Computer-Aided Evaluation of Anatomical Accuracy of Image Fusion between X-Ray CT and SPECT

Jingfeng Han\textsuperscript{a,*}, Harald Köstler\textsuperscript{b}, Christian Bennewitz\textsuperscript{c}, Torsten Kuwert\textsuperscript{c}, Joachim Hornegger\textsuperscript{a}

\textsuperscript{a}Chair for Pattern Recognition, Friedrich-Alexander University, Martensstrasse 3, D-91058 Erlangen, Germany
\textsuperscript{b}Chair for System Simulation, Friedrich-Alexander University, Cauerstrasse 6, D-91058 Erlangen, Germany
\textsuperscript{c}Clinic of Nuclear Medicine, Friedrich-Alexander University, Krankenhausstrasse 12, D-91054 Erlangen, Germany

Abstract

Hybrid scanners, which enable the performance of single photon emission computed tomography (SPECT) and X-ray computed tomography (CT) in one imaging session, have considerable diagnostic potential. However, evaluating the anatomical accuracy of image fusion inherent to these systems remains a challenge. This paper proposes a method for evaluating this variable with minimum user interaction. It focuses on measuring the distance between the centers of gravity of the SPECT hot spot and its counterpart in the CT image. A localized maximally stable extremal regions method is proposed to automatically segment SPECT hot spots, while the corresponding CT structures are segmented by the semi-automatic random walk method, based on a fast multigrid solver. Accuracy and reproducibility of the validation method have been preliminary confirmed by the test with 21 clinical data-sets.

Key words: SPECT/CT scanner, image fusion, image segmentation

* Corresponding author. Tel.: +49 9131 85 27826. Email: jingfeng@informatik.uni-erlangen.de

Preprint submitted to Elsevier Science 16 March 2008
CT SPECT/CT SPECT

Fig. 1. Two examples of misregistration of hybrid scanners. Upper row: Coronal views of somatostatin receptor SPECT and CT in a patient with a neuroendocrine carcinoma (NEC). Lower row: Transversal views of somatostatin receptor SPECT and CT in another NEC patient.

1 Introduction

Hybrid scanners combining SPECT and CT offer physicians the opportunity to acquire spatially correlated functional and morphological information in a single session. These hybrid systems have greatly improved the diagnostic accuracy and have therefore met with broad clinical acceptance (1; 2; 3; 4; 5; 6; 7; 8). However, the anatomical accuracy of this hardware-based fusion may be far from perfect. As illustrated by Fig. 1, Misregistrations in Fig.1 are due to artifacts caused by respiratory or cardiac motion as well as by patient movements occurring between the acquisition of the CT and SPECT data-sets. Phantom studies are not suitable for the validation of the hybrid scanners because it is technically difficult to simulate such complicated deformations. Due to our knowledge, the anatomical accuracy of SPECT/CT scanner has not been sufficiently validated. In a recent study (9), the accuracy of a SPECT/CT system has been preliminarily evaluated by measuring the distance between the centers of gravity of corresponding lesions in two modalities. However, reproducibility and accuracy of the validation method were not guaranteed, since the centers of gravity was interactively selected by the users. In this paper, we propose a more automatic and reproducible validation scheme for SPECT/CT hybrid scanners. A software tool has been developed to evaluate the quality of SPECT/CT fusion by measuring the distance between the centers of gravity of the hot spot in a SPECT volume and the corresponding structure in CT data-sets. A small distance between two centers of gravity indicates a high accurate fusion of SPECT/CT data-sets. The entire validation procedure has the following three steps:

(1) **Pre-selection.** The user manually selects a cubic region that contains both SPECT/CT hot spots from the fused multi-planar reformatting
Fig. 2. Segmented objects (in red) in SPECT/CT volumes.

(MPR) view. Pre-selection of the region of interest largely speeds up the following processing operations, but has no influence on the final validation results.

2) **Segmentation.** The hot spot of SPECT data is extracted by a fully automatic segmentation method based on localized maximally stable extremal regions (MSER) (10; 11; 12). The corresponding structure in CT data is segmented by a semi-automatic random walk method (13). Fig. 2 shows an example of segmented objects in SPECT/CT volumes.

3) **Measurement.** Computing the centers of gravity of the segmented SPECT/CT objects. Measuring the distance of two centers to evaluate the quality of the SPECT/CT fusion.

The validation tool has been successfully integrated into a commercial software for medical image analysis (Syngo, Siemens Medical Solutions). In the following we will introduce the segmentation methods for SPECT hot spots and for CT structures. Thereafter we will present the clinical evaluation of this validation tool.

## 2 Localized MSER segmentation for SPECT Hot Spots

The determination of the surface that separates different physiological features in functional images is difficult because of the low spatial resolution, the blurring of the edges and the high noise characteristics of functional images. Thresholding is one of the most widely used techniques to segment the volume of interest in functional image data. The threshold could have a fixed value, for example in (14) 25%, 40%, or 50% of the maximal gray level is used. The threshold could also be automatically computed for each individual image. A classic adaptive thresholding method, histogram based thresholding (15), has been widely used to segment the object in the SPECT volume (16). It deter-
mines the threshold value by maximizing the variance between the population of background voxels and object voxels. But the distinction between the background and the object class in the histogram is often too weak to find the optimal threshold. The study (17; 18) shows that the threshold of the SPECT lesions can be correctly calculated with knowledge of the size of lesions and the activity of the background. However, the prior knowledge about lesions is often not available in routine circumstances. In this paper, we present a fully automatic thresholding method for the segmentation of SPECT hot spots. The method is based on a localized MSER algorithm, which does not need any prior knowledge about the object.

The usage of mathematical notation of the algorithm is kept consistent with (10). A discrete three-dimensional image is defined as a mapping $I : D \in \mathbb{Z}^3 \rightarrow S$, where $D$ is the set of voxels and $S$ is the set of gray levels, e.g. $\{0, 1, ..., 255\}$ is a typical set for SPECT images. Region $Q$ is a continuous subset of $D$, for which each pair $(p, q) \in Q$ of pixels is connected by a path fully contained in $Q$. The algorithm starts from a seed $x_0$, typically specified by users, which is located within the hot spot. In this paper we define a localized extremal region $\mathcal{Q}_i : i \in S$ as a maximal region that contains the seed $x_0$ and $I(x) \geq i$ for all $x \in \mathcal{Q}_i$. We say $i$ is the base-level of the localized extremal region $\mathcal{Q}_i$. It is clear that the sequence of localized extremal regions $\mathcal{Q}_0, ..., \mathcal{Q}_i, \mathcal{Q}_{i+1}, ..., \mathcal{Q}_{I(x_0)}$ is nested, i.e. $\mathcal{Q}_i \supset \mathcal{Q}_{i+1}$ and $x_0 \in \mathcal{Q}_i$. For the purpose of segmentation, we are more interested in the extremal region $\mathcal{Q}_i^*$ that is maximally stable if the area variation

$$
\rho_i^\Delta = \frac{|Q_{i-\Delta}| - |Q_{i+\Delta}|}{|Q_i|}
$$

(1)

has a minimum at $i^*$. Here $|.|$ denotes cardinality and $\Delta \in \mathbb{Z}$ is a given parameter.

A straightforward way is to compute the area of all the extremal regions $\mathcal{Q}_0, ..., \mathcal{Q}_{I(x_0)}$, then to choose the region with minimum area variation. This method can successfully segment the “strong” hot spots, which have relatively cool backgrounds and have no other compatible hot objects in the neighborhood. The submandibular gland hot spot in the left of Fig. 3 belongs to this case.

However, in practice lesions frequently appear as “weak” hot spots, whose intensities are not so distinctive from the background, or some other hot objects are close to the hot spots lesion. This global maximally stable regions could not always correctly segment this class of weak hot spots. The liver lesion hot spot in the right of Fig. 3 is an example of weak hot spot. The plot of area variation with respect to various base-level clearly shows the problem: The minimum of area variation lies in the low base-level, which is far away from...
the true hot spot base-level. In order to avoid this problem, the algorithm
should not search the maximally stable region among all the extremal regions
but only among a set of localized extremal regions that always contain the
seed $x_0$ but have not yet merged with surrounding regions. The merge of the
localized extremal region with surrounding regions usually results in a dra-
matic increase of region area. It can be easily detected by the change rate of
area variation (see the last row of Fig. 3). Because the algorithm proceeds
from relatively high base-level $Q_{I(x_0)}$ to the lowest base-level, the change rate
of area variation is defined as

$$d_i^\Delta = \frac{\rho_i^\Delta - \rho_{i+1}^\Delta}{\rho_{i+1}^\Delta}. \quad (2)$$
The implementation of the method is summarized in Algorithm 1.

**Algorithm 1** Localized MSER for SPECT hot spot

1: Manually select a point $x_0$ within the SPECT hot spot as a seed
2: for $i = I(x_0), ..., 0$ do
3: compute the area of extremal region $|Q_i|$ 
4: if $i > 2\Delta$ then 
5: compute the area variation of region $\rho^\Delta_{i-\Delta}$ 
6: if $d^\Delta_{i-\Delta} > \xi$, set $i' = i - \Delta$ then break 
7: end if 
8: end for 
9: among the regions: $Q_{I(x_0)}, \ldots, Q_{i'}$, find the localized MSER $Q_{i^*}$ that has minimal area variation.

**Relationship with standard MSER.** The localized MSER is actually a simplified version of the standard MSER. In the standard MSER method, the stable analysis is performed on a region tree, which is efficiently calculated by a partition method (19). The merge of two regions has been implicitly encoded in this rooted tree structure. The additional merge detections are not necessary any more. While the localized MSER actually starts the connected thresholding from the given seed and seek a range of thresholds that leaves the peak of hot spot effectively unchanged. Although both methods theoretically can achieve almost the same segmentation of the hot spot, the localized MSER appears to be more efficient than the standard MSER method. The partition algorithm of standard MSER method has computational complexity of $O(n \log \log n)$, where $n$ is the number of voxel. While the computation time of localized MSER method is dependent on the size of the desired object. For the segmentation of a small hot spot among a large volume data, the localized MSER method becomes more efficient than the standard MSER method.

3 Random Walk for CT Structures

In this application, the largest challenge of segmentation of corresponding CT structures is the unclear boundary between the desired object and irrelevant surrounding tissues. This is demonstrated in Fig. 4.

Most fully automatic segmentation methods that do not use prior knowledge of the shape and size of the object cannot reliably locate such almost invisible boundaries. For example, the standard fast marching level set method iteratively evolves the level set interface (boundary) based on a speed function, which is typically a function of the gradient magnitude. As shown in Fig. 4b, the speed function cannot stop the level set interface crossing the invisible
Fig. 4. **a:** The sagittal view of a patient with a lesion in the neck. **b:** Gray levels represent the normalized speed function of the fast marching level set method. The red region represents the segmentation result after 100 iterations. **c:** The view of seed selection for the random walk method. The blue frame defines the region of interest. The blue lines (seeds), drawn by users, define the outside of the lesion, while the red ones define the inside of the lesion. **d:** The region segmented by the random walk method.

boundary. Given a perfect parameterization, some more advanced segmentation methods can satisfactorily segment the lesion in this example. However, it is very tricky for users to determine the optimal setting of multiple algorithm parameters for each segmentation task. For example, the well-known gradient vector flow method (20) requires the users to specify four parameters at the same time. Moreover, it is difficult to define objective stopping criteria for many fully automatic iterative methods. In practice, the user gives a fixed number of iterations or interactively stops the iterations. Due to the tedious parameterization and subjective stopping criterion of fully automatic methods, many medical researchers favor interactive segmentation, with the help of a mouse or a light pen. However, the manual selection of a three-dimensional boundary is not only time-consuming, but also non-reproducible.

Random walk segmentation (13) is chosen to detect CT structures in this application. This intuitive semi-automatic segmentation allows the user to select seeds to mark the inside and outside regions. Then the algorithm determines the optimal partition based on the image intensity and pre-selected seeds. Fig. 4c shows the selected seeds of the neck lesion and Fig. 4d shows the segmented lesion by random walk method.

The two-label (object and background) random walk is described here with a simple $3 \times 3$ image in Fig. 5. For the description of the general multi-label case, we refer to (13). In the random walk segmentation, an image is modeled as a graph of nodes and edges, where each node is a voxel and an edge connects the adjacent nodes. Assuming that the user selects node 2 and 7 as a seed of background and object respectively, for each unlabeled node $l$, the algorithm would determine the following: given a random walker starting from this node, what is the probability $u_l$ that it first reaches the object seed? Obviously the probabilities of object seed and background seed is 1 and 0. When the probabilities of unlabeled nodes are known, the boundary of the object (the curve in Fig. 5) can be approximated by interpolation between the
adjacent object nodes ($u_l > 0.5$) and background nodes ($u_l < 0.5$).

In the random walk algorithm, the image structure is represented by the edge weights:

$$w_{jk} = e^{-\beta(I(x_j) - I(x_k))^2},$$

where $I(x_j)$ and $I(x_k)$ are the intensity of node $j$ and $k$, $\beta \in \mathbb{R}$ is the only free parameter in the algorithm. Intuitively, if two adjacent voxels have a large variation of intensity, the random walker has a relatively low probability of crossing the edge between them. The probabilities of unlabeled nodes are determined by a combinatorial Dirichlet problem, in which the discrete Laplacian matrix $L$ is defined as:

$$L_{jk} = \begin{cases} 
-w_{jk} & \text{if node } j \text{ and } k \text{ are adjacent,} \\
\frac{d_j}{\sum_k w_{jk}} & \text{if } j = k, \\
0 & \text{otherwise.}
\end{cases}$$

The given seeds serve as known values ($u_2 = 0, u_7 = 1$), which can be moved into the right hand side vector. The corresponding linear system is
Generally, the system equation of the random walk segmentation is denoted as

\[ Lu = f. \] (5)

With respect to the retrospective methods, random walk segmentation has five advantages.

(1) If sufficient seeds could be correctly selected, the small objects with partially unclear boundaries can be satisfactorily segmented.

(2) Compared to a fully manual selection, the user interaction is minimal. From our experience, drawing the seeds on the middle slice in the axial, sagittal and coronal views is sufficient.

(3) The method does not require any stopping criteria. The algorithm only requires the solution of a large, sparse, symmetric positive-definite system of linear equations.

(4) The segmentation is almost real-time. A fast multigrid solver that has complexity \( O(n) \) is implemented to solve the linear system.

(5) The method has only one algorithm parameter that is kept constant in all the experiments.

Next we briefly introduce a fast node-based multigrid method to solve the linear system (5). For the definitions of mathematical notations and the detail of this multigrid solver, we refer to (21; 22) and our previous works (23; 24). Multigrid is based on the assumption that high frequency errors can be treated efficiently by an appropriate smoother on a fine scale and low frequency errors are approximated on coarser scales. Therefore we build up an image pyramid and construct the system (5) on each level. In Algorithm 2 we recursively define one multigrid iteration with initial guess \( u_h^{(0)} = 0 \).

As smoother we apply a line-wise red-black Gauss-Seidel method denoted by \( S_h \), where the parameter \( \nu \) specifies the number of performed Gauss-Seidel iterations and \( h \) is the current scale or level. Then we compute the residual \( r^h \)
Algorithm 2 Multigrid V-cycle: Compute $u_h^{(k+1)} = M_h(u_h^{(k)}, L^h, f^h, \nu_1, \nu_2)$

1: $\bar{u}_h^{(k)} = S_h^u(u_h^{(k)}, L^h, f^h)$ \{pre-smoothing\}
2: $r^h = f^h - L^h\bar{u}_h^{(k)}$ \{compute residual\}
3: $r^H = I_H r^h$ \{restrict residual\}
4: if number of coarse grid points $< \epsilon_{\text{min}}$ then
5:  Solve $L^H e^H = r^H$ exactly
6: else
7:  $e^H = M_{H}(0, L_H, r^H, \nu_1, \nu_2)$
8: end if
9: $e^h = I_h^H e^H$ \{interpolate error\}
10: $\tilde{u}_h^{(k)} = \bar{u}_h^{(k)} + e^h$ \{coarse grid correction\}
11: $u_h^{(k+1)} = S_h^u(\tilde{u}_h^{(k)}, L^h, f^h)$ \{post-smoothing\}

and restrict it to the next coarser level $H$ by a full weighting restriction operator $I_h^H$. Afterwards the so-called error equation $L^H e^H = r^H$ is solved on level $H$ recursively by multigrid. The coarse matrix $L^H = I_h^H L_h I_h^H$ is computed by Galerkin coarsening. Next the error $e^H$ is interpolated by a trilinear interpolation operator $I_h^H$ to the fine level $h$ and used there as a correction to the current solution $u^H$. Finally we apply again some post-smoothing Gauss-Seidel steps at level $h$. As stopping criterion we check, if the norm of the residual $\|r^h\|$ drops below a given threshold that we choose $10^{-8}$ in our implementation. To reduce the number of multigrid iterations we additionally use iterant recombination (25; 26) that is similar to preconditioning.

The solution of the system for a typical $40 \times 40 \times 40$ sub-volume, using the described multigrid solver with 2 pre- and 2 post-smoothing steps and 5 V-cycles, requires less than 2 seconds on an AMD Athlon 3200+ computer (2.20 GHz, 2.00 GB RAM).

4 Experimental Results

4.1 Patients

To evaluate the validation tool, 21 patients, 13 females and 8 males between 10-80 years old with the average age at 59.22, were examined by a SPECT/spiral CT scanner (Siemens Symbia system) between November 2006 and March 2007. The data-sets were selected where both the hot spot on SPECT and the corresponding structure on CT were clearly visible in the neck region. We chose adenomas of the parathyroid glands on 8 patients, the physiological accumulations of the submandibular gland on 10 patients, thyroid nodule on 1 patient, neuroendocrine tumor on 1 patient, thyroid carcinoma on 1 patient...
Fig. 6. Comparison of the ground truth shift \( (s_x, s_y, s_z) \) and the measured shift \( (d_x, d_y, d_z) \). Two shift parameters in X-, Y- and Z-directions are close to lines of identity.

for this study, where Tc-99m MIBI (18 patients), I-131-NaJ (2 patient) and In-111-SMS (1 patients) were used as tracers.

4.2 SPECT Hot Spot Segmentation

First we evaluate the localized MSER method for the segmentation of SPECT hot spots. The proposed method was compared with the histogram based thresholding (15), a widely used segmentation method for functional images (16). Two radiologists were asked to evaluate the segmented hot spots by the two methods. To avoid any bias in the ratings, the evaluation was carried out independently by the two radiologists. The segmentation results were scored between 0-2 by the radiologists: (1) If one segmentation method was obviously better than the other, this method would be scored 2 and the other method would be scored 0. (2) If both segmentation methods had comparable results, both methods would be scored 1. For the hot spots in the selected data-sets, the evaluation showed that the radiologists were much more satisfied with localized MSER segmentation: The average scores of the localized MSER method was 1.667 and 1.619, while the average scores of the histogram based thresholding was only 0.333 and 0.381, respectively.

4.3 Accuracy Test

The accuracy of the validation tool was evaluated as follows: Two operators perform the validations independently. One operator directly used the validation tool to measure the distances in X-, Y- and Z-direction \( (t_x, t_y, t_z) \) between the hot spot on SPECT and the structure on CT. In the same way, the second operator validated the SPECT/CT volumes, where the SPECT volume had been artificially shifted in X-, Y- and Z-directions. The shift parameters \( (s_x, s_y, s_z) \) were randomly generated between 5 mm and 10 mm or between
−5 mm and −10 mm. We denote the distances measured by the second operator as \((\hat{t}_x, \hat{t}_y, \hat{t}_z)\). The extent to which the ground truth shift \((s_x, s_y, s_z)\) and the measured shift \((d_x, d_y, d_z) := (\hat{t}_x - t_x, \hat{t}_y - t_y, \hat{t}_z - t_z)\) match, indicates the accuracy of the validation. As shown in Fig. 6, the experiment yielded a clear linear association between the ground truth and the measurement: The correlation coefficients are 0.9927, 0.9909 and 0.9853 in \(X\), \(Y\) and \(Z\)-directions, respectively. The anatomical inaccuracies, measured by the mean ± standard deviation of the absolute error, were reported to be 0.7189 ± 0.6298 mm in \(X\)-direction, 0.9250 ± 0.4535 mm in \(Y\)-direction and 0.9544 ± 0.6981 mm in \(Z\)-direction, respectively.

4.4 Reproducibility Test and Time Measurement

To evaluate the intra-observer reproducibility, the distances between the SPECT hot spot and CT structure were measured 20 times in five different patients, yielding a mean standard deviation of 0.2177 mm in the \(X\)-direction, 0.3039 mm in the \(Y\)-direction and 0.3350 mm in the \(Z\)-direction respectively. This indicated a high intra-observer reproducibility of the measurements of the \(X\)-, \(Y\)- and \(Z\)-distances. The mean time for a full validation process, including data loading and user operations, was less than 2 minutes on an AMD Athlon 3200+ computer (2.20 GHz, 2.00 GB RAM).

5 Summary and Discussion

The purpose of this paper is to introduce a novel approach to validate the anatomical accuracy of the SPECT/CT hybrid scanner, which is based on the segmentation of hot spot on SPECT and the corresponding structure on CT. The experimental results show that the measurement of this validation tool is sufficiently accurate and reproducible for clinical data-sets. The tool implemented in a plugin of Syngo software platform has been used by Clinic of Nuclear Medicine, Friedrich-Alexander University of Erlangen, for the purpose of research. Although we only tested the method with the data generated by the Siemens Symbia SPECT/spiral CT system in this work, the tool and the same principle can be also applied on the other SPECT/CT systems. In our future work, we plan to apply this validation tool to analyze the variation of the accuracy of hybrid scanners with respect to different patient positions, tracers or acquisition protocols.
Acknowledgment

The work is supported by Deutsche Forschungsgemeinschaft (DFG) under the grant SFB 603, TP C10. The authors thank HipGraphics Inc. for providing the software (InSpace) for volume rendering.

References

(MSVs), in: Proceeding of international conference on pattern recognition (ICPR), Hong Kong, China, 2006, pp. 63–66.


[26] I. Christadler, H. Köstler, U. Rüde, Robust and efficient multigrid techniques for the optical flow problem using different regularizers, in: Pro-