

# Medical Image Registration: A Survey of Modern Techniques

Jonathan Fomby

University of South Carolina, Computer Science & Engineering,  
301 Main Street, Columbia, SC 29208  
fombyj@cse.sc.edu

**Abstract.** Image registration is the process by which two different images or volumes, representing the same structure or data but from different perspectives, are translated and morphed as is necessary such that they can be overlaid to clearly depict the same structure from the same viewpoint. In medical imaging, this is often necessary when using information from different imaging techniques, such as MRI or CT. This is a difficult problem, with new research being done constantly in an effort to provide accurate and fast registration techniques. In this project I will discuss several recent techniques which aim at improving registration efficacy, analyzing their results and in general providing an idea as to the direction of up-to-date research in several key directions.

## 1 Introduction

Medical image registration remains an extremely difficult problem in computer science, and as a result, there is a large amount of ongoing research. This research ranges from broad topics covering general registration algorithms to much more focused studies which try to solve the problems for a single, detailed method, such as aligning images of a specific organ from two different modalities. Due to the copious and varied amounts of research and the large number of papers put forth each year, it can be difficult to keep up with all of the facets of modern image registration techniques. This is a rapidly changing field, with new hardware and software solutions constantly being put forth for the myriad problems which still exist for registration, in addition to the new ones which are being discovered. The original goal of this project was to comb over as much of the research as possible, gleaning from this information several common directions for the modern techniques. I was going to then present several papers representative of these paths, in an effort to provide a simpler, more manageable view of the current state of research. Over the course of my work I shifted the goal slightly. Instead of finding coherent research directions and techniques which were representative, I instead found it more manageable to identify several goals of the ongoing research, and present papers which clearly illustrate some attempts to reach these goals. In the end, I singled out three commonly represented goals. How these goals were hopefully achieved varied wildly, but many agreed they were worth pursuing. I decided upon three different goals: improving the success rate

of image registration, increasing the speed of registration, and improving non-rigid registration as much as possible, including both speed and accuracy. I chose a paper for each topic, and will present their methods and results in the following three sections. The paper will then conclude with an analysis of image registration as a whole, and then a brief conclusion.

## 2 Registration Success Rate

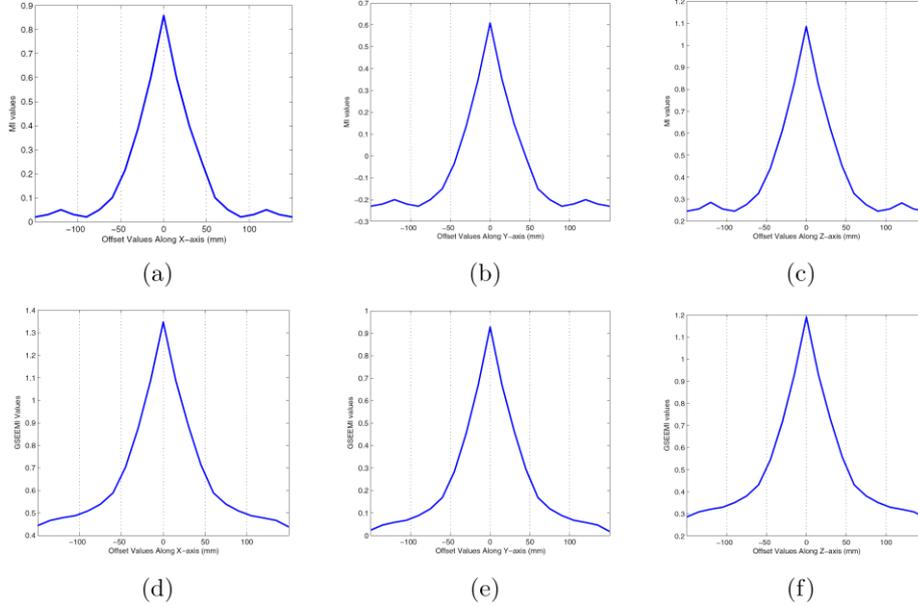
The success rate of image registration is important for obvious reasons. If the image is inaccurately registered, then it is of no use. It seems only natural that there would be research going into increasing the success rate of medical image registration in order to improve the field in general. To illustrate progress in this field, I chose a paper which uses a new similarity metric in order to improve success[1]. With a different method being used to check for similarity, we can potentially get more successful registrations. The results tend to show that this new method has a higher success rate than the traditional method, while having similar accuracy results for those images that are successfully registered

### 2.1 Methods

In this paper the authors compare two different similarity metrics. The first is the standard, often used Mutual Information (MI), based upon the classic Shannon's Entropy. In essence, the more information that the two images share, the more closely they are registered. By finding the transformation that would maximize the MI, one could presumably obtain a properly registered image. The authors propose the use of a different entropy for their mutual information measurement. Specifically, they propose the use of generalized survival exponential entropy (GSEE) in order to calculate GSEE mutual information (GSEE-MI).

Traditional Shannon entropy is based upon the density function of the data. The new proposed measure, GSEE, is instead based upon the cumulative distribution function, literally substituting that measurement in place of the density function in Shannon's definition. It should be noted that these two functions are related in that the density function is the derivative of the cumulative distribution function. Given that the functions and the entropies are so closely related, it might seem odd that one could be used to significantly better effect. However, the use of the distribution function and thus GSEE has several benefits, though one outweighs the others. Specifically, Shannon's entropy is based upon the density of a random variable  $p(X)$ , which may more may not exist, and if it does it is estimated. The problem is that this estimated value only converges sometimes, it cannot be guaranteed. As such, the authors use the new measure of mutual information based upon GSEE rather than Shannon's Entropy, thus resulting in GSEE-MI.

Part of the problem using Shannon's entropy is errors introduced by interpolation. When attempting to register the images, it is necessary to gather information such that the measure of similarity can be compared. This is done



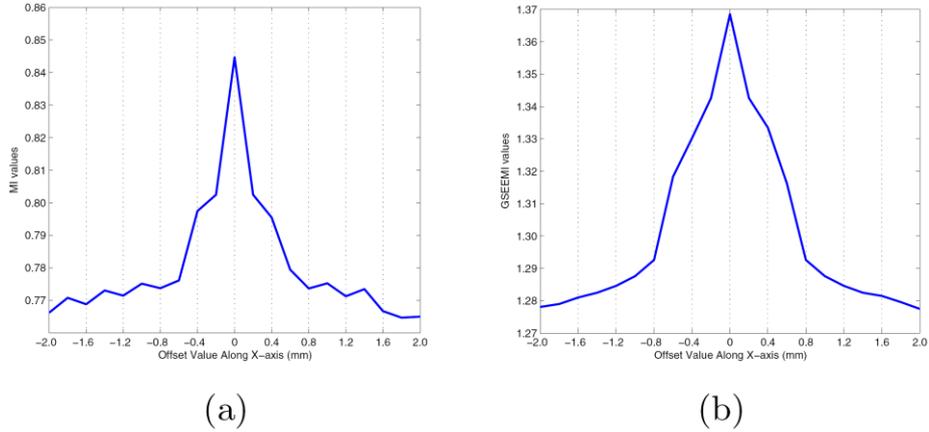
**Fig. 1.** MI: Shifting Probes Along (a) X-axis and (b) Y-axis and (c) Z-axis; GSEE-MI: Shifting Probes Along (d) X-axis and (e) Y-axis and (f) Z-axis

via the joint intensity distribution, which compares the intensity of the two images. The commonly used method of obtaining this data involved histogramming with partial volume interpolation. It is this interpolation which introduces artifacts as it estimates sub-voxel properties, which can cause further errors in registration.

Based upon both the traditional measure of MI and their new measure GSEE-MI, the authors ran a series of unsupervised registration experiments. Four pairs of CT and MRI volumes are used, each having already been successfully registered such that there were no errors greater than 1mm. As such, there was a solid ground truth registration already known for each pair. The authors then altered all six measurements, translating and shifting across all three axes in a uniformly distributed manner. These perturbations were repeated 100 times for each image, resulting in a total of 400 registrations to take place. Success was measured by a registration containing errors no more than  $2mm$  or  $2^\circ$ .

## 2.2 Results

A quick way to verify the new similarity metric graphically before any registration even takes place can be seen in Figure 1. Here we have the measure of MI using the two different techniques as an image is translated across the three axes. For the traditional measure of MI, we see local maxima as the alignment drifts farther from true registration. When the algorithm was trying to maximize MI, it could get caught at one of these local maxima, thinking that it was not possible to get any better. The new



**Fig. 2.** (a) MI Translational Probes Along X-axis in the range  $[-2,2]$  mm; (b) GSEE-MI Translational Probes Along X-axis in the range  $[-2,2]$  mm

GSEE-MI metric, on the other hand, contains no local maxima. Due to the continuous distribution function, we see a smooth curve all the way to the one desired peak, totally devoid of any local maxima. Figure 2 shows a smaller window on the X-axis only, but further serves to illustrate how the new technique is beneficial in its not being as negatively affected by PV interpolation. With the old technique, we again see total maxima, while even on this small of a scale the benefit of the GSEE-MI metric can be seen, as we have much smoother curves without the local maxima.

We can see the success rate of the two different methods in Table 1. Here, we can see a dramatic difference in the results using the two different methods. Success rate goes from mid-sixty percentages to mid-nineties. This is a substantial increase.

**Table 1.** Success rates of the conventional MI-based approach and the proposed GSEE-MI-based approach

Testing dataset	MI Success Rate	GSEE-MI Success Rate
#1	64%	95%
#2	63%	93%
#3	72%	97%
#4	65%	89%

In addition to the above results, which demonstrate the new method's greatly improved success rate, Table 2 gives the mean and standard deviation of the registration errors for those image pairs which were successfully registered. It should be noted *only* successful registration pairs are included in this table. For those that were deemed unsuccessful, it would not be worthwhile to include their errors; this table shows that for those registrations that were successful, they were comparable to existing techniques. As such, this method appears to be more successful than existing techniques, while maintaining comparable quality for correct registrations.

**Table 2.** Registration Accuracy of the conventional MI approach and the proposed GSEE-IM approach

Method	Translation ( $10^{-3} mm$ )			Rotation ( $10^{-3} degrees$ )		
	$\Delta t_x$	$\Delta t_y$	$\Delta t_z$	$\Delta \theta_x$	$\Delta \theta_y$	$\Delta \theta_z$
MI	0.64±1.84	-0.37±0.73	1.16±1.32	0.85±1.84	-1.05±1.47	0.86±1.63
GSEE-MI	0.85±1.64	-0.62±0.34	1.05±1.57	1.32±0.62	-0.83±1.58	0.52±1.15

### 2.3 Analysis

The results seem to indicate that the proposed method should definitely be considered as a replacement for the traditional use of Shannon’s entropy-based MI. The GSEE-MI method had a significantly higher success rate, while maintaining comparable accuracy.

There are some flaws with the paper however. For one, the authors give no information as to the running time of the different methods. Based upon their similarity I would assume that the run-time would be comparable, but some data would have been nice to have. In addition, they mention a different way to approximate the joint intensity distribution, namely Parzen windowing, but dismissed it due to it being more computationally expensive. Without providing any run-time information on their new method, it would be hard to compare it to existing implementations using this Parzen windowing technique. Again, it is probably obvious for those who have done much research in this area how the different methods would perform, and so they felt that they could leave it out, but for someone learning it is unclear.

Regardless of the flaws, the paper still stands as a good example as to how even extremely traditional methods, perhaps especially extremely traditional methods, may need to be reevaluated periodically in order to find out if there is not a better method out there. They have shown that there are still ways to increase the success rate of even the simplest of image registration problems.

## 3 Registration Speed

Image registration is often an extremely computationally expensive procedure. The 3D volume must often be translated and rotated in every possible direction in order to find the correct place to move for one individual step. Working with so many voxels which need to be compared to each other adds up quickly. Even with hardware getting progressively better, it cannot be relied upon alone to speed up image registration algorithms. As such, the issue of speeding up image registration is being greatly researched. To illustrate this topic, I am going to describe a paper which attempts to apply some novel algorithms to image registration, using the growing field of parallel computing as a way to speed up performance [2]. They show that there is potential in parallel computing when it come to image registration, but the algorithms have to

work with the parallelization in order to overcome the overhead associated with parallel computing

### 3.1 Methods

It seems that almost all of the new processors being developed for home use are currently dual-core, with quad-core processors on the horizon. Parallel computing is becoming the norm rather than the exception, and it only seems natural that such a computationally expensive task as medical image registration should make use of this available power. This is the task to which the authors have set themselves. However, parallel computing needs a cooperative algorithm if the hardware is going to be fully exploited. As such, the authors have decided to introduce some non-standard algorithms to the task of image registration. They compare some of the current popular techniques with several variations of their two new algorithms, analyzing the results of running them on anywhere from 1 to 12 processors.

The first new algorithm which they introduce is called DIRECT (DIviding RECTangles). This is a region-based optimization algorithm attempts to minimize/maximize some similarity function. In this paper, the author chose to use mutual information. The algorithm works by dividing the  $n$ -dimensional search space into an  $n$ -D rectangle. It then recursively divides this rectangle into smaller rectangles, attempting to find local minima on a global scale, and the “basin of attraction” on the local scale. In essence, it divides the rectangle into thirds along the longest sides, and proceeds to further subdivide the rectangles which are potentially optimum. Eventually it converges, when the rate of change for the similarity metric reaches a certain point. In all honesty, I still don’t fully understand the algorithm, but the important thing to remember is that due to its nature, it is highly parallelizable. It requires several rectangles to be analyzed at once, and this can be efficiently distributed to multiple processors or cores.

The other algorithm the authors introduce which they feel might benefit from parallelization is the Multipledirection Search (MDS) algorithm. Again, it is one for those with experience in number theory. MDS is an iterative simplex based optimization algorithm, employing a simplex with  $n+1$   $n$ -D vertices. As with the DIRECT method, I do not believe that I have the math background to fully understand as presented in the paper, and resources were hard to find without paying. As such, I will simply state that while the DIRECT method works better on the global level at establishing general registration, the MDS method is better on a local level, and so would hypothetically be better for finding local optimums. In essence, one could get a rough registration with DIRECT, and then fine-tune it with MDS, which is the intent of the authors.

As mentioned, the authors use mutual information as their similarity metric. However, instead of just using traditional Shannon-MI, they use a normalized form as well, conveniently called Normalized Mutual Information (NMI). The reason for this is that NMI is not as affected by how much overlap there is between the source and target images. When registering across modalities, one image may depict a larger part of the body than the other, and it is important that the similarity metric takes into ac-

count the actual similarity that is there and does not penalize for a smaller/less complete image. The authors also chose three different interpolation methods in order to test the performance of each. Since this paper is an attempt to increase the speed of the calculations, and the interpolation technique can greatly affect speed, they chose three fast ones, trilinear (L1), partial volume (PV), and jittered-blur (JB), in order to see how they would affect performance.

For test data, the authors got images from three different sources. First, synthetic brain images were obtained from a reputable Medical Imaging website, BrainWeb. Second, they used cardiac MRI volumes obtained from various stages in the beat cycle. Third, actual MRI brain volumes were obtained from a MS patient. Tests were run on all three sets of data. Volumes were translated and rotated across all three axes, and results were obtained for a large combination of different interpolations and optimization techniques. In order to compare performance of traditional algorithms on parallel systems, a very popular method called Powell's method was used. They also broke registration up into global and local steps in some phases. In all, they used five different registration techniques, with all three different interpolation techniques each. The registrations were done using DIRECT (D), DIRECT then MDS (D/M), DIRECT then Powell's (D/P), DIRECT then a locally-biased DIRECT (D/LD), and finally Powell's on its own. Registrations were considered to be successful if there were no errors greater than  $1mm$  or  $1^\circ$ .

### 3.2 Results

While the paper is primarily focused on the speed of these algorithms in parallel, speed is useless if the images cannot be successfully registered. As such, the authors carefully detailed their success rate across all three sets of data and all combinations of techniques. In order to preserve space, I am only going to present a portion of their findings in Table 3 below. Here, we see the results from the clinical brain images. It is worth noting that before speed is even taken into account, DIRECT performs better

**Table 3.** Clinical Brain Success Rate (All Misrotations)

$d_0$ (mm)		NMI					MI				
		5	10	20	30	mean	5	10	20	30	mean
D	PV	1.00	1.00	0.97	1.00	0.99	0.98	1.00	0.93	0.99	0.97
	LI	0.97	0.99	0.96	0.98	0.97	0.98	0.99	0.96	0.97	0.97
	JB	0.96	0.99	0.95	0.99	0.97	0.99	0.99	0.91	0.80	0.92
P	PV	0.86	0.79	0.55	0.31	0.63	0.93	0.91	0.71	0.71	0.81
	LI	0.93	0.84	0.54	0.33	0.66	0.97	0.91	0.79	0.69	0.84
	JB	0.74	0.68	0.46	0.26	0.53	0.90	0.91	0.83	0.58	0.80
D/ M	PV	0.73	1.00	0.79	0.93	0.86	0.77	0.96	0.85	0.90	0.87
	LI	0.66	0.79	0.69	0.77	0.73	0.73	0.86	0.78	0.83	0.80
	JB	0.70	0.66	0.59	0.74	0.67	0.83	0.72	0.63	0.86	0.76
D/ P	PV	0.71	0.89	0.70	0.81	0.78	0.78	0.96	0.85	0.89	0.87
	LI	0.69	0.86	0.72	0.73	0.75	0.74	0.97	0.88	0.90	0.87
	JB	0.61	0.66	0.68	0.70	0.66	0.83	0.94	0.84	0.86	0.87
D/ L	PV	0.88	1.00	0.88	0.98	0.93	0.85	0.99	0.80	0.92	0.89
	LI	0.89	0.99	0.91	0.97	0.94	0.82	0.98	0.82	0.91	0.88
	JB	0.91	0.92	0.76	0.97	0.89	0.91	0.81	0.92	0.91	0.89

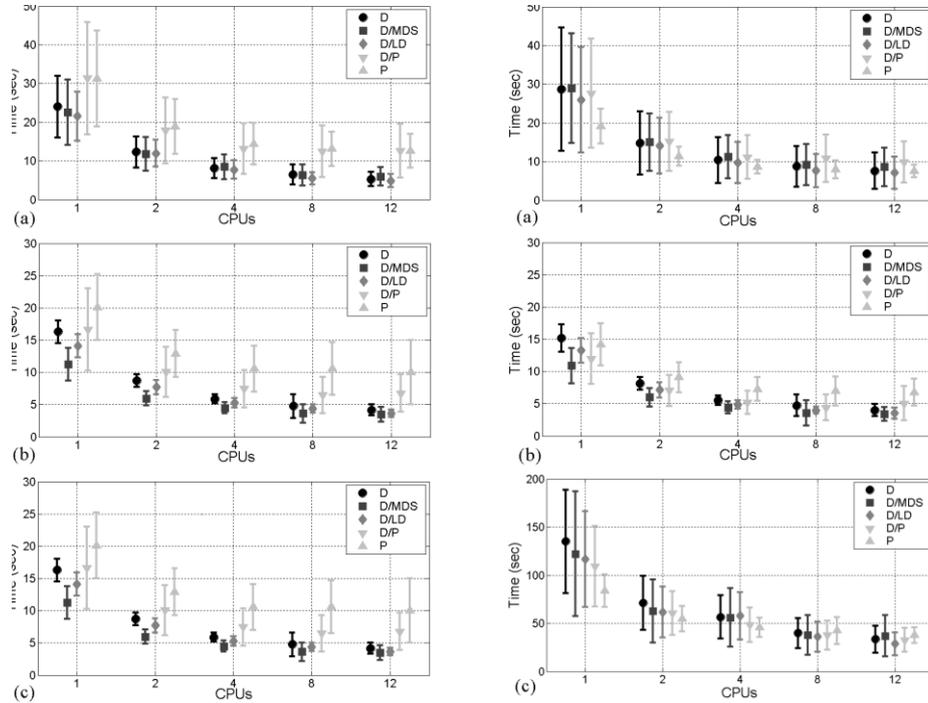


Fig. 3. Timing results using MI. (a) BrainWeb MRI PD-T1 (9% noise), PV, (b) Systole-diastole cardiac, JB. (c) Clinical brain, PV. Left side is with MI, right side is NMI

in terms of success rate than the traditional Powell's method. Indeed, it outperforms all of the methods. The similarity metric does not seem to have a dramatic effect in this case, nor does the interpolation method influence things too terribly. In some situations one might be better than the other, but on the large scale they seem about equal. However, this is just the success rate results, and the authors are also highly focused on performance. We can see these performance results in Figure 3, and they are somewhat surprising. Even when working on only one processor, the DIRECT method is often on par or faster than Powell's method. Adding a second processor speeds things up substantially across the board. After two processors, adding more does not speed things up as substantially, but this is to be expected. With parallel computing, there is some overhead associated with communicating and sharing memory between the processors. Eventually it gets to the point where adding another processor is of no real benefit due to this overhead, and we can see that in the diminishing returns as we add more processors. It is also interesting to note that, while the DIRECT/MDS implementation usually did better on only one processor, when more parallelism was available, DIRECT usually started to catch up. This suggests that DIRECT may be best used on its own, which was not expected due to the global nature of its search.

### 3.3 Analysis

The paper does a good job of demonstrating the potential speedup available with the use of parallel computing. It is somewhat surprising that it levels off so quickly, and that more processors than 12 would seem to offer very little benefit, however one can only stretch things so far. It might seem that the decreases in run-time are not terribly dramatic, however. Going from 30 seconds to 10 seconds is nothing terribly special, but it does get us that much closer to real time registration. Additionally, in an operating room setting every second can count.

The biggest increases in performance, however, would likely be shown in the performance of non-rigid registration. While rigid registration can be done in 15 seconds, non-rigid registration may take 6-7 hours. Here, cutting the time by  $2/3$  could have a dramatic difference. This seems to be where the true benefits of parallel computing will be seen, and indeed the authors mention that their next step is to expand their algorithms to non-rigid registration to see if they can improve matters there.

## 4 Non-Rigid Registration

One of the biggest fields of research in image registration right now is in dealing with non-rigid registration, in which a non-uniform transform is applied to the image being registered. Points relatively near each other may be shifted in different directions depending upon the transform. Many people consider rigid registration to be partially solved, though as we have seen there is still more work to be done. However, in comparison, non-rigid registration is a much larger problem. The complexity and computation time differences are dramatic. For this section I chose a paper which attempts to use an existing technology in a new way, using Nonuniform Rational B-Splines (NURBS) to aid in non-rigid registration [3]. The authors successfully implement this technique, which proves to be more robust than the current alternatives they used for comparison

### 4.1 Methods

Non-rigid registration usually begins with a rigid registration to get the images as aligned as possible before the non-rigid transformation takes place, and the same practice is used in this paper. The non-rigid part is the interesting bit. Free Form Deformations (FFD) are often used in non-rigid registration. With FFDs, the volume to be registered is placed inside of a flexible solid. The solid is then deformed in a manner to get the image volume to deform similarly to match up with the target image. This deformation allows interpolation to shift voxels around to hopefully get a fully aligned image. Traditionally this is done via a B-spline, with nodes placed regularly throughout the solid. This paper proposes the use of NURBS in place of the traditional B-spline.

NURBS are a specialized form of B-spline with several desirable qualities. First, the nodes can be placed non-uniformly in the solid. This means that we do not have to

have as many nodes in the corners of the volume where there is little useful information. Conversely, we can have a dense clustering of nodes in information dense portions of the volume. The more information there is, the more control we need over the deformation in order to increase the accuracy of any interpolation that is done. Another beneficial quality of NURBS is that they essentially have weights at nodes. These weights affect how strongly the deformation will be drawn to a particular node. By adjusting these weights, we can get more flexibility in our deformation.

The proposed algorithm starts with initialization points. In keeping with the theory that we want dense placement of nodes in information rich portions of the image, the two images are first segmented to differentiate background from relevant information, brain images in this paper. Based upon these two masks, control points are then placed on the image to be registered. They are sparsely placed in the background, densely placed in the brain portion. The algorithm then goes in an optimization cycle, working on subspaces in the image in an attempt to maximize NMI. It does this via Powell's method, analyzing the result of moving each control point along one axis at a time to find the best results. This is iteratively across all control points, dealing with subsets of points at a time, shifting weights as needed. There is also a step present in order to prevent tearing of the image, which would happen if one portion of the solid essentially broke through another part, resulting in an impossible deformation. This is a very expensive algorithm computationally, and it takes a long time to complete. Eventually, the algorithm will converge and optimization is reached.

The authors engaged in two experiments to test their method. First, they took an existing brain image, and deformed it with several known transforms created from thin plate splines. They then applied their algorithm to the original brain image, trying to register it to these known deformations. They compared their results to a similar experiment using traditional B-splines. Secondly, they took a group of 14 brain images, each from a different person, and attempted to register 13 of them to the one remaining one in order to provide a generic picture of the brain.

## 4.2 Results

The results show a decent improvement, though not as dramatic of one as might would be hoped. Table 4 shows the results of the first experiment where they attempted to recreated a known deformation, displaying correlation coefficients (CC) as well as NMI for both the traditional FFD based method and their new NFFD based method.

**Table 4.** CC and NMI values

Warp	CC		NMI	
	NFFD	FFD	NFFD	FFD
1	0.97875	0.94454	1.245	1.204
2	0.97007	0.93797	1.237	1.223
3	0.98754	0.94879	1.312	1.287
4	0.98478	0.95875	1.244	1.241
5	0.98587	0.95417	1.328	1.314

**Table 5.** Average overlap coefficients (in %) for different tissue classes after affine, FFD, and NFFD registration

	Affine	FFD	NFFD
Gray Matter	67.27	83.71	91.27
White Matter	71.68	87.43	90.68
Cerebral spinal fluid	56.44	71.29	80.74

As can be seen, the NFFD had slightly better results. However, the FFD method was already 95% accurate, so improving it to 98% does not seem to be that dramatic of an improvement. Mean and maximal error for this experiment were also provided, and showed similar results. The new NFFD method performed better, but not by a large margin.

More dramatic are the results from the second experiment, where 14 brain volumes were all registered together. The results of this experiment are showing in Table 5. The correct registration of the white matter was not improved much, but that is to be expected since the white matter in the brain is the most easily distinguished part, so registration there was already fairly good. On the more difficult parts, the gray matter and the spinal fluid, we see more dramatic improvements. The new algorithm is better at registering the more complicated parts, which makes sense given that there was probably much more dense node placement there with the NFFD method than with the FFD method.

### 4.3 Analysis

The new NFFD method is a definite improvement over the regular FFD method, if not necessarily a dramatic one. This helps to demonstrate the difficulty of the non-rigid registration process. Dramatic improvements are hard to come by. However, the use of NURBS might be a step in the right direction. Since there are so many data points which need to be tested for optimal deformation, taking up a great deal of time and computational energy, it is important for an algorithm not to spend too much time on those points which will not greatly affect the final deformation. This is part of the benefit of the NURBS method. It reduces the amount of data it has to work on by focusing on the more “important“ parts of the image being registered. Additionally, it allows for higher resolution in the information dense portions in order to more closely register them. In essence it does away with a generalized method which assumes all image volumes are alike, and allows for much more custom deformations to take place by being much more flexible with node placement. This seems to be a logical path to take with research. Try to find an algorithm which will specialize itself for the image volume at hand, not one that tries to force the volume to a generic standard.

Additionally, the authors included run-time information, which is extremely helpful. If the algorithm took twice as long, then it would be debatable whether or not the improvements would be worth the computation time. However, this turns out not to be the case. The average run time for the FFD method on their test system was about 6 hours, whereas the average run time for the NFFD method was about 7 hours.

It is somewhat discouraging that the new, more accurate method takes longer, but it does not take dramatically longer. This extra time is most likely spent recalculating node positions, especially for those areas with dense node placement. In all, this new method provides more accurate results in a very similar timeframe.

## 5 Conclusion

As has been seen, medical image registration is a topic rife with research possibilities, and indeed there are many people out there pursuing this research in all directions. It is such a broad field that it is hard to nail down any one set path, but I have demonstrated how common goals can exist across the majority of this research. The techniques need to be fast and successful, and these traits need to expand out to non-rigid registration. It is interesting to note that all three of the papers discussed had many similarities in them. Powell's method was mentioned often. Maybe there is a way to improve on it, or perhaps a way to make it run in parallel more effectively. All papers used some form of mutual information as a similarity metric. It would be interesting to see the results of using NMI in the first paper's experiments, or using GSEE-MI when working on non-rigid registration. One of the most obvious paths is the expansion of parallel computing algorithms to non-rigid registration, which the authors of the second paper claimed they intended to do. It is time to fully exploit the hardware out there when addressing such a computationally expensive procedure. It is easy to see how the results of one study can in turn affect those of many, many others. While the field is wide and varied, there are many common themes intertwined throughout, and new methods can easily ripple down the chain, spawning new ideas and methodologies all the time. As such, it is important for those doing the research not to get too embedded in one particular methodology. Survey papers like these, though much more expansive than this one, are important in that they can help tie together all of those working on these complicated problems, and breathe new life into old ideas.

## References

1. Chung, A., Liao, S.: Multi-modal Image Registration Using the Generalized Survival Exponential Entropy. In: Larsen, R., Nielsen, M., Sporring, J. (eds.): MICCAI 2006, LNCS 4191. Springer-Verlag, Berlin Heidelberg New York (2006) 964-971
2. Peters, T.M., Wachowiak, M.P.: High Performance Medical Image Registration Using New Optimization Techniques. IEEE Transactions on Information Technology in Biomedicine 10(2) (2006) 344-353
3. Wang, J. and Jiang, T.: Nonrigid registration of brain MRI using NURBS. Pattern Recogn. Lett. 28(2) (2007), 214-223