Automated analysis and visualization of preclinical whole-body microCT data

Baiker, M.

7

2D/3D Registration Of MicroCT Data to Multi-View Photographs Based On a 3D Distance Map

This chapter is based on:

2D/3D Registration Of MicroCT Data to Multi-View Photographs Based On a 3D Distance Map

Martin H. Wildeman, Martin Baiker, Johan H. C. Reiber, Clemens W. G. M. Löwik, Marcel J. T. Reinders, Boudewijn P. F. Lelieveldt

Proceedings of the IEEE International Symposium on Biomedical Imaging, 2009, pp. 987-990

Chapter 7

Abstract

In this work we present a method for registration of a CT-derived mouse skin surface to two or more 2D, geometrically calibrated, photographs of the same animal using a similarity transformation model. We show that by using a 3D distance map, which is reconstructed from the animal skin silhouettes in the 2D photographs, and by penalizing large angle differences between distance map gradients and CT-based skin surface normals, we are able to construct a registration criterion that is robust to silhouette outliers and yields accurate results for synthetic and real data (mean skin surface distance $0.12mm$ and $1.35mm$ respectively).

7.1 Introduction

Molecular Imaging comprises imaging of biological processes at a cellular level and at molecular resolution *noninvasively* and *in vivo* [3]. A broad spectrum of modalities exists for acquiring structural as well as functional data. If modalities are combined, this can results in new insights that could not have been gained by looking at the data separately.

In some cases, it can be useful to register datasets of a different modality *and* dimensionality. An example of this is the registration of 3D anatomical (e.g. CT or MRI) data and multiple 2D views of bioluminescence imaging (BLI) data of small animals. Since BLI images do not show anatomical references, they are usually combined with diffuse light photographs. If these photographs can be registered to 3D CT/MRI data, accurate bioluminescence source localization and quantification becomes possible in 3D using Bioluminescence Tomography, assuming that a 3D map with optical tissue properties can be derived from the CT/MRI dataset [8]. There are traditionally two categories of methods for solving 2D/3D registration problems, dependent on what the registration criterion is based on [157]: feature-based methods [158] and texture(intensity)-based methods [159]. In comparison, feature-based methods are usually faster than texture-based methods but have the disadvantage that the registration accuracy is dependent on the feature selection. To combine high processing speed with high accuracy, gradient-based methods have been proposed more recently [160]. Besides the choice of the registration criterion it is important to choose the registration domain (2D or 3D). While it is relatively easy to generate a 2D representation of a 3D object, the contrary may be complicated since typically the amount of 2D data is very limited. But registration in 3D has the advantage, that the time-consuming, repeated 3D to 2D mapping and the inherent loss of information can be avoided [158].

In recently published work, Markelj *et al.* [160] present an iterative method for rigid registration of a CT or MRI-derived dense gradient field with a sparse 3D gradient field, reconstructed using multiple 2D gradient images (conventional X-ray). The presented results show that their approach combines accuracy with robustness. A drawback of the method is that it needs to search for correspondence between the gradient fields in each iteration, which is time-consuming. Iwashita *et al.* [161] employ a correspondence-free 3D/2D approach for iterative registration of a 3D shape to pre-calculated distance maps, generated using estimated 2D contours of the object. However, the method is carried out in 2D and therefore requires an expensive 3D to 2D mapping in each iteration.

In this work, a 2D/3D registration approach is presented that does *not* need to establish correspondence during registration. On the one hand the method is based on data reduction by extracting features, but on the other hand it integrates feature gradients as well. The contributions of this work are:

- We introduce a registration criterion, based on a 3D distance map, which is reconstructed from a sparse set of 2D images. To ensure robustness and to increase accuracy, it includes angle penalties based on the direction of the gradients in this distance map as well as distance penalties based on the zero level set of the map.

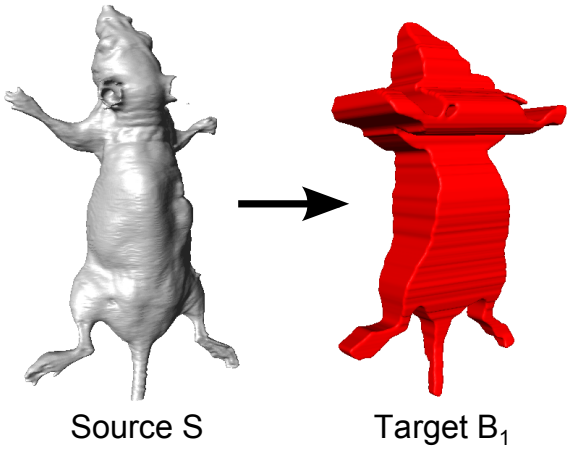

Figure 7.1: Examples of a CT-derived skin surface (source S) and a surface visualization of a target B_1, derived from two orthogonal photographic views.

- We apply this criterion to register CT-derived mouse skin surfaces to two or more 2D silhouettes of the animal skin using a similarity transformation model.

7.2 Methodology

7.2.1 Shape Representation and transformation type

Since the light photographs only show the animal exterior, the internal structural information in the CT data cannot be exploited for registration. Therefore the skin surface is extracted and represented as a triangular mesh. The target object is not fully determined in 3D because the photographs show projections of the animal and there is only a limited number of photographs available. Therefore an implicit shape representation in 3D is determined based on back-projection of the skin silhouettes that are derived from the 2D images. The target volume B_1 is then defined as the intersection of the back-projected silhouettes. An example of a skin surface extracted from CT and a target volume based on two photographs (top and side view) is given in Fig. 7.1. The size of the entire registration domain B_0, with $B_1 \in B_0$, is determined by the size of the 2D images. Since the CT dataset and the photographs are both acquired from the same animal but the two modalities are not calibrated, a similarity transformation model was chosen to describe the transformation from source to target.

7.2.2 Registration criterion

Given the source skin surface S with n vertices $v \in \mathbb{R}^3$ and the registration domain B_0, an energy function is defined as the sum of squared Euclidean distances (SSD) of all vertices in S to the boundary of the target volume B_1:

$$E_{\text{total}} = \sum_{v=1}^{n} E(v, \Theta)^2 \tag{7.1}$$

$$E(v, \Theta) = DM(B_0, \mathbf{x}) \tag{7.2}$$

$$\mathbf{x} = \lfloor T(v, \Theta) \rceil \tag{7.3}$$

In these equations, T represents the similarity transformation (translation, rotation and isotropic scaling), which is expressed as the parameter vector Θ. $DM(B_0, \mathbf{x})$ contains the Euclidean distance between $\mathbf{x}$ and the boundary of B_1 [162]. For computational purpose, DM was precalculated for all possible (rounded) vertex locations $\mathbf{x}$ by calculating the distance transform of B_0, using Danielsson's method [163]. Note that in the following, the boundary of B_1 is referred to as S_{B1}.

The registration criterion formulated in Eq. 7.1 yields accurate results if the match only receives data support on the tangential lines between S and S_{B1}, and only if the source and the target shape are very similar [160]. However, as depicted in Fig. 7.1, S and S_{B1} differ significantly from each other because one modality may contain information that is not available in the other (Fig. 7.2) and more importantly, because B_1 is based on a very limited number of 2D images. As a result, the minimum SSD yields an overestimation of S, because all vertices on S are considered to contribute equally to the SSD (Figs. 7.3, left and 7.3, middle). To reduce the influence of surface nodes on S, that do not determine the shape of S_{B1}, three steps are implemented:

1. The maximum distance in DM is limited, by introducing a $D_{\max}$ if $|DM(B_0, \mathbf{x})| \geq D_{\max}$:

$$DM_{\text{bound}}(B_0, \mathbf{x}) = max(min(DM, D_{\max}), -D_{\max}) \tag{7.4}$$

A bounded distance map reduces the influence of vertices that cause a large distance error, even if the solution is optimal.

2. Vertices on S that fall outside S_{B1} are penalized by multiplying the distance with a factor $\alpha > 1$:

$$DM_{\text{boundOP}}(B_0, \mathbf{x}) = \begin{cases} \alpha DM_{\text{bound}}, & DM_{\text{bound}} > 0 \\ DM_{\text{bound}}, & DM_{\text{bound}} \leq 0 \end{cases} \tag{7.5}$$

Adding a penalty to all vertices that fall outside B_1 reduces the overestimation of S. Note that these vertices are by definition positioned erroneously, since S_{B1} fully encloses S.

3. The surface normal of S (which closely corresponds to the direction of the steepest gradient in the CT volume) is compared to the distance map gradient (DMG) by adding a penalty if the angle difference r_{v} between the gradients is above a maximum angle $r_{\max}$ (Fig. 7.3, right):

Figure 7.2: A photograph of a mouse at 0° (left), the true silhouette (middle) and a simulated projection (right) based on a CT of the same subject.

$$E_{\text{angle}}(v,\Theta) = \begin{cases} DM_{\text{boundOP}}, & \text{if } r_v < r_{\max} \\ \alpha D_{\max}, & \text{if } r_v >= r_{\max} \end{cases} \tag{7.6}$$

$$r_{\text{v}} = \frac{360}{2\pi} cos^{-1}\left(DMG(B_0,\mathbf{x}) \cdot vertexnormal(T(v,\Theta))\right) \tag{7.7}$$

By integrating an angle penalty and iteratively decreasing $r_{\max}$, the influence of vertices on S, that are not 'contour'-vertices, is decreased step by step. Indeed, for $r_{\max} \to 0$, only vertices are taken into account, whose surface normal is identical to the distance map gradient (i.e. the tangential lines). In practice however, the final $r_{\max}$ will depend on the number of vertices, i.e. the sampling density of S (Fig. 7.4). The final energy function is:

$$E_{\text{final}} = \sum_{v=1}^{n} E_{\text{angle}}(v,\Theta)^2 \tag{7.8}$$

7.2.3 Minimization of the criterion function

For initialization, the Centers of Gravity (CoGs) of S and B_1 are aligned and an initial scaling parameter is derived from the dimensions of B_1. Subsequently the energy function is minimized using an iterative nonlinear regression method [101].

7.3 Experiments

7.3.1 Validation tests

To validate the proposed method, registration of a CT-derived skin surface to a reconstructed volume was performed. This volume was based either on two and four simulated 2D projections, or two real 2D photographs. For quantitative performance assessment, the transformation parameters (absolute translation of CoG, solid angle, scaling) after registration were compared to the true transformation parameters. In addition, the mean surface distance, and a mean Dice coefficient [164] were calculated between the ground truth and the registered source surface. For the simulated data, the transformation

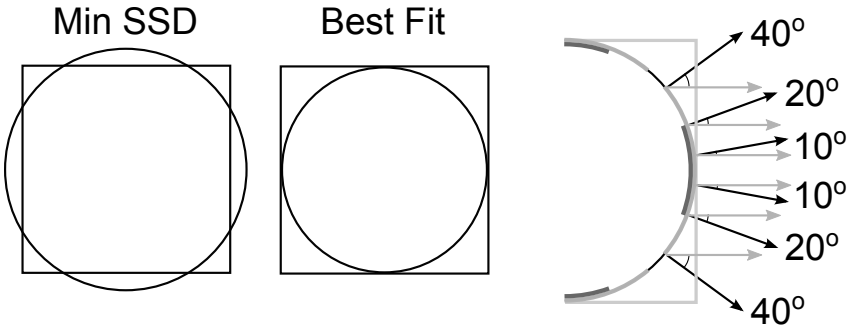

Figure 7.3: Demonstration of the source surface (represented as a circle) overestimation when all surface vertices are weighted equally. In this case, the transformation yielding the minimum SSD (left) does not correspond to the optimum solution (middle). The drawing on the right shows penalized vertices, depending on the angle difference (dark gray = 20°, light gray = 40°) between vertex normals (black) and the DMG (light gray).

Table 7.1: Results of the validation tests (voxel dimension $\hat{=}700\mu m$).

Data	Views	α	Dist. between Skin Surfaces [voxels]	Abs. Dist. of CoGs [voxels]	Scaling Error [%]	Solid Angle Error [degrees]	Dice Coefficient
Synth.	2	1	0.46 ± 0.18	0.64 ± 0.61	0.74 ± 0.54	0.59 ± 0.42	$0.995 \pm 0.75e^{-3}$
Synth.	4	1	0.54 ± 0.08	0.54 ± 0.35	1.33 ± 0.34	0.37 ± 0.27	$0.995 \pm 6.46e^{-4}$
Synth.	2	2.5	$\mathbf{0.17} \pm 0.08$	0.14 ± 0.12	-0.32 ± 0.22	0.60 ± 0.38	$\mathbf{0.998} \pm 5.10e^{-4}$
Synth.	4	2.5	$\mathbf{0.21} \pm 0.04$	0.56 ± 0.16	-0.07 ± 0.19	0.34 ± 0.22	$\mathbf{0.998} \pm 3.68e^{-4}$
Real	2	1	1.98 ± 0.59	5.59 ± 3.74	-1.65 ± 1.53	1.73 ± 1.22	$0.983 \pm 3.00e^{-3}$
Real	2	2.5	1.93 ± 0.66	4.51 ± 3.10	-3.40 ± 1.86	1.66 ± 1.03	$\mathbf{0.984} \pm 3.07e^{-3}$

parameters were known. For the real data, these were calculated, by minimizing the Euclidean distance between two sets of four manually determined anatomical landmarks in the CT data and the photographs respectively. To investigate the effect of α, all experiments were run using $\alpha = 1$ (no penalty) and $\alpha = 2.5$. $r_{\max}$ was first set to 180° (no angle penalty) and then iteratively reduced to 40° and 10°.

7.3.2 Data Acquisition

For validation, 10 mice (Balb/c) were individually placed on a holder. First, a MicroCT dataset was acquired (Skyscan 1178, Kontich, Belgium) *in vivo*, with the resolution $80 \times 80 \times 80$ μm^3. Subsequently, 2D photographs (top and side view with an angle difference of 90°) were taken from the same animal (Caliper IVIS 3D BLI system, Hopkinton, USA), and subsampled to 700×700 μm^2. The usage of the holder prevented posture changes during animal transfer between the modalities. For registration, the skin was segmented semi-automatically from the CT datasets and converted to a triangular mesh ($\approx$ 2000 vertices). The skin boundary in the 2D images was outlined manually in this study (Fig. 7.2, middle). The entire registration domain had the dimensions $52.5 \times 52.5 \times 105mm^3$ ($150 \times 150 \times 300$ voxels). The dimensions of the simulated data as well as the

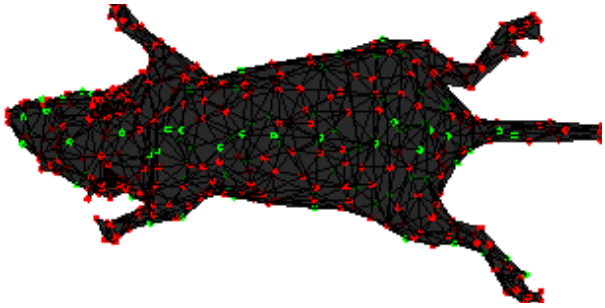

Figure 7.4: The effect of introducing an angle penalty. Vertices on S, that are penalized, are shown in red. Non-penalized vertices are shown in green. In this example, $r_{\max}$ was set to $10°$.

angles were chosen equal to the dimensions and the angles of the real data. For the four view case, two additional angles were used at $\pm 45°$.

7.4 Results and Discussion

A qualitative example of the ground truth as well as the registration result is shown in Fig. 7.5. Table 7.1 presents quantitative results for the experiments using synthetic and real data.

The experiments with the synthetic data reveal good performance of the algorithm. The mean surface error is less than a voxel dimension ($\approx 0.12mm$), the Dice coefficient is very high (0.998) and the transformation parameters could be recovered accurately. In addition, applying a penalty for overestimation ($\alpha = 2.5$) resulted in a better estimation of the scaling parameter than without using the penalty ($\alpha = 1$). Also the registration performance seems to be somewhat better for the 2 view case than for the 4 view case, which is unexpected. However, the variance is smaller in the 4 view case, which indicates that the algorithm performs more robustly, while still yielding accurate results. The other results (e.g. the surface distance) are comparable and clearly below a voxel dimension. This indicates that interpolation effects during generation of the projections at $\pm 45°$ might cause this error.

For the real data, the mean surface distance is less than two voxel dimensions ($\approx 1.35mm$) and the Dice coefficient is very high (0.984). Altogether, the results are less good than using the synthetic data. Reasons might be shape discrepancies (Fig. 7.2 and 7.5) and the non-standardized, manual 2D silhouette and 3D skin surface segmentation. Applying the penalty $\alpha = 2.5$ seems to cause results that are less good for some evaluation measures. This can be explained by the fact that the manually placed anatomical landmarks, which were used to derive the ground truth, have limited accuracy.

7.5 Conclusions and Future Work

We presented a novel method to register a 3D surface to multiple 2D photographs. The chosen registration criterion has proven to yield accurate results using synthetic and

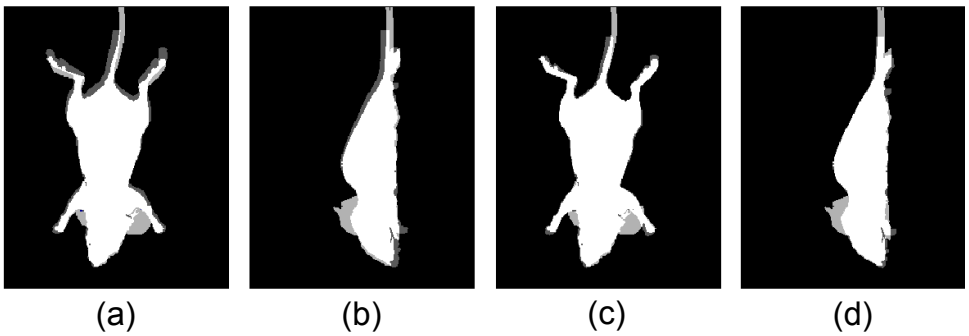

Figure 7.5: Qualitative registration result showing 2D silhouettes (light gray), CT-based projections of a skin surface (dark gray) and their overlap (white) for the ground truth (a, b) and the presented method (c, d). The method is robust against missing registration features (e.g. the ears and part of the tail).

real data and is robust to missing registration features. While the performance was demonstrated using two and four views, more views can be added easily, without adding computational overhead. Since the method does not rely on calculating correspondences, it can be implemented very efficiently.

Because the experiments using more than two (synthetic) views did not yield conclusive results, these experiments will be expanded in the future using real data. Above that, the skin silhouette and the CT skin segmentation should be automated. In addition, we plan to combine the algorithm with a method to approximate major organs from CT data [86] to provide a heterogeneous tissue model for bioluminescence tomography.