COMPUTER ANIMATION AND VIRTUAL WORLDS
*Comp. Anim. Virtual Worlds* 2014; **25**:69–81
Published online 15 July 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cav.1535

RESEARCH ARTICLE

# On multiuser perspectives in passive stereographic virtual environments

Paolo Tripicchio[†], Claudio Loconsole*[†], Andrea Piarulli, Emanuele Ruffaldi, Franco Tecchia and Massimo Bergamasco

Perceptual Robotics Laboratory, TECIP Institute, Scuola Superiore Sant'Anna, Pisa, Italy

## ABSTRACT

The use of stereographic systems is spreading out in modern society, from the revolution of cinematography to its adoption in high-tech products such as portable gaming devices or photo cameras. However, the fruition of immersive stereographic systems by more than one person at a time is still a research issue. In more detail, the class of passive stereo systems presents technological limitations of displaying correct multiuser perspectives. In fact, the stereo image projected onto the screen is usually rendered according to a unique point of view (PoV). Nevertheless, in multiuser systems, the selection of an appropriate PoV can minimize both optical discomfort and perspective distortion. This paper aims to evaluate which among existing PoV calculation methods provides the best performances in terms of projection realism, optical comfort and overall system usability in multiuser passive stereo systems. The performances have been evaluated in three different "distance" scenarios to take into account also the effects of binocular disparity in the PoV calculation. To accomplish this objective, we administered a questionnaire to nine couple subjects, evaluating each of the investigated PoV calculation methods for each of the three distance scenarios. Copyright © 2013 John Wiley & Sons, Ltd.

**KEYWORDS**
passive stereo vision; user study; virtual reality

***Correspondence**
Claudio Loconsole, Perceptual Robotics Laboratory, TECIP Institute, Scuola Superiore Sant'Anna, Pisa, Italy.
E-mail: c.loconsole@sssup.it

## 1. INTRODUCTION

Recently, more and more examples of virtual reality (VR) applications have entered both daily life, such as three-dimensional (3D) cinemas, gaming or interactive museums, and professional life, such as product design. To achieve better realism in such applications, the use of stereographic vision is advisable.

Stereoscopy indeed plays an important role in VR because it affects both the realism of visualization [1] and the sense of presence [2,3]. Furthermore, it allows users to perceive a virtual object as real, thanks to the simulation of human binocular disparity. According to this kind of simulation, appropriate different images are sent to each user's eye to reproduce a realistic stimulus. However, the choice of a stereographic projection technique can deeply affect user perception [4,5]. In more detail, it can affect the user's sensation of depth [6–8] and the user's performance in target acquisition tasks [9] and in haptic interaction [10,11]. The literature on stereographic vision for virtual environments (VEs) is wide, and a comprehensive recent review of display technology can be found in [12] and [13]. In the following, a description of passive stereographic solutions is presented to introduce the employed system architecture.

### 1.1. Stereographic Systems

All the techniques employed for the implementation of stereographic vision in large projection screens belong to two main categories, namely active and passive stereos. In this section, we focus only on the latter stereo techniques.

According to the technology used in the manufacture of glasses for passive stereo systems, glasses can be subdivided into the following categories:

- glasses suitable for anaglyph images
- glasses based on light polarization
- glasses that exploit frequency-shifting techniques

Frequency shifting currently represents the state-of-the-art technology for passive stereo system glasses, and one

[†]P. Tripicchio and C. Loconsole are both the first authors of this paper because of their equal contribution.

of the most widespread commercial solutions (interference filter technology) is given by the glasses produced by INFITEC [14]. Regarding physical structure, one of the most used architectures is the cave automatic VE (CAVE) [15] system: immersive VEs composed of multiple (usually more than three) very large stereographic screens that surround the user to give him a strong sense of presence and immersion. The realization of a CAVE requires a very large space, so it follows that there are a number of proposed solutions to simplify the CAVE structure while still maintaining its key characteristics. For example, it is very common to find four-wall CAVEs or just a single frontal stereo screen, commonly called powerwall. It is also usual to employ L-shaped systems, composed of just the frontal and bottom CAVE screens, to conjugate system simplicity and a wide enough view (Figure 1).

To enhance even more the immersion in VEs, head-tracking systems can be used in both active and passive stereo techniques. In this way, it is possible to adapt the perspective of the images displayed onto the screen by taking into account the real position of the user inside the environment to produce a faithful visualization. However, although the tracking system can acquire the position of all of the participants' points of view (PoVs), it is very difficult in passive stereo systems to let the users view different couple images whose perspective is adapted to their own eyes' position. Indeed, the main disadvantage of this systems is that only one participant at a time is able to perceive the right perspective.

For limited resolution and displaying surfaces, it is possible to have correct multiuser (MU) views using the so-called autostereoscopic display [16]. Anyway, the current technology limitations and the cost of such displays make this solution a nonoptimal choice for multistereo display. These limitations are significant because there are many applications where an MU approach is needed such as the following:

- Applications in the field of *cultural heritage* and the exploration of *virtual museums*. In these applications, the fruition of the exhibit space should be large enough to allow school students or groups of people seeking virtual visits.
- *Industrial prototyping* and *commerce*, where there is a possibility of displaying artifacts or products to a large public (or at least a few people together) or to allow a team to work on a *collaborative product design*.

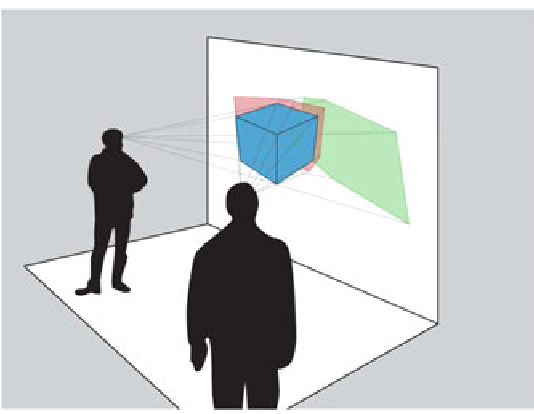


**Figure 1.** Schematic view of the experimental setup.

The fruition of consistent immersive screen projections by more than one user is surely an added value. Teaching activities could also benefit from cooperative immersive VR technologies.

For these reasons, this paper presents an analysis to quantitatively evaluate different PoV calculation methods in an L-shaped MU passive stereo system, to establish which of them offers the highest degree of visual realism, comfort and usability.

In this work, we conduct several experiments on different couple subjects using an L-shaped system projection screen and exploiting frequency-shifting INFITEC glass technology. Mainly, two factors are considered in the experiments:

- The *PoV calculation method*: Each method uses a different strategy for the common rendering PoV calculation (the used methods are presented and discussed in Section 3.1).
- The *range of the distances* of the displayed virtual objects from the common rendering PoV: The object distance deeply affects the simulation of the depth cue.

In Section 1.2, a brief overview of the stereographic vision theory is introduced to understand the rationale of our experiments. In Section 1.3, some of the possible approaches to MU stereo vision are presented.

### 1.2. Stereographic Vision Theory

Stereographic vision exploits a number of cues captured by the human vision apparatus that result in a perception of depth. Geometric perspective, lighting and shading, sizes of known objects, mutual occlusion and relative motion are all factors that contribute to the human perception of depth. Nevertheless, binocular disparity is considered one of the dominant depth cues [17]. For this reason, in a VE context, stereoscopic vision is consequently achieved by rendering two different perspective-correct images of the same virtual scenario for each eye and displaying them onto the same screen. This aims for an acceptable synthesis of the environment for the user's visual cortex. Indeed, the visual cortex accordingly fuses the two images and extracts depth information.

The most common techniques for stereo rendering (Figure 2) are the following:

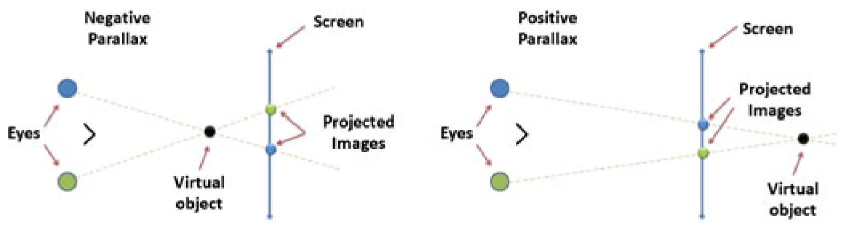



**Figure 2.** Resulting projections for stereo depth perception.

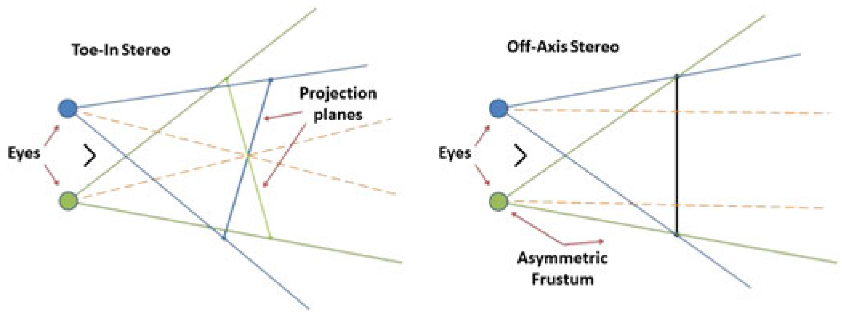



**Figure 3.** Possible implementation. (a) Toed in. (b) Asymmetric frustum projection.

- *Toed-in rendering method* (Figure 3a). It is an easy technique to be implemented and can be adapted to any visualization system [18].
- *Off-axis parallel projection* (Figure 3b). It is a more accurate technique to generate stereo images and is equivalent to lens shift in photography [19].

The VR framework chosen to carry out our experiments (Section 2) implements the off-axis parallel projection method.

## 1.3. Possible Multiuser Stereo Vision Approaches

In the context of an MU visualization system, the objective is to provide the best perspective for all users or, at least, to identify a way to improve the overall quality and system usability for the users [20,21].

The typical solution is the individual view that can be based on visual filtering [22–24], time multiplexing [25,26], personal head-mounted displays [27] or physical separation among the users [28,29]. When more than one user is involved in a VE, solutions based on a shared viewpoint are more suitable and, at the same time, less expensive. Some of these approaches use a common fixed PoV [30] or are based on deformation of the projection [25,31]. In case of shared space among different users, it is possible to adopt a technique that privileges a specific user, for example, according to his or her role in an interactive task [31,32]. The selection of a specific technique depends on the application, and it is important to evaluate the techniques in terms of performance [33–36] and consistency among the users' views [37].

The most suitable architectural solutions for an MU stereo vision application are the following:

- *multiple head-mounted display* solution [27]
- *CAVE-like visualization structure* solution

With a *multiple head-mounted display* solution, every user can see a different, strictly personal image, whose perspective could be adapted to his position, without affecting the images perceived by other users. Nevertheless, using this approach, the sense of presence could diminish because users would not be able to see their own body immersed inside the VE but rather a digital avatar representing them. For the same reasons, a great effort to assure the sense of copresence is needed. Furthermore, the adoption of this solution could be very expensive because each user must be equipped with a different stereographic screen.

A similar approach could be the adoption of one different separate screen for each user. The main drawback of this solution is the total loss of the sense of copresence resulting from the physical distance between the users. A possible solution in case of just two users is to orient the screens at 90° to each other, with an edge in common. However, while assuring the sense of copresence, this approach highly limits user's mobility.

The *CAVE-like visualization structure* solution, instead, is more appropriate to assure colocation and copresence, as it allows users to share the same physical space. There are several hardware approaches that may be used in such environments, but most of them are expensive and quite hard to realize.

Cheaper and more flexible solutions can be developed by adopting software approaches to the problem, thus with no need for special hardware. Although they do not guarantee a correct perspective to all the participants, the quality of the provided visualization can still be acceptable. An example of this kind of approach makes use of different rendering PoVs. Each subject perceives correctly only some components of the scene, typically the ones he is interacting with. So, in some circumstances, a common PoV can be found, depending on the PoVs of all the participants, to optimize the projection perceived by all the users and minimize potential image distortions and optical discomfort.

Starting from recent researches in shared viewpoints [36,38], this study compares six different methods to evaluate which common PoV calculation strategy maximizes three performance indexes (PIs): the projection realism [39], the optical discomfort [40] and the overall usability [41].

The methods we chose for the comparison are presented in detail in the experimental factors (Section 3.1).

In the next section, we report the experimental setup of the system, and in Section 3, the details of the experiments are given. Afterward, in Section 4, experimental results are presented and discussed to identify the best method for MU stereo VE for each of the conditions investigated in the study. Finally, Section 5 is dedicated to conclusions and future works.

## 2. EXPERIMENTAL SETUP

The purpose of the experiments is to evaluate which of the presented methods provide the “best” PIs in an MU passive stereo system both for different distance factors (far, medium distance and near objects) and also as overall evaluation.

The system used for experiments and tests is made up of four subsystems:

(1) The *L-shaped stereo system*
(2) Two *INFITEC glasses* equipped with Vicon reflective markers
(3) A *multiuser head-tracking system* based on Vicon technology
(4) The *extreme VR* (*XVR*) *framework*

The *L-shaped stereo system* (frontal and floor screens) is used as an immersive VR environment. Although such a system is not a complete CAVE, the immersion is good enough to perform the evaluation experiments. In Figure 4, the used architecture is shown. The projection system makes use of four projectors to display stereographic images on the two large screens. On each projector’s lens is mounted an *INFITEC* filter to properly decouple images for each users’ eye. To correctly convey the two images to the users, *INFITEC glasses* are equipped with reflective markers. These markers are needed by the *Vicon* [42] infrared capture system to properly track users’ head movements.

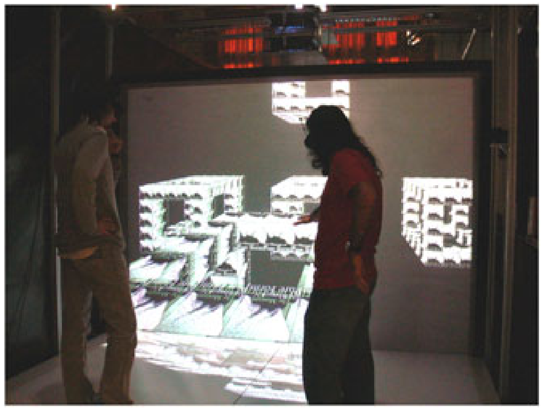


**Figure 4.** Near-environment experiment session running on the L-shaped system.

The *XVR framework* [43] has been adopted to generate the images that synthesize the VEs. The *XVR framework* engine provides basic facilities for stereographic vision and spatialized audio. With this framework, developers have access also to the low-level *OpenGL API* commands as well as the possibility to visualize complex multitextured animated models imported from the most common 3D modeling software.

To simplify the management of the L-shaped environment, we used an *XVR* module called the *XVR network renderer* [44]. Thanks to this software, it is possible to develop and run an *XVR* application without taking care of the physical structure used for the visualization. In fact, exploiting the cluster rendering technique, the network renderer automatically synchronizes the visualization of the VE on multiple screens with the developed *XVR* application.

The management of our system is carried out by four PCs. One of them runs the *XVR*-based application, whereas the other two are used to run *XVR network renderer* instances of the cluster. The last PC is dedicated to managing the head-tracking software. As mentioned earlier, to track the head, we adopt a commercial *Vicon* system solution. The tracker system is composed of seven cameras that track the positions of retro-reflective markers. These markers are placed over the stereo glasses in a particular geometry pattern to be clearly recognizable by the *Vicon* system. In this way, it is possible to retrieve the position and the orientation of the viewers’ head in real time.

Because of space constraints and environment exploration needs, only a couple of users at a time shared the L-shaped stereo system.

## 3. EXPERIMENT DESCRIPTION

For the sake of clarity, this section is subdivided into three subsections. In the first one, experimental factors are introduced. In the second subsection, information about the tested user group and experiment procedure is provided. Finally, in the third subsection, the experiment evaluation methodology is presented, and the influence of experimental factors on the users' experience is analyzed.

### 3.1. Experimental Factors

In this study, two experimental factors are investigated:

- The *common PoV calculation method*
- The *average distance* of the displayed virtual objects from the common rendering PoV

Now, we focus on both factors to highlight the advantages and the drawbacks of factor options.

When more than one user has to participate or cooperate inside the same VE, head-tracking issues arise. To solve this problem, we propose several *common PoV calculation methods* that can be used to determine which PoV has to be used for the VE rendering in an MU tracked stereo system. These methods are chosen among the most commonly used ones, and for each of them, we report the rationale for their choice.

The proposed *common PoV calculation methods* are the following (the users can freely move their head in all directions during the experimental sessions):

(1) *Fixed tracking (FT)*: In FT stereo systems, perspective is calculated by referring to a fixed PoV (Diagram 1 in Figure 5) that is placed in the middle of the system workspace at an average user height of 1.75 m. This method is the standard when both head tracking and perspective correction are not available [30], and it is also currently used in 3D cinemas. The lack of a head-tracking system results in a wrong perspective. When users are not in the exact fixed PoV, the system can produce misperception or even a projection that no longer results in a stereo image. A recent paper of Banks [45] discussed the perception side effects in the FT stereo perspective.

(2) *Manually switched single tracking (MSST)*: The system tracks only one of the two users at a time. The switch between the tracked users is manually performed (Diagram 2 in Figure 5 shows viewer B as the currently tracked user). This scheme is chosen for comparison purposes and in particular to

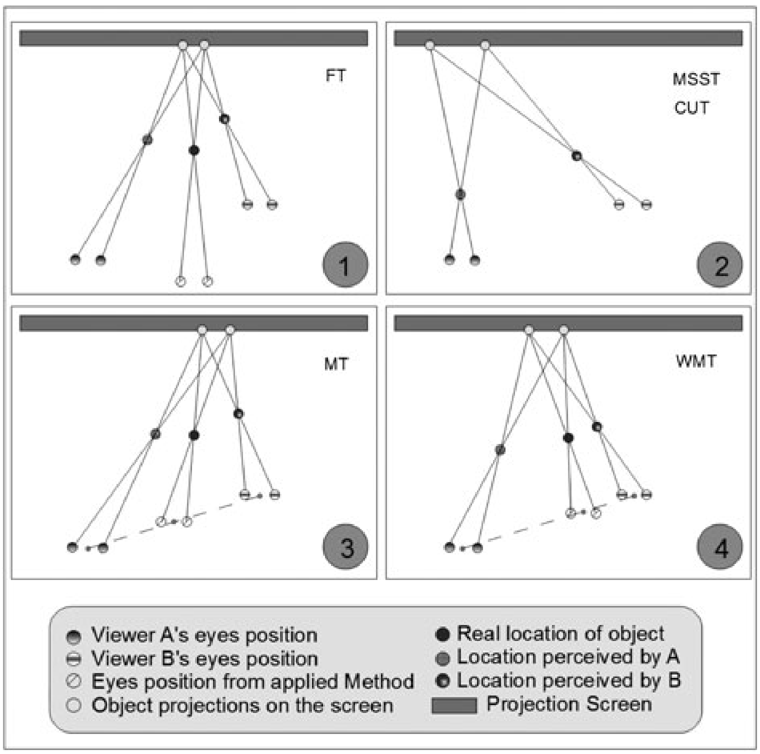



**Figure 5.** Schematic two-dimensional diagrams of proposed *common point-of-view calculation methods*. In the diagram, interocular distance for both viewers is assumed to be the same only for clarity purposes. Viewers A and B are located in the space at different distances from the screen. In the immersive environment, users can move freely. However, only in the pictures, for the sake of clarity, are the plane of the user's face and the eyes assumed parallel to the frontal projection screen. At the right side of each diagram, the acronym of the corresponding method is reported with the exception of weighted mean tracking with threshold because it behaves as a mixed solution between Diagrams 2 and 4. FT, fixed tracking; MSST, manually switched single tracking; CUT, closer-user tracking; MT, mean tracking; WMT, weighted mean tracking.

give the viewer a classic single-user experience in which he or she can evaluate the realism, the optical comfort and the usability of a VE in the MU context. Methods that privilege a specific user according to his or her role [31,32] can be considered equivalent to MSST.

(3) *Mean tracking (MT)*: In the MT method, the final position of the PoV is obtained through the vectorial mean between the positions of the two tracked users' head (Diagram 3 in Figure 5 shows the distortion perceived by both viewers). This method basically tries to lower the global distortion in the scene using a least square method. In particular, it works also when more than two users at a time share the same VE. This method for instance is employed in [38] where it is called the "view clustering" technique.

(4) *Closer-user tracking (CUT)*: In the CUT method, the system tracks only the user that is closer to the frontal screen (Diagram 2 in Figure 5). This is a variant of the MSST method with automatic changes between users' PoV. This choice is suggested by the consideration that, according to stereographic vision theory, the closer the user to the screen is, the greater is the disparity between the images that have to be conveyed to each users' eye. The main intention is to exactly assess if the abrupt switch between user controls could cause optical discomfort and if the closer user is effectively the one that more needs a proper perspective for the enjoyment of the virtual system.

(5) *Weighted mean tracking (WMT)*: In WMT, the final position of the PoV is obtained through the weighted mean between data related to the users. The closer the user is to the frontal screen, the greater the weight associated to that user (Diagram 4 in Figure 5). This method is based also on the assumption that the closer to the screen a user is, the more he or she is interacting with the environment. For this reason, we decided to test a method that gives some users more importance than others during a virtual interactive session.

(6) *Weighted mean tracking with threshold (WMTWT)*: The WMTWT method is similar to the WMT one. The only difference is that when the distance between the two users is above a certain threshold, only the user closer to the frontal screen is tracked (this is a variant between Diagrams 2 and 4 in Figure 5). This method mixes the WMT and CUT methods. The idea behind this method is the following: suppose a group (cluster) of users is watching the virtual scene and one of them leaves the group by going towards the screen to interact with some objects in the environment. So, it makes sense that if the control is based on a weighted mean, the projective distortion of the closer user will be greater than that of the other users in the group. Finally, with the introduction of a threshold, the modality can be switched in real time during the fruition of the virtual scene.

We define the common PoV as a suitable PoV that allows users to correctly enjoy the shared VE by minimizing the view disturbances. To identify the optimal common PoV in the MU stereo system, we perform several tests involving the previously proposed methods.

As mentioned before, the other experimental factor we investigate in the experiments is the *range of the distances* of the displayed virtual objects from the common rendering PoV in the VE. According to the distances involved in the VE, we define three types of scenarios (Figure 6):

(1) The *near-range* scenario: A synthetic environment composed by six hollow cubic figures (the distances from objects are in the range of few meters) is chosen as an example of an interactive environment where the distance between the user and the objects allows the manipulation of the objects themselves.

(2) The *medium-range* scenario: Presented is a 3D reconstruction of the artistic monumental cemetery of Pisa (Italy) in which both near and far objects (ranging from a few meters to a few dozens of meters) are present. Here, the architecture of the displayed building gives the users some hints about the correct projection, thanks to straight walls and lines of columns. Moreover, the presence of close and far objects in the same scene can make us better appreciate the different projection artifacts.

(3) The *far-range* scenario: An example is a virtual reproduction of the never-built Descamisado's monument landscape [46] (with user–object distances in the order of some hundred meters). In this case, the virtual representation presents objects that are really far from the users and has been chosen as an example of a virtual panoramic image often used in virtual tour simulations.

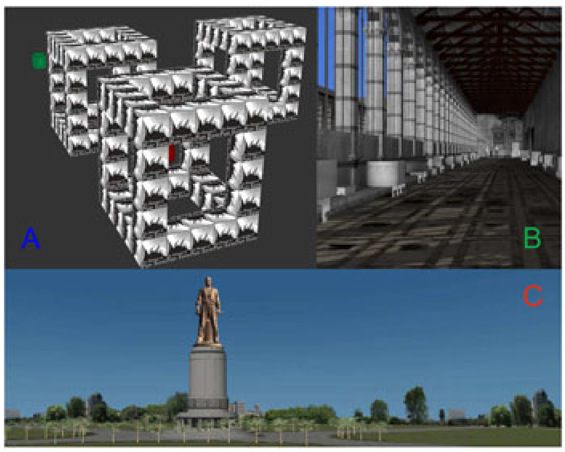



**Figure 6.** Scenarios employed during the experiment. Near-range (A), medium-range (B) and far-range (C) scenarios.

### 3.2. Tested User Group and Experimental Procedure Description

The experiments, the object of this work, are conducted on a group of nine couples (18 different subjects, the couples are composed by the same two people for every experiment). The six common PoV calculation methods presented in Section 3.1 are investigated (Table I) for each of the three scenarios for a total of 18 different combinations (called sessions). For every session, the couple of users performs an *explorative* (qualitative) *task* during which users are asked to explore the environment and to give a qualitative evaluation of the method they are experiencing in that moment for the specific scenario through a questionnaire. During each 5-min sessions, subjects are invited to move freely inside the VE and to deeply explore the scene (jumping, sitting on the floor, tilting the head, etc.). It follows that during each session, the spatial separation between the two users in the shared L-shaped passive stereo system is different both in width and in depth. In the transition among sessions, the couple subjects answer the questionnaire at the same moment independently in written form to avoid having the subjects influence each other. The order of session presentation to each couple is completely randomized to avoid the influence of a given repeated order of session presentation on subjects' evaluation.

However, a note is necessary about the switch among the common PoV calculation methods during the experiment. From the analysis of preliminary tests, we observe that an immediate switch between different PoVs is perceived by users as a heavy source of discomfort and uneasiness. For this reason, we decided to adopt a different PoV-switching procedure. This procedure performs a smooth transition between two PoVs in 1.5 s. In this way, the transition, although fast, becomes far less disturbing without requiring pauses in the protocol.

Regarding the experiment organization, the adopted timeline is the following:

- *before* the experiment:
    (1) Subjects are made aware of the structure involved and the modality of the experiment.
- *during* the experiment, for each combination of methods and scenario corresponding to a *session*:
    (1) A subject has to perform in pair with another subject the explorative task. At the beginning of each session, we remind the subjects to deeply explore the scene (jumping, sitting on the floor, tilting the head, etc.).
    (2) Subjects are asked to answer three questions with the aim of evaluating the PIs (Section 3.3); questions are answered while users are still immersed into the environment because the entire experiment for every couple of subjects lasts quite long (1.5 h). In this way, optical and cognitive flows of the participant, as well as the sense of presence, are not interrupted, and the different answers are coherent with the experience resulting from the particular combination of method and scenario. Moreover, as mentioned before, the couple subjects answer the questionnaire in the same moment and in written form to avoid having the subjects influence each other in answering the questions. So, the only potential between-subjects influence can rely on the movements performed in the exploration task by the same-couple subjects, owing to the need of sharing the same space in the L-shaped system. The questionnaire structure and the PIs are described in detail in the next section.

**Table I.** Acronym associated to each method.

| Acronym | Method |
|---|---|
| FT | Fixed tracking |
| MSST | Manually switched single tracking |
| MT | Mean tracking |
| CUT | Closer-user tracking |
| WMT | Weighted mean tracking |
| WMTWT | Weighted mean tracking with threshold |

### 3.3. Experimental Evaluation Methodology

For each session, each subject has to answer three questions (Table II) to evaluate the most suitable combination of method and scenario for an MU stereo vision. The answers to the three questions aim to evaluate three PIs:

PI1 *Projection realism*: that is, if subjects believe that the proposed method does not deform too much the perceived objects, presenting them as sufficiently similar to real objects
PI2 *Optical comfort*: to express the presence or absence of an unease sensation generated by the user visual apparatus during experiment sessions
PI3 *Overall method usability*

For each question, scores can vary from 1 to 5, where 1 represents a totally negative evaluation and 5 a very positive one (Column 3 of Table II). The entire experiment leads each user to answer a total of 54 questions (three questions for 18 experimental sessions). In the next section, a detailed statistical analysis of the scores given by the subjects is performed to identify the best common PoV calculation method's score according to a specific scenario.

## 4. RESULTS AND DISCUSSION

For each couple, six PoV visualization methods are administered, and the methods are tested for three different scenarios. The test is based on the statistical evaluation of three PIs (PI1, PI2 and PI3) on the basis of two factors:

**Table II.** Questions for the performance index evaluation.

| # | Question formulation | Possible answers |
|---|---|---|
| 1 | How much was the projection realistic? | Totally unrealistic (1) to totally realistic (5) |
| 2 | Have you experienced optical discomfort? | Painful (1) to comfortable (5) |
| 3 | According to you, is this method usable? | Not usable (1) to usable (5) |

- *method*, PoV estimation method
- *scenario*

All dependent variables ( PI1, PI2 and PI3) are tested for Gaussianity [47]. As all variable are non-normally distributed, analyses are conducted with the use of nonparametric statistics. For each couple and each PI, a paired Mann–Whitney test [48] is performed, to evaluate the null hypothesis of no significant difference in score distribution between individuals of the couples.

Dependent variables are then tested for couple effect (Kruskal–Wallis test [49] and between-factor *couple*), and significance levels are adjusted on the basis of permutation tests [50]. Permutation tests work as follows: given the null hypothesis that factor labeling is arbitrary (i.e., a *couple* has no significant effects), the significance of a statistic can be assessed by comparison with the distribution of values obtained when labels are permuted. Obviously, even for small datasets, not all permutations can be considered (the number of permutations and consequent tests for a dataset of say 20 elements reaches $10^{18}$). To overcome this limitation, given all possible permutations, a subset is chosen using a randomized sampling: for each of the three dependent variables, 10 000 permutations of the original dataset are extracted. A significant effect of *couple* is found for each of the three PIs ($p < 0.001$ for each index). To give equal weights to *couple* scores (thus removing the couple effect), distributions of scores for each PI and each *couple* are standardized, obtaining distributions with zero mean and equal range.

For each of the three standardized dependent variables, the same statistical design is applied: between factors are *scenario*, *method* and their interaction. Factor effects are estimated on the basis of the Scheirer–Ray–Hare test [51], which is the nonparametric equivalent of a multiple-factor analysis of variance. For each factor, significance levels are adjusted on the basis of permutation tests. For each significant factor, differences between factor levels are assessed via *post hoc* analyses [48]. Because dependent variables are not normally distributed, *post hoc* tests are performed using the Mann–Whitney test. *Post hoc* significance levels are adjusted using the Sidak correction for multiple comparisons.

Figure 7 reports descriptive nonparametric statistics for the dataset. Each subplot depicts the statistics of the six method scores for each standardized PI (Columns 1, 2 and 3 are related to PI1, PI2 and PI3, respectively) and for each *scenario* (Rows 1, 2 and 3 refer to near, medium and far ranges, respectively). Each boxplot graphically depicts the statistics of the underlying distribution through

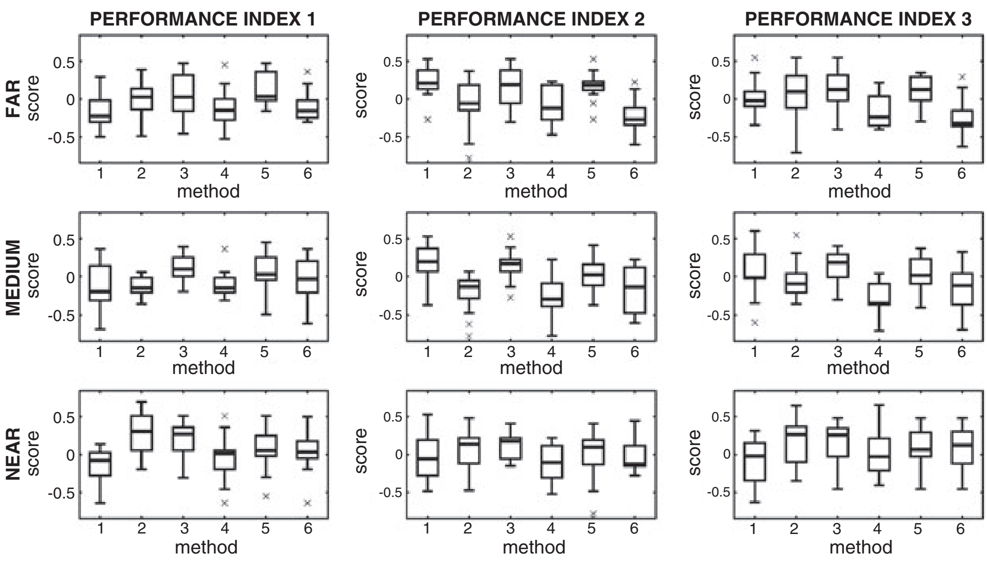



**Figure 7.** Distributions of *method* scores for each *scenario–performance index* couple. In this and in the following figures, each boxplot graphically describes distribution statistics: the box has lines at the lower quartile, median (thick line) and upper quartile values. Whiskers extend to values within 1.5 times the interquartile range from the end of the box. Outliers (i.e., data with values exceeding the ends of the whiskers) are displayed with an 'x' sign.


*Comp. Anim. Virtual Worlds* 2014; **25**:69–81 © 2013 John Wiley & Sons, Ltd.
DOI: 10.1002/cav

**Table III.** Degrees of freedom, chi-squared statistics and significance levels for Scheirer–Ray–Hare tests on the three performance indexes are presented.

| | | Performance Index 1 | | Performance Index 2 | | Performance Index 3 | |
|---|---|---|---|---|---|---|---|
| Factor | *df* | $\chi^2$ | Significance | $\chi^2$ | Significance | $\chi^2$ | Significance |
| Method | 5 | 42.28 | <0.001 | 59.74 | <0.001 | 44.00 | <0.001 |
| Scenario | 3 | 15.35 | <0.001 | 5.13 | <0.080 | 6.34 | <0.050 |
| Method ∗ scenario | 10 | 17.01 | <0.080 | 17.89 | <0.060 | 18.17 | <0.060 |

In this and in the following, only $p < 0.05$ will be considered significant. In the significance columns, *p*-values derived from permutation tests are presented.

**Table IV.** *Post hoc* results for *method* on the three performance indexes.

| | Significant comparisons | | |
|---|---|---|---|
| Method | Performance Index 1 | Performance Index 2 | Performance Index 3 |
| FT | — | $> 2, p < 0.050$<br>$> 4, p < 0.001$<br>$> 6, p < 0.001$ | — |
| MSST | $> 1, p < 0.010$ | — | $> 4, p < 0.005$ |
| MT | $> 1, p < 0.001$<br>$> 4, p < 0.001$<br>$> 6, p < 0.050$ | $> 2, p < 0.005$<br>$> 4, p < 0.001$<br>$> 6, p < 0.001$ | $> 1, p < 0.050$<br>$> 4, p < 0.001$<br>$> 6, p < 0.005$ |
| CUT | — | — | — |
| WMT | $> 1, p < 0.001$<br>$> 4, p < 0.005$ | $> 2, p < 0.050$<br>$> 4, p < 0.001$<br>$> 6, p < 0.050$ | $> 4, p < 0.001$<br>$> 6, p < 0.050$ |
| WMTWT | — | — | — |

For each *method* and each performance index, only comparisons with methods having significantly lower scores are reported.

FT, fixed tracking; MSST, manually switched single tracking; CUT, closer-user tracking; MT, mean tracking; WMT, weighted mean tracking.

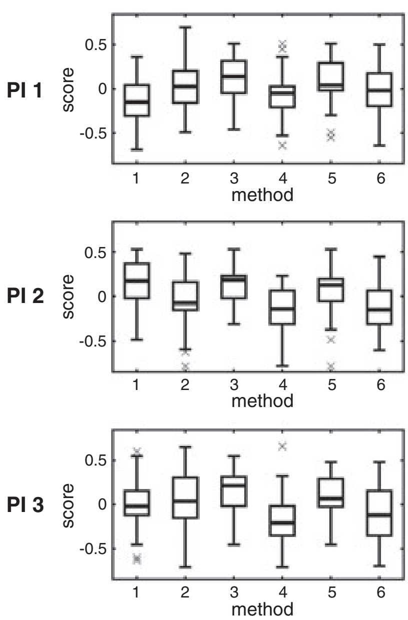



**Figure 8.** Distributions of *method* scores for each performance index (PI). This figure depicts significant *method* effects for each of the three PIs (*p*-values are all <0.001).

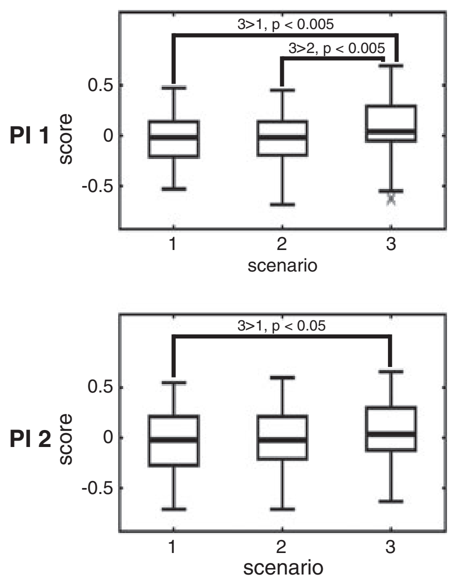



**Figure 9.** Distributions of *scenario* scores for PI1 and PI3. Significant *post hoc* levels are denoted by brackets. PI, performance index.


DOI: 10.1002/cav

five parameters: sample minimum, lower quartile, median, upper quartile and sample maximum.

No significant difference was found between scores of individuals in each couple (i.e., individuals in the same couple have a similar perception of the presented experimental condition), which holds for all three indexes. On the other hand, a significant *couple* effect is found consistently for the three PIs (i.e., differences in the scale of judgments exist between different couples). These data taken together, although not conclusive, seem to indicate the existence of a mutual influence of subjects concurrently undergoing the same experimental condition. On the contrary, scores are significantly different from one couple to another.

The *couple* effect is removed by standardizing the scores for each *couple* and each PI. Significant *method* effects are found consistently for each of the three PIs (all $p$-values are lower than 0.001, Table III). This means that some PoV estimation methods were perceived as significantly better than others in terms of projection realism, optical comfort and overall system usability. For PI1 and PI3, significant effects of *scenario* ($p < 0.001$ and $p < 0.05$, respectively) are also apparent, meaning that some of the presented scenarios are judged better than the others in terms of projection realism and overall usability, independently from the PoV estimation methods.

*Post hoc* analyses are carried out for *method* on the three PIs and for *scenario* on PI1 and PI3. The significance of the *method* between factors tells us, for example, that some of the administered PoV estimation methods were significantly better than the others in terms of projection realism, optical comfort and usability (PI1, PI2 and PI3, respectively) but does not tell us which ones. *Post hoc* analyses are conducted to precisely evaluate which are these methods (the same rationale also applies to *scenario*). Table IV reports significant differences between methods for each PI, whereas Figure 8 reports method scores for each of the three PIs. As apparent from both Figure 8 and Table IV, the best-rated PoV calculation methods on the basis of both PI1 and P3 are MSST, MT and WMT, whereas for PI2, FT, MT and WMT reach higher scores. From the intersection between the aforementioned findings, the methods perceived as most realistic by couples are MT and WMT. Figure 3 depicts *scenario* effects on PI1 and PI3: independently from the *method*, *far-range scenario* is perceived as more realistic than *near* ($p < 0.005$) and *medium-range* ($p < 0.005$) *scenarios*. Regarding the overall usability, *far-range scenario* reaches significantly higher scores than *near-range scenario* (Figure 9).

### 4.1. Discussion of Results

The analysis of the results shows the importance of head tracking for MU immersive stereo VEs. In more detail, the methods that include in the PoV calculation strategy a sort of "mean" of the PoVs of the two users maximize the PIs. However, the introduction of the "threshold" concept, or in general a switch between the PoVs of the two users, worsens both the projection realism (PI1) and the optical comfort, which causes also a low overall usability. In fact, considering all three PIs, the best score methods are *MT* and *WMT*. Relevant scores of PI2 are obtained also by the *FT* method, which in optical comfort obtains a score comparable with that of *MT*. This finding confirms that the FT method, even if introducing perspective distortion, does not introduce optical discomfort, and then its use in modern 3D cinemas is safe for the final user.

## 5. CONCLUSION AND FUTURE WORKS

The study presented a qualitative analysis of MU perspectives in passive stereographic VEs. A comparison among six PoV calculation strategy methods is conducted. In particular, the impacts of three PIs (projection realism, optical comfort and overall method usability) are taken into account to estimate which of the presented methods is more suitable to be employed in different VEs. Nine couple subjects, for a total of 18 users, performed 18 explorative tasks (resulting from the combination of the six PoV calculation strategy methods and three "distance" scenarios) in a passive L-shaped immersive stereo system. We find that the best score methods according to all the PIs are *MT* and *WMT*.

### 5.1. Future Works

Regarding future work, we have planned two main research activities. Firstly, we would like to test the investigated PoV calculation methods with more than two people at the same time to allow more realistic VR interaction for groups of people. Secondly, we intend to carry out an investigation of other feedback modalities with stereo properties, in particular involving aural stimuli.

In more detail, the latter investigation aims to compensate for the discomfort in the perception in passive MU stereo applications with the introduction of acoustic cues in addition to visual stimuli.

## REFERENCES

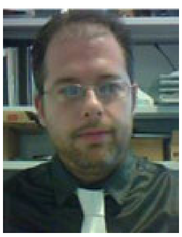
1. Lo CH, Chalmers A. Stereo vision for computer graphics: the effect that stereo vision has on human judgments of visual realism. In *Proceedings of the 19th Spring Conference on Computer Graphics*. Budmerice, Slovakia, ACM Press, New York, NY, USA, 2003; 109–117.
2. Hendrix CM. Exploratory studies on the sense of presence in virtual environments as a function of visual and auditory display parameters, *Ph.D. Thesis*, University of Washington, 1994.
3. IJsselsteijn WA, de Ridder H, Freeman J, Avons SE. Presence: concept, determinants and measurement, In *Proceedings of the SPIE, the International Society for*

*Optical Engineering*, San Jose, California, USA, 2000; 520–529.

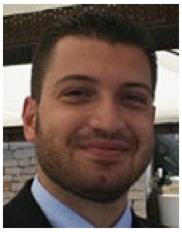
4. IJsselsteijn W, de Ridder H, Hamberg R. Perceptual factors in stereoscopic displays: the effect of stereoscopic filming parameters on perceived quality and reported eyestrain. *Proceedings of SPIE* 1998; **3299**: 282.
5. IJsselsteijn WA, de Ridder H, Vliegen J. Subjective evaluation of stereoscopic images: effects of camera parameters and display duration. *IEEE Transactions on Circuits and Systems for Video Technology* 2000; **10**(2): 225–233.
6. Wann J, Mon-Williams M, Rushton S. Natural problems for stereoscopic depth perception in virtual environments. *Vision Research* 1994; **35**(19): 2731–2736.
7. Ijsselsteijn W, de Ridder H, Hamberg R, Bouwhuis D, Freeman J. Perceived depth and the feeling of presence in 3DTV. *Displays* 1998; **18**(4): 207–214.
8. Patterson R, Moe L, Hewitt T. Factors that affect depth perception in stereoscopic displays. *Human Factors* 1992; **34**(6): 655–667.
9. Wall SA, Paynter K, Shillito AM, Wright M, Scali S. The effect of haptic feedback and stereo graphics in a 3d target acquisition task, In *Proceedings of Eurohaptics*, Edinburgh, Scotland, 2002; 23–29.
10. Bouguila L, Sato M, Ishii M. Effect of coupling haptics and stereopsis on depth perception in virtual environment, In *World Multiconference on Systemics, Cybernetics and Informatics (SCI)*, Orlando, Florida, USA, 2000; 406–414.
11. Chun K, Verplank B, Barbagli F, Salisbury K. Evaluating haptics and 3D stereo displays using Fitts' law, In *Proceedings of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications, HAVE*, Ottawa, Canada, 2004; 53–58.
12. Urey H, Chellappan K, Erden E, Surman P. State of the art in stereoscopic and autostereoscopic displays. *Proceedings of the IEEE* 2011; **99**(4): 540 –555.
13. Holliman N, Dodgson N, Favalora G, Pockett L. Three-dimensional displays: a review and applications analysis. *IEEE Transactions on Broadcasting* 2011; **57**(2): 362 –371.
14. http://www.infitec.net.
15. Cruz-Neira C, Sandin DJ, DeFanti TA. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Siggraph '93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. Anaheim, California, USA, ACM Press New York, NY, USA, 1993; 135–142.
16. Dodgson N. Autostereoscopic 3d displays. *Computer* 2005; **38**(8): 31 –36.
17. Qian N. Binocular disparity and the perception of depth. *Neuron* 1997; **18**: 359–368.
18. Yamanoue H. The differences between toed-in camera configurations and parallel camera configurations in shooting stereoscopic images. In *IEEE International Conference on Multimedia and Expo (ICME),* Toronto, Canada*,* 2006; 1701–1704.
19. Jones GR, Lee D, Holliman NS, Ezra D. Controlling perceived depth in stereoscopic images, In *Proceedings of SPIE*, Vol. 4297, San José, California, USA, 2001; 42–53.
20. Rani NS, VaithyaNathan V. Methodology for selecting the locatiton of an optimum viewing point for an ideal 3D perspective view of a given 3D environment. *Journal of Computer Science* 2006; **2**(8): 634–637.
21. Rule J. The geometry of stereoscopic pictures. *Journal of the Optical Society of America* 1941; **31**: 325–334.
22. Springer JP, Sladeczek C, Scheffler M, Hochstrate J, Melchior F, Frohlich B. Combining wave field synthesis and multi-viewer stereo displays, In *IEEE Virtual Reality Conference Proceedings*, Alexandria, Virginia, USA, 2006; 237–240.
23. Agrawala M, Beers AC, McDowall I, Fröhlich B, Bolas M, Hanrahan P. The two-user responsive workbench: support for collaboration through individual views of a shared space. In *SIGGRAPH '97: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. Los Angeles, California, USA, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1997; 327–332.
24. Frhlich B, Blach R, Stefani O, Hochstrate J, Hoffmann J, Klger K, Bues M. Implementing multi-viewer stereo displays, In *Proceedings of WSCG (full papers)*, Plzen, Czech Republic, 2005; 139–146.
25. Blom K, Lindahl G, Cruz-Neira C. Multiple active viewers in projection-based immersive environments, In *Immersive Projection Technology Workshop*, Orlando, Florida, USA, 2002.
26. Wozelka R. Implementation and deployment of a stereo projection system using low-cost components, In *CESCG 2006 Web Proceedings*, Wien, Austria, 2006.
27. Cakmakci O, Student Member, Rolland J, Associate Member. Head-worn displays: a review. *Journal of Display Technology* 2006; **2**(3): 199–216.
28. Bimber O, Encarnação LM, Schmalstieg D. The virtual showcase as a new platform for augmented reality digital storytelling, In *Proceedings of the Workshop on Virtual Environments 2003*, Zurich, Switzerland, 2003; 87–95.
29. Arthur K, Preston T, Taylor R, Brooks F, Whitton M, Wright W. Designing and building the pit: a head-

tracked stereo workspace for two users. *Technical Report*, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 1998.
30. McGinity M, Shaw J, Kuchelmeister V, Hardjono A, Del Favero D. Avie: a versatile multi-user stereo 360° interactive VR theatre. In *EDT '07: Proceedings of the 2007 Workshop on Emerging Displays Technologies*. San Diego, California, USA, ACM Press, New York, NY, USA, 2007; 2.
31. Naemura T, Kaneko M, Harashima H. Multi-user immersive stereo, In *Proceedings of the 1998 International Conference on Image Processing, 1998 ICIP*, Chicago, Illinois, USA, 1998; 903–907.
32. Heldal I, Schroeder R, Steed A, Axelsson A, Spante M, Widestrom J. Immersiveness and symmetry in copresent scenarios, In *Virtual Reality, 2005 IEEE*, Bonn, Germany, 2005; 171–178.
33. Wartell ZJ, Hodges LF, Ribarsky W. The analytic distortion induced by false-eye separation in head-tracked stereoscopic displays. *Technical Report*, Georgia Institute of Technology, 1999.
34. Arthur KW, Booth KS, Ware C. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems (TOIS)* 1993; **11**(3): 239–265.
35. Bolas M, McDowall I, Corr D. New research and explorations into multiuser immersive display systems. *IEEE Computer Graphics and Applications* 2004; **24**(1): 18–21.
36. Simon A. Usability of multiviewpoint images for spatial interaction in projection-based display systems. *IEEE Transactions on Visualization and Computer Graphics* 2007; **13**(1): 26–33.
37. de Haan G, Molenaar R, Koutek M, Post FH. Consistent viewing and interaction for multiple users in projection-based VR systems. *Computer Graphics Forum* 2007; **26**(3): 695–704.
38. Marbach J. Image blending and view clustering for multi-viewer immersive projection environments. In *Proceedings of the 2009 IEEE Virtual Reality Conference*. Lafayette, Louisiana, USA, 2009; 51–54.
39. Steinicke F, Bruder G, Kuhl S. Realistic perspective projections for virtual objects and environments. *ACM Transactions on Graphics* 2011; **30**(5): 112:1–112:10.
40. Pala S, Stevens R, Surman P. Optical cross-talk and visual comfort of a stereoscopic display used in a real-time application, In *Proceedings of SPIE XIV*, San José, California, 2007. 649011-649012.
41. Nielsen J. Usability engineering, 1993.
42. http://www.vicon.com.
43. Carrozzino M, Tecchia F, Bacinelli S, Cappelletti C, Bergamasco M. Lowering the development time of multimodal interactive application: the real-life experience of the XVR project. In *ACE '05: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. Valencia, Spain, ACM Press, New York, NY, USA, 2005; 270–273.
44. Marino G, Tecchia F, Bergamasco M. Cluster-based rendering of complex virtual environments, In *Intuition 2007: Proceedings of the 4th International Intuition Conference on Virtual Reality and Virtual Environments*, Athens, Greece, 2007.
45. Banks M, Held R, Girshick AR. Perception of 3-d layout in stereo displays. *Information Display* 2009; **25**(1): 12–16.
46. Carrozzino M, Evangelista C, Scucces A, Tecchia F, Tennirelli G, Bergamasco M. The virtual museum of sculpture. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*, DIMEA '08. ACM, New York, NY, USA, 2008; 100–106.
47. Lilliefors HW. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* 1967; **62**(318): 399–402.
48. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* 1947; **18**(1): 50–60.
49. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 1952; **47**(260): 583–621.
50. Manly BFJ. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, Boca Raton, Florida, USA, 2007.
51. Sokal RR, Rohlf FJ. *The Principles and Practice of Statistics in Biological Research*, 3 edn. W.H. Freeman, New York, USA, 1994.

## AUTHORS' BIOGRAPHIES



**Paolo Tripicchio** (Eng. Ph.D.) is a Teaching Assistant in Mechatronics at Pisa University since 2007. He received an M.S. in Automation Engineering from the University of Pisa in 2006 and a Ph.D. in Perceptual Robotics with honors from Scuola Superiore Sant'Anna, Pisa (Italy), in 2012. He is currently a Member of the Intelligent Automation Systems Division of the PERCRO laboratory of the TECIP Institute in Pisa and a Scientific Program Manager of the Gustavo Stefanini advanced robotics research center in La Spezia, Italy. His research concerns robotics, haptics, control systems, virtual reality, and AI. He is an author of more than 25 scientific publications in international journals and conference proceedings. He has been the Chair of the 1st International Workshop on "Intelligent Multimodal Interfaces Applied in Skills Transfer, Healthcare and Rehabilitation" and of the "Joint PRESENCCIA and SKILLS PhD Symposium." He was a Lecturer on "virtual reality" in an M.Sc. course in Electronics and Electrical Systems of the Orizaba Institute of Technology, Mexico, and on "virtual environments, design and applications" for the Researchers and Professors of Toluca autonomous University and Orizaba Institute of Technology in 2012.



**Claudio Loconsole** (Eng. Ph.D., S'11) is a Research Fellow at PERCRO Laboratory, Scuola Superiore Sant'Anna, Pisa, Italy, where he received his Ph.D. with honors in Innovative Technologies (curriculum Perceptual Robotics) in 2012. He received his B.S. and M.S., both with honors, in Computer Science Engineering from Politecnico di Bari, Italy, in 2007 and Politecnico di Torino, Torino, Italy, in 2009, respectively. His research interests include motion planning, neurorehabilitation, human-robot and human-computer interaction, and computer vision. He is the author/coauthor of many international scientific publications and the winner of some scientific international prizes.

**Andrea Piarulli** is currently a Ph.D. student at PERCRO Laboratory, Scuola Superiore Sant'Anna, Pisa, Italy. He received his master's degree in Electronic Engineering from the University of Pisa. His research interests include biomedical signal analysis (EEG, ECG, and MR), applied statistics, and electronics.

**Emanuele Ruffaldi** (Eng. Ph.D.) is an Assistant Professor in Applied Mechanics at the PERCRO Lab, Scuola Superiore S. Anna, Pisa, Italy. He has obtained his Ph.D. on Perceptual Robotics in 2006 from Scuola Superiore S.Anna, discussing a thesis on perceptually inspired haptic algorithms. His research interests are in the application of machine learning for modeling behavior of humans with a focus on skill training in virtual environments and integration with robot learning. He participated in the research activities of the European project SKILLS IP as a work package leader for the rowing demonstrator. Inside SKILLS, he has focused his research on two aspects: on one side, the issues of combining the modeling of the skill with the creation of appropriate training feedback and the other on the digital representation of human skills combined with issues of data management. He has contributed to previous EU projects such as ENACTIVE NoE, Decision in Motion, HAPTEX, CREATE, and PureForm. Emanuele has been the author of more than 70 papers published as proceedings of scientific conferences and journals. He has been the general chair of the ENACTIVE08 and the Program Chair of IEEE RO-MAN 2010, and he is currently the Dissemination Chair of IEEE TC on haptics.

**Franco Tecchia** is an Assistant Professor in Computer Science with a primary specialization on Software Engineering and Computer Graphics at Scuola Superiore Sant'Anna, Italy. He received his Ph.D. in Computer Science from the University College London, UK, where he was also a Research Fellow in VE and Computer Graphics. His research activities focus on the design and development of complex VR systems. At Scuola Sant'Anna, he is the Head of the Computer Graphics and Virtual Environment area at PERCRO Laboratory, working in the context of national and international research activities including the European Projects VIRTUAL, PURE FORM, CREATE, PRESENCCIA, SKILLS, VERE, BEAMING, and the NoE ENACTIVE and INTUITION.

**Massimo Bergamasco** received his M.Sc. in Mechanical Engineering from the University of Pisa in 1985. He has been a Full Professor of Applied Mechanics in the Faculty of Engineering, Scuola Superiore Sant'Anna, Pisa, Italy, since 2006, where he became an Assistant Professor and Associate Professor in 1987 and 1998, respectively. He is also the Responsible for the PERCRO Laboratory, which he founded in 1994. During his research activities, he has been the Responsible for many national and international projects in the field of design and realization of virtual environment systems. He was the Scientific Responsible for several European projects (EU IP Skills and Enactive NoE). He is and has been a Member of the Editorial Board of several scientific journals and a Permanent Member of the steering Committee of several international conferences. His scientific activity includes more than 200 scientific papers published in journals and/or international conference proceedings.