Computer Vision and Image Understanding 117 (2013) 1647–1658

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

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]

[a] The University of Tokyo, Ikeuchi Lab, Meguro-ku, Tokyo, 153-8505, Japan
[b] Medical Imaging Project, Canon Inc. Ohta-ku, Tokyo, 146-8501, Japan
[c] Nara 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.

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@cvl.iis.u-tokyo.ac.jp (B. Zheng).

1077-3142/$ - see front matter
http://dx.doi.org/10.1016/j.cviu.2013.01.015

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

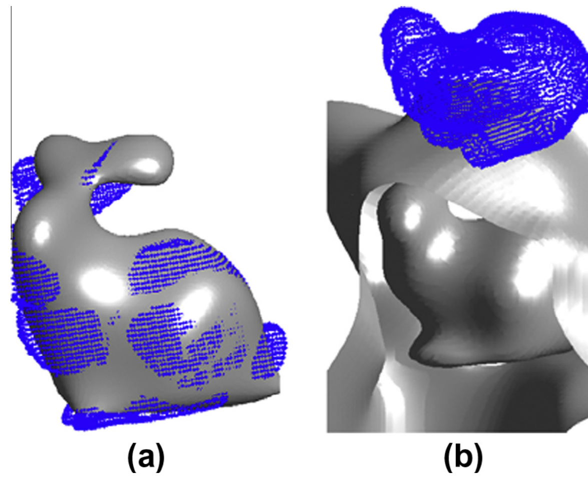

Fig. 1. Registration example (IP is shown in gray surface and target model is shown in blue points). (a) Inaccurate result caused by low degree IP with lack of detailed representation and (b) failure registration caused by high degree IP with undesired surfaces. (For interpretation of the references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

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-

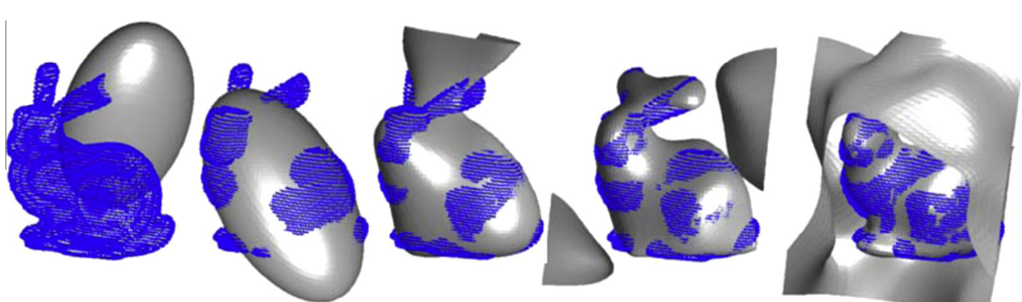

Fig. 2. Coarse-to-fine registration in general case driven by IP models from low degree to high degree (from left to right). Leftmost: the initial position for registration. The initial position for each step is determined by the result of previous step.