IMPROVING ROBUSTNESS FOR INTER-SUBJECT MEDICAL IMAGE REGISTRATION USING A FEATURE-BASED APPROACH

Linus Svärm* Olof Enqvist† Fredrik Kahl† Magnus Oskarsson*

* Centre for Mathematical Sciences, Lund University, Sweden  
{linus, magnuso}@maths.lth.se
† Department of Signals and Systems, Chalmers University of Technology, Sweden  
{olof.enqvist,fredrik.kahl}@chalmers.se

ABSTRACT
We propose new feature-based methods for rigid and affine image registration. These are compared to state-of-the-art intensity-based techniques as well as existing feature-based methods. On challenging datasets of brain MR and whole-body CT images, a significant improvement in terms of speed, robustness to outlier structures and dependence on initialization is shown.

Index Terms—Registration, features, robust estimation

1. INTRODUCTION
Image registration is at the core of many applications in medical imaging. It serves as a tool for performing motion correction, detecting anatomical changes and for fusing information from different modalities. It has also become an important component in atlas-based segmentation methods, whose success in, e.g., neuroimaging is largely due to the ability to accurately register brain MRIs. A number of issues need to be addressed in order to improve the applicability in other domains where the biological inter-variation is more significant. Specifically, we address the following challenges: (i) Efficiency. In atlas-based segmentation, one needs to register each atlas image to the target. (ii) Robustness to outlier structures. In many scenarios, one would like the registration method to ignore certain regions as they are not present in both images. (iii) Reliability. Many registration methods are dependent on a good initialization.

Our main contribution is to show that we can improve state-of-the-art in terms of the above-mentioned challenges via a feature-based approach. We present new techniques for fast outlier rejection, which in combination with random sampling, makes it possible to handle a huge amount of outliers at a reasonable computational cost. At present, we handle the problem of inter-subject variability, but do not consider registration across different modalities. We are also restricting our efforts to low-dimensional transformations (rigid and affine) and we do not focus on sub-pixel accuracy.

Intensity-based methods consider information from the whole image, while feature-based ones rely on geometrically defined interest points, see [1] and the references therein. Much effort has been spent on finding good interest point descriptors [2, 3], but it is still difficult to avoid incorrect correspondences. Some approaches are conservative in the matching process and rely on getting no false matches [4]. This may work well for particular applications, but not in general. Robust estimators based on iterative methods have been proposed, e.g., [5], but they are sensitive to initialization. The most similar work to ours is the feature-based method in [6]. They model both inliers and outliers in a statistical setting, and can thereby learn parameters. A drawback is that candidate transformations are generated based on a single correspondence which may severely limit the actual search space. Robustness has also been addressed in intensity-based methods, e.g., in [7], using the $L_1$-norm and [8], where another robust loss function is utilized. The downside is that the optimization relies on local refinement which is sensitive to initialization.

There are a number of standard registration toolkits publicly available and we will compare our approach to several of them, including IRTK [9], NIFTYREG [7] and the feature-based method in [6]. We also include the robust approach in [8], which has proved more robust than the registration methods FLIRT [10] and SPM [11].

2. FEATURES AND MATCHING
Rotation-invariant feature descriptors like SIFT [12] and SURF [13] are standard 2D tools, both of which have been adapted to 3D [14, 4]. The path we have chosen is similar, though not identical to either of SIFT and SURF.

(i) Interest point detection. This step identifies a limited number of so-called interest points for further processing. Like SIFT we obtain interest points from scale-space extrema of a difference-of-Gaussians operator, as originally proposed.
in [15]. This produces a set of interest points in scale space. The detected scale will be used to achieve scale invariance in feature matching.

(ii) Orientation assignment. We achieve rotational invariance by assigning a dominant orientation to each interest point which is common practice. In cases with known (or negligible) orientation differences between sources and targets, one can assign the canonical orientation to all interest points. This improves the discriminative power of the descriptor.

(iii) Descriptor computation. A patch is defined around the interest point in question. The patch is aligned to the orientation assignment and its size is proportional to the interest point scale. The patch is divided into $4 \times 4 \times 4 (= 64)$ regions and in each region, gradients are computed. These are used to compute the SURF descriptor being a list of 6 values for each region. These $64 \times 6 = 384$ values form the descriptor vector that can be used to recognize similar interest points.

(iv) Feature matching. There will always be outliers among the matches. This can to some extent be handled with random sampling techniques, e.g., RANSAC [16]. However, for large rates of outliers, these methods fail or become exceedingly slow. Next we will present our outlier removal scheme for handling extreme outlier rates.

3. TRANSFORMATION ESTIMATION

Our main contribution is a technique to deal with large rates of incorrect matches—outliers. Given $n$ feature points $x_j \in \mathbb{R}^3$ in the source image and corresponding $y_j \in \mathbb{R}^3$ in the target image, we seek a transformation $T$ that minimizes the truncated $L_2$ loss of the residuals

$$\ell(T) = \sum_{j=1}^{n} \min(r_j^2(T), \epsilon^2),$$

where $r_j(T) = |T(x_j) - y_j|$ is the Euclidean error in the target image. The function in (1) assigns a quadratic loss to inlier correspondences with a residual lower than $\epsilon$ (which should relate to the noise level) and a constant loss to outlier correspondences with a residual higher than $\epsilon$. We will consider transformations on the form $T(x) = Mx + t$. This is a rigid transformation if $M$ is a rotation matrix, and a general affine transformation if it is a positive definite matrix.

Our approach is a generalization of [17] which was developed for multi-view geometry. The first step is an outlier rejection algorithm with running time that is independent of the rate of outliers. In practice, it will remove most of the outliers while guaranteed not to remove any inlier correspondences. In the second step, we use RANSAC [16] to get rid of a few remaining outliers and estimate an accurate transformation.

3.1. Outlier Rejection

First, we will consider maximizing the number of inliers, and then adapt it to the truncated $L_2$ loss in (1).

Problem 1. Find a transformation given by $(M, t)$ such that $|Mx_j + t - y_j| \leq \epsilon$, for as many $j$ as possible.

The algorithm loops through all $n$ correspondences and performs an outlier test for each correspondence. If the test is positive, then the correspondence can be removed permanently since it is guaranteed not to be part of the optimal solution. Let $L$ be the number of inliers of the best solution found so far. We compute a bound of the following type. If correspondence $K$, $(x_K, y_K)$, is an inlier to the (unknown) optimal transformation, then there are no more than $U_K$ inliers. Hence, if $U_K < L$ we get a contradiction and correspondence $K$ must be an outlier. Essential is of course to be able to quickly compute the upper bound $U_K$ under the assumption that correspondence $K$ is an inlier. This issue will be the focus for the remainder of this section.

The first step is reducing the original problem to a simpler one by eliminating the translation $t$.

Problem 2. Given $K$, let $\tilde{x}_j = x_j - x_K$, $\tilde{y}_j = y_j - y_K$ and find $M$, such that $|M\tilde{x}_j - \tilde{y}_j| \leq 2\epsilon$ for as many $j$ as possible.

Proposition 1. If correspondence $K$ is an inlier to Problem 1, then Problem 2 has at least as many inliers as Problem 1.

Proof. Let $(M, t)$ be the solution to Problem 1. For any inlier $j$, we have $|Mx_j + t - y_j| \leq \epsilon$, and using the same $M$ in Problem 2, we get

$$|M\tilde{x}_j - \tilde{y}_j| = |M(x_j - x_K) - (y_j - y_K)| \leq |Mx_j + t - y_j| + |Mx_K + t - y_K| \leq 2\epsilon.$$

Hence, solving Problem 2 produces an upper bound $U_K$. As a by-product, we get candidate solutions that we can use to continuously improve the best solution found so far, which will make the outlier test more efficient as the value of $L$ increases.

3.2. Outlier Rejection for Rigid Registration

For rigid transformations ($M = R$), we also make use of the dominant directions assigned to each feature point by the descriptor. Let $u_j$ be the dominant direction of point $j$ in the source image and $v_j$ the corresponding direction in the target. As these are used to align the descriptors it is unlikely for a correct match not to satisfy the following angular error bound

$$\rho_j(R) = \angle(Ru_j, v_j) \leq \tau,$$

for a moderate threshold $\tau$. We will further simplify our problem to a 1D-rotation problem (rotation around one axis).

Problem 3. Find a rotation $R$ with $Ru_K = v_K$ such that the following constraints are satisfied for as many $j$ as possible.

$$||\tilde{x}_j| - |\tilde{y}_j|| \leq 2\epsilon,$$

$$\angle(Ru_j, v_j) \leq 2\tau.$$
\[ \angle(R\tilde{x}_j, \tilde{y}_j) \leq \tau + \alpha, \]  
where \( \alpha \) is obtained from \(|\tilde{x}_j|^2 + |\tilde{y}_j|^2 - 2|\tilde{x}_j||\tilde{y}_j| \cos \alpha = 4e^2\).

**Proposition 2.** If correspondence \( K \) is an inlier to the optimal transformation of Problem 1 that also satisfies (2), then Problem 3 has at least as many inliers as Problem 1.

**Proof.** Let \( R_0 \) be the optimal rotation to Problem 2 and consider the triangle with sides \( A = R_0\tilde{x}_j, B = \tilde{y}_j \) and \( C = A - B \). Then, \(|C| = |R_0\tilde{x}_j - \tilde{y}_j| \leq 2e\), and with the triangle inequality on \( ABC \), we get (3). Using the rule of cosines on the same triangle we get \( \angle(R_0\tilde{x}_j, \tilde{y}_j) \leq \alpha \) with \( \alpha \) as above. To find a suitable \( R \), we use the fact that \( \angle(Ru_K, v_K) = \rho_K(R, t) \leq \tau \). Let \( R_\Delta \) be a rotation with rotation angle \( \tau \) mapping \( R_0u_K \) to \( v_K \) and set \( R = R_\Delta R_0 \). Then \( \angle(R\tilde{x}_j, \tilde{y}_j) \leq |R_\Delta| + \angle(R_0\tilde{x}_j, \tilde{y}_j) \leq \tau + \alpha \) and using the triangular inequality for rotations

\[ \angle(Ru_j, v_j) \leq |R_\Delta| + \angle(R_0u_j, v_j) \leq 2\tau, \]

This shows that \( R \) is a solution to Problem 3 with at least as many inliers as the optimal solution to Problem 2. Applying Proposition 1 completes the proof. \( \square \)

Algorithm 1 shows how this result can be used to reject outliers. After the change of coordinates in Step 1, any rotation satisfying \( Ru_K = v_K \) will be a rotation about the first coordinate axis. Hence the subsequent task is to find a rotation angle, \( \phi \). We omit the details due to space limitations.

**Algorithm 1 Upper bound \( U_K \) for correspondence \((x_K, y_K)\)**

1. Change coordinates s.t. \( u_K = i_K = (1, 0, 0)^T \)
2. Remove any correspondence that violates (3).
3. For each remaining correspondence \((\tilde{x}_j, \tilde{y}_j)\):
   a. Compute interval \( I_\delta \) of rotation angles satisfying (4).
   b. Compute interval \( I_\delta \) of rotation angles satisfying (5).
   c. Intersect \( I_\delta \) with \( I_\delta \). Store the resulting 0-2 intervals.
4. Compute an angle \( \phi \) inside as many intervals as possible. Output the maximum number of intersecting intervals, \( U_K \).

Computationally, the most expensive part is Step 4 which includes sorting. Hence the complexity of Algorithm 1 is \( O(n \log n) \). If we repeat this for every \( K \) we get a total cost of \( O(n^2 \log n) \) for our outlier rejection scheme.

### 3.3. Outlier Rejection for Affine Registration

The method from the previous section can be used to perform outlier rejection whenever the deformation is known to be small, but some important applications require a more significant scaling to register the images, for example, registering a whole-body scan of a tall person to one of a short person. To handle such cases efficiently, we propose an outlier rejection scheme based on the assumption that the rotation between images is small. This is true for most medical 3D images, such as CT, MR or PET. The class of transformations that we will use in this case is \( T(x) = sx + t \), where \( s \) is a positive scale factor. Now consider Problem 2 for this transformation and some index \( K \). For each \( j \) we can compute an interval constraint on \( s \) for correspondence \( j \) to be an inlier. We find the interval by the following geometric reasoning. As \( s \) varies, the point \( s\tilde{x}_j \) moves along a line segment starting at the origin. We are interested in points such that \(|s\tilde{x}_j - \tilde{y}_j| \leq 2e \). It is easy to see that this is true for an interval of \( s \) and we can find the interval boundaries by using simple linear algebra

\[ 4e^2 = |s\tilde{x}_j - \tilde{y}_j| = s^2|\tilde{x}_j|^2 + |\tilde{y}_j|^2 - 2s\tilde{x}_j\tilde{y}_j, \]

and solving the obtained quadratic equation. As in rigid registration, we seek an \( s^* \) that intersects as many of these intervals as possible. This can be found by sorting all the interval boundaries and going through the sorted list once. The number of intervals intersecting at \( s^* \), \( U_K \) is an upper bound on the form: If correspondence \( K \) is an inlier, then there are no more than \( U_K \) inliers. The complexity of running this scheme for all of the \( n \) correspondences is again \( O(n^2 \log n) \).

### 3.4. Outlier Rejection for Truncated Least Squares

For the truncated \( L_2 \) loss in (1), we can use these methods in the following way: Assume that correspondence \( K \) is an inlier and compute a bound, \( U_K \), on the number of inliers. The truncated \( L_2 \) loss cannot be smaller than \( (n - U_K)c^2 \). If the best solution found so far has a lower loss, then we get a contradiction and correspondence \( K \) can be rejected.

### 4. EXPERIMENTAL RESULTS

This section presents experimental results for three different set-ups. The performance is compared to five other methods, namely IRTK, NIFTYREG, Feature-Based Alignment (FBA) from [6] (not applicable for affine registration) and the robust method in [8], hereafter called ROBUST. These methods were used with default settings including multi-resolution initializations, except ROBUST for which the outlier sensitivity was decreased for better results. Note that in [8], ROBUST was shown to outperform FLIRT [10] and SPM [11].

#### 4.1. Rigid Registration of Brain MR images

In the first experiment, we quantify the robustness to outlier structures in the images using 30 T1-weighted brain MR scans from brain-development.org with a resolution of \( 180^3 \). We use the same setup as in [8] for testing robustness. For each image, a random rigid transformation is applied to reflect possible head movements in a scanner: a random voxel translation and a rotation about a random axis. We add Gaussian noise (\( \sigma = 10 \)) and create outlier boxes of size \( 20^3 \) by copying a box from another image, see left of Fig. 1. We evaluate the dependence on rotation angle with 40 outlier boxes in each image and then the sensitivity to varying
the number of outlier boxes with a fixed amount of rotation (30 degrees). Our framework for rigid registration was used with an outlier threshold $\epsilon = 5$ voxels for the Euclidean error and $\tau = \pi/5$ for the angular error.

Results are shown in Fig. 2. We plot the number of successful registrations for each method, i.e., registrations with a rotation error less than 1 degree and a translational error less than 1 voxel. In both settings, our method significantly outperforms the competitors. IRTK, which is not designed to deal with outlier patches, failed consistently and we chose to leave it out of the comparison.

### 4.2. Affine Registration of Organ CT Images

In this experiment we used 10 whole-body CT images (resolution $512 \times 512 \times 800$) of different subjects to evaluate the robustness to initialization. There are 15 organs, e.g., kidneys, spleen, liver, that have been manually delineated [18]. We used cropped organ images from one subject in order to register them to the whole-body CT of a different subject using IRTK, ROBUST, NIFTYREG and our method with the outlier threshold $\epsilon$ at 30 voxels. The performance is measured by the distance in voxels between organ centres in the target and warped source images.

Table 1 reports the proportion of successfully registered organs. IRTK did not perform well with cropped images (as expected since it is dependent on a reasonable initialization), so we left it out of the comparison. Note that we are primarily interested in robustness (and not accuracy) and therefore we concentrate on rates of successful registrations. It is evident that ROBUST is dependent on a good initialization while NIFTYREG and our method are less so.

### 4.3. Affine Registration of Whole-Body CT

In this experiment we test the performance of pairwise registering whole-body CT images using the same data, settings and algorithms as in the previous experiment. We ran a number of inter-subject registrations and measured how many of the organs that were successfully registered as defined in the previous experiment. The results are given in Table 2. Even though the CT scans are roughly aligned from the beginning, the competing methods are less successful than ours.

### 4.4. Execution Times

Average execution times for the different methods and experiments are given in Table 3. Even for registering a single image pair, our approach is more efficient, but the real advantage appears in a multi-atlas setting. The most expensive step in our algorithm (and similarly for FBA) is computing features for the two images. In an atlas setting, the features can be precomputed for the atlas. As an example, the time to register one whole-body CT image to a multi-atlas of size $N$ is $20 + 1.4N$ seconds. With $N = 50$, our method would take 1.5 minutes, IRTK over an hour, NIFTYREG about 7 hours and ROBUST 22 hours on the same computer.

### 5. CONCLUSIONS

We have proposed a highly efficient and robust method for inter-subject registration of 3D medical images. It can handle large image variations and it does not require initialization. In terms of efficiency, it is especially well-suited to the multi-atlas setting. Our code for feature extraction and registration will be made publicly available. 1

---

1 [http://www.maths.lth.se/matematiklth/personal/fredrik/download.html](http://www.maths.lth.se/matematiklth/personal/fredrik/download.html)
6. REFERENCES


