This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

Review of 2D/3D reconstruction using statistical shape and intensity models and X–ray image synthesis: towards a unified framework

Cornelius J. F. Reyneke, *Member, IEEE*, Marcel Lüthi, Valérie Burdin, *Member, IEEE*, Tania S. Douglas, *Senior Member, IEEE*, Thomas Vetter and Tinashe E. M. Mutsvangwa, *Member, IEEE.*

***Abstract*—Patient–specific three–dimensional (3D) bone models are useful for a number of clinical applications such as surgery planning, postoperative evaluation as well as implant and prosthesis design. Two–dimensional–to–three–dimensional (2D/3D) reconstruction, also known as model–to–modality or atlas–based 2D/3D registration, provides a means of obtaining a 3D model of a patient's bone(s) from their 2D radiographs, when 3D imaging modalities are not available. The preferred approach to estimating both shape and density information (that would be present in a patient's CT data) for 2D/3D reconstruction makes use of digitally reconstructed radiographs and deformable models in an iterative, non–rigid, intensity–based approach. Based on a large number of state–of–the–art 2D/3D bone reconstruction methods, a unified mathematical formulation of the problem is proposed in a common conceptual framework, using unambiguous terminology. In addition, shortcomings, recent adaptations and persisting challenges are discussed along with insights for future research.**

***Index Terms*—2D/3D reconstruction, atlas, bone density, deformable 2D/3D registration, digitally reconstructed radiograph, intensity–based 2D/3D registration, non–rigid 2D/3D registration, model–to–modality, patient–specific model, statistical appearance model, statistical shape and intensity model, statistical shape model.**

I. Introduction

Three–dimensional (3D) reconstructions of patient–specific anatomical structures help medical professionals to better visualize and interact with the volumetric data from 3D imaging modalities such as computed tomography (CT) or magnetic resonance images (MRI) [1]. They have become an especially valuable tool for orthopaedic applications which include the detection of bone–related pathological deformations as well as the quantitative measurement of bone geometry and bone density for surgical planning, implant design, and postoperative evaluations [1], [2]. However, CT and MRI technologies are expensive and have long image acquisition times [3]–[6]. Furthermore, hospitals in resource–limited and remote settings often do not have access to these technologies. Those that do, have to consider the artefacts and distortions caused by the presence of metallic implants, as well as the prohibitive costs per scan, for either modality [1]. In order to overcome these challenges, researchers have investigated ways to obtain 3D models of patient–specific anatomical structures from two–dimensional (2D) imaging modalities such as X–ray, dual–energy X–ray (DXA), fluoroscopic and ultrasound images. This approach is known as two–dimensional–to–three–dimensional (2D/3D) reconstruction [7]. Patient–specific models obtained in this way are less accurate with regard to shape and appearance than those obtained from CT or MRI [8]. However, imaging costs and the dose of ionizing radiation to a patient (in the case of CT) is significantly reduced [1], [9], [10]. For example, in the case of preoperative imaging for total hip arthroplasty, the dose is 30% less for conventional 2D X–ray images than for a CT scan [3]. Low–dose imaging systems can further reduce a patient's exposure to ionizing radiation and, in some cases, allow patients to be scanned while they are standing up, enabling the assessment of both their load–bearing posture and musculoskeletal interactions [11]. These advantages make repetitive follow-up imaging and paediatric imaging applications possible [12]. Furthermore, when obtaining 3D patient–specific models from fluoroscopic sequences, in vivo motion tracking is made possible for applications such as 3D knee kinematics studies [9], [13]–[15].

Two–dimensional–to–three–dimensional reconstruction is accomplished using a deformable model, which encodes prior knowledge and assumptions about the typical 3D appearance of an anatomical structure [7]. These models can then be manipulated, using a set of parameters, to match the information gathered from a patient's 2D image(s) in a 2D/3D registration [2]–[4], [6], [9], [15]–[59]. An exemplary specimen (sometimes referred to as a template or atlas) is used to identify a set of characteristic features such as landmarks or contours. The mathematical relationships between these features and their variations are then defined either analytically [16], [17], [19], [22], [23], [25], [26], [28], [30], [31], [40], [41], [48], [54], or statistically [2]–[4], [6], [9], [15], [18], [20], [21], [24], [27], [29], [32]–[39], [42]–[47], [49]–[53], [55]–[59]. Analytical models (sometimes referred to as free–form models) rely on intuitions regarding the mathematical relationships, such as

Research supported by the SARChI programme of the South African National Research Foundation (grant no 98788) and by the South African Medical Research Council.

C. J. F. Reyneke, T. S. Douglas and T. E. M. Mutsvangwa are with the Division of Biomedical Engineering, University of Cape Town, South Africa (phone numbers: +27 21 650 1418; fax: +27 21 650 1418; e–mail: ryncor001@myuct.ac.za; tania.douglas@uct.ac.za & tinashe.mutsvangwa@uct.ac.za).

V. Burdin is with Département Image et Traitement de l'Information, IMT Atlantique, Brest, France, and LaTIM INSERM U1101 (e–mail: valerie.burdin@imt–atlantique.fr).

M. Lüthi and T. Vetter are with the Graphics and Computer Vision Group (GRAVIS), Basel University (UNIBAS), Switzerland (e–mail: marcel.luethi@unibas.ch & thomas.vetter@unibas.ch).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

those of medical experts. Statistical models, on the other hand, are learnt from a set of labelled examples. They are particularly well–suited to the rigid variability present in bony structures and are popular for their ability to represent objects robustly; while such a model is deformed its validity as a representation of the bone–of–interest is preserved (Figure 1 illustrates this concept) [1], [42], [60]. A more detailed explanation of statistical models such as these is provided in Section II-A. For a detailed discussion about deformable models, and their use in medical image registration, the reader is referred to [61]. In order to obtain a patient–specific reconstruction, the parameters of the deformable model are tuned according to information inferred from a patient's X–ray images, the target values of which are determined using one of two registration strategies [1].

Fig. 1. A statistical shape model (SSM), illustrating the first (left), second (middle) and third (right) principal components (PCs) of variation. The femur shape with the darkest shade corresponds to the mean configuration of the SSM, and is identical in each of the three depictions. The femur shapes represented in lighter shades correspond to SSM configurations at -3σ and $+3\sigma$. The reader should note that the principal components of an SSM are ordered according to the amount of variation that each accounts for.

The first strategy seeks to establish correspondence, and minimize the geometric distance, between features detected on both the deformable model and the patient's X–ray image(s). This is accomplished using a Kriging optimization (also known as a Gaussian process regression) [4], [9], [16], [17], [19], [20], [22], [23], [25], [26], [28], [30]–[32], [35]–[37], [40]–[44], [46], [48]–[51], [54], [62]. It is however first necessary to identify these corresponding features. Methods to do so vary from fully manual to fully automatic, but both are time–consuming and prone to error [4], [42], [45]. The corresponding features can either be projected onto the 2D space of the X–ray images or back–projected into the 3D space of the model. Methods of the former category search for the parameters in an iterative scheme which is terminated once the distance value is less than a pre-defined threshold [18], [20], [24], [32], [36], [44], [63]. Methods belonging to the latter category compute the parameters directly, but often do not provide sufficiently accurate reconstructions [54] (Accuracy requirements are discussed in more detail in Section II-A).

The second strategy makes use of simulated X–ray images obtained from the deformable model, and seeks to maximize the score of a similarity measure which performs a comparison of pixel intensity values in the 2D space of the X–ray images [2], [3], [6], [18], [21], [24], [27], [29], [33], [34], [39], [45], [47], [52], [53], [55], [56], [59]. Once again an iterative optimization scheme is adopted, and is concluded once the similarity measure value exceeds a pre–defined threshold (A basic diagram of the algorithm is shown in Figure 2). In this case no feature–detection is necessary as the entire 3D object is projected into 2D space. This projection, known as a digitally reconstructed radiograph (DRR), can be as simple a silhouette, thickness–projection or maximum–intensity–projection (MIP) [24], [27], [46], or can model the complexity of the full intensity distribution. Intensity–based approaches are known to be more accurate than feature–based methods, but have a longer computation time [7], [64]. The algorithm suffers from a computational bottleneck resulting from its iterative nature, because an accurate volumetric reconstruction often requires a few thousand DRRs to be rendered (this issue is discussed in more detail in Section II-B) [6]. In addition, methods such as these have a small range of similarity measure values that can be optimized reliably without getting caught in local maxima [7], [64]. Consequently the first instance (or configuration) of the deformable model must be initialized so that its corresponding DRR projection(s) are already somewhat similar to the target X–ray image(s) [7], [65]. This is often accomplished by employing pose–initialization methods before the registration (optimization and pose–initialization are discussed in greater detail in II-D) [64], [66], [67].

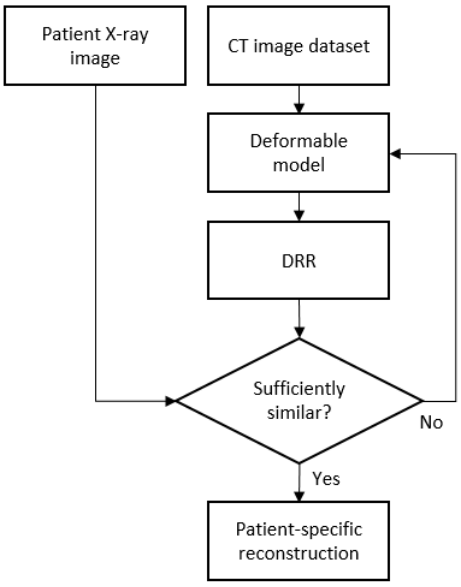


Fig. 2. Overview of the DRR–based 2D/3D bone reconstruction algorithm. The process is iterative and is concluded once the similarity measure value exceeds a pre–defined threshold.

Two–dimensional–to–three–dimensional bone reconstruction algorithms originally only accounted for shape and therefore only produced surface models [4], [9], [16]–[20], [22]–

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

[28], [30]–[33], [35]–[38], [40]–[44], [46], [48], [49], [51], [54], [56]. Authors have since sought to include bone density information in their models, which enable volumetric models as well as a number of additional clinical applications [2], [3], [6], [21], [29], [34], [39], [45], [47], [50], [52], [53], [55], [59]. A few clinical application examples include; distinction between cortical and trabecular bone for total hip arthroplasty (cortical regions on the outer hull of a bone absorb X–rays more readily than the trabecular regions which comprise the inner–bone) [3], [53]; estimation of the "local anchorage quality" of osteoporotic bone (sufficient anchorage quality enables prosthesis and implant stability), thereby contributing toward efficient surgical planning for femoral fracture fixations [34]; early identification of individuals at risk of developing osteoarthritis; measuring the progression of osteoarthritis and providing radiographic markers for individuals who require surgery [1], [68]–[70]; estimation of the potential risk factors for individuals with osteoporosis (prediction of potential bone fractures and detection of existing vertebral fractures) [71], [72]; aiding in the often problematic task of distinguishing bone deformities from bone fractures [71].

The scope of our review is limited to DRR–based 2D/3D bone reconstruction methods that make use of statistical models and biplanar X–ray images [2], [6], [21], [29], [34], [45], [53], [59]. More specifically, we focus on statistical models of both bone shape and density, which we will refer to as statistical shape and intensity models (SSIMs), and on those trained using CT data since DRR projections of MRI data are unreliable [7]. Statistical shape and intensity models are sometimes referred to as statistical appearance models (SAMs), and are distinguished from statistical shape models (SSMs) which account for shape information alone [1], [45].

Our survey of existing literature found that the terminology used by researchers is ambiguous and inconsistent, and that a unified mathematical formulation of the 2D/3D reconstruction problem has not yet been presented. We remedy this using a common conceptual framework and set of terminology, as well as a unified mathematical formulation of DRR–based 2D/3D reconstruction using SSIMs.

A number of adaptations to this approach have been proposed in an attempt to overcome the limitations of the algorithm. These are discussed according to which of the four main modules of the algorithm they are most relevant to; the SSIM (Section II-A), the DRR rendering method (Section II-B), the similarity measure (Section II-C), the optimization strategy (Section II-D), or the algorithm as a whole (Section II-E).

Finally, we provide a detailed discussion of key issues, persisting challenges and insights for future research (Section III).

II. 2D/3D BONE RECONSTRUCTION USING DRRS

A. Statistical models: construction, validation, requirements and variations

In order to better understand SSIMs which have been applied to DRR–based 2D/3D reconstruction we first describe the mathematical formulation of SSMs, specifically those that utilize principal component analysis (PCA). The methods used in their construction and validation, as well as the level of accuracy and maximum computation time that is typically required by current clinical applications, are then described.

1) Construction of SSMs: We closely follow the procedure described by [60] and [1]. The first step in constructing an SSM is to decide how the bone shape is described mathematically. A number of shape features have been proposed including landmark points, dense surface meshes such as Fourier surfaces and spherical harmonics, medial models such as m–reps and s–reps [73], [74], deformation fields and distance maps; the most commonly used descriptor for bone is landmark points [1]. The simplest means of labelling landmarks is to manually select bone shape features that can be easily identified by an anatomist or radiographer (these usually number in the order of tens) [1]. However, most modern approaches use a dense set of landmarks (which number in the order of thousands). A shape, Γ_n, is first defined as a set of points:

$$\Gamma_n = \{a_p^n \in \mathbb{R}^3, p = 1, ..., P\} \tag{1}$$

for the n^{th} shape, Γ_n, $a_p^n = (x_p^n, y_p^n, z_p^n)$ where x_p, y_p and z_p denote the Euclidean coordinates of the p^{th} point. The shape, Γ_n, can then be represented as a vector:

$$\mathbf{s}_n = (x_1^n, y_1^n, z_1^n, ..., x_P^n, y_P^n, z_P^n) \qquad \mathbf{s}_n \in \mathbb{R}^{3P} \tag{2}$$

Once the same number of points have been used to define all the shapes that comprise the dataset, $n = 1, ..., N$, it is necessary to ensure that the points of each shape are in correspondence with one another. This means any particular point $(a_p^1, ..., a_p^N)$ should describe the same anatomical feature across the dataset. A non-exhaustive list of methods for establishing correspondence among point-sets includes the Iterative Closest Point algorithm or variations thereof, Coherent Point Drift, and Robust Point Matching [75]–[77]. Additional information regarding such methods is provided in [78]. Furthermore, since SSMs are good regularizers themselves, they can also be used for landmark–based surface registration, as described by [79].

Next, it is necessary to align the bones so that variations resulting from anything other than shape (such as rotational and translational effects) are eliminated; the most common method for doing this is generalized Procrustes analysis (GPA) [80].

Once aligned, it is possible to model the shape variations using a normal distribution:

$$\mathbf{s} \sim \mathcal{N}(\bar{\mathbf{s}}, S) \tag{3}$$

where the mean, $\bar{\mathbf{s}}$, and covariance matrix, S, can be estimated as follows:

$$\bar{\mathbf{s}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{s}_n \tag{4}$$

$$S = \frac{1}{N-1} \sum_{n=1}^{N} (\mathbf{s}_n - \bar{\mathbf{s}})(\mathbf{s}_n - \bar{\mathbf{s}})^T \tag{5}$$

Finally, principal component analysis (PCA) can be performed to determine common configurations of bone features, represented as an average point distribution and its principal modes

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

of variation (also referred to as eigenmodes or eigenvectors) [1]. The concept of principal components is illustrated in Figure 1. This leads to a probabilistic representation for **s**:

$$\mathbf{s} = \bar{\mathbf{s}} + \sum_{m=1}^{M} w_m \sqrt{u_m} \mathbf{e}_m \qquad w_m \sim N(0,1) \tag{6}$$

where M is the number of principal components, u_m and $\mathbf{e}_m$ represent the m^{th} eigenvalue and eigenvector respectively, and w_m represents the m^{th} model parameter of **s**. A more detailed mathematical description of this process is provided by [1], [60], [81]. The eigenvectors are intrinsically sorted in descending order so that $u_m > u_{m+1}$ and the variance of **s** is equal to the sum of the individual component variances, since the parameters, $w_1, ..., w_m$, are uncorrelated [60]. Then if u_m decays rapidly, **s** can be accurately approximated using only the first G components:

$$\mathbf{s} \sim \bar{\mathbf{s}} + \sum_{m=1}^{G} w_m \sqrt{u_m} \mathbf{e}_m \tag{7}$$

The number of components to use are typically selected by increasing G until the ratio of accumulated variance to total variance, γ, reaches a threshold value:

$$\gamma = \frac{\sum_{m=1}^{G} u_m}{\sum_{m=1}^{M} u_m} \tag{8}$$

Commonly accepted values for γ are provided in [82].

2) Validation of SSMs: In order to assess the quality of an SSM, a number of statistics are typically computed: compactness, specificity and generality [83], [84].

Compactness is a model's ability to faithfully capture shape variance while using a minimum number of principal components [85]. This characteristic measures the cumulative variance that the model can account for, as a function of the number of principal components, M, that are utilized [83].

Specificity measures the ability of the model to generate instances similar to those available within the training set [84]. It is computed as the average distance (such as root–mean–square error (RMSE) or mean–absolute–distance (MAD)) between randomly generated, uniformly distributed, shapes and their nearest match in the training set [83]. The measure is plotted as a function of the number of principal components that are employed to generate the random shapes. The number of randomly generated shapes are typically large when compared with the number of training examples (often having multiple orders of magnitude), and are proportional to the number of points that comprise the model. An in–depth discussion of distance metrics for evaluating 3D image segmentations is provided in [86].

Generality is the ability of the model to generate instances not explicitly provided by the training set. A leave–one–out strategy is used when a limited number of training examples are available; all but one example is used to train a shape model, which is then fitted to the left–out training example [83]. This process is repeated so that each instance has a turn to be excluded and measured. Ultimately, the average reconstruction error for unseen examples is provided, as a function of the number of principal components used to approximate the left–out training example. Generality is an important characteristic of the model to measure because if the model is over–fitted to the training set, it may have excellent specificity but will hinder its ability to generalize to unseen examples [83].

3) Variations on SSMs: Recently, PCA–based SSMs have been reformulated as Gaussian process morphable models (GPMMs) by [60]. The advantage of this reformulation is that it does not restrict the covariance function to be the sample covariance matrix (obtained from the training examples), but any valid positive definite covariance function. The covariance function can thus be analytically defined to include prior knowledge and intuitions about a shape, even when training data are not available. Another advantage of this approach is that it is formulated in the continuous domain. This means that dense correspondence can be established for any required resolution, and can therefore be tailored to a specific application. The reader is directed to [60] for an in–depth explanation of GPMMs.

4) Construction of SSIMs: Digitally reconstructed radiograph–based 2D/3D reconstruction technically requires bone intensity information to be included in the deformable model and, since this is absent from 3D surface models (SSMs), such information is often estimated in a simple manner using projection techniques such as MIP. Efforts have since been made to extend SSMs and accurately encode the intensity information present in example CT datasets. By training the model in a similar fashion to SSMs, they can also account for the average bone intensity and bone intensity variation [17], [19], [22], [23], [25], [28], [30], [31], [35], [38], [87]. The construction of SSIMs involves a few extra steps compared with SSMs, the most important of which involves establishing correspondence not only on the surface of the shape, but also within the volume [1], [88]. Statistical shape and intensity models can be broadly categorized into two types of representation [88]. One representation models a segmented CT volume (of the bone–of–interest) as a mesh in which mesh–morphing techniques such as those proposed in [89] and [90] are used to establish correspondence. The second representation models the volume as an image using the original CT voxels. In this case correspondence is established using image registration algorithms such as free–form deformation, or the diffeomorphic Demons algorithm [88]. In a comparison of the two types of model representation (mesh and image–based SSIMs), the former was better at reproducing shape information, while the latter was better at reproducing intensity information [88].

Once correspondence has been established, each training instance is fitted to a target reference shape; the mean shape of the SSM is typically used. Then the intensity information is sampled from the mean, and concatenated into an intensity (texture) vector [1]:

$$\mathbf{i}_n = (i_1^n, ..., i_P^n) \qquad \mathbf{i}_n \in \mathbb{R}^P \tag{9}$$

The intensities are usually normalized to reduce the effects of global intensity variations [81]. Finally, the SSIM can either model the two types of information separately, known as an independent SSIM, or together as a combined SSIM [1], [91].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

Independent SSIMs have the advantage of enabling more accurate patient–specific reconstructions, but are more time–consuming to implement [1]. The intensity variation is modelled separately, using PCA, in a similar process to that used to obtain a shape model:

$$\mathbf{i} = \bar{\mathbf{i}} + \sum_{q=1}^{Q} w_q \sqrt{u_q}\mathbf{e}_q \qquad w_q \sim N(0,1) \tag{10}$$

where Q is the number of principal components of the intensity model, $\mathbf{i}$, u_q and $\mathbf{e}_q$ represent the q^{th} eigenvalue and eigenvector respectively, and w_q represents the q^{th} model parameter of $\mathbf{i}$. An independent SSIM was implemented by [45]; they constructed a separate shape model and intensity model, in a similar fashion to the method provided by [92], using a dataset containing 20 dry, cadaveric femurs. They used the diffeomorphic Demons algorithm [93] to establish dense correspondence between the femur volumes in an automated fashion before completing two distinct PCAs for shape and intensity.

Combined SSIMs are easier to implement, have fewer parameters to optimize, and require less memory than independent SSIMs, ultimately reducing the time taken to search for patient–specific model parameters [1]. They model the correlations between bone intensity and shape information which will affect the quality of the reconstructed model, however, the relationship between these subspaces is not well–studied and requires further investigation [2]. The set of shape and intensity model parameters are combined into a vector as follows [1]:

$$\mathbf{w}_{\text{combined}} = \begin{pmatrix} \mathbf{W}_m \mathbf{w}_m \\ \mathbf{w}_q \end{pmatrix} = \begin{pmatrix} \mathbf{W}_m \mathbf{E}_m^T (\mathbf{s} - \bar{\mathbf{s}}) \\ \mathbf{E}_q^T (\mathbf{i} - \bar{\mathbf{i}}) \end{pmatrix} \tag{11}$$

where $\mathbf{E}_m^T$ and $\mathbf{E}_q^T$ are the transposed eigenvector matrices of shape and intensity, respectively. A diagonal matrix of shape parameters, $\mathbf{W}_m$, is used to reconcile the shape and intensity model parameters, since they have different magnitudes [1]:

$$\mathbf{W}_m = \beta \mathbf{I} \tag{12}$$

where $\mathbf{I}$ is a unit matrix and β is the ratio of total intensity variation to the total shape variation [1]:

$$\beta = \frac{\sum_{m=1}^{M} u_m}{\sum_{q=1}^{Q} u_q} \tag{13}$$

A more compact model is then obtained by discarding the correlation between the shape and intensity model parameters, by applying a third PCA on $\mathbf{w}_{\text{combined}}$ [1].

An alternative approach to constructing a combined SSIM is to perform a single PCA on a volume vector:

$$\mathbf{v}_n = (x_1^n, y_1^n, z_1^n, i_1^n, ..., x_P^n, y_P^n, z_P^n, i_P^n) \qquad \mathbf{v}_n \in \mathbb{R}^{4P} \tag{14}$$

similarly to the process described in Section II-A1, where the n^{th} volume, $\mathbf{v}_n$, is represented by a discrete set of P landmark points, x_p^n, y_p^n and z_p^n denote a point's Euclidean coordinates, and i_p^n denotes its corresponding intensity value.

5) Validation of SSIMs: The method and means with which the quality of the newly included intensity information is validated is not discussed in great detail in the literature. This is most likely due to the fact that the level of fidelity that is required is application–specific. The authors of [88] and [94] made use of dice coefficients to measure the segmentation accuracy of different bone regions, such as cortical and trabecular bone. These provided a percentage of volume overlap between manually segmented ground–truth volumes and their model reconstructions. An in–depth review of metrics that have been used for evaluating 3D medical image segmentations, such as volume overlap, is provided by [86]. These, however, only evaluate shape information. In order to gauge the accuracy of the reconstructed intensity information, voxel–wise comparisons have been proposed [47], [51], [62]. In fact any intensity–based similarity metrics, such as those discussed in Section II-C, can be extended to 3D and used. Futhermore, SSIM analogs for model quality metrics (such as generality, specificity and compactness) have not been encountered in the literature.

6) Variations on SSIMs: A combined SSIM known as the "InShape" model was proposed by [95], which was trained with a dataset of 15 femur specimens. They made use of Euclidean distance maps to represent the surface of the model and an intensity model similar to the one proposed by [87] to represent the intensity information. These were then combined into a spatial–intensity distribution using a level–set segmentation. The resulting SSIM, therefore, essentially stores texture–based features [95].

The authors of [21] followed the original method provided by [87] for constructing an SSIM, but opted for a novel combined representation of shape and intensity information. A hierarchical tetrahedral mesh was used to describe bone shape. While this type of data structure is more involved, it provides a high degree of flexibility and is superior when adapting to local shape structures [96]. Once established, each tetrahedron in the tetrahedral mesh was assigned an analytical function in the form of a Bernstein polynomial using a barycentric coordinate system. Bernstein polynomials provide a more efficient means of representing intensity information compared with storing an intensity value for each voxel. The intensity functions are continuous, and in explicit form, which make them easier to deform or integrate and thus ideal for non–rigid registration as well as efficient DRR rendering [96]. The greater the number of voxels encapsulated by each tetrahedron the more memory-efficient the storage of the density model, the shorter the rendering time of a DRR [96]. However, when the number of tetrahedra used approaches the number of voxels in a CT image there is an improvement in the average intensity difference and standard deviation of intensity difference when compared to a ground-truth DRR. The advantages of representing the model in this way are therefore most apparent when a coarser sampling is used. In addition, the barycentric coordinate system is symmetric, and normalized, which causes the intensity functions to be shape invariant despite the combined nature of the SSIM [96]. This characteristic is advantageous when performing computations on, and deforming, the SSIM. This SSIM paradigm has since

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

become a popular method of incorporating CT intensity values into SSMs; it is used by [29], [53], and [6].

The authors of [29] extended the method proposed by [21] by altering the way in which the tetrahedral mesh is constructed so that it is more conducive to their novel DRR rendering approach. This is discussed in more detail in Section 2.2. They also only made use of the voxel intensity values of a single patient while the method in [96] performed statistical analysis on the intensity information of all the patients.

A summary and comparison of the SSIMs which have been applied to 2D/3D bone reconstruction is provided in Table I.

7) Clinical performance requirements for statistical models of bone: There is little information in the literature regarding the clinical performance requirements of 2D/3D registration using SSIMs. One example does provide some guidance with regard to shape information; after an analysis of common orthopaedic procedures the following specifications for the accuracy and robustness of a bone shape model were proposed by [97], in order to be useful for surgical guidance:

- Root–mean–squared registration error ranges of 1 mm – 1.5 mm (2 mm – 3 mm in the worst case).
- The registration is successful in its first attempt at least 95% of the time.
- The registration takes no more than 1 minute.
- Simple and minimal preoperative and intraoperative user interaction.

To the best of our knowledge, specifications for the accuracy and robustness of the estimation of intensity information using 2D/3D registration are not available in the literature and appear to be application–specific. For example, [52] note that the trabecular region is predominantly affected by osteoporosis, which is the primary cause of compression fractures. The accuracy of this region would therefore be prioritized when developing SSIMs for clinical applications aimed at osteoporosis.

B. Rendering of DRRs

Digitally reconstructed radiograph projection is a rendering technique that simulates the X–ray imaging process in order to produce a synthetic X–ray image from a CT volume or volumetric model (An example is shown in Figure 3). The most commonly used method to do so is "ray–casting" [98]. It is derived from the Beer–Lambert Law (Equation 15), which describes the attenuation of X–rays through matter. In this approach the value of each pixel of the DRR image is computed by evaluating each voxel of a CT volume (or instance from a SSIM) which is encountered along the path of a single ray between the current pixel and a center–of–projection (COP) (see Figure 4). When a voxel is evaluated, its corresponding linear attenuation coefficient (LAC), as well as the distance the ray has "travelled", is considered. A line integral along the path of a single ray is evaluated where I is the X–ray signal intensity, L is the path from the X–ray source (COP) to a detector pixel, $\mu(E, r)$ is the energy and position–dependent LAC of the material that the X–ray passes

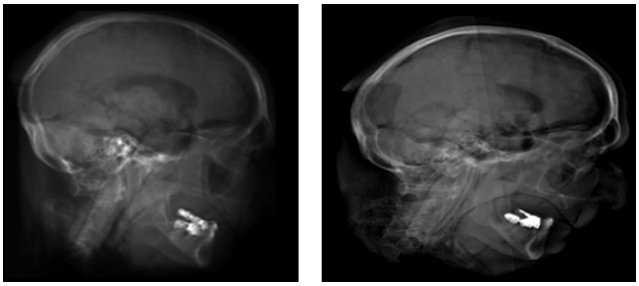


Fig. 3. Example of a DRR (left) obtained using ray–casting through a CT volume of a patient, contrasted against an actual X–ray image of the same patient (for the Lateral view).

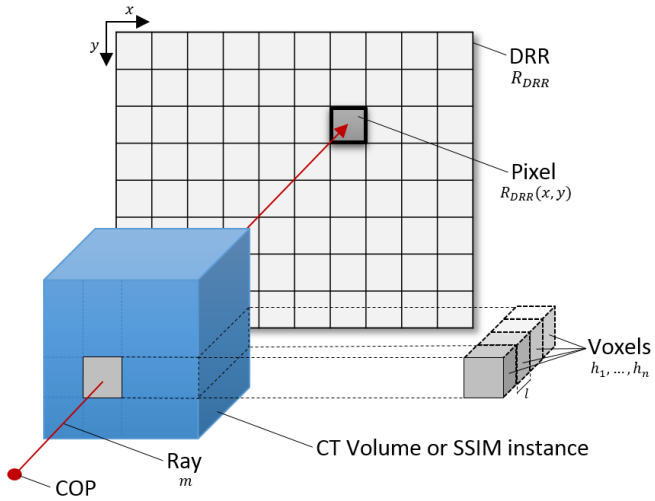


Fig. 4. An illustration of the ray–casting method. Each DRR pixel intensity value $R_{\text{DRR}}(x, y)$, is computed by evaluating the CT voxels encountered along a ray m, between a COP and the current pixel.

through and $I_0(E)$ is the incident X–ray intensity and energy spectrum [99]:

$$I = \int I_0(E) \times \exp(-\int_L \mu(E, r)dl)\ dE \tag{15}$$

This formulation is further simplified by modelling the X–ray source as monoenergetic; where more complex physical phenomena such as "scatter contamination", "beam hardening", and "veiling glare" are disregarded; and then discretized along the ray path to produce a voxelized representation [99]–[101]:

$$I = I_0 \exp(-\sum_n \mu_n l_{n,m}) \tag{16}$$

where n and m designate a specific voxel and ray, respectively; l is the intersection length and μ is the LAC. The corresponding LAC for a voxel's Hounsfield unit is obtained using Equation 17. This value is contingent on the relative LAC of water, for the particular energy spectrum of the incident rays [101].

$$h_n = \frac{(\mu_n - \mu_w) \times 10^3}{\mu_w} \tag{17}$$

where h_n is a CT Hounsfield unit, μ_w is the LAC of water for a specific CT energy, and μ_n is the LAC for the current voxel. The COP is either modelled as being close in a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

TABLE I
A COMPARISON OF SSIMS THAT HAVE BEEN APPLIED TO 2D/3D BONE RECONSTRUCTION

Author(s)	Nature of SSIM	Model representation	Type of bone	Novelty
[45]	Independent	Voxel–based	Femur	None
[95]	Combined	Voxel–based	Femur	Spatial–intensity (texture) distribution
[96]	Combined	Mesh–based	Femur and pelvis	Model intensities using Bernstein polynomials

perspective projection scheme, or infinitely far away in an orthographic projection scheme where the rays are parallel to one another (See Figure 5). The most commonly used scheme is perspective projection, which is more challenging since it involves more complex projection geometry and because it is possible for adjacent voxels to be traversed simultaneously. A number of sampling methods (also known as voxel traversal, ray traversal, or compositing methods) have been proposed to address this issue, often implementing some type of voxel interpolation. These are however beyond the scope of this paper, but are discussed in detail in [102] and [103]. The DRR

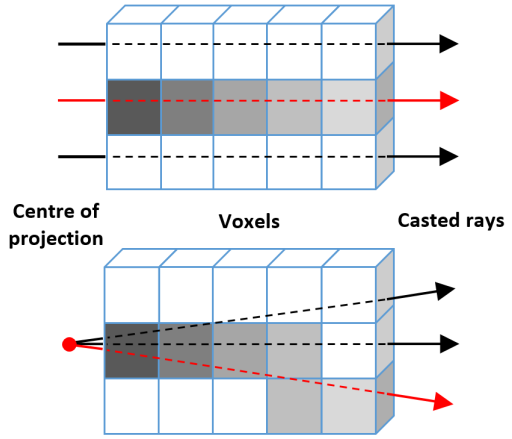


Fig. 5. An illustration of two DRR ray projection schemes. Orthographic (or orthogonal) projection models the COP as being infinitely far away (above) as opposed to perspective projection which models the COP as a point–source (below). Interpolation is typically used in order to overcome the simultaneous traversal of adjacent voxels when sampling along perspective–projected rays. The increasing greyscale values of the voxels indicate the order in which the voxels have been traversed.

rendering method can be viewed as a transform, $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$. A general formulation of the ray-casting projection through an SSIM, regardless of the sampling or projection strategy that is used, can then be described as follows [104]:

$$\mathcal{P}(\mathbf{v}(\mathbf{w})) = R_{\text{DRR}} \tag{18}$$

$$R_{DRR}(x, y) = I_0 \exp(-\int_{L_{x,y}} \mu_l \; dl) \tag{19}$$

where $\mathbf{v}(\mathbf{w})$ is an SSIM instance obtained with a specific set of model parameters $\mathbf{w}$, and $R_{\text{DRR}}(x, y)$ is a function that returns an accumulated intensity value along ray-path $L_{x,y}$ from the X–ray source (COP) to a detector pixel at the 2D coordinates, (x, y). For projection geometry-specific formulations of this equation the reader is directed to [104].

1) Variations on DRR rendering: DRR–based approaches are the most frequently reported type of 2D/3D registration methods in literature [7]. They are known to be more accurate than feature–based methods at the cost of having a longer computation time [7], [65]. The iterative nature of the algorithm causes a computational bottleneck to emerge as accurate volumetric reconstruction often requires a few thousand DRRs to be rendered [6], [64]. Efforts have since been made to reduce the time taken to obtain a DRR projection. These can be roughly categorized as hardware and software–based methods, but are not mutually exclusive distinctions [7].

Hardware–based approaches exploit parallel processing through the use of commercial graphics processing units (GPUs) to speed up the ray–casting algorithm, and consequently the registration process, in some cases by up to 25 times [7], [105]–[107].

Software–based approaches apply concepts from the computer graphics community to provide more efficient rendering algorithms. These take the form of image–order and object–order techniques. Image–order techniques involve backward projection and are typically performed pixel–by–pixel. These typically include adaptations to the ray–casting method (which itself is an image–order technique) either by reducing the number of rays, or number of samples that are considered as is the case with Monte Carlo volume rendering [108]. Object–order techniques involve forward projection and are typically performed voxel–by–voxel. These include alternatives to the standard ray–casting approach such as shear–warp factorization, wobbled splatting, Fourier volume rendering and cylindrical harmonics, among others [109]–[113]. Most object–order methods are geared towards rigid registration for purposes such as preinterventional patient alignment (image–guided registration therapy) [100]. Here it is assumed that the volume is static, while the camera pose/COP is searched for. It therefore makes sense that a popular strategy for object–order methods is to spend computational resources encoding the volume in such a way that, once initialized, can obtain a DRR from different COPs with ease. Non–rigid reconstruction, on the other hand, seeks to find a patient's specific volume by iteratively deforming an appearance model while the COP remains stationary. Image–order techniques are more amenable to this strategy.

Some methods, such as those proposed by [29], [53], and [6], have been specifically applied to the reconstruction paradigm, and utilize a combination of hardware and software–based approaches. The SSIM proposed by [21] (see Section II-A) was accompanied by a complimentary DRR rendering algorithm that adheres to the non–linear intensity distribution of its tetrahedra [53]. The algorithm proposed by [29] improved upon the DRR rendering method [96]. They noted that a X–ray attenuation value does not need to be accumulated according to the order of visibility of tetrahedra

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

along a line integral. Thus each tetrahedron can be processed independently using a projected tetrahedra approach (replacing the numerical evaluation rays intersecting the SSIM), making the algorithm both a closed–form solution and ideal for GPU implementation. This characteristic also compliments the method in [96] which uses barycentric coordinates, allowing for the simultaneous classification of the projected outline of the tetrahedra. The approach provided by [53] further improved that of [29], by focusing effort on the DRR rendering step of the reconstruction algorithm. The authors of [53] provided a way to combine the deformation and tetrahedral mesh projection steps thus enabling a full GPU implementation. They employed cell–based ray–casting which reuses the closed–form method for ray integral computation, but implemented a novel tetrahedron thickness calculation instead of the projected tetrahedra approach. This permits the advantages of a barycentric coordinate system while circumventing the need to classify the projected outline of the tetrahedra [6], [53]. A summary of notable DRR rendering methods is provided in Table II (note that $\mathcal{O}(n^2)$ represents quadratic time, $\mathcal{O}(n^3)$ represents cubic time and $\mathcal{O}(n^2\log[n])$ represents quadratic–logarithmic time).

Finally, the accuracy of the registration can be further improved through the generation of more realistic DRRs, which will inherently be more similar to a patient's X–ray image(s) [34], [95]. The ray–casting method only models the attenuation of primary photons while disregarding secondary effects such as beam hardening. Techniques to account for these effects were proposed by [101], which produce more realistic DRR renderings but also increases the already taxing computational complexity of the DRR rendering module.

C. Similarity measures

One of the primary factors influencing the accuracy of intensity–based 2D/3D reconstruction algorithms is the similarity measure, the choice of which can produce drastically different registration results [42], [118]. The DRR(s) and X–ray image(s), R_{DRR} and R_{XRAY}, are compared to one another with regard to some statistic or set of features and a numerical value is produced which quantifies the similarity of information between them. In some cases researchers refer to a cost, loss, error, merit, criterion, energy, or objective function, which the optimization strategy seeks to minimize (or maximize).

In the context of deformable image–based 2D/3D reconstruction the similarity measure can be framed as a Q–dimensional function where Q represents the number of parameters of the SSIM. The value of each location of a Q–dimensional parametric search space then corresponds to the similarity measure value for a specific set of model parameters, $\mathbf{w} = \{w_1, ..., w_Q\}$ (an instance of the SSIM) [7], [119].

The similarity measure values for each X–ray image view/COP can be consolidated as follows:

$$\text{sim}_{\text{tot}}(\mathbf{w}) = \sum_{c=1}^{C} \text{sim}(\mathcal{P}_c(\mathbf{v}(\mathbf{w})), R_{XRAY,c}) \tag{20}$$

where $\mathcal{P}_c$ represents the DRR rendering method for a COP c, and C is the total number of patient X–ray images.

1) Variations on similarity measures: The mutual information (MI) similarity measure is commonly used by 2D/3D reconstruction algorithms [118]. It compares the probability distribution of the pixel intensity values (histograms) in a DRR and X–ray image. While it is robust to noisy and occluded X–ray images, it only considers intensities, thus disregarding any spatial information [29]. A number of MI variations have since been proposed which incorporate spatial information. The asymmetric multi–feature MI similarity measure uses additional spatial features in the form of intensity gradients and is especially effective when reconstructing from a small number of patient X–ray images [7]. The performance of two novel variations of the MI similarity measure, distance coefficient MI and distance weighted MI, were compared by [120]. They found that these measures outperformed conventional MI without having a significant impact on computation time. The MI similarity measure was also adapted to include spatial information by [121], by incorporating the Kullback–Leibler bound into a Markov random field model. Their experiments show that their measure, which they have named the Maximization of Mutual Information, not only outperforms the MI similarity measure, but is also robust to occlusions present in X–ray images.

The researchers in [29] discuss the similarity measure-related complications that arise when the statistical model is trained using a dry-bone dataset. Images of live subjects often contain other organs, as well as soft tissue, which adds noise and reduces the contrast of the images. Despite the merits of MI they found that it did not perform well and therefore included a soft-tissue model in their SSIM in order to ameliorate these effects. Alternatively, the researchers in [122] proposed the use of normalized cross correlation (NCC) as an initial similarity measure and then automatically switched to using the variance weighted sum of localized normalized correlation (VLNC). The NCC similarity measure is not very accurate, but reduces the influence of regions containing foreign objects. The VLNC focuses on high-variance regions, assuming that these contain relevant information. In doing so they improve the accuracy of the final registration.

An extensive study of six similarity measures was completed in the context of intensity–based 3D–2D registration by [123]. These included cross correlation, entropy, mutual information, gradient correlation, pattern intensity and gradient difference. By comparing these against a gold standard (landmark–based registration), the measures were ranked with regard to how accurate and robust they were. The study concluded that the pattern intensity and gradient difference similarity measures were accurate and robust, even when interventional instruments, thin–line structures and soft tissue were present [42].

The authors of [7] have suggested that the similarity measure utilized by the optimization strategy be specifically adapted to the understanding of the image formation process as well as the relationship between the intensities of the DRR and X–ray images. One such similarity measure was proposed by [124], which was tailor-made to the statistics of CT and X–ray image acquisition. In this approach Poisson and Gaussian probability distributions were used to model the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

TABLE II
NOTABLE DRR RENDERING METHODS WHICH INCLUDE SOFTWARE, HARDWARE, AND SUBSPACE/SEGMENTATION–BASED APPROACHES

Author(s)	Name of technique	Voxel/tetrahedra–based	Time complexity (if available)	GPU–based acceleration	Comments
[114], [115]	Ray–casting	Voxel	$\mathcal{O}(n^3)$	None	- Used as a benchmark. - Computationally inefficient.
[29]	Projected tetrahedra	Tetrahedra	Not available	Partial	- Suitable for mesh–based 2D/3D reconstruction.
[116]	Light fields	Voxel	$\mathcal{O}(n^2)$	No	- Reduced time complexity. - The 4D space must be densely sampled to provide satisfactory images.
[112], [117]	Fourier volume rendering	Voxel	$\mathcal{O}(n^2\log[n])$	Full	- Large memory demands. - Independence of image quality on the sampling step size. - Perspective projection not supported.
[110]	Shear–warp factorization	Voxel	$\mathcal{O}(n^3)$	None	- Well suited rigid 2D/3D registration. - Reduced computations when recomputing projections.
[111]	Splatting	Voxel	$\mathcal{O}(n^3)$	None	- Very accurate, but difficult to implement. - Aliasing artefacts.
[113]	Cylindrical harmonics	Voxel	Not available	None	- Fast. - Reduced computations when recomputing projections.
[6], [53]	Cell–based ray–casting with novel tetrahedral thickness calculation	Tetrahedra	Not available	Full	- Suitable for mesh–based 2D/3D reconstruction.

intensity values of the two modalities (since the photon noise which is present in authentic X–ray images is Poisson-like in nature). Ultimately, new similarity measures were calculated from the assumed distributions, using maximum likelihood estimation, which were especially robust to image noise [7], [124]. Another such similarity measure was presented by [125], who established the feasibility of constructing a similarity measure using coefficients from an orthogonal set of base functions, and by decomposing X–ray and DRR images into orthogonal Zernike moments [7], [125]. The advantages of doing so include better robustness to histogram differences, invariance to in-plane rotation, and control over the level of detail that is considered. These bespoke similarity measures, however, have yet to be applied specifically to the 2D/3D reconstruction paradigm; their performance in this regard, as well as in comparison to other similarity measures, is therefore unclear.

D. Optimization strategies

The optimization strategy for a deformable image-based 2D/3D registration algorithm is iterative, and describes how a patient-specific bone reconstruction is searched for. This is accomplished by providing a procedure for how the model parameters for the next iteration are chosen based on the differences between the patient's X–ray image(s) and the DRR projection(s) for the current iteration. As previously mentioned, these differences are measured using one or more similarity measures, which the optimization strategy seeks to maximize:

$$\mathbf{w}_{\text{patient}} = \underset{\mathbf{w}}{\operatorname{argmax}}\, \text{sim}_{\text{tot}}(\mathbf{w}) + R(\mathbf{w}) \tag{21}$$

where $\mathbf{w}_{\text{patient}}$ is the final set of optimal model parameters selected to represent the patient's bone volume. The regularization term, $R(\mathbf{w})$, accounts for the nature and constraints of the transform, $\mathcal{P}_c$ [61]. These include prior knowledge regarding the bone's shape and intensity, ensuring smooth deformations and dampening the effects of outliers (for an interesting discussion regarding how the regularization term is related to the covariance matrix, S, the reader is referred to [126]). Figure 6 provides a useful overview of the DRR-based 2D/3D reconstruction algorithm, including important mathematical formulae.

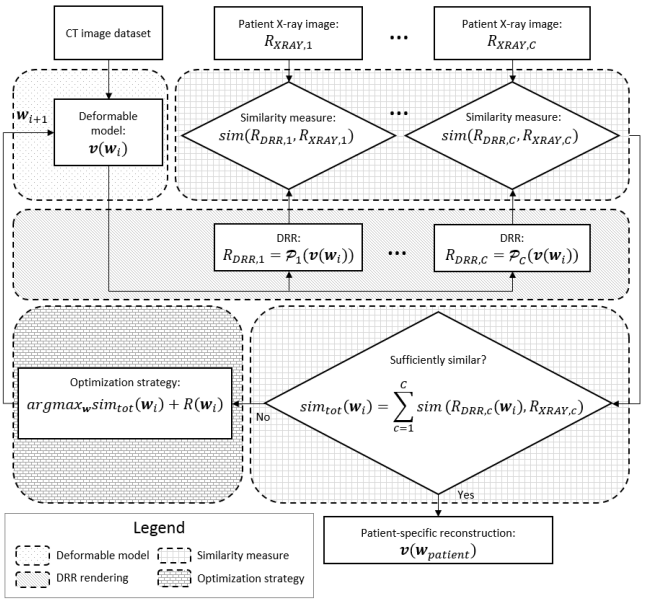


Fig. 6. A detailed overview of the DRR–based 2D/3D bone reconstruction algorithm (including a mathematical formulation). The modules which comprise it include the deformable model, of which an SSIM $\mathbf{v}(\mathbf{w}_i)$ is one type, DRR rendering R_{DRR}, similarity measure sim_{tot}, and optimization strategy. The process is iterative, where $\mathbf{w}_i$ is the set of model parameters for iteration i, and which is concluded once the value of sim_{tot} exceeds a pre–defined threshold.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

The similarity measure, when represented as a function of registration error, should ideally be a convex, monotonic function, having a sharp minimum (indicating the true registered position or best possible similarity value) and successively decreasing values the further away from the minimum they are [7], [119]. This ensures good registration results when implementing iterative, local optimization strategies such as Powell's method, the downhill–simplex method, gradient–descent, [7], [29], [34], [53].

However, as is the case with intensity–based 2D/3D registration, the registration error function is highly non–convex, making the optimization strategy susceptible to getting caught in local minima [61], [64]. The similarity measure therefore has a small capture–range, which is the subset of values around the global minimum which do form a convex, monotonic function (see Figure 7) [119].

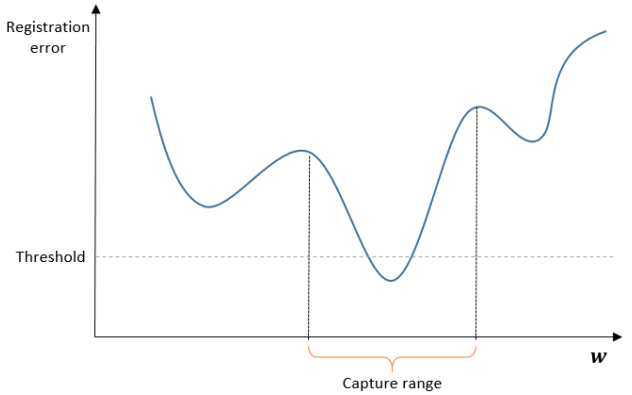


Fig. 7. A plot of intensity–based 2D/3D registration error as a function of the model parameter values. The range of values around the global minimum which form a convex, monotonic function is known as the capture–range.

1) Variations on the optimization strategy: In order to widen the capture range, and increase chances of convergence, researchers have implemented global optimization methods and heuristic search strategies such as simulated annealing and Monte Carlo random sampling [7], [127], [128]. While these increase the probability of finding the global maximum, the authors of [7] note that they are problem–specific. We are not aware of any DRR–based 2D/3D reconstruction methods that make use of these.

A simpler and more well established approach is to use a hierarchical multi–resolution and/or multi–scale optimization strategy, adopting a coarse–to–fine approach to the parameter search [7], [21], [129]. They first downsample the images to reduce their complexity (and avoid local maxima) and then successively introduce finer detail in stages. The complexity which is gradually increased refers either to that of the data, warp or model [129]. An extensive survey of multi–resolution approaches is provided by [130]. The authors of [21], [2] and [53] implemented similar multi–resolution optimization strategies. They downsampled the resolution of the X–ray images and reduced the number of model parameters which, during later iterations, were both increased to ensure an accurate reconstruction of the pelvis bone. These optimization strategies not only increase the capture range of the algorithm (by avoiding local maxima), but also improve its computation time. The authors of [7] noted that downsampling the image may suppress some image features while making other non–corresponding features more similar. However, they also noted that these approaches are especially suitable for 2D/3D reconstruction.

Linear optimization strategies such as the downhill simplex method, which was implemented and applied to the femur and pelvis bones by [29], are sometimes preferred to non–linear strategies as they are often computationally cheaper [46], [53]. However, from what is evident in intensity–based 2D/3D bone reconstruction literature, a number of non–linear alternative optimization strategies have been proposed which provide more accurate model reconstructions [131]. Furthermore, linear strategies are limited in that they have an inability to cope with large deformations [61]. Levenberg–Marquardt optimization is a non–linear strategy that is reported by [6], [45]; both applied it to the femur bone because it provided accurate reconstructions while still outperforming gradient–descent–based approaches with regard to computation time [6]. Levenberg–Marquardt optimization was also implemented by [59], who applied it to the pelvis.

Another way to avoid local maxima is to implement a pose initialization method before proceeding with the optimization strategy [64], [132]. Such methods aim to initialize the first SSIM instance in such a way that its initial DRR projection(s) are relatively close to that of the target X–ray images. This is typically accomplished with a manual landmarking process, which is time–consuming. An accurate initialization method requiring only three manually labeled landmarks has been developed by [133], which uses epipolar geometry to assist in localizing points between two X–ray perspectives. However, methods such as these are still prone to human error [4], [134]. Fully automatic methods have since been explored, but face many challenges; noise, poor image contrast and the superposition of bone structures at different depths make it difficult to distinguish the bone–of–interest [134]. Random forest regression was implemented by [3] for fully automated contour extraction and pose initialization. This method provided accurate results, but was only tested using simulated patient X–ray images. They also did not comment on the method's influence on the computation time of the overall algorithm.

Finally, supervised learning–based similarity measures have been demonstrated by researchers to widen the capture–range of intensity–based 2D/3D registration algorithms, but are far more computationally expensive [64]. We are also not aware of any DRR–based 2D/3D reconstruction methods that make use of these.

A summary of the optimization strategies discussed in this section, which have been specifically applied to DRR–based 2D/3D bone reconstruction, is provided in Table III.

E. Investigations on the algorithm as a whole

The use of two X–ray images enables a significant improvement in the accuracy of the reconstructed model than if only a single X–ray image is used. However, according to [21], the use of additional X–ray images has diminished

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

TABLE III
A SUMMARY AND COMPARISON OF OPTIMIZATION STRATEGIES WHICH HAVE BEEN USED FOR DRR–BASED 2D/3D BONE RECONSTRUCTION

Author(s)	Nature of SSIM	DRR rendering method	Optimization strategy	Similarity measure	Type of bone	Number of X–ray images	Average reconstruction error (mm)	Type of optimization
[21]	Combined	Not clear (probably ray–casting)	Novel multi–resolution non–linear optimization	(Normalized) Mutual information	Pelvis	2 or more	2.0	Non–linear
[29]	Combined	Projected tetrahedra	Downhill simplex algorithm	(Normalized) Mutual information	Femur and pelvis	2 or more	2.0	Non–linear
[45]	Independent	Ray–casting	Levenberg–Marquardt	Maximization of mutual information	Femur	2 or more	1.5	Non–linear
[53]	Combined	Cell–based ray–casting with novel tetrahedral thickness calculation	Gradient descent with X–ray sub–sampling	(Normalized) Mutual information	Pelvis	1	2.0	Linear
[6]	Combined	Cell–based ray–casting with novel tetrahedral thickness calculation	Levenberg–Marquardt	(Normalized) Mutual information	Femur	2	1.18	Nonlinear
[34]	Combined	Ray–casting	Evolutionary	Mean reciprocal square difference normalized mutual information	Femur	2 or more	4.24	Unclear
[59]	Combined	Ray–casting	Levenberg–Marquardt	Maximization of mutual information	Femur	2 or more	1.2	Non–linear

returns in terms of accuracy. This is arguably not worth the extra computation time or additional radiation exposure to the patient. It was demonstrated in [21] that the best registration accuracy was obtained when the angle between the two perspectives is orthogonal. This makes sense since the information provided by the two images in this case is the least correlated. However, it does not necessarily follow that this is always the case. A good example is the pelvis bone, which has a symmetric structure; an AP and LAT X–ray image combination will have a large superposition of structure - especially in regions containing joints. While orthogonality is important, the selection of the best imaging direction is dependent on the anatomical structure-of-interest.

The 2D/3D reconstruction of joint structures using an articulated statistical shape and intensity model (ASSIM) has been investigated by [135]. They reconstructed pelvic and femoral bones simultaneously in order to infer two patient-specific surgically relevant parameters from a single X–ray image. They extended their previous SSIM, presented in [53], to include both the pelvic and femoral bones, but constrained the 3D rotation of the proximal femur such that its center of rotation corresponds to the center of a sphere made to fit the acetabulum. Their method produced results close to the CT gold standard, but became less accurate when faced with outliers where some regions pertinent to the inference were not visible in the AP X–ray image. They proposed increasing the number of training examples in order to overcome this limitation [135].

The authors of [34] investigated the performance of a number of similarity measure–optimization strategy combinations for appearance–model–based 2D/3D registration of the human femur [1], [34], [95]. These included Powell optimization, 1+1 evolutionary optimization, mean reciprocal square difference metric, Matte's mutual information metric and the normalized mutual information metric [136]–[139]. Using the InShape SSIM, the following parameters were optimized: the six degrees of freedom of the rigid 3D transformation (three rotational and three translational) as well as the modes of shape and intensity variation. They found that the mean reciprocal square difference metric in combination with 1+1 evolutionary optimization, as well as normalized mutual information measure combined with the 1+1 evolutionary optimization strategy, resulted in the best performance.

III. DISCUSSION

During our literature review we noted a number of challenges and insights regarding DRR–based 2D/3D bone reconstruction, which may serve as future avenues of research. These are discussed according to the algorithm module to which they are most relevant.

A. Statistical shape and intensity models

While SSIMs are robust to artefacts and noise, they require a large amount of training data in order to obtain a model capable of expressing all possible target shapes. The generalization of the classical SSM approach, known as GPMMs, and introduced by [60], makes it possible to intuitively and analytically define covariance functions, which enable the construction of expressive shape priors even when only a few example shapes are available (essentially providing a framework for combining template and statistically based models, and allowing for dense point correspondence at any resolution). Their approach could be extended to include SSIMs, ultimately providing more accurate bone shape and intensity estimations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

Furthermore, it is unclear whether independent or combined SSIMs are superior for the purposes of non–rigid 2D/3D bone reconstruction; the relationship between the shape and intensity subspaces is not well studied and requires further investigation.

B. Digitally reconstructed radiographs

The original patient X–ray image can be regarded as the gold–standard, equivalent to a maximum similarity measure value. However, this is not realistically possible because of the presence of noise. Also, a DRR rendering algorithm's failure to account for secondary imaging effects such as veiling glare and beam–hardening further prevents the DRR from attaining a gold standard value [101]. We therefore define the "ideal" threshold, which takes this compromise into account. A further concession in accuracy is made when we consider that DRRs will actually be obtained from an SSIM instance and not the patient CT volume, since it is not available. Thus another value, the "best–case" threshold, is defined, which accounts for the consequent loss of information resulting from the PCA approximation. Finally, the "sufficient" threshold observes the fact that the best–case threshold may be unattainable since the model parameter search is limited to the 2D domain, as well as the performance of the chosen optimization strategy. It may also be the case that the reconstruction accuracy necessary for a particular application is attained prior to achieving the best–case threshold. An illustration of these concepts is provided in Figure 8.

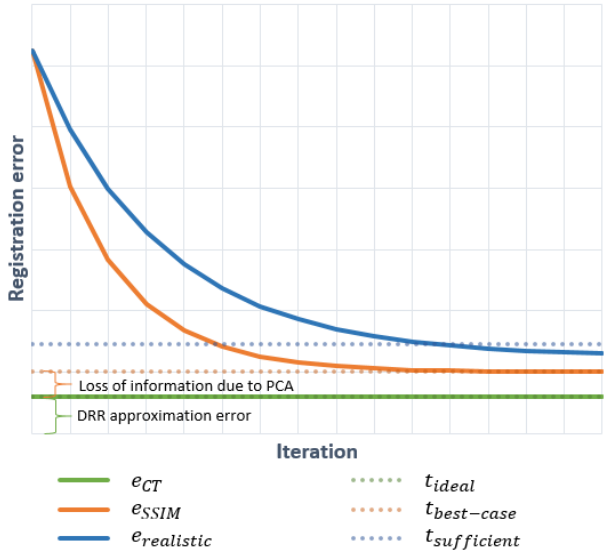


Fig. 8. Iterative 2D/3D reconstruction optimizations, depicted as decreasing registration error (increasing similarity). The best possible registration accuracy, varies under different circumstances: the ideal threshold corresponds to a DRR obtained directly from the patient CT volume ($e_{CT} = 1 - \text{sim}(R_{DRR,CT}, R_{XRAY})$); the best–case threshold corresponds to a DRR obtained from an SSIM instance, the model parameters of which were searched for using the CT volume ($e_{SSIM} = 1 - \text{sim}(R_{DRR,SSIM}, R_{XRAY})$); the sufficient threshold also corresponds to a DRR obtained from an SSIM instance, but accounts for the fact that the best–case threshold may not be reachable when the model parameters are searched for using only the patient's X–ray image(s).

Some researchers have established the feasibility of using partial DRRs to guide the model parameter search. However, the decision regarding which regions to select is made on intuition. Should there be access to a corresponding CT and X–ray image dataset for a number of patients, a potentially insightful study can be completed. Firstly, an SSIM can be built using the (hand–segmented) patient CT volumes. The ideal similarity threshold value (for a patient) can be computed using a DRR obtained from their CT volume, and their X–ray image. The best–case similarity value can be obtained in a similar fashion, using a near–perfect model fit (searched for in the 3D domain using the patient's CT volume). Then, by comparing the original X–ray image to the DRR obtained from the CT volume as well as the DRR obtained from the SSIM, the regions and/or features which should be focused on can be determined empirically.

While a number of DRR rendering methods have been specifically proposed for application to 2D/3D registration, only the ray–casting and tetrahedra–based projection methods have been applied to 2D/3D bone reconstruction. The performance of different DRR volume rendering methods in this context have not yet been compared. Object–order methods such as cylindrical harmonics, shear–warp factorization and frequency domain–based rendering spend computational resources encoding the volume in such a way that obtaining a DRR from a large number of COPs is made much simpler. This allows DRR projections to be iteratively re–rendered at a reduced computational cost until an optimal COP is found. While this is beneficial for 2D/3D registration (where the assumption is that the object does not change), this is not the case for 2D/3D reconstruction where the COP(s) are fixed and the object is iteratively deformed. However, these methods are of use for the initial pose estimation; it may serve a 2D/3D reconstruction algorithm well to include a two–part DRR rendering module where an image–order technique takes over from an object–order technique once the pose has been established. If additional accuracy is required, the methods provided by [101], which account for secondary physics effects, can be employed to further improve the registration. However, it is important to note that the DRR itself is not the goal of the 2D/3D reconstruction algorithm; it is only a means to determine a model reconstruction (optimizing a set of model parameters). Any additional DRR rendering accuracy, other than what would enable an accurate reconstruction, is unnecessary, and it would be beneficial to rather trade this accuracy for a faster computation time (See Figure 8).

As is shown in Table II, many rendering methods have a $\mathcal{O}(n^3)$ time complexity. These methods should be avoided (at least once the pose of the volume/X–ray source has been established).

C. Similarity measure

In the majority of cases, the mutual information similarity measure, or some variation thereof, has been used for image–based 2D/3D reconstruction (and is also popular for rigid 2D/3D registration). Newer similarity measures, such as those proposed by [42], [124] and [125] have been specifically

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

adapted to the context of X–ray images, but have not yet been applied to the context of reconstruction. These may prove to be superior to conventional similarity measures.

D. The algorithm as a whole

The 2D/3D reconstruction of an anatomical structure from a single X–ray image is a valuable area of research because clinical routines that require only one are far more common and cost less. However, since depth and scale are correlated, the precise estimation of either the size of the anatomical structure, or its depth, becomes more challenging when constrained to a single view [94]. The researchers in [37] first demonstrated such an algorithm using a feature-based approach to reconstruct a shape model. Their model was comprised of average templates of both 2D thickness images and height maps that were projected from 3D CT training images. The single X–ray image informed how the templates were aligned and warped using carefully selected landmarks, and then combined to obtain an estimated 3D shape. The researchers in [50] were the first to reconstruct both shape and intensity information from a single X–ray image. They extended [37]'s method by also warping the volumetric BMD, and normalized it, to match that of the input X–ray image. Furthermore, having used a calibration phantom, they were able to estimate actual BMD values. In [53] the researchers reconstructed both shape and intensity information, but made use of an SSIM and no calibration phantom. However, the focus of their research was on the computation time and accuracy of their proposed DRR renderer. In [135] they reported the estimation accuracy of surgically relevant pelvic measurements, but did not explicitly report on the intensity reconstruction accuracy - it was only used as a proxy to improve the pelvic shape estimation.

The aforementioned methods, however, all utilized an artificial X–ray image (DRR) to represent the patient data. The researchers in [42] were the first to demonstrate such an algorithm using a real X–ray image, but only reconstructed a shape model. They first proposed a surface–based iterative scaled rigid registration using a standard AP X–ray image, and then a surface–based iterative affine registration using 2D lateral fluoroscopy [42], [43]. They exploited two parameters, namely the image scale (in mm/pixel) and the distance from the COP to the detector (the imaging plane), which can be retrieved from a DICOM image (provided that the anatomical structure was imaged using a standard clinical procedure). The researchers in [47], [62], [140] and [55] have reported success when using a single image of actual patients, and also reported on the accuracy of the reconstructed shape and intensity information, but only when using DXA. Dual–energy X–ray images provides two advantages over conventional X–ray images. Firstly, DXA uses orthographic projection geometry which eliminates the need for camera calibration [94]. Spatial distortions, such as depth–dependent object magnification effects, which complicate shape estimation are thus avoided [140]. Secondly, DXA is the standard for BMD estimation and the use of a DXA image over a conventional X–ray image, along with a calibration phantom, allows for accurate BMD estimation [141]. However, DXA requires a specialized machine, to which access may be limited. It is not part of a typical clinical routine and has a limited field–of–view. Therefore, we believe that the 2D/3D reconstruction of both shape and intensity information of an anatomical structure from a single conventional X–ray image is still an unsolved problem, and worthy of future research efforts.

There is currently not enough data to objectively compare the performance of different DRR–based 2D/3D bone reconstruction algorithms (in order to determine how suitable they are) [1]. Researchers employ different datasets and are often unclear about the evaluation methodology that they use, nor does a standard methodology exist, to validate and benchmark these algorithms [7]. Furthermore, researchers either focus on different bones or use different types of modules (SSIMs, DRR rendering techniques, similarity measures and optimization strategies) within their algorithm. It is clear that the interactions between these modules are complex and greatly affect the overall accuracy of the model reconstructions, as well as the time spent computing them. It would therefore be worthwhile to formulate a standardized means of comparing and benchmarking 2D/3D bone reconstruction algorithms in order to aid future research efforts. To the best of our knowledge, only [34] have made an effort to compare different 2D/3D reconstruction algorithms; they implemented and compared the performance of different similarity measure–optimization strategy combinations. An ideal evaluation methodology should include a complete open–source algorithm, which would allow the different modules to be interchanged, and measured (this may not always be possible since algorithm modules such as those proposed by [21], [29] and [53] are interwoven), and would include a standard dataset (for at least one type of bone).

Of the reconstruction algorithms that we have investigated, only [29] and [135] have shown that their algorithm can be applied to more than one type of bone (femur and pelvis), and still yield a sufficiently accurate reconstruction. This is worth considering since unique complexities and challenges exist for different types of bone, such as bone structure superposition when dealing with scapulae, or intricate bone substructures as is the case with the skull. Therefore, the success of an algorithm when applied to one particular type of bone does not necessarily guarantee its success when applied to another.

IV. Conclusion

We have presented a review of the literature pertaining to DRR–based 2D/3D bone reconstruction from X–ray images using statistical shape and intensity models. This was completed using unambiguous terminology, and a unified mathematical formulation of the problem, in a common conceptual framework. We have also provided a discussion of the shortcomings, recent adaptations and persisting challenges of this approach along with insights for future research.

Ultimately, 2D/3D bone reconstruction methods have the potential to enable clinical procedures that normally require access to CT and MRI machines and would otherwise be impossible to perform given only a patient's X–ray images. This is especially pertinent to resource–limited settings where access to these 3D imaging technologies may not be possible.

The option of a relatively cheap and more easily accessible compromise will have a beneficial impact. In addition, by minimizing a patient's exposure to ionizing radiation (compared with CT) their implicit risk of cancer can be reduced. Imaging artefacts that are caused by the presence of metal implants can also be avoided. The reconstruction of 3D patient–specific bone models from X–ray images is therefore an important and powerful technique in medical imaging, and is worthy of future research efforts.

REFERENCES

[1] N. Sarkalkan, S. Nazli, W. Harrie, and A. A. Zadpoor, "Statistical shape and appearance models of bones," *Bone*, vol. 60, pp. 129–140, 2014.

[2] G. Chintalapani, O. Sadowsky, L. M. Ellingsen, J. L. Prince, and R. H. Taylor, "Integrating statistical models of bone density into shape based 2D–3D registration framework," in *MICCAI 2009 Workshop: Probablistic Models for Medical Image Analysis*, 2009, pp. 151–161.

[3] W. Yu, C. Chu, M. Tannast, and G. Zheng, "Fully automatic reconstruction of personalized 3D volumes of the proximal femur from 2D X–ray images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 9, pp. 1673–1685, Sep. 2016.

[4] Y. Chaibi *et al.*, "Fast 3D reconstruction of the lower limb using a parametric model and statistical inferences and clinical measurements calculation from biplanar X–rays," *Comput. Methods Biomech. Biomed. Engin.*, vol. 15, no. 5, pp. 457–466, 2012.

[5] L. Zollei, E. Grimson, A. Norbash, and W. Wells, "2D–3D rigid registration of X–ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 2. IEEE, 2001, pp. 696–703.

[6] O. Klima, P. Kleparnik, M. Spanel, and P. Zemcik, "Intensity–based femoral atlas 2D/3D registration using Levenberg–Marquardt optimisation," in *Proc. of SPIE Medical Imaging*, vol. 9788F, Mar. 2016, pp. 1–12.

[7] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image–guided interventions," *Med. Image Anal.*, vol. 16, no. 3, pp. 642–661, Apr. 2012.

[8] P. Galibarov, P. Prendergast, and A. Lennon, "A method to reconstruct patient–specific proximal femur surface models from planar pre–operative radiographs," *Medical engineering & physics*, vol. 32, no. 10, pp. 1180–1188, 2010.

[9] N. Baka *et al.*, "Statistical shape model–based femur kinematics from biplane fluoroscopy," *IEEE Trans Med Imag*, vol. 31, no. 8, pp. 1573–1583, Aug. 2012.

[10] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, "Strategies for reducing radiation dose in CT," *Radiologic Clinics of North America*, vol. 47, no. 1, pp. 27–40, 2009.

[11] M. Wybier and P. Bossard, "Musculoskeletal imaging in progress: the EOS imaging system," *Joint Bone Spine*, vol. 80, no. 3, pp. 238–243, 2013.

[12] R. Gheno *et al.*, "Three–dimensional measurements of the lower extremity in children and adolescents using a low–dose biplanar X–ray device," *European radiology*, vol. 22, no. 4, pp. 765–771, 2012.

[13] T. Jerbi, V. Burdin, J. Leboucher, E. Stindel, and C. Roux, "2D–3D frequency registration using a low–dose radiographic system for knee motion estimation," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 3, pp. 813–820, Mar. 2013.

[14] G. Sharma *et al.*, "Radiological method for measuring patellofemoral tracking and tibiofemoral kinematics before and after total knee replacement," *Bone and Joint Research*, vol. 1, no. 10, pp. 263–271, 2012.

[15] M. Valenti, "Statistical shape models based 2D/3D registration methods for knee orthopaedic surgery," Ph.D. dissertation, Technical University of Denmark, 2000. 125, 129, 2016.

[16] L. Caponetti and A. M. Fanelli, "3D bone reconstruction from two X–ray views," in *Proceedings of the Twelfth Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1990.

[17] C. E. Aubin, J. L. Descrimes, J. Dansereau, W. Skalli, F. Lavaste, and H. Labelle, "Geometrical modeling of the spine and the thorax for the biomechanical analysis of scoliotic deformities using the finite element method," *Ann. Chir.*, vol. 49, no. 8, pp. 749–761, 1995.

[18] M. Fleute, S. Lavallée, and R. Julliard, "Incorporating a statistically based shape model into a system for computer–assisted anterior cruciate ligament surgery," *Med. Image Anal.*, vol. 3, no. 3, pp. 209–222, Sep. 1999.

[19] S. Laporte, W. Skalli, J. A. De Guise, F. Lavaste, and D. Mitton, "A biplanar reconstruction method based on 2D and 3D contours: Application to the distal femur," *Comput. Methods Biomech. Biomed. Engin.*, vol. 6, no. 1, pp. 1–6, 2003.

[20] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise, "3D/2D registration and segmentation of scoliotic vertebrae using statistical models," *Comput. Med. Imaging Graph.*, vol. 27, no. 5, pp. 321–337, Sep. 2003.

[21] J. Yao and R. H. Taylor, "Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 1329–1334.

[22] A. Le Bras *et al.*, "3D reconstruction of the proximal femur with low–dose digital stereoradiography," *Comput. Aided Surg.*, vol. 9, no. 3, pp. 51–57, 2004.

[23] V. Pomero, D. Mitton, S. Laporte, J. A. de Guise, and W. Skalli, "Fast accurate stereoradiographic 3D–reconstruction of the spine using a combined geometric and statistic model," *Clin. Biomech.*, vol. 19, no. 3, pp. 240–247, Mar. 2004.

[24] T. S. Y. Tang and R. E. Ellis, "2D/3D deformable registration using a hybrid atlas," *Med. Image Comput. Comput. Assist. Interv.*, vol. 8, no. Pt 2, pp. 223–230, 2005.

[25] D. Mitton *et al.*, "3D reconstruction of the pelvis from bi–planar radiography," *Comput. Methods Biomech. Biomed. Engin.*, vol. 9, no. 1, pp. 1–5, Feb. 2006.

[26] G. Zheng, "Reconstruction of patient–specific 3D bone model from biplanar X–ray images and point distribution models," in *Image Processing, 2006 IEEE International Conference on*. IEEE, 2006, pp. 1197–1200.

[27] H. Lamecker, T. H. Wenckebach, and H. C. Hege, "Atlas-based 3D–Shape reconstruction from X–Ray images," in *18th International Conference on Pattern Recognition (ICPR'06)*, 2006, pp. 371–374.

[28] M. Gunay, M. B. Shim, and K. Shimada, "Cost– and time–effective three–dimensional bone–shape reconstruction from X–ray images," *Int. J. Med. Robot.*, vol. 3, no. 4, pp. 323–335, Dec. 2007.

[29] O. Sadowsky, G. Chintalapani, and R. H. Taylor, "Deformable 2D–3D registration of the pelvis with a limited field of view, using shape statistics," in *International Conference on Medical Image Computing and Computer–Assisted Intervention*. Springer, 2007, pp. 519–526.

[30] M. K. Lee *et al.*, "The study of femoral 3D reconstruction process based on anatomical parameters using a numerical method," *JBSE*, vol. 3, no. 3, pp. 443–451, 2008.

[31] R. Dumas *et al.*, "A semi–automated method using interpolation and optimisation for the 3D reconstruction of the spine from bi–planar radiography: a precision and accuracy study," *Med. Biol. Eng. Comput.*, vol. 46, no. 1, pp. 85–92, Jan. 2008.

[32] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache, "Articulated spine models for 3D reconstruction from partial radiographic data," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 11, pp. 2565–2574, Nov. 2008.

[33] A. Hurvitz and L. Joskowicz, "Registration of a CT–like atlas to fluoroscopic X–ray images using intensity correspondences," *International journal of computer assisted radiology and surgery*, vol. 3, no. 6, pp. 493–504, 2008.

[34] P. Steininger *et al.*, "Comparison of different metrics for Appearance–Model–Based 2D/3D–registration with X–ray images," in *Informatik aktuell*, 2008, pp. 122–126.

[35] S. Kadoury, F. Cheriet, and H. Labelle, "Personalized X–ray 3D reconstruction of the scoliotic spine from hybrid statistical and image–based models," *IEEE Trans. Med. Imaging*, vol. 28, no. 9, pp. 1422–1435, Sep. 2009.

[36] G. Zheng and S. Schumann, "3D reconstruction of a patient–specific surface model of the proximal femur from calibrated X–ray radiographs: A validation study," *Medical physics*, vol. 36, no. 4, pp. 1155–1166, 2009.

[37] C. Langton, S. Pisharody, and J. Keyak, "Generation of a 3D proximal femur shape from a single projection 2D radiographic image," *Osteoporosis international*, vol. 20, no. 3, pp. 455–461, 2009.

[38] L. Humbert, J. A. De Guise, B. Aubert, B. Godbout, and W. Skalli, "3D reconstruction of the spine from biplanar X–rays using parametric models based on transversal and longitudinal inferences," *Med. Eng. Phys.*, vol. 31, no. 6, pp. 681–687, Jul. 2009.

[39] O. Ahmad, K. Ramamurthi, K. E. Wilson, K. Engelke, R. L. Prince, and R. H. Taylor, "Volumetric DXA (VXA): a new method to extract 3D information from multiple in vivo DXA images," *Journal of bone and mineral research*, vol. 25, no. 12, pp. 2744–2751, 2010.

[40] B. Zhang, S. Sun, J. Sun, Z. Chi, and C. Xi, "3D reconstruction method from biplanar radiography using DLT algorithm: Application to the femur," in *2010 First International Conference on Pervasive Computing, Signal Processing and Applications*, 2010, pp. 251–254.

[41] T. Cresson, D. Branchaud, R. Chav, B. Godbout, and J. A. de Guise, "3D shape reconstruction of bone from two X–ray images using 2D/3D non–rigid registration based on moving least–squares deformation," in *Medical Imaging 2010: Image Processing*, 2010.

[42] G. Zheng, "Statistical shape model–based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard AP X–ray radiograph," *Med. Phys.*, vol. 37, no. 4, pp. 1424–1439, Apr. 2010.

[43] G. Zheng, L. Nolte, and S. J. Ferguson, "Scaled, patient–specific 3D vertebral model reconstruction based on 2D lateral fluoroscopy," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 6, no. 3, pp. 351–366, May 2011.

[44] Z. Zhu and G. Li, "Construction of 3D human distal femoral surface models using a 3D statistical deformable model," *J. Biomech.*, vol. 44, no. 13, pp. 2362–2368, Sep. 2011.

[45] G. Zheng, "Personalized X–ray reconstruction of the proximal femur via intensity-based non–rigid 2D–3D registration," in *International Conference on Medical Image Computing and Computer–Assisted Intervention*. Springer, 2011, pp. 598–606.

[46] H. Boussaid, S. Kadoury, I. Kokkinos, J. Y. Lazennec, G. Zheng, and N. Paragios, "3D model–based reconstruction of the proximal femur from low–dose biplanar X–ray images," in *The 22nd British Machine Vision Conference–BMVC 2011*, 2011, pp. 1–10.

[47] T. Whitmarsh, L. Humbert, M. De Craene, L. M. D. R. Barquero, and A. F. Frangi, "Reconstructing the 3D shape and bone mineral density distribution of the proximal femur from dual–energy X–ray absorptiometry," *IEEE Trans Med Imag*, vol. 30, no. 12, pp. 2101–2114, Dec. 2011.

[48] L. Seoud, F. Cheriet, H. Labelle, and J. Dansereau, "A novel method for the 3D reconstruction of scoliotic ribs from frontal and lateral radiographs," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 5, pp. 1135–1146, May 2011.

[49] S. P. Väänänen, H. Isaksson, P. Julkunen, J. Sirola, H. Kröger, and J. S. Jurvelin, "Assessment of the 3-D shape and mechanics of the proximal femur using a shape template and a bone mineral density image," *Biomechanics and modeling in mechanobiology*, vol. 10, no. 4, pp. 529–538, 2011.

[50] S. P. Väänänen, J. S. Jurvelin, and H. Isaksson, "Estimation of 3D shape, internal density and mechanics of proximal femur by combining bone mineral density images with shape and density templates," *Biomechanics and modeling in mechanobiology*, vol. 11, no. 6, pp. 791–800, 2012.

[51] G. Zheng, "3D volumetric intensity reconstruction from 2D X–ray images using partial least squares regression," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE, 2013, pp. 1268–1271.

[52] T. Whitmarsh, L. Humbert, L. M. D. R. Barquero, S. Di Gregorio, and A. F. Frangi, "3D reconstruction of the lumbar vertebrae from anteroposterior and lateral dual–energy X–ray absorptiometry," *Medical image analysis*, vol. 17, no. 4, pp. 475–487, 2013.

[53] M. Ehlke, H. Ramm, H. Lamecker, H. C. Hege, and S. Zachow, "Fast generation of virtual X–ray images for reconstruction of 3D anatomy," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2673–2682, Dec. 2013.

[54] V. Karade, K. Vikas, and R. Bhallamudi, "3D femur model reconstruction from biplane X–ray images: a novel method based on Laplacian surface deformation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 4, pp. 473–485, 2014.

[55] S. P. Väänänen, L. Grassi, G. Flivik, J. S. Jurvelin, and H. Isaksson, "Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image," *Medical image analysis*, vol. 24, no. 1, pp. 125–134, 2015.

[56] B. Aubert, C. Vergari, B. Ilharreborde, A. Courvoisier, and W. Skalli, "3D reconstruction of rib cage geometry from biplanar radiographs using a statistical parametric model approach," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 4, no. 5, pp. 281–295, 2016.

[57] X. Kang, W.-P. Yau, and R. H. Taylor, "Simultaneous pose estimation and patient-specific model reconstruction from single image using maximum penalized likelihood estimation (MPLE)," *Pattern Recognition*, vol. 57, pp. 61–69, 2016.

[58] P. Cerveri, C. Sacco, G. Olgiati, A. Manzotti, and G. Baroni, "2D/3D reconstruction of the distal femur using statistical shape models addressing personalized surgical instruments in knee arthroplasty: A feasibility analysis," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, no. 4, 2017.

[59] G. Zheng and W. Yu, "Statistical shape and deformation models based 2D–3D reconstruction," in *Statistical Shape and Deformation Analysis*. Elsevier, 2017, pp. 329–349.

[60] M. Lüthi, T. Gerig, C. Jud, and T. Vetter, "Gaussian process morphable models," *IEEE Trans. Pattern Anal. Mach. Intell.*, in press.

[61] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans Med Imag*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013.

[62] L. Humbert *et al.*, "3D reconstruction of both shape and bone mineral density distribution of the femur from DXA images," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, 2010, pp. 456–459.

[63] N. Baka *et al.*, "2D–3D shape reconstruction of the distal femur from stereo X–ray imaging using statistical shape models," *Med. Image Anal.*, vol. 15, no. 6, pp. 840–850, 2011.

[64] S. Miao, Z. J. Wang, and R. Liao, "A CNN regression approach for real–time 2D/3D registration," *IEEE Trans Med Imag*, vol. 35, no. 5, pp. 1352–1363, Feb. 2016.

[65] R. A. McLaughlin, H. John, D. J. Hawkes, N. J.Alison, J. V. Byrne, and C. Tim, "A comparison of 2D–3D intensity–based registration and feature–based registration for neurointerventions," in *Lecture Notes in Computer Science*, 2002, pp. 517–524.

[66] P. K. Bhunre, W. K. Leow, and T. S. Howe, "Recovery of 3D pose of bones in single 2D X-ray images," in *Applications of Computer Vision, 2007. WACV'07. IEEE Workshop on*. IEEE, 2007, pp. 48–48.

[67] J. Von Berg *et al.*, "Temporal subtraction of chest radiographs compensating pose differences," in *Medical Imaging 2011: Image Processing*, vol. 7962. International Society for Optics and Photonics, 2011, p. 79620U.

[68] T. Neogi *et al.*, "Magnetic resonance imaging–based three–dimensional bone shape of the knee predicts onset of knee osteoarthritis: data from the osteoarthritis initiative," *Arthritis & Rheumatism*, vol. 65, no. 8, pp. 2048–2058, 2013.

[69] D. J. Haverkamp, D. Schiphof, S. M. Bierma-Zeinstra, H. Weinans, and J. H. Waarsing, "Variation in joint shape of osteoarthritic knees," *Arthritis & Rheumatology*, vol. 63, no. 11, pp. 3401–3407, 2011.

[70] R. Agricola, M. Reijman, S. Bierma-Zeinstra, J. Verhaar, H. Weinans, and J. Waarsing, "Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: a prospective cohort study," *Osteoarthritis and Cartilage*, vol. 21, no. 4, pp. 559–564, 2013.

[71] M. Roberts, E. Pacheco, R. Mohankumar, T. Cootes, and J. Adams, "Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi–automatic segmentation," *Osteoporosis international*, vol. 21, no. 12, pp. 2037–2046, 2010.

[72] J. S. Gregory and R. M. Aspden, "Femoral geometry as a risk factor for osteoporotic hip fracture in men and women," *Medical engineering & physics*, vol. 30, no. 10, pp. 1275–1286, 2008.

[73] S. M. Pizer *et al.*, "Deformable M–reps for 3D medical image segmentation," *International journal of computer vision*, vol. 55, no. 2, pp. 85–106, 2003.

[74] ——, "Nested sphere statistics of skeletal models," in *Innovations for Shape Analysis*, 2013, pp. 93–115.

[75] P. J. Besl and N. D. McKay, "Method for registration of 3D shapes," in *Robotics–DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[76] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[77] H. Chui, J. Rambo, J. Duncan, R. Schultz, and A. Rangarajan, "Registration of cortical anatomical structures via robust 3D point matching," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 1999, pp. 168–181.

[78] K. C. Lam, X. Gu, and L. M. Lui, "Landmark constrained genus–one surface Teichmüller map applied to surface registration in medical imaging," *Medical image analysis*, vol. 25, no. 1, pp. 45–55, 2015.

[79] F. Bernard *et al.*, "Shape-aware surface reconstruction from sparse 3d point-clouds," *Medical image analysis*, vol. 38, pp. 77–89, 2017.

[80] A. Lorusso, D. W. Eggert, and R. B. Fisher, *A comparison of four algorithms for estimating 3D rigid transformations*. University of Edinburgh, Department of Artificial Intelligence, 1995.

[81] M. B. Stegmann, "Active appearance models," Ph.D. dissertation, Technical University of Denmark, 2000. 125, 129, 2000.

[82] T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," *Medical image analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[83] M. A. Styner *et al.*, "Evaluation of 3D correspondence methods for model building," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 2003, pp. 63–75.

[84] T. Mutsvangwa, V. Burdin, C. Schwartz, and C. Roux, "An automated statistical shape model developmental pipeline: application to the human scapula and humerus," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1098–1107, Apr. 2015.

[85] R. Haq, J. Cates, D. A. Besachio, R. C. Borgie, and M. A. Audette, "Statistical shape model construction of lumbar vertebrae and intervertebral discs in segmentation for discectomy surgery simulation," in *International Workshop on Computational Methods and Clinical Applications for Spine Imaging*. Springer, 2015, pp. 85–96.

[86] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC medical imaging*, vol. 15, no. 1, p. 29, 2015.

[87] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[88] S. Bonaretti, C. Seiler, C. Boichon, M. Reyes, and P. Büchler, "Image–based vs. mesh–based statistical appearance models of the human femur: implications for finite element simulations," *Medical engineering & physics*, vol. 36, no. 12, pp. 1626–1635, 2014.

[89] L. Grassi, N. Hraiech, E. Schileo, M. Ansaloni, M. Rochette, and M. Viceconti, "Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur," *Medical engineering & physics*, vol. 33, no. 1, pp. 112–120, 2011.

[90] R. Bryan, P. S. Mohan, A. Hopkins, F. Galloway, M. Taylor, and P. B. Nair, "Statistical modelling of the whole human femur incorporating geometric and material properties," *Medical engineering & physics*, vol. 32, no. 1, pp. 57–65, 2010.

[91] G. Chintalapani, L. M. Ellingsen, O. Sadowsky, J. L. Prince, and R. H. Taylor, "Statistical atlases of bone anatomy: construction, iterative improvement and validation," *Med. Image Comput. Comput. Assist. Interv.*, vol. 10, no. Pt 1, pp. 499–506, 2007.

[92] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3D statistical deformation models using non–rigid registration," in *Medical Image Computing and Computer–Assisted Intervention–MICCAI 2001*. Springer, 2001, pp. 77–84.

[93] A. Mishra, P. Mondal, and S. Banerjee, "2D/3D non–rigid image registration by an efficient demons approach," in *2014 IEEE 27th International Symposium on Computer–Based Medical Systems*, May 2014.

[94] W. Yu, M. Tannast, and G. Zheng, "Non–rigid free–form 2D–3D registration using a B–spline–based statistical deformation model," *Pattern Recognition*, vol. 63, pp. 689–699, 2017.

[95] K. D. Fritscher, G. Agnes, and S. Rainer, "3D image segmentation using combined shape–intensity prior models," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 1, no. 6, pp. 341–350, 2007.

[96] J. Yao and R. H. Taylor, "Construction and simplification of bone density models," in *Medical Imaging 2001: Image Processing*, Jul. 2001.

[97] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient–based 2D/3D rigid registration of fluoroscopic X–ray to CT," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1395–1406, Nov. 2003.

[98] Z. Mu, "A fast DRR generation scheme for 3D–2D image registration based on the block projection method," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 169–177.

[99] X. Jia and S. B. Jiang, *Graphics Processing Unit–Based High Performance Computing in Radiation Therapy*. CRC Press, 2015.

[100] O. M. Dorgham, S. D. Laycock, and M. H. Fisher, "GPU accelerated generation of digitally reconstructed radiographs for 2D/3D image registration," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2594–2603, Sep. 2012.

[101] D. Staub and M. J. Murphy, "A digitally reconstructed radiograph algorithm calculated from first principles," *Medical physics*, vol. 40, no. 1, 2013.

[102] R. Westermann and B. Sevenich, "Accelerated volume ray–casting using texture mapping," in *Proceedings of the conference on Visualization'01*. IEEE Computer Society, 2001, pp. 271–278.

[103] I. Herrera, C. Buchart, I. Aguinaga, and D. Borro, "Study of a ray casting technique for the visualization of deformable volumes," *IEEE Trans Vis Comput Graphics*, vol. 20, no. 11, pp. 1555–1565, Nov. 2014.

[104] A. Macovski, *Medical imaging systems*. Prentice Hall, 1983.

[105] D. A. LaRose, "Iterative X–ray/CT registration using accelerated volume rendering," Ph.D. dissertation, Citeseer, 2001.

[106] T. S. Tang *et al.*, "Hardware–assisted 2D/3D intensity–based registration for assessing patellar tracking," in *International Conference on Medical Image Computing and Computer–Assisted Intervention*. Springer, 2004, pp. 1095–1096.

[107] J. Spoerk, H. Bergmann, F. Wanschitz, S. Dong, and W. Birkfellner, "Fast DRR splat rendering using common consumer graphics hardware," *Medical physics*, vol. 34, no. 11, pp. 4302–4308, 2007.

[108] X. Li, J. Yang, and Y. Zhu, "Digitally reconstructed radiograph generation by an adaptive Monte Carlo method," *Physics in medicine and biology*, vol. 51, no. 11, p. 2745, 2006.

[109] X. Jia, W. Wei, and K. Jia, "A GPU–based DRR generation method using cubic window," in *Intelligent Information Hiding and Multimedia Signal Processing (IIH–MSP), 2012 Eighth International Conference on*. IEEE, 2012, pp. 403–406.

[110] P. Lacroute and M. Levoy, "Fast volume rendering using a shear–warp factorization of the viewing transformation," in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. ACM, 1994, pp. 451–458.

[111] W. Birkfellner *et al.*, "Wobbled splatting – a fast perspective volume rendering method for simulation of X–ray images from CT," *Physics in medicine and biology*, vol. 50, no. 9, pp. N73–N84, 2005.

[112] T. Malzbender, "Fourier volume rendering," *ACM Transactions on Graphics (TOG)*, vol. 12, no. 3, pp. 233–250, 1993.

[113] F. Wang, T. Davis, and B. Vemuri, "Real–time DRR generation using cylindrical harmonics," *Medical Image Computing and Computer–Assisted Intervention (MICCAI)*, pp. 671–678, 2002.

[114] R. L. Siddon, "Fast calculation of the exact radiological path for a three–dimensional CT array," *Medical physics*, vol. 12, no. 2, pp. 252–255, 1985.

[115] F. Jacobs, E. Sundermann, B. De Sutter, M. Christiaens, and I. Lemahieu, "A fast algorithm to calculate the exact radiological path through a pixel or voxel space," *CIT. Journal of computing and information technology*, vol. 6, no. 1, pp. 89–94, 1998.

[116] D. B. Russakoff, T. Rohlfing, D. Rueckert, R. Shahidi, D. Kim, and C. R. Maurer, "Fast calculation of digitally reconstructed radiographs using light fields," in *Proceedings of SPIE*, vol. 5032, 2003, pp. 684–695.

[117] E. Ntasis, W. Cai, G. Sakas, and K. S. Nikita, "Real time digital reconstructed radiograph (DRR) rendering in frequency domain," in *BMES/EMBS Conference, 1999. Proceedings of the First Joint*, vol. 2. IEEE, 1999, p. 1041.

[118] L. Wang, X. Gao, and Q. Fang, "A novel mutual information–based similarity measure for 2D/3D registration in image guided intervention," in *Orange Technologies (ICOT), 2013 International Conference on*, 2013, pp. 135–138.

[119] D. Škerl, D. Tomaževič, B. Likar, and F. Pernuš, "Evaluation of similarity measures for reconstruction–based registration in image–guided radiotherapy and surgery," *International Journal of Radiation Oncology Biology Physics*, vol. 65, no. 3, pp. 943–953, 2006.

[120] A. Khamene, P. Bloch, W. Wein, M. Svatos, and F. Sauer, "Automatic registration of portal images and volumetric ct for patient positioning in radiation therapy," *Medical Image Analysis*, vol. 10, no. 1, pp. 96–112, 2006.

[121] G. Zheng, "Effective incorporating spatial information in a mutual information based 3D–2D registration of a CT volume to X–ray images," *Comput. Med. Imaging Graph.*, vol. 34, no. 7, pp. 553–562, 2010.

[122] D. Knaan and L. Joskowicz, "Effective intensity-based 2D/3D rigid registration between fluoroscopic X-ray and CT," in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2003, pp. 351–358.

[123] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2D–3D medical image registration," *IEEE Trans Med Imag*, vol. 17, no. 4, pp. 586–595, Aug. 1998.

[124] R. Munbodh, Z. Chen, D. A. Jaffray, D. J. Moseley, J. P. Knisely, and J. S. Duncan, "Automated 2D–3D registration of portal images and CT data using line–segment enhancement," *Medical physics*, vol. 35, no. 10, pp. 4352–4361, 2008.

[125] S. Dong, J. Kettenbach, I. Hinterleitner, H. Bergmann, and W. Birkfellner, "The Zernike expansion–an example of a merit function for 2D/3D registration based on orthogonal functions," in *International*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2876450, IEEE Reviews in Biomedical Engineering

Conference on Medical Image Computing and Computer–Assisted Intervention. Springer, 2008, pp. 964–971.

[126] F. Steinke and B. Schölkopf, "Kernels, regularization and differential equations," *Pattern Recognition*, vol. 41, no. 11, pp. 3271–3286, 2008.

[127] M. R. Mahfouz, W. A. Hoff, R. D. Komistek, and D. A. Dennis, "A robust method for registration of three–dimensional knee implant models to two–dimensional fluoroscopy images," *IEEE Trans Med Imag*, vol. 22, no. 12, pp. 1561–1574, Dec. 2003.

[128] J. Dey and S. Napel, "Targeted 2D/3D registration using ray normalization and a hybrid optimizer," *Medical physics*, vol. 33, no. 12, pp. 4730–4738, 2006.

[129] W. Sun, W. J. Niessen, M. van Stralen, and S. Klein, "Simultaneous multiresolution strategies for nonrigid image registration," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4905–4917, Dec. 2013.

[130] H. Lester and S. R. Arridge, "A survey of hierarchical non–linear medical image registration," *Pattern recognition*, vol. 32, no. 1, pp. 129–149, 1999.

[131] I. Matthews and S. Baker, "Active appearance models revisited," *International journal of computer vision*, vol. 60, no. 2, pp. 135–164, 2004.

[132] A. Varnavas, T. Carrell, and G. Penney, "Fully automated 2D–3D registration and verification," *Medical image analysis*, vol. 26, no. 1, pp. 108–119, 2015.

[133] T. Mutsvangwa, W. Wasswa, V. Burdin, B. Borotikar, and T. Douglas, "Interactive patient–specific 3D approximation of scapula bone shape from 2D X–ray images using landmark–constrained statistical shape model fitting," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 1816–1819.

[134] C. Chen and G. Zheng, "Fully automatic segmentation of AP pelvis X–rays via random forest regression with efficient feature selection and hierarchical sparse shape composition," *Computer Vision and Image Understanding*, vol. 126, pp. 1–10, 2014.

[135] M. Ehlke, T. Frenzel, H. Ramm, M. A. Shandiz, C. Anglin, and S. Zachow, "Towards robust measurement of pelvic parameters from AP radiographs using articulated 3D models," in *Computer Assisted Radiology and Surgery (CARS)*, 2015.

[136] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*. Cambridge university press, 1996, vol. 2.

[137] I. Rechenberg, *Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, 1970.

[138] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET–CT image registration in the chest using free–form deformations," *IEEE Trans Med Imag*, vol. 22, no. 1, pp. 120–128, Jan. 2003.

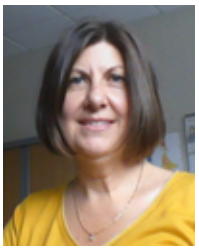
[139] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual–information–based registration of medical images: a survey," *IEEE Trans Med Imag*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[140] T. Whitmarsh *et al.*, "3D bone mineral density distribution and shape reconstruction of the proximal femur from a single simulated dxa image: an in vitro study," in *Medical Imaging 2010: Image Processing*, vol. 7623. International Society for Optics and Photonics, 2010, p. 76234U.

[141] W. H. Organization, *Prevention and management of osteoporosis: report of a WHO scientific group*. World Health Organization, 2003, no. 921.

Cornelius Reyneke (M'18) is a Ph.D. student in the Division of Biomedical Engineering at the University of Cape Town, South Africa. He received the B.Eng. degree in electrical/electronic engineering together with the B.Sc. degree in information technology from the University of Johannesburg, South Africa in 2013, and the M.Eng. degree in electrical/electronic engineering, specializing in computer vision, from the University of Johannesburg, South Africa in 2015. He is interested in the application of machine learning and computer vision methods to medical image analysis and computer–aided diagnosis.

Marcel Lüthi is a researcher at the University of Basel. Before completing his Ph.D. at the Computer Science from the University of Applied Science Bern, Switzerland and a Masters degree in Engineering Mathematics from Chalmers University of Technology, Sweden. His research interests are statistical shape modeling and its application to medical image analysis.



Valérie Burdin (M'98) received the M.Sc. degree in mathematics (1986) from the University of Saint–Etienne, Saint–Etienne, France and the Ph.D. degree in telecommunications and signal processing (1992) from the University of Rennes I, Rennes, France. She is with Mines Telecom Institute/IMT-Atlantique from 1997 as an Associate Professor, and then, from 2013 onwards as a Professor. She conducts research activities in the Laboratoire de Traitement de l'Information Mdicale (LaTIM INSERM UMR1101), in the team: "Imagine". Since 1990, she works closely with the Department of Orthopedics Surgery and, since 2005, with the Physical Medicine and Rehabilitation, Hospital of Brest. Her current research interests include motion analysis and 3D modeling of musculoskeletal systems with a focus between anatomical and kinematic data. This morpho-functional approach implies methodological aspects such as 3D data registration, segmentation, statistical shapes modeling and biomechanics. Valérie Burdin is author or co-author of more than 100 peer-reviewed publications and is a member of the IEEE Engineering in Medicine and Biology Society (EMBS), the French Society of Biomedical Engineering (SFGBM) and the French Society for Movement Analysis in Adults and Children (SOFAMEA).

Tania Douglas is a Professor in the Division of Biomedical Engineering at the University of Cape Town, where she also holds the DST/NRF Research Chair in Biomedical Engineering & Innovation. She completed degrees in electrical/electronic and biomedical engineering at the University of Cape Town, Vanderbilt University and the University of Strathclyde, and conducted postdoctoral research in image processing at the Japan Broadcasting Corporation. She also completed an MBA at the University of Cape Town. Her research interests include medical imaging and image analysis, the development of contextually appropriate health technologies, and heath innovation management. She is a fellow of the South African Academy of Engineering, a member of the Academy of Science of South Africa, and a Fellow of the International Academy for Medical and Biological Engineering.

Thomas Vetter studied mathematics and physics and received the Ph.D. degree in biophysics from the University of Ulm, Germany. As a Postdoctoral Researcher at the Center for Biological and Computational Learning, Massachusetts Institute of Technology, Cambridge, he started his research on computer vision. In 1993, he moved to the Max Planck Institute, Tübingen, Germany, and, in 1999, he became a Professor for computer graphics at the University of Freiburg, Germany. Since 2002, he has been a Professor of applied computer science at the University of Basel, Switzerland. His current research is on image understanding. He combines methods from machine learning, computer graphics and computer vision to implement analysis–by–synthesis systems for an automated image perception.

Tinashe Mutsvangwa (M'13) received the B.Sc. degree in electrical/electronic engineering and the M.Sc. and Ph.D. degrees in biomedical engineering all from the University of Cape Town, Cape Town, South Africa in 2003, 2005, and 2009, respectively. He is currently a Senior Lecturer in Biomedical Engineering and Health Innovation and Design, University of Cape Town. Previously, Dr Mutsvangwa received the Research-Based University Chair of Excellence Fellowship from 2012–2014. His research projects centre on the application of image and image analysis methods for computer–aided diagnosis. Topics span from 3D reconstruction of bone, X–ray imaging, statistical shape and appearance modelling, 3D morphometrics, diagnosis of TB via mobile phone imaging, computer–assisted Fetal Alcohol Syndrome (FAS) screening, facial characterisation of mental disorder patients to stereo–photogrammetry, 3D geometric morphometrics, machine learning, and pattern recognition methods as applied to analysis of the human form and the use of mixed-reality for aiding orthopaedic diagnosis.