

Multi-Object Segmentation using Coupled Nonparametric Shape and Relative Pose Priors

Mustafa Gökhan Uzunbaş^a, Octavian Soldea^a, Müjdat Çetin^a, Gözde Ünal^a, Aytül Erçil^a,
Devrim Unay^b, Ahmet Ekin^b, and Zeynep Firat^c

^a Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, 34956 Turkey;

^b The Video Processing and Analysis Group, Philips Research Europe, Eindhoven, The Netherlands;

^c The Radiology Department of the Yeditepe University Hospital, Istanbul, Turkey

ABSTRACT

We present a new method for multi-object segmentation in a maximum *a posteriori* estimation framework. Our method is motivated by the observation that neighboring or coupling objects in images generate configurations and co-dependencies which could potentially aid in segmentation if properly exploited. Our approach employs coupled shape and inter-shape pose priors that are computed using training images in a nonparametric multivariate kernel density estimation framework. The coupled shape prior is obtained by estimating the joint shape distribution of multiple objects and the inter-shape pose priors are modeled via standard moments. Based on such statistical models, we formulate an optimization problem for segmentation, which we solve by an algorithm based on active contours. Our technique provides significant improvements in the segmentation of weakly contrasted objects in a number of applications. In particular for medical image analysis, we use our method to extract brain Basal Ganglia structures, which are members of a complex multi-object system posing a challenging segmentation problem. We also apply our technique to the problem of handwritten character segmentation. Finally, we use our method to segment cars in urban scenes.

Keywords: segmentation, active contours, shape prior, relative pose prior, kernel density estimation, moments.

1. INTRODUCTION

The availability in recent years of a broad variety of 2D and 3D images has presented new problems and challenges for the scientific community. In this context, segmentation is still a central research topic¹⁻⁴. A significant amount of research was performed during the past three decades towards completely automated solutions for general-purpose image segmentation. Variational techniques^{4,5}, statistical methods^{6,7}, combinatorial approaches⁸, curve-propagation techniques⁹, and methods that perform non-parametric clustering¹⁰ are some examples.

In contour-propagation approaches, which our framework is also based on, an initial contour estimate of the structure boundary is provided and various optimization methods are used to refine the initial estimate based on the input image data. This approach, called active contours, is based on the optimization of an energy functional using partial differential equations. In the definition of the energy functional, earlier methods use the boundary information for the objects of interest^{9,11}. More recent methods use regional information on intensity

Further author information: (Send correspondence to O.S.)

G.U.: E-mail: uzunbas@su.sabanciuniv.edu, Telephone: ++1(609)5587016

O.S.: E-mail: octavian@sabanciuniv.edu, Telephone: +905457952874

M.C.: E-mail: mcetin@sabanciuniv.edu, Telephone: +902164839594

G.U.: E-mail: gozdeunal@sabanciuniv.edu, Telephone: +902164839553

A.E.: E-mail: aytulercil@sabanciuniv.edu, Telephone: +902164839543

D.U.: E-mail: devrim.unay@philips.com, Telephone: +31-40-274 6156

A.E.: E-mail: ahmet.ekin@philips.com, Telephone: +31-40-274 5848

Z.F.: E-mail: zfirat@yeditepe.edu.tr, Telephone: +902165784363

statistics such as the mean or variance of an area^{12,13}. In most recent active contour models, there has been an increasing interest in using prior models for the shapes to be segmented. The proposed prior models are based on distance functions, implicit representations, and relationships among different shapes, including pose and other geometrical relationships^{3,14–21}.

In this context, there are numerous automatic segmentation methods that enforce constraints on the underlying shapes. In Ref. 14, the authors introduce a mathematical formulation to constrain an implicit surface to follow global shape consistency while preserving its ability to capture local deformations. Closely related with Ref. 14, in Ref. 16 and 17, the authors employ average shapes and modes of variation through principal component analysis (PCA) in order to capture the variability of shapes. However, this technique can handle only unimodal, Gaussian-like shape densities. In Ref. 16, the image and the prior term are well separated, while a maximum *a posteriori* (MAP) criterion is used for segmentation. In Ref. 17, a region-driven statistical measure is employed towards defining the image component of the function, while the prior term involves the projection of the contour to the model space using a global transformation and a linear combination of the basic modes of variation. In Refs. 18,19, and 20, the authors use shape models that refer only to an average shape in an implicit form, and the prior terms refer to projection of the evolving contours via similarity transformations.

As an alternative solution to PCA limitations, Ref. 22 proposes a principal geodesic analysis (PGA) model. As another solution to the limitation of PCA and unimodal Gaussian distribution models, techniques based on nonparametric shape densities learned from training shapes have been proposed in Refs. 4,23. In these works, the authors assume that the training shapes are drawn from an unknown shape distribution, which is estimated by extending a Parzen density estimator to the space of shapes. The authors formulate the segmentation problem as a MAP estimation problem, where they use a nonparametric shape prior. In particular, the authors construct the prior information in terms of a shape prior distribution such that for a given arbitrary shape one can evaluate the likelihood of observing this shape among shapes of a certain category.

Simultaneous multi-object segmentation is an important direction of research, since in many applications the objects to be segmented are often highly correlated. This information can be used to impose further constraints on the boundary estimation problem. Although nonparametric priors have been successful in capturing nonlinear shape variability, until now they have not been used in multi-object segmentation techniques. In this work, we demonstrate the potential of nonparametric priors towards accurate multi-object segmentation, by modeling both the shapes and the inter-shape relationships among the components. An integration of these relationships into the segmentation process can provide improved accuracy and robustness^{24–26}. A limited amount of work has been performed towards automatic simultaneous detection and segmentation of multiple organs. In Ref. 24, a joint prior based on a parametric shape model is proposed to capture co-variations shared among different shape classes, which improves the performance of single object based segmentation. With a similar approach and using a Bayesian framework, in Refs. 25,26, joint prior information about multiple objects is used to capture the dependencies among different shapes, where objects with clearer boundaries are used as reference objects to provide constraints in the segmentation of poorly contrasted objects.

Among spatial dependencies between multiple objects, one basic aspect is inter-shape pose analysis²⁷. Neighboring objects usually exhibit strong mutual spatial dependencies. In this context, Ref. 28 proposes a solution for the segmentation problem in the presence of a hierarchy of ordered spatial objects. In Ref. 29, the authors model the shape and pose variability of sets of multiple objects using principal geodesic analysis (PGA), which is an extension of the standard technique of principal component analysis (PCA) into the nonlinear Riemannian space. In these works, joint analysis of the objects is advocated over individual analysis.

Bearing in mind that segmentation is equivalent to extracting the shape and the pose of the boundary of the object, prior information on both shape and pose would be helpful in segmentation. Moreover, the relative shape arrangements among these neighbors can be modeled employing statistical information from a training set. In this paper, we introduce such statistical prior models of multiple-objects into an active contour segmentation method in a nonparametric MAP estimation framework. In this framework, we define two prior probability densities: one on the shape and the other one on the inter-shape (or relative) pose of the objects of interest. Both of the densities are evaluated during the evolution of active-contours, aiming an energy functional minimization. Our multi-object, coupled shape prior computation is an extension of the work in Ref. 4 and 23, where nonparametric density estimates of only single object shapes are computed. We use multivariate Parzen

density estimation to estimate the unknown joint density of multiple object shapes. Our coupled shape prior allows for simultaneous multi-object segmentation. As compared to existing methods in Refs. 24, 26, which are based on multi-object priors, our approach takes advantage of nonparametric density estimates in order to capture non-linear shape variability. In addition to shape priors, we also introduce inter-shape pose priors into segmentation. We compute the probability distribution of inter-shape pose using nonparametric density estimation, again. For inter-shape pose representations, we use standard moments, which are intrinsic to shape and have natural physical interpretations³⁰. Standard moments describe, among other features, the size, the mass center, and the orientation of the analyzed objects. In addition, the evaluation of moments is computationally attractive. We observe that our inter-shape pose prior helps the active contours evolve towards more accurate boundaries. To the best of our knowledge, our approach is the first scheme of multi-object segmentation, which employs coupled nonparametric shape and inter-shape pose priors based on moment computations in a probabilistic framework. We present experiments in a number of applications involving medical, natural, and handwriting imagery.

2. SEGMENTATION BASED ON SHAPE AND POSE PRIORS

We propose a shape prior and an inter-shape pose prior model embedded in an active contour framework. We advocate the use of these tools, bearing in mind that shape and inter-shape priors can be efficiently learnt and modeled³¹ and the active contour framework is a convenient tool in managing the evolution of the segmenting curves and surfaces.

First, in Section 2.1, we introduce a general segmentation framework. Next, in Section 2.2, we describe our formulation for coupled shape prior based segmentation. In Section 2.3, we describe our inter-shape pose prior on top of coupled shape priors in the same framework. In Section 2.4, we summarize the overall segmentation algorithm and provide implementation details.

2.1 A Probabilistic Segmentation Framework Based on Energy Minimization

In a typical active contour model, the segmentation process involves an iterative algorithm for minimization of an energy functional. We define our energy (cost) functional in a maximum *a posteriori* (MAP) estimation framework as

$$E(\mathbf{C}) = -\log P(\text{data}|\mathbf{C}) - \log P(\mathbf{C}), \quad (1)$$

where \mathbf{C} is a set of evolving contours $\{C^1, \dots, C^m\}$ that represent the boundaries of m different objects. In the following, we will refer to Ref. 12 as *C&V*. We choose the likelihood term $P(\text{data}|\mathbf{C})$ as in *C&V*. $P(\mathbf{C})$ is a coupled prior density of multiple objects. In this work, we focus on building $P(\mathbf{C})$.

The coupled prior is estimated using a training set of N shapes of the objects $\{\mathbf{C}_1, \dots, \mathbf{C}_N\}$. The essential idea of using such a prior is that the set of candidate segmenting contours \mathbf{C} will be more likely if they are similar to the example shapes in the training set. We define the joint prior $P(\mathbf{C})$ in terms of the shape and pose parameters of multiple objects

$$P(\mathbf{C}) = P(\tilde{\mathbf{C}}, \mathbf{p}) = P(\tilde{\mathbf{C}}) \cdot P(\mathbf{p}|\tilde{\mathbf{C}}). \quad (2)$$

Here, \mathbf{p} is a vector of pose parameters (p_1, \dots, p_m) for each object. Each p_i ($i = 1, 2, \dots, m$) consists of a set of translation, rotation, and scale parameters and $\tilde{\mathbf{C}}$ represents the aligned version of \mathbf{C} with respect to \mathbf{p} . In particular, we have $\tilde{\mathbf{C}} = T[\mathbf{p}]\mathbf{C}$, where $T[\cdot]$ denotes an alignment operation. In this context, the coupled shape density $P(\tilde{\mathbf{C}})$ represents only shape variability and does not include pose variability. On the other hand, $P(\mathbf{p}|\tilde{\mathbf{C}})$ captures the joint pose variability of the objects. We decompose the pose information into global and inter-shape (i.e. relative) pose variables:

$$\mathbf{p} = (p_{glb}, \mathbf{p}_{int}) = (p_{glb}, p_{int}^1, \dots, p_{int}^m), \quad (3)$$

where p_{glb} denotes the overall pose of the objects of interest and $\mathbf{p}_{int} = (p_{int}^1, \dots, p_{int}^m)$ represents inter-shape pose information among these objects.

Substituting Equation (3) into (2), we have

$$P(\tilde{\mathbf{C}}, p) = P(\tilde{\mathbf{C}}) \cdot P(p_{glb}, p_{int}^1, \dots, p_{int}^m|\tilde{\mathbf{C}}) \quad (4)$$

We model p_{glb} and $\mathbf{p}_{int} = (p_{int}^1, \dots, p_{int}^m)$ as independent variables, since global pose of objects and inter-shape pose are two different pieces of information that are not usually related:

$$P(\tilde{\mathbf{C}}, p) = P(\tilde{\mathbf{C}}) \cdot P(p_{glb}|\tilde{\mathbf{C}}) \cdot P(\mathbf{p}_{int}|\tilde{\mathbf{C}}) \quad (5)$$

Here, $P(p_{glb}|\tilde{\mathbf{C}})$ is assumed to be uniform since all poses p_{glb} are equally likely. * Then, we can express $P(\mathbf{C})$ as

$$P(\mathbf{C}) = P(\tilde{\mathbf{C}}) \cdot \gamma \cdot P(\mathbf{p}_{int}|\tilde{\mathbf{C}}), \quad (6)$$

where γ is a normalizing scalar. Substituting $P(\mathbf{C})$ into Equation (1), we obtain

$$E(\mathbf{C}) = -\log P(\text{data}|\mathbf{C}) - \log P(\tilde{\mathbf{C}}) - \log \gamma - \log P(\mathbf{p}_{int}|\tilde{\mathbf{C}}) \quad (7)$$

Given Equation (7), the focus of our work is to learn and specify the priors $P(\mathbf{p}_{int}|\tilde{\mathbf{C}})$ and $P(\tilde{\mathbf{C}})$.

For the sake of simplicity of exposition, and without loss of generality, our development of these priors in the following two subsections is based on two objects (i.e. $m = 2$). However, the framework we develop is general enough to be applied to an arbitrary number of objects. We mention that the full derivations are available in Ref. 32.

2.2 Coupled Shape Prior for Multiple Objects

In this section, we construct a coupled nonparametric shape prior density $P(\tilde{\mathbf{C}})$ for two different classes of objects. We choose level sets as the representation of shapes³³ and we use multivariate Parzen density estimation³⁴ to estimate the unknown joint shape distribution. Consider $m = 2$ and define the joint kernel density estimate of two shapes as,

$$P(\tilde{C}^1, \tilde{C}^2) = \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^{m=2} k(d(\phi_{\tilde{C}^j}, \phi_{\tilde{C}_i^j}), \sigma_j) \quad (8)$$

where N is the number of training shapes and $k(\cdot, \sigma_j)$ is a Gaussian kernel with standard deviation σ_j . In Equation (8), $\phi_{\tilde{C}^j}$ is the candidate signed distance function (SDF) of the j th object, which is aligned to the training set, and $\phi_{\tilde{C}_i^j}$ is the SDF of the i th training shape of the j th object. Note that, given a distance measure $d(\cdot, \cdot)$, we can construct the kernel for joint density estimation, by multiplying separate kernels $k(\cdot, \sigma_j)$ for each object. Our nonparametric shape prior, which is defined in Equation (8), can be used with a variety of distance metrics. Following Ref. 23, we employ the L_2 distance d_{L_2} between SDFs. In order to specify the kernel size σ_j of the j th object, we use maximum likelihood kernel size with leave-one-out method (see Ref. 35).

When referring to the shape kernel, we use the shorthand notation, k_i^j for

$$k_i^j = k(d_{L_2}(\phi_{\tilde{C}^j}, \phi_{\tilde{C}_i^j}), \sigma_j) = \frac{\exp\left(-\frac{1}{2\sigma_j^2} \int \left(\phi_{\tilde{C}^j}(x) - \phi_{\tilde{C}_i^j}(x)\right)^2 dx\right)}{\sqrt{2\pi\sigma_j^2}}. \quad (9)$$

Next, we define a gradient flow for the joint shape prior in Equation (8). We use one contour for each object, which is represented implicitly by its corresponding SDF. Then, we compute the gradient flow in the normal direction that increases most rapidly for each object contour. Using the L_2 distance in kernels, we find that the gradient directions for the contours \tilde{C}^j , are

$$\frac{\partial \phi_{\tilde{C}^j}}{\partial t} = \frac{1}{\sigma_j^2} \sum_{i=1}^N \lambda_i(\tilde{C}^1, \tilde{C}^2) (\phi_{\tilde{C}_i^j}(x, y) - \phi_{\tilde{C}^j}(x, y)) \quad (10)$$

where $j = 1, 2$, $\lambda_i(\tilde{C}^1, \tilde{C}^2) = \frac{k_i^1 k_i^2}{N \cdot P(\tilde{C}^1, \tilde{C}^2)}$, and $\sum_{i=1}^N \lambda_i(\tilde{C}^1, \tilde{C}^2) = 1$. Note that $\phi_{\tilde{C}^j}$ is a function of iteration time t and $\phi_{\tilde{C}_i^j}$ is a shorthand notation for the evolving level set function $\phi_{\tilde{C}_i^j}(t)$. Equation (10) defines the evolution of

*In some applications where certain global poses are more likely *a priori*, a non-uniform density could be used.

the contours toward shapes at the local maximum of the coupled shape prior of two objects. Note that training shapes that are closer to the evolving contour influence the evolution with higher weights. Note also that the structure of $\lambda_i(\tilde{C}^1, \tilde{C}^2)$ conveys the coupled nature of the evolution. In particular, the closer the active contour corresponding to one of the objects is to a training sample, the higher the weight of this training sample on the second active contour evolution is.

2.3 Moment-Based Relative Pose Prior for Multiple Objects

Aiming multi-object segmentation, we model relative pose (i.e. the pose after global alignment) of each object by a four dimensional vector $p_{int}^j = [A, c_x, c_y, \theta]$, where A is the area, c_x and c_y are the coordinates of the object, and θ is the relative orientation of the object to the global ensemble. We compute the pose of the individuals as related to their common mass center (after global alignment, see Equation (3)) via moments. Following Ref. 30, the two-dimensional moment, m , of order $p + q$, of a density distribution function, $f(x, y)$, is defined as $m_{p,q} = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} x^p y^q f(x, y) dx dy$. The two-dimensional moment for a $(N \times M)$ discretized image, $f(x, y)$, is $m_{p,q} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} x^p y^q f(x, y)$. We compute moments of objects that are defined by their boundaries. These boundaries define domains of integration (or summation in the discrete case), which we denote by Ω . We adjust the support of the input functions to an implicit representation in which $f(x, y) = 1$ if (x, y) is inside the object and 0 otherwise. With these choices, we compute the two-dimensional moment, m , of order $p + q$, using the formula $m_{p,q} = \int_{\Omega} x^p y^q dx dy$.

Let $M_0 = \{m_{0,0}\}$, $M_1 = \{m_{1,0}, m_{0,1}\}$, and $M_n = \{m_{i,j} | m_{i,j} \in \mathcal{M}, i + j = n\}$ for any $n \geq 0$. In addition, define the complete set of moments of order up to order two by $\mathcal{M}_2 = M_0 \cup M_1 \cup M_2$. Following Ref. 36, define the inertia moments as $I_{xx} = m_{0,2}$, $I_{xy} = I_{yx} = m_{1,1}$, and $I_{yy} = m_{2,0}$. Let θ be the angle between the eigenvectors of $\bar{I} = \begin{pmatrix} I_{xx} & -I_{xy} \\ -I_{xy} & I_{yy} \end{pmatrix}$ and the coordinate axes. Then, we have $\theta(C) = \frac{1}{2} \arctan \left(\frac{2(m_{1,0}m_{0,1} - m_{1,1}m_{0,0})}{(m_{0,2} - m_{2,0})m_{0,0} + m_{1,0}^2 - m_{0,1}^2} \right)$.

We construct p_{int}^j as the set of internal pose parameters, where C^j is a contour defining object j , i.e. $p_{int}^j = \left[m_{0,0}, \frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}}, \theta \right]$. Here, $m_{0,0}$ corresponds to area, $\frac{m_{1,0}}{m_{0,0}}$, $\frac{m_{0,1}}{m_{0,0}}$ correspond to horizontal and vertical positions relative to the center of mass, and θ corresponds to the canonic orientation of the object j relative to the orientation of the ensemble. Bearing in mind that we model shapes as zero level sets of the SDFs, these zero level sets define the integration domain of the standard moments. Following Section 2.2, we estimate

$$P(\mathbf{p}_{int} | \tilde{\mathbf{C}}) = \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^2 k \left(d \left(p_{int}^j, p_{int}^{j^i} \right), \sigma_j \right), \quad (11)$$

using Parzen kernel density estimation, where k is a Gaussian kernel. Here, $d(p_{int}^j, p_{int}^{j^i}) = (p_{int}^j - p_{int}^{j^i})^T \cdot Q \cdot (p_{int}^j - p_{int}^{j^i})$, where Q is a diagonal weighting matrix. Note that we employ the weighting coefficients in order to balance the influence of different pose parameters in the distance computation. In the following, we use the shorthand notation k_i^j for the moment based kernel $k \left(d \left(p_{int}^j, p_{int}^{j^i} \right), \sigma_j \right)^\dagger$.

The gradient flow of Equation (11) is $\frac{\partial \phi_{\tilde{\mathbf{C}}^j}}{\partial t} = \frac{1}{P(\mathbf{p}_{int} | \tilde{\mathbf{C}}) \cdot N} \sum_{i=1}^N \frac{k_i^1 k_i^2}{-\sigma_j^2} MPF(j, i)$, where

$$MPF(j, i) = (m_{0,0}^j - m_{0,0}^{j^i}) + \sum_{m_{r,s}^j \in \mathcal{M}_2, r+s=1} \left(\frac{m_{r,s}^j}{m_{0,0}^j} - \frac{m_{r,s}^{j^i}}{m_{0,0}^{j^i}} \right) \frac{(x^r y^s m_{0,0}^j - m_{r,s}^j)}{(m_{0,0}^j)^2} + (\tilde{\theta}^j - \tilde{\theta}^{j^i}) \sum_{r=0}^2 \sum_{s=0}^{2-r} x^r y^s M_{r,s}^{\theta j} \quad (12)$$

for each $j \in \{1, 2\}$. Here, $m_{r,s}^j$ denotes moments of the globally aligned evolving contour and $m_{r,s}^{j^i}$ means the moments of the i th aligned training image, whereas the rotation angles θ follow similar conventions. In Equation (12), the term $M_{r,s}^{\theta j}$ depends on θ . The complete definitions and details of $M_{r,s}^{\theta j}$ can be found in Ref. 32.

[†]Although we use the same kernel notation for shape and inter-shape pose prior modeling, we use it in different sections only, and the distinction is clear from context.

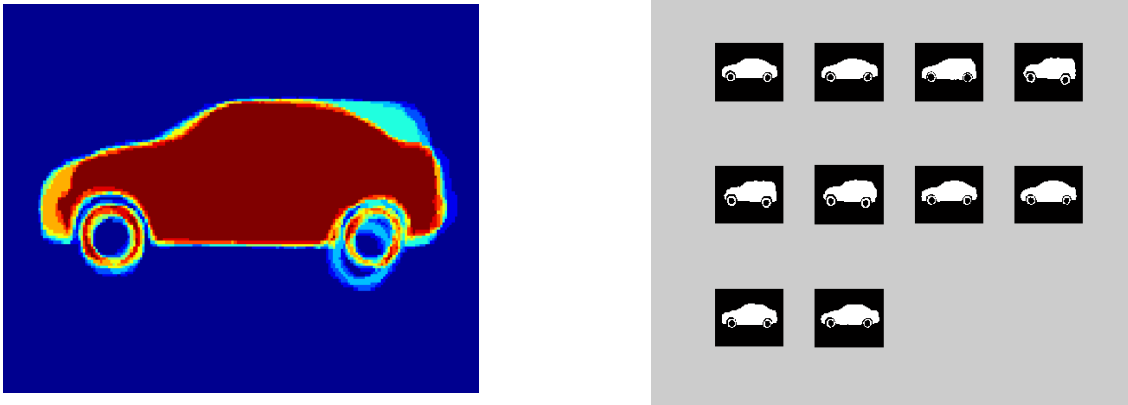


Figure 2. Aligned training data for the car segmentation experiment. The left image shows a set of ten binary car images super-imposed. The right image shows the binaries of the aligned images, which are super-imposed in the left image.

segmentation results in Figure 3. In Figure 3, the leftmost column is initialization, the second one shows results achieved using the C&V method, and the rightmost column shows the steady state result of our method.

We observe that the C&V method fails to produce reasonable segmentation results in these complicated scenes. This is due to two reasons. First, the intensity structure of the regions is much more complicated than the piecewise constant structure assumed by the C&V data term. Second, the C&V method uses only the curve length penalty to constrain the shapes, which is certainly not sufficient to drive the curves towards the boundaries of the objects of interest. Our approach provides improvements on this second issue through the shape and pose priors used in our framework. In particular, we note that our approach provides reasonable localization of the objects as well as fairly accurate estimation of a significant portion of the boundaries. This is a major improvement over C&V. However, our results are clearly not perfect. We believe the major reason is that the data term we use in segmentation (which is the C&V data term, and which is not the focus of this particular paper) is not very well matched to the object intensity patterns in these scenes (as noted in the first point above). For example, strong intensity contrasts caused by illumination variability apparent in the car surfaces (especially in the scene in the bottom row) have the effect of driving the curves towards illumination boundaries rather than the object boundaries. This issue can be resolved by using a more advanced data term built upon better statistical (e.g. learning-based) models of object intensity distributions. We do not pursue this extension here, as it is out of the scope of the current work. However, despite this particular issue related to the data term, our results still demonstrate the positive impact of our shape and pose priors on segmentation quality.

3.2 Experiments on Handwriting Data

Our database of handwriting images consists of two sets of ten images of the words “on” and “an”. The “an” set consists of two subsets of five images. The first subset consists of five capital letter words while the second one consists of small letters only. When segmenting one of the “on” words, we consider a training set that comprises all other nine images from its dataset. For “AN” and “an”, we present segmentation results on capitals or small letters disjointly, while using the capital or small letter subsets for training, respectively. When we select an image of a word for testing purposes, we build a training set of all other images, i.e. we use nine words for training “on” and four words for training “AN” or “an”. We show segmentation results in Figures 4 and 5. In all handwriting images, the leftmost column represents initializations, while the middle and the rightmost columns show the results of the C&V method and our proposed method, respectively, both in steady state.

In the first and second line in Figure 4, while the C&V method cannot distinguish the two letters, a result which is expected, our method provides distinct and accurate boundaries for the two letters involved. The third

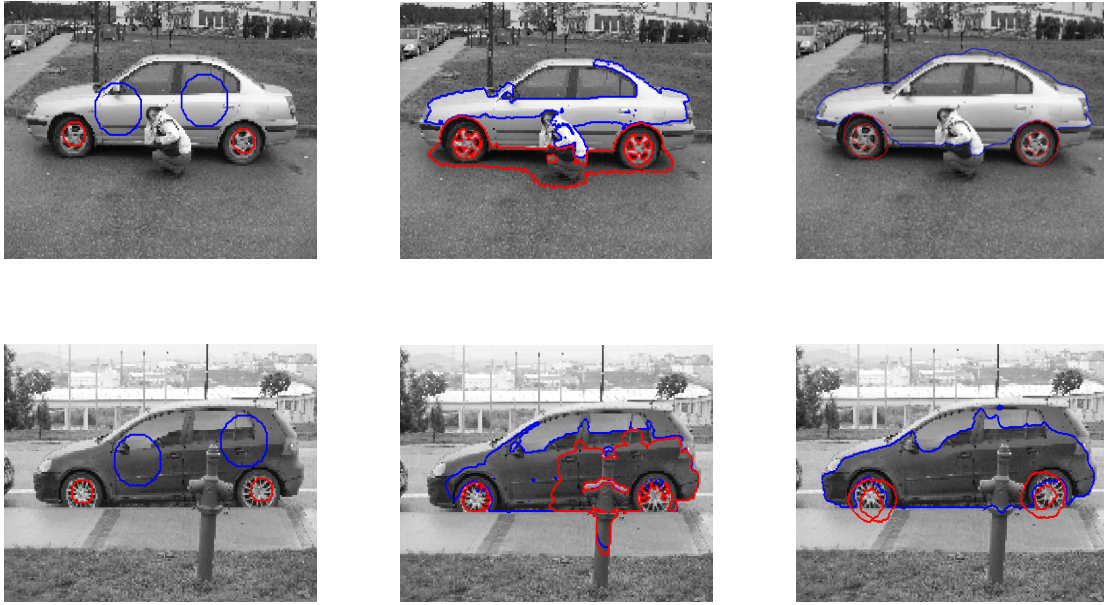


Figure 3. Segmentations of car images. The leftmost column illustrates initializations. The second column represents the result achieved by the C&V method. The rightmost column illustrates the final result achieved by our method.

and fourth lines represent experiments in the presence of occlusions. These experiments show the capability of our approach to recover the shapes of the letters despite occlusions. We believe the segmentation accuracy of our approach could be improved by using a richer set of training data than the one used in these preliminary experiments. In Figure 5, we present the results of a low-SNR scenario. Under these conditions, the C&V method does not provide reliable segmentations, while our approach can still recover the letter boundaries to a certain extent.

3.3 Experiments on Medical Data

In this section, we show and compare results of segmentation on real MR data. We present results on T2 and proton density (PD) MR images. Especially PD modality presents challenges due to its low contrast. We show the results using C&V and our method.

In particular, we present the segmentations of the complete head of Caudate Nucleus and Putamen. We demonstrate the results of this experiment in Figure 6. We show the input MR images, their ground truths, and the results obtained in steady state. We use a training set of twenty binary shapes that include the structures of interest. The C&V method results in inevitable leakages for both Caudate Nucleus and Putamen (see third column). However, the proposed coupled shape prior based approach segments both structures more effectively due to the coupling effect between shapes (see the rightmost column in Figure 6). The benefit of using a coupled prior is expected to be greater when the boundary of some objects is not well supported by the observed image intensity (see Putamen results in the second row, which is T2 MR modality).

4. CONCLUSION

In this paper, we have proposed a multi-object segmentation approach that employs coupled shape and inter-shape pose prior information. We employ an active contour framework towards evolving different contours in parallel. We employ training-based priors to estimate the coupled shape information as well as the inter-shape pose information among objects of interest. The priors are modeled using Parzen density estimation. We present segmentation results of cars in urban scenes, handwritten characters in words, and two Basal Ganglia

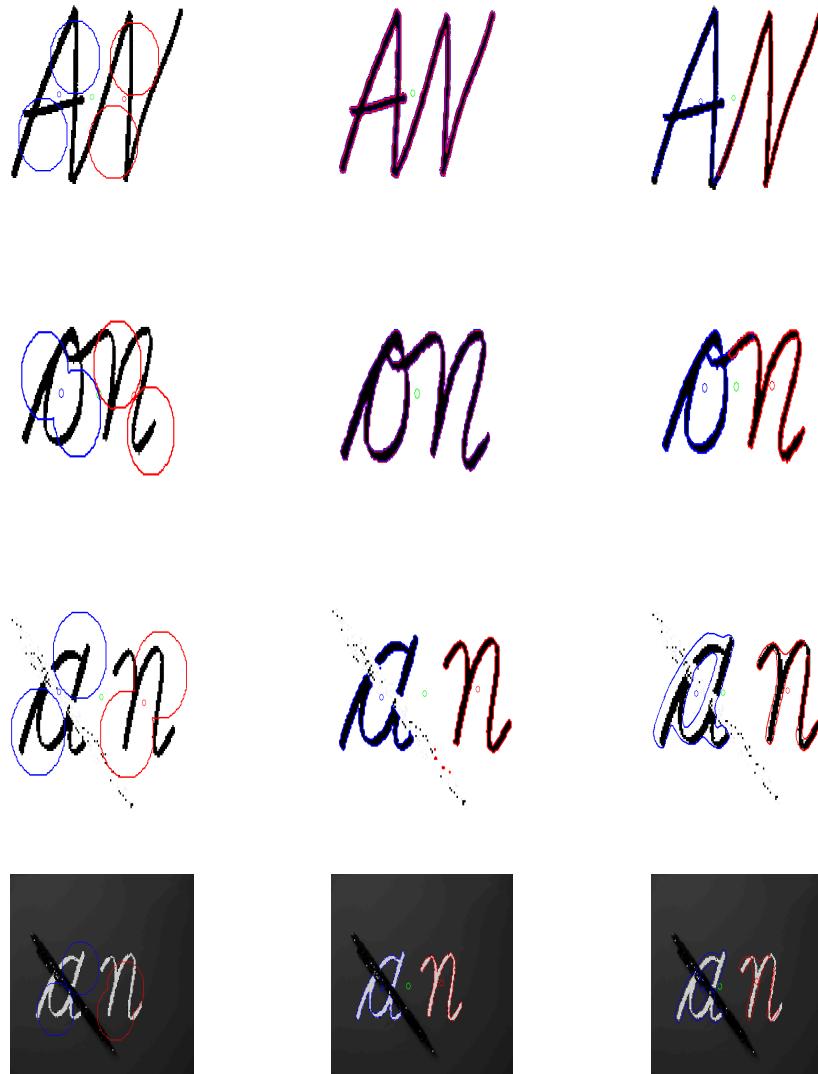


Figure 4. Handwriting segmentation experiment involving contiguous handwritten text and occlusions. The leftmost column shows the initializations. The middle column shows the results of the C&V method. The rightmost column shows the results of our method.

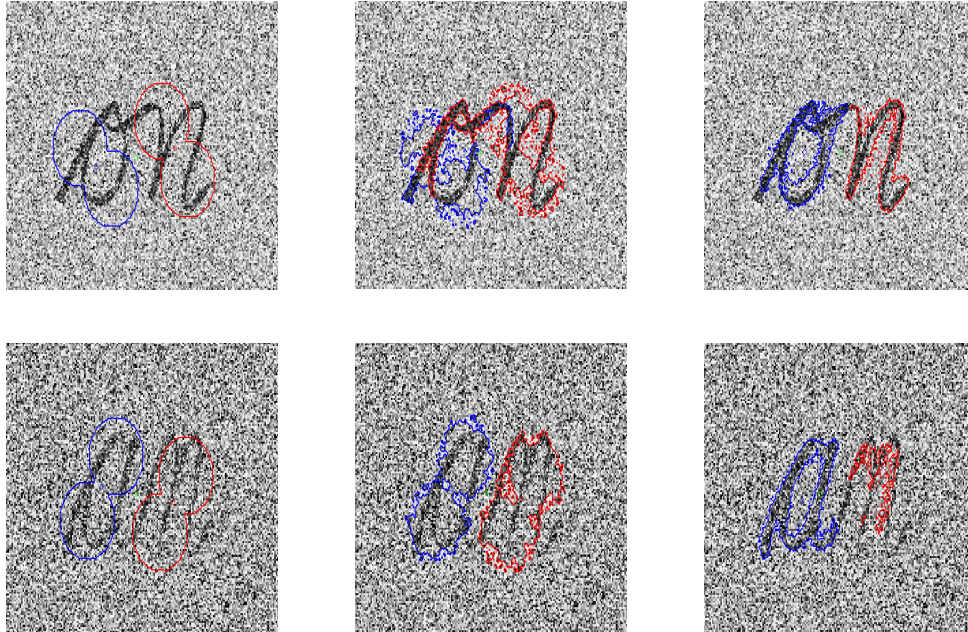


Figure 5. Handwriting segmentation experiment in the presence of severe noise. The leftmost column shows the initializations. The middle column shows the results of the C&V method. The rightmost column shows the results of our method.

structures, Caudate Nucleus and Putamen, captured in MR images. We have demonstrated our approach in several experiments, in which poorly contrasted difficult shapes are segmented. We have also experimentally shown the occlusion recovery capabilities of our approach.

ACKNOWLEDGMENTS

This work was partially supported by the European Commission under Grants MTKI-CT-2006-042717 (IRonDB), FP6-2004-ACC-SSA-2 (SPICE), MIRG-CT-2006-041919, and a graduate fellowship from The Scientific and Technological Research Council of Turkey (TUBITAK). The MR brain data sets were provided by the Radiology Center at Anadolu Medical Center and Yeditepe University Hospital. The second author thanks his wife, Diana Florentina Soldea, for helping in collecting images of cars in the Sabanci University campus.

REFERENCES

- [1] S. S. Jasit, S. Sameer, S. K. Setarehdan, S. Rakesh, B. Keir, C. Dorin, and R. Laura, “A note on future research in segmentation techniques applied to neurology, cardiology, mammography and pathology. Advanced algorithmic approaches to medical image segmentation: State-of-the-art application in cardiology, neurology, mammography and pathology,” *Springer-Verlag*, pp. 559–572, 2002.
- [2] K. van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, “A unifying framework for partial volume segmentation of brain MR images,” *IEEE Transactions On Medical Imaging* **22**, pp. 105–119, January 2003.
- [3] S. Dambreville, Y. Rathi, and A. Tannenbaum, “A framework for image segmentation using shape models and kernel space shape priors,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(8), pp. 1385–1399, 2008.
- [4] D. Cremers, J. S. Osher, and S. Soatto, “Kernel density estimation and intrinsic alignment for shape priors in level set segmentation,” *International Journal of Computer Vision* **69**(3), pp. 335–351, 2006.
- [5] D. Mumford and J. Shah, “Boundary detection by minimizing functionals,” *IEEE Conference on Computer Vision and Pattern Recognition*, 1985.

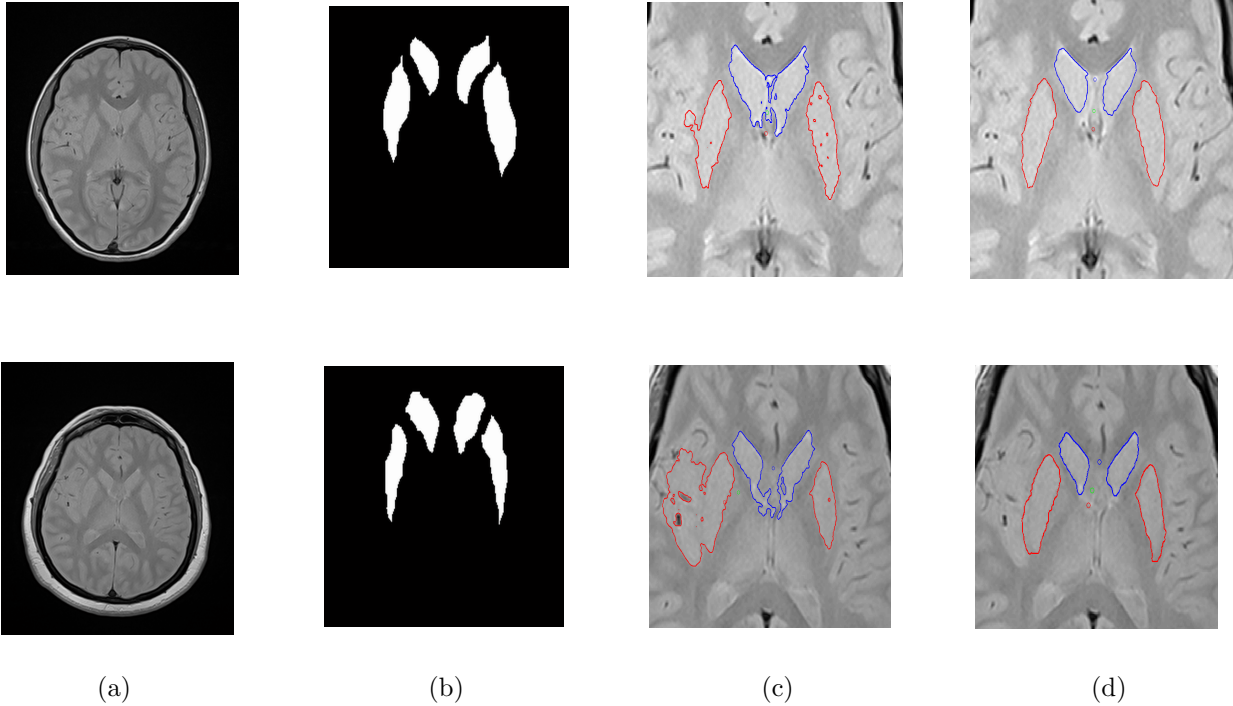


Figure 6. Segmentation results of the Caudate Nucleus and the Putamen in an MR slice: (a) MR images, (b) ground truths, (c) C&V method, (d) our method. The top and bottom rows represent PD and T2 MR modalities respectively. Columns (b), (c), and (d) represent images bounded by regions of interest.

- [6] J.-P. Wang, "Stochastic relaxation on partitions with connected components and its application to image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**(6), pp. 619–636, 1998.
- [7] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(9), pp. 884–900, 1996.
- [8] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(11), pp. 1222–1239, 2001.
- [9] M. Kass, A. Witkins, and D. Terzopoulos, "Snakes: Active contour models," *International Journal Computer Vision* **1**(4), pp. 321–331, 1988.
- [10] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(8), pp. 790–799, 1995.
- [11] V. Caselles, "Geometric models for active contours," in *IEEE International Conference on Image Processing*, **3**, pp. 9–12, IEEE Computer Society, (Washington, DC, USA), 1995.
- [12] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing* **2**, pp. 266–277, 2001.
- [13] A. Yezzi, Jr., A. Tsai, , and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," *IEEE International Conference on Computer Vision* **2**, pp. 898–903, 1999.
- [14] M. Rousson and N. Paragios, "Shape priors for level set representations," in *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part II*, pp. 78–92, Springer-Verlag, (London, UK), 2002.
- [15] Y. Chen, H. D. Tagare, S. Thiruvankadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision* **50**(3), pp. 315–328, 2002.
- [16] E. M. Leventon, L. E. W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," **1**, pp. 316–323, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000.

- [17] A. Tsai, A. Yezzi Jr., W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging* **22**(2), pp. 137–154, 2003.
- [18] Y. Chen, S. Thiruvankadam, F. Huang, K. S. Gopinath, and R. W. Brigg, "Simultaneous segmentation and registration for functional MR images," *International Conference on Pattern Recognition* **01**, pp. 747–750, 2002.
- [19] S. Jehan-Besson, M. Gastaud, M. Borland, and G. Aubert, "Region-based active contours using geometrical and statistical features for image segmentation," *IEEE International Conference in Image Processing* **2**, pp. 643–646, September 2003.
- [20] D. Cremers, C. Schnorr, and J. Weickert, "Diffusion-snakes: Combining statistical shape knowledge and image information in a variational framework," pp. 137–144, IEEE Workshop on Variational and Level Set Methods, (Washington, DC, USA), 2001.
- [21] F. Huang and J. Su, "Moment-based shape priors for geometric active contours," pp. 56–59, The 18th International Conference on Pattern Recognition, 2006.
- [22] K. Gorczowski, M. Styner, J. Y. Jeong, J. S. Marron, J. Piven, H. C. Hazlett, M. S. Pizer, and G. Gerig, "Discrimination analysis using multi-object statistics of shape and pose," *SPIE, Medical Imaging: Image Processing* **6512**, March 2007.
- [23] J. Kim, M. Çetin, and S. A. Willsky, "Nonparametric shape priors for active contour-based image segmentation," *Signal Processing* **87**, pp. 3021 – 3044, 2007.
- [24] A. Tsai, W. Wells, C. Tempany, W. E. Grimson, and A. Willsky, "Mutual information in coupled multi-shape model for medical image segmentation," *Medical Image Analysis* **8**, pp. 429–445, 2004.
- [25] J. Yang, L. Staib, and J. Duncan, "Neighbor-constrained segmentation with level set based 3-D deformable models," *IEEE Transactions on Medical Imaging* **23**(8), pp. 940–948, 2004.
- [26] J. Yang and J. Duncan, "Joint prior models of neighboring objects for 3-D image segmentation," **1**, pp. 314–319, IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, 2004.
- [27] C. Lu, S. M. Pizer, S. Joshi, and J.-Y. Jeong, "Statistical multi-object shape models," *International Journal of Computer Vision* **75**(3), pp. 387–404, 2007.
- [28] M. Rousson and C. Xu, "A general framework for image segmentation using ordered spatial dependency," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science* **4191**, pp. 848–855, 2006.
- [29] M. Styner, K. Gorczowski, T. Fletcher, J. Y. Jeong, S. M. Pizer, and G. Gerig, "Statistics of pose and shape in multi-object complexes using principal geodesic analysis," *International Workshop on Medical Imaging and Augmented Reality, Lecture Notes in Computer Science* **4091**, pp. 1–8, 2006.
- [30] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing* **54**(5), pp. 438–460, 1992.
- [31] P. Golland, L. E. W. Grimson, E. M. Shenton, and R. Kikinis, "Small sample size learning for shape analysis of anatomical structures," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science* **1935**, pp. 72–82, 2000.
- [32] M. G. Uzunbaş, "Segmentation of multiple brain structures using coupled nonparametric shape priors," *M. Sc. Thesis, Sabanci University*, 2008.
- [33] S. Osher and R. Fedkiw, "Level set methods and dynamic implicit surfaces," Springer, Berlin, 2003.
- [34] D. Erdogmus, R. Jenssen, N. Y. Rao, and C. J. Principe, "Gaussianization: An efficient multivariate density estimation technique for statistical signal processing," *The Journal of VLSI Signal Processing* **45**, pp. 67–83, 2006.
- [35] B. W. Silverman, "Density estimation for statistics and data analysis," Chapman Hall, London, 1986.
- [36] J. L. Meriam and L. G. Kraige, *Engineering Mechanics, Dynamics, Fifth Edition*, Wiley, 2001.
- [37] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "Labelme: a database and web-based tool for image annotation," *International Journal of Computer Vision* **77**, pp. 157–173, May 2008.