

Shape Statistics for Image Segmentation with Prior

Guillaume Charpiat
Olivier Faugeras
Renaud Keriven

Research Report 07-29
January 2007



Centre d'Enseignement et de Recherche
en Technologies de l'Information et Systèmes

CERTIS, ENPC,
77455 Marne la Vallée, France,
<http://www.enpc.fr/certis/>

Shape Statistics for Image Segmentation with Prior

Statistiques de formes pour la segmentation d'images avec a priori

Guillaume Charpiat¹
Olivier Faugeras²
Renaud Keriven³

¹DI, ENS, 45 rue d'Ulm 75005 Paris, France

²ODYSSEE, INRIA, 2004 route des Lucioles B.P. 93, 06902 Sophia-Antipolis Cedex, France

³CERTIS, ENPC, 77455 Marne la Vallee, France, <http://www.enpc.fr/certis/>

Abstract

We propose a new approach to compute non-linear, intrinsic shape statistics and to incorporate them into a shape prior for an image segmentation task. Given a sample set of contours, we first define their mean shape as the one that realizes the infimum of the sum of the square of its distance to each of the samples. We consider here the Hausdorff distance between shapes, or more exactly a differentiable approximation of it. The mean shape is then computed with a gradient descent framework. We perform statistics on the instantaneous deformation fields that the mean shape should undergo to move towards each sample. The application of principal component analysis (PCA) to the deformation fields leads to sensible characteristic modes of deformation that convey the shape variability. Contour statistics are then turned into a shape prior which is rigid-motion invariant. Image segmentation results with the shape prior outperform usual segmentations, which validates the approach.

Résumé

Nous proposons une nouvelle approche pour calculer des statistiques sur les formes qui soient non-linéaires et intrinsèques, et pour les intégrer dans un critère d'a priori sur la forme dans le cadre d'une tâche de segmentation d'images. Étant donné un ensemble d'apprentissage de contours, nous définissons tout d'abord leur forme moyenne comme la forme en laquelle l'infimum de la somme des carrés de sa distance à chacun des échantillons est atteint. Nous travaillons ici avec la distance de Hausdorff, ou plus exactement avec une approximation différentiable de celle-ci. La forme moyenne est alors calculée par descente de gradient. Nous exprimons ensuite des statistiques sur les champs de déformation instantanés que l'on devrait appliquer à la moyenne pour qu'elle se déplace vers chacun des échantillons. En appliquant une analyse en composantes principales (ACP) aux champs de déformation, on obtient des modes de déformation caractéristiques particulièrement pertinents, exprimant ainsi la variabilité de la forme. Ces statistiques sur les contours sont ensuite transformées en un critère d'a priori de forme qui est invariant par similitude. Les résultats de segmentation d'images sont bien meilleurs avec cet a priori que sans.

Contents

1	Introduction	1
2	Shape statistics	1
2.1	Empirical mean	1
2.2	Empirical covariance	2
3	Shape Priors	2
3.1	Context	2
3.2	Shape Probability	3
3.2.1	A simplistic method	3
3.2.2	Gaussian Eigenmodes (PCA on gradients)	3
3.2.3	Eigenmode Histograms	4
3.3	Pre-Computing	4
4	Toy Example	5
5	Rigid Registration	6
6	Gaussian Eigenmodes Examples	8
7	Conclusion	14

1 Introduction

Finding the contour of an object in an image is a very difficult and ill-posed task. We are here interested in the precise task of finding an object in an image with the knowledge of a set of examples, i.e. a set of images such that in each of them the object appears and has already been segmented. Usually, the only information retrieved from the set of examples comes from statistics on the intensity of the two regions (the inside of the object and the background), which does not carry any information about the shape of the object. Consequently, during an active-contour based evolution in order to segment a new image, the only restriction concerning shape will be brought by the regularity term which imposes the smoothness of the contour. We introduce here a way to take into account shape statistics into the standard active-contour algorithms for segmentation. The framework is presented in the context of planar curves but can be easily extended to boundary surfaces of 3D objects. Our approach is also designed to be rigid-motion invariant. In comparison to already existing techniques, it is mathematically well-posed, it does not require a dense sample set, it is not expensive computationally and the shape variability is well conveyed, even for very small datasets.

2 Shape statistics

2.1 Empirical mean

The first task is to define and compute the mean of a set of shapes in a rigid-motion invariant framework. *Shape* stands for any smooth contour, independently on its parameterization. A rigid motion R is a combination of a translation, a rotation and a scaling centered on the inertial center of the shape. Thus it can be represented by few real parameters (rotation angle, scaling factor and two translation parameters). Inspired by the work of Frchet [5], Karcher [6], Kendall [7], and Pennec [10], we provide the following

Definition 1 *Given N shapes $(\Gamma_1, \dots, \Gamma_N)$, and an energy $E(\cdot, \cdot)$ which expresses the distance between any two shapes, we define their empirical mean as any shape $\hat{\Gamma}$ that achieves a local minimum of the function μ defined by*

$$\mu : \Gamma \mapsto \frac{1}{N} \sum_{i=1, \dots, N} \inf_{R_i} E^2(\Gamma, R_i(\Gamma_i))$$

where for each shape Γ_i the infimum is taken over all rigid motions R_i .

To compute the empirical mean of N shapes, we initialize all rigid motions R_i automatically so that all rigidly-moved shapes $R_i(\Gamma_i)$ have the same center of

mass and average radius. We start from an initial shape $\Gamma(0)$, and solve the PDEs

$$\begin{aligned}\partial_t \Gamma(t) &= -\nabla_{\Gamma} \mu(\Gamma, R_1(\Gamma_1), \dots, R_N(\Gamma_N)) \\ \partial_t R_i &= -\nabla_{R_i} E^2(\Gamma, R_i(\Gamma_i))\end{aligned}\tag{1}$$

The derivative $\nabla_{\Gamma} \mu$ of the energy μ with respect to the shape Γ is a normal deformation field defined on Γ . The two PDEs are solved simultaneously. The evolution of Γ is done in a level-set framework whereas the parameters of the rigid motions follow a usual gradient descent in \mathbb{R} . Note that there may exist several means if μ has several local minima. In practice however, we have chosen for E the differentiable approximation of the Hausdorff distance proposed by Charpiat et al. in [3], in which the expression of its gradient is also given, and the mean appears to be unique.

2.2 Empirical covariance

Still inspired by [3], we now define something similar to the covariance matrix of a set of N shapes given their mean $\hat{\Gamma}$. Each shape Γ_i implies thanks to equation 1 a normal deformation field $\alpha_i = -\nabla_{\hat{\Gamma}} E^2(\hat{\Gamma}, R_i(\Gamma_i))$ on $\hat{\Gamma}$. This field α_i is the best instantaneous deformation that one should apply to $\hat{\Gamma}$ to make it closer to $R_i(\Gamma_i)$. These fields $\alpha_1, \alpha_2, \dots, \alpha_N$ belong to the same space, the tangent space of the mean curve $\hat{\Gamma}$, and they can be seen as functions from the points of $\hat{\Gamma}$ to \mathbb{R} . The correlation between any two fields is $\langle \alpha_i | \alpha_j \rangle_{L^2} = \int_{\hat{\Gamma}} \alpha_i(\mathbf{x}) \alpha_j(\mathbf{x}) d\mathbf{x}$. To express statistics on the deformation fields, we perform principal component analysis (PCA) on them and obtain new instantaneous deformation fields β_k (with associated standard deviations σ_k) which form an orthogonal basis of the previous set of deformations. As shown in the sequel, these characteristic modes of deformation are very sensible and convey the shape variability of the sample set of shapes.

3 Shape Priors

We now propose several shape priors based on different ways to introduce a distance to the shape distribution $(\Gamma_1, \Gamma_2, \dots, \Gamma_N)$.

3.1 Context

Let C be the evolving curve which we would like to fit the contour of the object in the new given image I . We can express any energy minimization as a probability

maximization, the two approaches being essentially the same. We would like to maximize the probability $P(C|I)$ with respect to C :

$$P(C|I) = P(I|C) \frac{P(C)}{P(I)} \simeq P(I|C) P(C).$$

Then $P(I|C)$ is given by the standard approach (based on intensity gradients, statistics on textures and so on), while $P(C)$ expresses the probability that C has such shape. In the sequel $P(C)$ will stand for any positive, upper-bounded energy depending on C , and its total mass will not necessarily be equal to one.

3.2 Shape Probability

3.2.1 A simplistic method

Given a distance or energy E between shapes, a new distance between C and the whole shape distribution $\mathcal{D} = (\Gamma_i)$ could be defined as $\sum_i E^2(C, \Gamma_i)$, or better, as:

$$\sum_i \inf_{R_i} E^2(C, R_i(\Gamma_i))$$

But the global minimum of this energy has already been studied and it is the mean shape $\hat{\Gamma}$, so that this energy is not really more interesting than $E(C, \hat{\Gamma})$. In particular the shape variations in \mathcal{D} are not taken into account.

Cremers et al. [4] have applied the Parzen method to \mathcal{D} and consider the shape probability $\sum_i \exp\left(-\frac{E^2(C, \Gamma_i)}{2\sigma^2}\right)$ but it suffers from highly dense sample requirements.

3.2.2 Gaussian Eigenmodes (PCA on gradients)

It is also possible to include second order statistics from section 2.2 into the design of the shape probability. The most significant modes β_k are the ones with highest associated standard deviation σ_k ; therefore you could take only the very first modes into account. However the issue to determine the number of modes of interest is not fundamental since the importance of each mode will be related to its standard deviation and consequently the consideration of some extra modes with low standard deviation will not change significantly the distribution.

The PCA decomposition supposes implicitly that linear combinations of instantaneous deformation fields α_i make sense (which is precisely the expectation of such an approach) and that their distribution is Gaussian, that is to say, that for any mode β_k , the distribution of the k -th principal component $\langle \beta_k | \alpha_i \rangle_{L^2}$ of the

observed fields α_i is Gaussian, which implies that the shape probability should be of the form

$$P(C) = P(\alpha) = \prod_k e^{-\frac{\langle \beta_k | \alpha \rangle_{L^2}^2}{2\sigma_k^2}} \times e^{-\frac{\|\text{Rem.}(\alpha)\|_2^2}{2\sigma_{\text{noise}}^2}}$$

where $\alpha = -\nabla_{\hat{\Gamma}} E^2(\hat{\Gamma}, R(C))$ with R being the best rigid motion that minimizes $E^2(\hat{\Gamma}, R(C))$ and where $\text{Rem.}(\alpha)$ stands for the remaining part $\alpha - \sum_k \langle \alpha | \beta_k \rangle_{L^2}$ that cannot be described by the eigenmodes. The parameter σ_{noise} stands for a standard deviation associated to this noise and can be chosen for instance equal to 0.01 times the lowest eigenvalue (σ_k). Note that the corresponding distance between a shape and the shape distribution is:

$$\sqrt{\sum_k \frac{1}{\sigma_k^2} \langle \alpha | \beta_k \rangle^2 + \frac{1}{\sigma_{\text{noise}}^2} \|\text{Rem.}(\alpha)\|_2^2}.$$

From a certain point of view, the level sets of this distance are ellipsoidal. This distance is a variation on the Mahalanobis distance.

3.2.3 Eigenmode Histograms

However the distribution is generally not Gaussian. Then the modes can be computed with ICA (independent component analysis) instead of PCA, and for each mode k the histogram h_k of the observed components $\langle \beta_k | \alpha_i \rangle$ can be drawn. The histograms may need some smoothing if the density of the distribution is low. Then the empirical probability is given by:

$$P(C) = P(\alpha) = \prod_k h_k(\langle \beta_k | \alpha \rangle).$$

3.3 Pre-Computing

Except for the first prior, α appears in the expression of the probability to maximize, which implies that the derivative of α with respect to C will have to be computed:

$$\nabla_C \alpha = -\nabla_C \nabla_{\hat{\Gamma}} E^2(\hat{\Gamma}, R(C))$$

Hence we need to compute the second order cross-derivative of E . In the case of the approximation of the Hausdorff distance, the computation was heroic but we were able to complete it; the resulting formula is too complicated to be reproduced here. It is available in the supplementary material.

The calculations happen to be sometimes much simpler if the energy is based not on a shape but on its signed distance function as in [11], that is to say if you

consider that the real object of interest is now a function defined on the whole space \mathbb{R}^n and not only its zero level, regardless of whether it is a distance function or not. The framework would then be similar to the one of Leventon et al.[8] which consists in the application of PCA to signed distance functions. However such an approach is questionable since a linear combination of signed distance functions is generally not a signed distance function and has sometimes a really unexpected zero level. That is why we keep on considering the Hausdorff distance approximation between shapes.

4 Toy Example

Starting with a toy example to show the strength of this approach, a small set of four similar rectangles with two kinds of outgrowths is considered (figure 1). For this particular example, there was no optimization concerning rigid motion. With the Gaussian eigenmode prior, the segmentation of a new rectangle that combines the two outgrowths as well as a third new one leads to a shape which can be described as a new combination of the two already observed outgrowths, ignoring the third one.

One interesting point of this method is that one does not need to choose the number of relevant eigenmodes since the importance of each of them is naturally described by its associated standard deviation. In order to show the interest of these characteristic modes of deformation, the standard deviation associated to the noise has been chosen to be equal to $5 \cdot 10^{-3}$ times the standard deviation associated to the highest eigenmode, that is to say that a “noisy” deformation field with null component on each mode will cost 200 times more than a field of same norm but collinear to the first mode. In order to be coherent, if there exist eigenmodes with eigenvalues smaller than the one associated to the noise, then they have to be forgotten and considered as noise. In the case of figure 1, the first two eigenmodes were found to have nearly the same eigenvalue and the two others were about a hundred times smaller, and indeed the segmentations with all modes or only the first two modes were the same.

Independently on the value of the weight of the shape prior providing this one is high enough, the qualitative behavior of the shape prior on the evolution is to “project” the evolving shape onto a linear combination of the eigenmodes, in the sense that the gradient of E from the mean $\hat{\Gamma}$ to the evolving shape C will progressively reduce its components on eigenmodes (and remaining noisy part) according to their standard deviation. As the distribution is here supposed to be Gaussian, any increasing of the weight of the shape prior will make the result a little nearer to the mean shape. However in the case of component histograms as proposed in section 3.2 this phenomenon would not appear since the shape with

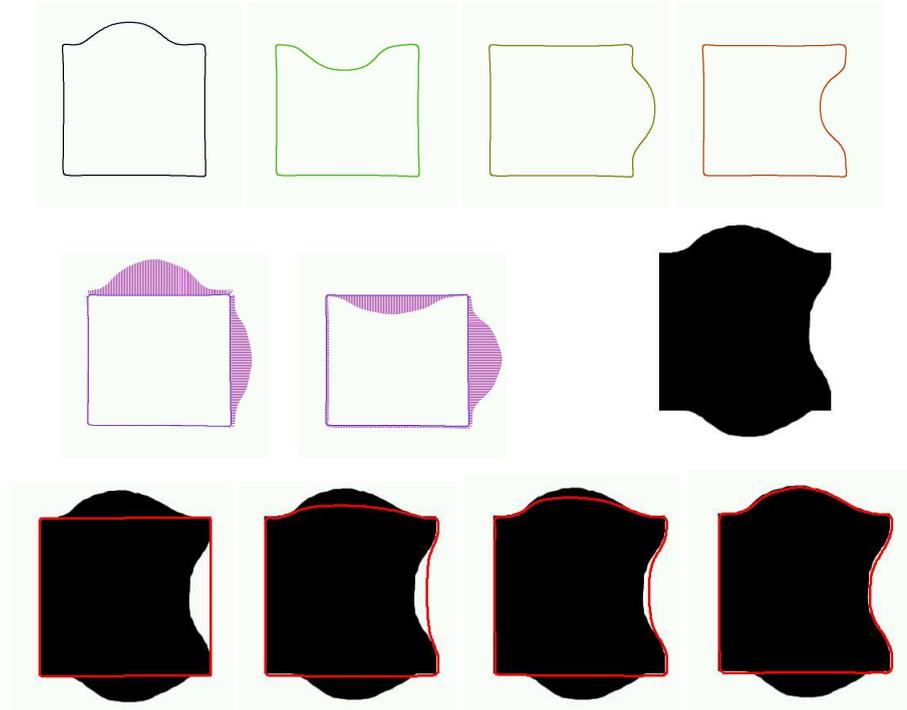


Figure 1: Top row: the four shapes that compose the sample set. Middle, left: the mean of these shapes (in blue), with the two first eigenmodes (in purple), successively. Middle, right: the new image to be segmented, built approximately as a new combination of the previously observed deformations added to a new non-observed deformation. Bottom: segmentation with the only knowledge of the mean, eigenmodes and eigenvalues, under the Gaussian distribution assumption which is obviously not satisfied (left: initialization; middle: some steps of the evolution; right: result at convergence).

highest *a priori* probability is not the mean anymore.

5 Rigid Registration

A set of 12 images of starfish² (see figure 2) has been segmented by hand. This could be automatized in the general case if the examples of the sample set are chosen so that they are easy images to segment with usual algorithms. The mean curve $\hat{\Gamma}$ of the set of starfish has been computed with the rigid-motion invariant framework proposed in section 2. The mean and its first eigenmodes are displayed

²found via Google Images

on figure 3 with amplitude proportional to their eigenvalue.



Figure 2: Some examples from the learning set of starfish.

We first consider the rigid shape prior $E^2(\hat{\Gamma}, R(C))$ in a new image segmentation task, where C is the current evolving shape and R a rigid motion. The shape C and the parameters of R are estimated simultaneously within a framework similar to the previous one.

There exist many approaches to image segmentation in the computer vision literature, for instance geodesic active contours [2] or region histograms [1, 9]. We have chosen here a region intensity histogram criterion. More precisely, as an image has “naturally” 256 grey levels, the intensity histograms of the inside h_I and outside h_O of the contour C are real-valued functions of an integer which can have 256 possible values. They associate to any grey level the number of such colored pixels in the corresponding region. The two histograms are supposed to be relatively homogeneous and as much different one from the other as possible. The segmentation criterion to minimize is a weighted sum of the length $|C|$ of the contour C plus the correlation between a slightly smoothed version of the intensity probability distributions of the two regions $p_I = h_I/|I|$ and $p_O = h_O/|O|$ (where $|I|$ is the area of the region I):

$$\frac{1}{256} \sum_{a \in \llbracket 0, 255 \rrbracket} (G_\sigma \star p_I)(a) (G_\sigma \star p_O)(a) + |C|$$

where $G_\sigma \star$ is the Gaussian smoothing with parameter σ .

The result of the combination of this criterion with the rigid shape prior is shown on figure 4: the location of the starfish is found, but of course the shape variability of the sample set has not been taken into account. Consequently, either the result of the segmentation is exactly the mean without any non-rigid deformation (if the weight of the shape prior is huge), or it is the mean plus any small deformation (without any priors on this deformation). This algorithm is simplistic since the prior is a fixed shape (up to rigid motion), but it helps finding a not-too-varying object as well as an occluded object (see figure 5 for comparison).

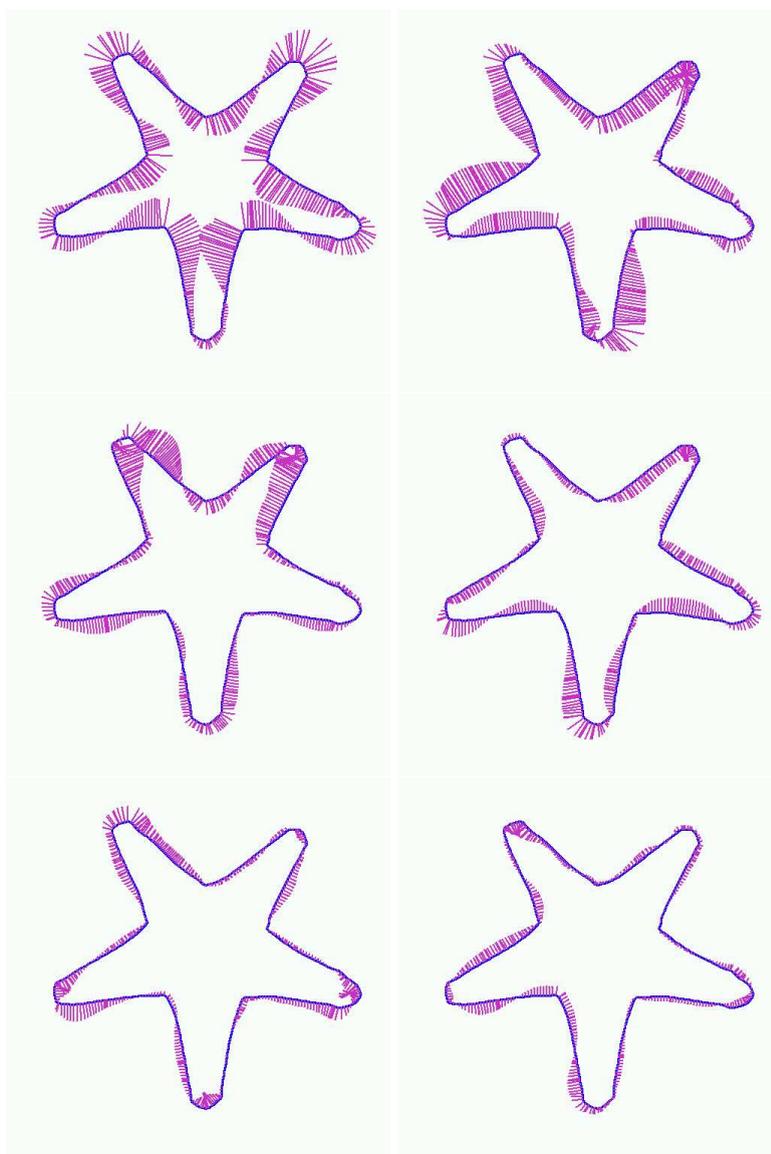


Figure 3: The mean of the set of starfish with its first six eigenmodes.

6 Gaussian Eigenmodes Examples

The results of the segmentation of the original image for different shape priors are shown in figure 6. On the top row are shown typical results for the previous segmentation criterion without any shape prior, for two different initializations. As the region based criterion has many local minima, the segmentation result depends strongly on the initialization. Not that the small white balls around the

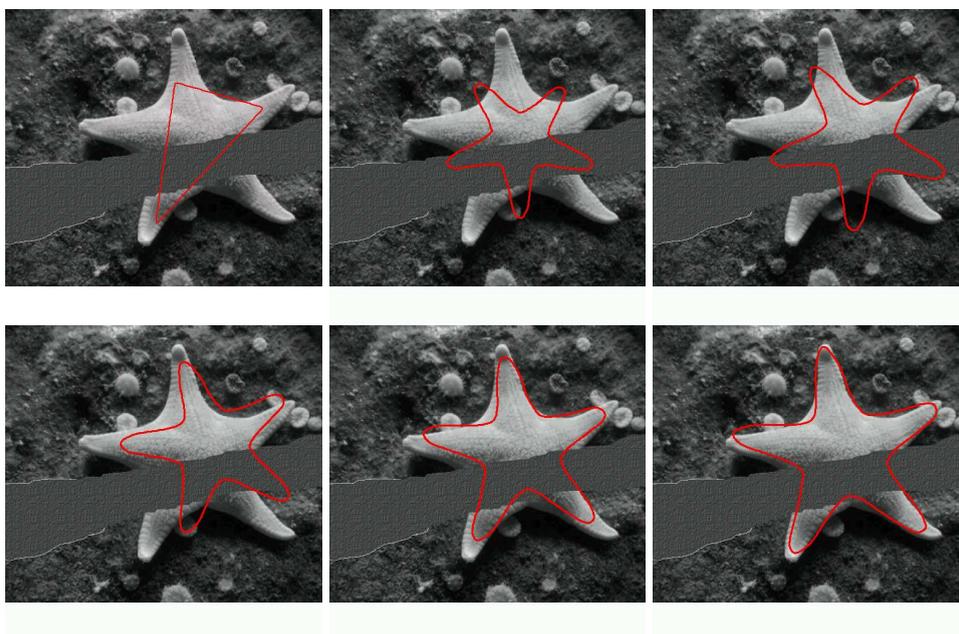


Figure 4: Top left: any reasonable initialization for the region intensity histogram criterion. A part of the image has been erased in order to increase the difficulty. Top middle: automatic change for the mean at the same location with similar size. Rest: some steps of the segmentation process with knowledge of the mean shape. This rigid criterion finds the location of starfish but lacks information about how to adapt the final shape. Indeed for a smaller weight of the prior, the result will include the small white balls connected to the starfish, as in figure 5.

starfish are difficult to distinguish from the starfish, and the shaded regions of the starfish have colors similar to the background. Then, on the middle row (left hand-side) is shown the result obtained with a rigid registration of the mean shape (as in figure 4). This result is much more “stable” in the sense that it can be obtained from any reasonable initialization. The reason for this is that the minimization is processed with respect to only few parameters (translation, orientation, scaling) instead of a whole shape (which is infinite-dimensional). Therefore in practice the space to be explored in order to find the solution is much smaller in the case of rigid registration. In order to allow some deformations around the registered mean shape, we start from the result of the rigid registration, and minimize the sum of the square of the distance to the mean shape and of the region histogram criterion. This minimization is computed with respect to both the evolving shape and the location parameters of the mean. For high values of the weight of the shape prior, the result is of course close to the one obtained by only rigid registration. For low

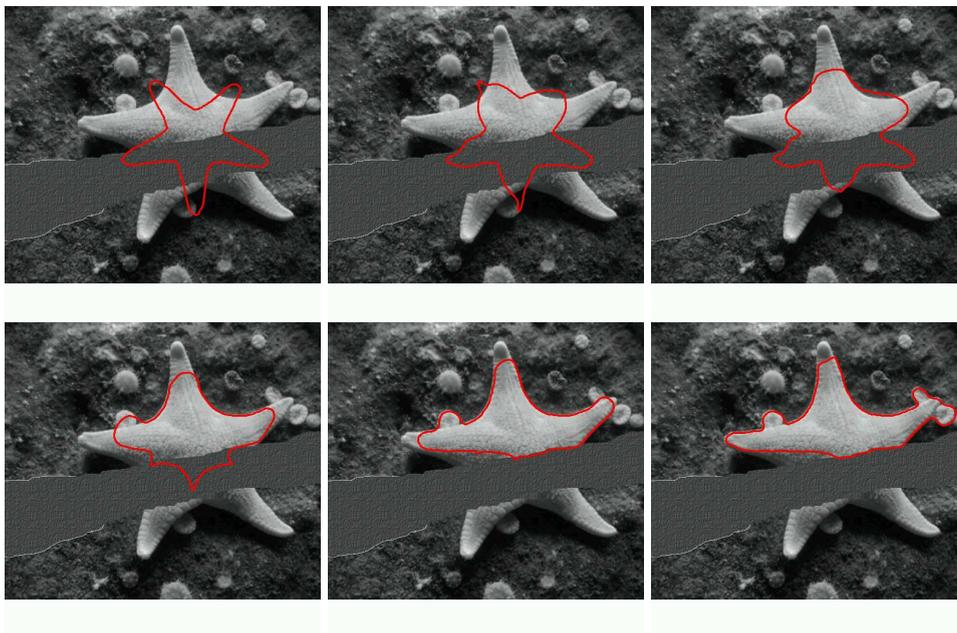


Figure 5: Top left: same initialization as before. Rest: some steps of segmentation process without any shape prior, for the same region criterion. The result lacks the global shape of the starfish and include the small white balls.

values, the result is close to the one obtained without shape prior, that is to say to one similar to the ones on the top row. A typical example is shown (middle row, right handside) for a middle value of the weight: as there is no prior on the deformations applied to the mean shape, the algorithm leads to outgrowths that are non-sensible for a starfish. For instance it includes the white balls within the starfish and let a deformation grow far inside the starfish in order to get rid of the shadow regions.

Finally, still starting from the result of the rigid registration, we minimize the sum of the Gaussian eigenmode shape prior and of the region histogram criterion. The evolving shape C is shown in red and the estimated mean shape location $R^{-1}(\hat{\Gamma})$ in blue. The deformations that are required for the inclusion of the white balls within the starfish have a heavy cost for the shape prior since they are not characteristic deformations of the mean shape. Consequently the algorithm finds globally the shape of the starfish, except a part of its shadowed regions which the region intensity criterion considers as included into the background (see top row). Note however that the deformations due to the shadow have been described as best as possible as resulting from a combination of eigenmodes. Therefore they have been reduced to a reasonable deformation that a starfish can undergo, and

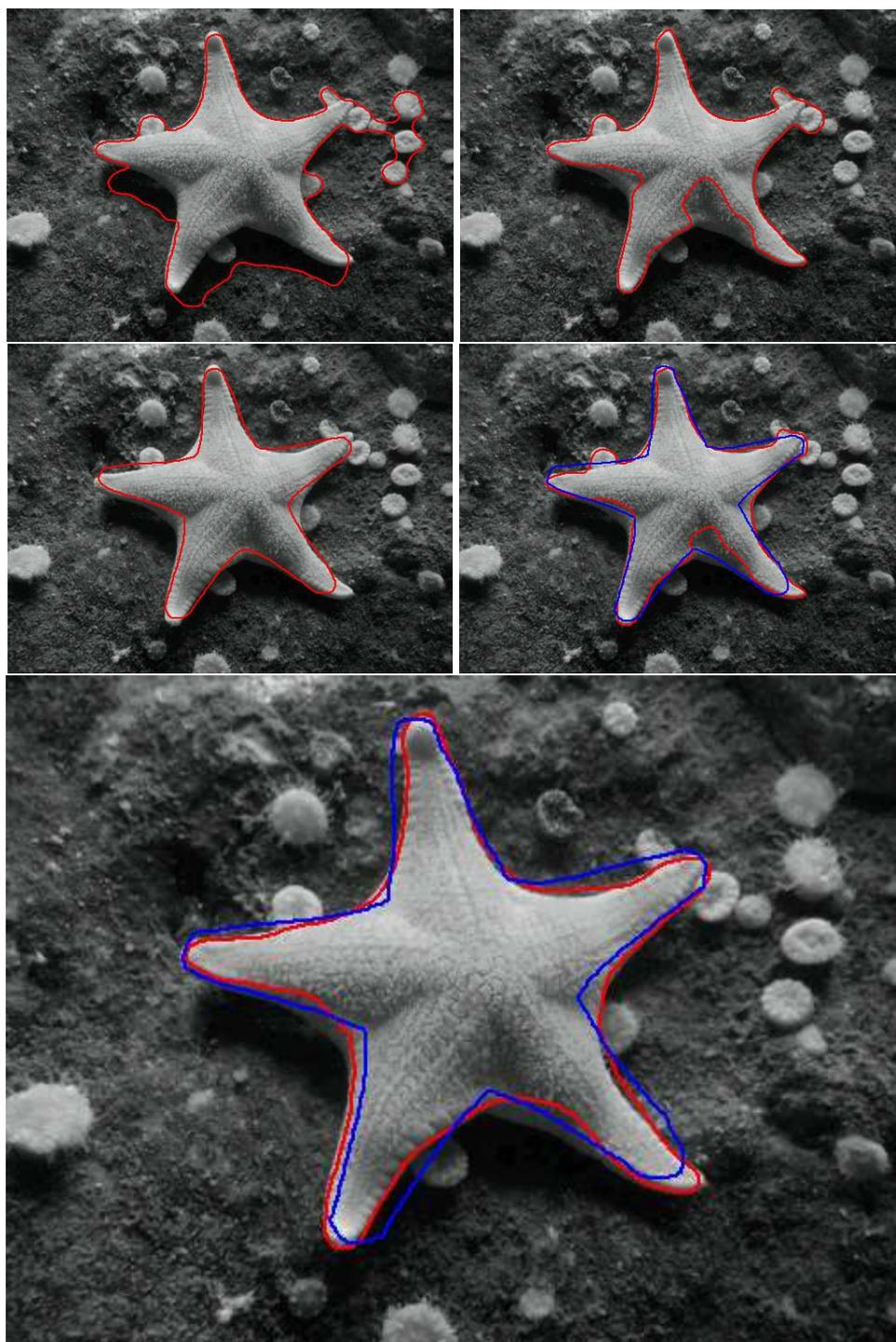


Figure 6: Segmentation results for different priors. See text for details. Top row: two examples of results obtained without any shape prior, for two different initializations. Middle row, left: result of rigid registration. Middle row, right: result of the segmentation (in red) with the (non-rigid) mean shape prior (whose estimated location is also shown, in blue). Bottom: result (in red) for the Gaussian eigenmode prior (with estimated location of the mean in blue).

they look far better than the observed ones in the other segmentation results.

Another example with a database of 14 boletus contours is shown. The mean and eigenmodes are shown on figure 7 and a segmentation task is performed on figure 8. The segmentation criterion is the similar to the previous one except that colors are considered instead of grey levels. Thus the color histogram is three dimensional. The segmentation process of a new image with this criterion is displayed on the first row of figure 8. With the same initialization but with also the rigid shape prior we obtain the evolution shown in the second row. Then the Gaussian eigenmode shape prior is used and leads to the final segmentation (bottom, right). The difference between the results is striking.

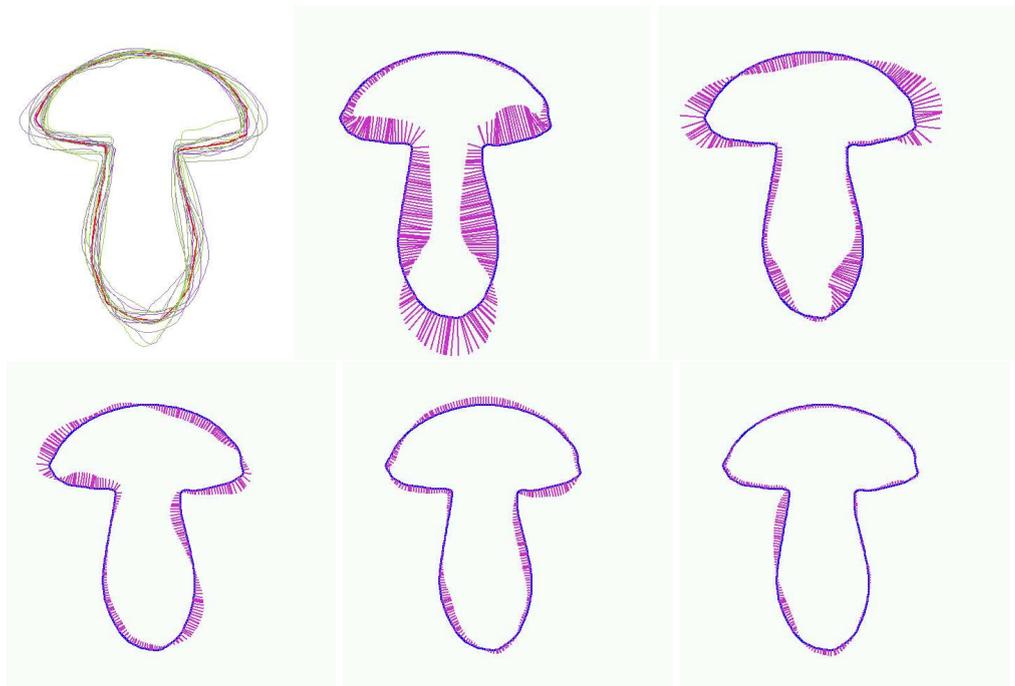


Figure 7: First image: automatic rigid alignment of the samples when computing the mean (thanks to the optimization with respect to the R_i). Then: mean of the set of mushrooms with its five first eigenmodes.

For all presented examples, the added computational cost due to the shape prior is reasonable. In these experiments, the total time cost with the prior was found to be about three to four times the total time cost without prior. Most of the time cost is due to the computation of a double integration needed by the approximation of the Hausdorff distance and could be reduced by optimizations.

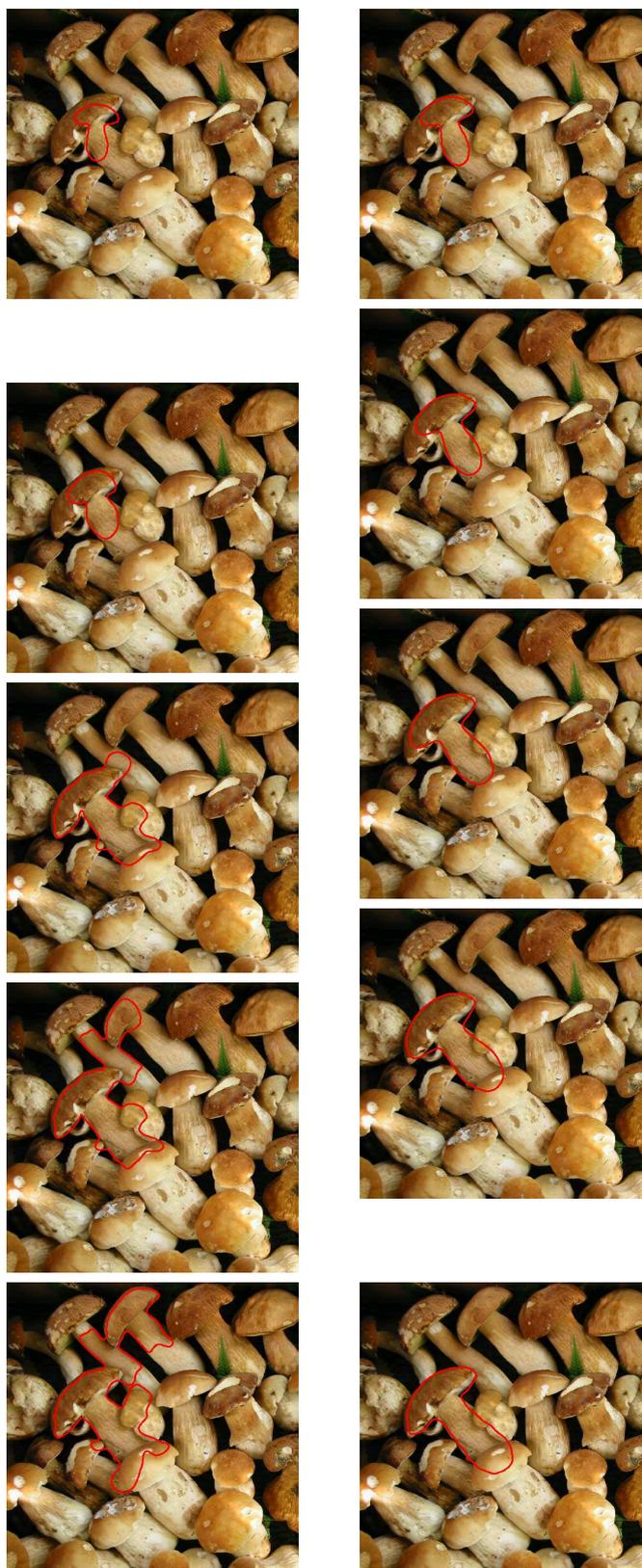


Figure 8: Top left: any reasonable initialization. Left row: without shape prior. Right row: rigid prior. Bottom right: with Gaussian eigenmodes. See text for details.

7 Conclusion

We have shown that it is feasible to define mathematically shape statistics that are relevant even on very small datasets and to turn them into a rigid-motion invariant shape prior for image segmentation at a reasonable computational cost. We have also shown that in practice, usual segmentation results are outperformed by the additional use of this shape prior.

References

- [1] Thomas Brox, Mikael Rousson, Rachid Deriche, and Joachim Weickert. Unsupervised segmentation incorporating colour, texture, and motion. In *10th International Computer Analysis of Images and Patterns*, number 2756 in Lecture Notes on Computer Science., pages 353–360. Springer Verlag, August 2003.
- [2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *The International Journal of Computer Vision*, 22(1):61–79, 1997.
- [3] G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, February 2005.
- [4] D. Cremers, T. Kohlberger, and C. Schnörr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, September 2003. Special Issue on Kernel and Subspace Methods in Computer Vision.
- [5] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, X(IV):215–310, 1948.
- [6] H. Karcher. Riemannian centre of mass and mollifier smoothing. *Comm. Pure Appl. Math*, 30:509–541, 1977.
- [7] W. Kendall. Probability, convexity, and harmonic maps with small image i: uniqueness and fine existence. *Proc. London Math. Soc.*, 61(2):371–406, 1990.
- [8] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 316–323, Hilton Head Island, South Carolina, June 2000. IEEE Computer Society.

- [9] N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *The International Journal of Computer Vision*, 46(3):223–247, 2002.
- [10] Xavier Pennec. *L'Incertitude dans les Problèmes de Reconnaissance et de Recalage – Applications en Imagerie Médicale et Biologie Moléculaire*. PhD thesis, Ecole Polytechnique, Palaiseau (France), December 1996.
- [11] Mikael Rousson and Nikos Paragios. Shape priors for level set representations. In *ECCV 2002*, volume 2, pages 78–92, May 2002.