IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 May 2023, accepted 3 June 2023, date of publication 14 June 2023, date of current version 21 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3285709

RESEARCH ARTICLE

Photorealistic True-Dimensional Visualization of Remote Panoramic Views for VR Headsets

SALVATORE LIVATINO[1], ALESSIO REGALBUTO[2], GIUSEPPE MORANA[1], GIOVANNI SIGNORELLO[3], GIOVANNI GALLO[4], ALESSANDRO TORRISI[5], GIANLUCA PADULA[6], KATARZYNA PELC[7], ALESSIO MALIZIA[8,9], AND GIOVANNI MARIA FARINELLA[4]

[1]School of Physics, Engineering and Computer Science (SPECS), University of Hertfordshire, AL10 9AB Hatfield, U.K.
[2]Magnopus, AL3 4EL St. Albans, U.K.
[3]Department of Agriculture, Food and Environment, University of Catania, 95125 Catania, Italy
[4]Department of Mathematics and Computer Science, University of Catania, 95125 Catania, Italy
[5]Fable Data, SE1 0AT London, U.K.
[6]Academic Laboratory of Movement and Human Physical Performance, Dynamolab, Medical University of Łódź, 90-419 Łódź, Poland
[7]Business School, University of Bedfordshire, LU1 3JU Luton, U.K.
[8]Department of Computer Science, University of Pisa, 56127 Pisa, Italy
[9]Faculty of Logistics, Molde University College, 6410 Molde, Norway

Corresponding author: Salvatore Livatino (s.livatino@herts.ac.uk)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Hertfordshire Health Science Engineering and Technology ECDA (Ethics Committee with Delegated Authority).

ABSTRACT Virtual Reality headsets have evolved to include unprecedented display quality. Meantime, they have become light-weight, wireless and low-cost, which has opened to new applications and a much wider audience. Photo-based omnidirectional imaging has also developed, becoming directly exploitable for VR, with their combination proven suitable for: remote visits and realistic scene reconstruction, operator's training and control panels, surveillance and e-tourism. There is however a limited amount of scientific work assessing VR experience and user's performance in photo-based environment representations. This paper focuses on assessing the effect of photographic realism in VR when observing real places through a VR headset, for two different pixel-densities of the display, environment types and familiarity levels. Our comparison relies on the observation of static three-dimensional and omnidirectional photorealistic views of environments. The aim is to gain an insight about how photographic texture can affect perceived realness, sense of presence and provoked emotions, as well as perception of image-lighting and actual space dimension (true-dimension). Two user studies are conducted based on subjective rating and measurements given by users to a number of display and human factors. The display pixel-density affected the perceived image-lighting and prevailed over better lighting specs. The environment illumination and distance to objects generally played a stronger role than display. The environment affected the perceived image-lighting, spatial presence, depth impression and specific emotions. Distances to a set of objects were generally accurately estimated. Place familiarity enhanced perceived realism and presence. They confirmed some previous studies, but also introduced new elements.

INDEX TERMS Extended reality, head mounted displays, photorealism, visualization.

I. INTRODUCTION

Telepresence is a fascinating concept that has stimulated human fantasy and inspired projects from a long time [1]. It often refers to tele-exploration and tele-operation and applies to many fields [2], [3]. There is today a formidable opportunity to take experience of remote places to an entirely new level thanks to latest progress in camera and display systems, and to the newly available Virtual Reality (VR) headsets (also known as Head Mounted Displays - HMDs),

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea F. Abate.

FIGURE 1. Conceptual illustration of photorealistic panoramic viewing using Virtual Reality headsets.

providing great mobility and visual performance at much lower cost. This, together with the new needs for telework and teleoperation, has brought a renewed interest towards more effective ways to enhance telepresence using VR technologies.

The new VR headsets feature wide field-of-view (FOV), high image quality, a great sense of isolation from the surrounding space, and can naturally benefit from using new 360 deg. camera views. VR headsets can be used to provide immersive observation of remote objects and locations, including cities, landscapes, houses, and merchandise. They can complement or replace the actual viewing experience, being used as well for pre-viewing and re-viewing scenes at any time. The figure 1 shows a conceptual illustration of VR-headset photorealistic viewing.

We know from the VR literature that the more immersive the telepresence experience is, the more effective a task performance is expected to be [5], e.g. in cases of psychomotor tasks [6], immersive analytics [7] and decision making [8]. However, it is not always straightforward how to maximize telepresence. It seems to be relevantly about the provided sense of presence, which in turn calls for a number of elements to contribute effectively to the provided experience.

A relevant element that plays a fundamental role in many applications, and towards presence too, is realism [9], [10]. This indicates similarity to everyday life experience. For VR systems, visual realism is of major relevance, referring to the natural viewing experience. A vivid, undistorted and correctly proportioned visual impression will enhance realism, but this objective can be hard to achieve as it calls for effective system and application design. Among main system elements that affect visual realism: image acquisition, processing algorithms, and visualization medium. Human factors also play a role and can be affected by those system elements as well as user previous experiences [11]. We find limited literature assessing immersive observations of photographed environments and providing guidelines. A main reason being that extensive omnidirectional high-resolution capture is a very recent achievement.

The aim of this work is to investigate the contribution provided by photographic images when observing real places through a VR headset, in terms of perceived image-lighting and human sensations (visual realism, presence, emotions, depth perception). We look at the role played by three specific elements: display, environment, and familiarity. For display we focus on the role played by pixel-density on perceived image lighting by experimenting with two different displays. For environment we focus on spatial perception and on the effects of illumination by experimenting with two different environments. For familiarity we focus on the role played by previous knowledge by experimenting with two groups of people.

Section II briefly introduces the relevant VR technologies and the state of the art. Section III presents the proposed investigation and measures. Section IV describes the design of our user studies including materials and procedures. The outcome of our experiments is presented and discussed in sections V, VI and VII, according to the research questions. Conclusions are drawn in section VIII.

II. VR AND VISUAL REALISM

A. VR HEADSETS AND 360 CAMERAS

There has been a great development in VR headsets since 2012 with newer systems featuring wider displays, higher resolution, lower cost, and wireless connection. Parallel to latest VR headsets development there has been that of 360 cameras. The reason being VR headsets naturally fit with omnidirectional viewing, through head-rotation. The most interesting type of such camera systems (and also the most expensive) are now capable of acquiring stereoscopic-3D (S3D) images, e.g. [12]. S3D is a viewing capability VR headsets naturally support through separate displays [13]. A 360-camera view can also be generated with standard (not 360) 2D/3D cameras, but this calls for acquisition and processing of several photographs [4], [15], [14]. Latest developments in 360/3D cameras have made easier the capture of compelling photorealistic views for VR use, with great potential towards providing a high-level of visual realism. This has made easier the adoption of photographs and video to represent VR environments, opening to a wide use of photorealistic VR environment representations. The use of photographic images in VR is in contrast with the so far mainstream use of computer graphics to generate VR images (through the processing of modelling and rendering). This has resulted in having most literature works assessing visual realism over synthetic images.

B. VISUAL REALISM IN VR

VR headsets have demonstrated being able to provide and sustain immersion and VR experience, e.g. in terms of presence and emotions, capitalizing on the portrayed image quality [16] and viewing setup [13], [17]. Literature works have focused on: (a) elements contributing to high-fidelity image reproduction and visual realism, e.g. the role played by display resolution [18] and illumination [19], and S3D viewing [20], [21]; (b) the effects of image-quality and visual realism in terms of presence, emotions, and depth perception [19] [22]. Interestingly, visual realism is also being investigated in relation to data visualization, e.g. geospatial information [23] and immersive analytics [7]. Experiments in literature are nowadays mainly performed on HMD systems. Alternatively, we find the use of 3D desktops, 3DTVs and wall screens. Visual realism is mainly assessed on synthetic images, while only few works use photographic texture.

1) CAMERA, DISPLAY, IMAGE QUALITY, REALISM

Visual realism is investigated by Janssen et al. [24] and Ijsselsteijn et al. [20] in terms of image quality, which is claimed to be determined by usefulness and naturalness, a concept only partially shared by Kuijsters et al. [19], who conclude that naturalness may not contribute to image quality [22], [25]. Ferwerda [26] and Hagen [27] analyze the role of visual and photographic realism towards achieving a more realistic response.

Several works relate realism to different image elements. Some authors investigate S3D camera parameters most related to viewing setup. Ijsselsteijn et al. [20] focus on camera baseline, convergence and focal-length / FOV, and discusses their contribution to naturalness; whereas Banks et al. [28] focus on the appropriate matching between the camera's focal-length and image sensor-size, with display's viewing distance and screen size, and they discuss their contribution to veridical visual perception (intended as faithful representation of dimensions).

Lens and display geometry are also regarded relevant factors, as optical flow changes when display peripheral geometry is deformed by the optics, causing viewing and perspective distortions [21].

Some authors investigate camera parameters most related to *image lighting*. Kuijsters et al. [19] focus on color, contrast and texture, and discuss their contribution both to image quality and depth perception. Tiiro [29] and Pardo et al. [30] focus on color, shadows, texture and definition, to bring realism into VR scenes. Bowman et al. [18] address resolution. Slater et al. [31], [32], Palad [33], Gu et al. [34], address videness and sharpness.

Many authors have more recently focused on automatic objective image quality assessments with works exploiting deep learning techniques applied to images and videos, e.g. [35]. This approach, typically aimed at better streaming, can run automatically and does not require user studies. It needs parameters training, though.

2) VISUAL REALISM AND PRESENCE

Visual realism often refers to different degrees of immersion and may contribute to a higher sense of presence. Literature works have discussed visual realism contribution to presence. Witmer et al. [9] discuss its role, whereas Shubert et al. [36] consider experienced realism a key element of the Igroup Presence Questionnaire (IPQ). Interrante et al. [37] consider presence the main reason for accurate distance estimation in their photorealistic virtual environment, whereas Ijsselsteijn et al. [38] discuss presence in relation stereoscopic-3D and camera parameters. More recently, Ling et al. [39] discuss 2D and 3D viewing, FOV, center-of-projection and vantage points, and their effect over perceived presence; whereas Hvass et al. [40] discuss the effect of geometrical realism over presence.

3) VISUAL REALISM AND EMOTIONS

Presence is also connected to visual realism through its effect over provoked emotions and sensations [41]. It seems that the effect of immersion over emotions may depend on the type of the emotion [16], with some authors relating immersion to arousing emotions such as fear and anxiety [17], and other demonstrating its positive effect over non-arousing emotions such as joy, relaxation, sadness and satisfaction [42]. The work of Banos et al. [42] also discuss the use of colors, reflections and natural sceneries to assess sadness, anxiety, joy, relaxation, and satisfactions, in realistically portrayed virtual environments. Seagull et al. [43], on the other hand, focus on assessing the perceived physical fidelity through quality of experience, satisfaction and enjoyability. Finally, Hakkinen et al. [44] underline how S3D image quality can affect emotions.

4) VISUAL REALISM AND DEPTH PERCEPTION

Vision is the main human sensory modality, the sensorial input we believe the most, and with a demonstrated ability towards perceiving object location in 3D-VR space [44], [45], [46]. Visual realism has therefore also been studied in terms of its contribution to distance perception and estimation, which are claimed to be key elements in realistic VR viewing [47], [48]. Image quality and its relation to depth perception have gathered wide interest in the last decade, thanks also to the development and marketing of 3DTVs. There has been a research focus on picture and depth quality and their effects to naturalness and presence. The work of Kuijsters et al. [19] discusses contribution of S3D to naturalness and depth perception. The work of Li et al. [49] focuses on the role of the FOV towards distance judgments, whereas the work of Li et al. [50] focuses on the role of human peripheral vision. Literature works have also focused on depth perception and distance estimation, and established connection between stereoscopic-3D viewing and user's performance in a number of applications of VR for telepresence and teleoperation [38], [51], [52].

5) VISUAL REALISM AND VIRTUAL VS REAL VIEWING

Some literature works focus on comparing real viewing to realistic visual replicas [53], [54], [55] and high-fidelity and faithfully-sized replicas [37], [56], and non-photorealistic replicas [57]. The work of Interrante et al. [37] discusses the effect of photorealism and realistically portrayed dimensions to distance estimation. It associates users' performance in estimating distances correctly to the perceived sense of presence and contribution of *visual calibration* (occurring when the correspondent real environments are also observed).

6) VISUAL REALISM AND SENSE OF PLACE

The user's perceived sense of place during observation is reinforced by prior knowledge of the shown environment. This has been confirmed in case of VR environments [11] and 3D geo-visualization [58]. There is a relation between emotion and sense of presence, which can alter place perception [59]. Julin et al. [60] suggest the sense of place offers an interesting point of view for assessing effectiveness of photorealistic 3D visualization. Virtanen et al. [61] find that 3D geo-visualization helps comprehension of spatial relations.

III. PROPOSED INVESTIGATION AND MEASURES

In his pioneer work about mixed reality taxonomy, P. Milgram et al. [62] defined reality as what can be observed through either *direct viewing* (defined as "the viewing experience of a live scene") or *indirect viewing* ("a live scene observed through a medium"). We are in this work addressing those two aspects of Milgram et al.'s *reality* because we wish to understand what makes VR-headset's *indirect viewing* closer to *direct-viewing*. This means in other words, to have images that resemble observations with the naked eye when viewed through the VR medium. We call it: realistic viewing.

We understand the above leads to a complex and extended investigation due to the many aspects and parameters a realistic viewing involves. We delimitate then our assessment by focusing on some of the elements indicated by the literature as relevant to a realistic visual experience. Literature works show that scene illumination and viewing setups affect visual realism, with consequence on presence, emotion, and depth perception. We decide then to investigate the effect of some display and environment elements to realistic viewing. We focus on: display pixel-density, and environment illumination and distance to objects. We add to our investigation a specific study looking at the role of previous knowledge, based on what we have learnt on sense of place from literature works.

A. RESEARCH QUESTIONS

- **Display.***How display pixel-density affects realistic viewing experience?*

The literature has shown that display quality is connected to its lighting characteristics (color, contrast, etc.), and these affect some human factors (realisms, presence, etc.). We focus on the effect of display pixel-density towards realistic viewing.

- **Environment.***How the perceived sense of realism varies for different types of environments?*

The literature has shown that the portrayed environment characteristics affect some human factors. We focus on the effect of environment illumination and spatial perception.

- **Familiarity.***How place familiarity affects realistic viewing experience?*

The literature has shown that previous knowledge may reinforce or alter perceived presence. We focus on its effect to visual realism.

We aim to answer the above research questions (representing independent variables) by asking users to judge on a number of factors (dependent variables) that can potentially affect them. Subjective ratings and measurements represent trustable indications for immersive VR systems due to the high user involvement and observable effects those systems provide [66]. The ratings are related to display perceived lighting and human sensations in terms of visual realism, presence, emotions, and depth impression. The measurements are related to the perceived distance to objects.

Our study is unprecedent and merely explorative. We hardly find previous research on contribution of visual realism when observing omnidirectional three-dimensional photographs of real places through a VR headset. We hope our study can raise awareness on the effect that photo-based VR representations have towards visual realism.

B. DISPLAY FACTORS

The difficulty in isolating a specific system element is typical when assessing VR systems. This is even more difficult nowadays as we predominately use off-the-shelf systems. We are aware this represents a limitation in our study but believe we can still provide meaningful results as seen in many literatures works. Bowman et al. [18] discuss this issue for image-resolution in support of multiple-component assessments.

We want to assess display quality by interrogating users on a number of *image lighting* characteristics. We choose the below listed elements, and for each element we ask users to rate the element's perceived *relevance to realistic viewing*.

- *Pixel-Density*. Overall pixels' number divided by screen-size. It is related to image resolution and provides an indication of the perceived degree of detail. It affects image illumination too.
- *Lightness*. Overall perceived light intensity. It typically depends on display luminance (amount of light radiance) and the set display brightness.
- *Color*. Combination of light's hue and intensity. It represents the wide variety of color shades.
- *Contrast*. Difference in lightness between pixels. It typically indicates the overall tonal range, which distinguishes between brightness and darkness. The perceived contrast in photographs depends on the location of where shadows and highlights occur [33].
- *Vividness*. Clarity, richness and liveness of the image. It typically represents the contrast of image mid-tones (leaving shadows and highlights unchanged) enhancing the appearance of overall details.
- *Sharpness*. Distinctiveness among pixels. It indicates how well pixel borders merge together and therefore the perceived focus (level of detail).

- *Definition*. Absence of blurs and pixelation. It typically indicates clarity of all represented objects.

C. HUMAN FACTORS

Human Factors are relevant because of their relation to visual impression and to display factors. We interrogate users on the perceived: visual realism, presence, emotions and depth impression. We also ask users to estimate distance to objects to assess its potential role in realistic viewing experience.

1) VISUAL REALISM

There is no standard approach to determine visual realism in VR applications [64], [65], but few inspiring examples mostly related to specific applications. A number of authors propose the use of image-quality metrics and subjective rating to assess visual realism, image quality and naturalness. Brackney et al. [66] indicate interaction, control and motion, as relevant elements that should be part of a realism questionnaire and proposes the use of the INASCL simulator [67], which contemplates 11 criteria including fidelity, defined as true-to-life experience. It divides realism as physical, conceptual and psychological, with all those aspects to contribute to engagement and to be included when assessing realism of VR simulations. Wilson [63] proposes the Realism Assessment Questionnaire, partially inspired by Hill [68], and applies it to colonoscopy studies.

We propose to interrogate users on four questions, taken from the *experienced realism* subscale of the Igroup Presence Questionnaire (IPQ), also referred as *realness* [36], [69]. We add then two questions related to photorealism. Questions are summarized below:

- *Realness. (a) Similarity to real world; (b) experience similar to real world; (c) similarity to imagined world; (d) excessive realism* [69].
- *Photorealism.* Overall level of photorealism.
- *Similarity to photo/video.* Experience similar to seeing a photo or video.

2) PRESENCE

Different methods have been proposed to assess the sense of presence, which are typically based on subjective ratings given through validated questionnaire. Among the most popular, Slater, Usoh and Steed [31], Witmer and Singer [9] and the above mentioned IPQ [36]. We propose to rely on the use of the IPQ. Therefore, our questions address the followings:

- *Overall Presence.* (P1) Sense of being there.
- *Spatial Presence*. (SP1) Sense of surrounding reality; (SP2) perceiving pictures only; (SP3) feeling present in virtual space; (SP4) sense of acting there; (SP5) feeling present in virtual environment.
- *Involvement*. (INV1) real world awareness; (INV2) real world unawareness; (INV3) attention to surrounding reality; (INV4) attention to VR world.

3) EMOTIONS

Authors have typically investigated this aspect through questionnaires targeting specific emotions and visual elements expected to elicit emotions, e.g. image characteristic such as light intensity, shadows and colors. We propose a number of potentially relevant emotions (selected after a pilot assessment) and ask users about their *current sensation towards the emotion*. At the end of the entire evaluation, we also ask users about their: *sensation of being back to reality*. The selected emotions are listed below.

- *Emotions*. Happiness; Enjoyment; Relaxation; Scariness; Sadness; Anxiety; Anger; Surprise; and Back to reality sensation.

4) DEPTH IMPRESSION

As seen in previous works, we ask users generic questions about the delivered sense of three-dimensional appearance. We also ask for users' opinion on contribution of depth impression towards other factors.

- *Depth Impression*. (a) Overall depth impression (delivered sense of three-dimensional appearance); (b) Speed to get 3D impression.
- *Depth Contribution*. (c) Depth impression contribution to realism; (d) Depth impression contribution to emotions; (e) LSC to 3D (lights, shadows and colors contributions to depth impression).

5) DEPTH ESTIMATION

Distance estimation and perceived depth have been measured in different works through quantified judgements reported by users [21], [54]. The methods typically rely on either interactive procedures or motionless observations.

In case of interactive procedures, distance is estimated by asking users to perform specific actions, e.g. when investigating FOV and *minification* in VR images. Some works estimate distances through blind walks [37], [47]; others through directed walks [48]; and through triangulated walks techniques [53].

In case of motionless observations, distance is estimated by asking users to perform observations statically. Relevant examples based on image comparison include the works of Hibbard et al. [70], which associates eye-disparity increase and realism of S3D views to depth judgement, and Baek et al. [71], which investigates the role played by different displays and related parameters (FOV, resolution, brightness, S3D, camera distance) to distance estimation.

In our assessment, we are most related to the case of motionless observations, being our observed scenery static with users not allowed to move except for turning their head around. We therefore follow what is a common approach for distance perception in this type of observation, which is based on motionless depth judgements. We ask users to estimate egocentric distance to 6 pre-selected scene elements and relative distance between 5 of them. We then estimate the average relative error as in [37] and depth perception accuracy based on Baek's formulation [71].

For our distance estimation, we consider a range where humans (under direct viewing) can generally benefit from binocular vision (0.3-10 meters) [72]. This is relevant because of its demonstrated positive contribution to realism and presence. The proposed assessment compares users' estimates to ground truth and between the two environments.

- *Distance Estimation*: (a) Egocentric distance estimation to 6 scene-elements; (b) Relative distance estimation between 5 scene-elements.

IV. EXPERIMENT DESIGN

For this study informed consent was obtained.

A. USER STUDY 1: DISPLAY AND ENVIRONMENT

We conducted a within-subject study with 20 test-users. Participants' age ranged between 22 and 53, with an average of 28. The study had the conditions described below.

- *Display*

We chose two displays that mainly differed in pixel-density, aiming to gain insight on how much this element can affect the considered display and human factors. More details about displays and their choice are in section IV-C.

A. *LG display*. It has higher pixel-density.
B. *iPhone display*. It has lower pixel-density.

- *Environment*

We chose two environments that mainly differed in illumination conditions and distance to objects, aiming to gain insight on how much those characteristics can affect the considered display and human factors.

1. *Island*. It portrays a wide range of light-intensities and colors. It has a high range of distances to objects.
2. *Cave*. It portrays a limited range of light-intensities and colors. It has a low range of distances to objects.

We tested all combinations of the above conditions related to display and environment by running pairwise comparisons on displays (A1 vs B1, A2 vs B2, A1+A2 vs B1+B2) and environments (A1 vs A2, B1 vs B2, A1+B1 vs A2+B2).

B. USER STUDY 2: FAMILIARITY

We conducted an in-between subject study with 40 test-users. Participants' age ranged between 25 and 58, with an average of 32. Participant had none to good experience with VR devices and computer games (uniformly distributed). The study had the conditions described below.

- *Familiarity*

We chose two user groups that differed for their prior knowledge of the shown environments, aiming to gain insight on how much familiarity with the observed place can affect visual realism and presence. The study had the conditions described below.

C. *Site-Familiar*. Users that knew well the observed site.
D. *Unfamiliar*. Users that had never been in the observed site.

We tested the two above conditions by running pairwise comparisons on the two environments. We tested on one display only for practical reasons. We had therefore the following combinations: C1 vs D1, C2 vs D2, C1+C2 vs D1+D2.

The site-familiar users were people that worked in the considered environment and visited the place on a weekly or monthly base. They were therefore ideal to assess faithful reproduction of environment objects as they knew them well. We thought their given scores would be particularly meaningful (a sort of "ground truth"). Unfamiliar users could still judge realism based on appearance of well-known objects such as trees and landscapes. We considered of interest to see how site-familiar users' scores would differ from those of unfamiliar users.

C. MATERIALS

1) 3D ACQUISITION AND VISUALIZATION

We focus on delivering a realistic *true-dimensional* visualization. We define it a viewing experience that let users perceive the actual space dimension, objects' size and distance. This is different to what we typically experience when looking at environment photographs or a show on TV. In these cases, wide-angle views are often used to include large space-portions on one single view. The perceived environment and object sizes become then typically magnified, which causes surprise or disappointment when one sees the portrayed place through own eyes (direct viewing).

Our system capitalizes on the use of high-resolution, omnidirectional and three-dimensional images. It relies on acquiring photographs or videos of a remote place and play them back appropriately to maximize realistic impression. The latter is proposed to be achieved from a system concept point of view through providing:

- *High-fidelity images*, i.e. deliver faithful representations through capturing and displaying high-resolution images, using quality lens and lossless processing algorithms.
- *Natural FOV*, i.e. deliver a viewing angle that resembles the one we natural perceive. There is therefore a need to carefully set acquisition FOV to that typical felt in humans and align to it the visualization FOV through a calibration process. The latter involves matching software, headset and human FOVs.
- *S3D vision*, i.e. deliver binocular three-dimensional visualization captured with human-like camera interocular distance (IOD) (as it would happen with direct eye-viewing), visualized with a human-like virtual-camera IOD (in software), and observed with an headset possibly allowing for IOD adjustment according to user's interpupillary distance (IPD).

TABLE 1. Technical specs of our displays (better values in bold).

Display Characteristics	iPhone	LG G-series
Max brightness in Nits (Higher is better)	**559**	379
Black levels in Nits (Lower is better)	**0.3647**	0.4337
Contrast Ratio 100% Brightness (Higher is better)	**1,534.0**	875.0
Avg. White point in K (Closer to 6504K is better)	**6,515**	7,244
Greyscale Accuracy (Lower is better)	**1.9683**	3.6935
Saturation Accuracy (Lower is better)	**1.19.29**	4.7599
Great MacBeth Color Accuracy (Lower is better)	**1.7645**	3.9702
Display Resolution (pixels) / Screen Size (in.)	1334x750/4.7	**2560x1440/5.5**
Pixel Density (ppi) / Pixel Size (mm)	326/0.078	**538/0.047**

- *Viewing by head-rotation*, i.e. deliver panoramic-view observation by rotating the head, similarly to the way we naturally see and discover the world around us.

Literature works have indicated that naturalness and realistic space perception are relevant for presence [24], [27], [36], [37], [38], therefore for our system we made a careful selection of off-the-shelf hardware and also tried to get our choices confirmed by running a number of pre-assessments. We started by having an image capture-visualization system that would feature high-resolution, omnidirectional, and correctly proportioned three-dimensional images [94].

For image acquisition we chose the consumer 3D camera *Fujifilm 3D FinePix W3* because, among those available in the market, it provides focal-length and camera stereo-objectives separation (baseline) closer to the average human inter-ocular distance IOD (6.5 cm). A dense set of images were needed to keep distortion low. Therefore, hundreds high-resolution images of the environment were captured at set locations by rotating the camera of approximately one degree, up to cover a 360-degree view. Single overlapping images merged through a stitching process into one high-resolution panoramic composition using our algorithm based on SIFT-like feature matching [15]. The obtained panoramas were organized after spherical configuration from single viewpoint, with stereo-couples going through the same stitching process to avoid mismatches that can affect 3D viewing comfort.

2) DISPLAY AND VR HEADSET

We opted for smartphone-based VR goggles because it allowed us to test with different displays, while keeping the same remaining headset's characteristics, such as lens and eye-display distance. We chose the *VR Shinecon* glasses. Our headset allowed users to adjust focus and IOD. We chose the *LG G-series* and the *Apple iPhone*(the display characteristics are in Table 1). Both displays were IPS LCD.

The two displays differed in size (the LG was 17% larger). However, the additional display surface was not visible to users because a display portion was hidden by the headset structure. This resulted in users observing the same FOV on both displays (we measured a FOV discrepancy below 5%). The LG display had greater pixel-density compared to the iPhone, whereas the iPhone display excelled in lighting characteristics such as: black levels, contrast, grey scale, saturation and color accuracy. We chose the system with lower pixel-density (iPhone) to have better lighting characteristics to ensure that a possible lower performance was generally not to be attributed to display lighting characteristics. Analogously, a better LG performance was not to be attributed to better display lighting characteristics. We made sure display brightness was set as equal as possible using a LUX meter. Figure 2 shows the VR headset. The Table 1 shows the display specifications.

FIGURE 2. The two environments observed in our experiments. The Lachea island [73] (top and middle left rows) representing richness of light-intensities with wide-color range and higher distance-range. The Monello cave [74] (bottom rows) representing low light-intensity and color ranges, and lower distance-range. The used VR headset is shown in the middle-right row.

3) ENVIRONMENT

We chose two environments with different illumination conditions. One depicted a wealth of light-intensities with wide-color and distance ranges. The other one had opposite characteristics, i.e. low ranges of light-intensities, colors and distances. The reasons for this choice were in the fact that the works in literature show that colors, as well as closer 3D views, are expected to increase visual realism, presence, and spatial perception. Therefore, if the first environment would have induced for example more realism, this should be mainly to be attributed to light intensities and colors. Analogously, a more realistic performance in the second environment would be mainly to be attributed to the closer distances (rather than the nearly monochromatic views).

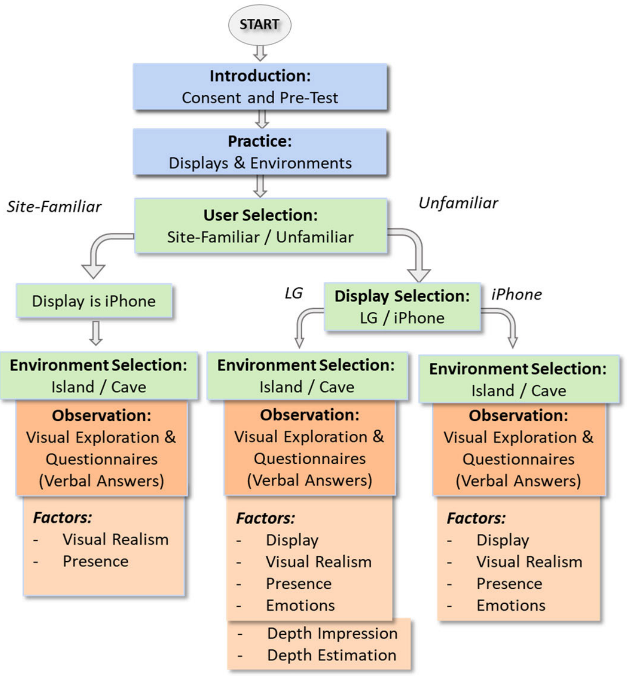


FIGURE 3. Main experimentation steps.

TABLE 2. Display and human factors assessed in the experiments.

Display and Human Factors	Display	Environment	Familiarity
Image lighting	X	X	
Visual Realism	X	X	X
Presence	X	X	X
Emotions	X	X	
Depth Impression		X	
Distance Estimation		X	

The chosen environments were two nature reserves located in Sicily (Italy). They were the *Lachea Island* [73] (further referred as *island*), and the *Monello Cave* [74] (further referred as *cave*). Figure 2 shows images of the two chosen environments [75].

D. TEST PROCEDURE

The test organization and procedure followed literature recommendations [76], [77]. Users executed under the same conditions the tasks described below. The figure 3 shows an illustration of the main experimentation steps. Table 2 shows what factors were assessed in each study.

- ***Introduction***. Initially, we provided participants with an information sheet, consent form and pre-test screening (background questionnaire and a vision test using Snellen chart to check if any sight issue. We carefully explained participants about the meaning of each evaluation factor and the related questions. As image-lighting characteristics have clear visual meaning, we supported questions with illustrations.
- ***Practice***.Participants were asked to familiarize with tasks and system through practice trials. Participants were asked to explore panoramic views while wearing the VR headset. Each participant ran trials on each display and on each environment.
- ***Selection***. The task and environment sequence were assigned to each participant according to a predetermined schedule (based on the Latin square root) to counterbalance task sequence and avoid fatigue and learning effects.
- ***Observation***. Participants correctly, firmly, and comfortably, wore the headset. They were then observing the panoramic view for as long as they needed. During the observation participants were questioned according to the set factors and provided answers verbally.
- ***Estimation***. Participants were asked to turn their head to see a specific portion of the panorama (at a pre-designed position and orientation). They could not walk, but they could turn their head around. Participants were then requested to verbally provide a set number of measurements (in meters), each including an integer and one-digit fractional part. They could observe for how long they needed. The distances asked to be estimated were between 2 and 9.2 meters. The locations of the 6 observation points were on each environment at approximately the same distance and orientation from viewers. The set locations also followed the same distribution pattern. Positions for distance estimates were randomly queried. The range of distances were set similar to that used in the work of Interrante et al. 2008 [36] to allow for data accuracy comparison. Figure 5 shows the set views on each environment and table 12 shows the ground truth distances.
- ***Questionnaires***. We conformed to the traditional approaches in terms of forms and questionnaires [77], [78], [79]. Questionnaires were designed according to the 7-point Likert's scale and scores range was -3 to 3. Forms included sections for reporting open comments through written feedback.
- ***Test Results***. We computed scores' median and standard error. We also measure statistical significance by estimating paired Student's T-test. There is some debate on whether Likert scale variables can be treated as ordinal or categorical, nonetheless some piece of research choose to see it as a continuous variable. In our studies (with 20 users in a within-subject setting and 40 users in between-subject setting) we verified a normal distribution of data, and we then computed the p-value according to the Student's T test [80]. We measured the effect of the different displays, environments, and familiarity on the dependent variables related to display and human factors. We sat that an alpha of 0.05 as p value, determined whether the result is judged statistically significant (tables' red numbers). Alpha values between 0.05 and 0.06 were referred as having a "tendency to significant" (tables' brown numbers). The

TABLE 3. Image lighting: LG display vs iPhone display (*p* values).

	Pixel-Density	Lightness	Color	Contrast	Vividness	Sharpeness	Definition
Isl. & Cave	**0.0482**	0.1068	**0.0594**	0.2824	0.2624	0.2736	0.3688
Island	**0.0306**	**0.0538**	0.0759	0.3658	**0.0248**	0.5000	0.2951
Cave	0.0658	0.1598	**0.0430**	0.1989	0.5000	**0.0473**	0.4425

results are presented and analyzed below according to each research question and related pairwise comparisons. Evaluation measurements are mentioned to perform: *high* for median score between 1.5 to 3; *low* for median score -3 to -1.5; and *medium* for median score -1 through 1.

V. USER STUDY 1 OUTCOME: DISPLAY

This section analyses results focusing on the role played by the visual display. Scores were gathered from users looking at the same environment (either the island or the cave) through different displays (LG and iPhone). Referring to what is described in section IV-A this means: A1 vs B1; A2 vs B2; and A1+A2 vs B1+B2. Figure 4 and tables 3 through 6 show the obtained results.

A. DISPLAY FACTORS

Results: Relevance to realistic viewing scored high on both displays on *image contrast*, *sharpness* and *definition*; and on the LG display only on *image lightness* and *color*. All other image characteristics scored medium. The LG display performed significantly better than iPhone the perceived relevance to realistic viewing provided by: *pixel-density* (scores of both environments combined led to $p = 0.0482$, whereas scores from the island only led to $p = 0.0306$); *color* (scores in the cave led to $p = 0.043$, and scores in both environments led to a tendency to significant, resulting in $p = 0.059$; *vividness*(scores in the island led to $p = 0.0248$); *sharpness* (scores in the cave led to $p = 0.0473$); and *lightness* (scores in the island led to a tendency to significant, resulting in $p = 0.0538$). Table 3 shows all *p*-values.

Analysis: The high scores on a few factors and the medium scores on the remaining one, indicate that users were generally satisfied with both the displays and considered them of high quality. The scores on *pixel-density* indicates the higher LG display's pixel-density was noted by the users, and has therefore potential to play a role towards other factors. As for *color* performing higher on LG display, this is contrary to what one would expect looking at display related specs. The above results and some literature works stating display resolution and pixel-density enabling for more detailed reproduction of color shades and contrast [18], [19], [81], make us think it is the LG higher *pixel-density* causing higher perception of *color* contribution to realistic viewing. The higher pixel-density appears therefore to outweigh the better iPhone's display color accuracy and higher specs in terms of: grayscale, saturation, black and white levels, and contrast ratio. As for *vividness*, *lightness* and *sharpness,* there is clear indication of their contribution to perceived realism.

TABLE 4. Visual Realism: LG display vs iPhone display (*p* values).

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
Isl.&Cave	**0.0514**	0.1537	0.1566	0.3377	0.2601	0.3934
Island	**0.0539**	0.2718	**0.0045**	0.2663	0.4310	0.2869
Cave	**0.0049**	**0.0357**	0.3086	0.4091	0.0892	0.5000

TABLE 5. Overall and Spatial Presence: LG vs iPhone displays (*p* values).

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
Isl.&Cave	0.1270	0.0775	0.0977	0.0959	0.0711	0.1695
Island	0.1872	**0.0211**	0.1435	**0.0458**	0.0978	**0.0193**
Cave	0.0668	0.1340	**0.0520**	0.1459	**0.0444**	0.3196

Overall, the outcome of display factors represents a new finding not specifically addressed in the literature. It shows a dominant role of display pixel-density compared to the display lighting characteristics.

B. HUMAN FACTORS

1) VISUAL REALISM

Results: The scores of *similarity to real world* and *experience similar to real world* were all high, with top values on LG in the cave (median 3 and 2.5 respectively). The first factor recorded a significant better performance of the LG display in the cave ($p = 0.0049$) and a tendency to significantly better performance in the island and on both environments ($p = 0.0539$ and $p = 0.0514$ respectively). The second factor recorded a significant better performance of the LG display in the island ($p = 0.0045$). The scores for *similarity to the imagined world* were medium to high with the LG display performing significantly better than iPhone in the cave ($p = 0.0357$). *Excessive realism* scored low (median -3 to -2.5) and similarly on both displays. *Photorealism* scored high on both displays, whereas scores for *similarity to photo/video* were slightly lower. There were no significant differences on both factors. Table 5 shows the *p*-values.

Analysis: *Realness* scored high-level and consistently across its four factors. Users commented of a remarkable similarity to reality, which was only undermined by the lack of environment dynamics and the limited movement options beside head-rotation. *Photorealism* was highly appreciated and often commented as "very impressive", "effective" and "definitely greater than that provided by a photograph". The latter explains the slightly lower scores of the *similarity to photo/video*. The overall outcome for *visual realism* confirmed effectiveness of our system, which indicates that image acquisition and processing did a good job in maintaining high quality and well combined with visualization. From users' comments and scores, we can state that the vivid and highly photorealistic images do not replace for the lack of dynamics and missing user's actions, but images still induce a very realistic visual experience.

Many users pointed that a correct depth impression was a key supporting element, which is in line with [20] and [24]. The higher visual realism of the display with higher

FIGURE 4. **Outcome for User Study 1 showing median values (with standard error): Image-Lighting (top-row); Visual Realism (2nd row from top); Presence (3rd row from bottom); Emotion (bottom-row). The displays are indicated as LG and iPhone, the environments as Island and Cave. For Display-related pairwise comparisons, pay attention to (A1 vs B1), (A2 vs B2) and (A1+A2 vs B1+B2). For Environment-related comparisons, pay attention to (A1 vs A2), (B1 vs B2) and (A1+B1 vs A2+B2).**

pixel-density (LG) is generally expected, but we have now shown this has taken place also against better display lighting specs.

2) PRESENCE

Results: Both displays scored high the *overall presence* (P1), *surrounding reality* (SP1), *feeling presence in virtual space* (SP3) and *feeling present in virtual environment*(SP5). The SP1, SP3 and SP5 saw a significantly better performance of the LG display over the iPhone in the island only ($p = 0.0211$, $p = 0.0458$ and $p = 0.0193$ respectively). Low were the scores of *perceiving pictures only* (SP2), with the LG display scoring a tendency to significant better performance in the cave ($p = 0.052$); whereas the *sense of acting there* (SP4) scored medium, with the LG display scoring significantly better in the cave ($p = 0.044$).

The *real world awareness* and *unawareness,* and *attention to the VR world*(INV1, INV2 and INV4 respectively) scored high (median 3 for INV1 and INV2); whereas the *attention to surrounding reality* (INV3) scored low (median -3). The scores for both displays were very similar, therefore no significant differences were recorded. Tables 5 and 6 show all *p*-values.

Analysis: The outcome clearly shows both displays gave users a strong sense of *overall* and *spatial presence*. The latter being undermined only by the static nature of the images, which affected the *acting there* (SP4) factor. Fourteen out of twenty users (70%) positively commented about VR

TABLE 6. **Involvement: LG display vs iPhone display (*p* values).**

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
Isl. &Cave	0.3980	0.3417	0.1947	0.4608
Island	0.2960	0.2994	0.2162	0.5000
Cave	0.5000	0.3840	0.1731	0.4217

TABLE 7. **Emotions: LG display vs iPhone display (*p* values).**

	Happiness	Enjoyment	Relaxation	Scariness	Sadness	Anxiety	Anger	Surprise	Back Reality
Isl.&Cave	0.3955	0.2386	0.1289	0.2388	0.1511	0.1331	0.3784	0.1440	0.4368
Island	0.5000	0.3397	0.1546	0.2330	0.1008	**0.0450**	0.3233	0.1490	0.4206
Cave	0.2911	0.1375	0.1032	0.2447	0.2013	0.2211	0.4334	0.1390	0.4530

TABLE 8. **Image lighting: Island vs Cave (*p* values).**

	Pixel-Density	Lightness	Color	Contrast	Vividness	Sharpeness	Definition
LG & iPhone	**0.0058**	**0.0212**	**0.0046**	0.4641	0.1507	**0.0020**	0.0680
LG	**0.0308**	**0.0286**	**0.0567**	0.3113	**0.0401**	**0.0020**	0.2145
iPhone	**0.0361**	0.1484	**0.0146**	0.2608	0.3960	0.1100	0.0941

headsets' suitability to scene exploration. In particular, they found natural with both displays to explore the scenes through head rotation and believed this highly contributed towards increasing spatial presence in the remote place. This is in line with the literature [9], [82], [92]. Twelve users (60%) were also particularly appreciative of the headsets' lightweight, portability and the provided sense of isolation from the surrounding environment, which they commented as also contributing to presence [32].

The significant better performance of *surrounding reality* (SP1) on the LG display only, was credited to the higher pixel-density, which is in line with some literature [39]. The low scores given to the *perceiving pictures only* factor (SP2) indicates effectiveness of presence and of visual realism too, (pictures as such were no longer noted). Despite scoring medium the *sense of acting there*(SP4), users commented the observation as ''very engaging'', ''rich of visible elements'' and ''showing great variation over different viewing directions''. Interestingly the SP4 significant higher scores in the cave environment of the LG display compared to the iPhone, were often commented as caused by the closer distances calling for frequent head rotations.

Regarding *real world awareness* and *unawareness* (INV1 and INV2), scores show that users clearly forgot about their actual premises once they wore the VR headset. HMDs are well-known for their isolation from surrounding space, which in our case appeared further enhanced by strong depth impression (leading to high spatial presence). The two displays scoring high and nearly identical demonstrate good quality. The *attention to surrounding reality* (INV3) and *attention to virtual world*(INV4) scoring opposite (low and high) are what is expected to demonstrate great involvement.

In summary, both displays seemed to be able to convey high presence despite the static content, with the LG display in most cases performing better than the iPhone in bringing a greater surrounding sensation (in the island) and a stronger sense of acting (in the cave).

3) EMOTIONS

Results: Users scored high on both displays *happiness*, *enjoyment* and *surprise*, and on LG display *relaxation*. Users scored medium all the other factors except for *anger*, which scored low. There were no significant differences between the displays, except for *anxiety* in the island where the LG display scored significantly higher ($p = 0.045$). Table 7 shows all p-values.

Analysis: The above outcome indicates emotions are triggered and most of them are positive, which goes perfectly along with what assessed in [43] regarding perceived physical fidelity through quality of experience, and in [44] regarding S3D image quality affecting emotions. As for the significantly greater *anxiety* when observing the island through the LG, this may be due to the greater pixel-density. The outcome for *back to reality* sensation after the VR experience, was varied. To some, it induced excitement, to others disappointment. This seemed therefore to be a subjective aspect.

VI. USER STUDY 1 OUTCOME: ENVIRONMENT

This section analyses results focusing on the role played by the shown environment. Scores were gathered from users looking through the same display (either LG or iPhone), at two different environments (the island and the cave). Referring to what is described in section IV-A this means: A1 vs A2; B1 vs B2; and A1+B1 vs A2+B2. Figures 4 through 7 and tables 8 through 11 show the obtained results.

A. DISPLAY FACTORS

Results: The contribution of *definition* to realistic viewing scored high in both environments. The contribution of *lightness*, *color*, *contrast* and *vividness* to realistic viewing scored high in the island only. The contribution of *sharpness* to realistic viewing scored high in the cave only. As for contribution of *pixel-density* to realistic viewing, it scored high in the island, but only when observed through the LG display.

We can observe that scores in the island were significantly higher compared to the cave: on both displays in terms of contribution to realism given by *pixel-density, lightness* and *color* (respectively $p = 0.0058$, $p = 0.0212$, $p = 0.0046$); and on each single display, (except for *color* on the LG display where only a tendency to significant difference was recorded, $p = 0.0567$). There was no significant difference for contribution of *lightness* on the iPhone. The contribution to realistic viewing of *vividness* scored significantly higher in the island ($p = 0.0401$), whereas the contribution of *sharpness* scored significantly higher in the cave on the LG display and on both displays (equally with $p = 0.002$). Table 8 shows all p-values.

Analysis: If we compare data from the two environments on both displays combined (first row in table 8) with data from each individual display (so either LG or iPhone, respectively second and third rows), we note the significant

TABLE 9. **Visual Realism: Island vs Cave display (*p* values).**

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
LG & iPhone	0.2456	0.2838	0.3638	0.5000	**0.0369**	0.3388
LG	0.2540	0.4196	**0.0400**	0.3233	**0.0397**	0.5000
iPhone	0.3706	0.1546	0.2392	0.3383	0.2594	0.2687

differences occurring when combining data of both displays still occur on each individual display only in case of *pixel-density* and *color*. This shows the environment plays a major role compared to display on those two factors.

According to users' comments, the contribution of *pixel-density* to realistic viewing was very appreciated in the island because objects were looking well defined even at the far distances this environment portrayed. The above indicated contribution of pixel-density is related to represented object distances. As for *color* contribution to perceived realism, it was very appreciated in the island because of the wider color-range and warmer tones. This indicates scene lighting plays a role towards contribution of *color* to perceived realism.

As for contribution of *lightness* and *sharpness*, the significant differences between environments when data from both displays are combined, were confirmed on LG only. The contribution of *vividness* was significantly different on the LG display only. The above outcome tells us the role played by the environment over *lightness*, *sharpness* and *vividness*, is subject to the role played by the display. In other words, we can say that we need the LG display higher pixel-density to trigger a significant difference in the contribution of those factors to perceived realism.

B. HUMAN FACTORS

1) VISUAL REALISM

Results: The two environments showed high median values over all visual realism factors except for the *excessive realism* which scored low. There were no significant differences in all the *realness* factors except for the *similarity to the imagined world,* where the contribution to realism in the island scored significantly higher when observed through the LG display only. The contribution of *Photorealism* scored significantly higher in the cave on both display and on the LG display only ($p = 0.0369$ and $p = 0.0397$ respectively). Table 9 shows all p-values.

Analysis: The overall outcome confirmed the high level of visual realism experienced by the users on both environments. The significant higher values related of the contribution to *photorealism* in the cave is surprising because of the nearly monochromatic appearance, which was supposed to undermine it, and of the portrayed closer distance to objects, expected to make it easier the discovery of deformations [83]. The higher LG's *pixel-density* appear to counterbalance the above aspects by providing sharper images, which as earlier mentioned were well appreciated on closer objects, and the cave has plenty of them.

TABLE 10. **Overall and Spatial Presence: Island vs Cave display (*p* values).**

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
LG & iPhone	0.4034	**0.0273**	0.3763	**0.0273**	**0.0096**	**0.0368**
LG	0.4362	**0.0436**	0.3117	**0.0382**	**0.0429**	**0.0141**
iPhone	0.2840	0.1618	0.5000	0.1567	**0.0478**	0.3287

TABLE 11. **Involvement: Island vs Cave display (*p* values).**

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
LG & iPhone	0.3461	0.2751	**0.0569**	**0.0001**
LG	0.2960	0.3932	0.1289	**0.0016**
iPhone	0.5000	0.2845	0.1337	**0.0074**

2) PRESENCE

Results: Both environments scored high the *sense of being there* (P1), *surrounding reality* (SP1), *feeling presence in virtual space* factor (SP3) and *feeling present in virtual environment*(SP5). The SP1, SP3 and SP5 saw significant higher values in the island when merging scores of both displays ($p = 0.0273$, $p = 0.0273$ and $p = 0.0368$ respectively), and on LG display only ($p = 0.0436$, $p = 0.0382$ and $p = 0.0141$ respectively). Low were the scores of the *perceiving pictures only* (SP2), and medium were the scores of the *sense of acting there* (SP4), with the cave scoring SP4 significantly higher than the island on all display combinations ($p = 0.0096$ on both display data, $p = 0.0429$ on LG display only, and $p = 0.0487$ on iPhone display only). Table 10 shows all p-values.

The *real world awareness* and *unawareness* (INV1 and INV2) scored high (median 3) and it scored similarly in both environments. The *attention to surrounding reality* (INV3) scored low with a tendency to significant higher values in the island on both displays combined ($p = 0.0569$). The *attention to the VR world* (INV4), scored high on both environments with the island performing significantly higher on both displays ($p = 0.0001$), on LG display only ($p = 0.0016$) and on iPhone display only ($p = 0.0074$). Table 11 shows all p-values.

Analysis: Both environments gave users strong *overall presence* and *spatial presence*. The significant better performance in the island compared to cave on SP1, SP3 and SP5, was clearly triggered by the LG display. If we add the considerations made in the previous section on SP1, SP3 and SP5 factors (LG significantly higher performance in the island only), we observe once again it is the combination island – LG display that triggers the difference.

With the support of users' comments, we can conclude that the island's enhanced sense of presence is due partially to the wider environment-views (confirming [39]), and partially to the good lit and wide-color spectrum (confirming [84]); whereas the cave's enhanced *sense of acting* seems mainly geared by the close-distance views calling for head rotation to discover the environment.

Interestingly, both in case of the island and cave, it is the use of the LG display that makes the differences significant.

TABLE 12. **Emotions: Island vs Cave display (*p* values).**

	Happiness	Enjoyment	Relaxation	Scariness	Sadness	Anxiety	Anger	Surprise	Back Reality
LG & iPhone	0.1773	**0.0488**	**0.0000**	0.0651	0.1078	**0.0031**	0.3402	0.1044	0.3491
LG	0.3561	**0.0279**	**0.0000**	0.1576	0.2474	0.0759	0.2906	0.2068	0.3327
iPhone	0.1787	0.3383	**0.0000**	0.1303	0.1410	**0.0071**	0.5000	0.1638	0.4571

Looking at the display specs we could argue it is the higher *pixel-density* triggering the difference.

The effective sense of isolation from the surrounding space the HMD provides is again confirmed on INV1 and INV2, whereas the tendency to significant higher performance on INV3 and INV4 in the island, appears due to the island's richer scenario.

Looking at the overall outcome for environment and display, it can be asserted it is the content playing a major role compared to display. E.g. the island's colorful landscape results more appealing to viewers than the cave monochromatic views, which affects *involvement* and most of the *spatial presence* factors.

3) EMOTIONS

Results: The scores in both environments were high for *happiness*, *enjoyment* and *surprise*. In the island scores were high on *relaxation* only, whereas in the cave scores were high on *scariness* and *anxiety*. The scores were low in both environments on *anger*. The most significant difference is on *relaxation* with island performing higher than cave ($p < 0.0001$ on both and on each display). The island also performed significantly higher on *enjoyment* on both displays and LG only ($p = 0.0488$ and $p = 0.0279$ respectively), whereas the cave performed significantly higher on *anxiety* on both displays ($p = 0.0031$) and on iPhone only ($p = 0.0071$). Table 12 shows all p-values.

Analysis: Emotions were largely triggered in both environments. The main reasons why the island scores significantly higher on *enjoyment* and *relaxation* were identified by users as due to its warm colors. This goes along with [42], addressing contribution of lighting in realistically portrayed virtual environments. As for *sadness*, *scariness* and *anxiety*, we noted higher values in the cave. Fifteen of our users (75%) commented these three types of emotions were elicited by the nearly monochromatic scenes and lack of daylight. It was also noted that *scariness* and *anxiety* appeared further enhanced by the cave's closer distances (e.g. walls and stones) triggering more rapid head-movements than in the island.

We observe a similar general trend between the *overall presence* (P1) and the emotion's *happiness* and *enjoyment* factors, which is supported by many literature works [16], [41], [85], [86], [82].

4) DEPTH IMPRESSION

FIGURE 5. **This island' and cave' set scenarios for users' distance estimation: egocentric distance to the center of while circles; relative distance among the circles (red arrows). The white circles and red arrows are only for reader's comprehension and were not shown to users. Users held a small red cross indicating the locations.**

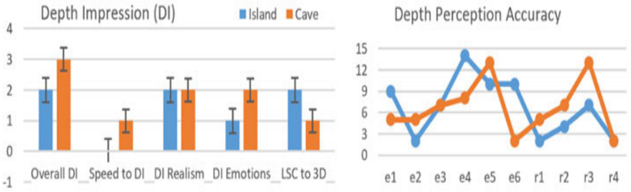


FIGURE 6. **Left: Outcome for: *Depth Impression*; indicating *median* and *standard error* values. Right: Depth Perception Accuracy in % is estimated as in [75]. A 5% error is considered neglectable.**

TABLE 13. **Depth Impression: Island vs Cave display (*p* values).**

	Overall DI	Speed to DI	DI Realism	DI Emotions	LSC to 3D
LG	**0.0377**	0.1394	0.1423	0.0967	**0.0171**

Results: Figure 6 shows the outcome of our Depth Impression questionnaire. The test was only conducted on the LG display for practical issues. The *overall depth impression* scored high on both environments, with the cave performing significantly ($p = 0.0377$).

Medium to high values were the scores of the other factors with no significant differences. We noted the *depth contribution to emotions* was generally higher in the cave. The *LSC to 3D* (benefits from light, shadows and colors towards depth impression) scored high in both environments, with the island performing significantly higher ($p = 0.0171$). Table 13 shows all p-values

Analysis: The *overall depth impression* factor scored high in both the environments, despite their difference in terms of light-intensities and distance-range. This shows our system successfully provides three-dimensional impression. Users commented to get a greater depth impression on the closer portrayed objects (e.g. cave stones and walls), which is in line to what happens with real environment observations [87], [88]. It confirms the relevant contribution of binocular depth-cue on short-medium distances (0.3-10 meters [69]), which occurs more relevantly in the cave because of the more objects at closer distances. Users indicated the strong depth impression contributed to *enjoyment* in both environments (and to perhaps to *scariness* too as commented regarding emotions). This is a fascinating aspect worth further studies.

Concerning the *LSC to 3D* factor, the illumination clearly played a role towards 3D impression, which is in general agreement with literature works on the contribution of light,

TABLE 14. Depth Estimation: Island vs Cave display (*p* values).

	e1		e2	e3	e4	e5	e6	r1	r2	r3	r4
LG	**0.019**		**0.020**	0.293	**0.045**	**0.058**	**0.002**	0.124	0.060	**0.021**	0.075

TABLE 15. Locations' *ground truth* value for *egocentric distance* and *relative distance* estimations (in meters).

e1	e2	e3	e4	e5	e6	r1	r2	r3	r4
2.0 **2.0**	3.5 **3.1**	4.3 **4.0**	5.0 **5.4**	6.1 **6.2**	9.2 **9.2**	2.4 **2.1**	1.2 **1.0**	2.1 **2.0**	4.8 **4.6**

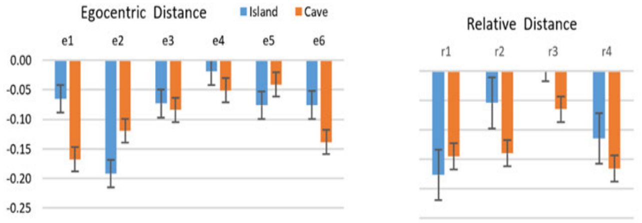


FIGURE 7. Average relative error (in meters) for *egocentric distance* and *relative distance* estimated by users for positions *e1* through *e6*, and for positions between *r1-r2, r2-r3, r3-r4* and *r5-r6*.

TABLE 16. Visual Realism: Site-Familiar vs Unfamiliar (*p* values).

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
Isl. &Cave	0.2538	0.1588	0.1628	0.2885	0.2202	0.2024
Island	0.2975	**0.0458**	0.2293	0.1792	**0.0285**	0.2577
Cave	0.2101	0.2718	0.0964	0.3978	0.4120	0.1470

shadows and colors to depth perception [31], [89], [84]. The *LSC to 3D* significantly higher performance in the island was according to the 90% of users, due to the many colors. We deem that colors particularly supported the monocular depth-cues induced at the higher distance-range (over 10 m.), compensating for the lack of binocular depth-cues [72]. Users also indicated that *shadows* were felt in the cave as main contributors to the perceived depth impression.

5) DISTANCE ESTIMATION

Results: Figures 8 shows users' performance for *egocentric distance estimation*. The test was only conducted on the LG display for practical reasons. Both in the island and cave, we could observe higher accuracy in our intermediate locations (*e3*, *e4*, *e5*, *r2, r3*). The table 15 shows ground truth values for island and cave locations.

The average relative error was below 5% for island's *r2* and *r3* and cave's *e4* and *e5*, and only 2% for the island's *e4*. The errors for island's *r3* and *e4* were significantly lower compared to the analogous cave's locations ($p = 0.021$, $p = 0.045$), whereas the error for cave's *e5* (compared to island's *e5*) was with a tendency of being significantly lower.

The average relative error was between 5% and 8% for island's *e1*, *e3*, *e5*, *e6* and cave's *e3* and *r3*; whereas it was between 12% and 19% for island's *e2 and* cave's *e1* and *e6*. The errors for island's *e1* and *e6* were significantly lower, and that for island's *e2* significantly higher, compared to cave's equivalent locations. The *p* values were respectively $p = 0.019$, $p = 0.002$, $p = 0.020$). There were otherwise no significant errors. Table 14 shows all *p*-values.

Analysis: The average relative error is contained, and it is also comparable to that measured in the work of Interrante et al. [37]. The above facts imply the proposed acquisition and visualization settings well support realistic distance estimation. The estimates' accuracy is overall comparable in both environments despite the differences in illumination and objects' distance. This indicates acquisition and visualization settings play a greater role than environment characteristics. Figure 7 diagrams also show users typically underestimated distances (negative error values), which is typical when observing synthetic scenes through an HMD [47], [49].

The good accuracy when observing the island (with the higher number of significantly better values), is sustained by many users' comments. They indicate that the cave's represented stalactites and stalagmites with relatively smooth surfaces of unknown size and shape, make more difficult to comprehend their precise 3D locations. On the other hand, the outdoor island's views portraying more complex and articulated objects' shapes, such as the green multi-directional prickly pears' blades, provide a well contrasted object's appearance that makes easier comprehend object's 3D positions and orientation. Furthermore, the island observed points represent smaller surfaces than in the cave, which also supports distance-estimation accuracy [71].

VII. USER STUDY 2 OUTCOME: FAMILIARITY

This section analyzes results focusing on the role played by place familiarity. Scores were gathered from both unfamiliar and site-familiar users looking at two different environments (island and cave) through the same display (iPhone). Referring to what is described in section IV-A this means: C1 vs D1, C2 vs D2 and C1+C2 vs D1+D2. Figure 8 and tables 16 through 18 show results.

A. HUMAN FACTORS

1) VISUAL REALISM

Results: Users scored *realness* high-level (i.e. high scores on all factors except for *excessive realism*). There were no significant differences except on *experience similar to real world*, where site-familiar users scored significantly higher in the island ($p = 0.0458$). *Photorealism* scored high with site-familiar users scoring it significantly higher than unfamiliar in the island ($p = 0.0285$). Table 16 shows all *p*-values.

Analysis: Visual realism scored high-level overall. It is of great interest and surprising to note the high-scores site-familiar users provided when observing the two environments. This is noted because the judgement of site-familiar users is based on correct environment knowledge, e.g. trees'

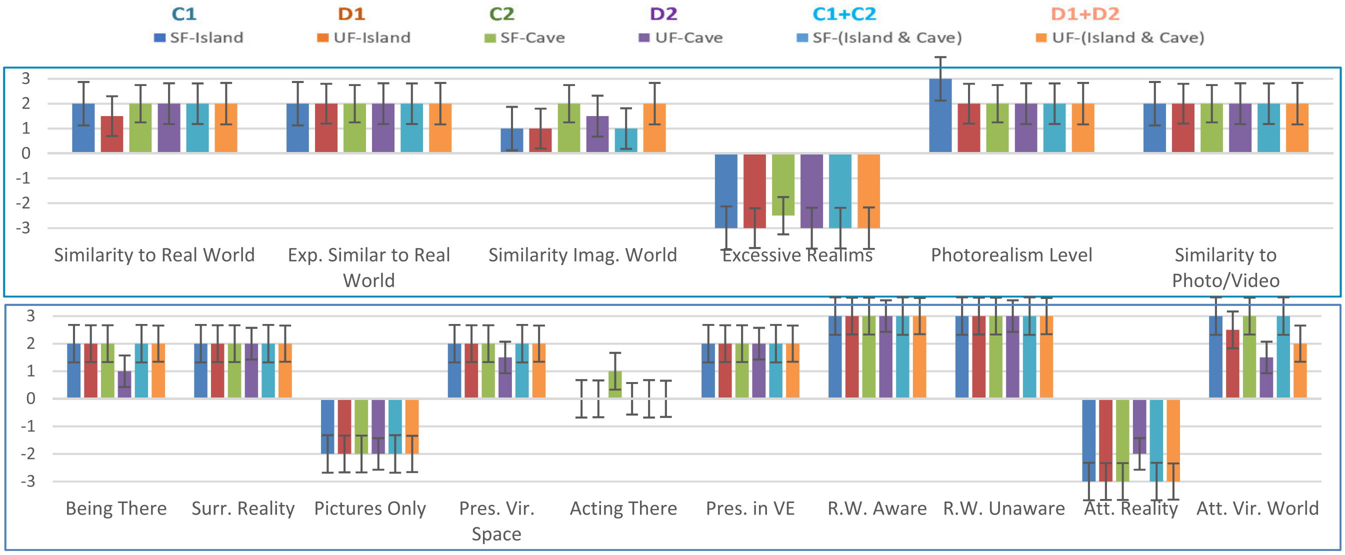


FIGURE 8. Outcome for User Study 2: Visual Realism (top-row)and Presence (bottom-row). Indicated values are median and standard error. Site familiarity is indicated by SF, whereas unfamiliarity to the site is indicated by UF. The two environments are indicated as Island and Cave.

and rocks' shapes. Therefore, they should be the best in spotting to unnatural appearance due to mismatches and deformations and therefore lowering their scores. It is also interesting to note that site-familiar and unfamiliar users' scores are generally similar.

Concerning *photorealism*, we know from the literature that it contributes to distance perception and depth-impression both for those familiar and unfamiliar to a place [56], [89], [90]. We can then observe that high photorealism seems to further support those site-familiars. However, this only happens in the island case. Users referred the island's higher scores were due to its image characteristics (illumination and colors) providing a more convincing effect. This would also justify the significant higher scores the same users give on the same environment to the *experience similar to real world* factor.

Overall, the comparison between site-familiar and unfamiliar users gave an outcome opposite to supposition. We expected unfamiliar users would overlook deformations or wrong details in scene elements, as they would not know the actual look of things, e.g. they would not notice trees that look taller or rocks with deformed shapes, whereas site-familiar users would be more critical. Rather, site-familiar users generally gave higher scores than unfamiliar users, and their comments were more appreciative. They typically commented seeing very realistically looking environments.

2) PRESENCE

Results: Site-familiar users scored presence generally higher than unfamiliar users. The scores of site-familiar users on *overall presence* (P1) were significantly higher than unfamiliar users in the island ($p = 0.0333$), and with a tendency to significant higher difference when considering both environments' scores combined ($p = 0.0599$). We found no significant differences on *spatial presence* (SP1-SP5).

TABLE 17. Overall, Spatial Presence: Site-Familiar vs Unfamiliar (*p* values).

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
Isl. & Cave	**0.0599**	0.2286	0.4575	0.1928	0.2469	0.1591
Island	**0.0333**	0.2719	0.5000	0.2189	0.1322	0.1690
Cave	0.0864	0.1853	0.4150	0.1668	0.3616	0.1491

TABLE 18. Involvement: Site-Familiar vs Unfamiliar (*p* values).

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
Island&Cave	0.2987	0.2729	0.2166	0.1440
Island	0.2960	0.3978	0.4096	0.2869
Cave	0.3015	0.1480	**0.0235**	**0.0010**

Site-familiar users scored *attention to surrounding reality* (INV3) and *attention to VR world* (INV4) significantly higher in the cave (respectively $p = 0.0235$ and $p = 0.001$). Tables 17 and 18 show all p-values.

Analysis: The generally higher scores of site-familiar users seemed connected to the more enthusiastic attitude these users had. According to their comments, the positive attitude came by seeing realistic visual reproductions of places they knew well, which brought memories. This is in line with some literature work indicating a correlation between familiarity and *presence* [91], [93].

Interestingly, the significantly better performance of site-familiar users when observing the island did not occur in the cave for P1, because unfamiliar users scored it higher. Based on users' comments, those higher scores were given because of the stronger depth impression the cave delivered, which enhanced *presence*.

VIII. CONCLUSION

This paper investigated the effect of photographic realism in VR when observing real places through a VR headset. We focused on the role played by: pixel-density (display); illumination and object distances (environment); previous knowledge (familiarity). Experimental data were gathered by interrogating users on the effect of a number of display and human factors (image lighting, visual realism, presence, emotions, depth impression and distance estimation), which were presented and analyzed according to the research questions. The main outcomes are summarized below.

- *Display*. The users appreciated the quality of both displays. They felt that *image lighting* factors, such as *color*, *lightness*, *vividness* and *sharpness*, contributed to realistic viewing. The contribution of higher *pixel-density* was positively felt and it prevailed over better lighting specs, leading to a significant improvement in some *realness* and *spatial presence* factors.
- *Environment.* The environment showed to affect the felt contribution to realistic viewing provided by *pixel-density*, *lightness*, *color*, *vividness*, and *photorealism*. It relevantly influenced *spatial presence* and in particularly the *sense of acting there* (in the visualized world), as well as *involvement* in terms of *attention to VR world*. Some *emotions* were clearly elicited. They were *enjoyment* and *relaxation* (while observing the island) and *anxiety* (while observing the cave). Both environment's illumination and distance to objects appeared to contribute towards *depth impression*, respectively in the island and cave. The cave scored a significantly higher *depth impression* because of the binocular depth-cues induced by the close distance to objects.
- *Display-Environment Combination.* The environment generally played a stronger role than the display, which proved that good quality displays are "transparent" to scene content. Nonetheless, the display characteristics were still able to further enhance *image-lighting*, *spatial presence* and some elements of *visual realism* and *emotions*, leading to significant score differences between one specific display-environment combination and all the others. This was the case for the LG-Island and the LG-Cave combinations. In case of *emotions*, the higher *pixel-density* display specifically enhanced *enjoyment* in the island and *anxiety* in the cave.
- *Depth Perception*. Depth sensation emerged as relevant element supported by environment illumination and distance to objects. It also seemed to possibly affect *anxiety* in the cave. The used system proved being able to provide effective 3D viewing and accurate distance estimation, with similar accuracy in both environments.
- *Familiarity.* The effectiveness of our systems was confirmed by site-familiar users scoring high *visual realism* and *presence*. Interestingly, site-familiar users scored some factors even significantly higher than users unfamiliar to the observed place, which proved that previous knowledge can positively enhance perceived *realism* and *presence*, e.g. by bringing memories. This represents a fascinating aspect worth future investigation.

The performed studies proved the overall effectiveness of the visual experience provided by omnidirectional photorealistic images, observed through a VR-headset, and displayed through three-dimensional and true-dimensional settings. Users' scores were generally high for *visual realism*, *presence*, *emotions* and *depth impression*, whereas *distance estimation* had accuracy acceptable for many applications.

We deem the outcome of both experiments was positive, particularly if we consider the limitations in terms of static images and a choice for system elements constrained by the use of off-the-shelf devices (camera and display). We hope this paper provides useful insight to VR system designers and helps understanding the potential VR headsets and photorealistic texture have towards several applications. Future work will include the use of dynamic sceneries and a comparison with synthetic image textures.

ACKNOWLEDGMENT

This work was initially supported by the CET Project (Sustainable Fruition of Eco-Tourist Circuits) [75] at Cutgana, and the Project MEGABIT–Programma Pia.ce.ri.–linea di intervento 2, University of Catania, Italy. The authors thank Cutgana's staff for great help with data acquisition and testing. They also thank Simona Pulvirenti (Sicilian Region, Italy) for insightful input, and to Kate Williams (University of Hertfordshire, U.K.) for rigorous revision help.

REFERENCES

[1] C. Bracken and P. Skalski, *Immersed in Media: Telepresence in Everyday Life*. Evanston, IL, USA: Routledge, 2010.

[2] H. Lee, Y. Kim, and A. Bianchi, "A survey on medical robotic telepresence design," *Arch. Des. Res.*, vol. 30, no. 1, pp. 61–71, 2017.

[3] S. Livatino, D. C. Guastella, G. Muscato, V. Rinaldi, L. Cantelli, C. D. Melita, A. Caniglia, R. Mazza, and G. Padula, "Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids," *IEEE Access*, vol. 9, pp. 25795–25808, 2021.

[4] S. Wang, C. Liu, and Y. Zhang, "Fully convolution network architecture for steel-beam crack detection in fast-stitching images," *Mech. Syst. Signal Process.*, vol. 165, Feb. 2022, Art. no. 108377.

[5] S. Kasahara and J. Rekimoto, "JackIn head: Immersive visual telepresence system with omnidirectional wearable camera for remote collaboration," in *Proc. 21st ACM Symp. Virtual Reality Softw. Technol.*, Nov. 2015, pp. 217–225.

[6] J. A. Stevens and J. P. Kincaid, "The relationship between presence and performance in virtual simulation training," *Open J. Model. Simul.*, vol. 3, no. 2, pp. 41–48, 2015.

[7] M. Kraus, N. Weiler, A. Diehl, and B. Bach, "Visualization in the VR-canvas: How much reality is good for immersive analytics in virtual reality?" in *Proc. 2nd Workshop Creation, Curation, Critique Conditioning Princ. Guidelines in Visualization (IEEE VIS)*, Jan. 2018, pp. 1–4.

[8] D. Patton, "How real is good enough? Assessing realism of presence in simulations and its effects on decision making," in *Human Performance and Decision-Making in Adaptive Systems*. Switzerland: Springer, 2014. [Online]. Available: https://link.springer.com/

[9] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, Jun. 1998.

[10] R. B. Welch, T. T. Blackmon, A. Liu, B. A. Mellers, and L. W. Stark, "The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence," *Presence, Teleoperators Virtual Environ.*, vol. 5, no. 3, pp. 263–273, Jan. 1996.

[11] K. Jaalama, N. Fagerholm, A. Julin, J.-P. Virtanen, M. Maksimainen, and H. Hyyppä, ''Sense of presence and sense of place in perceiving a 3D geovisualization for communication in urban planning—Differences introduced by prior familiarity with the place,'' *Landscape Urban Planning*, vol. 207, Mar. 2021, Art. no. 103996.

[12] *Insta360 One X*. Accessed: May 2023. [Online]. Available: https://insta360.com/product/insta360-onex

[13] S. Livatino, L. T. De Paolis, M. D'Agostino, A. Zocco, A. Agrimi, A. De Santis, L. V. Bruno, and M. Lapresa, ''Stereoscopic visualization and 3-D technologies in medical endoscopic teleoperation,'' *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 525–535, Jan. 2015.

[14] A. Regalbuto, S. Livatino, K. Edwards, and I. Mporas, ''Mobile VR headset usability evaluation of 2D and 3D panoramic views captured with different cameras,'' in *Proc. Int. Conf. Interact. Collaborative Robot.*, Sep. 2017, pp. 191–200.

[15] M. Brown and D. G. Lowe, ''Automatic panoramic image stitching using invariant features,'' *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, Aug. 2007.

[16] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, ''The impact of perception and presence on emotional reactions: A review of research in virtual reality,'' *Frontiers Psychol.*, vol. 6, p. 26, Jan. 2015.

[17] M. C. Juan and D. Pérez, ''Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE,'' *Presence, Teleoperators Virtual Environ.*, vol. 18, no. 3, pp. 232–248, Jun. 2009.

[18] D. A. Bowman and R. P. McMahan, ''Virtual reality: How much immersion is enough?'' *Computer*, vol. 40, no. 7, pp. 36–43, Jul. 2007.

[19] A. Kuijsters, W. A. Ijsselsteijn, M. T. Lambooij, and I. E. Heynderickx, ''Influence of chroma variations on naturalness and image quality of stereoscopic images,'' *Proc. SPIE*, vol. 7240, Feb. 2009, Art. no. 72401E.

[20] W. A. IJsselsteijn, H. de Ridder, and J. Vliegen, ''Subjective evaluation of stereoscopic images: Effects of camera parameters and display duration,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 225–233, Mar. 2000.

[21] F. E. Jamiy and R. Marsh, ''Distance estimation in virtual reality and augmented reality: A survey,'' in *Proc. IEEE Int. Conf. Electro Inf. Technol. (EIT)*, May 2019, pp. 63–68.

[22] P. Seuntiens, L. Meesters, and W. Ijsselsteijn, ''Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation,'' *ACM Trans. Appl. Perception*, vol. 3, no. 2, pp. 95–109, Apr. 2006.

[23] V. Juřík, L. Herman, D. Snopková, A. J. Galang, Z. Stachoň, J. Chmelík, P. Kubíček, and Č. Šašinka, ''The 3D hype: Evaluating the potential of real 3D visualization in geo-related applications,'' *PLoS ONE*, vol. 15, no. 5, May 2020, Art. no. e0233353.

[24] T. J. W. M. Janssen and F. J. J. Blommaert, ''A computational approach to image quality,'' *Displays*, vol. 21, no. 4, pp. 129–142, Oct. 2000.

[25] R. G. Kaptein, A. Kuijsters, M. T. M. Lambooij, W. A. IJsselsteijn, and I. E. J. Heynderickx, ''Performance evaluation of 3D-TV systems,'' *Proc. SPIE*, vol. 6808, pp. 443–453, Jan. 2008.

[26] J. A. Ferwerda, ''Three varieties of realism in computer graphics,'' *Proc. SPIE*, vol. 5007, pp. 290–297, Jun. 2003.

[27] M. A. Hagen, *Varieties of realism: Geometries of representational art, CUP Archive*. New York, NY, USA: Cambridge Univ. Press, 1986. [Online]. Available: https://archive.org/details/varietiesofreali0000hage

[28] M. S. Banks, D. M. Hoffman, J. Kim, and G. Wetzstein, ''3D displays,'' *Annu. Rev. Vis. Sci.*, vol. 2, pp. 397–435, 2016.

[29] A. Tiiro, *Effect of Visual Realism on Cybersickness in Virtual Reality*, vol. 350. Oulu, Finland: Univ. Oulu, 2018.

[30] P. J. Pardo, M. I. Suero, and Á. L. Pérez, ''Correlation between perception of color, shadows, and surface textures and the realism of a scene in virtual reality,'' *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 35, no. 4, p. B130, 2018.

[31] M. Slater, M. Usoh, and Y. Chrysanthou, ''The influence of dynamic shadows on presence in immersive virtual environments,'' in *Virtual Environments*. Cham, Switzerland: Springer, 1995, pp. 8–21.

[32] M. Slater, M. Usoh, and A. Steed, ''Depth of presence in virtual environments,'' *Presence, Teleoperators Virtual Environ.*, vol. 3, no. 2, pp. 130–144, Jan. 1994.

[33] V. Palad. (2019). *The Difference Between Clarity, Sharpness, and Contrast Sliders*. Accessed: Dec. 2020. [Online]. Available: https://pixelsandwanderlust.com/the-difference-between-clarity-sharpness-and-contrast-sliders/

[34] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, ''No-reference image sharpness assessment in autoregressive parameter space,'' *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015.

[35] A. F. R. Guarda, N. M. M. Rodrigues, and F. Pereira, ''Deep learning-based point cloud geometry coding: RD control through implicit and explicit quantization,'' in *Proc. IEEE Int. Conf. Multimedia Expo. Workshops (ICMEW)*, Jul. 2020, pp. 1–6.

[36] T. Schubert, F. Friedmann, and H. Regenbrecht, ''The experience of presence: Factor analytic insights,'' *Presence*, vol. 10, no. 3, pp. 266–281, Jun. 2001.

[37] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, and L. Anderson, ''Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments,'' *Presence*, vol. 17, no. 2, pp. 176–198, Apr. 2008.

[38] W. IJsselsteijn, H. D. Ridder, J. Freeman, S. E. Avons, and D. Bouwhuis, ''Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence,'' *Presence*, vol. 10, no. 3, pp. 298–311, Jun. 2001.

[39] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx, ''The effect of perspective on presence and space perception,'' *PLoS ONE*, vol. 8, no. 11, Nov. 2013, Art. no. e78513.

[40] J. S. Hvass, O. Larsen, K. B. Vendelbo, N. C. Nilsson, R. Nordahl, and S. Serafin, ''The effect of geometric realism on presence in a virtual reality game,'' in *Proc. IEEE Virtual Reality (VR)*, Mar. 2017, pp. 339–340.

[41] M. Huang and N. Alessi, ''Presence as an emotional experience,'' in *Medicine Meets Virtual Reality*. Amsterdam, The Netherlands: IOS Press, 1999.

[42] R. M. Baños, E. Etchemendy, D. Castilla, A. García-Palacios, S. Quero, and C. Botella, ''Positive mood induction procedures for virtual environments designed for elderly people?'' *Interacting With Comput.*, vol. 24, no. 3, pp. 131–138, May 2012.

[43] F. J. Seagull and D. M. Rooney, ''Filling a void: Developing a standard subjective assessment tool for surgical simulation through focused review of current practices,'' *Surgery*, vol. 156, no. 3, pp. 718–722, Sep. 2014.

[44] J. Häkkinen, T. Kawai, J. Takatalo, T. Leisti, J. Radun, A. Hirsaho, and G. Nyman, ''Measuring stereoscopic image quality experience with interpretation-based quality methodology,'' *Proc. SPIE*, vol. 6808, Jan. 2008, Art. no. 68081B.

[45] A. Ahrens, K. D. Lund, M. Marschall, and T. Dau, ''Sound source localization with varying amount of visual information in virtual reality,'' *PLoS ONE*, vol. 14, no. 3, Mar. 2019, Art. no. e0214603.

[46] W. D. Hairston, M. T. Wallace, J. W. Vaughan, B. E. Stein, J. L. Norris, and J. A. Schirillo, ''Visual localization ability influences cross-modal bias,'' *J. Cognit. Neurosci.*, vol. 15, no. 1, pp. 20–29, Jan. 2003.

[47] F. El Jamiy and R. Marsh, ''Survey on depth perception in head mounted displays: Distance estimation in virtual reality, augmented reality, and mixed reality,'' *IET Image Process.*, vol. 13, no. 5, pp. 707–712, Apr. 2019.

[48] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr, ''Effects of stereo viewing conditions on distance perception in virtual environments,'' *Presence*, vol. 17, no. 1, pp. 91–101, Feb. 2008.

[49] B. Li, R. Zhang, A. Nordman, and S. A. Kuhl, ''The effects of minification and display field of view on distance judgments in real and HMD-based environments,'' in *Proc. ACM SIGGRAPH Symp. Appl. Perception*, Sep. 2015, pp. 55–58.

[50] B. Li, J. Walker, and S. A. Kuhl, ''The effects of peripheral vision and light stimulation on distance judgments through HMDs,'' *ACM Trans. Appl. Perception*, vol. 15, no. 2, pp. 1–14, Apr. 2018.

[51] F. Kellner, B. Bolte, G. Bruder, U. Rautenberg, F. Steinicke, M. Lappe, and R. Koch, ''Geometric calibration of head-mounted displays and its effects on distance estimation,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 4, pp. 589–596, Apr. 2012.

[52] S. Livatino, G. Muscato, and F. Privitera, ''Stereo viewing and virtual reality technologies in mobile robot teleguide,'' *IEEE Trans. Robot.*, vol. 25, no. 6, pp. 1343–1355, Dec. 2009.

[53] E. Klein, J. E. Swan, G. S. Schmidt, M. A. Livingston, and O. G. Staadt, ''Measurement protocols for medium-field distance perception in large-screen immersive displays,'' in *Proc. IEEE Virtual Reality Conf.*, Mar. 2009, pp. 107–113.

[54] I. T. Feldstein, F. M. Kölsch, and R. Konrad, ''Egocentric distance perception: A comparative study investigating differences between real and virtual environments,'' *Perception*, vol. 49, no. 9, pp. 940–967, Sep. 2020.

[55] E. Brivio, S. Serino, E. N. Cousa, A. Zini, G. Riva, and G. De Leo, ''Virtual reality and 360° panorama technology: A media comparison to study changes in sense of presence, anxiety, and positive emotions,'' *Virtual Reality*, vol. 25, no. 2, pp. 303–311, Jun. 2021.

[56] V. Interrante, B. Ries, and L. Anderson, ''Distance perception in immersive virtual environments, revisited,'' in *Proc. IEEE Virtual Reality Conf. (VR)*, Mar. 2006, pp. 3–10.

[57] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall, "Does the quality of the computer graphics matter when judging distances in visually immersive environments?" *Presence*, vol. 13, no. 5, pp. 560–571, Oct. 2004.

[58] R. Newell, R. Canessa, and T. Sharma, "Visualizing our options for coastal places: Exploring realistic immersive geovisualizations as tools for inclusive approaches to coastal planning and management," *Frontiers Mar. Sci.*, vol. 4, p. 290, Sep. 2017.

[59] M. A. Davenport and D. H. Anderson, "Getting from sense of place to place-based management: An interpretive investigation of place meanings and perceptions of landscape change," *Soc. Natural Resour.*, vol. 18, no. 7, pp. 625–641, Aug. 2005.

[60] A. Julin, K. Jaalama, J.-P. Virtanen, M. Maksimainen, M. Kurkela, J. Hyyppä, and H. Hyyppä, "Automated multi-sensor 3D reconstruction for the Web," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 5, p. 221, May 2019.

[61] J.-P. Virtanen, M. Kurkela, T. Turppa, M. T. Vaaja, A. Julin, A. Kukko, J. Hyyppä, M. Ahlavuo, J. E. von Numers, H. Haggrén, and H. Hyyppä, "Depth camera indoor mapping for 3D virtual radio play," *Photogramm. Rec.*, vol. 33, no. 162, pp. 171–195, Jun. 2018.

[62] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. 77, no. 12, pp. 1321–1329, 1994.

[63] A. N. R. Chandra, F. El Jamiy, and H. Reza, "A review on usability and performance evaluation in virtual reality systems," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2019, pp. 1107–1114.

[64] E. Wilson, D. G. Hewett, B. C. Jolly, S. Janssens, and M. M. Beckmann, "Is that realistic? The development of a realism assessment questionnaire and its application in appraising three simulators for a gynaecology procedure," *Adv. Simul.*, vol. 3, no. 1, pp. 1–7, Dec. 2018.

[65] S. J. Hamstra, R. Brydges, R. Hatala, B. Zendejas, and D. A. Cook, "Reconsidering fidelity in simulation-based training," *Academic Med.*, vol. 89, no. 3, pp. 387–392, Mar. 2014.

[66] D. E. Brackney and K. Priode, "Back to reality: The use of the presence questionnaire for measurement of fidelity in simulation," *J. Nursing Meas.*, vol. 25, no. 2, pp. 66–73, Aug. 2017.

[67] L. Lioce, C. H. Meakim, M. K. Fey, J. V. Chmil, B. Mariani, and G. Alinier, "Standards of best practice: Simulation design standard IX," *Clin. Simul. Nursing*, vol. 11, no. 6, pp. 309–315, 2015.

[68] A. Hill, M. S. Horswill, A. M. Plooy, M. O. Watson, R. Karamatic, T. A. Basit, G. M. Wallis, S. Riek, R. Burgess-Limerick, and D. G. Hewett, "Assessing the realism of colonoscopy simulation: The development of an instrument and systematic comparison of 4 simulators," *Gastrointestinal Endoscopy*, vol. 75, no. 3, pp. 631–640, Mar. 2012.

[69] *Igroup Presence Questionnaire (IPQ)*. Accessed: 2021. [Online]. Available: http://www.igroup.org/pq/ipq/index.php

[70] P. B. Hibbard, A. E. Haines, and R. L. Hornsey, "Magnitude, precision, and realism of depth perception in stereoscopic vision," *Cognit. Res., Princ. Implications*, vol. 2, pp. 1–11, Dec. 2017.

[71] S. Baek and C. Lee, "Depth perception estimation of various stereoscopic displays," *Opt. Exp.*, vol. 24, no. 21, 2016, Art. no. 23618.

[72] J. C. A. Read, "Stereo vision and strabismus," *Eye*, vol. 29, no. 2, pp. 214–224, Feb. 2015.

[73] (Feb. 2023). *Cyclopean Isles*. Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Cyclopean_Isles

[74] (Feb. 2023). *Monello Cave (Grotta Monello)*. Italy. [Online]. Available: http://www.cutgana.unict.it/content/grotta-monello

[75] G. Signorello, G. Gallo, G. Farinella, S. Livatino, and A. Regalbuto. (Feb. 2023). *Cutgana CET-Project 3D Tours*. [Online]. Available: https://cutgana.unict.it/sites/cutgana.unict.it/VirtualTours3D/IsolaLachea/index.html and https://cutgana.unict.it/sites/cutgana.unict.it/VirtualTours3D/GrottaMonello/index.html

[76] D. J. Kasik, J. J. Troy, S. R. Amorosi, M. O. Murray, and S. N. Swamy, "Evaluating graphics displays for complex 3D models," *IEEE Comput. Graph. Appl.*, vol. 22, no. 3, pp. 56–64, May 2002.

[77] S. Livatino and K. Hochleitner, "Simple guidelines for testing VR applications," in *Advances Human Computer Interaction*, Rijeka, Croatia: InTech, 2008.

[78] J. Rubin, *Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests*. Hoboken, NJ, USA: Wiley, 1993.

[79] J. Nielsen, *Usability Engineering*. San Mateo, CA, USA: Morgan Kaufmann, 1993.

[80] J. D. Winter and D. Dodou, "Five-point Likert items: T test versus Mann-Whitney-Wilcoxon (*Addendum added October 2012*)," *Practical Assessment, Res., Eval.*, vol. 15, p. 11, Jan. 2010.

[81] T. Ni, D. Bowman, and J. Chen, "Increased display size and resolution improve task performance in information-rich virtual environments," in *Proc. Graphics Interface*, 2006, pp. 139–146.

[82] M. Slater, D.-P. Pertaub, and A. Steed, "Public speaking in virtual reality: Facing an audience of avatars," *IEEE Comput. Graph. Appl.*, vol. 19, no. 2, pp. 6–9, Mar./Apr. 1999.

[83] S. Livatino, G. Mattiolo, C. Castello, and L. Randazzo, "The role of photorealism in virtual reality games," in *Proc. IEEE Virtual Reality Int. Conf.*, Laval, France, Apr. 2007, pp. 31–38.

[84] R. Eggleston, W. Janson, and K. Aldrich, "VR system effects on size-distance judgements in a virtual environment," in *Proc. IEEE Virtual Reality*, Mar. 1996, pp. 139–146.

[85] F. Pallavicini, A. Pepe, A. Ferrari, G. Garcea, A. Zanacchi, and F. Mantovani, "What is the relationship among positive emotions, sense of presence, and ease of interaction in virtual reality systems? An on-site evaluation of a commercial virtual experience," *Presence*, vol. 27, no. 2, pp. 183–201, Feb. 2020.

[86] A. K. Seth, K. Suzuki, and H. D. Critchley, "An interoceptive predictive coding model of conscious presence," *Frontiers Psychol.*, vol. 2, p. 395, Jan. 2012.

[87] G. Bruder, F. Sanz, A. Olivier, and A. Lecuyer, "Distance estimation in large immersive projection systems," in *Proc. IEEE Virtual Reality (VR)*, Mar. 2015, pp. 27–32.

[88] J. S. Lappin, A. L. Shelton, and J. J. Rieser, "Environmental context influences visually perceived distance," *Perception Psychophysics*, vol. 68, no. 4, pp. 571–581, May 2006.

[89] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, "The perception of egocentric distances in virtual environments—A review," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–40, Nov. 2013.

[90] L. Phillips, B. Ries, V. Interrante, M. Kaeding, and L. Anderson, "Distance perception in NPR immersive virtual environments, revisited," in *Proc. Appl. Perception Graph. Vis.*, 2009, pp. 11–14.

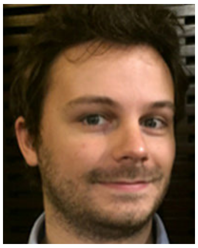
[91] S. M. Moore and M. N. Geuss, "Familiarity with teammate's attitudes improves team performance in virtual reality," *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0241011.

[92] C. Hendrix and W. Barfield, "Presence within virtual environments as a function of visual display parameters," *Presence: Teleoperators Virtual Environ.*, vol. 5, no. 3, pp. 274–289, Jan. 1996.

[93] R. A. Epstein, J. S. Higgins, K. Jablonski, and A. M. Feiler, "Visual scene processing in familiar and unfamiliar environments," *J. Neurophysiol.*, vol. 97, no. 5, pp. 3670–3683, May 2007.

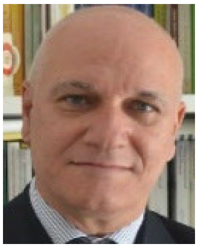
[94] Z. Gao, A. Hwang, G. Zhai, and E. Peli, "Correcting geometric distortions in stereoscopic 3D imaging," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0205032.

SALVATORE LIVATINO received the M.Sc. degree in computer science from the University of Pisa, Italy, in 1993, and the Ph.D. degree in computer science and engineering from Aalborg University, Denmark, in 2003. He is currently a Reader in Virtual Reality and Robotics with the University of Hertfordshire, U.K. His research interests include 3D visualization, virtual and augmented reality, with a focus on user interfaces for tele-exploration, teleoperation, and control panels.

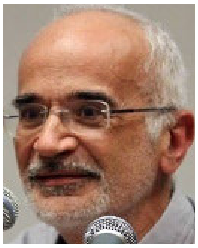


ALESSIO REGALBUTO received the M.Sc. degree in computer engineering from the University of Catania, Italy, in 2014, and the Ph.D. degree from the University of Hertfordshire, U.K., in 2019. He was with Ncam Technologies on virtual production and real-time AR systems for movies and live-events. He is currently a XR Specialist and a Senior Software Engineer with Magnopus, U.K., where he works on AR/MR/VR with Unity and Unreal Engine.

GIUSEPPE MORANA received the M.Sc. degree in computer science from the University of Pisa, Italy, in 1993, and the master's degree in management and university e-gov, in 2008. He is currently pursuing the Ph.D. degree in computer engineering with the University of Hertfordshire, U.K. He is a Senior Resource and Technical Manager with the University of Catania, Italy. His research interests include VR technology, computer graphics, and robotics.



GIOVANNI SIGNORELLO is currently a Full Professor with the University of Catania, Italy. Most of his research activity focuses on nonmarket valuation methods. His research interests include the use of computer science tools in the management and conservation of natural capital systems, and in the assessment of individual preferences.

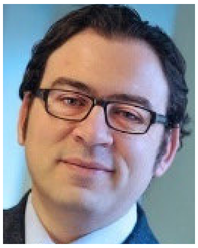


GIOVANNI GALLO received the M.Sc. and Ph.D. degrees in mathematics from the University of Catania, Italy, in 1984 and 1990, respectively, and the Ph.D. degree in computer science from New York University, USA, in 1992. He is currently a Full Professor of computer science with the Department of Math and Computer Science, University of Catania. His research interests include data science, computer graphics, and computer vision.

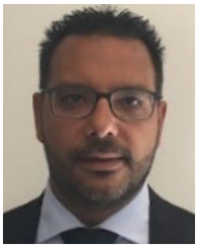
ALESSANDRO TORRISI received the M.Sc. and Ph.D. degrees in computer science from the University of Catania, Italy, in 2009 and 2013, respectively. He is currently a Senior Data Scientist with Fable Data, London, U.K. His research interests include machine and deep learning, computer vision, and natural language processing and understanding.

GIANLUCA PADULA received the M.Sc. degree in biomedical engineering from the Polytechnic University of Milan, Italy, in 2004, and the Ph.D. degree in biophysics and molecular biology from the Medical University of Łódź, Poland, in 2010. He is currently the Director of the Laboratory of Movement and Human Physical Performance, Dynamolab, Medical University of Łódź. He is a member of the Polish Academy of Science.

KATARZYNA PELC received the M.A. degree in economics from the University of Poznan, Poland, in 2003, and the M.A. degree in European studies from Aalborg University, Denmark, in 2004, where she is currently pursuing the Ph.D. degree in political science, with a focus on qualitative analysis. She is a Project Manager with the University of Bedfordshire, U.K., working with REAMIT, funded by Interreg NW Europe.



ALESSIO MALIZIA received the Ph.D. degree in computer science from the Sapienza University of Rome, Italy, in 2007. He joined the Department of Computer Science, University of Pisa, as an Associate Professor, in 2021. His previous experience includes the University of Hertfordshire as a Professor of user experience design (UX), Brunel University London, U.K., and Universidad Carlos III, Madrid, Spain.



GIOVANNI MARIA FARINELLA founded (2006) and currently directs the International Computer Vision and Medical Imaging Summer Schools (ICVSS and MISS), for which he received the PAMI Mark Everingham Prize, in 2017. He is currently a Professor with the Department of Mathematics and Computer Science, University of Catania, Italy. His research interests include computer vision and machine learning. He is an Associate Editor of IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and *Pattern Recognition*.

• • •