33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Hextuple registration of interim and follow-up PET-CT images for the accurate tracking of patient recovery after therapy

Norbert Zsoter, Peter Bandi, Ildiko Garai, Laszlo Papp

Abstract— **An extended registration framework is presented to accurately register follow-up PET-CT study triples. Since there are six images to register, sophisticated feature extraction and similarity measurement methods are proposed. An irregular sampling method is introduced to decrease the processing speed of the hextuple registration. The similarity measurement is based on a normalized hybrid extended SSD (Sum of Squared Differences) and and extended NMI (Normalized mutual Information). The method has been tested on a huge amount of simulated data to avoid observer specific results. Based on the validation, our method outperforms prior solutions in both speed and accuracy, hence it should be the subject of further investigations.**

I. INTRODUCTION

Positron Emission Tomography (PET) and Computed Tomography (CT) combined together bring great benefits for nuclear medicine physicians working in the field of oncology [1], [2], [3]. One of the most common related procedures is the detection of Hodgkin's lymphomas on PET-CT images acquired by a hybrid camera [2], [3]. Generally the first interim PET-CT study is followed by two follow-up PET-CT study pairs representing the stages of the patient after therapy [3], [4], [5], hence overall six images need to be processed. For the accurate comparison of these studies their spatial alignment is essential. Since lymph nodes are typically small and their size/uptake changes must be followed over the study pairs, the transformation applied to the follow-up studies must preserve the features of them. Hence non-linear distortions are not desired for these kind of images [4], [5]. The common way of evaluation is performed by placing VOIs (Volume of Interests) over the corresponding lymphomas. This step requires the presence of an experienced physician hence it is perfromed manually. Although a rigid transformation cannot eliminate all local misalignments, it is a great help for physicians to identify corresponding lymphomas by bringing them as near as possible.

The number of images makes it challenging to accurately register the studies. Prior solutions have shown that the global optima among more than two images can be achieved if all similar images are involved in the registration [5], [6]. Nevertheless the increasing number of images changes the way of calculating the similarity as well. There are several widely applied similarity measurements for both mono - and multi-modal cases. In prior works the Sum of Squared Differences (SSD) has been successfully applied between CT images [7], [8]. Other works proposed the usage of the Normalized Mutual Information (NMI) for both mono - and multi-modal cases [9], [10], [11].

In spite of the above conclusions there is still a lack of general knowledge regarding the optimal feature extraction and similarity measurement when more than two images need to be superimposed. The increasing number of images requires more computation for the registration, hence there are key factors that need to be identified: The number of iterations to superimpose the images is one of the most important factors to minimize the time of the registration. Another interesting question arises from the nature of our current medical data: Since the corresponding PET and CT images are performed by a hybrid camera, they are already registered by an intrinsic transformation. Therefore it is not clear whether all six images or only the subset of them (only CTs) need to be involved in an extended similarity measurement.

Our goal was to introduce proper feature extraction and similarity measurement methods to register PET-CT study triples and to compare them to prior solutions. We propose an extended hybrid similarity measurement performed among all images to achieve the global optimal alignment of them. The comparison and validation was performed on simulated data derived from real clinical PET-CT examples to minimize the possibility of physician specific results.

II. MATERIALS AND METHODS

A. Simulated data

Ten real PET-CT studies were collected representing the interim stage of Hodgkin's lymphoma cancer. The CT Hounsfield values were corrected to remove possible artifacts caused by earring and implant based on (1).

$$\forall v \in CT : v = \min\left(\max\left(v, -120\right), 3000\right) \qquad (1)$$

where $v \in \mathbb{Z}$ was a Hounsfield value in a CT image. Low and high threshold values were determined based on standard Hounsfield Scale to cover the fat - bone range.

All PET-CT studies were copied twice to simulate the presence of two follow-up PET-CT study pairs. For all 10 generated image sets 20 random T_1^g and T_2^g rigid transformations were generated to transform the first and second simulated follow-up studies respectively. Gaussian noise was applied to the transformed images to simulate a voxel content difference as presented in [7]. This way overall 200 simulated

N. Zsoter, P. Bandi and L. Papp are with Mediso Medical Imaging Systems Ltd., Baross str. 91-95, Budapest, Hungary norbert.zsoter@mediso.hu

I. Garai is with Medical University of Debrecen, Scanomed Ltd., Hungary garai@belklinika.com

cases were simulated. The generated transformation pairs were stored in $T^g_{1,i}$ and $T^g_{2,i}$ where $(1 \leq i \leq 200)$.

B. Methods

Four different registration methods have been investigated. The first two cases were previously proposed similarity measurements, while the last two were proposed by us. All four methods had common parameters regarding the way of sampling, multi-resolution technique, maximum number of iterations and the chosen optimization method. Overall three multi-resolution levels were defined, where the sampling was performed in equidistant spatial coordinates. The resolution of the sampling was stored in $R \subseteq \mathbb{N}_{>0}$ set where $R = \{r_1, r_2, r_3\}$. The maximum number of iterations were stored in $It \subseteq \mathbb{N}_{>0}$ set where $It = \{it_1, it_2, it_3\}$. For values of R and It see Tab. I.

Let $I = \{C_1, C_2, C_3, P_1, P_2, P_3\}$ denote one of the 200 generated cases where C_j, P_j image pairs represent a corresponding CT-PET pair and $(1 \leq j \leq 3)$. Let $\overline{o}_I \in \mathbb{R}^3$ and $\overline{e}_I \in \mathbb{R}^3$ denote the minimal origin and maximal end spatial positions of images in I.

Let $X_i \subseteq \mathbb{R}^3$ denote a coordinate set where $\overline{x} \in X_i$ if $\overline{o}_I < \overline{x} < \overline{e}_I$. Let furthermore assume that the elements of X_i characterize equidistant spatial coordinates where the distance between neighboring coordinates is $r_i \in R$.

Let $\overline{x}^T = T(\overline{x})$ denote that $\overline{x} \in \mathbb{R}^3$ spatial coordinate is transformed by T transformation and the result is $\overline{x}^T$. Let us define furthermore an *opt* optimization method which generates a (T_1, T_2) rigid transformation pair based on the similarity measurement calculated by one of the specific methods detailed below. The general framework of all four registration methods is presented in Tab. II.

Below the specific registration methods are detailed that define the way of feature extractions (if it was necessary) and the way of calculating $s \in \mathbb{R}$ similarity value.

1) eSSD: The *extended SSD* method involved all three C_1, C_2, C_3 images of the actual I set.

Let $\overline{v}^C = (C_1(\overline{x}), C_2(T_1(\overline{x})), C_3(T_3\overline{x})))$ denote a value vector which represents the sampled values of C_1, C_2 and C_3 images based on the given T_1 and T_2 transformations $(\overline{v}_C \in \mathbb{R}^3)$.

Let us define an $SC_{i,j} \subseteq \mathbb{R}^3$ set which contains all $\overline{v}^C$ samples determined in the i^{th} resolution level and in j^{th} iteration. Based on $SL_{i,j}$ the *extended SSD* measurement $eSSD(SC_{i,j})$ is defined by (2).

TABLE I
VALUES OF SAMPLING RESOLUTION AND MAXIMUM NUMBER OF ITERATIONS IN ALL RESOLUTION LEVEL

Resolution level (i)	r_i (mm)	it_i
1	16	300
2	8	200
3	4	200

where $r_i \in R$ and $it_i \in It$

TABLE II
ALGORITHMIC STEPS OF THE GENERAL REGISTRATION FRAMEWORK

1	$I := \{C_1, C_2, C_3, P_1, P_2, P_3\}$
2	*perform specific feature extraction steps*
3	*set* T_1, T_2 *identical*
4	$\forall r_i \in R$:
5	*calculate* X_i *based on* I *and* r_i
6	$\forall it_j \in It$:
7	$\forall \overline{x} \in X_i$:
8	$\overline{x}_{T_1} = T_1(\overline{x})$
9	$\overline{x}_{T_2} = T_2(\overline{x})$
10	*calculate specific* s *similarity measurement*
11	$(T_1, T_2) = opt(s)$

where *opt* was chosen to be the Nelder-Mead optimization method [12]

$$\sum_{\overline{v^C} \in SC_{i,j}} \frac{(v_1^C - v_2^C)^2 + (v_1^C - v_3^C)^2 + (v_2^C - v_3^C)^2}{n} \tag{2}$$

where $\overline{v^C} = (v_1^C, v_2^C, v_3^C)$ and $n = |SC_{i,j}|$.

2) eNMI: The *extended NMI* method involved all three C_1, C_2, C_3 images of the actual I set.

Assuming that the definition of $SC_{i,j}$ is the same as in the previous section, the *extended NMI* measurement $eNMI(SC_{i,j})$ is defined by (3).

$$\frac{\sum_{\overline{v} \in SC_{i,j}} \sum_{k=1}^{3} p(v_k) \log p(v_k)}{\sum_{\overline{v} \in SC_{i,j}} p(v_1, v_2, v_3) \log p(v_1, v_2, v_3)} \tag{3}$$

where $p(v_k) \in \mathbb{R}$ denotes the probability of the v_k value occurrence in the elements of $SC_{i,j}$ and $p(v_1, v_2, v_3) \in \mathbb{R}^3$ denotes the probability of the value vector $p(v_1, v_2, v_3)$ occurrence in $SC_{i,j}$.

3) eSSD Laplace: Current method built on the idea of operating with second derivative images having great differences and sign changes that were typically located at the edges. This way the contrast between a "good" and a "bad" fit could be increased which could led to a more accurate similarity measurement. The *extended SSD Laplace* method involved all three C_1, C_2, C_3 images of the actual I set.

Let $X^{C_k} \subseteq \mathbb{R}^3$ denote a set which contains all voxel coordinates of image C_k where $(1 \leq k \leq 3)$. The CT images were preprocessed by a three dimensional Laplacian operator [13] with a 3D $3 \times 3 \times 3$ voxel sized kernel as defined by (4) is order to provide a second derivative image. The result of the operator on C_j image was stored in C_j^L image where $(1 \leq j \leq 3)$.

$$\begin{aligned} &\forall (x, y, z) \in X^{C_k} : \\ &C_k^L(x, y, z) = 27 \cdot C_k(x, y, z) - \\ &\sum_{a=-1}^{1} \sum_{b=-1}^{1} \sum_{c=-1}^{1} C_j(x + a, y + b, z + c) \end{aligned} \tag{4}$$

where $\forall a,b,c : (x+a, y+b, z+c) \in X^{C_k}$ is true and $(1 \leq k \leq 3)$.

Let $\overline{v}^L = \left(C_1^L(\overline{x}), C_2^L(T_1(\overline{x})), C_3^L(T_2(\overline{x}))\right)$ denote a value vector which represents the sampled values of C_1^L, C_2^L and C_3^L images based on the given T_1 and T_2 transformations. Let us define an $SL_{i,j}$ set which contains all $\overline{v}^L$ samples determined in the i^{th} resolution level and in j^{th} iteration. Based on $SL_{i,j}$ the *extended SSD Laplace* measurement $eSSD_L(SL_{i,j})$ is defined by (5).

$$\sum_{\overline{v} \in SL_{i,j}} \frac{(v_1 - v_2)^2 + (v_1 - v_3)^2 + (v_2 - v_3)^2}{n} \tag{5}$$

where $\overline{v} = (v_1, v_2, v_3)$ and $n = |SL_{i,j}|$.

4) Hextuple: The *hextule* method involved all six $C_1, C_2, C_3,\ P_1, P_2, P_3$ images of the actual I set. The CT images were preprocessed by a three dimensional Laplacian operator as defined by (4) to generate C_j^L for the corresponding C_j images where $(1 \leq j \leq 3)$.

Current method measured the *extended SSD* over C_j^L images and the *extended NMI* over P_j images by (5) and (3) respectively.

Since the similarity values of an $eSSD_L$ and an $eNMI$ measurement were very different, they were normalized to properly include them into one similarity measurement. Assuming that the initial spatial alignment of the images gives worse similarity values than the optimal one, reference similarity values $eSSD(SCL_{i,0})$ and $eNMI(SP_{i,0})$ were calculated in all i resolution level.

The final similarity value $hextuple(SL_{i,j}, SP_{i,j})$ was calculated by (6).

$$\frac{eSSD_L(SL_{i,j})}{|eSSD_L(SL_{i,0})|} + \frac{eNMI(SP_{i,j})}{|eNMI(SP_{i,0})|} \tag{6}$$

Note that both $eSSD_L$ and $eNMI$ values were normalized by their values given in the first initial iteration. This way both values were present in the measurement with equal weights.

C. Validation

The above detailed methods have been performed on all 200 simulated image sets. The result (T_1, T_2) transformation pair of the given method have been compared to the reference (T_1^g, T_2^g) pair by subtraction. The parameter difference of the given (T_1, T_2) and (T_1^g, T_2^g) was stored in δ_i in mm $(1 \leq i \leq 12)$. Translation and rotation mean (μ_δ^{tr}), $\mu_\delta^{rot})$ and standard deviation (σ_δ^{tr}), (σ_δ^{rot}) of δ_i errors were measured for all methods. Furthermore the sum of the ϵ_j spatial volume corner distances were measured $(1 \leq j \leq 8)$ and their mean (μ_ϵ) and standard deviation (σ_ϵ) was calculated as well for all methods.

III. RESULTS

According to current results the normalized hybrid measurement provided by our *hextuple* method gave the best fit of the images, while the second best result was achieved by the *eSSD Laplace* method. The worst result was provided by *eSSD* which was significantly outperformed by the *eNMI* method. For detailed results see Tab. III, Tab. IV, Tab. V and Fig 2.

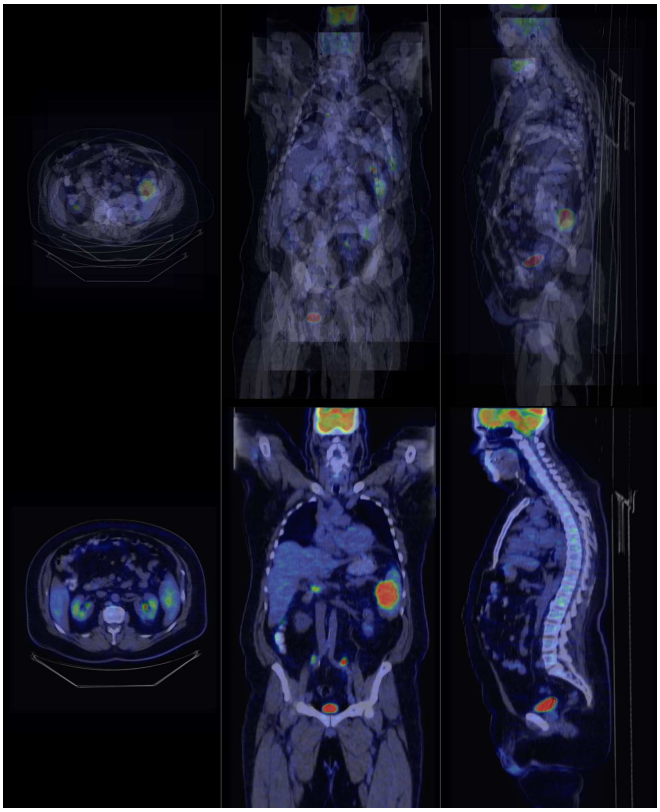

Fig. 1. Orthogonal fusion views of an example I set having the initial stage (top) and the final stage (bottom) provided by our *hextuple* method.

IV. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

We have compared our methods with the extended version of some well known similarity measurements that are gold standards in the field of medical image registration. Working with the above simulated data led us to the conclusion that the known behavior of prior methods may change if the number of images are more than two. This means that additional feature extraction steps are required to correctly handle the increasing number of images. Current research registered three and six images as well and successfully shown that the increasing number of information requires

TABLE III

TRANSLATION ERRORS OF DIFFERENT METHODS

	Resolution levels		
Method	r_1	r_2	r_3
$eSSD$	22.12± 28.47	17.57± 27.1	14.78± 26.27
$eNMI$	8.14± 7.35	6.45± 6.07	5.53± 5.74
$eSSD_L$	8.71± 6.75	5.42± 5.04	4.02± 3.91
$hextuple$	6.87± 5.63	4.24± 3.67	2.91± 2.74

Where the values in cells denote the $\mu_\delta^{tr} \pm \sigma_\delta^{tr}$ error in mm.

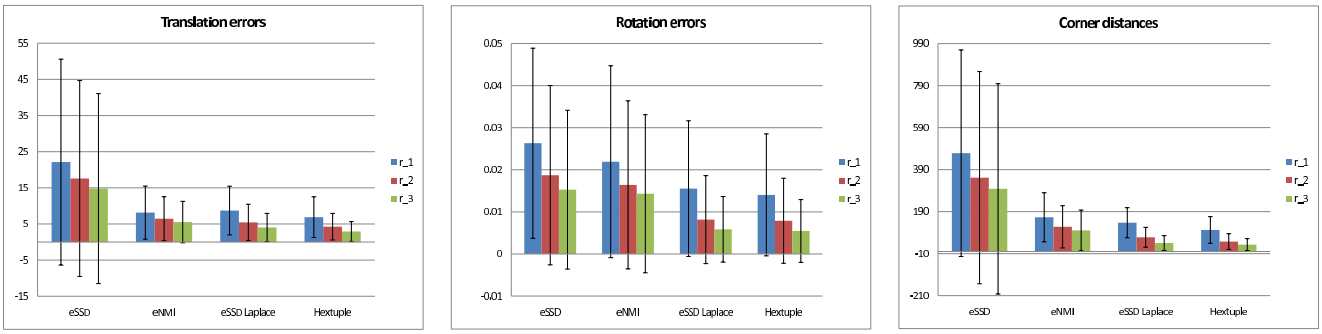

Fig. 2. Translation, rotation and corner distance errors represented for all methods in all r_i resolution level. The $eSSDLaplace$ and $hextuple$ measurements gave the best fit. Conventional $eSSD$ method worked with significantly greater error. The greatest differences appeared in the translational parameters.

TABLE IV

ROTATION ERRORS OF DIFFERENT METHODS

	Resolution levels		
Method	r_1	r_2	r_3
$eSSD$	0.026± 0.022	0.018± 0.021	0.015± 0.018
$eNMI$	0.021± 0.022	0.016± 0.019	0.014± 0.018
$eSSD_L$	0.015± 0.016	0.008± 0.010	0.005± 0.007
$hextuple$	0.014± 0.014	0.007± 0.010	0.005± 0.007

Where the values in cells denote the $\mu_\delta^{rot} \pm \sigma_\delta^{rot}$ error in mm.

TABLE V

CORNER DISTANCE ERRORS OF REGISTRATION METHODS

	Resolution levels		
Method	r_1	r_2	r_3
$eSSD$	468.33± 491.76	351.84± 505.31	298.69± 500.79
$eNMI$	163.02± 116.92	117.74± 100.59	100.76± 97.09
$eSSD_L$	137.21± 71.74	67.98± 47.54	40.46± 35.27
$hextuple$	102.91± 63.29	47.51± 37.36	32.89± 28.61

Where the values in cells denote the $\mu_\epsilon \pm \sigma_\epsilon$ error in mm.

more sophisticated, more complex similarity measurements. Our *hextuple* method provided the best results, hence it is a logical decision to involve all images into a more complex similarity measurement. Although the $hextuple$ method provided the best results, the second best $eSSD_L$ method gave acceptable results as well. Taking into consideration that the $hextuple$ method needed the most computation, specific solutions may choose our $eSSD_L$ method if the time of the whole process is a sensitive issue.

B. Future Works

Based on current results, we will collect more medical data where non-linear transformations are desired to evaluate our methods for those cases as well. We will also focus on the speed optimization of our *hextuple* method. An obvious way is the investigation of only those values located in a specific distance from edges where the Laplacian image has a sign change. Since the Laplacian image is available in our *hextuple* method, the localization of these positions does not require significantly more computations.

In theory all similarity measurements can be furthermore extended. Applications that build up phantoms from a huge amount of real medical data could take advantage from current results if multi-modal image series need to be registered.

REFERENCES

[1] L.S. Freudenberg, G. Antoch, P. Schtt, et al., FDG-PET/CT in re-staging of patients with lymphoma, *Eur J Nucl Med Mol Img*, vol. 3, 2004, pp. 325-329.

[2] E. Miller, U. Metser, G. Avrahami, et al., Role of 18F-FDG PET/CT in Staging and Follow-up of Lymphoma in Pediatric and Young Adult Patients, *J Comp As Tom*, vol. 30, 2006, pp. 689-694.

[3] W. C. Lavely, D. Delbeke, J. P. Greer, et al., FDG PET in the follow-up management of patients with newly diagnosed Hodgkin and non-Hodgkin lymphoma after first-line chemotherapy?, *Int J Rad Oncology*, vol. 57, 2003, pp. 307-315.

[4] C. S. Portlock, FDG-PET after two cycles of chemotherapy predicts treatment failure and Progression-free survival in Hodgkins lymphoma, *Curr. Oncol. Rep.*, vol. 8, 2006, pp. 356-357.

[5] S. K. Balci, P. Golland, M. Shenton, et al., Free-Form B-spline Deformation Model for Groupwise Registration, *Med Image Comput Comput Assist Interv, MICCAI*, vol. 10, 2007, pp. 23-30.

[6] C. Wachinger, N. Navab, Structural image representation for image registration, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010, pp. 23-30.

[7] Y. Papastavrou, D. Cash, D. Hawkes, et al., A Multi-component similarity measure for improved robustness of non-rigid registration of combined FDG PET-CT head and neck images, *4th European Conference of the International Federation for Medical and Biological Engineering, IFMBE Proceedings*, vol. 22, 2009, pp. 433-435.

[8] Y. Yin, C.-L. Lin and E. A. Hoffman, Mass preserving nonrigid registration of CT lung images using cubic B-spline, *Med. Phis*, vol. 36, 2009, pp. 42134222.

[9] M. M. Coselmon, J. M. Balter, D. L. McShan, et al., Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines, *Med Phys*, vol. 31, 2004, pp. 2942-2948.

[10] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, et al., Mutual-information-based registration of medical images: a survey, *IEEE Trans Med Img*, vol. 22, 2003, pp. 986 - 1004.

[11] H. Chen and P. K. Varshney, Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation, *IEEE Trans Med Img*, vol. 22, 2003, pp. 1111-1119.

[12] D. V. Sompel, M. Brady, Simultaneous reconstruction and registration algorithm for limited view transmission tomography using a multiple cluster approximation to the joint histogram with an anatomical prior, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009*, 2009, pp. 5733-5736.

[13] Z. Deng, T. H. Nishimura, Medical image noise reduction using Radon transform and Walsh list in Laplacian pyramid domain, *IEEE 13th International Symposium on Consumer Electronics, 2009. ISCE '09*, 2009, pp. 756-760.