Point based registration of high-resolution histological slices

Peter Leškovský

peter.leskovsky@tu-berlin.de

Olaf Hellwich

http://www.cv.tu-berlin.de/

Computer Vision & Remote Sensing
Technical University Berlin
Berlin, DE

Abstract

Automatic registration methods for histological images are presented. Point based approaches with outlier filtering by RANSAC and image factorisation are investigated. The algorithms are tested on high-resolution biopsy images of the gastrointestinal tract. Quantitative evaluation is performed at different resolution levels. The results reveal the superiority of Harris feature points over SIFT features, in our setting, as well as high dependence of the presented approaches on the image resolution. At low resolutions good registration results are obtained using only raw feature matching and N-nearest neighbour filtering. Outlier filtering is necessary to obtain stable results at higher resolutions.

1   Introduction

In histological pathology, the tissue samples are cut into thin slices during the preparation phase. These are then individually mounted on glass slides. At this stage, the correspondence between slices is lost. Demanding manual shifts of the glass slides and visual verification at different magnifications are necessary in order to identify the corresponding cell groups in the subsequent slices. However, the examination across the specimen sections is unavoidable in pathology, *e.g.* when tracking the propagation of malignant tumours. In virtual microscopy the histological slides are digitised and stored as images. It therefore suggests possible automation to the time demanding manual alignment of different slices in that the correspondence information can be precomputed and stored. Navigating across the slices is then done only by a button click. The main challenge here, however, is the generation of dense correspondences across the tissue sections. In this work we therefore evaluate automatic registration algorithms optimized for histological images.

Before the tissue sample is cut, it is embedded in paraffin. This procedure limits the extent of deformations applied on the tissue when cut. Little stretching and bending is observed. The sections rather tear and break into small pieces which can shift arbitrary on the glass slide. Some histologists [4] suggest this understanding of the underlying deformations. We therefore implemented registration schemes which assume piecewise rigid transformation. Each local rigid transformation should best describe the correspondences between particular regions (*e.g.* torn pieces of tissue). Such approach should handle cases where tissue parts are independently shifted within one slice (see *Fig.* 1, left) and also be able to partially compensate for local deformations of elongated parts (see *Fig.* 1, right).
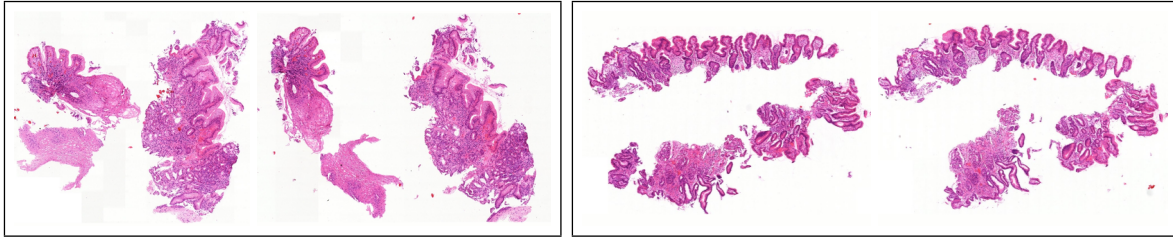
Figure 1: Examples of input images, presenting two cases with two slices each. The size of each slice image in the figure is approximately 1100×850 *px* at resolution level 4.

The registration can be driven by image or point based techniques (see *e.g.* [1, 4]). The former ones use differences in image intensities within overlapping regions to improve their alignment. They therefore require an initial overlap for the region of interest. No such initialisation is necessary for the point based methods. They align the images using salient point correspondences, robustly extracted from the images. However, these methods are sensitive to outliers present among the correspondences, arising during the matching of salient points.

In general it can not be assumed that a good initial alignment of all tissue parts, necessary for image based methods, will be possible. This is due to the arbitrary shifts of the torn tissue pieces. We therefore concentrate on point based methods. For these, it is important to select suitable feature detector, to decide on the resolution at which features should be extracted and to answer how or whether outliers in the correspondence set have to be detected.

2   Registration Methods

We experiment with Harris [6] and SIFT [3] feature detectors to extract salient points for the point based registration methods. While the first kind detects distinctive corner-like regions, the latter is sensitive to round, blob-like structures. In both cases, however, we use the 128 dimensional SIFT descriptor, captured at the most responsive scale. We obtain the initial correspondences similar as in [3], by selecting the nearest point in the descriptor's space and filtering the matches by thresholding the nearest neighbour distance ratio (NNDR). Note that the constant zoom at each resolution level allows us to restrict the matches only to descriptors of the same scale. At last, we enforce bijective matching between the point sets.

We apply two outlier removal methods: RANSAC [2] and image factorisation (IMF) [5]. In both cases, following the piecewise rigid transformation assumption, we expect to extract groups of correspondent points which define single rigid transformation (rotation and translation). These should trace the transformation of the torn pieces of tissue.

The RANSAC algorithm identifies the largest groups of pair correspondences which rigidly map onto each other. According to a specified transformation error, the correspondences are split into in- and out-liers. The outliers contain wrong matches along with true correspondences, describing the transformation between other tissue pieces. We thus recursively apply the RANSAC algorithm on the outliers until maximum admitted error is reached.

The IMF method extracts groups of rigidly matching point sets in one step. A mean shift algorithm in the transformation's space is used. As a set of basal rigid transformation functions, two correspondence pairs could be used to define the translation and rotation. Nevertheless, we obtained better results when estimating the basal transformations from a single correspondence pair, using the SIFT angles to define rotation.

Each algorithm outputs a group of rigid transformations, *i.e.* modes, including the corresponding point matches. These best represent the geometrical changes of the separate tissue parts. We further boost the number of obtained matches using the initial, not filtered point correspondences. We distribute them to the nearest mode, based on their transformation error. Finally, to obtain dense deformation field, we interpolate the mapping computed for each point match. For this we use distance weighted N-nearest neighbour based regularisation.

3 Evaluation

The evaluation has been run on a set of *ca.* 150 histological slices, containing biopsies of the gastrointestinal tract. Three dyes have been used to stain the samples: H&E, GI and PAS. Only intra-stain registrations of successively cut slices have been examined.

The samples were optically scanned at resolution of $0.23\mu m$ per pixel. After acquisition, an image pyramid of eight coarser resolutions was created by successively downsampling the images by a factor of two. The image levels range from values 0 (finest resolution) to level 7 where each pixel corresponds to a region of $2^7 \times 2^7$ pixels at the highest resolution.

We use point correspondences as ground truth (GT) data. We obtained them by manual selection at image resolution level 4. To increase the precision of the GT set, we run local image based registration at these points. Patches of size 50×50 pixels, at image level 4, were extracted around each correspondence pair for this purpose.

We have evaluated the basic point based registration method (see *Sec.* 2) with three outlier removal settings and two salient point detectors. The NNDR matching was used with none, RANSAC and IMF outlier removal methods. For each method, SIFT and Harris salient points have been tested. We refer to these settings as to NS, NH, RS, RH, FS and FH, defining the applied outlier and the feature detectors by the first and the second letter respectively.

The only parameters which have to be set for each method, are the matching ratio threshold for NNDR, the maximum distance error for RANSAC and the transformation kernel bandwidth for IMF. We obtained the best parameters for the methods empirically, tuning them for two representative virtual slides, shown in *Fig.* 1. The tuning has been performed at the resolution levels 5 and 6. Following these we set the NNDR to 0.8, the maximum distance error parameter to 0.4 and the kernel bandwidth to 0.04.
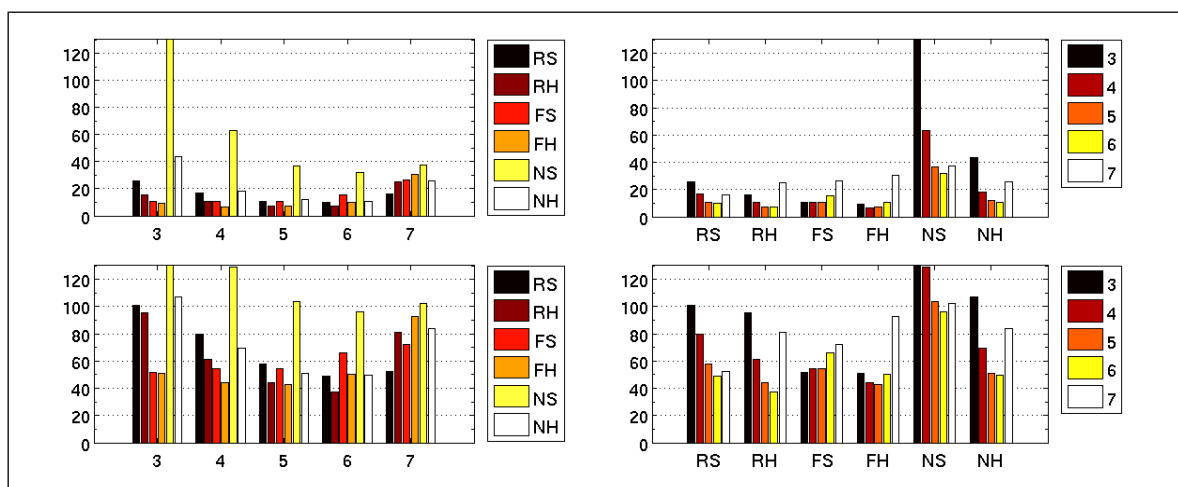

Figure 2: Evaluation results grouped by resolution level (left) and method (right). The mean (top) and the root mean squared (bottom) error, in pixels at resolution level 4, are plotted.

As the obtained parameters are specific to the deformation degree of the separate tissue pieces in world coordinates, no further tuning is theoretically necessary for other resolution levels. Using these parameters we run the evaluation on our set of histological slides, excluding those used for the parameter estimation. The results, running the algorithms at levels 3 to 7 are shown in *Fig.* 2. The registration quality was measured by the distance between the point correspondences and the GT. The mean and the root mean square error, in pixels at resolution level 4, is shown. The standard error of the mean is below 2 pixels for all methods.

3.1   Results

The quantitative evaluation (see *Fig.* 2) reveals that at coarser resolution levels (5 and 6) good results can be obtained with simple filtering and no outlier detection. This holds when Harris feature detector is used. In general, all methods score better at these resolutions, while little improvement can be observed at finer levels. This behaviour was expected. At coarse levels the features correspond to macroscopic structures which can be easily traced across slices. At fine resolution, cells become visible. The variation of the image gradients in small regions becomes higher and the SIFT descriptor not well discriminative. Many outliers arise from simple matching. This is also visible at the increasing error after, and including, level 4. The outlier detection algorithms can handle these difficulties to some extent. Nevertheless, after level 3 larger registration errors occur. Comparing the aligned images between different levels we found that at higher resolutions wrong transformation modes are detected. This also can be accredited to the increased amount of outliers. The results thus point out that at these levels the selected point descriptors are not optimal.

The superiority of Harris points is noticeable for all methods applied, through all levels. We observed that Harris features are more dispersed across the tissue, providing a dense tissue coverage. This is in contrast to SIFT features which tend to build clusters around salient structures and thus do not cover well all tissue parts.

At last, the RANSAC and the IMF performance is similar at levels 4 to 6. Only at level 3 large error variation for the RANSAC method is visible. We believe this problem stems from the inner parameters which inherently define maximum number of outliers.

In *Fig.* 3, registration results are shown for the methods FH, RH and NH. Good alignment of all tissue pieces was achieved by the FH method. The RH algorithm did not find enough correspondences for two tissue parts which are not matched. We believe this is due to non optimal parameter setting of the method. Following the FH method we know that correspondence groups for these tissue parts exist, they only were not identified by the RH algorithm. We observed similar results for lower resolution images, where few correspondences are detected. At last, the influence of the mismatches is visible for the NH method.
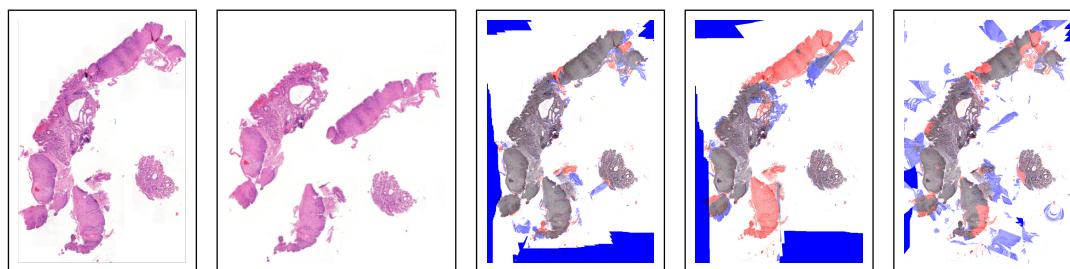

Figure 3: Registration results in difficult scenario. Two slices followed by alignments given by three methods (FH, RH and NH) are shown. The process was run at image level 4.

Although the registration result is overall worse than for the FH method, small regions exist where better alignment was achieved with this simple matching technique. It is therefore clear that the outlier detection algorithms apply strong filtering on the correspondence set. This can be finally relaxed, in order to increase the number of correctly detected matches.

4 Discussion and conclusions

We examined point based matching method for the registration of histological images. It was not clear whether general purpose point descriptors can provide enough discriminative power in order to be applied in this specific task. Our results show that salient feature detectors, like Harris and SIFT, combined with the SIFT descriptor are well adaptable for this purpose. In this context we identified Harris based corner features more representative for the tissue samples than the blob like SIFT features. Nevertheless, both of them should be applied only up to some resolution level. This is the level at which individual cells are small and do not influence much the variance of the descriptor. Going for finer resolutions, many matching faults will arise from the high similarity between all small regions.

We also evaluated two techniques which can be applied for filtering the outliers. They greatly improve the registration results when good parametrisation is chosen. How to obtain the best parameters, whether automatically or using a minimal user input, will be the point of our next research.

Acknowledgement

References

[1] L. Cooper, O. Sertel, J. Kong, et al. Feature-based registration of histopathology images with different stains: An application for computerized follicular lymphoma prognosis. *Computer Methods and Programs in Biomedicine*, 96(3):182 – 192, 2009.

[2] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24:381–395, 1981.

[3] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[4] A. Pitiot, E. Bardinet, P. M. Thompson, et al. Piecewise affine registration of biological images for volume reconstruction. *Medical Image Analysis*, 10(3):465 – 483, 2006.

[5] R. Subbarao and P. Meer. Nonlinear mean shift over Riemannian manifolds. *International Journal of Computer Vision*, 84:1–20, 2009.

[6] T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: a survey. *Found. Trends. Comput. Graph. Vis.*, 3:177–280, 2008.