Accurate registration of random radiographic projections based on three spherical references for the purpose of few-view 3D reconstruction.

Ralf Schulze* and Dan Bruellmann
Dept. of Oral Surgery (and Oral Radiology), Johannes Gutenberg-University

Ulrich Heil,† Oliver Weinheimer,† Daniel Gross,† and Elmar Schoemer†
Institute of Computer Science
Johannes Gutenberg-University

Eric Thomas‡ and Ulrich Schwanecke‡
Dept. of Design, Computer Science and Media
Wiesbaden University of Applied Sciences

Precise registration of radiographic projection images acquired in almost arbitrary geometries for the purpose of 3D reconstruction is beset with difficulties. We modify and enhance a registration method presented in[20] based on coupling a minimum amount of three reference spheres in arbitrary positions to a rigid object under study for precise a posteriori pose estimation. Two consecutive optimization procedures (a, initial guess; b, iterative coordinate refinement) are applied to completely exploit the reference's shadow information for precise registration of the projections. The modification has been extensive, i.e. only the idea of using the sphere shadows to locate each sphere in three dimensions from each projection was retained whereas the approach to extract the shadow information has been changed completely and extended. The registration information is used for subsequent algebraic reconstruction of the 3D information inherent in the projections.
We present a detailed mathematical theory of the registration process as well as simulated data investigating its performance in the presence of error. Simulation of the initial guess revealed a mean relative error in the critical depth coordinate ranging between 2.1% and 4.4%, and an evident error reduction by the subsequent iterative coordinate refinement. To prove the applicability of the method for real-world data, algebraic 3D reconstructions from few ($\leq$ 9) projection radiographs of a human skull, a human mandible and a teeth-containing mandible segment are presented. The method facilitates extraction of 3D information from only few projections obtained from off-the-shelf radiographic projection units without the need for costly hardware. Technical requirements as well as radiation dose are low.

I. INTRODUCTION

Established techniques for radiographic three-dimensional (3D) reconstruction such as computed tomography (CT) are based on an extensive set of projections from all around the object. Since the imaging geometry of each and every projection has to be precisely known a priori, large-scale scanners operating with sophisticated hardware technology are required. Nowadays the quality of CT images is excellent but the radiation dose is still very high since it is directly related to the number of projections and total exposure time. Furthermore CT scanners are expensive high-tech machines, which are not readily available everywhere. Recently, increasing interest has been directed towards three-dimensional reconstructions from few two-dimensional (2D) radiographic projections,[1–7] frequently also suffering from limited angular range exposure.[3–6] Although in theory some of the techniques may be extended to random geometries, however, the latter mostly had been constrained to a circular orbit and equidistant projection.[2, 4, 5, 8] We propose a technique opposing the CT-philosophy by using only very few 2D radiographic projections, which, in addition, may be freely positioned around the object in random geometries. This approach however, poses two major challenges. First, 3D reconstruction from projections requires the precise knowledge of the imaging geometry the projections have been acquired in.[5, 9] Second, 3D reconstruction from few views distributed over a limited angular range is an inherently ill-posed problem.[10] Recent results from implementation of knowledge about the objects under investigation seem promising with respect to suppression of few-view and limited-angle artifacts.[3–5] This issue, however does not fall within the scope of this work. We concentrate on registration of the projections, which remains problematic if the latter have been acquired in almost

*Dept. of Oral Surgery (and Oral Radiology), Johannes Gutenberg-University; Electronic address: rschulze@mail.uni-mainz.de; URL: www.klinik.uni-mainz.de/ZMK/Oralchir/schulze.html
†URL: www.informatik.uni-mainz.de/247.php
‡URL: www.mi.fh-wiesbaden.de/~schwan/Projects/3DReco/

arbitrary geometries. Although extensive work has been published on image registration in general (for detailed review see [11, 12]), the precise radiographic projection registration for the purpose of extracting 3D information has rarely been focused on lately. The registration of 2D radiographic projections is substantially more complex than registering 2D photographs. The main problem is the loss of information resulting from the many-to-one mapping in summation images. It is important to realize here, that the prerequisite to extract 3D information from few projections is their formation from different imaging geometries. This, vice versa necessarily yields very different images, particularly if the degrees of freedom in the imaging geometry are high. Together with the lack of point-like features in radiographs this fact makes accurate registrations of such projections a challenging task. Registration basing on the well-known general eight-point algorithm [13] to reconstruct a scene from planar projections had been adopted to radiographic situations.[14–16] Yet the technique is rather inaccurate due to its strong dependence on accurate input coordinates.[17] 3D reconstructions from real world radiographic projections have been presented for coronary vessels using internal landmarks in biplane images.[18] Here, however gantry information was used to supplement the geometric informations acquired from the projections. In nondestructive material testing a rather large reference object providing massive redundant information was used for the reconstruction of industrial objects from unconstrained radiographic views.[19] An interesting and novel approach estimating the geometry together with the 3D object by means of sophisticated application of Bayesian theory has been published recently.[8] Albeit being revolutionary in nature, however, it is computationally extremely demanding and, at this stage, still requires assumptions on the imaging geometry as well as a meaningful initial guess on the 3D object being reconstructed.

Except of the recent paper just introduced, we are not aware of other published solutions using intrinsic references capable to register radiographic projections precisely enough to allow for 3D reconstruction when no additional geometric information is available. Extrinsic reference bodies, however are inconvenient in application, may occlude image information or induce artifacts in the reconstruction. Hence, if necessary, their number should be limited to an absolute minimum and they should be rather small in size. An arbitrary radiographic projection system contains nine image-relevant degrees of freedom: six degrees of the rigid object (translations + rotations) plus three degrees of the source relative to the (flat) image-receptor. At least in theory, up to eight degrees of freedom may be registered by means of three reference spheres attached to an object under study in such, that their center-of-gravity points (COP) form an arbitrary triangle.[20] Using the elliptically distorted sphere shadow as input information, one obtains three points in space per projection, i.e. the three COPs. This is enough information for rigid pose estimation.[21] Since one obtains $3 \cdot 3$ instead of $3 \cdot 2$ coordinates from each projection for registration, this approach differs fundamentally from the common use[7, 9] of (spherical) fiducials, where only their 2D image-centroid is considered. We base our registration-technique on this principal idea, however, the mathematics to extract the information within the shadow(s) is completely different. Contrary to[20] we do not only consider the major elliptical axis of each sphere shadow but exploit the shadow information by means of optimization methods. To extract additional information, we do not only consider each shadow separately, but also the combination of i) the three shadows in one projection and ii) the shadows in all k projections simultaneously. This enhances the method in[20] to a stage where completely arbitrary object pose may be registered with sufficient accuracy for real world radiographic 3D reconstruction. The complete theory of our approach is introduced, which consists of two consecutive optimization procedures. The first makes use of the entire reference information within each projection, whereas the second processes the information from all projections available simultaneously. The robustness of the registration process is tested using simulated data. To demonstrate the usefulness of the method, in the sequel referred to as Reference-Sphere-Method (RSM), algebraic 3D reconstructions from few real world radiographic projections registered with the proposed algorithms are presented. Experimental radiographs of a human mandible segment, a human dry skull and a dry mandible exposed with varying degrees of freedom are used for the reconstructions. Although we include a section explaining the 3D reconstruction algorithms applied here, the general few-view and limited angle problem, however, fall not within the scope of this paper.

II. METHODS

A. Image segmentation and quantitative ellipse evaluation

The areas G_A,G_B and G_C covered by the shadows of the three reference spheres A, B, C are segmented by a circular Hough-Transform using a sobel operator followed by the computation of the barycenter $\mathbf{c}$ within each shadow. We search for the best fitting ellipse represented by the points belonging to each sphere shadow. The quantitative evaluation of the shadow's point-cloud has to be very precise to ensure accurate registration. Consequently, finding the best possible estimate for each ellipse is of crucial importance. Using $\mathbf{c}$ as starting point, n radial scan lines extending beyond the segmented shadow area are constructed and their individual intensity profile is evaluated to determine n boundary points of the ellipse with subpixel accuracy (see. Fig. 1). We obtained good results for $n = 256$.

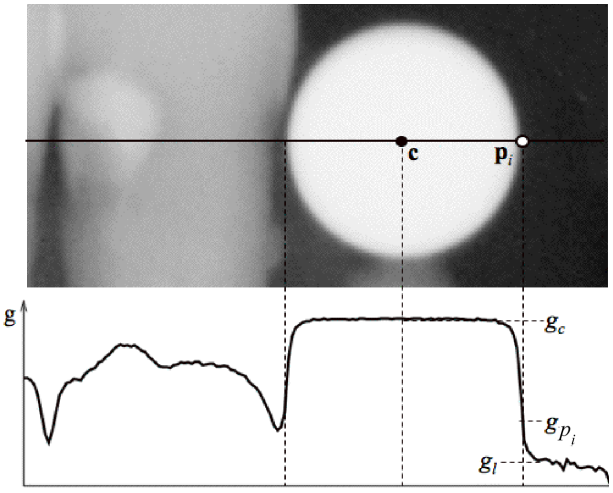

FIG. 1: Extract of a radiograph displaying one scanline, the intensity-profile of which is plotted underneath. Due to the scaling of the x-axis and the radiodensity of the steel sphere, the profile of the sphere shadow appears rather flat. The exact ellipse boundary point $\mathbf{p}_i$ is located somewhere between the gray value g_c at the barycenter $\mathbf{c}$ and the point with local minimum gray g_l. An estimate g_{p_i} is obtained from eq. (1). For details see text.

X-ray absorption decreases non-linearly towards the sphere's boundary due to its spherical geometry. Consequently the true sphere-image boundary is disguised by background noise and located further peripherally than the intensity profile would indicate at first sight. Taking this into account, we prepared one calibration series of nine images with known sphere positions altered in a step-wise fashion. It should be noted, that this calibration was carried out only once, i.e. for the dental ccd-sensor (see below). Using this series, a simple empirical approach was designed, modeling intrinsic system parameters. Assuming that the most accurate shadow definition is provided when it is surrounded by air only, we defined a global minimum gray value (air) $g_g > 0$ for each projection, whereas local minima on the scan lines are denoted by g_l where $g_l \geq g_g$. In the subpixel domain, the signal convolved with the point spread function may be approximated by a continuous signal composed of 2D-Gaussians. Consider a gray profile plotted along one scanline as in Fig. 1. Obviously, the true ellipse boundary point $\mathbf{p}_i$ (with gray g_{p_i}) must be located somewhere between $\mathbf{c}$ (with gray-value g_c) and the point, where a local minimum gray g_l is found. An estimate of g_{p_i} can be obtained from:

$$g_{p_i} = g_l + y\,(g_c - g_l)\,. \tag{1}$$

where, by modification of the well known "full width at half maximum" technique, we establish a "full width at y% of the maximum" criterion:

$$y = \max\left(\omega_{\min}, \omega_{\max} - \rho\left(\frac{g_l}{g_g} - 1\right)\right). \tag{2}$$

Here, $\omega_{\min}$, $\omega_{\max}$ and ρ are normalized scaling factors defining the interval, wherein the actual border $\mathbf{p}_i$ (indicated by a gray value g_{p_i}) of the ellipse is to be found. Although they were determined empirically from only one calibration series using the dental ccd-receptor specified below, we obtained good results for $\omega_{\min} = 0.1$, $\omega_{\max} = 0.4$ and $\rho = 0.03$ for all series shown in this work. It should be noted, however, that specific calibration to a certain radiographic system essentially means adapting these factors in a calibration process.

By piecewise linear interpolation between neighboring pixels along the scanline the individual boundary point $\mathbf{p}_i$ can be estimated with subpixel accuracy. In rare cases, where due to superimposition of highly dense objects $\mathbf{p}_i$ cannot be determined, that particular scanline is neglected. We now have defined a set of boundary points $\mathbf{p}_i$ with $i \in \{1, ..., n\}$ for each elliptical sphere shadow.

B. Computing the initial guess

The boundary points $\mathbf{p}_i$ represent the input for the following procedure designed to compute an initial guess on the 3D coordinates of each COP. This stage of the registration procedure utilizes the image information of each projection separately. The general radiographic system we consider here consists of a point-like x-ray source $\mathbf{e}$ located at a known distance d from the flat image-plane. As shown in,[20] a sphere of known radius R may be located in 3D by carefully evaluating its 2D elliptically distorted radiographic shadow. It is important to realize, that by locating three non-collinear spheres in 3D from each projection, the absolute minimum information for pose registration is provided: three points in space forming a non-degenerated triangle.[21] While the authors in[20] only considered the major ellipse axis, we select a more complete approach implementing conic sections theory. The x-rays tangential to the reference sphere's surface form a cone with symmetry axis $\mathbf{r}$, apex $\mathbf{e}$ and half-opening angle φ (see Fig. 2a) given by the standard cone equation

$$\mathbf{r}^{\mathbf{T}}\,(\mathbf{x} - \mathbf{e}) = \cos\varphi\,|\mathbf{x} - \mathbf{e}| \tag{3}$$

with side condition $|\mathbf{r}| = 1$. If we intersect the cone with the image-plane, we obtain an ellipse with diameters λ_1, λ_2. Assuming a known position $\mathbf{e}$ of the focal point, which is a reasonable assumption for the majority of medical radiographic devices, we search for a cone with known apex $\mathbf{e}$ optimally fitting into the set of ellipse boundary points $\mathbf{p}_i$ within the image-plane. The unknown symmetry axis $\mathbf{r}$ and half-opening angle φ of the cone can be determined by solving the equation

$$\sum_{i=1}^{n}\left(\mathbf{r}^{\mathbf{T}}\mathbf{q}_i - \cos\varphi\right)^2 \longrightarrow \min \tag{4}$$

where $\mathbf{q}_i = \frac{\mathbf{p}_i - \mathbf{e}}{|\mathbf{p}_i - \mathbf{e}|}$ represent the x-ray directions.

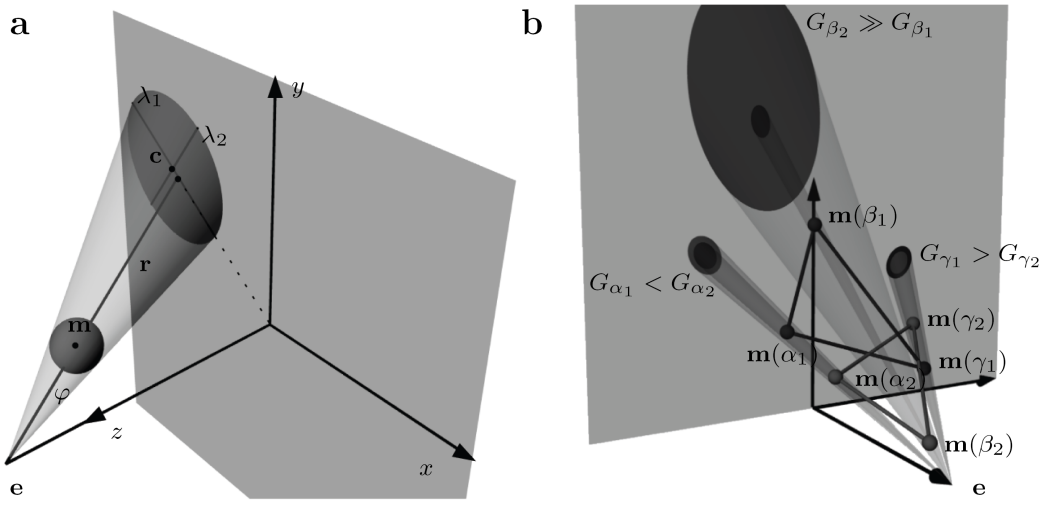

FIG. 2: a) Sketch of the exposure geometry in a Cartesian coordinate system. A sphere (COP: $\mathbf{m}$) exposed from a focal point located at a known position $\mathbf{e}$ fits into the cone of x-rays tangential to its surface, with half-opening angle φ and symmetry axis $\mathbf{r}$. It casts an elliptical shadow (barycenter: $\mathbf{c}$, axis lengths: λ_1, λ_2) defined by the intersection of the cone with the image-plane. b) The triangle spanned by the three reference sphere COPs $\mathbf{m}(\alpha), \mathbf{m}(\beta), \mathbf{m}(\gamma)$ at two arbitrary positions (indices: $1, 2$) yields different elliptical shadows (area: G). Note, that only for illustration purposes two possible triangle-positions are shown. The difference between measured (G_A,G_B,G_C) and computed (G_α,G_β,G_γ) areas covered by the sphere's shadow is minimized to obtain a best fit position.

Due to the normalization of $\mathbf{r}$ only the three parameters φ and two of the components of $\mathbf{r}$ have to be found. To account for the side condition $|\mathbf{r}| = 1$, the Lagrange-function L with factor μ is used:

$$L\left(\mathbf{r}, \cos\varphi\right) = \sum_{i=1}^{n} \left(\mathbf{r^T q}_i - \cos\varphi\right)^2 - \mu\left(\mathbf{r}^2 - 1\right). \quad (5)$$

The essential constraints for minimizing eq. (5) are: $\frac{\partial L}{\partial \cos\varphi} = 0$ and $\frac{\partial L}{\partial \mathbf{r}} = 0$. Thus, we receive:

$$\cos\varphi = \mathbf{r^T s} \quad \text{with} \quad \mathbf{s} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{q}_i \quad (6)$$

and

$$\mathbf{Q}\cdot\mathbf{r} = \mu\mathbf{r} \text{ with } \mathbf{Q} = \sum_{i=1}^{n} \mathbf{q}_i \left(\mathbf{q}_i - \mathbf{s}\right)^{\mathbf{T}} \in \mathbb{R}^{3\times 3}. \quad (7)$$

The matrix $\mathbf{Q}$ represents the sum of the dyadic products of the vectors $\mathbf{q}_i$ and $\mathbf{q}_i - \mathbf{s}$. From eq. (7) follows

$$\mathbf{Q} = \sum_{i=1}^{n} \mathbf{q}_i\mathbf{q}_i^{\mathbf{T}} - n\mathbf{s}\mathbf{s}^{\mathbf{T}} = \sum_{i=1}^{n} (\mathbf{q}_i - \mathbf{s})(\mathbf{q_i} - \mathbf{s})^{\mathbf{T}} = \mathbf{Q^T}. \quad (8)$$

Equations (6) and (8) indicate that $\mathbf{s}$ is the expected value of the x-ray directions $\mathbf{q}_i$ whereas $\mathbf{Q}$ represents their covariance matrix. Referring to the cone determination, we have reduced our problem to finding the eigenvectors of the covariance matrix $\mathbf{Q}$. The eigenvector with the smallest eigenvalue represents the symmetry axis $\mathbf{r}$ of our target cone (see Fig. 2a). Finally, the 3D position of the reference sphere's COP is given by $\eta\mathbf{r}$ with $\eta = R/\sin\varphi$, where φ is computed from equation (6). Again, it is important to realize, that from the initial guess we obtain independent coordinates of three points in space of which we know, that they de facto form a rigid triangle fixed somewhere to a (rigid) object.

C. Focal point determination

At least in theory, the position of the focal point may also be determined from analysis of the elliptical distorted shadows. It has been found, that the major ellipse axes necessarily intersect in the projection of the optical center on the image plane.[20, 22] This knowledge may be used to extend the algorithm to situations, where only the distance between focal spot and imaging plane is known, i.e. two more degrees of freedom are involved. Under this prerequisite, the search space for $\mathbf{e}$ is a two-dimensional plane parallel to the receptor-plane. We introduce a function Z measuring the consistency between calulated cones (eqs. (3) to (8)) and the observed boundary points $\mathbf{p}_i$ for a given focal point $\mathbf{e}$. Hence, by searching over the appropriately bounded two-dimensional search we solve

$$Z(\mathbf{e}) = \sum_{i=1}^{n}\sum_{v=1}^{3} \left(\mathbf{r}^{(v)\mathbf{T}}\mathbf{q}_i^{(v)} - \cos\varphi^{(v)}\right)^2 \longrightarrow \min, \quad (9)$$

where $\mathbf{q}_i^{(v)}$ define the x-ray directions on the vth ellipse. The solution of eq. (9) indicates a best fit location of $\mathbf{e}$. Preliminary results on real data seem promising, if elliptical distortion is large ($\lambda_1 \gg \lambda_2$). If $\lambda_1 \approx \lambda_2$, however, at this stage numerical instability renders determination of $\mathbf{e}$ impractical. Future work will focus on further investigation of this issue.

D. Coordinate-Refinement

From the initial guess we obtained estimated coordinates of three points in space (i.e. the COPs) for every single projection. These three points are the fundament on which we aim to compute the rigid transformation between the views. Due to error in the initial estimates, however, the triangle spanned between the points may vary considerable between the projections. On the other hand we securely know that in reality the triangle must be rigid (under the assumption of a rigid object). In other words, the knowledge that the triangle is indeed rigid may be used to generate a best fit rigid transformation from an erroneous, non-rigid one. In the following we formulate a fitting algorithm making use of exactly this a priori information.

1. *Triangle Fit – Theory*

Let $l_{\alpha\beta}$, $l_{\beta\gamma}$ and $l_{\alpha\gamma}$ be the side lengths of a given triangle. Note, that at this stage we assume the side lengths of the triangle are known and constant. We define unit direction vectors $\mathbf{r_a}, \mathbf{r_b}, \mathbf{r_c}$ pointing along the projection lines, i.e. the symmetry axes $\mathbf{r}$ of the cones (Fig. 2). The location of the triangle vertices are parameterized by $\mathbf{m}(\alpha) = \alpha\mathbf{r_a}$, $\mathbf{m}(\beta) = \beta\mathbf{r_b}$ and $\mathbf{m}(\gamma) = \gamma\mathbf{r_c}$. From the incidence constraints for the triangle vertices and the line equations we obtain three second order equations in the parameters α, β, γ:

$$\begin{aligned} (\alpha\mathbf{r_a} - \beta\mathbf{r_b})^2 - l_{\alpha\beta}^2 &= 0\,, \\ (\beta\mathbf{r_b} - \gamma\mathbf{r_c})^2 - l_{\beta\gamma}^2 &= 0\,, \\ (\gamma\mathbf{r_c} - \alpha\mathbf{r_a})^2 - l_{\alpha\gamma}^2 &= 0\,. \end{aligned} \tag{10}$$

Eq. (10) is well-established for pose estimation in photogrammetry and computer vision.[23] The Bézout bound of this polynomial system is $2 \cdot 2 \cdot 2 = 8$ solutions. Through resultant elimination of β and γ, we receive an univariate polynomial in α of degree 8. As all impair coefficients of the polynomial disappear, the solutions for α^2 correspond to a polynomial of degree 4. Thus, at maximum 8 solution triples (α, β, γ) are feasible. Observe that we only have to consider positive triples solving (10). The solution set for image i with $i \in \{1, \ldots, k\}$ is given by

$$S_i := \{(\alpha, \beta, \gamma) | \alpha, \beta, \gamma > 0 \text{ solutions of (10) for image } i\}\,.$$

By fitting the triangle into the three projection lines we obtain all possible sphere locations. Fig. 2b exemplarily shows two possible solution sets. By comparing the fitted shadow areas $G^{\alpha\beta\gamma} := (G_\alpha, G_\beta, G_\gamma)$ and the measured areas $G^{ABC} := (G_A, G_B, G_C)$, we obtain the optimal $(\alpha_o, \beta_o, \gamma_o)$ among the eight possible solutions of eq. (10) for image i from

$$(\alpha_o, \beta_o, \gamma_o) = \arg \min_{(\alpha,\beta,\gamma)\in S_i} \left\| G^{\alpha\beta\gamma} - G^{ABC} \right\|^2 . \tag{11}$$

2. *Triangle fit – Coordinate refinement*

We now extend the TF to the real situation, where due to errors in the initial guess (section II B) registration, the triangles differ between the projections. For this purpose, we introduce an optimization function F, which measures the squared differences between measured G_A,G_B,G_C and computed G_α,G_β,G_γ elliptical shadow areas for the given triangle side lengths $l_{\alpha\beta}$, $l_{\beta\gamma}$, $l_{\alpha\gamma}$:

$$F(l_{\alpha\beta}, l_{\beta\gamma}, l_{\alpha\gamma}) = \sum_{i=1}^{k} \min_{(\alpha,\beta,\gamma)\in S_i} ||G^{\alpha\beta\gamma} - G^{ABC}||^2 . \tag{12}$$

From the different triangles obtained from the initial guess, we compute a mean triangle by calculating the center of mass points between the sphere positions for each vertex. Around this starting triangle, the search space is limited to an appropriate interval, e. g. the mean $\pm$ standard deviation. For each given triangle, the squared area difference between measured values G_A,G_B,G_C and computed values G_α,G_β,G_γ is recorded. Due to the limited search space brute force methods perform quite well and protect against erroneous local minima, if the variation in the triangle lengths is high. In other words, starting from a mean triangle its side lengths are varied in such, that the area-error between measured and computed sphere shadows is minimized over k projections. The minimum indicates an optimal position of the projections relative to one another, i.e. a best fit geometric transformation between the views. Hence, TF does not necessarily produce accurate 3D coordinates of the sphere's COP in a world coordinate system, rather it enhances accuracy of relative positions.

E. 3D Reconstruction

Algebraic reconstructions (ART, for examples refer to[24]) were obtained by iteratively solving the sparse linear system $\mathbf{Ax} = \mathbf{b}$. Here $\mathbf{A}$ represents the matrix modeling the projection parameters, $\mathbf{b}$ the vector containing the projection data and $\mathbf{x}$ defines the volume to be reconstructed. It should be noted, that we refer to ART as a generic term for algebraic methods iteratively solving the sparse matrix system. We used a single-step

TABLE I: Parameters of the two scenarios selected to reflect a typical dental or medical radiographic situation. R: sphere radius; d: source-to-receptor distance

scenario	d (mm)	R (mm)	pixel size (mm)
dental	250	1.5 or 2.5	0.039
medical	1000	3.0 or 5.0	0.143

iterative solver based on an adapted conjugate gradients algorithm (CG),[25] using the zero-vector as starting point. To avoid high frequency artifacts, distance-driven backprojection[26] was used. 3D reconstructions were carried out on the graphics processing unit (GPU; GeForce 7900 GT, NVIDIA Corp., USA) interfaced with a standard desktop PC (Intel Dual Core 2.4 GHz). All real world data presented here display the raw reconstruction data. We only use manually determined opacity transfer functions to control which parts of the data are visible. These functions map the reconstructed densities to suitable gray values as well as real numbers between 0 and 1. The latter represent the opacity associated with the data, with 0 being transparent and 1 being opaque.

F. Simulated data

1. Initial Guess

Simulations were performed with a tool allowing for projection of a sphere, the absorption of the which was defined by the Lambert-Beer law. Two typical projection scenarios specified in Tab. I were evaluated.

We investigated the accuracy of increasing depth coordinates ($0.04\,d$ to $0.20\,d$) with respect to different levels of Gaussian noise (standard deviations ranging between 0% to 20%). Absolute errors e_a (true (z_t) minus computed (z_c) coordinates) and relative errors ($\frac{e_a}{z_t}$) were defined. Depth accuracy was also evaluated for partly superimposed sphere shadows. Partly ($\frac{1}{2}$ to $\frac{2}{3}$) covering the sphere shadow, increasing intensities were superimposed in a stepwise fashion consisting of 11 steps each of which increasing the superimposed gray by 10%. The mean gray value within the uncovered sphere shadow was taken as starting point.

2. Triangle Fit

For illustration purposes, we selected a simplified dental scenario of two synthetically created radiographs of three spheres providing a rigid triangle. They were related geometrically by a known rigid rotation ($\mathbf{R}_t$) only, with no additional translation. Error was established by shifting one sphere's COP by $\leq$40 mm along its projection line in only one radiograph. This resembles a real-world situation, where e.g. due to local superimposition of a highly dense object, the initial guess returnes a large error for one sphere. After calculation of the rotation matrices with ($\mathbf{R}_{TF}$) and without ($\mathbf{R}_o$) TF, we computed the overall angular error by $\mathbf{R}_t^{-1}\mathbf{R}_{TF}$ or $\mathbf{R}_t^{-1}\mathbf{R}_o$, respectively. Ideally, the products would yield the identity matrix $\mathbf{I}$, thus deviations from $\mathbf{I}$ define the actual error.

G. Experimental data based on real-world radiographs

All experimental data were acquired in radiographic systems with a fixed transversal relation between focal point and image receptor. The distance d between these components was variable but known. A dry human mandible segment containing three teeth was exposed using an optical bench and a commercial dental ccd-sensor (Full Size, Sirona Dental Systems GmbH, Bensheim, Germany; pixel size: 0.039 mm). Three steel spheres (R: 2.5 mm) were glued to the specimen in a random triangular geometry. By freely positioning the specimen in the system roughly rotating it at 2π, tilting it between rotations ($\pm 20°$) plus varying d ($\pm 20\%$), eight projections of the specimen were acquired. In addition, a human dry skull was exposed in nine freely selected positions on a medical amorphous selenium detector (R: 5.0 mm; pixel size: 0.143 mm). No extra calibration to the particular radiographic system was performed for this series. To investigate the versatility of the method, we decided to expose a third series (entire human mandible, nine projections, R: 5.0 mm; pixel size: 0.075 mm) with an indirect digital system (storage phosphor plate; Duerr Dental GmbH & Co. KG, Bietigheim-Bissingen, Germany). The image information captured on storage phosphors has to be read out by a laser-scanning system (here VistaScan, Duerr Dental GmbH & Co. KG, Bietigheim-Bissingen, Germany) posing an additional (unknown) spatial transfer function on the data. However, so far we used the image data after read-out as is, i.e. without any further geometrical correction. For processing the image data of all series were encoded into DICOM-format.

III. RESULTS

A. Simulated Data

1. Initial Guess

Mean relative depth error ($\pm$ standard deviation) in the medical scenario was 2.1 % ($\pm$ 5.2 %), versus 4.4 % ($\pm$ 8.2 %) in the dental simulation. Depth errors were positively correlated with increasing depth (Fig. 3). Since absolute depth errors are more instructive with respect to registration accuracy, they are displayed in Fig. 3.

As expected, a significant ($p \leq 0.05$) positive correlation (dental: $r = 0.904$, medical $r = 0.613$) of Gaussian

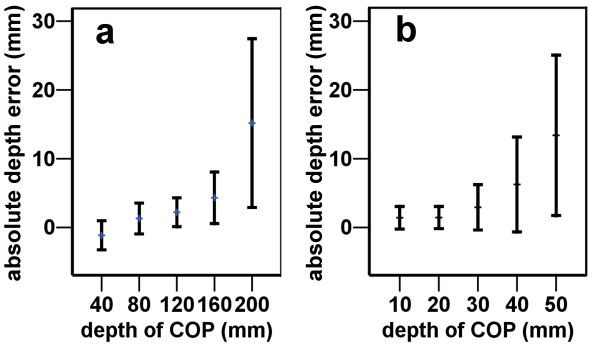

FIG. 3: Initial Guess: absolute depth error (mm) for increasing depth of the sphere's COP for a typical noise free a) medical and b) dental radiographic situation.

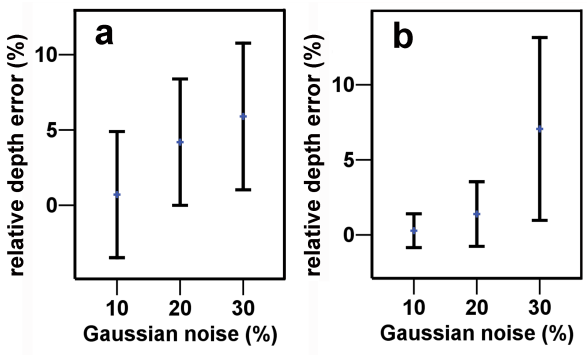

FIG. 4: Initial Guess: relative error (%) over all depths with respect to varying Gaussian noise for a typical a) medical and b) dental radiographic situation.

noise and depth accuracy was found (Fig. 4). Not surprisingly, the simulated data revealed increasing depth errors with increasing intensities superimposed over the shadow (Fig. 5). It should be noted here, that 100% means an intensity twice as high as within the uncovered shadow.

2. Triangle Fit

In the simplified simulation scenario displayed in Fig. 6, error reduction by the TF-algorithm is evident. While the rotational error without TF rapidly increases, it remains stable and close to zero for up to ≤15 mm after TF correction. Note, that the peak following by roughly 23 mm simulated depth error is due to a sudden switch to another possible triangle position (see also Fig. 2b). The exact location of such peaks depends on the geometrical situation of the projection scenario and on the triangle itself.

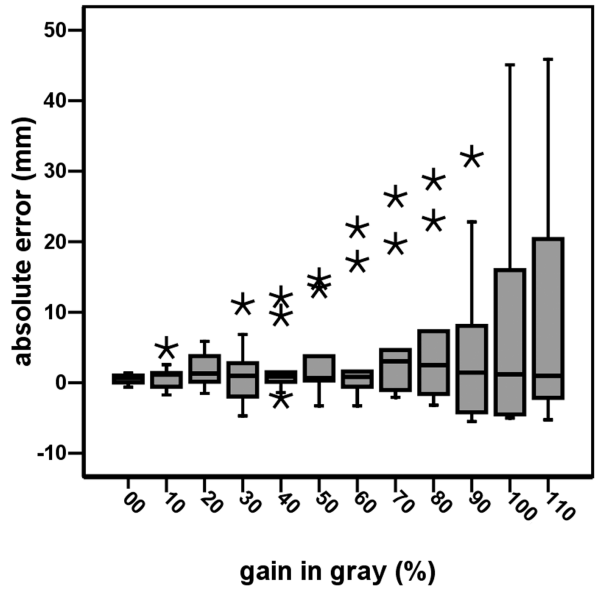

FIG. 5: Initial Guess: absolute error (mm) with respect to superimposed shadows of increasing intensity, starting from the mean within the uncovered sphere shadow (0%). The boxes contain 50 % of the data lying within the interquartile range (IQR) between the 75th and 25th percentiles, with the median indicated as bold line. Outliers and extreme values lying more than 1.5 IQR from the end of the box are displayed as asterisk.

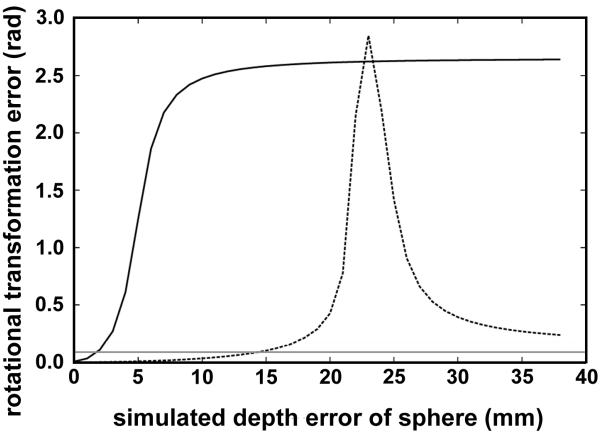

FIG. 6: For a given depth-error ≤40 mm simulated for one sphere in one out of two projections, the resulting rotational error ($\leq \pi$) with (dashed line) and without TF (solid line) is plotted. The rotational error after TF-correction remains underneath 5° (= 0.087 rad, gray horizontal line) for depth errors up to ≤15 mm. The peak in the TF graph at roughly 23 mm depth error is due to a switch to another possible triangle position at that point (see also Fig. 2b).

B. Experimental Results

1. Performance

Registration time on the CPU ranged between 132 s and 343 s. Roughly over 90 % of that time is consumed

by the automated shadow detection (Tab. II). Due to restrictions in the current GPU-implementation of our reconstruction algorithm, large-scale images were downsampled for reconstruction (Tab. II). Reconstruction time for a 300^3 voxel volume ranged between 64 s and 104 s (Tab. II). Without going into detail here, reconstruction time mainly depends on image size, the number of projections as well as volume size and number of iterations.

2. *Real world reconstructions*

Slices through a teeth-containing human mandible segment (300^3 voxels, three iterations) obtained from eight projections are presented in Fig. 7. The projections were obtained in an approximately circular orbit with additional angulations of the specimen of appr. $\pm 20°$ and varying d ($\pm 20\%$). Fig. 8 displays the human mandible (300^3 voxels, three iterations) reconstructed from nine projections distributed around the object approximating a random spherical orbit. This was carried out by placing the mandible in nine positions on a aluminum-cassette containing the storage phosphor plate (18 cm $\times$ 24 cm). Without calibration to that indirect digital system, RSM-registration yielded sufficient precision for 3D reconstruction. Internal structures of the mandible such as an impacted retained root with circular translucency or extrinsically added steel wires are clearly reconstructed (Fig. 8). The human skull (Fig. 9) was radiographed on an amorphous selenium detector in nine varying positions. Again, no specific calibration for the radiographic system was carried out for RSM registration. By application of an opacity transfer function, a volume rendered display of the surface was obtained (Fig. 9 d to f).

IV. DISCUSSION

Despite a huge body of literature dealing with image registration (for extensive review, see[11, 12]), only a limited number of papers concentrate on registration for 3D reconstruction from few radiographic projections.[14–17, 27] Intrinsic point-based registration methods, where corresponding point pairs have to be identified, have limited accuracy simply because the underlying anatomic landmarks produce very different images when exposed under significantly varying geometries. The latter, however is the fundamental prerequisite for acquisition of 3D information. Other registration methods based on voxel properties, to date, are still relatively inaccurate.[12] Except of a technical paper[19] dealing with radiographic nondestructive testing and the very recent work in,[8] we are not aware of published solutions for 3D reconstructions from few projections acquired in arbitrary geometries. In the latter work a new solution to the problem has been introduced.[8] The authors propose to solve the registration together with the 3D reconstruction without any reference points at all using non-linear algebraic solver in combination with Bayes theory. Because of the high complexity, at present computation time will be a limiting factor. Also, the real world data presented in[8] were acquired in a relatively constrained (affine) geometry from rather many (23) equally spaced projections from around the object. Furthermore, a good initial guess is necessary to avoid local minima solutions and the output is a stack of 2D slices. Contrary to[8] our approach enables precise modeling of the realistic perspective radiographic imaging geometry. Our reconstructions are completely three-dimensional and, due to the nature of the registration process, inherently free of scaling ambiguities. Since the overall amount of information inherent in the projections limits the innovative approach in,[8] the extension to 3D as proposed by the authors will be an extremely challenging task.

The elementary idea to localize a spherical reference body in 3D from a single projection originates from.[20] Contrary to the original method, we did not only consider the major axes of the sphere's elliptical shadow, rather we completely exploited the shadow information using conic section theory. By supplementing the image analysis with mathematical a priori knowledge we obtain a more reliable initial guess on the sphere's depth. Nevertheless, we still observed that the latter is strongly dependent on noise and other disturbances.

The definition of the shadow boundary is a crucial first step, as our simulation experiments show. It is well-known in point-based registration, that error in the image input coordinates drastically affects the computed depth-coordinate.[15, 17, 28] Hence, an additional minimization function F (eq. (12)) was designed processing shadow information from all projections simultaneously. A minimum obtained from eq. (12) indicates the best possible geometric correspondence of the projections to one another, which is the essential information for the backprojection process. Yet it should be realized that the TF-corrected COP-coordinates are not necessarily closer to truth in a world coordinate system. Rather their relative position to one another is optimized. This is clearly indicated by our simulation results and the real-world reconstructions. Almost in all series, 3D reconstructions carried out without TF correction were inferior in quality or, even worse, did not produce meaningful reconstructions at all.

Although extrinsic reference bodies may seem somewhat out-dated,[12] the method described here has some distinct advantages listed below. Reference-based registration techniques have been successfully used in clinical routine, e.g. for tomosynthetic reconstructions.[29] These, however, were obtained from an imaging geometry constrained in a way that the object was rigidly coupled to the image receptor. We have demonstrated experimentally, that the RSM-algorithm is capable to register arbitrary radiographic projections accurately

TABLE II: Registration (CPU) and reconstruction (GPU) time for the real world data displayed in Fig. 7, 8 and 9

specimen	projections	image size	registration time	downsampled size	reconstruction time	total time
mandible segment	8	664×872	132 s	664×872	104 s	236 s
entire mandible	9	2800×2000	343 s	750×500	94 s	437 s
skull	9	2200×1600	223 s	550×400	64 s	287 s

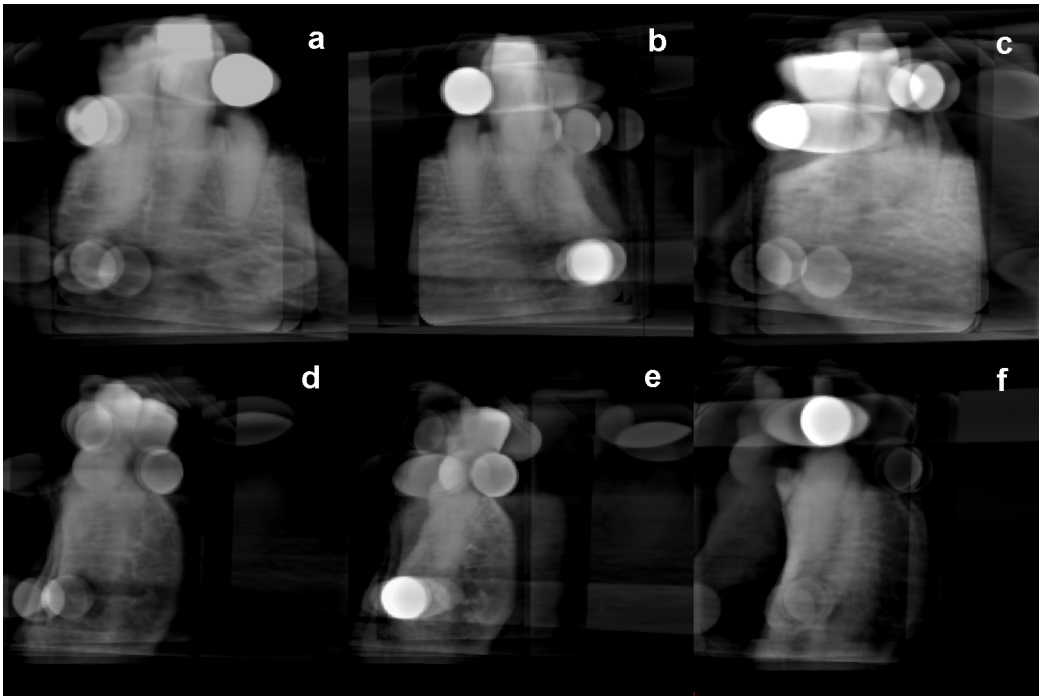

FIG. 7: Unprocessed reconstruction from eight projections of a dry human mandible segment containing three teeth. Images display frontal (a to c) and transversal (d to f) slices. The shadows of the reference spheres are clearly visible, however, no attempt to suppress them or other few-view and limited-angle artifacts has been carried out.

enough to obtain 3D reconstructions. One specific advancement of our approach is, that due to the 3D nature of the points used for registration, in the absence of error, the registration is inherently free of distortion and scaling ambiguities. The latter is clearly illustrated by the series of the mandible segment, where d varied significantly ($\pm 20\%$). Also, in contrast to others[9, 19] we register six degrees of freedom with a minimum amount of rather small reference objects, the position of which relative to the object may be freely chosen, as long as their COPs are not collinear. No assumptions on the geometry have to be made, except that at the current stage, the spatial relation between source and image-receptor has to be known a priori. In theory, the $x-, y-$ relation of the latter components may also be computed from the elliptically distorted sphere shadows by intersecting the major ellipse axes.[20, 22] Equation (9) has been designed to provide exactly these coordinates. From preliminary results we observed promising stability for excentric ellipses ($\lambda_1 \gg \lambda_2$), as for instance produced by oblique projections or a large angle of the beam-cone. Yet in cases of low excentricity, results returned from (9) are not numerically stable enough for accurate registration. This problem is currently being further investigated.

Compared to GPU reconstruction time, at present registration is relatively time consuming. It should be noted, however that the current implementation of the sphere shadow detection has not been optimized with respect to speed. We are currently working on that issue and preliminary results suggest that registration time can be reduced drastically.

The reconstruction process itself was not the primary scope of the paper, since the reconstructions were only included to illustrate the practical applicability of the registration procedure. For the same reason, we did not discuss typical few-view and limited-angle reconstruction problems also present in our data.[3, 4, 27] By implementation of a priori knowledge in the algebraic reconstruction process, however the quality of the reconstructed volumes may be enhanced considerably.[4]

Advantages of RSM-registration in practical application include low hardware cost and a very low radiation dose for 3D reconstruction. If costly hardware is not available or, due to technical reasons such as object size simply not

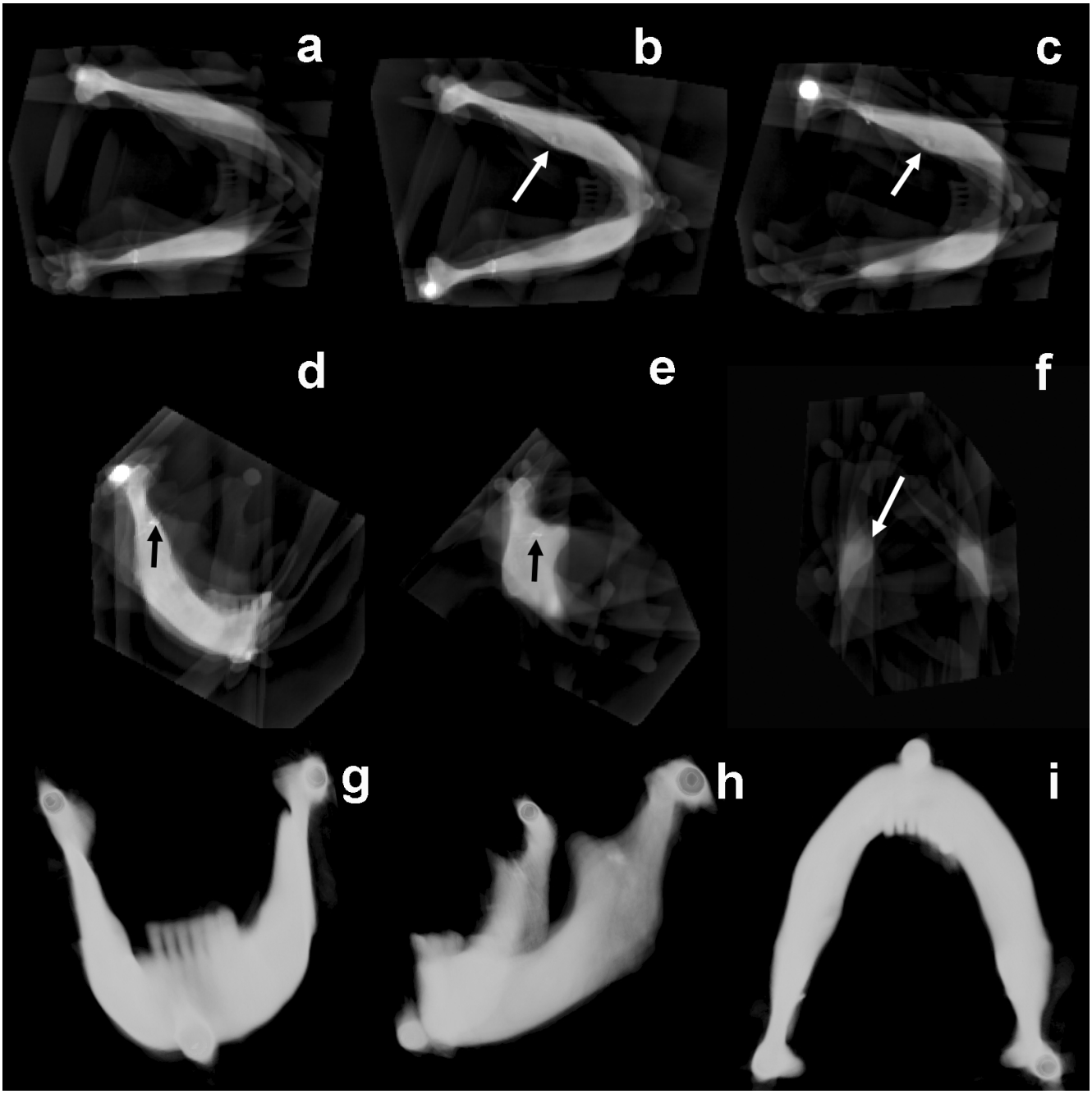

FIG. 8: Unprocessed reconstruction of an entire human mandible placed in nine varying positions on a storage phosphor plate. Axial slices are presented in images a to c, whereas images d and e show sagittal slices. Image f represents a frontal slice through the region of a retained impacted root tip, which is obviously surrounded by a circular translucency (white arrows, b, c and f). Another fine reconstruction detail is the thin copper wire (black arrows, d and e). Volume rendered views computed by application of an appropriate opacity transfer function are shown in images g to i.

applicable, projection images of the objects from common radiographic devices may be used as alternative to generate 3D information. Obviously, the overall radiation dose of only a few digital radiographic projections is very low as compared to conventional computed tomography. Few skull radiographs registered with the RSM-procedure, for instance, could be used to obtain a reliable 3D bone surface, e.g. for preoperative planning procedures. Future work in this ongoing research project will focus on further enhancement of the registration as well as the 3D reconstruction procedure. Integration of more mathematical and physical knowledge in the sense of a more sophisticated calibration procedure could help to further advance RSM-registration. By inclusion of a priori information, i.e. implementation of non-linear constraints,[3, 4] reconstruction artifacts could be reduced.

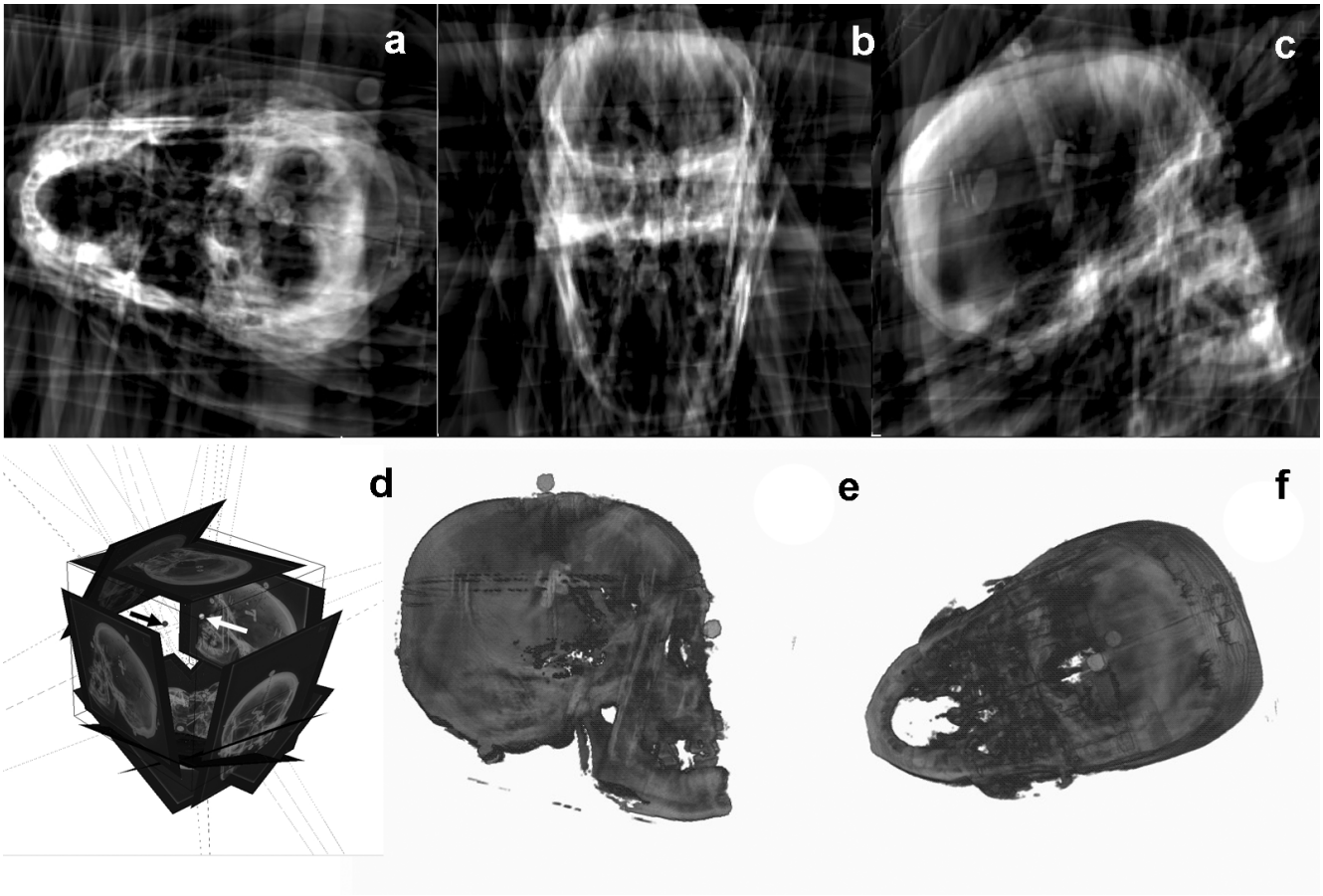

FIG. 9: Unprocessed axial a), frontal b) and sagittal c) slice through a human skull CG-reconstruction (300^3 voxels) from nine projections on a selenium detector. d) RSM-registered imaging geometry (arrows indicate two sphere positions); e) and f) display a volume rendered reconstruction after suppression of gray values outside the reconstructed skull (see text).

V. CONCLUSION

The reference-based procedure describe here enables accurate registration of radiographic projections acquired in almost arbitrary geometries. Only a minimum number of three rather small extrinsic references is required that are fixed to an object in an arbitrary triangular geometry. The registration process facilitates reconstruction of the 3D information inherent in the projections. Radiographic 3D reconstructions acquired with our approach may be useful, where established 3D techniques are not available, not applicable or imply an unacceptable high radiation dose.

Acknowledgments

The authors would like to thank the Department of Anatomy and Cell Biology at the Johannes Gutenberg-University Mainz for the donation of the human mandible segment. This work was funded by the German Research Foundation DFG, under grant # SCHU1496/1-2 and by a research fund of the Johannes Gutenberg-University.

[1] P. Kuchment, K. Lancaster and L. Mogilevskaya, "On local tomography," Inverse Probl. **11**, 571-589 (1995).

[2] M. Li, H. Yang and H. Kudo, "An accurate iterative reconstruction algorithm for sparse objects: application to 3D blood vessel reconstruction from a limited number of projections," Phys. Med. Biol. **47**, 2599-2609 (2002).

[3] S. Siltanen, V. Kolehmainen, S. Järvenpää, J. P. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä and E. Somersalo, "Statistical inversion for medical x-ray tomography with few radiographs: I. General theory," Phys. Med. Biol. **48**, 1437-1463 (2003).

[4] V. Kolehmainen, S. Siltanen, S. Järvenpää, J. P. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä and E. Somersalo, "Statistical inversion for medical x-ray tomography with few radiographs: II. Application to dental radiology," Phys. Med. Biol. **48**, 1465-1490 (2003).

[5] V. Kolehmainen, A. Vanne, S. Siltanen, S. Järvenpää, J. P. Kaipio, M. Lassas and M. Kalke, "Parallelized Bayesian inversion for three dimensional dental x-ray imaging," IEEE T. Med. Imaging **25**, 218-228 (2006).

[6] M. Rantala, S. Vänskä, S. Järvenpää, M. Kalke, M. Lassas, J. Moberg and S. Siltanen, "Wavelet-based reconstruction for limited-angle x-ray tomography," IEEE T. Med. Imaging **25**, 210-216 (2006).

[7] R. L. Webber, R. A. Horton, D. A. Tyndall and J. B. Ludlow, "Tuned-aperture computed tomography (TACT). Theory and application for three-dimensional dentoalveolar imaging," Dentomaxillofac. Radiol. **26**, 53-62 (1997).

[8] S. S. Brandt and V. Kolehmainen, "Structure-from-motion without correspondence from tomographic projections by Bayesian inversion theory," IEEE T. Med. Imaging **26**, 238-248, (2007).

[9] S. B. Robinson, P. F. Hemler and R. L. Webber, "A geometric problem in medical imaging," Proc. SPIE **4121**, 208-217 (2000).

[10] A. I. Katsevich, "Local tomography for the limited-angle problem," J. Math. Anal. Appl. **213**, 160-182 (2003).

[11] D. L. G. Hill, P. G. Batchelor, M. Holden and D. J. Hawkes, "Medical image registration," Phys. Med. Biol. **46**, R1-R45 (2001).

[12] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," Med. Image Anal. **2**, 1-36 (1998).

[13] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," Nature **293**, 133-135 (1981).

[14] C. E. Metz and L. E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory," Med. Phys. **16**, 45-51 (1989).

[15] K. R. Hoffmann, C. E. Metz, and Y. J. Chen, "Determination of 3D imaging geometry and object configurations from two biplane views: An enhancement of the Metz-Fencil technique," Med. Phys. **22**, 1219-1227 (1995).

[16] S.-Y. J. Chen and C. E. Metz, "Improved determination of biplane imaging geometry from two projection images and its application to three-dimensional reconstruction of coronary arterial trees," Med. Phys. **24**, 633-654 (1997).

[17] L. E. Fencil and C. E. Metz, "Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation," Med. Phys. **17**, 951-961 (1990).

[18] S.-Y. J. Chen and J. D. Caroll, "3-D reconstruction of coronary arterial tree to optimize angiographic visualization," IEEE T. Med. Imaging **19**, 318-336 (2000).

[19] C. Lehr and C.-E. Liedke, "3D Reconstruction of volume defects from few x-ray images," Lect. Notes Comput. Sc. **1689**, 275-284 (2006).

[20] R. Schulze, D. D. Bruellmann, F. Roeder and B. d'Hoedt, "Determination of projection geometry from quantitative assessment of the distortion of spherical references in single-view projection radiography, " Med. Phys. **31**, 2849-2854 (2004).

[21] R. M. Haralick, C.-N. Lee, K. Ottenberg and M. Noelle, "Review and analysis of solutions of the three point perspective pose estimation problem," Int. J. Comput. Vision **13**, 331-356, (1994).

[22] G. Shivaram and G. Seetharaman, "A new technique for finding the optical center of cameras," Proc. IEEE Int. Conf. Image Process., 167-171 (1998).

[23] R. M. Haralick, C.Lee, K.Ottenberg, and M.Nölle, "Analysis and solutions of the three point perspective pose estimation problem," Proc. IEEE Conf. Computer Vision and Pattern Recognition, Maui, USA, 592-598, (1991).

[24] F. Natterer, "Numerical methods in tomography," Acta Numerica **8**, 107-143 (1999).

[25] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," http://www.cs.cmu.edu/ quake-papers/painless-conjugate-gradient.pdf March, (1994).

[26] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," Phys. Med. Biol. **49**, 2463-2475 (2004).

[27] C. J. Henri, D. L. Collins and T. M. Peters, "Analysis of projection geometry for few-view reconstruction of sparse objects," Med. Phys. **20**,1537-1547 (1993).

[28] K. Tabushi, I. Susumu and M. Sakura, "Two-radiograph reconstruction using six geometrical solution sets and least-square method," Med. Phys. **19**, 1307-1310, (1992).

[29] R. L. Webber and J. K. Messura, "An in vivo comparison of diagnostic information obtained from tuned-aperture computed tomography and conventional dental radiographic imaging modalities," Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. **88**, 239-247 (1999).