

Human-PnP: Ergonomic AR Interaction Paradigm for Manual Placement of Rigid Bodies

Fabrizio Cutolo¹(✉), Giovanni Badiali³, and Vincenzo Ferrari^{1,2}

¹ Department of Translational Research and New Technologies in Medicine and Surgery, EndoCAS Center, University of Pisa, Pisa, Italy

fabrizio.cutolo@endocas.org

² Information Engineering Department, University of Pisa, Pisa, Italy

³ School in Surgical Sciences, University of Bologna, Bologna, Italy

Abstract. The human perception of the three-dimensional world is influenced by the mutual integration of physiological and psychological depth cues, whose complexity is still an unresolved issue per se. Even more so if we wish to mimic the perceptive efficiency of the human visual system within augmented reality (AR) based surgical navigation systems. In this work we present a novel and ergonomic AR interaction paradigm that aids the manual placement of a non-tracked rigid body in space by manually minimizing the reprojection residuals between a set of corresponding virtual and real feature points. Our paradigm draws its inspiration from the general problem of estimating camera pose from a set of n -correspondences, i.e. perspective- n -point problem. In a recent work, positive results were achieved in terms of geometric error by applying the proposed strategy on the validation of a wearable AR system to aid manual maxillary repositioning.

Keywords: Augmented reality and visualization · Computer assisted intervention · Interventional imaging

1 Introduction

In the context of image-guided surgery (IGS), augmented reality (AR) technology represents a promising integration between navigational surgery and virtual planning.

In 2012 Kersten-Oertel et al. [1] proposed a taxonomy of mixed reality visualization systems in IGS and defined the three major components based on which they then presented a systematic overview of the trends and solutions adopted in the field [2]. The acronym for the taxonomy (DVV) derives from its three key components: Data type, Visualization Processing and View. According to the taxonomy, for classifying and assessing the efficacy of a new AR system for IGS, we must focus our attention on the particular surgical scenario in which the visualization system aim to be integrated. The surgical scenario affects each of the three DVV factors, namely the type of data that should be displayed at a specific surgical step, the visualization processing technique implemented to provide the best pictorial representation of the augmented scene

and how and where the output of the visualization processing should be presented to the end-user.

Several visualization processing techniques have been adopted to allow a more immersive viewing experience for the surgeon and a more precise definition of the spatial relationships between real scene and visually processed data along the three dimensions. The human visual system exploits several physiological and psychological cues to deal with the ill-posed inverse problem of understanding a three-dimensional scene from one retinal image. However, monocular and binocular cues are not always sufficient to infer the spatial relationships between objects in the three-dimensional scene. Therefore, a full comprehension of the mechanisms underpinning depth perception is not a completely resolved issue per se in a real scene and it results even more complex within an augmented scene [3]. In this regard, among the suggested visualization processing techniques, researchers have tried to improve the perceptive efficiency by modeling and contextually rendering the virtual content in a photo-realistic manner, and/or by using pixel-wise transparency maps and “virtual windows” [4] to recreate occlusions and motion parallax cues. Some of the proposed techniques for enhancing depth perception comprise high-fidelity texturing [5] or colour coding methods, whereas others consist in lighting and shading cues and/or on the adoption of an interactive “virtual mirror” [6, 7]. Alternatively, depth perception can be improved by relying on standard stereopsis and two-view displays or on more complex full parallax multi-view displays. In any case, to the best of our knowledge, hitherto there are no visualization processing techniques that provide the user with useful information able to improve the postoperative outcome for those specific surgical tasks that involve the accurate manual placement of rigid anatomies in space.

Many surgical procedures in the field of orthopedic surgery or maxillofacial surgery, involve the task of reducing displacements or correcting abnormalities between rigid anatomical structures, i.e. bones, on the basis of a pre-operative planning. The direct tracking of all the rigid anatomies involved in the procedure would yield a measure of the six-degrees-of-freedom displacements between them and it would aid the correct performance of the surgical task, yet it is not always feasible for technical and logistic reasons. In case of single object tracking, the pointer of a standard surgical navigator can be employed by the surgeon to compare the final positions of clearly detectable reference points, over the repositioned anatomy, with those of their counterparts from the surgical planning. Nevertheless, this approach does not allow the assessment of all of the six-degrees-of-freedom at the same time.

AR seems the optimal solution to aid this kind of surgical tasks. Yet, the traditional AR interaction technique featuring the superimposition of a semi-transparent virtual replica of the rigid anatomy in a position and orientation (pose) defined during planning, is not very effective in aiding the surgeon in the correct performance of those procedure. In this regard, it is more beneficial and intuitive for the surgeon to deal with task-oriented visualization techniques, more than with complex reproductions of the virtual anatomies through photorealistic rendering, transparencies and/or virtual windows.

The goal of this work is to present a novel and ergonomic AR interaction paradigm based on a simple visualization processing technique that aims at aiding the accurate manual placement of a non-tracked rigid object in space. Our strategy relies on the tracking of a single object in the scene (e.g. the patient’s head), namely on the real-time

estimation of the geometric relation between a scene reference system (SRS) and the camera reference system (CRS), e.g. performed by means of a video based registration approach. In this scenario, the AR guide aids the surgeon in placing other non-tracked rigid bodies (e.g. bones fragments) at a planned pose relative to the CRS. Our paradigm draws its inspiration from the general problem of estimating camera pose from a set of n -correspondences, i.e. perspective- n -point problem. The key idea is that manually minimizing the distance, in the image plane, between a set of corresponding real and virtual feature points is sufficient to aid the accurate placement of a non-tracked rigid body in space.

2 Methods

Perspective- n -Point Problem. The task of estimating the pose of a camera with respect to a scene object given its intrinsic parameters and a set of n world-to-image point correspondences is known as the Perspective- n -Point (PnP) problem in computer vision and exterior orientation or space resection problem in photogrammetry.

This inverse problem concerns many fields of applications (structure from motion, robotics, augmented reality, etc.) and it was first formally introduced in the computer vision community by Fishler and Bolles in 1981 [8], albeit already used in the photogrammetry community before then. According to Fishler and Bolles the PnP problem can be defined as follows (distance-based definition):

Given the relative spatial locations of n control points $\mathbf{P}_i, i = 1, \dots, n$, and given the angle to every pair of these points from an additional point called the center of perspective \mathbf{C} , find the lengths $D_i = |\mathbf{CP}_i|$ of the line segments joining \mathbf{C} to each of the control points.

The constraint equations are:

$$D_i^2 + D_j^2 - 2D_i D_j \cos \theta_{ij} = d_{ij}^2, i \neq j \quad (1)$$

Where $D_i = |\mathbf{CP}_i|$, $D_j = |\mathbf{CP}_j|$ are the unknown variables, $\theta_{ij} = \widehat{\mathbf{P}_i \mathbf{C} \mathbf{P}_j}$ and $d_{ij} = |\mathbf{P}_i \mathbf{P}_j|$ are the known entries (Fig. 1). In computer vision θ_{ij} are determined finding the correspondences between world-to-image points and knowing the intrinsic camera parameters, while d_{ij} are established by the control points.

Following this definition, once each distance D_i is computed, the position of the points \mathbf{P}_i can be expressed in the CRS. Therefore, being the position of each \mathbf{P}_i in the SRS known, the problem of estimating camera pose with respect to the SRS is reduced to a standard absolute orientation problem whose solution can be found in closed-form fashion through quaternions [9] or singular value decomposition (SVD) [10].

The same problem is also known under the transformation-based definition [11] which can be formalized as:

$$\lambda_i \widehat{\mathbf{P}}_i = [\mathbf{K} | 0] \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \widehat{\mathbf{P}}_i, i = 1, \dots, n \quad (2)$$

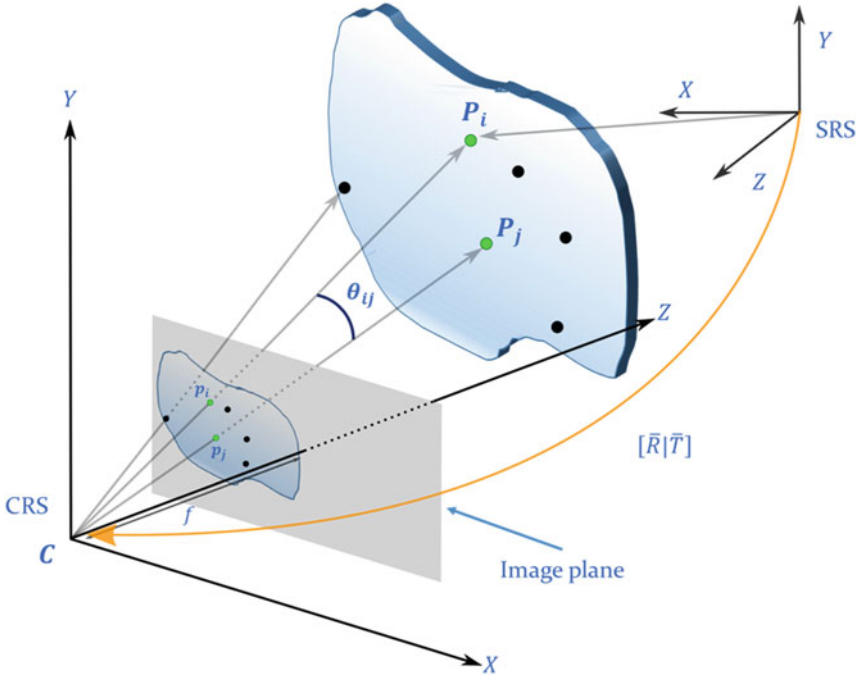


Fig. 1. Geometry of the PnP problem.

Where the scene and image points \hat{P}_i and \hat{p}_i are represented in homogeneous coordinates and the equation is up to a scale factor λ_i . Hence, according to this definition, the PnP problem aims at determining the pose (in terms of a rotation matrix R and a translation vector T) given a set of n world-to-image correspondences and known the intrinsic camera parameters encapsulated by the matrix K .

The PnP problem has been extensively studied by several groups, which have proposed different iterative, closed-form for solving it.

Closed-form methods [12–18], directly provide an estimation of the camera pose but they are usually less accurate and more susceptible to noise than iterative methods. Iterative non-linear optimization methods solve the PnP problem by iteratively minimizing a cost function generally related to the geometric (reprojection residuals) or algebraic error but they need a good initial guess and yield only one solution at a time [19–21]. A useful overview of the state-of-the-art methods can be found in [17] and in [22].

In terms of geometric reprojection residual, the non-linear cost function can be formulated as the sum of the squared measurement errors (d_i):

$$\begin{aligned} \bar{R}|\bar{T} &= \arg \min \sum_{i=1}^n d(\mathbf{p}_i, \hat{\mathbf{p}}_i)^2 \\ &= \arg \min \sum_{i=1}^n \|\mathbf{p}_i - \hat{\mathbf{p}}_i(K, \hat{R}, \hat{T}, P_i)\|^2 \end{aligned} \quad (3)$$

Where p_i are the measured image points, and \hat{p}_i are the calculated projections of the corresponding control points as a function of K, \hat{R}, \hat{T} .

The other important research direction on the PnP problem is the study of the multi-solution phenomenon of the PnP problem [23], principally when $n = 3$ (P3P) [24, 25], being three the smallest subset of control points that yields a finite number of solutions. P3P problem yields at most four solutions which can be disambiguated using a fourth point, and it is the most studied case since it can be used as first step to reduce the complexity of the computation of a PnP problem, e.g. in a RANSAC scheme by removing the outliers.

AR Video-Based Camera Registration. Regardless of the method adopted for solving the PnP problem, an immediate application of the PnP problem is to locate the pose of a calibrated camera with respect to an object, given the 3D position of a set of n control points rigidly constrained to the object and the 2D position of their correspondent projections onto the image plane.

For a correct registration of computer-generated elements to the real scene in AR-based surgical navigation systems, the image formation process of the virtual camera must perfectly mimic the real camera one. In mostly all the AR applications the estimation of the intrinsic camera parameters is the result of an off-line calibration process whereas the extrinsic camera parameters are determined online, e.g. solving a PnP problem in real-time. This video-based camera registration method suggested us the implementation of an ergonomic AR interaction paradigm for positioning and orienting a non-tracked rigid object in space.

Human- PnP . As written in the introduction, many surgical procedures in the field of orthopedic surgery or maxillofacial surgery, involve the task of manually placing rigid anatomies on the basis of preoperative planning. In that case, let us assume that we can rely on a robust and accurate registration of the surgical planning onto the real scene, by means of the tracking of at least one rigid body (e.g. the head). The six-degrees-of-freedom pose of an additional and non-tracked rigid anatomy in relation to the SRS, can be retrieved by physically placing it as to minimize the geometric distance, on the image plane, between a set of real and virtual feature points. For brevity, from now on, we shall refer to these structures as “tracked anatomy” for the former and “non-tracked anatomy” for the latter, while the proposed method will be referred to as the human-perspective- n -point problem (h PnP).

From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the PnP problem is formulated. As mentioned in the previous section, the main goal of the PnP problem is to infer useful information on the real 3D scene, based on 2D observations of it. In an AR application, this spatial information is used to geometrically register the virtual elements onto the real scene. Thus, as a general rule and regardless of the method adopted for solving the PnP problem, a robust and accurate registration should minimize in the image plane the geometric reprojection residuals between measured and estimated projections (see Eq. 3). Similarly, the goal of our h PnP interaction paradigm is to achieve the desired placement of a non-tracked anatomy by manually minimizing the reprojection residuals

between correct/planned projections \bar{p}_i of virtual landmarks, and observed projections \hat{p}_i of real landmarks.

The correct/planned projections \bar{p}_i are rendered on the image plane according to the real-time estimation of the camera pose $[\bar{R}, \bar{T}]$ relative to the tracked anatomy reference system (SRS) and assuming the intrinsic camera parameters, encapsulated by matrix K , are determined offline, e.g. through the Zhang's method [26]. The position of each virtual landmark P_i in the SRS is established during surgical planning.

The real projections \hat{p}_i are associated with the pose, encapsulated by $[\hat{R}, \hat{T}]$, between viewing camera and non-tracked anatomy reference frame (ARS): this resulting pose varies according to the manual placement of the rigid body relative to the camera:

$$\begin{aligned} \tilde{R}|\tilde{T} &= \arg \min \sum_{i=1}^n d(\bar{p}_i, \hat{p}_i)^2 \\ &= \arg \min \sum_{i=1}^n \left\| \bar{p}_i(K, \bar{R}, \bar{T}, P_i) - \hat{p}_i(K, \hat{R}, \hat{T}, P_i) \right\|^2 \end{aligned} \quad (4)$$

In this way, we wish to obtain $[\tilde{R}|\tilde{T}] \approx [\bar{R}|\bar{T}]$ (see Fig. 2), namely we seek to positioning and orienting the ARS as coincident with the planned and registered SRS (non-tracked anatomy reference frame \approx planning reference frame).

To implement this strategy, we add simple virtual elements (e.g. virtual asterisks, crosses, etc.) to the virtual scene during the surgical planning: one element for each of the clearly detectable physical landmarks on the rigid body. The landmarks may consist of a series of distinguishable feature points over the surface of the anatomy or rigidly

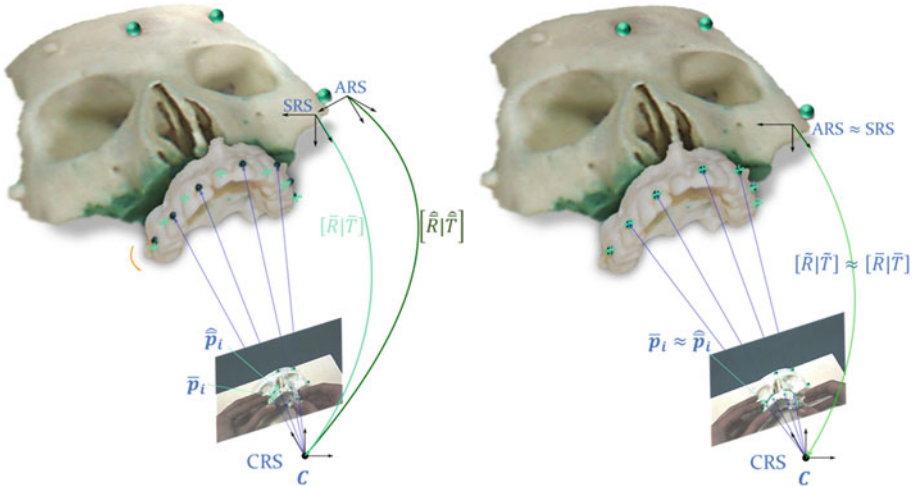


Fig. 2. Geometry of the hPnP: minimizing the reprojection residual between registered projections \bar{p}_i and real projections \hat{p}_i is sufficient to aid the accurate placement of a rigid body (the maxilla in the image) in space.

constrained to it. Under such AR guidance, the user moves the non-tracked rigid body up to obtain a perfect overlapping between real and virtual landmarks, hence manually minimizing the reprojection residuals on the image plane: $\bar{p}_i \approx \hat{p}_i \forall i$ (Fig. 3). The theoretical assumptions underpinning the PnP problem ensure that if $\bar{p}_i \approx \hat{p}_i \forall i$, the non-tracked anatomy is placed in the correct pose as planned in SRS.

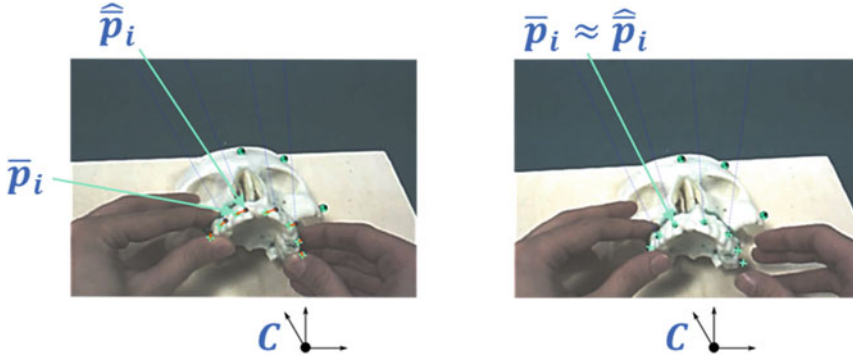


Fig. 3. Detail of the image plane with the minimization of the reprojection residuals. Here the virtual information consists of a cyan-colored asterisk for each physical landmark clearly detectable over the maxilla, e.g. coloured landmarks fixed on the brackets of the orthodontic appliance (Color figure online).

3 Results

In a recent work [27], the described strategy was applied in the validation of a wearable AR system to aid maxillary repositioning. AR system consisted of a stereoscopic video see-through head mounted display equipped with two external USB cameras placed in a quasi-orthoscopic position [28, 29]. The video see-through paradigm of the system is implemented as follows (Fig. 4): real-world views are grabbed by a pair of calibrated external cameras; the captured frames, after compensation of the radial distortion, are screened as backgrounds of the virtual scene onto the corresponding display; the virtual elements, defined during planning, are added to the real scene and observed by a pair of virtual cameras whose processes of image formation mimic those of the real cameras in terms of intrinsic and extrinsic camera parameters. Zhang’s method is used to calibrate the two cameras. The estimation of the extrinsic parameters, allowing the real-time registration of the virtual elements to real scene, is achieved through a marker-based video-registration method [29].

In the study, manual repositioning of the upper maxilla following LeFort 1 osteotomy was chosen as test procedure. The test was conducted on a CT-scanned/3D-printed replica of a cadaveric human skull. The planned pose of the maxilla, as defined during preoperative planning, acts as a guide for the surgeon during the intervention performed in-vitro. The traditional AR interaction technique, featuring the superimposition of a semi-transparent virtual replica of the maxilla, as dictated by the surgical planning, did

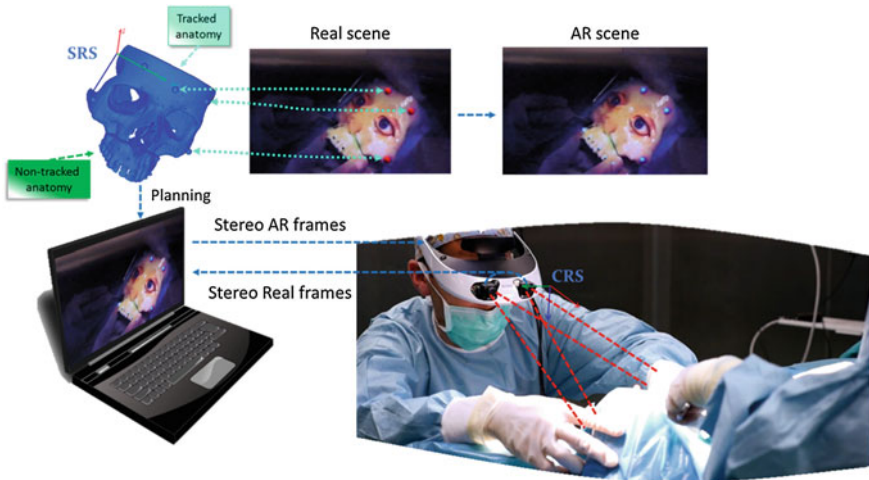


Fig. 4. Video see-through paradigm of the stereoscopic head mounted display used to aid maxillary repositioning.

not prove to be very effective in aiding the surgeon in manually repositioning the upper maxilla. This was mostly due to the surgeon's limited perception of the relative distances of objects within the AR scene owing to the presence of unnatural occlusions between the real and the virtual maxilla. Conversely, a more ergonomic form of visualization consisted in the use of an interaction paradigm which actualizes the above described hPnP approach: physical landmarks onto the maxilla and corresponding to coloured landmarks fixed on the brackets of the orthodontic appliance usually applied prior to this kind of interventions, were designated as reference markers for the AR view modality. The repositioning of the maxilla is assisted by visually aligning small virtual asterisks, drawn in positions defined during planning (relative to the SRS), with the corresponding real landmarks.

The upper surface of the maxilla (corresponding to the post-osteotomy surface) was covered with highly malleable plasticine so to be fixed to the upper skull once the surgeon performed the repositioning. The surgical accuracy was validated with the aid of an optical navigation system that recorded the coordinates of three reference points on the non-tracked maxilla after repositioning. Six surgeons and three unskilled engineers were involved in the testing, each of whom was asked to manually reposition the maxilla as dictated by three surgical plannings of variable complexity. Results in terms of linear distances between the real positions of the reference holes and the expected positions (defined during planning) were very promising: mean error was 1.70 ± 0.51 mm. The axial errors were 0.89 ± 0.54 mm on the sagittal axis, 0.60 ± 0.20 mm on the frontal axis, and 1.06 ± 0.40 mm on the cranio-caudal axis. Such results were obtained without the tracking of the maxilla but just relying on the ergonomics of the chosen AR interaction paradigm: the overlapping on the image plane between virtual feature points and real landmarks, visible over the non-tracked anatomy, proved to be sufficient to aid the accurate repositioning of the maxilla.

4 Discussion

It is important to note that the chosen AR interaction paradigm was not bound to the particular video-based tracking modality exploited in the cited study, neither to the use of a specific wearable stereoscopic system. Howbeit, the user can enhance the accuracy in object placement by checking consistency of real and virtual landmarks from different viewpoints. In this regard, the ergonomics of the proposed method may benefit from the adoption of a wearable AR system. Moreover, the choice of such instance of visualization data was, in that work, empirically inspired by the authors' endeavor of defining a modality that were ergonomic for the surgeon and that provided the smallest perceived parallax error: no further discussion was held on the theoretical hypotheses behind such interaction paradigm which are here discussed for the first time.

5 Conclusion

In this work, we proposed a novel and ergonomic AR interaction paradigm that aims at obtaining the accurate placement of a rigid body in space without the need for multiple objects tracking and/or complex visual representations of the virtual guide. From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the PnP problem in computer vision is formulated. This approach, represented by the acronym hPnP, could be of help in those tasks, also not specifically surgical, where the AR guide aims at aiding the placement of a rigid body in space. The key-principle behind this interaction paradigm can be exploited in many different AR-based navigation systems: it can be integrated with different end-products of the visualization process in terms of display technology and perception location and/or it could be realized in conjunction with various tracking modalities.

To increase robustness and applicability of the proposed AR interaction paradigm in a real clinical scenario, redundancy in choosing the set of landmarks must be granted. Further, the presence of line-of-sight occlusions caused by soft-tissues, surgeon's hands or surgical instrumentation may be restricted by conveniently selecting the position of the landmarks in relation to the surgical field.

Acknowledgments. This work was funded by the Italian Ministry of Health grant SThARS (Surgical training in identification and isolation of deformable tubular structures with hybrid Augmented Reality Simulation, 6/11/2014–5/11/2017). Grant “Ricerca finalizzata e Giovani Ricercatori 2011–2012” Young Researchers – Italian Ministry of Health.

References

1. Kersten-Oertel, M., Jannin, P., Collins, D.L.: DVV: towards a taxonomy for mixed reality visualization in image guided surgery. In: Liao, H., Edwards, P.J., Pan, X., Fan, Y., Yang, G.-Z. (eds.) MIAR 2010. LNCS, vol. 6326, pp. 334–343. Springer, Heidelberg (2010)
2. Kersten-Oertel, M., Jannin, P., Collins, D.L.: The state of the art of visualization in mixed reality image guided surgery. *Comput. Med. Imaging Graph.* **37**, 98–112 (2013)

3. Sielhorst, T., Bichlmeier, C., Heining, S.M., Navab, N.: Depth perception - a major issue in medical AR: evaluation study by twenty surgeons. In: Larsen, R., Nielsen, M., Sparring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 364–372. Springer, Heidelberg (2006)
4. Bichlmeier, C., Wimme, F., Heining, S.M., Navab, N.: Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2007, pp. 129–138 (2007)
5. Haouchine, N., Dequidt, J., Berger, M.O., Cotin, S.: Single view augmentation of 3D elastic objects. In: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 229–236 (2014)
6. Bichlmeier, C., Heining, S.M., Feuerstein, M., Navab, N.: The virtual mirror: a new interaction paradigm for augmented reality environments. *IEEE Trans. Med. Imaging* **28**, 1498–1510 (2009)
7. Bichlmeier, C., Euler, E., Blum, T., Navab, N.: Evaluation of the virtual mirror as a navigational aid for augmented reality driven minimally invasive procedures. In: 2010 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 91–97 (2010)
8. Fischler, M.A., Bolles, R.C.: Random sample consensus - a paradigm for model-fitting with applications to image-analysis and automated cartography. *Commun. ACM* **24**, 381–395 (1981)
9. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* **4**, 629–642 (1987)
10. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of 2 3-D point sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**, 699–700 (1987)
11. Wu, Y.H., Hu, Z.Y.: PnP problem revisited. *J. Math. Imaging Vis.* **24**, 131–141 (2006)
12. Haralick, R.M., Lee, C.N., Ottenberg, K., Nolle, M.: Review and analysis of solutions of the 3-point perspective pose estimation problem. *Int. J. Comput. Vis.* **13**, 331–356 (1994)
13. Quan, L., Lan, Z.D.: Linear N-point camera pose determination. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 774–780 (1999)
14. Fiore, P.D.: Efficient linear solution of exterior orientation. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 140–148 (2001)
15. Gao, X.S., Hou, X.R., Tang, J.L., Cheng, H.F.: Complete solution classification for the Perspective-Three-Point problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 930–943 (2003)
16. Ansar, A., Daniilidis, K.: Linear pose estimation from points or lines. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 578–589 (2003)
17. Lepetit, V., Moreno-Noguer, F., Fua, P.: EPnP: an accurate O(n) solution to the PnP problem. *Int. J. Comput. Vis.* **81**, 155–166 (2009)
18. Hesch, J.A., Roumeliotis, S.I.: A direct least-squares (DLS) method for PnP. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 383–390 (2011)
19. Haralick, R.M., Joo, H., Lee, C.N., Zhuang, X.H., Vaidya, V.G., Kim, M.B.: Pose estimation from corresponding point data. *IEEE Trans. Syst. Man Cybern.* **19**, 1426–1446 (1989)
20. Lowe, D.G.: Fitting parameterized 3-Dimensional models to images. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 441–450 (1991)
21. Lu, C.P., Hager, G.D., Mjolsness, E.: Fast and globally convergent pose estimation from video images. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 610–622 (2000)
22. Garro, V., Crosilla, F., Fusiello, A.: Solving the PnP problem with anisotropic orthogonal procrustes analysis. In: Second Joint 3dim/3dpvt Conference: 3d Imaging, Modeling, Processing, Visualization & Transmission (3dimpvt 2012), pp. 262–269 (2012)

23. Hu, Z.Y., Wu, F.C.: A note on the number of solutions of the noncoplanar P4P problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 550–555 (2002)
24. Zhang, C.X., Hu, Z.Y.: Why is the danger cylinder dangerous in the P3P problem. *Zidonghua Xuebao/Acta Automatica Sinica* **32**, 504–511 (2006)
25. Wang, T., Wang, Y.C., Yao, C.: Some discussion on the conditions of the unique solution of P3P problem. *IEEE ICMA 2006: Proceeding of the 2006 IEEE International Conference on Mechatronics and Automation*, vols. 1–3, Proceedings, pp. 205–210 (2006)
26. Zhang, Z.Y.: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 1330–1334 (2000)
27. Badiali, G., Ferrari, V., Cutolo, F., Freschi, C., Caramella, D., Bianchi, A., Marchetti, C.: Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning. *J. Craniomaxillofac. Surg.* **42**, 1970–1976 (2014)
28. Ferrari, V., Megali, G., Troia, E., Pietrabissa, A., Mosca, F.: A 3-D mixed-reality system for stereoscopic visualization of medical dataset. *IEEE Trans. Biomed. Eng.* **56**, 2627–2633 (2009)
29. Cutolo, F., Parchi, P.D., Ferrari, V.: Video see through AR head-mounted display for medical procedures. In: 2014 IEEE International Symposium on Mixed and augmented reality (ISMAR), pp. 393–396 (2014)