Tractable and Reliable Registration of 2D Point Sets

Erik Ask[1], Olof Enqvist[2], Linus Svärm[1], Fredrik Kahl[1,2], and Giuseppe Lippolis[3]

[1] Centre for Mathematical Sciences, Lund University, Lund, Sweden
[2] Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
[3] Department of Clinical Sciences, Division of Urological Cancers, Skåne University Hospital, Lund University, Malmö, Sweden

Abstract. This paper introduces two new methods of registering 2D point sets over rigid transformations when the registration error is based on a robust loss function. In contrast to previous work, our methods are guaranteed to compute the optimal transformation, and at the same time, the worst-case running times are bounded by a low-degree polynomial in the number of correspondences. In practical terms, this means that there is no need to resort to ad-hoc procedures such as random sampling or local descent methods that cannot guarantee the quality of their solutions.

We have tested the methods in several different settings, in particular, a thorough evaluation on two benchmarks of microscopic images used for histologic analysis of prostate cancer has been performed. Compared to the state-of-the-art, our results show that the methods are both tractable and reliable despite the presence of a significant amount of outliers.

1 Introduction

Image registration is a classical problem in computer vision and it appears as a subroutine for many imaging tasks. For example, it is a prerequisite for shape analysis and modeling [6] and for automated analysis of multi-modal microscopy images [17]. It is also an important component in image guided surgery where often fiducial markers are used for estimating the transformation [14,7]. In this paper, we are interested in estimating rigid image transformations under less controlled situations where there may be a substantial number of mismatches and where it is important to obtain reliable results. For example, the method should not be dependent on a good initialization.

Naturally, the registration problem has been studied in depth. When choosing the method of preference, one is often faced with the following dilemma. Using a simplified, mathematical model of the problem enables efficient computations, but sacrifices realism. While using a more realistic model incurs the computational cost of hard inference. As an example, consider the case of feature-based registration under the assumption that measurement noise in the target image can be modeled by independently distributed Gaussian noise. This is in fact the standard Procrustes problem which can be solved in closed form. However, the model is not so realistic as there are typically erroneous measurements - *outliers* - among the feature correspondences. This makes the registration estimates very unreliable. On the other hand, modeling outliers leads

D. Fleet et al. (Eds.): ECCV 2014, Part I, LNCS 8689, pp. 393–406, 2014.

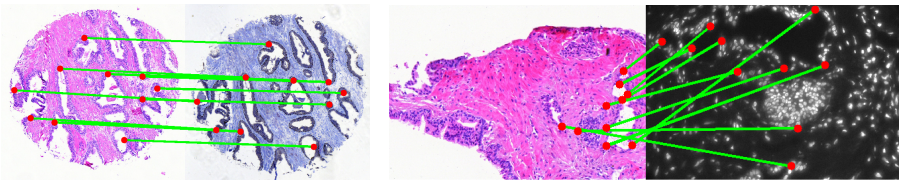

Fig. 1. Examples from our two benchmarks with 10 manually marked correspondences. *Left:* Prostate tissue stained with H&E and p63/AMACR. *Right:* Prostate tissue stained H&E and TRF (fluorescent). The goal is to find a rigid transformation that aligns the two images using features from an automated method such as SIFT.

to a much more complicated optimization problem and solving this problem exactly is sometimes dismissed as infeasible. Heuristic methods based on random sampling and expectation maximization dominate the field. We show that one can achieve a method which is both efficient (in terms of speed) and reliable (with respect to outliers).

Our own interests stem from the study of automated methods in medical imaging. In particular, we seek to develop robust registration procedures for combining information from different sources and modalities. The images may be degraded and have limited/varying fields of view. We present experimental results from two different applications. In our first setting, the objective is to perform histologic analysis of biopsies. Prostate cancer is the second most common cancer in men worldwide [16] and whose gold standard of diagnosis and prognosis is based on histologic assessment of tumours in images stained with Hematoxylin and Eosin (H&E). Several automatic pattern recognition prototypes exist [20,8]. In order to improve the accuracy in clinical practice, considerable research efforts have been directed to complement the analysis with additional types of stainings and imaging modalities [17]. One example is given in the left of Fig. 1 where two adjacent tissue sections have been stained with H&E and antibodies directed against p63/AMACR, respectively. Another example is given in the right of Fig. 1 with one H&E staining and one Time Resolved Fluorescence (TRF) image measuring the Androgen Receptor (AR) obtained from the same section. This type of images is quite challenging for any automated approach because reliable feature correspondences are hard to obtain and there are image degradations due to imperfect acquisition.

In our second setting, we are dealing with images of the human brain and the goal is to study the perfusion of blood flow through small vessels, so-called capillaries in the white and gray matter regions of the brain. This is important for patients with hydrocephalus which are treated by placing a drainage tube (shunt) between the brain ventricles and the abdominal cavity to eliminate the high intracranial pressure, see `wikipedia/hydrocephalus` [24]. To capture the anatomy of the region of interest, MR-Flair images have been obtained. The perfusion data is obtained via contrast-enhanced CT images taken at one second apart during a two-minute session. To acquire good temporal resolution, only a couple of slices can be captured at each time instant. The challenge here is to register single slices from the CT image to the full 3D volume of the MR image. As the head of the patient is in an upright position, the mapping from one CT slice to the corresponding (but unknown) slice in the MR-Flair volume is well described by a rigid 2D transformation after having adjusted for known scale differences.

In this paper, we develop two new robust methods for feature-based image registration based on the L_1-norm of the residual functions. As we shall see, from a statistical point of view, this model is well-suited for dealing with outliers. The methods are compared and extensively evaluated on two benchmarks of prostate tissue samples. The focus of our evaluation is on two important desiderata that a satisfactory solution should possess, namely *tractability* and *reliability*. The first term refers to the computational complexity. We investigate both the performance in practice and derive theoretical complexity bounds as a function of the number of feature correspondences. The second one concerns the reliability of the estimate. We are interested in methods that produce provably optimal estimates under a robust loss function. If the registration fails, then it can either due to lack of good correspondences or the algorithm's inability to find a good solution. In our approach, the latter source of error is removed from the process.

Our main contributions can be summarized as follows.

- Two new registration methods based on the L_1-norm and the truncated L_1-norm with worst-case complexity $\mathcal{O}(n^3)$ and $\mathcal{O}(n^3 \log(n))$, respectively, where n is the number of correspondences[1].
- An extensive experimental evaluation and comparison with other registration methods on benchmarks with 88 and 103 image pairs, respectively.

Note that the set of algorithms we propose is restricted to rigid point set registration in the plane, and other settings are not considered in this paper.

2 Related Work

Closed form solutions to the standard Procrustes alignment problem have been known for a long time [15] and used in various settings, for instance, in surface alignment [3]. However, as the estimate is based on least-squares (minimum of L_2-errors), outliers will have a large influence and that makes the approach unreliable. Already in [21], it is emphasized that robustness is a key issue and a multi-scale approach is proposed that integrates local measures to obtain an estimate of a rigid transformation. The method is applied to the problem of registering serial histological sections. In [19], a probabilistic method is developed that explicitly models outliers and which regards the registration problem as an inference problem. Inference is performed via expectation-maximization. The method in [13] proposes to use the Huber kernel as a residual function to make the registration less sensitive to outliers. Levenberg-Marquardt iterations are performed in order to minimize the loss function. In [22] deterministic annealing is proposed in order to optimize a robust loss function for the registration of autoradiograph slices. Yet another example is [9], where meta-heuristics is applied for the optimization step of the registration of angiograms. See also the registration survey [2]. All of these local optimization techniques are dependent on a good initial estimate and they are susceptible to local optima. Hence, they cannot guarantee the quality of their solutions.

Another popular approach for dealing with outliers is RANSAC [12]. It works by hypothesize-and-test: Pick a random minimal subset of correspondences, compute a

[1] The algorithms will be made publicly available to promote further research.

hypothetical transformation and check how many of the other correspondences are consistent with this transformation. The method is by nature random (which can be remedied by exhaustively examining all possible subsets). Still, the estimator has no guarantee of finding the optimal solution which makes the method unreliable. This will be empirically demonstrated in our evaluation.

Several works have focused on optimal estimators based on branch-and-bound. One of the first algorithms was developed in [5] and it finds the rigid transformation that maximizes the number of inliers. In [11], a robust estimator based on a vertex cover formulation is proposed and in [18], a formulation based on integer programming is given. The methods are independent of initialization and converge to a global optimum. However, as they are based on branch-and-bound, the computational complexity of the algorithm is exponential. The most closely related work to ours is [10,1], where a truncated L_2-norm algorithm is derived with complexity $\mathcal{O}(n^4)$. However, the runtime tends to be prohibitive (see experimental section), making it a less tractable alternative.

3 Choice of Loss Function

It is a common and reasonable assumption that there exist *correct* but noisy point correspondences as well as complete mismatches or outliers. The errors in the positioning of correct correspondences follow approximately a normal distribution, whereas the outliers are uniformly spread over the image. In [4] it is shown that in order to find a maximum likelihood estimate, a sum of loss functions of the following type

$$\ell(r) = -\log\left(c_1 + \exp\left(-r^2/c_2\right)\right) \tag{1}$$

should be minimized, where r is the residual error for one correspondence and the constants depend on the amount of inlier noise as well as on the rate of outliers; see Fig. 2. An approximation which is commonly used is obtained by truncating the squared error. However, the quality of this approximation depends heavily on the rate of outliers in data. At higher rates the loss function levels out much more slowly. In this case a truncated L_1-loss can be a more appropriate choice.

All these loss functions lead to a non-convex problem with many local minima. One may even wrongly conclude that the problem is intractable to solve optimally, that is, that no polynomial-time algorithm exists.

4 Fast Optimization of the Truncated L_1-Norm

Given corresponding point coordinates in two images, $\mathbf{x}_i = (x_i, y_i)^T$ and $\mathbf{x}'_i = (x'_i, y'_i)^T$, $i = 1, \ldots, n$, consider the following problem

$$\min_{R,\mathbf{t}} \sum_{i=1}^{n} \ell(||R\mathbf{x}_i + \mathbf{t} - \mathbf{x}'_i||_1) \tag{2}$$

where R is a 2×2 rotation matrix and $\mathbf{t}$ a translation vector, parameterized as

$$R(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \text{ and } \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix},$$

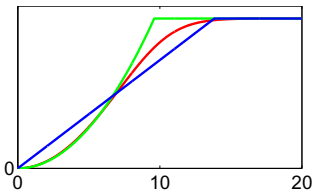

Fig. 2. The robust loss function (red) suggested in [4], the truncated L_2-error (green) and the truncated L_1-error (blue) that can be optimized using the proposed framework

respectively and where ℓ is the loss function $\ell(r) = \min\{r, \epsilon\}$ for some given threshold ϵ, that is, the truncated L_1-norm.

The following observation allows us to simplify the problem.

Lemma 1. *For any fixed rotation R, consider the minimization of (2) over* $\mathbf{t}$

$$\min_{\mathbf{t}} \sum_{i=1}^{n} l\left(|x_i \cos\alpha - y_i \sin\alpha + t_1 - x_i'| + |x_i \sin\alpha + y_i \cos\alpha + t_2 - y_i'|\right). \quad (3)$$

Then there exist two indices j *and* k *in* $\{1, \ldots, n\}$ *such that*

$$t_1^* = x_j' - x_j \cos\alpha + y_j \sin\alpha \text{ and } t_2^* = y_k' - x_k \sin\alpha - y_k \cos\alpha \quad (4)$$

is an optimal choice of $\mathbf{t}$.

In order to get a geometric intuition why the above lemma is true, consider the graph of the loss function in (3). Note that it is piecewise linear in $\mathbf{t}$ and a global minimum can be found by examining all break points, that is, points which are non-differentiable in all directions. There are two different causes for non-differentiability in our objective function. One is due to truncation and one is due to taking absolute values. Our proof shows that break points that are also local minima are given by (4). This means that break points caused by truncation need not be examined since all local minima are due to taking absolute values.

Proof. The optimal $\mathbf{t}^*$ to the truncated L_1-loss, denoted $L(\mathbf{t}^*)$, is also a global minimizer to the L_1-loss on the set of optimal inlier correspondences (those that have residuals less than ϵ). To see this, let $L_{inliers}(\mathbf{t}^*)$ be the optimal loss on the inliers and $L_{outliers}(\mathbf{t}^*)$ the loss on the outliers. Assume that there exists a different solution $\mathbf{t}$ with $L_{inliers}(\mathbf{t}) < L_{inliers}(\mathbf{t}^*)$. Clearly, $L_{outliers}(\mathbf{t}) \leq L_{outliers}(\mathbf{t}^*)$ as this is already maximal. Hence $L(\mathbf{t}) = L_{inliers}(\mathbf{t}) + L_{outliers}(\mathbf{t}) < L_{inliers}(\mathbf{t}^*) + L_{outliers}(\mathbf{t}^*) = L(\mathbf{t}^*)$ which is a contradiction.

This shows that an optimal $\mathbf{t}^*$ is a local optimum to the L_1-loss. The formula for the L_1-loss is given by

$$\sum_{i=1}^{n} |x_i \cos\alpha - y_i \sin\alpha + t_1 - x_i'| + |x_i \sin\alpha + y_i \cos\alpha + t_2 - y_i'|.$$

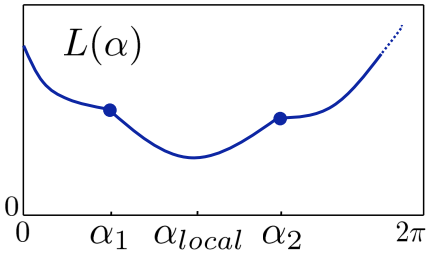

Fig. 3. Sketch of the objective function in (5), denoted $L(\alpha)$, which is piecewise smooth

As no absolute value contains both t_1 and t_2 we can write this as a function of t_1 plus a function of t_2 and the minimization with respect to t_1 and t_2 can be analyzed separately. Consider the t_1-part. We have a piecewise linear function that tends to infinity as $|t_1|$ tends to infinity and thus a minimizer of this function is at a break point. The break points are due to the absolute values - there is a break point whenever one of the absolute values is exactly zero. Hence a minimizer exists for which at least one absolute value is zero, so $t_1^* = x'_j - x_j \cos\alpha + y_j \sin\alpha$ for some j as stated in the lemma. The same argument for t_2 proves the lemma. □

The lemma shows that if the two indices j and k are given (for example, by exhaustively trying all possibilities), we can reduce the problem via substitution of $\mathbf{t}^*$ in (4) to a one-dimensional search over rotation angle α,

$$\min_{\alpha} \sum_{i=1}^{n} \ell(|\delta x_{ij} \cos\alpha - \delta y_{ij} \sin\alpha - \delta x'_{ij}| + |\delta x_{ik} \sin\alpha + \delta y_{ik} \cos\alpha - \delta y'_{ik}|), \quad (5)$$

where $\delta x_{ij} = x_i - x_j$, $\delta y_{ij} = y_i - y_j$, etc. Let us denote the resulting, piecewise smooth objective function in (5) by $L(\alpha)$, see Fig. 3 for an illustration. It has optimum either at a break point or at a stationary point. The break points are places where the derivative $L'(\alpha)$ is discontinuous and occur when an absolute value is exactly zero or the number in an input to ℓ is exactly ϵ. Hence the number of break points grows linearly with n. Given the break points $\alpha_1, \alpha_2, \ldots, \alpha_M$, consider an interval $[\alpha_i, \alpha_{i+1}]$ of $L(\alpha)$. It can be described by

$$L(\alpha) = w_1 \cos\alpha + w_2 \sin\alpha + w_3, \quad (6)$$

for some constants w_1, w_2 and w_3. By examining all intervals, we can compute the optimal rotation angle α^* using Algorithm 1.

4.1 Complexity

There are two important things to note here. First, that each time we compute w_1, w_2 and w_3 for $[\alpha_i, \alpha_{i+1}]$ in (6), we can take advantage of the constants from the previous interval $[\alpha_{i-1}, \alpha_i]$. Only the coordinates $\mathbf{x}_i$ and $\mathbf{x}'_i$ that gave rise to α_i are required for computing the update. Second, that there is only one local minimum to

$$w_1 \cos\alpha + w_2 \sin\alpha + w_3,$$

Algorithm 1. Finding the rotation angle

Set $L^* := \infty$.
Compute all break points of $L(\alpha)$ for $\alpha \in [0, 2\pi)$.
Sort the break points $\alpha_1, \alpha_2, \ldots, \alpha_M$.
for $i = 1, \ldots, M$
    Compute $L(\alpha_i)$ and compare with L^*.
    Compute w_1, w_2 and w_3 of (6) for $[\alpha_i, \alpha_{i+1}]$.
    Compute local minimum α_{local} of (6).
    if $\alpha_{local} \in [\alpha_i, \alpha_{i+1}]$,
        compute $L(\alpha_{local})$
        compare with L^*.

being

$$(\cos\alpha, \sin\alpha) = \pm(w_1, w_2)/\sqrt{w_1^2 + w_2^2}, \tag{7}$$

given by the minus sign. Hence each step in the for-loop of Algorithm 1 is $\mathcal{O}(1)$ so the computationally heaviest step is the sorting. Given the indices j and k, we can find an optimal α^* in $\mathcal{O}(n \log n)$. If we consider all possible index pairs j and k exhaustively, the total complexity is $\mathcal{O}(n^3 \log n)$. Note that the most complex arithmetic operations in the algorithm consists of computing square roots.

4.2 Fast Outlier Rejection

To increase the speed even more we propose a fast outlier rejection step as preprocessing, inspired by the work in [1]. For this we need a variant of Algorithm 1 that works with the zero-one loss (denoted by L_0), that is, counting the number of outliers rather than truncated L_1-norm. First note that the zero-one loss has the same break points as truncated L_1 and that the loss function only changes values at these break points. There, it either increases with one or decreases with one. Algorithm 2 lists the details.

Algorithm 2. Upper bound on inliers

Initialize best loss, $L_0^* = \infty$.
Compute all break points of $L_0(\alpha)$ for $\alpha \in [0, 2\pi)$.
Sort the break points $\alpha_1, \alpha_2, \ldots, \alpha_M$.
Compute $L_0(\alpha_1)$ and update L_0^*.
for $i = 2, \ldots, M$
    Depending on the type of α_i
    Set $L(\alpha_i) = L(\alpha_{i-1}) \pm 1$ and update L_0^*.

We will use this algorithm together with the following observation.

– Assume that for the optimal transformation $(R^*, \mathbf{t}^*)$, correspondence k is an inlier and there are N outliers, i.e. residuals larger than ϵ. If we change the translation to

$\mathbf{t}$ so that $r_k(R^*, \mathbf{t}) = 0$, then, since $||\mathbf{t} - \mathbf{t}^*|| \leq \epsilon$, the error on inliers has increased with at most ϵ so there are at most N residuals larger than 2ϵ.

This means that we can use Algorithm 2 with threshold 2ϵ to produce a bound of the following kind: *If correspondence k is an inlier, then there are at least N outliers.* This also yields a bound on the truncated L_1 loss, as if N residuals are $> \epsilon$, then the truncated L_1 loss is at least $N\epsilon$. If this is a higher loss than one we have already found, we can discard correspondence k from further consideration.

Algorithm 3. Fast Outlier Rejection

Given an upper bound L_c on the optimal loss.
for $i = 1, \ldots, n$
  Set $\mathbf{t} = \mathbf{x}'_i - \mathbf{x}_i$
  Use Algorithm 2 with threshold 2ϵ to compute L_0^*
  (The output L_0^* is a bound on the number of outliers)
  if $L_0^*\epsilon > L_c$,
    discard correspondence i

A value for L_c can be found by running Algorithm 3 using ϵ in place of 2ϵ and simply storing the best loss function value rather than discarding points. As the dominating cost inside the loop is the sorting in Algorithm 2 running this scheme to remove outliers costs only $\mathcal{O}(n^2 \log n)$ and can be used as a preprocessing step while keeping guaranteed optimality.

5 Fast Optimization of the L_1-Norm

Optimizing the L_1-norm is a simpler problem compared to the truncated case. In fact, one can set $\epsilon := \infty$ and use the same algorithm, but we can do better. Lemma 1 still applies, so we can eliminate the translation and only consider the rotation problem, which simplifies to

$$\min_{\alpha} \sum_{i=1}^{n} |\delta x_{ij} \cos\alpha - \delta y_{ij} \sin\alpha - \delta x'_{ij}| + |\delta x_{ik} \sin\alpha + \delta y_{ik} \cos\alpha - \delta y'_{ik}|. \quad (8)$$

An important difference here is that we can compute the break points for the first term and the second term *independently*. This means that we can precompute and sort all the break points for $j, k = 1, \ldots, n$ in $\mathcal{O}(n^2 \log(n))$ and then use the for-loop of Algorithm 1 to find the optimal α^*. Now, the heaviest part is no longer the sorting. The total time complexity is $\mathcal{O}(n^3)$ since the for-loop is $\mathcal{O}(n)$ and exhaustively trying all combinations of j and k is $\mathcal{O}(n^2)$.

6 Experiments

The proposed methods have been evaluated on two challenging registration tasks.

6.1 Registering Histology Sections

The first set of experiments is concerned with the registration of histology sections of prostate tissue, and also serves as a quantitative evaluation. We used one dataset with 88 image pairs of adjacent slides of prostate tissue, stained using H&E and p63/AMACR, respectively. Another dataset consists of 103 images of H&E stained slides, in which sub-parts are also analyzed with TRF. Examples can be seen in Fig. 1. The size of the stained images are on the order of 1100x1100, while the TRF images are 368x546.

We used SIFT features as the basis of our point-to-point correspondences. Matching was restricted to the same scale octave and we used Lowe's ratio criterion with a threshold at 0.9 to discard poor matches. This yielded 800-1500 matches for the first dataset, and, due to TRF images being smaller, 40-500 matches in the second dataset. The inlier rate varies from 1% to 40% with a 10% average for the H&E-p63/AMACR set and from 4% to 54% for the H&E-TRF set with a 28% average.

The proposed algorithms were compared to the algorithm for truncated L_2-norm from [1] as well as standard L_2-minimization and RANSAC followed by L_2-minimization on the inlier set. For each problem instance, 10 correspondences were manually picked by an expert and used to compute an optimal transformation under the L_2-loss. Reported results are compared to the rotation and the translation of this estimate. We have also selected two failure criteria based on these comparisons. The first being that the rotation error is larger than 5°, the second that the translation error is larger than 25 pixels. The percentage of results that fail according to these criteria are presented.

Table 1. The results for the H&E - p63/AMACR benchmark. In the left column, the inlier threshold ϵ is varied. Then, for each of the methods (RANSAC with varying number of iterations, and the truncated L_1- and L_2-norms), three numbers are reported: average rotation error (degrees), average translation error (pixels) and failure rate. A failure case is one with error in rotation larger than 5° or in translation larger than 25 pixels. When $\epsilon = \infty$, no truncation takes place.

	RANSAC									Truncated norms					
ϵ	100 iter.			500 iter.			1000 iter.			L_1			L_2		
1p	10.7°	204p	48%	8.38°	221p	27%	8.14°	105p	28%	2.72°	61p	11%	2.47°	58p	8%
5p	14.4°	280p	42%	2.85°	53p	9%	1.94°	28p	5%	1.21°	7p	3%	0.43°	6.4p	2%
10p	7.80°	158p	28%	1.23°	42p	6%	2.24°	43p	6%	0.29°	4.8p	1%	0.28°	4.6p	1%
20p	3.96°	78p	18%	2.43°	34p	8%	0.91°	23p	3%	0.27°	4.0p	0%	0.26°	3.9p	0%
∞		-			-			-		2.43°	6.5p	5%	6.54°	94p	69%

The experimental results on H&E-p63/AMACR are shown in Table 1. The most accurate results are obtained by the truncated L_2-method. Truncated L_1-norm performs poorly on the lowest threshold, but at more reasonable levels for this task performance is similar to truncated L_2. None of the methods based on RANSAC succeeds on all examples, although the accuracy is good at higher thresholds with 1000 iterations. We also note that regular L_1-norm (marked ∞) succeeds much more frequently than L_2-norm and with better accuracy than a majority of the RANSAC variants—on a dataset

with only 10% inliers on average. For the highest threshold level (20p), we have also performed a test exhaustively trying all possible hypotheses ransac migth get. The accuracy is slightly worse but comparable to the L_1- and L_2- truncated methods. However, as the time complexity is $\mathcal{O}(n^3)$, close to the complexity of truncated L_1 with more expensive operations and no fast rejection method, it is in practice as slow or slower as the L_1- method while having no theoretical guarantees.

Table 2. The results for the H&E - TRF benchmark. See Table 1 for explanation.

	RANSAC									Truncated norms					
ϵ	100 iter.			500 iter.			1000 iter.			L_1			L_2		
1p	2.53°	30.6p	6%	0.40°	6.2p	1%	0.31°	2.7p	0%	0.34°	2.9p	0%	0.31°	2.6p	0%
2p	2.27°	31.2p	5%	0.29°	2.6p	0%	0.27°	2.7p	1%	0.29°	2.6p	0%	0.28°	2.6p	0%
3p	1.75°	23.4p	4%	0.29°	2.7p	0%	0.29°	2.7p	0%	0.28°	2.6p	0%	0.28°	2.6p	0%
4p	0.66°	8.5p	3%	0.28°	2.6p	0%	0.28°	2.6p	0%	0.28°	2.6p	0%	0.27°	2.6p	0%
5p	1.14°	7.4p	2%	0.26°	2.5p	0%	0.26°	2.5p	0%	0.27°	2.5p	0%	0.26°	2.6p	0%
10p	0.76°	7.8p	1%	0.27°	2.4p	0%	0.27°	2.4p	0%	0.26°	2.5p	0%	0.26°	2.4p	0%
∞		-			-			-		15.6°	173p	57%	33.6°	341p	100%

Results from the benchmark experiment on H&E-TRF registration are shown in Table 2. This dataset has significantly fewer matches per image pair and higher inlier ratios, making it more suitable for RANSAC. With 1000 iterations, RANSAC performs on par with truncated L_1-norm and truncated L_2-norm, but with fewer iterations there are still some failures. The poor results for regular L_1-norm and L_2-norm show that for this task, aligning a sub-image to a larger image, using truncated norms is essential.

We also tested the intensity-based Image Regsitration Toolkit [23], using normalized mutual information. For the first dataset, the toolkit failed to produce a correct registration (less than 5° rotation and 25 pixels translation error) in 86% of the experiments. For the second dataset it failed to produce any corrrect results. The poor results are not surprising as these methods often are sensitive to initialization and to outlier structures in the images.

6.2 Registering CT to MR-Flair

This experiment is a demonstration of the applicability of the method. For more quantitative results, see Sections 6.1 and 6.3. The dataset consists of 44 image slices captured using the MR-Flair methodology and 4 image slices from a CT-scan of one single subject. To correlate the information provided by the different modalities, one would like to register each of the CT slices to the MR-Flair volume. As the CT slices are roughly aligned with the slices of the MR-Flair volume, we try to register each of the CT slices to each of the MR-Flair slices and then try to find the sequence of four MR-Flair slices that best match the four CT slices. We use standard 2D SIFT to obtain correspondences. To improve the matching performance, all descriptors were extracted at a fixed scale instead of using the estimated scale from the difference of Gaussians detector. The motivation is that in very noisy images the scale estimation tends to be uncertain.

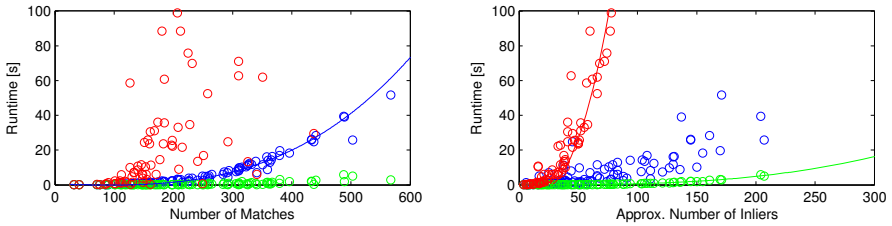

Fig. 4. *Left:* Runtime as a function of number of matches is graphed for truncated L_1-norm (green), truncated L_2-norm (red) and regular L_1-norm. The L_1-method follows closely a $\mathcal{O}(n^3)$-curve (blue). *Right:* Runtime as a function of number of inlier matches is graphed. The truncated methods are more correlated to the number of inliers, see the $\mathcal{O}(n^4)$-curve (red) and the $\mathcal{O}(n^3 \log(n))$-curve (green), respectively. ($\epsilon = 10$ pixels.)

Due to the small dataset we only present qualitative results for our truncated L_1-approach. Different slice-matches have different number of potential inliers, making the truncated L_1-cost skewed. However this is easily rectified by using a modified cost c_i defined as $c_i = N_i\epsilon - l_i^*$ where N_i is the combined number of correspondences for subsequence i, $epsilon$ the truncation level, and l_i^* the combined optimal truncated L_1-solution. Using this criterion and $\epsilon = 10$ the best subsequence evaluated at $c_{19} = 397$, with closest runners up $c_{18} = 362$, $c_{17} = 367$. All other sequence-matchings had significantly lower score. We show the found matchings for the best matching in Fig. 5.

The frequently used intensity-based method called NIFTYREG [21] using mutual information was also tested, but without any reasonable registration results at all. Note that this method was also developed to cope with outlier structures by using robust estimation techniques.

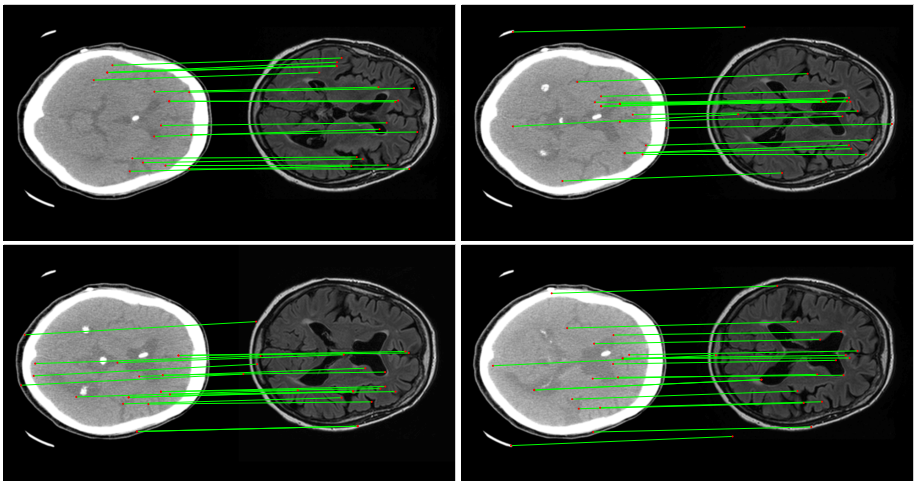

Fig. 5. The found inliers for the best subsequence obtained using our truncated L_1 algorithm

6.3 Speed

The theoretical worst time complexities are stated in Table 3. In practice RANSAC is not run exhaustively but with a fixed number of k iterations, giving a complexity of $\mathcal{O}(nk)$. For average-size problems (280 matches) and $k = 1000$, RANSAC required 73 ms. The fastest (but worst-performing) method is the closed-form L_2-method with a typical runtime of 0.2 ms. For the remaining methods timing plots are shown in Fig. 4. Because of the fast outlier rejection scheme discussed in Section 4.2, runtimes of the truncated L_1-norm and L_2-norm depend mainly on the size of the inlier sets. The full L_1-method has no such advantage. These numbers clearly show the advantage in runtime for the truncated L_1-method over both the regular L_1-norm and the truncated L_2-norm. However, on datasets consisting of a majority of inliers, the lower complexity of the L_1-norm would give faster runtimes as all operations are identical apart from the sorting strategies. The timing statistics is from experiments on H&E-TRF, though the same analysis holds for H&E-p63/AMACR.

Table 3. Characteristics of the algorithms presented or discussed in the paper. Note that the stated complexity for RANSAC is for exhaustive selection of all minimal subsets which can be thought of as a worst time complexity bound.

Algorithm	*complexity*	*tractability*	*reliability*	*reference*
RANSAC	$\mathcal{O}(n^3)$	high	medium	[12]
Truncated L_1-norm	$\mathcal{O}(n^3 \log(n))$	high	high	this paper
L_1-norm	$\mathcal{O}(n^3)$	high	medium	this paper
Truncated L_2-norm	$\mathcal{O}(n^4)$	medium	high	[1]
L_2-norm	$\mathcal{O}(n)$	high	low	[15]

7 Discussion

So what is the right way to attack feature-based image registration in presence of outliers? The literature provides us with a vast amount of choices, but many of these are based on local optimization and require a reasonable starting solution, which means

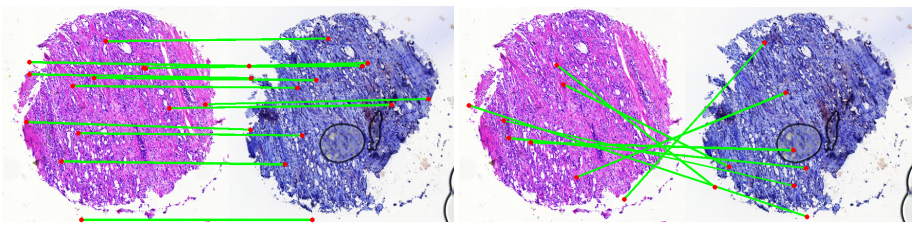

Fig. 6. *Left:* 13 inliers among 1179 hypothetical SIFT matches of the truncated L_1-method (*success*). *Right:* 8 inliers of RANSAC with 1000 iterations (*failure*). This was the hardest case to register among all pairs. ($\epsilon = 20$ pixels.)

that the outlier problem is already more-or-less solved. To handle really difficult outlier problems, RANSAC-type algorithms are the standard against which others are measured. However, as our experiments show, they are sub optimal both in terms of accuracy and with respect to the risk of failure. Some of the failures could be avoided by increasing the number of iterations - even up to exhaustively searching all the minimal subsets. But, that will increase the complexity to $\mathcal{O}(n^3)$, being practically the same as the algorithms proposed here (Table 3). More importantly, even then there is no guarantee as to the solution quality (Fig. 6). Hence, we would only recommend RANSAC when the amount of outliers is known to be low and the available runtime is very limited.

This contrasts sharply to the typical setting for medical image registration where the process is performed offline. With different image modalities, the rates of outliers are usually high. In these cases the increased reliability of optimizing a truncated norm is valuable and the L_1-based methods, although slower than RANSAC, should be efficient enough for most applications. Our experiments indicate only a small gain in accuracy for the truncated L_2-norm, so using the truncated L_1-norm would be the general recommendation.

In many applications, the actual improvement in terms of accuracy and failure rates of these methods might not be huge. This is compensated by the value of removing a possible error source and not having to tune the parameters of the algorithm. We believe that the choice between a tractable, reliable algorithm with guaranteed high-quality solutions and a fast algorithm with no guarantees whatsoever should be an easy one.

References

1. Ask, E., Enqvist, O., Kahl, F.: Optimal geometric fitting under the truncated L_2-norm. In: Conf. Computer Vision and Pattern Recognition, Portland, USA (2013)
2. Audette, M.A., Ferrie, F.P., Peters, T.M.: An algorithmic overview of surface registration techniques for medical imaging. Med. Image Anal. 4(3), 201–217 (2000)
3. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Analysis and Machine Intelligence 14(2), 239–256 (1992)
4. Blake, A., Zisserman, A.: Visual Reconstruction. MIT Press (1987)
5. Breuel, T.M.: Implementation techniques for geometric branch-and-bound matching methods. Computer Vision and Image Understanding 90(3), 258–294 (2003)
6. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Analysis and Machine Intelligence 23(6), 681–685 (2001)
7. Datteri, R.D., Dawant, B.M.: Estimation and reduction of target registration error. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 139–146. Springer, Heidelberg (2012)
8. Doyle, S., Madabhushi, A., Feldman, M., Tomaszeweski, J.: A boosting cascade for automated detection of prostate cancer from digitized histology. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 504–511. Springer, Heidelberg (2006)
9. Dreo, J., Nunes, J.C., Siarry, P.: Robust rigid registration of retinal angiograms through optimization. Comput. Med. Imaging Graph. 30(8), 453–463 (2006)
10. Enqvist, O., Ask, E., Kahl, F., Åström, K.: Robust fitting for multiple view geometry. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part I. LNCS, vol. 7572, pp. 738–751. Springer, Heidelberg (2012)

11. Enqvist, O., Josephson, K., Kahl, F.: Optimal correspondences from pairwise constraints. In: Int. Conf. Computer Vision, Kyoto, Japan (2009)
12. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting. Commun. Assoc. Comp. Mach. 24, 381–395 (1981)
13. Fitzgibbon, A.W.: Robust registration of 2D and 3D point sets. Image Vision Comput. 21, 1145–1153 (2003)
14. Fitzpatrick, J.M., West, J.B., Maurer, C.R.: Predicting error in rigid-body point-based registration. IEEE Trans. Medical Imaging 17(5), 694–702 (1998)
15. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternion. Journal of the Optical Society of America A 4(4) (1987)
16. Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., Forman, D.: Global cancer statistics. CA Cancer J. Clin. 61(2), 69–90 (2011)
17. Kwak, J.T., Hewitt, S.M., Sinha, S., Bhargava, R.: Multimodal microscopy for automated histologic analysis of prostate cancer. BMC Cancer 11(62) (2011)
18. Li, H.: Consensus set maximization with guaranteed global optimality for robust geometry estimation. In: Int. Conf. Computer Vision, Kyoto, Japan (2009)
19. Myronenko, A., Song, X.: Point-set registration: Coherent point drift. IEEE Trans. Pattern Analysis and Machine Intelligence 32(12), 2262–2275 (2010)
20. Nguyen, K., Sabata, B., Jain, A.K.: Prostate cancer grading: Gland segmentation and structural features. Pattern Recognition Letters 33(7), 951–961 (2012)
21. Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N.: Reconstructing a 3D structure from serial histological sections. Image Vision Comput. 19, 25–31 (2001)
22. Rangarajan, A., Chui, H., Mjolsness, E., Pappu, S., Davachi, L., Goldman-Rakic, P., Duncan, J.: A robust point-matching algorithm for autoradiograph alignment. Med. Image Anal. 1(4), 379–398 (1997)
23. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 32(1), 71–86 (1999)
24. Wikipedia, `http://www.wikipedia.org`