Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation

Mikael Rousson and Daniel Cremers

Department of Imaging and Visualization
Siemens Corporate Research, Princeton, NJ

{mikael.rousson,daniel.cremers}@siemens.com

Abstract. We propose a nonlinear statistical shape model for level set segmentation which can be efficiently implemented. Given a set of training shapes, we perform a kernel density estimation in the low dimensional subspace spanned by the training shapes. In this way, we are able to combine an accurate model of the statistical shape distribution with efficient optimization in a finite-dimensional subspace. In a Bayesian inference framework, we integrate the nonlinear shape model with a nonparametric intensity model and a set of pose parameters which are estimated in a more direct data-driven manner than in previously proposed level set methods. Quantitative results show superior performance (regarding runtime and segmentation accuracy) of the proposed nonparametric shape prior over existing approaches.

1   Introduction

Originally proposed in [5, 11] as a means to propagate interfaces in time, the level set method has become increasingly popular as a framework for image segmentation. The key idea is to represent an interface $\Gamma \subset \Omega$ in the image domain $\Omega \subset \mathbb{R}^3$ implicitly as the zero level set of an embedding function $\phi : \mathbb{R}^3 \to \Omega$:

$$\Gamma = \{ x \in \Omega \mid \phi(x) = 0 \}, \tag{1}$$

and to evolve Γ by propagating the embedding function ϕ according to an appropriate partial differential equation. The first applications of this level set formalism for the purpose of image segmentation were proposed in [1, 10, 7]. Two key advantages over explicit interface propagation are the independence of a particular parameterization and the fact that the implicitly represented boundary Γ can undergo topological changes such as splitting or merging. This makes the framework well-suited for the segmentation of several objects or multiply-connected objects.

When segmenting medical images, one commonly has to deal with noise, missing or misleading image information. For certain imaging modalities such as ultrasound or CT, the structures of interest do not differ much from their background in terms of their intensity distribution — see Figure 1. Therefore they can no longer be accurately segmented based on the image information

| Cardiac ultrasound | Histograms | Prostate CT | Histograms |
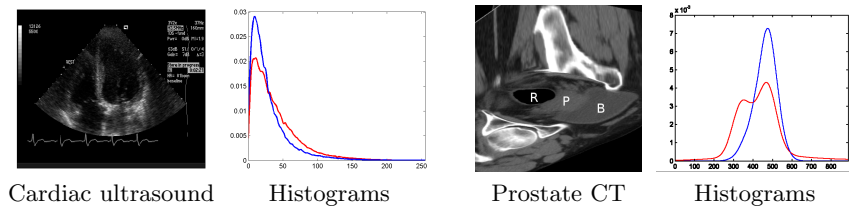
Fig. 1. Segmentation challenges and estimated intensity distributions.
The two curves on the right correspond to the empirical probability of intensities inside and outside the left ventricle (for the ultrasound image) and the prostate (for the CT image). The region-based segmentation of these structures is a challenging problem, because objects and background have similar histograms. Our segmentation scheme optimally exploits the estimated probabilistic intensity models.

alone. In recent years, researchers have therefore proposed to enhance the level set method with statistical shape priors. Given a set of training shapes, one can impose information about which segmentations are *a priori* more or less likely. Such prior shape information was shown to drastically improve segmentation results in the presence of noise or occlusion [9, 16, 3, 14, 4, 6]. Most of these approaches are based on the assumption that the training shapes, encoded by their signed distance function, form a Gaussian distribution. This has two drawbacks: Firstly, the space of signed distance functions is not a linear space, therefore, the mean shape and linear combinations of eigenmodes are typically no longer signed distance functions. Secondly, even if the space were a linear space, it is not clear why the given set of sample shapes should be distributed according to a Gaussian density. In fact, as we will demonstrate in this work, they are generally not Gaussian distributed. Recently, it was proposed to use nonparametric density estimation in the space of level set functions [3] in order to model nonlinear[1] distributions of training shapes. While this resolves the above problems, one sacrifices the efficiency of working in a low-dimensional subspace (formed by the first few eigenmodes) to a problem of infinite-dimensional optimization.

In the present paper, we propose a framework for knowledge-driven level set segmentation which integrates three contributions: Firstly, we propose a statistical shape prior which combines the efficiency of low-dimensional PCA-based methods with the accuracy of nonparametric statistical shape models. The key idea is to perform kernel density estimation in a linear subspace which is sufficiently large to embed all training data. Secondly, we propose to estimate pose and translation parameters in a more data-driven manner. Thirdly, we optimally exploit the intensity information in the image by using probabilistic intensity models given by kernel density estimates of previously observed intensity distributions.

---

[1] The term *nonlinear* refers to the fact that the manifold of permissible shapes is not merely a linear subspace.

2   Level Set Segmentation as Bayesian Inference

The goal of level set segmentation can be formulated as the estimation of the optimal embedding function $\phi\colon \Omega \to \mathbb{R}$ given an image $I\colon \Omega \to \mathbb{R}$. In the Bayesian framework, this can be computed by maximizing the posterior distribution

$$\mathcal{P}(\phi \,|\, I) \propto \mathcal{P}(I \,|\, \phi)\, \mathcal{P}(\phi). \qquad (2)$$

The maximization of (2) results in a problem of infinite-dimensional optimization. Given a set of training shapes encoded by their signed distance functions $\{\phi_i\}_{i=1..N}$, Tsai et al. [16] proposed to reduce the segmentation problem to one of finite-dimensional optimization by constraining the optimization problem to the finite-dimensional subspace spanned by the training shapes.

In this paper, we make use of this compact representation of the embedding function. Given the distance d on the space of signed distance functions defined by: $d^2(\phi_1, \phi_2) = \int_\Omega \left(\phi_1(x) - \phi_2(x)\right)^2 dx$, we align the set of training shapes with respect to translation and rotation. Subsequently, we constrain the level set function ϕ to a parametric representation of the form:

$$\phi_{\boldsymbol{\alpha}, h, \theta}(x) = \phi_0(R_\theta x + h) \,+\, \sum_{i=1}^{n} \alpha_i\, \psi_i(R_\theta x + h), \qquad (3)$$

where $\phi_0(x) = \frac{1}{N}\sum_{i=1}^{N}\phi_i(x)$ represents the mean shape, $\{\psi_i(x)\}_{i=1..n}$ are the eigenmodes of the distribution, and $n < N$ is the dimension of the subspace spanned by the N training shapes. The parameter vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ models shape deformations, while the parameters $h \in \mathbb{R}^3$ and $\theta \in [0, 2\pi]^3$ model translation and rotation of the respective shape.[2]

The infinite-dimensional Bayesian inference problem in (2) is therefore reduced to a finite-dimensional one where the conditional probability

$$\mathcal{P}(\boldsymbol{\alpha}, h, \theta \,|\, I) \propto \mathcal{P}(I \,|\, \boldsymbol{\alpha}, h, \theta)\, \mathcal{P}(\boldsymbol{\alpha}, h, \theta), \qquad (4)$$

is optimized with respect to the shape parameters $\boldsymbol{\alpha}$, and the transformation parameters h and θ. In the following, we will assume a uniform prior on these transformation parameters, i.e. $\mathcal{P}(\boldsymbol{\alpha}, h, \theta) = \mathcal{P}(\boldsymbol{\alpha})$. In the next section, we will discuss three solutions to model this shape prior.

3   An Efficient Nonparametric Statistical Shape Model

Given a set of aligned training shapes $\{\phi_i\}_{i=1..N}$, we can represent each of them by their corresponding shape vector $\{\boldsymbol{\alpha}_i\}_{i=1..N}$. In this notation, the goal of statistical shape learning is to infer a statistical distribution $\mathcal{P}(\boldsymbol{\alpha})$ from these sample shapes. Two solutions which have been proposed are based on the assumptions

---

[2] In our applications, where the scale of objects is known, a generalization to larger transformations groups (e.g. similarity or affine) did not appear useful.

Uniform density   Gaussian density   Kernel density
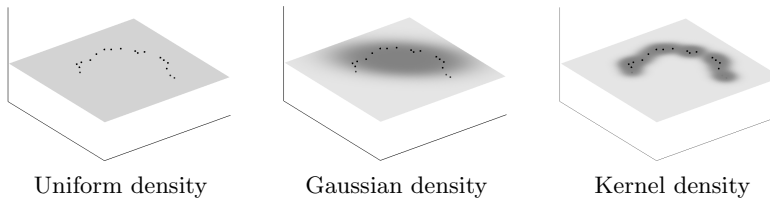
Fig. 2. Schematic plots of different density estimates within a subspace.
Darker shading indicates areas of high probability density for the respective models.
The kernel density estimator adapts to the training data more flexibly since it does
not rely on specific assumptions about the shape of the distribution.

that the training shapes can be approximated by a **uniform distribution** [16,
14]: $\mathcal{P}(\boldsymbol{\alpha}) = \text{const.}$, or by a **Gaussian distribution** [9]:

$$\mathcal{P}(\boldsymbol{\alpha}) \propto \exp\left(-\boldsymbol{\alpha}^\top \Sigma^{-1} \boldsymbol{\alpha}\right), \quad \text{where } \Sigma = \frac{1}{N} \sum_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^\top. \tag{5}$$

In the present paper, we propose to make use of nonparametric density es-
timation [13] to approximate the shape distribution within the linear subspace.
We model the shape distribution by the **kernel density estimate**:

$$\mathcal{P}(\boldsymbol{\alpha}) = \frac{1}{N\sigma} \sum_{i=1}^{N} K\left(\frac{\boldsymbol{\alpha} - \boldsymbol{\alpha}_i}{\sigma}\right), \quad \text{where } K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right). \tag{6}$$

There exist various methods to automatically estimate appropriate values for
the width σ of the kernel function, ranging from k-th nearest neighbor estimates
to cross-validation and bootstrapping. In this work, we simply set σ to be the
average nearest neighbor distance: $\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \min_{j \neq i} |\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j|^2$.

In the context of level set based image segmentation, the kernel density esti-
mator (6) has two advantages over the uniform and Gaussian distributions:

– The assumptions of uniform distribution or Gaussian distribution are gener-
  ally not fulfilled. The kernel density estimator, on the other hand, is known to
  approximate arbitrary distributions. Under mild assumptions, it was shown
  to converge to the true distribution in the limit of infinite sample size. We
  refer to [15] for a proof.

– The space of signed distance functions is known to not be a linear space.
  Therefore, neither the mean shape ϕ_0 nor a linear combination of eigenmodes
  as in (3) will in general be a signed distance function. As a consequence, the
  functions $\phi(x)$ favored by the uniform or the Gaussian distribution cannot
  be expected to be signed distance functions. The kernel density estimator
  (6), on the other hand, favors shape vector α which are in the vicinity of the
  sample shape vectors α_i. By construction, these vector correspond to signed
  distance functions. In fact: **In the limit of infinite sample size, the
  distribution inferred by the kernel density estimator (6) converges
  towards a distribution on the manifold of signed distance functions.**

Figure 2 shows schematic plots of the three methods for a set of sample data spanning a two-dimensional subspace in $\mathbb{R}^3$. The kernel density estimator clearly captures the distribution most accurately.

In analogy to the shape learning, we make use of kernel density estimation to learn the conditional probability for the intensity function I in (4) from examples. A similar precomputation of intensity distributions by means of mixture models was proposed in [12]. Given a set of presegmented training images, the kernel density estimate of the intensity distributions p_{in} and p_{out} of object and background are given by the corresponding smoothed intensity histograms. This has two advantages: Firstly, the kernel density estimator does not rely on specific assumptions about the shape of the distribution. Figure 1 shows that the intensity distributions for ultrasound and CT images are not well approximated by Gaussian or Laplacian models. Secondly, in contrast to the joint estimation of intensity distributions (cf. [2, 8]), this simplifies the segmentation process which no longer requires an updating of intensity models. Moreover, we found the segmentation process to be more robust to initialization in numerous experiments.

4 Energy Formulation and Minimization

Maximizing the posterior probability in (2), or equivalently minimizing its negative logarithm, will generate the most probable segmentation of a given image. With the nonparametric models for shape and intensity introduced above, this leads to an energy of the form

$$E(\boldsymbol{\alpha}, h, \theta) = -\log \mathcal{P}(I|\boldsymbol{\alpha}, h, \theta) - \log \mathcal{P}(\boldsymbol{\alpha}), \tag{7}$$

The nonparametric intensity model permits to express the first term and equation (6) gives exactly the second one. With the Heaviside step function H and the short hand $H_\phi = H(\phi_{\boldsymbol{\alpha}, h, \theta}(x))$, we end up with:

$$E(\boldsymbol{\alpha}, h, \theta) = -\int_\Omega H_\phi \log p_{in}(I) + (1 - H_\phi) \log p_{out}(I) dx - \log\left(\frac{1}{N\sigma} \sum_{i=1}^{N} K\left(\frac{\boldsymbol{\alpha} - \boldsymbol{\alpha}_i}{\sigma}\right)\right),$$

With $e(x) = \left[\log \frac{p_{out}(I(x))}{p_{in}(I(x))}\right]$, $K_i = K\left(\frac{\boldsymbol{\alpha} - \boldsymbol{\alpha}_i}{\sigma}\right)$, and $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_n)$, we obtain the following system of coupled gradient descent equations:

$$
\begin{cases}
\dfrac{d\boldsymbol{\alpha}}{dt} = \displaystyle\int_\Omega \delta(\phi_{\boldsymbol{\alpha}, h, \theta}(x))\, \boldsymbol{\psi}(R_\theta x + h)\, e(x)\, dx + \dfrac{1}{\sigma^2} \dfrac{\sum_{i=1}^{N} (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}) K_i}{\sum_{i=1}^{N} K_i}, \\[4mm]
\dfrac{dh}{dt} = \displaystyle\int_\Omega \delta(\phi_{\boldsymbol{\alpha}, h, \theta}(x))\, \nabla\phi_{\boldsymbol{\alpha}, h, \theta}(x)\, e(x)\, dx, \\[4mm]
\dfrac{d\theta}{dt} = \displaystyle\int_\Omega \delta(\phi_{\boldsymbol{\alpha}, h, \theta}(x))\, (\nabla\phi_{\boldsymbol{\alpha}, h, \theta}(x) \cdot \nabla_\theta Rx)\, e(x)\, dx.
\end{cases}
\tag{8}
$$

In applications, we solve these equations by initializing the shape $\boldsymbol{\alpha}$ with the mean shape ($\boldsymbol{\alpha} = 0$) and the transformation parameters h and θ with some

| Initialization | No prior | Uniform prior | Kernel prior |
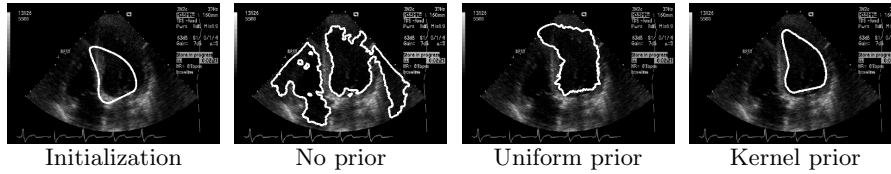
Fig. 3. Model Comparison. Level set segmentations obtained without prior, with a uniform prior in the subspace and with a kernel prior in the subspace. In contrast to the uniform prior, the nonparametric prior accurately constrains the segmentation to a submanifold of familiar shapes (90% correctly classified, 2.7% false positives).

reasonable estimates. Subsequently, we discretize the above partial differential equations by a standard finite difference scheme.

Note that in all equations, the Dirac delta function δ appears as factor inside the integrals over the image domain Ω. This allows to restrict all computations to a narrow band around the zero crossing of ϕ. While the evolution of translation and pose parameters h and θ are merely driven by the data term $e(x)$, the shape vector $\boldsymbol{\alpha}$ is additionally drawn towards each training shape with a strength that decays exponentially with the distance to the respective shape.

5   Experimental Results and Validation

Heart segmentation from ultrasound images

Figure 3 shows experimental results obtained for the segmentation of the left ventricle in 2D cardiac ultrasound sequences, using shape priors constructed from a set of 21 manually segmented training images. In contrast to the segmentation with uniform prior (top row), the nonparametric statistical shape prior allows to accurately constrain the segmentation (bottom row). This becomes particularly apparent in areas where the data term is too weak. As a quantitative evaluation we computed the percentage of correctly classified object pixels and that of mis-classified ones. During energy minimization, the percentage of correctly classified pixels increases from 56% to 90% while the percentage of false positives decreases from 27% to 2.7% by using the kernel prior. Using the uniform prior, we attain 92% correctly classified, yet the percentage of false positives increases to 42%: Merely constraining the boundary evolution to the linear subspace spanned by the training shapes is insufficient to provide for accurate segmentation results.

Prostate segmentation from 3D CT images

We built a nonparametric 3D shape model of the prostate using 12 manually extracted prostates (with seminal vesicles) collected from two different patients. In contrast to existing work, we subsequently used a single shape model for the segmentation of images from different patients.

We employed a leave-one-out strategy by removing the image of interest from the training phase. Figure 5 shows 2D cuts of a few results obtained using this
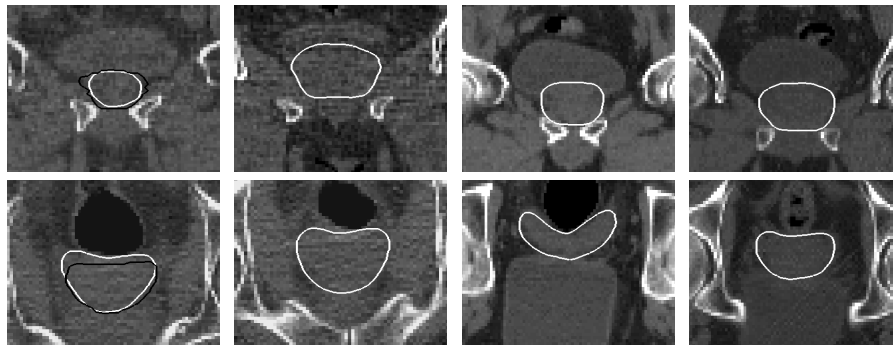
Fig. 4. Prostate segmentation for 2 **patients with the same shape model.**
Each column shows coronal and axial slices of the same segmentation, for the first
patient (left two columns) and the second one (last two). The first column also
shows the manual segmentation (black contour).

3D view   Kernel/Uniform   Kernel/Gaussian   Kernel/Manual
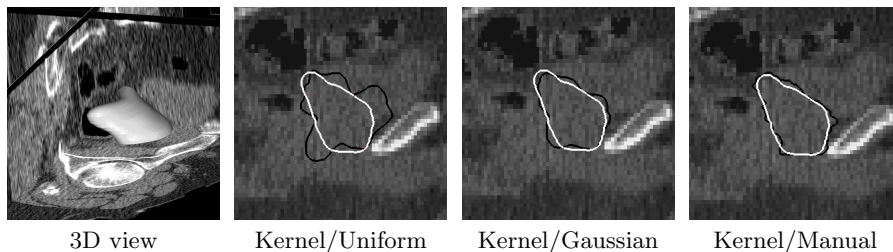
**Fig. 5. Comparison of the segmentations obtained with the kernel prior
(white) and with alternative approaches (black).**

strategy. With a one-click initialization inside the organ, the algorithm lead to a
steady-state solution in less than 20 seconds. We obtained 86% successfully clas-
sified organ voxels and 11% mis-classified organ voxels. This compares favorably
to the intra-patients results reported in [6]. Figure 4 provides qualitative com-
parisons to the manual segmentation, as well as to the segmentations obtained
with uniform and Gaussian approximations of the shape distribution.

6   Conclusion

We proposed an efficient and accurate statistical shape prior for level set seg-
mentation which is based on nonparametric density estimation in the linear
subspace spanned by the level set surfaces of a set of training shapes. In addi-
tion, our segmentation scheme integrates nonparametric estimates of intensity
distributions and efficient optimization of pose and translation parameters. We
reported quantitative evaluation of segmentation accuracy and speed for cardiac
ultrasound images and for 3D CT images of the prostate. These indicate that
the proposed nonparametric shape prior outperforms previously proposed shape
priors for level set segmentation.

Acknowledgments

References

1. V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numer. Math.*, 66:1–31, 1993.
2. T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.
3. D. Cremers, S. J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. In *Pattern Recognition*, volume 3175 of *LNCS*, pages 36–44. Springer, 2004.
4. E.B. Dam, P.T. Fletcher, S. Pizer, G. Tracton, and J. Rosenman. Prostate shape modeling based on principal geodesic analysis bootstrapping. In *MICCAI*, volume 2217 of *LNCS*, pages 1008–1016, September 2004.
5. A. Dervieux and F. Thomasset. A finite element method for the simulation of Raleigh-Taylor instability. *Springer Lect. Notes in Math.*, 771:145–158, 1979.
6. D. Freedman, R.J. Radke, T. Zhang, Y. Jeong, D.M. Lovelock, and G.T. Chen. Model-based segmentation of medical imagery by matching distributions. *IEEE Trans Med Imaging*, 24(3):281–292, March 2005.
7. S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi. Gradient flows and geometric active contour models. In *Proc. IEEE Intl. Conf. on Comp. Vis.*, pages 810–815, Boston, USA, 1995.
8. J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *Int. Conf. on Image Processing*, volume 3, pages 797–800, 2002.
9. M. Leventon, W. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, volume 1, pages 316–323, Hilton Head Island, SC, 2000.
10. R. Malladi, J. A. Sethian, and B. C. Vemuri. A topology independent shape modeling scheme. In *SPIE Conference on Geometric Methods in Computer Vision II*, volume 2031, pages 246–258, 1994.
11. S. J. Osher and J. A. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton–Jacobi formulations. *J. of Comp. Phys.*, 79:12–49, 1988.
12. N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. of Computer Vision*, 46(3):223–247, 2002.
13. F. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27:832–837, 1956.
14. M. Rousson, N. Paragios, and R. Deriche. Implicit active shape models for 3d segmentation in MRI imaging. In *MICCAI*, pages 209–216, 2004.
15. B. W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, London, 1992.
16. A. Tsai, A. J. Yezzi, and A. S. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. on Medical Imaging*, 22(2):137–154, 2003.