Linear Feature Extraction Based On Grouping Factors

Jing Xiao, Zhiyong Yang, Qifa Ke, Songde MA
National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences
P. O. Box 2728, Beijing 100080, P.R. China
E-mail: xiaoj@prlsun1.ia.ac.cn

Abstract–*Human vision has marvelous ability in extracting linear features from images, such as roads, rivers and so on. In this paper we present a new method to simulate this ability. Our method is based on some general grouping factors arising at two levels. At the first level, grouping factors are identified as direct bar-bar interaction and orientation interaction. Bar-bar interaction is short-ranged and homogeneous. Orientation interaction is locally oriented and mediated by statistics of local visual context. At the second level, grouping factors are global geometric binding effects which arise from geometric redundancy reduction and thus are global effects. Based on them, we present an energy model. Then the extraction of linear features is generally formulated as combinatorial optimization. Since local, global interactions and local context effects all are included in it, the model may capture partially grouping ability of human vision systems. The experiments show that, without selecting any original points, our method can extract linear features from images robustly and quickly.*

1 Introduction

In the past, much effort has been put in automatic extraction of linear features, such as roads, from images. Most of the existing methods, such as "Duda-Road-Operator"[1], generic edge detectors[2, 3], crest detectors and morphological operators[4] , are local methods since only the local information of the current terminal point is used to decide whether it belongs to a linear feature. In the presence of noise, these strategies would result in wrong decision. Kalman filter or other linear prediction methods can be used to reduce the wrong decision. But in real images, a linear object is not always locally linear, so the linear prediction methods are not always successful. In [5], a tracking method is presented which uses more global imformation. Donald Geman etc[6] proposed a model-based method for road tracking. Given the starting point and direction, this method can track the main road based on a statistical model. Brief review and comparision of the existing methods can be found in [6].

Human beings have marvelous ability to perceive linear features in images even heavily dirtied with noises. This can not be simply explained by local filtering, but involves perceptual organization. Recently neuroscience and psychophysics convergently support that this organization arises at the early vision stages without intelligent mechanisms. In this paper we suggest some general grouping factors to realize this organization. Based on these factors, we present a new and efficient method for the extraction of linear features from images. Our method includes the following two parts: 1) An edge detector is applied to transfer the original image to binary edge map(EM). After edge tracing, we obtain the edge chains of all the structures in the background. We can see from the EM that all the linear features are immerged in the background and they are, in general, broken into many segments. 2) Based on the grouping factors, we present an energy model to group the segments into line contours. Then the linear features emerge as global optimum solutions of this energy model. We can see that our method is very efficient without heuristic search.

We discribe the general grouping factors in the second section, and we present an energy model based on them in the third section, which is the basis of our method. In the fourth section, we give the main steps of our method and in the fifth section we show some experiments. Finally we give the conclusion.

2 Some General Grouping Factors

Considering our task, in the EM image, the edges of linear features are broken into many segments, so we must group them to reconstruct the line contours. We apply some general grouping factors to attain this goal.

Referring to the recent findings in neuroscience and psychophysics, we identify the general grouping factors at four levels: local direct factors, global grouping factors due to geometrical arrangement of a chain of bars,global grouping factors due to higher structures formed by the bars, and global grouping factors due to functional or semantic meanings of the structures formed by the bars. We also include the effects of local visual context on the local grouping factors. The global grouping factors are much less affected by the local context. But unluckily, at this stage we can only formulate the first two kinds of grouping factors.

2.1 Direct Interactions

Let's consider first the local grouping factors, i.e. local bar-bar interaction and orientation interaction. Bar-bar interaction is simply proposed as (See $Fig.1$)

$$V_p = -W(f(x_\parallel, x_\perp) + f(y_\parallel, y_\perp)) \quad (1)$$

$W > 0$ is a constant. Orientation interaction is suggested as(See $Fig.1$)

$$V_o = -J\Phi(g(x_\parallel, x_\perp) + g(y_\parallel, y_\perp))\cos(\theta) \quad (2)$$

θ is the turning angle. $\cos(\theta)$ is dipole interaction representing orientation tuning. Φ is a spatial smoothing factor related to local statistics of bar-bar connection. $J > 0$ is tuning strength. $f(\bullet,\bullet)$ and $g(\bullet,\bullet)$ are two anisotropic shape functions. We simply choose them as

$$f(x_\parallel, x_\perp) = g(x_\parallel, x_\perp) = exp(-\frac{x_\parallel^2}{2\mu_\parallel^2})exp(-\frac{x_\perp^2}{2\mu_\perp^2}) \quad (3)$$

Φ is in fact the context effect. Considering in 4 cases, we choose $\Phi = 1,\ 1.5,\ 1.5,\ 0.5$ individually. Case 1 is "clean limit": there are no other bars in a proper window between the two bars. Case 2 is "all inhibition": all orientations formed by bar-bar connection are inhibited in the window. In case 3, there is a coherent orientation, formed by bar-bar connection, different from the orientation formed by the considered bars. Both case 2 and 3 lead to detectability enhancement. In case 4, there are more than one un-correlated orientations formed by bar-bar connection, which degrades contour perceptual saliency. The shape function and context effect are absent in usual theories. Also note that the orientation interaction contains explicitly the position coordinate and thus the interplay between V_p and V_o is included in a natural way.

The direct interactions are in fact a way of mapping local correlation in the visual environment. There are many kinds of regularities in visual environment. A hypothesis is that one of the goals of the human vision system is to encode the visual environment in a way making efficient use of these regularities[7]. These kinds of correlation can be obtained by the study of statistics of images and can be used to derive the responding properties of the visual cortex[8].

2.2 Geometrical Binding Effect

If a chain of bars are aligned on a line of constant curvature, there is some geometrical redundancy with the chain. Due to the redundancy, two bars separated by a finite distance are correlated, and thus there is grouping tendency between them. So we suggest geometrical binding energy for the chain of bars[9]

$$B_1 = -(p - q \times |c|^{2\phi}) \times L^\alpha \quad (4)$$

$p > 0, q > 0$ are constants. L is the total length of the bar chain, including the gaps between any two neighbour bars. When the chain consists of only one bar, B_1 is exactly the energy of one bar. c is the curvature. ϕ, α are exponents. We set $0 < \alpha < 1, 0 < \phi < 1$ and choose proper length unit and short-distance cut-off to ensure that $p - q \times |c|^{2\phi} > 0$. So we always have $B_1 < 0$, which means that the geometrical binding effect is always attractive. Since the binding effect is associated with the whole emergent line contour, it's a global effect. Note that B_1 is designed in such a way that it has the following two important properties

$$\frac{\partial B_1}{\partial |c|} > 0; \frac{\partial B_1}{\partial L} < 0; \frac{\partial^2 B_1}{\partial^2 L} > 0 \quad (5)$$

The bigger $|c|$, the bigger $B_1 (B_1 < 0)$ and the smaller the binding tendency. The bigger L, the smaller B_1 and the greater the binding tendency. Thus, there is greater binding tendency with long, straight line contours formed by bars.

If a chain of points are aligned on a line contour whose curvature changes greatly from bar to bar, we suggest the geometrical binding effect for this chain of bars as

$$B_2 = 0 \quad (6)$$

In this case, there is no geometrical redundancy to be reduced. The local property of the line is much "unexpected" for vision system and thus all the changes should be represented. So we suggest the total geometrical binding effect of a chain of bars as

$$B = \sum B_1 \quad (7)$$

Note that this geometrical binding effect is related to the curvature scale used and thus there is interesting connection between it and the curvature scale space in shape representation.

3 Energy Model and Line Contour Perception

From above, we can see that the energy of a set of ungrouped segments is

$$H_1 = \sum B_1 = \sum -(p - q \times |c|^{2\phi}) \times L_1^\alpha \quad (8)$$

Where L_1 is the length of one of the segments. So H_1 is the sum of geometrical binding energy of all the segments.

If we group these segments into a line contour, its total energy is

$$H = \sum B_1 + \sum V_p + \sum V_o \quad (9)$$

Where B_1 is the geometrical binding energy of a chain of bars from the bar set, which are aligned on a line of constant curvature.

When $H < H_1$, according to marginal effect, all the gaps are filled due to the reduction of energy and then all the bars are grouped as a line contour.

Then we define the average energy of unit length as the global saliency of the line contour

$$S_G = \frac{H}{L} \tag{10}$$

Where L is the total length of the contour. The smaller $S_G(S_G < 0)$ a line contour has, the more salient perceptually the line contour. So we select the line contours with small S_G as perceptually salient ones. Note this explains why a set of bars should be grouped into line contour. First, line contours are more favored energetically than a set of isolated bars since the direct interactions and the geometrical binding effect are attractive. Second, line contours are more salient than a set of isolated bars. Due to both the local and global grouping factors, line contours can be represented with less code length than a set of isolated bars.

4 Description of Our Algorithm

Our algorithm can be deviden into two parts: 1)binary edge map production, edge tracing and contour segmentation. 2) grouping the segments to extract linear features via our energy model.

In the first step, considering the computational cost and performance, we choose the zero-crossing detector *Difference of Exponential(DOE)*[13] to produce the edge map:

$$DOE = k\left(e^{-\frac{|x|}{t\sigma}} - e^{-\frac{|x|}{\sigma}}\right) \tag{11}$$

In [13], it has been proved that with the scale ratio t of 0.3, this edge detector performs best. So we choose $t = \frac{1}{3}$. This operator can be implemented recursively, so its computational cost is very low.

According to Eq.(8), we must first decompose the edge chains into segments of constant curvature. There are many algorithms to fulfil this task of contour segmentation[10, 11, 12]. We adopt the algorithm in [10], which is robust to noise and insensitive to reasonable changes in parameters. Those who interest in the details of this algorithm are refered to [10].

Now we obtain the edge chains which have been decomposed into constant curvature segments. Without given starting bars, using our energy model, we can group the segments to extract the most salient linear teatures. The details can be found in [14].

5 Experiment Results

In what follows, we'll present some experiments.

Left in $Fig.2$ is a synthetic image with a broken sinusoid curve immersed in a background of bar noise. We can easily extract it by our algorithm, as shown in the right of $Fig.2$.

$Fig.3a$ is a satellite image with some roads(linear features) immerged in a complicated background. $Fig.3b$ is the result of contour segmentation and $Fig.3c$ is the most salient 12 linear features. It takes about 3 seconds in Sparc10 station to produce $Fig.3c$ from the edge map. We find that almost all of the roads are successfully extracted. $Fig.4a$ is an aerial image with curve-like rivers and shores in it. $Fig.4b$ is the result of contour segmentation and $Fig.4c$ is the final result using our method. This result is satisfactory, although it is not very accurate.

As can be seen from all these experiments, our method can successfully extract linear features from images quickly and robustly without given starting bars. And it is insensitive to reasonable changes in parameters.

6 Conclusion

In this paper, referring to the recent findings in neuroscience and psychophysics, we present some general grouping factors, including local bar-bar interaction and orientation interaction and global geometrical binding effect. Containing these local and global grouping factors, an energy model is introduced. Based on this model, we propose an efficient method to extract linear features from images. Some experiments show that our method is quick, robust and very effective.

References

[1] R.O. Duda and P.E. Hart, *Pattern Calssification and Scene Analysis*, New York: John Wiley & Sons, 1973.

[2] R. Neviata, Locating structures in aerial images, *IEEE Trnas. PAMI*, pp. 476-484, Sept. 1982.

[3] F. Wang and R. Newkirk, A knowledge-based system for high-way network extraction, *IEEE Trans. Geosicence and Remote Sensing*, Vol. 26, Sept. 1988.

[4] I. Estival and H. Le Men, detection of linear networks on satellite images, *Proc. Conf. Pattern Recognition*, pp. 856-858, Oct. 1986.

[5] M. Fischler, H. Tenenbaum, and H. Wolf, Detection of roads and linear structures in low-resolution aerial imagery using a multi-source knowledge integration technique, *Computer Graphics and Image Processing*, Vol. 15, pp. 201-223, 1981.

[6] Donald Geman and B. Jedynak, An active testing model for tracking roads in satellite images, *IEEE Trnas. PAMI*, Vol. 18, pp.1-14, 1996.

[7] H. B. Barlow, T. P. Kaushal, & G. J. Michison, Finding minimum entropy codes. *Neural Comput.* **1**, 412-423, 1989.

[8] B. A. Olshausen, & D. J. Field, Emergence of simple cell receptive field properties by learning a sparse code for natural images. *Nature* **381** 607-609, 1996.

[9] Yang Zhiyong and Ma Songde, A phenomenological approach to salient maps and illusory contours, To appear in *Network: Computation in Neural Systems*.

[10] Daniel M. Wuescher and K.L. Boyer, Robust contour segmentation using a constant curvature criterion, *IEEE Trans. PAMI*, Vol. 13, pp. 41-51, 1991.

[11] M. Fischler and R.C. Bolles, Perceptual organization and curve partitioning, *IEEE Trans. PAMI-8*, pp. 100-105,1986.

[12] Paul L. Rosin and Geoff A. W. West, Nonparametric segmentation of curves into various representations, *IEEE Trans. PAMI*, Vol. 17, pp.1140-1153, 1995.

[13] Song De Ma, Bingcheng Li, Derivative Computation by Multiscale Filters , *Visual Image Computing*, December, 1995.

[14] Jing XIAO, Zhiyong YANG, Qifa KE and Songde MA, Linear Feature Extraction Based On Grouping Factors, *Technical Report of National Lab of Pattern Recognition*, 1998.

Figure 1: Direct interaction: Bar-bar interaction and orientation interaction.

Figure 2: Left: A sinusoid curve of bars immersed in a background of bars. Right: The sinusoid extracted by our model.

Figure 3: a: A satellite image with some roads immerged in a complicated background. b: The result of edge detection and contour segmentation. c: The most salient 12 linear features extracted by our method.

Figure 4: a: An aerial image with curve-like rivers and shores. b: The result of edge detection and contour segmentation. c: The final result by our method.