Geostatistical Medical Image Registration

Conference Paper *in* Lecture Notes in Computer Science · November 2003

DOI: 10.1007/978-3-540-39903-2_109 · Source: DBLP

5 authors, including:

Juan Ruiz-Alzola
Universidad de Las Palmas de Gran Canaria

Eduardo Suarez-Santana
Instituto Tecnológico de Canarias

Carlos Alberola-López
Universidad de Valladolid

Simon K Warfield
Harvard University

Geostatistical Medical Image Registration*

J. Ruiz-Alzola[1,2], E. Suarez[1], C. Alberola-Lopez[3], S.K. Warfield[2], and C.-F. Westin[2]

[1] Medical Technology Center
Universidad de Las Palmas de Gran Canaria, Spain
[2] Dep. of Radiology, Brigham & Women's Hospital and Harvard Medical School, USA
[3] ETSI Telecomunicacion, Universidad de Valladolid, Spain
{jruiz,eduardo}@ctm.ulpgc.es; westin@bwh.harvard.edu; caralb@tel.uva.es
http://www.ctm.ulpgc.es

Abstract. We propose a novel approach to landmark-based medical image registration based on the geostatical method of Kriging prediction. Our method exploits the spatial statistical relation between two images, as estimated using general-purpose registration algorithms, in order to construct an optimum predictor of the displacement field. High accuracy is achieved by using an estimated spatial model of the displacement field directly from the image data, while practically circumventing the difficulties that prevented Kriging from being widely used in image registration.

1 Introduction

In this paper we propose a geostatistical framework for the registration of medical images. Our motivation is to provide the highest possible accuracy to computer-aided clinical systems in order to estimate the geometric (coordinate) transformation between two multidimensional, possibly multimodal, datasets. Registration of medical (both 2D and 3D) images, from the same or different imaging modalities, is needed by computer-aided clinical systems for diagnosis, pre-operative planning, intra-operative procedures and post-operative follow-up. Registration is also needed to perform comparisons across a population, for deterministic and statistical atlas construction and to embed anatomic knowledge in segmentation algorithms. Good reviews of current state of the art for medical image registration can be found elsewhere (see for example [7]).

Our framework is based on the reconstruction of a dense arbitrary displacement field by interpolating the displacements measured from control points. Several schemes have been proposed in the past to interpolate sparse displacement fields for medical image registration. The most popular ones are based on thin-plate splines, which usually make an independent interpolation for each of the components of the vector field. Interpolating or smoothing thin-plate splines [3,11,12] are used depending on whether the sparse displacements are considered to be noiseless or not. The former need the order of the spline to be specified in advance while the latter also need the regularization parameter to be specified. Adaptiveness can be obtained by spatially changing the spline order and the regularization term. The bending term in the spline energy functional could, in principle,

* This work has been partially funded by the Spanish Government (MCyT TIC-2001-3808-C02)

R.E. Ellis and T.M. Peters (Eds.): MICCAI 2003, LNCS 2879, pp. 894–901, 2003.

also be modified to account for non-isotropic behavior and even a set of co-variables could also be added to the coupling term of the functional. None of these improvements are usually implemented, possibly because of the difficulty of obtaining an objective design from data. Other popular schemes are based on dynamic mechanical models [2] stated as a set of partial differential equations where the sparse displacements are associated with actuating forces. The mechanical model provides an ad-hoc regularization to the interpolation problem that produces a physically feasible result. The assumption that the physical difference between the source and the target image can be actually represented by mechanical models is by no means evident. Moreover, mechanical properties must also be endowed to the anatomic structures in order to obtain a proper model.

Our framework departs from the previous two approaches by adopting the geostatistical method of Kriging. Related work on the field of statistical shape analysis has been previously reported in [6]. The underlying idea is to use an experimental approach that makes the least a priory assumptions by statistically analyzing the available data, i.e., the displacements of the control points. Our method consist of locally applying the so-called Ordinary Kriging estimator [5], to obtain the best linear unbiased estimator (BLUE) of the displacement at every point from the displacements initially obtained at the control points. Central to this approach is the estimation of the second order characterization of the displacement field, now modeled as a vector random process. The estimated variogram [5] (a statistics related to the spatial covariance function or covariogram) plays the role of the thin-plate spline kernels, though now they are directly obtained from data and not from an a priory thin-plate dynamic model. Remarkably, thin-plate splines can be considered as a special case of Kriging estimation [8]. A novel approach to the estimation of the variogram the displacement field, which circumvents the practical limitations of Kriging, is presented here.

The next two sections describe the registration approach reported in this paper. Some illustrative results and conclusions are presented afterwards. An appendix on variogram estimation including some geostatistical terms is intended to make the paper self-contained.

2 Our Approach to Landmark-Based Registration

We will consider that the deformation that puts into correspondence the source and target images is a realization of a vector random field. The global component of the deformation corresponds to the trend (mean) of the random field, whereas the local deformation is modeled by an intrinsically stationary random field. The field is sampled by means of landmark correspondences, i.e., to each landmark in the source image corresponds a landmark in the target one, which are then used to reconstruct the whole realization of the random deformation field.

The geostatistical method tries to honour actual observations by estimating the spatial variability model directly from available data. This essentially consists of estimating the variogram of the field, which is a difficult problem specially if it is to be done from landmarks displacements. This has possibly prevented Kriging from being used in landmark-based registration. Here we propose a practical way to circumvent these difficulties by splitting the approach into three steps:

1. Image-based global registration: Unbiased estimation of the variogram requires to detrend the field. We propose to make an intensity-based global (i.e., rigid or affine) registration to remove the trend effect, with a variety of algorithms being available. For example, rigid registration by maximization of mutual information is a well-known algorithm [14], which can be used when image intensities in both images are different.
2. Model estimation: Estimating the variogram structure of the detrended displacement field is still a difficult task. The number of available landmarks in most practical applications is almost never enough to make good variogram estimations, and trying to extract a significant number of them from the images would render the method unpractical. We propose to use a fast general-purpose intensity-based nonrigid registration algorithm, in order to obtain an approximate dense displacement field suitable for variogram estimation. Again a number of algorithms are available, though we are using with excellent results a regularized block-matching scheme with mutual information (and others) similarity measure, developed by our team [13].
3. Landmark-based local registration: Landmarks are extracted from the globally registered image pair, and the detrended deformation field is reconstructed using Ordinary Kriging with the estimated variogram structure.

3 Displacement Field Reconstruction

The reconstruction of the displacement field can be cast as the optimal prediction of the displacement at every location $\boldsymbol{x_i}$ from a set of observations obtained by measuring the displacement between pairs of point landmarks extracted from both images. The observations can therefore be modeled as

$$\begin{aligned} \boldsymbol{Z}(\boldsymbol{x_i}) &= \boldsymbol{X'_i} - \boldsymbol{X_i} \\ &= \boldsymbol{D}(\boldsymbol{x_i}) + \boldsymbol{N_i}, \end{aligned} \tag{1}$$

where $\boldsymbol{X_i}$ and $\boldsymbol{X'_i}$ are independent random vectors with corresponding landmark positions, and the $\boldsymbol{N_i}$ are independent Gaussian random vectors with covariance matrices $\boldsymbol{C_{N_i}}(\boldsymbol{x_i}) = \boldsymbol{C_{X_i}} + \boldsymbol{C_{X'_i}}$. The covariances of the landmark positions ($\boldsymbol{C_{X_i}}$ and $\boldsymbol{C_{X'_i}}$) is obtained from the images using the Cramer-Rao lower bound proposed in [11,12]

$$\Sigma_{CR}(x) = \frac{\sigma_n^2}{m} \left(\sum_{N(m)} \nabla I(x) \nabla I(x)^t \right)^{-1}, \tag{2}$$

where $N(m)$ is a neighborhood around the landmark with m elements.

3.1 Ordinary Kriging Prediction of Displacement Fields

The local mean for each component of the field is assumed to be an unknown constant. We have found that this is a very convenient model, even after the global pre-registration that should render zero-mean values for the resulting displacement components. The

reason is that usually a locally varying mean structure can model much of the local deformation. Therefore, in this case we will not use all the samples but a limited number in a neighborhood around the prediction location. This has the added benefit of reducing the computational burden.

The ordinary co-kriging (i.e., multivariate Kriging) estimator of the displacement at point $\boldsymbol{x}$ takes the form

$$\begin{aligned}\hat{\boldsymbol{D}}(\boldsymbol{x}) \equiv \begin{bmatrix}\hat{D}^1(\boldsymbol{x})\\ \vdots \\ \hat{D}^d(\boldsymbol{x})\end{bmatrix} &= \begin{bmatrix}\boldsymbol{k_1^1}^t(\boldsymbol{x},\boldsymbol{O}) & \dots & \boldsymbol{k_d^1}^t(\boldsymbol{x},\boldsymbol{O})\\ \vdots & \ddots & \vdots \\ \boldsymbol{k_1^d}^t(\boldsymbol{x},\boldsymbol{O}) & \dots & \boldsymbol{k_d^d}^t(\boldsymbol{x},\boldsymbol{O})\end{bmatrix}\begin{bmatrix}\boldsymbol{Z^1}(\boldsymbol{O})\\ \vdots \\ \boldsymbol{Z^d}(\boldsymbol{O})\end{bmatrix}\\ &= \begin{bmatrix}\boldsymbol{k^1}^t(\boldsymbol{x},\boldsymbol{O})\\ \vdots \\ \boldsymbol{k^d}^t(\boldsymbol{x},\boldsymbol{O})\end{bmatrix}\boldsymbol{Z}(\boldsymbol{O})\\ &= \boldsymbol{K}(\boldsymbol{x},\boldsymbol{O})\boldsymbol{Z}(\boldsymbol{O}),\end{aligned} \tag{3}$$

where d is the number of components and $\boldsymbol{Z^r}(\boldsymbol{O})$ is a vector with the stacked observations from the r-th component in the neighborhod. If there is no second-order probabilistic dependence among the fields, each of them is dealt with independently leading to a block-diagonal $\boldsymbol{K}(\boldsymbol{x},\boldsymbol{O})$ matrix and resulting the conventional ordinary Kriging predictor for each component.

The ordinary Kriging coefficients must minimize the mean square prediction error

$$\begin{aligned}MSPE^r(\boldsymbol{x},\boldsymbol{O}) &= E\left(D^r(\boldsymbol{x}) - \boldsymbol{k^r}^t\boldsymbol{Z}(\boldsymbol{O})\right)^2\\ &= E\left(D^r(\boldsymbol{x}) - \mu_{D^r}(\boldsymbol{x}) - \boldsymbol{k^r}^t(\boldsymbol{Z}(\boldsymbol{O}) - \mu_Z(\boldsymbol{O}))\right)^2\\ &= \sigma^2_{D^r}(\boldsymbol{x}) - 2\boldsymbol{k^r}^t\boldsymbol{c}_{\boldsymbol{Z}D^r}(\boldsymbol{O},\boldsymbol{x}) + \boldsymbol{k^r}^t\boldsymbol{C_Z}(\boldsymbol{O})\boldsymbol{k^r},\end{aligned} \tag{4}$$

subject to the unbiasedness constraint

$$E\left(\hat{\boldsymbol{D}}(\boldsymbol{x})\right) = E\left(\boldsymbol{D}(\boldsymbol{x})\right). \tag{5}$$

Closed-form equations for the coefficients and for the achieved squared error can be readily obtained after some algebra (see for example [5]). Due to space constraints we only present the coefficients equation, expressed in terms of covariances. The matrix Λ is block diagonal, with each diagonal block equal to a column vector of ones, and the vector λ^r is a zero row vector with a single one in the r position:

$$\begin{aligned}\boldsymbol{k^r} = \boldsymbol{C_Z}^{-1}(\boldsymbol{O})\{&\boldsymbol{c}_{\boldsymbol{Z}D^r}([\boldsymbol{O},\boldsymbol{x}]) - \boldsymbol{\Lambda}\left(\boldsymbol{\Lambda}^t\boldsymbol{C_Z}^{-1}(\boldsymbol{O})\boldsymbol{\Lambda}\right)^{-1}\\ &\left(\boldsymbol{\Lambda}^t\boldsymbol{C_Z}^{-1}(\boldsymbol{O})\boldsymbol{c}_{\boldsymbol{Z}D^r}([\boldsymbol{O},\boldsymbol{x}]) - \boldsymbol{\lambda^r}\right)\}.\end{aligned} \tag{6}$$

Extensions of ordinary Kriging are possible by incorporating more complex mean structure models. Though this could seem in principle appealing, it has the serious drawback of hindering the estimation of the spatial variability model since the mean structure has to be filtered out before the covariance structure can be estimated. Notice that estimating the variogram does not require to pre-estimate the mean as far as this is constant.

4 Results

We are currently using the proposed framework in a number of applications. In order to better ilustrate its behavior, we have selected two simple experiments. Figure (2.a) shows a T1w MRI axial slice of a multiple sclerosis patient, and Fig. (2.b) a corresponding T2w axial slice of a different patient. Ellipsoids representing landmark covariances have been overlayed (seven landmarks in the brain and four on the skull). Figures (2.d) and (2.e) show two T1w mid-sagittal slices of MS patients, also with covariance landmark ellipsoids overlayed (eleven landmarks in the brain and three on the skull). In each case, the second image is to be warped onto the first one.

In both cases the images are first globally registered. Then a forward displacement field is obtained for each one using our general purpose general registration scheme [13], in order to estimate the variograms. Sample variograms and their weighted least squares fit to theoretical models (linear combination of Gaussian and Power models) are shown in Fig. (1). 5000 displacements are sampled for this purpose, which make the estimation highly accurate.

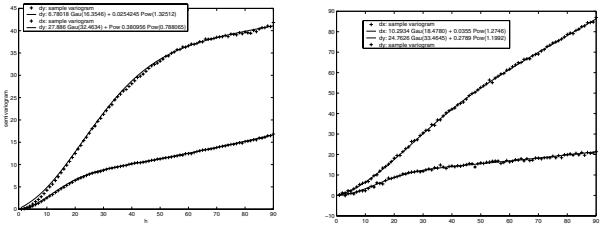

Fig. 1. Displacement variograms (left) axial (right) sagittal

Registration results are shown in Figs. (2.c) and (2.f) by ordinary Kriging estimation of the displacement field, using only the displacements from the landmarks on the images. Notice how even with so few landmarks a good result is achieved, specially in areas closer to the landmarks, due the proper estimation of the random displacement field. The open source software Gstat [9] has been used in these experiments.

5 Discusion and Conclusions

We have presented a practical approach to the statistical estimation of displacement fields from pairs of landmarks. The method is grounded on the solid theory of Ordinary Kriging and it also provides a way to estimating the spatial dependence models from image data, which circumvents some of the hurdles found when using Kriging. The fact that the statistical relation between both geometries is succesfully used makes the method highly accurate and particularly well suited for image registration and shape analysis applications. It is remarkable to note that thin-plate splines can be considered a particular case of Kriging and, in this sense, our approach generalizes this popular registration method.

Appendix: Geostatistical Spatial Modeling

Consider a random field $Z^r(\boldsymbol{x})$ (the superindex r is meant to consider several random fields, such as the components of a vector random field) such that

$$var(Z^r(\boldsymbol{x_i}) - Z^r(\boldsymbol{x_j})) = 2\gamma_{Z^r}(\boldsymbol{x_i} - \boldsymbol{x_j}), \quad \forall \boldsymbol{x_i}, \boldsymbol{x_j} \in \boldsymbol{\Omega}. \tag{7}$$

The function $2\gamma_{Z^r}(\boldsymbol{h})$ with $\boldsymbol{h} = \boldsymbol{x_i} - \boldsymbol{x_j}$ is called the *variogram* of the random field $Z^r(\boldsymbol{x})$ and, assuming it exists, is the central parameter to model the spatial dependence of the random field in the geostatistical method. $\gamma_{Z^r}(\boldsymbol{h})$ (without the 2 factor) is usually called *semivariogram*. The variogram can be easily related to the variance and covariance from the relation

$$var(Z^r(\boldsymbol{x_i}) - Z^r(\boldsymbol{x_j})) = \sigma^2_{Z^r}(\boldsymbol{x_i}) + \sigma^2_{Z^r}(\boldsymbol{x_j}) - 2C_{Z^r}(\boldsymbol{x_i}, \boldsymbol{x_j}). \tag{8}$$

The shape of a variogram is summarized by the following parameters:

- *Nugget:* it is the size of the discontinuity of the semivariogram at the origin. The presence of nugget effect is usually attributed to measurement noise and to a very local random component of the field that appears as uncorrelated at the working resolution.
- *Sill:* if the variogram is bounded, the sill is the value of the bound. A sill indicates total uncorrelation as, for example, with white noise. Usually random fields become uncorrelated for big lags, reaching a sill.
- *Partial sill:* it is the difference between the sill and the nugget.
- *Range:* it is the lag for which the sill is reached, of course assuming there is a sill in the variogram.

Approaches to construct valid theoretical variogram models are available [5,4,10,1, 9]. Most often existing variogram models such as nugget (white field), spherical, linear, exponential, power, etc are used as building blocks in a linear combination of valid variogram models, making use of the convexity of the set of valid variograms.

The variogram can be extended for the multivariate case [5]. The *cross-variogram* function is defined as

$$2\gamma_{Z^rZ^s}(\boldsymbol{x_i} - \boldsymbol{x_j}) = cov(Z^r(\boldsymbol{x_i}), Z^s(\boldsymbol{x_j})). \tag{9}$$

Intrinsic Stationarity: The scalar random field $Z^r(\boldsymbol{x})$ is said to be *intrinsically stationary* if it has a constant mean $E(Z^r(\boldsymbol{x})) = \mu_{Z^r}$ and its variogram exists. Moreover, any conditionally negative-definite function $2\gamma(\boldsymbol{h})$ is the variogram of an intrinsically stationary random field. The variogram of an intrinsic random field $Z^r(\boldsymbol{x})$ is

$$\begin{aligned} 2\gamma_{Z^r}(\boldsymbol{x_i} - \boldsymbol{x_j}) = E\left(Z^r(\boldsymbol{x_i}) - Z^r(\boldsymbol{x_j})\right)^2 \equiv \\ 2\gamma_{Z^r}(\boldsymbol{h}) = E\left(Z^r(\boldsymbol{x} + \boldsymbol{h}) - Z^r(\boldsymbol{x})\right)^2. \end{aligned} \tag{10}$$

Relation Between Intrinsic and Second-order Stationarities: Note that the family of intrinsic stationary fields is larger than the second order stationary one. In particular unbounded valid variograms, i.e., variograms without a sill (see below), do not have a correspondent auto-covariance function. For second-order stationary fields there is a simple relation between the variogram and the auto-covariance, i.e.,

$$2\gamma_{Z^r}(\boldsymbol{h}) = 2\left(C_{Z^r}(\mathbf{0}) - C_{Z^r}(\boldsymbol{h})\right). \tag{11}$$

It is clear that in the common situation for second order stationary fields where the covariance approaches to zero for large space lags, the sill of the variogram is $2C_{Z^r}(\mathbf{0})$.

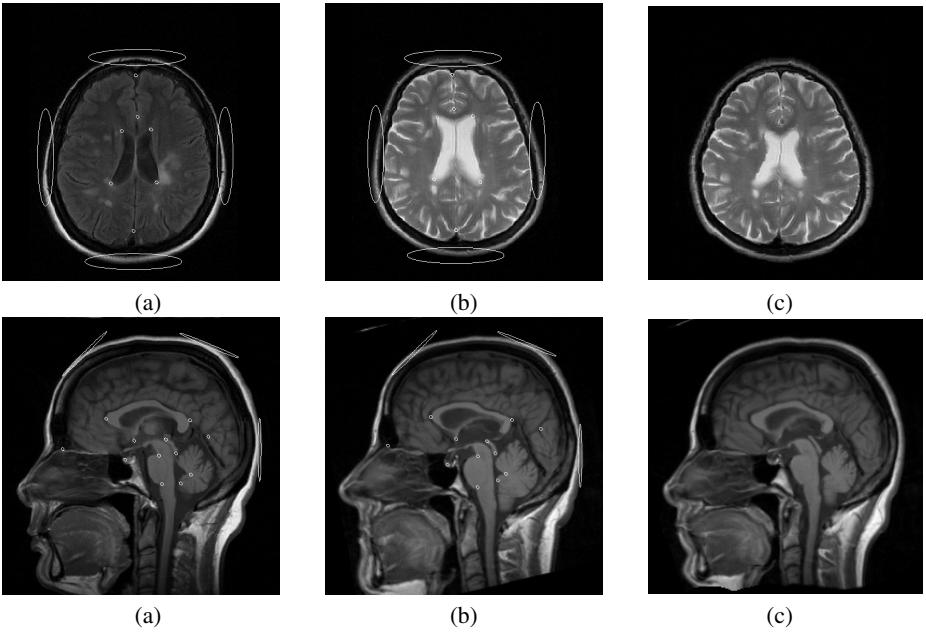

Fig. 2. Registration results (see text). (a) axial T1 (b) axial T2 (c) warped axial T2. (d) first sagittal T1 (e) second sagittal T1 (f) warped second sagittal

Variogram Estimation

The variogram is estimated under the assumption of intrinsic estationarity, i.e., the mean must be constant. Should this not be the case, a trend model must be pre-estimated in order to be substrated from the field prior to estimating the variogram. This process is undesirable since it introduces bias in the variogram estimation due to its inherent circularity: the probabilistic characterization of the random component of the field must be known in order to estimate the trend, but the trend must also be known in order to estimate the probabilist characterization of the random component. Nevertheless, this

issue is present in any model with a trend and a random component and, in fact, estimating the sample variogram instead of the sample auto-covariance has several advantages [5] from this point of view:

1. If the mean value of the field is an unknown constant, it is not necessary to pre-estimate it because the variogram sample estimator is based on differences. Hence, in this case, the sample variogram can be estimated unbiasedly.
2. The sample variogram estimator is more robust against mean model mismatch than the sample auto-covariance one.
3. The sample variogram estimator is less biased than the sample auto-covariance one, when the mean model is pre-estimated and substracted from the field realization in order to make the spatial dependence model estimation.

References

1. S. L. Arlinghaus and D. A. Griffith, editors. *Practical Handbook of Spatial Statistics*. CRC Press, revised edition, 1995. ISBN: 0849301327.
2. R. Bajacsy and S. Kovacic. Multiresoltion elastic matching. *Computer Vision, Graphics, and Image Processing*, (46):1–21, 1989.
3. F. L. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
4. J.-P. Chiles and P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty.* Wiley Series in Probability and Statistics. Applied Probability and Statistics. Wiley-Interscience, march 199.
5. N. A. C. Cressie. *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, 1993.
6. I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, 1998.
7. J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, Apr. 1998.
8. G. Matheron. *Down-to-Earth Statistics: Solutions Looking for Geological Problems*, chapter Splines and kriging: their formal equivalence, pages 77–95. Syracuse University Geological Contributions, Syracuse, 1981.
9. E. J. Pebesma and C. G. Wesseling. Gstat: a program for geostatistical modelling, prediction and simulation. *Computers & Geosciences*, 24(1):17–31, 1998.
10. B. D. Ripley. *Statistical Inference for Spatial Processes*. Cambridge University Press, reprint edition, september 1991. ISBN: 0521424208.
11. K. Rohr. Image registration based on thin-plate splines and local estimates of anistropic landmark localization uncertainties. In *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), Massachusetts Institute of Technology (MIT), Cambridge, MA*, number 1496, pages 1174–1183, Heidelberg, october 11-13 1998.
12. K. Rohr. *Landmark-based image analysis (using geometry and intensity models)*, volume 21 of *Computational Imaging and Vision*. Kluwer Academic Publishers, 2001.
13. E. Suarez, C.-F. Westin, E. Rovaris, and J. Ruiz-Alzola. Nonrigid registration using regularized matching weighted by local structure. In *Proceedings of the Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'02), Tokio, Japan*, number 2489, pages 581–589, Heidelberg, september 2002.
14. W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.