

Robust and staining-invariant elastic registration of a series of images from histologic slices

Stefan Wirtz^a, Nils Papenberg^a, Bernd Fischer^a and Oliver Schmitt^b

^aInstitute of Mathematics, University of Lübeck,
Wallstraße 40, 23560 Lübeck, Germany

^bInstitute of Anatomy, University of Rostock,
Gertrudenstraße 9, 18055 Rostock, Germany

ABSTRACT

In image registration of medical data a common and challenging problem is handling intensity-inhomogeneities. These inhomogeneities appear for instance in images of serially sectioned brains caused by the histological staining process or in medical imaging with contrast agents. Beneath this, natural outliers (for instance cells or vessels) produced by the underlying material itself may be mistaken as noise. Both image registration applications have in common that the well known *sum of squared differences* (SSD) measure would detect false differences. To deal with these kinds of problems, we supplement the common SSD-measure with image derivatives of higher order. Additionally we introduce a non-quadratic penalizer function to the distance measure leading to robust energy. The concepts are well known in optical flow. Overall, we present a variational model which combines all of these properties. This formulation leads to a fast and efficient algorithm. We demonstrate its applicability at the problems described above.

Keywords: Registration, histologic images, robust energy, higher order similarity measures, staining inhomogeneous

1. INTRODUCTION

In registration of histologic images, for instance the registration of images of a serially sectioned brain,¹ it is always difficult to handle the inhomogeneities in the intensity values resulting from the staining process. Furthermore, sometimes the underlying material itself does produce natural outliers, which may be mistaken as noise and are challenging to deal with.

Due to the nature of distortions observed in histologic images, we have chosen the well-known elastic registration scheme,¹⁻⁶ which mimics the elastic behavior of the histologic sections. Furthermore we apply techniques from optical flow⁷⁻⁹ to receive robustness under noise and independence from the staining inhomogeneities.

The introduced techniques are state of the art in optical flow applications. The first ideas were presented by *Uras et al*¹⁰ in 1988 and by *Tistarelli*¹¹ in 1994. They described a method dealing with image derivatives to handle the so called aperture problem in optical flow estimation. In 2004, *Brox et al*⁹ proposed a non-linear method that uses the gradient as an indicator for the similarity of the given images. For the first time a variational model was described that can handle changes of the intensity values of an image sequence using its derivatives. The benefit of this new approach turned out to be impressive. Since the variational approaches of optical flow estimation and image registration are closely related, it is possible to transfer these ideas easily to image registration.

Another idea from optical flow was the introduction of a non-quadratic penalizer-function to the variational model. This leads to so-called robust energies,^{12,13} i.e. the influence of outliers to the distance measure decreases

Further author information: (Send correspondence to Stefan Wirtz)

Stefan Wirtz: E-mail: wirtz@math.uni-luebeck.de, Telephone: +49 451 7030 414

Nils Papenberg: E-mail: papenber@math.uni-luebeck.de, Telephone: +49 451 7030 418

Bernd Fischer: E-mail: fischer@math.uni-luebeck.de, Telephone: +49 451 7030 434

Oliver Schmitt: E-mail: schmitt@med.uni-rostock.de, Telephone: +49 381 494 8408

by using non-quadratic penalizer-functions. In histologic brain sectioning, colored cells are in high contrast to their surrounding tissue and have no coherent and equivalently colored spot in the neighbored slices. So these cell-spots may be seen as outliers, resp. "material noise". They are very important from an anatomical point of view, but may mislead the registration process, if not handled accordingly, i.e. by a penalizing process.

2. METHODS

To compensate for rotational and translational differences and to get a better starting position for the non-linear process, a linear pre-registration is performed. For the non-linear part of the registration process, we use a variational formulation in order to come up with a fast implementation.¹⁴

Each digitized slice from a series of M slices can be represented by a mapping

$$R^{(\nu)} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \nu = 1, \dots, M.$$

Let the whole series of images denoted by

$$R := \left(R^{(1)}, \dots, R^{(M)} \right).$$

Then the goal of registration is to find transformations $\varphi^{(\nu)} := \varphi^{(\nu)}(x) = x - u^{(\nu)}(x)$, based on the displacement fields

$$u := \left(u^{(1)}(x), \dots, u^{(M)}(x) \right), \quad u^{(\nu)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \nu = 1, \dots, M,$$

such that neighbored regions become coherent again. In a variational setting this amounts in the minimization of a functional

$$\mathcal{J}[R; u] := \mathcal{D}[R; u] + \alpha \mathcal{S}[u], \quad (1)$$

consisting of a distance- or similarity-measure (for instance the SSD, extended to image series) and a regularizer (smoother),

$$\mathcal{S}[u] := \sum_{\nu=1}^M \int_{\Omega} \frac{\mu}{4} \sum_{j,k=1}^2 \left(\partial_{x_j} u_k^{(\nu)} + \partial_{x_k} u_j^{(\nu)} \right)^2 + \frac{\lambda}{2} \left(\operatorname{div} u^{(\nu)} \right)^2 dx, \quad (2)$$

based on the linear elastic potential of the displacement u . Its influence is scaled by a positive factor α . The scalars λ and μ are material constants which code the elastic behavior of the underlying tissue.

As mentioned before, the goal of this paper is to supplement the common SSD similarity measure

$$\mathcal{D}^0[R, u] := \frac{1}{2} \sum_{\nu=2}^M \int_{\Omega} \left(R^{(\nu)} \circ \varphi^{(\nu)} - R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 dx \quad (3)$$

by two distance measures \mathcal{D}^1 and \mathcal{D}^2 , such that the scheme is more robust towards noise and is invariant with respect to intensity-inhomogeneities. \mathcal{D}^0 measures the difference of the intensity values in the image. In regions where the intensities vary due to illumination or staining, a pure SSD based measure would detect false differences. To overcome the problem we introduce image derivative based SSD measures \mathcal{D}^1 and \mathcal{D}^2 , respectively. More precisely,

$$\mathcal{D}^1[R, u] := \frac{1}{2} \sum_{\nu=2}^M \int_{\Omega} \left| \nabla R^{(\nu)} \circ \varphi^{(\nu)} - \nabla R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right|^2 dx, \quad (4)$$

where $|\cdot|$ is the Euclidean norm and

$$\mathcal{D}^2[R, u] := \frac{1}{2} \sum_{\nu=2}^M \int_{\Omega} \left(\Delta R^{(\nu)} \circ \varphi^{(\nu)} - \Delta R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 dx, \quad (5)$$

with Laplacian Δ .

These different measures can be combined to a single distance measure, which then inherits the advantages of all participating quantities. Each part of the combined measure is weighted by a positive factor γ_i , $i = 0, 1, 2$. For normalization reasons and to fulfill the properties of a measure, the factors have to sum up to one $\sum_{i=0}^2 \gamma_i = 1$. This leads to the *higher in order sum of squared differences measure*

$$\begin{aligned} \mathcal{D}[R, u] := & \frac{1}{2} \sum_{\nu=2}^M \int_{\Omega} \gamma_0 \left(R^{(\nu)} \circ \varphi^{(\nu)} - R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 \\ & + \gamma_1 \left| \nabla R^{(\nu)} \circ \varphi^{(\nu)} - \nabla R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right|^2 \\ & + \gamma_2 \left(\Delta R^{(\nu)} \circ \varphi^{(\nu)} - \Delta R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 dx. \end{aligned} \quad (6)$$

In order to arrive at a functional (1) that is robust under noise, we introduce a non-quadratic increasing penalizer-function $\Psi^{12,13}$ within the new measure (6)

$$\begin{aligned} \mathcal{D}[R, u] := & \frac{1}{2} \sum_{\nu=2}^M \int_{\Omega} \Psi \left[\gamma_0 \left(R^{(\nu)} \circ \varphi^{(\nu)} - R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 \right. \\ & + \gamma_1 \left| \nabla R^{(\nu)} \circ \varphi^{(\nu)} - \nabla R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right|^2 \\ & \left. + \gamma_2 \left(\Delta R^{(\nu)} \circ \varphi^{(\nu)} - \Delta R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 \right] dx. \end{aligned} \quad (7)$$

In our case, we have chosen $\Psi(s^2) := \sqrt{s^2 + \varepsilon^2}$, which leads to a quasi L_1 -minimization. Note, the factor $\varepsilon \ll 1$ is introduced only for numerical reasons and is not an additional parameter. Other penalizer-functions like Perona/Malik¹⁵ are also feasible. If we take a closer look at (7) and choose $\gamma = (\gamma_0, \gamma_1, \gamma_2) = (1, 0, 0)$ and $\Psi(s^2) = s^2$ the robust-higher-order-SSD reduces to the common SSD measure.

Now the functional \mathcal{J} is completely defined. To minimize $\mathcal{J}[R; u]$ with respect to u , we simply compute a stationary point of its Gâteaux-derivative, i.e. $d\mathcal{J}[R; u] = 0$. This results in a system of non-linear partial differential equations, the Navier-Lamé equations (NLE) for serially connected slices (see Wirtz, et al¹)

$$\sum_{\nu=1}^M \left(f^{(\nu)} - \alpha \left(\mu \Delta u^{(\nu)} + (\lambda + \mu) \nabla \operatorname{div} u^{(\nu)} \right) \right) = 0. \quad (8)$$

Equation (8) describes the elastic deformation of an object subject to the force $f^{(\nu)}$. Here, $f^{(\nu)}$ is nothing but the derivative of the chosen distance measure \mathcal{D} ,

$$\begin{aligned} f^{(\nu)} := & \tilde{\Psi} \cdot \\ & \left(\gamma_0 \nabla R^{(\nu)} \circ \varphi^{(\nu)} \left(R^{(\nu-1)} \circ \varphi^{(\nu-1)} - 2R^{(\nu)} \circ \varphi^{(\nu)} + R^{(\nu+1)} \circ \varphi^{(\nu+1)} \right) \right. \\ & + \gamma_1 \nabla \nabla R^{(\nu)} \circ \varphi^{(\nu)} \left(\nabla R^{(\nu-1)} \circ \varphi^{(\nu-1)} - 2\nabla R^{(\nu)} \circ \varphi^{(\nu)} + \nabla R^{(\nu+1)} \circ \varphi^{(\nu+1)} \right) \\ & \left. + \gamma_2 \nabla \Delta R^{(\nu)} \circ \varphi^{(\nu)} \left(\Delta R^{(\nu-1)} \circ \varphi^{(\nu-1)} - 2\Delta R^{(\nu)} \circ \varphi^{(\nu)} + \Delta R^{(\nu+1)} \circ \varphi^{(\nu+1)} \right) \right), \end{aligned} \quad (9)$$

where $\nabla \nabla$ denotes the Hessian (second order derivative) and

$$\begin{aligned} \tilde{\Psi} := & \Psi' \left[\gamma_0 \left(R^{(\nu)} \circ \varphi^{(\nu)} - R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 + \gamma_1 \left| \nabla R^{(\nu)} \circ \varphi^{(\nu)} - \nabla R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right|^2 \right. \\ & \left. + \gamma_2 \left(\Delta R^{(\nu)} \circ \varphi^{(\nu)} - \Delta R^{(\nu-1)} \circ \varphi^{(\nu-1)} \right)^2 \right]. \end{aligned}$$

For the chosen penalizer-function $\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}$, the derivative is easy to compute, $\Psi'(s^2) = 2s/\sqrt{s^2 + \varepsilon^2}$. Now the role of ε is obvious.

We linearize the non-linear partial differential equation (8) by a fixed-point-type iteration method and approximate the NLE by finite differences, which then turns into a linear system of equations with $2N$ unknowns, where N denotes the number of pixels of a single slice. This linear system has to be solved for every slice in each iteration step. The changes in the similarity-measure and the extension to image-series do not affect the numerical minimization scheme presented in *Fischer & Modersitzki*^{6, 14} and can be implemented straight forward. The main and only difference in the scheme is the calculation of the force. It is just the right hand side of the linear system of equations and does not affect the system-matrix. In conclusion, the algorithm still has an asymptotic complexity of $\mathcal{O}(MN \log N)$ by linear (in N) memory usage. For additional speed-up and in order to handle large deformations we used on top a multi-resolution approach based on a Gaussian pyramid.^{16, 17}

3. RESULTS

As test-data for the new approach we used two different data types. Firstly, we registered two two-dimensional MRI images from a human calf and secondly, we registered a subset of images from a series of histologic images of a mouse brain.

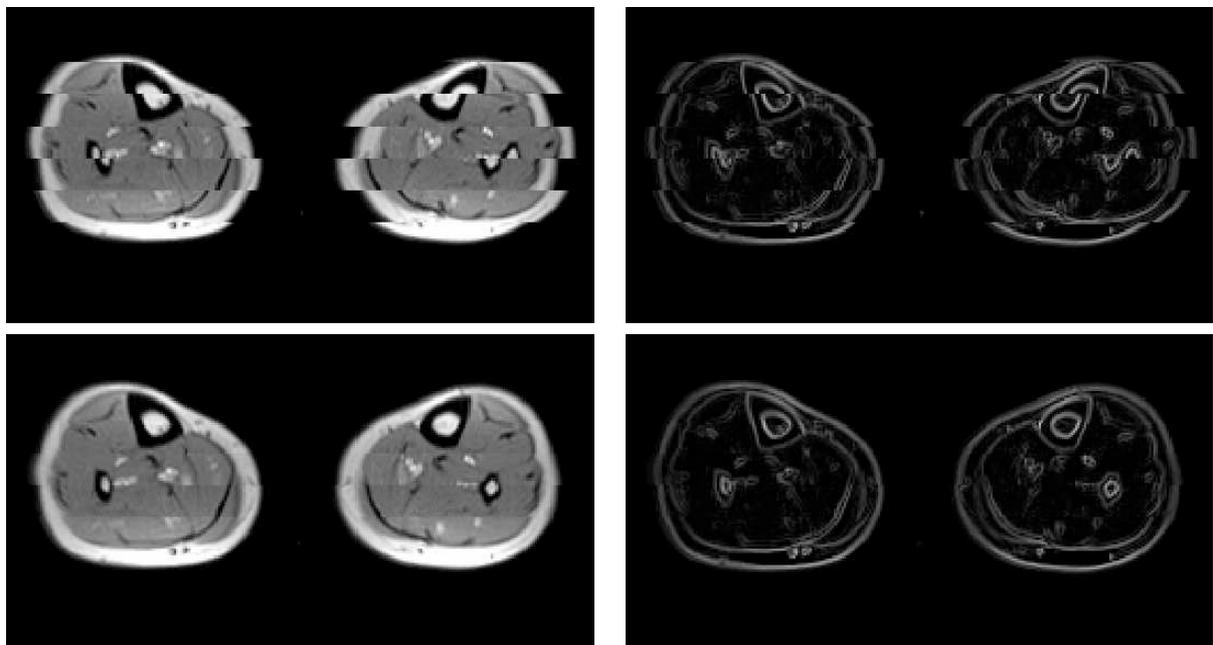


Figure 1. *upper left:* template and reference image *upper right:* Euclidean norm of gradient of template and reference image *lower left:* deformed template and reference image after registration *lower right:* Euclidean norm of gradient of template and reference after registration

The challenge in the dataset from the human calf was the contrast agent. The used reference and template are illustrated in figure 1 in the upper left. To enforce the difference both images are presented in a checkerboard-like plot. In the reference image, no contrast agent was given, so especially the vessels were barely visible. Beside this, in the template image the vessels were in high contrast because of the contrast agent. Furthermore the mean intensity value of, for instance, in the muscle tissue, differs between reference and template, so that common SSD measure does not fulfill the expectations. These differences disappear in the gradient images (see figure 1, upper right), because the muscle tissue is still a homogeneous region in spite of the contrast agent.

The results of the registration is illustrated in the lower row of figure 1. We used the gradient distance measure \mathcal{D}^1 ($\gamma = (0, 1, 0)$) to receive these images, so only gradient information drove the registration process. Figure 1 lower right shows the gradient of the deformed template and the reference after the registration. The

result fulfills the expectations of an expert. In the gray-value-plot, one can observe that the intensity of the muscle tissue is preserved during the registration process.

In registering a series of images from histologic slices the challenge is not only the huge amount of data but the inhomogeneities from staining the tissue (see figure 2 first row). Coherent areas in different slices are represented with completely different intensity values. The like is within a slice. Here again, the common SSD measure works not as expected. It caused large artificial geometry changes like torsions (see figure 2 second row). When using common SSD, the dataset has to be homogenized (see figure 3 first column), which is very time-consuming.

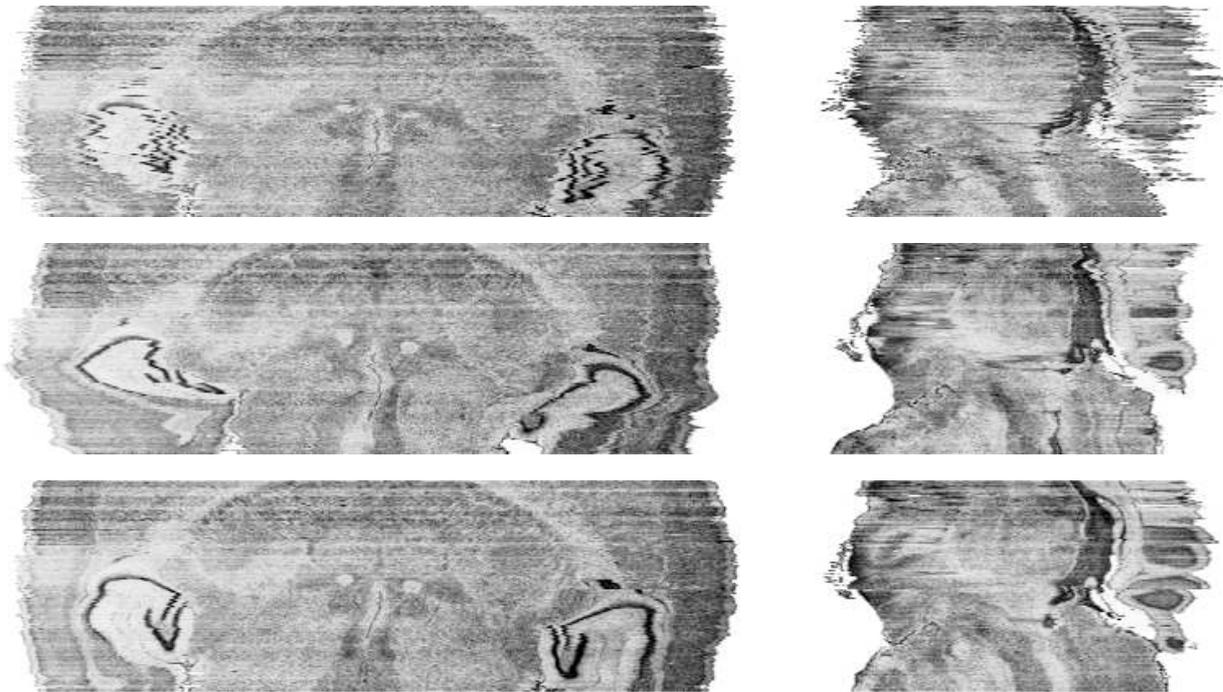


Figure 2. Orthogonally to the direction of sectioning resampled slices from the non-homogenized data. A column of this image represents a line of a slice. First row before registration, second row after registration with common SSD, third row after registration with extended SSD using $\gamma = (1, 100, 0)$ and the penalizer-function. Note, the initial fuzzy looking areas are transformed into recognizable structures. Common SSD caused large artificial geometry deformations. Cortex much smoother and geometry more preserved in using extended SSD. Results are not optimized. Only 44 steps in registration process done.

One of the main advantages of the introduced scheme is that a homogenization of the intensity values is not necessary. Especially the cortex and light in contrast inner structures are well reconstructed (see figure 2 third row and figure 3 second row). The cortex of the brain is one of the most difficult parts in registering images from histologic slices, because in the image acquisition process most deformations from cutting and staining act very badly on the cortex.

Initially the surface of the brain is extremely fuzzy. White and grey matter in the cortex are very close together and intersect from one slice to another. The often and high change in contrast hinders the reconstruction with common SSD. The used image-gradient-based distance measure with the penalizer-function could handle these contrast-changes much better. It led to a much smoother surface and structures from the cortex were coherent again.

On homogenized data the extended SSD lead to advancements especially in the cortex. Geometry deformations were reduced and the brain surface became smoother (see figure 3).

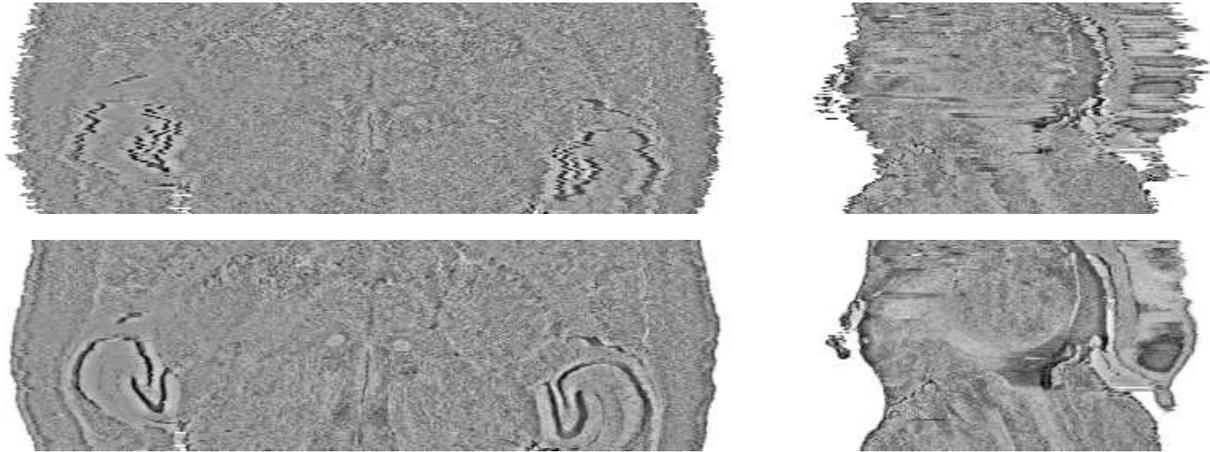


Figure 3. Orthogonally slices from homogenized data. First row before, second row after registration. Registration with $\gamma = (0.9, 0.1, 0)$. If the data is first homogenized, registration quality increases. Extended SSD caused better results especially in cortex and light inner structures.

4. CONCLUSION

To our best knowledge, we have presented in this paper the first registration scheme based on higher order image derivatives. Furthermore we transferred the idea of robust measures from motion analysis to image registration. This scheme enables us to handle both, a huge amount of data and inhomogeneities in the intensity values. We proposed a specific variational technique accompanied by strategies from optical flow based upon image-derivatives.

In using the variational formulation, a fast and efficient algorithm was derived. The applicability of the introduced scheme was proven for registration of images with contrast agents and the registration of a histologically stained image series. The received results are very promising and indicate that the outlined approach is capable for these kinds of problems. In future work this has to be investigated continually.

ACKNOWLEDGMENTS

This work is partly funded by the Deutsche Forschungsgemeinschaft (DFG) via Graduate School 357: "Efficient Algorithms and Multiscale Methods". We thank MeVis, Center for Medical Diagnostic Systems and Visualization, Bremen, Germany, for providing the data from the human calf.

REFERENCES

1. S. Wirtz, J. Modersitzki, B. Fischer, and O. Schmitt, "Super-fast elastic registration of histologic images of a whole rat brain for three-dimensional reconstruction," in *SPIE Medical Imaging 2004*, 2004.
2. C. Broit, *Optimal Registration of Deformed Images*. PhD thesis, Computer and Information Science, University of Pennsylvania, 1981.
3. R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Comp. Vision, Graphics and Image Process.* **46**, pp. 1–21, 1989.
4. G. Christensen, *Deformable shape models for anatomy*. PhD thesis, Sever Institute of Technology, Washington University, 1994.
5. J. Gee, D. Haynor, L. Le Briquer, and R. Bajcsy, "Advances in elastic matching theory and its implementation," in *CVRMed*, pp. 63–72, 1997.
6. J. Modersitzki, *Numerical Methods for Image Registration*, Oxford University Press, 2004.
7. H.-H. Nagel and W. Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**, pp. 565–593, 1986.

8. L. Alvarez, J. Weickert, and J. Sánchez, “Reliable estimation of dense optical flow fields with large displacements,” **39**, pp. 41–56, Aug. 2000.
9. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” in *Computer Vision – ECCV 2004, Part IV*, T. Pajdla and J. Matas, eds., **3024**, pp. 25–36, Springer, Berlin, 2004.
10. S. Uras, F. Girosi, A. Verri, and V. Torre, “A computational approach to motion perception,” **60**, pp. 79–87, 1988.
11. M. Tistarelli, “Multiple constraints for optical flow,” in *Computer Vision – ECCV ’94*, J.-O. Eklundh, ed., **800**, pp. 61–70, Springer, Berlin, 1994.
12. M. J. Black and P. Anandan, “The robust estimation of multiple motions: parametric and piecewise smooth flow fields,” **63**, pp. 75–104, Jan. 1996.
13. E. Mémin and P. Pérez, “A multigrid approach for hierarchical motion estimation,” in *Proc. Sixth International Conference on Computer Vision*, pp. 933–938, Narosa Publishing House, (Bombay, India), Jan. 1998.
14. B. Fischer and J. Modersitzki, “Fast inversion of matrices arising in image processing,” *Numerical Algorithms* **22**, pp. 1–11, 1999.
15. P. Perona and J. Malik, “Scale space and edge detection using anisotropic diffusion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**, pp. 629–639, 1990.
16. B. Jähne, *Digital Image Processing*, Springer, 4th ed., 1997.
17. J. Kybic and M. Unser, “Multidimensional elastic registration of images using splines,” in *ICIP*, **II**.