Evaluation of Medical Image Registration
by Using 3D SIFT and Phase-Only Correlation

Zisheng Li¹, Tsuneya Kurihara¹, Kazuki Matsuzaki¹, and Toshiyuki Irie²

¹ Central Research Laboratory, Hitachi, Ltd.,
1-280 Higashi-Koigakubo, Kokubunji-shi, Tokyo 185-8601, Japan
² Department of Radiology, Hitachi General Hospital,
2-1-1 Jonan-cho, Hitachi-shi, Ibaraki 317-0077, Japan
{zisheng.li.fj,tsuneya.kurihara.vn,kazuki.matsuzaki.gv,
toshiyuki.irie.rq}@hitachi.com

Abstract. An effective method for quantitatively evaluating rigid and non-rigid
image registration without any manual assessment is proposed. This evaluation
method is based on feature point detection in reference images and correspond-
ing point localization in registered floating images. For feature point detection,
a 3D SIFT keypoint detector is applied to determine evaluation reference points
in liver vessel regions of reference images. For corresponding point
localization, a 3D phase-only correlation approach is applied to match reference
points and their corresponding points. Distance between the reference points
and the correspondences can be used to estimate image registration errors. With
the proposed method, users can evaluate different registration algorithms using
their own image data automatically.

Keywords: Image registration, evaluation, feature point detection, corresponding point localization.

1 Introduction

Image registration aims to find a spatial transformation that maps points from one
image to corresponding points in another image. Medical image registration is
fundamentally used in many applications, such as diagnosis, planning treatment, guiding
treatment, and monitoring disease progression. Thus, it is necessary to validate
whether a rigid/non-rigid registration algorithm satisfies the needs of an image
processing application with high accuracy, robustness, and other performance criteria.

The most straightforward method for estimating image registration error is to com-
pare a given registration transformation with a “gold standard” transformation [1],
whose accuracy is high. However, the lack of a gold standard prevents any automatic
assessment of registration accuracy. Even if individuals trained to interpret medical
images are involved in an experimental validation of registration algorithm, it is diffi-
cult to provide a method for consistently and accurately assessing individual images.
As a result, few attempts have been made to evaluate and compare the performance of
image registration algorithms. Even so, two projects that stand out in this regard are
the “Retrospective Image Registration and Evaluation (RIRE) Project” [2] (for evaluating rigid registration accuracy) and the “Non-rigid Image Registration Evaluation Project (NIREP)” [3]. The RIRE project used bone-implanted fiducial markers to obtain a marker-based rigid transformation as the gold-standard transformation. Registration error was measured by calculating the error relative to the gold standard over a set of specified regions. In contrast, instead of providing a gold-standard transformation for accuracy evaluation, NIREP provided four metrics for evaluating the performance of non-rigid image registration. Both projects required manual annotation and segmentation to create evaluation databases, and the evaluation data only included brain images.

In our previous work, for clinical applications such as computer-aided diagnosis and observation of treatment progress [7, 12], rigid and non-rigid image registration algorithms are developed. The present study aims to quantitatively evaluate the accuracy and robustness of these methods by using clinical data, especially abdominal CT images for diagnosis of hepatic tumor. Since we do not have any manually annotated image data as ground truth for this evaluation task, it is necessary to develop an automatic method for evaluating the accuracy of the previously developed registration algorithms. To satisfy that necessity, the following method is proposed. In abdominal CT images, liver vessel regions of reference images are segmented, and feature points in the segmented regions are detected using a 3D SIFT (scale-invariant feature transform) keypoint detector [4]. In floating images after registration, correspondences of the SIFT feature points are searched for and localized using a 3D phase-based image-matching method [5, 13]. By calculating distance between the feature point pairs, it is possible to obtain a numerical-registration error without any manual assessment of the registration algorithm. A brief description of the proposed method is given in Section 2, and the fiducial-point detection and the correspondence matching for evaluating registration accuracy are described in Sections 3 and 4, respectively. Results of experiments on medical image registration are presented in Section 5, and concluding remarks are given in Section 6.

## 2 Proposed Method

To quantitatively measure registration error, distance between landmarks or regions in reference images and their correspondences in floating images after registration should be accurately estimated. It is therefore necessary to develop an accuracy-evaluation framework that mainly consists of two stages: fiducial point detection and corresponding point matching. The framework of the proposed method is illustrated in Fig. 1, and some data examples generated by different procedures for accuracy evaluation are shown in Fig. 2.

As for landmark detection, to extract reference points, scale-space extrema detection, i.e., keypoint detection in 3D SIFT features [4], is adopted. Since anatomic landmark points of vessel regions are considered as appropriate reference points for evaluating the accuracy of medical image registration [6], liver vessel regions are segmented from our clinical data, and SIFT feature points are extracted from the