

# Contour Continuity in Region Based Image Segmentation

Thomas Leung and Jitendra Malik

Department of Electrical Engineering and Computer Sciences  
University of California at Berkeley, Berkeley, CA 94720, USA  
{leungt,malik}@cs.berkeley.edu

**Abstract.** Region-based image segmentation techniques make use of similarity in intensity, color and texture to determine the partitioning of an image. The powerful cue of contour continuity is not exploited at all. In this paper, we provide a way of incorporating curvilinear grouping into region-based image segmentation. Soft contour information is obtained through orientation energy. Weak contrast gaps and subjective contours are completed by contour propagation. The normalized cut approach proposed by Shi and Malik is used for the segmentation. Results on a large variety of images are shown.

## 1 Introduction

Grouping is a very important problem in visual perception. To humans, an image is not just a random collection of pixels; it is a meaningful arrangement of regions and objects. Figure 1 shows a variety of images — (a) an artificial image showing strong subjective contours in the Kanizsa triangle; (b) a painting which combines line and paint and; (c) a natural photograph. Despite the large variations of these images, humans have no problem interpreting them. We can agree about the different regions in the images and recognize the different objects.

Human visual grouping was studied extensively by the Gestalt psychologists in the early part of the century [28]. They identified several factors that lead to human perceptual grouping: similarity, proximity, continuity, symmetry, parallelism, closure and familiarity. In computer vision, these factors have been used as guidelines for many grouping algorithms. The most studied version of grouping in computer vision is image segmentation. Image segmentation techniques can be classified into two broad families—(1) region-based, and (2) contour-based approaches. Each of these has characteristic advantages and disadvantages.

Region-based approaches try to find partitions of the image pixels into sets corresponding to coherent image properties such as brightness, color and texture. This gives rise to proto-surfaces. This is important for a number of reasons: (1) surfaces are the natural units of perception [19]; (2) one can compute texture/color descriptors for each region, which is necessary for applications such as content-based image querying.

Region-based techniques usually involve defining a global objective function (for example, Markov random fields [7] or other variational formulations [18]).



**Fig. 1.** Some challenging images for a segmentation algorithm. The left image is the Kanizsa triangle. Subjective contours induced in this image lead us to interpret it as a triangle occluding 3 circles instead of 3 “pacman” figures. The middle image is a combination of line drawing and paint. The right image is a photograph with several low contrast edges, e.g. between the man’s jacket and the woman’s blouse. Our goal is to develop a single grouping procedure which can deal with all these types of images.

The advantage of having a global objective function is that hard decisions are made only when information from the whole image is taken into account at the same time. The major drawback of region techniques is that curvilinear continuity, a very powerful constraint, is not exploited.

Contour-based approaches usually start with a first stage of edge detection, followed by a linking process that seeks to exploit curvilinear continuity. Examples include dynamic programming [14], relaxation approaches [21] and saliency networks [23]. The major advantage of contour-based approaches is that the grouping factor of curvilinear continuity can be treated very naturally. The disadvantages are: (1) contour-based techniques do not give us closed connected regions; so the grouping problem of finding closed connected surfaces is not really solved. As mentioned earlier, surfaces are desirable entities. Thus a post-processing step remains. This may take the form of explicitly enforcing closure of contours as in [4, 11]. (2) Contour-based techniques cannot deal with textured regions easily. Basically, for textured regions, we have to find feature vectors and look for differences — that is more natural in region approaches. (3) Hard decisions are made locally and prematurely. To detect an extended contour of very low contrast, a very low threshold has to be set for the edge detector. This will cause random edge segments being found everywhere in the image. The linking process then has to deal with all the spurious edge segments everywhere. In other words, there is no recognition of the fact that extended contours, though with low contrast, are perceptually more significant than short random edges in the image.

Our goal in this paper is to incorporate curvilinear grouping in a region-based setting. In all region approaches, one needs to define a distance function

— a measure of dissimilarity between pairs of pixels. Then, some procedure is needed to find regions where similarity between pixels in a given region is high and similarity between pixels in different regions is low. We propose a distance function which takes contours into account softly through orientation energy. Similarity is low if the orientation energy between two pixels is high, hinting at the possible presence of a contour between them. Low contrast gaps along a contour are completed by the propagation of orientation energy. We would like to emphasize that subjective contours, those enabling us to perceive the Kanizsa triangle in Figure 1(a), are simply the extreme case of low contrast contours (contrast is zero). Thus, any algorithm exploiting curvilinear continuity should be able to deal with subjective contours naturally. With the distance function defined, the normalized cut procedure proposed by Shi and Malik in [25] is employed to produce the segmentation.

The outline of the paper is as follows. In Section 2, previous work in curvilinear grouping will be reviewed. How we exploit curvilinear continuity in defining pixel similarity is described in Section 3. The overall image segmentation algorithm is presented in Section 4. A review of the normalized cuts approach for image segmentation will be given in Section 4.1. In Section 5, we will show results on various kinds of images like those shown in Figure 1. We conclude in Section 6.

## 2 Previous Work on Curvilinear Grouping

Psychophysical studies on curvilinear grouping date back to the early part of the century. Wertheimer pointed out the factor of good continuation in perceptual grouping in [28]. Subjective contours, first discovered by Schumann [22], were studied extensively by Kanizsa [12]. Kanizsa was the first to point out that subjective contours are simply gradientless edges — the extreme case of low contrast edges. More recently, Kellman and Shipley [13] and Field et al [5] studied the formation of curvilinear groups psychophysically.

In computer vision, the problem of contour completion has been approached in many different ways. Most techniques employ some form of edge detection followed by a linking process. The linking process can take the form of: dynamic programming [14], relaxation [21], saliency networks [23], voting [8] or sequential tracking in a Bayesian framework [2]. In general, these methods can bridge small gaps along an edge to produce an extended contour.

Another direction of work is the computation of subjective contours. Most methods involve first the computation of key-points, such as junctions, corners and line ends. Different models of subjective contour completion from these key-points are proposed: variational models [6, 20], random walk [29] or a grouping field [9]. Occlusion reasoning is also studied in [9, 20]. Most of these approaches have only been applied to artificial images, while some have been tested on very simple real images. Since subjective contours are just the extreme case of low contrast contours, a useful model should be able to work on realistic photographs.

We believe the inadequacy of previous work is partly due to the dependence on key point detection, which is rather unreliable in realistic images.

### 3 Curvilinear Continuity

The main idea of region-based segmentation is to compute local similarity and then optimize the segmentation over the whole image through a global criterion. For segmentation using intensity alone, local similarity is invariably some measure of how much the intensities of two pixels are alike. In this section, we describe how information about curvilinear continuity can be incorporated into the similarity measure between two pixels.

#### 3.1 Soft Contour Information

Most previous curve detection methods rely on a first step of edge detection, followed by edge linking. Edge detection is a local process. This has the important drawback that hard decisions are made prematurely. We propose that contour information be computed “softly” and hard decisions be made only when the image is considered as a whole. Information about the strength of a contour can be obtained through *orientation energies* [15, 16]. Let  $\mathcal{F}_1(x, y)$  be the second derivative of an elongated gaussian kernel and  $\mathcal{F}_2(x, y)$  be the Hilbert transform of  $\mathcal{F}_1(x, y)$ . More precisely:

$$\begin{aligned}\mathcal{F}_1(x, y) &= \frac{d^2}{dy^2} \left( \frac{1}{C} \exp\left(\frac{y^2}{\sigma^2}\right) \exp\left(\frac{x^2}{\lambda^2 \sigma^2}\right) \right) \\ \mathcal{F}_2(x, y) &= \text{Hilbert}(\mathcal{F}_1(x, y))\end{aligned}$$

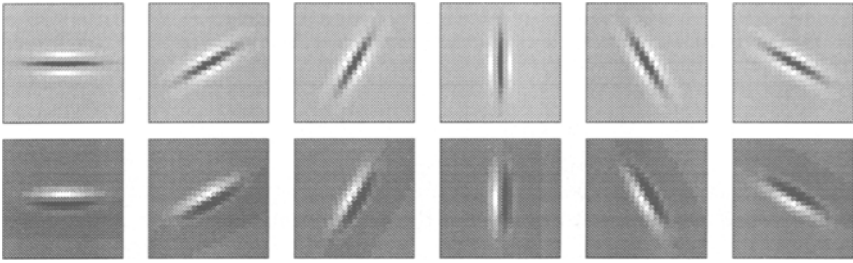
where  $\sigma$  is the scale and  $\lambda$  is the elongation of the filter.  $C$  is a constant. The orientation energy at angle  $0^\circ$  is defined as:

$$OE_{0^\circ} = (I * \mathcal{F}_1)^2 + (I * \mathcal{F}_2)^2$$

$OE_{0^\circ}$  has maximum response for horizontal contours. Rotated copies of the two filter kernels will be able to pick up edge contrast at various orientations. These kernels are shown in Figure 2. At each pixel, we can define the orientation energy and the orientation as:

$$\begin{aligned}OE_{con}(x, y) &= \max_{\phi} OE_{\phi}(x, y) \\ \phi(x, y) &= \arg \max_{\phi} OE_{\phi}(x, y)\end{aligned}\tag{1}$$

The orientation energy defined in Equation 1 has the following nice properties: (1) the second derivative of the gaussian and the Hilbert transform of it are a quadrature pair [15, 16]. Quadrature pairs are phase independent and they are not affected by exact localization of the edges. (2) The filters are insensitive to linear intensity variations, which are usually caused by smooth shading,



**Fig. 2.** Top: rotated copies of the second derivative of an elongated gaussian kernel.  $\sigma = 1.0$  and the elongation ratio is 4. Bottom: rotated copies of the Hilbert transform of the second derivative kernel. In our implementation, we use 12 orientations uniformly sampled in the range from  $0^\circ$  to  $180^\circ$ . This quadrature pair is chosen because both filters are insensitive to linear variations of intensity, which are usually caused by smooth shading, rather than object boundaries. The orientation energy resulting from the quadrature pair is phase independent and is not affected by the precise localization of the contour.

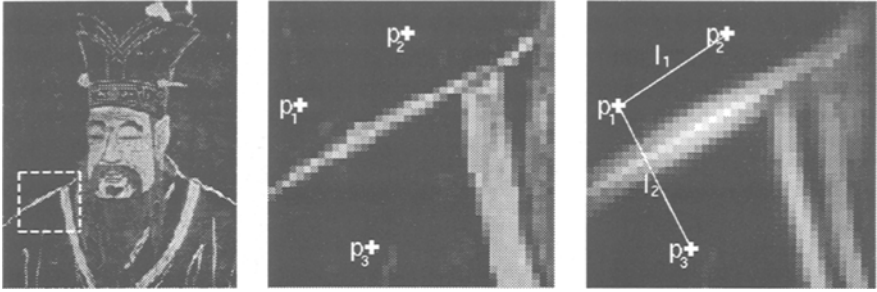
rather than object boundaries. (3) The filters are elongated, thus information is integrated along the edge. Long contours will produce stronger responses than shorter ones. Extended low contrast contours will also have significant response because of the integration. (4) Our filters do not respond well to high curvature contours. This property agrees with human perception that curvilinear grouping does not occur over high curvature contours.

From the orientation energy, we can compute the dissimilarity between two pixels due to the factor of curvilinear continuity. Intuitively, two pixels belong to two different groups if there is a contour separating them. The dissimilarity is stronger if the contour separating them is extended. Figure 3 is a good illustration of this intuition. On the left is the image. The middle figure shows a magnified part of the original image. On the right is the orientation energy. Pixels  $p_1$ ,  $p_2$  and  $p_3$  all have similar intensity values. Based purely on intensity (normally the only cue for region based segmentation), all three locations will have strong similarities. However, there is an extended contour separating  $p_3$  from  $p_1$  and  $p_2$ . Thus, we expect  $p_1$  to be much more strongly related to  $p_2$  than  $p_3$ . This intuition carries over in our definition of dissimilarity between two pixels: if the orientation energy along the line between two pixels is strong, the dissimilarity between these pixels should be high. Formally, we define the dissimilarity,  $d_{edg}(p_1, p_2)$ , between two pixels as follows:

$$d_{edg}(p_1, p_2) = OE_{con}(\hat{x}) - \text{avg}(OE_{con}(p_1), OE_{con}(p_2); \hat{x}) \quad (2)$$

where

$$\hat{x} = \arg \max_{x \in l} OE_{con}(x)$$



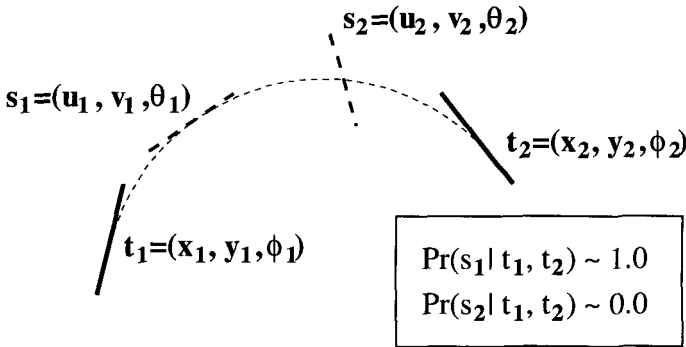
**Fig. 3.** Left: the original image. Middle: part of the image marked by the box. The intensity values at pixels  $p_1$ ,  $p_2$  and  $p_3$  are similar. However, there is an edge in the middle, which suggests that  $p_1$  and  $p_2$  belong to one group while  $p_3$  belongs to another. Just comparing intensity values at these three locations will mistakenly suggest that they belong to the same group. Right: orientation energy. The orientation energy somewhere on  $l_2$  is strong which correctly proposes that  $p_1$  and  $p_3$  belong to two different partitions, while orientation energy along  $l_1$  is weak throughout, which will maintain the hypothesis that  $p_1$  and  $p_2$  belong to the same group.

$$\text{avg}(OE_{con}(p_1), OE_{con}(p_2); \hat{x}) = \frac{|p_2 - \hat{x}|}{|p_1 - p_2|} OE_{con}(p_1) + \frac{|\hat{x} - p_1|}{|p_1 - p_2|} OE_{con}(p_2) \quad (3)$$

$l$  is the straight line between  $p_1$  and  $p_2$ ;  $\hat{x}$  is the location where the orientation energy is maximum on  $l$ ; and  $\text{avg}(OE_{con}(p_1), OE_{con}(p_2); \hat{x})$  is the weighted average orientation energy of  $p_1$  and  $p_2$ . It is worth noting that in our definition, we considered the *relative* difference in orientation energy. When images are noisy, orientation energy along the path may be high merely due to noise. In this case, orientation energies at  $p_1$  and  $p_2$  will be high as well. Thus, the relative difference is a better indicator of the presence of a contour.

### 3.2 Contour Propagation

The orientation energy defined in the previous section will be high at a sharp contour. When there are gaps along a contour where contrast is low (or even absent in the case of subjective contours), the orientation energy will not do a very good job. However, it is worth pointing out that due to the integration of the filter kernels along a contour, orientation energy will still be reasonable when the gaps are small. It will fail when gaps are large. In this section, we provide a more explicit way of enhancing the orientation energy at low contrast gaps by propagation from neighboring pixels along an extended contour. Before we proceed, let us repeat that subjective contours are just the extreme case of a low contrast contour (the contrast is zero). Thus, our treatment here will be able to capture subjective contours in the same way as we enhance the orientation energy at low contrast gaps.



**Fig. 4.** Consider pixel location  $(u_i, v_i)$ . We are interested in the probability  $P_{cp}(s_i | t_1, t_2)$  that there is an image curve originating from  $t_1$  and terminating in  $t_2$  which passes through  $s_i$ . This probability can be related to the minimum energy of all curves passing through the three points. Notice that this probability will depend heavily on the orientation  $\theta_i$ . In this case,  $P_{cp}(s_1 | t_1, t_2)$  will be high while  $P_{cp}(s_2 | t_1, t_2)$  will be close to 0.

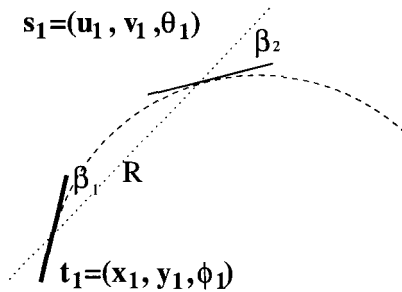
Consider the oriented element  $(u_1, v_1, \theta_1)$  where  $(u_1, v_1)$  are the coordinates and  $\theta_1$  is the orientation. What is the probability that there is an image contour originating at  $(x_1, y_1, \phi_1)$ , passing through  $(u_1, v_1, \theta_1)$  and ending up at  $(x_2, y_2, \phi_2)$ ? There are infinitely many curves that satisfy our requirement. One popular choice is the curve that minimizes the *elastica* functional, i.e. the image curve with the minimum energy.<sup>1</sup> The elastica was already introduced in Euler's time. Ullman [27] and Horn [10] are among the first to apply it to contour completion. For a discussion of the application of the elastica to computer vision, readers are referred to [17]. Using the elastica, the probability in which we are interested can be related to the energy of the minimizing curve:  $P_{cp}(s_1 | t_1, t_2) = \exp(-E_{elastica})$ , where  $s_i = (u_i, v_i, \theta_i)$ ,  $t_i = (x_i, y_i, \phi_i)$  and  $E_{elastica}$  is the minimum energy. Likewise, the probability that there is no image curve passing through  $(u_1, v_1, \theta_1)$  is simply  $1 - P_{cp}(s_1 | t_1, t_2)$ . First notice that  $P_{cp}(s_1 | t_1, t_2)$  should depend heavily on  $\theta_1$ : in Figure 4,  $P_{cp}(s_1 | t_1, t_2)$  should be close to 1, while  $P_{cp}(s_2 | t_1, t_2)$  will be close to 0.

Computing the elastica energy exactly is not necessary and not computationally advisable. Sharon et al [24] showed that we can have a simple approximation by assuming that  $P_{cp}(s_i | t_1, t_2)$  can be factored into two terms:  $P_{cp}(s_i | t_1, t_2) = f(s_i, t_1)f(s_i, t_2)$  with  $f(\cdot)$  as:

$$f(s_i, t_i) = \exp\left(-\frac{R}{\sigma_R} - \frac{D_\beta}{\sigma_\beta}\right) \quad (4)$$

$$D_\beta = \beta_1^2 + \beta_2^2 - \beta_1\beta_2$$

<sup>1</sup> Energy is defined as the integral of the square of the curvature along the curve.



**Fig. 5.** Two oriented elements  $t_1$  and  $s_1$  separated by a distance  $R$ .  $\beta_1$  is the angle between the orientation at  $t_1$  and the line joining the two pixels. Similarly,  $\beta_2$  is the angle between the orientation at  $s_1$  and the line joining the two pixels.

where  $\beta_1$ ,  $\beta_2$  and  $R$  are defined in Figure 5.  $\sigma_R$  and  $\sigma_\beta$  specify the amount of variations allowed in the distance and the agreement of the angles.

Using,  $P_{cp}(s_1|t_1, t_2)$ , the propagation of orientation energy can be defined as:

$$OE_{cp}(u, v) = \max_{\theta} \max_{(x_1, y_1), (x_2, y_2) \in \mathcal{N}} P_{cp}((u, v, \theta)|t_1, t_2) \cdot \overline{OE_{con}}$$

where  $\overline{OE_{con}} = \frac{1}{2}(OE_{con}(x_1, y_1) + OE_{con}(x_2, y_2))$  and  $\mathcal{N}$  is a local neighborhood of  $(u, v)$ . Notice that this definition maintains the intuitive requirement that the propagated orientation energy is never larger than the average energy of the propagating sites. Moreover, the propagation is biased towards extended low curvature contours. Combining orientation energy due to contour contrast ( $OE_{con}$ ) and contour propagation ( $OE_{cp}$ ), we have, at each pixel:

$$OE_{tot}(x, y) = \max(OE_{con}(x, y), OE_{cp}(x, y))$$

Let us discuss what the above equation implies. If the orientation energy due to contour contrast at  $(u, v)$  is high ( $OE_{con}$  is large), we do not change anything. However, when the energy is low, caused by low contrast, we are enhancing the energy through the propagation from neighboring elements. In the case when energy at  $(u, v)$  is zero (subjective contours), we can complete the curve by propagating from neighboring contours.

With the combined orientation energy, we can update our dissimilarity measure (Equation 2) as:

$$d_{edg}(p_1, p_2) = OE_{tot}(\hat{x}) - \text{avg}(OE_{tot}(p_1), OE_{tot}(p_2); \hat{x}) \quad (5)$$

with  $\hat{x} = \arg \max_{x \in I} OE_{tot}(x)$  and  $\text{avg}(\cdot, \cdot; \cdot)$  defined in Equation 3.

## 4 Partitioning Images

All region-based image segmentation techniques start with local measurements of similarity. Local information is combined over the image through a global



criterion. The “correct” segmentation is the one that maximizes this global criterion. Equation 5 tells us the dissimilarity between two pixels based on contour information alone. We combine contour information with intensity and/or color to produce the overall similarity measure between pixels  $p_i$  and  $p_j$  as follows:

$$W_{ij} = \exp(-\alpha_{edg}d_{edg}(p_i, p_j) - \alpha_{int}d_{int}(p_i, p_j) - \alpha_{col}d_{col}(p_i, p_j)) \quad (6)$$

$d_{int}(p_1, p_2) = |I_1 - I_2|$  and  $d_{col}(p_1, p_2) = |H_1 - H_2| + |S_1 - S_2|$  where  $I_i, H_i, S_i$  are the intensity, hue and saturation at pixel  $i$ . The  $\alpha$ 's are weighting coefficients for the different factors. One can think of the  $\alpha$ 's as prior knowledge relating to the relative importance of the different factors. In this paper, we provide a way of adjusting these parameters for each image. This will be discussed in section 4.2. First, we will review the normalized cuts algorithm [25] which we used to segment the image based on the similarity measure in Equation 6.

#### 4.1 Image Segmentation using Normalized Cuts

Let  $W \in \mathbb{R}^{N \times N}$  be the association matrix, i.e.  $W_{ij}$  is the weight between pixels  $i$  and  $j$  (Equation 6). We proceed to apply the normalized cut criterion proposed by Shi and Malik in [25] to partition the image. Shi and Malik formulate visual grouping as a graph partitioning problem. The nodes of the graph are the entities that we want to partition, for example, in image segmentation, they will be the pixels. The edges between two nodes correspond to the *strength* that these two nodes belong to one group, again in image segmentation, the edges of the graph will correspond to how much two pixels agree in intensity, color, etc. Intuitively, the criterion for partitioning the graph will be to minimize the sum of weights of connections *across* the groups and maximize the sum of weights of connections *within* the groups.

Let  $G = \{\mathbf{V}, \mathbf{E}\}$  be a weighted undirected graph, where  $\mathbf{V}$  are the nodes and  $\mathbf{E}$  are the edges. Let  $\mathbf{A}, \mathbf{B}$  be a partition of the graph:  $\mathbf{A} \cup \mathbf{B} = \mathbf{V}, \mathbf{A} \cap \mathbf{B} = \emptyset$ . In graph theoretic language, the similarity between these two groups is called the *cut*:

$$cut(\mathbf{A}, \mathbf{B}) = \sum_{u \in \mathbf{A}, v \in \mathbf{B}} w(u, v)$$

where  $w(u, v)$  is the weight on the edge between nodes  $u$  and  $v$ . Shi and Malik proposed to use a *normalized* similarity criterion to evaluate a partition. They call it the *normalized cut*:

$$Ncut(\mathbf{A}, \mathbf{B}) = \frac{cut(\mathbf{A}, \mathbf{B})}{assoc(\mathbf{A}, \mathbf{V})} + \frac{cut(\mathbf{B}, \mathbf{A})}{assoc(\mathbf{B}, \mathbf{V})} \quad (7)$$

where  $assoc(\mathbf{A}, \mathbf{V}) = \sum_{u \in \mathbf{A}, t \in \mathbf{V}} w(u, t)$  is the total connection from nodes in  $\mathbf{A}$  to all the nodes in the graph. For discussions on this criterion, please refer to [25].

One key advantage of using the normalized cut is that a good approximation to the optimal partition can be computed very efficiently.<sup>2</sup> Let  $\mathbf{W}$  be the association matrix, i.e.  $W_{ij}$  is the weight between nodes  $i$  and  $j$  in the graph. Let  $\mathbf{D}$  be the diagonal matrix such that  $D_{ii} = \sum_j W_{ij}$ , i.e.  $D_{ii}$  is the sum of the weights of all the connections to node  $i$ . Shi and Malik showed that the optimal partition can be found by computing:

$$\begin{aligned} \mathbf{y} &= \arg \min Ncut \\ &= \arg \min_{\mathbf{y}} \frac{\mathbf{y}^T (\mathbf{D} - \mathbf{W}) \mathbf{y}}{\mathbf{y}^T \mathbf{D} \mathbf{y}} \end{aligned} \quad (8)$$

where  $\mathbf{y} = \{a, b\}^N$  is a binary indicator vector specifying the group identity for each pixel, i.e.  $y_i = a$  if pixel  $i$  belongs to group  $\mathbf{A}$  and  $y_j = b$  if pixel  $j$  belongs to  $\mathbf{B}$ .  $N$  is the number of pixels. Notice that the above expression is a Rayleigh quotient. If we relax  $\mathbf{y}$  to take on real values (instead of two discrete values), we can optimize Equation 8 by solving a generalized eigenvalue system. Efficient algorithms with polynomial running time are well-known for solving such problems. For details of the derivation of Equation 8, please refer to [25].

Transforming the vector  $\mathbf{y}$  to a discrete bipartition  $\hat{\mathbf{x}}$  can simply be done by finding the threshold  $\tau$  such that if we let  $\hat{x}_i = 1$  when  $y_i > \tau$  and  $\hat{x}_i = -1$  otherwise, the normalized cut value in Equation 7 is minimized. In our implementation, we try  $\tau$  at 10 values uniformly distributed between the maximum and minimum of  $\mathbf{y}$ . Given a bipartition of the image, we recursively bipartition each segment until either a pre-defined number of steps is reached, or when the normalized cut value resulting from the bi-partitioning is too large. When normalized cut value is too large, it means there is no more segmentation necessary for this group.

## 4.2 Parameter Estimation

One natural question to ask is that now we have the various factors combined at the same time, how do we evaluate the importance of each of the them in an image. In other words how do we determine the  $\alpha$ 's in Equation 6? One may look at this problem as finding prior weights for the various factors. However, we argue that these weights should change from image to image. For example, in a line drawing, all information comes from the contours. Intensity and color should be de-emphasized. Thus,  $\alpha_{edg}$  should be much larger than  $\alpha_{int}$  and  $\alpha_{col}$ . On the other hand, in an image with regions of distinctly different colors, but similar intensities,  $\alpha_{col}$  should dominate.

Since the goodness of each partition is given by the normalized cut value (the smaller the better), we can simply use gradient descent on the parameters to minimize this value. The gradient of  $Ncut$  with respect to  $\alpha_i$  is:

$$\frac{\partial Ncut}{\partial \alpha_i} = \frac{Ncut(\alpha_i + \delta_i) - Ncut(\alpha_i)}{\delta_i}$$

<sup>2</sup> Finding the true optimal partition is an NP-complete problem.

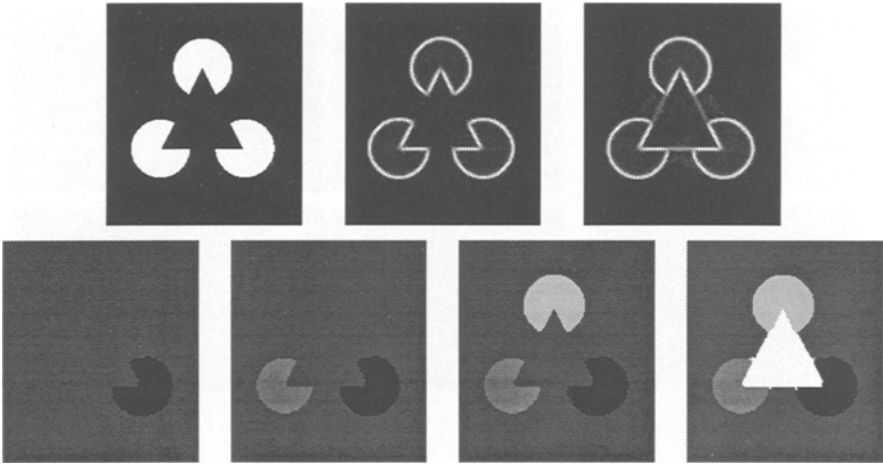
We update the parameters by moving them in the negative gradient direction of  $Ncut$ . Notice that computing  $Ncut(\alpha_i + \delta_i)$  is very efficient. When  $\delta_i$  is small, the partition for  $Ncut(\alpha_i + \delta_i)$  is similar to the one already found for  $Ncut(\alpha_i)$ . Therefore, we can use the optimal partition for  $Ncut(\alpha_i)$  as an initial estimate for solving the generalized eigenvalue system in Equation 8 with the new parameters. This can save at least an order of magnitude in computation time.

## 5 Results

In this section, we will demonstrate some of the results on various kinds of images: (1) images with strong subjective contours; (2) line-drawings; (3) paintings which combine both line-drawing and paint; and (4) gray-level images. In all the examples shown in this paper, there was no manual tuning of thresholds or parameters. The large variety of images presented in this paper illustrate the wide applicability of our algorithm.

Since subjective contours are just the extreme case of low contrast contours, all curvilinear grouping algorithms should be general enough to deal with this special case. The Kanizsa triangle and the Kanizsa “2-fish” images [12] shown in Figures 6 and 7 are good examples to demonstrate how our algorithm performs in the presence of subjective contours. Orientation energies due to contours alone ( $OE_{con}$ ) and after contour propagation ( $OE_{tot}$ ) are also shown. Notice the spurious completion in the orientation energy after contour propagation,  $OE_{tot}$ . These are the results of propagation between points in the middle of two different contours. In some previous work, this problem is avoided by detecting key-points (junctions, corners and line ends) and then performing subjective contour completion only from the key-points. However, we argue that detecting these features in real images is a problem in itself. We believe that such features cannot be reliably detected purely from local information. In other words, in realistic images, we have to live with such artifacts and rely on a global criterion to sort out what is significant and what is not. The segmentations of the Kanizsa figures are also shown in Figures 6 and 7. Notice that the results agree with our percepts.

The result on a 3D line drawing is shown in Figure 8. Region-based segmentation usually has problems dealing with these line-drawings because the intensity values in the different regions are the same. The only information comes from the contours. Results of the segmentation on graylevel images are shown in Figure 9. The original image is shown on the left; the segmentations are shown in the middle. The right column shows the edges of the segmentation. The edges are shown just because the contrast between two different segments in the middle figure may not be high enough to be seen. As we have argued earlier, regions are more useful than just edges. The reader is urged to check the edge image while examining the segmentation. Notice the extended weak contrast edges in the images: (1) between the bodies of the penguins and the snow; (2) between the suit and the pants in the third image; (3) between the man’s jacket and the woman’s blouse in the fourth image. Our algorithm is able to pick them up

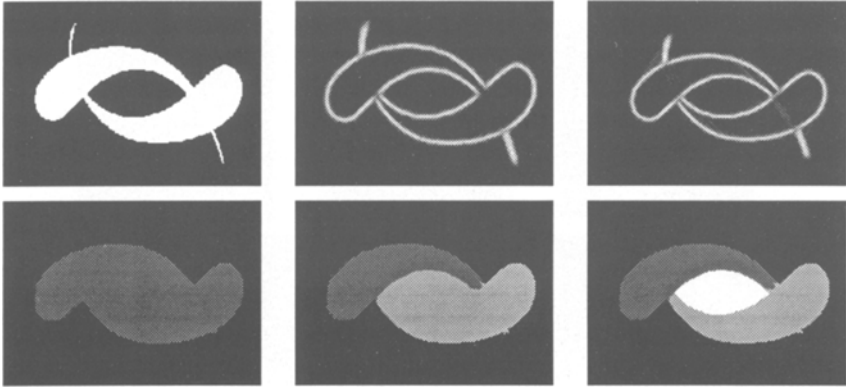


**Fig. 6.** Top left: the Kanizsa triangle. While we perceive a triangle occluding 3 circles, any region segmentation technique based just on intensity will invariably find 3 “pacman” figures. Middle: the orientation energy from contours ( $OE_{con}$ ). Right: orientation energy after contour propagation ( $OE_{tot}$ ). Notice that contour propagation also produces spurious extensions around the figure. We argue that they are unavoidable and the only way to resolve the problem is through a global criterion. Bottom: The sequence of segmentation as obtained from our recursive partitioning. Notice that the spurious contour propagation does not affect our segmentation. The reason is that the global criterion is able to gather information from the whole image and realize that the spurious contour propagation is not significant.

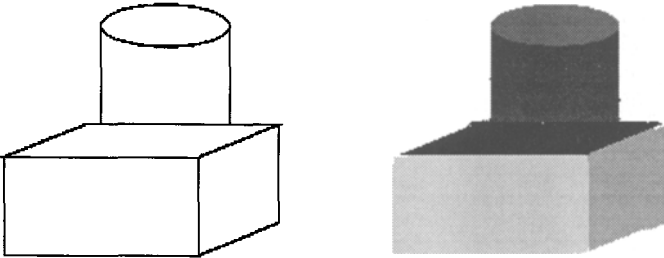
without producing lots of random edges everywhere, as would be expected from an edge detector. The fifth image demonstrates a combination of line drawing and paint, which is segmented by our algorithm as well. Notice once again that the same parameter settings were used for all segmentation results shown in Figures 6 to 9. Running times of the algorithm are on the order of 3-5 minutes for  $200 \times 200$  images on a Pentium II.

## 6 Discussions

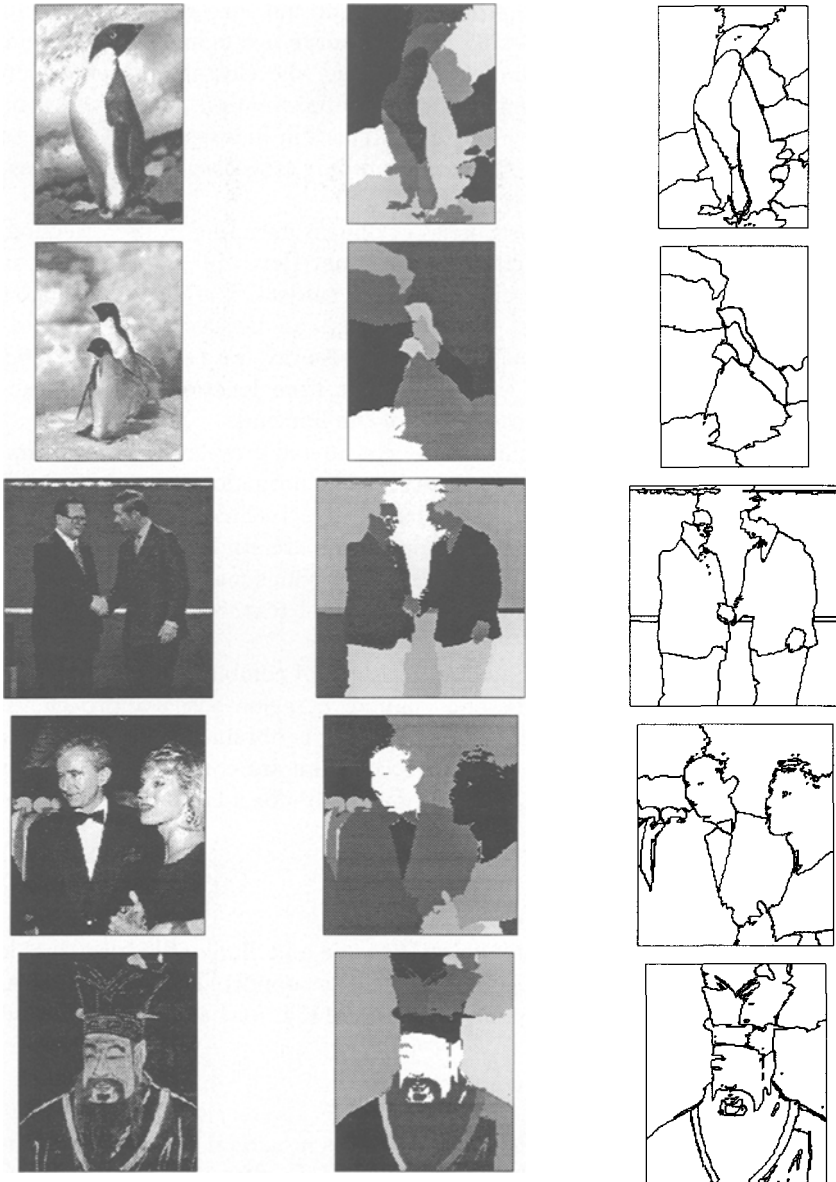
In our study, we have not investigated the issue of occlusion. Previous work on subjective contours has tried to reason about occlusions through junctions, corners and line ends at the same time contours are formed. However, we believe that local detection of key-points is unreliable. Instead, once a reliable segmentation is produced, junctions can be easily detected as the intersection of regions. Moreover, junctions in the image does not always correspond to occlusion events. For example, any objects with surface markings will produce lots of “junctions”



**Fig. 7.** Top left: the Kanizsa “2-fish” image — two mutually occluding fishes with the bodies of each overlapping the tails of the other. Middle and Right: orientation energy due to contours ( $OE_{con}$ ) and after propagation ( $OE_{cp}$ ). Notice the spurious completion around the figure. Bottom: the sequence of recursive partitioning. The “tails” are missed because they are too small.



**Fig. 8.** A linedrawing and its segmentation into regions found by our algorithm. Traditional region-based segmentation techniques would be completely useless on this image because the intensity or color in the different regions are identical.



**Fig. 9.** The segmentation of gray-level images. Original image in the first column, segmentation in the middle, and the edges of the segments in the third. The edges are shown to indicate the boundaries between the segments in cases where the contrast in the middle figure may not be high enough to be seen. Notice the weak contrast edges in the images and our segmentation result. The same parameter settings were used for all the results shown in Figures 6 to 9.

in the image. We believe the problem of occlusion, or more generally figure-ground segregation is best tackled when contour belongingness and junctions are considered at the same time. For example, the Gestalt grouping factor of symmetry is a useful cue for figure-ground determination. A pair of symmetry contours suggest that the contours and the space in between are likely to belong to the same object, thus the “figure”. Contours, together with junctions, can provide better figure ground segregation.

In our work, we have only studied curvilinear grouping in the direction parallel to the curve. However, Heitger and von der Heydt [9] showed that contour completion also occurs at line ends, inducing subjective contours orthogonal to the direction of the line ends. Our algorithm can be easily extended to take this into account. Together with orientation energy, we can compute *end-stop energy*. Contour propagation will then occur from locations of high end-stop energy perpendicular to the orientation of the line ends.

Note that in this paper, we have only considered gray-level and color images. However, real images will also have texture. The normalized cuts framework can incorporate texture information very naturally. Techniques for doing texture segmentation in the normalized cuts framework are studied in [1, 26]. However, the effect of contour continuation in textured regions may have to be treated in a more subtle fashion, depending on the class of texture [3]. We are exploring this in ongoing work.

To conclude, we have presented a technique of combining information about curvilinear continuity, intensity and color in a region-based approach to perform image segmentation. Contour information is obtained through orientation energy. Low contrast gaps or subjective contours are completed by the propagation of orientation energy. Segmentation results for a large varieties of images are demonstrated.

## Acknowledgement

The authors would like to thank Serge Belongie and Jianbo Shi for valuable discussions related to this work. This research was supported by (ARO) DAAH04-96-1-0341, the Digital Library Grant IRI-9411334, and a Berkeley Fellowship.

## References

1. S. Belongie and J. Malik. Finding boundaries in natural images: A new method using point descriptors and area completion. In *Proc. 5th Euro. Conf. Comp. Vision*, Freiburg, Germany, June 1998.
2. I.J. Cox, J.M. Regh, and S. Hingorani. A bayesian multiple-hypothesis approach to edge grouping and contour segmentation. *Int. J. of Comp. Vision*, 11(1), 1993.
3. B. Dubuc and S.W. Zucker. Indexing visual representations through the complexity map. In *Proc. 5th Int. Conf. Computer Vision*, Cambridge, MA, U.S.A., June 1995.
4. J.H. Elder and S.W. Zucker. Computing contour closures. In *Proc. Euro. Conf. Computer Vision*, volume I, Cambridge, England, Apr 1996.
5. D.J. Field, A. Hayes, and R.F. Hess. Contour integration by the human visual system: evidence for a local “association field”. *Vision Research*, 33(2), 1993.

6. D. Geiger and K. Kumaran. Visual organization of illusory surfaces. In *Proc. Euro. Conf. Computer Vision*, volume I, Cambridge, England, Apr 1996.
7. S. Geman and D. Geman. Stochastic relaxation, gibbs distribution, and the bayesian retonation of images. *IEEE Trans. PAMI.*, 6, Nov. 1984.
8. G. Guy and G. Medioni. Inferring global perceptual contours from local features. *Int. J. of Computer Vision*, 20(1-2), Oct. 1996.
9. F. Heitger and R. von der Heydt. A computational model of neural contour processing. In *Proc. Int. Conf. Computer Vision*, Berlin, Germany, May 1993.
10. B.K.P. Horn. The curve of least energy. *ACM Trans. on Math. Soft.*, 9(4), 1983.
11. D. Jacobs. Robust and efficient detection of salient convex groups. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(1), Jan. 1996.
12. G. Kanizsa. *Organization in vision: essays on gestalt perception*. Praeger, 1979.
13. P.J. Kellman and T.F. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23, 1991.
14. U. Montanari. On the optimal detection of curves in noisy pictures. *Comm. Ass. Comput.*, 14, 1971.
15. M.C. Morrone and D.C. Burr. Feature detection in human vision: a phase dependent energy model. *Proc. R. Soc. Lond. B*, 235, 1988.
16. M.C. Morrone and R.A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6, 1987.
17. D. Mumford. Elastica and computer vision. In Chandrajit Bajaj, editor, *Algebraic Geometry and Its Applications*. Springer Verlag, 1994.
18. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions, and associated variational problems. *Comm. Pure Math.*, 1989.
19. K. Nakayama and S. Shimojo. Experiencing and perceiving visual surfaces. *Science*, 257, Sept 1992.
20. M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation and Depth*. Springer-Verlag, 1993.
21. P. Parent and S.W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(8), Aug. 1989.
22. F. Schumann. Contributions to the analysis of visual perception — first paper: some observations on the combination of visual impressions into units (a. hogg, trans.). In S. Petry and G. Meyer, editors, *The perception of illusory contours*. Springer-Verlag, 1987. (original work published 1900).
23. A. Sha'ashua and S. Ullman. Structural saliency: the detection of globally salient structures using a locally connected network. In *Proc. 2nd Int. Conf. Computer Vision*, Tampa, Fl., U.S.A., 1988.
24. E. Sharon, A. Brandt, and R. Basri. Completion energies and scale. In *Proc. IEEE Conf. Comp. Vision and Pat. Rec.*, San Juan, Puerto Rico, June 1997.
25. J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 1997.
26. J. Shi and J. Malik. Self inducing relational distance and its application to image segmentation. In *Proc. 5th Euro. Conf. Comp. Vis.*, Freiburg, Germany, June 1998.
27. S. Ullman. Filling-in the gaps: the shape of subjective contours and a model for their generation. *Biological Cybernetics*, 25, 1976.
28. M. Wertheimer. Laws of organization in perceptual forms (partial translation). In W.B. Ellis, editor, *A sourcebook of Gestalt Psychology*. Harcourt Brace and Company, 1938.
29. L.R. Williams and D.W. Jacobs. Local parallel computation of stochastic completion fields. *Neural Computation*, 9(4), May 1997.