Contour-HOG: A Stub Feature based Level Set Method for Learning Object Contour

Zhi Yang
zhiyang@buffalo.edu

Yu Kong
yukong7@buffalo.edu

Yun Fu
raymondyunfu@gmail.com

Department of ECE and College of CIS
Northeastern University
Boston, MA, USA
Department of CSE
State University of New York
Buffalo, NY, USA

An object can be effectively characterized by its contour. Caselles *et al.* [1] introduced the concept of geodesic active contours, which applies the energy reducing form to acquire contours. Shape priors are great helpful to obtaining more accurate contours. Leventon [6] utilized the curvature prior as the shape prior for different classes of objects to guide contour evolution. Etyngier *et al.* [3] proposed a non-linear manifold learning method for learning shape prior. Another line of work uses edges to describe objects which provide local perspective of an object and are robust when part of the object is occluded. However, since the global perspective is missing, the arrangement of the edge features, such as the pairwise interactions between edge features [2, 5] or the relative positions of edge features with respect to the centroid of the shape [7], is exploited to improve the edge based models.

We propose a novel edge-based method for learning objects. Given an image, our method first detects edgelet feature as a rough contour for an object. Edgelet feature indicates potential positions for the contour and may stop curve evolution. These positions are referred to as *stub features*. Object contour is adaptively refined by the level set method. The evaluation criteria for contour evolution is defined by the similarity between the evolving contour and the target contour computed by their HOG features. Therefore the curve evolution method is referred to as the *Contour-HOG* method. We formulate the joint distribution of the edgelet feature, the HOG feature and the curvature of the evolved contour in a probabilistic model, and perform classification by computing the posterior of the evolved contour conditioned on the three types of features. Compared with previous methods, our method uses stub features to roughly localize a target object. This allows us to accurately capture the contour of the object. Moreover, the method fuses both local and global features to better describe the contour and thus improves the recognition accuracy.

Our method begins by detecting edgelet feature [8]. We use this feature to roughly find an object in an image. With the detected stub feature, we compute their similarities with the stub feature in training data under a predefined edge mask $M_{k,s}$. The similarity is computed as

$$p(x,s,k) = p_o(x,s,k)p_g(x,s,k), \tag{1}$$

where x is the 2D coordinate for a stub feature, $p_g(x,s,k)$ represents the likelihood of a local gradient and a mask having the similar magnitude, and $p_o(x,s,k)$ is the distribution of a gradient sharing the same orientation with the mask. We define $p_g(x,s,k)$ as a Gaussian distribution:

$$p_g(x,s,k) = \frac{1}{\sqrt{2\pi}|\Sigma_g|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}[g(x)-\mu]^T \Sigma_g^{-1}[g(x)-\mu]\right). \tag{2}$$

Here, μ and Σ_g are the mean and standard deviation of gradient magnitude, respectively. $g(x)$ is the magnitude of gradient vector derived by operating a mask $M_{k,s}$ on an input image I. The similarity of two orientation vectors is computed by

$$p_o(x,s,k) = \frac{\|v(I_q(M_{k,s}(x)), I_t(M_{k,s}))\|_2^2}{NT^2}, \tag{3}$$

where $v(o_1,o_2)$ is the distance between two quantized directions, N denotes the number of pixels selected by the mask, T represents the number of orientations, $I_q(M_{k,s}(x))$ and $I_t(M_{k,s})$ are the orientation vectors for a testing image and a training image, respectively.

After stub feature detection, we run a contour evolution method to obtain the contour of an object. We adopt the Elliptic Fourier descriptor (EFD) [4] to represent the object contour.

Our evolution model utilizes curvature force to guide contour evolution. The curvature force is defined as the ratio between the curvature of the shape prior κ_p and the curvature of the evolving contour κ_c: $f_\kappa = \frac{\kappa_p}{\kappa_c}$.

To obtain the global perspective of the evolving contour, we compute the HOG feature of an evolving contour and measure its similarity with the contours of classes. The using of the global similarity measure of contours allows us to accurately obtain an object contour.

In our work, an object is classified by the similarity between the evolved contour of the object and the contour of a target object in the training dataset. The similarity is evaluated based on their curvatures, stub features and HOG features. We compute the similarity as

$$p(\kappa_c, \kappa_t, S_{\text{hog}}, S_{\text{stub}}) = p(\kappa_c|\kappa_t, S_{\text{hog}}, S_{\text{stub}})p(\kappa_t, S_{\text{hog}}, S_{\text{stub}}), \tag{4}$$

where κ_c and κ_t denote the curvature of an evolved contour and a target contour, respectively. S_{hog} is the affinity of contour-HOG features between an evolved contour and a target contour, and S_{stub} is the affinity of stub features between them. In our work, we assume κ_t and S_{hog} are independent of S_{stub}. Then Eq.(4) can be given by

$$p(\kappa_c, \kappa_t, S_{\text{hog}}, S_{\text{stub}}) = p(\kappa_c|\kappa_t, S_{\text{hog}}, S_{\text{stub}})p(\kappa_t, S_{\text{hog}})p(S_{\text{stub}}). \tag{5}$$

Here, we define $p(\kappa_c|\kappa_t, S_{\text{hog}}, S_{\text{stub}})$ as a Gaussian distribution:

$$p(\kappa_c|\kappa_t, S_{\text{hog}}, S_{\text{stub}}) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{||\kappa_c - \kappa_t||_2^2}{2\sigma^2}), \tag{6}$$

where σ is the standard deviation. In Eq.(5), distribution $p(\kappa_t, S_{\text{hog}})$ is the similarity of HOG features between an evolved contour and a contour of a class. $p(S_{\text{stub}})$ is the similarity of an edge segment and a predefined edge mask computed in Eq.(1).

We use a likelihood function $\Lambda(Y;\theta)$ to measure the likelihood of a particular model with N training samples. We define the joint likelihood function as a Gaussian mixture model (GMM):

$$\Lambda(Y;\theta) = \prod_{n=1}^{N} \sum_{k=1}^{K} w_k G(y;\mu_k,\delta_k). \tag{7}$$

Our model is learned by maximizing likelihood estimation (MLE):

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\{\Lambda(Y;\theta)\}. \tag{8}$$

[1] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1): 1–31, 1993.

[2] S. M. Ali Eslami, Nicolas Heess, and John Winn. The shape boltzmann machine: a strong model of object shape. In *CVPR*, 2012.

[3] Patrick Etyngier, Florent Ségonne, and Renaud Keriven. Active-contour-based image segmentation using machine learning techniques. In *MICCAI*, pages 891–899, 2007.

[4] F. P. Kuhl and C. R. Giardina. Elliptic fourier features of a closed contour. *Computer Graphics and Image Processing*, 18(3):236–258, 1982.

[5] M. Leordeanu, M. Hebert, and R. Sukthankar. Beyond local appearance: Category recognition from pairwise interactions of simple features. In *CVPR*, 2007.

[6] Michael E. Leventon, W. Eric L. Grimson, Olivier Faugeras, and William M. Wells III. Level set based segmentation with intensity and curvature priors. In *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 4–11, 2000.

[7] A. Opelt, A. Pinz, and A. Zisserman. A boundary-fragment-model for object detection. In *ECCV*, 2006.

[8] Bo Wu and Ram Nevatia. Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors. In *ICCV*, pages 90–97, 2005.