object User Interfaces?

Virtual reality is an imaginary digital space that can implement any types of fancy user interfaces in a way that the physical world would never be able to realise. So, we analysed how people interact with information that is distributed across physical object surfaces and started to design cross-object user interfaces that are spatially distributed with virtual objects in a virtual reality museum. And given the eye tracking dataset captured from the study, we continued to develop a deep learning-based model to predict users' visual attention, to better support cross-object user interface delivery in the virtual reality museum.

What are your plans for future developments in personalised virtual reality touring?

Our future plan for Cross-Object User Interfaces is in two directions. Firstly, we will continue to improve the deep learning-based prediction model for higher prediction accuracy in versatile virtual reality scenes. This will lead Cross-Object User Interfaces into generalising user interfaces. Secondly, we are currently studying techniques and methods of automatic scene construction in the virtual reality. This will enable generative construction of virtual reality museums that can automatically build scenes and objects with the anticipation of users' personal interests. Ultimately, we aim to transfer the techniques and methods to generalising environments such as augmented reality.