Sensing Immersive 360º Mobile Interactive Video

João Ramalho and Teresa Chambel
LaSIGE, Faculty of Sciences, University of Lisbon, 1749-016 Lisboa, Portugal

Keywords: Video, Immersion, Presence, Perception, Auditory, Visual, Tactile Sensing, 360º, 3D Audio, Mobility, Movement, Speed, Orientation, Wind.

Abstract: Video has the potential for a strong impact on viewers, their sense of presence and engagement, due to its immersive capacities. Multimedia sensing and the flexibility of mobility may be considered as options to further extend the video's immersive capacities. Mobile devices are becoming ubiquitous and the range of sensors and actuators they incorporate is ever increasing, which creates the potential to capture and display 360º video and metadata and to support more powerful and immersive video user experiences. In this paper, we explore the immersion potential of mobile interactive video augmented with visual, auditory and tactile multisensing. User evaluation revealed advantages in using a multisensory approach to increase immersion and user satisfaction. Also, several properties and parameters that worked better in different conditions were identified, which may help to inform design of future mobile immersive video environments.

1 INTRODUCTION

Immersion is the subjective experience of being fully involved in an environment or virtual world. It may be defined as a feature of display technology determined by inclusion, surround effect, sensory modalities and vividness through resolution (Slater and Wilbur, 1997); (Douglas and Hargadon, 2000).

Immersion is associated with presence, which relates to the viewer's conscious feeling of being inside the virtual world (Slater and Wilbur, 1997), may include perceived self-location in the virtual world (Wirth et al, 2007), and benefit from realism, that can be enhanced through photo realistic images and spatial audio. Video allows great authenticity and realism, and it is becoming ubiquitous, in personal capturing and display devices, on the Internet and iTV (Neng and Chambel, 2010; Noronha et al, 2012). Immersion in video has a strong impact on the viewers' emotions, and especially arousal, their sense of presence and engagement (Visch et al., 2010). 360º videos could be highly immersive, by allowing the user the experience of being surrounded by the video. Wide screens and CAVEs, or domes, with varying angles of projection, possibly towards full immersion, are privileged displays for immersive video view for their shapes and dimensions, but they are not very handy, and especially CAVEs are not widely available. On the other hand, mobile devices are commonly used and represent, by the sensors and actuators they are increasingly incorporating, a wide range of opportunities to capture and display 360º and HD video and metadata (e.g. geo-location and speed) with the potential to support more powerful and immersive video user experiences. Actually, mobile devices could be flexible enough to allow users to actually turn around, as if they hold in their hands a window to the video where they are immersed in, while watching and sensing it, as if they were there, and bring this experience with them everywhere. As second screens, mobile devices may also be used to help navigation in a video that is projected or displayed outside in a wider screen, and even to decide to catch the current video on that screen (e.g. TV) and go on watching it on the move, for an increased sense of immersion and flexibility.

In this paper, we explore the immersion potential of mobile interactive video augmented with visual, auditory and tactile multisensing, through the design and evaluation of new features in Windy SS (WSS). This is a mobile application for the capture, search, visualization and navigation of georeferenced 360º immersive interactive videos (through hypervideo), along trajectories, designed to empower users in their immersive video experiences, both accessing other users' videos and sharing their own. The focus of this paper is on perceptual sensing and its impact on immersion, especially in an increased sense of presence and realism, through the feeling of being inside the video, viewing and experiencing movement speed and orientation. Different conditions were tested, varying: types of video, viewing modes, spatial sound and tactile sensing approaches, mostly based on wind. Results confirmed advantages in using a multi-sensory approach to increase immersion, and identified which properties and parameters worked better and are more satisfying and impactful in different conditions, that may help to inform design of future mobile immersive video environments.

After this introduction, section 2 presents most relevant related work, section 3 presents sensing features of Windy SS that are evaluated in section 4 with a special focus on the immersive experience. The paper concludes in section 5 with conclusions and perspectives for future work.

2 RELATED WORK

The work presented in this paper builds on our previous work on 360º hypervideo (Neng and Chambel, 2010), developed for PCs, which evolved to allow capturing, sharing and navigating georeferenced 360º videos and movies, synchronized with maps, and crossing trajectories (Noronha et al, 2012), allowing to 'travel' in other users 'shoes'. Briefly, related work (see (Neng and Chambel, 2010); (Noronha et al., 2012)) concerns to hypervideo and immersive environments (mainly VR and AR, images, like Google Street View, seldom video), georenferencing and maps, orientation, cognitive load, and filtering. Bleumers et al., (2012) found that certain genres are more suitable for 360º video from a user perspective (e.g. hobbies, sports, or situations with little progress, inviting for exploration). On mobiles, relevant related work concerns to navigation (Neng and Chambel, 2010), recent PanoramaGL lib for 360º photos viewing, second screens (Courtois and D'heer, 2012), and the use of sensors and actuators, e.g. in art installations where user's movement influences wind (fans) blowing trees (Mendes, 2010).

Moon et al., (2004) showed that the use of wind output increased the sense of presence in VR, but their application did not allow user interaction, as the movement occurred in a pre-defined animation path. Cardin et al., (2007) presented a head mounted wind display for a VR flight simulator application. In the experience, participants determined the wind direction with a variation of 8.5 degrees. Lehmann et al., (2009) evaluated the differences between visual only, stationary and head mounted wind, with considerable increase in presence in stationary and head mounted wind prototypes. But these approaches target VR - not video, nor mobile environments - as they require heavy and very specific equipment. Furthermore, none of them presents methods to capture wind metadata and couple it with video as a way to increase realism and immersion.

Sound may be used to convey information related to movement, like speed and orientation, but for an intuitive mapping, it is necessary to take human perception into account. Very recently, Merer et al., (2013) addressed the question of synthesis and control of sound attributes from a perceptual point of view, based on a study to characterize the concept of motion evoked by sounds. This concept is not straightforward, involving actual physical motion and metaphoric descriptions, as used in music and cartoons. They used listeners questionnaires and drawings (that can be associated with control strategies as continuous trajectories, to be used in applications for sound design or music), focusing on aspects like shape, direction, size, and speed, and using abstract sounds, for which the physical sources cannot be easily recognized.

3D-sound can significantly enhance realism and immersion, by trying to create a natural acoustic image of spatial sound sources within an artificial environment, most development been made by the film industry (Dobler and Stampfl, 2004). Approaches like (Namezi and Gromala, 2012) address sound mapping in virtual environments (VE), with a concern on affective sound design to improve the level of immersion. They present a model, addressing the use of procedural sound design techniques to enhance the communicative and pragmatic role of sound in VE, concluding e.g. that soundwalks using headphones produce the most realistic representation of sound because they provide feedback such as distance, elevation, and azimuth. Most of these approaches address audio in general, production of film soundtracks or tend to address virtual reality scenarios, but not scenarios of augmenting immersion in videos.

3 SENSING IN WINDY SS

WSS (Ramalho and Chambel, 2013) is an interactive system capable of capturing, publishing, searching, viewing videos and synchronizing with interactive TVs in new ways. During video capture, while a

Sony Bloggie video camera captures the 360º videos, a smartphone registers several metadata relative to the video, such as the geo-references, speed, orientation, through GPS, and weather conditions (including wind speed and orientation), through the OpenWeatherMap(.org) webservice. After published in the community, videos can be searched either by using a set of keywords and filters, or by selecting map areas and drawing paths on a map.

When viewing videos, they can be navigated not only in time, but also through their geographic position, using a map. For instance, when using the application in an "Interaction with TVs and Wider Screens" mode (Ramalho and Chambel, 2013), in which the video is reproduced in a TV and the mobile device is used as a Second Screen, the mobile device shows a map with the route traversed by the video is depicted, and a marker indicates the current geographical position of the video being viewed. Routes correspondent to videos recorded on the proximities are also shown in the map. Dragging the reproduction marker to other points of the route, or to other routes enables users to navigate through the video or between other videos. Having the capture, search and navigation of videos been addressed in our previous work (Ramalho and Chambel, 2013), and in order to explore the immersion potential of mobile video, Windy SS was augmented with new features for visual, auditory and tactile multi-sensing, whose design rational is presented next.

3.1 Visual Sensing in 360º Video

With the aim of increasing immersion, one of the main challenges regarding 360º video relates to the way video is viewed and interacted with. Our Research Question 1 (RQ1) was defined as: "Would a full screen pan-around interface increase the sense of immersion 'inside' the 360º video?".

Two designs were conceived to address this RQ1. In both, videos are displayed in full screen on an Android tablet. 360º videos are mapped onto a transitional canvas that is in turn rendered around a cylinder, to represent the 360º view and allow the feeling of being surrounded by the video. As design 1: taking advantage of the compass within the tablet, and building upon an idea by Amselem (Amselem, 1995), by moving the tablet around, the user can continuously pan around the 360º video in both left and right directions, as if it was a window to the 360º video surrounding the user. However, although the option to pan the video by moving the device can be a very realistic and immersive approach to pan around 360º videos, there might be some situations where the user is not willing to move the device, such as when the user is seated on a couch. In order to suit both scenarios, design 2 was conceived: users can pan around the video without having to move, by making the entire screen consist of a drag interface. By swiping to the left or right with one finger over the video view, the video angled is panned accordingly (Fig.1).

3.2 Tactile Sensing through Wind

Striving to increase immersion through sensing the experience, we focused on the following Research Question: "Does wind contribute to increasing realism of sensing speed and direction in video viewing?" (RQ2). Thus, a Wind Accessory was developed (Fig.1, 4). This prototype is based on the Arduino Mega ADK; it is mounted on the back of the tablet and controls two fans generating a maximum combined air flow of 180 CFM (Cubic Feet per Minute). The purpose of this device is to blow wind to the viewer during video reproduction, creating a more realistic perception of speed and movement, and thus increasing the sense of presence and immersion. Therefore, the Wind Accessory operates its fans in real time according to messages received from the Windy Sight Surfers application.

3.2.1 Communication with the Wind Accessory

When a video is to be reproduced, a new Wind Accessory communication session is initiated, and the application sends messages to the Wind Accessory specifying the frequency the fans are to be rotated (one message per second). These values directly affect the RPM (Revolutions Per Minute) of the fans, and are calculated according to information contained in the video's metadata file. More specifically, the wind values take into account the wind speed and orientation during the video's recording (so that, while the camera is facing against the wind orientation, the fans' RPM are much higher than when the camera is aligned with the wind orientation), the speed the user was travelling during the recording, and the angle of the video being viewed during video reproduction. The way these factors are taken into account to calculate the values that will be sent to the Wind Accessory involves a three-step normalization process, which is described next.

3.2.2 Three-step Normalization

The first step is the normalization of the speed values contained in the video's metadata file according to the wind speed registered by the OpenWeatherMap web service (in MPS – Meters per Second). Taking into account the MPS wind speed values scale, and the respective "Effects on Land" scale, which are comprised in the Beaufort Scale (http://www.unc.edu/~rowlett/units/scales/beaufort.html), a pairing was established between the wind speed values, and a factor (0.8, 0.9 or 1), by which the wind values are multiplied. More specifically, if the wind speed value is less than 8, than the wind factor is 0.8; if the wind speed value is between 8 and 17, than the wind factor is 0.9; if the wind speed value is greater than 17, than the wind factor is 1. Before the video starts playing all the values in the video's metadata file are normalised by this factor value. This step is done in order to take into account the wind speed during the video's recording.

An example of benefit of this step is the case where two similar videos of the exact same path are recorded at the exact same speed in different occasions. In the first, it was a sunny day with very low wind speed values, whereas in the second video it was a very windy day, and thus with high wind speed values. In this situation, this filter enables the user to notice that one of the videos was recorded in a windy environment.

The second step normalizes the values obtained in the first step. This step is performed just to convert the values to the Pulse Width Modulation (PWM) range of the Arduino platform (0-255).

The third, and last, step occurs during video reproduction and relates to the angle the user is viewing at each moment while viewing the video. If in the real (recording) situation the person is moving against the wind direction, the user will feel much more wind resistance when compared to the situation where the person is moving along the wind direction. This situation also happens when the person moving turns their head around: as the human hears' shape allows sound coming from the from to be much more audible than sounds coming from the back, when a person turns his head against the wind direction, the hearing perception is that the wind is much more strong than when the head is turned to the wind direction. In order to mimic this characteristic, with the intent of increasing the immersiveness of the experience, before each message is sent to the Wind Accessory, the value obtained in the second step that is about to be sent goes through a last normalization. Before the message is sent, the angle of the video being viewed is taken into account so that if the user is viewing the angle of the video that corresponds to the wind orientation during recording, than the value is multiplied by 1; if the user is viewing the angle of the video that is the opposite to the wind orientation during recording, than the value is multiplied by 0.6; if the user is viewing an angle of the video that is approximately between 90º of the wind orientation during recording, than the value is multiplied by 0.8.

After this three-step normalization process, during video reproduction each value is sent to the Wind Accessory, thus creating a wind perception of the video being viewed.

Figure 1: Drag interface being used while viewing a 360º video. Wind accessory coupled with the tablet.

3.3 Auditory Sensing: Spatial Audio

When a video is shot, the sound is usually recorded accordingly to the orientation of the camera. This can create orientation difficulties for users, as the sound does not match the 360º characteristics of the video. Therefore, the following question arises: "Does a 3D mapping of the video sound allow for easier identification of the video orientation while it is being reproduced?" (RQ3) This experiment was accomplished using JavaScript's Web Audio API (http://www.w3.org/TR/webaudio/), being that any standard set of stereo headphones can reproduce the changes accordingly created effect, although high quality headphones are able to increase the realism of the referred effect. In this sound space, the sound source's position is associated with the video trajectory direction and, therefore changes in accordance with the angle of the video being visualized. That is, if the user is visualizing the front angle of the video, the sound source will be located in front of the user's head; if the user is visualizing the back angle of the video, the sound source will be located in the back of the user's head. As videos are 360º, the sound source's location changes over a virtual circle around the user's head (Fig.2, 4).

During user evaluation, users provided feedback on the best values for the virtual "distance" between users' head and the sound sources.

Figure 2: 3D Audio: source location changing around the 360º video viewing. Grey stripe on top represents video trajectory direction.

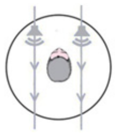

Figure 3: Doppler Effect: Audio changes cyclically as in grey paths.

3.4 Auditory Sensing: Cyclic Doppler Effect

The Doppler Effect can be described as the change in the observed frequency of a wave, occurring when the source and/or observer are in motion relatively to each other. As an example, this effect is commonly heard when a vehicle sounding a siren approaches, passes, and recedes from an observer. Given the fact that people inherently associate this effect to the notion of movement, we experimented to see if a controlled use of the Doppler Effect could increase the movement sensation of users while viewing videos (RQ inherent in RQ4 and RQ5, below). In order to do so, a second sound layer was added to the video, which cyclically reproduces the Doppler Effect in a controlled manner. This sound layer was also implemented using JavaScript's Web Audio API, and therefore the sound corresponding to the Doppler Effect is mapped onto a 3D sound space. In the basis of this sound layer is a sound that is reproduced cyclically and approaches, passes, and recedes the users' head (from the front to the back) (Figure 3).

Figure 4: Moving the tablet to view the 360º video. The connected headphones provide a 3D audio space.

Regarding the Doppler Effect, there are several aspects that influence the intensity of the movement sensation. It is especially affected by the intensity (volume) of the sound, and the rate at which it is reproduced. Also, the sound itself used to reproduce the effect can be of great importance, as some sounds might be more effective (create a stronger movement sensation), but also more intrusive (interfere with the main sound layer). With respect to the rate at which the sound is played, this value is set while playing and it varies during playback, according to the speed values stored while capturing the video (the value is updated every three seconds). In other words, the higher the speed, the higher the intensity of the Doppler Effect. Concerning the sound used to reproduce the Doppler Effect, as it will be described in section 4 (User Evaluation), several experiments were conducted with the intent to find out the right parameters. Several types of sounds were experimented, aiming to find the sounds that create a stronger movement sensation, while not being intrusive.

In this context, there is the need to answer the question: "What are the most effective sound categories to provide movement sensation, based on the Doppler effect?" (RQ4). This proven to be complex problem, as some of the most effective sounds were also considered the most intrusive, which resulted in a trade-off situation that is not easily resolved and might be grounds for further research. Also, the threshold level of the volume of the Doppler Effect sound layer's sound sources was measured.

3.4.1 Doppler Effect's High-pass Filter

One of the side effects of this approach might be to try to alert the user for movement when there is little/no movement. This can dramatically change the effectiveness of this feature by turning it into something obtrusive rather than beneficial. Therefore, this problem was also analysed with the intent to find if there is a minimum amount of movement required for the Doppler Effect to become beneficial, translating into the research question: "In which circumstances (movement degree) does the Doppler Effect increase immersion?" (RQ5). As results shown (section 4), there is a minimum amount of movement required, which led to the development of a high-pass filter that added the requirement for a minimum amount of movement in order for the Doppler Effect simulation to execute.

3.4.2 Used Sounds

Regarding the sounds used in the Cyclic Doppler

Effect feature, different sounds can drastically change the impact of this feature. Therefore, and given the strong emphasis that was put on the sound source's nature when implementing this component, different sounds, with different characteristics were experimented. The first sound to be experimented was the sound of wind. This is one of the sounds that may create a stronger movement sensation in the user. However, due to the fact that, from a conceptual point of view, the wind sound is quite complex (it does not consist of any waveform but rather consists of the combination of a large amount of waveforms), it can also be one of the most intrusive sounds. Therefore, several other sounds were experimented. Low frequency sounds (sounds with a low pitch) are known to be less intrusive than other sounds. This led to the experimentation of different sounds that reflect these characteristics.

In order to create the referred sounds, an analog synthesiser was used. The sounds were recorded in a computer directly running the synthesiser's output through an audio interface, being that these recorded and experimented sounds were created with the intention to reproduce the four main sound waveforms and investigate which, given their simplicity, are more suitable for the purpose of this application. Namely, it was designed one sound based on each of the four most basic waveforms: Sine Wave (which stands as the purest waveform, being the most fundamental building block of sound), Sawtooth Wave (characterised by having a strong, clear, buzzing sound; can be obtained by adding to a base Sine Wave a series of Sine Waves with different frequencies and volume levels (amplitudes) - refered to as Harmonics of the base Sine Wave), Square Wave (rich sound with a bright and rich timbre; not quite as buzzy as a sawtooth wave, but not as pure as a sine wave), and Triangle Wave (between a sine wave and a square wave; softer timbre when compared to square or sawtooth waves).

4 USER EVALUATION

We conducted a user evaluation of WSS's Experience Sensing features to investigate whether and in which conditions they contribute to a more immersive experience to the user. For each task, we learned about their perceived Usefulness, users' Satisfaction and Ease of use (USE) (http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html). Also, to test dimensions related to Immersion, we used self assesment approaches: the Self-Assessment Manikin (SAM,http://irtel.uni-mannheim.de/pxlab/demos/index_SAM.html), which measures emotion (pleasure, arousal and dominance) often associated with immersion; and additional parameters of Presence and Realism (PR) we found relevant. We evaluated global immersiveness in the WSS sensing experience, through a pre and post-event self-assessment Immersive Tendencies and Presence Questionnaires (Slater, 1999); (Witmer and Singer, 1998).

4.1 Method

We performed a task-oriented evaluation based mainly on Observation, Questionnaires and semi-structured Interviews. After explaining the purpose of the evaluation and a short briefing about the concept behind Windy SS, demographic questions were asked, followed by a task-oriented activity. Errors, hesitations and performance were observed and annotated. At the end of each of the twelve tasks, users provided a 1-5 USE rating, a 1-9 SAM rating and a 1-9 PR rating, and users' comments and suggestions were annotated. Slater states that Presence is a human reaction to Immersion (Slater et al., 2009). Therefore, by evaluating presence, one can tell about immersion capabilities of the system. To do so, users completed an adapted version of the seven-point scale format Immersive Tendencies Questionnaire (ITQ) before the experiment, and an adapted version of the Presence Questionnaire (PQ) after the experiment, with 28 questions each. At the end of the session, users were asked to rate the overall application in terms of USE dimensions, and to state which feature was their favourite.

The evaluation had 17 participant users (8 female, 9 male) between 18-34 years old (mean 24). In terms of literacy, all users had at least finished high school, they were all familiar with the concept of accessing videos on the Internet, but only 5 had previously interacted with 360º videos, and only 6 had heard about 3D audio. The foreseen time for the completion of the 12 tasks was 40 minutes, which was met by all users.

4.2 Results

Results are divided in two subsections, concerning: the perceptual sensing features, evaluated in terms of perceived usability, by USE, SAM and PR; followed by the Immersive Tendencies and the Presence Questionnaires results, and global overall comments. Results are commented along the corresponding tasks, features and global evatuation, highlighting

Mean and Std. Deviation in tables 1-5.

4.2.1 Perceptual Sensing Features

This subsection presents the evaluation results of the perceptual sensing features concerning Visual and Tactile Sensing, Auditory Sensing: Spatial Audio, and Auditory Sensing: Cyclic Doppler Effect categories, summarized in tables 1-3.

4.2.1.1 Visual and Tactile Sensing

Users were asked to: move around a 360º video by moving the tablet around (T1) and by using the drag interface (T2); and view a 360º video with the wind accessory and identifying the wind direction (T3). Users appreciated the tested features, especially the video navigation by moving the tablet around, which they reported to be a more natural approach when compared to the touch interface; and the wind accessory, which allowed a more realistic sense of speed in video viewing, as the PR results show (T3: PR: 8.9; 8.9), thus confirming RC2. Despite users favouring the "moving the tablet around" feature for the sense of immersion, the consensus among users was that there are situations where the drag interface can be more suitable, for its flexibility, answering RQ1, and reinforcing the idea that both interfaces are needed and complement each other.

Table 1: USE evaluation of Windy SS (scale: 1-5).

Features in Task:	Usefulness		Satisfaction		Ease of Use	
	M	σ	M	σ	M	σ
Pan Around the 360º video:						
T1 move around	4.8	0.3	4.7	0.3	4.8	0.8
T2 drag interface	4.8	0.4	4.5	0.5	4.8	0.4
Wind Accessory:						
T3 wind	4.3	0.6	4.8	0.7	4.9	0.3
Spatial Audio:						
T4 stereo	4.8	0.7	4.5	0.7	4.8	0.5
T5 3D	4.8	1.1	4.7	1	4.8	0.5
Cyclic Doppler Effect:						
T6 wind sound	2.7	0.6	2.5	0.5	4.9	0.5
T7 low freq sound	4.4	0.9	4.6	0.8	4.9	0.3
T8 high movement	4.8	0.4	4.7	0.4	4.9	0.4
T9 med movement	4.4	0.4	4.5	0.4	4.9	0.5
T10 low movement	2.8	0.4	3.1	0.4	4.9	0.4
T11 no Doppler	4.7	0.5	4.5	0.6	4.9	0.5
T12 custom Doppler	4.7	0.6	4.7	0.6	4.9	0.3
Overall	**4.3**	**0.6**	**4.3**	**0.6**	**4.9**	**0.5**

Table 2: SAM (Pleasure, Arousal, Dominance) evaluation of Windy SS (scale: 1-9).

Features in Task:	SAM					
	Pleasure		Arousal		Dominance	
	M	σ	M	σ	M	σ
T1 move around	8.2	0.5	8.7	0.8	8.0	0.7
T2 drag interface	7.9	0.6	7.4	0.8	8.8	0.5
T3 wind	8.3	0.8	8.6	0.9	8.6	0.8
T4 stereo	7.9	0.7	8.1	0.7	8.8	0.7
T5 3D	8.3	1.2	8.5	1.3	8.6	1
T6 wind sound	3.9	1	6.1	0.8	7.4	1.2
T7 low freq sound	8.2	0.9	7.9	0.8	8.4	1
T8 high movement	8.7	1	8.7	0.5	8.5	0
T9 med movement	8	1.7	7.8	0.8	8.1	0.6
T10 low movement	4.4	0.7	5.2	0.5	6.2	0.4
T11 no Doppler	7.6	1	7.7	1.3	8.6	1.2
T12 custom Doppler	8.8	1.2	8.6	0.8	8.7	0.9
Overall	**7.5**	**0.9**	**7.8**	**0.8**	**8.2**	**0.8**

Table 3: PR (Presence, Realism) evaluation of Windy SS (scale: 1-9).

Features in Task:	PR			
	Presence		Realism	
	M	σ	M	σ
T1 move around	8.8	0.8	8.8	0.7
T2 drag interface	8.3	0.9	8.4	0.8
T3 wind	8.9	0.5	8.9	0.5
T4 stereo	8.5	0.9	8.4	0.8
T5 3D	8.7	1	8.8	0.9
T6 wind sound	4	1.6	3.7	1.4
T7 low freq sound	8.2	0.6	8.2	0.5
T8 high movement	8.8	0.5	8.9	0.6
T9 med movement	8	1.8	7.9	1.5
T10 low movement	3.9	0.6	3.3	0.8
T11 no Doppler	8.3	1.1	8.4	1
T12 custom Doppler	8.9	1	9	0.7
Overall	**7.8**	**0.9**	**7.7**	**0.9**

4.2.1.2 Auditory Sensing: Spatial Audio

In order to test the Spatial Audio feature, users were asked to: view a 360º video with headphones being that the video's sound was standard stereo sound (T4); and view the same video with the 3D sound capability (T5). Regarding T5, users were asked to vary the virtual "distance" of the simulated speakers to the users' head through the manipulation of a seekbar (the seekbar value was relative to the radius of the virtual circle associated with the distance of the sound sources to the user's head), and find the optimal virtual distance between users' head and the sound sources. In respect to RQ3, users stated that this feature provided them with a better sense of orientation, and they preferred the 3D sound version, with the restriction that the the speakers must be located between 1 and 3 meters of the users' head (in the virtual sound space).

4.2.1.3 Auditory Sensing: Cyclic Doppler Effect

In order to test the Cyclic Doppler Effect feature, users were asked to view videos with the Doppler Effect feature activated, being that sounds with different characteristics were used in each video to create the Doppler Effect. In the first video, a wind sound was used (T6). In the second video, the created and recorded low frequency sound waves (described in section 3.4) were used (T7). Users needed to choose their preferred sound being that, in T7 users were asked to vary the Doppler Effect sound by choosing their preferred of the four created sounds from a radio button group. Also users were asked to vary the sound volume through a seekbar and identify the optimal value for the Cyclic Doppler Effect feature.

Next, users viewed three videos with the Doppler Effect feature activated and were asked to state in which of them they liked the Doppler Effect the most. The three videos presented situations where the degree of movement was: 1) high (T8); 2) medium (T9); and 3) little/no movement (T10). The order in which the videos were viewed was randomized for each user. The order in which the videos were viewed was randomized for each of the users. Lastly, taking into consideration all the user's preferences regarding the Doppler Effect feature (in T6-T10), users viewed a video twice: once without (T11) and once with (T12) the "custom" Doppler Effect feature, and were asked to state whether they felt the Doppler Effect feature increased the movement sensation. Answering RQ4, almost all users preferred the low frequency sound, as they stated it was much less obstrusive than the wind sound.

With regards to the volume, users tended to set the Doppler Effect volume level between 7% and 18% of the main video sound volume. In respect to RQ5, according to the users' feedback, the more movement there is in a video, the more satisfying the Doppler Effect becomes. Users referred to T8 (high degree of movement) to express a situation where they particularly enjoyed the Doppler Effect (T8: USE: 4.8; 4.7; 4.9; PR: 8.8; 8.9). On the other side, in videos with no movement, the viewing experience is better without this feature, as the USE and PR results show (T10: USE: 2.8; 3.1; 4.9; PR: 3.9; 3.3). This result confirms the need for the filtering feature that establishes the minimum movement amount for the Doppler Effect feature to be activated. When viewing the video with all the preferences adjusted, all users declared that the Doppler Effect feature increased the movement sensation, which is supported by the SAM and PR values (T12: SAM: 8.8; 8.6; 8.7; PR: 8.8; 9).

4.2.2 Global Presence and Immersion Evaluation

The Immersive Tendencies Questionnaire revealed a slightly above average score, whereas the Presence Questionnaire showed a high degree of self-reported presence in the application (tables 4 and 5). As Presence is a human reaction to immersion, the PQ score reveals the global immersiveness of the tested features. Moreover, the improvement from the ITQ to the PQ reveals that WSS surpassed user's immersive expectations of the system.

Table 4: Immersive tendencies questionnaire.

Tendency to: (scale: 1-7)	**M**	**σ**
maintain focus on current activities	4.2	1.3
become involved activities	4.3	1.7
view videos	5.0	1.2

Table 5: Presence questionnaire.

Major factor category (scale: 1-7)	**M**	**σ**
Control factors	6.1	0.8
Sensory factors	6.7	0.7
Distraction factors	5.3	1.0
Realism factors	6.3	0.8
Involvement/Control	6.2	1.1
Natural	6.1	0.9
Interface quality	6.2	0.8

As a final appreciation, users found Windy SS innovative, very fun, useful, easy to use and fairly easy to understand, preferring the wind accessory and the move around video interface.

5 CONCLUSIONS AND PERSPECTIVES

We presented the motivation and challenges, and described the design and user evaluation, of the sensing features of WSS towards increased immersive experiences. It is a mobile application that uses a wind accessory, 3D audio and a Doppler Effect simulation for the visualization and sensing of georeferenced 360º hypervideos. The user evaluation showed that the designed features increase the sense of presence and immersion, and that users appreciated them, finding all of them very useful, satisfactory and easy to use. Users showed great interest in the wind device, indicating this is a very effective way to improve the realism of the environment. Using 3D audio is a clear advantage and it is an approach that does not require a new infrastructure, as any pair of stereo headphones will suffice and the sensors commonly available in mobile devices allow detecting movement. According to our tests, the sound sources distance to the user' head in the virtual sound space, which should be a value comprised between 1 and 3 meters. The use of the Doppler Effect simulation, when carefully manipulated as it was described, can increase the users' movement sensation, especially in videos where there is a high degree of movement.

Next steps include: refining and extending our current solutions, exploring further settings for higher levels of immersion, like the CAVE and wide screens, and other modalities to increase users' engagement; exploring 3D audio capture to further increase realism in the spatial audio feature; considering georeferencing, ambient computing and augmented reality scenarios, e.g. as access points to videos shot in the same place at a different time, to compare them 'overlaid', or access videos with similar speed as the current speed experienced by the user (e.g. while traveling on a train). This concept can be extended for further filters (e.g. access videos in same time of day, or with similar weather conditions), relying on reality aid in finding videos and feeling more immersed in the virtual video being experienced.

ACKNOWLEDGEMENTS

This work is partially supported by FCT through LASIGE Multiannual Funding and the ImTV research project (UTA-Est/ MAI/0010/2009).

REFERENCES

Amselem D. (1995). A Window on Shared Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 4(2):130–145.

Bleumers, L., Broeck, W., Lievens, B., Pierson, J. (2012) Seeing the Bigger Picture: A User Perspective on 360° TV. In *Proc. of EuroiTV'12*, 115-124

Cardin, S., Thalmann, D., and Vexo, F. (2007) Head Mounted Wind. In *Proc. of Computer Animation and Social Agents*. Hasselt, Belgium, Jun 11-13. 101-108.

Courtois, C., D'heer, E. (2012) Second screen applications and tablet users: constellation, awareness, experience, and interest. *EuroiTV'12*, 153-156.

Dobler, D., and Stampfl, P. (2004) Enhancing Three-dimensional Vision with Three-dimensional Sound. In *ACM Siggraph'04* Course Notes.

Douglas, Y., and Hargadon, A. (2000) The Pleasure Principle: Immersion, Engagement, Flow. *ACM Hypertext'00*, 153-160.

Ekman, P. (1992) Are there basic emotions? *Psychological Review*, 99(3):550-553.

Lehmann, A., Geiger, C., Wöldecke, B. and Stöcklein, J. (2009) Poster: Design and Evaluation of 3D Content with Wind Output. *In Proc. of 3DUI*.

Mendes, M. (2010) RTiVISS | Real-Time Video Interactive Systems for Sustainability. *Artech*, 29-38.

Merer, A., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2013) Perceptual Characterization of Motion Evoked by Sounds for Synthesis Control Purposes. In *Trans. on Applied Perception*, 10(1),1-23.

Moon, T., Kim, G. J., and Vexo, F. (2004) Design and evaluation of a wind display for virtual reality. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*.

Mourão, A., Borges, P., Correia, N., Magalhães, J. (2013) Facial Expression Recognition by Sparse Reconstruction with Robust Features. In Image Analysis and Recognition. *ICIAR'2013*, LNCS, vol 7950, pp.107- 115.

Namezi, M., and Gromala, D. (2012) Sound Design: A Procedural Communication Model for VE. In *Proc. of 7th Audio Mostly Conference*, Greece, 16-23.

Neng, L., and Chambel, T. (2010) Get Around 360º Hypervideo. In *Proc. of MindTrek*, Tampere, Finland. 119-122.

Noronha, G., Álvares, C., and Chambel, T. (2012) Sight Surfers: 360º Videos and Maps Navigation. In *Proc.of Geo MM'12 ACM Multimedia'12*. Nara, Japan. 19-22.

Oliveira, E., Martins, P., and Chambel, T. (2011) iFelt: Accessing Movies Through Our Emotions. In *Proc. of EuroITV'2011*, Lisbon, Portugal, Jun 29-Jul 1. 105-114.

Ramalho, J. and Chambel, T. (2013) Windy Sight Surfers: sensing and awareness of 360º immersive videos on the move. In *Proc. of EuroITV 2013*, Como, Italy. 107-116.

Ramalho, J. and Chambel, T. (2013) Immersive 360° mobile video with an emotional perspective. In *Proc. of the 2013 ACM international workshop on Immersive media experiences at ACM Multimedia 2013*, Barcelona, Spain. 35-40.

Slater, M., Lotto, B., Arnold, M.M., Sanchez-Vives, M.M. (2009) How we experience immersive virtual environments: the concept of presence and its

measurement, *Anuario de Psicología*, 40(2), Fac. Psicologia, Univ. Barcelona, 193-210.
Slater, M. (1999) Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire, Presence 8(5), Oct. 560-565.
Slater, M., & Wilbur, S. (1997) A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence*, 6, 603-616.
Visch, T., Tan, S., and Molenaar, D. (2010) The emotional and cognitive effect of immersion in film viewing. *Cognition & Emotion*, 24: 8, pp.1439-1445.
Wirth, W., Hartmann, T., Bocking, S., Vorderer, P.,Klimmt, Ch., Schramm, H., et al. (2007) A process model of the formation of spatial presence experiences. *Media Psychology*, 9, 493-525.
Witmer, B., Singer, M. J. (1998) Measuring Presence in Virtual, Environments: A Presence Questionnaire. *Presence*, 7(3), Jun. (1998), MIT, 225–240.