Chapter 3

CT similarity before and after registration

As mentioned in the previous chapter, this master's thesis aims at using the transformations obtained between CT images for correcting PET images. In this chapter, CT volumes obtained from different respiratory phases will be co-registered, and the similarity measures prior and after registration analyzed.

Both the AIR [43] and ITK [44] based programs described in the User Guide in the appendix are utilized. ITK is a powerful and versatile package for testing registration algorithms based on different deformation models, cost functions, optimization algorithms and interpolation methods. That is the reason why, apart from AIR (that includes some polynomial models that ITK does not), no other packages were tested. The results are analyzed with Matlab.

3.1 Materials

This chapter's work is based on the set of images provided by [36]. The images were acquired by a 12-slice CT scanner with a time resolution of approximately 0.75 seconds until 15 different samples were taken (with their corresponding spirometer measurements). This time is enough to cover approximately 3 respiratory cycles. The couch was then shifted as much as the width of 12 slices and the scanning was repeated. The slice thickness was 1.5 mm, leading to a total couch shift of 18 mm each time. The two-dimensional pixel size within each slice was 0.78 mm for one of the cases and 0.94 mm for the rest of them.

As discussed in [36], it is possible to stack scans at similar breathing stages in order to compose a whole thorax image. One looks for 12-slice sets with close spirometer measurements and same respiratory phase (inhale / exhale), and stack the corresponding CT images one on top of each other to obtain a larger volume. Images from 5 patients were available. For the first two, the whole thorax was scanned (228 slices, 19 couch positions). For the other three, only part of the lungs are available (6, 3 and 3 12-slice sets for patients 3,4 and 5 respectively).

The first problem arising was at which volumes the whole thorax images shall be "reconstructed". An initial possibility would be to choose the volumes one is interested in and pick up the closest data set available for every bed position. This approach's main advantage is the possibility of choosing an arbitrary tidal volume. The problem is the possibility of falling into a bad-conditioned air volume value: the data set may lack values close to the chosen spirometer value.

Another option, which is actually the one considered in this thesis, prioritizes the image quality, trying to stack data sets acquired at air volumes as close as possible. Since newer CT machines can acquire several slices at the same time, and in a smaller time than the one that acquired the available data set, this second option seems more reasonable, as it can be assumed that the algorithms described in this thesis would be applied on high-quality pictures. The error e made when choosing a certain study volume V was arbitrarily defined as:

$$e(V) = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(V - C_i(V))^2}$$

where n is the number of couch position and $C_i(V)$ represents the closest available volume to V in the same breathing phase and in couch position i.

The reason for using this error measure is that the squared sum penalties the most those points that are quite far from available volumes, making the "stacked" image look discontinuous (figure 3.1).

Figure 3.2 illustrates the error as a function of the volume for a certain respiratory phase. The volume values located at the curve's valleys may be chosen in order to minimize the artifacts in the image due to the stacking process.

In order to make simple tests with the algorithms to ensure that they work properly before applying them to real data, some simple geometrical figures were also generated: a cube, a sphere, an ellipsoid and a cylinder (figure 3.3).

Figure 3.1: Artifact due to stacking images from different tidal volumes (sagittal view)

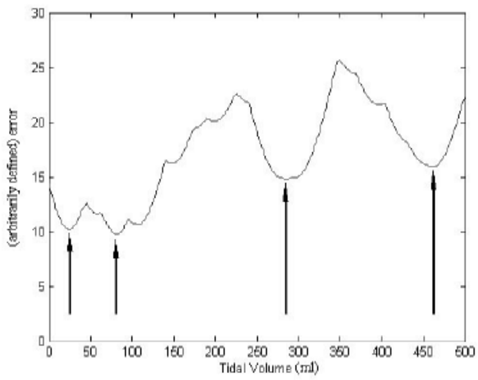

Figure 3.2: Error for every volume and chosen "reconstruction" points (patient 2, inhaling)

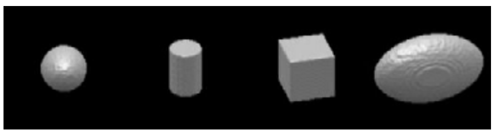

Figure 3.3: Some simple images to test algorithms

3.2 Similarities between CTs in different phases

In this first experiment, the differences between CT images in different breathing phases will be analyzed, both quantitatively and qualitatively, and the results compared. Qualitative measures are important because they involve human perception (which plays an important role when a physician makes a diagnose). Quantitative measures are especially useful when it comes to extensively compare image pairs and evaluation by humans results unpractical. It will be assumed that two images with a high similarity measure are also qualitatively similar. This assumption will be further investigated by the means of a survey.

3.2.1 Methods

For the quantitative measures, the squared sum of voxel value differences (SSD), the normalized cross-correlation (NCC), the difference image's entropy and the images' mutual information were used. Since the CT voxel intensities are integers between 0 and around 3500, 12 bits (0-4191) are enough to represent each voxel. However, 16 bits are generally used instead in order to comply to a "standard" computer format (usually *unsignedshort*).

In order to reduce the memory requirements when computing entropy measurements, the [0,4095] range was mapped to [0,255] by simply bit-wise shifting the numbers four positions to the right. This involves losing some precision, but is necessary in order not to exceed the used computer's memory capabilities. This study can afford this loss of information because the interest is centered in relative and not absolute results. After this bit shift, a random image with uniformly distributed intensities between 0 and 4095 would have for according to this criteria an entropy of 8 bits.

When two images are subtracted, the range is extended to [-4095, 4095]. When calculating the entropy of such an image, [-4095, 4095] was mapped to [0,255]. This time, 4095 was added to the value before shifting it 5 times to the right. To sum it all up, 8 bits is thus the maximum relative information that images may contain, according to the defined criteria.

In order to confirm the assumption of qualitative and quantitative measures producing similar results, a survey where 54 people took part was performed. The participants were asked if they were professionals used to working with CT images (12 of them did), and to rate the similarity of 13 pairs of images from 0 to 10. The above mentioned numerical comparison methods were applied as well, and the results compared to the subjective ones. Mean opinions from professionals, non-professionals, and all of them together are considered.

Image pair	Margin
1	0.7422
2	0.9508
3	0.9040
4	0.8292
5	0.8535
6	0.5400
7	0.8015
8	0.8889
9	0.6603
10	1.0090
11	0.6609
12	0.8803
13	0.7903

Table 3.1: Error margin for a confidence of 95% in every image pair in the survey. The grade range is 0-10.

Assuming a normal distribution, the error margin for a confidence of 95% is shown in table 3.1, reaching one point in the worst case. The following conclusions were extracted:

1. The grades from physicians and researchers and those from the others were quite similar. Professionals gave higher grades in general (in 12 out of 13 cases, with differences between a +0.2 and a +1.7 points, 4.33 against 4.48 in the exceptional case). The important fact is that the similarity ranking was almost the same. It can be seen in figure 3.4 that the first three bins' lengths are almost equal for every pair, which means that both groups ordered the image pairs in almost the same way.

2. Mutual information is not a reliable measurement. It is very useful when comparing images from different modalities (as discussed in the introduction), but produces inconsistent results in these experiments. It gives for example a bad score to pair 13 (check figure 3.5) when it is composed by the same slice with the same tidal volume in different time instants and earns the highest score according to the other 6 criteria.

3. Apart from the mutual information criteria, it can be appreciated that the numerical methods order the image pairs approximately in the same way as the subjective impressions. Even if it is complicated to compare quantitative and qualitative measures, this gives an indication of that

	NCC	SSD	Difference Image Entropy
Same slice,same respiratory phase (but in different time instants)	0.997	4000	1.15
Same slice, different respiratory phases (full exhale and full inhale)	0.92	40000	1.6

Table 3.2: Typical numerical values when using numerical criteria to compare image pairs in the CT data set. These values will serve as a reference in future comparisons.

they usually order image pairs by similarity in approximately the same way.

4. The numerical criteria and the visual ones differ the most in pairs 10, 11, and 12. In pair 10 (figure 3.6-a), stacking artifacts penalty the visual impression much more that the numerical measures. In pair 11 (3.6-b), the visual impression is not that bad but there is a shift that spoils the numerical measures. This can be appreciated in the difference image (3.6-c). In pair 12 (3.6-d), the vessel distribution in the lungs makes the visual impression bad, while such small details do not to affect the numerical results significantly.

This all can be summed up by saying that the numerical criteria (excepting mutual information) usually give a good idea of how similar two images are and can almost always replace an human opinion (which is usually similar from doctors and non-doctors), excepting in the cases described above.

Some representative numerical values are worth to be mentioned, in order to have a reference for comparing later results (only valid within this data set, other images may have different typical values). The similarity measure for two images from the same slice in the same respiratory phase, but in different time instants (figure 3.5), and another two from the same slice in different volumes, are shown in table 3.2. The cyphers from the first case could be a ambitious but reachable goal when registering images.

3.2.2 Results

The similarity between individual, corresponding transverse slices from CT volumes of the same patient at different breathing phases, depending on the coordinate along the couch movement, is represented in figure 3.7. CTs for three different air volumes (located around 0%, 50% and 100% of the total capacity) are compared for each of first two patients (as a complete thorax

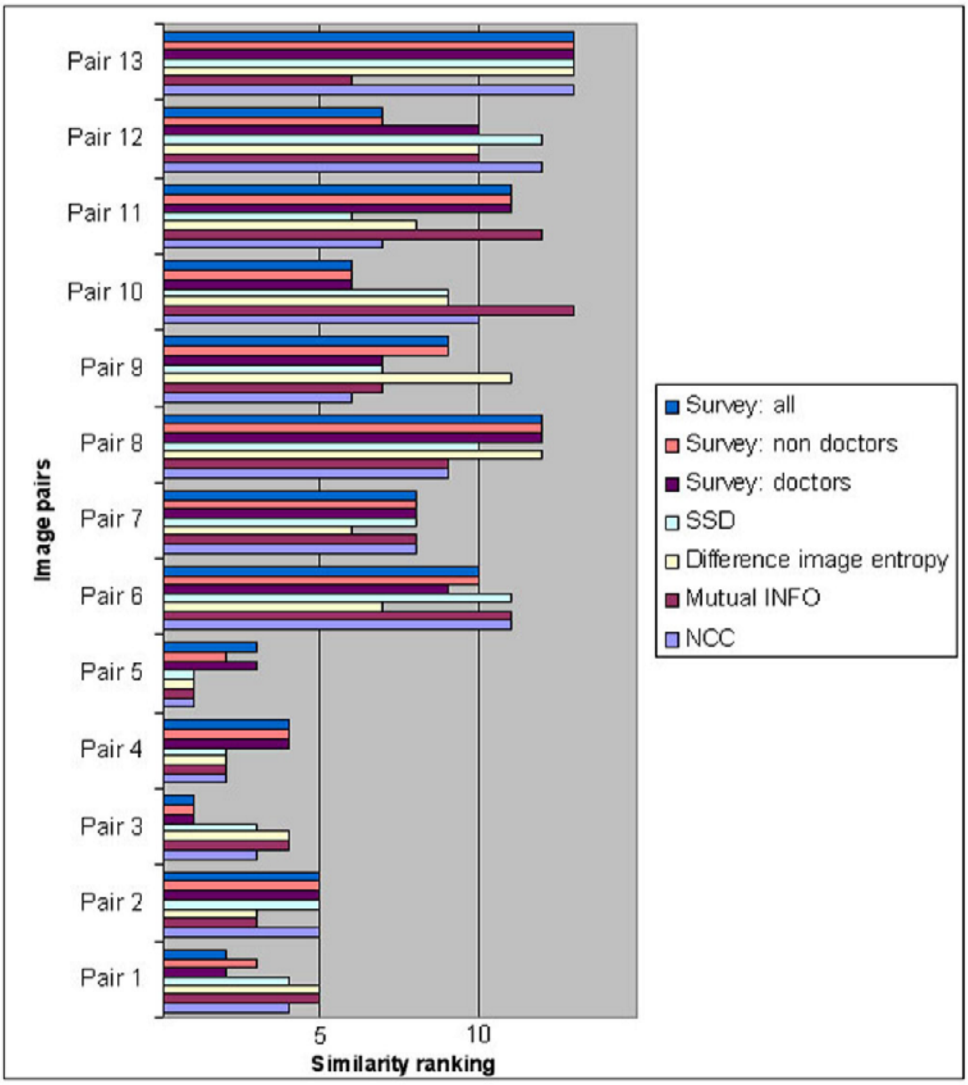

Figure 3.4: Similarity ranking for every image pair in the survey according to the different criteria. The scores are between 1 (least similar) to 13 (most similar).

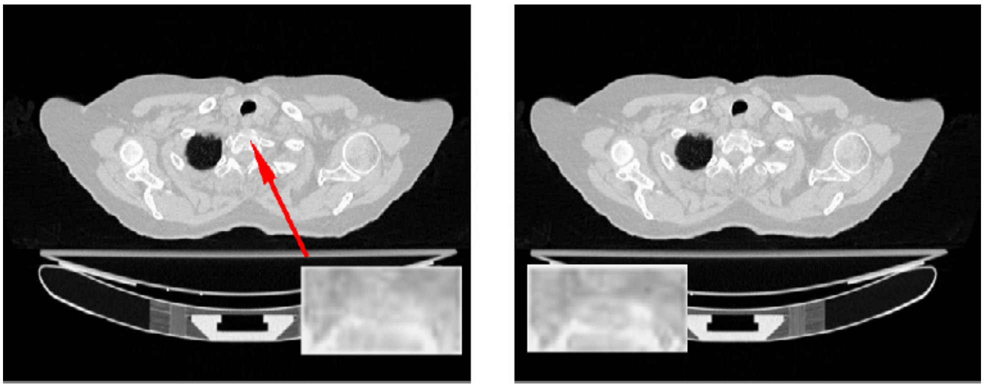

Figure 3.5: Two images corresponding to the same slice, approximately the same tidal volume, but different time instants. It is almost impossible for the human eye to find any differences without zooming in quite much.

scan is available for them). Sagittal views are shown in the same scale so that it is easy to notice the thorax areas with larger displacements.

The figure reveals how the three measures behave in a similar manner, as expected. It can also be appreciated that there are rapid changes every 12 slices due to stacking errors.

3.2.3 Conclusions

As all of the curves have maximums or minimums (maximum if they are dissimilarity measures, minimum if they are similarity measures) around the same regions of the z-axis (defined as in figure 3.8), it is possible to identify the areas where the patient moves the most.

The patients do not move that much around the waist and hip (the very left of the images in figure 3.7), while the amplitude of the movements grows as higher slices until the most complicated area is reached: around the lungs' lower lobes. Afterwards, the effects from the patient's movements begin to decline again.

A consequence of this observed behavior is that it would be possible to avoid taking so many CTs at many different respiratory phases around the regions where the patients does not move much, saving radiation to the patient. The respiratory movements could be described with less temporal frames. The similarity measures around the hip are for example so high that it would not be worth to compensate for the breathing movements, as its effects are negligible. The analysis of this possibility is outside the scope of this thesis, though.

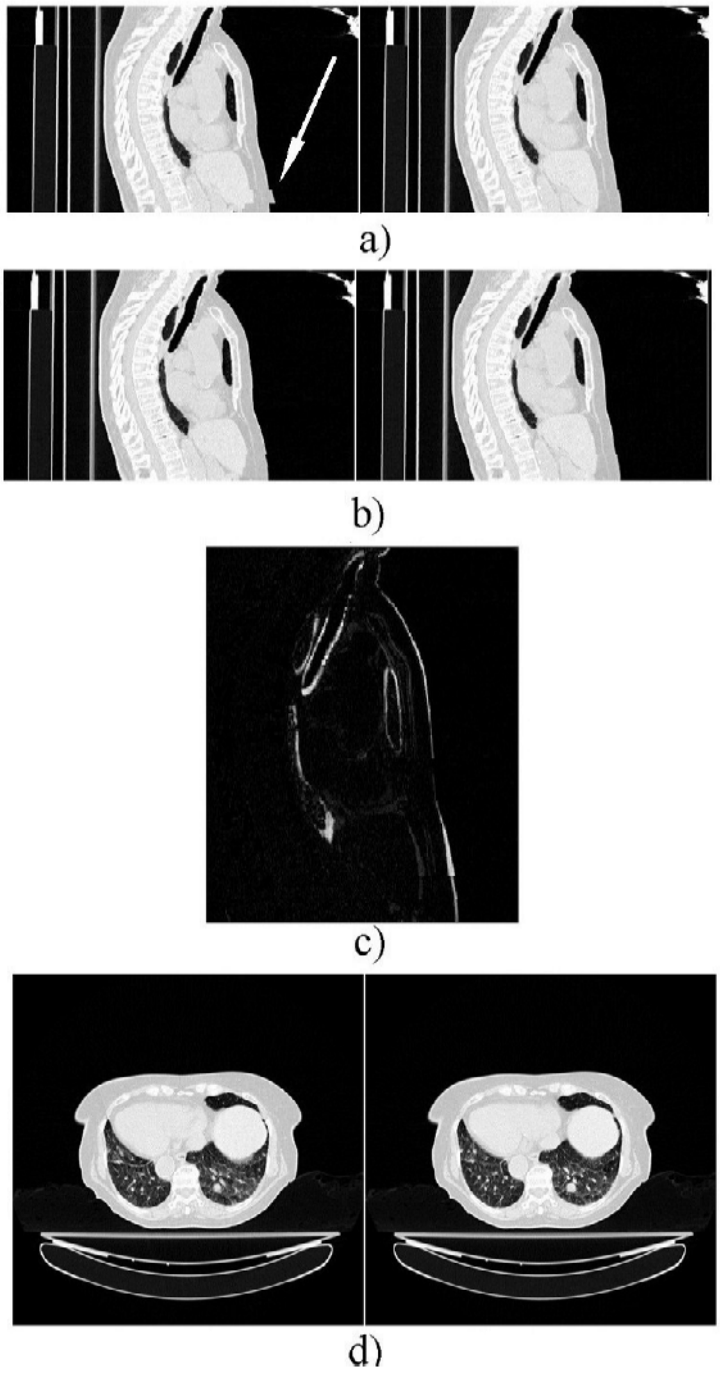

Figure 3.6: Image pairs where the numerical measures and human opinions differ the most. a) Artifacts due to stacking errors, b) one image is a shifted version of the other, c) difference image to clarify the shift and d) the vessels are quite different in the two images.

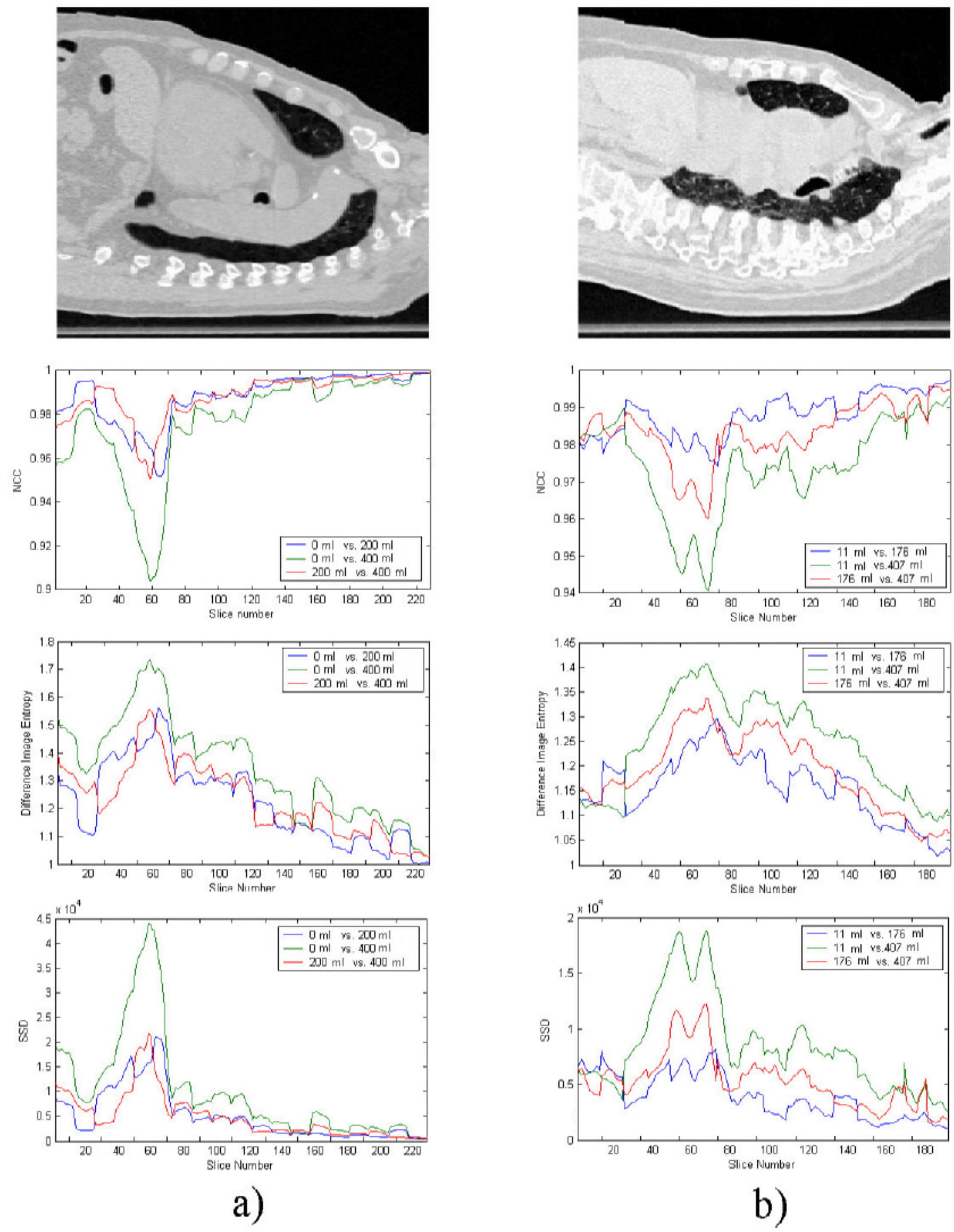

Figure 3.7: NCC, difference image entropy and SSD similarity measures depending on the transverse slice number for patients 1 (a) and 2 (b). The blue line compares the minimum volume position with the intermediate volume one, the green one the minimum with the maximum, and the red one the intermediate with the maximum (always in a slice by slice basis).

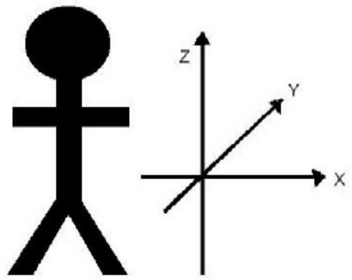

Figure 3.8: Spatial coordinates reference. We can imagine that the patient is looking at us.

3.3 Registering CTs from different phases

3.3.1 Methods

The two complete thorax volumes available were used in the experiments. In order to decrease the registration time, the stacked images were downsampled along the x and y axis (check figure 3.8). An spatial low-pass filter was applied first in order to avoid aliasing. No downsampling was performed in the z direction, as the resolution in such axis was approximately one half of what it was in the other two. The obtained data was then a volume four times smaller and with a similar resolution in the three spatial axis.

Three different tidal volumes were chosen for each of these two patients: the largest one, the smallest one, and another one with a value around a 50% of the maximum. The larger volume was always registered towards the smaller one in the experiments. This leads to three different registrations for each patient (100%-50%, 100%-0%, 50%-0%), six in total.

Nine different registration techniques were compared. Ordered in increasing complexity:

1. Linear registration with a rigid body 6 parameter model. The program was configured to run a maximum of 25 iterations per pyramidal level, stopping if no progress in the cost function (least squares) was made in 8 iterations. Voxels with a value of 10 or less were ignored when computing the cost function and its derivatives in order not to involve the dark area around the patient's body in the process.

2. Linear registration with a global rescale 7 parameter model, with the same constraints as above.

3. Linear registration with an affine 12 parameter model. Same configuration.

4. Polynomial registration, second order 30 parameter model. Same configuration.

5. Polynomial registration, third order 60 parameter model. Same configuration.

6. Polynomial registration, forth order 105 parameter model. Same configuration.

7. Polynomial registration, fifth order 168 parameter model. Same configuration.

8. Demons algorithm. Pyramidal processing with 14, 10 and 6 iterations used. A gaussian filter with a standard deviation of 1 mm (as proposed in [25]) was applied after each iteration to make the vectors maintain a certain coherence between them.

9. Improved demons algorithm. Pyramidal processing was used with 7, 5 and 3 iterations. Gaussian filter applied as above.

Linear interpolation was used for all techniques, as it is less computationally expensive than using BSplines, windowed sincs or other more complex methods, while the results are still satisfactory.

The choice of the threshold level for computing the cost function in the linear and polynomial methods is based on the images' histograms (figure 3.9); a threshold of 10 separates the darkest components due to the patient's surroundings from the rest of the image, without cropping important image components.

The number of iterations for the AIR programs is based on visual inspection of the cost function evolution. 10-15 iterations are sufficient for the polynomial algorithms (figure 3.10) to reach a stable value. 25 iterations are used in order to leave a safety margin. The same could be said about the demons algorithm, where 14,10, 6 (original) and 7,5,3 (improved) iterations were performed, respectively (figure 3.11).

Two different methods are used for evaluating the registration process. The first one is to calculate the similarity measures of the fixed and registered images. One problem with this measures is the difficulty for the reader to interpret the resulting cyphers, even if they are expressed relatively to a known case.

A second way of evaluating a registration technique consists of locating equivalent points (landmarks) in the fixed, moving and registered images. The landmark concept is illustrated in figure 3.12.

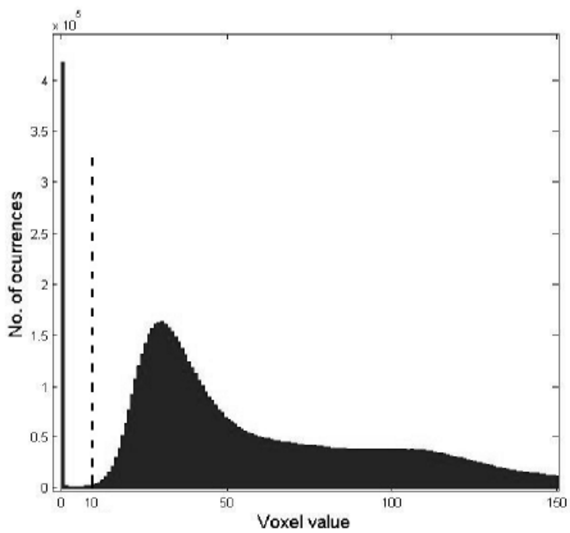

Figure 3.9: Beginning of a typical histogram for a CT volume.

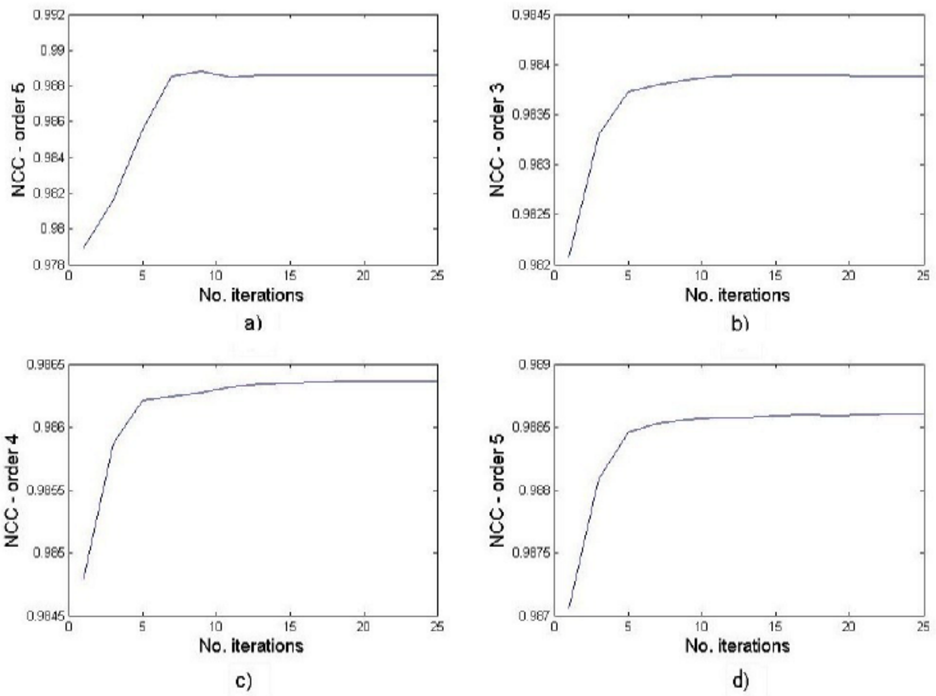

Figure 3.10: Normalized crossed correlation for the fixed - registered image pair depending on the number of iterations. a) Second order polynomials b) Third order c) Forth order d) Fifth order.

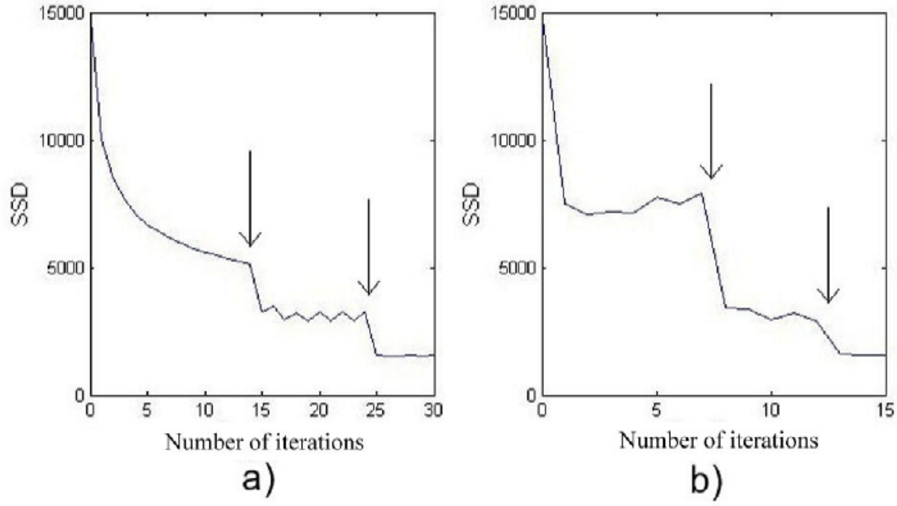

Figure 3.11: SSD metric evolution with the iteration number for both the original and improved demons algorithm in the first patient's full inhale to full exhale registration problem. The marked sudden jumps are due to resolution jumps in pyramidal processing.

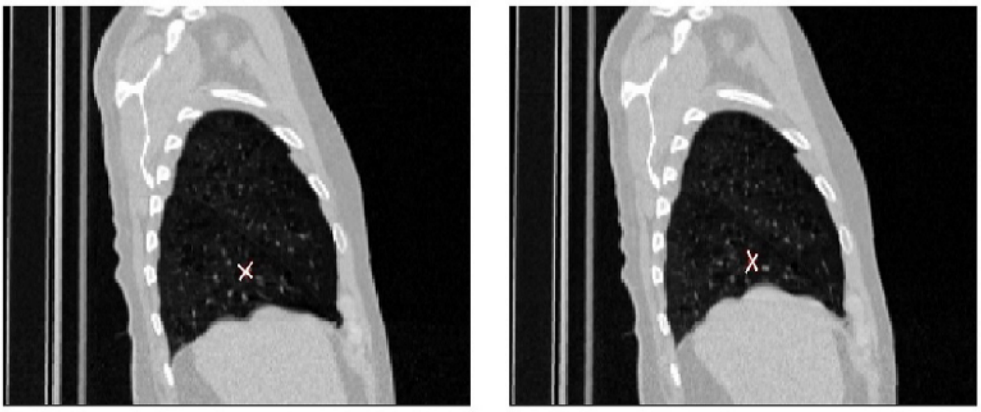

Figure 3.12: Landmark placed in equivalent positions on sagittal views from different images.

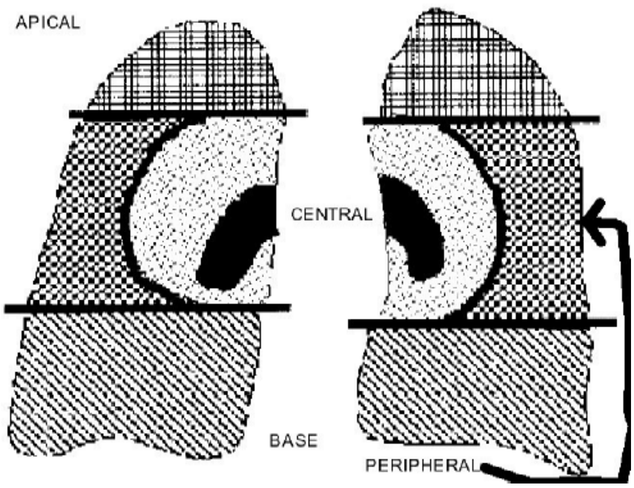

Figure 3.13: Lung regions. Apical corresponds to the top quarter of the lung, base to the lower quarter, central is defined as the region lung hila up to half of the distance between hila and the lateral border, and peripheral is the rest.

The landmarks' positions can be compared to inspect if the registration process has been successful. The distance decrement between landmarks from the fixed-moving to the fixed-registered image pair is a measure of how successful the registration has been and, being a distance, is a result very easy for the reader to interpret.

25 landmarks were identified with the help of a physician in the full exhale, full inhale and registered images (registered from full inhale phase to to the full exhale one). The landmarks were placed in locations where the image had prominent details that could be easily distinguished (for instance, thick vessels). Different gray-scale to pseudo-color mappings were used when visualizing the images in order to facilitate this task.

The lungs were divided into four regions as shown in figure 3.13. 9 of the landmarks were located in the base region (the most interesting one, as the movement amplitude is bigger), 6 in the peripheral one, 3 in the apical one and 7 in the central one. The registration was performed with the same methods as described in the previous sections. The distance between the landmarks prior to and after registration was measured.

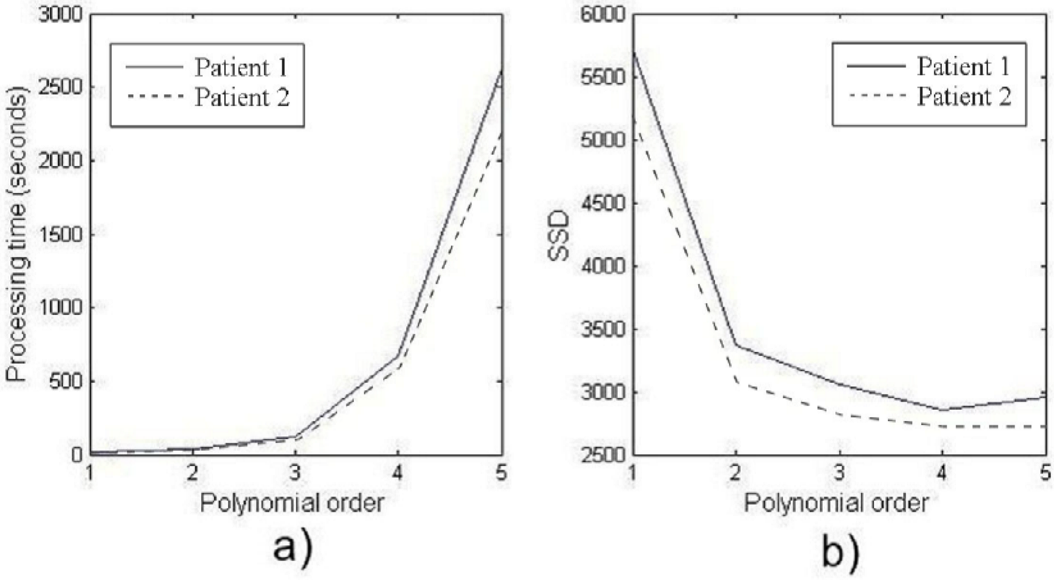

Figure 3.14: a) Computing time and b) SSD error depending on the order for registration with polynomials for two different patients.

3.3.2 Results

Similarity measures

In figure 3.14, the evolution of the errors and of the computing times with the polynomial order in two certain problems (patients 1 and 2, registration from full inhale to full exhale) is shown. The figure shows that the error decreases rapidly while the registration time increases slowly at the beginning. However, when the polynomial order becomes larger, the error hardly decreases while the processing time increases exponentially, apparently. Hence, it can be stated that using higher order polynomials does not contribute to a much higher resulting similarity, in general.

The average SSD, NCC and difference image entropy results for the registration processes are represented in figure 3.15. As it usually happens that the first and last slices of the registered images require data from outside the original volume (and thus unavailable), it was decided that the first four and last four slices would not be taken into consideration when calculating the similarity measures.

The figure reveals that linear algorithms do not improve and even worsen the results as compared to the non-registered case. Considering that "doing nothing" is a sub-case of these registration techniques, a perfect optimization algorithm could have stayed in the no-transform point and given a better

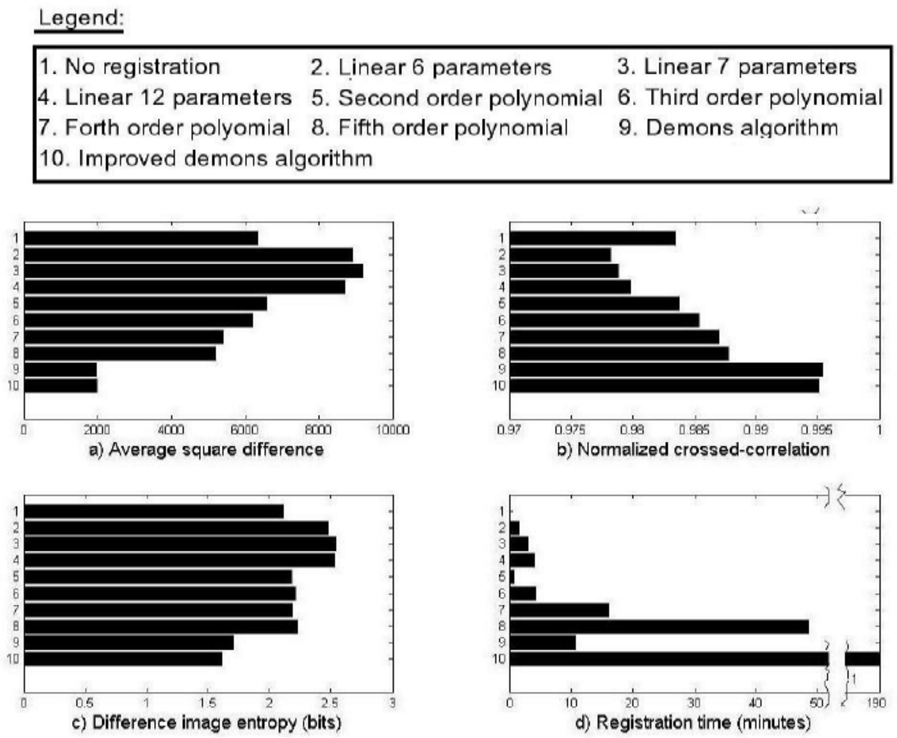

Figure 3.15: Average SSD, NCC and difference image entropy similarity measures for the registered volumes. The processing time is also shown.

result. The errors due to interpolation and the optimizing algorithm's imperfections can be blamed.

One can also notice that even a second order polynomial (12 parameters) does not contribute to any better result, taking even less time than the linear methods. The third order polynomials (30 coefficients) take as much time as the 12 parameter linear model, but improves the results substantially. The tendency continues with the forth and fifth order polynomials: the similarity measures improve (very clear from the SSD and NCC measures; the difference image entropy does not decrease that much...), but at the cost of computing time.

A marked behavior change arises with the demons algorithms. The error decreases dramatically, and even if the registration time increases considerably for the improved algorithm, it is only around 10 minutes for the original one. Both the original and the improved symmetrical version give similar results, but the improved one requires a 50% less iterations, as can be observed in figure 3.11. It must be pointed out that each iteration takes much longer than twice the time as the original algorithm in the ITK implementation. The original article [25] claims to achieve a 40% speed improvement, though.

As it has already been explained, the difference between Thirion's original algorithm and the improved one is that the latter considers not only the fixed image's gradient but also the moving image's one:

$$\vec{u} = \frac{(m-s)\vec{\nabla}s}{|\vec{\nabla}s|^2 + (m-s)^2}$$

for the original algorithm, while

$$\vec{u} = (m-s)\left(\frac{\vec{\nabla}s}{|\vec{\nabla}s|^2 + (m-s)^2} + \frac{\vec{\nabla}m}{|\vec{\nabla}m|^2 + (m-s)^2}\right)$$

for the improved one.

Even if the improved algorithm requires two gradients instead of one, the computational cost for each iteration should be less than the double as compared to the original algorithm, as the terms $(m-s)$ and $(m-s)^2$ are already calculated. Hence, it can be stated that the ITK implementation is not very time efficient.

In table 3.3, the results for the similarity measures are compared to the typical values for two equivalent slices from the same patient taken in approximately the same tidal volume but in different time instants (from last section; they are the values that were set as a goal). The table shows that these values are even improved by the demons algorithm, while forth and fifth order polynomials get very close (a 30%-35% increment in the SSD measure).

	NCC	SSD	Norm. NCC	Norm. SSD
Same slice,same respiratory phase (but in different time instants), typical	0.997	4000	1.0	1.0
Improved demons algorithms	0.995	1997	0.998	0.50
Demons algorithms	0.995	1967	0.998	0.49
Fifth order polynomial	0.988	5195	0.991	1.30
Forth order polynomial	0.987	5406	0.990	1.35
Third order polynomial	0.985	6195	0.988	1.55
Second order polynomial	0.984	6597	0.987	1.65

Table 3.3: Comparison between the similarities between two registered images using different registration techniques. Values for two slices from the same patient in the same respiratory phase but in different time instants (an ambitious goal to aspire to) are used as normalization factors.

Landmarks

The distances between landmarks before and after registering are shown in table 3.4. The landmarks in the apical region hardly move, and no improvement is observed after the registration (the demons algorithm is the only one that actually does not increase the error). For the base region, the movements are in general much more prominent: approximately 1 cm in average. While fifth order polynomials provide good results, the improved demons algorithm's performance is outstanding, reducing the error from 9 mm to only 1 mm. The results show how the other algorithms are clearly outperformed.

In the central and peripheral regions, the movements have an intermediate amplitude. The demons algorithms are still superior, followed by the fifth order polynomials.

The averaged values show that the improved demons algorithm offers an improvement as twice as large as any of the others, being able to reduce the landmark distance from 4.7 mm to 0.6 mm (that is, in a 87%). The polynomials give poorer results, as expected. The higher order ones can still reduce the error in an about a 50%.

Just to finish this section with some graphical results, a transverse slice at the lung's lower lobes level from patient one, registered from full inhale to full exhale, is shown in figure 3.16.

3.3.3 Conclusions

One conclusion is that similarity and landmark tests produce consistent results, ordering the performance of the different registration algorithms in the

	Apical (mm)	Base (mm)
Improved demons	0.78 - 0.26 = 0.52	8.94 - 0.81 = 8.13
Original demons	0.78 - 0.47 = 0.31	8.94 - 2.21 = 6.73
Fifth order polynomials	0.78 - 1.13 = -0.35	8.94 - 3.65 = 5.29
Forth order polynomial	0.78 - 1.56 = -0.78	8.94 - 4.38 = 4.56
Third order polynomial	0.78 - 1.56 = -0.78	8.94 - 5.30 = 3.65
Second order polynomial	0.78 - 1.00 = -0.22	8.94 - 6.78 = 2.17

	Central (mm)	Peripheral (mm)
Improved demons	2.22 - 0.68 = 1.54	3.17 - 0.39 = 2.79
Original demons	2.22 - 0.79 = 1.43	3.18 - 1.41 = 1.77
Fifth order polynomials	2.22 - 1.11 = 1.12	3.18 - 2.53 = 0.64
Forth order polynomial	2.22 - 1.81 = 0.41	3.18 - 2.72 = 0.45
Third order polynomial	2.22 - 1.76 = 0.47	3.18 - 2.18 = 0.99
Second order polynomial	2.22 - 1.87 = 0.35	3.18 - 3.19 = -0.01

	Average (mm)
Improved demons	4.70 - 0.61 = 4.09
Original demons	4.70 - 1.86 = 2.84
Fifth order polynomials	4.70 - 2.37 = 2.33
Forth order polynomial	4.70 - 2.93 = 1.77
Third order polynomial	4.70 - 3.11 = 1.59
Second order polynomial	4.70 - 3.85 = 0.85

Table 3.4: Landmark deviation improvement from the moving image's position to the registered image's one, in mm: before - after = improvement.

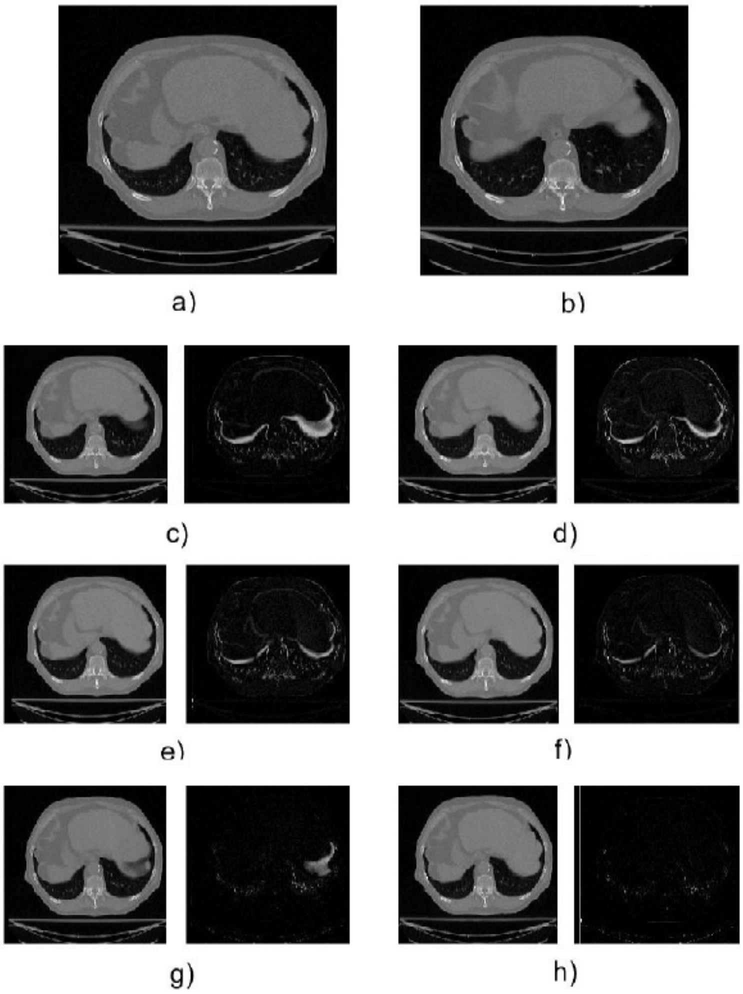

Figure 3.16: Patient one's transverse slice number 53, a) fixed image, b) moving image, and registered images along with their errors (difference image absolute vale) for: c) Second order polynomial d) third order polynomial e) forth order polynomial f) fifth order polynomial g) original demons algorithm and h) improved demons algorithm.

same way.

While linear algorithms do not improve and even worsen the similarity between the images, polynomial-based methods increase the similarity measure between image pairs. Polynomials of order five are very computationally expensive, and do not outperform polynomials of order four by a large margin.

Polynomials lack complexity to handle the complex deformations inside the patients thorax. The demons algorithms require in general much more computing time, but provide much better results: the similarity measures increase to the levels of those ones between two images from the same thorax and with the same air volume but acquired at different time instants. Such similarity level were the original goal that were hoped to be achieved.

Having reached the aimed values, the registration process can be regarded as successful. If the resulting deformation fields a good approximation of the real movement field, it must be possible to apply them on the corresponding PET data and compensate for the patient's respiratory motion.

It is also worth mentioning that there is a problem associated with evaluating with landmarks. Even if the landmarks were perfectly placed (which is a false assumption), and even if different color maps were used in order to be able to lay landmarks in different regions in the lungs, it was still impossible to set any of them in regions without much detail (especially the air inside the lungs). It could be possible that the polynomial methods worked better around these areas, as they vary in a smoother way.

3.4 Registering CTs with preprocessing

It may be possible to improve the registration results by preprocessing the data. In this preprocessing step, a filter or transformation is applied to both the fixed and moving image. They are then registered and the resulting deformation applied to the original moving image.

3.4.1 Methods

In order to evaluate the preprocessing step to see if it improves the co-registration, the same registration methods as in the previous section were tested on the preprocessed images. However, only the registration from the full inhale position to the full exhale one for patient 1 was investigated.

As it was explained in the introductory chapter, one has to be cautious when using preprocessing filters. It might be possible to improve the results for a specific problem through tuning the filter parameters. On the other

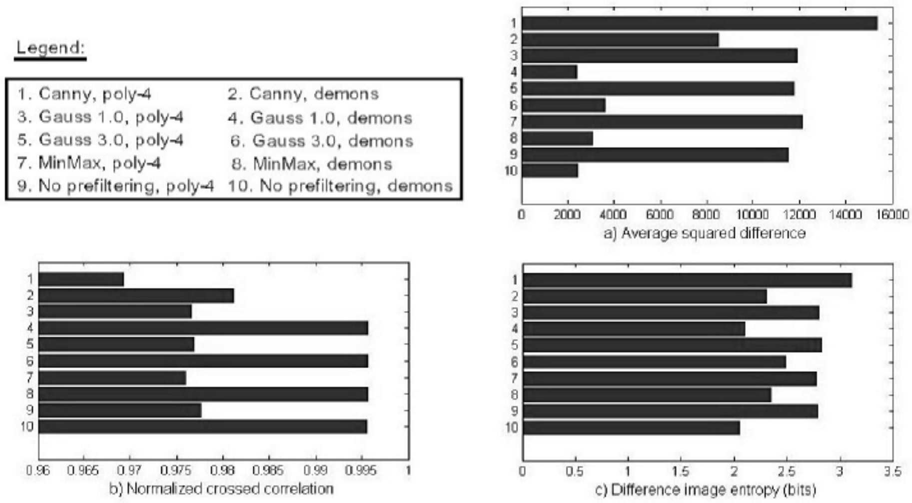

Figure 3.17: Average SSD, NCC and difference image entropy similarity measures for the registered prefiltered volumes.

hand, it is sometimes difficult to have a fully automatic tuning (and the result of the preprocessing step must be monitored).

The applied filters, configured as recommended in the ITK guide [26], were:

1. Canny edge detection filter, with a variance parameter of 2.0.

2. Discrete gaussian filter, with standard deviation $\sigma = 1mm$.

3. Discrete gaussian filter, with standard deviation $\sigma = 3mm$.

4. MinMax curvature flow filter (edge preservation blurring), with the typical values for the parameters: 10 iterations, a time step of 0.0625 and a neighborhood radius of just 1.

3.4.2 Results

The difference image entropy, NCC, and SSD similarity results are presented in figure 3.17. Only two representative registration strategies are represented for the sake of clarity: the original demons algorithm and a forth order polynomial. It can be noticed that prefiltering does not help to improve the numerical similarity results.

The results for the edge detection filter are especially bad. The noise in the images creates many false edges that are a source of mismatch between images.

An attempt of applying an edge preservation blurring filter prior to registration was also performed. The purpose was to eliminate the high frequency components while not removing the edges to be detected. The result of this attempt was not very successful either.

More promising results were achieved when prefiltering with Gaussian low-pass filters. The results were slightly better for the one with $\sigma = 1$ mm than for the one with $\sigma = 3$ mm, but still neither of them reaches the similarity measures achieved without prefiltering, neither for the demons algorithm nor for the polynomial transformation. Similar results were obtained for the MinMax curvature flow filter. The similarity measures are approximately the same as for the Gaussian filtering; the edge preservation properties of the MinMax filter did not help to improve the results.

3.4.3 Conclusions

It can be concluded that prefiltering the images before the registration process does not increase the similarity between the fixed and registered images in the available image set. However, this does not necessarily mean that prefiltering is useless in every individual case; it can improve the results when working with more noisy images. But, as this is not the case, and the study of this subject is outside the scope of this thesis, the possibility of prefiltering will not be considered from now on.