

Assessment of Similarity Measures for Accurate Deformable Image Registration

Yuji Yaegashi*, Kunihiro Tateoka, Kazunori Fujimoto, Takuya Nakazawa, Akihiro Nakata, Yuichi Saito, Tadanori Abe, Masaki Yano and Koichi Sakata

Department of Radiation Oncology and Medical Physics, Graduate School of Medicine, Sapporo Medical University, Japan

Abstract

Purpose: Deformable image registration is widely used in radiation therapy applications. There are several different algorithms for deformable image registration. The purpose of this study was to evaluate the optimal similarity measures needed to obtain accurate deformable image registration by using a phantom.

Methods: To evaluate the optimal similarity measures for the deformable image registration, we compared several similarity measures, including the normalized correlation coefficient, the mutual information, the dice similarity coefficient, and the Tanimoto coefficient. In this study, the mutual information was normalized to have a value of 1 when there is complete correspondence between the images in order to compare it with other similarity measures. First, a reference image was acquired with the phantom located in the center of the field of view of a computed tomography. The phantom consisted of two sections a Teflon sphere and four samples of various electron density values. Then, to acquire the moving images, the phantom was scanned for various displacement values as it was moved to the left (range: 1.00-30.0 mm). Second, images for various Teflon sphere diameters (range: 0-25.4 mm) were acquired with the CT scanner. The image similarity for each condition was compared with the reference image by using several similarity measures.

Results: In the moved phantom study, although the normalized correlation coefficient, dice similarity coefficient, and Tanimoto coefficient showed the same tendency of sensitivity for measuring image similarity, the mutual information showed significant sensitivity for both of the two distinct sections of the phantom. In the study in which the phantom sphere diameter was varied, the mutual information also showed the best performance among the tested similarity measures.

Conclusions: Mutual information appears to have an advantage over other similarity measures for accurate deformable image registration.

Keywords: Deformable image registration (DIR); Mutual information (MI); Similarity measures; Phantoms

Introduction

Image-Guided Radiation Therapy (IGRT) is the most advanced technology for localizing targets. IGRT is the use of frequent imaging during a course of radiation therapy to improve the precision and accuracy of the delivery of treatment. The use of daily images in the radiotherapy process leads to Adaptive Radiation Therapy (ART), in which the treatment is evaluated periodically, and the plan is modified in an adaptive manner for the remaining course of radiation therapy. The images obtained from Cone-Beam Computed Tomography (CBCT) at the time of treatment delivery also provide information on the changes that can occur in the patient anatomy during a course of radiation therapy, including therapeutic response of the tumor or normal tissue, internal motion, and weight loss.

Recently, Deformable Image Registration (DIR) has been a very important component in ART [1,2]. For instance, organ contours have been transferred from the planning CT images to the daily CBCT images by using DIR such as auto-segmentation [3,4]. DIR is otherwise used for four-Dimensional (4D) treatment optimization [5-9] and dose accumulation [3,10]. Several different DIR algorithms have been proposed, including B-spline [11], thin-plate spline [12], Thirion's demon [13,14], and viscous fluid [15,16]. The B-spline method is the transformation of a point which computer from the control points using a defined grid between two images. The thin-plate spline method is a physically motivated interpolation scheme for arbitrarily spaced tabulated data. Thirion's demons method uses gradient information from a static reference image to determine the 'demons' force required

to deform the 'moving' image [14]. Viscous fluid registration is considered to be embedded in viscous fluid, the motion of which is determined by Navier-Stokes equations for conservation of momentum [16]. For comparisons of the DIR accuracy of different algorithms, several similarity measures exist such as the Normalized Correlation Coefficient (NCC), Mutual Information (MI), the Dice Similarity Coefficient (DSC), and the Tanimoto Coefficient (TC). However, it is not clear which similarity measures are suitable for assessing the accuracy of deformable image registration. Therefore, it is necessary to investigate quantitatively the sensitivity of the image similarity measurement for each of the similarity measures. The purpose of this study is to evaluate the optimal similarity measures needed to obtain accurate DIR by using a phantom.

Materials and Methods

To determine which similarity measures have the best performance

***Corresponding author:** Yuji Yaegashi, Department of Radiation Oncology and Medical Physics, Graduate School of Medicine, Sapporo Medical University, Japan, E-mail: y-yaegashi@sapmed.ac.jp

Received September 25, 2012; **Accepted** November 17, 2012; **Published** November 27, 2012

Citation: Yaegashi Y, Tateoka K, Fujimoto K, Nakazawa T, Nakata A, et al. (2012) Assessment of Similarity Measures for Accurate Deformable Image Registration. J Nucl Med Radiat Ther 3:137. doi:10.4172/2155-9619.1000137

Copyright: © 2012 Yaegashi Y, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

for DIR, the accuracy of the similarity measures were estimated by using the phantom under some set of conditions. The phantom study was performed only under simple conditions, in order to measure as quantitatively as possible. The similarity measures we considered were the NCC, MI, DSC, and TC.

Phantom study

In this phantom study, the ISIS QA-1 (TGM², Clearwater, FL) was used to evaluate the accuracy of the similarity measures (Figure 1). The ISIS QA-1 has been developed for quality assurance for CT-simulators and treatment accelerators. The phantom is composed of two sections: a 25.4-mm Teflon sphere located in the center of the phantom and four built-in known electron-density values of bone, water, and lung at inhale and exhale. The ISIS QA-1 was scanned with a 4-slice GE Lightspeed RT wide-bore CT scanner (GE Medical Systems, Waukesha, WI). All images were acquired under the same CT scanner settings (kV, mA, slice thickness, etc.).

First, a reference image was acquired with the phantom located in the center of the CT field of view. Then, to acquire the moving images, the phantom was scanned for various displacement values as it was moved to the left (range: 1.0-30.0 mm) of both the Teflon sphere section and the four electron density values (Figure 2). Then, the ISIS-QA1 was scanned at various Teflon sphere diameters (range: 0-25.4 mm). The reference image was defined as 25.4 mm sphere diameter image. The image similarity compared with the reference image was computed for each condition by using similarity measures (Figure 3).

Similarity measures

To assess the optimal similarity measures for accuracy of different DIR algorithms, we evaluated four similarity measures as follows.

Normalized correlation coefficient

Cross-correlation can be used as a measure for calculating the degree of similarity between two images. The advantage of the NCC over cross-correlation is that it is less sensitive to linear changes in the amplitude of grayscale values in the two compared images. Furthermore, the NCC is confined in the range between -1 and 1. If the two images correspond completely, the value of NCC is 1. Its mathematical definition is

$$NCC = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A(i,j) - \bar{A})(B(i,j) - \bar{B})}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A(i,j) - \bar{A})^2 \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (B(i,j) - \bar{B})^2}}$$

Where $A(i,j)$ and $B(i,j)$ are the moving image and the reference image of the coordinate (i,j) , respectively. N and M represent the dimensions of the image matrix $N \times M$.

Mutual information

Mutual information is an information theory measure of the statistical dependence between two random variables, which represents an entropy measure. The most commonly used measure of information in image processing is the Shannon-Wiener entropy measure. The entropy of the image can be thought of as a measure of dispersion in the distribution of the image grayscale values. Maximization of MI indicates complete correspondence between two images.

The MI is defined as follows:

$$MI(A,B) = H(A) + H(B) - H(A,B)$$

$$= H(B) + H(B|A)$$

$$= H(A) + H(A|B)$$

where $H(A)$ and $H(B)$ are the entropies of images A and B , respectively, and $H(A,B)$ is their joint entropy. MI measures the amount of information gained about B when A is known, and vice versa. If A and B are independent, then $MI(A,B) = 0$. MI increases as the dependence between A and B increases.

The entropies and joint entropy can be computed from the following equations:



Figure 1: For estimation of the similarity measures, images were obtained from CT using the ISIS QA-1 phantom. The phantom was set on the CT couch with an adjustable alignment bar.

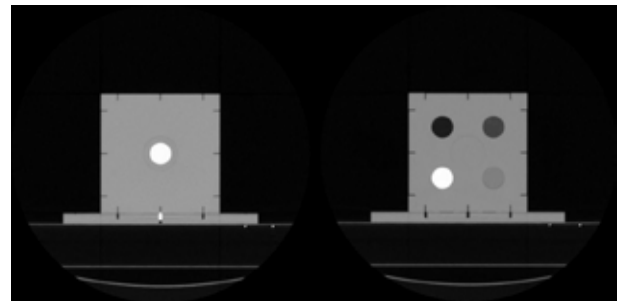


Figure 2: The phantom contains a Teflon sphere (25.4 mm) centered in the insert and four different Hounsfield Unit (HU) materials. The four inserts are: Bone (+800), lung at inhale (-800), lung at exhale (-500) and water (0).

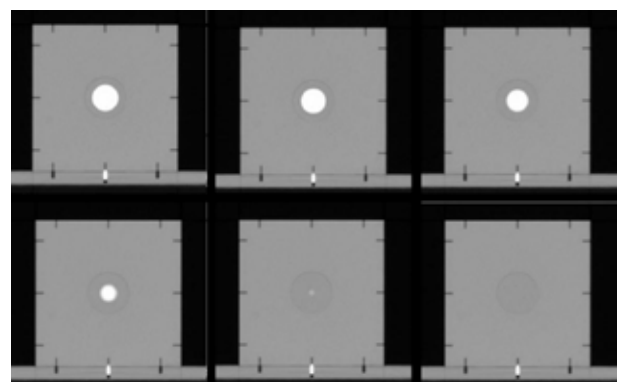


Figure 3: CT images of the Teflon sphere as the diameter was changed from 25.4 mm to 0 mm. Similarity measurements were performed using the image with the sphere diameter of 25.4 mm as the reference image.

$$H(A) = -\sum_a P_A(a) \cdot \log P_A(a)$$

$$H(A,B) = -\sum_{a,b} P_{A,B}(a,b) \cdot \log P_{A,B}(a,b)$$

$$H(A|B) = -\sum_{a,b} P_{A,B}(a,b) \times \log P_{A,B}(a|b)$$

where $P_A(a)$ and $P_B(b)$ are the marginal probability mass functions and $P_{A,B}(a,b)$ is the joint probability mass function. The MI measures the degree of dependence between A and B by measuring the distance between the joint distribution $P_{A,B}(a,b)$ and the distribution associated with the case of complete independence $P_A(a) P_B(b)$. The probability mass function $P_{A,B}(a,b)$ can be calculated using the joint histogram of two images.

The MI for grayscale values a and b at equivalent locations in two images A and B is defined as;

$$MI(A,B) = \sum_{a,b} P_{A,B}(a,b) \cdot \log \frac{P_{A,B}(a,b)}{P_A(a) \cdot P_B(b)}$$

Here, the MI was normalized to have a value of 1 when there is complete correspondence between the images in order to compare it with other similarity measures.

Dice similarity coefficient

The DSC is a similarity measure between images A and B , which ranges from 0 for no correspondence between the images to 1 for complete correspondence. The DSC is defined as

$$DSC = \frac{2 \cdot |A \cap B|}{|A| + |B|}$$

Tanimoto coefficient

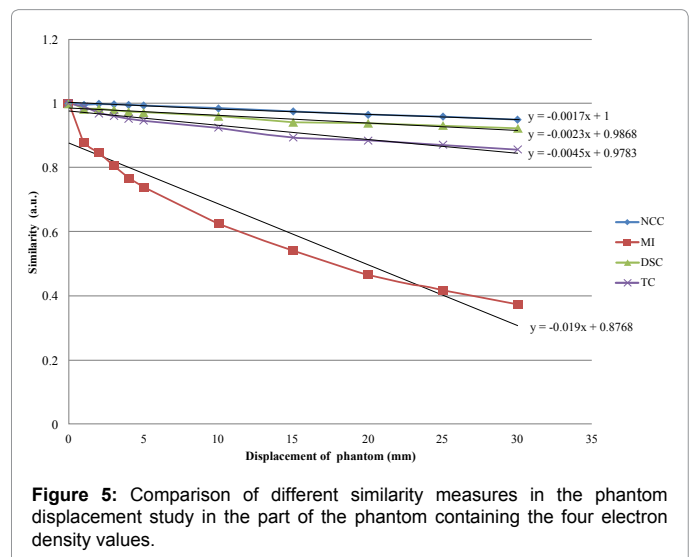
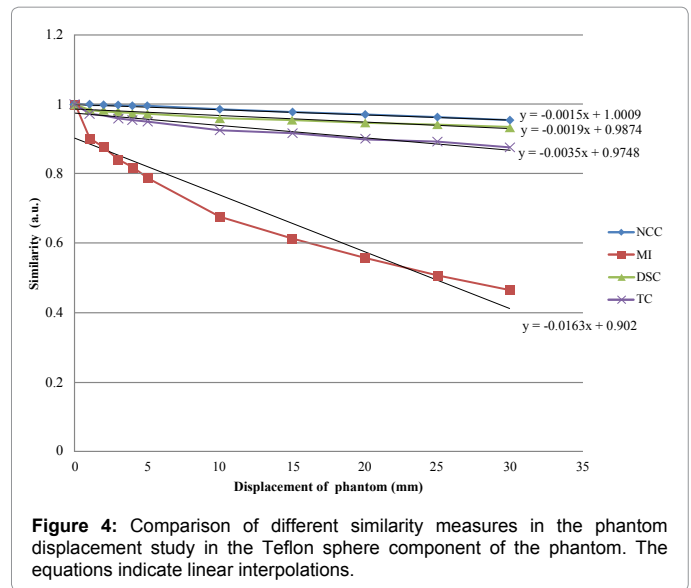
The TC (also known as the extended Jaccard coefficient) is another measure of the similarity between images A and B . A higher TC indicates a better correspondence between the images. A value of 1 indicates complete correspondence, and a value of 0 means that there is no correspondence at all. The TC is defined as

$$TC = \frac{|A \cap B|}{|A \cup B|}$$

$$= \frac{A \cdot B}{|A| + |B| - A \cdot B}$$

Results

Figure 4 shows the image similarity of the various displacements of the phantom at the section of the Teflon sphere. The mean rates of change of the image similarities with displacement can be obtained from the linear fits to the data and the corresponding slope, which are 0.0015 (NCC), 0.0019 (DSC), 0.0035 (TC), and 0.0163 (MI). These values also indicate the sensitivity of the image similarity measurement. The image similarities indicated by the NCC, DSC, and TC all show a similar slight decrease with increasing phantom offset in position. Compared to other similarity measures, the MI showed a significant decrease in image similarity with phantom offset. Figure 5 shows the image similarity using the part of the phantom containing the four electron density values, which has a more complex Hounsfield Unit (HU) for the CT image used to make figure 5 than that of the Teflon sphere section. The mean rates of change of the image similarities

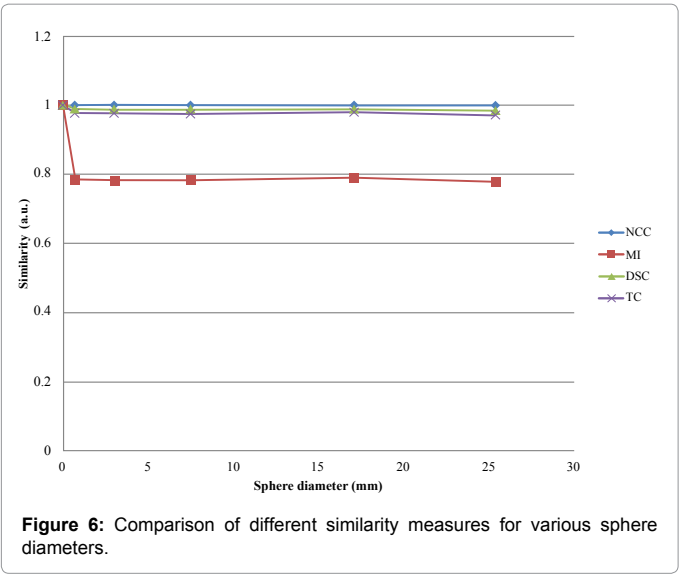


with displacement in figure 5 were found to be 0.0017 (NCC), 0.0023 (DSC), 0.0045 (TC), and 0.0190 (MI). The results were similar under both phantom conditions. Therefore, the MI showed the highest sensitivity of all the image similarity measures in the various phantom displacement studies.

Figure 6 shows the image similarity for the various Teflon sphere diameters. Despite the fact that the Teflon sphere diameter gradually decreased to 0 mm, the mean rates of change of the image similarities with respect to sphere diameter indicate almost no change for the NCC, DSC, and TC. Although the image similarity calculated from the MI initially shows a decrease, it indicates a constant value thereafter.

Discussion

For the two phantom displacement studies, the image similarity sensitivities calculated from all of the similarity measures demonstrated the same tendency. Consequently, the MI image similarity sensitivity is higher than other that of similarity measures; it is not dependent on the complexity of the HUs in the CT images. The reasons for this



are as follows: the NCC indicates the linearity of the image similarity between two images in a pixel-by-pixel manner, while the DSC and TC simply express the rate of change of the image similarity between two images with a change in displacement. Thus, the NCC, DSC, and TC demonstrate that when there are many pixel values that match between the CT images, the differences in image similarity may cancel out. However, the MI demonstrates good sensitivity to the image similarity for the phantom displacement case, because MI is not calculated pixel by pixel but instead uses the joint histogram of the grayscale values of the two images. The joint histogram is used to estimate a joint probability distribution of their grayscale values by dividing each entry in the histogram by the total number of entries.

For the study involving various Teflon sphere diameter values, none of the similarity measures showed any significant differences. The results suggested that there are many grayscale values that match between the two images. To verify this explanation, we drew regions of interest (ROIs) around each of the various Teflon sphere diameter images as shown in figure 7, and the image similarities in each ROI were compared with the reference image (25.4 mm sphere diameter image) by using several similarity measures as shown in figure 8. The image similarity sensitivities became higher for all the similarity measures. In particular, the NCC showed a negative correlation as compared with the reference image when the Teflon sphere diameter was zero. Similarly, in figure 8, the image similarity calculated by the MI initially shows a decrease, and thereafter, it has a constant value. Based on these results, it seemed that the similarity measures other than MI can be used for image similarity measurement when it is possible to define an ROI.

To estimate similarity measures for clinical images, we tested patient data obtained from four-dimensional computed tomography (4DCT) images (Figure 9), because the 4DCT images have complete correspondence in the locations for each respiratory phase image. Therefore, the error caused by the difference in the location between images can be excluded. Figure 10 shows a comparison of the differences in similarity measures for the lung cancer images obtained with 4DCT for different respiratory phases. The 4DCT dataset used comprised 10 respiratory phases. The end-inhalation phase was typically defined as the 0% phase, and the end-exhalation phase was defined as the 50% phase. In this study, the similarity measures of the

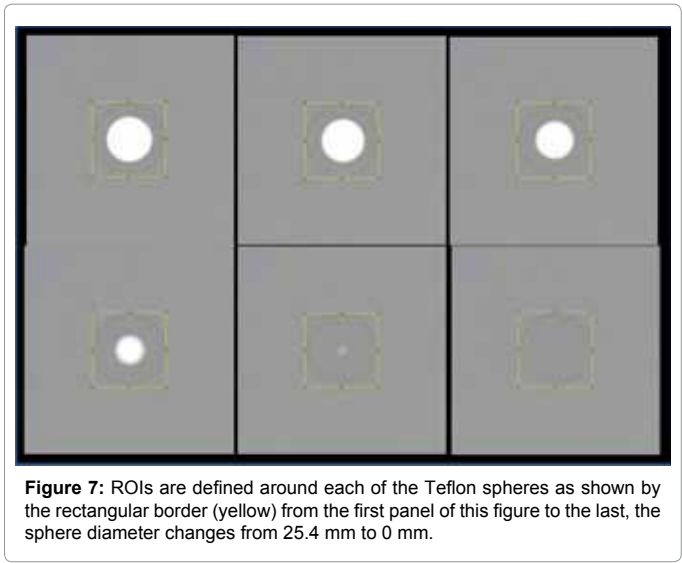
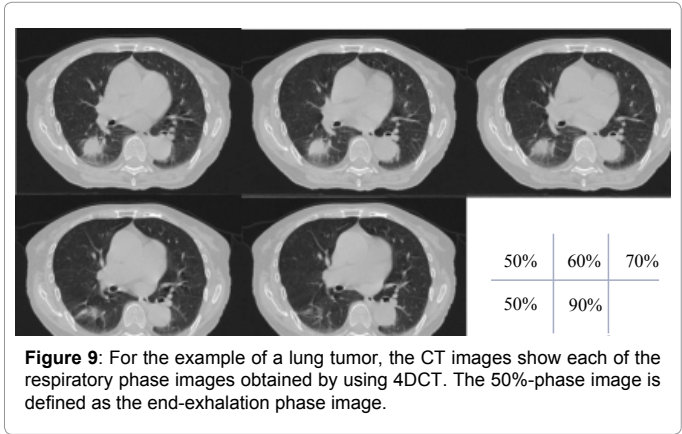
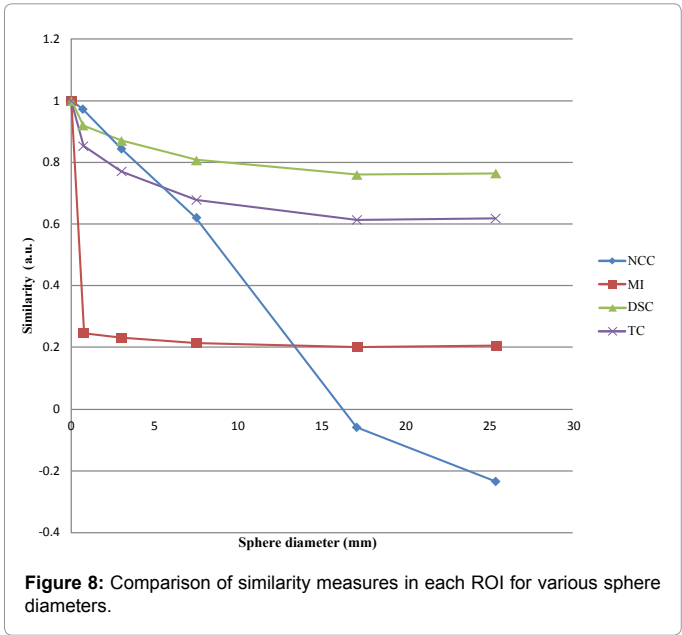


Figure 7: ROIs are defined around each of the Teflon spheres as shown by the rectangular border (yellow) from the first panel of this figure to the last, the sphere diameter changes from 25.4 mm to 0 mm.



10. Schaly B, Kempe JA, Bauman GS, Battista JJ, Van Dyk J (2004) Tracking the dose distribution in radiation therapy by accounting for variable anatomy. *Phys Med Biol* 49: 791-805.
11. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, et al. (1999) Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Trans Med Imaging* 18: 712-721.
12. Venugopal N, McCurdy B, Hnatov A, Dubey A (2005) A feasibility study to investigate the use of thin-plate splines to account for prostate deformation. *Phys Med Biol* 50: 2871-2885.
13. Thirion JP (1998) Image matching as a diffusion process: An analogy with Maxwell's demons. *Med Image Anal* 2: 243-260.
14. Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, et al. (2005) Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Phys Med Biol* 50: 2887-2905.
15. D'Agostino E, Maes F, Vandermeulen D, Suetens P (2003) A viscous fluid model for multimodal non-rigid image registration using mutual information. *Med Image Anal* 7: 565-575.
16. Christensen GE, Rabbitt RD, Miller MI (1996) Deformable templates using large deformation kinematics. *IEEE Trans Image Process* 5: 1435-1447.