A Multiscale Approach to Mutual Information Matching

J.P.W. Pluim,
J.B.A. Maintz,
and M.A. Viergever

UU-CS-1998-21
August 1998

ISSN: 0924-3275

A multiscale approach to mutual information matching

Josien P.W. Pluim, J.B. Antoine Maintz and Max A. Viergever

Image Sciences Institute, Utrecht, The Netherlands

ABSTRACT

Methods based on mutual information have shown promising results for matching of multimodal brain images. This paper discusses a multiscale approach to mutual information matching, aiming for an acceleration of the matching process while considering the accuracy and robustness of the method. Scaling of the images is done by equidistant sampling. Rigid matching of 3D magnetic resonance (MR) and computed tomography (CT) brain images is performed on datasets of varying resolution and quality. The experiments show that a multiscale approach to mutual information matching is an appropriate method for images of high resolution and quality. For such images an acceleration up to a factor of around 3 can be achieved. For images of poorer quality caution is advised with respect to the multiscale method, since the optimisation method used (Powell) was shown to be highly sensitive to the local optima occurring in these cases. When incorrect intermediate results are avoided, an acceleration up to a factor of around 2 can be achieved for images of lower resolution.

Keywords: multimodality registration, mutual information matching, multiscale image analysis, brain images, magnetic resonance, computed tomography

1. INTRODUCTION

The recent diversification of medical imaging techniques has stimulated research into methods for integrating different types of images of a single patient. Matching (registration) of images is the process of bringing these images into spatial correspondence. This may serve various purposes. One example is the registration of functional brain information, obtained by a PET or fMRI scanner, and anatomical information, captured in an MR image.[1–3] With integrated visualisation of the registered images, the anatomical information can act as a reference framework for the functional information.[4,5] Another application for matching is the determination of the patient position during treatment, e.g. in neurosurgery or radiotherapy, with respect to images taken previously.[6–9] Registration of these pre-operative images and real-time intra-procedural images relating to the current patient position will aid the clinician in several ways: movements of probes or surgical instruments can be tracked and displayed in the high quality pre-operative images during the treatment or can even be computer guided towards a predefined location within the patient and treatment plans based on the pre-operative data can be checked.

Over the years, research into multimodality registration has produced a wealth of different methods; surveys with a classification of approaches can be found in.[10–12] Important issues concerning the choice of a matching method are accuracy, robustness, speed, reproducibility, patient friendliness and the amount of user interaction required.

Matching methods can be classified into *extrinsic* and *intrinsic* methods with a further subdivision of the intrinsic methods into *segmentation based* methods and *voxel property based* methods.

Extrinsic methods align registration points that have been specifically added for this purpose, such as fiducials applied to either a stereotactic frame or the patient's skin or skull. These markers are constructed in such a manner that they are clearly visible in the different types of images. Alignment of the markers in the images produces the desired registration.[13] Matching of frame and screw markers is still considered one of the most accurate methods up-to-date.[14] The drawbacks of such methods are that they are time-consuming, can be burdensome to the patient and they can not be applied retrospectively.

Intrinsic methods use information that is inherent to the scanned object. Segmentation based intrinsic methods extract features from the images and employ only this extracted information to match the images. One such

Corresponding author:
E-mail: josien@cv.ruu.nl
Address: Image Sciences Institute, University Hospital Utrecht, Heidelberglaan 100, Room E01.334, 3584 CX Utrecht, The Netherlands

segmentation based registration routine is that involving anatomical landmarks.[1,6,15] These anatomically distinct points are usually defined by a user, which makes the issues of accuracy and reproducibility user-dependent. Other segmentation based methods match image structures, such as curves,[16,17] surfaces[2,18–21] or combinations thereof.[22–24] Drawbacks of these methods are that user-interaction is often required for the extraction of the structures, which may introduce inaccuracies. Moreover, these methods may be very application-specific.

Voxel property based intrinsic methods use the voxel grey values of the images for their registration, with varying degrees of preprocessing. Preprocessing steps can entail the extraction of features, such as ridges or edges,[25–27] or the application of filters to increase the similarity in voxel intensities.[28] Statistical measures are usually employed for the registration of the (preprocessed) images. Popular measures include cross-correlation of grey values,[29] entropy[3,30] and feature space dispersion.[21,31] Since the grey values of the images are used to compute a measure of registration, these methods are very general and require little to no user intervention.

Some of the more recent voxel property based matching methods make use of *mutual information*. Mutual information is a concept from information theory, measuring the degree of grey value dependency between images. The dependency is assumed to be maximum when the images are matched. The use of mutual information in image registration has yielded excellent results.[3,32–36] Unfortunately, mutual information based matching is a time-consuming method, taking (potentially) the entire image content into account in each step of the matching process.

In this paper, the embedding of mutual information matching in a multiscale approach is considered. The aim is to increase the speed of matching, without significantly compromising the method's accuracy and robustness. Several scaled versions of the images are constructed and these are matched in a coarse-to-fine manner, using the solution found at one level as the initialisation for the next level. The results are compared to direct, non-multiscale matching of the original datasets to investigate the effects of the multiscale approach on accuracy and robustness. We have tested this on three-dimensional CT and MR datasets of different resolution and quality, to emulate the variety in image quality of current medical imaging techniques. The experiments are restricted to rigid transformations, which are determined by six parameters: one translation and one rotation along each of the three axes. Although a rigid transformation is not satisfactory in cases when a brain shift occurs (owing to, for example, opening of the skull and drug induced reduction of the brain's water content),[37] it does suffice in many applications that do not require opening of the head as, for example, radiation therapy of brain tumours.

2. MULTISCALE MATCHING

The general idea behind a multiscale approach to image matching is that rough estimates of a desired solution are found using coarsely scaled images and the fine-tuning of the solution is done at finer scales. The purpose of the multiscale approach in this paper is to accelerate the matching process. In theory, this should not adversely affect the accuracy of the method if the multiscale hierarchy ends with the original images. The robustness of the multiscale approach depends on the robustness of the optimisation method used with respect to local optima in the registration function. Whether these theories hold true in practice will be the subject of discussion in this paper.

Downscaling of the images can be achieved in a number of ways: from relatively simple techniques, like down-sampling (i.e. sampling only one out of each group of voxels) or averaging the grey values of a group of voxels, to more refined methods as e.g. Gaussian blurring.[38] In the multiscale approach, matching is started with downscaled versions of the images. The resulting transformation serves as the initialisation for the registration at the next level and this continues until the finest level has been reached. Theoretically, this should reduce the overall computation time of matching methods whose complexity depends largely on the number of voxels contained in the images, as is the case with mutual information matching. The downscaling of the images results in a reduction in both the number of voxels and the information content, which will yield matching results faster. At the finer scales, matching should also be performed more rapidly, since the starting position in the search space is in the vicinity of the final solution. Ideally, the sum of the registration times at the various scales will be considerably less than the time required to directly match the original images.

The downscaling method chosen in this paper is equidistant sampling. We have experimented with averaging of voxel grey values over a cubic neighbourhood, but this approach rendered the matching of blurred images less accurate. Moreover, with this downscaling technique the computation time of multiscale matching proved considerably longer: largely due to the time required to compute the hierarchy of averaged images, but also to the slower convergence of

the registration process. More advanced downscaling techniques, such as Gaussian filtering, suffer less from negative effects of the blurring method on the accuracy of the results. Whether this outweighs the considerable computation time is currently being investigated.

3. MUTUAL INFORMATION

Registration of the images is done by maximisation of the *mutual information* of these images. Mutual information is a concept that originates from information theory.

The mutual information I of two images M and N is derived from the probability distributions of their grey values, m and n, as follows:

$$I\,(M,N) = \sum_{m \in M, n \in N} p_{MN}(m,n)\ \log \frac{p_{MN}(m,n)}{p_M(m)\ p_N(n)}, \tag{1}$$

where $p_M(m)$ and $p_N(n)$ are the marginal probability distributions of the grey values of images M and N respectively and $p_{MN}(m,n)$ is the joint probability distribution of the images' grey values.

In words, mutual information can be described as a measure for the amount of information two images contain about each other. Or, as a measure for the dependency between the grey values in the images. This dependency between the images is assumed maximum when the two images are matched.

The strength of mutual information lies in the fact that it is a general measure which does not usually require preprocessing of the images. It is more generally applicable than other statistical measures, such as cross-correlation. These measures rely on an *equivalence* relation between the intensities of two images, whereas mutual information depends merely on the existence of a statistical relation. Mutual information is therefore more suitable for multimodality image matching than measures such as cross-correlation or grey value differences, which pose stricter demands on the relation between the images.

In our approach, a rigid transformation is sought, traversing the six-dimensional parameter space by *Powell's method.*[39] This method iterates a series of one-dimensional maximisations for each dimension. Having found an optimum in one direction, the maximisation is continued in the next direction, starting from the current position. Once all six parameters have been optimised, the loop is repeated until the improvement achieved in the most recent iteration is within pre-defined boundaries. The one-dimensional maximisations are done with *Brent's method.*[39] First, an interval containing a maximum is determined (bracketing): around the current position, a small interval is defined and this interval is then moved and enlarged until it holds a maximum. Next, a parabola is fit through the end points of the interval and the maximum. The function value at the parabola's maximum is compared to the maximum found so far. The higher of the two becomes the new maximum; the lower of the two becomes a new end point of the interval together with the old end point on the other side of the new maximum, thereby reducing the interval. The descriptions above capture the basic ideas behind Powell's method and Brent's method; more elaborate descriptions can be found in Press et al.[39]

The mutual information value of the two images is calculated for the overlapping parts of the images only. In most cases, the transformed position of a voxel in one image will not coincide exactly with a voxel position in the other image. Therefore, an interpolation is required to determine which entry of the joint histogram of image grey values (from which the joint probability distribution of the images' grey values is deduced) is to be increased. Our choice is *partial volume interpolation*: for every voxel v in the transformed image, the eight nearest neighbours u_i in the other image are determined and the histogram entries of (v, u_i) are enlarged by a certain weight (these weights being inversely proportional to the distance between v and u_i). This method allows for subvoxel accuracy (in contrast to nearest neighbour interpolation) and does not introduce new, interpolated grey values.

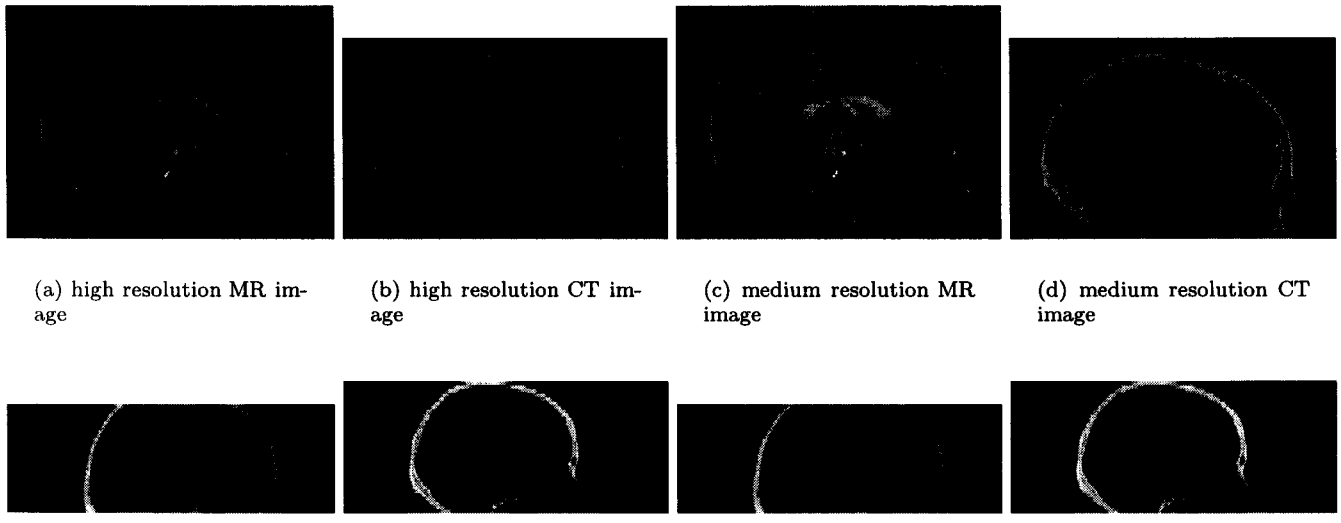
4. MATERIALS

4.1. Datasets

In our experiments, we use two pairs of high resolution datasets and seven pairs of low resolution datasets, with each pair containing a 3D CT image and a 3D MR image of the head. All images were obtained from actual patients and they are a realistic representation of clinical CT/MRI registration problems. The datasets include a variety of imaging artefacts and pathological abnormalities occurring in normal clinical practice.

For two purposes, additional high resolution data were obtained by resampling the low resolution images (using trilinear interpolation) to an isotropic grid of 1.0 *mm* spacing: for more high resolution experiments and also to investigate whether resampling can help overcome the problems with the performance of the matching method, which are caused by the relatively low information content in the slice direction. Low information content can cause local optima in the registration function in that direction and Powell's method is not guaranteed to surpass local optima in favour of the global optimum. Although the information content cannot be raised, we can attempt to smooth the registration function by creating these isotropic images.

Because a clear difference between the degree of acceleration achieved for images of high and low resolution was noted, images of an intermediate resolution were created by averaging the slices of the genuine high resolution images: the grey values of two or three slices were averaged into a single slice. An overview of the datasets can be found in Table 1. Examples are shown in Figure 1.



(a) high resolution MR image

(b) high resolution CT image

(c) medium resolution MR image

(d) medium resolution CT image

(e) low resolution MR image

(f) low resolution CT image

(g) artificial high resolution version of image (e)

(h) artificial high resolution version of image (f)

Figure 1. Examples of the different images used in the experiments. The images consist of transversal slices, but were reformatted to a sagittal view to clearly show the differences in slice thickness. Figures (a) and (b) show an MR and a CT image from one of the genuine high resolution sets. In Figures (c) and (d) the same images, averaged to a slice thickness of 3.0 *mm*, are shown. Figures (e) and (f) contain examples of low resolution images. In Figures (g) and (h) the results of resampling the previous images to an isotropic grid (i.e. the artificial high resolution images) are shown.

4.2. Equipment

Unless stated otherwise, all experiments were conducted on an SGI challenge XL, having a main memory of 1024 Mb and a 194 MHz R10000 processor.

5. REGISTRATION EXPERIMENTS

A series of multiscale registration experiments has been conducted on the datasets described above. In all cases, the MR image was matched to the CT image, because interpolation errors are likely to be smaller for the relatively homogeneous CT images. The images were not pre-registered other than having their centres and corresponding axes aligned. The software provides a means of equidistant sampling: setting a *sampling factor* to x results in only one out of every x voxels along an image axis being used in the computations. The sampling factors can be defined separately for each dimension. In the experiments, all the sampling factors used were natural numbers. In the following, the

Datasets			
Images	Modality	Image dimensions	Voxel dimensions
High resolution 1 (*)	MRI	256 x 256 x 180	0.98 x 0.98 x 1.0
	CT	256 x 256 x 100	0.94 x 0.94 x 1.55
High resolution 2 (*)	MRI	256 x 256 x 100	0.90 x 0.90 x 1.5
	CT	320 x 320 x 145	0.71 x 0.71 x 1.5
High resolution 3-9	MRI	320 x 320 x [80-104]	1.0 x 1.0 x 1.0
(from low resolution 1-7)	CT	334 x 334 x [108-132]	1.0 x 1.0 x 1.0
Medium resolution 1	MRI	256 x 256 x 90	0.98 x 0.98 x 2.0
(from high resolution 1)	CT	256 x 256 x 50	0.94 x 0.94 x 3.1
Medium resolution 2	MRI	256 x 256 x 60	0.98 x 0.98 x 3.0
(from high resolution 1)	CT	256 x 256 x 50	0.94 x 0.94 x 3.1
Medium resolution 3	MRI	256 x 256 x 50	0.90 x 0.90 x 3.0
(from high resolution 2)	CT	320 x 320 x 72	0.71 x 0.71 x 3.0
Low resolution 1-7 (*)	MRI	256 x 256 x [20-26]	1.25 x 1.25 x 4.0
	CT	512 x 512 x [27-34]	0.65 x 0.65 x 4.0

Table 1. Description of the datasets used. The image dimensions are measured in voxels and the voxel dimensions in millimetres. The numbers between brackets denote the range of the third image dimension. The genuine datasets are marked by (*).

original scale refers to scale 1 (i.e. the images sampled by factors 1). In a multiscale experiment, the term *finest scale* is used to denote the scale in that hierarchy which has been sampled with the smallest factor. The finest scale does not necessarily equal the original scale; it can also be one of the coarser scales (except for the coarsest scale).

The high resolution images were sampled with a factor of either 1, 2, 3 or 4, referred to as scales 1 to 4 respectively. The sampling was identical in each dimension. The experiments conducted consist of direct matches at every scale and of multiscale matches with every possible coarse-to-fine combination of scales (including every combination not ending in the original scale), resulting in 4 direct matches and 11 multiscale matches per dataset.

For the medium and low resolution images it was not realistic to use equal sampling in all dimensions, owing to the combined effect of larger slice thickness and relatively few slices. The sampling factors for the slice dimensions were tuned to those chosen for the in-plane dimensions in such a manner that the resampled data most closely resemble an isotropic image. The medium resolution images were sampled (in-plane) by factors 1, 2, 3 and 4 and the low resolution images by factors 1, 3 and 6. Again, the images were registered both directly at every scale and also via every possible coarse-to-fine combination of scales (including every combination not ending in the original scale). This led to 4 direct matches and 11 multiscale matches for the medium resolution datasets and 3 direct and 4 multiscale experiments for the low resolution datasets.

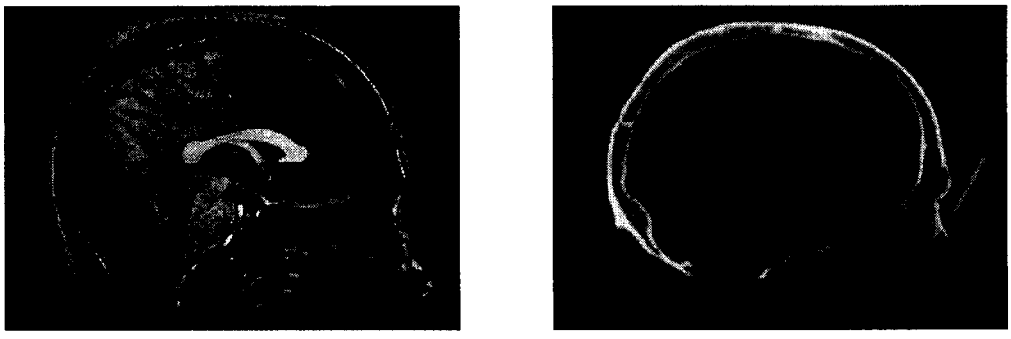
To judge the results of the direct mutual information registration of the original images, the results were inspected visually. These results then served as the standard for the other direct matches and for the multiscale experiments. The accuracy and robustness of the multiscale matches was compared to those of the direct matches. A measure of *maximum distance* between two transformations was defined as the maximum deviation between the transformations calculated on a sphere located at the image centre and having a diameter of 20 *cm*.

Visual inspection entailed the overlaying of the skull contours, obtained by intensity thresholding of the CT image, onto 2D MR images, in transversal, sagittal and coronal views. An example of a sagittal view is shown in Figure 2.

6. RESULTS AND CONCLUSIONS

6.1. High resolution images

The multiscale registration yielded accurate results for the genuine high resolution data. This was deduced from visual inspection (e.g. see Figure 2) and by comparison of the results to those for direct matching. On average, a

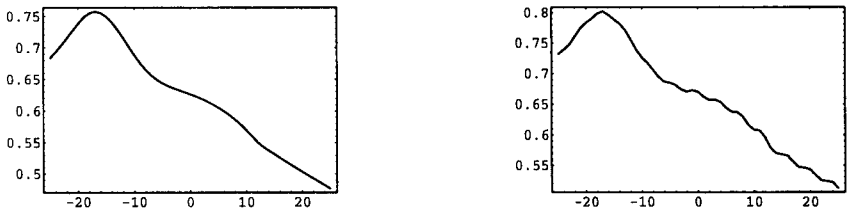


(a) high resolution MR image, with CT skull contours overlaid

(b) high resolution CT image

Figure 2. An example of the matching results for two high resolution images. The skull contours were obtained by thresholding the CT image and then superimposed on the MR image.

multiscale experiment obtained a result equivalent to that of a direct match of the finest (not necessarily the original) scale in the hierarchy. In 1 of the 22 experiments an exception occurred because incorrect intermediate solutions were involved: mismatches appeared at coarse scales. These mismatches are a result of the scaling of the images, since this will cause the registration function to loose its smoothness and the maximisation method will return a local optimum instead of the global optimum. An example of the effect of scaling on the registration function can be seen in Figure 3. When a mismatch occurred at a coarse scale, the match at the next scale - finer, but still coarse - could not always (completely) rectify the incorrect solution. Only the multiscale matches including the original scale always yielded the same solutions as the direct matches. This is an indication of how sensitive the optimisation method is to the position from which the process is started. Sometimes, when starting from the incorrect solution obtained at the previous scale, the registration did not yield the correct results, while a direct match at that scale (which is usually initiated at a different position in the search space) did succeed.



(a) Scale 1

(b) Scale 4

Figure 3. Examples of registration functions. The figures show the mutual information value (vertical axis) as a function of one of the transformation parameters (horizontal axis). The figure on the left shows this function for high resolution set 1 at the original resolution. On the right, the function is depicted for the same data, but at scale 4 (i.e. with one out of every 4 voxels sampled in each dimension). Clearly, the correct solution for this parameter is approximately -18, but at scale 4 the optimisation method returned the local maximum at -1.

For the artificial high resolution images, i.e. the resampled low resolution data, the results were slightly poorer than for the two genuine high resolution datasets. The resampled data were more prone to mismatches at the coarse scales and these misregistrations permeated the hierarchy further and more frequently.

The desired acceleration of the matching process was achieved by the multiscale approach, although not in all cases. When the finest scale in the multiscale hierarchy was the original scale, an acceleration by a factor of 2.5 to 3.5 was accomplished. However, the improvement in the computation time was decreased in those cases where mismatches occurred at coarse scales. The larger the error of the solution and the further it permeated the hierarchy, the smaller the acceleration achieved.

It is interesting to note that a larger acceleration can be achieved on architectures with less internal memory. When matching images that do not fit completely into internal memory, there is an extra cost of memory swapping. Therefore, by reducing the number of computations performed on the large, i.e. high resolution, images, an additional acceleration is obtained. For the two genuine high resolution datasets, the experiments described were also performed on an HP 9000/755 (main memory: 192 Mb, 100MHz PA-RISC 1.1 processor) and an SGI O2 (main memory: 64 Mb, 180 MHZ IP32 Processor) to compare to the results on the SGI challenge XL (main memory: 1024 Mb, 194 MHz R10000 processor). The acceleration on the HP computer was a of factor of 3 to 5.5 (versus 2.5 to 3.5 for the SGI challenge XL) and of a factor of 2.5 to 6 on the SGI O2.

6.2. Medium resolution images

Similar results as above were found for the medium resolution images. The solutions found by a multiscale match are identical to those obtained with a direct match of the finest scale in the hierarchy. Furthermore, all matches including the finest scale yielded results that were very similar to those of matching the genuine high resolution images. The average (over all experiments) of the maximum distances was 0.4 *mm*.

The registrations of the medium resolution data did not produce more mismatches than those of the high resolution data. Still, the occurrence of incorrect intermediate solutions reduced the acceleration achieved. Without (large) mismatches all multiscale matches including the finest scale were 1.5 to 2.5 times faster than the direct match of the medium resolution images at the finest scale.

6.3. Low resolution images

It was harder to draw conclusions from the experiments with the low resolution data, since the results were less consistent. Only for some datasets did multiscale matches achieve an acceleration in comparison to the direct matches. For the other datasets incorrect solutions were found at (nearly) all coarser scales. However, there were instances when the mismatches at the coarser scales were not considerable (i.e. the maximum distance, between the solution found and the direct match of the original images, was small) and still no significant acceleration was obtained. This is caused by the fact that the registration function behaves considerably poorer (i.e. it is not a smooth function, monotonically decreasing from a distinct global optimum) than is the case for high resolution images. Consequently, Brent's algorithm takes longer to converge to a solution. Nevertheless, in those cases where mismatches are avoided and the multiscale approach works, the multiscale registration will perform 1.5 to 2.0 times faster than direct matching.

Further experiments with the low resolution images showed the optimisation method to be highly sensitive to the order in which the transformation parameters are optimised. Changing this order slightly, and thus taking a different path through the search space, could mean the difference between a complete mismatch and a correct solution. Applying a different parameter order to all cases of incorrect registrations of both the high and low resolution data could solve those mismatches in every single case. Unfortunately, the optimal order would differ from one dataset to another.

Some remarks can be made about the matching results for the low resolution images compared to those for the artificial high resolution images. The transformation parameters found differed only slightly; so much so that is was not possible to determine which were more accurate by visualisation of the results. Although, judging by the number of mismatches and the success rate, the supersampled images seemed to perform better, the fact remains that multiscale matches were much slower than direct matches on three occasions and that one misregistration occurred at the original scale. The acceleration achieved for the artificial high resolution sets was larger than that for the low resolution data, but nonetheless the registration of the low resolution images was significantly faster. All in all, the resampling of anisotropic data did not seem beneficial in these experiments.

It is important to note that, for the low resolution data, even direct matching of the original images sometimes failed. For higher resolution images, mismatches only occurred at coarse scales.

6.4. Summary of results

In Tables 2 and 3 an overview is presented of the results for direct and multiscale matching, respectively. In Table 2 'Accuracy' denotes the quality of the matches at the original scale, as verified by visual inspection. The column 'Speed' gives the computation time required by the direct matches of the original scale. 'Mismatches' contains the number of misregistrations out of ('/') the number of experiments, both for the finest and for coarser scales.

For the results in Table 3 only the multiscale experiments *including the original scale* were taken into account. To indicate how well a multiscale match approaches the direct match, 'Accuracy' gives the averaged value of the maximum distances between the multiscale solutions and the transformations found with direct matches at the original scale. 'Failure rate' shows the number of times multiscale experiments performed considerably worse than the direct matches; meaning those cases when either a misregistration occurred or the computation time increased. For a multiscale match to appear in the 'Success rate' column, it had to achieve a reasonable acceleration (with a lower boundary half way between no acceleration and the maximum acceleration found) and have a maximum distance to the direct match of 1 *mm* or less. All matches not classified as either a failure or a success did not impair nor significantly improve the registration process: their registration results were good, but the acceleration achieved was negligible. 'Intermediate mismatches' contains the number of experiments *not* including the original scale (whether direct or multiscale) that returned incorrect solutions. The results of these experiments are the starting points for the experiments considered in Table 3. This column serves to put the success rate into perspective. For the genuine high resolution images, for example, five out of fourteen intermediate results were incorrect, meaning that only nine of the experiments including the original scale had a reasonable chance of achieving an acceleration. The last two columns give the averaged acceleration factors, both for all multiscale experiments including the original scale as well as for the successes only.

For the direct matches of the low resolution images, no mismatches occurred with the optimisation order of the transformation parameters used. However, when this order was changed slightly, three out of the seven matches returned incorrect solutions.
For the multiscale matches of the seven artificial high resolution images, 4 out of the 49 experiments were denounced failures. Three experiments had a much larger computation time than the direct matches and one experiment resulted in a mismatch. The maximum distance of the misregistration was excluded from the calculation of the accuracy measure, since it would have blown the measure out of proportion.

Direct matching				
			Mismatches	
Images	Accuracy	Speed (sec)	finest scale	coarser scales
high resolution 1+2	++	1000-1500	0/2	3/6
high resolution 3-9	+	4000-5000	0/7	14/21
medium resolution	+	500-600	0/3	2/9
low resolution	+	250-450	0/7	8/14

Table 2. Results for direct matching. For explanation, see text.

Multiscale matching						
		Failure	Success	Intermediate	Acceleration	
Images	Accuracy	rate	rate	mismatches	total	successes
high resolution 1+2	1.3e-2	0/14	8/14	5/14	1.9	2.6
high resolution 3-9	0.96	4/49	16/49	21/49	2.0	3.2
medium resolution	1.2e-2	0/21	8/21	3/21	1.7	2.2
low resolution	0.14	0/21	5/21	11/21	1.1	1.8

Table 3. Results for multiscale matching. For explanation, see text.

7. DISCUSSION

The overall conclusion to be drawn from the experiments described above, is that a multiscale approach to mutual information registration can significantly accelerate the matching of images. However, the extent of the acceleration achieved is dependent on the accuracy of the results found at the intermediate scales. With the optimisation method used, this accuracy in turn depends on the smoothness of the registration function. For less smooth functions, the order in which the transformation parameters are optimised and the position from which the search is started, are of a large influence to the outcome.
A higher degree of acceleration is found for the high resolution images than for lower resolution images. This can probably be attributed to the fact that a faster registration of large images at their original scale has the added advantage of requiring less memory swapping. This need for swapping also explains why a larger acceleration factor is found on computers with less internal memory.

Mismatches at coarse scales should be avoided as they have a negative effect on the improvements possible with a multiscale approach. Especially when the mismatch is not immediately corrected at the next scale, but higher up in the hierarchy, the advantages of multiscale matching are lost. All the mismatches encountered could be corrected by performing the same registration, but with a different order of optimisation for the transformation parameters. This shows the sensitivity of the optimisation method to the path it is taking through the parameter space. The optimisation method was also shown to be sensitive to the position from which the search is initiated: an incorrect starting position (as defined by an intermediate solution) sometimes led to a mismatch at a scale for which a perfectly good solution was found with a direct match. The optimisation method is suitable for a smooth registration function with one, clearly distinguishable maximum. However, the registration functions of scaled or anisotropic images do not comply with these demands. Determining a "correct" parameter order beforehand does not seem viable: it is very data-dependent. For example, for images with a considerable slice thickness, it would seem a logical idea to first optimise all parameters that are independent of the slice thickness (i.e. translations along the coronal and sagittal axes and rotation around the transversal axis) and then those that are dependent of the slice thickness. Still, this partitioning alone is not sufficient since the order of the three parameters that are dependent of the slice thickness also matters greatly. The solution for correctly matching lower quality images seems to lie in less sensitive optimisation methods, such as, for example, simulated annealing.[40]

A general advice is that a multiscale approach can safely be used where high quality images are concerned, thereby achieving a significant acceleration. With images of poorer quality, the loss in robustness is considerable and only minor accelerations can be gained. Therefore, caution is advised with the use of multiscale approaches for low resolution images.

ACKNOWLEDGEMENTS

Several of the images were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN. We furthermore thank the Laboratory for Medical Imaging Research in Leuven (especially Frederik Maes) for kindly supplying us with their software for mutual information based registration and the corresponding registration results.

REFERENCES

1. A. Evans, S. Marrett, L. Collins, and T. Peters, "Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems," in *Medical Imaging: Image Processing*, R. Schneider, S. Dwyer III, and R. Jost, eds., vol. 1092 of *Proc. SPIE*, pp. 264–274, SPIE Press, Bellingham, WA, 1989.
2. C. Pelizzari, G. Chen, D. Spelbring, R. Weichselbaum, and C. Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain," *Journal of Computer Assisted Tomography* **13**(1), pp. 20–26, 1989.
3. C. Studholme, D. Hill, and D. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures," *Medical Physics* **24**(1), pp. 25–35, 1997.

4. R. Stokking, K. Zuiderveld, H. Hulshoff Pol, P. van Rijk, and M. Viergever, "Normal fusion for three-dimensional integrated visualization of SPECT and magnetic resonance brain images," *Journal of Nuclear Medicine* **38**(4), pp. 624–629, 1997.
5. M. Viergever, J. Maintz, R. Stokking, P. van den Elsen, and K. Zuiderveld, "Matching and integrated display of brain images from multiple modalities," in *Medical Imaging: Image Processing*, M. Loew, ed., vol. 2434 of *Proc. SPIE*, pp. 2–13, SPIE Press, Bellingham, WA, 1995.
6. J. Bijhold, "Three-dimensional verification of patient placement during radiotherapy using portal images," *Medical Physics* **20**(2), pp. 347–356, 1993.
7. K. Gilhuijs and M. van Herk, "Automatic on-line inspection of patient setup in radiation therapy using digital portal images," *Medical Physics* **20**(3), pp. 667–677, 1993.
8. K. Gilhuijs, P. van de Ven, and M. van Herk, "Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data," *Medical Physics* **23**(3), pp. 389–399, 1996.
9. J. Weese, T. Buzug, C. Lorenz, and C. Fassnacht, "An approach to 2D/3D registration of a vertebra in 2D x-ray fluoroscopies with 3D CT images," in *Proceedings Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery*, J. Troccaz, E. Grimson, and R. Mösges, eds., vol. 1205 of *Lecture Notes in Computer Science*, pp. 119–128, Springer-Verlag, Berlin, 1997.
10. S. Lavallée, "Registration for computer-integrated surgery: methodology, state of the art," in *Computer Integrated Surgery*, R. Taylor, S. Lavallée, G. Burdea, and R. Mösges, eds., ch. 5, pp. 77–97, MIT Press, Cambridge, MA, 1996.
11. J. Maintz, "A survey of medical image registration," *Medical Image Analysis* **2**(1), 1998.
12. P. van den Elsen, E. Pol, and M. Viergever, "Medical image matching– a review with classification," *IEEE Engineering in Medicine and Biology* **12**(1), pp. 26–39, 1993.
13. C. Maurer, Jr., J. Fitzpatrick, M. Wang, R. Galloway, Jr., R. Maciunas, and G. Allen, "Registration of head volume images using implantable fiducial markers," *IEEE Transactions on Medical Imaging* **16**(4), pp. 447–462, 1997.
14. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, Jr., R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, J. Maintz, M. Viergever, G. Malandain, X. Pennec, M. Noz, G. Maguire, Jr., M. Pollack, C. Pelizzari, R. Robb, D. Hanson, and R. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *Journal of Computer Assisted Tomography* **21**(4), pp. 554–566, 1997.
15. A. Evans, D. Collins, P. Neelin, and T. Marrett, "Correlative analysis of three-dimensional brain images" in *Computer Integrated Surgery*, R. Taylor, S. Lavallée, G. Burdea, and R. Mösges, eds., ch. 6, pp. 99–114, MIT Press, Cambridge, MA, 1996.
16. A. Liu, E. Bullitt, and S. Pizer, "Surgical instrument guidance using synthesized anatomical structures," in *Proceedings Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery*, J. Troccaz, E. Grimson, and R. Mösges, eds., vol. 1205 of *Lecture notes in computer science*, pp. 99–108, Springer-Verlag, Berlin, 1997.
17. J. Feldmar, G. Malandain, N. Ayache, S. Fernández-Vidal, E. Maurincomme, and Y. Trousset, "Matching 3D MR angiography data and 2D x-ray angiograms," in *Proceedings Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery*, J. Troccaz, E. Grimson, and R. Mösges, eds., vol. 1205 of *Lecture Notes in Computer Science*, pp. 129–138, Springer-Verlag, Berlin, 1997.
18. A. Collignon, T. Géraud, D. Vandermeulen, P. Suetens, and G. Marchal, "New high-performance 3D registration algorithms for 3D medical images," in *Medical Imaging: Image Processing*, M. Loew, ed., vol. 1898 of *Proc. SPIE*, pp. 779–788, SPIE Press, Bellingham, WA, 1993.
19. C. Maurer, Jr., G. Aboutanos, B. Dawant, R. Margolin, R. Maciunas, and J. Fitzpatrick, "Registration of CT and MR brain images using a combination of points and surfaces," in *Medical Imaging: Image Processing*, M. Loew, ed., vol. 2434 of *Proc. SPIE*, pp. 109–123, SPIE Press, Bellingham, WA, 1995.
20. W. Grimson, G. Ettinger, S. White, T. Lozano-Pérez, W. Wells III, and R. Kikinis, "An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization," *IEEE Transactions on Medical Imaging* **15**(2), pp. 129–140, 1996.

21. D. Hill, D. Hawkes, N. Harrison, and C. Ruff, "A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics," in *Information Processing in Medical Imaging*, H. Barrett and A. Gmitro, eds., vol. 687 of *Lecture Notes in Computer Science*, pp. 182–196, Springer-Verlag, Berlin, 1993.
22. F. Betting and J. Feldmar, "3D-2D projective registration of anatomical surfaces with their projections," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., pp. 275–286, Kluwer Academic Publishers, Dordrecht, 1995.
23. J. Feldmar, N. Ayache, and F. Betting, "3D-2D projective registration of free-form curves and surfaces," *Computer Vision and Image Understanding* **65**(3), pp. 403–424, 1997.
24. S. Lavallée, R. Szeliski, and L. Brunie, "Anatomy-based registration of three-dimensional medical images, range images, x-ray projections, and three-dimensional models using octree-splines," in *Computer Integrated Surgery*, R. Taylor, S. Lavallée, G. Burdea, and R. Mösges, eds., ch. 7, pp. 115–143, MIT Press, Cambridge, MA, 1996.
25. J. Maintz, P. van den Elsen, and M. Viergever, "Evaluation of ridge seeking operators for multimodality medical image matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(4), pp. 353–365, 1996.
26. J. Maintz, P. van den Elsen, and M. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Medical Image Analysis* **1**(2), pp. 151–161, 1996.
27. P. van den Elsen, J. Maintz, E. Pol, and M. Viergever, "Automatic registration of CT and MR brain images using correlation of geometrical features," *IEEE Transactions on Medical Imaging* **14**(2), pp. 384–396, 1995.
28. P. van den Elsen, E. Pol, T. Sumanaweera, P. Hemler, S. Napel, and J. Adler, "Grey value correlation techniques used for automatic matching of CT and MR brain and spine images," in *Visualization in Biomedical Computing*, R. Robb, ed., vol. 2359 of *Proc. SPIE*, pp. 227–237, SPIE Press, Bellingham, WA, 1994.
29. L. Lemieux, R. Jagoe, D. Fish, N. Kitchen, and D. Thomas, "A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs," *Medical Physics* **21**(11), pp. 1749–1760, 1994.
30. A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3D multi-modality medical image registration using feature space clustering," in *Computer Vision, Virtual Reality, and Robotics in Medicine*, N. Ayache, ed., vol. 905 of *Lecture Notes in Computer Science*, pp. 195–204, Springer-Verlag, Berlin, 1995.
31. R. Woods, J. Mazziotta, and S. Cherry, "MRI-PET registration with automated algorithm," *Journal of Computer Assisted Tomography* **17**(4), pp. 536–546, 1993.
32. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., pp. 263–274, Kluwer Academic Publishers, Dordrecht, 1995.
33. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging* **16**(2), pp. 187–198, 1997.
34. C. Meyer, J. Boes, B. Kim, P. Bland, K. Zasadny, P. Kison, K. Koral, K. Frey, and R. Wahl, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Medical Image Analysis* **1**(3), pp. 195–206, 1997.
35. P. Viola and W. Wells III, "Alignment by maximization of mutual information," in *International Conference on Computer Vision*, E. Grimson, S. Shafer, A. Blake, and K. Sugihara, eds., pp. 16–23, IEEE Computer Society Press, Los Alamitos, CA, 1995.
36. W. Wells III, P. Viola, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," in *Medical Robotics and Computer Assisted Surgery*, pp. 55–62, Wiley, New York, 1995.
37. D. Hill, C. Maurer, Jr., M. Wang, R. Maciunas, J. Barwise, and J. Fitzpatrick, "Estimation of intraoperative brain surface movement," in *Proceedings Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery*, J. Troccaz, E. Grimson, and R. Mösges, eds., vol. 1205 of *Lecture Notes in Computer Science*, pp. 449–458, Springer-Verlag, Berlin, 1997.
38. L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever, "Scale and the differential structure of images," *Image and Vision Computing* **10**(6), pp. 376–388, 1992.
39. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 1992.
40. P. Laarhoven and E. Aarts, *Simulated annealing: theory and applications*, D. Reidel Publishing Company, Dordrecht, 1987.