Non-parametric Image Registration Using Generalized Elastic Nets

Andriy Myronenko, Xubo Song, Miguel A. Carreira-Perpinan

Non-parametric Image Registration Using Generalized Elastic Nets

Andriy Myronenko, Xubo Song, and Miguel Á. Carreira-Perpiñán

Dept. of Computer Science & Electrical Engineering
OGI School of Science & Engineering, Oregon Health & Science University
20000 NW Walker Road, Beaverton, OR 97006, USA
{myron,xubosong,miguel}@csee.ogi.edu

Abstract. We introduce a novel approach for non-parametric non-rigid image registration using generalized elastic nets. The concept behind the algorithm is to adapt an elastic net in spatial-intensity space of one image to fit the second image. The resulting configuration of the net, when it achieves its minimum energy state, directly represents correspondence between images in a probabilistic sense and recovers underlying image deformation, which can be arbitrary. Representation of elastic net in the spatial-intensity space with specific priors that enforce natural elastic deformation is introduced. Efficient algorithm for optimization of elastic net energy is developed. The accuracy and effectiveness of the method is demonstrated on different medical image registration examples with locally non-linear underlying deformations.

1 Introduction

Image registration is an important component in medical image analysis. It is a process of determining a geometric transform that relates the contents of two images in a meaningful way and establishes the correspondence between them. Applications of image registration include combining images of the same subject from different modalities, aligning temporal sequences of images to compensate for motion of the subject between scans, image guidance during interventions and aligning images from multiple subjects in cohort studies [1].

Non-rigid image registration is the most interesting and challenging work in registration today. Many non-rigid registration techniques have been proposed during last 20 years [2]. Most of them build a parameterized model that constrains the form of allowed deformations and then optimize a similarity function to find an approximation of a real underlying deformation. Study of non-parametric registration has focused on variational approaches [3]. In this paper we introduce a non-parametric registration method that can deal with non-rigid deformation of arbitrary complexity, using a probabilistic model known as the elastic net (EN). The elastic net is a net of connected points which jointly and smoothly move in a high-dimensional space to model a data set. An energy function can be defined to trade off accuracy of the net fitting the data (fitness term) vs net continuity (tension term). The elastic net was originally introduced as a

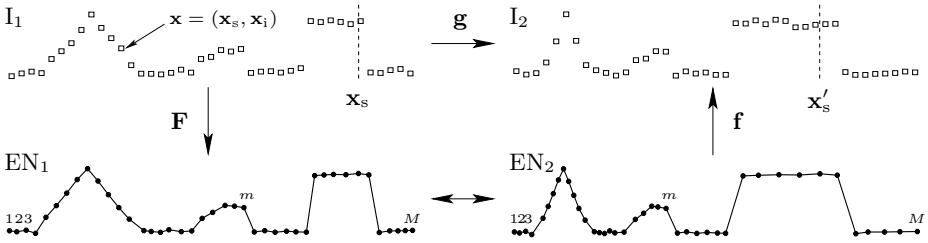

Fig. 1. Illustration of the alignment method (for 1D images). I_1 represents a 1D intensity image in spatial (x-axis) and intensity (y-axis) space (each pixel is marked as a small □). I_2 represents the same image with local, nonlinear spatial distortion and intensity noise. EN_1 is an elastic net fitted to I_1 (in spatial-intensity space, with centroids marked •) and EN_2 is adapted from EN_1 to fit I_2. Since the centroids in EN_1 and EN_2 correspond one-to-one ($1 \leftrightarrow 1$, etc.), and the elastic net allows to define mappings between image points and centroids (see section 2), we can map any spatial location $\mathbf{x}_s$ in I_1 to a spatial location $\mathbf{x}'_s$ in I_2 through the elastic nets, thus aligning I_1 to I_2.

continuous optimization method for the traveling salesman problem [4, 5] and has also been successfully applied to modeling maps of primary visual cortex. However it has had a limited use in computer vision. A generalization of elastic nets to arbitrary quadratic tension terms was investigated in [6]. Here we adapt the generalized elastic net to represent image deformations. The intuition is to position a net according to the first image and then deform it to align with the second image. The deformation produced by elastic net, when its energy is minimized, directly represents the deformation field between the images.

This is illustrated more specifically in Fig. 1. We consider an image as a noisy 2D manifold in the spatial-intensity space, i.e., each pixel is represented by a point $\mathbf{x} = (\mathbf{x}_s, \mathbf{x}_i) \in \mathbb{R}^3$ of spatial location $\mathbf{x}_s \in \mathbb{R}^2$ and intensity $\mathbf{x}_i \in \mathbb{R}$. We model this manifold in a probabilistic way with an elastic net EN_1, which allows to map any image point onto the net, and vice versa. We then adapt EN_1 for a given image I_1 to a new image I_2 in the spatial-intensity space; again this allows to map a net point onto image space and vice versa. The alignment mapping which maps a spatial location in I_1 to another spatial location in I_2 is obtained through the deformed elastic net. We describe the method of generalized elastic net and its adaptation to image registration problem in detail in section 2, give experimental results in section 3 and discuss them in section 4.

2 Image Registration with Generalized Elastic Nets

Generalized elastic nets (GEN) The *elastic net* is a Gaussian mixture with a quadratic prior on its centroids [4–6]. The centroids implicitly represent a nonlinear, low-dimensional manifold that probabilistically models a high-dimensional data set $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ (expressed as a $D \times N$ matrix). Specifically, given a collection of M D-dimensional centroids $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_M)$ (ex-

pressed as a $D \times M$ matrix) and a scale parameter $\sigma \in \mathbb{R}^+$, consider a Gaussian-mixture density $p(\mathbf{x}) = \sum_{m=1}^{M} \frac{1}{M} p(\mathbf{x}|m)$ with $\mathbf{x}|m \sim \mathcal{N}(\mathbf{y}_m, \sigma^2 \mathbf{I}_D)$. A smoothing or neighborhood-preserving prior on the centroids is defined as $p(\mathbf{Y}; \beta) \propto \exp\left(-\frac{\beta}{2} \sum_m \|\mathbf{y}_{m+1} - \mathbf{y}_m\|^2\right)$ where β is a regularization hyperparameter. Without the prior, the centroids could be permuted at will with no change in the model, since the variable m is just an index. The elastic net minimizes the energy function

$$E(\mathbf{Y}, \sigma) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\left\|\frac{\mathbf{x}_n - \mathbf{y}_m}{\sigma}\right\|^2} + \frac{\beta}{2} \sum_m \|\mathbf{y}_{m+1} - \mathbf{y}_m\|^2 \qquad (1)$$

which is derived from the log posterior $\log p(\mathbf{Y}|\mathbf{X}, \sigma)$ of the full model (i.e., maximum-a-posteriori estimation). We call the first term the *fitness term*, arising from the Gaussian mixture $p(\mathbf{X}|\mathbf{Y}, \sigma)$, and the second term the *tension term*, arising from the prior $p(\mathbf{Y})$. The elastic net was generalized in [6, 7] to accommodate general quadratic priors. The prior can be used to convey the topological (dimension and boundary conditions) and geometric (e.g. curvature) structure of a manifold implicitly defined by the centroids. The *generalized elastic net (GEN)* minimizes the energy function

$$E(\mathbf{Y}, \sigma) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\left\|\frac{\mathbf{x}_n - \mathbf{y}_m}{\sigma}\right\|^2} + \frac{\beta}{2} \operatorname{tr}\left(\mathbf{Y}^T \mathbf{Y} \mathbf{S}\right). \qquad (2)$$

Quadratic priors are considered of the form $\mathbf{S} = \mathbf{D}^T \mathbf{D}$, so that $\operatorname{tr}\left(\mathbf{Y}^T \mathbf{Y} \mathbf{S}\right) = \left\|\mathbf{D}\mathbf{Y}^T\right\|^2$ in terms of the Frobenius norm. The matrix $\mathbf{D}$ represents a discretized differential operator. For example (for a 1D net for simplicity, and using forward differences [6]), a first-order derivative results in a sum of squared lengths $\left\|\mathbf{D}\mathbf{Y}^T\right\|^2 = \sum_m \|\mathbf{y}_{m+1} - \mathbf{y}_m\|^2$ and approximates a penalty $\int \|\nabla \mathbf{y}\|^2$ over a continuous net $\mathbf{y}$ (with an infinite number of centroids). This corresponds to a matrix $\mathbf{D}$ where each row is a shifted version of $(-1\ 1\ 0\ 0 \ldots 0)$, and it was the tension term used in the original elastic net (Eq.(1)), penalizing stretching of the net. A second-order derivative results in $\sum_m \|\mathbf{y}_{m+2} - 2\mathbf{y}_{m+1} + \mathbf{y}_m\|^2$, etc. By choosing $\mathbf{S}$ as an appropriate combination of differential operators we can impose a desired type of smoothness on the GEN (see [7] for a discussion of the effect of different derivatives on the maps of primary visual cortex). The resulting $\mathbf{S}$ has a sparse, banded structure. We consider open boundary conditions at the image boundaries. Fig. 1 schematically shows a 1D elastic net.

Adaptation of the GEN Although it is possible to derive an EM algorithm to estimate $\mathbf{Y}$ and σ jointly, the GEN is usually trained with a deterministic annealing algorithm in order to obtain good local minima. This minimises E over $\mathbf{Y}$ for fixed σ, starting with a large σ and tracking the minimum to a small value of σ. For constant σ, [6] used a fixed-point iteration to find stationary points of E:

$$\frac{\partial E}{\partial \mathbf{Y}} = -\frac{1}{\sigma^2}(\mathbf{X}\mathbf{W} - \mathbf{Y}\mathbf{G}) + \beta \mathbf{Y}\left(\frac{\mathbf{S} + \mathbf{S}^T}{2}\right) = \mathbf{0} \quad \Longrightarrow \quad \mathbf{Y}\mathbf{A} = \mathbf{X}\mathbf{W} \qquad (3)$$

with weight matrix $\mathbf{W} = (w_{nm})$ and invertible diagonal matrix $\mathbf{G} = \mathrm{diag}\,(g_m)$

$$w_{nm} = \frac{e^{-\frac{1}{2}\left\|\frac{\mathbf{x}_n - \mathbf{y}_m}{\sigma}\right\|^2}}{\sum_{m'=1}^{M} e^{-\frac{1}{2}\left\|\frac{\mathbf{x}_n - \mathbf{y}_{m'}}{\sigma}\right\|^2}} \qquad g_m = \sum_{n=1}^{N} w_{nm} \qquad \mathbf{A} = \mathbf{G} + \sigma^2\beta\left(\frac{\mathbf{S} + \mathbf{S}^T}{2}\right).$$

The weight w_{nm} is the responsibility $p(m|\mathbf{x}_n)$ of centroid y_m for generating point $\mathbf{x}_n$, g_m is the total responsibility of centroid $\mathbf{y}_m$, and the matrix $\mathbf{XW}$ is a list of average centroids. We solve for $\mathbf{Y}$ in the system of eq. (3) and iterate, since $\mathbf{W}$ and $\mathbf{G}$ depend on $\mathbf{Y}$. In [6], the system (3) was solved using Cholesky factorisation. While this is robust and efficient (since it takes advantage of the sparsity structure of $\mathbf{S}$), here we use a different method based on linear conjugate gradients (CG) [8]. Linear CG solves an $M \times M$ positive definite linear system in at most M steps, each costing $\mathcal{O}(M^2)$ (actually less since $\mathbf{A}$ is sparse), and has two important advantages: (1) we can initialize the linear CG from the previous $\mathbf{Y}$ value (which will be close to the solution) rather than solving each system anew, as Cholesky does; (2) we can run only a few linear CG steps and obtain an approximate but good enough solution rather than an exact, costly one. This considerably accelerates the overall annealing algorithm without sacrificing accuracy. A further acceleration is obtained by truncating the Gaussian kernel so that most weights w_{mn} are zero and can be ignored.

Registration We now show how the framework of elastic net can be adapted for the problem of image registration. First, we represent two images I_1 and I_2 in the spatial-intensity space. Then we construct an elastic net with as many centroids as pixels in image I_1. This net $\mathbf{Y}$ is initialized with each centroid representing the spatial-intensity value of one pixel in I_1 (i.e., $\mathbf{Y} = \mathbf{X}_1$). The net is adapted by adjusting the centroids to fit data $\mathbf{X}_2$, representing image I_2 in spatial-intensity space. This is done by minimizing the energy in Eq. (2). The intuition for using the same number of centroids as there are pixels in I_1 is that the final centroid locations, when the energy function is optimized, directly shows the displacement of each pixel in I_1 when it is deformed into I_2. As a result, no interpolation is needed. It also provides the maximum level of deformation complexity. In general, we can choose to have more or fewer centroids than pixels. In this case the displacement of a pixel $\mathbf{x}$ in I_1 can be found by interpolation using the probabilities $p(m|\mathbf{x})$ and $p(\mathbf{x}|m)$ provided by the GEN.

We assume that the deformation between two images is only spatial, not in intensity. This translates to constraining the intensity components in the centroid vectors to be constant. In other words, the free parameters for centroid $\mathbf{y}_m = (\mathbf{y}_{ms}, \mathbf{y}_{mi})$ are $\mathbf{y}_{ms}$ only, and the optimization updates only apply to $\mathbf{y}_{ms}$. Doing so is important to produce only spatial deformation for I_1 when fitting it to I_2. In general, intensity variations across images can be accommodated by updating the complete $\mathbf{y}_m = (\mathbf{y}_{ms}, \mathbf{y}_{mi})$. We use the following penalty matrix: $\mathbf{S} = \beta_1\mathbf{D}_1^T\mathbf{D}_1 + \beta_2\mathbf{D}_2^T\mathbf{D}_2$, where $\mathbf{D}_1$ and $\mathbf{D}_2$ are first- and second-order derivatives, and their relative strengths are controlled by hyper-parameters β_1 and β_2; practically useful values for them can be obtained manually for a given type of images.

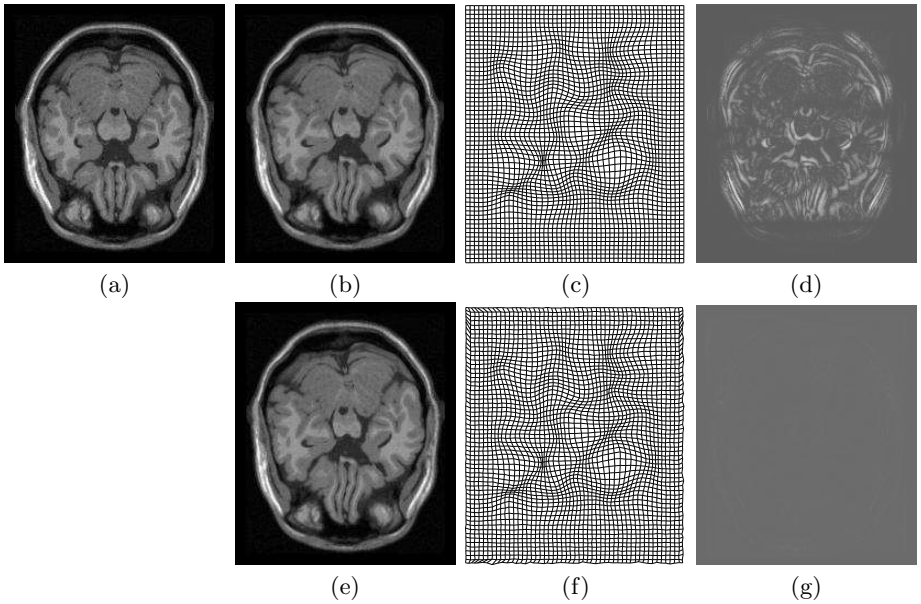

Fig. 2. (a) Original MR slice; (b) deformed original image according to control points using thin plate splines; (c) deformation field represented by control point; (d) absolute value of intensity difference between original and deformed images; (e) registration result of the algorithm (original image is registered onto the deformed one); (f) deformation field found by the algorithm (almost equal to the true one); (g) absolute value of intensity difference between registered and original deformed images (almost zero).

3 Experimental Results

In all experiments, the image intensities are first re-scaled to allow the use of a single σ for all dimensions, and the images are coarsely aligned using cross-correlation to eliminate rigid translation. The resulting data sets $\mathbf{X}_1, \mathbf{X}_2 \subset \mathbb{R}^3$ were used to adapt the elastic net. The resulting, aligned dataset $\mathbf{X}_1'$ (obtained from the spatial deformation given by the GEN and the original intensity values) was post-processed with bi-cubic interpolation to produce the aligned image. The prior parameters β_1 and β_2 were set manually for each type of image. We ran 10 annealing iterations from $\sigma = 3$ to $\sigma = 0.5$ pixels.

We show the performance of the algorithm on artificial data with known non-linear deformations and on two real-life examples. The algorithm was implemented in Matlab with subroutines coded in C, and tested on Pentium4 CPU 3.5GHz with 4Gb RAM. The test images are gray-scale images of size 250 $\times$ 250, and the registration process takes about 20 minutes for each image pair.

Brain MRI 2D images with and without known deformation A slice of MRI brain image was artificially deformed using known deformation field

Table 1. Experimental results for different deformation levels.

Deformation STD	Transformation RMSE	Intensity RMSE
1.0	0.3135	0.0044
1.5	0.5124	0.0047
2.0	0.9753	0.0053
2.5	1.1152	0.0060
3.0	1.0962	0.0059

(Fig. 3). We define a uniform grid of control points in the original image, randomly move them and use the thin plate spline technique to create a locally nonlinear deformed image. Our algorithm is applied to align original image (a) onto deformed one (b). The final absolute image difference (g) is so small it is hardly visible, demonstrating the high accuracy of the method. Table 1 shows the value of root mean square error (RMSE) between true and estimated deformation as well as the intensity RMSE between original and registered images, as a function of spatial distortion level controlled by the standard deviation (STD) of control points perturbation measured in pixels. The transformation error is at most of the order of one pixel.

Figure 3 shows the registration of images (a) and (b) from two patients. Image (b) is registered onto (a) resulting in (d). Panels (c) and (e) are color composite views of the two images before (c) and after (e) registration, where image (a) is coded with green and (b) with red color. Visual inspection clearly reveals much improved alignment, even when the two original images have significantly different intensity ranges.

Microscopic iris images We stabilize a video sequence of microscopic iris images through frame-by-frame registration. This is necessary to remove the severe jitter and deformation across frames in order to be able to track the leukocyte motion. The deformation between frames is highly nonlinear. Our algorithm proves to be accurate and effective for these images, as demonstrated in Fig. 4. Ideal registration in this case should lead to an absolute difference image after alignment with background intensity close to zero and bright blobs corresponding to the moving leukocytes, which is exactly the case in Fig.4(e).

4 Conclusion

We have developed a probabilistic approach for non-parametric non-rigid image registration based on the generalized elastic net (using first- and second-order differential priors). The method is able to accurately register images even with highly non-linear local deformations, and we have designed a new, more efficient optimization algorithm based on linear conjugate gradients. When the elastic net is initialized with one centroid for each pixel in image I_1, the resulting deformed net will provide directly the displacement for each pixel. When the number

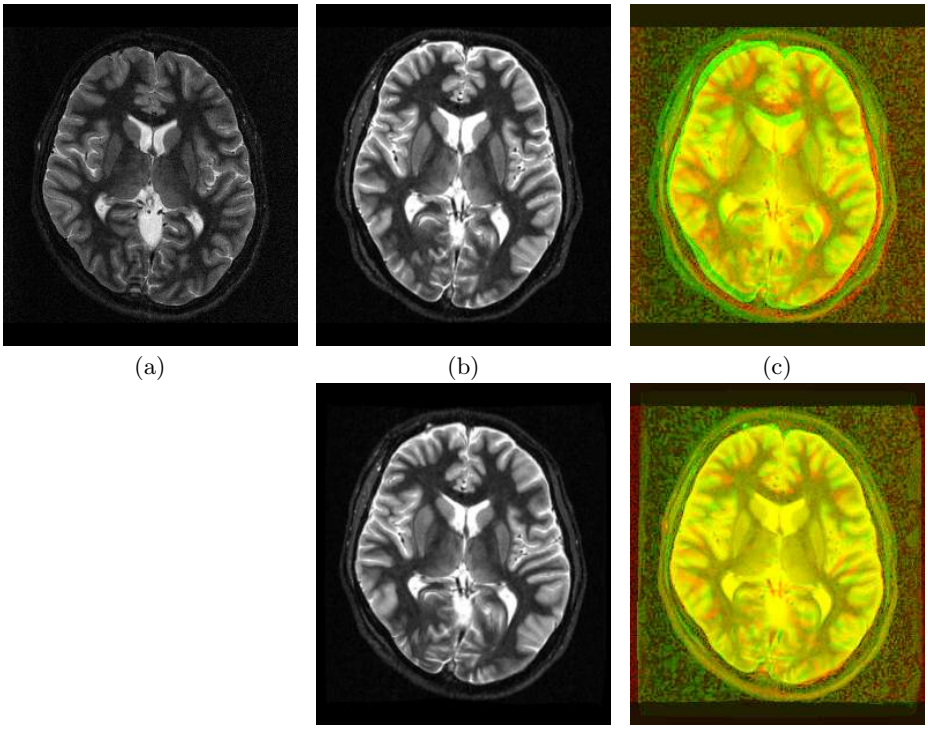

Fig. 3. (a) First person; (b) second person; (c) their composite view; (d) registration of (b) onto (a); (e) composite view of (a) and (d).

of centroids goes to infinity in the limit, the mapping approximates continuous mapping with continuous derivative, however it is still implicitly defined by finite collection of centroids. In general, the deformation complexity can be controlled by using an arbitrary number of centroids. In this case, determining the displacement of a pixel in I_1 can be interpolated by relating probabilistically the pixel and the elastic net. With the image deformation represented as the motion of net centroids, we do not need any image interpolation on each iteration, unlike most other registration methods. While we have focused on intensity features, the method easily accommodates arbitrary features (e.g. gradient information and color components) and images of different spatial resolutions. The generalization of the method for 3D images is straightforward, however the computational time is large at present. One possible way to reduce the computational complexity is to use fewer centroids. The method is also well suited for continuous tracking of the centroids over consecutive frames in an image sequence, by successively adapting the net to each image.

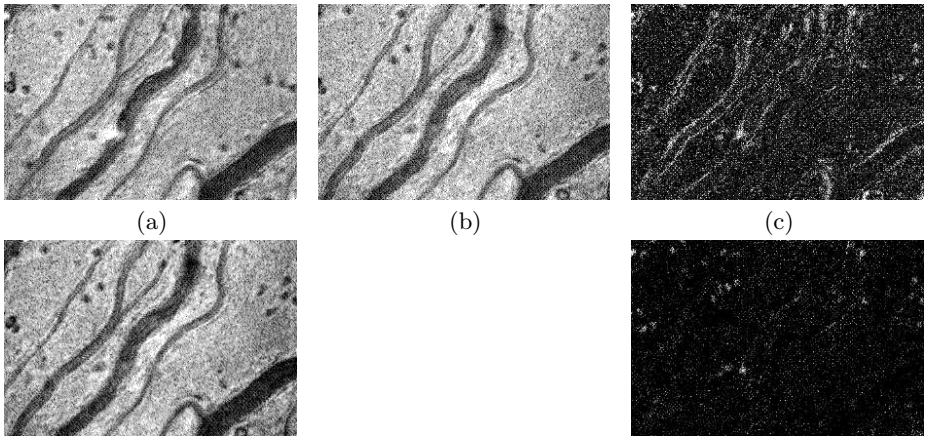

Fig. 4. Microscopic video of iris: (a) frame 1, (b) frame 37; (c) absolute intensity difference between the two frames before registration; (c) registration result of the algorithm: image (b) is aligned with image (a); (e) absolute intensity difference between the two frames after registration.

Future research includes varying the number of centroids for different deformation complexity, adaptive choice of the regularization hyper-parameter, and local adaptation of centroid variance σ for different dimensions to deal better with local image properties and intensity variations.

References

1. Hill, D.L.G., Batchelor, P.G., Holden, M., Hawkes, D.J.: Medical image registration. Physics in Medicine and Biology **46**(3) (2001) R1–R45
2. Crum, W.R., Hartkens, T., Hill, D.L.G.: Non-rigid image registration: Theory and practice. Brit. J. of Radiology **77**(special issue) (2004) S140–S153
3. Thirion, J.P.: Image matching as a diffusion process: An analogy with Maxwell's demons. Medical Image Analysis **2**(3) (1998) 243–260
4. Durbin, R., Willshaw, D.: An analogue approach to the traveling salesman problem using an elastic net method. Nature **326**(6114) (1987) 689–691
5. Durbin, R., Szeliski, R., Yuille, A.: An analysis of the elastic net approach to the traveling salesman problem. Neural Computation **1**(3) (1989) 348–358
6. Carreira-Perpiñán, M.Á., Dayan, P., Goodhill, G.J.: Differential priors for elastic nets. In Gallagher, M., Hogan, J., Maire, F., eds.: Proc. of the 6th Int. Conf. Intelligent Data Engineering and Automated Learning (IDEAL'05). Volume 3578 of Lecture Notes in Computer Science., Springer-Verlag (2005) 335–342
7. Carreira-Perpiñán, M.Á., Goodhill, G.J.: Influence of lateral connections on the structure of cortical maps. J. Neurophysiol. **92**(5) (2004) 2947–2959
8. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer-Verlag, New York (1999)