Benchmarking of image registration methods for differently stained histological slides

Jiří Borovec* Arrate Muñoz-Barrutia†‡ Jan Kybic*

* Department of Cybernetics, Faculty of Electrical Engineering, CTU in Prague, Czechia
† Instituto de Investigación Sanitaria Gregorio Maranón, Madrid, Spain ‡ Universidad Carlos III de Madrid, Leganés, Spain

jiri.borovec@fel.cvut.cz http://cmp.felk.cvut.cz/∼borovji3/

Introduction

Image registration is a common task for many biomedical analysis applications. The present work focuses on the benchmarking of registration methods on differently stained histological slides. This is a challenging task due to the differences in the appearance model, the repetitive texture of the details and the large image size, between other issues. Our benchmarking data is composed of 616 image pairs at two different scales — average image diagonal 2.4k and 5k pixels. We compare eleven fully automatic registration methods covering the widely used similarity measures. For each method, the best parameter configuration is found and subsequently applied to all the image pairs. The performance of the algorithms is evaluated from several perspectives — the registrations (in)accuracy on manually annotated landmarks, the method robustness and its computation time.

Materials

Whole slide microscopy images

- 32 sets of consecutive sections — breast tumor, lung tumor, rat kidney
- stained by a different dye — Cytokeratin, CC10, proSPC, H&E, Ki67, CD31, CNEU, ER, PR, Podocin, Negative

Evaluation measures

- **Target Registration Error (TRE)** is mean Euclidean distances in pixels between the positions of the landmarks in the reference and the transformed images normalized by the size of the image diagonal.
- **Robustness** is the ratio of cases where registration improved the initial relative TRE.
- **Execution time** is measured on a computer using a single CPU/thread.

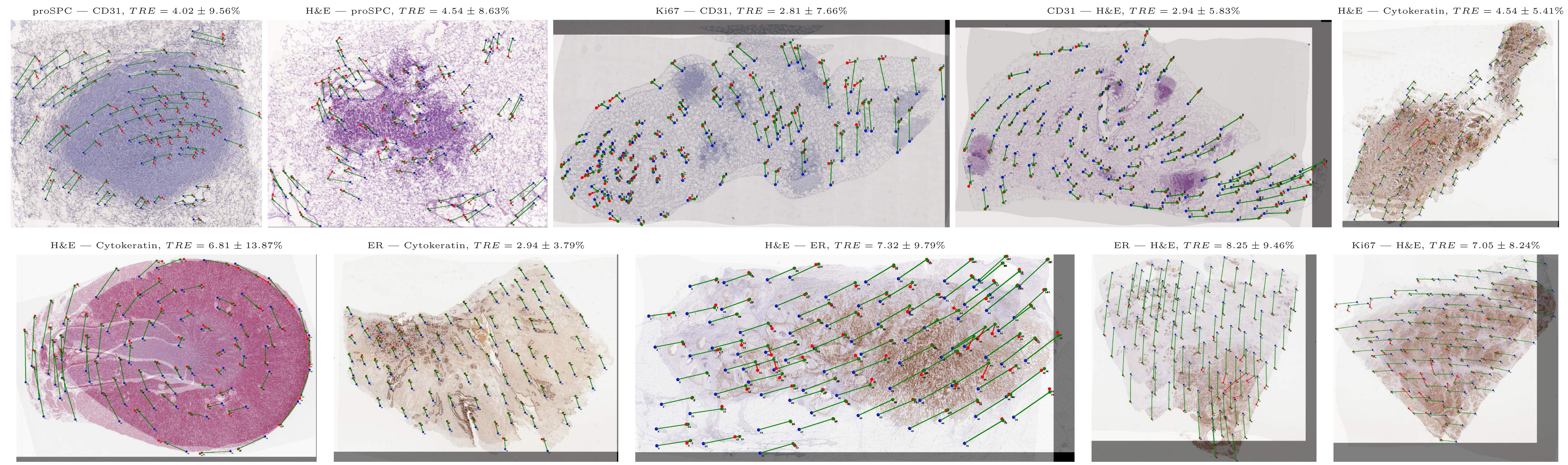
Numerical results

Mean TRE before registration $4.25\% \pm 5.98\%$, median=2.79%. Mean TRE between experts 0.4%.

	Methods	Scale	Linear robust. [%]	Linear TRE [%] mean	Linear TRE [%] median	Linear time [s] mean	Linear time [s] median	Elastic robust. [%]	Elastic TRE [%] mean	Elastic TRE [%] median	Elastic time [s] mean	Elastic time [s] median
Features	OpenCV	S	**96.92**	0.83	***0.28***	52.85	16.93	-	-	-	-	-
		M	***96.92***	0.91	***0.27***	247.3	101.1	-	-	-	-	-
	TrakEM2 [1]	S	93.67	***0.72***	0.34	**6.13**	***5.72***	93.83	2.98	1.52	24.92	27.07
		M	**97.56**	**0.66**	0.30	**11.5**	**11.41**	91.40	3.36	2.38	***49.54***	***50.52***
Similarity	Elastix [2]	S	***96.27***	1.43	0.30	664.9	674.0	**97.56**	2.23	0.42	1127	1140
		M	94.64	2.89	1.95	787.9	782.0	***96.27***	3.21	1.71	1247	1246
	ANTs [3]	S	**96.92**	1.07	**0.25**	26.67	24.80	***96.75***	1.00	**0.22**	72.95	69.53
		M	96.27	1.05	**0.24**	128.5	126.6	**96.59**	***1.08***	**0.22**	434.4	417.2
	NiftyReg [4]	S	79.22	2.04	1.25	***6.48***	**2.70**	70.13	1.61	0.81	**5.92**	**3.65**
		M	72.40	2.86	1.92	***28.72***	***29.0***	58.12	2.21	1.67	**17.55**	**16.57**
	bUnwarpJ [5]	S	-	-	-	-	-	79.06	1.52	0.58	125.7	91.71
		M	-	-	-	-	-	80.84	2.68	1.54	686.9	445.8
Hybrid	OpenCV + Elastix [6]	S	***96.27***	**0.67**	0.29	290.6	172.7	94.64	**0.75**	0.48	325.1	198.6
		M	96.75	***0.76***	0.29	575.7	394.1	94.38	**0.85**	0.45	15397	5350
	DROP [7]	S	95.62	1.12	0.52	67.49	62.54	92.26	***0.86***	***0.32***	299.4	137.4
		M	96.27	1.12	0.51	303.9	270.6	54.81	1.17	0.62	17179	13095
	RVSS [5]	S	-	-	-	-	-	56.53	3.04	2.37	31.81	27.46
		M	-	-	-	-	-	63.34	2.66	2.04	153.1	100.4
Labels	ASSAR [8]	S	88.31	0.90	0.66	105.1	103.5	83.28	0.94	0.55	210.7	207.7
		M	86.69	1.04	0.71	309.9	275.0	81.66	1.16	0.76	605.9	558.0
	SegReg [9]	S	92.37	1.56	0.44	16.09	14.97	93.34	1.34	0.41	***16.21***	***14.99***
		M	89.94	1.40	0.45	67.03	57.65	90.26	1.41	***0.42***	75.88	66.85

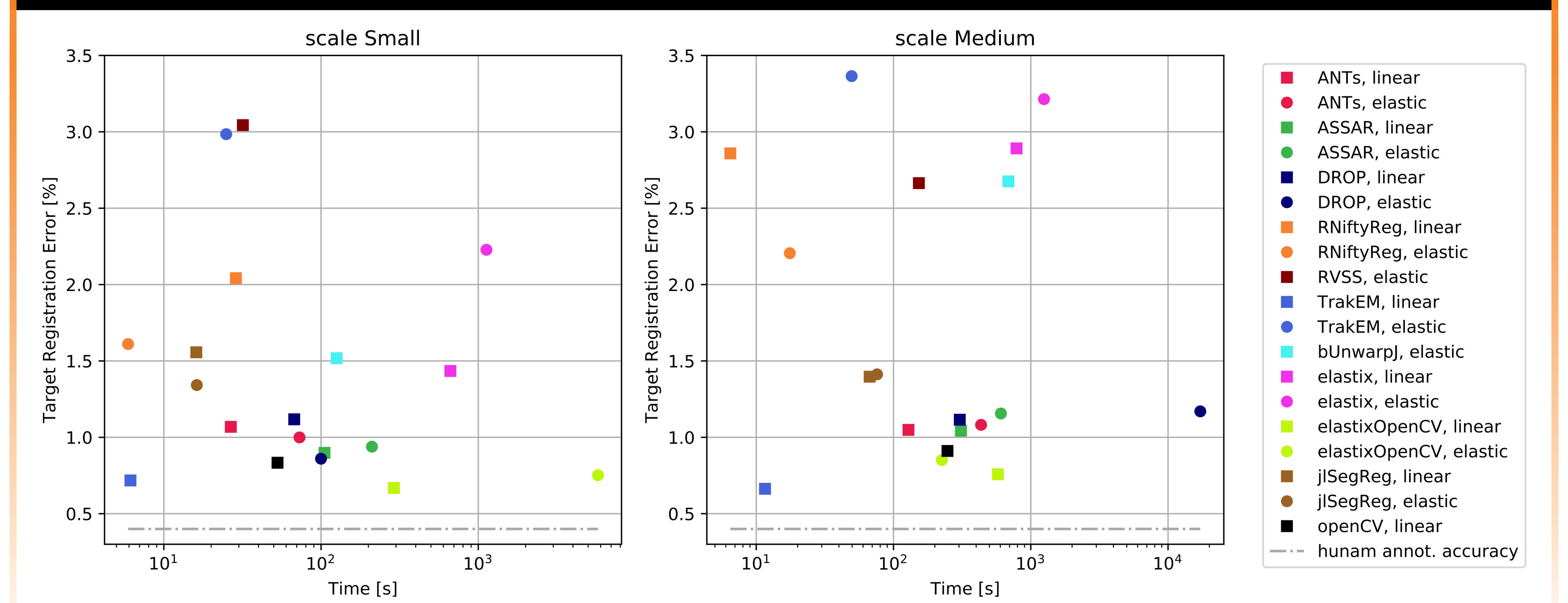
Experimental results on both small $[2k \times 2k]$ (S) and medium $[4k \times 4k]$ (M) datasets and for linear and elastic (free-form) transformations. Some implementations do not support both transformations (denoted by '-'). We mark best scores (**first** and ***second***) for each metric and scale.

Illustration of registered image pairs



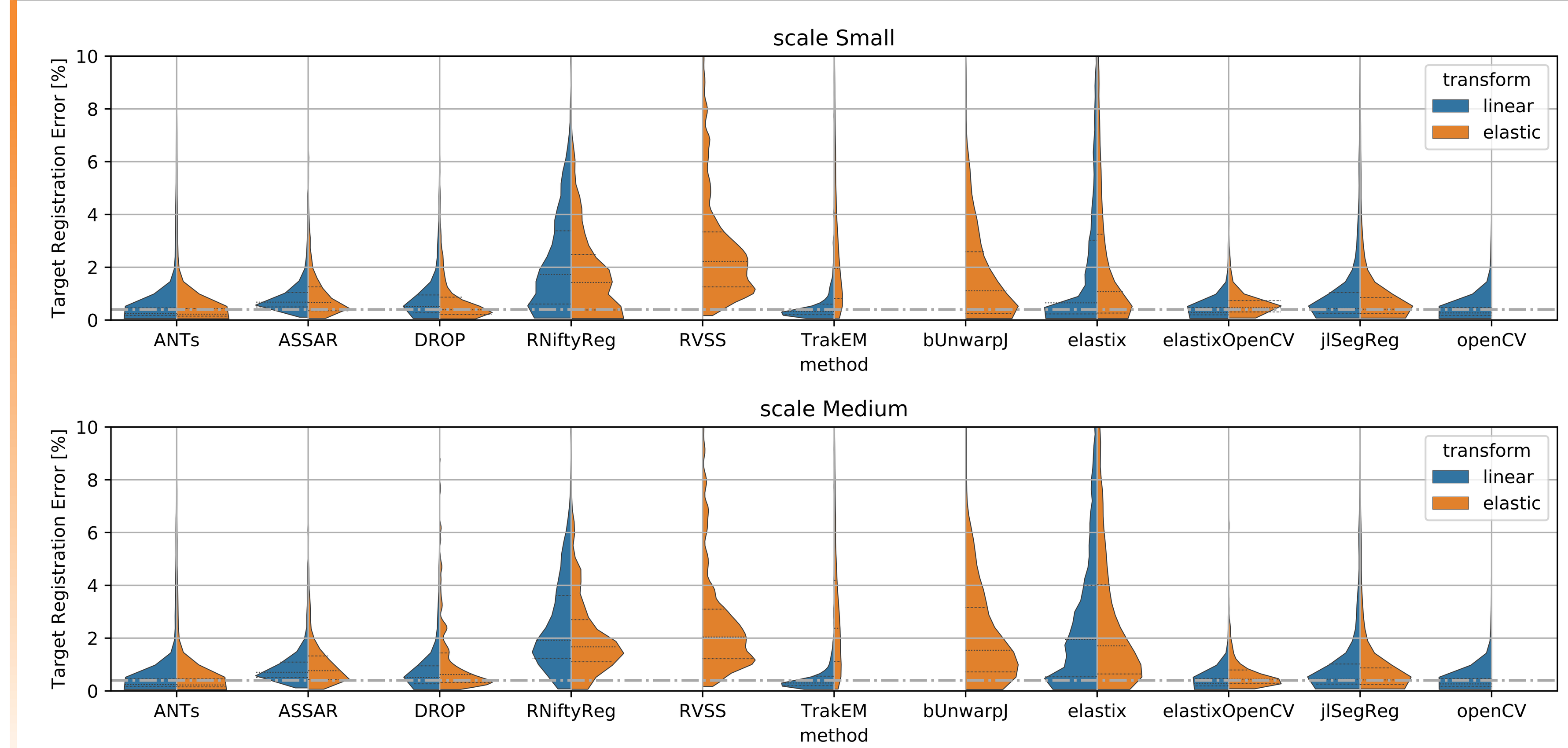
Overlap of moving and warped image after registration with marked landmarks: moving—initial (blue), reference—target (red) and warped (green). The red line represents the registration error.

Relation between TRE and execution time



Dependency of time to TRE for all methods and evaluated with respect to used transformation.

Distributions of TRE



Distribution of TRE for all methods with respect to used transformation.

Configuration of registration methods

type	Method	Criterion	Optimization
Features	OpenCV	SURF & MSER	RANSAC
	TrakEM2 [1]	SIFT	RANSAC
Similarity	Elastix [2]	MMI	L-BFGS
	ANTs [3]	MMI / CC	LPF
	NiftyReg [4]	NMI	conjug. gradient
	bUnwarpJ [5]	SSD	LM + BFGS
Hybrid	OpenCV + Elastix [2]	SURF & MSER + MMI	RANSAC+ L-BFGS
	RVSS [5]	SIFT + SSD	LM
	DROP [7]	Gabor + SAR	linear prog. MRF
Labels	ASSAR [8]	MIL	BP
	SegReg [9]	MIL	BOBYQUA

Similarity: Intensity-based registration methods; **Features**: Feature-based registration methods; **Hybrid**: Feature and intensity-based registration methods; **Labels**: Segmentation based methods.

Conclusion

- Selected registration methods cover the most common similarity criteria.
- Execution time of some methods is reasonably good (suitable for practical usage).
- Performances as measured show that the task is still not fully solved.
- In many cases the liner transformation performs better then the elastic.
- Image dataset — http://cmp.felk.cvut.cz/˜borovji3/?page=dataset

References

[1] A. Cardona, S. Saalfeld, J. Schindelin, and Al. TrakEM2 software for neural circuit reconstruction. *PloS one*, 7(6):e38011, jan 2012.
[2] S. Klein, M. Staring, and K. Murphy. Elastix: a toolbox for intensity-based medical image registration. *Medical Imaging, IEEE*, 29(1), 2010.
[3] B.B. Avants, C.L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, 2008.
[4] M. Modat, G. Ridgway, Z. Taylor, and Al. Fast Free-Form Deformation Using Graphics Processing Units. *Computer Methods and Programs in Biomedicine*, 98(3):278–284, 2010.
[5] I. Arganda-Carreras, C. Sorzano, R. Marabini, and Al. Consistent and elastic registration of histological sections using vector-spline regularization. In *Computer Vision Approaches to Medical Image Analysis*, volume 4241, pages 85—-95, 2006.
[6] J. Borovec, J. Kybic, and Al. Registration of multiple stained histological sections. In *International Symposium on Biomedical Imaging, IEEE*, pages 1034–1037, San Francisco, 2013.
[7] B. Glocker, A. Sotiras, and Al. Deformable Medical Image Registration: Setting the State of the Art with Discrete Methods. *Annual Review of Biomedical Engineering*, 13(1):219–244, 2011.
[8] J. Kybic and J. Borovec. Automatic simultaneous segmentation and fast registration of histological images. In *International Symposium on Biomedical Imaging, IEEE*, pages 774 – 777, 2014.
[9] J. Kybic, M. Dolejsi, and J. Borovec. Fast registration of segmented images by normal sampling. In *Bio Image Computing (BIC) workshop at CVPR*, pages 11–19, 2015.