

Brownian Warps: A least committed prior for non-rigid registration

M. Nielsen	P. Johansen	A. D. Jackson, B. Laustrup
IT U.	DIKU	NBI
Copenhagen	Copenhagen	Copenhagen
Denmark	Denmark	Denmark

Abstract

Non-rigid registration requires a smoothness or regularization term for making the warp field regular. Standard models in use here include b-splines and thin plate splines. In this paper, we suggest a regularizer which is based on first principles, is symmetric with respect to source and destination, and fulfills a natural semi-group property for warps. We construct the regularizer from a distribution on warps. This distribution arises as the limiting distribution for concatenations of warps just as the Gaussian distribution arises as the limiting distribution for the addition of numbers. Through an Euler-Lagrange formulation, algorithms for obtaining maximum likelihood registrations are constructed. The technique is demonstrated using 2D examples.

1 Introduction

In any non-rigid registration algorithm, one must weigh the data confidence against the complexity of the warp field mapping the source image geometrically into the destination image. This is typically done through spring terms in elastic registration [3, 8, 7], through the viscosity term in fluid registration [5] or by controlling the number of spline parameters in spline-based non-rigid registration [1, 20].

If non-rigid registration algorithms, symmetric in source and destination, can be constructed, many problems in shape averaging and shape distribution estimation can be avoided. The regularizer is not symmetric with respect to source and destination in the methods mentioned above. While symmetric regularizers can be constructed in most cases simply by adding a term for the inverse registration [6], this solution is not theoretically satisfactory. The aim of the present paper is to construct a regularizer that exhibits this symmetry inherently and that is also least committed in the same way as the Gaussian distribution is for the addition of numbers. This will be made precise later.

Section 2 gives more concise definitions, motivation, and states the principles of the problem. Section 3 contains the solution to the problem formulated

in Section 2 and some of its properties. Section 4 describes a gradient descent method for finding the optimal warp given a set of landmark matches. In this way, this paper offers both theoretical considerations and their application. Subsequent development will demonstrate this on real medical data.

2 Definitions and motivation

A non-rigid registration may be modeled by a warp field $W : \mathbb{R}^D \mapsto \mathbb{R}^D$ mapping points in one D -dimensional image into another D -dimensional image. We give the definition:

Definition 1 (Warp Field) *A warp field $W(x) : \mathbb{R}^D \mapsto \mathbb{R}^D$ maps all points in the source image $I_S(x) : \mathbb{R}^D \mapsto \mathbb{R}$ into points of the destination image $I_D(x) : \mathbb{R}^D \mapsto \mathbb{R}$ such that $I_S(W(X))$ is the registered source image. W is invertible and differentiable (i.e., a diffeomorphism) and has everywhere a position Jacobian $\det(\partial_{x_i} W^j)$*

Here, we have made the assumption that non-rigid registrations are invertible and differentiable. This seems valid in cases where images are created from similar structures. In some cases, such as separated bone fractures, this conjecture is not appropriate. However, in nearly all medical cases, a non-rigid registration is made on the basis of anatomical structures of identical topology, and the above definition will apply. A diffeomorphism will always have the same sign of the Jacobian everywhere. Our choice of positive Jacobian applies to those cases where the object is not geometrically reflected.

The identification of a warp field on the basis of images is a matter of inference. Below we will apply the Bayes inference machine [13], but a similar formulation should appear when using information theoretic approaches such as the minimum description length principle [17].

We wish to determine the warp field W that maximizes the posterior

$$p(W|I_S, I_D) = \frac{1}{Z} p(I_S, I_D|W) p(W)$$

where Z is a normalizing constant (sometimes denoted the partition function), $p(I_S, I_D|W)$ is the likelihood term, and $p(W)$ is the prior. The likelihood term is based on the similarity of the warped source and destination image and may, in this formulation, be based on landmark matches [4], feature matches [15, 18], object matches [2], image correlation [15], or mutual information [21]. The subject of this paper is to address the prior $p(W)$ that expresses our belief in the regularity of the warp field prior to identifying the images. In specific medical applications, this may be based on active shape models [9, 16]. However, to construct such models, homology must be created as an in principle dense field, and the present work may also be used in this context.

We wish the prior $p(W)$ to exhibit the specific properties that it is:

- Derived from first principles,

- Least committed,
- Symmetric with respect to source and destination,
- Invariant with respect to warps.

In the following section we will formalize these properties and derive a prior on warps that we will denote Brownian Warps in analogy to Brownian motion.

3 Brownian warps

We seek that distribution of warps which is the analogue of Brownian motion. We wish this distribution to be independent of warps performed earlier (i.e., invariant with respect to warps). This property is of fundamental importance particularly when determining the statistics of empirical warps, creating mean warps etc. In such cases it is required by consistency in order to avoid the use of a fiducial pre-defined standard warp. We may formulate this as:

$$p(W = W_2 \circ W_1) = \int p(W_2) dW_1.$$

This corresponds to the semi-group property of Brownian motion: The distribution of positions after two moves corresponds to two independent moves and, through the central limit theorem, leads to a Gaussian distribution of positions. Since this also holds for a concatenation of many warps, we can construct a warp as

$$W_B = \lim_{N \rightarrow \infty} \prod_{i=0}^N \circ W_i,$$

where the W_i are independent warps. This corresponds exactly to the definition of Brownian motion if the concatenation product is replaced by an ordinary sum.

In order to find this limiting distribution when all W_i are independent, we investigate motion in the neighborhood of a single point following along all the warps and make the following lemma:

Lemma 1 (Local structure) *Let $J_W = \partial_{x_i} W^j$ be the local Jacobian of W . Then, the Jacobian of a Brownian warp*

$$J_{W_B} = \lim_{N \rightarrow \infty} \prod_{i=0}^N J_{W_i}$$

Proof This is obviously true due to the chain rule of differentiation. □

Assume that an infinitesimal warp acts as the infinitesimal independent motion of points. In this case, all entries in the local Jacobian are independent and identically distributed. Hence, we may now model

$$J_{W_B} = \lim_{N \rightarrow \infty} \prod_{i=0}^N \left(I + \sigma \frac{1}{\sqrt{N}} H_i \right), \quad (1)$$

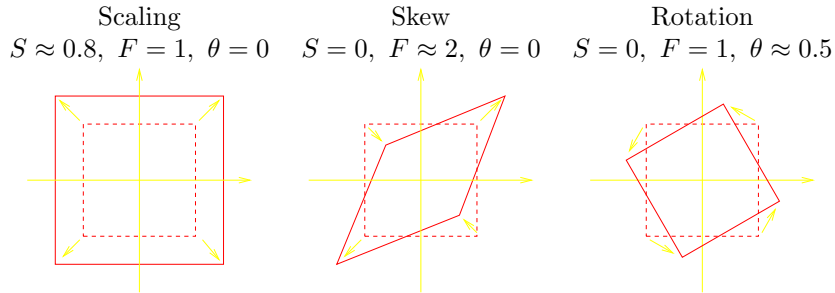


Figure 1: The independent action of the parameters on a unit square.

where H_i is a $D \times D$ matrix of independent identically distributed entries of unit spread. The denominator \sqrt{N} is introduced to make the concatenation product finite, and σ is the spread or the “size” of the infinitesimal warps.

To summarize, the limiting distribution of Eq. 1 is the distribution of the Jacobian of a Brownian Warp. In turn this defines the Brownian distribution on warps, as we have no reason to assume other structure in the distribution.

Unfortunately, the solution to Eq. 1 is not given in the literature on random matrices. Gill and Johansen [10] solve the problem for matrices with positive entries and Högnäs and Mukherjea [12] solve, among other cases, the situation when the matrices are symmetric. Recently, we have solved the case for two dimensions [14] and are presently considering the solution for three. Here, we present only the result.

Theorem 1 (2D Brownian Jacobian) *The limiting distribution of Eq.1 where H_i have independent entries of unit spread and $W : \mathbb{R}^2 \mapsto \mathbb{R}^2$, is given as*

$$p(J_{W_B}) = G(S/\sigma) \sum_{n=0}^{\infty} g_n(F/\sigma) \cos(n\theta), \quad (2)$$

where G is the unit spread Gaussian, g_n are related to the Jacobi functions, and the parameters are given as follows:

$$\text{Scaling} \quad S = \log(\det(J_{W_B}))$$

$$\text{Skewness} \quad F = \frac{1}{2 \det(J_{W_B})} \|J_{W_B}\|_2^2$$

$$\text{Rotation} \quad \theta = \arctan\left(\frac{j_{12} - j_{21}}{j_{11} + j_{22}}\right)$$

It is shown in [14] that the limiting distribution does not depend on features of the infinitesimal distribution other than its spread, σ . This limiting distribution is thus least committed in the sense that it arises from the sole assumption of invariance under warps. The parameter σ may be viewed as a measure of rigidity. The effects of the parameters are shown in Fig. 1.

Now, we prove that the above Brownian warp distribution is invertible and symmetric with respect to source and destination. Evidently, this is true by construction; one can simply invert the infinite multiplication sequence since the final distribution depends only on the spread of the independent infinitesimal warps. However,

Theorem 2 (Invertability) *The distribution of warps given as spatially independent Jacobians each distributed according to Eq. 2 has with probability 1 no folds.*

Proof A fold implies that the local Jacobian of the warp is zero or less. The above distribution has a positive Jacobian with probability 1. \square

Theorem 3 (Symmetry) *The distribution of warps given as spatially independent Jacobians each distributed according to Eq. 2 is invariant under inversion of the warp.*

Proof The inversion of a warp $W \mapsto W^{-1}$ makes the local Jacobians undergo an inversion, too: $J \mapsto J^{-1}$. Under this, the distribution parameters map as $S \mapsto -S$, $F \mapsto F$, and $\theta \mapsto -\theta$. Since the distribution is even in S and θ , it is unaltered under inversion of J . \square

Theorem 4 (Euclidean invariance) *The distribution of warps given as spatially independent Jacobians each distributed according to Eq. 2 is invariant under Euclidean coordinate transformations of source and destination.*

Proof The individual Jacobians will transform as $J \mapsto RJR^{-1}$, where R is a rotation matrix under simultaneous and identical rotation and scaling of source and destination. This transformation leaves all three parameters S, F, θ invariant. \square

One should notice that this invariance holds under *simultaneous* and identical scaling and rotation of source and destination. If one wants to incorporate independent similarity invariance, it is necessary to introduce uniformly distributed global bias parameters in S and θ as done by Glasbey and Mardia [11].

For computational purposes it may be convenient to approximate the above distribution by a distribution which is also independent in F and θ . This can be done in many ways without losing the symmetry and Euclidean invariance. However, warp invariance will no longer hold exactly. We suggest the following approximation.

$$p(J) \approx G_\sigma(S)G_{\sigma/\sqrt{2}}(\theta)e^{-(F/\sigma)^{0.67}} \quad (3)$$

where G_σ is a Gaussian of spread σ . This approximation has a relative error at less than 3% for all reasonable values of S, θ, F when $\sigma > 0.4$.

In Figure 2 the joint distribution of F and θ is illustrated using this approximation and is compared to the analytical expression approximated up to $n = 14$ for $\sigma = 0.3, 0.6, 1.0$. The primary oscillating error seen for small σ is due to the cut off at $n = 14$ in the analytical expression.

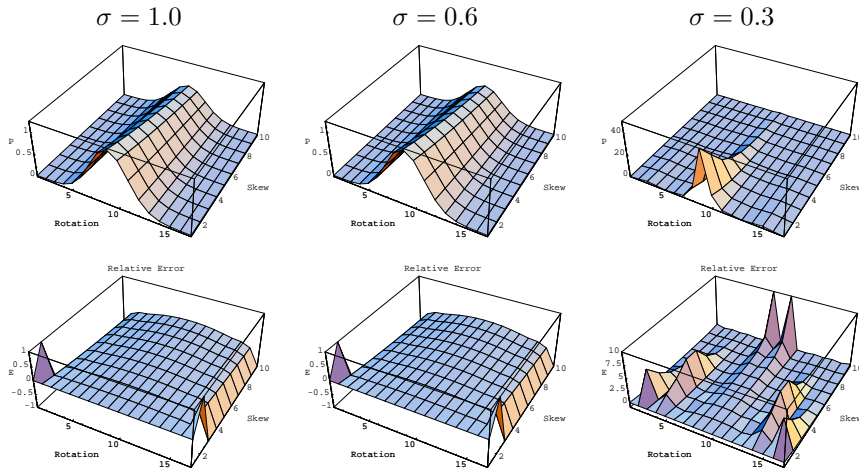


Figure 2: The joint distribution of F and θ for the approximation (top) and pointwise relative difference to the analytical expression approximated up to $n = 14$ (below) for $\sigma = 0.3, 0.6, 1.0$.

4 Implementation

In this section we show how the above distribution can be used for maximum a posteriori (MAP) estimation of the most probable warp given a set of landmark matches.

We reformulate the MAP problem as an energy minimization approach by:

$$E(W) = -\log p(W) + c = \int_{\Omega} S^2 + 2\theta^2 + 2\sigma^{1.33} F^{0.67} d\tilde{x},$$

where c is an arbitrary irrelevant constant and $\tilde{x} = x\sqrt{\det(J)}$ are integration variables invariant under the warp chosen to ensure global as well as local warp invariance. Unfortunately, the related Euler-Lagrange equation is neither linear nor separable, and simple tricks such as eigenfunction expansions and derived linear splines are not possible. Therefore, we treat the energy minimization problem using a gradient back-projection scheme [19].

$$\text{for } x \in \Omega' : \partial_t W = -\frac{\delta E}{\delta W},$$

where Ω' is the image domain excluding the matched landmark points. This may easily be relaxed to matched curves without identified landmarks as in geometry-constrained diffusion [2].

We see from the energy formulation that the rigidity parameter determines the relative weight of the skewness term to the scaling and rotation terms. For illustration of the independent terms, see Fig. 3. For large deformations, the difference to spline-based methods, becomes obvious as for example thin plate splines can introduce folds in the warping (see Fig. 4).

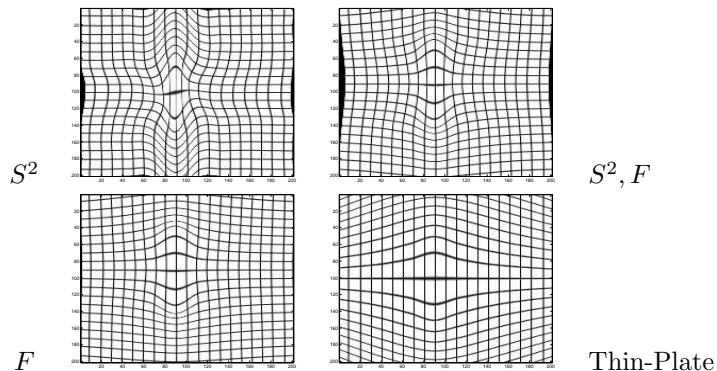


Figure 3: Illustration of deformation of a regular grid. Two points in the center have been moved up and down respectively, while the corners are kept fixed. We see that the scaling term (top left) aims at keeping the area constant. The skewness term (bottom left) aims at keeping the stretch equally large in all directions. Top right is a combination of scaling and skewness ($\sigma = 1$). Bottom right is a thin plate spline for comparison.

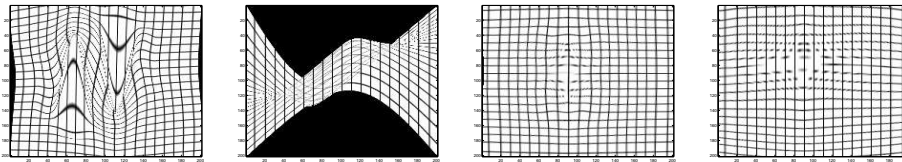


Figure 4: Leftmost are two images of large deformations: Left is the maximum likelihood Brownian warp, right is a thin plate spline. Rightmost two images are two consecutive warps where landmark motions are inverse: Left is Brownian warps, right is thin plate spline. Brownian warps do not give the exact inverse due to numerical imprecision, but closer than the thin plate spline.

5 Conclusion

We have introduced a prior for warps based on a simple invariance principle under warping. This distribution is the warp analogue of Brownian motion for additive actions. An estimation based on this prior guarantees an invertible, source–destination symmetric, and Euclidean-invariant warp. When computational time is of concern, approximations can be made which violate the basic warp invariance while maintaining invertibility, source–destination symmetry, and Euclidean invariance. We suggested one such approximation being where close to the true distribution. For extremely fast implementations, we recommend an approximation including only the skewness term, as this has nice regularizing properties. We have shown computational examples on synthetic data.

Future works includes applications to medical data, the development of algorithms using a mutual information data term, use as void hypothesis in shape

deformations, comparisons to flows in chaotic fluid dynamics, extensions to three dimensions, extensions to spatially higher-order correlated priors, and extensions to fraction Brownian warps.

References

- [1] A.A. Amini, R.W. Curwen, and J.C. Gore, *Snakes and splines for tracking non-rigid heart motion*, ECCV96, 1996, pp. II:251–261.
- [2] Per R. Andresen and Mads Nielsen, *Non-rigid registration by geometry-constrained diffusion*, Medical Image Analysis **6** (2000), 81–88.
- [3] R. Bajcsy and S. Kovacic, *Multiresolution elastic matching*, CVGIP **46** (1989), 1–21.
- [4] F.L. Bookstein, *Morphometric tools for landmark data: Geometry and biology*, Cambridge University Press, 1991.
- [5] M. Bro-Nielsen and C. Gramkow, *Fast fluid registration of medical images*, Proc. Visualization in Biomedical Imaging (VBC'96) (1996), 267–276.
- [6] P. Cachier and D. Rey, *Symmetrization of the Non-Rigid Registration Problem using Inversion-Invariant Energies: Application to Multiple Sclerosis*, Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI 2000) (Pittsburgh, Pennsylvania USA) (A.M. DiGioia and S. Delp, eds.), Lectures Notes in Computer Science, vol. 1935, Springer, octobre 11-14 2000.
- [7] G.E. Christensen and J. He, *Consistent nonlinear elastic image registration*, MMBIA01, 2001, pp. xx–yy.
- [8] G.E. Christensen, M. I. Miller, and M. Vannier, *A 3d deformable magnetic resonance textbook based on elasticity*, AAAI Spring Symposium Series (1994), 153–156, Stanford University.
- [9] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, *Active shape models: Their training and application*, CVIU **61** (1995), no. 1, 38–59.
- [10] R. D. Gill and S. Johansen, *A survey of product-integration with a view toward application un survival analysis*, The annals of statistics **18** (1990), no. 4, 1501–1555.
- [11] C. A. Glasbey and Mardia K. V., *A penalized likelihood approach to image warping*, J. R. Statist. Soc. B **63** (2001), 465–514.
- [12] G. Högnäs and A. Mukherjea, *Probability measures on semigroups*, Plenum Press.
- [13] E. T. Jaynes, *Probability theory: The logic of science*, <http://omega.albany.edu:8008/JaynesBook.html>, Fragmentary Edition of June 1994.
- [14] B. Lautrup, A. Jackson, P. Johansen, and M. Nielsen, *Random maps*, Tech. report, Niels Bohr Institute, University of Copenhagen, 2001.
- [15] J. Maintz and M. Viergever, *A survey of medical image registration*, 1998.
- [16] B.S. Morse, S.M. Pizer, and A. Liu, *Multiscale medial analysis of medical images*, IVC **12** (1994), no. 6, 327–338.
- [17] Jorma Rissanen, *Stochastic complexity in statistical enquiry*, World Scientific Publishing Company, Singapore, 1989.
- [18] K. Rohr, *Landmark-based image analysis: Using geometric and intensity models*, Kluwer, 2001.
- [19] J. B. Rosen, *The gradient projection method for nonlinear programming. part I. linear constraints*, SIAM **8** (1960), no. 1, 181–217.
- [20] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, *Nonrigid registration using free-form deformations: application to breast mr images*, MedImg **18** (1999), no. 8, 712–721.
- [21] P.A. Viola and W.M. Wells, III, *Alignment by maximization of mutual information*, Ph. D., 1995.