Combining Local and Global Cues for Closed Contour Extraction

Vida Movahedi
http://www.cse.yorku.ca/~vida
James H. Elder
http://elderlab.yorku.ca/~elder/

Centre for Vision Research
York University
Toronto, Canada

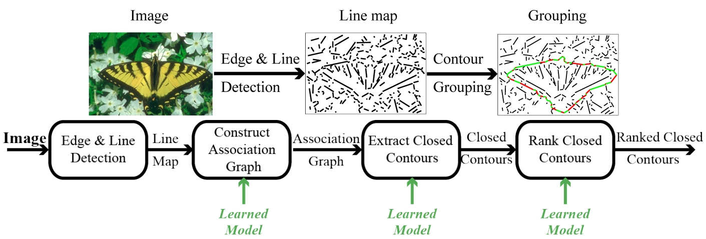

Figure 1: Overview of algorithm. The output consists of multiple non-intersecting closed polygons comprised of visible line segments (green) connected by linear interpolants (red).

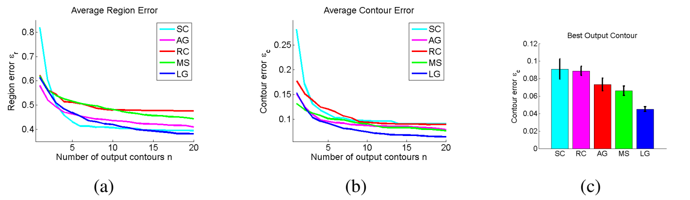

Figure 2: Quantitative evaluation on SOD test dataset. (a) Minimum region error and (b) Minimum contour error as a function of the number of output contours allowed. (c) Minimum contour error over all output contours.

We address the problem of extracting the simple closed contours that bound the salient objects in a natural image[4]. The method proposed in this paper falls in the category of *contour grouping methods*, which search for the optimal cycle of local oriented primitives (*e.g*. edgels, line segments) forming the boundary.

The first stage of processing in our method extracts from the image a sparse graph representation of the oriented structure in the image. Local oriented edges detected in the image are grouped locally into subpixel-localized line segments of variable length. As in prior work, each of the resulting segments forms a vertex in our association graph, and each edge in this graph represents a grouping hypothesis between specified endpoints of two segments.

Contour grouping algorithms are generally based upon classical Gestalt cues such as proximity, good continuation and similarityWhile prior methods [1, 2] have approximated these local cues as independent and modelled the marginals parametrically, we use EM to learn Gaussian mixture models for the 4-dimensional likelihoods, which allows us to capture small dependencies between the cues. Our sparse graph is then formed by representing only the k outgoing edges with highest likelihood for each vertex.

Algorithms for computing closed contours are generally based upon these local Gestalt cues relating pairs of oriented elements, and a Markov assumption to then group these elements into chains. Without additional global constraints, these algorithms generally do not perform well on general natural scenes. Such global cues could include symmetry, shape priors or global colour appearance. A key challenge is to combine these local and global cues in a statistically optimal way. A main contribution of our paper is a novel, effective method for combining local and global cues, both at the stage of forming new closed contour hypotheses, and at the stage of evaluating and ranking these hypotheses.

We develop and test this method using global colour contrast cues in both a greedy constructive search for closed contours, and in the final ranking of closed contours. We adopt a regression approach: learn a regressor that will predict the distance of a candidate path from ground truth based on both local and global cues. We take a nonparametric approach to the regression problem, binning the local cue $f_l\left(c_k\right)$ and global cue $f_g\left(c_k\right)$, and then estimating the mean $\hat{\varepsilon}_l$ and $\hat{\varepsilon}_g$ and variance σ_l^2 and σ_g^2 of predicted error for each. Assuming independence between the local and global cues, the least-squares error prediction ε_{ML} (ML for normal distributions) is then given by

$$\hat{\varepsilon}_{ML} = \frac{\hat{\varepsilon}_l/\sigma_l^2 + \hat{\varepsilon}_g/\sigma_g^2}{\sigma_l^{-2} + \sigma_g^{-2}} \tag{1}$$

A second challenge faced by contour grouping algorithms stems from the exponential nature of the search space, which tends to yield many very similar high probability hypotheses, while neglecting other more diverse hypotheses with slightly lower probability. This tendency lowers performance as partial hypotheses that look slightly less promising are weeded out too early. A second contribution of the present paper is a novel method for promoting diversity in the formation and ranking of contour hypotheses, leading to substantial improvements in performance.

We evaluate our results on a standard public dataset, the Salient Object Dataset (SOD) [6], which in turn is based upon the Berkeley Segmentation Dataset (BSD). We compared our local/global (LG) method with the Ratio Contour (RC) algorithm of Stahl & Wang [7], the Adaptive Grouping (AG) method of Estrada & Jepson [3], the Multiscale (MS) method of Estrada & Elder [2], and the Superpixel Closure (SC) method of Levinshtein & Dickinson [5], in each case using the parameters recommended by the authors.

We used two evaluation measures: i) A region error ε_r based on the popular intersection to union measure, and a contour error ε_c based on a normalized version of the contour mapping (CM) measure [6], defined as the average distance between corresponding pixels on contours A and B, normalized by the perimeter of the ground truth boundary B.

Figure 2 shows the quantitative results. In panels (a) and (b) we report the average region error ε_r and contour error ε_c, respectively, as a function of the number of contours n the algorithm is allowed to report. We see that our LG method outperforms other methods by the region measure when the number of output contours is 12 or greater, and by the contour measure when the number of output contours is 2 or greater.

Panel (c) shows how the best contours computed by each method compare, using the contour error measure ε_c. Our LG method performs better than other methods when the number of hypotheses is not limited. This suggests that further improvements to the LG algorithm can potentially be obtained through better methods for ranking the closed contours extracted by the algorithm.

[1] J. H. Elder and R. M. Goldberg. Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision*, 2(4): 324–353, 2002.

[2] F. J. Estrada and J. H. Elder. Multi-scale contour extraction based on natural image statistics. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06*, page 183, 2006.

[3] F. J. Estrada and A. D. Jepson. Robust boundary detection with adaptive grouping. In *Computer Vision and Pattern Recognition Workshop, CVPRW '06*, volume 2006, pages 184–184, 2006.

[4] F. Ge, S. Wang, and T. Liu. Image-segmentation evaluation from the perspective of salient object extraction. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 1:1146–1153, 2006.

[5] A. Levinshtein, C. Sminchisescu, and S. Dickinson. Optimal contour closure by superpixel grouping. *Proceedings of European Conference on Computer Vision*, 2010.

[6] V. Movahedi and J.H. Elder. Design and perceptual validation of performance measures for salient object segmentation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 49 –56, 2010.

[7] J. S. Stahl and S. Wang. Edge grouping combining boundary and region information. *IEEE Transactions on Image Processing*, 16(10): 2590–2606, 2007.