ISBN: 978-972-8924-97-3 © 2009 IADIS

A PARALLEL APPROACH FOR ALIGNMENT OF MULTI-MODAL GRID-BASED DATA

Egor Dranischnikow, Elmar Schömer
Institute of Computer Science, Johannes Gutenberg-University, Mainz

Ulrich Schwanecke
Dept. of Design, Computer Science and Media, Hochschule RheinMain, University of Applied Sciences, Wiesbaden Ruesselsheim Geisenheim

Ralf Schulze, Dan Brüllmann
Dept. of Oral Surgery (and Oral Radiology), University Medical Center of the Johannes Gutenberg-University, Mainz

ABSTRACT

Multi-modal registration is still a big challenge in image processing. In this article we present a variation of the well-known fast Fourier transform (fft) accelerated methods for finding the alignment between two datasets, i.e. the rigid transformation consisting of a rotation as well as a translation mapping all regions in both datasets belonging together in an optimal manner. Our method can be applied to such multi-modal registration problems as computer tomography (CT)/positron emission tomography (PET) matching, or matching CT data with data obtained by magnetic resonance imaging (MRI).
We reformulate the alignment problem into an optimization problem concerning a metric measure. The particular form of the proposed objective function can be exploited to fft-accelerate the translational part of the alignment-problem. Thus the reduced problem can be solved for the three remaining degrees of freedom of the rotatory part using standard optimizers, such as downhill-simplex or Powell's method.
A further advantage of our approach is the straight forward parallelization of the objective function's computation. Our implementation on a graphic processing unit (GPU) yielded a speedup factor between 5 and 25 depending on the size of the data. The results show, that the application of a GPU can be highly rewarding for all fft accelerated algorithms.

KEYWORDS

Alignment, cone-beam data, multi-modal registration, parallel implementation, GPGPU.

1. INTRODUCTION

The most common approach in multi-modal registration is based on the concept of mutual information (MI). Albeit MI based metric measure has been shown to be robust and generally applicable (see Maes, F. et al (1997), Mattes, D. et al (2003), and Viola and Wells (1997)), a series of drawbacks are known. In particular the following two facts are problematic: The distributions of the random variables needed for computation of mutual information are unknown and their estimation can consume a lot of processor resources; furthermore, the resulting objective function possesses many local maxima (Haber and Modersitzki (2005)).

If both datasets have been recorded with the same technology, it is possible to use the meansquare error as a metric. Its main advantage is the possibility to fft-accelerate the optimization process in some degrees of freedom. A fast method for the solution of the pure translational alignment problem was presented in Cowtan (2001), and a combination of translational and rotatory degrees of freedom was introduced in Kovacs et al (2003). Fourier methods, in particular the phase correlation method, are pretty well known and play a prominent role in 2D image processing and pattern analysis (see Castro and Morandim (1987), Pan et al (2009), Reddy and Chatterji (1996)).

Trying to match CT and PET datasets, we face the following dilemma: the MI based general approach is too slow, because it cannot be fft-accelerated, and the methods which can be accelerated are not applicable to

multi-modal data. In this article we present an approach that is not as universal as the conventional MI method but then can be fft-accelerated. Our alignment method is based on metrics of the form

$$\varphi(f,g) = \sum_{\vec{x}\in\Omega} w(\vec{x}) \mid f(\vec{x}) - g(\vec{x}) \mid^2 ,$$

with Ω - a domain in N^3, w - a weight function and f,g - arbitrary scalar functions. The weight function is the most important difference to the cases considered in Reddy and Chatterji (1996) and Castro and Morandim (1987). Both papers assume that one function of the two to be aligned is a rotated and translated replica of the other one, but for the multi-modal data this is almost never true. The presence of dissimilar parts has a negative impact on the reliability of the phase correlation method. Therefore we are going to use an idea similar to that proposed in Cowtan (2001), with a weight function as a fragment mask.

The remainder of this paper is structured as follows. In the next section we introduce relevant definitions and mathematics as well as the core algorithm of our method. Section 3 covers the parallelization of the algorithm on a GPU, whereby we are also going to discuss the performance of the parallelized version. Experiments with synthetic data are presented in section 4, whereas section 5 contains some widespread applications, for which the proposed method yields good results.

2. MATHEMATICS

We consider a three-dimensional dataset as a periodic function $f:\Omega\to R$. The set of these functions will be denoted by $Abb(\Omega,R)$. To evaluate the distance between two elements f,g from $Abb(\Omega,R)$, we use the pseudo metric

$$d_w(f,g) = \sqrt{\sum_{\vec{x}\in\Omega} w(\vec{x}) \mid f(\vec{x}) - g(\vec{x}) \mid^2} ,$$

with $w(\vec{x}) \geq 0$. d_w is a metric if and only if $w(\vec{x}) > 0 \forall \vec{x} \in \Omega$. Using a pseudo metric for distance estimation enables us to neglect parts of the data, which are for sure dissimilar in the multi-modal datasets.

A rigid body transformation $T(\vec{\alpha},\vec{\zeta};\bullet)$ between elements of Ω, which depends on three rotatory degrees of freedom $\vec{\alpha}$ and three translational degrees of freedom $\vec{\zeta}$, generates the operator $\hat{T}$ defined by

$$\hat{T}\left[\vec{\alpha},\vec{\zeta}\right](f) := \tilde{f}\left(T^{-1}(\vec{\alpha},\vec{\zeta};\bullet)\right).$$

Thereby $\tilde{f}$ is the continuation of the data record f from Ω to continuous space R^3 which can be computed by means of trilinear, nearest neighbor or any other interpolation scheme. In our experiments both methods yielded similar results.

With the help of d_w we can define the objective function

$$\varphi(f,g;\vec{\alpha},\vec{\zeta}) := d_w(f',g)^2$$

with $f' := \hat{T}\left[\vec{\alpha},\vec{\zeta}\right](f)$ beeing the transformed dataset. φ measures the quality of a coordinate transform or rather the quality of a given data matching.

Now we can consider the alignment problem as the following optimization problem.

Problem 1 *(General alignment-problem)*

Given f,g from $Abb(\Omega,R)$, find

$$(\vec{\alpha}^*,\vec{\zeta}^*) := \arg\min_{\vec{\alpha},\vec{\zeta}\in R^3} \varphi(f,g;\vec{\alpha},\vec{\zeta}) .$$

Initially we consider only the translational part of the alignment problem, i.e. the problem of finding the best translational movement for a fixed rotation. Therefore we obtain the following reduction of the general alignment problem:

Problem 2 *(Translational alignment-problem)*

Given f,g from $Abb(\Omega,R)$, evaluate

$$\varphi_R(f,g;\vec{\alpha}) := \min_{\vec{\zeta}\in R^3} \varphi(f,g;\vec{\alpha},\vec{\zeta}) \approx \min_{\vec{\zeta}\in\Omega} \varphi(f,g;\vec{\alpha},\vec{\zeta}).$$

From now on we only consider the integer translational movements, establishing the "$\approx$" in the equation above. The objective function for a fixed $\vec{\alpha}$ given by

$$\varphi^d(\vec{\zeta}_d) := \varphi(f,g;\vec{\alpha},\vec{\zeta}_d) = \sum_{\vec{x}\in\Omega} w(\vec{x}) \mid f'(\vec{x}-\vec{\zeta}_d) - g(\vec{x}) \mid^2,$$

with $f' := \hat{T}\left[\vec{\alpha},\vec{0}\right](f)$ and $\vec{\zeta}_d \in \Omega$, can be rewritten as:

$$\varphi^d(\vec{\zeta}_d) = \sum_{\vec{x}\in\Omega} w(\vec{x}) f'(\vec{x}-\vec{\zeta}_d)^2 - 2\sum_{\vec{x}\in\Omega} w(\vec{x}) f'(\vec{x}-\vec{\zeta}_d) g(\vec{x}) + \sum_{\vec{x}\in\Omega} w(\vec{x}) g(\vec{x})^2.$$

By using the mirror operator $\hat{S}(f)(\bullet) := f(-\bullet)$ and convolution for periodic functions

$$(f * g)(\vec{y}) := \sum_{\vec{x}\in\Omega} f(\vec{y}-\vec{x}) g(\vec{x})$$

we get

$$\hat{S}\varphi^d = f'^2 * \left(\hat{S}w\right) - 2 f' * \left(\hat{S}(g\cdot w)\right) + U_0$$

with U_0 a constant originating from the third term. The above equation holds, because of the following calculation for two periodic functions $f_1, f_2 \in Abb(\Omega, R)$:

$$\sum_{\vec{x}\in\Omega} f_1(\vec{x}-\vec{y}) f_2(\vec{x}) \underset{\vec{x}'=-\vec{x}}{=} \sum_{\vec{x}'\in\Omega} f_1(-\vec{x}'-\vec{y}) f_2(-\vec{x}') = \sum_{\vec{x}'\in\Omega} f_1((-\vec{y})-\vec{x}') \hat{S} f_2(\vec{x}') = \left(f_1 * \hat{S} f_2\right)(-\vec{y})$$

It is well-known, that the convolution of two functions f_1 and f_2 can be efficiently computed using fft by:

$$f_1 * f_2 = N\hat{F}^{-1}\left(\hat{F}(f_1)\cdot\hat{F}(f_2)\right),$$

whereby $\hat{F}$ denotes the Fourier transform and the factor N depends on the size of the domain Ω.

This leads us to the following form of the objective function (see Cowtan (2001) or Dranischnikow (2008) for more details):

$$\varphi^d = N\hat{S}\left(\hat{F}^{-1}\left(\hat{F}(f'^2)\cdot\hat{F}(\hat{S}w)\right) - 2\hat{F}^{-1}\left(\hat{F}(f')\cdot\hat{F}(\hat{S}(w\cdot g))\right)\right) + U_0.$$

By using fft we can calculate the best translation in $O(n\log n)$ whereas the naïve implementation would result in an $O(n^2)$ time complexity. It is easy to transform these formulae into an efficient parallel algorithm, which we will discuss in section 3.

We split the general alignment problem 1 into the translational alignment problem 2 just mentioned and the problem of finding the best rotational movement assuming that we had solved problem 2. This leads to the following subproblem.

Problem 3 *(Rotatory alignment-problem)*

Given f, g from $Abb(\Omega, R)$, evaluate

$$\min_{\vec{\alpha}\in R^3} \varphi_R(f,g;\vec{\alpha}).$$

To solve problem 3 we resort to conventional optimizers such as downhill-simplex from Nelder and Mead (1965) or Powell's method (Press et al (2002)). Our experiments clearly indicate, that the downhill-simplex method is more suitable (see table 2). In comparison to the initial problem, the search-space has been reduced to three dimensions. The main advantage of this procedure is that the objective function $\varphi_R(f,g;\vec{\alpha})$ possesses, according to experience, considerably less local extrema.

3. PARALLELIZATION

The evaluation of the objective function $\varphi_R(f, g; \vec{\alpha})$ consumes the lion's share of the computation time and consists of the following steps:

1. Computation of the rotated dataset by means of trilinear interpolation
2. Solving the translational alignment problem:
   (a) Squaring the values of the rotated dataset
   (b) Fourier transform
   (c) Multiplication (point wise) and addition of vectors in the frequency domain
   (d) Inverse Fourier transform
   (e) Search for the minimum

Steps 1, 2a and 2c do not demand any communication between the processes and are therefore easy to parallelize. It is known, that the Fourier transform (2b and 2d) possesses a quite high potential for parallelization (Gupta and Umar (1993), Tong and Swarztrauber (1987), and Swarztrauber (1987)). Searching for a minimum (2e) is a typical example for reduction (compare Zomaya (1996) or Harris (2007)) and can be efficiently implemented on most parallel architectures.

The architecture of our choice for the parallelization of the algorithm was a NVIDIA GPU. The CUDA-framework from NVIDIA makes it possible to use a GPU without casting the problem into a graphical API. Despite the simple handling, a careful implementation is necessary. Careless memory access can be devastating for the performance because of the absence of a cache. However, with texture and shared memory the programmer has enough tools on his disposal to evade most of the bottle necks caused by the relative slow global memory. Furthermore, with CUFFT (NVIDIA Corporation (2007)) an efficient implementation of the necessary fft routines is available.

We are using textures for computation of the rotated dataset because otherwise it is quite challenging to optimize the global memory access. The efficient global access for steps 2a and 2c is straight forward, because the simple data layout allows the simultaneous memory accesses to be coalesced into a single memory transaction. The optimal size of the blocks can be found in fftw manner (Frigo and Johnson (2005)) during a testing phase. This ensures that the best possible performance can be reached on different GPUs.

When using real-to-complex in-place Fourier transforms a padding is necessary (see NVIDIA Corporation (2007) for more details). Due to the SIMD architecture of the GPU it is more efficient to treat those padded elements the same way as the other "normal" elements.

Table 1. Comparison of the run times between GPU with trilinear interpolation, CPU with trilinear interpolation and CPU with nearest neighbor interpolation. Because of the textures the GPU has a kind of intrinsic advantage over the CPU, therefore we also offer a comparison with a CPU implementation, that applies the less costly nearest neighbor interpolation. We used an Intel Core2 Duo 2.4 GHz with NVIDIA 8800GTX (CUDA2.1) for this benchmarking test

Kind of Computation	Number of voxels in the dataset: 2^N.							
	N=18	N=19	N=20	N=21	N=22	N=23	N=24	N=25
CPU (time in ms)	99	220	410	870	1800	3500	7100	14000
CPU Nearest Neighbor (time in ms)	42	100	190	400	850	1700	3500	7000
GPU (time in ms)	22	25	40	63	86	150	290	500
Speedup CPU/GPU	4.5	8.8	10.3	13.8	20.1	23.3	24.5	28.0

Table 1 shows the gains, which can be achieved through parallelization on a GPU. The biggest speedup can be obtained for embarrassingly parallel tasks. The computation of the rotated volume profits enormously from the use of textures, which are faster than the normal global memory. The fft computation does not yield the same speedup because the main part of the communication must pass through the slow global memory. However the size of the CPU cache is insufficient for fft on bigger datasets, therefore the speedup factor of fft increases with the size of the dataset.

4. EXPERIMENTS WITH SYNTHETIC DATA

Before going to the real life examples, we investigate the ability to recover the rotation with help of synthetic data. There are two major questions to answer. The first one is which method downhill-simplex or Powell's method yields better results and is more suitable and the second one is, what is the probability of finding the right rotation.

We used two kinds of data: Rotations uniformly distributed over the whole range and only rotations up to 45° but otherwise uniformly distributed. Both 128x128x128-datasets were cut out from a bigger CT volume and share an identical region, which is precisely known. This scenario is similar to section 5.3, compare also figure 2. A matching was considered successful, when the inaccuracy of the calculated rotation did not exceed 1°.

Table 2: Experimental results with synthetic data. Success rate and the number of function evaluations per case are averaged over 500 runs

method	all rotations		rotation up to $\pi/4$	
	success rate	# evaluations	success rate	# evaluations
downhill-simplex	0.338	82	0.864	83
Powell	0.264	349	0.784	338

The experimental results in table 2 suggest that downhill-simplex turns out to be a better choice. The higher success rate and the lower computation cost are two powerful arguments. The high success rate implies, that a few starting configurations are sufficient to find the global minimum. It also turns out that an additional Gaussian noise does not have any notable impact on the success rate of the algorithm.

5. APPLICATIONS

In this section we are going to describe some real-world alignment problems of multi-modal data for which our method can be successfully used.

5.1 PET/CT-Matching

The research project FUNMIG (Fundamental processes of radionuclide migration) studies the behavior of inserted radioactive elements in diverse rocks. The results of computer simulations, which predict the migration of the radioactive substance, should be compared with experimental results in order to evaluate correctness and accuracy of the underlying model.

The inner structure of a rock has been investigated by means of CT, whereas PET was used for monitoring the diffusion of the radioactive elements. Because the two measurements were carried out from within two different coordinate systems (and as a matter of fact in two different cities), there is the necessity to find the right transformation in order to be able to compare the data (Ulenkampff et al (2008)).

The CT-record of the rock is available as a Boolean model. The hollow voxels are denoted by 1, the massive voxels by 0. This dataset must be matched with 5 PET records. Each record describes the configuration of the radioactive elements in the rock for a fixed point in time. The values of the voxels can vary between 0 (nonradioactive) and 2^{16} (very radioactive). All PET-measurements have been carried out from within the same coordinate system.

We assume that high radioactivity (the term "high" should be seen in relationship to the radioactivity of the other voxels) can be reached only in hollow voxels, because only there a large accumulation of radioactive elements is possible. It is indeed thinkable, that some thin fissures have not been registered by CT, but still can contain a non-negligible amount of radioactive atoms. However, the radioactivity of such voxels should not be very high. On the other hand, the radioactive material does not spread over the whole crevice; therefore hollow voxels exist that are not contaminated.

These considerations lead us to the objective function

$$\varphi(\vec{\alpha},\vec{\zeta}) := \sum_{p(\vec{x})>v} p(\vec{x})\,|\,c(\vec{x})-1|^2\,, \tag{1}$$

where c denotes the transformed CT data, p denotes the PET data and v the threshold from which on we consider a voxel as being radioactive. The objective function punishes voxels, that are considered to be radioactive (i.e. $p(\vec{x})>v$) but are not hollow. According to our assumption, that higher radioactivity increases the probability that a voxel is hollow, the punishment is severer for massive voxels with high radioactivity. This simple approach is sufficient to accomplish a successful matching, as the left hand side of figure 1 illustrates. Further details and results can be found in Dranischnikow (2008).

5.2 MRI/CT Matching

The next application of our registration method arises in medical science. The lung of a patient has been examined with the help of two different methods. Conventional CT acquires the tissue of the lung whereas inhaled ^{3}He gas made visible by means of magnetic resonance imaging (see Kauczor et al (1997)) depicts the parts of the lung with proper function. Merging the two datasets would help to diagnose the state of a disease more precisely and choose a more adequate treatment.

A certain resemblance to the problem above is evident. Comparing this scenario to the above, the lung plays the role of the crevice and the ^{3}He gas the role of the radioactive material. With this interpretation the objective function (1) can be reused and again delivers successful matchings, as figure 1, right-hand side, illustrates. More details and results are provided in Dranischnikow (2008).

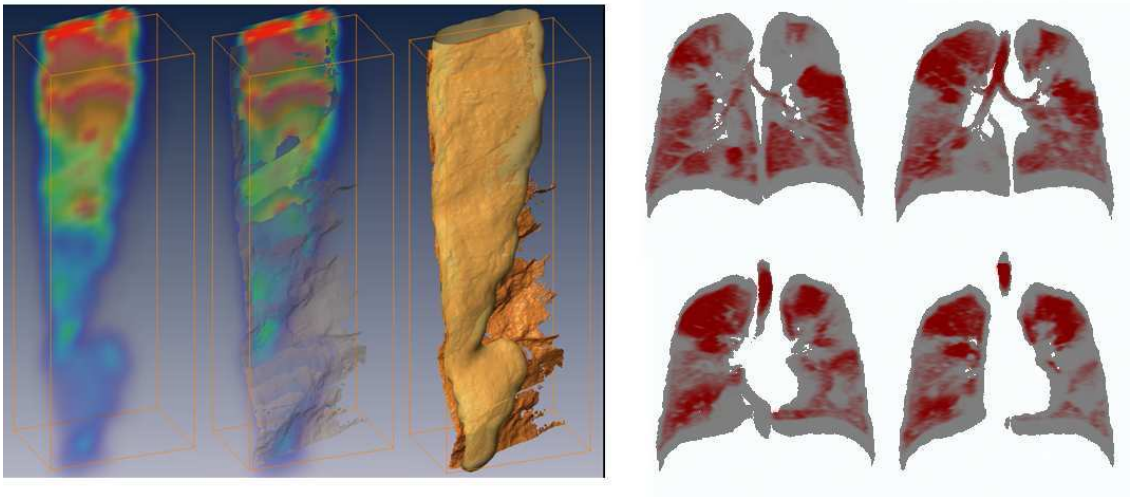

Figure 1. On the left-hand side: The CT and PET (after 60 minutes) datasets are successfully matched. First volume shows the PET data, where red color corresponds to high radioactivity. Second volume shows the matching with the CT data, where the crevice is depicted brown. Third volume shows the surface of the radioactive voxels matched with the CT data. On the right-hand side: Four slices through the lung of a patient: gray color displays the lung in the CT data, the more intense the red color the higher the concentration of ^{3}He.

5.3 CT-Fusion

The last application of our method copes with a task from the field of dentistry. Here sometimes an object (usually the head of a patient) is too large for an available cone beam CT with small or medium size field of view. Therefore it is necessary to perform two or more measurements. Afterwards the small datasets should be merged into one large dataset. Figure 2 (left-hand side) depicts, how this task can be reduced to the alignment problem of the shared region.

We choose $f :=$ first dataset, $g :=$ second dataset and

$$w(\vec{x}) := \begin{cases} 1, \vec{x} \in A \\ 0, \vec{x} \notin A \end{cases}$$

whereby A denotes the (estimated) shared region of two datasets. These definitions allow us to use our algorithm in order to merge the small volumes into a bigger one. The results are shown in figure 2 on the right-hand side.

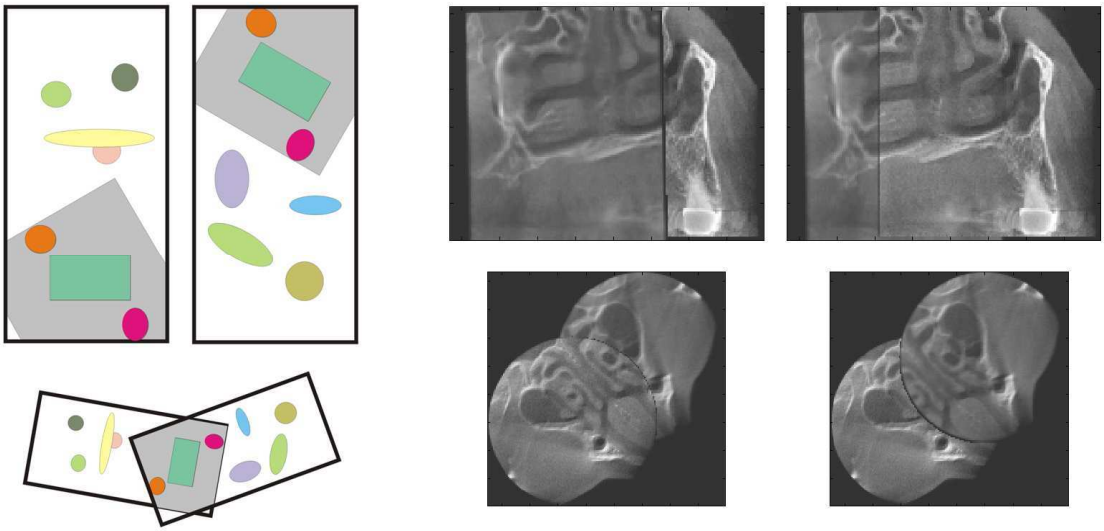

Figure 2. Left-hand side: The solution for the alignment problem of the shared region enables the required merging of two datasets. Right-hand side: The result of merging two small datasets into one big record of a jaw. Two possibilities to expand the data are shown. The left volume completed with parts of the right volume and the other way around. The accuracy of the merging as well as the junctions is clearly visible.

6. CONCLUSION

In this paper we presented a new method to register multimodal data. Our approach is based on a special metric that makes it possible to neglect subsets of the data that are for sure dissimilar. The proposed metric has been successfully used for matching multi-modal datasets. Experiments showed that this approach is applicable for PET/CT- and MRI/CT-matchings. The algorithm also operates on single-modal data and can be applied to register an object/pattern in an n-dimensional space.

The main advantage of our approach is the possibility to fft-accelerate the search in some degrees of freedom. In this paper we considered the acceleration of the search for the translational movement, which resulted in fewer local minima of the new objective function. This leads to a better performance in terms of success rate of matching as well as in terms of run time.

A further improvement can be achieved by using parallel hardware such as a GPU. A speedup factor up to 25 could be achieved using modern commodity graphics hardware. Thus, the utilization of GPUs for computation, because of the impressive performance that GPUs offer at a very modest financial cost can be an alternative to the use of expensive hardware such as combined PET/CT- or MIR/CT-scanner.

The proposed strategy for parallelization can also be used for other fft-accelerated algorithms such as the phase correlation method and therefore can substantially speed up those applications.

ACKNOWLEDGEMENT

We gratefully acknowledge Frieder Enzmann (Institut fuer Geowissenschaften der Johannes Gutenberg-Universitaet Mainz) and Oliver Weinheimer (Department of Radiology Johannes Gutenberg-University) for providing data and insight into the problems discussed in section 5.

REFERENCES

Castro de, E. and Morandim C., 1987. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5), pp. 700–703.

Cowtan, K., 2001. Fast fourier feature recognition. *Acta Cryst.*, D57, pp. 1435–1444.

Dranischnikow, E, 2008. *Parallele Algorithmen zum Alignment von Gitterbasierten Daten.* Diploma Thesis, Johannes Gutenberg-Universitaet Mainz.

Frigo, M. and Johnson, S, 2005. The design and implementation of fftw3. *Proceedings of the IEEE*, 93(3), pp. 2–231.

Gupta, A. and Umar, V. K., 1993. The scalability of fft on parallel computers. *IEEE Transactions on Parallel and Distributed Systems*, 4(8), pp. 922–932.

Haber, E. and Modersitzki, J., 2007 Intensity gradient based registration and fusion of multi-modal images, Methods of *Information in Medicine*, 46(3), pp. 292–299.

Harris, M. 2007, *Parallel Prefix Sum (Scan) with CUDA.* NVIDIA Corporation, 2007.

Kauczor, H.-U. et al, 1997. Imaging of the lungs using 3he mri: Preliminary clinical experience in 18 patients with and without lung disease. *J. Magn. Res. Imag*., 7, pp. 538–543.

Kovacs, J.A. , et al, 2003. Fast rotational matching of rigid bodies by fast fourier transform acceleration of five degrees of freedom, *Acta Cryst., D*, pp. 71–76.

Maes, F. et al, 1997. Multimodality image registration by maximization of mutual information. *IEEE Trans. Medical Imaging*, (4), pp. 187–198.

Mattes, D. et al, 2003. Pet-ct image registration in the chest using free-form deformation. *IEEE Trans. on Medical Imaging*, 22(1), pp. 0–8.

Nelder, J. A. and Mead, R. A., 1965. A simplex method for function minimization, *Computer Journal*, pp. 308–313.

NVIDIA Corporation, 2007. *CUFFT Library*.

Pan, W. et al, 2009. An adaptable-multilayer fractional fourier transform approach for image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 31(3), pp. 400– 414.

Press W.H. et al, 2002 . *Numerical Recipes in C++. The Art of Scientific Computing.* Cambridge University Press.

Reddy, B.S. and Chatterji, B.N., 1996. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8), pp. 66–71.

Swarztrauber, P. N., 1987. Multiprocessor ffts. *Parallel Computing.* 5, pp. 197–210.

Tong, C. and Swarztrauber, P. N., 1987. Ordered fast fourier transforms on a massively parallel hypercube multiprocessor. *J. Parallel and Dist. Comput.*, pp. 50–59.

Ulenkampff, J. K et al, 2008. Evaluation of positron-emission-tomography for visualization of migration processes in geomaterials. *Phys. Chem. Earth*, 33, pp. 937–942.

Viola, P. and Wells, W. M. III, 1997. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2), pp. 7–154.

Zomaya, A.,1996. *Parallel and Distributed Computing Handbook*, McGraw-Hill, 1996.