



A coarse-to-fine IP-driven registration for pose estimation from single ultrasound image



Bo Zheng^{a,*}, Ryo Ishikawa^b, Jun Takamatsu^c, Takeshi Oishi^a, Katsushi Ikeuchi^a

^aThe University of Tokyo, Ikeuchi Lab, Meguro-ku, Tokyo, 153-8505, Japan

^bMedical Imaging Project, Canon Inc. Ohta-ku, Tokyo, 146-8501, Japan

^cNara Institute of Science and Technology (NAIST), Ikoma City, Nara 630-0192, Japan

ARTICLE INFO

Article history:

Received 16 March 2012

Accepted 1 January 2013

Available online 24 July 2013

Keywords:

Ultrasound (US) image

Pose estimation

3D image registration

Implicit polynomial

ABSTRACT

A fast registration making use of implicit polynomial (IP) models is helpful for the real-time pose estimation from single clinical free-hand Ultrasound (US) image, because it is superior in the areas such as robustness against image noise, fast registration without enquiring correspondences, and fast IP coefficient transformation. However it might lead to the lack of accuracy or failure registration.

In this paper, we present a novel registration method based on a coarse-to-fine IP representation. The approach starts from a high-speed and reliable registration with a coarse (of low degree) IP model and stops when the desired accuracy is achieved by a fine (of high degree) IP model. Over the previous IP-to-point based methods our contributions are: (i) keeping the efficiency without requiring pair-wised correspondences, (ii) enhancing the robustness, and (iii) improving the accuracy. The experimental result demonstrates the good performance of our registration method and its capabilities of overcoming the limitations of unconstrained freehand ultrasound data, resulting in fast, robust and accurate registration.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Motivation

To support medical diagnosis, various imaging modalities, such as computed tomography (CT) scan, MRI, PET, and ultrasound (US), are widely used in clinics. Among these modalities, US has beneficial characteristics such as free-hand manner, non-invasiveness, compactness, low cost, and synchronization of operations and imaging. Thus US is attractive for assistance with surgical operations and real-time diagnosis of problems with the circulatory system, abdomen, breast, prostate gland, etc.

However, US images are notorious for the poor image quality, due to speckle noises, low signal-to-noise ratio, occlusions, and uniform brightness. And field of view (FOV) in US imaging is very limited; in severe cases, only 2D cross-sectional images are obtained. These aspects confuse the doctors in making right decisions for diagnosis.

In order to solve these issues, some recent literature advocates the fusion-of-modality techniques. For example, before the surgical operation, 3D models of target parts are obtained by rich but time-consuming modalities such as CT, MRI, and PET. By superimposing US images obtained during the operation on these 3D

models, the result will provide rich information to help a doctor's diagnosis. To achieve this, the key for superimposing is to estimate the pose of US images related to the images derived from other modalities.

1.2. Related work

The pose estimation can be viewed as a registration problem for two models: a source model (preoperative 3D model) and a target model (2D/3D US image). To do this, a class of methods such as [1,2] bind the optical position sensors to a US probe, and measure the relative US position to 3D models; For enhancing robustness, the methods in [3–5] combine the information from position sensors and image features.

Without position sensors, Penney et al. [6] propose to register the surface points manually selected from US images to a preoperative 3D shape model by MRI segmentation; similarly, Amin et al. [7] register the bone boundaries in US images to a shape model segmented from CT image by a modified ICP method; Lange et al. [8] take advantage of 3D-Power Doppler to extract vessel shapes from intraoperative 3D US and register with preoperative models in liver surgery; to enhance the robustness, mutual information is advocated as measuring image similarities, such as [9,10]; Wein et al. [11] achieve the CT-ultrasound registration by simulating the US image sequence from CT image, and using a new similarity metric: linear correlation of linear combination; other methods such as

* Corresponding author.

E-mail address: zheng@cvi.iis.u-tokyo.ac.jp (B. Zheng).

[12,13,10] estimate the relative positions according to the image features or intensity and gradient information of US images and preoperative 3D models. Although each of the methods has its effectiveness, they suffer from expensive computation caused either by the intensity-based similarity calculation or point-to-point ICP-based registration and thus they are difficult to work in real time.

Regardless of the data type, the registration problem is solved basically by three families of methods: (i) ICP-based methods: the iterative closest point method (ICP) first proposed by Besl and McKay [14] or its accelerated variations such as [15] for 3D range data, (ii) point-model methods: e.g., Fitzgibbon [16] encodes the Euclidean distance field by fast distance transformation and employs the robust estimation to remove the outliers; Huang et al. [17] proposed new similarity measurement using information theory to achieve the robust non-rigid registration; and (iii) the approach relying on algebraic/geometric invariant features, e.g., moment features is described in [18], and IP global features are proposed by Taubin et al. [19,20]. The first family of methods can achieve fine registration, but requires time-consuming computation of point-to-point/surface correspondences; The second family of methods can achieve the registration efficiently but needs huge memory, especially in dense 3D cases, to preserve the distance field; and the third family of methods can achieve fast registration, but cannot deal with registration in the case of partially overlapping the target objects [18].

In our previous work [21,22], we propose to approximate the Euclidean distance with the algebraic formulation using implicit polynomials (IPs) and speed up the registration. The advantages of this method over the prior methods are that: i) unlike the ICP-based methods, it avoids the extra computation for point-wise correspondences; ii) unlike the point-model method of preserving a discrete distance field, it needs very little memory space for preserving a few IP coefficients, and the algebraic model can generate an infinite distance field to support registration in a wider space; iii) unlike the coarse registration methods, it supports partial-overlapped registration. A recent work proposed by Rouhani and Sappa [23] improves the optimization by Levenberg–Marquardt algorithm which leads to a faster convergence. These methods adopt a single IP based registration which remains an essential issue: a moderate IP model is really difficult to generate and thus to be obstacle to an accurate and robust registration.

The previous studies in [24,25] pointed out two issues frustrating the IP fitting: i) An IP of low degree loses local accuracy for object representation, whereas ii) an IP of high degree might be globally unstable (the undesired surfaces appear in the fitting result). However the former may lead to the lack of accuracy for registration, but the latter may lead to a failure registration. Fig. 1 shows an example when an IP (gray surface) is registered to scattered points (blue points). While Fig. 1 (a) shows the registration result losing much accuracy due to coarse IP of low degree,

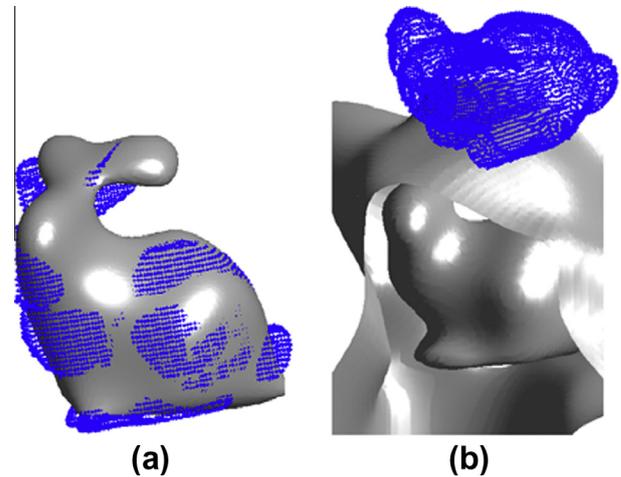


Fig. 1. Registration example (IP is shown in gray surface and target model is shown in blue points). (a) Inaccurate result caused by low degree IP with lack of detailed representation and (b) failure registration caused by high degree IP with undesired surfaces. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 1 (b) shows failure registration due to the global instability problem of IP fitting with high degree.

1.3. Overview and contributions

Our method inherits IP's merits: neither time-consuming process of correspondence searching nor huge memory for storing the discrete distance field is required. In addition, over the previous methods in [21–23] that use single IP for registration, we propose a coarse-to-fine IP registration. As illustrated in Fig. 2 leftmost, it starts from a low degree IP (ellipsoid) to achieve a robust initial guess. Second, after a rough registration covered by the IPs of low degree, the higher degree IPs can drive to a more accurate position, even if the IP is not stably modeled (extra zero sets appear around the desired zero set), as shown in Fig. 2 rightmost. Our method improves the robustness and accuracy. The robustness is guaranteed by the coarse estimation with an IP of low degree, whereas the high accuracy can be achieved by an IP of high degree given the appropriate initial guess.

Compared to the global IP matching methods, such as [19,20], our method overcomes the partially overlapped problem. Such merits make it possible to be applied for the registration between a 3D shape and a 2D US image plane. We adopt boundary information which is independent of the types of modalities. As illustrated in Fig. 3 (a), our method supposes the 3D model has been obtained in an advance which is desired to registered with the online ultrasound image shown in Fig. 3 (b). In an online process, e.g., dur-

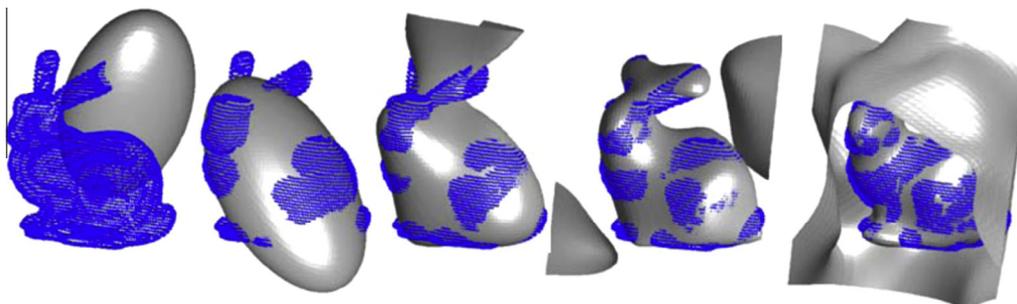


Fig. 2. Coarse-to-fine registration in general case driven by IP models from low degree to high degree (from left to right). Leftmost: the initial position for registration. The initial position for each step is determined by the result of previous step.

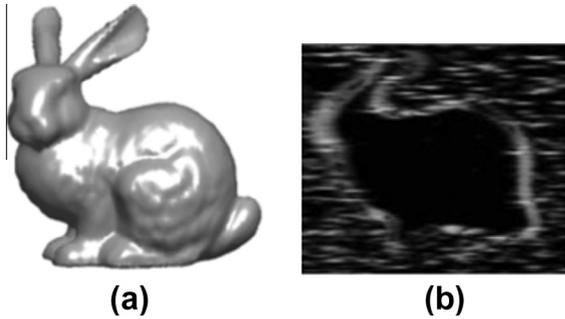


Fig. 3. (a) Offline obtained 3D prior model and (b) online obtained ultrasound image.

ing a surgical operation, it fast aligns a 3D IP models to a 2D US image, as the coarse-to-fine approach shown in Fig. 4. Then the desired relative pose information between 3D model and the 2D US image (associated with the probe position information of ultrasound device) is obtained.

This paper is organized as follows: Section 2 introduces the mathematics of IP modeling and its properties; In Section 3, we present the registration technique using IP, formulated in both a general case and a US image case, and based on coarse-to-fine approach; Section 4 reports experimental results followed by conclusion in Sections 5. In addition, we present our symbolic computational transformation of IP in Appendix A.

2. Implicit polynomial

We adopt implicit polynomials for modeling the preoperative 3D images captured by a modality such as MRI, CT-SCAN, or 3D US, supposing that the 3D boundary has been obtained from the segmentation result of the captured volume images.

2.1. Model

Suppose model \mathcal{M} obtained from 3D boundary data represents a closed surface, and by contracting and expanding it defines two 3D layers (see [26]): the inner layer Γ_- and outer layer Γ_+ . Then the model is represented by its distance transform:

$$\Phi(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{M} \\ +d(\mathbf{x}, \mathcal{M}), & \mathbf{x} \in \Gamma_+, \\ -d(\mathbf{x}, \mathcal{M}), & \mathbf{x} \in \Gamma_- \end{cases} \quad (1)$$

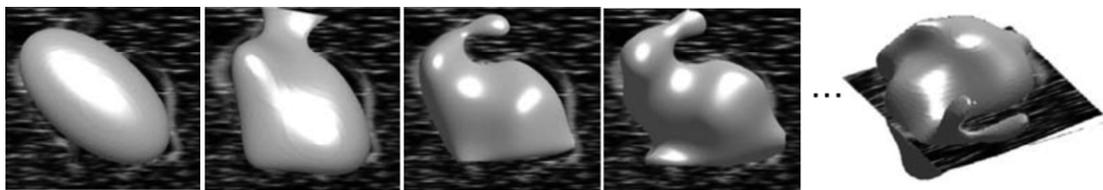


Fig. 4. Online coarse-to-fine registration in US image case driven by IP models from low degree to high degree.

where $\mathbf{x} = (xyz)$ is the location of one data point in Cartesian coordinates, and $d(\mathbf{x}, \mathcal{M}) = \min_{\mathbf{x}_M \in \mathcal{M}} \|\mathbf{x} - \mathbf{x}_M\|$ is the Euclidean distance from \mathbf{x} to the model consisting of the points \mathbf{x}_M . We approximate Φ with an implicit polynomial (IP) as:

$$\begin{aligned} \Phi \approx f_n(\mathbf{x}) &= \sum_{0 \leq i,j,k; i+j+k \leq n} a_{ijk} x^i y^j z^k \\ &= \left(\underbrace{1 \ x \ \dots \ z^n}_{\mathbf{m}(\mathbf{x})^T} \right) \left(\underbrace{a_{000} \ a_{100} \ \dots \ a_{00n}}_{\mathbf{a}} \right)^T, \end{aligned} \quad (2)$$

where n is the degree of an IP, and the IP can be represented as an inner product between the monomial vector and the coefficient vector as: $\mathbf{m}^T \mathbf{a}$. The indices $\{i,j,k\}$ are sorted by the inverse *lexicographical order* (see Table 1 in [27]). This problem can be viewed as the linear least-squares problem solved by 3L method [26].

2.2. Properties

IP representation has several good properties such as robust against noise, easy calculation for partial derivatives and easy transformation, but also has bad properties such as global unstable fitting. These properties are described in the followings:

2.2.1. Robust against noise

The previous studies, such as [24,28], pointed that IP fitting is robust against noise, missing data and occlusions, and in [25] there shows an example that IP representation does not vary largely when adding noise to the data set to some extent. Note this property is best matched to the modeling for medical images, since many images captured by the modalities, such as ultrasound device, are often with poor image quality due to speckle noises, low signal-to-noise ratio and occlusions.

2.2.2. Easy calculation for partial derivatives

Given an IP, its partial derivatives can be simply calculated by a symbolical transformation. For example, if an IP of n -degree is associated with coefficient vector $\mathbf{a} = \{a_{ijk}\}$, then its first partial derivative of y , $\frac{\partial f_n}{\partial y}$, is an IP of $(n-1)$ -degree and can be calculated as $\mathbf{a}' = \{a'_{ijk}\}$, where

$$a'_{i,j-1,k} = \begin{cases} j a_{i,j,k}, & j \geq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Fig. 5(a) illustrates the gradients $\nabla f_n = \left(\frac{\partial f_n}{\partial x}, \frac{\partial f_n}{\partial y}, \frac{\partial f_n}{\partial z} \right)^T$, where each entry is the first order derivatives.

Table 1

Comparison of registration results. (ICP: iterative closest point [14], SI low: single IP method with low degree [21], SI high: single IP method with high degree [21], CtF: coarse to fine method (proposed method), Acc.: accuracy measured with mean squared errors and CT: CPU time (s).)

Case	ICP		SI low		SI high		CtF	
	Acc.	CT (s)	Acc.	CT (s)	Acc.	CT (s)	Acc.	CT (s)
Fig. 7(a)	5.6e−3	1.1	9.8e−3	0.04	7.8e−3	0.08	7.8e−3	0.12
Fig. 7(b)	3.5e−3	10.7	1.2e−2	0.13	N/A	N/A	4.8e−3	0.19
Fig. 7(c)	3.2e−3	9.3	2.3e−2	0.12	N/A	N/A	4.2e−3	0.18

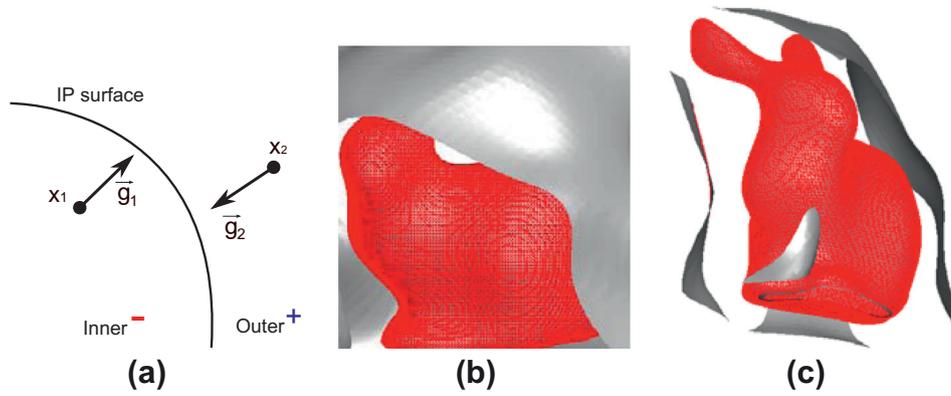


Fig. 5. Properties: (a) IP gradients of inner point x_1 and outer point x_2 , (b) IP fit of degree 8 (global unstable with the extra IP surfaces in gray and the desired surface in red), and (c) different view of (b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2.3. Easy transformation

An IP model $f(x) = \mathbf{m}(x)^T \mathbf{a}$ can be easily transformed by rotation or/and translation, as it is a matrix–vector multiplication operation on the coefficients as:

$$\mathbf{a}' = V(\mathbf{p})\mathbf{a}, \tag{4}$$

where V is a function mapping 3D transformation parameters $\mathbf{p} = (\mathbf{t}, R) \in \mathbb{R}^3 \times SO(3)$ to a square matrix with same dimension to coefficient vector \mathbf{a} . To our knowledge, although prior literature, such as [20], have proved the existence of matrix V , the computational implementation is not explicitly described. Tarel et al. [19] proposed a tensor-based transformation for IP, but it still suffers from heavy computational cost. Therefore, let us present our symbolic computational in **A** based on Taubin and Cooper’s theory [20].

2.2.4. Difficult control for global stability

No matter how robust against noise, IP still suffers from the representation problem on global instability [24,25], when fitting with high degree IP to the object. The example shown in Fig. 5(b) and (c) implies that high degree IP is locally accurate but globally unstable, which might lead to a failure registration.

3. Registration

In this section, before we present our coarse-to-fine registration method, let us first consider the case driven by single IP model. The objective of IP-driven registration is to find a transformation that makes the IP zero set to be “best” matched with the given data set. It can be formulated as an energy minimization problem in a

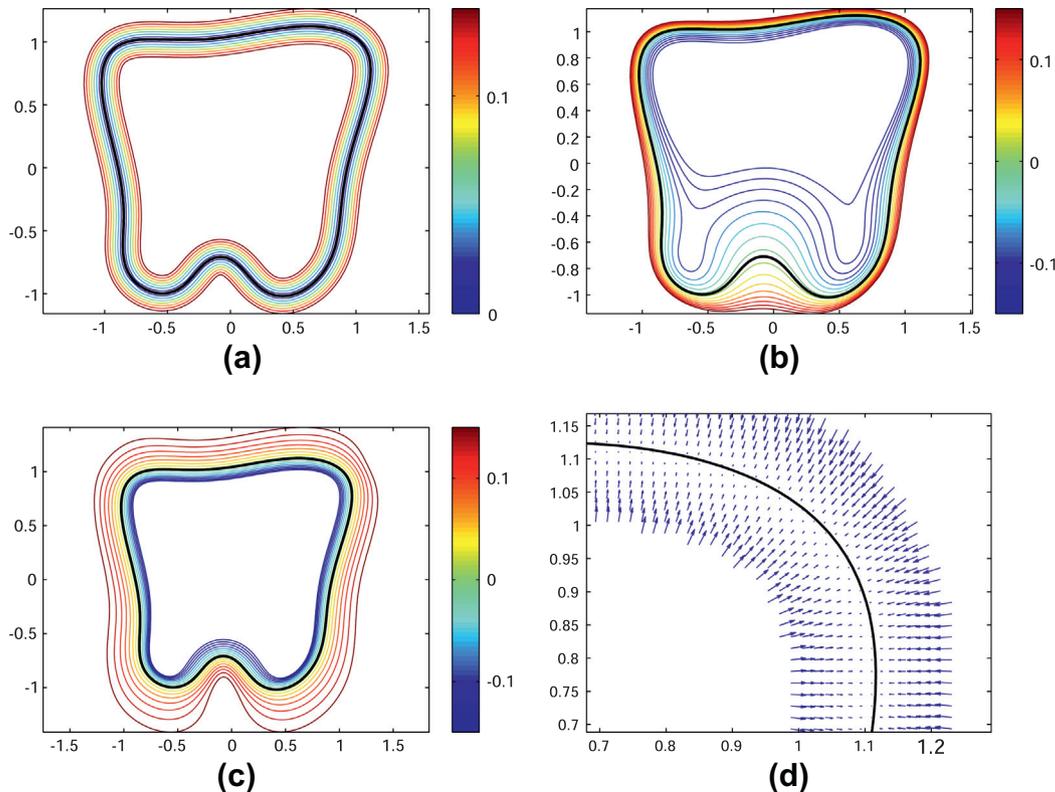


Fig. 6. 2D contours for illustrating distance fields: (a) Euclidean distance calculated by closest points searching. (b) Algebraic distance f in Eq. (2). (c) Algebraic distance calculated by $\frac{f}{|\nabla f|}$ in Eq. (5). (d) Gradient field calculated by Eq. (6).

general case or US image case as introduced in Section 3.2. To this end, the moderate measurement for the distance between data set and IP is required first.

3.1. Distance metric

We define function $d(x_i, f_n)$ for measuring the algebraic distance between a discrete point x_i and an IP model f_n as:

$$d(x, f_n) = \frac{f_n(x)}{\|\nabla f_n(x)\|}, \forall x \in \Omega, \quad (5)$$

The magnitude of this algebraic distance d has been proved to close Euclidean distance in [29]), but note that d is a kind of signed distance since if x in the inner region of IP, then $d < 0$. Fig. 6(a)–(c) illustrate the difference of three distance fields generated by Euclidean distance, Eqs. (2) and (5).

We define the gradient vector with length of distance $G: \mathcal{R}^3 \rightarrow \mathcal{R}^3$ as

$$G(x) = d(x, f_n) \frac{\nabla f_n(x)}{\|\nabla f_n(x)\|}, \forall x \in \Omega. \quad (6)$$

This can be viewed as a resized version of the original gradient vectors with distance $d(x)$. Fig. 6(d) illustrates the gradient field for example.

3.2. Energy functions for general case and US case

Given the definition of a distance between a certain data set and an IP, the next problem is how to predict the transformation for the IP that can minimize distance to the data set.

In our method, the minimization is done through the following two steps: for accelerating the convergence, first, it minimizes the function without any constraint in the transformation. This means every point can move freely during the first minimization. Next, it determines the transformation parameters to better describe the first minimization result. These two steps are repeated until convergence.

Let us describe the registration between an IP and a partial object, in a general case and a US image case.

3.2.1. General case

In general case, registration can be simply formulated to minimize the energy functional of the integration of distances as follows:

$$E = \int_{\Omega} d^2(x, f_n(x)) d\Omega, \quad (7)$$

where Ω represents the partial surface of an object, and $x \in \Omega$. Then, by calculus of variations [30], the Gateaux derivative (first variation) of the functional E to point x can be approximately formulated as

$$\frac{\partial E}{\partial x} \approx 2G(x), \quad (8)$$

if we consider ∇f_n as a constant value for computational convenience.

Therefore, we need to minimize this functional to satisfy the Euler–Lagrange equation $\frac{\partial E}{\partial x} = 0$. Thus the steepest descent process is executed in the following gradient flow for each point x :

$$\frac{\partial x}{\partial t} = -2G(x). \quad (9)$$

3.2.2. US image case

For the task of pose estimation of the US probe, the US images from the 2D probe that image the organ slice boundaries are usually expected to be matched with the corresponding planar intersection of the 3D model.

Furthermore, US images are heavily noise-contaminated. Since precise extraction of the organ’s boundary is therefore very difficult, maybe impossible, simple boundary-based registration described in the general case is not sufficient for US images. But fortunately, the organ’s inside region is clearly seen in US images as shown in Fig. 3(b). These two points are key for improving the general method for US images.

For solving the first point, we use a Gaussian smoothed edge indicator defined by [31]. Let $I(x)$ be a 2D US image function in 3D space, and pixel position $x(\in \mathcal{R}^3)$ be in the 3D plane Ω . Let $g(x)$ be the indicator defined as follows:

$$g = \frac{1}{(1 + |\nabla G_{\sigma} * I|/k)^2}, \quad (10)$$

where G_{σ} is a Gaussian filter with standard deviation σ , and $*$ denotes convolution.

For making use of the second point, we compose the energy function with combining a boundary constraint $\mathcal{L}(x)$ and an inside constraint $\mathcal{A}(x)$, such as

$$\mathcal{E}(x) = \alpha \mathcal{L}(x) - \beta \mathcal{A}(x), \quad (11)$$

where α and β are constants, and the terms $\mathcal{L}(x)$ and $\mathcal{A}(x)$ are defined by

$$\mathcal{L}(x) = \int_{\Omega} \delta(d(x, f_n) \cdot g(x)) d^2(x, f_n) d\Omega, \quad (12)$$

and

$$\mathcal{A}(x) = \int_{\Omega} H(-d(x, f_n) \cdot g(x)) d^2(x, f_n) d\Omega, \quad (13)$$

respectively; Ω corresponds to the image plane. δ is the univariate Dirac function defined as:

$$\delta(x) = \begin{cases} 0, & |x| > \kappa \\ \frac{1}{2\kappa} [1 + \cos(\frac{\pi x}{\kappa})], & |x| \leq \kappa \end{cases} \quad (14)$$

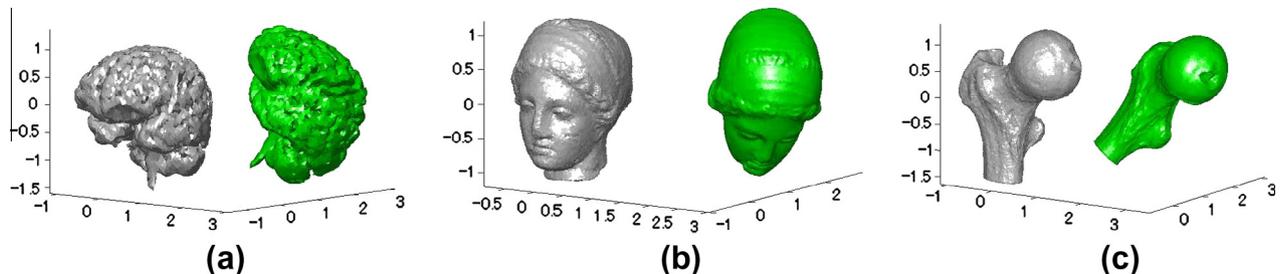


Fig. 7. Initial position for registration. Target models are shown in green. (a) “brain” object with 11,162 vertices, (b) “statue” object with 33,587 vertices, and (c) “bone” object with 34,267 vertices.

where κ is constant and practically set with $\kappa = 2$ in our experiments. H is the Heaviside function defined as:

$$H(x) = \begin{cases} 0, & x < 0 \\ x + x^2/2, & x \geq 0 \end{cases} \quad (15)$$

Let us explain the meaning of these terms. Minimize the energy functional \mathcal{E} in (11), which is equivalent to minimizing \mathcal{L} in (12), and maximize \mathcal{A} in (13). First, the energy function with respect to the boundary \mathcal{L} is equivalent to the integral of edge indicator values along the intersected curve between IP and US image plane because the IP model should be converged around an edge-like part in US images. Second, the energy function with respect to the inside region \mathcal{A} is the integral of edge indicator values over

the inner region surrounded by the intersected curve. It is expected to be as large as possible. To summarize, minimizing \mathcal{L} and $-\mathcal{A}$ achieves the registration.

By calculus of variations, the Gateaux derivative (first variation) of the functional \mathcal{E} in (11) can be approximately written as

$$\frac{\partial \mathcal{E}}{\partial \mathbf{x}} \approx (\alpha \delta(d \cdot \mathbf{g}) - \beta H(-d \cdot \mathbf{g}))G. \quad (16)$$

The steepest descent process for minimization of the functional \mathcal{E} is the following gradient flow:

$$\frac{\partial \mathbf{x}}{\partial t} = -(\alpha \delta(d \cdot \mathbf{g}) + \beta H(-d \cdot \mathbf{g}))G. \quad (17)$$

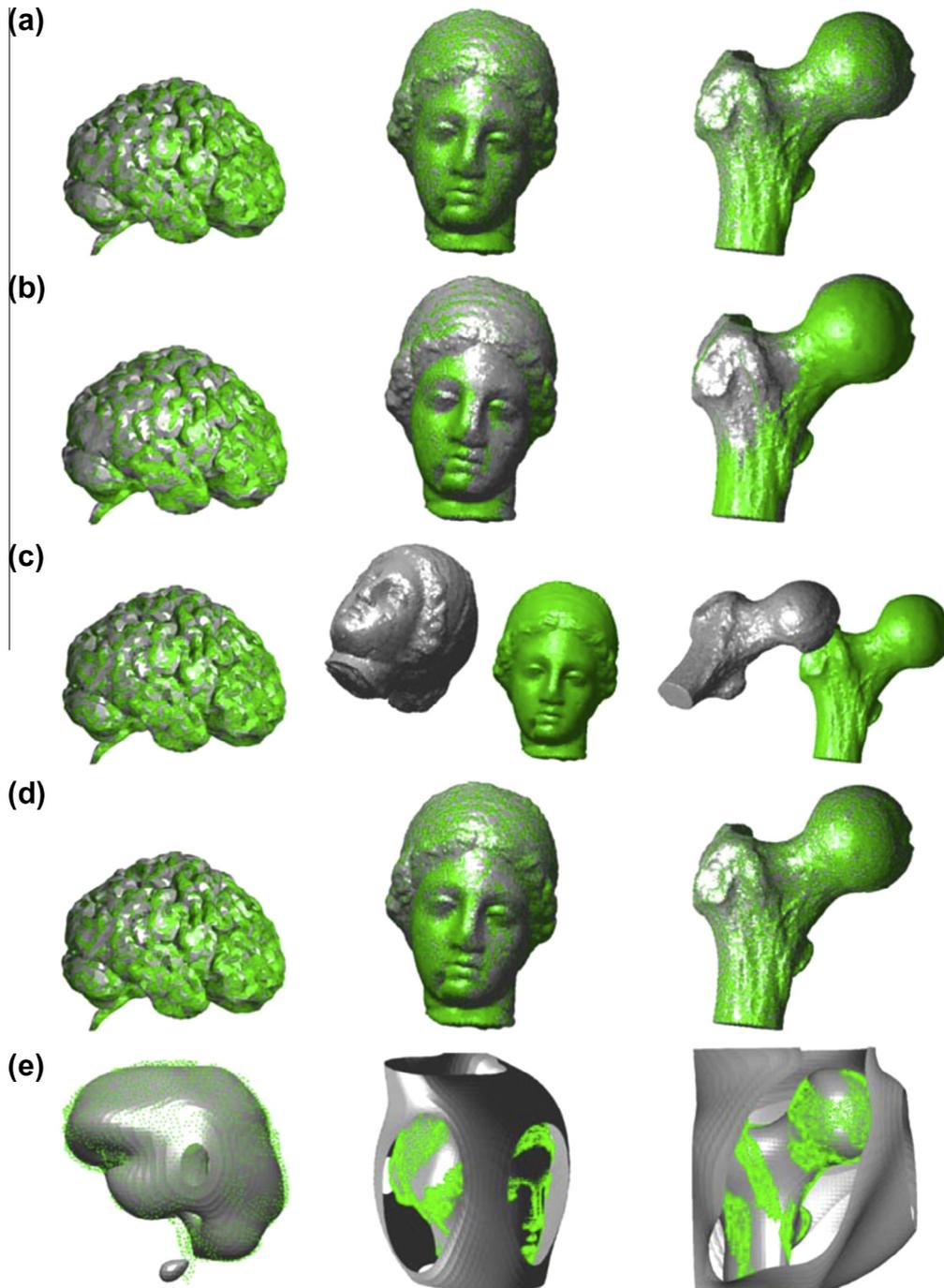


Fig. 8. Registration result. (a) ICP [14], (b) registration with single IP of 4° [21,23], (c) registration with single IP of 8° [21,23], (d) proposed coarse-to-fine method, and (e) IP-to-points representation corresponding to (d).

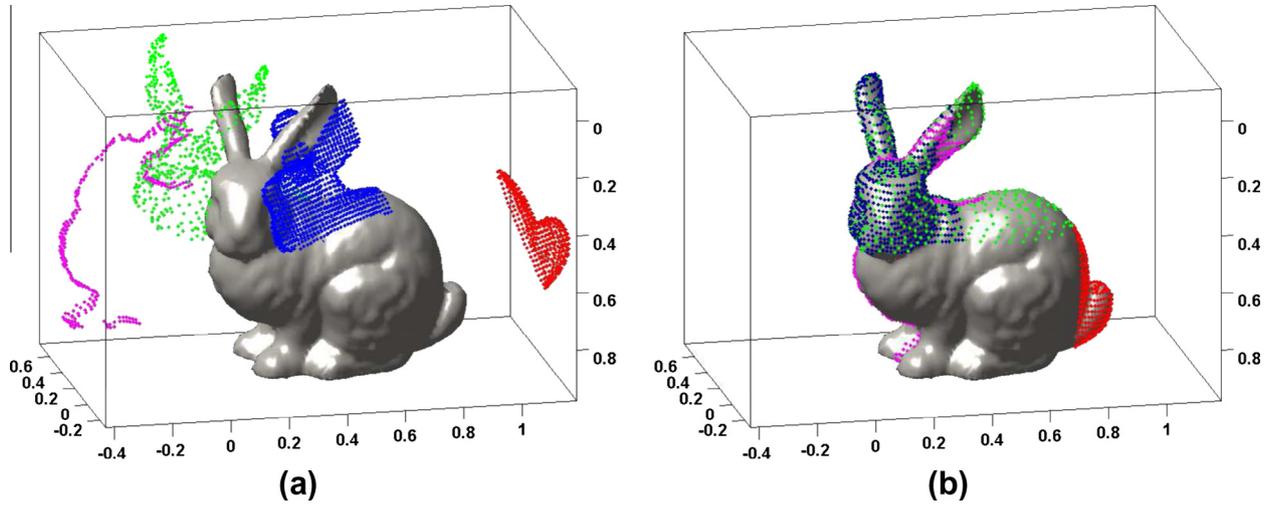


Fig. 9. Registration for partially overlapped data. Four parts of bunny: “head”, “tail”, “sparse head” and “3D-plane curve” are shown in blue, red, green and pink respectively. (a) Initial position of each part for the registration. (b) Result of our method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3. Operation for minimization at each step

The minimization can be done through the following two steps: for accelerating the convergence, first, it minimizes the function without any constraint in the transformation. This means every point can move freely using Eqs. (9) and (17). Next, it determines the transformation parameters to better describe the first minimization result. These two steps are repeated until convergence.

Let $X \in \mathcal{R}^{N \times 3}$ preserve N 3D data points. Then the approximation of (9) and (17) by the above difference scheme can be simply written as:

$$\frac{\partial X}{\partial t} = X^{k+1} - X^k \quad (18)$$

It is the approximation of discrete data set transformation by the above spatial difference scheme in (9) and (17).

$$A = (X^k - \bar{X}^k)^T (X^{k+1} - \bar{X}^{k+1}), \quad (19)$$

where \bar{X} is a matrix in which each row consists of the mean value (center point) of X , and if A is decomposed with singular value decomposition (SVD) algorithm as $A = USV^T$, then the transformation is given as:

$$R = UV^T, \quad \mathbf{t} = \bar{X} - \bar{X}'R^T \quad (20)$$

where R and \mathbf{t} are rotation and translation parameters respectively.

Now, we have two choices of operation at each step that makes the iteration convergent:

Point-to-IP Operation: Each point x moves toward IP, by updating its position as:

$$x^{k+1} = Rx^k + \mathbf{t}; \quad (21)$$

IP-to-point Operation: IP moves toward the point set by updating the coefficient vector as:

$$\mathbf{a}^{k+1} = V(R, \mathbf{t})\mathbf{a}^k, \quad (22)$$

by the algorithm proposed in A.

Both the two operations are mathematically equivalent, but however they can be selected in different cases for accelerating the computation. When the point set is small, e.g., less than 6800 points, point-to-IP operation is faster than the IP-to-point operation of an 8-degree IP. But on the other hand, e.g., in the US case, since the point set extracted from pixels is often huge (a 320×240 image can own 76,800 points), then the IP-to-point operation is definitely required for the acceleration.

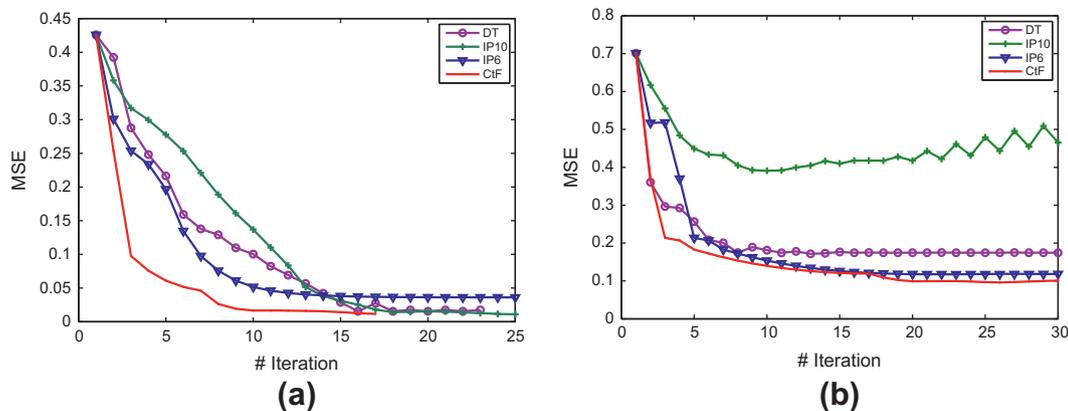


Fig. 10. Comparable result based on graph of MSE vs. iteration number. Four methods are employed: (1) DT: Distance transformation method in [16], (2) IP10: Single IP registration [21,22] using 10-degree IP, (3) IP6: Single IP registration [21,22] using 6-degree IP and (4) CtF: proposed Coarse-to-fine method. (a) And (b) correspond to the data sets “sparse head” part in green and the “3D-plane curve” in pink respectively, see Fig. 9. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

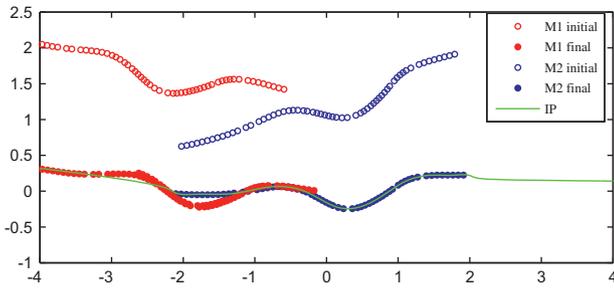


Fig. 11. Open curve registration: two models, M1 and M2 shown in red and blue respectively, register to an IP fit curve shown in green line. Initial positions are shown circle points for each model and the final positions are shown in dot points. IP fit (in green line) is resulted by fitting a source model which is close to the blue dot points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.4. Coarse-to-fine registration

Given the registration approaches both for general and US cases derived by single IP, let us present our coarse-to-fine registration with coupled two sub-processes: (1) coarse-to-fine representation and (2) coarse-to-fine transformation.

The coarse-to-fine representation is based on our previous method, incremental fitting [25]. This approach can fit a complex model in an incremental manner, for achieving the coarse-to-fine representation. It works efficiently because the upper-triangular linear system is solved incrementally until the desired fitting accuracy can be satisfied.

However this incremental method cannot be directly used for registration, because at each step it cannot guarantee to obtain a complete degree form of IP which is necessary for calculating the transformation matrix (See A). To make the incremental scheme suitable for registration, we modify the incremental scheme to be a degree-by-degree incremental manner; that is, it increase with an integral degree at each step.

Then our registration can be conducted through a coarse-to-fine transformation. As shown in Fig. 2, transformation is carried out at each step after the degree is increased. Each transformation starts with the position obtained from the result of lower IP at previous step, and stops at the position when the registration residual is small enough. The approach can be described by the following algorithm.

Algorithm 1. Algorithm: Coarse-to-fine registration

- (1) Input: given source model S and target model T ,
Output: transformation parameter \mathbf{p}
Initialization: set \mathbf{p}_i with null rotation and translation,
- (2) Fitting S with the i th degree IP f_i ;
- (3) Updating coefficient as: $\mathbf{a}_i \leftarrow V(\mathbf{p}_i)\mathbf{a}_i$;
- (4) Estimating \mathbf{p}_i by the registration between f_i and T ;
- (5) If the desired fitting accuracy can be satisfied, then stop;
otherwise, increasing degree as: $i \leftarrow i + 1$ and goto (2)

4. Experimental results

In this section, we report results of experiments dealing with some synthetic data sets to evaluate the method on computational performance. All the experiments were implemented in Matlab 8 combined by C++ code with a PC having an Intel core 2 CPU, 2.4 GHz, and 2 GB memory.

4.1. Registration for general case

4.1.1. Globally overlapped registration

In the experiments of registration for general case, we first test the robustness and accuracy for global overlapped registration using the 3D shape models shown in Fig. 7, where the source models are shown in gray and the target models are shown in green; we added Gaussian noise to the source model with standard deviations 0.01; and we initially rotated the target model by 30° to each axis, and translated by unit 2 along the X and Y-axis. Note, All the original data sets are regularized by centering the data-set center of mass at the origin of the coordinate system and scaling it by dividing each point by the average length from point to origin.

Then we tested three methods for comparison: (i) standard ICP method [14], (ii) correspondences free methods with single IP [21,23] and (iii) proposed method. Fig. 8 shows the registration results for each method.

Through the results in Fig. 8, we can see that, (i) all the results for “brain” object look acceptable, since the IP fits are stable even to the high degree, the registration with single IP works well and (ii) however, for the other two objects, single IP methods perform worse as losing accuracy with low degree IP, and failure registration with high degree IP due to the fits of high degree is unstable.

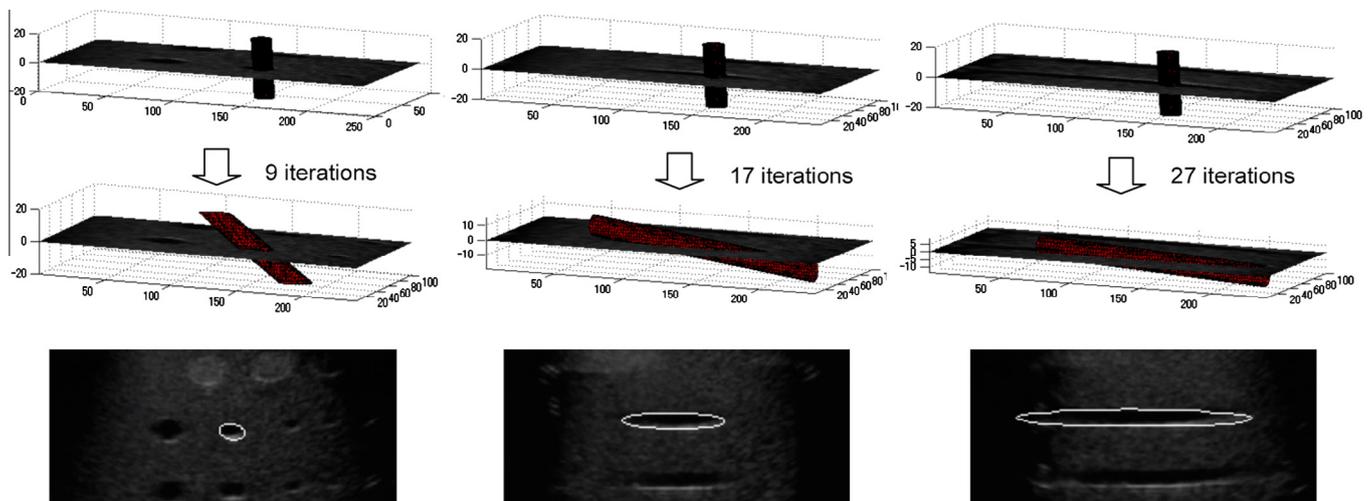


Fig. 12. Pose estimation for phantom ATSS14. First row: initial position; Second row: final position and Third row: cross-section contour of IP and US image (white points).

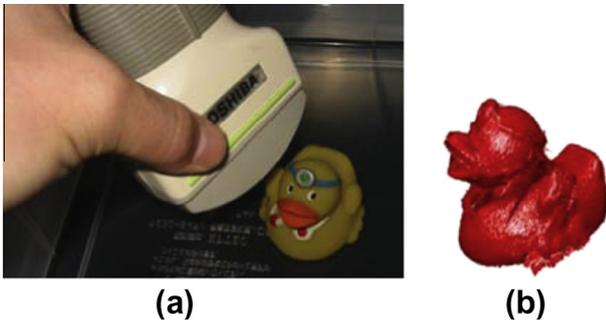


Fig. 13. (a) Measuring duck toy. (b) 3D duck model obtained by range sensor.

Table 1 shows the comparison of registration result. We can see that although the ICP method always keep highest accuracy, but it is time-consuming due to the heavy computation for correspondence searching, while the correspondences free registration with single IP shows the low accuracy and instability in the second and third cases.

4.1.2. Partially overlapped registration

Second, we synthesized a set of data for testing the performance on partially overlaps. As shown in Fig. 9, the following four parts of the bunny object in different colors are registered to the whole bunny model: the “head” in blue and “tail” in red are cut from the original point cloud; the “sparse head” in green is generated by down-sampling 10% of original points and adding Gaussian noise with standard deviation of 0.1; and “3D-plane curve” part in pink is selected from the points on the cross-section between a 3D plane and the bunny object, and added with Gaussian noise in standard deviation of 0.1. Fig. 9(a) shows the initial position for registration for each part and (b) shows the final result obtained by our method.

We show a comparable result based on graph of MSE vs. iteration number in Fig. 10, where two noisy data set of “sparse head”

and “3D-plane curve” are adopted for the registration in Fig. 10(a) and (b) respectively. The result implies that (i) single IP registration [21,22] of 6-degree is stable but lose accuracy while the registration of 10-degree might lead to slow or failed convergence; (ii) in contrast, our method is often with fast convergence when using the IP of low degree and with high accuracy when using the IP of high degree; (iii) compare to the distance transformation method [16] which requires $100 \times 100 \times 100$ grid in memory for storing the distance field, IP only needs very small memory for storing several coefficients.

A discussion on partially overlapped registration. However, we do not suggest IP should be applied for the registration of open curves or surfaces (e.g., the range images captured from different views) due to the reason: the IP fit of an open curve/surface might infinite extend in the space if there is no constraint to make it closed, which may lead the registration to depend on the initial position guess very much. To illustrate this, a simple example is shown in Fig. 11 where an IP fit (shown in green curve) is priorly obtained by fitting to a target model which is close to the blue dots points (M2’s final position), but it extends much more longer than this target model in 2D space. Then two source models M1 and M2 starting at red and blue circle points to register to the IP. Although M2 successfully register to the correct position, M1 got failure result caused by the influence under the extended extra part of the IP fit.

4.2. Registration for US image case

Fig. 12 shows pose estimation process for a US image, where the US images are obtained by measuring phantom AT5514 [32] consisting of the cylinder-like shapes inside in different materials. Then we use a cylinder-like IP model (2-degree IP shown in red in Fig. 12)) to register the US images. The initial position is set as: the cylinder-like IP model crosses the US image plane with an angle of 85°. In the process the motion of the IP is driven to fit to the hole-like region in the image.

Fig. 12 shows three US image frames with the same initialization and in the second row of Fig. 12 the iteration number required

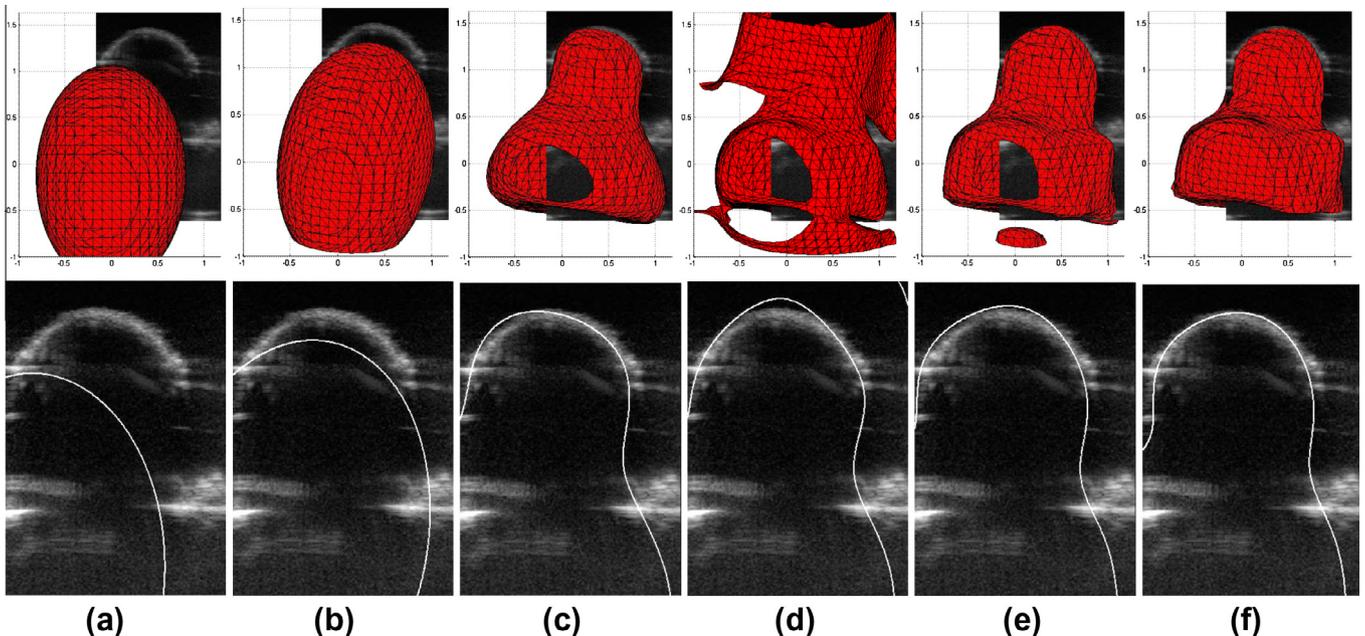


Fig. 14. Pose estimation for duck toy. First row: Relative poses of coarse-to-fine IPs and US images at each iteration. Second row: cross-section IP contour at selected iteration. From (a) to (f): IP increases as 2°, 4°, 6°, 8° and 10°.

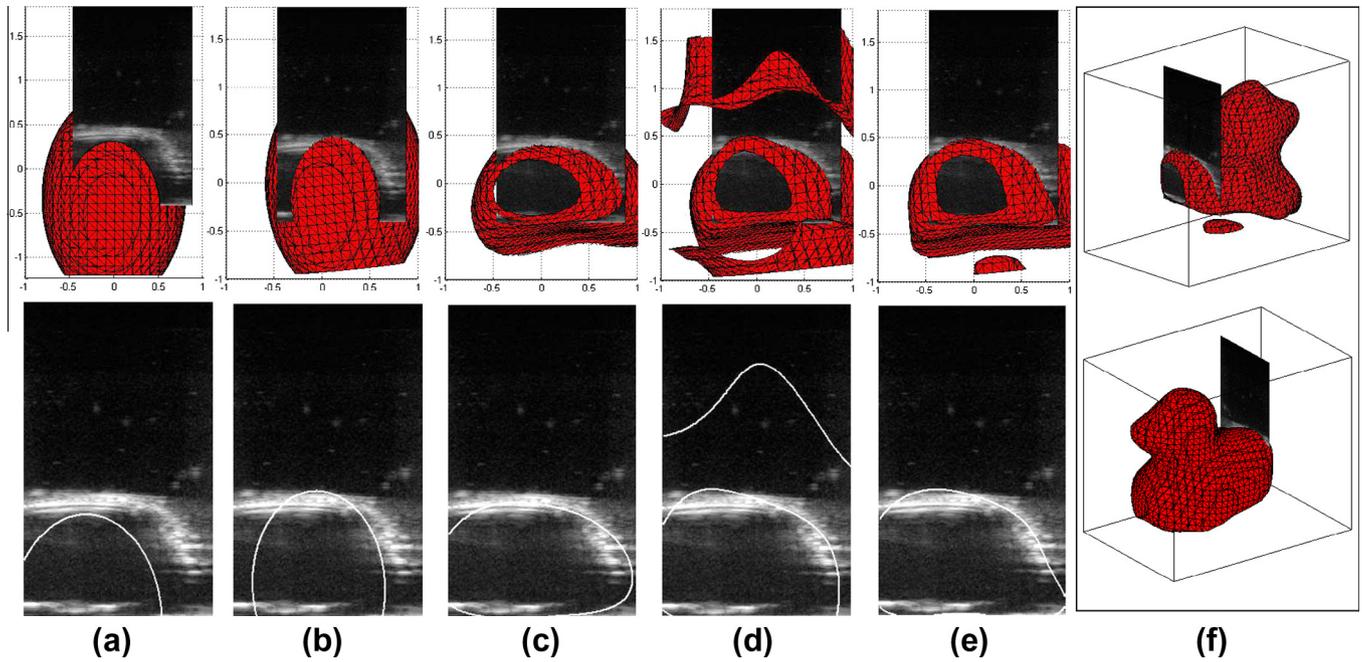


Fig. 15. From (a) to (f): same to Fig. 14. Rightmost: registration result shown in different views.

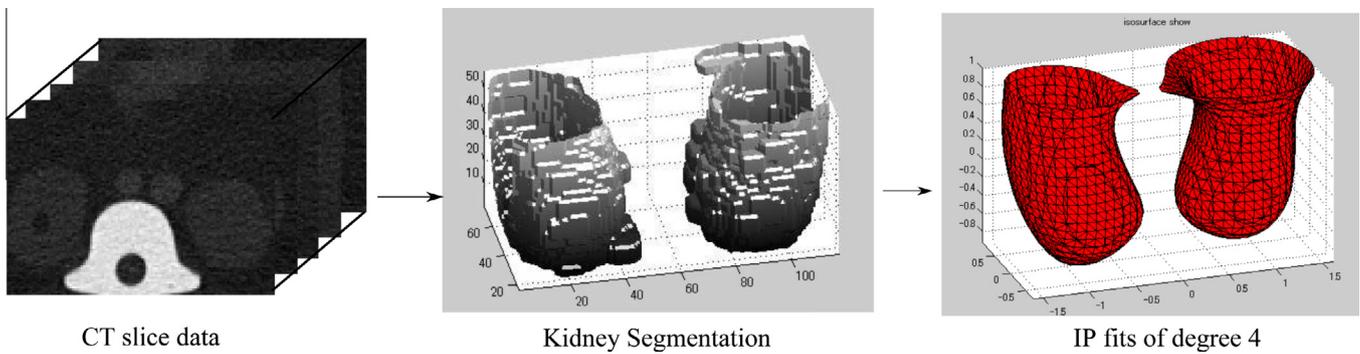


Fig. 16. Segmentation of kidney data from the CT scans and modeling with a 4-degree IP.

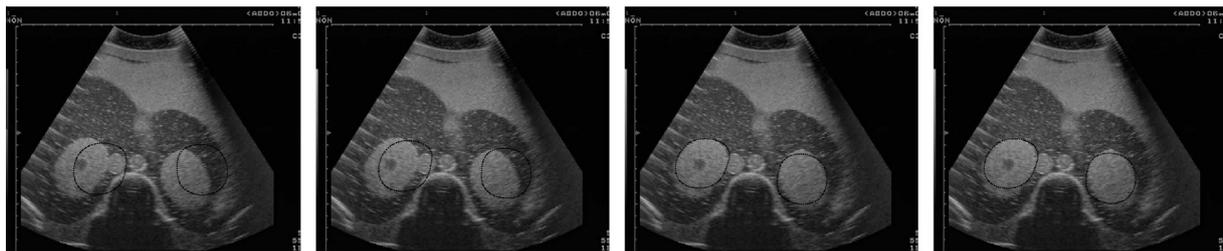


Fig. 17. US pose estimation for CT data with 4-degree IP. Cross-section contours of frames are shown in black points.

and final relative pose are shown. The third row shows the cross-section contour between the IP and the US plane at the final iteration. For each case, the consumed CPU time is within 30 ms.

4.2.1. Pose estimation for duck toy

Fig. 14 shows another result of the US image pose estimation where the images are obtained by scanning a duck toy made of rubber in the cistern shown in Fig. 13 left. The duck model scanned by 3D range sensor without denoising is shown in Fig. 13 right. Top

rows of Figs. 14 and 15 show the coarse-to-fine IP models and its relative positions to US images, and bottom rows show the cross-section contour resulting from the registration at the selected steps. In both of the cases the consumed CPU time is within 100 ms.

4.2.2. Pose estimation and tracking

The third example is to test by solving a registration problem between real CT data and a US image. To do this, we first seg-

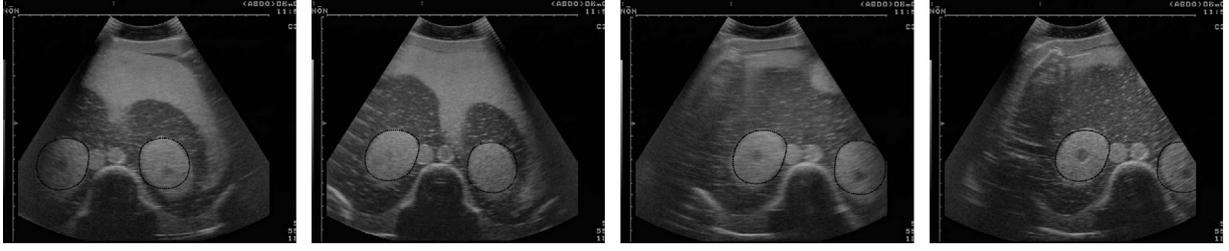


Fig. 18. Tracking for US image sequence. Cross-section contours of frames are shown in black points.

mented the CT data to obtain the desired organ object by modern segmentation methods such as Graph Cut (see Fig. 16). Note, since the organ object is simple, we model it only by a 4-degree IP as shown in Fig. 16 right. Finding the relative pose to CT data for a US image now becomes finding the relative pose to the IP model.

The first example is pose estimation for a single US image shown in Fig. 17 where we show the registration process with relative pose and cross-section contour at each selected iteration.

We also tried a tracking problem for a US image sequence by tracking the position of the kidney object with the above IP model. Results are shown in Fig. 18, where we show the cross-section contours of IP for selected frames.

5. Conclusions

In this paper, we extend our previous method, correspondences free registration with single IP model, to a new coarse-to-fine registration driven by the multiple IPs of incremental degrees. The better performance is achieved by two aspects: (i) The registration robustness and computational efficiency are improved by initial coarse registration with IP of low degree, since no extra IP surface appears and low cost of transformation for the IP and (ii) the registration accuracy is improved by IP of high-degree, due to the local details are representable for the IP.

In general, the proposed method is has the potential for real-time US image pose estimation because of four main advantages: (i) IP representation is robust against US image noise. (ii) IP transformation can be performed faster than transformation of data sets, since the number of data sets is often large extracted from US images. (iii) Thanks to IP gradient flows, the registration procedure can be performed by simple steepest descent-based minimization without time-consuming calculation of point-wise correspondences. (iv) The coarse-to-fine algorithm improves the robustness and accuracy than single IP registration.

Acknowledgments

This work was partially supported by Canon Inc., under the project: Physics-based vision theories for next-generation medical image processing.

Appendix A. A symbolical method for IP transformation

A.1. Pure rotation transformation

First, let us explain some notations for polynomial operations which were first introduced by Taubin and Cooper (see [20]). Let a coefficient a_{ijk} of an IP in Eq. (2) be presented as $\frac{\Phi_{ijk}}{\sqrt{ijk!}}$, and a vector

$\Phi_{[l]} = \left(\frac{\Phi_{000}}{\sqrt{0!0!0!}}, \frac{\Phi_{1-1,1,0}}{\sqrt{(l-1)!1!0!}}, \dots, \frac{\Phi_{00l}}{\sqrt{0!0!l!}} \right)^T$ be according to the l th form of IP.

An operation on this vector is defined as: $\Phi_{[l,m]} = \Phi_{[l]} \star \Phi_{[m]}^T$, where \star represents the classic matrix multiplication, with the difference

that the individual elements $\frac{\Phi_{ijk}}{\sqrt{ijk!}}$ and $\frac{\Phi_{abc}}{\sqrt{abc!}}$ in each vector are multiplied as $\frac{\Phi_{i+a,j+b,k+c}}{\sqrt{ijk!}\sqrt{abc!}}$. For example,

$$\Phi_{[1]} = \left(\frac{\Phi_{100}}{\sqrt{1!0!0!}}, \frac{\Phi_{010}}{\sqrt{0!1!0!}}, \frac{\Phi_{001}}{\sqrt{0!0!1!}} \right)^T, \quad (\text{A.1})$$

and

$$\Phi_{[1,1]} = \Phi_{[1]} \star \Phi_{[1]}^T = \begin{pmatrix} \Phi_{200} & \Phi_{110} & \Phi_{101} \\ \Phi_{110} & \Phi_{020} & \Phi_{011} \\ \Phi_{101} & \Phi_{011} & \Phi_{002} \end{pmatrix}. \quad (\text{A.2})$$

In [20], it was pointed out that under a non-singular coordinate transformation A , e.g., $\Phi'_{[l]} = A_{[l]} \Phi_{[l]}$, the transformed coefficient matrix is given by

$$\Phi'_{[l,m]} = A_{[l]}^{-T} \Phi_{[l,m]} A_{[m]}^{-1}, \quad (\text{A.3})$$

where $A_{[l]}$ is a non-singular $h_l \times h_l$ transformation matrix (assuming h_l be the number of monomial in the l th form), and $A_{[m]}$ is the same. From this equation, we know that if $A_{[l]}$ and $A_{[m]}$ are given, then a linear relationship between $\Phi'_{[l,m]}$ and $\Phi_{[l,m]}$ can be found; that is, the element-wise correspondence can be linearly expressed as

$$\Phi'_{ijk} = \sum_{\beta:i,j,k \leq l+m} a_{\alpha\beta} \Phi_{ijk} = \mathbf{a}'_2^T \Phi_{[l+m]}. \quad (\text{A.4})$$

Then since all of the elements in vector $\Phi_{[l+m]}$ are contained in matrix $\Phi_{[l,m]}$, and similarly all of the elements in $\Phi'_{[l+m]}$ are contained in $\Phi'_{[l,m]}$, a new linear correspondence can be constructed between $\Phi_{[l+m]}$ and $\Phi'_{[l+m]}$ by arranging the necessary elements in the right order into $\Phi'_{[l+m]}$ as

$$\begin{aligned} \Phi'_{[l+m]} &= (\Phi'_{l+m,0,0}, \Phi'_{l+m-1,1,0}, \Phi'_{l+m-2,2,0}, \dots)^T \\ &= (\mathbf{a}'_1^T \Phi_{[l+m]}, \mathbf{a}'_2^T \Phi_{[l+m]}, \mathbf{a}'_3^T \Phi_{[l+m]}, \dots)^T = A_{[l+m]} \Phi_{[l+m]}, \end{aligned} \quad (\text{A.5})$$

if the $(l+m)$ th transformation matrix is constructed as: $A_{[l+m]} = [\mathbf{a}'_1^T, \mathbf{a}'_2^T, \mathbf{a}'_3^T, \dots]^T$.

As a simple example, the elements of $\Phi_{[2]}$ ($= (\frac{\Phi_{200}}{2}, \Phi_{110}, \Phi_{101}, \frac{\Phi_{020}}{2}, \Phi_{011}, \frac{\Phi_{002}}{2})^T$) are contained in $\Phi_{[1,1]}$ shown in (A.2), and the elements of $\Phi'_{[2]}$ are contained in $\Phi'_{[1,1]}$. From the equation $\Phi'_{[1,1]} = R_{[1]} \Phi_{[1,1]} R_{[1]}^T$ (supposing $R_{[1]}$ is a pure rotation matrix), it is easy to find out the relationship of a linear expansion for the elements, e.g., the first element of $\Phi'_{[1,1]}$ is expanded as: $\Phi'_{200} = \mathbf{r}_1 \Phi_{[1,1]} \mathbf{r}_1^T = r_{11}^2 \Phi_{200} + 2r_{11}r_{12} \Phi_{110} + 2r_{11}r_{13} \Phi_{101} + r_{12}^2 \Phi_{020} + 2r_{12}r_{13} \Phi_{011} + r_{13}^2 \Phi_{002}$, where \mathbf{r}_1 is the first row of $R_{[1]}$ and r_{ij} is an element of $R_{[1]}$. Then since all the elements in $\Phi_{[1,1]}$ are contained in $\Phi_{[1]}$, we can find a linear relationship $\frac{1}{2} \Phi'_{200} = \mathbf{a}'_1^T \Phi_{[2]}$, where $\mathbf{a}'_1 = (r_{11}^2, r_{11}r_{12}, r_{11}r_{13}, r_{12}^2, r_{12}r_{13}, r_{13}^2)^T$ is viewed as the first row of $A_{[2]}$.

Therefore, transformation matrices for different forms can be calculated in an incremental manner as.

Algorithm 2. Algorithm

(1)	Initialization: given $A_{[1]}$;
(2)	$\Phi'_{[1,1]} = A_{[1]}^{-T} \Phi_{[1,1]} A_{[1]}^{-1} \rightarrow A_{[2]}$;
(3)	$\Phi'_{[1,2]} = A_{[1]}^{-T} \Phi_{[1,2]} A_{[2]}^{-1} \rightarrow A_{[3]}$;
	\vdots
(n)	$\Phi'_{[1,n-1]} = A_{[1]}^{-T} \Phi_{[1,n-1]} A_{[n-1]}^{-1} \rightarrow A_{[n]}$.

Here, \rightarrow represents constructing the transformation matrix for the next step, after finding out the linear mapping relationship between $\Phi'_{[l,j]}$ and $\Phi_{[1,l]}$. Note if the initial argument $A_{[1]}$ represents the pure rotation, then $A_{[l]}$ is an orthogonal matrix (see [20]), and thus Eq. (A.3) degenerates to

$$\Phi'_{[l,m]} = A_{[l]} \Phi_{[l,m]} A_{[m]}^T.$$

A.2. Affine transformation

By using homogeneous coordinates, an IP of n variables in Euclidean space can be described in projective space by a corresponding homogeneous IP of $n + 1$ variables. To convert a ternary (i.e. 3D) IP of degree d

$$f_{3D}^d(x, y, z) = \sum_{0 \leq i, j, k; i+j+k \leq d} a_{ijk} x^i y^j z^k \quad (\text{A.6})$$

into its homogeneous representation, a new component $t = 1$ is added to the 3D IP as

$$f_{4D}^d(x, y, z, t) = \sum_{0 \leq i, j, k, l; i+j+k+l=d} a_{ijkl} x^i y^j z^k t^l. \quad (\text{A.7})$$

Therefore a homogeneous polynomial corresponding to a 3D IP of degree d is a form of degree d in a 4D IP. Thus the procedure mentioned in the last section can be used to transform the homogeneous IP (4D IP). The different is that the incremental algorithm starts from a 4×4 affine transformation matrix $A_{[1]}$, e.g., in the Euclidean case

$$A_{[1]} = \begin{pmatrix} R_{3 \times 3} & \mathbf{t} \\ \mathbf{0}_{1 \times 3} & 1 \end{pmatrix},$$

where $R_{3 \times 3}$ and \mathbf{t} are a pure 3D rotation matrix and a translation vector. As a result, a 3D IP of degree d can be affine-transformed only by $A_{[d]}$, once it is worked out. Furthermore, the incremental scheme can be modified for saving the computational cost, e.g., for calculating $A_{[9]}$ we can select the incremental order as:

$$A_{[1]} \rightarrow A_{[2]} \rightarrow A_{[4]} \rightarrow A_{[8]} \rightarrow A_{[9]}.$$

References

- [1] N. Pagoulatos, S. Member, W.S. Edwards, D.R. Haynor, Y. Kim, Interactive 3-D registration of ultrasound and magnetic resonance images based on a magnetic position sensor, *IEEE Trans. Inform. Technol. Biomed.* 3 (4) (1999) 278–288.
- [2] J.W. Trobaugh, W.D. Richard, K.R. Smith, R.D. Bucholz, Frameless stereotactic ultrasonography: method and applications, *Comput. Med. Imag. Graph.* 18 (1994) 235–246.
- [3] J. Blackall, G. Penney, A. King, D. Hawkes, Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion and deformation, *Trans. Med. Imag.* 24 (11) (2005) 1405–1416.

- [4] G. Penney, D. Barratt, C. Chan, M. Slomczykowski, T. Carter, P. Edwards, D. Hawkes, Cadaver validation of intensity-based ultrasound to CT registration, *Med. Image Anal.* 10 (3) (2006) 385–395.
- [5] X. Huang, N.A. Hill, J. Ren, G. Guiraudon, T.M. Peters, Intra-cardiac 2d us to 3d ct image registration, in: *Proceedings of the SPIE International Symposium on Medical Imaging 2007* (6509-85).
- [6] G. P. Penney, J. M. Blackall, D. Hayashi, T. Sabharwal, A. Adam, D.J. Hawkes, Overview of an ultrasound to ct or mr registration system for use in thermal ablation of liver metastases, in: *MUAF01*, 2001, pp. 65–68.
- [7] D.V. Amin, T. Kanade, A.M. DiGioia, B. Jaramaz, Ultrasound registration of the bone surface for surgical navigation, *Comput. Aid. Surg.* 8 (1) (2003) 1–16.
- [8] T. Lange, S. Eulenstein, M. Hunerbein, H. Lamecker, P. Michael Schlag, Augmenting intraoperative 3d ultrasound with preoperative models for navigation in liver surgery, in: *Proc. MICCAI 2004*, 2004, pp. 534–541.
- [9] J.P.W. Pluim, J.B.A. Maintz, M.A. Viergever, Mutual-information-based registration of medical images: a survey, *IEEE Trans. Med. Imag.* (2003) 986–1004.
- [10] W. Wein, B. Roper, N. Navab, Automatic registration and fusion of ultrasound with CT for radiotherapy, *Proc. MICCAI2005 2* (2005) 303–311.
- [11] W. Wein, S. Brunke, A. Khamene, M.R. Callstrom, N. Navab, Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention, *Med. Image Anal.* 12 (5) (2008) 577–585.
- [12] A. Leroy, P. Mozer, Y. Payan, J. Troccaz, Rigid registration of free-hand 3D ultrasound and CT-scan kidney images, *Proc. MICCAI2004* 3216 (2004) 837–844.
- [13] A. Roche, X. Pennec, G. Malandain, N. Ayache, Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information, *IEEE Trans. Med. Imag.* 20 (10) (2001) 1038–1049.
- [14] P. Besl, N. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern. Anal. Mach. Intell.* (TPAMI) 14 (2) (1992) 239–256.
- [15] T. Oishi, A. Nakazawa, R. Kurazume, K. Ikeuchi, Fast simultaneous alignment of multiple range images using index images, in: *Proc. The 5th International Conference on 3-D Digital Imaging and Modeling (3DIM 2005)*, 2005, pp. 476–483.
- [16] A.W. Fitzgibbon, Robust registration of 2d and 3d point sets, *Image Vis. Comput.* 21 (13–14) (2003) 1145–1153.
- [17] X. Huang, S. Member, N. Paragios, D.N. Metaxas, S. Member, S. Member, Shape registration in implicit spaces using information theory and free form deformations, *IEEE Trans. Pattern Anal. Mach. Intell.* (TPAMI) 28 (2006) 1303–1318.
- [18] J. Salvi, C. Matabosch, D. Fofi, J. Forest, A review of recent range image registration methods with accuracy evaluation, *Image Vis. Comput.* 25 (5) (2007) 578–596.
- [19] J.-P. Tarel, H. Civi, D. B. Cooper, Pose estimation of free-form 3D objects without point matching using algebraic surface models, in: *Proceedings of IEEE Workshop Model Based 3D Image Analysis*, 1998, pp. 13–21.
- [20] G. Taubin, D. Cooper, *Symbolic and Numerical Computation for Artificial Intelligence*, Computational Mathematics and Applications, Academic Press, 1992. Chapter 6.
- [21] B. Zheng, R. Ishikawa, T. Oishi, J. Takamatsu, K. Ikeuchi, 6-dof pose estimation from single ultrasound image using 3d ip models, in: *Computer Vision Pattern Recognition Workshop on OTCVBS 0*, 2008, pp. 1–8.
- [22] B. Zheng, R. Ishikawa, T. Oishi, J. Takamatsu, K. Ikeuchi, A fast registration method using IP and its application to ultrasound image registration, *IPSJ Trans. Comput. Vis. Appl.* 1 (2009) 209–219.
- [23] M. Rouhani, A. Sappa, Correspondence free registration through a point-to-model distance minimization, in: *2011 IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 2150–2157.
- [24] T. Tasdizen, J.-P. Tarel, D.B. Cooper, Improving the stability of algebraic curves for applications, *IEEE Trans. Imag. Proc.* (TIP) 9 (3) (2000) 405–416.
- [25] B. Zheng, J. Takamatsu, K. Ikeuchi, An adaptive and stable method for fitting implicit polynomial curves and surfaces, *IEEE Trans. Patt. Anal. Mach. Intell.* (TPAMI) 32 (3) (2010) 561–568.
- [26] M. Blane, Z.B. Lei, D.B. Cooper, The 3L algorithm for fitting implicit polynomial curves and surfaces to data, *IEEE Trans. Patt. Anal. Mach. Intell.* (TPAMI) 22 (3) (2000) 298–313.
- [27] B. Zheng, J. Takamatsu, K. Ikeuchi, Adaptively determining degrees of implicit polynomial curves and surfaces, in: *Proc. Conf. Asian Conf. on Computer Vision (ACCV) LNCS 4844*, 2007, p. 289–300.
- [28] T. Sahin, M. Unel, Fitting globally stabilized algebraic surfaces to range data, in: *Proc. IEEE Conf. Int. Conf. on Computer Vision (ICCV)*, vol. 2, 2005, pp. 1083–1088.
- [29] G. Taubin, Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.* (TPAMI) 13 (11) (1991) 1115–1138.
- [30] L. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, 1998.
- [31] F. Catte, P.-L. Lions, J.-M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* 29 (1) (1992) 182–193.
- [32] <http://www.atlabs.com>.