Fast Registration by Boundary Sampling and Linear Programming

Jan Kybic$^{(\boxtimes)}$ and Jiří Borovec

Faculty of Electrical Engineering, Czech Technical University in Prague,
Prague, Czech Republic
kybic@fel.cvut.cz, jiri.borovec@fel.cvut.cz

Abstract. We address the problem of image registration when speed is more important than accuracy. We present a series of simplification and approximations applicable to almost any pixel-based image similarity criterion. We first sample the image at a set of sparse keypoints in a direction normal to image edges and then create a piecewise linear convex approximation of the individual contributions. We obtain a linear program for which a global optimum can be found very quickly by standard algorithms. The linear program formulation also allows for an easy addition of regularization and trust-region bounds. We have tested the approach for affine and B-spline transformation representation but any linear model can be used. Larger deformations can be handled by multiresolution. We show that our method is much faster than pixel-based registration, with only a small loss of accuracy. In comparison to standard keypoint based registration, our method is applicable even if individual keypoints cannot be reliably identified and matched.

Keywords: Image registration · Linear programming

1 Introduction

Image registration [1] is one of the key image analysis tasks, especially in medical imaging. There are many scenarios, when image registration needs to be fast — consider matching preoperative and intraoperative images during surgery, interactive change detection of CT or MRI data for a busy radiologist, deformation compensation or 3D alignment of large histological slices for a busy pathologist, or processing large amounts of images from today's high-throughput imaging methods. On the other hand, sub-pixel or even pixel-level accuracy is not always required, which gives us the possibility to trade accuracy for speed. In this work, we shall present such a method.

Feature-based methods (e.g. [2]) are not always suitable for biomedical images, since there are few reliable and distinguishable features (e.g. corners) and weak constraints on the deformation field. The other class of registration

methods, based on minimizing a pixel-based image similarity criterion, are often slow. Luckily, it turns out that criterion evaluation can be simplified, without compromising registration accuracy too much. Only a subset of the pixels can be used to evaluate the criterion [3] or its gradient [4], possibly with more weights given to pixels with a high gradient [5]. In the extreme, only edge pixels would be sampled, which leads to the idea of registering images by their segmentations, which can be done e.g. by descriptor matching [6], or region boundary matching [7].

We assume that a correspondences can be found reliably between regions but not between points in their interior. We represent the criterion as a sum of contributions from a set of sampling points placed sparsely on the region boundaries, as in [8]. Our main new insight is that if the region boundary moves in the normal direction, the criterion change is piecewise linear. This allows to formulate the optimization as a linear program (LP), which can be solved very efficiently. We present two variants of our registration method: LPSEG based on a segmentation, and LPNOSEG based on points of high gradient.

1.1   Other Related Work

Taylor and Bhusnurmath [9] create a global piecewise linear approximation of the similarity criterion, considering all pixels and all possible displacements, which is a lower bound of the true criterion. This is robust but slow. Ben-Ezra et al. [10] start by selecting a small set of keypoints using optical flow and a motion model is fitted using linear programming. This method is capable of running at several images per second but requires the motion to be small and the motion model to be simple. Linear programming can also be used for keypoint matching [11].

2   Method

2.1   Similarity Criterion

A pixel dissimilarity measure ϱ induces an image dissimilarity criterion

$$J_c(f, g) = \int_{\mathbf{x} \in \Omega} \varrho\big(f(\mathbf{x}), g(\mathbf{x})\big) \, d\mathbf{x} \tag{1}$$

where $\Omega \subseteq \mathbb{R}^d$ is the image domain and f, g can be intensities, but also texture features, or segmentation labels. This formulation includes directly criteria such as SSD or SAD, while other popular dissimilarity measures such as mutual information can be represented approximately, with ϱ depending on the images and updated occasionally. During registration, we calculate the criterion $J_c(f, g')$ between a reference image f and a transformed version $g'(\mathbf{x}) = \big(g \circ T\big)(\mathbf{x}) = g\big(T(\mathbf{x})\big)$ of the moving image g. Following [8], we assume to be given a set of M points $\mathbf{p}_i$ on the boundaries between regions in the image f, with normals $\mathbf{n}_i$. We are also given an a priori estimate T_0 of T, which is
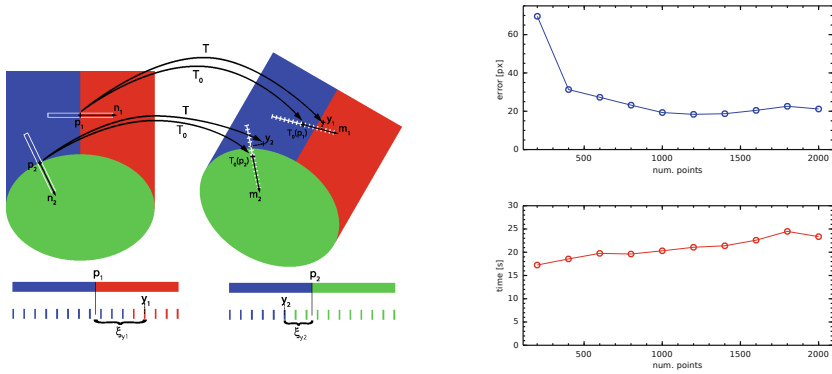
Fig. 1. *Left:* Illustration of the sampling points $\mathbf{p}_i$, their normals $\mathbf{n}_i$, and transformations T_0 and T. *Right:* The dependency of the registration error *(top graph)* and time *(bottom graph)* on the number of sampling points for LPNOSEG.
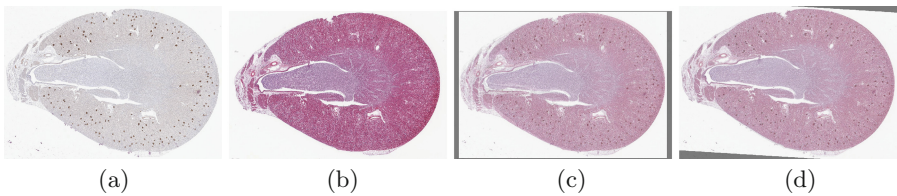

Fig. 2. Rat-kidney. Reference image (H&E) *(a)*, moving image (Podocin) *(b)*, overlay before *(c)* and after *(d)* — look e.g. at the top edge for differences.

locally close to linear. If $T_0 \approx T$, then in g there is also a boundary close to a point $\mathbf{v}_i = T_0(\mathbf{p}_i)$ with a scaled normal $\mathbf{m}_i = \big(\nabla T_0(\mathbf{p}_i)\big)\mathbf{n}_i$. The displacement along the boundary can be neglected, since both f and g are supposed to change only in the normal direction. Hence, the transformation T can be approximated by its normal projection ξ_i at points $\mathbf{p}_i$ (see Fig. 1, left)

$$T(\mathbf{p}_i) \approx \mathbf{v}_i + \mathbf{m}_i \xi_i \tag{2}$$

$$\text{where} \qquad \xi_i = \big\langle T(\mathbf{p}_i) - \mathbf{v}_i, \tilde{\mathbf{m}}_i \big\rangle \quad \text{with} \quad \tilde{\mathbf{m}}_i = \mathbf{m}_i \, / \, \|\mathbf{m}_i\|^2 \tag{3}$$

and the criterion J_c can be approximated as a sum of contributions at $\mathbf{p}_i$,

$$J_c(f, g') \approx J(T) + \text{const}, \qquad \text{where} \quad J(T) = \sum_{i=1}^{M} D_i(\xi_i) \tag{4}$$

The individual contributions are calculated by integration along the normals

$$D_i(\xi) = \sigma_i \int_{-h_{\max}}^{h_{\max}} \varrho\big(f(\mathbf{x}), g'(\mathbf{x})\big)\mathrm{d}h \quad \text{with} \quad \mathbf{x} = \mathbf{p}_i + \mathbf{n}_i h \tag{5}$$

with σ_i being the area corresponding to $\mathbf{p}_i$ and $h_{\max}$ the region width.

In the NLSEG method, a simple greedy strategy can provide sampling points $\mathbf{p}_i$ on the class boundaries and their normals [8], given a pixel- or superpixel-based segmentations. The NLNOSEG method differs in choosing points of high gradient, pruned with non-maxima suppression, and with $\mathbf{n}_i$ being the gradient direction. See Fig. 3ab for examples.

2.2 Piecewise Linear Approximation

Using the formula $T(\mathbf{p}_i + \mathbf{n}h) \approx \mathbf{v}_i + \mathbf{m}_i(\xi_i + h)$, the integral (5) can be approximated by a sum over h, sampling at 1 pixel intervals:

$$D_i(\xi) = \sigma_i \sum_{h=-h_{\max}}^{h_{\max}} \varrho\left(f(\mathbf{p}_i + \mathbf{n}_i h), g(\mathbf{v}_i + \mathbf{m}_i(\xi_i + h))\right) \tag{6}$$

If $\mathbf{p}_i$ is at a boundary of a segmentation f, then $f(\mathbf{p}_i + \mathbf{n}_i h)$ is going to be equal to some f^- for $h < 0$ and otherwise to f^+. Similarly, we assume that $g(\mathbf{v}_i + \mathbf{m}_i h)$ is equal to some g^- for $h < \zeta_i$ and otherwise to g^+, where ζ_i is the unknown normal shift at $\mathbf{p}_i$ due to the difference between T_0 and the true transformation.

The (continuous) contribution $D_i(\xi_t)$ is a convolution of these step functions:

$$D_i(\xi_i) = \begin{cases} u_i^0 + h_{\max} u_i^+ & \text{if } h_{\max} \leq t \\ u_i^0 + t u_i^+ & \text{if } 0 \leq t \leq h_{\max} \\ u_i^0 + t u_i^- & \text{if } -h_{\max} \leq t \leq 0 \\ u_i^0 - h_{\max} u_i^- & \text{if } t \leq -h_{\max} \end{cases} \tag{7}$$

$$\text{with} \quad t = \zeta_i - \xi_h, \qquad u_i^0 = h_{\max}\left(\varrho(f^-, g^-) + \varrho(f^+, g^+)\right)$$

$$u_i^+ = \varrho(f^+, g^-) - \varrho(f^+, g^+), \qquad u_i^- = \varrho(f^-, g^-) - \varrho(f^-, g^+)$$

For $|t| > h_{\max}$, the contribution $D_i(\xi_i)$ is constant and does not bring any information. This is avoided by choosing a suitable $\xi_{\max}$, ensuring that $D_i(\xi_i)$ is only evaluated for $|\xi_i| \leq \xi_{\max}$, assuming that the true shift also satisfies $|\zeta_i| \leq \xi_{\max}$ and choosing $h_{\max} \geq 2\xi_{\max}$.

Since the parameters in (7) are in general not known, we shall estimate them as follows: Evaluate $D_i(\xi)$ for all shifts $\xi = -\xi_{\max} \ldots \xi_{\max}$ using (6). Keeping the minimum and find the slope by least-squares fitting:

$$\hat{\zeta}_i = \arg\min_\xi D_i(\xi), \qquad \hat{u}_0 = D_i(\hat{\zeta}_i) \tag{8}$$

$$\hat{u}_i^+ = \arg\min_{u_i^+} \sum_{\xi=\zeta_i+1}^{\xi_{\max}} \left(D_i(\xi) - \hat{D}_i(\xi)\right)^2 = \frac{\sum_\xi \left(D_i(\xi) - \hat{u_0}\right)(\xi - \hat{\zeta}_i)}{\sum_\xi (\xi - \hat{\zeta}_i)^2}$$

and similarly for $\hat{u}_i^-$. This is both faster and more robust than fitting all four parameters. Note that the value of u_i^0 is not needed and can be dropped.

2.3   Geometric Model and Regularization

A geometrical transformation $T : \mathbb{R}^d \to \mathbb{R}^d$ is represented as a linear combination

$$T(\mathbf{x}) = \varphi_0(\mathbf{x}) + \sum_{j=1}^{N} c_j \varphi_j(\mathbf{x}) \tag{9}$$

with some basis functions φ_j and N scalar coefficients c_j. This includes many practically used transformation functions, such as an affine transformation, radial basis functions etc. For nonlinear (elastic) transformations, we shall use uniformly spaced B-splines [12], The image registration problem is then:

$$\mathbf{c}^* = \arg\min E(\mathbf{c}) \tag{10}$$
$$E(\mathbf{c}) = J(\mathbf{c}) + R(\mathbf{c}) \tag{11}$$

with the data criterion $J(\mathbf{c}) = J\big(T(\mathbf{c})\big)$ defined by (4) and (9). A regularization R is needed because the image is usually not completely covered by the sampling points $\mathbf{u}_i$ and some coefficients c_j might therefore not be completely determined. We have chosen to penalize the ℓ_1 norms of the coefficients and their first-order finite differences along coordinate axes [13]:

$$R(\mathbf{c}) = \gamma \left| \Delta\mathbf{c} \right|_{\ell_1} + \lambda \left| \mathbf{c} \right|_{\ell_1} \tag{12}$$

2.4   Linear Program

We can now proceed to formulate the linear program, which will help us recover the optimal transformation parameters c_j, given the contribution approximation parameters $\hat{u}_i^+$, $\hat{u}_i^-$, $\hat{\zeta}_i$ (8). The criterion (11) is written as

$$\min \left[\sum_{i=1}^{M} D_i + \gamma \sum_{(j,k)\in\mathcal{N}} r_{jk} + \lambda \sum_{j=1}^{N} s_j \right] \tag{13}$$

where the absolute values from (12) were replaced using inequalities

$$c_j - c_k \leq r_{jk}, \quad c_j - c_k \leq r_{jk}, \qquad c_j \leq s_j, \quad -c_j \leq s_j \tag{14}$$

where (j,k) are pairs of indices of neighboring B-spline coefficients c. Similarly, we replace the piecewise linear model (7) by

$$D_i \geq (\xi_i - \xi_i^0)u_i^+ \qquad\qquad \xi_i \leq \xi_{\max} \tag{15}$$
$$D_i \geq (\xi_i - \xi_i^0)u_i^- \qquad\qquad -\xi_i \leq \xi_{\max} \tag{16}$$

It remains to eliminate ξ_i from the above equations, using the linear relationship between the shifts ξ_i and transformation coefficients c_j from (3) and (9)

$$\xi_i = \left\langle \varphi_0(\mathbf{x}_i) - \mathbf{v}_i, \tilde{\mathbf{m}}_i \right\rangle + \sum_{j=1}^{N} c_j \left\langle \varphi_j(\mathbf{x}_i), \tilde{\mathbf{m}}_i \right\rangle \tag{17}$$

The resulting LP with approximately $(2+d)N + M$ variables (neglecting the edge cases), namely D_i, c_j, r_{jk} and s_j, is solved using the simplex or interior point methods.

Table 1. *Left:* Average runtime in seconds and geometric error in pixels. *Right:* Time for solving one LP instance with 1200 sampling points and 128 B-spline coefficients for different LP solvers.

Method	Time	Error
bUnwarpJ [14]	676	129
elastix (B-splines) [15]	1286	125
FRSEG [8]	105	18
RVSS [14]	91	151
RNiftyReg (GPU)	22	67
DROP [13]	403	27
LPSEG (new)	36	50
LPNOSEG (new)	20	19

LP method	time
Gurobi simplex	54 ms
Gurobi interior point	79 ms
GLPK simplex	145 ms
GLPK interior point	failed

2.5 Iteration and Multiresolution

Iteration: When the LP is solved, the resulting sampling point positions $T(\mathbf{p}_i)$ are compared with the sampled positions used in (6). If more than a given percentage (e.g. 10 %) of the points are considered unacceptable, the procedure is repeated with the initial transformation T_0 replaced by T. A point $T(\mathbf{p}_i)$ is unacceptable if its associated shift ξ_i is close to the assumed maximum amplitude $\xi_{\max}$ (usually 5–10 pixels) or if the angular difference with $T_0(\mathbf{p}_i)$ is larger than a given threshold (e.g. 60°).

Multiresolution: At each level, the image size is reduced by a factor of two and the resulting transformation is used as the initial transformation T_0 at the next finer level. The number of sampling points is also reduced at each level by taking every second point, unless a desired minimum number of points is achieved. The final multiresolution aspect is that we usually start by an affine transformation and progressively shift to B-spline models with more and more parameters. This is controlled by the desired knot spacing.

3   Experiments

The first example in Fig. 2 shows histological slices and their alignment by the LPSEG algorithm. For these images of $\approx 800 \times 1100$ pixels, the complete registration process takes about 2 s. However, most of the time is taken by the segmentation; once sampling points are extracted, the registration itself takes only 0.2 s.

Images in Fig. 3 are already much larger ($\approx 5000 \times 3000$ pixels) and the running time of the LPSEG method has grown to 36 s, with 33 s taken by the segmentation and only 3 s by the registration itself. The segmentation-less LPNOSEG is faster, requiring around 20 s, with 15 s to identify the sampling points and 5 s for the registration itself. In this case, the registration by LPNOSEG is better (see the bottom edge of Fig. 3hi).
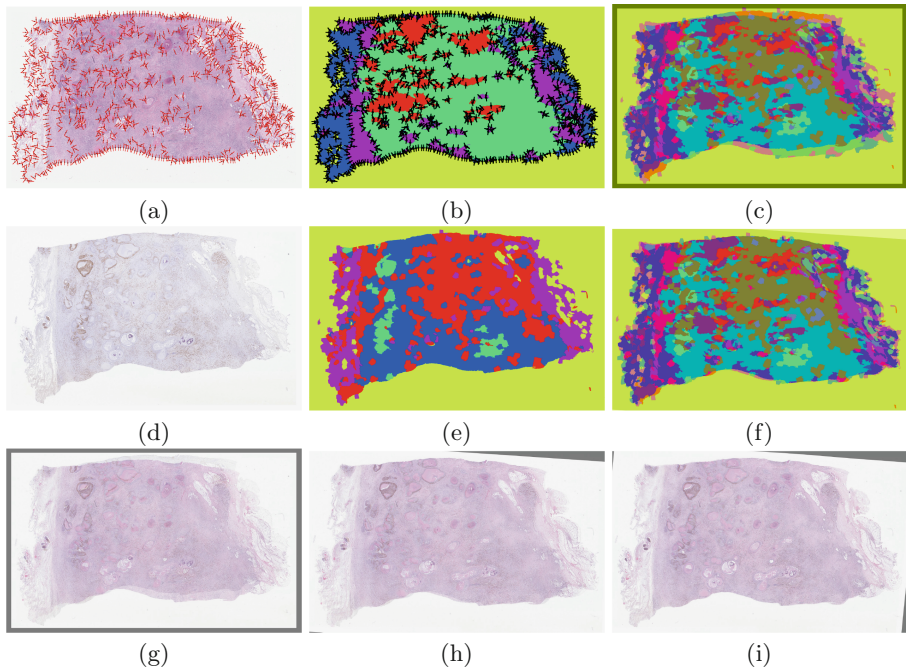
Fig. 3. Reference image with sampling points *(a)* and its segmentation *(b)* with independently extracted sampling points, segmentation overlay before registration *(c)*. Moving image *(d)*, its segmentation *(e)*, and segmentation overlay after registration *(f)*. Image overlay before registration *(g)*, image overlay after LPSEG registration using segmentation *(h)* and using LPNOSEG without segmentation *(i)*.

The graphs in the right part of Fig. 1 show the dependency of the running time and the mean registration error (measured using manually identified landmarks [8]) on the number of sampling points for the LPNOSEG algorithm. Note that the error decreases with an increasing number of sampling points and then it stagnates, as further points are probably not sufficiently reliable or accurate. The running time increases with the number of points but only slowly, the dominant part seems to be the preprocessing.

Table 1 (right) compares the GLPK[1] and Gurobi[2] LP solvers. Finally, the left part of Table 1 compares the speed and accuracy of our two proposed methods (LPSEG, LPNOSEG) with the alternatives mentioned in [8] on a subset of the same dataset with images of size around 5000×3000 (as in Fig. 3). FRSEG [8] (fast registration of segmented images) is similar to LPSEG but does not use the LP formulation. We see that LPSEG is fast and LPNOSEG even faster. The only other method with comparable speed is the GPU-accelerated version

---

[1] https://www.gnu.org/software/glpk/.
[2] http://www.gurobi.com.

of RNiftyReg. At the same time, our methods are accurate, outperformed only by the slower FRSEG, which is based on the same ideas.

4   Conclusions

We have presented two very fast image registration methods based on simplifying images using segmentations (LPSEG method) and representing the similarity criterion using contributions at a set of sampling points. The novelty here is representing the problem as a linear program, allowing an efficient and robust optimization. We have also shown that segmentation can be sometimes advantageously replaced by edge-finding (LPNOSEG), with a further increase in speed. The bottleneck is currently the segmentation or edge-finding but we hope it should be possible to avoid it by using a faster segmentation algorithm or a GPU implementation.

References

1. Zitová, B., Flusser, J.: Image registration methods: a survey. Image Vis. Comput. **21**, 977–1000 (2003)
2. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
3. Thévenaz, P.: Halton sampling for image registration based on mutual information. Sampl. Theory Signal and Image Process. **7**(2), 141–171 (2008)
4. Klein, S., Staring, M., Pluim, J.P.W.: Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. IEEE Trans. Image Proc. **16**(12), 2879–2890 (2007)
5. Sabuncu, M., Ramadge, P.: Gradient based nonuniform subsampling for information-theoretic alignment methods. In: Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1683–1686 (2004)
6. Domokos, C., et al.: Nonlinear shape registration without correspondences. IEEE Trans. Pattern Anal. Mach. Intell. **34**(5), 943–958 (2012)
7. Droske, M., Ring, W.: A Mumford-Shah level-set approach for geometric image registration. SIAM Appl. Math. **66**(6), 2127–2148 (2005)
8. Kybic. J., et al.: Fast registration of segmented images by normal sampling. In: CVPRW: BioImage Computing Workshop, pp. 11–19, June 2015
9. Taylor, C.J., Bhusnurmath, A.: Solving image registration problems using interior point methods. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5305, pp. 638–651. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88693-8_47
10. Ben-Ezra, M.: Real-time motion analysis with linear programming. Comput. Vis. Image Underst. **78**(1), 32–52 (2000)
11. Jiang, H., et al.: Matching by linear programming and successive convexification. IEEE Trans. Pattern Anal. Mach. Intell. **29**(6), 959–975 (2007)

12. Unser, M.: Splines: a perfect fit for signal and image processing. IEEE Signal Processing Magazine **16**(6), 22–38 (1999)
13. Glocker, B.: Dense image registration through MRFs and efficient linear programming. Med. Image Anal. **12**(6), 731–741 (2008)
14. Arganda-Carreras, I., Sorzano, C.O.S., Marabini, R., Carazo, J.M., Ortiz-de-Solorzano, C., Kybic, J.: Consistent and elastic registration of histological sections using vector-spline regularization. In: Beichel, R.R., Sonka, M. (eds.) CVAMIA 2006. LNCS, vol. 4241, pp. 85–95. Springer, Heidelberg (2006). https://doi.org/10.1007/11889762_8
15. Klein, S., Staring, M., Murphy, K.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29**(1), 196–205 (2010)