

The Role of Camera Convergence in Stereoscopic Video See-through Augmented Reality Displays

Fabrizio Cutolo

University of Pisa

Dep. of Information Engineer & EndoCAS Center

Via Caruso 16, 56122, Pisa

fabrizio.cutolo@endocas.unipi.it

Vincenzo Ferrari

University of Pisa

Dep. of Information Engineer & EndoCAS Center

Via Caruso 16, 56122, Pisa

vincenzo.ferrari@unipi.it

Abstract—In the realm of wearable augmented reality (AR) systems, stereoscopic video see-through displays raise issues related to the user’s perception of the three-dimensional space. This paper seeks to put forward few considerations regarding the perceptual artefacts common to standard stereoscopic video see-through displays with fixed camera convergence. Among the possible perceptual artefacts, the most significant one relates to diplopia arising from reduced stereo overlaps and too large screen disparities. Two state-of-the-art solutions are reviewed. The first one suggests a dynamic change, via software, of the virtual camera convergence, whereas the second one suggests a matched hardware/software solution based on a series of predefined focus/vergence configurations. Potentialities and limits of both the solutions are outlined so as to provide the AR community a yardstick for developing new stereoscopic video see-through systems suitable for different working distances.

Keywords—Augmented reality and visualization; stereoscopic display; stereo overlap; video see-through

I. INTRODUCTION

Human eyes are placed frontally about 6-7 cm apart (interpupillary distance = 6-7 cm) so they both perceive the same scene but from slightly different viewpoints (i.e. with an horizontal parallax). In other words, through the crystalline lenses, the two retinas receive slightly different views of the same three-dimensional (3D) scene. The positional differences between the two retinal images are defined as binocular or retinal disparities. Specialized neurons (binocular neurons) in the visual cortex of the brain, process those disparities to generate a sort of depth map of the observed scene. We commonly refer to this mechanism as stereopsis from the Greek words $\sigma\tau\epsilon\rho\epsilon\omicron$ (stereo-meaning solid and $\omicron\psi\iota\varsigma$ (opsis meaning appearance, sight, and we define this depth cue as binocular parallax. The goal of stereoscopic 3D displays is hence to create an illusion of depth perception by providing consistent binocular disparity information in the recorded images delivered to the left and right eyes [1].

Depth cueing through stereoscopy is an essential feature of head-mounted displays (HMDs) for augmented reality (AR). Most of the AR HMDs fall into two categories according to the see-through paradigm they implement: video see-through HMDs and optical see-through HMDs. Typically, in optical see-through systems, the user’s direct view of the real world is augmented with the projection of virtual information on a beam combiner and then into the user’s line of sight [2].

Differently, in video see-through systems the virtual content is merged with camera images captured by a stereo camera rig rigidly fixed on the 3D display.

The pixel-wise video mixing technology that underpins the video see-through paradigm can offer high geometric coherence between virtual and real content. Nevertheless, the industrial pioneers, as well as the early adopters of AR technology properly considered the camera-mediated view typical of video see-through devices as drastically affecting the quality of the visual perception and experience of the real world [2], [3].

In stereoscopic video see-through HMDs, double vision (diplopia) for the user may arise if the fixation point, determined by the intersection of the optical axis of the stereo camera pair, leads to reduced stereo overlap between the two images delivered to the eyes through the HMD. This stereo conflict happens because a large portion of the scene is not represented on both images (the left part of the left view is not represented in the right view and vice versa), and therefore the visual cortex is not able to fuse the two views. This perceptual conflict is due to the fixed configuration of the stereo setting and it heavily constraints the working distance on where the stereoscopic video see-through HMD yields a comfortable visual result. Two possible solutions for coping with this limit are here reviewed and brought as example. This work is inspired by the need for assessing the role of camera convergence in stereoscopic video see-through AR displays and establishing a yardstick for designing new solutions that allow users of such systems to work comfortably at close distance ranges.

II. BINOCULAR PARALLAX, CONVERGENCE AND ACCOMMODATION IN HUMAN VISUAL SYSTEM

Binocular parallax, horopter and Panum’s area

Binocular parallax is the most effective relative depth cue at close distances, namely in an individual’s personal space or at arm’s length [4], [5]. The equation that links the theoretical depth discrimination threshold (i.e. human depth resolution) ΔZ_h to the angular retinal disparity $\Delta\alpha$ can be trivially derived from geometrical relations (see Fig. 1). In particular, for a given a fixation point, associates the convergence angle of the eyes θ to the absolute depth of the fixation point (Z) and to the interpupillary distance I :

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731974-VOSTARS project www.vostars.eu

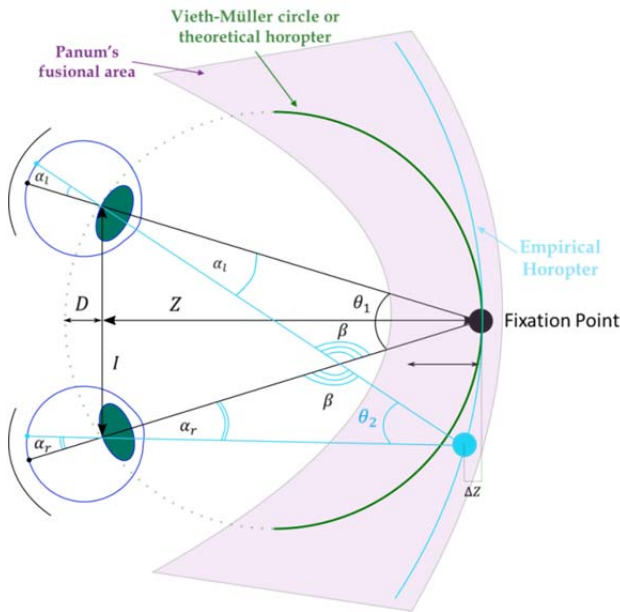


Fig. 1. Binocular horizontal disparity and geometry of the stereo perception in human visual system. Human depth resolution can be approximately expressed as a function of the distance between fixation point and observer Z , and of the interpupillary distance I .

$$\theta = 2 \tan^{-1}(I/2Z) \quad (1)$$

The human depth resolution formula [6] is the result of the derivative $dZ/d\theta$:

$$\Delta Z_h = -\frac{Z^2}{I} \left(1 + \frac{I^2}{4Z^2}\right) \Delta\theta \quad (2)$$

Whose simplified form is:

$$\Delta Z_h \approx -\frac{Z^2 * \Delta\theta}{I} \quad (3)$$

It is worth noting, from trivial geometrical considerations, that taking into account the angle α :

$$\begin{aligned} \Delta\theta &= \theta_1 - \theta_2 = (\pi - \beta - \alpha_l) \\ &\quad - (\pi - \beta - \alpha_r) = \alpha_r - \alpha_l \\ &= -\Delta\alpha \end{aligned} \quad (4)$$

Hence,

$$\Delta Z_h \approx \frac{Z^2 * \Delta\alpha}{I} \quad (5)$$

So, and as explained with more details in [7], (5) associates the depth resolution of the human visual system to the retinal angular difference $\Delta\alpha$, to the interpupillary distance I , and to the distance between fixation point and observer Z [8].

When retinal disparities are too high, they produce diplopia that is, itself, a depth cue for the human visual system. The fixation point has 0 retinal disparity, as well as all the corresponding retinal points. The circle formed passing through the fixation point F and the two nodal points of the eyes O_1 and O_2 , is defined Vieth-Müller circle or theoretical horopter (from the Greek words *ὄρος* meaning boundary, + *ὀπίτηρ* meaning observer). Any points belonging to this circle stimulate geometric corresponding points on the retinae of the two eyes, hence they bring 0 disparity exactly as F ($\alpha_l = \alpha_r$ in Fig. 1 for construction).

In reality, the empirical horopter for any observer is less convex than a circle and the Hering-Hillebrand deviation gives a measure of the deviation of the empirical horopter from the Vieth-Müller circle [8], [9]. According to such deviation (referred to as H), the relation that fits the empirical measurements on the real horopter based on the disparity between the retinal angles α_l and α_r is (see Fig. 1):

$$H = \cot \alpha_l - R \cot \alpha_r \quad (6)$$

that leads to [10]:

$$\alpha_r = \tan^{-1} \frac{R \tan \alpha_l}{1 - H \tan \alpha_l} \quad (7)$$

with R = relative magnification between right eye and left eye. Thus the empirical deviation from the theoretical horopter is measured by disparity D between α_l and α_r and it is modeled as follows:

$$D = \alpha_l - \tan^{-1} \frac{R \tan \alpha_l}{1 - H \tan \alpha_l} - D_0 \quad (8)$$

with D_0 encapsulating the effect of the Helmholtz shear. The conditions for the theoretical horopter are: $D_0 = 0$, $H \approx 0$, and $R \approx 1$. In that case, the empirical disparity of the points belonging to the horopter is null. Within a special visual space around the fixation point (except the points belonging to the horopter), the corresponding points on the retina produce disparities whose values are processed by the visual cortex of the brain to estimate depth relations in such area around the fixation point (Fig. 2). These disparities, provided they are sufficiently small, can be either positive, if the points under observation are behind the horopter, or negative if the points are in front of the horopter. This visual space is called Panum's fusional area. In this area, the two stereoscopic retinal images are fused into one. Outside the Panum's area, the objects are perceived as diplopia.

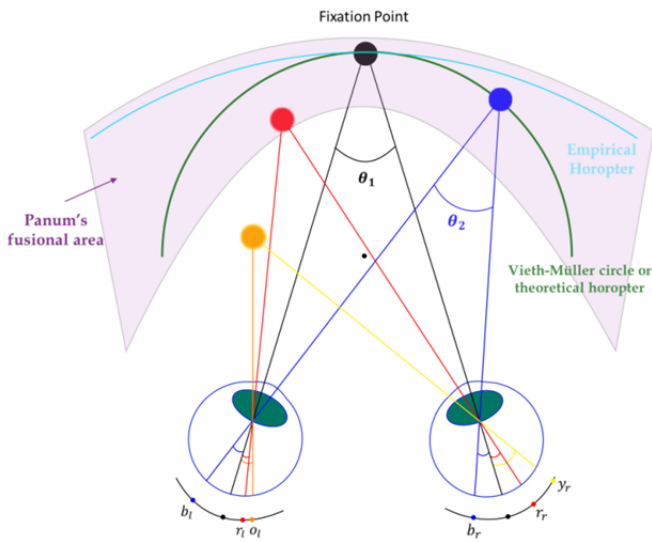


Fig. 2. At a given distance of the fixation point, any point belonging to the horopter is imaged on geometric corresponding retinal points ($\mathbf{b}_l \equiv \mathbf{b}_r$). Within a special visual space around the fixation point (other than the points belonging to the horopter), the projected points on the retina produce disparities that are mapped into depth perception around the fixation point. This visual space surrounding the horopter is called Panum's fusional area. In this area, the two stereoscopic retinal images are fused into one by the brain. Outside the Panum's area, the objects are perceived as diplopic (θ_1 cannot be fused with γ_r).

Convergence and Accommodation

In human visual system, convergence is that disjunctive movement of the eyes that allows the minimization of the targeted visual information projected on the two retinas [11]. The coordinated action of the extra-ocular muscles, when focusing on the same object, causes the mutual rotation of the optical axes of the eyes and, therefore, helps in perceiving depth/distance in combined interaction with accommodation. As well as the accommodation cue, this cue is powerful within the personal space [12].

III. STEREOSCOPIC VIDEO SEE-THROUGH

In stereoscopic video see-through HMDs, the goal is to create an illusion of depth perception by providing consistent binocular disparity information in the images delivered to the left and right eyes by the two displays of the visor.

Any stereoscopic video see-through display comprises two stages whose specifics are to be matched one another in order to provide a consistent illusion of depth perception to the viewer: acquisition stage and viewing stage [13]. The stereo rig anchored to the visor has the task of capturing the real scene (i.e. acquisition task), whereas the two internal displays of the visor have the task of delivering the stereoscopic augmented information to the viewer (i.e. visualization task).

In his work, Kyto has extensively addressed all the factors that influence depth perception through stereoscopic video see-through displays [5], [14]. In particular the authors proposed a useful theoretical comparison between human depth resolution through stereo displays and stereo camera depth resolution [15]. By carefully evaluating the results of

that analysis, we can get an idea of the requirements in terms of disparity accuracy Δd , focal length, camera sensor width S_w , and baseline for the external stereo camera pair that allow the achievement of human-like viewing conditions. Nonetheless, in most applications, a tradeoff is to be sought between accuracy in stereo depth measurements (at least comparable to that of the human visual system), and the quasi-ortho-stereoscopic depth perception through the video see-through HMD [14]; for example a changing of the baseline and/or of the focal length may lead to an improved depth resolution out of stereo triangulation, but at the expense of introducing unwanted perceptual artefacts to the viewer [13], [16].

Particularly in image guided surgery, the quality of an AR experience depends on how well the virtual content is integrated into the real world spatially, photometrically and temporally [17]. In this context, wearable AR systems offer the most ergonomic solution for those medical tasks manually performed under user's direct vision (open surgery, introduction of biopsy needle, palpation, etc.) since they intrinsically provide the user with an egocentric viewpoint of the surgical scene. They contextually integrate the surgeon's visual perception of the real scene with useful AR-based visualization modalities (derived from radiological images). Different embodiments of video see-through HMDs have been proposed in minimally invasive surgery [18], [19], laparoscopic surgery [20]-[22], orthopedic surgery [23], [24], in neurosurgery [25], and in maxillofacial surgery [26], [27]. In assessing the efficacy and reliability of such devices, the understanding of all the physiological and psychological mechanisms that underpin depth perception is of particular importance. In this regard, unreliable modalities of AR visualization can in fact bring cognitive overload and perceptual conflicts causing misinterpretation and hindering clinical decision-making [28].

A comprehensive overview of all the possible perceptual artefacts that arise in the acquisition or in the visualization stage in stereoscopic video systems is presented in [13]. Among all the possible perceptual artefacts, as anticipated, diplopia may arise if the fixation point, determined by the intersection of the optical axes of the stereo camera pair, leads to reduced stereo overlap. In the next subsection, we shall briefly describe two possible solutions for coping with this problem and we shall contextually point out the strengths and weaknesses of both the approaches. It is worth mentioning that both the solutions that we shall review were properly designed for specific medical/surgical applications in which the user is asked to interact with the augmented scene at varying working distances (however at close range), during procedures demanding for high hand-eye coordination. This task-oriented requirement increases the need for stereoscopic video see-through systems that allow sufficient stereo overlap when viewing close objects, although at odds with a desired ortho-stereoscopy [29]. Both solutions in fact feature a non-negligible eye-camera offset, so in rigorous terms, the ortho-stereoscopy of both the systems was not ensured from the very beginning and in contrast with the assertions made by Takagi et al. [30].

IV. VIDEO SEE-THROUGH HMD WITH DYNAMIC VIRTUAL CONVERGENCE

State et al. [16] proposed a software solution with dynamic control of the virtual convergence of the display frustum to allow users to work comfortably at different depths. A narrower selection of the wide-angle cameras imaging frustum is cropped dynamically as a function of a heuristic estimation of the working distance. For each eye, the modified augmented image is then delivered to the corresponding internal monitor of the video see-through HMD. No recalibration of the stereo camera rig is needed since there are no moving parts in the AR system. Nothing is said regarding the camera focuses, so we assume that both were kept fixed for each working distance.

Their solution offers two possible methods for managing the geometry of the display frustum: sheared frustums and rotated frustums. Sheared display frustum does not add unwanted vertical disparity to the stereo images which could bring geometrical artefacts in perceiving depth relations (e.g. keystone depth plane distortion [13], [30]). Unfortunately sheared frustums, especially at close working distances, bring a more pronounced disparity-vergence conflict if compared to the rotated solution. Rotated frustums, albeit introducing vertical disparity between corresponding features in stereo images, especially at the corners, is able to more consistently simulate the physical rotation of the displays and hence to stimulate the user's eyes to converge. This fact, besides reducing the disparity-vergence conflict, provides an additional depth cue.

Another interesting aspect of their work was the real-time and automatic control of the virtual convergence as a function of an estimation of the working distance. The control of the virtual convergence was implemented through three distinct approaches: an image-based method based on the maximization of the mutual information among paired views of the virtual content. A second method, a pixel-wise inspection of the z-buffer of the stereo images is used to provide depth estimation. A third approach, in which a depth estimate is computed by working on geometry data instead than on a pixel-wise mapping of the rendered images. The third method can work on the current frame, before being rendered on the display, and therefore provide an instantaneous managing of the convergence. On the contrary, the first two methods are suitable for predicting convergence for the subsequent frame.

The main drawback of this technique is that it drastically reduces the resolution of the images acquired by the stereo cameras. Further, the idea of dynamically changing the virtual convergence of the stereo camera pair through a real-time estimation of the operative depth, albeit appealing, is prone to possible perceptual conflicts for the user and it may lead to incorrect depth perception or discomfort during use if not properly managed.

V. VIDEO SEE-THROUGH HMD WITH ADJUSTABLE CAMERA VERGENCE

In 2014 Ferrari et al. [31] proposed a matched hardware/software solution that entails the adjustment of the degree of convergence of the stereo camera pair established as a function of the working distance on a per-session basis.

In more details, to restore stereo overlap, and reduce image disparities well within the binocular fusional area, the degree of convergence of the stereo camera pair is made adjustable so to be adapted at different and predefined working distances. In this way the fixation point is moved closer to the observer and the visual disparities between left and right images are reduced (Fig. 3). To implement this idea, an ad hoc version of a previously presented video see-through system [32]-[35] based on a HMZ-T2 Sony HMD, was assembled (Fig. 4). The system comprises two supports equipped with adjuster screws for modifying the stereo camera vergence, and the camera focus can be coherently adjusted with the working distance thanks to motorized mechanisms.

For each set of predefined focus/vergence configuration, the intrinsic and extrinsic camera parameters associated with it are to be determined offline as a result of a one-time calibration routine, with the calibration data stored for subsequent reuse. A two-stage video-based pose estimation algorithm, allows sub-pixel registration accuracy in the augmented scene without requiring additional work to the user (i.e. no further calibrations are required). More specifically:

1) 3D localization through stereo triangulation is correctly performed solely relying on the sets of predetermined intrinsic and extrinsic calibration data associated to the specific stereo camera vergence configuration.

2) The initial estimation of the camera pose, computed in closed-form through a standard SVD-based method [36], may result intrinsically inaccurate, given the uncertainty in the estimation of the stereo camera parameters, but it sets up a good initial guess for the subsequent pose refinement step.

The proposed method was tested on a set of 3 predefined configurations of the stereo camera vergence. For each configuration, the two adjuster screws were set to move the fixation point at 30, 100 and 170 cm. Therefore, an estimation of the corresponding convergence angle was computed by substituting I and Z in (1).

For each configuration and before use, the focus of both cameras was adjusted to focus at the fixation point, hence as a function of a set of expected working distances. Accordingly, three sets of intrinsic and extrinsic camera parameters had to be estimated and stored following three offline calibration routines.

However, also this solution has its own weaknesses. In rigorous terms, the converging of the optical axes of the stereo

cameras alone, without a simultaneous and coupled converging of the optical axes of the displays, goes against a desired ortho-stereoscopy, and therefore it might cause perceptual artifacts [30].

As properly hypothesized by State et al. and Ferrari et al. [31], their experience suggests that the distortion of the perceived 3D space is not too severe to hinder the correct use of the stereoscopic video see-through display; obviously this holds true if we consider the really constrained and task-oriented working distances for which the preset vergence configurations were set. Furthermore, and unlike the solution proposed by State et al., the system has “moving parts” and therefore needs for regular recalibrations to cope with the degradation of the stereo calibration over time.

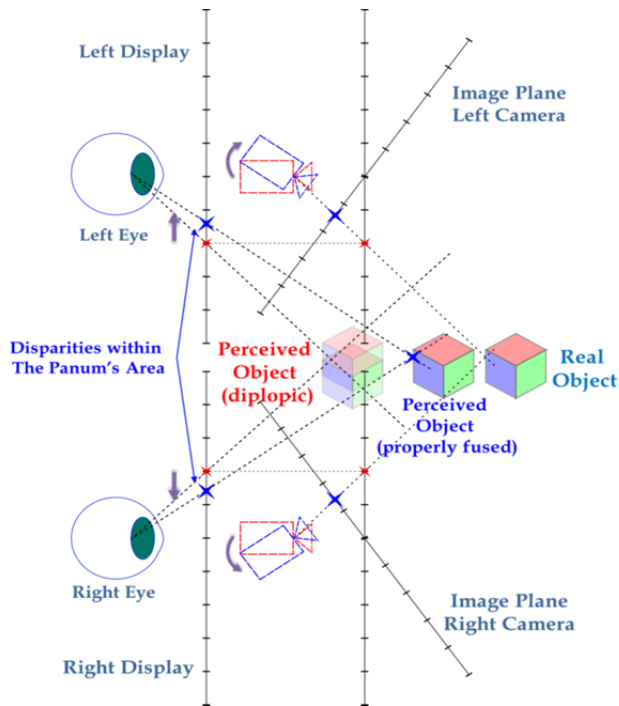


Fig. 3. Adjustment of the degree of convergence of the stereo camera pair to increase stereo overlaps and reduce visual disparities between images on the left and right display of the HMD. The disparity between the blue crosses on the two displays is lower than the disparity between the red crosses.



Fig. 4. HMD prototype embedded with adjuster screws for stereo camera vergence control.

VI. DISCUSSION AND CONCLUSION

The use of stereoscopic video see-through HMDs in case of AR assistance during manually performed tasks (e.g. surgery) in which the user is asked to interact with the scene at close working distances, and during procedures demanding for high hand-eye coordination (e.g. medical/surgical applications), is heavily hindered by the occurrence of diplopia.

Typically, in these systems, cameras and displays are preset at a fixed convergence angle on the basis of assumptions made on the average working distance. Thereby, in these systems, stereo conflicts may arise if the fixation point, determined by the intersection of the optical axis of the stereo camera pair, leads to reduced stereo overlap between the two images delivered to the eyes through the HMD. This occurrence heavily limits the actual distance on where the stereoscopic video see-through HMD can yield a comfortable visual result. In this paper, two possible solutions for coping with this limit were reviewed and brought as example, one purely software and the other matched hardware/software.

The solution suggested by State et al. features a dynamical change of the virtual convergence of the stereo camera pair based on a real-time estimation of the operative depth. This solution does not comprise any moving parts within the HMD, hence is theoretically calibration-free but this is at the expense of a drastic reduction of the resolution of the images acquired by the stereo cameras. Furthermore, having the camera focuses fixed (also to avoid further calibrations) may produce blurred images if the system were used at working distances far from the focus.

The solution by Ferrari et al. entails the physical adjustment of the degree of convergence and of the associated camera focuses of the stereo camera pair, established as a function of the working distance on a per-session basis. In this case, a set of calibration routines (intrinsic and extrinsic) has to be performed before but, in this way, all the camera frustum is viable and the camera images are properly on focus for each working distance established.

Based on our experience and as clearly stated in both the reviewed works, make the augmented scene as stereo-perceivable at close distances is key for those applications in which the user is asked to interact with the augmented scene at close working distances (i.e. in the personal space, at arm's length) and during procedures that demands for high dexterity as a surgical tasks. It is our conviction that this requirement, once achieved, may fully compensate for the increased distortion of the perceived 3D space, due to the dynamic change of the convergence of the cameras (be them virtual or real) [37]. In our opinion, from a functional standpoint, resolving diplopia has a higher priority than dealing with the perceptual artefacts caused to the non-rigorous orthoscopy of the stereoscopic display. Nevertheless, we also believe that the technology should move towards the implementation of parallax-free video see-through HMDs that entails an automatic and coupled management of the display and camera convergence (as a function of a real-time depth estimation algorithm).

REFERENCES

- [1] S. A. Benton, *Selected Papers on Three-dimensional Displays*: SPIE Optical Engineering Press, 2001.
- [2] J. P. Rolland, R. L. Holloway, and H. Fuchs, "A Comparison of Optical and Video See-through Head-Mounted Displays," *Telemanipulator and Telepresence Technologies*, vol. 2351, pp. 293-307, 1994.
- [3] R. T. Azuma, "A survey of augmented reality," *Presence-Teleoperators and Virtual Environments*, vol. 6, pp. 355-385, Aug 1997.
- [4] J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth."
- [5] M. Kyto, A. Makinen, T. Tossavainen, and P. Oittinen, "Stereoscopic depth perception in video see-through augmented reality within action space," *Journal of Electronic Imaging*, vol. 23, Jan-Feb 2014.
- [6] S. Reichelt, R. Haussler, G. Futterer, and N. Leister, "Depth cues in human visual perception and their realization in 3D displays," *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics Iv*, vol. 7690, 2010.
- [7] P. Howard, *Perceiving in depth, volume 1: basic mechanisms*: Oxford University Press, 2012.
- [8] M. Banks, "The vertical horopter is not adaptable and is not adaptive for viewing along the ground," *I-Perception*, vol. 2, 2011.
- [9] K. N. Ogle, "Precision and validity of stereoscopic depth perception from double images," *J Opt Soc Am*, vol. 43, pp. 907-13, Oct 1953.
- [10] K. M. Schreiber, J. M. Hillis, H. R. Filippini, C. M. Schor, and M. S. Banks, "The surface of the empirical horopter," *Journal of Vision*, vol. 8, 2008.
- [11] S. Lamantia, D. Purves, G. J. Augustine, D. Fitzpatrick, and W. C. Hall, *NeuroScience: Sinauer Associates Incorporated*, 2011.
- [12] M. J. Tové, *An Introduction to the Visual System*: Cambridge University Press, 1996.
- [13] L. Xing, "Towards Reliable Stereoscopic 3D Quality Evaluation: Subjective Assessment and Objective Metrics," 2013.
- [14] M. Kytö, "Depth Perception of Augmented and Natural Scenes through Stereoscopic Systems," 2014.
- [15] M. Kyto, M. Nuutinen, and P. Oittinen, "Method for measuring stereo camera depth accuracy based on stereoscopic vision," *Three-Dimensional Imaging, Interaction, and Measurement*, vol. 7864, 2011.
- [16] State, J. Ackerman, G. Hirota, J. Lee, and H. Fuchs, "Dynamic virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo overlap throughout a close-range work space," *Ieee and Acm International Symposium on Augmented Reality, Proceedings*, pp. 137-146, 2001.
- [17] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced Medical Displays: A Literature Review of Augmented Reality," *Journal of Display Technology*, vol. 4, pp. 451-467, Dec 2008.
- [18] State, M. A. Livingston, W. F. Garrett, G. Hirota, M. C. Whitton, E. D. Pisano, et al., "Technologies for augmented reality systems: realizing ultrasound-guided needle biopsies," presented at the Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996.
- [19] Bichlmeier, E. Euler, T. Blum, and N. Navab, "Evaluation of the virtual mirror as a navigational aid for augmented reality driven minimally invasive procedures," in *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, 2010, pp. 91-97.
- [20] H. Fuchs, M. A. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, et al., "Augmented reality visualization for laparoscopic surgery," *Medical Image Computing and Computer-Assisted Intervention - Miccai'98*, vol. 1496, pp. 934-943, 1998.
- [21] V. Ferrari, G. Megali, E. Troia, A. Pietrabissa, and F. Mosca, "A 3-D mixed-reality system for stereoscopic visualization of medical dataset," *IEEE Trans Biomed Eng*, vol. 56, pp. 2627-33, Nov 2009.
- [22] Pietrabissa, L. Morelli, M. Ferrari, A. Peri, V. Ferrari, A. Moglia, et al., "Mixed reality for robotic treatment of a splenic artery aneurysm," *Surg Endosc*, vol. 24, p. 1204, May 2010.
- [23] Y. Abe, S. Sato, K. Kato, T. Hyakumachi, Y. Yanagibashi, M. Ito, et al., "A novel 3D guidance system using augmented reality for percutaneous vertebroplasty," *Journal of Neurosurgery-Spine*, vol. 19, pp. 492-501, Oct 2013.
- [24] F. Cutolo, S. Carli, P. D. Parchi, L. Canalini, M. Ferrari, M. Lisanti, et al., "AR interaction paradigm for closed reduction of long-bone fractures via external fixation," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 2016, pp. 305-306.
- [25] F. Sauer, A. Khamene, B. Bascle, S. Vogt, and G. J. Rubino, "Augmented reality visualization in iMRI operating room: System description and pre-clinical testing," *Medical Imaging 2002: Visualization, Image-Guided Procedures, and Display*, vol. 4681, pp. 446-454, 2002.
- [26] G. Badiali, V. Ferrari, F. Cutolo, C. Freschi, D. Caramella, A. Bianchi, et al., "Augmented reality as an aid in maxillofacial surgery: Validation of a wearable system allowing maxillary repositioning," *Journal of Cranio-Maxillofacial Surgery*, vol. 42, pp. 1970-1976, Dec 2014.
- [27] F. Cutolo, G. Badiali, and V. Ferrari, "Human-PnP: Ergonomic AR Interaction Paradigm for Manual Placement of Rigid Bodies," in *Augmented Environments for Computer-Assisted Interventions*, ed: Springer International Publishing, 2015, pp. 50-60.
- [28] Ware, *Information visualization: perception for design*: Elsevier, 2012.
- [29] P. Milgram and M. Kruger, "Adaptation Effects in Stereo Due to Online Changes in Camera Configuration," *Stereoscopic Displays and Applications Iii*, vol. 1669, pp. 122-134, 1992.
- [30] Takagi, S. Yamazaki, Y. Saito, and N. Taniguchi, "Development of a stereo video see-through HMD for AR systems," *Ieee and Acm International Symposium on Augmented Reality, Proceeding*, pp. 68-77, 2000.
- [31] V. Ferrari, F. Cutolo, E. M. Calabro, and M. Ferrari, "HMD Video See Through AR with Unfixed Cameras Vergence," *2014 Ieee International Symposium on Mixed and Augmented Reality (Ismar) - Science and Technology*, pp. 265-266, 2014.
- [32] F. Cutolo, P. D. Parchi, and V. Ferrari, "Video See Through AR Head-Mounted Display for Medical Procedures," *2014 Ieee International Symposium on Mixed and Augmented Reality (Ismar) - Science and Technology*, pp. 393-396, 2014.
- [33] F. Cutolo, C. Freschi, S. Mascioli, P. Parchi, M. Ferrari, and V. Ferrari, "Robust and Accurate Algorithm for Wearable Stereoscopic Augmented Reality with Three Indistinguishable Markers," *Electronics*, vol. 5, p. 59, 2016.
- [34] F. Cutolo, M. Siesto, S. Mascioli, C. Freschi, M. Ferrari, and V. Ferrari, "Configurable Software Framework for 2D/3D Video See-Through Displays in Medical Applications," in *Augmented Reality, Virtual Reality, and Computer Graphics: Third International Conference, AVR 2016, Lecce, Italy, June 15-18, 2016. Proceedings, Part II*, L. T. De Paolis and A. Mongelli, Eds., ed Cham: Springer International Publishing, 2016, pp. 30-42.
- [35] S. Parrini, F. Cutolo, C. Freschi, M. Ferrari, and V. Ferrari, "Augmented reality system for freehand guide of magnetic endovascular devices," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 490-493.
- [36] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Squares Fitting of 2 3-D Point Sets," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 699-700, Sep 1987.
- [37] K. Matsunaga, T. Yamamoto, K. Shidoji, and Y. Matsuki, "The effect of the ratio difference of overlapped areas of stereoscopic images on each eye in a teleoperation," *Stereoscopic Displays and Virtual Reality Systems Vii*, vol. 3957, pp. 236-243, 2000.