Computerized Medical Imaging and Graphics 58 (2017) 62–74

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

Automated detection of bone metastatic changes using serial CT scans

Jihun Oh[a,*], Gyehyun Kim[a], Jaesung Lee[a], Minsu Cheon[a], Yongsup Park[a], Sewon Kim[b], Jonghyon Yi[b], Ho Yun Lee[c]

[a] DMC R&D Center, Samsung Electronics, Seoul, Republic of Korea
[b] Health and Medical Equipment Business, Samsung Electronics, Suwon, Gyeonggi-do, Republic of Korea
[c] Department of Radiology and Center for Imaging Science, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 24 February 2016
Received in revised form
10 November 2016
Accepted 11 November 2016

Keywords:
Bone metastases
Follow-up CT scan
Longitudinal CT scans
Quantitative imaging
Registration

ABSTRACT

Bone metastases resulting from a primary tumor invasion to the bone are common and cause significant morbidity in advanced cancer patients. Although the detection of bone metastases is often straightforward, it is difficult to identify their spread and track their changes, particularly in early stages. This paper presents a novel method that automatically finds the changes in appearance and the progress of bone metastases using longitudinal CT images.

In contrast to previous methods based on nodule detection within a specific bone site in an individual CT scan, the approach in the present study is based on the subtraction between two registered CT volumes. The volumes registered using the proposed weighted-Demons registration and symmetric warping were subtracted with minimizing noise, and the Jacobian and false positive suppressions were performed to reduce false alarms.

The proposed method detects the changes in bone metastases within 3 min for entire chest bone structures covering the spine, ribs, and sternum. The method was validated based on 3-fold cross validation using the radiologists' markings of 459 lesions in 24 subjects and was performed with a sensitivity of 92.59%, a false positive volume of 2.58%, and 9.71 false positives per patient. Note that 113 lesions (24%) missed by the radiologists were identified by the present system and confirmed to be true metastases. Indeed, three patients diagnosed initially as normal, having no metastatic difference, by radiologists were found to be abnormal using the proposed system.

Automatic detection method of bone metastatic changes in the entire chest bone was developed. Weighted Demons, symmetric warping, following false positive suppressions, and their parallel computing implementation enabled precise and fast computation of delicate changes in serial CT scans. The cross validation proved that this method can be quite useful for assisting radiologists in sensing minute metastatic changes from early stage.

1. Introduction

The use of longitudinal computed tomography (CT) has increased in a prospective and follow-up CT study of the nodule changes and metastatic tumor. Bone metastases, which are the spread of cancer from one part of the body to the bone, are quite common and critical issues in staging cancer, causing significant morbidity and mortality, such as pain, fractures, or spinal cord compression. The majority of patients with breast cancer (90%) or prostate cancer (75%) develop bone metastases. In addition, spinal metastases secondary to breast cancer develop toward mixed lesions with a corresponding decrease in lytic lesions (Skrinskas et al., 2009). Although the detection of bone metastases is often straightforward, it is difficult to identify the spread of bone metastases and track their changes, particularly at the early stages. For quantitative readouts of bone metastases from longitudinal CT, practitioners commonly detect the changes in metastatic tumors by comparing each longitudinal CT scan and calculating their volume size manually. Several approaches have been proposed for the detection of metastases. Huang et al. (Huang and Chian, 2012) proposed an automated computer-aided detection (CAD) system to detect vertebral metastases of breast cancer using the texture features and an artificial neural network. In this method, vertebral metastases were identified in the trabecular centrum of the verte-

* Corresponding author.
E-mail address: oj9040@gmail.com (J. Oh).

http://dx.doi.org/10.1016/j.compmedimag.2016.11.007
0895-6111/

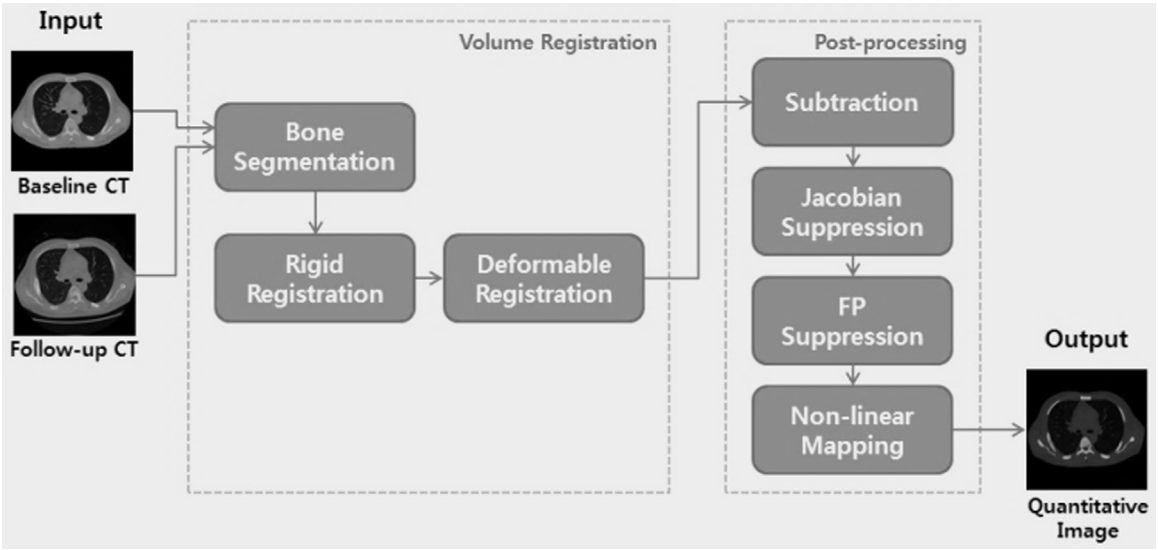

Fig. 1. Flow chart of bone metastatic changes detection algorithm.

bral body detected for each slice, resulting in 86% sensitivity. Wiese et al. (Wiese et al., 2011) introduced a CAD system to aid radiologists in finding sclerotic bone metastases in the spine. This method segments potential lesion candidates in each two-dimensional axial CT image, and eliminates false candidates using a feature filter, resulting in 77% sensitivity at an average of 9 false positives (FP) per case. Yao et al. (Yao Jianhua et al., 2012) suggested a CAD system for detecting sclerotic metastatic bone lesions of the ribs on routine CT studies, with a sensitivity of 75.4% at an average of 5.6 FP per case. Burn et al. (Burns, 2013) used a watershed and merging algorithm to detect lesions and a support vector machine (SVM) classifier to filter out the false positives caused by degenerative changes and partial volume averaging of the vertebral endplates. These approaches required a lesion segmentation technique of a specific bone site in each CT scan. Studies based on multi-modal PET-CT imaging (Beheshti, 2009; Beheshti et al., 2008; Tateishi, 2008) have also been performed, but they are beyond the scope of the present study.

In addition, Hardisty et al. (Hardisty et al., 2007) developed a semi-automated segmentation method to extract the vertebral body and the trabecular centrum in a healthy or metastatic spine using a composition of atlas-based demons registration and level set segmentation, hence enabling a quantitative assessment of vertebral metastases. For visual enhancement, Toth et al. (Toth et al., 2014) recommended the use of a combination of cancellous bone reconstruction and multi-planar reconstructions (MPR) that increased the detection of bone metastases significantly and decreased the interpretation time (a sensitivity of 74% compared to 35% for MPR alone). None of the previous studies could show a sensitivity of 90% or higher at a 10 FP rate. The aim of this study was to enhance the sensitivity rate to more than 90% with the FP rate, which would allow the detection of more true bone metastases and was agreed by clinicians for use in clinical practice. Furthermore, the detection area is limited to the spinal column, which is the most common site of bone metastases, but should include the ribs and sternum as other common sites.

This paper proposes a new automated registration-based method (Fig. 1) to localize changes in sclerotic (blastic) or lytic bone metastases in not just the spine but ribs and sternum using longitudinal CT scans. The method consists of volume registration between the baseline and follow-up scans and post-processings to compute the metastatic changes and to alleviate the false positives.

2. Material and methods

2.1. CT acquisition protocols and subjects

Serial chest inspiration CT scans of 24 patients were used for building a binary classifier to detect bone metastatic changes. The serial CT scans were acquired in a normal dose helical scan (120 kV, 100–330 mAs, standard filter) from a variety of manufacturers and consisted of the baseline and follow-up scans with a gap of 18 ± 12 months. The in-plane resolution was 0.6 mm x 0.6 mm and the slice thickness was 2.5 mm. The image size was 512×512 and the mean number of slices was 120 with a maximum of 160.

2.2. Bone segmentation

Bone segmentation is an important preprocessing step for detecting bone metastatic changes. The bone segmentation was initiated with thresholding (> 1000 Hounsfield unit (HU)), where the HU scale used in the CT numbers was obtained from a linear transformation of the attenuation coefficients measured by a penetrating x-ray beam, and the scale lies in the specific range depending on the penetrated substance (air = −1000 HU, water = 0 HU, bone = 1000 HU), exploiting the high density characteristics of the bone and is followed by selecting the largest 3-D connected region. Bone marrow, which is an internal area of cortical bone, may not be covered with the given threshold of 1000 HU because of its lower HU range. When using a lower threshold for including up to the bone marrow, unwanted soft tissue of the organs of which the HU range is overlapping may also be included. Therefore, to easily enclose the bone marrow, it is filled in by 2-D closing, i.e. dilation followed by erosion, with a 3×3 disk-shaped structuring element and subsequent hole filling morphological operations in each slice. The missed bone with the appearance of lytic metastases with a low intensity range can be enclosed by the aforementioned morphology method, but some cases seen as large indentation were difficult to include. To correct the missed segmentations, the continuity of the segmented voxels in the slice direction was inspected. If the voxels are not segmented but the nearby voxels at the same XY locations in its neighbor slices were segmented, they were deemed to be missing bone segments, and otherwise not bone segments. This process was repeated for up to three neighbor slices above and below.

2.3. CT volume registration

A patient at a follow-up CT scan cannot maintain the identical position and respiration state at a baseline CT scan. This difference causes mismatch of the anatomical structures between the two scans. In this research, image registration in longitudinal thoracic CT scans is necessary for spatial alignment of the images, and enables tracking the changes in the bone metastases in the corresponding location in time series. Previous studies (Baluwala Habib et al., 2011a, 2011b) on thoracic image registration focused on the inhomogeneous regularization across chest image structures in accordance with the difference in their physical properties. Because our research goal is to track metastatic changes just in bone regions, we are free from the concern and can explore straightforward image registration focusing on bony structures.

Each bony structure in the thoracic cage is rigid. On the other hand, the cartilage in the front center and vertebrae joints in the cage provide flexibility in the connections between the sternum and ribs. With the flexibility, during a breathing cycle, the sternum pulls forward or backward and each rib is raised up or down, which allows the volume inside the thoracic cage to expand or contract. The vertebrae may bend sideways to different degrees during the scan times. As such, an alignment of the entire thoracic cage between the two scans needs to be handled with global rigid registration and non-rigid registration.

2.4. Rigid registration

The point clouds used for rigid registration were sampled uniformly from the region encompassing a specific distance of 150 mm from the top of the lung, which correspond to the overlapped region in most chest scans. Herein the lung can be segmented roughly using a simple thresholding (−1000 HU) due to very low density. A pair of point sets (10^4) is sampled in both scans, and the rigid transformation is searched to minimize the sum of the distances between the sampled point sets. As shown in Fig. 2, the 3-D coordinates of the sampled points were utilized in the iterative closest point (ICP) algorithm (Besl Paul and McKay, 1992), which finds an optimal transformation matrix, consisting of rotation and translation components, by minimizing the sum of Euclidean distances between the two sets of points.

2.5. Deformable registration: weighted demons

A modified version of Demons was used to correct for the remaining motion error, which is regarded as non-rigid motion. The original Demons algorithm was introduced by Thirion (Thirion, 1998) in 1998 to compute the velocity field, which is induced while minimizing the sum of squared difference (SSD) dissimilarity measure. In this study, however, the computation needs to be weighted just on the bone and its adjacent surroundings. This is because the lung tissue and the chest wall outside of the chest bone have disparate motion from the chest bone depending on the respiration state in serial CT scans. Rather than hard masking, as shown in Fig. 3, soft masking allows gradual application of the surrounding intensities to derive the formation of smooth displacement field around a target region. This also protects the decline of registration accuracy due to possibly under-segmented bone.

The main contribution is to propose a new weighted-Demons algorithm that registers images focusing on the target region, but also involving the surroundings with gradually decreasing weight adjusted by the Gaussian kernel. To cope with this, a new dissimilarity measure of the Gaussian-weighted sum of squared differences is defined in Eq. (1). S is a static image and M is a moving image, and S_m and M_m are binary masks, representing the targets in each image, which corresponds to the baseline and follow-up CT scans and their bone masks. In addition, $G(\cdot)$ represents the Gaussian filtering and the transformation vector $\vec{T}$ deforms the moving image. Internally, the mask is transformed to a distance map where Gaussian weighting, adjusting the range of weighting with a standard deviation, was applied to the map through which the importance of each voxel for registration was determined.

$$E\left(S, M\circ\vec{T}\right) = \sum_{p}\left(G(S_m[p])S[p] - G\left(M_m[p]\circ\vec{T}[p]\right)M[p]\circ\vec{T}[p]\right)^2 \quad (1)$$

The Gaussian function and its derivative were used for applying the registering weights to the regions around the target. Here, the weights are isotropic so the influence around a pixel p to neighbors is identical depending only on their distance to p. In addition, the weights are decreasing gradually to reflect the influence of the remote points is small compared to close points.

Eq. (2) presents the velocity vector $\vec{v}$ that is derived from a bi-directional version (Wang, 2005; Rogelj and Kovacic, 2006) of the Demons algorithm. This enables the algorithmic acceleration of registration based on the law of action and reaction.

$$\vec{v} = -\frac{2E_M\nabla E_M}{\|\nabla E_M\|^2 + 4\alpha^2E_M^2} + \frac{2E_S\nabla E_S}{\|\nabla E_S\|^2 + 4\alpha^2E_S^2}, \quad (2)$$

where the active force, E_M, proposed is $\left(G(S_m)S - G\left(M_m\circ\vec{T}\right)M\circ\vec{T}\right)^2$ and the passive force, E_S, proposed is $\left(G\left(S_m\circ\vec{T}\right)S\circ\vec{T} - G(M_m)M\right)^2$. The analytical gradients with respect to the respective transform vectors are derived in Eqs. (3) and (4):

$$\nabla E_M = \frac{\partial E_M}{\partial\vec{T}} = -2\left(G(S_m)S - G(M_m)M\right)\cdot\left(G'(M_m)\nabla M\cdot M + G(M_m)\nabla M\right) \quad (3)$$

And

$$\nabla E_S = \frac{\partial E_S}{\partial\vec{T}} = 2\left(G(S_m)S - G(M_m)M\right).\left(G'(S_m)\nabla S.S + G(S_m)\nabla S\right). \quad (4)$$

The fluid-like regularization constrains the computed velocity$\vec{v}$, which was added to the displacement field by accumulation for its fast and simple computation. The diffusion-like regularization constrains the displacement vector$\vec{u}$. Those regularizations were implemented with Gaussian smoothing. The strength of the constraint was determined by controlling the standard deviation of the Gaussian function.

$$\vec{v} \longleftarrow K_{fluid} * \vec{v},\ \vec{u} \longleftarrow K_{diffusion} * \left(\vec{u} + \vec{v}\right) \quad (5)$$

The follow-up image and its Gaussian distance map, which was transformed according to the updated displacement field during the dissimilarity measure, was minimized.

$$M \longleftarrow M\circ\vec{T}\left(\vec{x} + \vec{u}\right),\ G(M_m) \longleftarrow G(M_m)\circ\vec{T}\left(\vec{x} + \vec{u}\right) \quad (6)$$

2.6. Symmetric warping

It is conventional to apply the final displacement vector to warping the moving image, but maintaining the static image fixed although symmetric registration, measuring the image similarity in both registration direction, is performed. However, this asymmetric warping would produce a blurring effect just to the moving image so the subtraction between the static image and the warped moving image may result in spike noise across the image. Because the noise prevalence determines the performance of false positives, it is important to lower the noise before a post-processing step.

To address this issue, both static and moving images as the final results are warped, thereby counterbalancing the blurring effect. In detail, to warp both images and match them in the middle, the given computed parameters of the rigid transform–rotation $R\left(\vec{\theta}\right)$ and

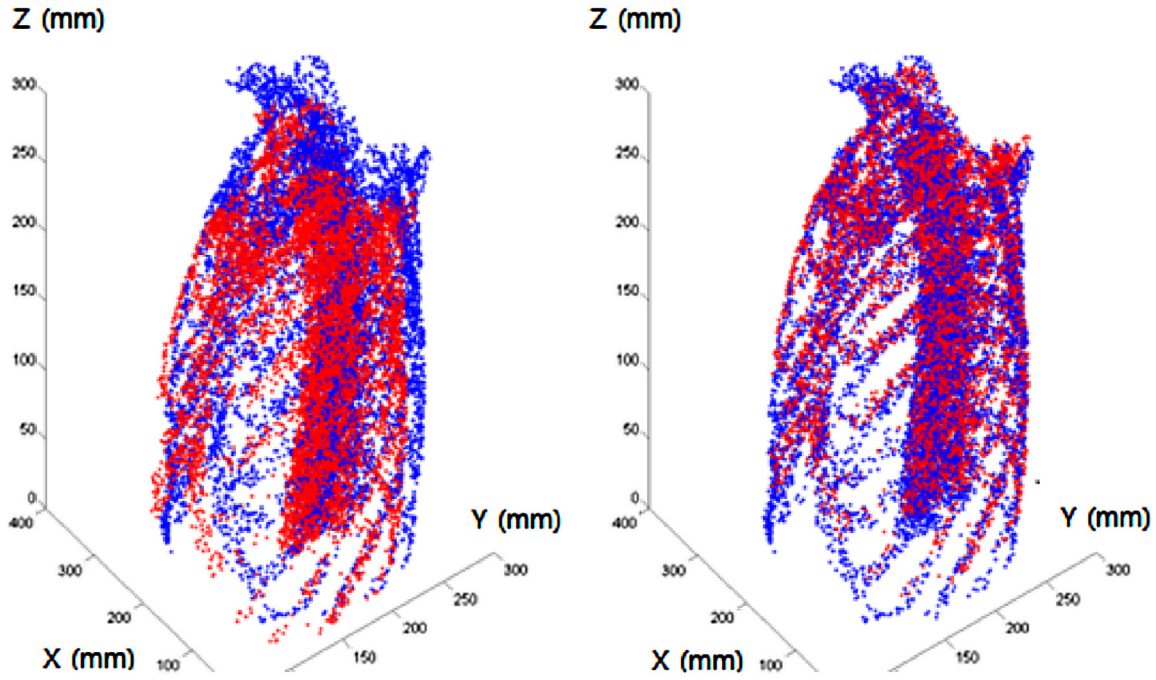

Fig. 2. Point-set based rigid registration using the ICP algorithm; before (left) and after (right) the registration of the point sets uniformly sampled from the baseline bone (blue) and follow-up bone (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. CT scan (a) and a bone mask (b) for each baseline and follow-up are inputs to the weighted demons, where the mask is transformed to a Gaussian-weighted distance map from the closest nonzero value (c).

translation$\vec{t}$ –to match the follow-up $F\left(\vec{x}\right)$ to the baseline $B\left(\vec{x}\right)$ at a voxel$\vec{x}$, symmetric rigid-transformed baseline $B^{*}\left(\vec{x}\right)$ and follow-up $F^{*}\left(\vec{x}\right)$ are aligned in the midway as follows:

$$B^{*}\left(\vec{x}\right) = B\left\{R\left(-\frac{\vec{\theta}}{2}\right)\vec{x} + R\left(-\frac{\vec{\theta}}{2}\right)\cdot\left(-\frac{\vec{t}}{2}\right)\right\} \tag{7}$$

and

$$F^{*}\left(\vec{x}\right) = F\left\{R\left(\frac{\vec{\theta}}{2}\right)\vec{x} + R\left(\frac{\vec{\theta}}{2}\right)\cdot\left(\frac{\vec{t}}{2}\right)\right\}. \tag{8}$$

Afterward, the symmetric non-rigid deformation of the baseline $B^{**}\left(\vec{x}\right)$ and follow-up $F^{**}\left(\vec{x}\right)$ scans can be obtained by

$$B^{**}\left(\vec{x}\right) = B^{*}\left(\vec{x} + 0.5u^{-1}\left(\vec{x}\right)\right) \text{and} F^{**}\left(\vec{x}\right) = F^{*}\left(\vec{x} + 0.5u\left(\vec{x}\right)\right), \tag{9}$$

where the inverse transform of the non-rigid displacement field can be computed numerically through n times recursive compositions of the 2^{n} down-scaled negative values in accordance with the diffeomorphic transformation (Vercauteren et al., 2009, 2008). Here, three scales and a maximum of 100 iterations per scale are used for the registration

The cropped spine regions in Fig. 4 apparently show the positive effect of symmetric warping compared to asymmetric warping. The subtraction images after asymmetric warping have spike noise around the bone boundary other than metastatic changes inside the bone. Symmetric warping has an effect on reducing the noise, but maintaining the metastatic changes.

2.7. Post-processing

2.7.1. Subtraction

To quantify the metastatic changes between two image volumes, subtraction (S) of the symmetrically registered baseline (B^{**}) from the follow-up (F^{**}) is performed as shown in Eq. (10).

$$S = F^{**} - B^{**} \tag{10}$$

The subtraction values are then masked on the bone region, which is defined as the union domain of two bone segments aligned.

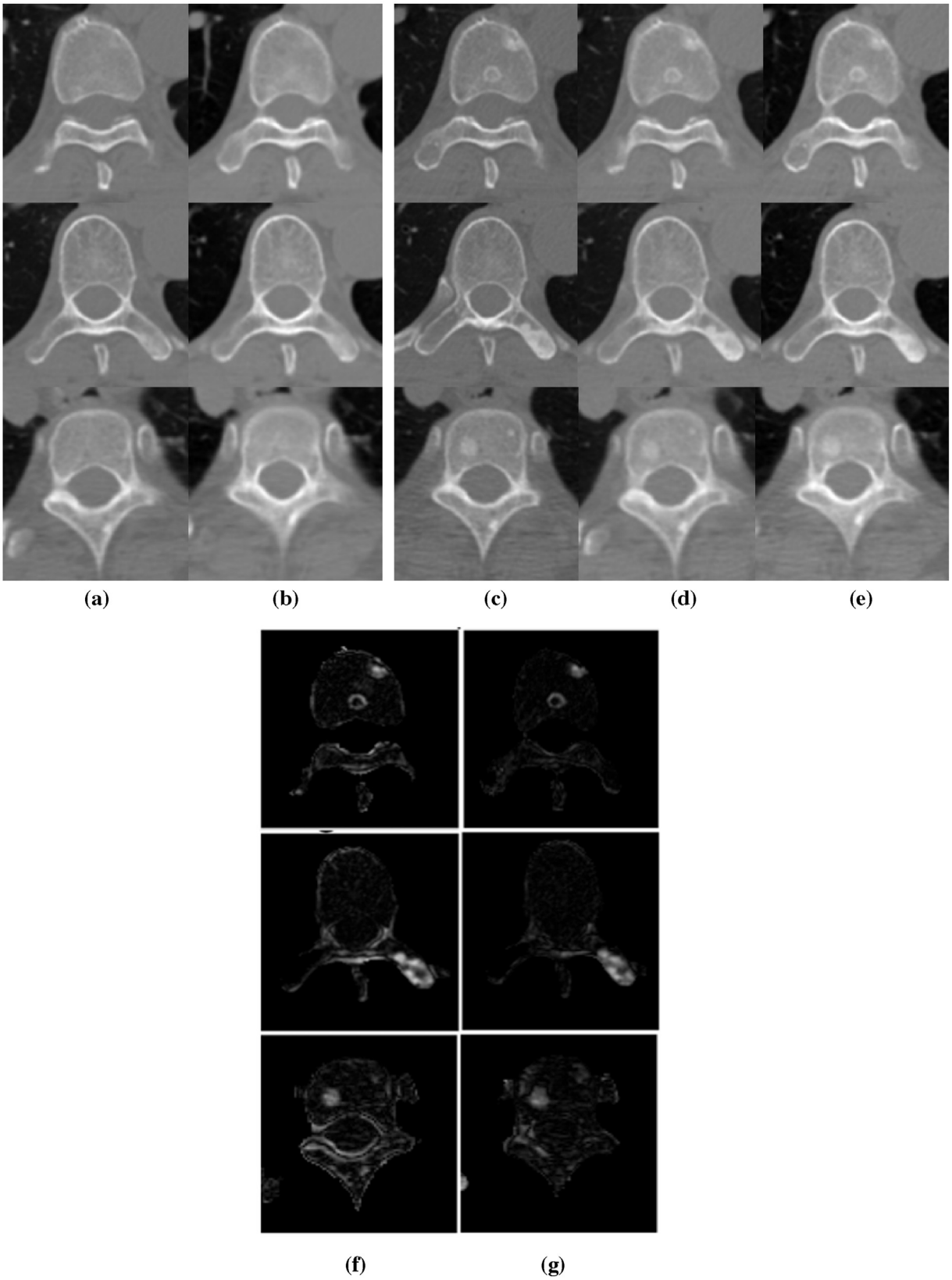

Fig. 4. For three spine regions (1st to 3rd row), warped images and subtraction images after warping are shown: (a) baseline; (b) symmetric warped baseline; (c) follow-up; (d) asymmetric warped follow-up; (e) symmetric warped follow-up; (f) subtraction of asymmetric warped scans (a) from (d); (g) subtraction of symmetric warped scans (b) from (e). Symmetric warping scheme (g) enables deformation of both the baseline and follow-up, thereby reducing the subtraction noise compared to the asymmetric warping scheme (f).

This is followed by one-voxel erosion to remove the unwanted spike noise that appeared along the bone boundary. The cartilage region is excluded from the region of interest.

2.7.2. Suppression

As one of the indirect methods for measuring the reliability in matching, the Jacobian determinant (J) (Bistoquet et al., 2008) in Eq. (12) of the transformation field in Eq. (11), shows how the local volume is changing around each voxel; a Jacobian determinant of 1 indicates that the local volume remains the same, a value greater than 1 indicates that the local volume expands, and a value smaller than 1 indicates that the local volume contracts. Because the bone is almost volume-preserving, the local volume should be maintained almost the same along the time series. Although bone metastases are present, if most regions surrounding that are normal, it can abide by the volume-preserving condition. The subtraction values in the region, where the Jacobian determinant far from 1 is unconvincing, are suppressed by weighting the unnormalized Gaussian function of Jacobian centered at one ($N(1, 0.2^2)$), as expressed in Eq. (13).

$$\vec{x'} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} u_x(x,y,z) \\ u_y(x,y,z) \\ u_z(x,y,z) \end{pmatrix}, \tag{11}$$

$$J\left(\vec{x'}\right) = \det \begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial y'}{\partial x} & \frac{\partial z'}{\partial x} \\ \frac{\partial x'}{\partial y} & \frac{\partial y'}{\partial y} & \frac{\partial z'}{\partial y} \\ \frac{\partial x'}{\partial z} & \frac{\partial y'}{\partial z} & \frac{\partial z'}{\partial z} \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 + \frac{\partial u_x}{\partial x} & \frac{\partial u_y}{\partial x} & \frac{\partial u_z}{\partial x} \\ \frac{\partial u_x}{\partial y} & 1 + \frac{\partial u_y}{\partial y} & \frac{\partial u_z}{\partial y} \\ \frac{\partial u_x}{\partial z} & \frac{\partial u_y}{\partial z} & 1 + \frac{\partial u_z}{\partial z} \end{pmatrix}, \tag{12}$$

$$S' = S * (G(J)), \tag{13}$$

where $\vec{x'}$ represents a field of transformation vector with x', y', and z' axial components, and which is a combination of regular grids in the x, y, and z axes and their displacements u_x, u_y, and u_z. Each of them has the same size with the scan volume.

In addition, weighting suppression on the low subtraction values diminishes the false positives. The suppress factor used is determined using the quadratic function, as expressed in Eq. (14) and its profile is shown in Fig. 5. This factor weighted the subtraction values as Eq. (15).

$$\textit{suppress factor} = \frac{1.75}{\max(S')^2} * S'^2 + 0.75 \tag{14}$$

$$S' \longleftarrow \textit{suppress factor} * S' \tag{15}$$

This weighting scheme cancels out the subtraction values close to zero while increasing the contrast for higher subtraction values. In addition the region detected as the lytic metastasis ($S' < 0$) is suppressed by a 0.357 wting, which is obtained heuristically throughout the experimental studies. A scapula, which is a part of the shoulder bone, is subject to many false positives because it moves aside from the chest so the registration error could exist more than in the chest region. For this reason, the scapula is extracted by a series of morphological processes, including erosion, searching for the largest 3-D connected region, closing, and dilation. Subsequently, the subtraction values are suppressed and smoothed out with the Gaussian filter to reduce the noise.

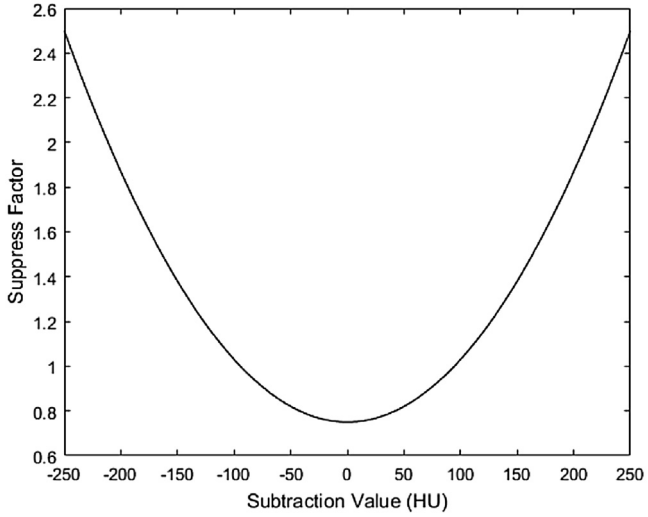

Fig. 5. Suppression factors determined by the quadratic function described in Eq. (14).

2.8. Registration and false positives reduction

The proposed weighted-Demons registration method was compared with the original Demons algorithm. The parameters used in the weighted Demons algorithm, alpha and Gaussian standard deviation for smoothing the velocity field, displacement field, and Gaussian-weighting were 0.5, 3 mm, 3 mm, and 6 mm, respectively, which were acquired experimentally. Identical parameters except for the Gaussian-weighting parameter were used in the Demons algorithm. Fig. 6 presents a visual comparison of the registration results in the axial slice, where the Demons algorithm yielded a large mismatch error in the shoulder bone and spine (b), whereas the weighted-Demons algorithm does not (c). In the 3-D overlap of the registered bones (d, e), weighted-Demons exhibited higher matching performance, particularly in the scapula and lower ribs regions. Aligning those regions is challenging because registration is performed to the torso overall, but not separately to the scapula. In addition, the lower ribs were prone to misalignment due to the large motion of the liver nearby. In the dice similarity coefficient (DSC) measurement, which represents an overlap ratio between two bone volumes, the weighted Demons with an average of 93 (std = 1.5)% outperformed the Demons algorithm with an average of 88 (std = 2)%. As validation of symmetric warping, for example, the metastatic change between the baseline and follow-up scans was not confirmed by clinicians in the case shown in Fig. 7. In the experimental tests, however, the quantitative result from conventional asymmetric warping (c) had a false positive with large magnitude, but could be mitigated by the symmetric scheme that was adopted (d).

In addition, the Jacobian map was devoted to lowering the false positive errors in Fig. 8. The low difference value is suppressed nonlinearly by the quadratic function, leading to a decrease in the number of false positive errors. In addition, cartilage, which is a connective tissue, not bone, found between the sternum and ribs is ruled out from the area of interest. To exclude the cartilage region, on the axial slice the voxels located within 70 mm to 30 mm to the left and right sides each from the centroid of the entire chest bone

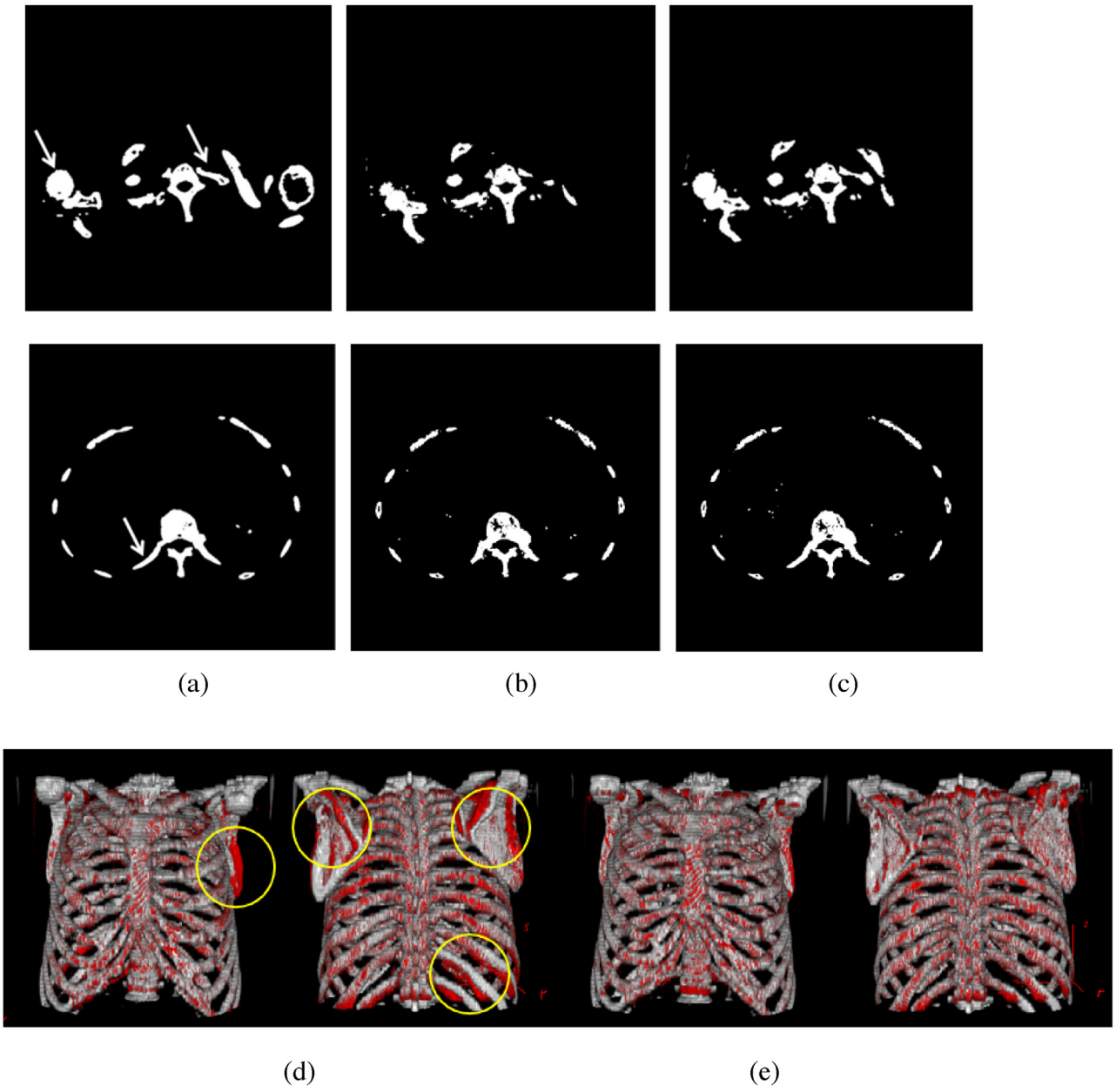

Fig. 6. Visual comparison of registration results; (a) baseline, (b) registered follow-up using Demons, (c) registered follow-up using w-Demons. Note the regions highlighted with the white arrow indicate better matching performance of w-Demons. The overlap volumes (white: baseline, red: registered follow-up) are rendered in (d) and (e) by the Demons and w-Demons, respectively.

and within 30 mm from the top of it were experimentally supposed as cartilage, thereby were masked out. The left, right scapulars are prone to error, thus being extracted and suppressed, as shown in Fig. 9. To segment it, the entire segmented bone is eroded iteratively until being disconnected into the thoracic and scapular bones. Furthermore, with the exception of the top three largest connected regions, which were assumed to be the spine and two scapulars, the rib bones and others are removed. Finally the segment located closest to the center of the image, which was assumed to be the spine, was eliminated and then only the scapulars remained. Lytic metastases were more sensitive to noise in the difference image, and were suppressed. Moreover, Gaussian filtering of the difference image alleviated the peak difference noise, resulting in the greatest effect among the false positives reductions above. The kernel size of the Gaussian filter was $5 \times 5 \times 1$, taking into account of the anisotropic voxel size. Gaussian filtering is a typical choice for the smoothing signal, i.e., retaining the low frequency component while reducing the high frequency component, which outperformed the other low pass filters, such as mean or median filters. Each of the false positive reduction techniques and corresponding reduction rates

Table 1
False positive reduction techniques and their performances.

	Reduction rate of false positive volume (%)
Jacobian map	0.2
Low diff value suppression	1
Cartilage exclusion	0.1
Lytic suppression	1.9
Scapular suppression	0.71
Gaussian smoothing	5.24

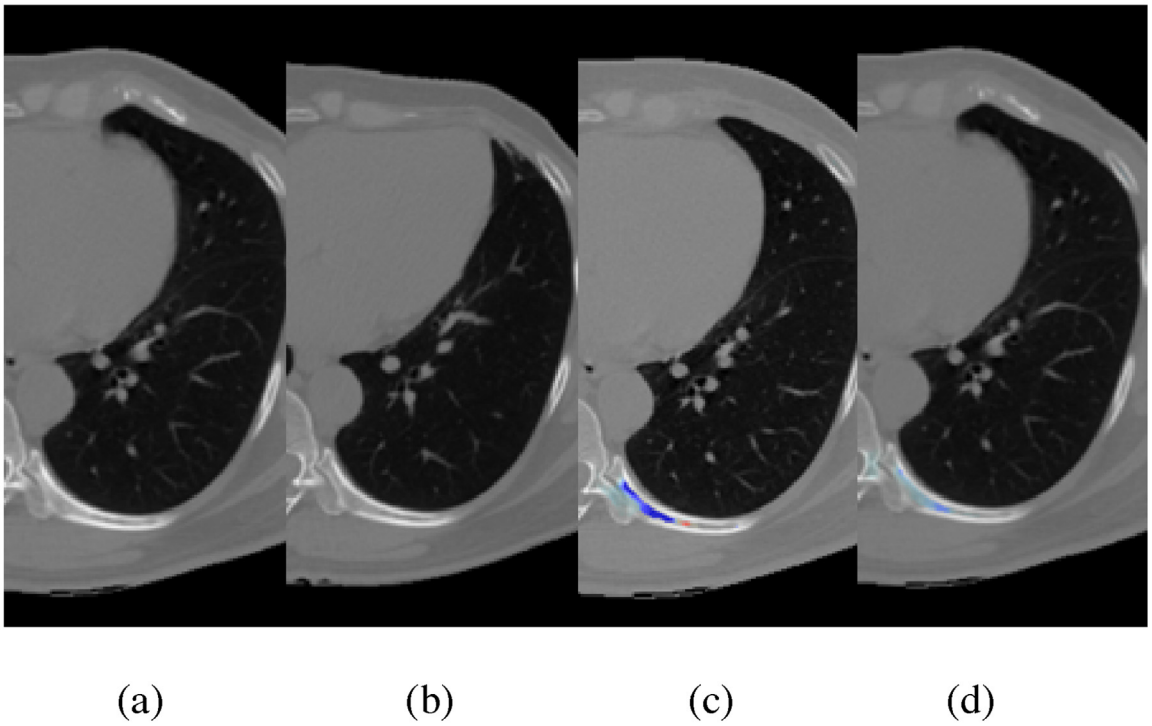

Fig. 7. (a) Baseline CT, (b) Follow-up CT, (c) False positive error by asymmetric warping, (d) Effect of symmetric warping.

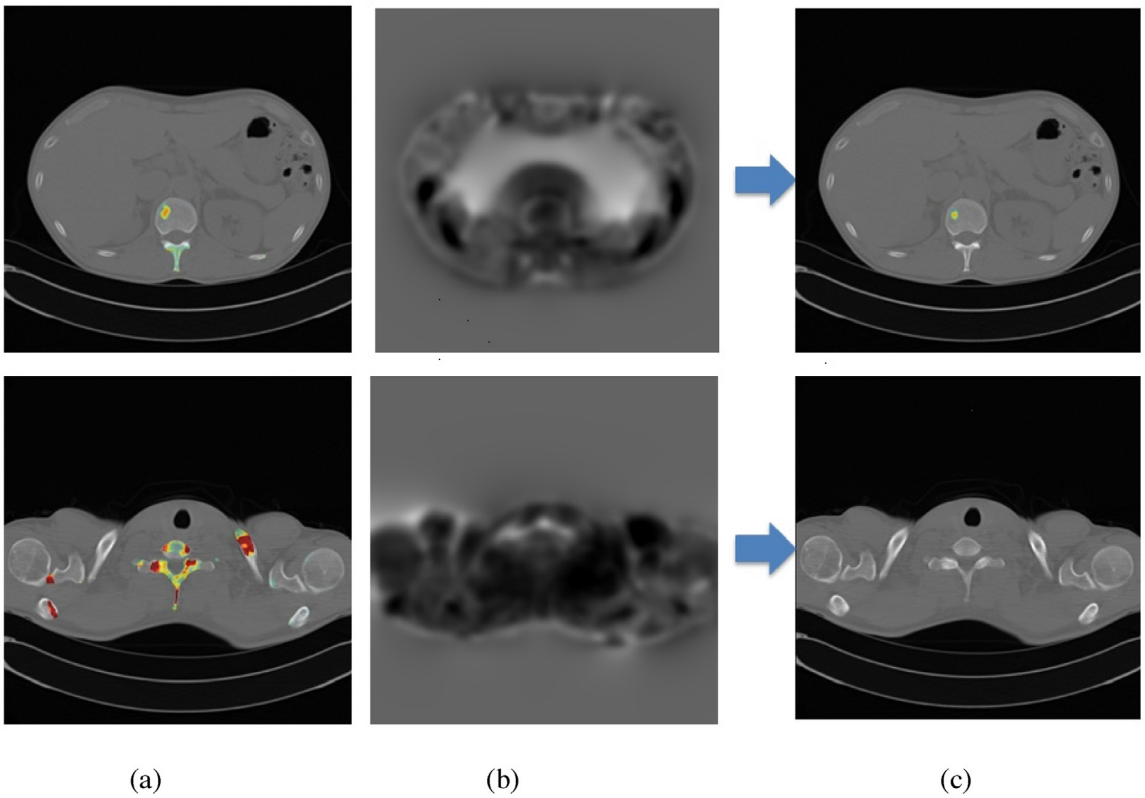

Fig. 8. (a) Bone metastatic changes detection without Jacobian-based suppression (b) Jacobian determinant map (a Jacobian of 1 is gray, greater than 1 is white, and lower than 1 is black) (c) with Jacobian-based suppression.

is listed in Table 1. The total reduction rate of the false positive volume is 9.15%.

To validate this method, the clinicians were asked to manually delineate suspicious lesions with bone metastatic changes as ground truth for 24 subjects. Clinicians localized a center point of each suspicious region comparing the baseline and synchronized follow-up slices through a visual inspection. The clinicians marked only the lesions having metastatic differences in size or density, but

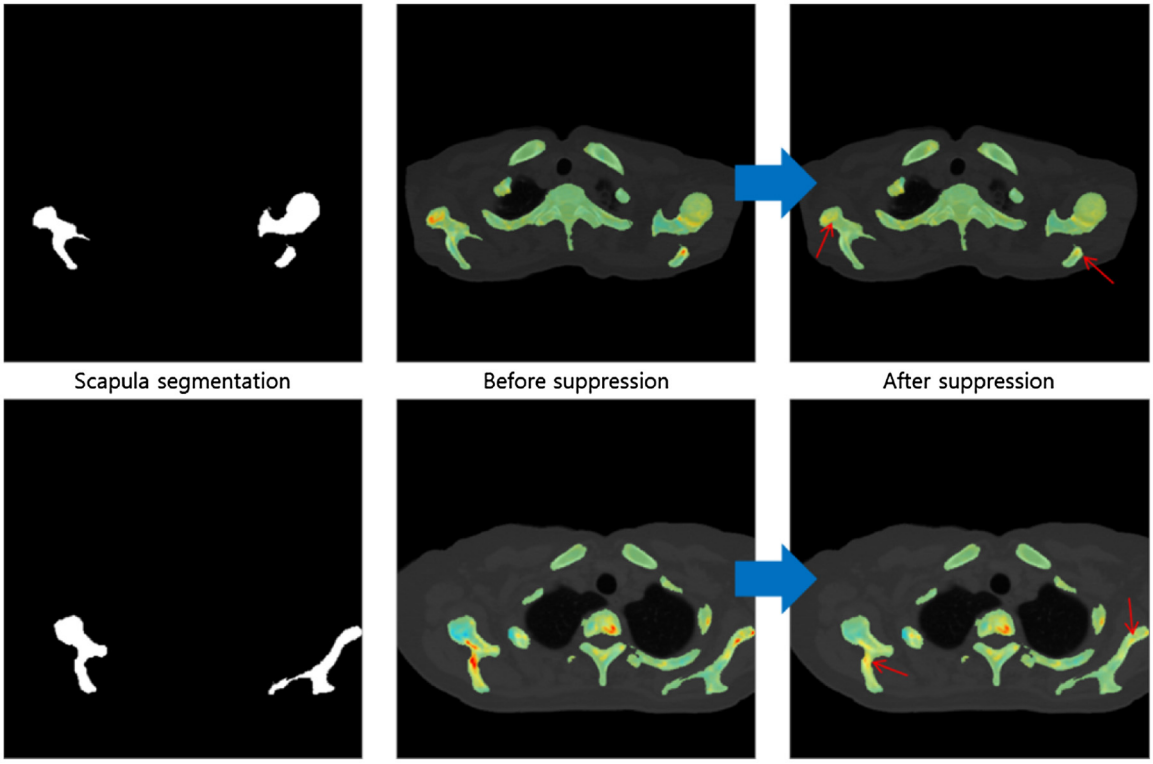

Fig. 9. Suppression on the scapular region.

Table 2
Data partitions into three groups for 3-fold cross validation.

	Group1	Group2	Group3
Subject No.	#1, #4, #5, #12, #13, #16, #18, #22	#2, #3, #7, #9, #10, #14, #23, #24	#6, #8, #11, #15, #17, #19, #20, #21
# of Lesion	175	144	140

did not include the lesions with little change. That is because we aimed to localize the deteriorating metastatic legion. Besides the marked ground truth, the ones found by the proposing classifier were added after confirmation by the clinicians. The total number of lytic and blastic metastatic lesions was 459. As to validation, the 24 subjects from whom lesions are gathered are not big enough to completely separate that into training and test sets. In addition, as the variance in the lesion numbers distributed over patients is 21.5 (Table 4), which is larger than the average 19.1, the trained classifier can be so biased depending on the selection of patients as train data. Thus, for fair validation the 3-fold cross validation approach was applied. The whole subjects were partitioned into three groups exclusively (Table 2) keeping the size of lesions in balance. Of the three groups, each validation used two groups as the training data and one group as the test data. This cross validation process is then repeated three folds, with each of the data groups used exactly once as the test data. At each round, in training, the accuracy in terms of sensitivity and the false positive rate of the binary classifier on a set of different thresholds was computed and then the optimal value was determined based on the receiver operating characteristics (ROC) analysis. A test using the learned threshold value was performed.

To determine if the computed metastasis belongs to the reference marked by clinicians, the post-processed subtract image was thresholded with a set of values. Particles with an absolute difference greater than the threshold value are regarded as candidates of metastases. To cluster the nearby candidates, morphological dilation with a structure size of [20205] taking the anisotropic voxel size into account was applied to the threshold image, and the clusters were labeled based on the 3-D connectivity. Statistical analysis was then performed in terms of sensitivity, the number of false positives (FPs), false positive percentage, and the number of false negatives (FNs). To illustrate this, for example, as shown in Fig. 10, one of two markings, but not for the other, is located inside one of the two detected metastases regions; the sensitivity is defined as the ratio of the true positive (TP) to the total number of markings (TP + FN) = 0.5. The FP was measured in the perspectives of the counting number and the percentage of detected metastases out of the total bone area. In the figure, the number of FP is one and the FP volume, which is a portion of the FP volume in the overall bone volume is 3% as an example. The number of FN is one because one reference is not included in the detected region.

3. Results

3.1. Statistical analysis

Receiver operating characteristic (ROC) analysis shows the sensitivity and FP performance graphically as the discrimination threshold is varied. It seeks to determine the optimal value among a set of thresholds ranging from 5 to 100 with a space of 5, which is used for a binary classifier system. As a performance goal, the designed classifier aims to satisfy that the sensitivity is greater than 90%, the number of FPs is less than 10, and the FP volume is less than 3% per case. The ROC curve with respect to the number of FPs was not increasing as smoothly as that with respect to the FP vol-

Table 3a
Training and test results in 3-fold cross validation. Of three partitioned data groups, each validation used two groups as training data and one group as test data. The optimal binary threshold was learned among a set of 5 to 100 with a space of 5 in training, and then the determined threshold was used for testing the classifier.

N-fold Cross Validation (N=3)		Val. 1	Val. 2	Val. 3	Average
Training Datasets		Group2 + Group3	Group1 + Group3	Group1 + Group2	
Test Datasets		Group1	Group2	Group3	
Training Results	AUC(%)	98.38	99.2	98.34	n/a
	Optimal Threshold	55	40–65 (55)	55	
	Sensitivity(%)	91.48	95.36	90.94	
	FP vol. (%)	2.90	1.75	3.10	
	# of FP	10.0	9.0	10.13	
	# of FN	1.43	1.19	1.88	
Test Results	Sensitivity(%)	94.82	87.05	95.91	92.59
	FP vol. (%)	1.95	4.25	1.55	2.58
	# of FP	9.13	11.13	8.88	**9.71**
	# of FN	1.63	2.13	0.75	**1.50**

Table 3b
Computation time and accuracy in different computing environments.

Computer spec	GPU spec	Scheme	Time	Sensitivity/case	FP/case	FP_vol/case	FN/case
Intel i7, 3.5 GHz, 32GB	Geforce GTX680 1536 cores 4GB memory	Full Resolution	2 min	92.6%	9.71	2.58%	1.50
Intel Xeon, 3.4 Ghz, 32 GB	Quadro K600 192 cores 1 GB memory	Down-sampling(1/4 size) Partitioning(4 partitions)	3 min 16 min	92.5% 92.6%	12.13 9.67	2.86% 2.58%	1.54 1.50

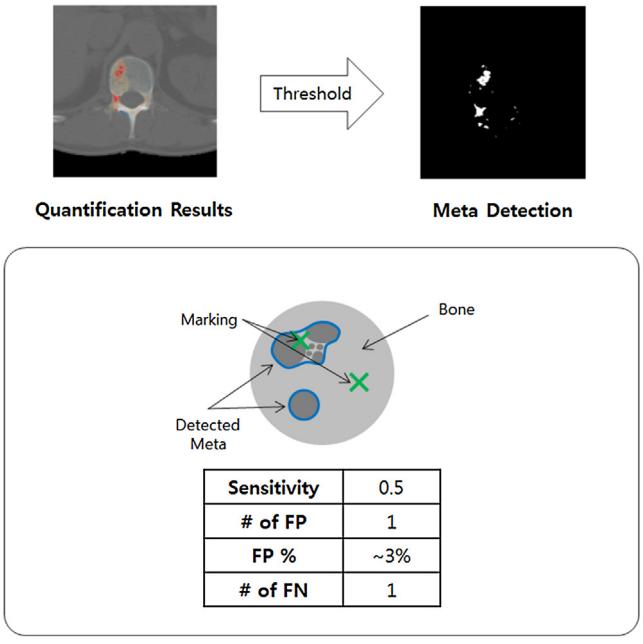

Fig. 10. Example of evaluation metric of bone metastatic changes detection.

ume. This is because as the threshold decreases, the sensitivity and FP both are supposed to increase, but the FP particles apart from each other at a low threshold value can congregate through the neighborhood clustering described above. Therefore, the number of FP can decrease. In this case, despite a low threshold, the number of FPs can decrease so the ROC curve does not appear smooth in some range. On the other hand, in the perspective of the FP volume the curve increases monotonously as the threshold decreases. Fig. 11 shows three ROC curves resulted from trainings in 3-fold cross validation. The area under the curves (AUC), which is averaged accuracy over all queries and reported as a single value, were 98.38%, 99.20%, and 98.34% (Tables 3a and 3b), which represents the excellent performance of the classifier. The threshold satisfying our goal was only 55 for 1st and 3rd trainings, but 40 through 65 for 2nd training. The averaged value through three trainings, 55, was determined as the optimal threshold. The black squares in Fig. 11 indicate the performance at the optimal threshold.

As test results, the 1st and 3rd ones satisfied the performance goal—a sensitivity of 94.82%, 1.95% FP volume, and 9.13 FPs and a sensitivity of 95.91%, 1.55% FP volume, and 8.88 FPs, respectively. However, the 2nd one was a little off from the goal—a sensitivity of 87.05%, 4.25% FP volume, and 11.13 FPs. That is because the test data of 2nd group included Case #9 which had low sensitivity compared to the other cases (refer to Table 4). It arose from Gaussian smoothing as one of false positive reduction techniques relieved even the weak metastases but effectively reducing the overall false positives. The classifier produced the averaged test result such that the sensitivity is 92.59%, the number of FPs per case is 9.71, the FP volume is 2.58%, and the number of FNs is 1.5 per case. Indeed, out of 123 lesions also identified using this method, 113 lesions (91.8%) were missed by radiologists and the other 10 lesions turned out to be degenerative. The three subjects (#14, #16, and #24) misjudged as normal, that is having no metastatic difference, by the radiologists were identified correctly as abnormal by the proposed method, which were re-verified by them. Fig. 12 shows the registered baseline and follow-up scans, and resulting detection map representing the degree of metastatic changes in colormap and herein the blue indicates lytic metastases and the red indicates the blastic metastases. As comparing between the baseline and follow-up scans in the figure, it is not easy to localize the metastatic changes, particularly in the small cross-section area like ribs, and furthermore, it is hard to identify them in early stages. However, the detection map can relieve radiologists of those difficulties by enabling fast, easy, and precise interpretation of bone metastases ultimately.

3.2. Computation performance

For implementation, a PC with Intel i7, 3.5 GHz and 32GB memory, running in Window 7, was used. The graphics processing unit (GPU) was Nvidia GeForce GTX 680 supporting compute capability 3.0, which integrates 1536 CUDA cores and 4 GB of global memory. The program was built with C/C++ and CUDA parallel comput-

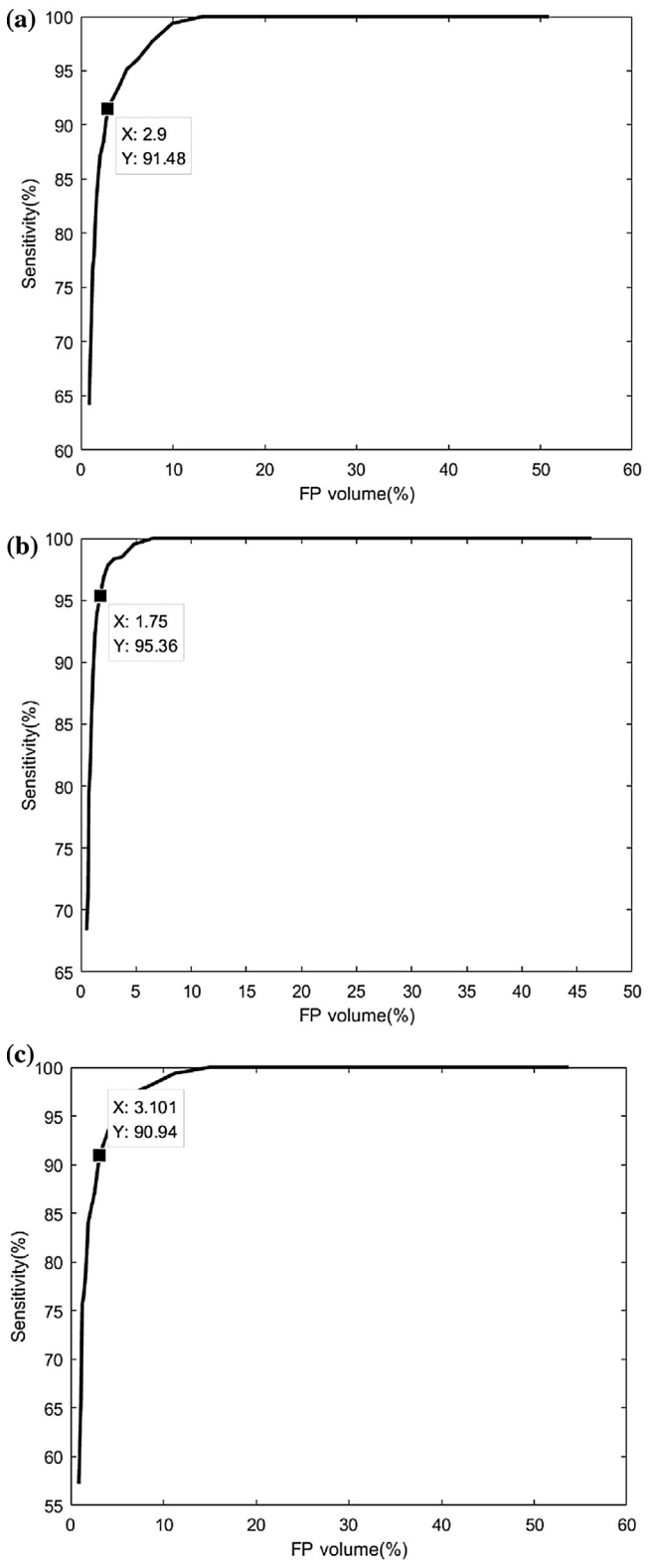

Fig. 11. ROC curves acquired from trainings of three-fold cross validation: the X axis denotes the FP volume (right). The black squares at a threshold of 55 indicate the optimal result satisfying the performance goal (Sensitivity > 90%, FP volume <3% (or # of FP < 0)).

ing language in Microsoft Visual Studio 2013. For a scan size of $512 \times 512 \times 127$, the computation time was 2 min for the entire process, as indicated in Tables 3a and 3b.

The study was also tested in a workstation with Intel Xeon, 3.4 GHz and 32GB memory. In contrast to the PC used first, this server has a relatively lower-end GPU, Nvidia Quadro K600, in which 192 CUDA cores and 1 GB of global memory are integrated. Because the required GPU memory space in the deformable registration was approximately 3 GB, two approaches to addressing the out of memory issue were attempted. One approach was to down-sample the volume by half in each X and Y axes, but maintain the full size in the Z axis due to anisotropic voxel size, finally reducing the volume size to 1/4. The resulting translation vector was up-sampled to the original size and warped the original volume. This approach achieved a comparable speed of 3 min for the entire process but could not guarantee similar accuracy to the one from using the full resolution (a sensitivity of 92.5% and 12.13 FPs). The other approach was to separate the volume into 4–5 partitions along the z axis with a margin of 10 slices in the upper and down sides each and register the partial volumes sequentially. Using exclusive combination of transformation vectors excluding the margin, the entire volume was warped. In this environment, the computation time was 16 min for the entire process without any loss of accuracy (a sensitivity of 92.6% and 9.67 FPs). Comparing the computation time with the original Demons in GPU, the proposed weighted Demons method was slightly slower (by 5%) due to the greater complexity in computing the gradient of dissimilarity measure. In detail, for high-end GPU acceleration, the original Demons and weighted Demons exhibited 1.9 min and 2 min in the full resolution approach, and for lower-end GPU acceleration 2.85 min and 3 min in the down-sampling approach, and 15.2 min and 16 min in the partition approach, respectively.

4. Discussion

The main objectives of this paper were to propose a method that automatically detects the changes in appearance and the progress of bone metastases using two longitudinal CT images. In contrast to previous methods based on nodule detection within a spine in an individual CT scan, the approach in the present study is to subtract between two registered CT volumes and detect the metastatic changes in more expanded bone sites, including not just the spine but ribs and sternum, with a high accuracy. The volumes registered using the proposed weighted-Demons registration and symmetric warping were subtracted, and the Jacobian and false positive suppressions were performed. Throughout the experimental studies for 459 lesions in 24 patients, the performance of the proposed method was validated with a sensitivity of 92.6% and a false positive of 9.7 based on the cross-validation. Above all, 113 lesions (24%) missed by radiologists were identified by this system and re-confirmed to be true metastatic changes. Indeed, three patients from whom the metastatic changes were not detected by radiologists were found to have changes by the proposed system, and which is confirmed as true by radiologists. In this study, we tested on two GPU computing power environments with Nvidia GeForce GTX 680 and a relatively lower-end GPU, Nvidia Quadro K600. The overall computation time was 2 min in the high-end GPU, and 16 min in the low-end GPU if maintaining the same accuracy but 3 min if taking a slight loss in the false positive rate.

5. Conclusion

This paper presented a novel method that automatically detects the changes in bone metastases in longitudinal CT images. In contrast to most studies based on nodule detection techniques of spinal

Table 4
Test results showing # of ground truth (GT) (Additional ground truth by CAD), sensitivity, the number of FPs, FP volume, and the number of FNs for 24 subjects each and their average and standard deviation (optimal threshold=55).

Case	# of GT	Sens. (%)	# of FPs	FP volume (%)	# of FNs
#1	27 (4)	92.6	9	2.5	2
#2	50 (6)	98.0	15	8.5	1
#3	41 (6)	92.7	13	10.6	3
#4	3 (0)	100.0	7	0.8	0
#5	9 (3)	88.9	9	1.8	1
#6	13 (5)	76.9	10	0.8	3
#7	22 (7)	81.8	10	2.0	4
#8	4 (1)	100.0	5	0.1	0
#9	13 (3)	46.2	11	1.8	7
#10	4 (2)	100.0	9	9.7	0
#11	43 (7)	95.3	7	0.9	2
#12	11 (3)	90.9	13	1.6	1
#13	3 (2)	100.0	12	5.6	0
#14	3 (3)	100.0	13	0.6	0
#15	3 (1)	100.0	12	0.8	0
#16	1 (1)	100.0	3	0.4	0
#17	2 (0)	100.0	7	0.6	0
#18	35 (9)	94.3	11	2.2	2
#19	20 (8)	95.0	12	2.8	1
#20	49 (25)	100.0	10	3.3	0
#21	6 (1)	100.0	8	3.1	0
#22	86 (14)	91.9	9	0.7	7
#23	9 (0)	77.8	5	0.3	2
#24	2 (2)	100.0	13	0.5	0
Avg. ± Std.	19.1 (4.7)±21.5 (5.4)	**92.6 ± 12.2**	**9.7 ± 2.9**	**2.6 ± 3.0**	**1.5 ± 2.1**

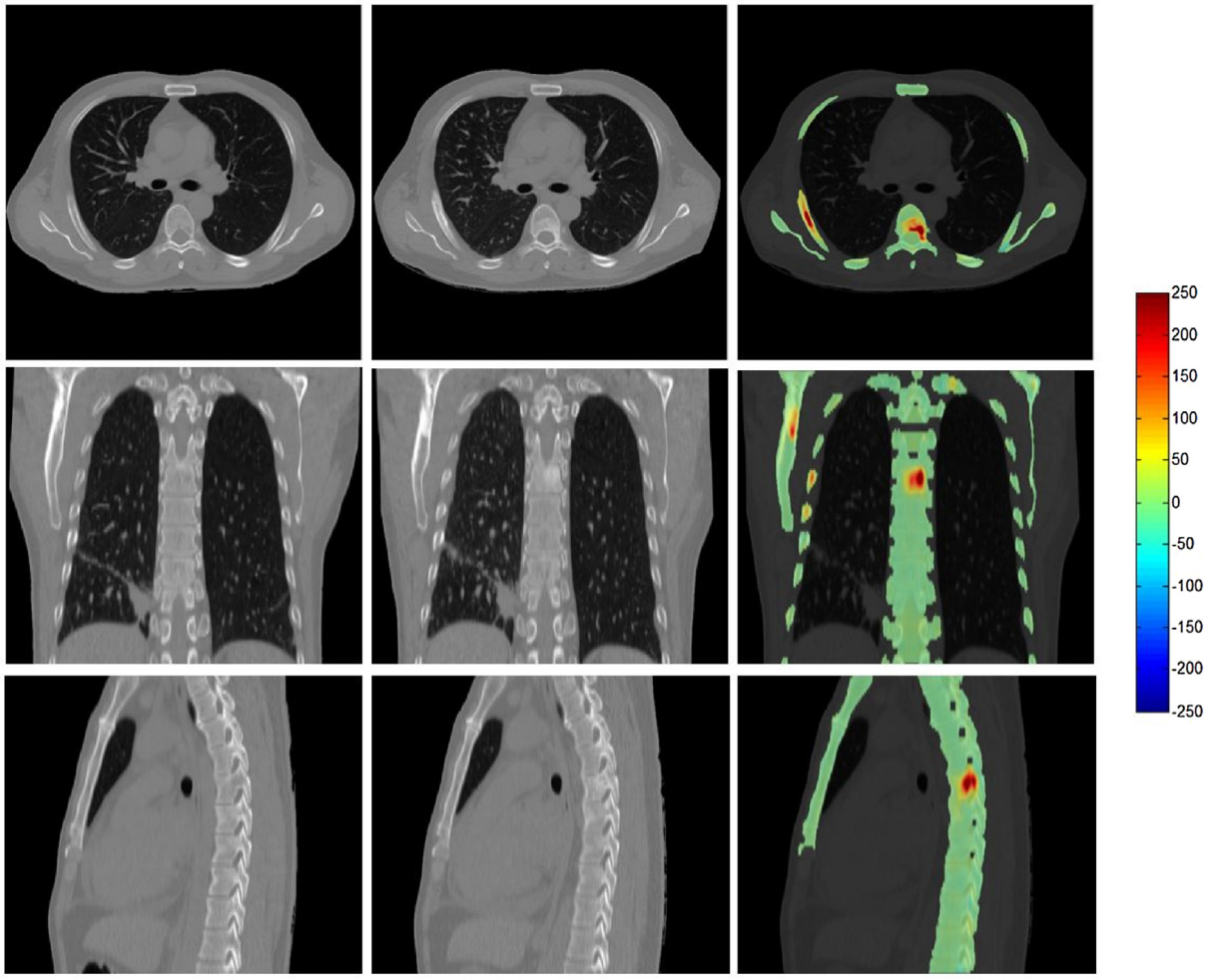

Fig. 12. Test result visualization in axial (1st row), coronal (2nd row), and sagittal (3rd row) planes: registered baseline (1st column) and follow-up (2nd column) CT scans and colormaps (3rd column) showing the degree of metastatic changes (red: blastic, blue: lytic, green: normal).

column bone site in each CT scan, the present approach performs the subtraction between two CT volumes, followed by false positive reduction techniques. The approach then detects the changes in the sclerotic and lytic bone metastases in not only the spine, but also in the ribs and sternum based on the density change between the baseline and follow-up scans. The 3-fold cross validation on 24 patient subjects showed that the proposed method successfully tracked the bone metastatic changes. In addition, parallel computing using graphics cards made it practically feasible in routine clinical use. Therefore, the proposed work can be quite useful for assisting radiologists in sensing even minute metastatic changes. In the future work, this work will be applied to other bone site.

Acknowledgements

This work was funded by Samsung Electronics Co., Ltd.. The authors thank Samsung Medical Center, Sungkyunkwan University School of Medicine for providing the image data and lesion markings as reference. This paper is free from any and all potential conflicts of interest.

References

Baluwala Habib, Y., Kinda Saddi, A., Julia Schnabel, A., 2011a. Non-rigid chest image registration with preservation of topology and rigid structures. Ann. BMVA 5, 1–15.

Baluwala Habib, Y., Kinda Saddi, A., Julia, A., 2011b. Schnabel Elastic registration of chest CT images with log un-biased deformations and rigidity constraint. Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium On. IEEE.

Beheshti, M., et al., 2008. Detection of bone metastases in patients with prostate cancer by 18F fluorocholine and 18F fluoride PET?CT: a comparative study. Eur. J. Nucl. Med. Mol. Imaging 35, 1766–1774.

Beheshti, M., 2009. The use of F-18 choline PET in the assessment of bone metastases in prostate cancer: correlation with morphological changes on CT. Mol. Imaging Biol 11, 446–454.

Besl Paul, J., McKay, N.D., 1992. Method for registration of 3-D shapese. Int. Soc. Opt. Photonics, 586–606.

Bistoquet, A., Oshinski, J., Skrinjar, O., 2008. Myocardial deformation recovery from cine MRI using a nearly incompressible biventricular model. Med. Image Anal. 12, 69–85.

Burns, J.E., 2013. Automated detection of sclerotic metastases in the thoracolumbar spine at CT. Radiology 268, 69–78.

Hardisty, M., Gordon, L., Agarwal, P., Skrinskas, T., Whyne, C., 2007. Quantitative characterization of metastatic disease in the spine. Part I. Semiautomated segmentation using atlas-based deformable registration and the level set method. Med. Phys. 34, 3127–3134.

Huang, Sheng-Fang, Chian, K.-H., 2012. Automatic detection of bone metastasis in vertebrae by using CT images. Proc. World Congr. Eng. 2.

Rogelj, P., Kovacic, S., 2006. Symmetric image registration. Med. Image Anal. 10, 484–493.

Skrinskas, T., Clemons, M., Freedman, O., Weller, I., Whyne, C., 2009. Automated CT-based analysis to detect changes in the prevalence of lytic bone metastases from breast cancer. Clin. Exp. Metastasis 26, 97–103.

Tateishi, U., 2008. Bone metastases in patients with metastatic Breast cancer: morphologic and metabolic monitoring of response to systemic therapy with integrated PET/CT 1. Radiology 247, 189–196.

Thirion, J.-P., 1998. Image matching as a diffusion process: an analogy with Maxwell's demons. Med. Image Anal. 2, 243–260.

Toth, D.F., Töpker, M., Mayerhöfer, M.E., Rubin, G.D., Furtner, J., Asenbaum, U., Karanikas, G., Weber, M., Czerny, C., Herold, C.J., Rapid, R.H., 2014. Detection of bone metastasis at thoracoabdominal CT: accuracy and efficiency of a new visualization algorithm. Radiology 3, 825–833.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2008. Symmetric log-domain diffeomorphic Registration: a demons-based approach. Med. Image Comput. Comput. Assist. Intervention 11, 754–761.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: efficient non-parametric image registration. Neuroimage 45, S61–S72.

Wang, H., 2005. Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. Phys. Med. Biol. 50, 2887–2905.

Wiese, T., Burns, J., Yao, J., Summers, R.M., 2011. Computer-aided detection of sclerotic bone metastases in the spine using watershed algorithm and support vector machines. Biomed. Imaging From Nano to Macro, 2011 IEEE Int. Symp., 152–155.

Yao Jianhua, Burns, Joseph, E., Summers, R.M., 2012. Sclerotic rib metastases detection on routine CT images. Biomed. Imaging (ISBI), 2012 9th IEEE Int. Symp., 1767–1770.