WoBo: Multisensorial travels through Oculus Rift

Stefano Fibbi  
Dept. Mathematics and Computer Science  
University of Cagliari  
Via Ospedale 72, 09024, Cagliari, Italy  
stefano.fibbi@gmail.com

Fabio Sorrentino  
Dept. Mathematics and Computer Science  
University of Cagliari  
Via Ospedale 72, 09024, Cagliari, Italy  
fabio.sorrentino@unica.it

Lucio Davide Spano  
Dept. Mathematics and Computer Science  
University of Cagliari  
Via Ospedale 72, 09024, Cagliari, Italy  
davide.spano@unica.it

Riccardo Scateni  
Dept. Mathematics and Computer Science  
University of Cagliari  
Via Ospedale 72, 09024, Cagliari, Italy  
riccardo@unica.it

Abstract  
WoBo (World in a Box) aims to provide a new experience for travellers, allowing them to visit distant or hardly reachable places through the exploitation of consumer cameras and a head mounted display. The experience consists in watching a 360-degrees video with 3D audio in a dedicated cabin. The user can select videos shot in different places, which have been created with six consumer cameras. We describe the proposed experience, the hardware and the software used for a first prototype.

Author Keywords  
Presence, Multisensorial, Head Mounted Displays

ACM Classification Keywords  
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

Introduction  
In many science fiction movies, one of the most desired possibilities is breaking the limit of physical distances for reaching far places, envisioning high speed vehicles, teletransportation or telepresence mechanisms. With WoBo (World in a Box), we propose a new way of interpreting the concept of travelling, able to provide the user with the possibility of visiting distant or hardly reachable places, simply by wearing a head-mounted display inside an
appropriate cabin. In this way, the users have the impression to be carried away in a new reality, and it will be as if the whole world could fit in the cabin. We describe in this paper the experience, the technical set-up of a first prototype and the results of a preliminary evaluation, showing an encouraging acceptance by end-users.

WoBo

World in a Box (WoBo) provides immersive and realistic virtual journeys using the latest VR technologies, allowing a fast and cheap world (virtual) travelling. Both virtual and real world destinations are reachable entering in a special box that we called WoBoX. We envision the installation of several WoBoXes around the world, all connected to each other. For instance, it would be possible for the user to step into one located in New York and watch what is happening around the WoBoX in Seoul. In addition, it would be possible to have WoBoXes also on beautiful but hardly reachable places, such as mountain tops (e.g. the Everest), deserts etc. Each WoBoX is both able to transmit and receive signals to and from other boxes. Therefore, a WoBoX consists of two parts: the first one is a 360 degrees video recording equipment, for creating an immersive and realistic visualization of the area around the WoBoX. The second part is a cabin where the user can, wearing a head mounted display and a set of earphones, experience the travel.

Video and audio acquisition

We acquire a single 360-degree video stitching six different ones captured from six cameras. We position two cameras along each one of the three orthogonal axes in a 3D space, looking at opposite directions. We created a small mount for securing the camera position while recording, following the schema in Figure 1. In our idea, such cube should be positioned on the top of each WoBoX, in order to capture in real-time what is happening around a specific place.

We used six GoPro Hero3+ in our prototype implementation. For having a complete coverage of the space around the cube mount, we shoot 4:3 videos having a resolution of 1440p at 30 fps. The 16:9 ratio was not an option, since it does not cover the entire space around the cube, considering the two horizontal cuts of the sensor image.

The six recorded videos are merged into a single panoramic video that can be mapped over a sphere as texture. This merging operation is done using a third party software (Kolor Autopano video) that uses a stitching algorithm to merge and create the final 360 degrees video. The algorithm uses a feature-based approach for the video stitching, identifying the different points where it is possible to combine the frames coming from the different cameras. More precisely, the software exploits the SIFT features [1] for calculating the salient points and computing the matchings between two camera frames. In the final WoBoX setting, we envision an almost real time stitching and broadcast of the 360-degrees video, trying to create a fast version of the algorithm that exploits the particular configuration of the mount cube. We consider acceptable also a delay of a few seconds between the video shooting and the user visualization, provided the user has no means
for modifying the surrounding environment and, consequently, for noticing such delay during the interaction.

In order to record a three-dimensional sound of the scene we need to use the six GoPros built-in microphones using the same concept of the binaural recording method. In our work we can not adopt the exact binaural method due to the fact that the user can move her head around. To overcome this issue we have decided to record the surrounding sounds and then replay them accordingly to the user’s head position in the 3D space. In this way we are able to give a 3D stereo sound sensation to the user of actually being in the place where the sound have been recorded. Such recording method is intended for replay using exclusively headphones.

![Figure 2: 3D Scene with video texturing](image)

The prototype application for watching the 3D video has been developed using Unity 3D. In order to set-up the environment, we inserted into the world a high-quality model of a sphere. The 360-degrees video is mapped as a texture on the sphere, while the world camera, corresponding to the user’s head, is positioned in its center. The world has one light source, positioned again in the center of the sphere. We reversed its normals for viewing the video texture from the inside. The configuration of the 3D world is shown in Figure 2. As explained in the audio acquisition section, we record six sound sources using the six GoPros built-in microphones. In the Unity framework we can add all these sounds positioning them in the 3D scene and, in addition, we can stream two different audio tracks: one for the right ear and the other one for the left ear.

**Travel experience**

In order to exemplify the travelling experience, we built a first prototype of a WoBoX, with the aim to obtain something close to telephone booths, but without interfering with the user movements. The prototype consists of a wooden box, whose dimensions are 1.5m x 2m x 2m. To simulate the availability of different WoBoXes in different places, we recorded three videos in different sea places in Sardinia, Italy. Once a user enters in the WoBoX, she can select one of the available destinations. Then, she can start the virtual visit. The prototype is equipped with an Oculus Rift and high quality headphones, connected to a desktop computer running the Unity 3D application. Since it was not possible to avoid using a wired connection for the Oculus Rift, we tried to reduce the interference with the user’s movements fixing the wire at the top of the cabinet.

---

We conducted a small-scale user study in order to evaluate the proposed user experience. The study was organised as follows: After completing a small demographic questionnaire, the users watched for a minute a seascape video in the WoBoX, with the head mounted display and 3D audio. After that, the user filled a Presence questionnaire as in [2]. We removed the questionnaire items related to the Control Factors (see [2] for a description of the factors and their mapping on the questions) since they were not meaningful considering that the users can control only the camera point of view, but they are not able to modify the virtual world configuration. The item ratings exploit a 1 to 7 Likert scale. Twelve users participated to the test, 9 males and 3 females, aged between 22 and 29 years old ($\bar{x} = 23.33, s = 2.46$) and the test results were encouraging. First of all, the sensory engagement was highly rated considering the overall experience (q4: $\bar{x} = 5.67, s = 0.71$), the visual (q5 $\bar{x} = 5.83, s = 0.71$) and audio experience (q6: $\bar{x} = 5.83, s = 1.19$). In addition, the user did not find inconsistencies in the information coming from the different senses (q11: $\bar{x} = 2.42, s = 1.73$, 1 means no inconsistencies) and they were not confused after the session (q22: $\bar{x} = 2.33, s = 1.67$). The debriefing session after the test highlighted a major usability problem for the prototype represented by the wires (for both the Oculus Rift and the headphones) that, physically interfering with the exploration, degrades the feeling of being into another place. We plan in the future to exploit devices that embed a smartphone or a tablet inside the head mounted display (such as the Samsung Gear VR\textsuperscript{2}) in order to avoid using wires.

Conclusions
In this paper we introduced WoBoX, a new experience for travellers which aims to connect different locations reachable immediately for a virtual journey. A WoBoX is able to record a 360-degree video of the surrounding environment, which can be viewed in any other WoBoX in the world. We described the hardware and software setting for a first prototype and the results of a preliminary user evaluation of the proposed experience, demonstrating an high level of sensory engagement.

Acknowledgements
Lucio Davide Spano gratefully acknowledges Sardinia Regional Government for the Financial support (P.O.R. Sardegna F.S.E., European Social Fund 2007-2013 - Axis IV, Obj l.3, LoA 1.3.1)

References

\textsuperscript{2}http://www.samsung.com/global/microsite/gearvr/