### 3.4.4 Additional results for single 2–D pairs of retinographies

In addition to these tests with *SLO* sequences, we have done an additional experiment with pairs of images belonging to other modalities. The experiment is similar to that with CT and MR volumes described in page 58: first, we register a given pair of images. Then, we misalign one images by known parameters $T_T$, and the registration algorithm is applied again. To measure the distance from the given to the recovered parameters $T_T$, we have taken the mean distance after applying the two corresponding transformations to all the non-void pixels of the creaseness image, this is, those detecting vessels. We applied 100 different transformations to three pairs of images, taken randomly within the range of translations of $\pm 25\%$ the size of the image in pixels, rotations of $\pm 25$ deg and scaling $\pm 10\%$.

We have tested three different configurations (we recall their definition for clearness)

- *no search* The iterative optimization is set to run without any previous search, this is, with nil translation and rotation seed values.
- *1S* A single seed is chosen after the exhaustive seach, which is done with the original not sampled images.
- *3M* Severals seeds are taken from the search at level three of the hierarchic pyramid.

This choice is significative because we try to demonstrate that some sort of search has to be done (thus first method should not work), and that the time to compute it can be reduced without loss of robustness (thus third method should take less time for similar results).

Table 3.7 shows the results, which corroborate the conclusions taken from the previous experiment. In initial tests, the same trials where done with transformation which did not include scaling. For them, hierarchical methods were almost 100% successful. For all tests, the recovered transformations had a mean error lower than 0.2 pixels, which is another evidence that our alignment measure is properly defined. The mean error is defined in the same way as was with the experiment with CT and MR volumes: mean distance for non-void pixels after applying trial and recovered transformations.

The explained scheme worked successfully for all pairs of images detailed in table 3.2; even without the initial exhaustive seed step, all except one actually converged. We have registered successfully more than 100 pairs of images of several modalities: 49 stereo and 48 temporal pairs ( 1 year apart, size $460 \times 416$) of angiographies and 10 green images to angiographies.

The method ran for all images unmodified for all but two parameters, which are related to the creaseness extraction: the scale of the crease and the smoothness of the image, and it had a typical registration time of less than 10 seconds. The visual inspection of the results, which was easy to perform because creases overlap when properly registered (see figure 3.15), was found satisfactory for all the sets.

Working with long sequences of images has the advantage that all tests and trials are immediately statistically relevant. For other modalities, the same relevance would

| | Set D Angiography – retinography. 370 × 278 | | Set E Retinography – green 250 × 202 | | Set F Stereo angiography 460 × 416 | |
|---|---|---|---|---|---|---|
| Method | Rec (%) | Time | Rec (%) | Time | Rec (%) | Time |
| No ini | 6 | 4,6 | 16 | 1,7 | 9 | 6,1 |
| 1S | 94 | 61,4 | 98 | 44,3 | 93 | 66,2 |
| 3M | 94 | 9,7 | 80 | 4,0 | 92 | 10,3 |

**Table 3.7:** Results of the registration algorithm for 100 random trial transformations for 3 pairs of images. The second image is partially occluded, as can be seen in figure 3.15. The Rec column represents the percentage of transformations successfully recovered (mean error lower than 5 pixels); the time is in seconds. Search methods as numbered are explained in page 91.

need either a very large database, or some artifitial experiment with a lower number of images. Therefore, we feel we have validated our method under requirements far more demanding than others in published papers, and results are still excellent.

The testbench has stated clearly that our registration scheme, i.e. creaseness+ translation search + Optimization, is suitable for registration of ophthalmologic images. Furthermore, the hierarchical approach is successful at lowering time about 95%, while keeping robustness only to 5% to 20% worse.

Next section is devoted to two further refinements of the method, with the aim of further decresing computation time and increasing accuracy.

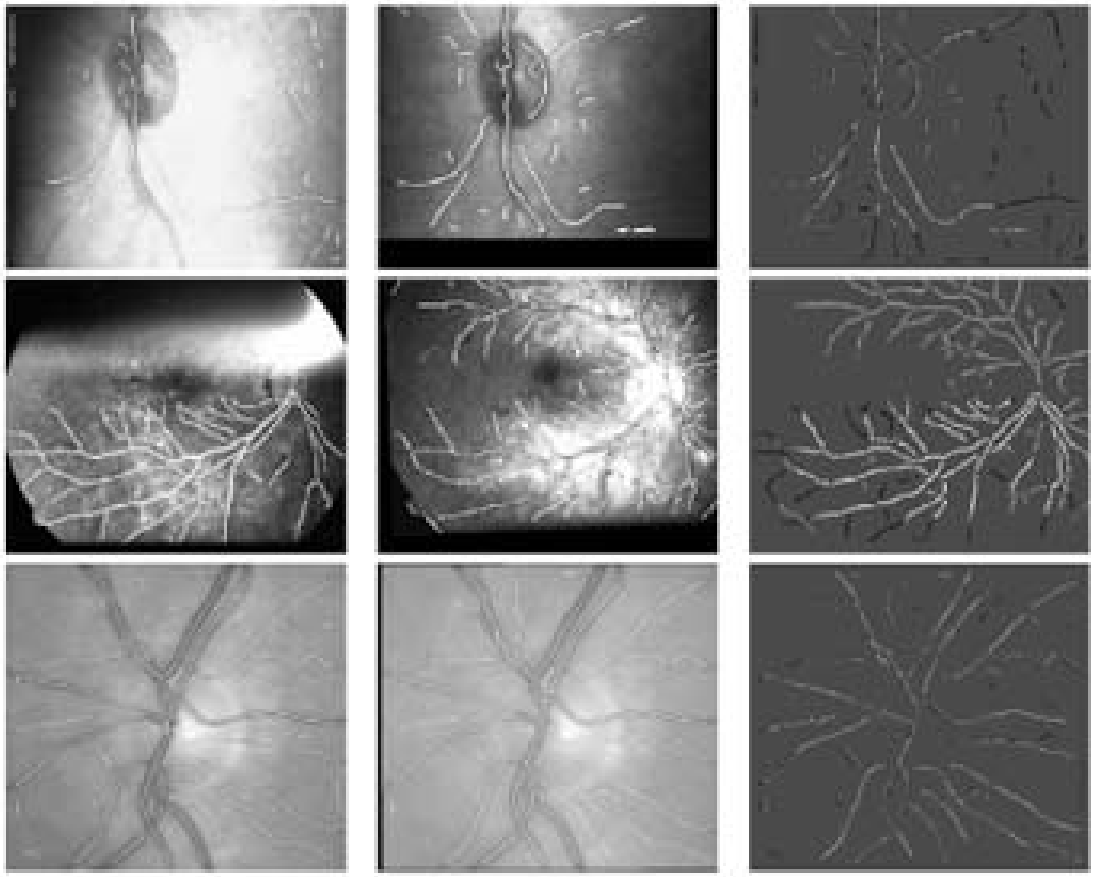

**Figure 3.15:** Datasets D–F. First and second columns: original images with creases superimposed in white. Third column: overlapping creases appear in white, non-overlapping in black and gray. From top to bottom: *SLO –SLO* , retinography to green, and retinography to retinography.

## 3.5 Refining the search

The methods from the previous testbench registered each frame in the sequence independently one to the other. But a realistic assumption is that consecutive frames ought to have similar transformations. Therefore, it makes sense to use the results from one frame as seeds for the next, thus avoiding the costly initial search in Fourier.

Of course this strategy will not work for all frames. For some, the translation is too large for the method to converge, so a secondary full search has to be set. This secondary search will be activated when:

- Previous frame was invalid or its convergence value was too low.
- Search using seeds achieved a convergence value below a threshold.

We have evaluated two methods under the same conditions of those of the previous section: methods *GM* and *GS* (for global multiple seed and global single seed), are based upon *4M* and *4S* to use single or multiple search as secondary search. They are defined in table 3.8. We have examined its performance for the same patients as the previous. Table 3.10 shows the results, but before commenting on them we would like to introduce an additional improvement.

| Method | Secondary search | Tranformation | Window size |
|---|---|---|---|
| GM | 4M | Global | −− |
| GS | 4S | Global | −− |
| L80 | 4S | Local | 80 |
| L160 | 4S | Local | 160 |

**Table 3.8:** Configuration parameters for final registration methods. They use for each frame the results from the previous as seed for the search, and the labeled 'Secondary search' method when the search has failed.

When examining visually the alignment of long sequences of *SLO* images, we realized that at some intervals images seemed to vibrate from one frame to the next. Small missregistrations for some parts of the images produced this effect, regardless of the global correctness of the registration. Illustrating figures are drawn in 3.16, numerical values of tranformations in table 3.9. An accurate inspection reveals that although all match visually, one frame, number 900, is using the top left region (square-like shape) to align, while 950 and 1000 converge using other creases. Numeric results agree with this observation.

The explanation of this oscillation must take into account the low signal-to-noise ratio but also the optics of the eye. Indeed, the acquisition procedure is very different from other medical images, rather resembling satellite observation of the earth. The fundus of the eye is a concave surface, observed through a set of lenses. Any movement between the camera and the fundus will produce not only translations in the acquired image, but also as more complex changes. The scheme is shown in figure 3.17. Deguchi and colleagues [7] model this geometry with the aim of fundus pattern reconstruction, by taking several views with a calibrated camera.

| NFrame | Rotation (deg) | $T_x$ (Pixels) | $T_y$ | $S_x$ | $S_y$ | Correlation | NC |
|---|---|---|---|---|---|---|---|
| 900 | 0.2 | -183.4 | -43.0 | 0.979 | 1.023 | 1086 | 0.17 |
| 950 | -2.3 | -190.4 | -21.5 | 1.020 | 1.014 | 763 | 0.14 |
| 1000 | -2.5 | -181.6 | -34.0 | 1.019 | 1.014 | 736 | 0.16 |

**Table 3.9:** Similar frames may convergence to different solutions.

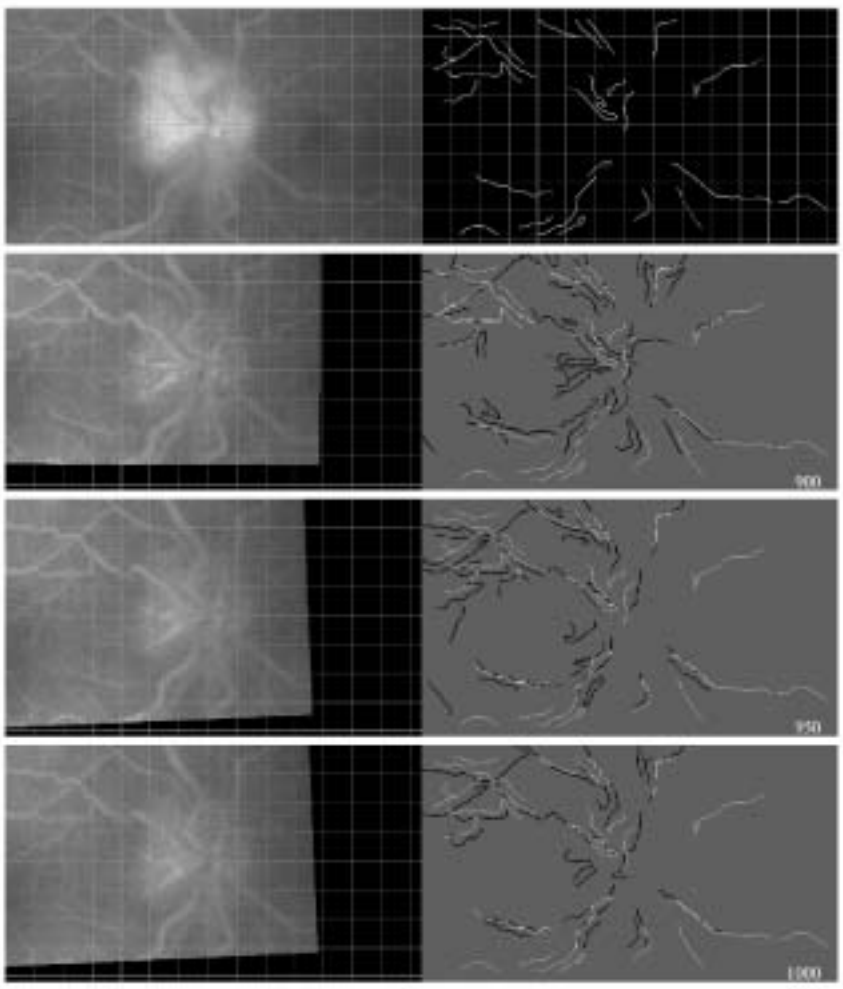

**Figure 3.16:** Similar frames may convergence to different solutions. On top, the reference frame for patient A. Frame 900 uses top left region to converge, while frames 950 and 1000 do not use it.

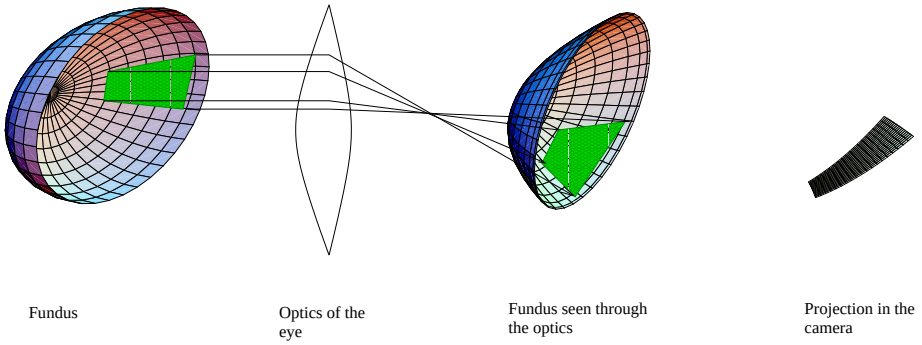


**Figure 3.17:** Model of the eye

This is the reason why adding the scaling parameters improved so much the alignment for some images. The actual pixel size derived from the acquisition parameters does not change, and neither does the distance between the eye and the camera. But small rotations of the sample plane can be modeled much better with these two extra parameters. Scaling may improve registration up to a point, but since the actual underneath model is much more complex, for some slices it is unstable and presents two solutions alike, thus becoming numerically ill posed.

Since our camera is not calibrated, Deguchi's approach is unfeasible for us. Instead, we propose a further extension of our algorithm, which consists on dividing the image into a regular set of square regions to be aligned independently ⑥. The method to align each pair of rectangles is the same we use for the global transformation.

The final algorithm is:

1. Register the images $S$ and $D$ with a global transformation $T$.
2. Transform the dinamic image with global result to produce another image $D_T$
3. Divide $S$ and $D_T$ into two sets of rectangles arranged in grids $S_i, D_i$, $i = 1 \ldots n$.
4. For each $S_i$ and $D_i$
   (a) if $S_i$ or $D_i$ is empty, set resulting transformation to nil.
   (b) Register $S_i$ with $D_i$. Let's call $T_i$ the result.
   (c) Discard $T_i$ if its normalized correlation $NC$ is lower than a threshold.
5. Apply $T$ and $\{T_i\}$ to the dinamic image to compound the final solution.

After all regions have been registered, we must combine the set of local transformations $T_i$ to produce the aligned dynamic image. We have implemented a method commonly found in satellite images, namely, a local deformation model [69]: we consider the center of each rectangle as a landmark, and compute its local transformation.

| | Method | Success (%) | Failures (%) | Correlation | NCor | Time (sec) |
|---|---|---|---|---|---|---|
| Patient A | **4M** | 86.8 | 9.3 | 1288 | 0.22 | 4.9 |
| | GM | 89.7 | 6.3 | 1205 | 0.20 | 3.6 |
| | **4S** | 86.1 | 10.0 | 1290 | 0.22 | 3.3 |
| | GS | 89.9 | 6.2 | 1203 | 0.20 | 3.0 |
| | L80 | 89.9 | 6.2 | 1203 | 0.24 | 3.0 + 7.7 |
| | L160 | 89.9 | 6.2 | 1191 | 0.23 | 3.0 + 5.4 |
| Patient B | **4M** | 94.5 | 4.3 | 2532 | 0.21 | 6.1 |
| | GM | 95.4 | 3.4 | 2318 | 0.19 | 4.1 |
| | **4S** | 90.8 | 8.0 | 2578 | 0.21 | 3.8 |
| | GS | 92.6 | 6.1 | 2390 | 0.19 | 3.1 |
| | L80 | 92.6 | 6.1 | 2385 | 0.22 | 3.1+8.5 |
| | L160 | 92.6 | 6.1 | 2374 | 0.21 | 3.1 + 7.15 |
| Patient C | **4M** | 59.3 | 39.8 | 1539 | 0.20 | 5.0 |
| | GM | 75.2 | 23.9 | 1308 | 0.17 | 4.0 |
| | **4S** | 57.9 | 41.2 | 1563 | 0.20 | 3.2 |
| | GS | 69.0 | 30.0 | 1374 | 0.18 | 3.0 |
| | L80 | 69.0 | 30.0 | 1530 | 0.19 | 2.6+7.2 |
| | L160 | 69.0 | 30.0 | 1530 | 0.18 | 2.6+5.5 |

**Table 3.10:** Registration statistics for global and local methods. We have copied results for methods **4M** and **4S** to make comparison easier. GM is the based on 4M, and 4S, L80 and L160 are based on 4S. The time for local methods is shown as: time for global registration + time for local adjustment.

Then, at any other point, the mapping is computed by two bilinear interpolations of the $x$ and $y$ coordinates of the nearest four rectangle centers.

When we first ran this experiment, we found the values of $T_i$ where often very large for windows where enough creaseness information did not exist, which means that the algorithm had converged to some random solution. Although the threshold $T_i$ discarded them, still much time was spent in this faulty searchs. The best way to avoid it was to optimize $T_i$ not centered in the actual window, but centered in the full image. Then, rotations became more constrained and the solutions more stable.

A parameter of this method is the size of the grid. We have set two configurations with size 80 and 160, and included them in the same tests as the previous proposals. They are specified in table 3.8.

After all rectangles have been registered, we need to combine the initial transformation $T$ with the set of local tranformations $T_i$ to compound the aligned dinamic image. We have implemented a method commonly found in satelite images: we consider the center of each rectangle as a landmark, and for each we compute its local transformation. Then, the coordinates of this set of transformed landmarks is bilinealy interpolated for the rest of points. We give full details of the method in appendix D.

We have applied to these methods the same battery tests as described in table

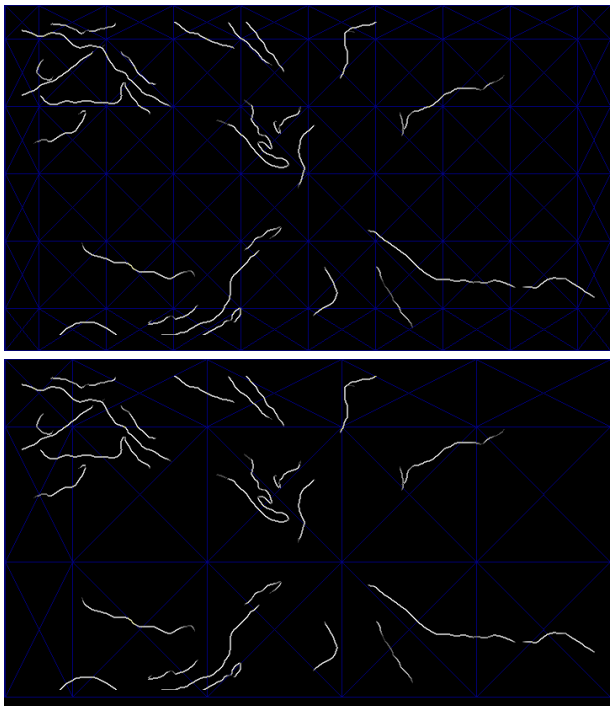

**Figure 3.18:** Images are divided in grids to be aligned independently. Grid size for method *L80* (top) and *L160* (bottom). To compound the set of transformations, the center of each rectangle will be considered as a landmark and its transformed coordinates are weighted with neighbours in a linear fashion. See appendix D for technical details.

3.4, page 92. Table 3.10 presents the statics for the four new methods, together with the two original methods they were based upon.

We have come to the following conclusions:

- Statistics show that the methods usings previous seeds achieve the goal of reducing the registration time. Probably the time could be further cut had we been less conservative in the requirements to accept a registration: when an optimization based on a seed has a value lower than a threshold, a full search is started, which does not always lead to better results. This is particularly true for patient C, which needed the full search for a larger number of slices (about one fifth), thus making the mean time similar to the non-seed based method. It remains for future tests to see whether a lower threshold value improves the time without loss of robustness.

- We were surprised to see that each seed-based method slightly improves results with respect to its memory-less version. This is particularly true for patient C, which had worse statistics. Also, the mean robustness shows that our choice of

four-level methods as source for the seed-based methods might be problematic for some sequences. At the price of increasing the processing time, a choice based on three-level methods would be more robust. However, this is a small difference which has to be investigated for a bigger number of sequences.

- As we could expect from the results of the previous section, the multiple-seed method set does not work better than single seed method, but the time spent is only slightly larger. Therefore it is up to the final requirements, and mainly up to the average expected rotation, whether to use one or the other.
- Improvements with local refinements are not clearly reflected in the statistics (NM, NC). We believe the cause is that the changes are too small compared to the total volume of the image.

The figures in these tests reveal that the seed-based methods are the choice when working with large volumes of data: with similar or better robustness, they converge faster. On the other hand, the measures employed so far do not reveal whether local registration improves the results. The following section is devoted to quantify it with the aid of other measures.