# Transform Learning – Registration of medical images using self organization

Dietlind Zühlke
Fraunhofer Institute of Applied Information Technology FIT
Mailing address
email: dietlind.zuehlke@fit.fraunhofer.de

Keywords: model based image registration, 3D ultrasound, self organization

***Abstract*—** A network model is introduced that allows a multimodal registration of two images. It provides a model-model registration. The application of the network in the registration of medical 3D ultrasonic image data is introduced. Results on artificial and real ultrasound image data sets are discussed.

## 1 Introduction

One objective of our research team is to develop an intervention assistant[1] for navigation supporting the local tumor therapy in liver tumors. A near real time registration between a pre-interventional tomographic image and an interventional 3D ultrasound (US) image data set is crucial for this purpose.

Registration in this field means transforming the coordinate space of one image into the coordinate space of the other image [1]. After this transformation both images can be aligned or even merged. This way information in one image modality can be joined with the information available in the other. In the given application this means, that the registration procedure has to merge the data sets accurately enough that the physician can intuitively map radiological information onto pertinent regions in the US.

According to Aylward et al. [2] there are three basic forms of image registration:

- image-image registration
- image-model registration
- model-model registration

The choice of the procedure depends on features of the image data, e.g. the visibility of anatomical structures.

The current clinical procedure is based on an interactive landmark based registration of external bony landmarks. This method is error prone and slow if many adjustments have to be made. The objective was to develop a registration method that is more stable, more accurate and faster.

The following considerations influenced the design of the new registration process:

- Image-image registration is rather slow, because much basic image information, in the majority of cases color- or gray values, is used for the algorithm. This approach is used for images with similar image features, such as similar gray values in same image modalities. It is not suitable for images of low quality (noisy and with artifacts), because convergency is slow and the probability to hang in a local minimum is high.
- Image-model and model-model registration both use reduced image information. These approaches are faster if this reduced information is already available. These registration methods yield unacceptable results if information loss is produced by an inadequate reduction.

To apply the image-model or model-model principle approach, a stable and geometrical valid model of one (or both) of the image data sets has to be produced in a first step. Therefore only structures can be exploited, that can be seen in both image data sets and that are stable with respect to deformation. In our case, the blood vessels are a structure that is in this sense useful for registration of soft tissue images. The topological structure defined by the branching points as well as the radius of the vessels are the two features that are not influenced by soft tissue shifts. They are therefore pivotal features. Other features such as length or direction are affected by tissue shifts.

The vascular tree models have to be derived from two different image modalities:

(1) tomographic image data sets (MRI, CT)

(2) 3D Doppler or contrast enhanced ultrasonic image data sets

The blood vessels are extracted from the tomographic data using a segmentation-based method [3]. We developed an extraction method for 3D Doppler or contrast enhanced ultrasound image data sets. The Vessel Extracting Gas (VEG) [4] keeps pivotal features of the vessels and suppresses noise.

[1] the LOCALITE SonoNavigator

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

# 2 Methods

## 2.1 Background

Conventionally, model-model registration is based on a graph or tree matching algorithm (see e.g. [5]). The matching is used to calculate the registration transform. These algorithms perform good if the graphs or trees being matched are very similar in structure as well as in size. An inappropriate matching, caused e.g. by trees of different sizes, leads to bad registration results.

In the given application the following vascular tree models are extractable:

- *Ultrasonic model:* partial vascular tree, low degree of detail
- *Tomographic model:* complete vascular tree, high degree of detail

In the given context registration methods based on tree matching are unsuitable because of those differences in size and degree of detail.

Another common registration method in model-model registration is the Iterative Closest Points method (see e.g. [6] and [7]). It iteratively matches two surface models given as point sets. In one iteration a registration is calculated that provides an immediate target jump to minimize the squared differences. To function properly, the algorithm needs the given point sets to overlap as much as possible.

The raw data given in our application are tree models. When using these tree models, the immediate target jump in the ICP is likely to get caught in a local minimum. Thus the idea was to develop an alternative registration method – called Transform Learning – that carefully approaches the real transformation in a stepwise manner. Together with a coarse initial alignment, interactively produced by the interventionalist, this idea reduces the probability to hang in a local minimum of registration quality.

## 2.2 Transform Learning

The input of the Transform Learning are two point sets. No initial correspondences between these points are needed. In the given application the tomographic vascular tree model shall be matched to the vascular tree model extracted from the ultrasound. Therefore the positions of the branching points of the tomographic vascular tree model are passed to the method as well as the positions of the branching points of the ultrasonic vascular tree model.

Transform Learning can additionally consider an initial alignment between the two point sets. In the medical application an interactively realized preregistration is used as initial alignment.

The learning process iteration consists of two major steps:

(1) **The "virtual" adaption of the vertices:** A training vector is presented to the network. The nearest vertex in the net is "virtually" adapted. The adaption is based on the "Winner-takes-all" principle [8] as well as on the adaption in the Growing Neural Gas by Fritzke [9]. It is called "virtual" because the position of the vertex does not change directly, but the old position and the "virtual" new position respectively are used as corresponding landmarks.

(2) **The adaption of the transformation:** If the number of training vectors that were presented is an integer multiple of the parameter $\kappa$, the corresponding landmarks calculated by the "virtual" adaptions are passed to a landmark based registration method. In our case a simple Least-Square-Estimation was used. The registration method returns a transformation adjustment $\Delta T$ that is appended to the last valid transformation.

This whole iteration is done until the virtual movement of the vertices falls below a certain specified threshold. In this case, the last valid transformation is applied to the complete model or even to the complete image from which the model has been derived.

### 2.2.1 Elements of the Transform Learning

The network builds up as follows:

- set of vertices $M$ with $m \in M$ has:
  - n-dimensional reference vector $w_c \in R^n$
- vector for saving old position values $V_{p_o}$
- vector for saving new position values $V_{p_n}$
- $4 \times 4$-transformation matrix $T$
- set $P$ of n-dimensional training or input signals $p \in P$

### 2.2.2 Process steps of the Transform Learning

Transform Learning algorithm:

(1) Initialize $M$ of the net with the vertices $w_M$ in $R^n$ given in the tomographic model.

(2) If a preregistration is given initialize the transformation matrix $T(t = 0)$ using the given preregistration matrix, else initialize $T(t = 0)$ using the identity matrix.

(3) Chose a training vector $p$ from $P$.

(4) Find the nearest vertex $s$, with
$s = \min \arg(\|T(t) \cdot w_s - p\|)$,
and add its reference vector multiplied with $T(t)$ to the vector of old position values $V_{p_o}$.

(5) Add the adapted reference vector
$w_{s_{adaptiert}} = T(t) \cdot w_s + \Delta w_s$
to the vector of new position values $V_{p_n}$ , where:
$\Delta w_s = \epsilon(p - w_s)$

(6) If the number of presented input signals is an integer multiple of the parameter $\kappa$, calculate the adapted transformation as follows:

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

- Input the vectors of position values to a landmark based registration method, which then outputs a transformation change $\Delta T(t)$.
- Then the adapted Transformation is: $T(t+1) = \Delta T(t) \cdot T(t)$.

(7) If at least for one reference vector the distance of the new and the old position exceeds $min$, increment $t$ and go back to step 3.

## 2.3 Determining an appropriate parameter set

As in any other learning method, the choice of the parameters has a crucial impact on the performance of the method. Thus the determination of an appropriate parameter set is very important.

To optimize the parameter set, a test environment[2] was built that permits qualitative evaluations of the results. It shows image volumes and vascular tree models respectively from the tomographic data as well as from the ultrasound data. The image volumes are displayed as derived surface models with an adjustable threshold (see yellow model in figure 1(a)). The vascular tree models are displayed as combinations of spheres and cylinders (see blue model in figure 1(a)). It is possible to qualitatively evaluate the registration by simultaneously displaying the surface model and the preregistered vessel model as well as the registered vessel model or by switching between the three of them.

The parameter optimization tests were first run on artificial ultrasound image data sets. To create this test data, a set of vessel tree models was taken, that was extracted from an MRI of a proband study by MeVis GmbH. The models are given as sets of vertices and edges with extracted diameters.

From these models artificial 3D Doppler ultrasound volumes were created as follows: The model was transformed to fit into a volume data set of a fixed size. Then gray values corresponding to typical Doppler ultrasound gray values were introduced into the artificial volume data set at the positions specified by the transformed model. The artificial data did *not* include the following characteristics of normal 3D Doppler or contrast enhanced ultrasound (see e.g. [10]):

- noise
- speckle
- artifacts

By initializing the Transform Learning approach the following free adjustable parameters have to be known:

- threshold for vessels $\theta$
- number of points presented in one iteration $\kappa$
- learning rate for vertices to be adapted $\epsilon$
- minimal "virtual" movement of vertices $min$

[2] based on the LOCALITE SonoNavigator

The threshold for the vessels determines the set of vessel voxels.

To identify a good set of parameters, the procedure was first started with the following values:

- learning rate $\epsilon = 0.2$ the value Fritzke suggests for GNG
- minimal "virtual" movement $min = 0.5$ in the range of desired accuracy
- number of points presented in one iteration $\kappa = 10$ to keep running time short

Varying the parameters influences the performance of the net as summarized in table 1. The following set of parameters yielded good results on artificial ultrasound image data sets checked in the test environment described above: $\kappa = 30$, $\epsilon = 0.196$, $min = 1.1$.

| parameter | too small | too large | recommended value |
|---|---|---|---|
| $\theta$ – threshold for vessels | much noise, net is adapted towards noisy areas | few vessel voxels, unrepresentative, inaccurate especially on small vessels | |
| $\kappa$ – number of points presented in one iteration | transformation jumps | long running time, because calculation seldom | 30.000 |
| $\epsilon$ – learning rate for vertices to be adapted | long running time, because movements are small | | 0.196 |
| $min$ – minimal "virtual" movement of the vertices | long running time or oscillation without convergency | premature convergency with bad registration | 1.100 |

Table 1: Results of parameter tests for the Transform Learning approach in test environment

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

# 3 Results

There are two criteria for the evaluation of the registration:

- mean accuracy
- run time performance

**Mean accuracy of the registration** For artificial ultrasonic image data sets this can directly be measured because the registration transformation is known. For real data this judgment is not as easy. It can at least be done qualitatively. Bad registrations can be characterized by visible deviations (see figure 1(a)). Good registrations will be identified if the structures are congruently aligned with each other (see figure 1(b)). It is justified to stay with this qualitative evaluation because in medical application the results of the method are used for visual orientation. If the alignment of the image data sets presented to the interventionalist (see section 1) is in this sense good enough his orientation in the data sets will improve.

**Run time performance of the registration** In the application described in chapter 1 the calculation has to be nearly real-time. The Transform Learning was tested on a PC with Intel Pentium D CPU (two processor kernels with 3.00 GHz resp.), 2.0 GB RAM under Windows XP with Service Pack 2. The given running times apply to this PC.

## 3.1 Results on artificial data

First tests were carried out on the artificial data described in chapter 2.3. To test the performance of Transform Learning, a transformation was introduced between the original positioning of the vascular tree and the artificial ultrasonic image data set. Many different geometrical shifts were tested, three samples are given here:

(1) a translation of 1 cm towards x-direction, of 0 cm towards y-direction, of 3 cm towards z-direction and a rotation of 20 around the x-axis, of 20 around the y-axis, of 20 around the z-axis
(2) no translation and a rotation of 20 around the x-axis, of 0 around the y-axis, of 10 around the z-axis
(3) translation of 2 cm towards x-direction, of 10 cm towards y-direction, of 0 cm towards z-direction and no rotation

The registration quality of Transform Learning using different sets of parameters on different geometrical shifts was evaluated with respect to the criteria mentioned above.

In first tests the parameter set determined in chapter 2.3 was applied.

In the following the results of two selected sets of parameters will be presented. The selected sets of parameters are:

(1) optimal set of parameters: $\epsilon = 0.2$, $\kappa = 30$, $min = 1$
(2) worse set of parameters: $\epsilon = 0.2$, $\kappa = 30$, $min = 2$

In table 2 the mean accuracy and the running time of Transform Learning for the geometrical shifts (1)-(3) and the parameter sets (1)-(2) are shown. The listed running times include the extraction of the ultrasound model.

| set of parameters | transformation | mean accuracy | running time |
|---|---|---|---|
| 1 | 1 | 1.125mm | 1.766s |
| 1 | 2 | 4.441mm | 1.578s |
| 1 | 3 | 9.883mm | 1.609s |
| 2 | 1 | 17.504mm | 0.578s |
| 2 | 2 | 10.515mm | 0.609s |
| 2 | 3 | 22.113mm | 0.531s |

Table 2: Examples of mean accuracy and running time for Transform Learning

The parameter set that yielded best results in the tests (1) was further tested with different geometrical shifts, that comprised rotations as well as translations. The running time of Transform Learning increased with increasing distance between the two models given whereas the results were generally good.

Transform Learning in the model-model approach seems to give very accurate results for geometrical shifts that are expectable for typical registration problems when radiological and ultrasound data have to be merged. The run time performance is acceptable.

Figure 1 shows the results of Transform Learning with the best set of parameters (1) and the geometrical shift (1) on an artificial 3D ultrasonic image data set. Figure 1(a) gives the geometric shift that was introduced between the model and the artificial ultrasound image. In figure 1(b) the registered vascular tree model shown in green lies inside the set of vessel voxels. In this parts only the yellow surface of the vessel voxels is seen. The green branches show parts where the registration has not perfectly aligned the vascular tree model and the vessel voxels.

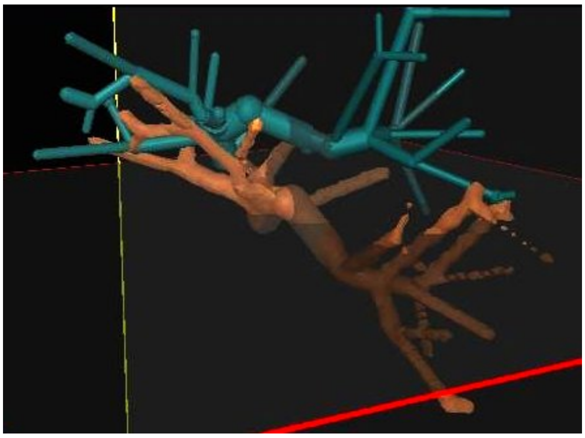
## 3.2 Results on real data

First preliminary tests were run on different real data sets. The data sets have been 3D ultrasound image (Doppler or contrast enhanced) and a vascular tree model derived from a 3D MRI or CT data set. A preregistration between the tomographic and the ultrasonic data set had been introduced by the interventionalist using external bony landmarks.

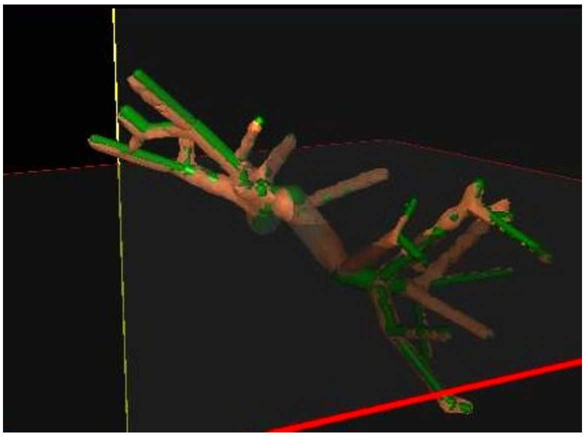
For first tests the set of parameters that yielded best results on artificial data was applied. The following set of parameters gave good results on real ultrasound image data sets checked in the test environment described above: $\epsilon = 0.2$, $\kappa = 30$ and $min = 3$.

An example result yielded with these parameters is shown in figure 2. The result of the registration method

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de



(a) model and vessel voxels aligned with geometrical shift (1), *blue:* tomographic vascular tree model extracted from a proband MRI, *yellow:* surface of the set of vessel voxels from an artificial 3D ultrasonic image data set



(b) model aligned with vessel voxels after Transform Learning registration, *green:* registered tomographic vascular tree model extracted from proband MRI, *yellow:* surface of the set of vessel voxels from an artificial ultrasonic image data set

Figure 1: Results of the Transform Learning in model-model approach

currently used in the clinical application is displayed in figure 2(a). It took the interventionalist about 15 min to achieve this interactively driven registration. In figure 2(b) the result of the Transform Learning is given. This registration took 20.976s including the extraction of the ultrasonic vascular tree model when the preregistration was known.

Another example result is given in figure 3. It was calculated using the same parameters as mentioned above. Again the result of the registration method currently used in the clinical application is displayed for comparison, see figure 3(a). This interactive method took about 20 min. In figure 3(b) the result of the Transform Learning is given. This registration took 1,312s including the extraction of the ultrasonic vascular tree model when the preregistration was known.

The result of Transform Learning is a better alignment of the structures that results in a better orientation for the interventionalist. The run time performance of the Transform Learning is acceptable for the intended application.

## 4 Discussion

According to first results on real data Transform Learning, as model-model approach in combination with VEG [4], is able to cope with the requirements of the real world:

- *Stability:* Compared to conventionally used registration methods, the whole process is more stable against the artifacts that are typical, e.g. noise and speckle [10], because of the previous extraction of a model from the ultrasound using VEG [4].
- *Run time performance:* The run time performance was acceptable throughout all tests.

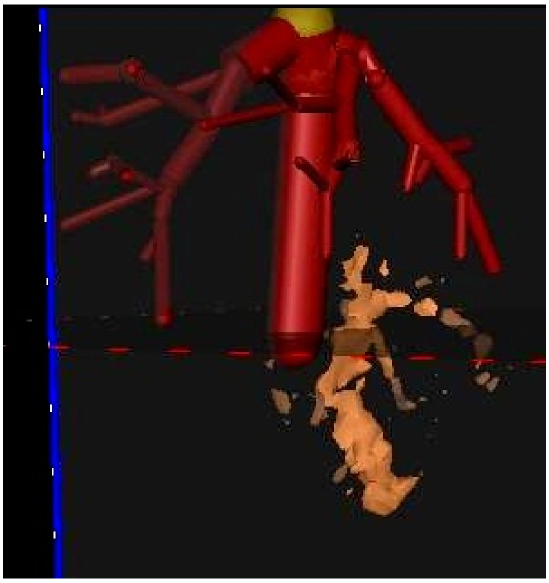
Further testing on real data is needed. The successful application of this method for real medical image registration will require that one fixed parameter set can cope with most of the data. Otherwise the parameters have to be adjusted for every single image. In this case it is important to know what adjustments have to be made for what kind of image data. First tests indicate, that the presented set of parameters yields acceptable results for a wide range of real data.

The registration quality must be further evaluated in the medical work flow. Here the important criteria are relevant improvement compared to the interactive landmark based registration and a reduction of registration time.

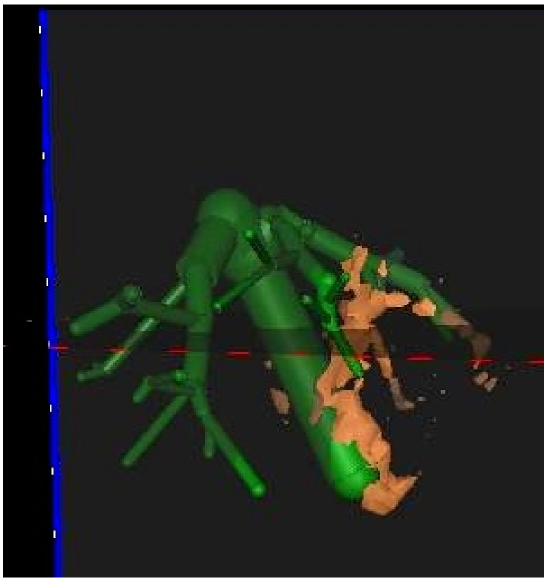
## Acknowledgements

The presented work took place in the context of the project FUSION – Future Environment for Gentle Liver Surgery Using Image-Guided Planning and Intra-Operative Navigation – funded by the Federal Ministry of Education and Research (BMBF). We want to thank the LOCALITE GmbH – Biomedical Visualization Systems – for the support, especially concerning the LOCALITE SonoNavigator. We would also like to thank MeVis GmbH, a partner in

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

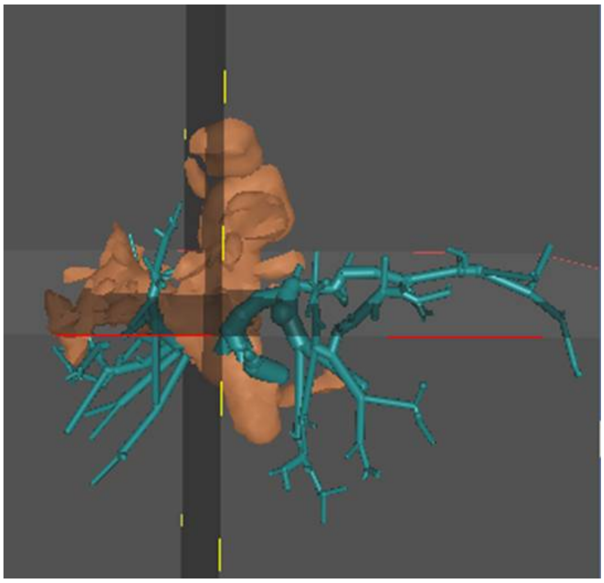


(a) model aligned with vessel voxels after conventional interactive registration, *red:* tomographic vascular tree model extracted from a proband MRI, *yellow:* surface of the set of vessel voxels from a proband's 3D ultrasonic image data set

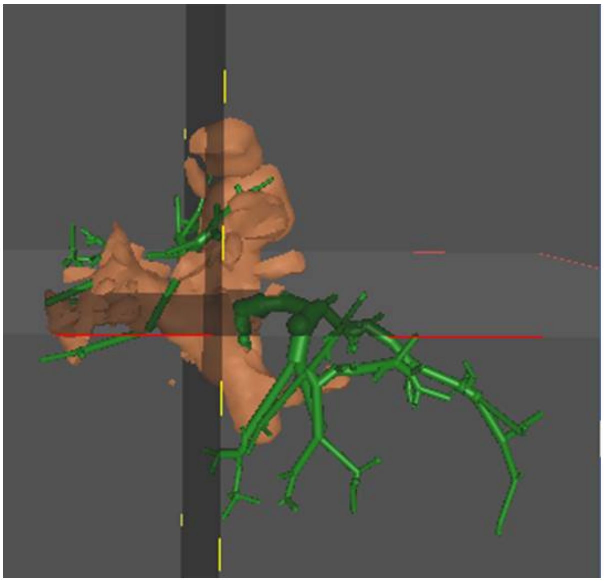


(b) model and vessel voxels after registration, *green:* registered tomographic vascular tree model extracted from a proband MRI, *yellow:* surface of the set of vessel voxels from a proband's 3D ultrasonic image data set

Figure 2: Results of the Transform Learning



(a) model aligned with vessel voxels after conventional interactive registration, *blue:* tomographic vascular tree model extracted from a proband MRI, *yellow:* surface of the set of vessel voxels from a patient's 3D ultrasonic image data set



(b) model and vessel voxels after registration, *green:* registered tomographic vascular tree model extracted from a patient's MRI, *yellow:* surface of the set of vessel voxels from a proband 3D ultrasonic image data set

Figure 3: Results of the Transform Learning

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

the FUSION project, for the extraction of the vessels from tomographic data sets. Furthermore we would like to thank our project partner – the Düsseldorf University Hospital – especially Dr. med. Ralf Kubitz for image data and expert evaluation of achieved results.

## References

[1] Torsten Rohlfing and Jürgen Beier and Oliver Schulte and Norbert Hosten and Roland Felix, "Multimodale Registrierung mit effizienten Lernverfahren für neuronale Netze", *Bildverarbeitung für die Medizin*, 1998.

[2] Stephen R. Aylward and Julien Jomier and Jean-Philippe Guyon and Susan Weeks, "Intra-operative 3D ultrasound augmentation", *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2002.

[3] S. Weber and G. Prause and T. C. Lueth and H.-O. Peitgen, "Image-guided planning and intra-operative navigation in open liver surgery", *Biomedizinische Technik / Biomedical Engineering*, 2005.

[4] Dietlind Zühlke, "Vessel Extracting Gas – Using self organization in the extraction of vascular trees", *WSOM07*, 2007, accepted.

[5] Arnaud Charnoz and Vincent Agnus and Grgoire Malandain and Luc Soler and Mohamed Tajine, "Tree Matching Applied to Vascular System", *Lecture Notes in Computer Science: Graph-Based Representations in Pattern Recognition*, Volume 3434, 2005.

[6] Paul J. Besl and Neil D. McKay, "A method for registration of 3-d shapes" , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) pp: 239-256, February 1992.

[7] D.Chetverikov and D. Stepanov, "The Trimmed Iterative Closest Point Algorithm", *Proc. of 16th International Conference on Pattern Recognition 3*, pp. 545-548, August 2002.

[8] Simon Haykin, *Neural Networks: A Comprehensive Foundation*, 2004.

[9] Bernd Fritzke, "A growing neural gas network learns topologies", *Advances in Neural Information Processing Systems 7*, pp. 625-632, 1995.

[10] F. Forsberg and JB Liu and PN Burns and DA Merton and BB Goldberg. "Artefacts in ultrasonic contrast agent studies" *J Ultrasound Med*, Vol. 13, pp. 357-365, 1994.