

ALGORITHM FOR SPLINE-BASED ELASTIC REGISTRATION IN APPLICATION TO CONFOCAL IMAGES OF GENE EXPRESSION

C.O.S. Sorzano^{1,2}, M. Blagov³, P. Thévenaz¹, E. Myasnikova³, M. Samsonova³, M. Unser¹

¹ Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland

² Escuela Politécnica Superior, Universidad San Pablo-CEU, Madrid, Spain

³ St. Petersburg State Polytechnical University, St. Petersburg, Russia

We perform elastic registration by an algorithm based on a pixelwise and regularized optimization criterion. We express the deformation field thanks to B-splines, which allows us to deal with a rich variety of deformations. The algorithm is able to handle soft landmark constraints, which is particularly useful when parts of the images contain very little information or when it is unevenly distributed. We solve the problem by minimizing a distance between the target image and the warped source. We regularize this minimization problem by divergence and curl. We apply the proposed algorithm to the registration of the confocal scanning microscopy images of *Drosophila* embryos.

Introduction

Putting two images into registration can be restated as finding a function (also called a deformation field) that performs a backward mapping of a target image onto a source image [1, 2]. This is a typical problem for various molecular-biology disciplines. The proposed method is a generalization of some recent works on the registration of biological images [3, 4, 5]. Representing the deformation field by B-splines enables us to produce nonlinear elastic deformation models [6] and provides high-quality interpolation [7]. The deformation field is estimated through a minimization problem that includes the energy of the error between both images, a smoothness regularization term, and the error resulting from the inexact match of landmarks that are used to handle any potential information mismatch.

In the present study, we illustrate the registration problem in the context of analysing the gene-expression images obtained by means of immunofluorescence histochemistry and by confocal scanning microscopy [8]. These technologies enable us to visualize and quantify molecules in cells with high spatial and temporal resolution.

We apply our registration algorithm to confocal images of fruit fly *Drosophila* embryos published in [9]. The registration is necessary to build a precise map of the gene expression. Each gene in this map is represented by its integrated expression pattern, which we obtain by averaging individual patterns after they have been registered. High-quality registration is of primary importance, as averaging of unregistered or of inaccurately registered images leads to a blurred average image that doesn't possess the essential characteristic features of the whole class of individual patterns.

Registration Method

Our registration methodology is based on the minimization of an energy functional that incorporates a similarity measure E_{img} between the target image and the warped source image, soft landmark constraints through the term E_{μ} , and *a priori* knowledge about the deformation field through two independent measures E_{div} and E_{rot} that are related to the gradients of both divergence and curl of the deformation field. Thus, the energy to minimize is the linear combination of these energy terms given by

$$E = w_i E_{\text{img}} + w_{\mu} E_{\mu} + w_d E_{\text{div}} + w_r E_{\text{rot}}. \quad (1)$$

The goal of image registration is to find a function $\mathbf{g}(\mathbf{x}) : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ which maps coordinates from a target image I_t onto a source image I_s so that $I_s(\mathbf{g}(\mathbf{x}))$ (a warped version of the source image) resembles as much as possible $I_t(\mathbf{x})$. A standard way of evaluating this resemblance is by measuring the energy of the difference image

$$E_{\text{img}} = \int_{\mathbf{x} \in \mathbf{R}^2} (I_t(\mathbf{x}) - I_s(\mathbf{g}(\mathbf{x})))^2 d^2 x. \quad (2)$$

Image representation. Note that, in (2), the target image is always evaluated at integer positions. However, this is not the case for the source image I_s which needs to be evaluated at non-integer positions $\mathbf{g}(\mathbf{x})$. For this purpose, we use cubic B-spline interpolation as follows: We express the source image as

$$I_s(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbf{Z}^2} c_{\mathbf{k}} \beta^3\left(\frac{x_2}{h_2} - k_2\right) \beta^3\left(\frac{x_1}{h_1} - k_1\right), \quad (3)$$

where β^3 is the B-spline of degree 3, where the coefficients $c_{\mathbf{k}}$ are given by [10], and where h_1 and h_2 are parameters that control the degree of detail of the representation.

Deformation representation. Likewise [3, 4, 5], we also express the deformation field as the linear combination of B-splines given by

$$\mathbf{g}(\mathbf{x}) = (g_1(x_1, x_2), g_2(x_1, x_2)),$$

where, for $i \in \{1, 2\}$, we have that

$$g_i(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbf{Z}^2} c_{i,\mathbf{k}} \beta^3\left(\frac{x_2}{h_{i,2}} - k_2\right) \beta^3\left(\frac{x_1}{h_{i,1}} - k_1\right). \quad (4)$$

Cubic B-splines constitute a Riesz basis of L_2 and have a fourth order of approximation. Therefore, any deformation field can be represented using a fine-enough scale. Moreover, the use of cubic B-splines guarantees the continuity of the derivatives of the deformation up to the second order.

Use of landmarks. Landmarks are used to impose soft constraints to the deformation field. We have chosen to deal with soft constraints because the landmark positions may be contaminated by noise; therefore, there is no point in trying to satisfy them exactly. A second reason for using soft landmark constraints is that

landmarks may happen to be distributed in such a way that there exists no geometric deformation in the search space that can completely satisfy all of them.

Let us assume that K pairs of corresponding landmarks $(\mu_s^{(k)}, \mu_t^{(k)})$, $k \in \{1, 2, \dots, K\}$, are available. Then, we handle the soft landmark constraints by an energy term of the form

$$E_{\mu} = \frac{1}{K} \sum_{k=1}^K \|\mu_t^{(k)} - \mathbf{g}(\mu_s^{(k)})\|^2.$$

A priori knowledge about deformation smoothness. The smoothness of the deformation field is a useful regularization term for the minimization problem, especially when little information is available. The authors of [11] propose two regularizing terms that fully exploit the vector nature of the data

$$w_d E_{\text{div}} + w_r E_{\text{rot}} = w_d \int_{\mathbf{R}^2} \|\nabla \text{div} \mathbf{g}\|^2 d^2 x + w_r \int_{\mathbf{R}^2} \|\nabla \text{rot} \mathbf{g}\|^2 d^2 x, \quad (5)$$

where $\text{div} \mathbf{g} = \partial_{x_1} g_1 + \partial_{x_2} g_2$ represents the divergence of the 2D vector field \mathbf{g} , where $\text{rot} \mathbf{g} = -\partial_{x_2} g_1 + \partial_{x_1} g_2$ gives the length of the unique component of the curl of \mathbf{g} , and where $\nabla f = (\partial_{x_1} f, \partial_{x_2} f)$ is the gradient of a scalar function f . The divergence of a vector field is related to the existence of sinks and sources in the vector field, while the curl is related to the rotation within the field. Therefore, the proposed regularization penalizes changes in the local structure of the divergence and curl.

Thanks to the fact that B-splines are piecewise polynomials, the integrals in (5) can be precomputed exactly using closed formulæ; therefore, the evaluation of the regularization term (5) is very fast and efficient.

Optimizer. To minimize (1) with respect to the parameters of the transform \mathbf{g} in (4), we use an optimization algorithm that is inspired by the Levenberg-Marquardt nonlinear regression [12] and that is enhanced by the BFGS (Broyden-Fletcher-Goldfarb-Shanno) estimation of the Hessian [13]. This algorithm achieves smooth transi-

tions between quasi-Newton and gradient-descent steps, depending on the goodness of fit of a local model of the function being minimized.

Implementation and Results

We have applied our registration method to confocal images of gene expression patterns in *Drosophila* embryos. First, we have carefully cropped the unregistered raw images so that embryos fit tightly into the image frame. As a result, the images differ in size since they reflect the actual size of an embryo. To ease registration, we proceed by bringing them to the same size.

The registration is implemented iteratively:

Step 0: Center the unregistered images and pad them with zeros to reach the size of the widest image, and of the highest image too.

Step 1: Compute a new average image.

Step 2: Align all images against the average embryo from Step 1.

Step 3: Alternate between Step 1 and Step 2.

This process can be iterated until the accuracy of registration is satisfactory. We monitor this accuracy by visual inspection of the registered images (Fig.1). We stopped after two iterations. The weight parameters for the registration algorithm were $w_i = 1$, $w_\mu = 0$, $w_d = w_r = 1$ (i.e., no landmarks were used). Sometimes, however, acceptable registration accuracy could not be achieved in automated mode (e.g., in the presence of a large displacement relative to the average embryo); in those cases, the use of landmarks proved extremely helpful.

After several alignment iterations, even the widest or highest embryos (those that determined the size of the first average at Step 0) shrink a little and are not filling the whole image space anymore. As a result, the registered images have an empty rim around them, which propagates to the average image (Fig.2(a)). These margins should be removed to enforce a tight fit of the average embryo into the image frame. We compute the cropping coordinates as follows: For each individually transformed

image, we track the coordinates of the original left, right, upper, and lower bounds of an embryo (as shown in Fig.1). Then, we compute its average size by pooling together the transformed x -coordinates of the left and right edges, and the transformed y -coordinates of the upper and lower edges. After cropping, the average image is brought to the same scale as the raw individual ones (Fig.2(b)). Finally, we use the same coordinates to crop all individually transformed embryos and present all the registered embryos at the same scale.

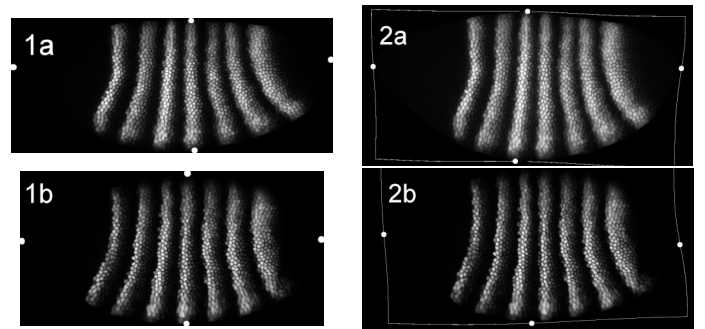


Fig.1. Images of two embryos before registration (1a and 1b), and after registration (2a and 2b). The transformed frames of the source images are drawn in the right panels. The extreme points of embryos are marked by small circles.

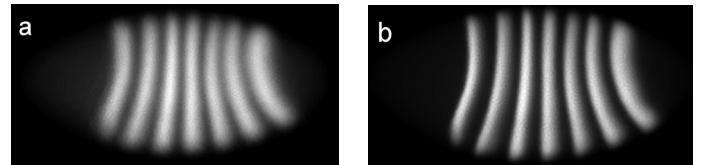


Fig.2. Average images obtained before registration (a) and after registration (b). The image (b) is cropped to the average scale.

Registration accuracy. The expression of the genes under consideration (segmentation genes) is largely a function of the position along the horizontal axis of the embryo body. Hence, one-dimensional data extracted from narrow strips cut out from an image along the x -axis can be used to give a fair representation of the registration accuracy. In Fig.3, we provide graphs of the data extracted from the central 10% strip before and after registration. In doing so, we have excluded the empty image areas. As a result of this segmentation, a raw image is transformed into a series of data records that describe groups of non-empty pixels, separated from other groups by empty pixels. The coordinates of non-empty group centroids

are transformed according to (4), and the registered one-dimensional patterns are presented in Fig.3(b). Thanks to the high quality of registration, the resulting integrated pattern successfully preserves the characteristic features of the individual patterns.

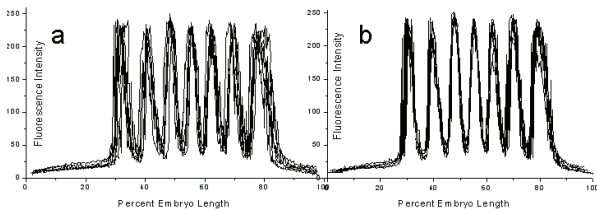


Fig.3. Five expression patterns presented as one-dimensional graphs before registration (a) and after registration (b).

Conclusion

We have presented an elastic registration algorithm based on B-splines that improves on state-of-the-art nonlinear registration algorithms. Our algorithm incorporates a pixelwise and regularized optimization criterion which proves to offer a very efficient framework. In particular, it allowed us to insert *a priori* knowledge about the deformation class through regularization weights. We have proposed a novel methodology to incorporate the landmark information in the optimization process from its very initialization. We have improved the performance of the optimizer by the introducing a BFGS estimator of the Hessian. Finally, we have performed the optimization in a multiresolution fashion to increase the robustness and efficiency of the algorithm.

We have successfully registered confocal images of gene expression patterns in *Drosophila* embryos. The registration shows a high accuracy, as evidenced by the coincidence of the shape and location of gene expression domains in the individual patterns. As a result, the average image gives a fair representation of the essential features of a whole class of individual embryos, which is very important for the further analysis and modeling of data.

References

1. Brown, L.G., 1992. A survey of image registration techniques. *ACM Computing Survey* 24, 325-376.
2. Zitova, B., Flusser, J., 2003. Image registration methods: A survey. *Image and Vision Computing* 21, 977-1000.
3. Kybic, J., Unser, M., 2003. Fast parametric elastic image registration. *IEEE Trans. Image Processing* 12, 1427-1442.
4. Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W., 2003. PET-CT image registration in the chest using free-form deformations. *IEEE Trans. Medical Imaging* 22, 120-128.
5. Musse, O., Heitz, F., Armspach, J.P., 2003. Fast deformable matching of 3D images over multiscale nested subspaces. *Application to atlas-based MRI segmentation. Pattern Recognition* 36, 1881-1899.
6. Szeliski, R., Coughlan, J., 1997. Spline-Based image registration. *Intl. J. Computer Vision* 22, 199-218.
7. Unser, M., Aldroubi, A., Eden, M., 1991. Fast B-spline transforms for continuous image representation and interpolation. *IEEE Trans. Pattern Analysis & Machine Intelligence* 13, 277-285.
8. Myasnikova, E., Samsonova, A., Kozlov, K., Samsonova, M., Reinitz, J., 2001. Registration of the expression patterns of *Drosophila*: Segmentation genes by two independent methods. *Bioinformatics* 17, 3-12.
9. Kozlov, K., Myasnikova, E., Pisarev, A., Samsonova, M., Reinitz, J., 2002. A method for two-dimensional registration and construction of the two-dimensional atlas of gene expression patterns *in situ*. *In Silico Biology* 2, 125-141. <http://www.bioinfo.de/isb/2002/02/0011>
10. Unser, M., Aldroubi, A., Eden, M., 1993. B-Spline signal processing: Part II—Efficient design and applications. *IEEE Trans. Signal Processing* 41, 834-848.
11. Amodi, L., Benbourhim, M. N., 1991. A vector spline approximation. *J. of Approximation Theory* 67, 51-79.
12. Thévenaz, P., Ruttiman, U.E., Unser, M., 1998. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Processing* 7, 27-41.
13. Press, W., Teukolsky, S., Vetterling, W., Flannery, B., 1992. *Numerical recipes in C*, 2nd Edition. Cambridge University Press.