

Evaluation of Non-Rigid Image-Registration Algorithms Using Discrepancy Distance Between Organ Contours

Y. Saito^{1,2}, K. Tateoka^{1,3}, K. Shima², T. Nakazawa¹, A. Nakata¹, T. Abe¹, M. Yano¹, K. Fujimoto³, M. Someya¹, K. Nakata¹, M. Hori¹, M. Hareyama³, and K. Sakata¹

¹Department of Radiation Oncology and Medical Physics, Sapporo Medical University, Japan

²National Hospital Organization Hokkaido Cancer Center, Japan

³Proton Therapy Center, Social Medical Corporation Teishinkai

Abstract

Purpose: Non-rigid image registration (NIR) is useful for adaptive radiotherapy. However, no method has been established for evaluating the quality of the algorithms used in NIR. To remedy this situation, we demonstrate herein a novel method to evaluate NIR algorithms.

Methods: We define the NIR error as the discrepancy distance between (i) the organ contours obtained from computed tomography (CT) images acquired during the treatment period (reference contours) and (ii) the contours obtained from the treatment-planning CT images that are constructed by automated propagation during the treatment period (evaluation contours). However, the continuous positional relationship between the points where the reference contour intersects the evaluation contour is assumed to be maintained. In addition, we adapt the proposed method so that it can be applied to the contours of complex organs such as spherical and tubular organs. To demonstrate this method, we measure the contours of the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum. The obtained NIR error presented in two-dimensional (2D) discrepancy maps.

Results: The 2D discrepancy maps show the difference between the reference and evaluation contours from CT images. The proposed method measures the difference between the contours of spherical and tubular organs and evaluates the NIR error based on the positional relationship between the points constituting the contours.

Conclusions: This study accounts for and measures the continuous positional relationship between corresponding points in the contours of complex-shaped spherical and tubular organs with irregularities and evaluates NIR algorithms based on these organ contours.

Keywords: Non-rigid image registration; Three-dimensional quantitative evaluation; Adaptive radiotherapy

Introduction

Intensity-modulated radiotherapy (IMRT), which uses a steep dose gradient, can be administered in high doses to target organs and minimizes exposure to the surrounding normal organs [1]. Currently, IMRT is planned based on treatment-planning computed tomography (pCT) images taken before treatment. However, during the radiotherapy period, a treatment plan based on pCT images does not sufficiently consider the anatomical changes (deformation and displacement) in the target organ and surrounding normal organs. Previous reports suggest that this approach to IMRT exposes patients to a risk of excess or insufficient dose administration to the target and surrounding normal organs because of anatomical changes occurring in the patient during the treatment period [2,3].

Regarding anatomical changes during the treatment period, Michel et al. [4] reported that an appropriate dose can be administered to the target and surrounding normal tissues by using adaptive radiotherapy, in which treatment re-planning is promptly implemented. However, this method requires the radiation oncologist to manually create the contours of the target and normal organs on the repeat CT (rCT) images acquired during the treatment period, which involves considerable time and effort.

In recent years, nonrigid image registration (NIR) has been used to address the issue of anatomical changes that occur during the radiotherapy period. NIR adopts anatomically equivalent voxels between the pCT and rCT images by using deformation vector fields (DVF) obtained from an NIR algorithm. Therefore, tracking

the anatomical displacement, deformation, and cumulative dose in each voxel during the treatment period has now become possible. In addition, NIR automatically propagates (auto-propagation) the displaced and deformed contours created from the pCT images to the new contours on the rCT images. However, if the DVFs obtained from the NIR algorithm are not sufficiently precise, the tracking results for anatomical displacement and deformation and the cumulative dose are inaccurate, as are the auto-propagated contours. Thus, NIR algorithms must be evaluated to ensure their accuracy. However, such an evaluation is reported to be difficult because the exact voxel-to-voxel correspondence between the two images in an NIR algorithm is unknown [5].

A current, widely used method to evaluate NIR algorithms is the dice similarity coefficient [6,7], which uses the volume overlap between the contours of any organ manually drawn by an oncologist (these are the reference contours) and the contours automatically generated by the

***Corresponding author:** Saito Y, Department of Radiation Oncology and Medical Physics, Sapporo Medical University, Japan, Tel: +81-11-611-2111; E-mail: ysaitoh@sapmed.ac.jp

Received December 16, 2015; **Accepted** September 16, 2016; **Published** September 23, 2016

Citation: Saito Y, Tateoka K, Shima K, Nakata A, Nakazawa T, et al. (2016) Evaluation of Non-Rigid Image-Registration Algorithms Using Discrepancy Distance Between Organ Contours. J Nucl Med Radiat Ther 7: 307. doi:10.4172/2155-9619.1000307

Copyright: © 2016 Saito 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.

NIR algorithm by DVFs (these are the evaluation contours). However, determining the precise discrepancy between contours at any given point is difficult with this evaluation method.

Pevsner et al. [8] improved the method of Remeijer et al. [9], which involves measuring the discrepancy between prostate contours on magnetic resonance images and CT images and presenting a three-dimensional (3D) evaluation in which the NIR error is defined as the discrepancy between the reference and evaluation contours. In the method of Pevsner et al. [8], the control direction follows a radial line extending from the center of gravity of the reference contour of the lung (a spherical organ), and in each control direction, the farthest point from the center of gravity is determined where the control-direction line intersects the reference contour (these points are called the reference-contour intersection points). The magnitude of the NIR error is defined as the length of the shortest 3D path from the reference-contour intersection point to the evaluation contour.

Similar to Pevsner et al. [8], Rodriguez-Vila et al. [10] defined the control direction to be radial from the reference point of the reference contour containing the seminal vesicle and prostate (spherical organs), and then determined, in each control direction, the two points where this line intersects the reference and evaluation contours (the latter points are called the evaluation-contour intersection points). They determined the point at the shortest 3D distance from the reference-contour intersection point and the distances from the evaluation-contour intersection point to the evaluation and reference contours. The level of the NIR error is defined as the longest (positive or negative) distance between the two points.

To evaluate NIR algorithms, Rodriguez-Vila et al. [10] and Pevsner et al. [8] used methods based on the shortest distance from any given point on the contour to another point. However, because an organ contour is an ordered set of points, the positional relationship must be maintained between the points on the modified outline contour. Accordingly, a method is sought to evaluate the NIR error that considers the positional relationship between points on the reference and evaluation contours. However, with the methods of Rodriguez-Vila et al. [10] and Pevsner et al. [8], examining complex contours of irregular shapes leads to situations where the positional relationship of the contour points is not maintained.

In addition, currently available methods to evaluate NIR algorithms have used spherical organ contours. Rodriguez-Vila et al. [10], however, proposed a different method to evaluate the contours of tubular organs. Unfortunately, they are yet to report the detailed method and results. In addition, no report has yet appeared detailing a method to evaluate NIR algorithms that is appropriate for contours of both spherical and tubular organs.

To address this shortfall, we propose herein a different method to evaluate NIR algorithms. The method is demonstrated by measuring the size and direction of the discrepancy between reference contours drawn manually by a radiation oncologist and evaluation contours generated automatically by DVFs obtained from NIR algorithms. The proposed method accounts for the continuous positional relationship between corresponding points in the reference and evaluation contours of complex-shaped spherical and tubular organs.

Method

We measure the discrepancy between reference and evaluation contours in the same way as Pevsner et al. [8] and Rodriguez-Vila et al. [10] and evaluate the NIR algorithm. However, the evaluation methods

of Pevsner et al. [8] and Rodriguez-Vila et al. [10] involve spherical contours and do not consider the contours of tubular organs. To adapt the proposed evaluation method to treat contours of both spherical and tubular organs, we create reference points on every contour of each CT slice and determine the discrepancy between the reference and evaluation contours.

Reference contours and evaluation contours

A reference contour is the contour of a given organ on an rCT image. An evaluation contour is the contour of the same organ taken from pCT images and auto-propagated via DVFs onto rCT images. In this study, the contours on the pCT and rCT images were generated by the same oncologist. NIR was performed by initially registering rigid images and then matching the coordinates of pCT and rCT images and deforming and displacing a given organ on pCT images, and then verifying their positions using the rCT images.

Positional relationship of contour points

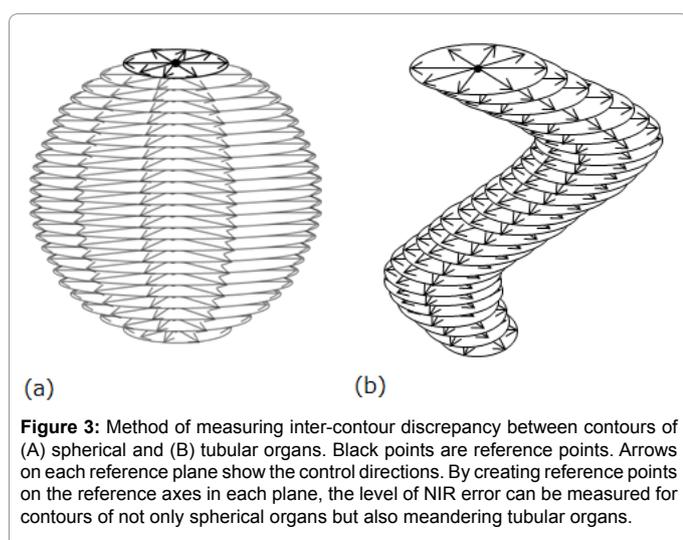
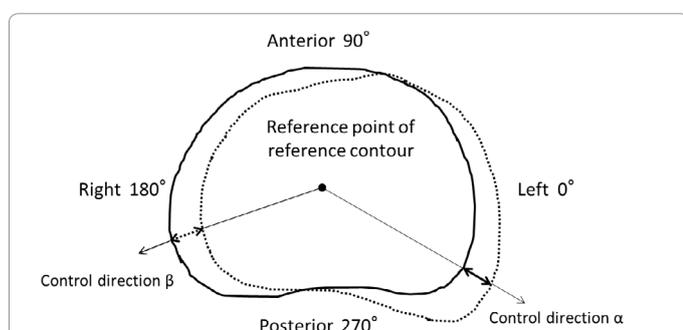
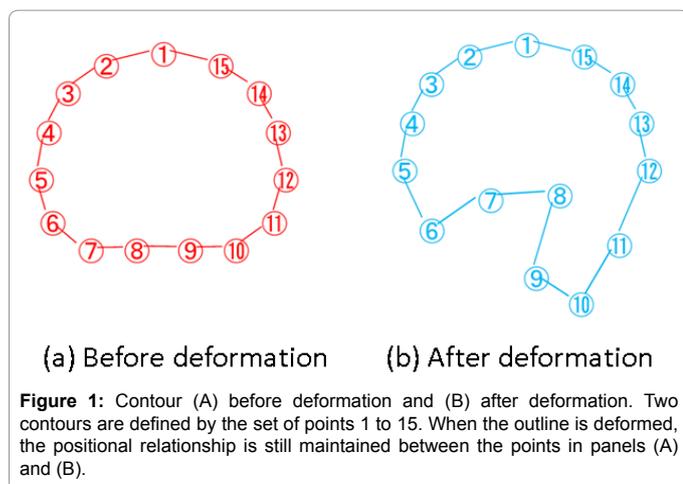
The contour of an organ is considered to be composed of a set of points. Although the shapes of the reference and evaluation contours of a given organ can differ, the positional relationship between the points constituting the contour is assumed to remain unchanged. Figure 1 shows a contour composed of 15 points before and after deformation. Although the shape is deformed, the positional relationship of the points constituting the contour is maintained (in other words, the ordering of the points along the contour is unchanged). To measure the NIR error based on Figure 1, the distance between the same numerals should be measured in panels (A) and (B). Previous methods allow the distances (A) ⑧ to (B) ③ and (A) ③ to (B) ⑧ to be measured, which means that the positional relationship between the contour points is not maintained.

Evaluation of nonrigid image-registration algorithms

The reference and evaluation contours on rCT images obtained as described in section 2.A are treated by software developed in-house in C⁺⁺. The coordinates of the reference and evaluation contours comprise the coordinates of each reference plane (Z coordinate) of each axial slice of the rCT image and the two-dimensional (2D) coordinates of each contour (X, Y coordinates). Subsequently, the in-house software defines the control direction to extend radially (at a given angle) from the center of gravity of the reference contour (i.e., from the reference point) on each reference plane. Next, the software determines the points where the reference and evaluation contours intersect the control direction (i.e., the reference- and evaluation-contour intersection points).

To evaluate NIR algorithms, the NIR-error level is defined as the distance between these reference- and evaluation-contour intersection points. The NIR-error level is defined as positive (negative) if the reference contour is closer to (farther from) the reference point than the corresponding evaluation contour (Figure 2). These definitions are similar to those proposed by Rodriguez-Vila et al. [10] along the reference axis.

However, for tubular organs with meandering contours, such as the rectum, the reference axis is sometimes located outside the reference contour. Therefore, in this study, we create reference points for each reference plane (Figure 3). In the body-axis direction, the reference and evaluation contours may have different sizes, no reference contour or no evaluation contour may appear in the reference plane, irregularly shaped reference and evaluation contours may appear, or the reference and evaluation contours may intersect the control direction at multiple points. In these cases, the calculation methods described below are used.



Reference and evaluation contours with different sizes in body-axis direction

Coordinate transformation is performed by stretching the

evaluation contours in the body-axis direction and then matching the size of the reference and evaluation contours in the body-axis direction. Next, the points where the reference and evaluation contours intersect the control direction are calculated in every reference plane. Finally, the coordinates of the evaluation-contour intersection point are returned to their pre-stretched values, and the distance between the two points (i.e., the level of NIR error) is determined (Figure 4).

Irregularly shaped reference and evaluation contours

In the method of Rodriguez-Vila et al. [10], the level of NIR error is defined as the longest of the shortest distances between contours in each control direction. However, in some situations, this method does not sufficiently maintain a continuous positional relationship between corresponding points on the reference and evaluation contours (Figure 5).

In the proposed method, we calculate the angle with respect to the control direction which multiple intersection points exist with the reference and evaluation contours; the reference and evaluation contours that traverse this angular direction are then split into the same number of uniform-length segments. The positional relationship between the points constituting the reference and evaluation contours can be accounted for by measuring the length of these segments. The level of NIR error is defined as the distance between the points that divide the reference and evaluation contours (Figure 6). An inter-contour discrepancy measurement, which considers the continuous positional relationship between corresponding points on the reference and evaluation contours, is obtained by dividing the reference and evaluation contours.

CT images of the contours of the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum of 15 prostate-cancer patients were used to evaluate NIR algorithms. An Optima 660 CT scanner was used for this study (GE Healthcare UK, Ltd.). NIR was performed using Velocity AI (Velocity Medical Solutions, USA).

Results

Figure 7 shows the reference and evaluation contours of the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum on each reference plane for prostate-cancer patients involved in this study. Figures 8A-8E show the discrepancy between the reference and evaluation contours of the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum, respectively, in a cylindrical projection view (2D discrepancy map).

Figure 8A uses a 2D discrepancy map to show the NIR error between the reference and evaluation contours in the prostate. The largest discrepancy level of +4 mm appears between the boundary region and the caudal rectum ($Z=-10.0$, $\theta=290^\circ$).

The largest measured discrepancy level of -2 mm occurs between the reference and evaluation contours of the right seminal vesicle in the bladder boundary ($12.5 \leq Z \leq 20.0$, $30^\circ \leq \theta \leq 120^\circ$) and in the boundary with the prostate in the lower-right seminal vesicle ($Z=0.0$, $30^\circ \leq \theta \leq 120^\circ$) (Figure 8b).

The largest measured discrepancy level of -2 mm occurs between the reference and evaluation contours of the left seminal vesicle in the bladder boundary ($12.5 \leq Z \leq 20.0$, $60^\circ \leq \theta \leq 150^\circ$) and in the boundary with the prostate in the lower-left seminal vesicle ($Z=0.0$, $60^\circ \leq \theta \leq 150^\circ$) (Figure 8c).

A large negative discrepancy (maximum of -6 mm) occurs

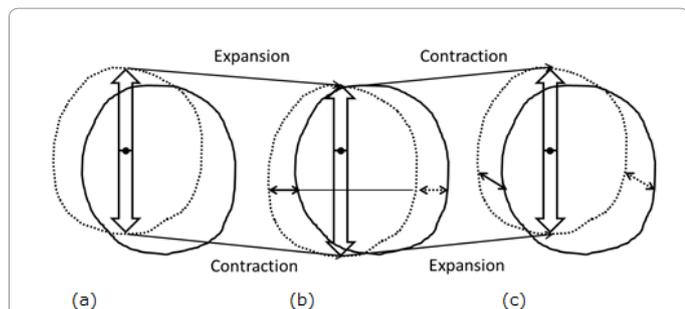


Figure 4: Method of measuring inter-contour discrepancy for different sizes of reference and evaluation contours in the body-axis direction. Solid and dashed lines represent reference and evaluation contours, respectively; black points represent the center of gravity of the evaluation contours in the sagittal plane. (A) Contours of the uppermost and lowermost sides in the body-axis direction of the reference and evaluation contours are not in the same axial slice. Coordinates are transformed by stretching (white arrow) the caudal and cranial sides of the contours using the center of gravity of the evaluation contours (black point) as a reference. (B) The positions of the cranio-caudal sides of the reference and evaluation contours are matched, and the points where the reference and evaluation contours intersect the control direction are calculated using the method shown in Figure 1 (solid line and dashed arrow). (C) Evaluation contours are returned to their original shape, the coordinates of the intersection point of the evaluation contours calculated in panel (B) are recalculated, and the distance between the two points is measured (solid line and dashed arrow).

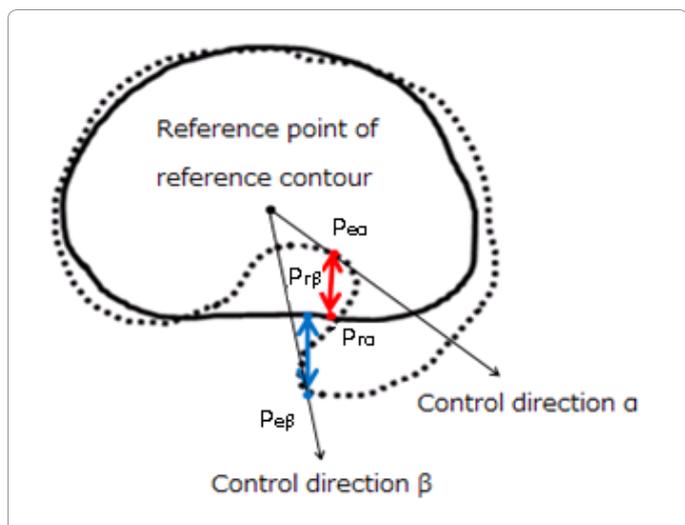


Figure 5: Measuring inter-contour discrepancy for irregularly shaped contours by using the method of Rodriguez-Vila et al. [10] Solid and dotted lines represent the reference and evaluation contours, respectively. Red and blue arrows indicate the two-point distances used to calculate the inter-contour discrepancy in the control directions α and β , respectively. The points where the level of NIR error measured in the control direction α are P_{ra} and P_{ea} on the reference and evaluation contours, respectively. The points where the level of NIR error measured in the control direction β are P_{rb} and P_{eb} on the reference and evaluation contours, respectively. Because the method of Rodriguez-Vila et al. [10] does not consider the continuity of the contours, P_{ra} is to the right of P_{rb} on the reference contours and P_{ea} is to the left of P_{eb} on the evaluation contours.

between the reference and evaluation contours of the bladder, i.e., at the right ventral area of the upper bladder ($62.5 \leq Z \leq 65.0$, $90^\circ \leq \theta \leq 225^\circ$) and the right dorsal area of the medial bladder ($Z=50$, $225^\circ \leq \theta \leq 270^\circ$). However, the discrepancy level was positive at the left and right boundaries of the seminal vesicle ($12.5 \leq Z \leq 20.0$, $225^\circ \leq \theta \leq 315^\circ$) (Figure 8d).

Figure 8e shows a 2D discrepancy map of the discrepancy between

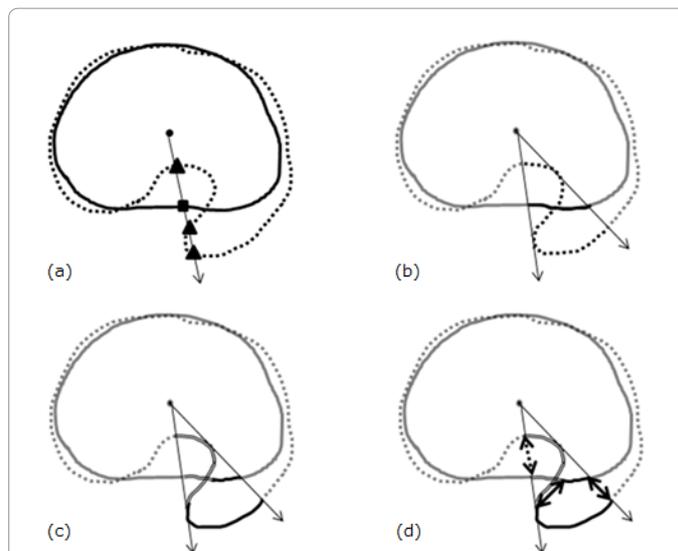


Figure 6: Method of measuring inter-contour discrepancy for irregularly shaped contours. Solid and dashed lines represent the reference and evaluation contours, respectively. The black point represents the center of gravity in the axial plane of the reference contour. (A) The control direction (arrow) intersects with the reference contour at one point (square), whereas it intersects the evaluation contour at three points (triangles). (B) In the angular domain, we select reference and evaluation contours that do not have a one-to-one relationship between the points where the reference and evaluation contours intersect the control direction (bold solid and dashed lines between the two arrows). (C) The reference and evaluation contours selected in panel (B) are each divided into two equal parts (bold and light solid lines between arrows). (D) The level of NIR error is defined as the distance between the points where the reference and evaluation contours are cut by the control directions.

the reference and evaluation contours of the rectum (a tubular organ). The discrepancy level of the prostate boundary area in the caudal rectum ($Z=-10.0$, $\theta=80^\circ$) is -4 mm. The discrepancy level in the 2D discrepancy map between the prostate contours in the same area is +4 mm, as described above.

The NIR errors in the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum have almost the same contour discrepancies as those shown on the CT images in Figure 7 (shown above as the Z coordinate).

Discussion

In this study, we evaluate NIR algorithms by measuring the size and direction of the discrepancy between the reference contours manually drawn by a radiation oncologist and the evaluation contours automatically generated using DVFs for complex-shaped spherical and tubular organs. However, because the positional relationship between the points constituting the reference and evaluation contours is maintained, the proposed approach also accounts for the continuous positional relationship between corresponding points in the two contours.

Figure 8 shows 2D discrepancy maps created from inter-contour discrepancies in the prostate, right seminal vesicle, left seminal vesicle, urinary bladder, and rectum (spherical and tubular organs). These results indicate that the discrepancy between the reference and evaluation contours shown on the CT images in Figure 7 can be precisely measured.

In this study, to evaluate NIR algorithms using the contours of

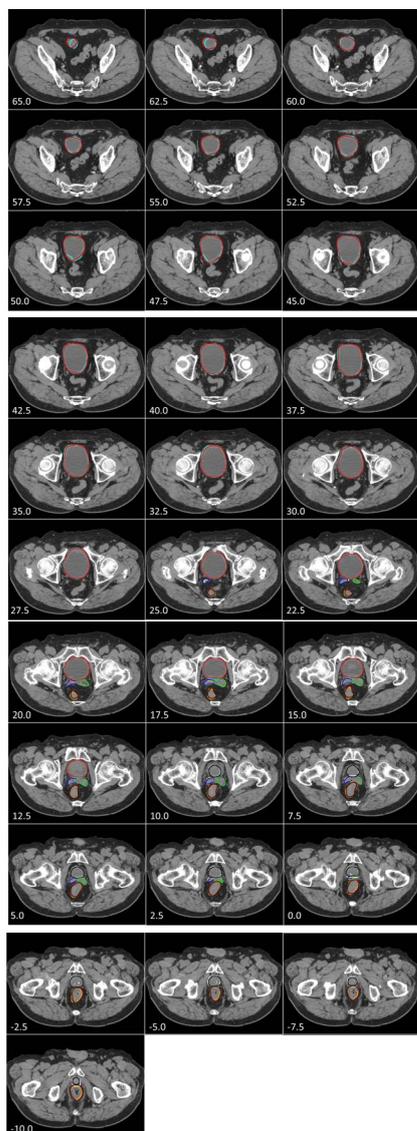


Figure 7: Reference contours (black, green, blue, red, and brown lines) and evaluation contours (white, magenta, yellow, cyan, and orange lines) of the prostate, left seminal vesicle, right seminal vesicle, urinary bladder, and rectum. The bottom-left number is the Z coordinate in the cranio-caudal direction.

spherical organs and meandering tubular organs such as the rectum, we determine on each reference plane a control direction that extends radially from the center of gravity (i.e., the reference point) of the reference contour. Finding the points where the reference and evaluation contours intersect the control direction is useful for evaluating NIR algorithms. In addition, because the spherical organs used in this study did not have complex shapes, we expect to obtain results similar to those obtained for the method of Rodriguez-Vila et al. [10] and no meaningful discrepancies with previous reports.

However, if the reference and evaluation contours have highly irregular shapes, discrepancies may appear between the results of the method of Rodriguez-Vila et al. [10] and those of the method proposed herein. Figure 9 shows the points used for measuring the discrepancy between the reference and evaluation contours in the

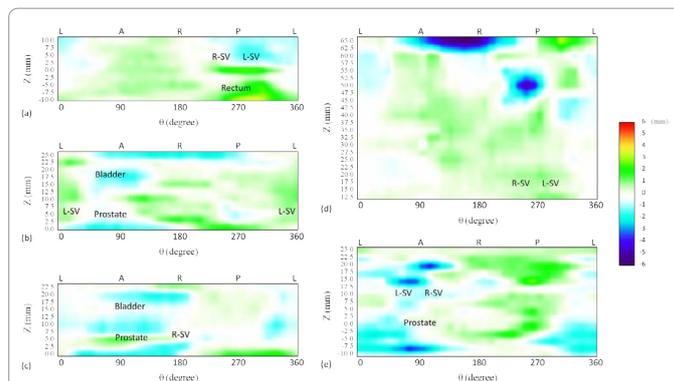


Figure 8: 2D discrepancy maps of (A) prostate, (B) right seminal vesicle, (C) left seminal vesicle, (D) bladder, and (E) rectum. False-color discrepancy levels are given by the color bar in the right. The horizontal axis is the control-direction angle θ and the vertical axis is the cranio-caudal distance (Z coordinate), where 0 is the center of the prostate reference plane.

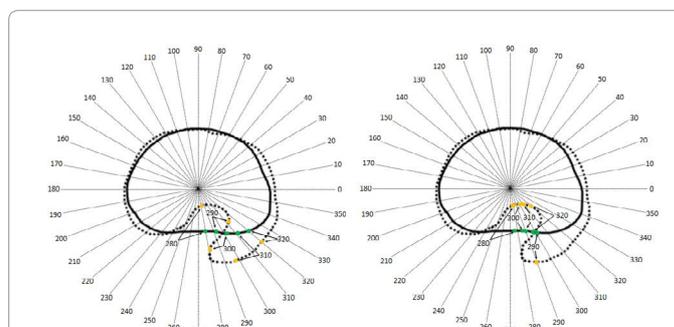


Figure 9: Comparison of points used to measure inter-contour discrepancy on the reference contours (solid lines) and evaluation contours (dotted lines) of organs with irregular shapes using the proposed method (left) and the method of Rodriguez-Vila et al. [10] (right). Green and yellow dots indicate the points used to calculate the reference and evaluation contours, respectively, and their control directions are indicated by arrows.

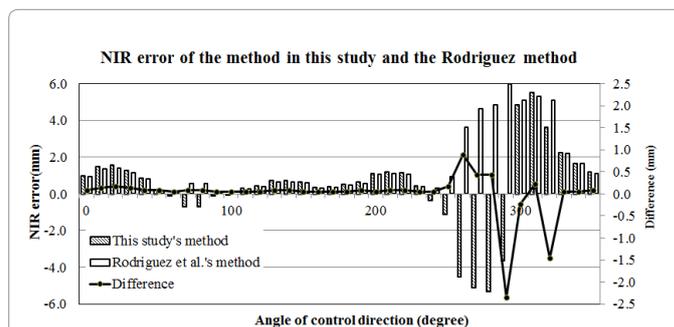


Figure 10: Level of NIR error in each control direction measured by the proposed method and by the method of Rodriguez-Vila et al. [10].

method of Rodriguez-Vila et al. [10] and in the proposed method (i.e., the inter-contour discrepancy points). Figure 10 shows the discrepancy levels measured between the contours in Figure 9 by using both the method of Rodriguez-Vila et al. [10] and the proposed method. For control directions from 260° to 320° , a large difference exists between the discrepancy levels obtained by the method of Rodriguez-Vila et al. [10] and those obtained by the proposed method; however, almost no other differences appear.

Except between 280° and 320°, the reference and evaluation contours intersect each control direction only at a single point. However, because the evaluation contours are irregularly shaped in the control directions from 280° to 320°, multiple intersection points exist in these regions. The method of Rodriguez-Vila et al. [10] measures the shortest distance between contours; therefore, to evaluate irregular contours such as those shown in Figure 9, the inter-contour discrepancy points from 280° to 320° on the reference contours and those from 280° and 300° to 320° on the evaluation contours are concentrated into a single small area. Consequently, unmeasured contours exist in irregular areas. In addition, because the shortest distance between contours is measured, the positional relationship of the points in the reference and evaluation contours is not considered, e.g., for the inter-contour discrepancy point at 290° on the reference and evaluation contours (Figure 9).

In the proposed method, however, sections of contours that intersect the control direction multiple times are divided into equal intervals, and the distance between the dividing points on these contours is calculated. Thus, as shown in Figure 9, the points used to measure the inter-contour discrepancy between the reference and evaluation contours from 280° to 320° appear at equal intervals. In addition, by dividing the contours into equal intervals, a continuous positional relationship between corresponding points on the reference and evaluation contours is also maintained in contours with irregular areas. Thus, at the 290° and 320° control points, a discrepancy of approximately 2 mm exists between the results obtained with the method of Rodriguez-Vila et al. [10] and those obtained with the proposed method.

However, conventional methods for evaluating NIR algorithms use the shortest distance between contours, which is a scalar quantity. In the proposed method, the discrepancy direction is chosen by establishing the control direction to extend radially from the reference point of the contour. This approach allows the size and direction of the error in the NIR algorithm to be evaluated within the target and critical organ boundary when conducting radiotherapy with a steep dose gradient.

Conclusion

We propose herein a new method to evaluate NIR algorithms that should facilitate optimization of IMRT. The method measures the discrepancy between the reference and evaluation contours of complex-

shaped spherical and tubular organs with irregularities. In addition, the proposed method accounts for the continuous positional relationship between corresponding points in the reference and evaluation contours.

Acknowledgments

We express our appreciation to Professor Jun Takada and Assistant Professor Kenichi Tanaka of the Radiation Protection Laboratory of Sapporo Medical University for useful discussions on this manuscript. We also thank the radiation oncologists of the Hokkaido Cancer Center and radiological technologists of the Hokkaido Cancer Center for providing us with excellent assistance and advice.

References

1. Hunt MA, Zelefsky MJ, Wolden S, Chui CS, LoSasso T, et al. (2001) Treatment planning and delivery of intensity-modulated radiation therapy for primary nasopharynx cancer. *Int J Radiat Oncol Biol Phys* 49: 623-632.
2. Huang E, Dong L, Chandra A, Kuban DA, Rosen II, et al. (2002) Intrafraction prostate motion during IMRT for prostate cancer. *Int J Radiat Oncol Biol Phys* 53: 261-268.
3. Chow JC, Jiang R, Markel D (2009) The effect of interfraction prostate motion on IMRT plans: a dose-volume histogram analysis using a Gaussian error function model. *J Appl Clin Med Phys* 10: 79-95.
4. Michel G, Di Y, Alvaro M (2010) Adaptive radiation therapy for prostate cancer. *Semin Radiat Oncol* 20: 130-137.
5. Christensen GE, Geng X, Kuhl JG, Bruss J, Graboswki TJ, et al. (2006) Introduction to the non-rigid image registration evaluation project (NIREP). *Biomed Image Regist* 4057: 128-135.
6. Fallone BG, Rivest DR, Riauka TA, Murtha AD (2010) Assessment of a commercially available automatic deformable registration system. *J Appl Clin Med Phys* 11: 101-123.
7. Zambrano V, Furtado H, Fabri D, Caucig LC, Góra J, et al. (2013) Performance validation of deformable image registration in the pelvic region. *J Radiat Res* 54: i120-i128.
8. Pevsner A, Davis B, Joshi S, Hertanto A, Mechalakos J, et al. (2006) Evaluation of an automated deformable image matching method for quantifying lung motion in respiration-correlated CT images. *Med Phys* 33: 369-376.
9. Remeijer P, Rasch C, Lebesque JV, van Herk M (1999) A general methodology for three-dimensional analysis of variation in target volume delineation. *Med Phys* 26: 931-940.
10. Rodriguez-Vila B, Gaya F, Vicente GF, Gomez EJ (2010) Three-dimensional quantitative evaluation method of non-rigid registration algorithms for adaptive radiotherapy. *Med Phys* 37: 1137-1145.