Improved elastic medical image registration using mutual information

Konstantin Ens[a,b], Hanno Schumacher[a], Astrid Franz[b] and Bernd Fischer[a]

[a]Institute of Mathematics, University of Luebeck, Wallstrasse 40, 23560 Luebeck, Germany;
[b]Philips Research Laboratories, Roentgenstrasse 24-26, 22335 Hamburg, Germany

ABSTRACT

One of the future-oriented areas of medical image processing is to develop fast and exact algorithms for image registration. By joining multi-modal images we are able to compensate the disadvantages of one imaging modality with the advantages of another modality. For instance, a Computed Tomography (CT) image containing the anatomy can be combined with metabolic information of a Positron Emission Tomography (PET) image. It is quite conceivable that a patient will not have the same position in both imaging systems. Furthermore some regions for instance in the abdomen can vary in shape and position due to different filling of the rectum. So a multi-modal image registration is needed to calculate a deformation field for one image in order to maximize the similarity between the two images, described by a so-called distance measure.

In this work, we present a method to adapt a multi-modal distance measure, here mutual information (MI), with weighting masks. These masks are used to enhance relevant image structures and suppress image regions which otherwise would disturb the registration process. The performance of our method is tested on phantom data and real medical images.

Keywords: medical image registration, statistical methods, mutual information, weighted distance measure, elastic transformations

1. INTRODUCTION

Registration of medical images is an active field of current research and still constitutes one of today's most challenging image processing problems.[1–4] In basic terms, registration is the process of finding a geometric transformation between two or more images such that corresponding image structures align correctly. These images may have been acquired with the same or different imaging modalities, at the same or different times, from one or several patients. Accurate image registration is a necessary prerequisite for many diagnostic and therapy planning procedures where complementary information from different images has to be combined.

As soon as different imaging modalities are involved, corresponding structures may be difficult to identify and match, especially when images stemming from modalities that are based on totally different physical imaging principles have to be registered onto each other. An example is the registration of Computed Tomography (CT) or Magnetic Resonance (MR) images with Positron Emission Tomography (PET) images. Whereas the CT and MR images show the anatomic structure, the PET image indicates regions of high metabolism, which may correspond to tumors, but does not contain detailed information about the anatomy. In order to correctly localize the tumor, the PET image has to be correctly aligned with a CT or MR image in order to decide, for instance, whether the tumor is located in the lung parenchyma or has penetrated the diaphragm, an information that may substantially change staging and further patient management.

But even if images of one and the same modality have to be registered, the solution strongly depends on whether the images have been acquired simultaneously or at different times. X-ray images taken simultaneously, but from different projection angles will show identical structures in different shapes and geometrical arrangements. CT follow-up studies consisting of images taken weeks or months apart to monitor tumor growth or treatment response pose the problem of registering structures changing in shape due to growth or shrinkage as

Further author information: (Send correspondence to Konstantin Ens)
E-mail: ens@math.uni-luebeck.de, Telephone: +49 451 7030 419

well as in position due to different patient positioning. Even when sampled more frequently in a cine acquisition to monitor the uptake of a contrast agent by a tumor, the structure of interest changes in shape and size due to enrichment of the contrast agent as well as in position due to patient and organ motion.

These very few, but characteristic examples may suffice to give a feeling for the complexity and variety of registration tasks in the medical field and indicate why no single registration method has succeeded in solving all of them so far. As of yet, each application demands for a dedicated registration algorithm. An optimal registration requires the incorporation of characteristics of the underlying application. The ideas, worked out in this paper, arose from the requirements of lung cancer diagnostics. For a successful and minimal as possible invasive treatment such as radiation therapy, finding the exact location of a tumor in chest is crucial. Due to the nature of the creation process of PET images they do not display detailed anatomical features and have only a low resolution. Therefore, one tries to combine PET images with CT images to use their detailed anatomic information. But in most cases a simple combination of these images is inaccurate, for example due to respiratory motion a discrepancy between the tumor position in the CT image and the tumor position in the PET image is possible. More registration errors produced in combined PET-CT images can be found in the literature.[5] It is the goal of this paper to devise techniques to overcome this intrinsic problem. The idea is to incorporate certain masks within the distance measure of a parameter-free registration procedure. Another application is the registration of images where one is only given incomplete (post-surgery image).[6]

Due to the high variability of soft tissue, nonlinear registration techniques are employed. More precisely, our starting point are nonlinear registration schemes which are based on a variational principle[1] using a parameter-free transformation description. These schemes minimize the following functional

$$J(\mathbf{u}; \mathbf{R}, \mathbf{T}) = \alpha \cdot S(\mathbf{u}) + D(\mathbf{R}, \mathbf{T_u}) \tag{1}$$

with respect to the sought-after transformation $\mathbf{u}$. The functional J in (1) consists of two main building blocks:

- The distance measure D describes the similarity of the reference image $\mathbf{R}$ and the image $\mathbf{T_u}$, which is the template image $\mathbf{T}$, transformed with the displacement field $\mathbf{u}$. The goal of image registration is to make the two images as similar as possible. The distance measure has to be chosen with respect to the underlying medical application. For monomodal registration tasks the sum of squared grey value differences (SSD) can be applied, since it assumes the same grey value structure in both images. If this prerequisite is not fulfilled, but at least a linear dependency between the grey values can be assumed, the cross correlation (CC)[7] can be used. If even this linear dependency is not given, as it is the case in multimodal registration tasks, entropy-based measures as the mutual information (MI)[8] have to be chosen.

- When only optimizing the image similarity, the problem (1) is ill-posed, since a transformation may consist of independent transformation vectors for each image pixel. In clinical applications suitable transformations have to fulfill certain smoothness constraints. Therefore a smoother S is incorporated in (1). Widely used smoothers include the so-called diffusive,[9] elastic,[10] and the curvature functional.[11,12] The smoother has to be chosen with respect to the imaged anatomical region, as different organs can undergo different kinds of deformations.

The regularization parameter α controls the impact of the two individual ingredients with respect to the overall functional.

This paper is based on the elastic smoother as it reflects physical properties of soft tissue. However, the developed techniques are applicable to any other choice as well. Furthermore, we restrict our attention to the MI based distance measure for registration of multimodality images. The weighted SSD based distance measure was investigated previously.[13] Images from two applications are considered. As a first application we consider CT and ultrasound (US) images of a phantom. As second application we investigate the registration of independently acquired PET and CT images (see fig. 4 and 6) for multimodality cancer diagnostic.

The idea is to supplement the MI distance measure with image dependent weights in order to achieve a suitable registration. Throughout this paper, it is assumed that the needed segmentation is at hand. Note, that the main work in any scheme, based upon the variational formulation (1), is the repeated solution of a linear

system of equations $A\mathbf{u} = f$, were A corresponds to the Gâteaux derivative of the smoother S and f to the derivative of the distance measure D, respectively. Thus changing the distance measure does not change the coefficient matrix A. Consequently, we may take advantage of any fast solution schemes[14, 15] developed for the matrix A.

Note, that the introduction of weighting masks is more powerful than using point landmarks. With a weighting mask over a segmented image area, the registration result in this whole area can be influenced, for instance in such a way that no displacements occur in this whole area. Point landmarks on the other hand could only avoid a displacement at exactly the landmark positions, but would allow displacements at inter-landmark locations.

The paper is organized as follows. Section 2 presents a mathematical formulation of the underlying problem which is used to derive an appropriate scheme. The following section 3 illustrates the power of our new scheme with an academic example as well as with a real life example. Finally, some conclusions are drawn in section 4.

2. METHOD

Given a reference image $\mathbf{R}$ and a template image $\mathbf{T}$ we are searching for a deformation field $\mathbf{u}$ to minimize the difference between both images using (1). For the ease of presentation, we restrict ourselves to the elastic smoother

$$S(\mathbf{u}) = \int_{\Omega} \frac{\mu}{4} \sum_{i,j=1}^{d} \left(\partial_{x_i}\mathbf{u}_j + \partial_{x_j}\mathbf{u}_i\right)^2 + \frac{\lambda}{2}(\text{div } \mathbf{u})^2 \; dx, \tag{2}$$

where d is the dimension of the images, μ, λ are so-called Lamé-constants describing the elasticity, and $\Omega =]0,1[^d$ is a compact support of both images $\mathbf{R}, \mathbf{T} : \Omega \to \mathbb{R}$. The distance measure D is chosen as mutual information (MI),[8] a well known concept from information theory and the standard distance measure in multi-modal image registration. Thus

$$D(\mathbf{u}; \mathbf{R}, \mathbf{T}) = -\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{p}_{\mathbf{RT_u}}(r,t) \log\left(\frac{\mathbf{p}_{\mathbf{RT_u}}(r,t)}{\mathbf{p}_{\mathbf{R}}(r)\mathbf{p}_{\mathbf{T_u}}(t)}\right) \; dr \; dt, \tag{3}$$

where $\mathbf{p}_{\mathbf{R}}(\cdot)$, $\mathbf{p}_{\mathbf{T_u}}(\cdot)$ are marginal probability distributions of the gray values in the images $\mathbf{R}$ and $\mathbf{T_u}$ and $\mathbf{p}_{\mathbf{RT_u}}(\cdot,\cdot)$ stands for the joint gray-value probability distribution of both images. In order to weight or insert important image structures and/or to suppress irrelevant image parts, weighting masks $\mathbf{m_R}$ and $\mathbf{m_T}$ for the reference and template are incorporated into the distance measure

$$D(\mathbf{u}; \mathbf{R}, \mathbf{T}, \mathbf{m_R}, \mathbf{m_T}) = -\int_{\mathbb{R}^2} \left(\mathbf{p}_{\mathbf{RT_u}}(r,t) + \widehat{\mathbf{p}}_{\mathbf{m}_{\mathbf{RT_u}}}(r,t)\right) \log\left(\frac{\left(\mathbf{p}_{\mathbf{RT_u}}(r,t) + \widehat{\mathbf{p}}_{\mathbf{m}_{\mathbf{RT_u}}}(r,t)\right))}{(\mathbf{p}_{\mathbf{R}}(r) + \widehat{\mathbf{p}}_{\mathbf{m_R}}(r))(\mathbf{p}_{\mathbf{T_u}}(t) + \widehat{\mathbf{p}}_{\mathbf{m}_{\mathbf{T_u}}}(t))}\right) \; dr \; dt. \tag{4}$$

Here $\widehat{\mathbf{p}}_{\mathbf{m_R}}(\cdot)$, $\widehat{\mathbf{p}}_{\mathbf{m}_{\mathbf{T_u}}}(\cdot)$ and $\widehat{\mathbf{p}}_{\mathbf{m}_{\mathbf{RT_u}}}(\cdot,\cdot)$ are modified distributions in such a way that the standardizing of the histogram is not hurt, and that

$$\int_{\mathbb{R}^2} \widehat{\mathbf{p}}_{m_{\mathbf{RT}}}(\mathbf{r},\mathbf{t}) \; \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{t} = \int_{\mathbb{R}} \widehat{\mathbf{p}}_{m_{\mathbf{R}}}(\mathbf{r})\mathrm{d}\mathbf{r} = \int_{\mathbb{R}} \widehat{\mathbf{p}}_{m_{\mathbf{T}}}(\mathbf{t})\mathrm{d}\mathbf{t} = 0 \text{ and}$$
$$\int_{\mathbb{R}^2} |\widehat{\mathbf{p}}_{m_{\mathbf{RT}}}(\mathbf{r},\mathbf{t})|\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{t}, \int_{\mathbb{R}} |\widehat{\mathbf{p}}_{m_{\mathbf{R}}}(\mathbf{r})|\mathrm{d}\mathbf{r}, \int_{\mathbb{R}} |\widehat{\mathbf{p}}_{m_{\mathbf{T}}}(\mathbf{t})|\mathrm{d}\mathbf{t} \leq 1$$

holds. Thus for $\widehat{\mathbf{p}}_{\mathbf{m_{Image}}}(\cdot)$ also negative values can be used. The definition of the masks in (4) are as follows. Let $\mathbf{I} : \Omega \to \mathbb{R}$ be an image then we define the weighting mask for $\mathbf{I}$ as $p_{\mathbf{m_I}} : \mathbb{R} \longrightarrow \mathbb{R}$, which is a mask in the histogram domain. To clarify the definition of the masks $\widehat{\mathbf{p}}_{\mathbf{m_{Image}}}(\cdot)$ an example is presented in fig. 1. This example also shows that additive masks in the histogram domain are equipotent with additive masks in the image domain. Due to this a mask can be defined in the image domain and transfered to the histogram domain afterwards. The construction of the masks $\mathbf{m_R}$ and $\mathbf{m_T}$ requires accurate prior knowledge of the images. With this knowledge one can fade out image structures to avoid local sub optima during the optimization process or weight structures to amplify their relevance for the registration process. Due to the fact that we are working with additive masks one way to select a mask in the image domain by use the difference between the actual gray value of the given pixel of the image and the needed gray value at this pixel. Afterwards the mask can

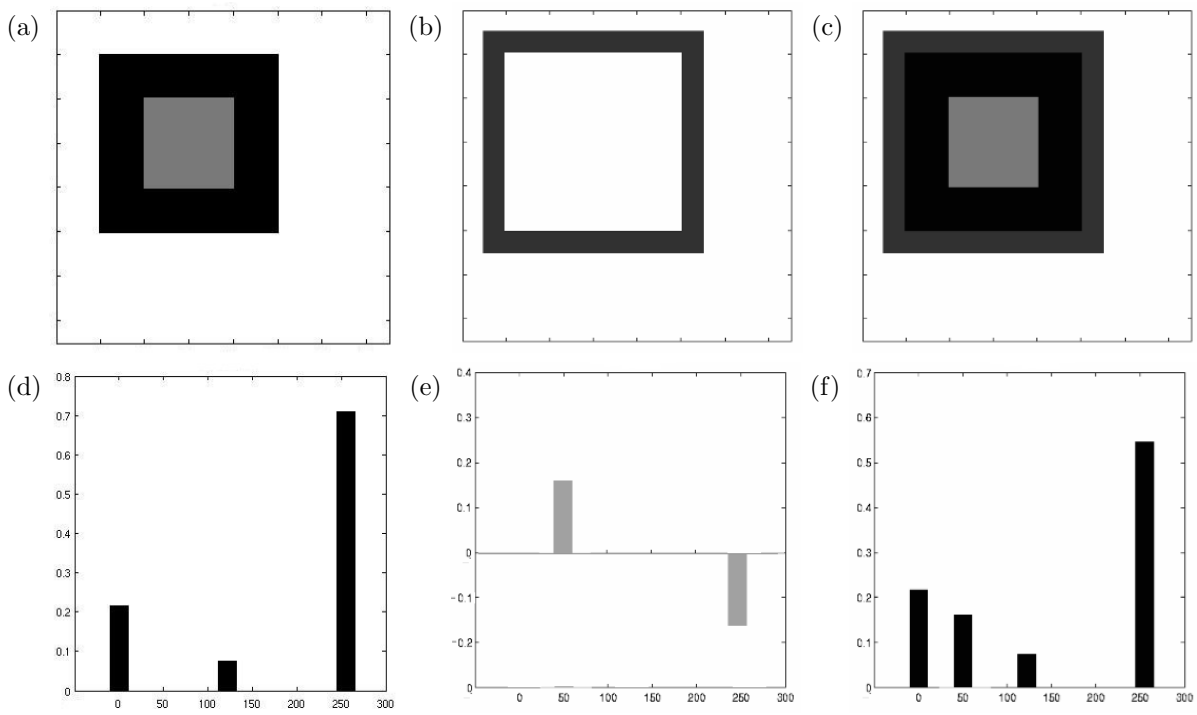

Figure 1. Masking an image can be also understood as a masking of its pertinent histogram. To clarify this point an image $\mathbf{I}$ (a) is masked with $\mathbf{m_I}$ (b). The resulting image is shown in (c). The same masking process can also be done in the histogram domain. Therefore the mask (b) is transfered to the histogram domain (e) and added to the histogram of image $\mathbf{I}$ (d). The result can be seen in (f). One should note that (f) is the histogram of (c).

be transfered from the image domain (gray values change) to the histogram domain (change of quantity of gray values for one image).
The weighted distance measure (4) leads to the modified functional

$$J(\mathbf{u};\mathbf{R},\mathbf{T}) = \alpha \cdot S(\mathbf{u}) + D(\mathbf{u};\mathbf{R},\mathbf{T},\mathbf{m_R},\mathbf{m_T}). \tag{5}$$

which has to be minimized with respect to the displacement field $\mathbf{u}$. To obtain a necessary condition for a minimizer, the Gâteaux derivative of the functional (5) is computed. The derivative of the improved distance measure (4) looks like

$$\mathbf{f}[\mathbf{x},\mathbf{u}(\mathbf{x})] = (G_\beta * L_{MI,\mathbf{u}}(\mathbf{R},\mathbf{T}))\,\nabla\mathbf{T}((\mathbf{x})-\mathbf{u}(\mathbf{x})) + (G_\beta * L_{MI,\mathbf{u}}(\mathbf{m_R},\mathbf{m_T}))\,\nabla\mathbf{m_T}((\mathbf{x})-\mathbf{u}(\mathbf{x})) \tag{6}$$

where

$$L_{MI,\mathbf{u}}(\mathbf{R},\mathbf{T},r,t) = \left(\frac{\partial_2 \mathbf{p_{RT}}(r,t)}{\mathbf{p_{R,T}}(r,t)} - \frac{\mathbf{p'_T}(t)}{\mathbf{p_T}(t)}\right),\ L_{MI,\mathbf{u}}(\mathbf{m_R},\mathbf{m_T},r,t) = \left(\frac{\partial_2 \mathbf{p_{m_{RT}}}(r,t)}{\mathbf{p_{m_R,m_T}}(r,t)} - \frac{\mathbf{p'_{m_T}}(t)}{\mathbf{p_{m_T}}(t)}\right),$$

and G_β is a Gaussian window with variance $\beta > 0$. The function $\mathbf{f}$ is often called the outer forces and drives the registration process. So, with the Gâteaux derivative $\mathbf{f}$ of the weighted distance measure and the Gâteaux derivative $\mathbf{A}$ of the smoother S we are coming up with the partial differential equation,

$$\mathbf{A}\mathbf{u} = \mathbf{f}. \tag{7}$$

Finally to work with our new weighted distance measure in the overall registration problem, we make use of a discretization with finite differences of eq. (7) to solve it for example with a fix-point iteration.[1]

3. RESULTS

In this section we compare our new improved approach of image registration with a standard unmasked method of image registration. We start out with an academical example in section 3.1, followed by the investigation of phantom data in section 3.2. Afterwards we move on to a real life example in section 3.3 given by PET and CT images recorded from the same patient.

3.1. An academical example

First we discuss an academic example. Figure 2 (a) illustrates the reference and figure 2 (b) the template image, respectively. We use the reference image without a mask in both cases. The inner structure of this image has gray value 255, the outer one 100, and the background 0. For the weighted registration a considered template image is used. Note that the inner black structure of this image has a gray value of 255, the outer dark gray structure of 245 and the white background has grey value 0. Hence, the grey value difference between the inner and the outer structure is negligible as compared to the grey value difference between the outer structure and the image background. The achievement of unmasked registration is deficient (see Figure 2 (d)). One can see

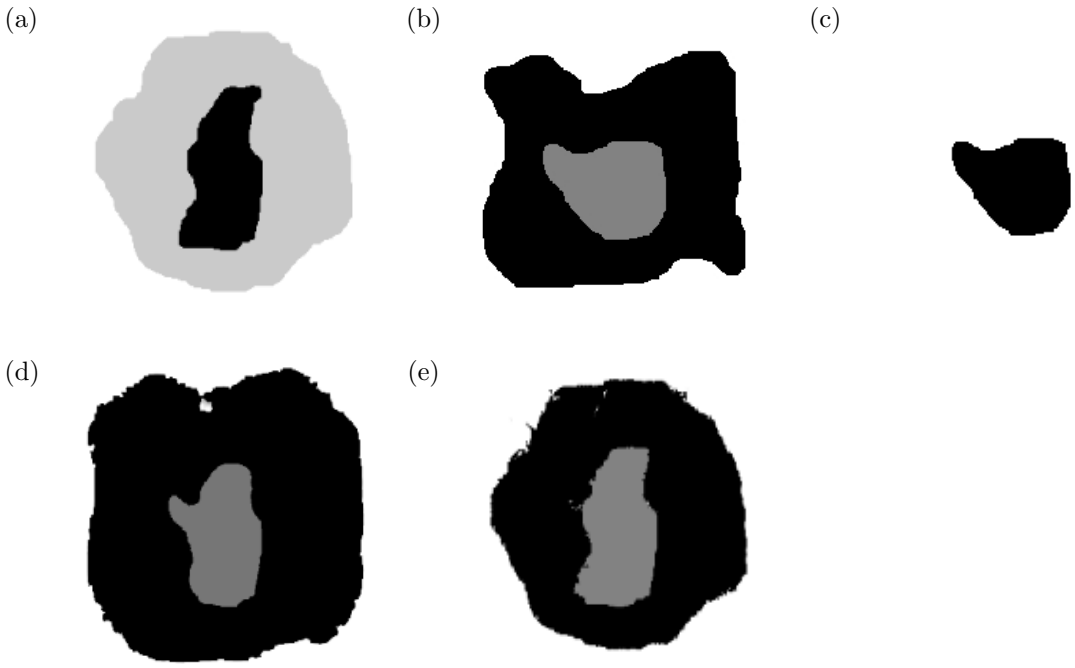

Figure 2. The unweighted and weighted registration of an academical example. The reference image is shown in (a) it consists of gray value 255 at the inner structure, 100 at the outer structure, and 0 in the background. The template image in (b) has gray value 245 at the inner structure, 255 at the outer structure, and 0 in the background. The mask of template image is presented in (c) with a factor -120 and 0 in the background. The result of an elastic registration is demonstrated in (d), the result of a weighted elastic registration in (e).

in the result image that the inner structures not moving to correct position and outer structures is bigger as in reference image. The mask of template image (see fig. 2(c)) must aid the difference between gray-values of inner and outer structures. We used the mask with value -120 for inner structure and leaved the outer structure and background without a mask. A masked registration is able of calculating an almost perfect result (see 2 (e)). Both registrations use the parameter set $\alpha = 10^{-4}$, $\lambda = 1$, $\mu = 2$, $\beta = 7$ and 20 bins. They were terminated after 50 steps using a fix-point iteration.[1]

3.2. A phantom example

The second example illustrates the application of weighting masks for the comparison of CT and ultrasound (US) images of a phantom (figs. 3(a) and (c)). The US image (fig. 3(c)) contains black colored regions corresponding to hidden image areas behind bones. These regions are visible in the CT image (fig. 3(a)). Thus the registration is affected negatively. In order to avoid this misleading one can mask the CT image, as it can be seen for example in fig. 3(b). This is done by masking out the bones and insert reflection artifacts at the border of the bones in

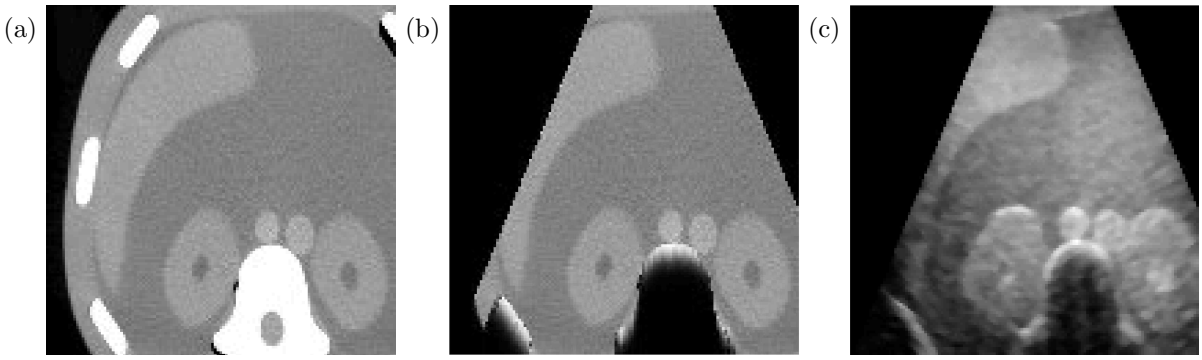

Figure 3. CT and US images of a phantom. US image (c) contain (e.g. behind the bones) black colored regions. These regions are missing in the CT image (a). Thus the registration is affected negatively. Around to avoid one can mask the CT image. As it is to be seen for example in the figure (b).

the CT image such that those areas correspond to the shadow areas and reflection artifacts in the US image. Additionally image part are fade out in the CT image, that can not be seen in the field of vision of the ultrasound sensor head. Due to the fact that these two images were generated from the same phantom, the registration is very easy and not illustrated here.

3.3. A clinical example

To demonstrate the performance of our new approach for real clinical problems, we applied an elastic registration with the weighted MI to register CT (fig. 4(a)) and PET (fig. 4(b)) thorax images used for cancer diagnostic. This allows to use the anatomical information of the CT image and the physiological characteristics of a tumor gained from a PET image. The unweighted registration of these images is shown in fig. 4(c). One can see in the result image that the lungs are clearly smaller, compared to the reference (fig. 4(a)), and the heart is deformed, compared to the template (fig. 4(b)). In order to improve this image registration, weighting masks are used.

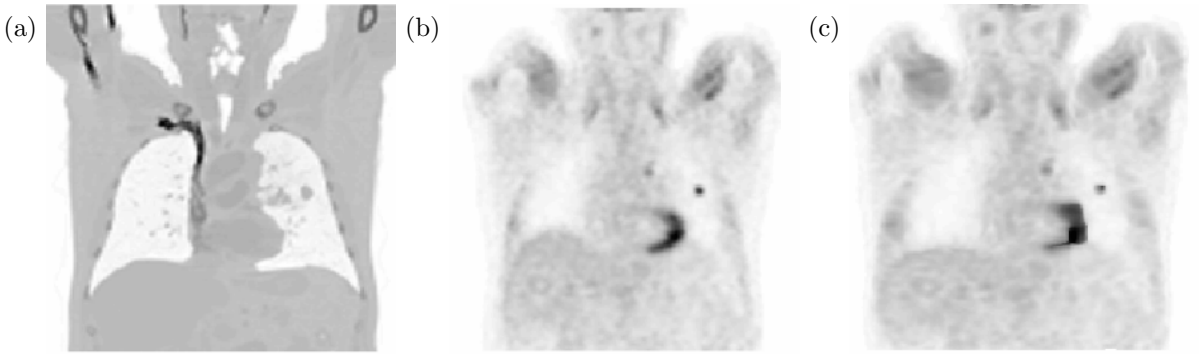

Figure 4. The unweighted registration of thorax CT and PET images for cancer diagnostic. The unmasked reference image **R** (CT image) is presented (a), the unmasked template image **T** (PET image) in (b). The achievement of registration is shown in (c). Patient data courtesy of Ludwig-Maximilians-University of Munich.

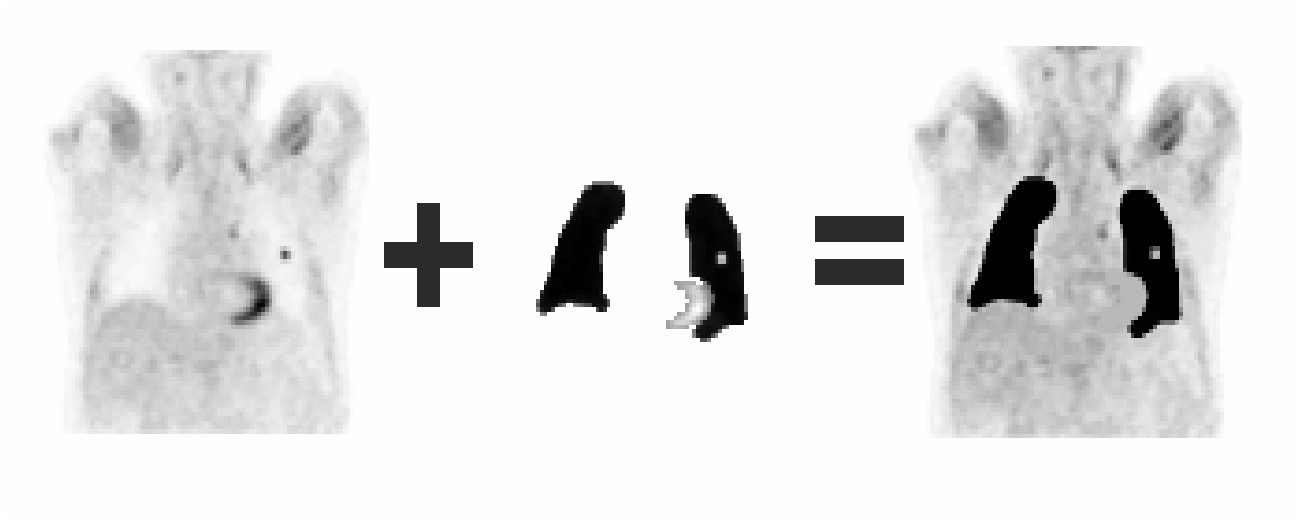

Figure 5. Unmasked template image (left), hand-segmented mask(centre) and masked template (right).

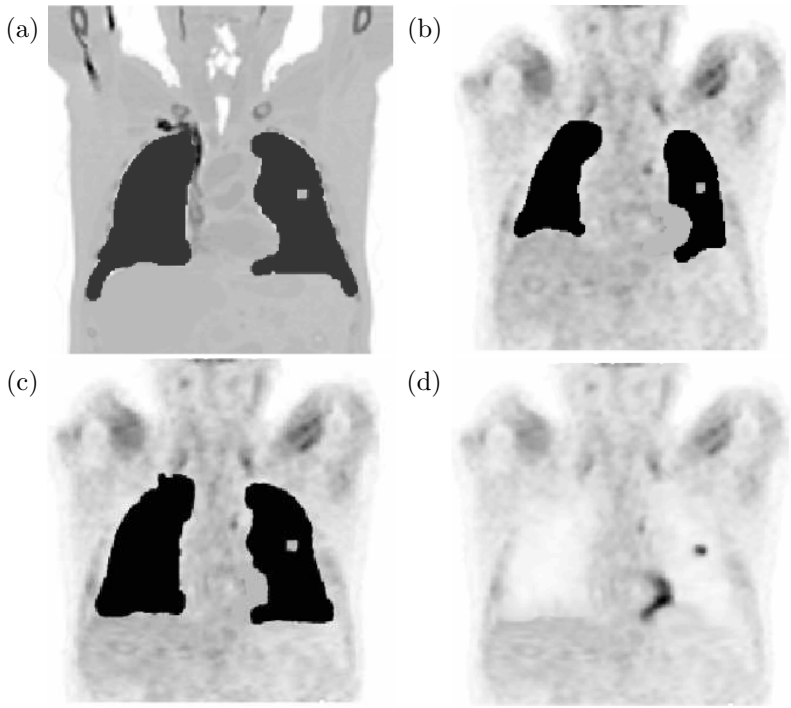

Figure 6. The improved registration of thorax CT and PET images for cancer diagnostic. The masked reference image **R** (CT image) is presented in (a), the masked template image (PET image) **T** in (b). The result of registration is shown in (c). The forces, those in the improved employment to develop were applied to unmasked template. The result is displayed in (d).

The masked images are displayed in figs. 6(a) and (b). A closer look on masking the PET image (fig. 4 (b)) is given in fig. 5. The masks were selected according to the following principle:

1. The lungs in both images were masked by segmentation to obtain the same volume of the lung in the CT image and the registered PET image. This is not possible without masking (see fig. 4 (c) in comparison with fig. 6 (d)).

2. The accumulation of tracer in the apex of the heart is masked out, because it is not substantial for this registration.

3. The tumor was masked, so that its position in the registered PET image agrees with the position in the CT image.

In comparison with the unweighted result in fig. 4(c), the weighted heart and show a more accurate size of the lungs. Both registration schemes use the parameter set $\alpha = 10^{-6}$, $\lambda = 80$, $\mu = 40$, $\beta = 50$ and 128 bins. They were terminated after 50 steps using a fix-point iteration.[1]

4. CONCLUSIONS

Previously, a weighted non-parametric image registration was only investigated using the sum of squared differences in mono-modal registration. This work extends the weighting approach to the distance measure MI. We developed a technique for including prior knowledge into non-parametric image registration based on mutual information. Due to this all advantages of a weighted registration can be used for the first time in a multi-modal image registration. We demonstrated that this technique considerably improves registration results for situations where corresponding structures in the images are not comparable. The drawback of this method is the need of masks. At the moment the masks are generated by hand. The approach was tested with the help of CT and US images of a phantom and CT and PET images of a patient. The evaluation has demonstrated the diversity, applicability and robustness of the proposed approach. A next step will be the investigation of automatic mask generation schemes.

REFERENCES

1. J. Modersitzki, *Numerical methods for image registration*, Oxford University Press, 2003.
2. J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, *Medical image registration*, CRC Press, Boca Raton, 2001.
3. J. B. A. Maintz and Viergever, "A survey of medical image registration," *Medical Image Analysis* (2(1)), pp. 1–36, 1998.
4. B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vision and Computing, 21 (11)* , pp. 977–1000, 2003.
5. W. Sureshbabu and O. Mawlawi, "Pet/ct imaging artifacts," *Journal of nuclear medicine technology, 33* , 2005.
6. S. Henn, L. Hömke, and K. Witsch, *A generalized image registration framework using incomplete image information - with applications to lesion mapping*, Springer-Verlag, springer series in mathematics in industry ed., 2004.
7. R. C. Gonzales and R. E. Woods, *Digital image processing*, Addison-Wesley, 1993.
8. P. A. Viola, "Alignment by maximization of mutual information," Tech. Rep. AITR-1548, 1995.
9. B. Fischer and J. Modersitzki, "Fast diffusion registration," *Contemporary Mathematics 313, Inverse Problems, Image Analysis, and Medical Imaging, AMS* .
10. C. Broit, *Optimal Registration of Deformed Images*. PhD thesis, Computer and Information Science, Uni Pensylvania, 1981.
11. B. Fischer and J. Modersitzki, "Curvature based image registration," *JMIV* **18(1)**, 2003.
12. B. Fischer and J. Modersitzki, "A unified approach to fast image registration and a new curvature based registration technique," *Linear Algebra and its Applications* **380**, pp. 107–124, 2004.
13. H. Schumacher, B. Fischer, and A. Franz, "Weighted medical image registration with automatic mask generation," *Proceedings of SPIE 2006, Medical Imaging, San Diego* , 2006.
14. B. Fischer and J. Modersitzki, "Fast inversion of matrices arising in image processing," *Numerical Algorithms 22* , pp. 1–11, 1999.
15. B. Fischer and J. Modersitzki, *A super fast registration algorithm*, pp. 169–173. Springer, 2001.