

An Image Registration Method with Radial Feature Points Sampling: Application to Follow-Up CT Scans of a Solitary Pulmonary Nodule

Masaki Ishihara^(✉), Yuji Matsuda, Masahiko Sugimura, Susumu Endo, Hiroaki Takebe, Takayuki Baba, and Yusuke Uehara

Fujitsu Laboratories Ltd., 1-1, Kamikodanaka 4-chome,
Nakahara-ku, Kawasaki 211-8588, Japan

{ishihara-masaki,matsuda-yuji,sugimura,yuehara}@jp.fujitsu.com,
endou.susumu-02,
takebe.hiroaki,baba-t,yuehara}

Abstract. In order to support radiologists' follow-up task of two CT scans captured in the past and in the present, we aimed to develop a system that displays both a region of interest (ROI) in one image selected by a radiologist and a corresponding ROI in another image. In this paper, we propose a registration method for the system. A typical registration method identifies several pairs of matched feature points (i.e., matching pairs) between two images within the range of a predefined distance from the ROI's center point (i.e., interest point) to correct a positional shift of an organ caused by heartbeats and breathing. However, low accuracy of registration is often observed because of biased distribution or a small number of matching pairs, depending on the sampling range. We developed a novel registration method that radially and evenly searches for several nearest matching pairs around the interest point and then estimates a translation vector at the interest point as a weighted average of these nearest pairs using a weighting factor based on its distance from the interest point. This method was based on the assumption that the transformation of an interest point work with the transformation of a near point since the lung is a continuum. The results of a comparative evaluation of the existing method and the proposed method on the basis of 15 cases showed that the accuracy of the proposed method was higher than that of the existing method in 13/15 cases. We analyzed the association between the accuracy and the range of sampling and found that the accuracy of the proposed method was similar to the best performance of the existing method with an ideal range of sampling the matching pairs. Finally, we showed evidence that the new method was reasonably consistent in terms of giving the best performance.

Keyword: Image registration

1 Introduction

It is important to monitor temporal changes in the solitary pulmonary nodule (SPN) to enable optimum diagnosis and follow-up treatment. Comparative

reading of medical images by radiologists is indispensable. In this paper, the term ‘tumor’ is used to describe the SPN, and ‘follow-up CT scans’ is used to describe the comparative reading task.

It is now possible to visualize each tumor more distinctively through advances in X-ray Computed Tomography (CT) scanning equipment. Accordingly, along with the increase of medical images required for diagnosis, the burden of radiologists has become an issue. For radiologists to expedite diagnosis, computer aided diagnosis (CAD) systems are required. In this paper, we aim to develop a support system for radiologists’ follow-up CT scans (such as that in Fig. 1) that satisfies the following requirements:

1. Display of multiple CT images side-by-side.
2. Display of close-up of both of a selected region by a radiologist and the corresponding region in another image.
3. No image deformation except translation and expansion.

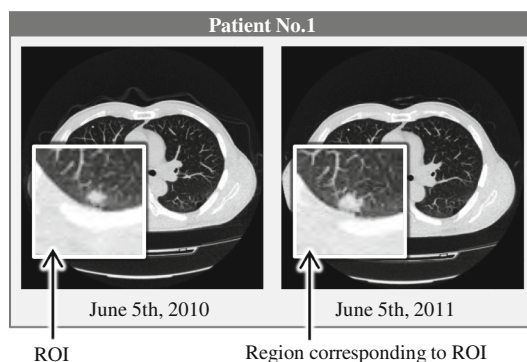


Fig. 1. A proposed support system for follow-up CT scans.

First requirement is needed to compare two CT images captured in the past and present in order to confirm the temporal change in tumor development. At this time, a positional shift of an organ occurs owing to heartbeats, breathing, and body position during CT scanning. For this reason, registration of the same tumor requires an understanding of temporal changes. Second requirement is needed to confirm in detail the change of the tumor and its surroundings. Therefore, the purpose of this paper is not registration of whole CT images, but registration of a region of interest (ROI) selected by the radiologist. Third requirement is needed since incorrect deformation impedes successful diagnosis. Hence, we are not concerned with the deformable image registration or the rigid registration which includes rotation. A simple method of rigid registration which includes translation and expansion in order to align a center of tumor is needed.

Since there are many types of lung cancer, efficient follow-up CT scans is needed. For this reason, we have chosen to target chest CT images. In the lung

cancer screening, registration error should be less than 2.5 mm, since tumors larger than 2.5 mm in radius are considered therapeutic objectives [1]. Hence, we set the target of each registration error below 2.5 mm. In consideration of waiting time, processing time is required to be less than two seconds.

The image registration methods can be roughly divided into two approaches, the intensity based approach and the feature based approach [2]. In the intensity based approach [3, 4], an image deformation that maximizes the similarity (like normalized cross correlation (NCC)) with iterative operation is computed. Therefore, convergence of iterative operations tends to take a long time. In the feature based approach [5, 6], the method detects feature points from two images to be compared and extracts image descriptors in each feature point. Then, the method computes the correspondence of feature points (called a matching pair), which maximizes the similarity of descriptors. Finally, the method obtains a registration result using a matching pair. Xia et al. [5] used phase congruency, and Yang et al. [6] used a salient region as a feature point detector. They commonly used SIFT as a feature descriptor.

The feature based approach is advantageous in the processing speed compared with the intensity based approach. Thus, to meet the target of processing speed, we employ the feature based approach.

2 Outline of a Support System for Follow-Up CT Scans

An outline of a support system for follow-up CT scans is described in this chapter. The procedures of our system are shown in Fig. 2.

First of all, input data of our system are two series of CT images captured in the present and the past and a ROI on a slice image of the present CT images. The ROI means information of a position and size of the tumor, which a user needs to indicate manually. Then three procedures (feature point detection, feature description, and matching) are executed in sequence. As a result, matching pairs of images are obtained for each pair of the slice image of the present and one of several slice images of the past.

The system performs image alignment in two steps by using the obtained matching pairs. First, vertical (superior-inferior) alignment results in finding the slice image of the past corresponding to the slice image of the present. In vertical alignment, a slice image of the past that has the maximum similarity is searched for. Specifically, similarity between slice images is calculated as a number of matching pairs that have higher similarity than a threshold. Next, two-dimensional horizontal (left-right or anterior-posterior) alignment is achieved for two corresponding slice images of the present and the past from their matching pairs. This alignment results in finding a vector of parallel translation between the two images.

Finally, the system outputs the position and size of the region in the past CT images corresponding to the ROI in the present CT images. Our system can provide zooming views of these regions to a user as shown in Fig. 1.

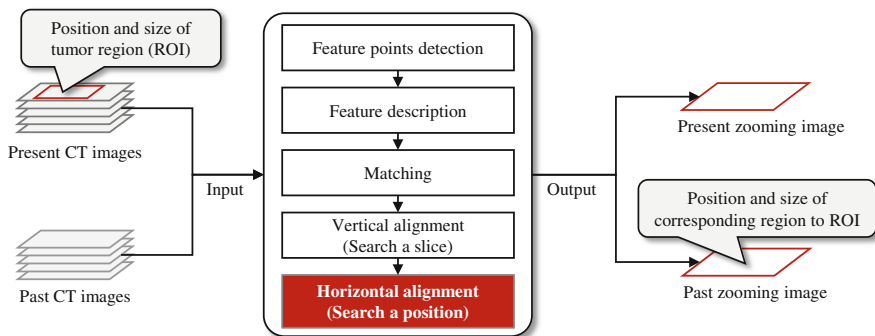


Fig. 2. System architecture of our system.

3 Existing Method of the Feature Based Image Registration and Its Problem

In this chapter, we show an existing method of the feature based image registration and its problem.

3.1 Existing Method of Feature Based Image Registration

In this paper, we define registration as obtaining a matching pair between an organ indicated by an interest point in one image and the same organ in another image. The interest point means a center point of a ROI selected by a radiologist. However, when the interest point is in a tumor, the detection of a matching pair often fails. Since there are temporal changes in the tissues between a pair of images, a matching pair of points is often not found.

Therefore, we assume that the matching pairs in the interest point can be estimated by interpolation of the other matching pairs of the feature points around the interest point [5]. On the basis of this assumption, our method selects several matching pairs of feature points in a sampling region centered at the interest point, and then aggregates these matching pairs as a vector, such as a mean vector (Fig. 3). In order to differentiate an estimated matching pair in the interest point from the other matching pairs around the interest point, we call this estimated relation a translation vector.

3.2 Problems of the Existing Method

In the pilot study, the performance of the existing method was evaluated with several sample cases described below. As a result, it was found that the registration error was 4.4 mm. This result means that the existing method was not acceptable since it did not satisfy the performance requirement (< 2.5 mm). The cause of the problem related to registration errors was analyzed. It was reasonable to suppose that the translation vector was obtained by aggregating several

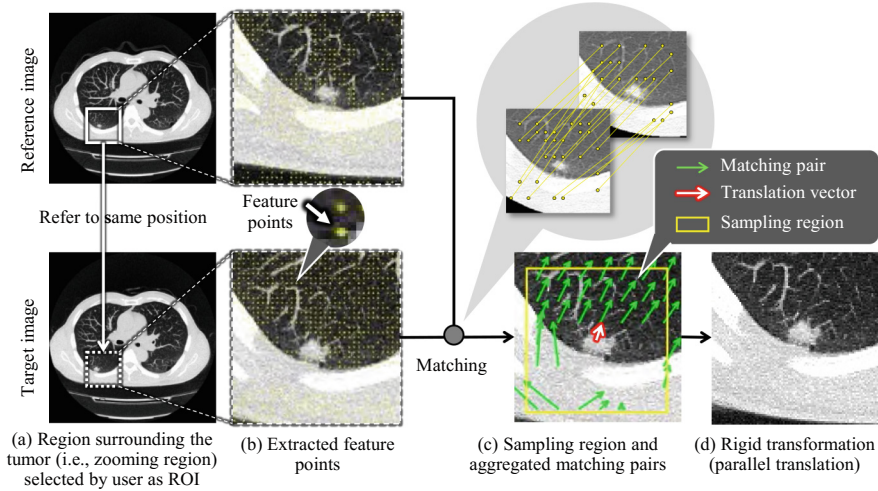


Fig. 3. Illustration of image registration based on feature points matching.

matching pairs in the sampling region, such as a rectangle or a circle range of a predefined distance from the interest point. When the obtained translation vector was applied to the registration of a user's selected region, several large errors could be observed depending on the region.

Several reasons for this registration error are described below. One of the reasons was noise errors due to matching pairs not being distributed equally in the lung region. In a similar way, these pairs were also not equally distributed in the sampling region. Therefore, if only a few matching pairs were extracted in the sampling region, the translation vector was susceptible to several incorrect pairs (i.e., noises) included in these pairs (Fig. 4(b)). Another reason was biased errors. If the distribution of the matching pairs in the sampling region was biased, the translation vector was dominantly influenced by several localized pairs (Fig. 4(c)).

These two errors resulted in a wrong translation vector, leading to the registration error. To estimate the correct translation vector, it was found suitable to sample and aggregate the matching pairs using the following two terms:

1. Sampling a sufficient number of matching pairs near the interest point.
2. Sampling evenly distributed matching pairs near the interest point.

4 Proposed Method

In this chapter, we propose a novel sampling method for matching pairs that is robust against bias distribution of the feature points around the interest point. First, regions in the lung that are suitable for feature point matching are explained in Sect. 4.1. Then, our proposed method is described in Sect. 4.2.

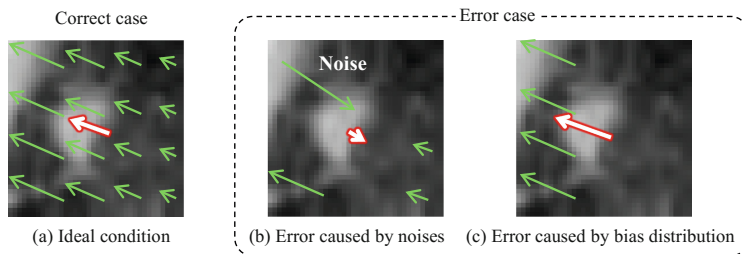


Fig. 4. Illustration of some problems of the basic method.

4.1 Selecting Feature Points






We here explain the method to extract feature points and match these points between pairs of images. It is necessary to extract feature points as evenly and minimally as possible, since the calculating cost for matching increases exponentially with the number of these points. Therefore, these points are extracted at regular intervals (i.e., dense sampling) in the range of a predefined distance from the interest point. This will enable us to control the number and the distribution of feature points. Moreover, several points that would obviously yield only low similarity in the matching process must be eliminated from the candidates of feature points, since these low similarity points often result in noise errors or unfound pairs.

According to our early analysis, we can assume that a feature point suitable for matching satisfies the following requirements:

1. There is a sharp contrast region near the feature point (i.e., characterized point by a pattern of luminance change).
2. There are few feature points similar to the feature point in the image (i.e., matching error-resistant point).
3. The feature point does not change much temporally between two images captured in the past and present (i.e., the feature point can be extracted stably over time).

We show some examples of typical areas in CT scan images and summarize the relationship between the areas and the above three requirements in Table 1. A bright line in a dark region is acceptable for all of the requirements and indicates horizontal vessels or bronchial tubes in a lung region on a slice image. In order to extract these tissue regions, edge detection is first applied to a slice image. We apply the Canny edge detector. Next, only points near the edge are extracted from dense points, and the other points are eliminated. After extraction of feature points, a luminance change pattern is extracted from each region near these points as a local feature. We adopted a binary descriptor BRIEF [7] as the local feature, which is used to calculate similarity in the matching process.

Table 1. Examples of several visual patterns near feature points in terms of each suitable requirement.

#	Visual patterns of feature point	Intensity gradient	Multiple similarity	Temporal change
1	 Dark and flat (e.g., Air region in lung)	False	False	False
2	 Bright and flat (e.g., Adipose tissue)	False	False	False
3	 Local bright point in dark region (e.g., Vertical vessels or bronchial tubes)	True	False	False
4	 Local dark point in bright region (e.g., Hole in adipose tissue)	True	False	False
5	 Bright line in dark region (e.g., Horizontal vessels or bronchial tubes)	True	True	True

4.2 Sampling Matching Pairs

An important part of our proposed method is a sampling method. We supposed that the method of sampling matching pairs satisfied the 1st term as described in Sect. 3.2. Setting a large sampling region is reasonable in order to sample as many matching pairs as possible. However, if a translation vector is estimated from the large sampling region, the assumption that the matching pair of the interest point is identical to the aggregation of the other pairs in the region is not true. In this case, the translation vector averaged all of the pairs in this region, causing a registration error. We should consider that the transformation of the interest point works with the transformation of the near point since the lung is a continuum. Therefore, it is reasonable to assume that the translation vector should be strongly reflected by the pairs near the interest point. For example, we adopted a weighted average according to the distance from the interest point to the nearest feature point.

Next, we suppose that the method of sampling matching pairs satisfies the 2nd term as described previously. If two matching pairs are observed, a matching pair (i.e., a translation vector) of any point on the line through the two points is estimated by a linear interpolation. We can obtain the translation vector in the interest point by sampling at least two nearest matching pairs. One pair is in one direction and the other pair is in an opposite direction as viewed from the interest point. Ideally, if matching pairs in all directions viewed from the interest point are evenly sampled, the translation vector can be estimated precisely by these pairs.

Based on the above concept, we formulated a method for calculating the translation vector v_t :

$$\begin{cases} v_t = \sum_{i=0}^{N-1} v_i \cdot w_i \\ w_i = r_i^{-1} \left(\sum_{j=0}^{N-1} r_j^{-1} \right)^{-1} \end{cases} \quad (1)$$

In this formulation, v_i indicates the nearest matching pair from the interest point and w_i indicates a weight for v_i . The weight w_i is the inverse of the distance r_i from the interest point to a source point of the nearest matching pair, and they are normalized from 0 to 1 (i.e., $\sum_{i=0}^{N-1} w_i = 1$). This formulation is defined in a polar coordinate system as the interest point is the origin. The upper bound of summation N is a predefined division number of the sampling region.

The sampling method of v_i is described in detail below. First, in a reference image, a radiologist selects a region for zooming and registration. The center point of the ROI is defined as the interest point (Fig. 5(a)). Next, feature points in the range of a predefined distance R_{max} from the interest point are detected and matched. After matching, matching pairs are obtained. The predefined distance is defined as the maximum range of motion of a lung. Next the exact region is radially and evenly divided to N sub-regions around the interest point (Fig. 5(b)). The nearest matching pairs v_i are sampled from each of these sub-regions (Fig. 5(c)). If v_i is not obtained in a sub-region, the v_i is not taken into formulation 1 (i.e., $v_i = 0, r_i^{-1} = 0$).

Based on the above procedure, we formulated a method for calculating the sampled matching pair v_i :

$$\left. \begin{array}{ll} \text{given} & N, R_{max} \\ \text{find} & v_i(r_i, \theta_i) \\ \text{that minimizes} & r_i \\ \text{subject to} & 0 \leq r_i < R_{max}, \\ & \frac{2\pi}{N}i \leq \theta_i < \frac{2\pi}{N}(i+1), \\ & i = 0, 1, \dots, N-1 \end{array} \right\} \quad (2)$$

The formulation 2 is defined in a two-dimensional polar plane for the purpose of illustration.

5 Experiments

To evaluate the performance, we compared registration errors by two methods. One is a method of estimating the translation vector using a fixed range sampling (called existing method); the other is our proposed method.

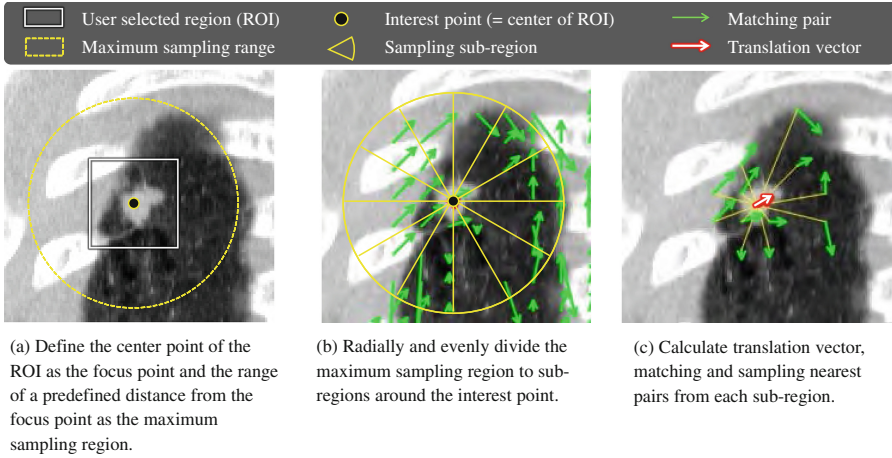


Fig. 5. Illustration of the sampling algorithm of the proposed method.

5.1 Evaluation Data and Evaluation Method

Chest CT images of five lung cancer patients were used. Fifteen cases in total, consisting of three planes (axial, coronal, and sagittal) for each patient, were used. A case indicates a pair of CT image series, such as previous and current ones, for comparative reading. All cases included tumors in both images for comparison. Capture intervals between images could be several months to several years (about one year on average).

We assumed that a radiologist selects two slice images from two CT image series and selects a region surrounding the tumor as an ROI in a reference image (i.e., previous image). That is, the vertical alignment is done manually due to evaluate only performance of the horizontal alignment. Next, our system provided a zoom view of the regions in both the reference image and the target image (Fig. 1). The accuracy of the registration was evaluated for the local region in the target image corresponding to the ROI in the reference image. We applied feature based registration with the proposed method and the existing method to these CT images. On the assumption of actual usage, the interest point was set to the centroid of the tumor. Table 2 shows details of CT images and the parameters of our proposed method.

First, the original CT images are converted to 16-bit gray scale images. Then each slice image is generated by the maximum intensity projection (MIP) of ± 10 slice images. Since MIP image provides the continuous structures such as the blood vessels or the bronchial tubes, the result of edges detection is effectively used for filtering the feature points. Note that before sampling of matching pairs, we applied a method for reducing noise pairs. First, we divided the image into blocks of fixed intervals (e.g., 12×12 pixels). We then aggregated these pairs in every block with a 2D histogram, consisting of a distance and an angle made by the matching pairs. The mode of each block was given from the 2D histogram.

We defined the mode as a matching pair. The ground truth for the experiment was obtained manually by registration of the centroid of the tumor on the basis of a radiologist’s annotation. The registration error was evaluated using Euclidean distance between an estimated coordinate and a ground truth coordinate.

5.2 Experimental Results

Figure 6 shows registration errors by the existing method and by our proposed method. An average error is shown in the bar chart, and minimum and maximum errors are shown in the bar. In this experiment, the sampling range for the existing method was changed from 10 to 250 pixels at intervals of 10 pixels with the interest point in the center.

Our proposed method achieved lower registration error than the average error by the existing method in 13/15 cases. Moreover, in nine cases, our proposed method achieved a lower registration error than 2.5 mm, which is our target accuracy. The processing time of our proposed method was 0.7 s in total. The processing time of the existing method is similar to the time of our proposed method. The processing time includes 0.5 s for feature extraction and matching and 0.2 s for aggregation of matching pairs and image display. This processing speed satisfies the required performance. For this experiment, a Xeon 3.5 GHz machine with 8 GB RAM was used.

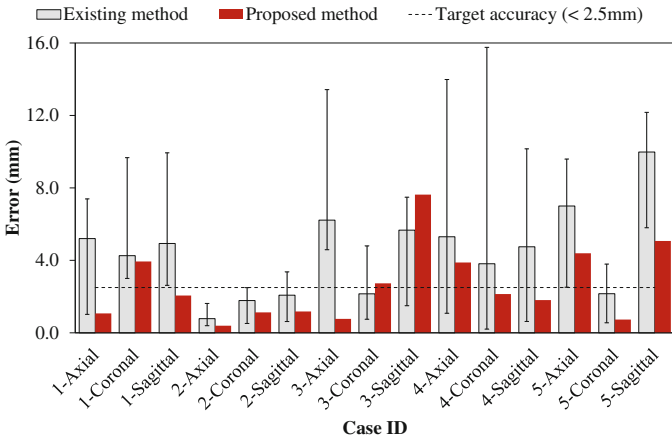


Fig. 6. Registration results of the proposed method and the existing method. The horizontal axis means case ID, and the vertical axis means registration errors. The dashed line means the upper limit of acceptable registration error (2.5 mm). The error bar means the registration error range of trials with gradually changed sampling regions. Therefore, the lower limit of error bar indicates the best performance, and the bar chart of existing method means the average of these trials.

5.3 Discussion

First of all, we discuss the availability of our proposed method. Figure 7 shows the registration errors of case ID 1-Axial, 1-Coronal, and 1-Sagittal by the existing method with varying sampling ranges. If the sampling range was narrow, the registration error tended to increase due to noises and deviations of the distribution of matching pairs. On the other hand, even if the sampling range was

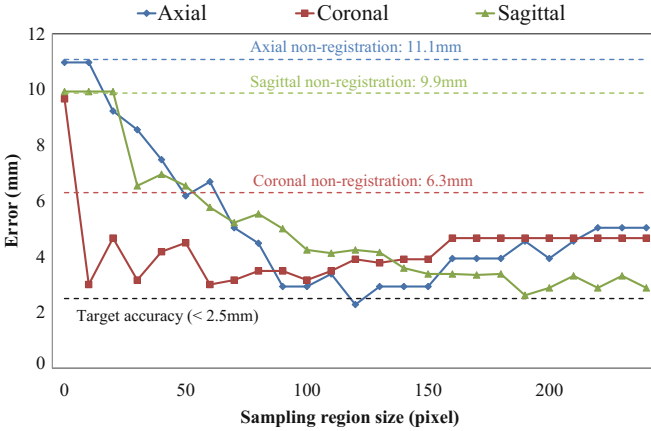






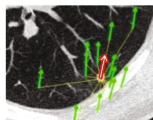




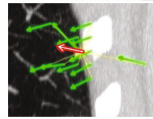
Fig. 7. The registration error of the existing method with varying sampling ranges in case ID 1-Axial, 1-Coronal and 1-Sagittal. The vertical axis shows the registration error and the horizontal one rectangle size of sampling range. The broken line shows the initial error of each plane and our target accuracy (< 2.5 mm).

Table 2. Experimental conditions.

A. Capture conditions	
Image size	512 × 512 pixels
Pixel spacing	0.781–0.625 mm/pixel
Slice thickness	1 mm
Patient position	Supine position
B. CT image reconstruction conditions	
Gray level range	16-bit (0–65,535)
Reconstruction method	Maximum intensity projection (±10 slice images)
C. Feature point detection conditions	
Sampling method	Dense sampling (3 × 3 pixels)
Filtering points	Around edge and inside of ROI
D. Sampling matching pairs conditions	
Sampling range (radius)	40 pixels
Angle interval	30 degrees

wide, the registration error tended to increase due to loss of locality. Figure 7 exactly shows some problems of the existing method described as in Sect. 3.2. Thus, highly accurate registration required an optimal sampling range for the registration.

Table 3. Several sample images of cases 3-Axial and 3-Sagittal, and registration results of the proposed method and the existing method. The far-right images show several matching pairs (green thin arrows) and the translation vector (red bold arrow) obtained by the proposed method.

CaseID	(a)Nodule region image (50×50 pixels)		(b)Diff image between reference and registered target image (50×50 pixels)		(c)Sampling range and vectors
	Reference	Target	Existing	Proposed	
3-Axial					
3-Sagittal					

We compared the difference in the sampling range, which minimizes the registration error in each case. Each plane's histograms that aggregated the optimal sampling range in all cases are shown in Fig. 8. As seen in Fig. 8, there was no significant trend in the sampling range. The optimal sampling range depended on the case. As a result, we confirm that the performance of the proposed method in terms of automatically estimating the optimal sampling range is as good as that of the existing method.

Second, we discuss the advantages and disadvantages of our proposed method with several typical examples. The advantages are explained with case 3-Axial. As seen in Table 3(c) 3-Axial, the proposed method samples a wider range on the side of chest cavity than on the side of chest wall. There are not many vessels or bronchi as landmarks around the interest point in this case. Since the wide range of sampling on the side of chest cavity contributes to search for these landmarks, the registration performance of our proposed method is better than that of the existing method.

Next, the disadvantages are explained with case 3-Sagittal. As seen in Table 3(c) 3-Sagittal, there is mismatching pair that has a very different direction than the other sampled matching pairs. We also analyze other cases that did not achieve the target accuracy. As the result, we found that the mismatching often

occurs at the area near the chest wall. Since the edge appears at the boundary of the lung, feature points are extracted in the proposed method. Features extracted from the boundary of the lung are similar to each other and cannot be obtained with unique correspondence. In case 3-Sagittal, we consider that the mismatching pair is strongly reflected in the translation vector since the interest point (i.e. the position of tumor) is near the chest wall.

To solve the problem of mismatching near the lung walls, it is assumed that the registration error is reduced using the corner point detector instead of the feature point near the edge. Thus, we consider integrating a corner point detector such as FAST [8]. In addition, the movable range of the organization within the lung due to heartbeat and breathing depends on the location [9]. Registration errors are expected to be reduced by limiting the range of feature matching using this knowledge.

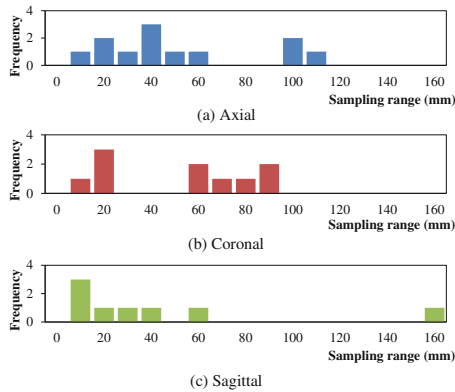


Fig. 8. The frequency of sampling ranges that minimize the registration error in each case. The interval of sampling range is set to 10 mm.

6 Conclusion

In order to support radiologists' follow-up task of two CT scans captured at two different time points, we aimed at developing a system that displays both the ROI in one image selected by a radiologist and the corresponding region for this particular ROI in another image. In this paper, we propose a registration method for the system.

A typical registration method identifies several pairs of matched feature points between two images to correct the positional shift of organs caused by heartbeat and breathing. To determine the vector for transformation of the ROI, the existing method samples several matching pairs in the range of a predefined distance from the interest point. However, low accuracy of registration is often observed due to biased distribution or a low number of matching pairs, depending on the sampling range.

We developed a novel registration method that radially and evenly searches for several nearest matching pairs around the interest point and subsequently estimates a translation vector at the interest point as a weighted average of these nearest pairs using a weighting factor based on its distance from the interest point. The results of comparative evaluation of the existing method and our proposed method using 15 cases showed that the accuracy of our proposed method was better than the accuracy of the existing method in the 13 cases, and the registration error in the nine cases was less than 2.5 mm, which is our target accuracy. We analyzed the association between the accuracy and the range of sampling, and found the accuracy of our proposed method was similar to the best performance of the existing method with an ideal sampling range. Finally, we showed that the proposed method was reasonably consistent in terms of giving a stable performance.

References

1. Guidelines for the management of pulmonary nodules detected by low-dose CT lung cancer screening version 3. The Japanese Society of CT Screening (2013)
2. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: a survey. *IEEE Trans. Med. Imaging* **32**(7), 1153–1190 (2013)
3. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* **18**(8), 712–721 (1999)
4. Werner, R., Hermann, S.: High accuracy optical flow for 3D medical image registration using the census cost function. In: Klette, R., Rivera, M., Satoh, S. (eds.) *PSIVT 2013*. LNCS, vol. 8333, pp. 23–35. Springer, Heidelberg (2014)
5. Xia, R., Zhao, J., Liu, Y.: A robust feature-based registration method of multimodal image using phase congruency and coherent point drift. In: *MIPPR 2013: Pattern Recognition and Computer Vision*, SPIE 8919, pp. 891903–891903-8 (2013)
6. Yang, J., Blum, R.S., Williams, J.P., Sun, Y., Xu, C.: Non-rigid image registration using geometric features and local salient region features. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.* **1**, 825–832 (2006)
7. Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., Fua, P.: BRIEF: computing a local binary descriptor very fast. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(7), 1281–1298 (2012)
8. Rosten, E., Porter, R., Drummond, T.: FASTER and better: a machine learning approach to corner detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(1), 105–119 (2010)
9. Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J.V., Miyasaka, K.: Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **53**(4), 822–834 (2002)