

A 3-D Mixed-Reality System for Stereoscopic Visualization of Medical Dataset

Vincenzo Ferrari*, Giuseppe Megali, Elena Troia, Andrea Pietrabissa, and Franco Mosca

Abstract—We developed a simple, light, and cheap 3-D visualization device based on mixed reality that can be used by physicians to see preoperative radiological exams in a natural way. The system allows the user to see stereoscopic “augmented images,” which are created by mixing 3-D virtual models of anatomies obtained by processing preoperative volumetric radiological images (computed tomography or MRI) with real patient live images, grabbed by means of cameras. The interface of the system consists of a head-mounted display equipped with two high-definition cameras. Cameras are mounted in correspondence of the user’s eyes and allow one to grab live images of the patient with the same point of view of the user. The system does not use any external tracker to detect movements of the user or the patient. The movements of the user’s head and the alignment of virtual patient with the real one are done using machine vision methods applied on pairs of live images. Experimental results, concerning frame rate and alignment precision between virtual and real patient, demonstrate that machine vision methods used for localization are appropriate for the specific application and that systems based on stereoscopic mixed reality are feasible and can be proficiently adopted in clinical practice.

Index Terms—Biomedical image processing, biomedical imaging, machine vision, stereo vision, virtual reality.

I. INTRODUCTION

RECENT developments in Computed Tomography (CT) and MRI equipment and the introduction of new contrast medium allow the acquisition of volumetric datasets describing human anatomy, functionality, and pathology, with high level of detail. Medical datasets are usually composed by many phases, each of which consists of hundreds of images. Radiologists build a “mental model” of the specific anatomy examining, slice by slice, the volumetric dataset. The reading of medical datasets is quite difficult. In fact, the interpretation process requires a deep knowledge concerning how the contrast medium flows inside the anatomical structures and requires geometrical skills to visualize “mentally” the 3-D development of target anatomies starting from sequences of 2-D images. The detailed information contained in a volumetric dataset are fully used during the diag-

Manuscript received January 12, 2009; revised April 27, 2009. First published July 31, 2009; current version published October 16, 2009. This work was supported by the European Community’s Seventh Framework Programme (FP7/2007-2013) under Grant 224565 (ARAKNES Project). *Asterisk indicates corresponding author.*

*V. Ferrari is with the EndoCAS Center, Department of Oncology, Transplantation and Advanced Technologies in Medicine, University of Pisa, 56124 Pisa, Italy (e-mail: vincenzo.ferrari@endocas.org).

G. Megali, E. Troia, A. Pietrabissa, and F. Mosca are with the EndoCAS Center, Department of Oncology, Transplantation and Advanced Technologies in Medicine, University of Pisa, Pisa 56124, Italy.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2009.2028013

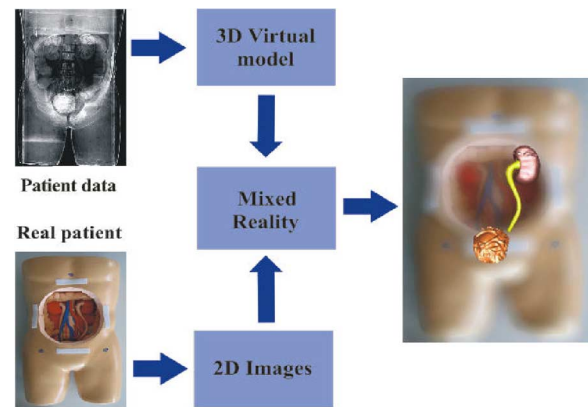


Fig. 1. Functional scheme of a mixed-reality system.

nostic phase, but are partially lost passing from the radiological department to the surgical room. In fact, generally, surgeons plan interventions just using limited information provided by the radiologist and consisting in the textual diagnosis coupled with few 2-D significant images selected from the volumetric dataset.

The application of the “computer-assisted” model to the patient workflow, consisting of computer-aided diagnosis (CAD) and computer-aided surgery (CAS) technologies, allows the optimal use of medical dataset and to overcome the aforementioned limitations of the current clinical practice. The 3-D visualization of patient specific virtual models of anatomies [1], [2], extracted from medical dataset, drastically simplifies the interpretation process of exams and provides benefits both in diagnosing and in surgical planning phases. Computer-assisted technologies allow augmentation of real views of the patient, grabbed by means of cameras, with virtual information [3]. This augmented reality, or, in general, mixed-reality techniques [4], introduces many advantages for each task where the physicians have to interact with the patient (palpation, introduction of biopsy needle, catheterization, intervention, etc.).

We developed a simple, light, and cheap 3-D visualization system based on mixed reality usable by physicians in their department or ambulatory to see preoperative radiological exams. The system allows to see “augmented images,” which are created by mixing virtual preoperative information obtained by processing radiological images (CT or MRI) with real patient live images (see Fig. 1) in a natural way.

In a previous work [5], we implemented a monoscopic mixed-reality viewer, running on an additional monitor in the operative room, which overlays virtual models of anatomies to laparoscopic camera images. Experimental validation of this viewer demonstrates that, although it improves orientation and greatly

enhances the physician understanding of the operative scenario, it introduces limitations in perceiving distance along the view axis [6]. Depth perception can be drastically increased using head-mounted stereoscopic devices [7] that allow one to evaluate object depth dislocation, like in the natural binocular view. The use of localized head-mounted display (HMD) allows one to see a synthetic scene from a point of view aligned with the real user's point of view (tracked in real time).

For mixed-reality implementation, the video see-through approach, based on the acquisition of real images by means of external cameras, is preferable to the optic see-through approach that projects virtual information on semitransparent glasses. This is due to the fact that tracking of eye movements, strictly required for the optical see-through approach, is very difficult to perform with sufficient precision [8], [9]. On the contrary, head tracking, required for video see-through approach, can be performed with high precision using the external localizer based on different technologies [10], [11].

The adoption of external localization systems introduces a lot of problems concerning system setup, large footprint, and necessities of frequent calibrations. Further, the high cost of commercial trackers limit the diffusion on a large scale of mixed-reality technologies.

We implemented a head-mounted stereoscopic video see-through system that does not require the use of an external localizer to track head movements. Our system implements mixed-reality aligning in real-time virtual and real scene just using geometric information extracted by segmenting colored markers, attached on the patient's skin, directly from couples of camera images [12]. The system has been tested and results, in terms of alignment accuracy and refresh rate, demonstrate that machine vision methods are appropriate for real-time localization and that mixed-reality systems can be proficiently adopted in clinical practice.

II. MATERIALS AND METHODS

In this section, after an overview of the whole system, some technical details concerning the methods implementing the mixed-reality mechanism are given.

A. System Overview

The system is based on an HMD equipped with two external cameras that allows physicians to see the preoperative radiological information superimposed to the real scene (Fig. 2). Cameras are mounted on the helmet in correspondence of the user's eyes. Couples of real images are grabbed in real time and showed on the relative liquid crystal display (LCD) internal monitor of the HMD restoring the natural binocular view. Preoperative information, in the form of 3-D virtual models, are projected onto real images in the corresponding position, giving to the user the sensation to see-through the patient. A schematic representation of the proposed stereoscopic mixed-reality system, comprised of the helmet and the mixed-reality mechanism (running on a workstation), is given in Fig. 3.

The helmet comprises an HMD (nVisor SX produced by NVIS, Inc., www.nvisinc.com) with two internal SXGA LCD



Fig. 2. System setup in the OR (left image) and mixed-reality views on the top of HMD (images on the right).

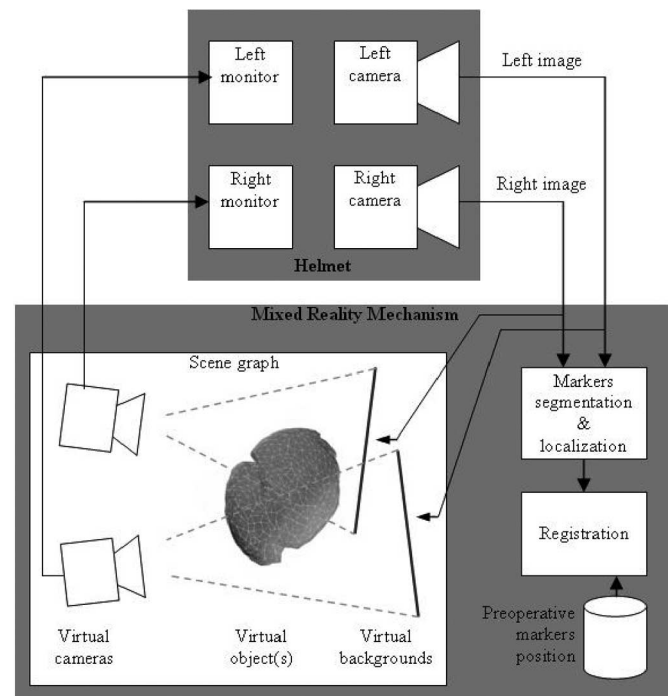


Fig. 3. Schematic representation of the stereoscopic mixed-reality system.

monitors and two external color USB SXGA cameras (uEye UI-1646LE produced by IDS, www.ids-imaging.de) with $1/3$ in image sensor. The chosen optic of cameras, having a focal length of 9 mm, reproduces realistic perception of distances in the typical workspace (50–80 cm far from the point of view). Cameras are fixed to the HMD by means of a mechanical support, showed in Fig. 4, which allows one to adjust camera vergence and, hence, to assure stereo perception at specific distances from the focused object.

The mixed-reality mechanism allows superimposing volumetric images (the medical dataset) and/or surface models (3-D reconstruction of the anatomies) to real images grabbed by cameras.

We reconstruct patient-specific surfaces of the abdominal organs and save them as Virtual Reality Modeling Language (VRML) files, starting from multidetector computer tomography (MDCT) dataset. This process, performed with a

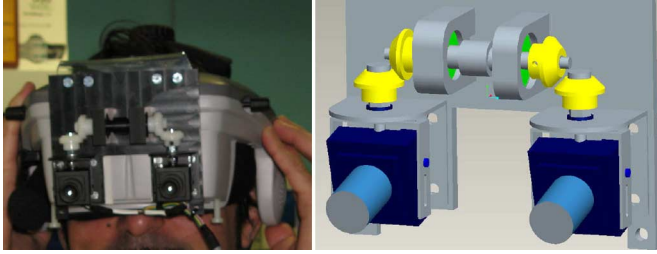


Fig. 4. Mechanical support mounting the cameras: (left) prototype and (right) CAD drawings.

segmentation software which we developed in collaboration with radiologists [13], requires less than 30 min.

The stereoscopic mixed-reality mechanism, developed on the top of the EndoCAS navigator platform modules [5], uses the software framework OpenSG 1.8 (www.opensg.org) for scene graph implementation and management.

Live images, grabbed by means of the real cameras, are used as backgrounds for the two viewports that render the virtual scene and are shown on the corresponding monitor of the HMD (Fig. 3). The mixed-reality mechanism is implemented projecting coherently the virtual objects on the viewports and, hence, superimposing virtual images on the live images of the background. The mixed-reality mechanism requires the definition of a virtual camera model that exactly reproduces the real one, and the alignment of live and virtually reconstructed images by means of a registration step. In our system, this step is based on the geometrical alignment of the real 3-D space (the cameras reference frame) and the virtual 3-D space (the medical dataset reference frame).

The alignment is obtained by applying a marker-based rigid registration method [14]. It requires the identification of the 3-D position of at least three points (markers) both on the medical dataset reference frame and on the real cameras reference frame. For this purpose, we attach on the skin of the patient three or more radio opaque markers before the acquisition of the medical dataset and identify their position on the medical images during their segmentation phase. Real-time segmentation of markers on live video images and their localization in the camera reference frame are performed by means of machine stereoscopic vision routines using the Halcon 7.1 software library developed by MVTech (www.mvtech.com). The whole system runs on a graphic workstation Dell Precision 650 (www.dell.com).

B. Camera Model

In order to mix coherently the virtual information with the real information, each virtual camera has to be modeled identically to the corresponding real camera, and the virtual objects have to be aligned to the real ones.

Let C be the reference system fixed to a camera (C_l for left camera and C_r for right camera) and W be the reference system fixed to the world, and a real point P can be represented, in homogeneous coordinates, in the world reference system W by means of P_w and in the camera reference system C by means of

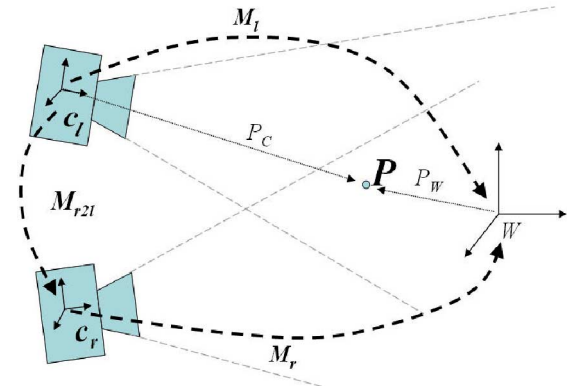


Fig. 5. Geometric transformation involved in the helmet.

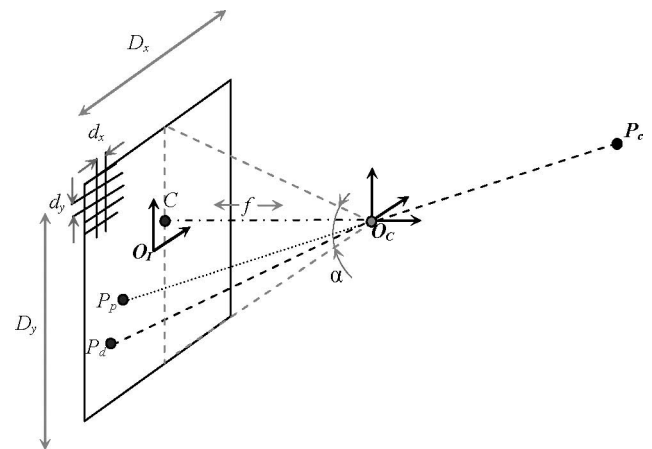


Fig. 6. Schematic representation of the pinhole camera model: the generic point P_c is ideally projected on the image sensor of the camera (the plane with origin O_I) through the projection center O_c (where we fixed the origin of the camera reference frame).

P_c . The rototranslation matrix M such that

$$P_c = MP_w \quad (1)$$

represents the geometric relation existing between the two reference systems C and W , and defines the external camera parameters.

In the system based on two moveable cameras (Fig. 5), at each instant, we need to compute two transformation matrices defined as in (1): one for the left camera (M_l) and one for the right camera (M_r). Since the two cameras are fixed one to another, this can be done using the geometric relation M_{r2l} describing the relative pose of the right camera with respect to the left one.

Each camera has been modeled using the pinhole camera model depicted in Fig. 6, where a generic 3-D point $P_c = [x, y, z, 1]^T$, given in the camera reference system (with origin in O_c), is ideally projected in the corresponding 2-D point $P_p = [u, v, 1]^T$ in the image reference system (with origin in O_I , the center of the image sensor).

The perspective projection matrix M_p such that

$$P_p = M_p P_c \quad (2)$$

is defined starting from the internal camera parameters (f , C_x , C_y) as

$$M_p = \begin{bmatrix} -f & 0 & C_x & 0 \\ 0 & -f & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3)$$

where f is the focal distance and (C_x, C_y) are the coordinates of the projection of O_c on the image reference frame (with origin in O_I). The pixelization process is defined by the pixel dimensions d_x and d_y and image sensor dimensions D_x and D_y . These camera internal parameters allow one to convert measurements done on the image (in pixels) in real measurements (in millimeters) and *vice versa*. The last internal camera parameters parameterize the model of the radial distortion, introduced by common lens, by means of which the projected point P_p is deviated on P_d .

All internal camera parameters and the relative pose of the cameras expressed by the matrix M_{r2l} are determined in a calibration phase acquiring some images of a known object in different positions with fixed camera configuration (in terms of vergence and camera focus) and using calibration routines [15]. Camera calibration procedure, requiring just few minutes, has to be performed each time camera focus or vergence is modified.

C. Implementation of Camera Projection Model

The real internal camera parameters, determined with the calibration process, have been used to model virtual cameras as off-center perspective cameras (in the scene graph libraries OpenSG). The technical implementation requires the definition of the vertical field of view angle α (see Fig. 6). Starting from the focal length f and the projection (C_x, C_y) of the camera reference frame origin O_c on the image plane, the angle α is determined by the equation

$$\alpha = \arctan\left(\frac{C_y}{f}\right) + \arctan\left(\frac{D_y - C_y}{f}\right). \quad (4)$$

The horizontal field of view is set automatically on the basis of image sensor resolution and dimensions.

The implementation of the off-center perspective camera in OpenSG requires the change of the reference system and normalization of the projection of O_c on the image reference frame. This is done by means of the equations

$$C'_x = \frac{2C_x - D_x}{D_x} \quad C'_y = -\frac{2C_y - D_y}{D_y}. \quad (5)$$

The misalignment between real and virtual projections due to radial distortion has been removed compensating the deformation on the real camera images. This step is necessary to obtain exact alignment and realistic mixed-reality representation especially with cameras having great distortions. Radial distortion correction has been implemented with OpenSG directly adjusting the virtual background relative to each camera. The original version of OpenSG (in the 1.8 release) does not allow correcting of images directly, but it simulates radial distortion that maps the image as texture on a triangulated background plate, where triangle vertexes are moved with respect to a regular grid. Regular grid deformation implemented in OpenSG for radial distortion

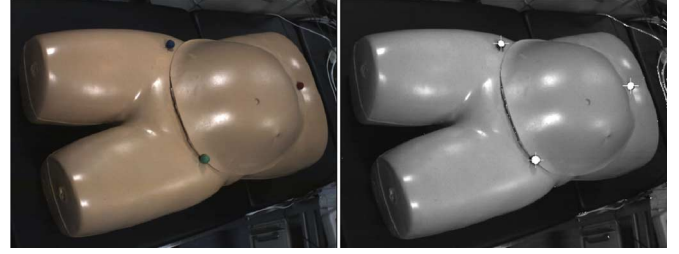


Fig. 7. Three markers consisting in colored felt balls attached on the phantom: (left) original and (right) elaborated image with segmented markers and debug crosses.

simulation is a polynomial function, usually used in machine vision [16]. The radial distortion model used by Halcon is totally different from the OpenSG model and is given by the equation

$$P_d = \frac{2}{1 + \sqrt{1 - 4k \|P_p - O_i\|^2}} (P_p - O_i) + O_i \quad (6)$$

where k is the radial distortion factor. Starting from (6), it is possible to recover analytically the nondistorted P_p point for each distorted P_d point

$$P_p = \frac{1}{1 + k \|P_d - O_i\|^2} (P_d - O_i) + O_i. \quad (7)$$

Equation (7) cannot be analytically translated in a polynomial function, so it is not compatible with the model implemented in OpenSG. Therefore, we implemented, on the top of OpenSG structures, a function for grid deformation based on (7), obtaining the correction of the radial distortion of the live images grabbed by cameras. This solution is very simple and fast since OpenSG uses graphical processing unit (GPU) to perform computations.

D. Fiducial Markers Localization and Registration

The alignment (superimposition) of the virtual world with the real world is made updating in real time the rototranslation matrices M_l and M_r (Fig. 5) that describe the geometrical relation between the cameras reference frame (i.e., the helmet reference frame) and the scene graph reference frame (i.e., the medical dataset reference frame).

To solve the alignment problem, we used three artificial markers fixed on the skin of the patient before the acquisition of the medical images. Matrices M_l and M_r are defined registering the position of fiducial markers segmented on the medical dataset (preoperative fiducial markers position: F_1, \dots, F_n), determined preoperatively and given in medical dataset reference frame, to their real position, measured in real time using stereoscopic vision routines on pair of live images (left and right camera images).

We used artificial markers consisting in 1-cm-diameter felt balls colored with three different uniform colors [red, green, and blue (RGB)], placed as shown in Fig. 7. Felt material allows for sensibly reducing reflection artifacts.

The real-time tracking of markers involves a segmentation process, performed on the grabbed images, consisting of two

steps: color segmentation and circular shape recognition. Color segmentation, based on thresholding, is performed using the hue saturation value (HSV) representation. This allows reduction of the segmentation errors due to the artifacts introduced by the change of the illumination level. For each colored ball, H and S thresholds values were defined after a training phase. Transformation of the images from the RGB camera color space to the HSV space is performed using a lookup table in order to speed up the frame rate [17]. After thresholding, the circularity shape factor (CSF) is computed for each connected region of the color segmented image. The CSF is defined as follows:

$$\text{CSF} = \frac{A}{D_{\max}^2 \pi} \quad (8)$$

where A is the area of the region and D_{\max} is the maximum distance from pixels to region barycenter (note that for a circular connected region, $\text{CSF} = 1$). Fiducial marker is then chosen as the biggest region having a $\text{CSF} > 0.5$. This empirical method identifies fiducial markers on the image plane with very low computational time. We evaluated that the target identification fails in less than 2% of cases. This result is sufficient for our purposes since the registration is computed a lot of time in a second. Obviously, the method definitely fails if large circular objects having the same color of markers are inserted in the scene, but this situation can be simply avoided in a real clinical scenario.

After marker segmentation, on the image plane, fiducial markers are localized in the 3-D space applying stereoscopic vision routines. First of all, the marker position is identified on the image plane with a single pixel consisting in the barycenter of the corresponding segmented region. Knowing the internal camera parameters, for each marker position P_d , in the image plane, the relative projection line in the 3-D world, defined as the line l passing through the camera center of projection O_c and lies on the point P_c , is determined. These steps, performed both on left and right images, identify two projection lines l_l and l_r , respectively. Knowing the relative pose of the right camera to the left camera (M_{r2l}), the 3-D position of each marker is then defined as the intersection point between l_l and l_r . Since l_l and l_r do not intersect (due to pixelization process and noise in marker identification), the 3-D marker position is approximated with the position of the point that is nearest to both projection lines.

In conclusion, stereo routines provide the 3-D position of each fiducial marker $i = 1, \dots, n$ in the reference frame of the left camera (F_{C11}, \dots, F_{C1n}). Having the two clouds of points composed, respectively, by fiducial markers in the radiological scanner reference frame F_i and fiducial markers in the reference frame of the left camera F_{C1i} , M_l is chosen in order to align as good as possible each couple of fiducials

$$M_l F_i \cong F_{C1i} \quad \forall i \quad (9)$$

using the marker-based rigid registration method described in [14], and M_r is obtained applying the fixed geometric relation M_{r2l} between cameras

$$M_r = M_{r2l} M_l. \quad (10)$$

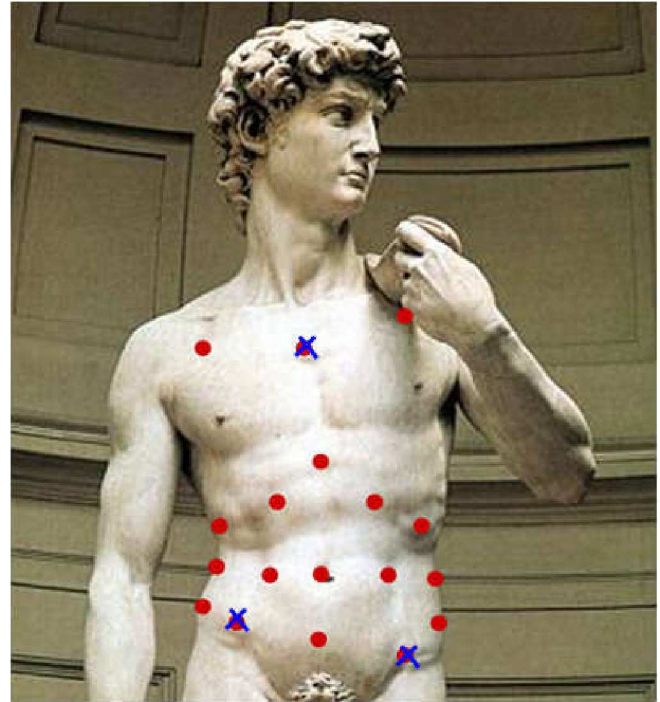


Fig. 8. Marker configuration on the subject skin (points) used for the motion estimation, and the most stable positions during breathing (crosses) chosen for the registration.

The evaluation of the fiducial registration error (FRE) [18] allows one to individuate fault conditions in fiducial markers segmentation. In fact, high values of FRE can be just due to an error in the stereoscopic localization of markers.

To reduce computational time required for marker segmentation and to obtain 25 frames per second (fps), the new marker position in the images is searched in a subimage square region centered in the previous marker position and having an area proportional to FRE. When a localization error occurs (a too large FRE), the entire image has to be processed one more time.

III. RESULTS

System evaluation was performed using a trunk phantom for laparoscopy produced by Sawbones (www.sawbones.com). We attached on the phantom external surface three radio opaque fiducial markers in correspondence of the sternum (one marker) and iliac spines (two markers). We used this marker configuration on the base of our experience in the rigid registration of the abdomen [5], [19]. We chose the most stable positions during normal breathing, on the base of experiments consisting in the acquisition of the trajectories of many markers attached on the skin of a subject by means of an optical tracker (Fig. 8).

We scanned the phantom with CT and extracted surfaces of fiducial markers (for registration) and anatomical structures (for mixed-reality visualization). Barycenters of segmented markers are used as fiducials F_i , while the corresponding F_{C1i} are obtained localizing with the stereoscopic routines the colored 1-cm-diameter felt balls attached in the corresponding positions.

The quality of the superimposition of the virtual anatomies to the real scene depends on the accuracy of localizing fiducial markers on the cameras' images sensors and on the cameras configuration. We estimated the error in localizing markers on image sensor as $\Delta p = \pm 3$ pixel. The major component of the localization error due to camera configuration is measured along the z -axis and is defined analytically by the distance resolution formula [20]

$$\Delta z = \frac{z^2}{fb} \Delta d \quad (11)$$

where b is the distance between cameras, z represents the working distance, f is the focal length, and Δd is given by $\Delta p * \text{pixel dimension}$. It allows estimation of the localization error for each configuration of the cameras and the markers. In our testing setup, where $b = 70$ mm (anthropometric value), $f = 9$ mm, and $\Delta d = \pm 11 \mu\text{m}$, we obtain a $\Delta z = \pm 4.4$ mm at a working distance $z = 500$ mm. This error is principally due to the short distance between the cameras with respect to the working distance. Even if the distance resolution error is numerically significant, in our mixed-reality visualization using the trunk phantom, it is perceived as a negligible alignment error from the user. Further, we experienced that the use of more precise localization systems, also in other markers configurations, during test sessions on some patients, does not introduce significant visualization improvements that are perceived as substantial by the user.

We evaluated the accuracy of our system using stereoscopic localization and compared it with using commercial localizers for the registration of the phantom used for the prototype testing.

In a previous work [5], we obtained very precise alignment using the NDI Optotrak Certus (the golden standard in commercial localizers) and sensorizing both the camera and the real object (a calibrated grid representing conceptually the patient). The precision obtained under these ideal conditions significantly decreases in a real clinical scenario for two main reasons. First of all, the patient is not rigid, and second, localizable fiducial markers cannot be fixed steadily on the patient. The last condition requires human actions for the repositioning of localizable markers on the patient, or the acquisition of fiducial markers position by means of a digitizer. For these reasons, we evaluate the accuracy of our system taking into account not just the precision of the localization system, but also the whole registration procedure required to align real and virtual world. To this purpose, we measured the fiducials registration error (FRE).

In our system, the marker localization, done by stereoscopic vision routines, does not require user intervention (automatic segmentation). To determine FRE, we placed the helmet in several positions of the workspace commonly covered by doctors during medical exam, and at each position, we performed localization and registration with preoperative markers positions (given in the CT reference frame). Registrations with an FRE > 10 mm (value that guarantees the success of marker segmentation and determined empirically watching debug crosses placed in correspondence of the estimated positions on the grabbed images) have been automatically removed from the FRE evaluation set.

TABLE I
FRE (IN MILLIMETERS) OBTAINED WITH OUR SYSTEM AND WITH
COMMERCIAL LOCALIZERS

Localizer	Mean FRE	Std FRE
Our system	3.47	1.78
FASTRAK Polhemus	3.00	0.41
NDI Aurora	2.72	0.62
NDI Optotrak Certus	2.01	0.54

To determine FRE in using commercial localizers, we asked five people to digitize for ten times, using different localization systems (FASTRAK Polhemus, NDI Aurora, NDI Optotrak Certus), the three fiducial markers consisting in signs on the phantom. Registering markers position acquired by the subjects, with markers positions given in the CT reference frame, we obtained the FRE evaluation for each localization system. The mean values of FRE and standard deviations are reported in Table I.

To obtain a localization refresh rate of 25 fps, we reduced the segmentation region of each marker to a subimage of 200×200 pixel. This box dimension allows detection of markers under typical head movement conditions and to maintain the same number of localization faults obtained segmenting the whole image (about 2% in an ambulatory room with artificial illumination). The system allows registration of the virtual anatomy if all the three markers are in the images (left and right). From our experience, the percentage of fault registration, for too high FRE (> 10 mm), does not increase changing the view angles (if the markers are not covered). This is due to the type of used markers, whose outer surface has an unstructured pattern, which introduce considerable variations in the identification of their barycenter on the image sensor independently of the view angles.

IV. CONCLUSION

In the paper, we described a system we developed for 3-D mixed-reality visualization of medical exams. The study demonstrates the feasibility and the potentialities of the use of mixed-reality techniques in the clinical and surgical domains.

Methods we implemented to avoid the use of external localizer have been described and evaluated. Results demonstrate that the stereoscopic localization approach, adopted in our system, even if it introduces an FRE bigger than that of the more common commercial localizers, does not compromise the usability of the system.

The system has been tested in a real clinical scenario by two surgeons of our department. Both surgeons demonstrated enthusiasm for the system and real interest in its integration in clinical practice. Their opinion about the system, and more generally about the integration of the proposed technologies in the clinical scenario, is that it simplifies the reading of medical images and introduces advantages especially for the tasks where the physician has to interact with the patient. Moreover, in their opinion, the only drawback is represented by the weight of the system, which should be reduced in the future.

Localization precision using stereo cameras can be drastically improved reaching submillimetric precision [21], [22]. The implementation of these methods requires very accurate segmentation of markers. Our future work is focused on the study of materials and structured shapes for marker fabrication in order to eliminate reflection artifacts and to achieve subpixel segmentation accuracy.

REFERENCES

- [1] T. M. Peters, "Image-guided surgery: From x-rays to virtual reality," *Comput. Methods Biomech. Biomed. Eng.*, vol. 4, no. 1, pp. 27–57, 2000.
- [2] T. M. Peters, "Image-guidance for surgical procedures," *Phys. Med. Biol.*, vol. 51, no. 14, pp. 505–540, Jul. 2006.
- [3] J. H. Shuhaiber, "Augmented reality in surgery," *Arch. Surg.*, vol. 139, pp. 170–174, 2004.
- [4] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. 77, no. 12, pp. 1321–1329, 1994.
- [5] G. Megali, V. Ferrari, C. Freschi, B. Morabito, F. Cavallo, G. Turini, E. Troia, C. Cappelli, A. Pietrabissa, O. Tonet, A. Cuschieri, P. Dario, and F. Mosca, "Endocas navigator platform: A common platform for computer and robotic assistance in minimally invasive surgery," *Int. J. Med. Robot. Comput. Assisted Surg.*, vol. 4, pp. 242–251, 2008.
- [6] C. Bichlmeier, T. Sielhorst, S. M. Heining, and N. Navab, "Improving depth perception in medical AR: A virtual vision panel to the inside of the patient," presented at the Bildverarbeitung für die Medizin, Munich, Germany, 2007.
- [7] L. Johnson, E. Philip, G. Lewis, and D. Hawkes, "Depth perception of stereo overlays in image-guided surgery," *Proc. SPIE Med. Imag.*, vol. 5372, pp. 263–272, 2004.
- [8] H. Hua, P. Krishnaswamy, and J. P. Rolland, "Video-based eyetracking methods and algorithms in head-mounted displays," *Opt. Exp.*, vol. 14, pp. 4328–4350, 2006.
- [9] E. C. Lee and K. R. Park, "A robust eye gaze tracking method based on a virtual eyeball model," *Mach. Vis. Appl.*, vol. 20, pp. 319–337, 2009.
- [10] Y. Genc, F. Sauer, F. Wenzel, M. Tuceryan, and N. Navab, "Optical see-through HMD calibration: A stereo method validated with a video see-through system," in *Proc. Int. Symp. Augmented Reality*, 2000, pp. 165–174.
- [11] Y. Baillot and S. J. Julier, "A tracker alignment framework for augmented reality," in *Proc. 2nd IEEE ACM Int. Symp. Mixed Augmented Reality*, 2003, pp. 142–150.
- [12] M. Kanbara, T. Okuma, H. Takemura, and N. Yokoya, "A stereoscopic video see-through augmented reality system based on real-time vision-based registration," in *Proc. IEEE Virtual Reality 2000*, pp. 255–262.
- [13] V. Ferrari, C. Cappelli, G. Megali, and A. Pietrabissa, "An anatomy driven approach for generation of 3D models from multi-phase CT images," *Int. J. Comput. Assisted Radiol. Surg.*, vol. 3, supp. 1, pp. 271–273, Jun. 2008.
- [14] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, Sep. 1987.
- [15] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
- [16] C. McGlone, E. Mikhail, and J. Bethel, *Manual of Photogrammetry*, 5th ed. Bethesda, MD: Amer. Soc. Photogrammetry Remote Sens., 2004.
- [17] V. Kravtchenko, "Thesis title: Tracking color objects in real time," Master's thesis, Univ. British Columbia, Vancouver, BC, Canada, 1999.
- [18] J. M. Fitzpatrick, J. B. West, and C. R. J. Maurer, "Predicting error in rigid-body point-based registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 694–702, Oct. 1998.
- [19] V. Ferrari "Improving daily clinical practice with abdominal patient specific 3D models," Ph.D. dissertation, EndoCAS Center, Università di Pisa, Pisa, Italy, Feb. 2009.
- [20] *Halcon Application Guide 7.1*, MVTex Software GmbH, Munich, Germany, 2005.
- [21] J. Amat, M. Frigola, and A. Casals, "Selection of the best stereo pair in a multi-camera configuration," in *Proc. ICRA*, 2002, pp. 3342–3346.
- [22] H. Sahabi and A. Basu, "Analysis of error in depth perception with vergence and spatially varying sensing," *Comput. Vis. Image Understanding*, vol. 63, no. 3, pp. 447–461, 1996.

Author's photographs and biographies not available at the time of publication.