Discriminant snakes for 3D reconstruction in medical images*

X.M. Pardo
Dept. Electrónica e Computación
Universidade de Santiago de Compostela
15706 Santiago, Spain. pardo@dec.usc.es

P. Radeva
Computer Vision Centre
Universitat Autònoma de Barcelona
089193 Bellaterra, Spain. petia@cvc.uab.es

Abstract

In this work we propose a new statistic deformable model that we call discriminant snake for 3D reconstruction in volumetric images. Our discriminant snake generalises the classical snake attracted by edge points, it deforms due to a generalised contour representation. The snake selects and classifies image features by a parametric classifier and each snaxel deforms to minimise the dissimilarity between the learned and found image features inside the feature space. We apply our statistic snake to segment anatomical organs and the results are very encouraging.
Keywords: *statistic snake, principal component analysis, Fisher linear discriminant analysis, supervised learning, 3D reconstruction.*

1. Introduction

The 3D reconstruction of objects from volume images is incresingly used in medicine and industrial applications. Manual delineation of the region of interest on volume images is, at least, fatigous and very time consuming, which has motivated the search for optimal computational techniques. Among the wide variety of image segmentation techniques, deformable models are receiving an special attention, mainly in medical imagery, due to their ability to interpret sparse set of features (e.g. edge points) and link them to obtain object contours applying general assumptions about the contour shape [12, 9, 11].

The classical snake is an image feature technique that uses energy terms defined from gradient features of global interest to find the desired contour [7]. In most real applications such assumption is too strong. Different authors suggest to combine the gradient-based potential with valley and crest maps [16, 3]. However, on one hand, the best way of integrating different features remains to be an open problem and in the other hand, these features are not selective enough yet. This leads to heuristical combinations that enhance too many feature points that do not belong to the object of interest, meanwhile others go unnoticed.

Constraints on shape have been proposed to compensate this lack of selectivity. *Hand-crafted* parameterized templates, with few degrees of freedom have been used, for example, for modeling features of faces [16]. More general methods, as Fourier descriptors, have been used for representing shapes in medical images [14]. Alternative approaches based on modal analysis have also been proposed to constraint the model to deform only in ways implied by the training set of shapes [4, 3].

Although shape models convey important information, they are not the panacea; high accuracy techniques must make the most of grey level information too. In line with this idea, shape models and appearance models are combined in face recognition [8].

Our approach based on the statistics ot the image features (texture) offers an alternative to those approaches based on shape statistics. We propose to use a bank of Gaussian derivative filters of different scales as the generalization of edge, crest and valley detectors, whose goal is to increase the selectivity on the description of the target object. At this point, the description of target image features is still too general. It is neccessary to locally decide (learn) the best way of combinig derivative degrees and scales in the description of each contour part by means of a classification vector. A supervised learning in conjunction with discriminant analysis is carried out to characterize each contour patch. Then a classifier is assigned to each snake patch and the external energy is represented by the distance of each snake point to the cluster of its target contour in the corresponding feature space.

Our statistic snake approach is particularly of interest for segmentation and tracking of objects in temporal or spatial image sequences. The snake is able to learn the changes in contour features inside each image as well as along the image sequence in an adaptive way.

The paper is organized as follows: in section 2 we give

*This work is supported by CICYT and EU grants TIC98-1100 and 2FD97-0220, and Xunta de Galicia grant PGIDT99PXI20606B.

a formulation of the snake as an energy-minimization technique, in section 3 we introduce our statistic snake based on principal component and Fisher linear discriminant analysis. In section 4 we give the results of applying our snake model to the problem of segmentation in medical images and finish the article by conclusions.

2. Snake formulation

A snake is an elastic curve ($u(s)$) that evolves from its initial shape and position as a result of the combined action of external and internal forces [7]. The internal forces model the elasticity of the curve, whereas external forces push the snake towards features of the image [7]. The external energy is generally defined from a potential field P:

$$E_{ext}(u) = \int P(u(s))ds.$$

A typical potential for a snake attracted to image edge points is given by [7]:

$$P(u(s)) \propto -|G_\sigma * \nabla I(u(s))| \quad (1)$$

where $I(u)$ represents the intentisity value, and G_σ is a Gaussian smoothing function of size σ.

The potential must define a contour which minima correspond, as accurately as posible, to the image features of interest. The total energy of the snake is the sum of the external and internal energies. The smallest energy correspond to the desired contour. The minimization of the energy function is generally performed using variational principles and finite difference techniques [7, 1]. Practical computations demand discretization over time and space. In finite difference approximations, the curve $u(s)$ is sampled at certain points where computations are done. Methods like finite elements and B-splines produce curves of high degree of continuity and the features of each curve vertex are now evaluated in its corresponding curve patch.

In the classical implementation, the external energy defines global coarse features of interest which rest especificity. Next, we describe the formulation of the new generalized and locally defined external energy.

3. Supervised feature learning

The feature spaces must be capable of representing any image features of interest and each snaxel must be able to distinguish between its corresponding contour target and other structures (in particular, other parts of the contour or contours of near objects). To that aim, the features are extracted by applying a bank of filters to the image and then a learning process yields the relevant features to characterize each contour configurations and discriminate between them. The different steps involved in the learning process are described in the following sections.

3.1. Feature extraction and potential field

We use a bank of Gaussian derivative filters to characterize the objects of interest, that contains derivatives up to degree three (variance of higher-order filters tends to be highly correlated to the outputs of lower order filters [13]). Additionally, we consider different scales to a number sufficient of characterizing all possible configurations, and of allowing the snake to follow the traslation of the object of interest in image sequences.

In general, it is useful to determine the response of filters at arbitrary orientations. As the directional derivative operator is steerable, we use a set of basis filters $\{G^d(x, y, \sigma, \theta_k)\}_{k=1}^{d}$ for defining the derivative of Gaussian of degree d at arbitrary angle ϕ, $G^d(x, y, \sigma, \phi)$. Derivative of degree d can be obtained by the interpolation of $d+1$ equiangular orientations. Given the basis functions for each filter degree, filters at arbitrary orientations can be synthesized by means of interpolant functions [6].

We define $\mathcal{G}_{\mathcal{D}_\Sigma}$ as the filter bank with derivatives until $\mathcal{D}$ degree and N_Σ scales $\sigma \in \{2^0, 2^1, \ldots, 2^{N_\Sigma - 1}\}$. The dimension of the bank of filters is:

$$d_\mathcal{G} = dim(\mathcal{G}_{\mathcal{D}_\Sigma}) = N_\Sigma \sum_{d=0}^{\mathcal{D}} (d+1). \quad (2)$$

Note, for example, that if $\mathcal{D} = 1$ we have an edge detector.

Due to the high dimension of the feature space generated by the bank of filters is necessary to reduce the dimensionality by eliminating non discriminat filter responses and to weight the contribution of the remaining filter responses to the classifier, as described below.

The relative importance of derivative features depends on the task domain. It seems natural to perform a self-training for reduction of the space and weighting the features. We apply a technique commonly used for dimensionality reduction based on PCA [15].

Given the problem of segmentation, a set of N sample image feature vectors on object-contour and non object-contour $\{s_1, s_2, \ldots, s_N\}$ are chosen, taking values in an $d_\mathcal{G}$-dimensional space, where $d_\mathcal{G}$ is the dimension of the bank of filters applied to the original image. Each component of s_i corresponds to the response of a filter (image convolution by a Gaussian derivative). Looking for a certain contour k, each image feature j can be classified to one of two clases $\{C_k, \bar{C}_k\}$, representing the pixels belonging to the contour k and the complementary class (complementary contour parts and remainder scene). Our goal is to obtain an optimal linear transformation that maps the original $d_\mathcal{G}$-dimensional space into an 1-dimensional feature space where the classification of image features is applied by measuring the distance to the class center of the corresponding learned configuration.

3.2. External forces by statistic classifiers

To map the original $d_{\mathcal{G}}$-dimensional space into an 1-dimensional feature space we use Fisher linear discriminant functions [5]. Applying FLDA the new scalar feature r_j is defined by the following linear transformation:

$$\mathcal{F}_k : (p_{j1}, \ldots, p_{jd_{\mathcal{G}}}) \in \mathcal{P} \longrightarrow r_j = V_k^T \mathcal{P}_j \in \mathbf{R}$$

where $V_k \in \mathbf{R}^{d_{\mathcal{G}}}$ is a vector that projects the features of pixel j in $\mathbf{R}$ to give the similarity to the contour class k. In order to find an optimal projection into a reduced feature space where the distance between samples of class C_k and the remainder samples is maximized, we have to obtain the best discriminant function.

PCA is used to reduce the dimension of the feature spaces (one per patch) from $\mathbf{R}^{d_{\mathcal{G}}}$ to $\mathbf{R}^m$. After PCA, we only maintain the m eigenvectors with largest eigenvalues that retain the 95% of the sample variance.

The optimal projection V_{k_opt} is defined as the vector which maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the within-class scatter matrix of the projected samples [2, 5], which can be obtained by means of a FLDA. As a result, we obtain for each contour part C_k:

$$V_{k\_opt}^T = V_{k\_fld}^T V_{pca}^T$$

where:

$$V_{pca} = \arg\max_{V_k} |V_k^T S_T V_k|$$

$$V_{k\_fld} = \arg\max_{V_k} \frac{|V_k^T V_{pca}^T S_{bc} V_{pca} V_k|}{|V_k^T V_{pca}^T S_{wc} V_{pca} V_k|},$$

where S_{bc}, S_{wc} and S_T are the between class, within class and total scatter matrices, respectively.

Each patch k of the snake curve has its own classifier $\mathcal{V}_{C_k}$ that defines the image features it is looking for. The scalar product of the mean feature vector μ_{C_k} of the patch and the classifier vector gives the center of the class $\mathcal{C}_{C_k}$ of points in the feature space that corresponds to the contour patch of interest:

$$\mathcal{V}_{C_k} = V_{k\_opt}$$
$$\mathcal{C}_{C_k} = \mathcal{V}_{C_k}^T \mu_{C_k}.$$

We define the local external energy of the snake by measuring the similarity of the actual image features $\mathcal{P}_f$, in the current location of the snake, to the desired contour configuration in terms of the distance from the projection of the image feature vectors $\mathcal{P}_f$ to the class center:

$$D_{C_k} = \left(\mathcal{V}_{C_k}^T (\mathcal{P}_f - \mu_{C_k})\right)^2 = \left(\mathcal{V}_{C_k}^T \mathcal{P}_f - \mathcal{C}_{C_k}\right)^2.$$

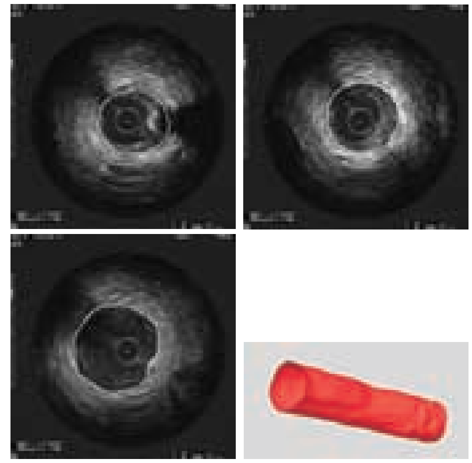

Figure 1. Samples of the segmentation of coronary vessels in a IVUS image sequence.

As a result each snaxel interprets in different way the image filter responses depending of the goal it carries (the contour type it has learned in its previous stage).

Given the problem of the 3D reconstruction, when the contour is delineated for a slice, its features are learned and the curve and the classifiers are moved to the next slice. In this way, the features of each class are updated and each patch searches for specific features discriminating between different objects and contour patches.

4. Results

In order to validate our approach we test our statistic snake on IVUS images. As an illustrative example, Fig. 1 shows the segmentation in intermediate slices of a sequence of 400 IVUS images of a coronary vessel. In these images, some contour parts of the blood vessel are very difficult of delineating even for a human operator. In that case we use a model shape information that avoids severe changes between contours of adjacent slices when there are not cues about the location of some contour parts.

To quantitatively evaluate the new approach, 5 experts $(e_1, e_2, e_3, e_4, e_5)$ manually segmented images from two different sequences of IVUS images. Model contours (M) were computed as the average hand-crafted contours and used as the ground truth segmentation. For each test image (j), we computed the distance between each pair of contours (e_i^j, M^j) and (S^j, M^j), where S^j represents the contour fitted by the snake approach in image j. The distance is characterized by its mean and variance $(\bar{d}, s^2)$ in pixels. The average distance given but the snake is similar to the ones given by manual segmentations (see Table 1). To objec-

Table 1. t-test between the segmentations provided by several experts and the snake.

e_1	e_2	e_3	e_4	e_5	S
$\bar{d}$, s^2	$\bar{d}$, s^2	$\bar{d}$, s^2	$\bar{d}$, s^2	$\bar{d}$, s^2	$\bar{d}$, s^2
IVUS 1					
1.77, 1.92	1.49, 1.28	1.51, 1.28	1.71, 1.43	1.43, 1.21	2.29, 2.59
IVUS 2					
1.36, 0.96	0.97, 0.56	1.22, 0.69	1.22, 0.69	1.14, 0.69	1.63, 1.17

IVUS 1					
	e_1/S	e_2/S	e_3/S	e_4/S	e_5/S
t	0.495	0.832	0.817	0.598	0.907
P-value	> 0.2	> 0.2	> 0.2	> 0.2	> 0.2
H_0	true	true	true	true	true

IVUS 2					
	e_1/S	e_2/S	e_3/S	e_4/S	e_5/S
t	0.576	1.594	0.956	0.948	1.335
P-value	> 0.2	> 0.1	> 0.2	> 0.2	0.2
H_0	true	true	true	true	true

tively compare the segmentations of our approach with the manual ones, we perform a statistical test. The hypothesis to be tested, the null hypotesis (H_0), is that there is non significative difference between the manual and the automatic segmentations. We use the Student's t test [10].

The average distance and variance from each contour point to the model contour were computed in each image sequence and compared to the results provided by the snake. We obtain the P-values from tabulated values [10] and decide on the hypotesis truthfulness using the usual significance level $\alpha = 0.05$. In all the cases, the null hypotesis was true, so there are not significant differences between expert and snake segmentations (see Table 1). The average distance from the final contours, given by the experts and our approach, to the ground truth segmentation and its variance are very similar. The same discrepances found between *ideal* and the automatic segmentation, are also found between the different manual segmentations.

5. Conclusions

The goal of this work is the generalization of the classical feature-based snake technique to reconstruct anatomical organs in images of different modality. Given the fact that object contour is rarely well defined by high gradient magnitude location we integrate gradient and texture image features chracterizing the contour obtained by a set of derivtive filters with different scales (multivalued potential). As a result, our model is a natural generalization of contour-, valley- and crest-based snakes.

The presented snake applies statistic methods to learn different contour configurations. The snake selects and self-trains the image features that are locally best representing the object contour. A parametric external energy of the snake is designed according to the fact that different patches of the snake are looking for different parts of the contour. As a result, the snake is more selective and robust and avoids ambiguities in case of two or more, even similar, objects close each other. We apply our snake to segment medical images obtaining better results than original snakes even in case of presence of near anatomical structures and changes in the morphology.

References

[1] A. A. Amini, T. E. Weymouth, and R. C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans. Patt. Anal. Machine. Intell.*, 12(9):855–867, 1990.

[2] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. Patt. Anal. Machine. Intell.*, 7:711–720, 1997.

[3] A. Blake and Isard. *Active contours*. Springer-Verlag, 1998.

[4] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models–their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[5] R. O. Duda and P. E. Hurt. *Pattern Clasification and Scene Analysis*. Wiley-Interscience, New York, 1973.

[6] W. Freeman and E. Adelson. The design and use of steerable filters. *IEEE Trans. Patt. Anal. Machine. Intell.*, 13(9):891–906, 1991.

[7] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.

[8] A. Lanitis, C. J. Taylor, and T. F. Cootes. Automatic face identification system using flexible appearance models. *Image and Vision Computing*, 13(5):393–401, 1995.

[9] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[10] W. Mendenhall and R. J. Beaver. *Introduction to probability and statistics*. PWS-KENT Publishing Company, Boston, 1991.

[11] J. Pardo, D. Cabello, and J. Heras. A snake for model-based segmentation of biomedical images. *Pattern Recognition Letters*, 18(14):1529–1538, 1997.

[12] P. Radeva, J. Serrat, and E. Marti. A snake for model–based segmentation. In *Proceedings of International Conference on Computer Vision (ICCV'95)*, MIT, USA, 1995.

[13] R. Rao and D. Ballard. Natural basis functions and topographic memory for face recognition. In *Proc. International Joint Conf. on Artificial Intelligence*, pages 10–17, 1995.

[14] L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14(11):1061–1075, 1992.

[15] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1), 1991.

[16] A. Yuille, P. Hallinan, and D. Cohen. Feature extraction from faces using deformable templates. *IJCV*, 8(2), 1992.