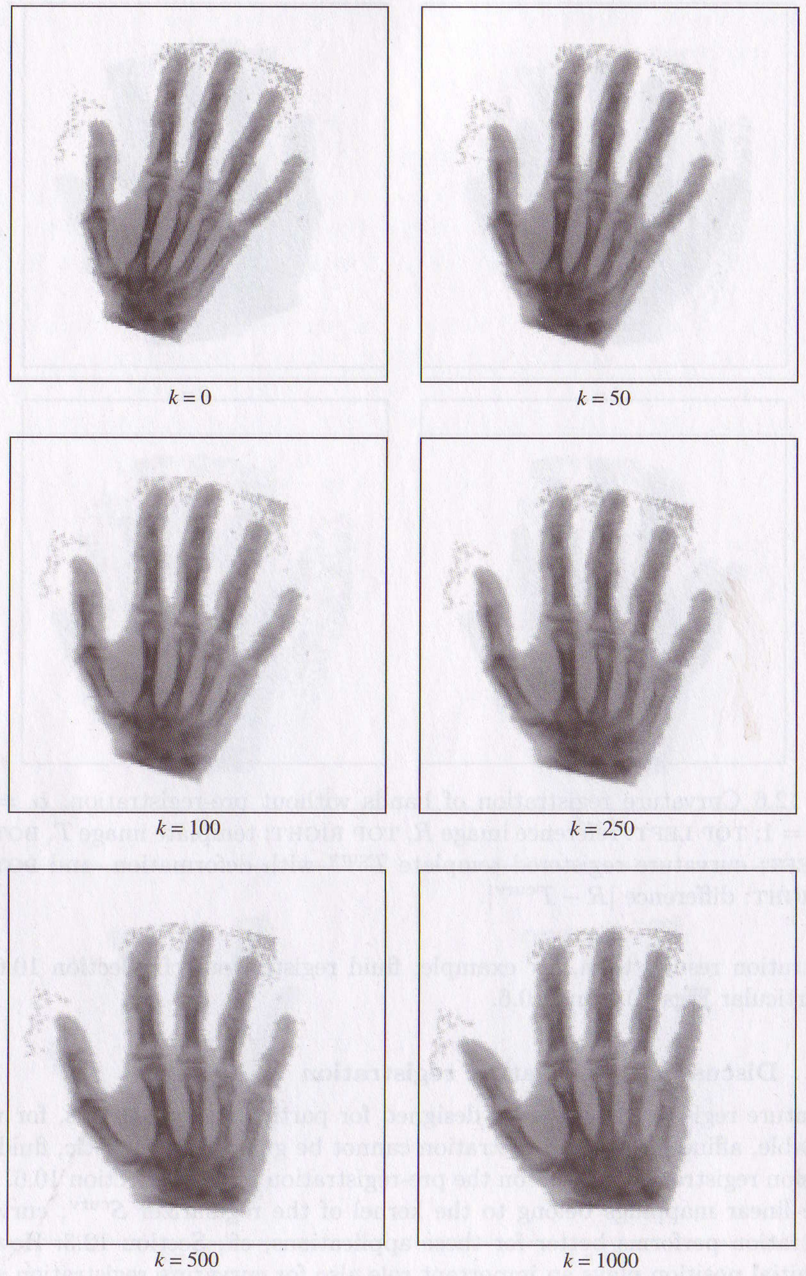

FIG. 12.7 Curvature registration of hands without pre-registration ($\alpha = 50$, $\tau = 0.5$), intermediate results for $k = 0, 50, 100, 250, 500, 1000$.

13

CONCLUDING REMARKS

The purpose of this chapter is to briefly summarize the proposed non-parametric registration approach as well as the presented schemes and to briefly discuss some of the open questions. The unified approach to non-parametric image registration is presented in Chapter 8. The basic ingredients are a distance measure $\mathcal{D}$ and a regularizer $\mathcal{S}$. The distance measure is the driving force of the registration and the regularizer is used to restrict the transformation to an appropriate class. The required displacement u is a solution of

$$\mathcal{D}[R, T; u] + \alpha\mathcal{S}[u] = \min,$$

with some additional boundary conditions. Here, we used

$$\mathcal{D}[R, T; u] := \mathcal{D}[R, T; u] := \tfrac{1}{2}\,\|T_u - R\|_{L_2(\Omega)}\,;$$

cf., Definition 6.1 (see also Theorem 8.1). For the regularizer $\mathcal{S}$, we discussed four particular choices which then lead to the elastic, fluid, diffusion, and curvature registration schemes.

The numerical treatment presented is based on the corresponding Euler–Lagrange equations, which lead to a system of non-linear partial differential equations, $\mathcal{A}[u] = f(\,\cdot\,, u)$; cf., eqn (8.5). A fixed-point or time marching algorithm is used in order to circumvent the non-linearity of f with respect to u. After appropriate finite difference approximations of the time-discrete equations, we ended up with an iterative scheme, where in each step a large system of linear equation has to be solved. Particular solution techniques are designed to reduce the numerical complexity of the overall iterative schemes.

In Section 13.1, we briefly summarize the elastic, fluid, diffusion, and curvature registration schemes. In Section 13.2, we comment on timings obtained in a particular environment. A remark on how to compare the different approaches is given in Section 13.4. We also comment on the choice of the parameters λ, μ and α, τ, respectively (cf., Section 13.5), on further acceleration techniques of the schemes (cf., Section 13.6), and on extensions (cf., Section 13.7).

13.1 Summary of non-parametric image registration

Elastic registration

- Regularization is based on the elastic potential,

$$\mathcal{S}^{\text{elas}}[u] = \mathcal{P}[u] = \int_\Omega \frac{\mu}{4} \sum_{j,k=1}^{d} (\partial_{x_j} u_k + \partial_{x_k} u_j)^2 + \frac{\lambda}{2} (\operatorname{div} u)^2 \, dx;$$

$$\mathcal{A}^{\text{elas}}[u] = \mu \Delta u + (\lambda + \mu) \nabla \operatorname{div} u.$$

- Two parameters, the Lamé constants μ and λ describe material properties, (α may be incorporated), and typically $\lambda = 0$ (maximal expansion of the elastic body).
- Physically meaningful transformation.
- Transformation is restricted to small, local deformations.
- Fast $\mathcal{O}(N \log N)$ implementation based on FFT-type factorization for periodic boundary conditions.

Fluid registration

- Regularization is based on the elastic potential of the time derivative of the displacement, $\mathcal{S}^{\text{fluid}}[u] = \mathcal{P}[\partial_t u]$;

$$\mathcal{A}^{\text{elas}}[v] = f(\,\cdot\,, u), \quad v = \frac{d}{dt} u = \partial_t u + \nabla u \; v.$$

- Two parameters, the Lamé constants μ and λ describe material properties, (α may be incorporated), and typically $\lambda = 0$ (maximal expansion of the fluid).
- Additional Euler step.
- Physically meaningful transformation.
- Extremely flexible.
- Fast $\mathcal{O}(N \log N)$ implementation based on FFT-type factorization for periodic boundary conditions.

Diffusion registration

- Regularization is based on first order spatial derivatives of the displacement,

$$\mathcal{S}^{\text{diff}}[u] = \frac{1}{2} \sum_{\ell=1}^{d} \int_\Omega \langle \nabla u_\ell, \nabla u_\ell \rangle_{\mathbb{R}^d} \, dx;$$

$$\mathcal{A}^{\text{diff}}[u] = \Delta u.$$

- Two parameters, the regularizing parameter α and time-step τ.
- Transformation is restricted to small deformations.

- Can be combined with the fluid idea to allow for larger deformations.
- $\mathcal{O}(N)$ implementation based on an AOS scheme.

Curvature registration

- Regularization based on second order spatial derivatives of the displacement,

$$\mathcal{S}^{\text{curv}}[u] = \frac{1}{2} \sum_{\ell=1}^{d} \int_\Omega (\Delta u_\ell)^2 \, dx;$$

$$\mathcal{A}^{\text{curv}}[u] = \Delta^2 u.$$

- Two parameters, the regularizing parameter α and time-step τ.
- Kernel contains affine linear transformations.
- Transformation is restricted to small deformations.
- Can be combined with the fluid idea to allow for larger deformations.
- Fast $\mathcal{O}(N \log N)$ implementation based on DCT-type techniques.

Note that the regularization of elastic, diffusion, and curvature registration can be applied to the displacement as well as to the update of the displacement. For elastic and fluid registration one may also use a time marching implementation whereas for diffusion and curvature registration one may also use a fixed-point iteration.

13.2 Timings for non-parametric registrations

The timings for two-dimensional image registrations of images of various sizes are summarized in Table 13.1, see also Fig. 13.1. To be precise, we present timings obtained for the solution of one linear system of equations for the respective registration schemes. Note that this is the time consuming part of the computations. The different implementations are by no means optimized. However, the implementations lead to a fair comparison of the different solution schemes.

For the linear systems of equations arising in the elastic registration we implemented four different solution schemes: a direct solver using MATLAB's \\-operator (cf., Math Works (1992)), the conjugate gradient (CG) method of Hestenes Stiefel (1952), a V-cycle of a multigrid method (MG; cf., e.g., Hackbusch (1993, §10.4)), and the FFT-type solution scheme developed in Section 9.5. In our MG implementation, the coarsest grid is 1-by-1, and two pre-smoothing steps and one post-smoothing step with a Jacobi smoother and relaxation parameter $\theta = 2/3$ are performed.

As is apparent from Table 13.1 and Fig. 13.1, the direct solution scheme is not applicable for images of more than 64×64 pixels. The reason is that this direct solution schemes requires additional storage for the triangular factor of an LU decomposition of A; cf., eqn (9.24). Since A is a $2N$-by-$2N$ matrix with $N = n_1 n_2$, the U factor requires $16N^2$ Bytes storage, e.g., 4096 MBytes for

Table 13.1 Timings for the solution of one linear system of equations arising in non-parametric registrations. The timings are given in CPU seconds and are obtained on a PC with Intel Pentium III, 1000 MHz, 256 KBytes cache, and 256 MBytes RAM. The operating system used is SuSE LINUX 8.0 and the computations were performed under MATLAB 6.0.0.88 (R12); $*$ indicates out of memory or computation stopped after about 30 minutes; see also Fig. 13.1.

Image	Elastic registration				Diffusion registration			Curvature registration	
size	Direct	CG	MG	FFT	Direct	DCT	AOS	Direct	DCT
16^2	0.0710	0.0930	0.0510	0.0010	0.0060	0.0050	0.0030	0.0130	0.0060
32^2	0.8490	0.1210	0.0670	0.0020	0.0260	0.0020	0.0120	0.0740	0.0020
64^2	11.1880	0.2460	0.1050	0.0120	0.1570	0.0160	0.0490	0.6520	0.0110
128^2	$*$	1.0620	0.2770	0.0680	1.3430	0.0960	0.2240	6.4390	0.0920
256^2	$*$	4.4710	1.0760	0.3740	10.9040	0.6230	0.9360	72.2980	0.6240
512^2	$*$	17.1200	4.1980	1.5100	$*$	4.1290	3.7940	$*$	4.1850
1024^2	$*$	67.1450	16.8190	6.3980	$*$	29.2050	14.8570	$*$	28.7030

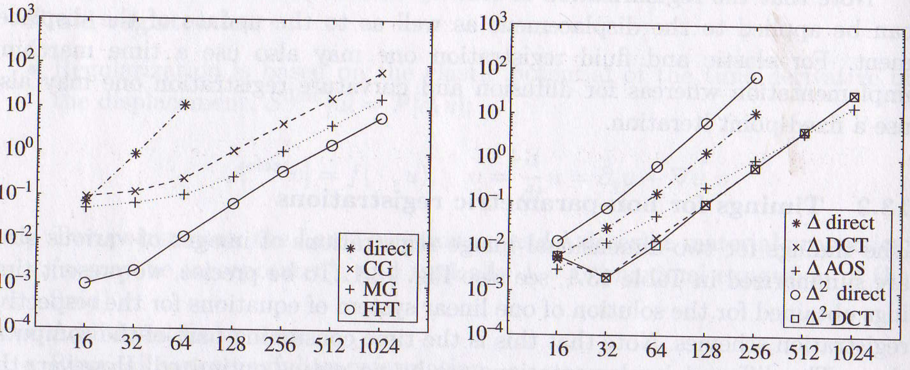

FIG. 13.1 Timings for the solution of one linear system of equations arising in non-parametric registrations. Timings are given in CPU seconds; LEFT: elastic registration with direct, CG, MG, and FFT solver, RIGHT: diffusion and curvature registration with direct, DCT-, and AOS-based solution schemes; see also Table 13.1.

$N = 128^2$. In addition, the direct solver does not work with Neumann or periodic boundary conditions, because of the singularity of the linear systems.

The CG iteration was stopped when the relative residual was brought below a user-supplied threshold ($\mathrm{tol}_{\mathrm{CG}} = 10^{-2}$), or the number of iterations increases beyond a user-supplied number ($k_{\max} = 100$). Unfortunately, in all our experiments the first criterion never applied. With this crude stopping rule, the complexity of CG is $\mathcal{O}(N)$; however, the constant is quite large.

For the linear MG, we applied only one V-cycle. The complexity of the linear MG is also $\mathcal{O}(N)$. However, the constant is better than the one for the CG iteration.

Finally, the FFT-based solution scheme has only complexity $\mathcal{O}(N \log N)$. Note that for small values of N (including all values presented in Table 13.1) this scheme is superior to the others. It is also worthwhile noticing that the FFT-based solution scheme produces exact solutions for the linear systems of equations, while the CG and MG schemes give only approximations. For the FFT-based scheme we have an additional initialization phase for the computation of the entries of the pseudo-inverse of $F^{\mathrm{H}}AF$; cf., Theorem 9.9.

For diffusion registration, we implemented a direct solver using MATLAB's \-operator (cf., Math Works (1992)), the DCT-based factorization (cf., Section 11.2), and an AOS-based solution scheme (cf., Section 11.3). Note that the system matrix for the diffusion registration is $I_d \otimes A^{\mathrm{diff},d}$ (cf., eqn (11.4)) and thus block diagonal. Moreover, from eqn (11.8) we have that $A^{\mathrm{diff},d}$ is a block tridiagonal matrix with a small bandwidth. However, the bandwidth grows with image size.

The direct solver is only applicable for small image sizes (up to 256^2) because of memory requirements. For the DCT-based solution scheme we have an initialization phase for the computation of $D^{(d)}$; cf., Theorem 11.3. For small values of N (up to $N = 512^2$) the $\mathcal{O}(N \log N)$ implementation performs faster than the AOS-based $\mathcal{O}(N)$ scheme.

Finally, we implemented a direct solver using MATLAB's \-operator (cf., Math Works (1992)), and a DCT-based factorization (cf., Section 12.1) for curvature registration. Since $A^{\mathrm{curv},d} = (A^{\mathrm{diff},d})^{\top} A^{\mathrm{diff},d}$, the bandwidth of the curvature matrix is larger than the bandwidth of the diffusion matrix. Thus, in accordance with our measurements, we expect that more time is required for the curvature than for the diffusion direct solution scheme. On the other hand, for the DCT-based solution scheme we expect exactly the same time for the diffusion and curvature registration schemes. For both registration schemes, a DCT, a multiplication with a diagonal matrix, and an inverse DCT have to be computed. The only difference is the entries of the diagonal matrices. Our numerical experiments almost reflect this expectation.

Summarizing, it turned out that for elastic and fluid registration the FFT-based direct solution scheme provides an attractive alternative to iterative solution schemes like CG and MG. In particular, for moderate image sizes, the FFT-based technique outperforms its competitors.

For diffusion registration, the DCT-based solution scheme is preferable for small image sizes while the AOS-based solution scheme is the fastest scheme for large-scaled images. The DCT-based solution scheme is at present the only competitive scheme for curvature registration.

Note that the FFT-based scheme is faster than the AOS-based solution scheme for all image sizes considered in Table 13.1. This is partly related to the excellent implementation of the FFT routine provided by Matlab (1992).

13.3 A competitive example: hands

In this section we present competitive results for elastic, fluid, diffusion, and curvature registration. We register the reference image and the linearly pre-registered template shown in Fig. 9.10.

For elastic registration, we chose $\mu = 500$ and $\lambda = 0$ and used the FFT-based solution scheme for a fixed-point-type implementation. For fluid registration we chose $\mu = 500$ and $\lambda = 0$, and used the FFT-based solution scheme for the velocity and the Euler step presented in Section 10.4.1. For diffusion registration we chose $\alpha = 5$ and $\tau = 0.25$, and for curvature registration we chose $\alpha = 50$ and $\tau = 0.5$. For diffusion registration we used a time marching AOS-based implementation and for curvature registration a time marching DCT-based implementation.

All registrations started with $u^{(0)} = 0$ and were stopped as soon as

$$\frac{\mathcal{D}^{\mathrm{SSD}}[R, T; u^{(k)}]}{\mathcal{D}^{\mathrm{SSD}}[R, T]} \le 0.6.$$

This stopping criterion was met for the values $k^{\mathrm{elas}} = 33, k^{\mathrm{fluid}} = 19, k^{\mathrm{diff}} = 12$, and $k^{\mathrm{curv}} = 11$. Note that these values depend on the choice of the parameters and therefore cannot be compared directly.

Figure 13.2 illustrates the displacements for the four techniques and Fig. 13.3 displays some details. As is apparent from these figures, all four registration schemes produced visually pleasing results. The differences between the four schemes are rather small in this example. In particular, the differences between elastic and fluid registration are hardly visible. As expected, diffusion registration yields a less smooth displacement field, because the regularizer is first order derivative based. Note that the results of the registration schemes may differ considerably for other examples; see, e.g., the • to C registration. However, if the images are very similar, i.e., show only minor differences, the results of the registration schemes are also very similar.

13.4 How to compare apples with pears

It has already been pointed out that image registration is an ill-posed problem. In particular, in Section 12.3 we presented two different registrations of two images of a square. Since the difference between the reference and the deformed template is zero for both registrations, both approaches are reasonable and equally likely. Without additional, external information, we are not able to decide which of the two registrations is to be preferred. In fact both are equally wrong, because the template image is a rotated copy of the reference (see also the discussion in Section 12.3).

Thus, without additional, external knowledge, registration is a risky task. At this point regularization enters into the picture. Regularization not only is a mathematical technique to overcome ill-posedness of a problem, but in

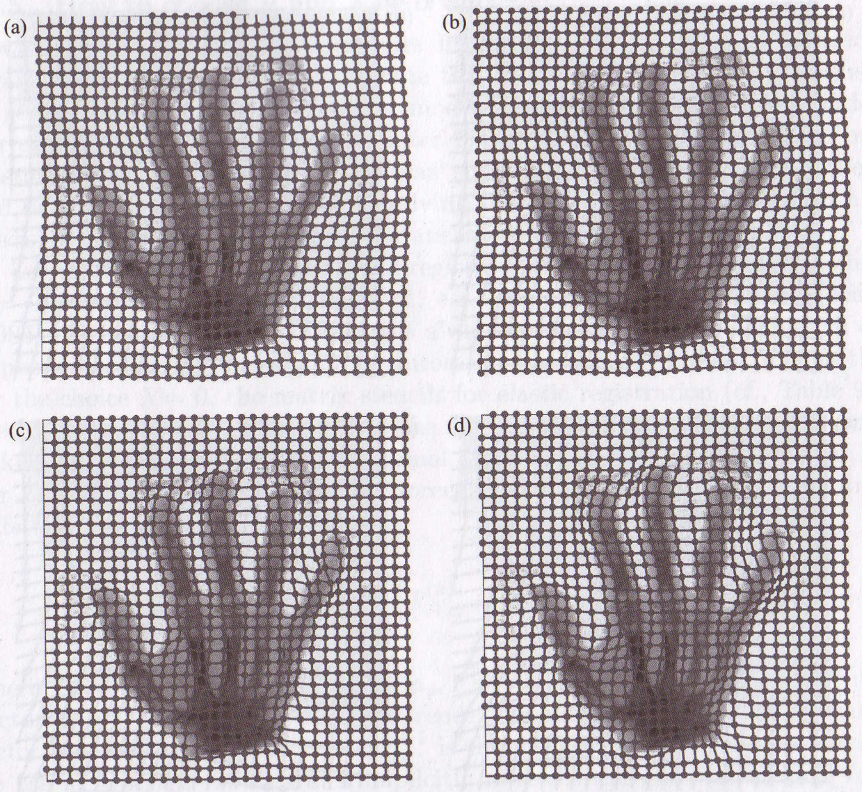

FIG. 13.2 Results of elastic (a), fluid (b), diffusion (c), and curvature (d) registration.

image registration, it also presents a way of providing external information. For example, if we know that the distortions of an image are mostly due to elastic deformations, we should use a regularizer based on the elastic potential. However, if we only have a vague idea about the source of the observed distortions, the elastic registration scheme may turn out to be the wrong choice. In this situation, a straightforward idea would be to apply the most flexible scheme in the hope of reducing the distance measure at most. Note that this flexibility enables one to register very different images; cf., e.g., Section 10.7. Though this truly is a very attractive option for some applications, this capability is frightening for others. For example, it is easy to obtain visually pleasing results for the registration of two different human brains even when one of them has been rotated by 90 or 180 degrees by accident.

Though the differences between the different regularizers might be large for some registrations (i.e., the • to C registration), they are in general small for

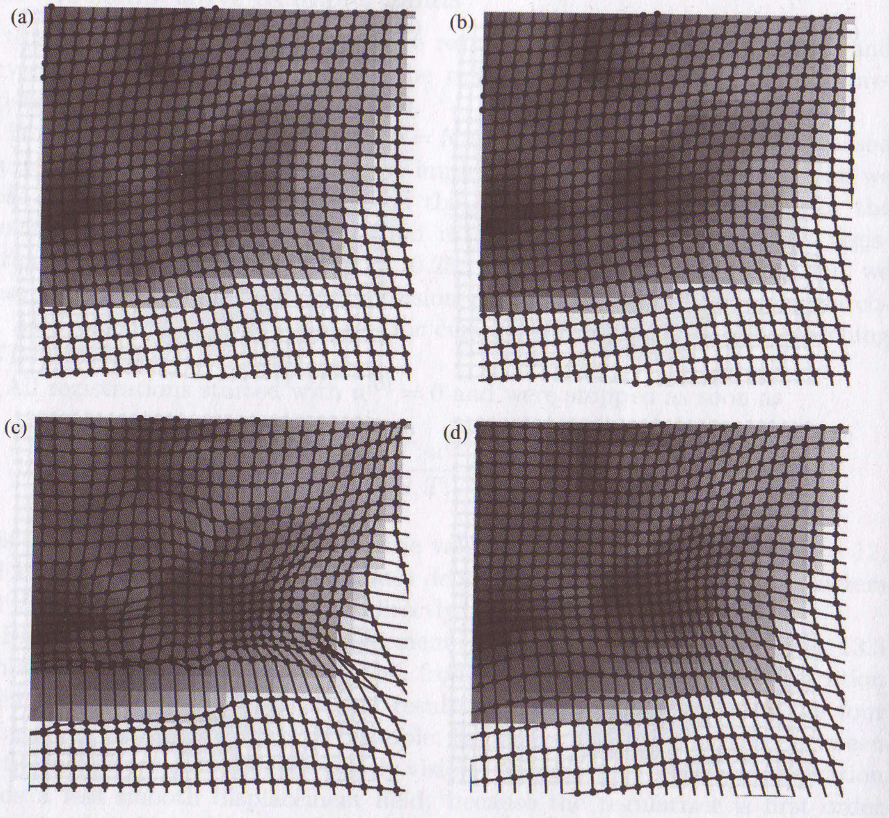

FIG. 13.3 Results of elastic (a), fluid (b), diffusion (c), and curvature (d) registration, details.

closely related images; cf., Section 13.3. Thus, one may argue that it is just a matter of taste which regularizer is to be used. However, depending on the application even small differences may be of importance and one always should use the best available model.

Another important question is, which distance measure should be used? Again the answer depends strongly on the particular application. Here, using a particular and appropriate distance measure can also help to reduce unwanted local minima; see also the discussion in Section 6.6. Moreover, one may also use the distance measure as an important modeling tool in order to provide additional, external knowledge to the registration procedure.

The solution to all these questions is probably to provide a variety of different registration techniques and to enable one to choose the most promising technique for a particular application.

13.5 How to choose μ and λ or α and τ

How does one choose the parameters in non-parametric image registration? Again this is a delicate question. Note that in image registration the focus is in general not on a perfect physical model but on a meaningful deformation. For example, in the HNSP (cf., Chapter 2) one has to deal not only with pure elastic deformations of the paraffin wax embedded tissue but also with therm and dehydration processes. Hence, solving the Navier–Lamé equations to high precision is probably not the appropriate way.

Let us start with elastic and fluid registration. In our experiments, we chose $\lambda = 0$, which is a common choice; cf., e.g., Broit (1981) or Bajcsy & Kovačič (1986). The choice of μ in this book is always explicitly stated. However, in our implementation we generally use an automatic procedure to choose μ. Note that for the choice $\lambda = 0$, the matrix stencils for elastic registration (cf., Table 9.1) are linear with respect to μ. Since the scaling of the force field is in general unknown, there is always one additional degree of freedom in our system. For our particular implementation, the forces are computed using a central finite difference approximation, i.e.,

$$F_i^{k,\ell} = \gamma_\ell\Big(R_i - T_i^{(k)}\Big)\Big(T_{i+e_\ell}^{(k)} - T_{i-e_\ell}^{(k)}\Big), \quad 1 \le \ell \le d,$$

where $i = (i_1, \ldots, i_d) \in \mathbb{N}^d, 1 \le i_p \le n_p, 1 \le p \le d$, and e_ℓ denotes the ℓ^{th} unit vector in $\mathbb{R}^d$. For a central finite difference, the standard choice for the scaling factor is $\gamma_\ell = (2h_\ell)^{-1}$, where $h_\ell = n_\ell^{-1}$ is the meshsize. In our implementation, we use $\gamma_\ell = n_\ell/(2 \cdot 255^2)$ which implicitly re-scales the gray value range from $[0, 255]$ to $[0, 1]$. Note that the displacement is computed from the force field via a linear system of equations. Thus, we are able to provide a heuristic for a clever choice of μ which we briefly describe.

To estimate μ, we define a tolerance for the maximal displacement of a pixel (we use $u_{\max} = 2$ in our computations). In the first step, we set $\mu = 1$ and $\lambda = 0$. We then compute the forces, solve the linear system of equations for a displacement $\hat{u}$, and compute the norm of this displacement. The estimator for μ is simply $\mu = u_{\max}/\|\hat{u}\|_V$ (cf., eqn (10.7)). We re-scale the matrix A^{elas} and set $u^{(1)} = \mu\hat{u}$. With this choice and the re-scaling of the displacement, we have $\|u^{(1)}\|_V = u_{\max}$. We take this value of μ during the computations but monitor the norm of the displacements. If $\|u^{(k)}\|_V > u_{\max}$, we re-scale the actual displacement such that the norm of the re-scaled displacement becomes $u_{\max}$. However, the value of μ is not altered. In our experiments (not contained in this book), this procedure works surprisingly well.

In principle the same remarks and ideas apply also to the time marching implementation for diffusion and curvature registration. Here, roughly speaking, the linear systems

$$(I - \tau\alpha A)\vec{U}^{(k+1)} = \vec{U}^{(k)} + \tau\vec{F}^{(k)}$$

are to be solved. Note that on the right hand side τ compromises between the current displacements and the current forces. Thus, a re-scaling of the forces also affects the choice of τ. Unfortunately, the matrix $I - \tau\alpha A$ is not simply linear in α or τ. Thus, a direct application of the above strategy cannot be recommended.

Another interesting strategy is to relax the regularizer during iteration; see also Henn (1997). For example, for the time marching implementation, we may use $\alpha = \alpha(t)$ where $\alpha(t) \to 0$ for $t \to \infty$. The motivation is that the regularizer should guide the registration to the required minimizer of the distance measure. Getting closer to this minimizer reduces the impact of the regularizer. Of course, it is also possible to take $\alpha = \alpha(x,t)$, which has in principle already been discussed in Section 11.6.2 for the demons registration. Note that for this choice, the matrices arising from the discrete formulation cannot be factorized explicitly.

13.6 Further accelerations of the schemes

Multi-resolutions and parallelizing techniques present further possibilities for a reduction of computation time. The first step of a multi-resolution procedure is to provide a pyramid of images with decreasing resolution. Starting with the coarsest resolution, one computes the transformation with relatively low computational costs. The coarse resolution transformation is prolongated onto the next finer grid and serves as a typically excellent starting point for the registration at the finer resolution. Because of this good starting point, one expects lower computational costs for corrections on the finer grid. This procedure is continued until the finest resolution is reached. Note that this is also the basic idea behind multigrid techniques; cf., e.g., Hackbusch (1993, §10.4), Henn (1997), or Schormann et al (1996).

For the three-dimensional reconstruction of the human brain presented in Section 9.10, we used a Gaussian pyramid approach. With this approach the computational time for a two-dimensional registration of two 1024^2 pixel images could be reduced from about 20 minutes to about 2 minutes.

Furthermore, a multi-resolution approach not only reduces the computational costs but also serves as an additional regularizer. The resolution hierarchy also introduces a hierarchy for the images differences. Major differences are visible in all resolutions while minor differences appear only at finer scales. In this sense, the registration process is guided. Essential differences are removed on the corse grid which also prevents the scheme stagnating in local minima.

Parallelizing techniques provide a further possibility for accelerating the computations. Unfortunately, these techniques have to be adapted to the particular computational environment as well as to the particular schemes. Thus, a description is in general very technical and therefore we refer to the literature: for elastic registration see, e.g., Christensen et al (1996) or Modersitzki et al (2001), for FFT-based elastic registration see Böhme et al (2002), and for AOS-based diffusion registration see Bolten et al (2003).

It should be noted that also higher order approximations for the non-linear forces (cf., eqn (8.11)) can be considered. But, as has already been pointed out in Chapter 6, the numerical computation of higher order derivatives can be critical in image processing.

13.7 Extensions

The proposed unified approach to image registration can be extended in various directions. For example, a symmetrized distance measure

$$\mathcal{D}^{\text{sym}}[R,T;\varphi] = \mathcal{D}[R,T;\varphi] + \mathcal{D}[T,R;\varphi^{-1}]$$

(see, e.g., Christensen & Johnson (2001)) may be treated analogously. This can be of interest if the template image is not a deformed version of an ideal (as, e.g., in the HNSP; cf., Chapter 2) but reference and template images are on a par. Moreover, one may also introduce ψ as an approximation to φ^{-1}. A variational formulation might be based on the functional

$$\mathcal{D}[R,T;\varphi] + \mathcal{D}[T,R;\psi] + \alpha\mathcal{S}[\varphi,\psi],$$

where $\mathcal{S}[\varphi,\psi]$ penalizes the deviation between ψ and φ^{-1}, e.g.,

$$\mathcal{S}[\varphi,\psi] = \int_\Omega \Big(1 - \det(\varphi \circ \psi(x))\Big)^2 \, dx,$$

or

$$\mathcal{S}[\varphi,\psi] = \int_\Omega \|\varphi(x) - \psi^{-1}(x)\|^2_{\mathbb{R}^d} + \|\varphi^{-1}(x) - \psi(x)\|^2_{\mathbb{R}^d} \, dx.$$

For other applications, like, for example, the registration of MRI scans from different phases of a contrast agent uptake, it can be of particular importance to maintain the volume of the tissue. This goal can be achieved, for example, by using the additional regularizer

$$\mathcal{S}^{\text{vol}}[\varphi] = \int_\Omega \det(\nabla u) \, dx.$$

Note that $\det(\nabla\varphi) = 1 + \det(\nabla u)$ and the transformation maintains the volume, if $\det(\nabla\varphi) = 1$.

Recent research is also concerned with the integration of interpolation constraints to non-parametric image registration. Here, the idea is to supply a number of landmarks which are of particular interest for the underlying application. The required transformation should minimize the joint functional subject to some interpolation constraints or extended by an additional penalizing term which particularly enforces a match of the landmarks.

REFERENCES

Abram, K. A. (2000). Registration of images with dissimilar contrast using a hybrid method employing correlation and mutual information. Technical report, Dept. of Computer Science, Dartmouth College, Hannover.

Aldroubi, A. and Gröchening, K. (2001). Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Review 43*(3), 585–620.

Alpert, N. M., Bradshaw, J. F., Kennedy, D., and Correia, J. A. (1990). The principal axes transformation – A method for image registration. *Journal of Nuclear Medicine 31*(10), 1717–1722.

Amit, Y. (1994). A nonlinear variational problem for image matching. *SIAM Journal of Scientific Computing 15*(1), 207–224.

Bajcsy, R. and Kovačič, S. (1986). Toward an individualized brain atlas elastic matching. Technical Report MS-CIS-86-71 Grasp Lap 76, Dept. of Computer and Information Science, Moore School, University of Philadelphia.

Böhme, M., Hagenau, R., Modersitzki, J., and Siebert, B. (2002). Non-linear image registration on PC-clusters using parallel FFT techniques. Preprint, Institute of Mathematics, University of Lübeck, Germany.

Bolten, M., Modersitzki, J., and Papenberg, N. (2003). A parallel implementation of the AOS-based diffusion registration. Preprint, in preparation, Institute of Mathematics, University of Lübeck.

Boor, C. De (1978). *A practicle guide to splines.* Springer, New York.

Braess, D. (1980). *Nonlinear approximation theory,* Volume 7 of *Springer Series in Computational Mathematics.* Springer.

Brewer, J. W. (1978). Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems 25,* 772–780.

Broit, C. (1981). *Optimal Registration of Deformed Images.* Ph.D. thesis, Computer and Information Science, University of Pensylvania.

Bro-Nielsen, M. (1996). *Medical Image Registration and Surgery Simulation.* Ph.D. thesis, IMM, Technical University of Denmark.

Bro-Nielsen, M. and Gramkow, C. (1996). Fast fluid registration of medical images. *Lecture Notes in Computer Science, 1131,* 267–276, Springer, Berlin and Heidelberg.

Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys 24*(4), 325–376.

Buck, R. C. (1978). *Advanced Calculus.* McGraw-Hill, New York.

Čapek, M. (1999). Optimisation strategies applied to global similarity based image registration methods. In *WSCG '99: the 7th International Conference in Central Europe on Computer Graphic,* pp. 369–374.

Christensen, G. E. (1994). *Deformable Shape Models for Anatomy.* Ph.D. thesis, Sever Institute of Technology, Washington University.

Christensen, G. E. and Johnson, H. J. (2001). Consistent image registration. *IEEE Transactions on Medical Imaging 20*(7), 568–582.

Christensen, G. E., Miller, M. I., Grenander, U., and Vannier, M. W. (1996). Individualizing neuroanatomical atlases using a massively parallel computer. *IEEE Computer*, 32–38.

Collignon, A., Vandermeulen, A., Suetens, P., and Marchal, G. (1995). 3d multi-modality medical image registration based on information theory. *Computational Imaging and Vision 3*, 263–274, Kluwer Academic, Dordrechth.

Collins, D. L. and Evans, A. C. (1997). Animal: Validation and applications of nonlinear registration-based segmentation. *International Journal of Pattern Recognition and Artificial Intelligence 11*(8), 1271–1294.

Curtis, P. C. (1958). n-parameter families and best approximation. *Pacific Journal of Mathematics 58*, 1013–1027.

Davis, P. J. (1979). *Circulant Matrices.* Chelsea Publishing, New York.

Duchon, J. (1976). Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *RAIRO Analyse Numérique 10*, 5–12.

Duda, R. and Hart, P. (1973). *Pattern Classification and Scene Analysis.* Wiley.

Dümbgen, L. (2001). Private communication.

Fischer, B. and Modersitzki, J. (1999). Fast inversion of matrices arising in image processing. *Numerical Algorithms 22*, 1–11.

Fischer, B. and Modersitzki, J. (2003). Curvature based image registration. *Journal of Mathematical Imaging and Vision 18*(1), 81–85.

Fletcher, R. (1987). *Practical Methods of Optimization.* Wiley.

Folland, G. B. (1995). *Introduction to partial differential equations* (2nd edn). Princeton University Press, Princeton, New Jersey.

Gaens, T., Maes, F., Vandermeulen, D., and Suetens, P. (1998). Non-rigid multimodal image registration using mutual information. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI '98* (ed. W. M. Wells, A. Colchester, and S. Delp), pp. 1099–1106. Springer.

Golub, G. H. and van Loan, C. F. (1989). *Matrix Computations* (2nd edn). The Johns Hopkins University Press, Baltimore, Maryland.

Gonzalez, R. C. and Woods, R. E. (1993). *Digital Image Processing.* Addison-Wesley, Reading, Massachusetts.

Gurtin, M. E. (1981). *An Introduction to Continuum Mechanics.* Academic Press, Orlando, Florida.

Gurtin, M. E. (1983). *Topics in Finite Elasticity.* SIAM, Philadelphia, Pennsylvania.

Hackbusch, W. (1987). *Partial Differential Equations.* Teubner, Stuttgart.

Hackbusch, W. (1993). *Iterative Solution of Large Sparse Systems of Equations.* Springer, New York.

Henn, S. (1997). *Schnelle elastische Anpassung in der digitalen Bildverarbeitung mit Hilfe von Mehrgitterverfahren.* Ph.D. thesis, Heinrich-Heine-Universität, Düsseldorf.

Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards 49*, 409–436.

Horn, B. K. P. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence 17*, 185–204.

Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis.* Cambridge University Press, Cambridge.

Horn, Roger A. and Johnson, Charles R. (1991). *Topics in Matrix Analysis.* Cambridge University Press, Cambridge.

Keeling, S. L. and Ring, W. (2002). Medical image registration and interpolation by optical flow with maximal rigidity. SFB Report No. 248, Insitute of Mathematics, Karl-Franzens-Universität Graz.

Kent, J. T. and Tyler, D. E. (1988). Maximum likelihood estimation for the wrapped cauchy distribution. *Journal of Applied Statistics 15*(2), 247–254.

Kent, J. T. and Tyler, D. E. (1991). Redescending m-estimates of multivariate location and scatter. *Annals of Statistics 19*(4), 2102–2119.

Kim, B., Boes, J. L., Frey, K. A., and Meyer, C. R. (1997). Mutual information for automated unwarping of rat brain autoradiographs. *Neuroimage 5*, 31–40. Article No. N1960251.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics 22*, 79–86.

Light, W. A. (1995). Variational methods for interpolation, particularly by radial basis functions. *Numerical Analysis*, 94–106.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging 16*(2), 187–198.

Maintz, J. B. A. and Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis 2*(1), 1–36.

Mairhuber, J. C. (1956). On Haar's theorem concerning Chebychev approximation problems having unique solutions. *Proceedings of the AMS 7*, 609–615.

Mardia, K., Kent, J., and Bibby, J. (1979). *Multiavariant Analysis.* Academic Press, New York.

MathWorks (1992). *MATLAB User's Guide.* MathWorks, Natick, Massachusetts.

Maurer, C. R. and Fitzpatrick, J. M. (1993). A review of medical image registration. In *Interactive Image-Guided Neurosurgery*, pp. 17–44. American Association of Neurological Surgeons, Park Ridge.

Meihe, X., Srinivasan, R., and Nowinski, W. L. (1999). A fast mutual information method for multi-modal registration. In *Information processing in medical imaging* (ed. A. Kuba, M. Šámal, and A. Todd-Pokropek), pp. 466–471. Springer Berlin and Heidelberg.

Modersitzki, J., Lustig, G., Schmitt, O., and Obelöer, W. (2001). Elastic registration of brain images on large PC-clusters. *Future Generation Computer Systems 18*, 115–125, Elsevier, Amsterdam.

Nocedal, J. and Wright, S. J. (1999). *Numerical optimization.* Springer, New York.

Pennec, X., Cachier, P., and Ayache, N. (1999). Understanding the "demon's algorithm" 3D non-rigid registration by gradient descent. In *Medical image computing and computer assisted intervention* (ed. C. Taylor and A. Colchester), pp. 597–605. Springer, Berlin and Heidelberg.

Piegl, L. and Tiller, W. (1997). *The NURBS book* (2nd edn). Springer.

Roche, A., Malandain, G., Ayache, N., and Prima, S. (1999). Towards a better comprehension of similarity measures used in medical image registration. In *Medical image computing and computer assisted intervention* (ed. C. Taylor and A. Colchester), pp. 555–566. Springer, Berlin and Heidelberg.

Rohr, K. (2001). Landmark-based image analysis. *Computational Imaging and Vision.* Kluwer Academic, Dordrecht.

Schaback, R. (1997). Reconstruction of multivariate functions from scattered data. `http://www.num.math.uni-goettingen.de/schaback/teaching.html`.

Schmitt, O. (2001). *Die multimodale Architektonik des menschlichen Gehirns.* Habilitation, Insitute of Anatomy, Medical University of Lübeck.

Schormann, T., Henn, S., and Zilles, K. (1996). A new approach to fast elastic alignment with applications to human brains. *Lecture Notes in Computer Science 1131*, 337–342, Berlin and Heidelberg.

Schormann, T. and Zilles, K. (1997). Limitations of the principal-axes theory. *IEEE Transactions on Medical Imaging 16*(6), 942–947.

Shannon, C. E. (1949). Communications in the present of noise. *Proceedings of the IRE 37*, 10–21.

Studholme, C., Hill, D. L. G., and Hawkes, D. J. (1996). Automated 3-D registration of MR and CT images of the head. *Medical Image Analysis 1*(2), 163–175.

Thévenaz, P., Blu, T., and Unser, M. (2000). Image interpolation and resampling. In *Handbook of Medical Imaging, Processing and Analysis* (ed. I. Bankman), pp. 393–420. Academic Press, San Diego, California.

Thévenaz, P., Ruttimann, U. E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing 7*(1), 27–41.

Thirion, J.-P. (1995). Fast non-rigid matching of 3D medical images. Technical Report 2547, Institut National de Recherche en Informatique et en Automatique, France.

Thirion, J.-P. (1998). Image matching as a diffusion process: An analogy with Maxwell's demons. *Medical Image Analysis 2*(3), 243–260.

Thomas, L. H. (1949). Elliptic problems in linear difference equations over a network. Technical report, Watson Scientific Computing Laboratory, Columbia University, New York.

van den Elsen, P. A., Pol, E.-J. D., and Viergever, M. A. (1993). Medical image matching – a review with classification. *IEEE Engineering in Medicine and Biology*, 26–38.

Viola, P. A. (1995). *Alignment by Maximization of Mutual Information*. Ph.D. thesis, Massachusetts Institute of Technology.

Weickert, J. (1998). *Anisotropic diffusion in image processing*. Teubner, Stuttgart.

Wells III, W. M., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R. (1996). Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis 1*(1), 35–51.

Wollny, G. and Kruggel, F. (2002). Computational cost of non rigid registration algorithms based on fluid dynamics. *IEEE Transactions on Medical Imaging, 21*(8), 946–952.

GLOSSARY

Norms and such

$[x]$	$[x] := \min\{j \in \mathbb{Z} \mid j - 0.5 \le x < j + 0.5\}$.
$\lceil x \rceil$	$\lceil x \rceil := \min\{j \in \mathbb{Z} \mid j \ge x\}$.
$\lfloor x \rfloor$	$\lfloor x \rfloor := \max\{j \in \mathbb{Z} \mid j \le x\}$.
$\lVert \cdot \rVert_2$	For $A \in \mathbb{C}^{n\times n}$, $\lVert A \rVert_2 := \max\{\sqrt{\lambda} : \lambda$ is eigenvalue of $A^{\mathrm{H}} A\}$.
$\lVert \cdot \rVert_f$	norm on a feature space.
$\lVert \cdot \rVert_{\mathbb{R}^d}$	norm on $\mathbb{R}^d$, $\lVert x \rVert_{\mathbb{R}^d} := \sqrt{\langle x, x\rangle_{\mathbb{R}^d}}$.
$\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$	inner product on $\mathbb{R}^d$, $\langle x, y\rangle_{\mathbb{R}^d} := y^\top x$.
$\langle \cdot, \cdot \rangle_0$	$\langle f, g\rangle_0 := \int_{\mathbb{R}^d} f(x)g(x)\,dx$.
$\langle \cdot, \cdot \rangle_{L_2}$	$\langle f, g\rangle_{L_2} := \int_{\mathbb{R}^d} f(x)g(x)\,dx$.
$\langle \cdot, \cdot \rangle_q$	semi-inner product, $\langle f, g\rangle_q := \sum_{\lvert\kappa\rvert = q} c_\kappa \, \langle D^\kappa f, D^\kappa g\rangle_0$.

Attributes

$\dagger$	$B^\dagger$ Moore–Penrose pseudo-inverse.
H	Hermitian, for $A = (a_{j,k})_{j,k} \in \mathbb{C}^{n\times n}$, $A^{\mathrm{H}} = (\overline{a_{k,j}})_{j,k}$.
$\perp$	$x \perp y \iff \langle x, y\rangle_{\mathbb{R}^d} = 0$.
$\top$	transpose, for $A = (a_{j,k})_{j,k} \in \mathbb{R}^{n\times n}$, $A^\top = (a_{k,j})_{j,k}$.
$\vec{\ }$	$\vec{X} \in \mathbb{R}^N$, recollection of $X \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ in lexicographical ordering, $N = n_1 \cdots n_d$.

Derivatives

∇^2	Hessian matrix $\nabla^2 f = (\partial_{x_j, x_k} f)_{j,k} \in \mathbb{R}^{d\times d}$.
∇	$\nabla f = (\partial_{x_1} f, \dots, \partial_{x_d} f)^\top$ for functions $f : \mathbb{R}^d \to \mathbb{R}$, $\nabla \varphi = (\partial_{x_k} \varphi_j)_{j,k} \in \mathbb{R}^{d\times d}$ for vector fields $\varphi : \mathbb{R}^d \to \mathbb{R}^d$.
∂_{x_j}	partial derivative with respect to x_j.

Operators

$*$	convolution.
$\otimes$	Kronecker product.
$\times$	Cartesian product of spaces, also cross-product in $\mathbb{R}^3$.

Sets

$\partial\Omega$	boundary of $\Omega =]0, 1[^d$.

A

AOS	additive operator splitting.

C	
$\mathbb{C}$	complex numbers.
c_B	center of an image, $c_B := \mathbb{E}_B[x] \in \mathbb{R}^d$.
CCD	charge-coupled device.
C_m	cosine matrix, $C_m := \left(\cos \frac{(2j+1)k\pi}{2m} \right)_{j,k=0,\dots,m-1} \in \mathbb{R}^{m\times m}$, also: basic circulant matrix, $C_m \in \mathbb{R}^{m\times m}$.
Corr	correlation, $\mathrm{Corr}_{R,T}(y) := \int_{\mathbb{R}^d} R(x)T(x-y)\,dx$.
Cov$_B$	covariance, $\mathrm{Cov}_B := \mathbb{E}_B\left[(x-c_B)(x-c_B)^\top\right] \in \mathbb{R}^{d\times d}$.
D	
$d\mathcal{F}[\varphi;\psi]$	Gâteaux derivative.
d	spatial dimension.
$\mathcal{D}^{\mathrm{corr}}$	correlation, $\mathcal{D}^{\mathrm{corr}}[R,T] := \left\langle \frac{R-\mu(R)}{\sigma(R)}, \frac{T-\mu(T)}{\sigma(T)} \right\rangle_{L_2}$.
DCT	discrete cosine transformation.
d/dt	total time derivative.
Δ	Laplace operator, $\Delta f = \sum_{j=1}^{d} \partial_{x_j,x_j} f$, $\Delta u = (\Delta u_1, \dots, \Delta u_d)^\top$ for a vector field u.
δ_x	point evaluation functional at x, $\delta_x[f] = \int_{\mathbb{R}^d} f(y)\delta_x(y)\,dx = f(x)$.
det	determinant.
div	divergence, $\operatorname{div} v = \langle \nabla, v\rangle_{\mathbb{R}^d} = \sum_{j=1}^d \partial_{x_j} v_j$.
D^κ	$D^\kappa f := \left(\frac{\partial}{\partial x_1}\right)^{\kappa_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\kappa_d} f$, $\kappa \in \mathbb{N}_0^d$.
$\mathcal{D}^{\mathrm{LM}}[\varphi]$	landmark-based distance measure, $\mathcal{D}^{\mathrm{LM}}[\varphi] := \sum_{j=1}^m \lVert \mathcal{F}(R,j) - \varphi(\mathcal{F}(T,j)) \rVert_f^2$.
D_m	$D_m := 2\,\mathrm{diag}(\cos \frac{k\pi}{m},\ k = 0,\dots,m-1) \in \mathbb{R}^{m\times m}$.
$\mathcal{D}^{\mathrm{MI}}$	mutual information, $\mathcal{D}^{\mathrm{MI}}[R,T] = H(\rho_R) + H(\rho_T) - H(\rho_{R,T})$.
d_q	$d_q := \dim(\Pi_{q-1}(\mathbb{R}^d))$.
$\mathcal{D}^{\mathrm{SSD}}$	sum of squared differences, $\mathcal{D}^{\mathrm{SSD}}[R,T] := \frac{1}{2}\lVert T-R\rVert_{L_2}^2 = \frac{1}{2}\int_{\mathbb{R}^d} \left(T(x) - R(x)\right)^2\,dx$.
E	
e_j	j^{th} column of the identify matrix I_d.
$\mathbb{E}_B[\cdot]$	expectation value, $\mathbb{E}_B[f] := \frac{\int_{\mathbb{R}^d} f(x)\,B(x)dx}{\int_{\mathbb{R}^d} B(x)dx}$.

F	
FBS	flat bed scan.
FFT	fast Fourier transformation.
F_n	Fourier matrix, $F_n := n^{-1/2}(\omega_n^{(j-1)(k-1)})_{j,k=1,\dots,n} \in \mathbb{C}^{n\times n}$.
$F(S,j)$	j^{th} feature of an image S.
G	
$\gamma(R,T;y)$	correlation coefficient, $\gamma(R,T;y) := \left\langle \frac{R-\mu(R)}{\sigma(R)}, \frac{T_y-\mu(T_y)}{\sigma(T_y)} \right\rangle_{L_2}$.
H	
HNSP	Human NeuroScanning Project.
$H(\rho)$	entropy, $H(\rho) := -\mathbb{E}_\rho[\log\rho] = -\int_{\mathbb{R}^q} \rho \log \rho\, dg$.
I	
I_d	identity matrix in $\mathbb{R}^{d\times d}$.
i.i.d.	independently and identically distributed.
Img(d)	set of images, $\mathrm{Img}(d) := \{\, b : \mathbb{R}^d \to \mathbb{R} \mid b \text{ is } d\text{-dimensional image} \,\}$.
L	
$L_2(\Omega)$	$L_2(\Omega) := \{\, f : \Omega \to \mathbb{R} : \int_\Omega \lvert f(x)\rvert^2\,dx < \infty \,\}$.
M	
MRI	magnetic resonance imaging.
$\mu(B)$	mean gray value, $\mu(B) := \lvert\Omega\rvert^{-1}\int_\Omega B(x)dx$.
N	
$\mathbb{N}$	$\mathbb{N} = \{1, 2, \dots\}$.
N	$N = n_1 \cdots n_d$, number of grid points.
$\vec{n}$	outer normal vector on a boundary, $\vec{n} = (n_1,\dots,n_d)^\top$, $\lVert\vec{n}\rVert_{\mathbb{R}^d} = 1$.
$\mathbb{N}_0$	$\mathbb{N}_0 = \{0, 1, 2, \dots\}$.
O	
$\mathcal{O}$	shortcut for a function belonging to a class of certain decay, $f : \mathbb{R} \to \mathbb{R}$ is $\mathcal{O}(h^p)$ if $\lim_{h\to 0} f(h)h^{-p-1} = 0$.
Ω	$\Omega =]0,1[^d$.
ω_n	n^{th} root of unity, $\omega_n := e^{-2\pi i/n}$.
Ω_n	$\Omega_n := \mathrm{diag}(\omega_n^0, \dots, \omega_n^{n-1})$.

P

PAT principal axes transformation.

φ transformation $\varphi : \mathbb{R}^d \to \mathbb{R}^d$.

$\Pi_q^d(\mathbb{R}^d)$ d-variate d-dimensional polynomials of degree q
$\Pi_q^d(\mathbb{R}^d) := \{ \varphi : \mathbb{R}^d \to \mathbb{R}^d \mid \varphi_\ell \in \Pi_q(\mathbb{R}^d),\ \ell = 1, \dots, d \}$.

$\Pi_q(\mathbb{R}^d)$ d-dimensional polynomials of degree q
$\Pi_q(\mathbb{R}^d) := \left\{ \psi : \mathbb{R}^d \to \mathbb{R} \mid \psi(x) = \sum_{|\kappa| \le q} \alpha_\kappa x^\kappa,\ \alpha_\kappa \in \mathbb{R} \right\}$.

ppi parts per inch.

$\mathcal{P}[u]$ elastic potential,

$$\mathcal{P}[u] = \int_\Omega \frac{\mu}{4} \sum_{j,k=1}^{d} (\partial_{x_j} u_k + \partial_{x_k} u_j)^2 + \frac{\lambda}{2} (\operatorname{div} u)^2 dx,$$

λ, μ Lamé constants.

R

R reference image.

$\mathbb{R}$ real numbers.

S

$\sigma(B)$ standard deviation, $\sigma(B) := \mu((B - \mu(B))^2)$.

$\Sigma(x,t)$ stress tensor, $\Sigma(x,t) := (\sigma_{j,k}(x,t))_{j,k}$.

SSD sum of squared differences.

$\mathcal{S}^{\mathrm{TPS}}$ $\mathcal{S}^{\mathrm{TPS}}[\psi] := \frac{1}{2} \langle \psi, \psi \rangle_q^2$.

T

T template image.

T_φ deformed template $T_\varphi(x) = T \circ \varphi(x) = T(\varphi(x))$.

TPS thin plate spline.

trace trace of a matrix $A \in \mathbb{R}^{n \times n}$, $\operatorname{trace}(A) = \sum_{j=1}^n a_{j,j}$.

T_u deformed template $T_u(x) = T(x - u(x))$.

U

u displacement, $u(x) = \varphi(x) - x$.

V

v velocity, $v(P,t) = \partial_t u(P,t)$.

V_m $V_m := C_m \operatorname{diag}(\sqrt{1/m}, \sqrt{2/m}, \dots, \sqrt{2/m}) \in \mathbb{R}^{m \times m}$.

$V(x,t)$ strain tensor, $V(x,t) := (\varepsilon_{j,k}(x,t))_{j,k}$.

AUTHOR INDEX

SUBJECT INDEX