AUTOMATIC SIMULTANEOUS SEGMENTATION AND FAST REGISTRATION OF HISTOLOGICAL IMAGES

Jan Kybic, Jiří Borovec

Czech Technical University in Prague, Czech Republic

ABSTRACT

We describe an automatic method for fast registration of images with very different appearances. The images are jointly segmented into a small number of classes, the segmented images are registered, and the process is repeated. The segmentation calculates feature vectors on superpixels and then it finds a softmax classifier maximizing mutual information between class labels in the two images. For speed, the registration considers a sparse set of rectangular neighborhoods on the interfaces between classes. A triangulation is created with spatial regularization handled by pairwise spring-like terms on the edges. The optimal transformation is found globally using loopy belief propagation. Multiresolution helps to improve speed and robustness. Our main application is registering stained histological slices, which are large and differ both in the local and global appearance. We show that our method has comparable accuracy to standard pixel-based registration, while being faster and more general.

Index Terms— image registration, image segmentation, mutual information, loopy belief propagation

1. INTRODUCTION

The motivation of our work is registration of large histological slices (see Fig. 1,2). Many different stains exist to visualize different aspects of the cell structure and the presence of various proteins and other biomarkers. Since only a limited number of stains can be used in a single slice, to obtain a complete information about a specific tissue location, the slices must be registered [1]. There can be large differences in the global appearance. Not only the color is different but also different objects are visible. Moreover, because of distance between slices, the small details small-scale details do not correspond either. Other applications include fast multimodal registration in medical imaging or remote sensing.

Our strategy is first to simultaneously segment the two images by learning two independent classifiers so that the segmentation they produce is as similar as possible. Then, we quickly register the two segmentations by aligning the class contours using a sparse set of control points. The segmentation and registration can be viewed as alternative minimizations of the same criterion, the mutual information on labels (MIL). The algorithm is fast and parallelisable.

1.1. Related work

Methods used for aligning histology slices include global pixel-based methods [1–3], local search methods [4,5], and features-based methods [6]. However, these approaches assume the same appearance of all images.

The project was partially supported by Czech Science Foundation under project P202/11/0111 and the Grant Agency of the Czech Technical University in Prague, grant No.SGS12/190/OHK3/3T/13.

Methods combining segmentation and registration find object contours by level set methods [7,8], or by thresholding [9], formulate a joint criterion minimized by GraphCut [10], or register segmented objects based on shape descriptors [11].

Image registration by discrete minimization was proposed by [12] and others on a dense grid.

2. METHOD

Given a reference image F and a moving image G on a pixel grid Ω, we want to find a transformation T, such that a point $\mathbf{r}$ in image F corresponds to point $\mathbf{r}' = T(\mathbf{r})$ in image G. Instead of registering directly the images F and G, we register their segmentations f and g. We alternate segmentation (Section 3) and registration (Section 4) until convergence. In both steps, we minimize the same criterion J, the mutual information between the two segmentations (Section 2.1):

$$J(f, g, T) = -\operatorname{MIL}(f, T \circ g) + \mathcal{R} \tag{1}$$

with $f = \Psi_f F, \quad g = \Psi_g G$, where f and g are soft segmentations, i.e. $f_k(j)$ is a probability that a pixel j in image F belongs to class k, and Ψ_f is a parameterized class probability model (Section 3.1). When needed, hard segmentation is obtained as $\check{f}_k(j) = [\![k = \arg\max_l f_l(j)]\!]$. The symbol $\mathcal{R}$ represents regularization (see equations 5,11). Image G should be treated analogously.

2.1. Mutual information of labels

Mutual information of labels (MIL) is similar to standard MI [13,14]

$$\operatorname{MIL}(f, g) = \sum_{k,l} p_{k,l} \log \frac{p_{k,l}}{p_k p_l} \tag{2}$$

with $p_k = \sum_l p_{k,l}$ and $p_l = \sum_k p_{k,l}$. The $p_{k,l}$ is a probability that at the same location, the class in F is k and the class in G is l.

$$p_{k,l} = \frac{1}{|\Omega|} \sum_{j \in \Omega} f_k(j) g_l(j) \tag{3}$$

Compared to standard MI, the number of bins (classes) is much smaller. Therefore, MIL is more robust, faster to evaluate and needs less input samples. We choose to consider only integer displacements during registration, $T(\mathbf{x}) \in \Omega$, so no interpolation is needed.

3. SEGMENTATION

The segmentation assigns each pixel a class $\{1, 2, \ldots, L\}$, where L is given. It starts by calculating SLIC superpixels [15], $S_i \subset \Omega$. See Fig.1 for examples. In the second step, a descriptor vector $\tilde{\mathbf{x}}_i$ is calculated for each superpixel S_i. The descriptors are application dependent; we found simple means $\tilde{\mathbf{x}}_i = \frac{1}{|S_i|} \sum_{j \in S_i} F(j)$ to be sufficient. Note that we work with colour, three channel images. For further convenience, we set $\mathbf{x}_i = \begin{bmatrix} 1 & \tilde{\mathbf{x}}_i \end{bmatrix}$.

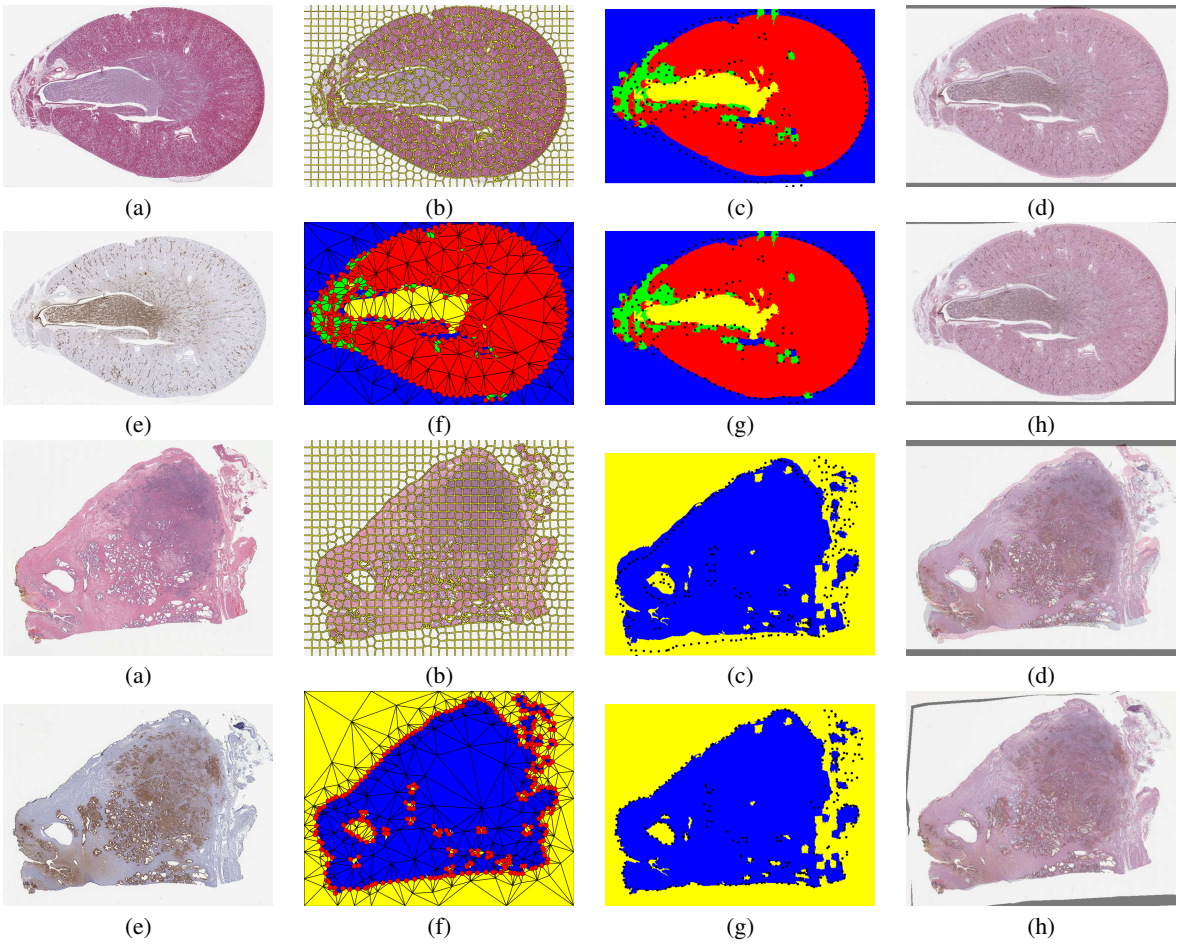

Fig. 1. Top two rows: Histology slices of rat kidney stained with H&E and PanCytokeratin. Bottom two rows: Histology slices of human prostate [1] stained with H&E and PSAP. We show the input images (a,e), the superpixels (b, yellow lines) and the segmentation and triangulated mesh (f) of (a) with big red dots for control points C. Images (c,g) show the position of the control points C in the moving image before and after the first iteration, overlaid over the segmentation of (e). Images (d,h) show the overlaid images before and after registration.

3.1. Softmax regression

The class probability model $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^L$ takes a descriptor vector $\mathbf{x}_i$ and produces the probabilities $z_{i,k}$ for each class k and superpixel i, corresponding to a soft pixel-level segmentation $f_k(j) = z_{i,k}$, with $j \in S_i$. We use a linear softmax regression [16], parameterized by vectors of coefficients $\mathbf{a}_k \in \mathbb{R}^d$ for class k.

$$z_{i,k} = \frac{\exp\left(\mathbf{a}_k^T \mathbf{x}_i\right)}{\sum_{l=1}^{L} \exp\left(\mathbf{a}_l^T \mathbf{x}_i\right)} \tag{4}$$

3.2. Optimization

We minimize (1) by the L-BFGS algorithm with respect to the parameters $\mathbf{a}$. Regularization is added to avoid non-uniqueness [16], yielding

$$J_S(\mathbf{a}_{FG}) = -\operatorname{MIL}(f, T \circ g) + \frac{\beta}{2L}\|\mathbf{a}_{FG}\|^2 \tag{5}$$

with $\mathbf{a}_{FG} = \begin{bmatrix}\mathbf{a}_F & \mathbf{a}_G\end{bmatrix}$, where $\mathbf{a}_F$, $\mathbf{a}_G$ are class model parameters (from (4)) for images F and G, respectively. The probability $p_{k,l}$ (3) to find MIL is evaluated directly on superpixels as

$$p_{k,l} = \frac{1}{|\Omega|} \sum_{(i,j)} s_{i,j} z_{i,k}^F z_{j,l}^G$$

with superpixel overlaps $s_{i,j} = \sum_{\mathbf{r} \in S_i} [\![T(\mathbf{r}) \in S_j]\!]$.

3.3. Initialization

Because the criterion (5) is not convex, a reasonable initialization for the parameters $\mathbf{a}$ is required. We choose randomly L positions $\mathbf{v}_1, \ldots \mathbf{v}_L$ in the image, and find sets of superpixels U_k^F and U_k^G within a small radius r_{init} from $\mathbf{v}_k$ in F and from $T(\mathbf{v}_k)$ in G, respectively. Using L-BFGS, we find such coefficients $\mathbf{a}_{FG}$, so that the superpixels in U_k^F and U_k^G are in class k by minimizing a sum of standard softmax cost functions for the two images

$$J_I(\mathbf{a}_{FG}) = \mathcal{L}^F + \mathcal{L}^G + \frac{\beta}{2L}\|\mathbf{a}_{FG}\|^2 \tag{6}$$

$$\mathcal{L}^\bullet = \frac{1}{\sum_{k=1}^{L} \sum_{i \in U_k^\bullet} w_i^F} \sum_{k=1}^{L} \sum_{i \in U_k^\bullet} w_i^\bullet \log z_{i,k}^\bullet \tag{7}$$

where $w_i = |S_i|$ is the superpixel size. Once the initial $\mathbf{a}_{FG}$ is found, we continue by minimizing (5). We repeat the random initialization N_{init} times and the best $\mathbf{a}_{FG}$ in terms of J (1) is retained.

4. REGISTRATION

The registration minimizes a sum of the MIL data criterion and a smoothness term using loopy belief propagation. The key observation is that as T varies, the only contributions to the change of $\text{MIL}(f, T \circ g)$ come from the boundaries between classes. Furthermore, as T is smooth, it is sufficient to describe it by its value $\mathbf{y}_i = T(\mathbf{r}_i)$ at a sparse set of control points $\mathbf{r}_i \in C$. Control points are located at class boundaries and are pruned to be at least ε pixels apart.

Once the new control point positions $\mathbf{y}_i = T(\mathbf{r}_i)$ are found, the transformation T can be interpolated everywhere using either bilinear interpolation on each triangle, or the Clough-Tocher scheme [17], if a C^1 continuity is required.

4.1. Data criterion

The contribution D_i of small neighborhoods Ω_i of size $h \times h$ pixels around each control point is

$$D_i(\mathbf{y}_i) = -\sum_{k,l} \tilde{p}_{k,l} \log \frac{p_{k,l}}{p_k p_l} \tag{8}$$

where $p_{k,l}$ is calculated from (3) in the whole image Ω, while $\tilde{p}_{k,l}$ is calculated from the corresponding neighborhood Ω_i.

We allow only integer $\mathbf{y}_i$ and limit the maximum displacement, $\|\mathbf{y}_i - \mathbf{r}_i\|_\infty \leq d$. Then D_i is precalculated for all $\mathbf{r}_i \in C$. For convenience, we define $D_i = 0$ for $\mathbf{r}_i \notin C$.

4.2. Regularization

We triangulate[1] the points C (see Fig.1b), providing an augmented set of control points $C' \supseteq C$ and a set of undirected edges $E \subseteq C' \times C'$. The regularization consists of pairwise spring-like terms [6, 18], penalizing differences in relative control point displacements

$$U = \frac{1}{2} \sum_{(i,j)\in E} \omega_{ij} \|(\mathbf{y}_i - \mathbf{r}_i) - (\mathbf{y}_j - \mathbf{r}_j)\|^2 \tag{9}$$

For small deformations such spring model approximates the behavior of a thin membrane [18] with a Poisson ratio $\nu = 1/3$ if we set the contribution to the weights ω_{ij} from each triangle as

$$\omega_{ij} = \lambda \frac{\|\mathbf{r}_i - \mathbf{r}_j\|^2}{8A} \left(3 \cot^2 \alpha + \frac{1}{2}\right) \tag{10}$$

where λ is the Lamé's first parameter, A is the triangle area and α is the angle oposite to the edge ij.

4.3. Belief propagation

The registration minimizes a sum of contributions (8) with the regularization (9) with respect to $\mathbf{y}_i$ for $\mathbf{r}_i \in C'$.

$$J_R = \sum_{\mathbf{r}_i \in C} D_i(\mathbf{y}_i) + \frac{1}{2} \sum_{(i,j)\in E} \omega_{ij} \|(\mathbf{y}_i - \mathbf{r}_i) - (\mathbf{y}_j - \mathbf{r}_j)\|^2 \tag{11}$$

[1] Using the MeshPy library, `http://mathema.tician.de/software/meshpy/`

We use belief propagation [19], which sends messages

$$\mu^t_{i\to j}(\mathbf{y}_j) = \min_{\mathbf{y}_i} \Big(\frac{\omega_{ij}}{2} \|(\mathbf{y}_i - \mathbf{r}_i) - (\mathbf{y}_j - \mathbf{r}_j)\|^2 + D_i(\mathbf{y}_i) + \sum_{s\neq j} \mu^{t-1}_{s\to i}(\mathbf{y}_i) \Big) \tag{12}$$

for $(i, j) \in E$ until convergence. Then the estimated displacement is extracted as $\mathbf{y}^*_j = \arg\max_{\mathbf{y}_j} \left(D_j(\mathbf{y}_j) + \sum_i \mu^t_{i\to j}(\mathbf{y}_j) \right)$

The computational complexity of belief propagation is $O(|E|d^2)$ per iteration, which can be prohibitive for large d. A speed-up can be achieved by a coarse to fine multiresolution strategy, which requires filtering and downsampling around the control points and a modification of the way the messages are calculated to incorporate the starting point from the previous level.

4.4. Affine initialization

If the deformation to be found is large, it is preferable to first find an affine transformation T_A minimizing the MIL data criterion. A simple non-iterative method works sufficiently well: We find the locally optimal displacements $\mathbf{y}^*_i = \arg\min_{\mathbf{y}} D_i(\mathbf{y})$ and calculate a least squares fit minimizing $\sum_i \|T_A(\mathbf{r}_i) - \mathbf{y}^*_i\|^2$ by solving a linear system of equations.

Second, we warp the image $G' = T_A \circ G$ (only the superpixel map S^G needs to be warped) and perform the procedure as described above to find an elastic transformation T_E between F and G'. The final transformation is a composition $T = T_E \circ T_A$.

5. RESULTS

Figure 1 shows the ASSAR intermediate and final results for two examples—differently stained nearby histological slices of rat kidney and human prostate of about 2000×2000 pixels.

We use the following parameters in both examples: $N_{\text{init}} = 10$, $r_{\text{init}} = 5$, $K = 4$, $\lambda = 10$, $h = 10$, $d = 3$, $\beta = 10^{-3}$, $\varepsilon = 20$, and 1000 superpixels. Several iterations of the ASSAR algorithm are performed, each taking about 30 s. The resulting deformation is usable already after the first iteration, further improvements are small.

We have compared the speed, robustness and accuracy of ASSAR with B-spline based elastic image registration implemented in bUnwarpJ [9] using sum of squared differences (SSD) and Elastix [20] using mutual information (MI). We have also applied a coarse feature-based affine pre-registration using SURF and RANSAC as implemented in OpenCV; this method is denoted OpenCV+Elastix. The comparison is done on 34 histology image pairs with 40∼90 manually identified landmarks per image (see Figure 2). We calculate the landmark registration error and consider registration a success if the median error decreases by at least 50 %. We report the execution time, mean and median landmark error and the percentage of successful runs in Table 1. We see that ASSAR has the highest success rate among all tested methods and it is also the fastest one, in spite of being currently implemented in Python. Only OpenCV+Elastix is more accurate but also much slower.

6. CONCLUSIONS

We have presented a new general image registration method for images with very different local and global appearances. The method is competitive with traditional pixel-based methods in terms of accuracy. This indicates that the class label probably captures most

method	time	mean err.	median err.	success rate
bUnwarpJ	401	64	50	50%
Elastix	515	54	45	67%
OpenCV+Elastix	764	23	12	88%
ASSAR	130	36	17	91%

Table 1. Comparison of registration methods. We show the execution time in seconds, the mean and median registration error in pixels, and percentage of successful runs.

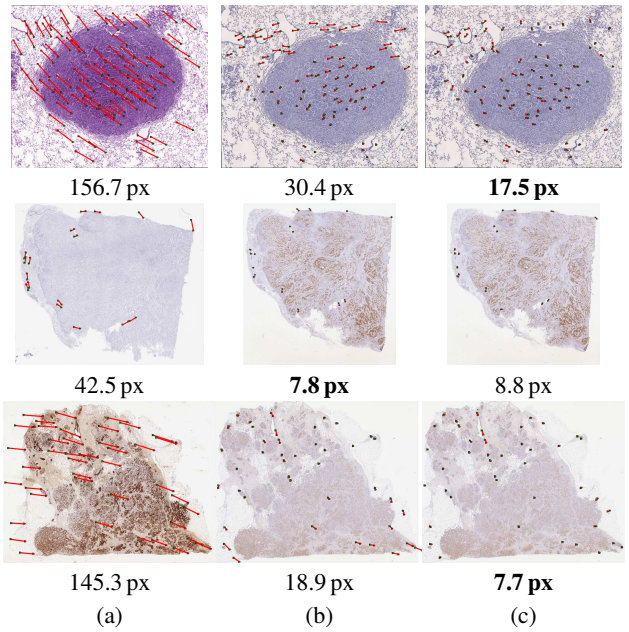

Fig. 2. Reference image with manually identified landmark positions in both images connected by red lines (a); the target image with true and calculated landmark positions by ITK (b), and ASSAR (c). Means of the landmark registration errors in pixels are shown below each image; bold denotes the best result.

of the information common between the images. ASSAR is also robust, with a large basin of attraction. Most importantly, by considering only a small fraction of the image pixels, an optimized implementation of ASSAR has the potential of being many times faster than standard pixel-based method. ASSAR could also be combined with existing methods by for example providing a fast initialization to a more accurate but less robust and slower standard pixel-based method.

References

[1] G. Metzger, S. Dankbar, J. Henriksen, A. Rizzardi, N. Rosener, and S. Schmechel, "Development of multigene expression signature maps at the protein level from digitized immunohistochemistry slides," *PLoS ONE*, vol. 7, no. 3, pp. e33520, 2012.

[2] M. Auer, P. Regitnig, and G. Holzapfel, "An automatic nonrigid registration for stained histological sections.," *IEEE Trans. Im. Processing*, vol. 14, no. 4, pp. 475–86, 2005.

[3] Jiří Borovec, Jan Kybic, Michal Bušta, Carlos Ortiz-de Solorzano, and Arrate Muñoz-Barrutia, "Registration of multiple stained histological sections," in *Proceedings of 2013 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 3 Park Avenue, 17th Floor, New York, USA, April 2013, IEEE, pp. 1022–1025, IEEE Computer Society Press.

[4] M. Chakravarty, G. Bertrand, C. Hodge, A. Sadikot, and L. Collins, "The creation of a brain atlas for image guided neurosurgery using serial histological data," *NeuroImage*, vol. 30, no. 2, pp. 359 – 376, 2006.

[5] H. Hufnagel, X. Pennec, G. Malandain, H. Handels, and N. Ayache, "Non-linear 2D and 3D registration using block-matching and B-splines," in *Bildverarbeitung für die Medizin*, pp. 325–329. Springer, 2005.

[6] S. Saalfeld, R. Fetter, A. Cardona, and P. Tomancak, "Elastic volume reconstruction from series of ultra-thin microscopy sections," *Nature Methods*, , no. 9, pp. 717–720, 2012.

[7] A. Yezzi, L. Zöllei, and T. Kapur, "A variational framework for integrating segmentation and registration through active contours," *Med. Im. Anal.*, , no. 7, pp. 171–185, 2003.

[8] F. Wang and B. C. Vemuri, "Simultaneous registration and segmentation of anatomical structures from brain MRI," in *MICCAI*, 2005, pp. 17–25.

[9] I. Arganda-Carreras, R. Fernández-González, A. Muñoz Barrutia, and C. Ortiz-De-Solorzano, "3D reconstruction of histological sections: Application to mammary gland tissue.," *Microscopy research and technique*, vol. 73, no. 11, pp. 1019–29, 2010.

[10] D. Mahapatra and Y. Sun, "Joint registration and segmentation of dynamic cardiac perfusion images using mrfs," in *MICCAI*, 2010, pp. 493–501.

[11] H. Gonçalves, J. A. Gonçalves, and L. Corte-Real, "HAIRIS: a method for automatic image registration through histogram-based image segmentation.," *IEEE Trans. Im. Processing*, vol. 20, no. 3, pp. 776–789, 2011.

[12] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through MRFs and efficient linear programming," *Med. Im. Anal.*, vol. 12, no. 6, pp. 731–741, 2008.

[13] P. Viola and W. M. Wells, "Alignment by Maximization of Mutual Information," *Int. J. Comp. Vision*, vol. 24, no. 2, pp. 134–154, 1997.

[14] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[15] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pat. Anal. Mach. Intel.*, vol. 34, no. 11, pp. 2274–2282, 2012.

[16] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2001.

[17] R. W. Clough and J. L. Tocher, "Finite element stiffness matrices for analysis of plates," in *Proceedings of the conference on matrix methods in structural mechanics*, 1965.

[18] H. Delingette, "Triangular springs for modeling nonlinear membranes," *IEEE Trans. Vis. Comp. Graphics*, vol. 14, no. 2, 2008.

[19] P. Felzenszwalb and D. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comp. Vision*, vol. 70, no. 1, pp. 41–54, 2006.

[20] S. Klein, M. Staring, and K. Murphy, "Elastix: a toolbox for intensity-based medical image registration," *Medical Imaging, IEEE*, vol. 29, no. 1, 2010.