Contour Energy Features for Recognition of Biological Specimens in Population Images

Daniel Ochoa[1,2], Sidharta Gautama[1], and Boris Vintimilla[2]

[1] Department of telecommunication and information processing, Ghent University,St-Pieters Nieuwstraat 41, B-900, Ghent, Belgium
[2] Centro de Vision y Robotica, ESPOL University, Km 30.5 via perimetral, Guayaquil, Ecuador
{dochoa,sid}@telin.ugent.be, boris.vintimilla@espol.edu.ec

Abstract. In this paper we present an approach to perform automated analysis of nematodes in population images. Occlusion, shape variability and structural noise make reliable recognition of individuals a task difficult. Our approach relies on shape and geometrical statistical data obtained from samples of segmented lines. We study how shape similarity in the objects of interest, is encoded in active contour energy component values and exploit them to define shape features. Without having to build a specific model or making explicit assumptions on the interaction of overlapping objects, our results show that a considerable number of individual can be extracted even in highly cluttered regions when shape information is consistent with the patterns found in a given sample set.

Keywords: feature extraction, statistical shape analysis, segmentation, recognition.

1 Introduction

Many biological methods rely on molecular, biochemical and microbiological analysis of communities. One of the most studied families is the nematoda phylum given its well-described nervous system, complete genome sequence and sensitivity to environmental changes that makes it attractive for biotechnology research and development. Spatial and temporal distribution in nematode populations can be used as bio-indicators for soil management, variations between conventional and new genetically modified plants, expressions of disease symptoms in crops, pesticide treatments and lately for measuring the impact of the expected global warming.

To study populations and their evolution, care has to be taken to resort to non destructive methods to avoid killing the individuals and collecting a considerable amount of specimens from different samples and control sets. Once the specimens are under the microscope the technician collects data related to length, area, spatial distribution that then are correlated to the rate of growth, biomass, feeding behavior, maturity index and other time-related metrics that are used to support or discard hypothesis about the sample set under consideration.

M. Kamel and A. Campilho (Eds.): ICIAR 2007, LNCS 4633, pp. 1061–1070, 2007.

As high resolution camera systems has become affordable for research labs, the increasing amount of digital image data in biological studies will require efficient and robust image analysis tools to generate accurate and reproducible quantitative results. In contrast to medical images, where imaging condition and sampling methods are highly controlled, biological images are inherently difficult to analyze because of sample variation, noise and clutter [1]. These problems can distort the shape measurements of the detected specimens if for instance overlapping specimens are regarded as one.

In early papers, images containing single nematodes are examined [2]. After background correction, the image is thresholded and skeletonized, after which contour curvature patterns are used to identify the head and tail of the nematode. In a first step towards classifying C.Elegans behavioral phenotypes quantitatively, in [3] motion patterns are identified by means of a one-nematode tracking system, morphological operators and geometrical related features. Images of nematode populations were used in [4] to describe how to apply the scale space principles to linear object detection but no attempt is done to extract single specimens from the population.

In practice much of the work is still very labour intensive. Using digitalization software, the user marks points along the nematode body and linear segments are interpolated. The live-wire [5] approach can make the manual process easier since by following the contour a line is attracted to the nematode body but problems remains in cluttered regions where line information vanishes. In any case, the bulk of the recognition task is still done by hand for every single nematode. Certainly a discouraging scenario for researchers considering that a data set might consists of massive amounts of image data with possibly hundreds of specimens.

Consequently the need of high-throughput screening of bio-images to fully describe biological processes on a quantitative level is still very much in demand [6]. Given the nature of these images, extracting reliable shape information for object identification with a restricted amount of image data, clutter and structural noise is a challenging task. However we consider that effective recognition is a necessary step before any post-processing task, in particular if a computer vision based software tool is to be incorporated to derive statistical data from population samples [7] where accurate measurements are needed to provided truly meaningful information to bio-researchers.

Previous work on nematode population samples resort to intensity thresholding followed by filling, drawing and measuring operations in a semi-automatic fashion [8] and specimens are carefully placed apart from each other to prevent occlusion. Applying image processing techniques when dealing with several biological specimens that despite of belonging to the same class differ in shape and appearance makes parameter setting a complicated issue [9] and narrows the scope of possible applications.

Unlike previous efforts aimed at extracting shape clues from a set of single nematode images integrating them into a model and then finding the best possible fit on the image data [10]. This paper explore whether shape information can be captured by the utilization of a population of active contours. We believe that despite the inherent variability of nematode shapes in population images common patterns can play an important role in recognition not only in still images as those used in our experiments but in the video sequence when disambiguation methods in occluding scenarios are required.

This paper is organized as follows. Section 2 shape models and the active contour approach are discussed. Shape statistics of detected nematodes are proposed and used for classification in Section 3. Results are shown in Section 4. Finally conclusions and further work are presented in Section 5.

2 Active Contour Segmentation

In general, nematodes in an image can be thought as lines of varying width along their length, wide in the center and narrow near both ends. Although an important research body about linear structure detection particularly for vessel/neurite segmentation has been developed in the past [11], parameter setting involves a trade off between the image-content coverage and conciseness [12] a critical problem when dealing with populations because as specimens overlap line information vanishes at junctions and structural noise appears when nematodes' internal organs become visible.

Graph based search was proposed to integrate line evidence in [13] to detect networks of lines but recognition of individual objects require additional post processing steps, that given the lack of salient contour points make common shape representations less suitable for recognition. Capturing shape variation for instance by means of appearance/shape models [14] is a complex task in of worm-like objects given the absence of discriminant landmark points. Moreover, complex motion patterns prevent the use of linear systems to create a simple shape model. Although nonlinear systems has been devised [10] the complete range of nematode body configurations is still far from being model.

Spatial arrangement of feature points at different scales were exploited in [15] to search a rigid wiry object in cluttered environment. In a similar vein but on non-rigid objects in this paper we propose the utilization of active contours energies to capture relevant statistical shape information for recognition applied to nematode detection in population images.

Active contours introduced by Kass with a model called Snake [16] had drawn attention due to their performance in various problems. Segmentation and shape modeling in single images proved effective by integrating region-based information, stochastic approaches and appropriate shape constrains [17,18]. Active contours merge image data and shape modeling through the definition of an linear energy function consisting of two terms: a data driven component (external energy), which depends on the image data, and a smoothness-driven component (internal energy) which enforces smoothness along the contour.

$$E_{snake} = \alpha \cdot E_{\text{int}} + \beta \cdot E_{ext} \qquad (1)$$

In parametric contours, the internal energy can be decomposed further in tension and bending energies, they report higher values as the contour stretches or bends during the evolution process. The goal is to minimize the total energy iteratively using gradient descent techniques while energy components balance each other.

$$E_{\text{int}} = \int_0^P e_t(p) + e_b(p)dp \ , \quad E_{ext} = \int_0^P e_{ext}(p)dp \tag{2}$$

The rationale behind the proposed approach is that given convergence of the active contour being mostly data-driven, appearance and geometrical information can be recover from their energy component values. We consider the analysis of energy-based derived features a natural way to explore the range of possible shape configurations in nematodes population images without: a) having to build an specific model and b) making explicit constrains about the interaction of occluding objects [19].

In our experiments we chose the ziplock snake model [20]. This model is designed to work with open contours and initialization is limited to locating the contours' end points. Optimization is carried out from the end points towards the center of the contour so that the initial control points will be located progressively on the object surface and increases the probalilities of a correct segmentation. Being parametric it can encode shape information [21] explicitly and provide faster convergence than geodesic snakes.

$$e_{ext} = I(x, y), e_t = \sqrt{\dot{x}^2 + \dot{y}^2}, e_b = \left|\frac{(\dot{x} \cdot \ddot{y} - \ddot{x} \cdot \dot{y})}{(\dot{x}^2 + \dot{y}^2)^{3/2}}\right| \tag{3}$$

The tension energy applied corresponds to the point distance distribution, the bending energy to the local curvature and a normalized version of the intensity image I was used as energy field. The number of control points n, should be selected big enough to capture the global variability of the nematode's shape without losing the discriminant power to recognize them individually.

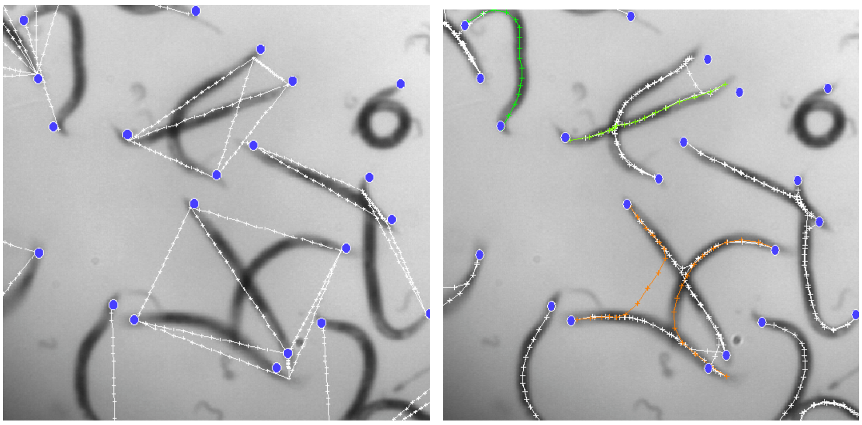

Fig. 1. Segmented contours in a nematode population image. Left: Initial end points and contours. Right: After convergence some contours are located on a single nematode others lie partially on background and sections of different nematodes.

Given a set of potential end points, either manually or automatically detected [3] contours without any particular initialization can be proposed and optimized until convergence. As the energy components interact with each other an number of contours will find the minimun energy on nematodes whereas others will be trapped in clutter regions. Unless explicit and detailed analysis of every or at least most common cases is done, a particularly challenging task in non-rigid objects, there is no suitable way to distinguish between nematodes and clutter contours Fig.1. Hence the suggested solution in presented in the following section.

3 Shape Classification Using Contour Energy Features

To select those contours corresponding to nematodes, we exploit the shape data available through the energy terms of the active contours. We define at global and local levels energy features distributions for both clutter *Cl* and nematode *Nl* contours classes. Let *S* be the set of contours, our *general measure Gi* is defined as the expected value of the observed energy component *e* for a given contour *s* in *S*.

$$G_i^s = E\left[e_i\right]_{s\in S}, i \in \{e_t, e_b, e_{ext}\} \tag{4}$$

A more *local measure Li* is defined as a n-dimensional vector formed by the local value of the energy component estimated for every control point in given a contour *s* in *S*:

$$L_i^s = (e_i^1.,..,e_i^n)_{s\in S}, i \in \{e_t, e_b, e_{ext}\} \tag{5}$$

From data distributions obtained experimentally we found no clear feature clusters and seems difficult that any isolated feature would be able separate nematode from clutter contours. However some patterns in the image become evident, in particular for local feature distributions. Regarding the shape of the nematodes seems that as the nematode gets thicker -in the central region- becomes less flexible. Looking at the external energy features is also apparent that distributions gradually displaces toward lower mean energy values since the nematode tend to be is darker in the middle than on both ends.

To combine features in a statistical framework we applied Bayes rule to classify contours in nematode and clutter classes. The ratio of the a posteriori probabilities of nematode to clutter class detection was define as discriminant function.

$$D = \frac{P(Nl \mid s)}{P(Cl \mid s)} = \frac{P(s \mid Nl)\cdot P(Nl)}{P(s \mid Cl)\cdot P(Cl)} \tag{6}$$

Assuming independence[1] between energy components, conditional distributions can be readily defined for $w \in \{Nl, Cl\}$ class using *Gi* and *Li* based feature sets. Regarding object features the prior was considered homogeneous and consequently discarded, for local features prior probabilities were defined by taking into account the distance d_c from a control point *c* to the closest end point.

[1] This is a crude assumption because one could expect some correlation between energy terms.

$$P(s \mid \mathrm{w}) = \prod_i P(\mathrm{G_i} \mid \mathrm{w}), i \subseteq \{e_t, e_b, e_{ext}\} \tag{7}$$

$$P(s \mid \mathrm{w}) = \prod_c \prod_i P(\mathrm{L_i} \mid \mathrm{c, w}), i \subseteq \{e_t, e_b, e_{ext}\} \tag{8}$$

$$P(\mathrm{w}) = \prod_c P(\mathrm{d_c}) \tag{9}$$

In order to measure the discriminant power of the derived feature sets different combinations of energy feature were tested. For local features classification the experiments were repeated without prior probabilities to estimate the impact of prior information. Also classification with different number of control points was performed by discarding an increasing number of segments on both ends of the contour.

4 Experimental Results

The developed approach was tested on set consisting of 8 high resolution time-lapse images depicting populations of adult nematodes with approximately 200 specimens laying freely on agar substract. Straight contours were placed between every pair of end points manually segmented from ground truth images within a distance equal to the maximum expected size of a nematode and iteratively optimized until convergence Fig. 1. Conditional distributions were derived from a training set of 30 randomly selected nematode and clutter contours. Given the non-gaussian nature of our data, conditional and prior distributions were fitted using gama and weibull pdfs respectively.

Classification results are summarized in a table 1. The general features *Gi*, seem to have the lower discrimination power and in general true positive *TP* rate tend to increases as more energy feature types are added. But surprisingly results improves as the number of points is reduced, this is indicative that nematode and clutter contours have similar average energy distributions and only when the central part of the contour considered the difference is large enough to allow better classification mainly because the central part of clutter contours they tend to fall on the background.

When considering local features *Li*, the tension energy proved being the most discriminant. Spatial distribution of control points can be explained by looking in the interaction between energies. Since the external energy is lower in the middle of the nematode control points tend to gather in that area, but as they move toward the center tension energy increases near both contour ends and pulls control points in the opposite direction. Therefore, the distance between the control points varies depending on the region a control point is located. Usually this region correspond to morphological structures inside the nematode.

It must be noted that only by combining several energy types the false positive *FP* rate can be consistently reduced. In particular, the bending energy allow us to filter out contours with sharp turns and the external energy. We also found that contours with spatial intensity distribution too different from those common to nematode or whose control points are located mostly on the image background can be rejected by considering in the external energy distribution. When the experiments were repeated without prior information becomes manifest the need of more control points to improve the accuracy of the results.

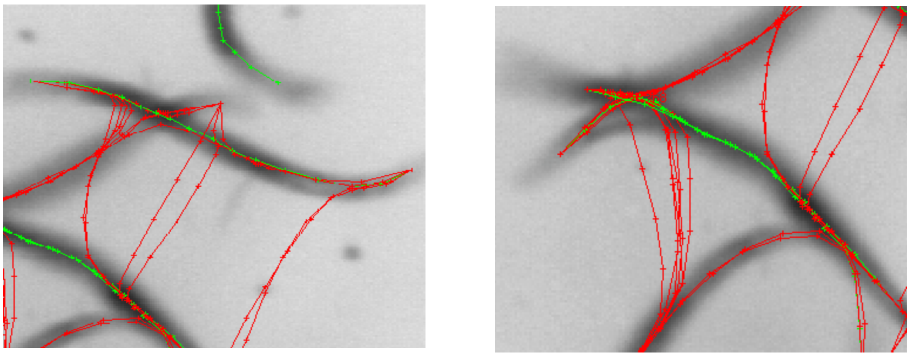

Fig. 2. Contour classification examples: green contours classified as nematode and red as clutter. Left: Correct nematode contour recognition in presence of occlusion. Right: misclassified clutter contour that runs over a true nematode.

Table 1. Classification results for *Gi* and *Li* feature combinations, the shadowed cells correspond to the best *Tp,Fp* pair

		Gi		Li		Li(no prior)	
n	Energy set	Tp	Fp	Tp	Fp	Tp	Fp
8	{Et}	0.804	0.272	0.918	0.360	0.913	0.335
8	{Eb}	0.707	0.348	0.918	0.545	0.908	0.520
8	{Eext}	0.832	0.457	0.848	0.552	0.826	0.515
8	{Et,Eb}	0.728	0.217	0.924	0.330	0.935	0.332
8	{Et,Eext}	0.804	0.245	0.913	0.341	0.897	0.306
8	{Eb,Eext}	0.755	0.272	0.918	0.456	0.886	0.405
8	{Et,Eb,Eext}	0.761	0.201	0.908	0.306	0.924	0.284
12	{Et}	0.266	0.082	0.908	0.317	0.918	0.316
12	{Eb}	0.364	0.223	0.870	0.477	0.897	0.510
12	{Eext}	0.641	0.364	0.793	0.466	0.832	0.515
12	{Et,Eb}	0.223	0.082	0.902	0.284	0.940	0.307
12	{Et,Eext}	0.288	0.065	0.891	0.291	0.891	0.295
12	{Eb,Eext}	0.348	0.103	0.853	0.364	0.864	0.392
12	{Et,Eb,Eext}	0.239	0.054	0.918	0.255	0.924	0.266
16	{Et}	0.114	0.038	0.880	0.335	0.924	0.296
16	{Eb}	0.255	0.141	0.826	0.477	0.880	0.488
16	{Eext}	0.522	0.245	0.788	0.456	0.821	0.473
16	{Et,Eb}	0.082	0.049	0.897	0.305	0.918	0.280
16	{Et,Eext}	0.087	0.038	0.875	0.309	0.875	0.274
16	{Eb,Eext}	0.190	0.054	0.848	0.360	0.875	0.366
16	{Et,Eb,Eext}	0.065	0.022	0.913	0.273	0.902	0.248

Misclassified nematode contours fall in two broad categories: contour close to the image border where the contrast between foreground and background is poor and appearance information is lost and those located in the perimeter of the sample affected by optical distortion that produce unusual shape configurations different from common patterns in the image set. Clutter contours can be mistakenly classified as nematode when most of their control points converge towards a real nematode, for instance when two parallel nematodes are close to each other, also when in the presence of heavy overlapping a clutter contour manages to cover parts of different nematodes.

Finally it should be noted that in occlusion cases, the more occluded is a nematode the less its discriminant function value. Structural noise introduced at junction regions due to the change of relative optical density affects negatively the convergence of contours and the feature values. Still we can recover a number of nematode from cluttered regions when enough shape information is retained. According to our experience in a local set overlapping of contours – for instance derived from common end point - nematode have discriminant function value consistently higher than those corresponding to clutter contours.

5 Conclusions

In this paper we had studied a feature sets aimed at improving recognition of individual specimens in populations images, where structural noise, intensity variations, different shape configurations and occlusion are present. The inherent similarity of nematodes in populations provides valuable information for recognition that is capture during the evolution of well known parametric contour models, allowing us to use the shape and image data encapsulated in the energy terms directly for classification. By this approach we avoid building a complex shape model and take advantage of image and shape statistics to narrow the range of possible appearance and geometrical configurations to those commonly present in a given sample set.

Our features based on energy component distributions were tested on manually segmented images in the framework of bayesian inference. Experimental results show how nematode and clutter contours can have similar statistics and only when less articulated nematode sections are considered rejection of significant number of clutter contours is obtain while retaining most of the nematode contours. Local features do not require prior knowledge of the location of stable object sections as the spatial distribution of contour points and its associated energy components effectively encode shape information.

Known problems of contour based segmentation although still present do not prevent correct nematodes recognition when they differ from background or lie in relative isolation. Recognition in cluttered regions is as expected more difficult but as long as meaningful detected sections of the nematode body are linked by a contour a positive classification is possible. We must point out that in practice only a representative sample of the population is needed to derive statistical data, our aim is to explore how to build such sample in a reliable way by exploiting image and shape regularities.

Future work will focus on extending our findings to video sequence in tracking of moving nematodes specifically when two or more nematode temporarily overlap and nematode identification is uncertain.

Acknowledgments. This work was supported by the VLIR-ESPOL program under the component 8, the images were kindly provided by Devgen corporation.

References

1. Bengtsson, E., Bigun, J., Gustavsson, T.: Computerized Cell Image Analysis: Past, Present and Future. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, Springer, Heidelberg (2003)
2. Fdez-Valdivia, J., Perez De la Blanca, N., Castilllo, P., Gomez-Barcina, A.: Detecting Nematode Features from Digital Images. Journal of Nematology 24, 289–298 (1992)
3. Wei, G., Cosman, P., Berry, C., Zhaoyang, F., Schafer, W.R.: Automatic tracking, feature extraction and classification of C. elegans phenotypes. IEEE transactions in Biomedical Engineering 51, 1811–1820 (2004)
4. Van Osta, P., Geusebroek, J.M., Ver Donck, K., Bols, L., Geysen, J., ter Haar Romeny, B.M.: The Principles of Scale Space Applied to Structure and Colour in Light Microscopy. Proceedings Royal Microscopical Society 37, 161–166 (2002)
5. Meijering, E., Jacob, M., Sarria, J.-C.F., Unser, M.: A Novel Approach to Neurite Tracing in Fluorescence Microscopy Images. Signal and Image Processing. 399, 96–148 (2003)
6. Meijering, E., Smal, I., Danuser, G.: Tracking in Molecular Bioimaging. IEEE Signal Processing Mag. 3, 46–53 (2006)
7. Moller, S., Kristensen, C., Poulsen, L., Cartersen, J., Molin, M.: Bacterial Growth on Surfaces: Automated Image Analysis fo Quantification of Rate-Related Parameters. Applied and Environmental Microbiology 6(1), 741–748 (1995)
8. Baguley, J., Hyde, L., Montagna, P.: A Semi-automated Digital Microphotographic Approach to Measure Meiofaunal Biomass. Limnology and Oceanography Methods 2, 181–190 (2004)
9. Tomankova, K., Jerabkova, P., Zmeskal, O., Vesela, M., Haderka, J.: Use of Image Analysis to Study Growth and Division of Yeast Cells. Journal of Imaging Sicience and Technology 6, 583–589 (2006)
10. Twining, C.J., Taylorn, C.J.: Kernel Principal Component Analysis and the Construction of Non-Linear Active Shape Models. In: British Machine Vision Conference, pp. 26–32 (2001)
11. Kirbas, C., Quek, F.: Vessel Extraction Techniques and Algorithms: A Survey. In: Proceedings 3th IEEE Symposium on BioInformatics and BioEngineering, pp. 238–246. IEEE Computer Society Press, Los Alamitos (2003)
12. Aylward, S.R., Bullitt, E.: Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. IEEE Transactions in Medical Imaging 21, 61–75 (2002)
13. Geusebroek, J., Smeulders, A., Geerts, H.: A minimum cost approach for segmenting networks of lines. International Journal of Computer Vision 43, 99–111 (2001)
14. Hicks, Y., Marshall, D., Martin, R.R., Rosin, P.L., Bayer, M.M., Mann, D.G.: Automatic landmarking for biological shape model. Proceedings IEEE International Conference on Image Processing 2, 801–804 (2002)

15. Carmichael, O., Hebert, M.: Shape-based recognition of wiry objects. Pattern Analysis and Machine Intelligence 26, 1537–1552 (2004)
16. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 4, 191–200 (1997)
17. Foulonneau, A., Charbonnier, P., Heitz, F.: Geometric shape priors for region-based active contours. In: Proceedings IEEE International Conference on Image Processing, vol. 3, pp. 413–416. IEEE Computer Society Press, Los Alamitos (2003)
18. Tsechpenakis, G., Rapantzikos, K., Tsapatsoulis, N., Kollias, S.: A snake model for object tracking in natural sequences. Signal Processing Image Communitations 19, 219–238 (2004)
19. Zimmer, C., Olivo-Marin, J.-C.: Coupled parametric active contours. IEEE Transactions on Pattern Analysis and Machine Intelligence 27, 1838–1842 (2005)
20. Neuenschwander, W.M., Fua, P., Iverson, L., Székely, G., Kubler, O.: Ziplock snakes. International Journal of Computer Vision 23, 191–200 (1997)
21. Jiankang, W., Xiaobo, L.: Guiding ziplock snakes with a priori information. IEEE Transactions on Image Processing 12, 176–185 (2003)