# Validation of a stereo camera system to quantify brain deformation due to breathing and pulsatility

Carlos Faria[a)]
*MARLab - Department of Industrial Electronics and Centre Algoritmi, Universidade do Minho, 4800-058 Guimarães, Portugal*

Ofri Sadowsky
*Medical Robotics Laboratory - Mechanical Engineering Department, Technion - Israel Institute of Technology, 32000 Haifa, Israel*

Estela Bicho
*MARLab - Department of Industrial Electronics and Centre Algoritmi, Universidade do Minho, 4800-058 Guimarães, Portugal*

Giancarlo Ferrigno
*NearLab - Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milano, Italy*

Leo Joskowicz
*School of Engineering and Computer Science, Hebrew University of Jerusalem, 91904 Jerusalem, Israel*

Moshe Shoham
*Medical Robotics Laboratory - Mechanical Engineering Department, Technion - Israel Institute of Technology, 32000 Haifa, Israel*

Refael Vivanti
*School of Engineering and Computer Science, Hebrew University of Jerusalem, 91904 Jerusalem, Israel*

Elena De Momi
*NearLab - Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milano, Italy*

(Received 4 May 2014; revised 11 August 2014; accepted for publication 15 September 2014; published 20 October 2014)

## 1. INTRODUCTION

Brain shift[1–3] represents a major source of error in image-guided neurosurgery. The skull-fixation system that assures the rigid transformation between preoperative images and the patient coordinated space[4] does not account for the brain deformation relative to the skull. In frameless neurosurgery, the registration of the intraoperative space to the imaging scans relies on external landmarks (skin surface or fiducialmarkers).[5]

There are two sources for brain tissue deformation reported in open skull neurosurgery. The first concerns brain shift and results from a combined effect of tissue resection, pathology, pharmacologic responses, gravity, edema, and loss of cerebrospinal fluid.[6] Hill *et al.*[1] measured the brain surface

deformation in 21 patients under a craniotomy and reported a mean cortical surface displacement of 5.6 mm and a maximum displacement over 10 mm.

Another source of brain deformation noticeable in open skull surgery is due to systematic blood pressure variations, breathing, and vasomotor oscillations, which result in a pulsating brain motion.[7,8] Britt and Rossi[9] quantified the brain pulsation motion of a mammal brain and divided its components into arterial component with a displacement 110–266 $\mu$m and a motion frequency between 2–3 Hz, and a respiratory component with a displacement of 300–950 $\mu$m and a frequency between 0.1–0.2 Hz. Scarce information was found related to the characteristics of pulsation/breathing motion of the human brain. Wagshul *et al.*[8] present a frequency analysis of brain pulsatility based on intracranial pressure and photoplethysmography, whose main components range between 1–4 Hz. Poncelet *et al.*[10] reported a pulsating displacement amplitude to be at most 0.5 mm.

Brain shift is a progressive condition with a very low displacement velocity, thus not requiring a high sample rate for its quantification. Several approaches have been proposed to measure brain shift: magnetic resonance,[1,11–14] computed tomography,[15] ultrasonography,[16–20] biomechanical/predictive models,[6,21–24] or methods that use *a priori* knowledge of a brain biomechanical model to constrain the registration of 3D intraoperative imaging.[25] However, the computation times, cost of technology, and necessary infrastructures as well as the cumbersomeness of user-interface rendered most of these solutions inapt.[26] In neurosurgeries involving craniotomy, cortical surface can be used to infer brain deformation. Two main research areas stand out: laser range scanning[27–29] and stereoscopic systems.[30,31] However, the low sample rate of laser range scanner systems enables its use to measure brain shift but not brain pulsation deformation.

The process of tracking a surface with a stereoscopic system is a two step process: (1) extraction of 3D positions of salient visual features on the surface and (2) tracking of these features in video frames at different times. Sun *et al.*[30] presented in 2005 a system able to measure brain deformation and pulsation with a stereoscopic microscope (capturing 15 frames/s) achieving a point-localization average error of 1.02 mm when compared to a optical stylus probe of a Polaris system and capturing the cortical surface shape of a phantom model with an average error inferior to 1.2 mm when compared to a CT scan. In 2011, Ji *et al.*[32] reported a stereo vision system to measure cortical surface strain based on differentiation of brain motion using optical flow, with high resolution (within 0.05–1 mm) although no explanation is given about how sensitivity is calculated. Later in 2013,[33] Ji *et al.* presented a noninvasive method, based on the approach reported in 2011,[32] to produce a dense map of cortical surface deformation due to blood pressure and respiratory cycles in patients undergoing brain tumor resections via craniotomy. The breakthrough from the previous work is based on the use of an average cortical surface deformation to compute the relative cortical deformation instead of considering an arbitrary measurement as the reference state. Ji *et al.*[33] also presents a fast Fourier transform (FFT) analysis of brain deformation from a clinical case. These papers also report surface cortical displacement up to 1 and 2 mm but none conducted system accuracy or validation tests.

We present a validation method to evaluate the static and dynamic accuracy of a stereo camera. A simple phantom is simultaneously tracked by the stereo camera and by a commercial high-accuracy optical tracking system used as gold-standard reference. Results provide further information about the reliability of stereoscopic systems—not only—in the context of brain shift/deformation (as usually reported in surface tracking papers) but more important, for measurements of brain pulsatility.

To measure the brain motion, we use a stereo camera to extract, and robustly and continuously track features on the cortical surface. Three brain motion acquisitions were recorded intraoperative. This paper is organized as follows: Sec. 2 presents in detail the stereo vision system setup, in Secs. 3 and 4 the materials and methods to validate the stereoscopic system are described, Sec. 5 shows the validation results, Sec. 6 presents brain motion data acquired in a neurosurgery, and Sec. 7 includes the discussion and conclusion.

## 2. STEREO VISION SYSTEM

The stereo vision system includes: a camera module, communication channels, a personal computer, and the tracking software. The stereo camera module consists of two mvBlueFox 223C CCD cameras (MATRIX VISION GmbH, Oppenweiler, Germany, 38.8×38.8×58.5 mm body dimensions) each equipped with SM25WI 2/3″ lenses (Rainbow CCTV, Irvine, Canada, 25 mm 1:1.4 C-Mount, manual iris/focus). The cameras are able to record 1360×1024 resolution, at a maximum frame rate of 20 frames/s, although the USB 2.0 connection limits the data transfer of full resolution images to 10 frames/s. The computer with the tracking software runs Microsoft Windows 7 64-bit (Intel i7-3770S CPU, 8 GB of RAM), see Fig. 1.

The tracking software consists of an image acquisition and a tracking module. The tracking module can be split into several components: (i) feature detection and stereo matching, (ii) stereo video tracking, and (iii) 3D reconstruction.

Our software includes the following features:

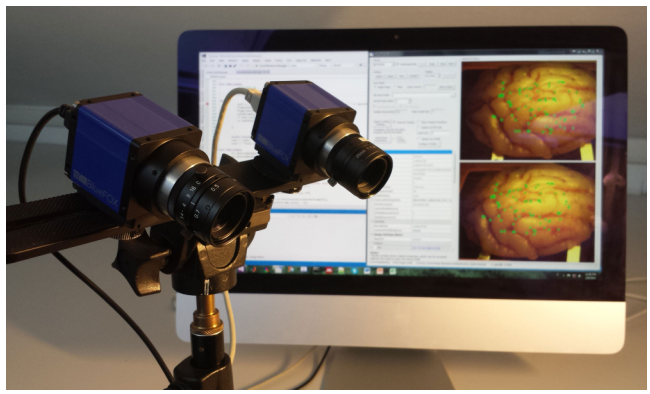

FIG. 1. Stereo vision system setup. Stereo camera fixed on a bracket, connected to a computer via USB 2.0.

- Simultaneous operation of multiple cameras (two, in this work).
- A user-controlled synchronized image capture cycle, which is critical for correct tracking.
- Saving the acquired images in video files.
- Encoding a timestamp into the acquired images to facilitate off-line synchronization.
- Integration with feature tracking and reconstruction algorithms.

### 2.A. Camera calibration

A calibration of the stereo camera is required prior to its use. The underlying camera model is taken from the OpenCV library and includes a pin-hole projection matrix, radial and tangential lens distortion, and a rigid transform relating the two cameras in a stereo system. To ease the calibration process, we use a coded checkerboard pattern from the CALTag toolkit[34] in which the coordinates of every corner are automatically identified from embedded binary codes, and is thus robust to rotation, occlusion, or partial views. Snapshots of the pattern are taken at several different poses (typically 15 to 20) and processed automatically to compute all calibration parameters.

### 2.B. Tracking software

The first step of the tracking algorithm is feature detection. To accelerate the process, we start by asking the user to specify approximately corresponding rectangular regions of interest (ROIs) on the left and right images. We then execute multiple filters on the acquired left image:

(1) Convert the image to a single channel, e.g., from color to gray.
(2) Apply an intensity threshold above which pixels are suspected as specular reflections and not used for features to prevent tracking false visual items.
(3) Select "interesting" objects which stand out relative to their neighborhood. For example, cortical blood vessels are typically darker than the surrounding brain tissue, and therefore we seek for pixels darker than their neighborhood.
(4) To locate specific features, we use the Harris corner detector[35] on the left camera image. It finds features with high pixel intensity derivatives in two orthogonal directions.

To avoid clustering all features in a small region, we subdivide the original ROI and attempt to locate several features (e.g., 4) in each rectangular subregion, which increases the area of coverage within the ROI. The total number of tracked features is determined by a specified ceiling amount (typically 100) and an acceptance threshold on the sensitivity of the Harris corner detector filter. The sensitivity parameter can be lowered to increase the number of detected corners present in the image, up to specified ceiling. Compared to the optical flow approach in Ji *et al.*,[33] which reconstructs a dense displacement map of the surface, we use a relatively sparse set of features. The small number of features allows real time tracking. Notice that normalized cross correlation (NCC) is robust to variations in the intensity gain of the patch, especially when compared to the constant intensity of a feature presumed by the optical flow method in Ji *et al.*[33] This is relevant when a shadow, not an occlusion, is cast over the tracked region.

Every feature is identified by a small patch (e.g., $15\times15$) of pixels in its neighborhood. This patch serves as a template for searching the feature in another image. The matching computes the NCC[36] with the template patch over a search region which is described below. The location of maximal NCC score is chosen as a first approximation. Since the physical motion of the tracked feature is not limited to integer pixel units, we refine this approximation by fitting a quadratic surface to a $3\times3$ neighborhood of the first guess and taking its apex (location of zero gradient) as the position of the match.

Feature matching is used in two stages of the algorithm: between the first left-camera image (image index $i=0$ in the sequence) and the first right-camera image (we call these "stereo matches"), and between subsequent images (indexes $i$ and $i+1$) from the same camera—"video matches" (see Fig. 2). For each stage, we define a different search region. For stereo matching search, we use the presumed correspondence between left and right ROI rectangles plus a user-specified volume depth estimation to allow for some tolerance. We reconstruct by triangulation (Sec. 2.C) the four corners and then add and subtract the depth parameter from the Z coordinate to receive eight corners of a box. These are reprojected to the image planes and the maximal offsets between these corresponding projections define the search range. The video matching search region is a (small) neighborhood of the current feature position, based on the assumption that the motion of the feature between frames is limited.

False matches, or outliers, must be removed from the data to achieve an accurate reconstruction of the visible surface. We identify outliers with the following criteria. First, image similarity measures of NCC and normalized mutual information (NMI)[37] are computed between the template and the match candidate, and a minimal score acceptance criterion is used. Empirically, we found that thresholding the NCC alone is insufficient to detect some occlusion cases, and the NMI criterion adds strength to it. Then, for stereo matching, two geometric constraints are tested: (1) the distance of the right-image feature from the epipolar line that is the projection of the left-image feature on the right image plane and (2) a parabolic surface is fitted to the collection of 3D reconstructed features (see below) as a rough approximation of the brain surface shape, and points too far from this surface are considered as outliers.

### 2.C. 3D reconstruction

The pin-hole projection model is based on the following equation:

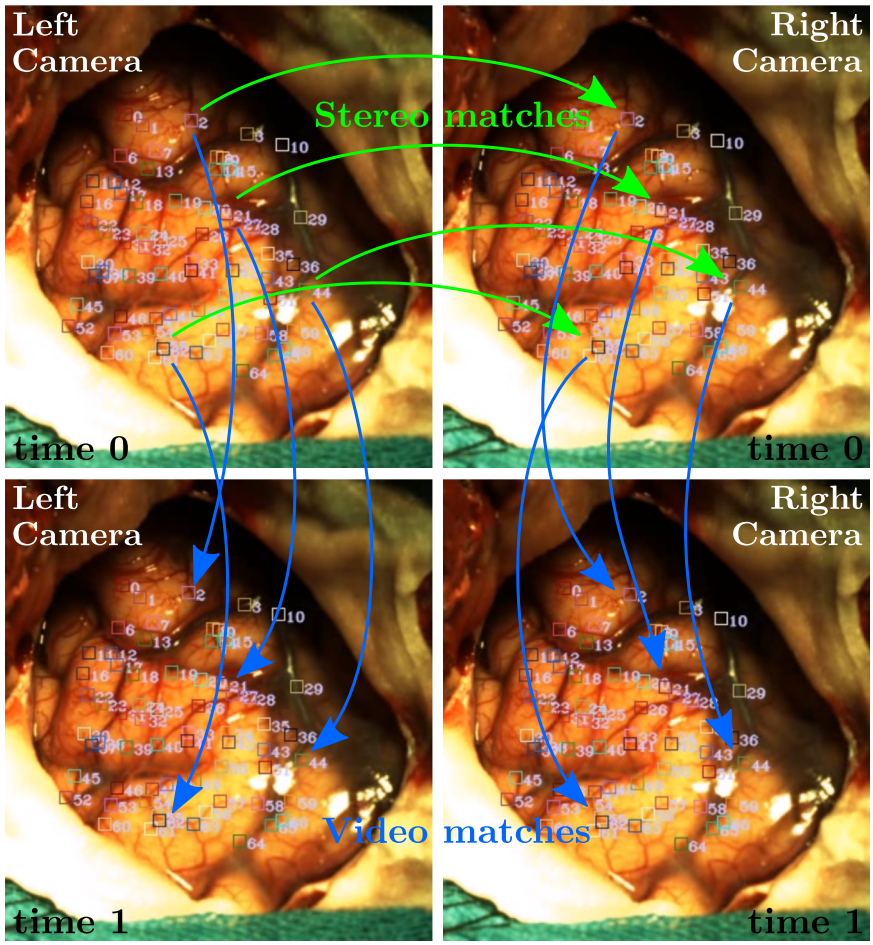


FIG. 2. Stereo and video feature matching. Each row shows a video frame, and columns stand for the left and right cameras. Stereo matches are depicted by green arrows from the left to the right camera at time 0, and blue arrows show video matches for both cameras from time 0 to time 1. Each feature neighborhood is depicted as a small numbered square.

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \mathbf{A}\left(\mathbf{R}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t}\right), \tag{1}$$

where $\lambda$ is an arbitrary scale; $u$ and $v$ are the coordinates of a pixel; $\mathbf{A}$ is a $3\times3$ projection matrix (with nonzero elements in the diagonal and the last column); $\mathbf{R}$ is a $3\times3$ rotation matrix; and $\mathbf{t}$ is a translation. $\mathbf{A}$ is constant per camera, and $\mathbf{R}$ and $\mathbf{t}$ express the pose of the camera relative to the imaged object which is acquired in a calibration process and can involve the stereo relation between the cameras.

The back-projection of a pixel (line $L$) is henceforth computed as

$$L = -\mathbf{R}^{-1}\cdot\mathbf{t} + \lambda\mathbf{R}^{-1}\mathbf{A}^{-1}\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \tag{2}$$

where the first term is the position of the camera's principal point relative to the imaged object, and the second is the direction of the ray that pierces the pixel. In short notation, we will write $L = \mathbf{p} + \lambda\mathbf{g}$ and assume that $\mathbf{g}$ is a unit-norm vector.

Given two pairs of pixel coordinates, $(u_1,v_1)$ and $(u_2,v_2)$, which presumably are projections of the same point $\mathbf{x}$ from an object, the 3D position of $\mathbf{x}$ is reconstructed by intersecting the two rays. Since the rays usually do not intersect (because of localization ambiguity and numerical roundoff), we seek for a point that is exactly midway along the shortest line segment that connects the two rays. This line segment is also perpendicular to both rays. The perpendicular projection of $\mathbf{x}$ on a line $L = \mathbf{p} + \lambda\mathbf{g}$ is

$$(\mathbf{x}-\mathbf{p}) - \big((\mathbf{x}-\mathbf{p})^T\mathbf{g}\big)\mathbf{g} = (I - \mathbf{g}\mathbf{g}^T)(\mathbf{x}-\mathbf{p}).$$

And the midway point satisfies the equation

$$(I - \mathbf{g_1}\mathbf{g_1}^T)(\mathbf{x}-\mathbf{p_1}) + (I - \mathbf{g_2}\mathbf{g_2}^T)(\mathbf{x}-\mathbf{p_2}) = \mathbf{0} \tag{3}$$

which is easily solved.

To address the radial and tangential lens distortion of the OpenCV camera model, we use the solution of the linear equation as an initial guess in Levenberg–Marquardt optimization to minimize the squared difference between the projection of $\mathbf{x}$ on the respective image and the located pixel coordinates of the feature.

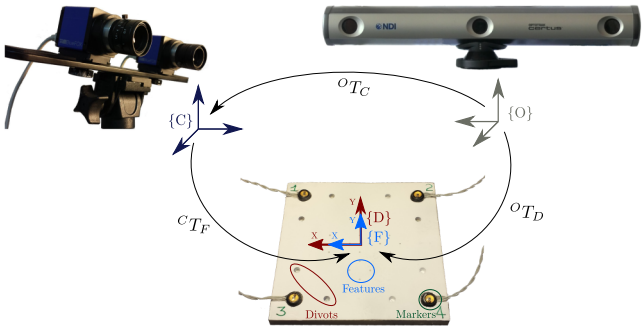


FIG. 3. Validation setup. The three reference frames: stereo camera $\{C\}$, optical tracking system $\{O\}$, and custom designed plate which includes the features' reference frame $\{F\}$ and the divots' reference frame $\{D\}$, which are overlapped to each other. ${}^{C}\mathbf{T}_{F}$ represents the transformation from $\{F\}$ to $\{C\}$, ${}^{O}\mathbf{T}_{D}$ from $\{D\}$ to $\{O\}$, and ${}^{O}\mathbf{T}_{C}$ from $\{C\}$ to $\{O\}$.

## 3. MATERIALS

To validate the stereoscopic system (Fig. 3), a custom designed plate was synchronously tracked by the stereo camera and an industrial optical tracking system considered as the gold-standard reference for this paper: Optotrak CERTUS (NDI, Ontario, Canada). The choice of the optical tracker was held based on the reported 3D tracking accuracy 0.1 mm (system datasheet).

### 3.A. Synchronization logic

To validate the stereo camera against the reference system, both systems must track the same or a related object synchronously. The synchronization between both systems was achieved through a logic hardware component connecting each system's IO modules and modulating the output signal from CERTUS to trigger the stereo camera (Fig. 4).

The short clock acquisition pulse from CERTUS occasionally caused the cameras not to trigger. A monostable multivibrator (HCF4047BE) was included in the logic component to increase the duty cycle of $C_{\text{trg}}$ and avoid not-detection problems.

### 3.B. Custom plate

The custom designed plate (Fig. 5) was built from a 80 $\times$80 mm plexiglass and machined with a PowerWASP EVO 3D printer (WASProject, Massalombarda, Italy), which has a 0.1 mm motion resolution.

The plate includes three groups of elements:

- Features, small drilled conical holes of 0.5 mm diameter and depth. These details have high pixel intensity derivatives along two orthogonal axes, which make them salient visual features easily recognized and tracked by the stereo camera's corner-detector algorithm.
- Markers (or active-LEDs) are the elements fixed on the corners of the plate (Fig. 3), tracked by CERTUS. A rigid body or dynamic reference frame (DRF) {M} was created with the markers using the 6D Architect software (NDI, Ontario, Canada) and is recorded, instead of the position of each single marker.
- Divots are 3 mm etched hemispherical holes with the center coplanar to the plate's DRF $xy$-plane, and are used to calibrate the plate's DRF $\{P\}$ (Fig. 5) to the markers' DRF.

The relation between features tracked by the stereo camera and markers tracked by CERTUS was inferred through the plate layout. To avoid the marker localization error (distance between its true position and the layout position on the plate's DRF $\{P\}$) that depends on the fixation of each marker to the plate, divot sockets were used to compute the transformation from $\{P\}$ to the markers' DRF, since divot positions can be directly measured and they are in the $xy$-plane of $\{P\}$ like the features (cf., Sec. 4.A).

The nominal homogeneous coordinates ($\mathbf{p}_i = [x_i \;\; y_i \;\; z_i \;\; 1]$) of divots in $\{P\}$ are (in mm)

$$\mathbf{p}_{\text{ND}} = \begin{bmatrix} 30 & 15 & 30 & 15 & -15 & -30 & -15 & -30 \\ 15 & 30 & -15 & -30 & 30 & 15 & -30 & -15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{T}. \quad (4)$$

A group of six features were selected to be tracked. The nominal homogeneous coordinates of the chosen features in the plate's DRF $\{P\}$ are (in mm)

$$\mathbf{p}_{\text{NF}} = \begin{bmatrix} 10 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10 & 5 & -5 & -10 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{T}. \quad (5)$$

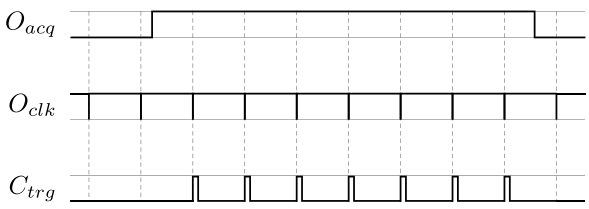


FIG. 4. External signals from CERTUS: (1) $O_{\text{act}}$ is high while the markers' positions are being recorded and (2) $O_{\text{clk}}$ goes low for 1 $\mu$s with each acquisition of the marker's positions. The input signal of the stereo camera $C_{\text{trg}}$ consists of a high-level pulse synchronous to the CERTUS clock ($O_{\text{clk}}$) when it is recording ($O_{\text{act}}$ is high).

## 4. METHODS

To assess the accuracy of the stereo vision system based on the accuracy of the reference system, the transformation between both tracking systems (${}^{O}\mathbf{T}_{C}$) and the transformation between the tracked elements (${}^{D}\mathbf{T}_{F}$) must be known.

### 4.A. Plate calibration

Initially, a calibrated 4 Marker Digitizing Tool (Optotrak CERTUS accessory, NDI) recorded the position of each divot

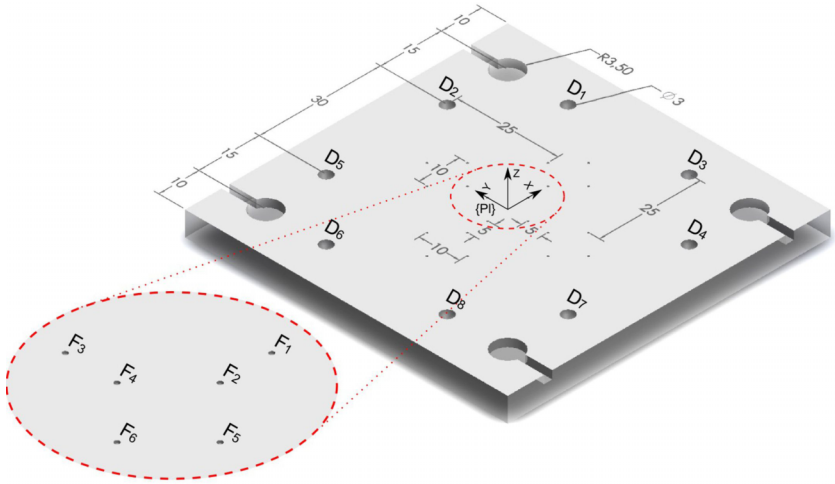


FIG. 5. Custom designed plate symmetrical along the $x$ and $y$-axis of its dynamic reference frame, denominated $\{P\}$. Plate dimensions depicted with a "zoom in" of the region with the six tracked features. The divots' $\{D\}$ and features' DRF $\{F\}$ are assigned to be coincident to the plate's DRF $\{P\}$.

($\mathbf{p}_D$) in the CERTUS's static reference frame (SRF) $\{O\}$, following the order depicted in Fig. 5, from 1 to 8. The position and orientation of the markers' DRF in $\{O\}$ was also recorded (${}^O\mathbf{T}_M$). The divots' positions in the markers' DRF ($\mathbf{p}_{D,M}$) were computed:

$$\mathbf{p}_{D,M} = ({}^O\mathbf{T}_M)^{-1} \cdot \mathbf{p}_D. \tag{6}$$

Taking advantage of the plate shape and to minimize the divots' localization error, $\mathbf{p}_{D,M}$ were projected to a plane fitted to them by least squares, and thus the projections of the divots on that plane ($\mathbf{p}_{D,P}$) were computed.

Knowing $\mathbf{p}_{D,P}$ and the markers' DRF (${}^O\mathbf{T}_M$) in the CERTUS SRF $\{O\}$, we calculated the divots' positions in $\{O\}$, $\mathbf{p}_{D,O}$, for each acquisition.

$$\mathbf{p}_{D,O} = {}^O\mathbf{T}_M \cdot \mathbf{p}_{D,P}. \tag{7}$$

To simplify the notation, a reference frame was assigned to $P_{D,O}$ to be coincident to plate DRF $\{P\}$ (see Figs. 3 and 5). The divots' DRF $\{D\}$ $x$-axis is denoted as $\mathbf{d}_{xx}$, $y$-axis as $\mathbf{d}_{yy}$, and $z$-axis as $\mathbf{d}_{zz}$, and are unit vectors ($\hat{\mathbf{d}}$):

$$\mathbf{d}_{xx} = \left( \sum_{i=1,2,3,4}^{i} \mathbf{p}_{D,O}(i) - \sum_{j=5,6,7,8}^{j} \mathbf{p}_{D,O}(j) \right) \tag{8}$$

$$\mathbf{d}'_{yy} = \left( \sum_{i=1,2,5,6}^{i} \mathbf{p}_{D,O}(i) - \sum_{j=3,4,7,8}^{j} \mathbf{p}_{D,O}(j) \right) \tag{9}$$

$$\mathbf{d}_{zz} = \mathbf{d}_{xx} \times \mathbf{d}'_{yy} \tag{10}$$

$$\mathbf{d}_{yy} = \mathbf{d}_{zz} \times \mathbf{d}_{xx}. \tag{11}$$

The origin of the divots' DRF ($\mathbf{d}_0$) is computed:

$$\mathbf{d}_0 = \frac{1}{8} \sum_{i=0}^{8} \mathbf{p}_{D,O}(i). \tag{12}$$

From a stereo camera acquisition, we got the position ($\mathbf{p}_F$) of the six selected features in the stereo camera SRF (see Fig. 5). Likewise to the divots, a reference frame was assigned to $\mathbf{p}_F$ coincident to the plate DRF $\{P\}$ (see Figs. 3 and 5). Since $\{D\}$ and $\{F\}$ are assigned coincident to the $\{P\}$, the transformation between the tracked elements (${}^D\mathbf{T}_F$) is the identity matrix,

$${}^D\mathbf{T}_F = I_4. \tag{13}$$

Being the DRF $\{F\}$ $x$-axis denoted as $\mathbf{f}_{xx}$, $y$-axis as $\mathbf{f}_{yy}$ and $z$-axis as $\mathbf{f}_{zz}$, and being unit vectors ($\hat{\mathbf{f}}$)

$$\mathbf{f}_{xx} = \left( \frac{1}{2}(\mathbf{p}_F(1) + \mathbf{p}_F(2)) - \mathbf{p}_F(6) \right) \tag{14}$$

$$\mathbf{f}'_{yy} = \left( \frac{1}{2}(\mathbf{p}_F(3) + \mathbf{p}_F(4)) - \mathbf{p}_F(5) \right) \tag{15}$$

$$\mathbf{f}_{zz} = \mathbf{f}_{xx} \times \mathbf{f}'_{yy} \tag{16}$$

$$\mathbf{f}_{yy} = \mathbf{f}_{zz} \times \mathbf{f}_{xx}. \tag{17}$$

The origin of the features' DRF ($\mathbf{f}_0$) is computed:

$$\mathbf{f}_0 = \frac{1}{4} \sum_{i=2,4,5,6}^{i} \mathbf{p}_F(i). \tag{18}$$

With the divots' $\{D\}$ and features' DRF $\{F\}$, at each synchronous acquisition from the stereo camera and CERTUS, we know the transformation from $\{D\}$ to the CERTUS SRF $\{O\}$ - ${}^O\mathbf{T}_D$ - and the transformation from $\{F\}$ to the stereo camera SRF $\{C\}$ - ${}^C\mathbf{T}_F$ (see Fig. 3).

### 4.B. Registration

To close the loop and compare both tracking systems, the transformation from the stereo camera SRF to the CERTUS SRF (${}^O\mathbf{T}_C$) was computed. The stereo camera and CERTUS remained immobile during the whole test.

The plate was placed at ten different positions and orientations, exploring the stereo camera depth of field and field

TABLE I. Test performed.

| | Static analysis | Dynamic analysis |
|---|---|---|
| Number of sets | 4 | 4 |
| Acquisitions per set | 10 | 10 |
| Frames per acquisition | 1 | 50 |

of view—keeping the features visible by the stereo camera and the markers visible by the CERTUS. It means a plate motion range volume of $0.05\times0.05\times0.1$ m$^3$ and a rotational motion within 60° around the $x$ and $y$-axis, and 360° around the $z$-axis of the stereo camera's SRF $\{C\}$. The plate motion range will henceforth be referred to as, *plate workspace*.

For each plate position, a single snapshot acquisition was recorded by both systems. In the end, we had ten DRFs of the divots in the CERTUS's SRF (${}^{O}\mathbf{T}_{D,i}$) and the ten corresponding DRFs of the features in the stereo camera's SRF (${}^{C}\mathbf{T}_{F,i}$). As seen in Fig. 3, and given (13),

$$ {}^{O}\mathbf{T}_{D,i} = {}^{O}\mathbf{T}_{C}\,{}^{C}\mathbf{T}_{F,i} \quad i = 1,\ldots,10. \tag{19}$$

An adaptation to the algorithm presented in De Momi *et al.*[38,39] was applied to compute the ${}^{O}\mathbf{T}_{C}$ transformation, which is expressed as six independent parameters: three for translation ($x$,$y$, and $z$) and three for rotation. The iterative algorithm is used to minimize the cost function F with respect to the six translation and rotation unknowns, with Levenberg–Marquardt optimization algorithm

$$ F = \| {}^{O}\mathbf{T}_{D,i}^{-1} \cdot {}^{O}\mathbf{T}_{C}\,{}^{C}\mathbf{T}_{F,i} \| \quad i = 1,\ldots,10. \tag{20}$$

In each iteration step, the translational and rotational parameters are computed against the translation and rotation errors, which are scaled to the same order of magnitude.

## 4.C. Acquisitions

A static and a dynamic analysis were performed, and the number of samples for each analysis is presented in Table I. For both static and dynamic analysis, four sets of ten acquisitions were recorded. In each static/dynamic set of acquisitions, the plate was moved in different depths relative to the stereo camera (see Fig. 6), within its workspace.

In a different study, we performed acquisitions on 12 drug resistant epileptic patients and measured the craniotomy areas, and obtained the distribution quartiles: $Q_{25} = 2107.62$ mm$^2$, $Q_{50} = 2332.75$ mm$^2$, and $Q_{75} = 2552.12$ mm$^2$. Generally, the craniotomy was ellipsoid in shape, but approximating to a circumference it leads to a craniotomy diameter between 50 and 60 mm. The SM25WI 2/3″ lenses have an intrinsic angle field of view of $20.1°\times15.1°$. At a distance of 40 cm, one camera has a field of view of $141.78\times106.03$ mm and at 50 cm, the camera field of view is $177.23\times132.54$ mm. Given that both cameras field of view is as overlapped as possible, the distance between the stereocamera and the surgical field was chosen in order for the field of view to cover a typical craniotomy.

### 4.C.1. Static acquisitions

Likewise to the registration process (cf., Sec. 4.B), the plate was placed at ten different positions and orientations within the plate workspace. For each acquisition, one snapshot of the plate is taken from the stereo camera (of features' DRF) and from CERTUS (of divots' DRF), respectively, ${}^{C}\mathbf{T}_{F}$ and ${}^{O}\mathbf{T}_{D}$.

### 4.C.2. Dynamic acquisitions

For each acquisition, the plate is manually moved within its workspace and is tracked by both stereo camera and CERTUS for 5 s at 10 frames/s. Due to the velocity requirements to track the cortical surface, we limited the linear and angular speed of the plate during the test. The linear velocity norm (referred as linear speed) and the angle from the "axis–angle representation" of the plate motion throughout all acquisitions have mean values of 5 mm/s and 1.471 deg/s and a maximum value of 25 mm/s and 9.699 deg/s, respectively.

## 4.D. Data analysis

For each static and dynamic acquisition and according to the registration from the stereo camera SRF to the CERTUS SRF, an error matrix $\mathbf{E}$ was computed:

$$ \mathbf{E} = \| {}^{O}\mathbf{T}_{D,a}^{-1} \cdot {}^{O}\mathbf{T}_{C}\,{}^{C}\mathbf{T}_{F,a} \| \quad a = \text{ each acquisition}. \tag{21}$$

In both static and dynamic analysis, the translation error $\mathbf{E}_T$ is computed as the norm of the translation component of the error matrix $\mathbf{E}$:

$$ \mathbf{E}_T = \sqrt{x^2 + y^2 + z^2}. \tag{22}$$

To represent the rotation error ($\mathbf{E}_R$), the angle from angle–axis representation is used.

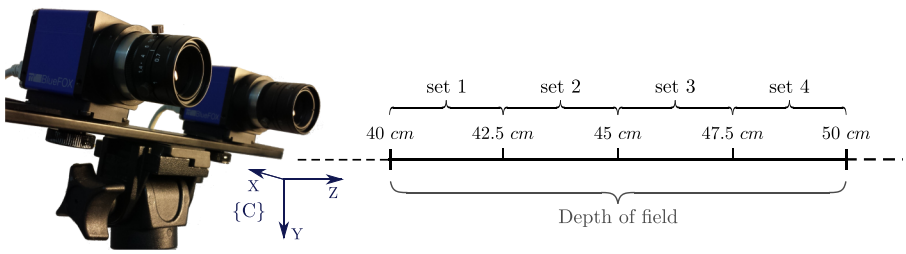


FIG. 6. Distance from the stereo camera at which the plate was tracked in each test set.

#### 4.D.1. Static analysis

A nonparametric test (Kruskal–Wallis $p < 0.05$ and Bonferroni correction) was performed to test whether the stereo camera translation and rotation errors—with an immobile target—are significantly different between each test set (different regions of the depth of field).

#### 4.D.2. Dynamic analysis

In the dynamic analysis, the tracking error was analyzed as a function of the linear and angular speed of the plate ($v$ and $\omega$). To the translation and rotation errors in each frame, it is associated the linear and angular speed of the plate, based on the CERTUS acquisitions.

Each "error-speed" measurement from each frame was grouped according to the linear and angular speed of the plate in three populations. The speed limits that define each population were assigned in order to keep the same number of samples in each population. Thus, the error-speed measurements for linear speed are grouped in (a) low linear speed ($v < 3$), (b) medium linear speed ($3 \leq v \leq 6$) and high linear speed ($v > 6$) mm/s. While the error-speed measurements for angular speed are grouped in (a) low angular speed ($\omega < 0.5$), (b) medium angular speed ($0.5 \leq v \leq 1.5$) and high angular speed ($v > 1.5$) deg/s.

A nonparametric test (Kruskal–Wallis $p < 0.05$ and Bonferroni correction) was performed to test for the null hypothesis that the error is significantly different for each population of low, medium and high linear and angular speed, in all depth of field.

## 5. RESULTS

The plane fitting error is the Euclidean distance from the detected position to the plane projected position of the feature or divot. The divot plane fitting was only performed once (cf., Sec. 4.A) and the errors present a mean, standard deviation, and maximum for the eight divots of: 0.042, 0.030, and 0.089 mm. The feature plane fitting procedure was performed in each frame of each acquisition, and presented a mean, standard deviation, and maximum error for the six features of: 0.053, 0.039, and 0.193 mm. Given the shape of the features in the plate, there is an uncertainty of detection of ±0.25 mm.

The registration matrix ${}^{O}\mathbf{T}_{C}$ minimizes the function (20). The residual errors of the iterative registration method are computed from the error matrices of all calibration acquisitions. The descriptive statistic analysis of these errors across all acquisitions is presented in Table II.

As shown in Fig. 7, the stereo camera detection error of an immobile plate across all frames is inferior to 0.6 mm and 1.6° (Fig. 7). The translation error between sets 2 and 4 are significantly different (∗ means "populations significantly different").

Besides the difference in sets 2 and 4, for the significance level of the nonparametric test performed, the null hypothesis was not rejected for either translation or rotation error, which means that there is no statistically significant impact on the error by detecting an object nearer or farther in the depth of field.

TABLE II. Registration residual errors.

| | | $\bar{x}$ | $\sigma$ | max($x$) |
|---|---|---|---|---|
| Translation (mm) | X | 0.131 | 0.096 | 0.331 |
| | Y | 0.288 | 0.183 | 0.594 |
| | Z | 0.162 | 0.105 | 0.349 |
| Rotation (deg) | $\alpha$ | 0.497 | 0.293 | 1.078 |
| | $\beta$ | 0.552 | 0.324 | 0.948 |
| | $\gamma$ | 0.573 | 0.301 | 0.913 |

As depicted in Fig. 8, the tracking error of the stereo camera for all acquisitions is inferior to 2 mm and 2.5°. The translation error median value is higher as the plate linear speed increases. According to the nonparametric statistical test performed, there is a significant difference in the translation error when the plate is moving at a different linear speed.

The statistical test also rejects the null hypothesis that the rotation error for different angular speeds fits the same distribution. Although, there is no statistically significant difference in the distribution of the rotation error between low and medium angular speeds, higher angular speed produces significantly greater rotation errors.

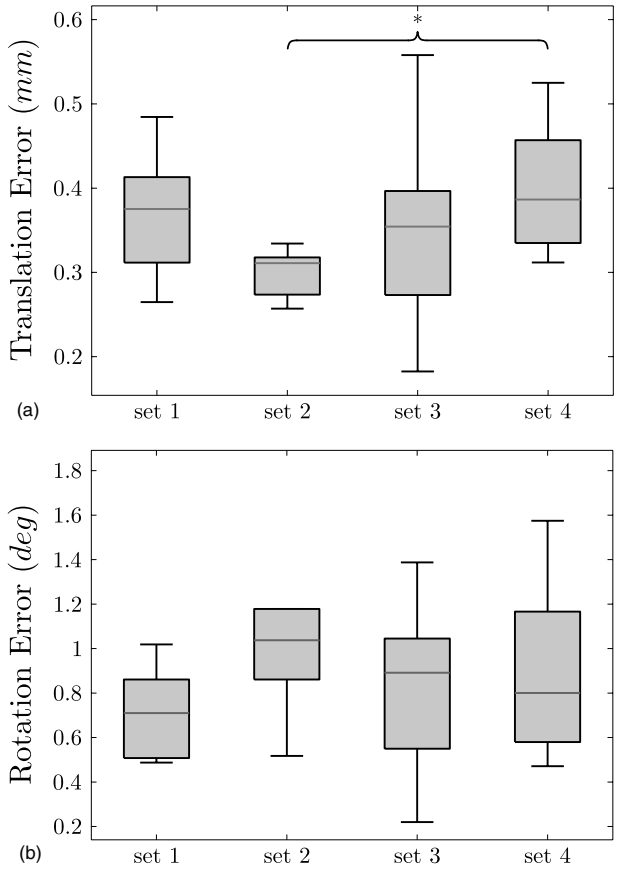


FIG. 7. Static analysis—Detection error (translation and rotation component) of an immobile object for each test set.

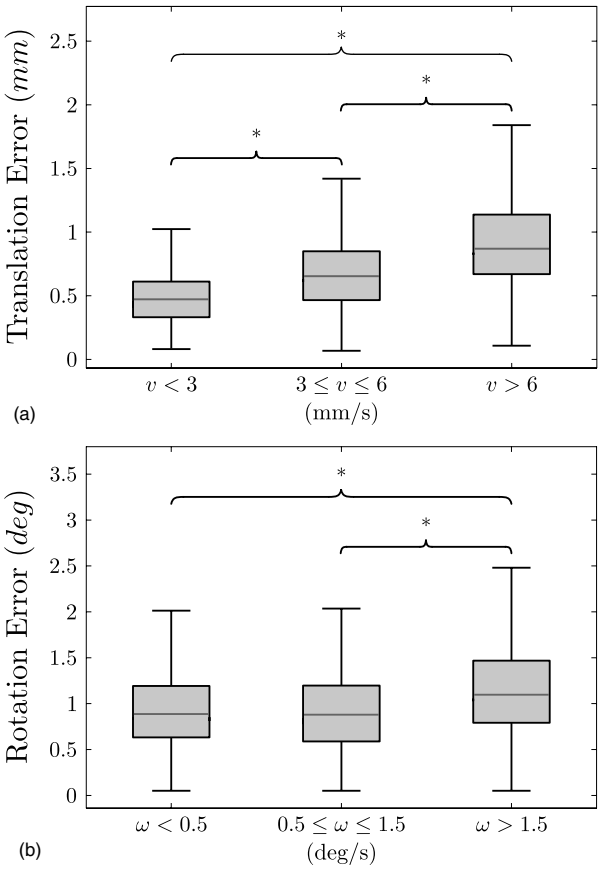


FIG. 8. Dynamic analysis—Tracking error of a moving object across all test sets. Low, medium and high linear and angular speeds depicted in left and right figures. Statistically significant difference of median translation error among all angular speeds. Statistically significant difference of median rotation error among all angular speeds, except between low and medium angular speeds.

## 6. BRAIN DATA ACQUISITION

For clinical data, three stereo image acquisitions were recorded from 53 yr old male patient, with cryptogenic epilepsy, who underwent left antero-mesial temporal lobectomy. After the craniotomy, the stereo camera was positioned on a tripod at 40 cm from the skull opening and pitched 45° downward to avoid touching the sterile drapes. Cortical surface motion was recorded in three videos of 25, 33, and 58 s at 10 frames/s.

In each acquisition, features are selected from the brain surface salient visual points and their positions in the stereo camera's SRF are tracked. In the first acquisition, 91 features were tracked, in the second video 99, and in the last video 57. To account for brain motion, the combined motion of the cloud of features was considered rather than the motion of singular features. Occluded features are dropped from the analysis.

Due to the limited hard drive access time, some frames were not recorded. Using the timestamp of each frame (cf., Sec. 2), the missing frames were identified and a cubic interpolation method was applied to reconstruct signal of the position of each feature along time. The displacement of the

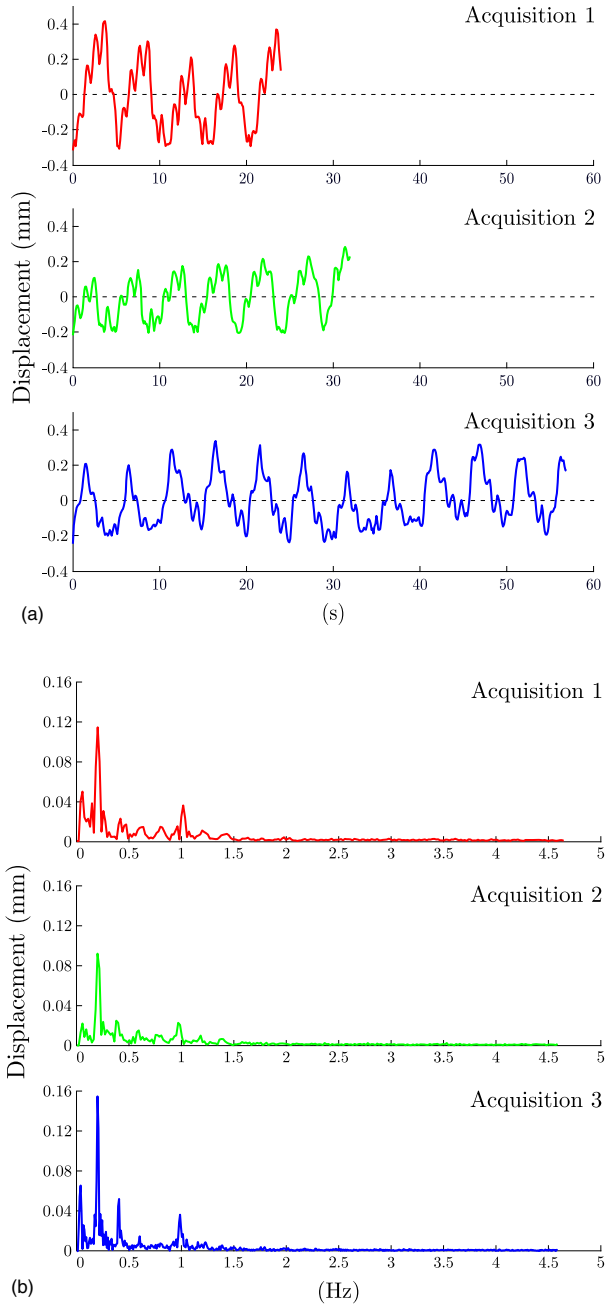


FIG. 9. Time and frequency analysis of brain motion in the three acquisitions, 25, 33, and 58 s. Left plots depict the displacement of the cloud of features' center of mass throughout the acquisition. Continuous component of the signal was subtracted. Right plots represent the FFT analysis of the displacement of the cloud of features' center of mass.

cloud of features center of mass in each acquisition frame was analyzed in the time and frequency domain, see Fig. 9.

The combined displacement of the cloud of features on the brain surface is inferior 0.8 mm in all three acquisitions. In the last acquisition frequency analysis (Fig. 9 acquisition 3), four peaks between 0 and 1 Hz are clearly visible. The first peak between 0.03 and 0.05 Hz can be attributed to pressure or temperature control (sympathovagal balance). The second peak has the biggest displacement amplitude with 0.2 Hz and is caused by breathing. After 0.2 Hz, there are

breathing harmonics with lower amplitude, which are found at integer multiples of the breathing frequency. A pronounced peak at 1 Hz is coherent with the heart rate, and is caused by the blood pressure variations. Comparing to the results presented by Ji *et al.*,[33] both show respiration motion component (around 0.2 Hz) as the major cause for brain motion followed by the blood pressure variations component (around 1 Hz).

We also evaluated each features' linear velocity norm between acquisitions. The maximum registered speed was 1.119 mm/s with a mean and standard deviation of 0.184 and 0.137 mm/s, respectively. These features velocities fit the low linear speed population of the dynamic validation analysis.

## 7. DISCUSSION

This paper presents an approach to track cortical surface shift and pulsatile deformation in real time, based on stereoscopy. The stereo tracking algorithm detects and efficiently tracks up to 100 features selected from the cortical surface, and is robust against false stereo matched points, occlusion, and outliers. Having a 10 frames/s acquisition rate, the stereo camera is able to measure not only the brain surface sag motion along time but also the brain pulsatile motion caused mainly by blood pressure variations and breathing. The ability to track the pulsatile motion is one of the breakthroughs in open skull-surgery brain tracking when compared to image based systems or laser-range scanned applications.

From the stereo vision systems for brain surface tracking able to measure the pulsatile motion presented in literature,[30,32] only Sun *et al.*[30] provided some details toward the system accuracy and validation. Their system has a point-localization average error of 0.87 mm and an average error of capturing the cortical surface shape of 1.2 mm.

A method for objective validation of a stereo vision system toward open-skull neurosurgery was also presented. The phantom-based study performs a static and dynamic evaluation of the detection and tracking capabilities of the stereo vision system, respectively, against a gold-standard optical tracking system of reference, the Optotrak CERTUS. The validation setup was thoroughly explained since the synchronization of both vision systems to the design of the phantom and relation between tracked elements.

According to the nonparametric statistic test performed, the translation and rotation error distribution was not significantly affected by the position of the plate within the stereo camera's depth of field. The dynamical analysis shows a statistically significant increase of the translation error as the linear speed of the plate increases. Only when comparing low or medium to high angular speed, the rotation errors show a significant increase.

Within the stereo camera's depth of field, which means tracking an object at a distance of 40 to 50 cm, the system proved to detect the plate position and orientation with an error inferior to 0.5 mm and 1.6° which is inferior to the reported point-localization accuracy of.[30] Moreover, considering the maximum feature speed of 1.119 mm/s, which is sensible compared to the maximum brain motion values reported by Greitz *et al.*[11] of 1.5 mm/s, both fit low linear speed population of the dynamic validation analysis, which presents a median linear error of 0.5 mm and the maximum 1.03 mm.

The stereo vision system was tested in the operating room. Cameras were placed at a safe distance from the patient, and three acquisition videos were acquired during 25, 33, and 58 s at 10 frames/s. The cortical pulsatile motion amplitude quantified after the center of mass of the features was inferior to 0.8 mm. Breathing is the largest contribution to cortical brain motion amplitude with frequency of 0.2 Hz.

Due to the short duration of the video recordings, it was not possible to quantify the brain shift. An integration of the stereo vision system with the surgical equipment would enable the quantification of brain shift and pulsatility, which could potentially be applied into surgical field stabilization and robotic motion compensation.

Tracking brain shift deformation, before the surgeons perform brain resection, could benefit the intraoperative functional mapping phase in which the surgeon stimulates brain cortex. Although not reliable as 3D imaging techniques like CT and MRI to measure the deformation of deep brain structures,[25] cortical surface information can be used by biomechanical models to update the brain model displayed on the navigation system. As further development, we plan to incorporate a real-time volumetric deformation model that extends the estimation of the deformation from the cortical surface to brain deep structures.[40] It will allow an easier correlation between the functional and the anatomical mapping. Nonetheless, once the tissue is removed, the features cannot be tracked anymore.

Another limitation is due to the possible interference of the cameras to the surgeon movements on the operating field, this could be solved using a microscope, with 3D cameras option. Since this can be displaced during the surgical procedure, it should be continuously tracked with an external localization system.

## ACKNOWLEDGMENTS

This work has been partially financed by the ACTIVE Project (FP7-ICT-2009-6-270460), NETT Project (FP7-PEOPLE-2011-ITN-289146), and FCT PhD grant (ref. SFRH/BD/86499/2012).

[a)]Author to whom correspondence should be addressed. Electronic mail: Electronic mail: cfaria@dei.uminho.pt

[1]D. Hill, C. Maurer, R. Maciunas, J. Barwise, J. Fitzpatrick, and M. Wang, "Measurement of intraoperative brain surface deformation under a craniotomy," Neurosurgery **43**, 514–526 discussion 527–528 (1998).

[2]N. Hata, A. Nabavi, S. Warfield, W. Wells, R. Kikinis, and F. Jolesz, "A volumetric optical flow method for measurement of brain deformation from intraoperative magnetic resonance images," in *Medical Image Computing and Computer-Assisted Intervention*, edited by C. Taylor and A. Colchester (Medical Image Computing and Computer-Assisted Intervention, Cambridge, 1999), pp. 928–935.

[3]D. Roberts, A. Hartov, F. Kennedy, M. Miga, and K. Paulsen, "Intraoperative brain shift and deformation: A quantitative analysis of cortical displacement in 28 cases," Neurosurgery **43**, 749–758, discussion 758–760 (1998).

[4]P. Grunert, K. Darabi, J. Espinosa, and R. Filippi, "Computer-aided navigation in neurosurgery," Neurosurg. Rev. **26**, 73–99 (2003).
[5]G. Eggers, J. Mühling, and R. Marmulla, "Image-to-patient registration techniques in head surgery," Int. J. Oral Maxillofac. Surg. **35**, 1081–1095 (2006).
[6]J. Cohen-Adad, P. Paul, X. Morandi, and P. Jannin, "Knowledge modeling in image-guided neurosurgery: Application in understanding intraoperative brain shift," in *Proceedings of the Medical Imaging: Visualization, Image-Guided Procedures and Display*, edited by K. R. Cleary and R. L. Galloway Jr., (SPIE, San Diego, 2006), Vol. 6141.
[7]A. Linninger, C. Tsakiris, D. Zhu, M. Xenos, P. Roycewicz, Z. Danziger, and R. Penn, "Pulsatile cerebrospinal fluid dynamics in the human brain," IEEE Trans. Biomed. Eng. **52**, 557–565 (2005).
[8]M. Wagshul, P. Eide, and J. Madsen, "The pulsating brain: A review of experimental and clinical studies of intracranial pulsatility," Fluids Barriers CNS **8**, 5 (2011).
[9]R. Britt and G. Rossi, "Quantitative analysis of methods for reducing physiological brain pulsations," J. Neurosci. Methods **6**, 219–229 (1982).
[10]B. Poncelet, V. Wedeen, R. Weisskoff, and M. Cohen, "Brain parenchyma motion: Measurement with cine echo-planar MR imaging," Radiology **185**, 645–651 (1992).
[11]D. Greitz, R. Wirestam, A. Franck, B. Nordell, C. Thomsen, and F. Ståhlberg, "Pulsatile brain movement and associated hydrodynamics studied by magnetic resonance phase imaging," Neuroradiology **34**, 370–380 (1992).
[12]P. Black, E. Alexander, C. Martin, T. Moriarty, A. Nabavi, T. Wong, R. Schwartz, and F. Jolesz, "Craniotomy for tumor treatment in an intraoperative magnetic resonance imaging unit," Neurosurgery **45**, 423–431, discussion 431–433 (1999).
[13]C. Nimsky, O. Ganslandt, S. Cerny, P. Hastreiter, G. Greiner, and R. Fahlbusch, "Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging," Neurosurgery **47**, 1070–1079, discussion 1079–1080 (2000).
[14]M. Hadani, R. Spiegelman, Z. Feldman, H. Berkenstadt, and Z. Ram, "Novel, compact, intraoperative magnetic resonance imaging-guided system for conventional neurosurgical operating rooms," Neurosurgery **48**, 799–807, discussion 807–809 (2001).
[15]L. Lunsford, D. Kondziolka, and D. Bissonette, "Intraoperative imaging of the brain," Appl. Neurophysiol. **66**, 58–64 (1996).
[16]J. Koivukangas, Y. Louhisalmi, J. Alakuijala, and J. Oikarinen, "Ultrasound-controlled neuronavigator-guided brain surgery," J. Neurosurg. **79**, 36–42 (1993).
[17]A. Jödicke, W. Deinsberger, H. Erbe, A. Kriete, and D. Böker, "Intraoperative three-dimensional ultrasonography: An approach to register brain shift using multidimensional image processing," Minimally Invasive Neurosurg. **41**, 13–19 (1998).
[18]R. Comeau, a. F. Sadikot, A. Fenster, and T. Peters, "Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery," Med. Phys. **27**, 787–800 (2000).
[19]V. Tronnier, M. Bonsanto, A. Staubert, M. Knauth, S. Kunze, and C. Wirtz, "Comparison of intraoperative MR imaging and 3D-navigated ultrasonography in the detection and resection control of lesions," Neurosurg. Focus **10**, E3 (2001).
[20]G. Unsgaard, A. Gronningsaeter, S. Ommedal, and T. Nagelhus Hernes, "Brain operations guided by real-time two-dimensional ultrasound: New possibilities as a result of improved image quality," Neurosurgery **51**, 402–411, discussion 411–412 (2012).
[21]O. Skrinjar, A. Nabavi, and J. Duncan, "Model-driven brain shift compensation," Med. Image Anal. **6**, 361–373 (2002).
[22]H. Sun, H. Farid, K. Rick, A. Hartov, D. Roberts, and K. Paulsen, "Estimating cortical surface motion using stereopsis for brain deformation models," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, edited by R. Ellis and T. Peters (Springer Berlin Heidelberg, Montréal, 2003), Vol. 2878, pp. 794–801.
[23]P. Jannin, C. Grova, and C. Maurer, "Model for defining and reporting reference-based validation protocols in medical image processing," Int. J. Comput. Assisted Radiol. Surg. **1**, 63–73 (2006).
[24]A. Wittek, K. Miller, R. Kikinis, and S. Warfield, "Patient-specific model of brain deformation: Application to medical image registration," J. Biomech. **40**, 919–929 (2007).
[25]M. Ferrant, A. Nabavi, B. Macq, F. Jolesz, R. Kikinis, and S. Warfield, "Registration of 3-D intraoperative MR images of the brain using a finite-element biomechanical model," IEEE Trans. Med. Imaging **20**, 1384–1397 (2001).
[26]P. Paul, X. Morandi, and P. Jannin, "A surface registration method for quantification of intraoperative brain deformations in image-guided neurosurgery," IEEE Trans. Inf. Technol. Biomed. **13**, 976–983 (2009).
[27]T. Sinha, B. Dawant, V. Duay, D. Cash, R. Weil, R. Thompson, K. Weaver, and M. Miga, "A method to track cortical surface deformations using a laser range scanner," IEEE Trans. Med. Imaging **24**, 767–781 (2005).
[28]M. Miga, T. Sinha, D. Cash, R. Galloway, and R. Weil, "Cortical surface registration for image-guided neurosurgery using laser-range scanning," IEEE Trans. Med. Imaging **22**, 973–985 (2003).
[29]A. Cao, R. Thompson, P. Dumpuri, B. Dawant, R. Galloway, S. Ding, and M. Miga, "Laser range scanning for image-guided neurosurgery: Investigation of image-to-physical space registrations," Med. Phys. **35**, 1593–1605 (2008).
[30]H. Sun, D. Roberts, H. Farid, Z. Wu, A. Hartov, and K. Paulsen, "Cortical surface tracking using a stereoscopic operating microscope," Neurosurgery **56**, 86–97 (2005).
[31]P. Paul, O. Fleig, and P. Jannin, "Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: Methods and performance evaluation," IEEE Trans. Med. Imaging **24**, 1500–1511 (2005).
[32]S. Ji, X. Fan, D. Roberts, and K. Paulsen, "Cortical surface strain estimation using stereovision," Med. Image Comput. Comput.-Assisted Intervention **14**, 412–419 (2011).
[33]S. Ji, X. Fan, D. Roberts, A. Hartov, and K. Paulsen, "Tracking Cortical Surface Deformation Using Stereovision," in *Conference Proceedings of the Society for Experimental Mechanics Series* (Springer, New York, NY, 2013), pp. 169–176.
[34]B. Atcheson, F. Heide, and W. Heidrich, "CALTag: High precision fiducial markers for camera calibration," in *Vision, Modeling, and Visualization* (Eurographics Association, Siegen, 2010), pp. 41–48.
[35]C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the Alvey Vision Conference 1988* (Alvey Vision Club, Manchester, 1988), pp. 23.1–23.6.
[36]J. Lewis, "Fast template matching," Vision Interface (1995).
[37]C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," Pattern Recognit. **32**, 71–86 (1999).
[38]E. De Momi, P. Cerveri, E. Gambaretto, M. Marchente, O. Effretti, S. Barbariga, G. Gini, and G. Ferrigno, "Robotic alignment of femoral cutting mask during total knee arthroplasty," Int. J. Comput. Assisted Radiol. Surg. **3**, 413–419 (2008).
[39]M. D. Comparetti, A. Vaccarella, I. Dyagilev, M. Shoham, G. Ferrigno, and E. De Momi, "Accurate multi-robot targeting for keyhole neurosurgery based on external sensor monitoring," Proc. Inst. Mech. Eng., Part H **226**, 347–359 (2012).
[40]I. Rasin, Z. Pekar, O. Sadowsky, A. Forte, S. Galvan, D. Dini, M. Shoham, and L. Joskowicz, "Real-time modeling of intra-operative brain shift based on video tracking," in *Proceedings of The Hamlyn Symposium on Medical Robotics, London* (2014).