

Constructing Coherent Boundaries

T. Rachidi* and L. Spacek
Department of Computer Science,
University of Essex,
Colchester C04 3SQ
U.K.

Abstract

This paper presents a new approach to the problem of gap bridging and junction detection. Perceptual groupings and evidence from existing boundaries are combined to produce joins which are consistent with the initial image structure. In particular, a new definition of co-curvilinearity, which distinguishes between true and false co-curvilinearity, is given. Structural quantities are computed for all potential joins. A local non-iterative algorithm selects the best joins and/or junctions which satisfy specific structural conditions. No assumption, domain restriction, or model is needed. The current implementation is fully presented together with the results obtained.

Keywords: *Gap bridging, Perceptual grouping, Junction detection, Connectivity.*

1 Introduction

Since the earliest attempts at computer vision and pattern recognition, it was realised that reducing an image into a set of connected boundaries helps with compressing visual data and eliminating noise, whilst retaining topological and geometric properties. Connected boundaries play fundamental roles in inferring shapes [9], recognising objects, and 3D interpretation of scenes [2].

When trying to extract boundary representations of images, the main problem encountered is the large number of discontinuities (or gaps) between boundary fragments of the same physical contour and near junction points. These gaps originate from noise distortion, weak contrast, and the intrinsic nature of conventional smoothing operators such as $\nabla^2 G$, which blur details of image structure near junction points.

Our goal is to reconstruct junctions and boundaries so that they are coherent and in accord with the initial image structure. The method should be suitable for general purpose computer vision systems, *i.e.* it should be knowledge-free. Such a process should, of course, incorporate perceptual organisation capabilities akin to

*This research was supported in part by an ORS award.

humans [13, 17, 8, 3, 10]. Many existing techniques for reconstructing junctions and bridging gaps already include these features. However, these techniques lack the crucial stage of checking the plausibility of the new constructions with the evidence locally present in the existing contours and in the raw image. It is our primary task to present a local algorithm, which combines perceptual grouping factors and specific structural conditions to bridge gaps and reconstruct junctions.

2 Related work

Many existing techniques are based on perceptual grouping, such as proximity, continuity, connectivity and co-curvilinearity [17, 8, 3, 10]. It is agreed that such capabilities must also be present in machine vision systems if they are to have generality and function in real scenes.

Co-curvilinearity, in particular, was identified as one of five viewpoint-invariant features which are necessary and sufficient for inferring shape and obtaining robust curve descriptions [7, 11]. It accounts for most, if not for all, groupings done by human visual system. In [6], angles are detected and grouped into junctions.

In [11], curves that terminate in a local neighbourhood are candidates for linking with each other by line segments. The neighbourhood is defined in terms of the acceptable distance d between the ends of the curves; d is taken as a fraction of the curve length with a fixed minimum length. A local non-iterative process selects the most co-curvilinear (or least bent) joins among neighbouring curves.

Mohan & Nevatia extended the perceptual grouping factors to account for parallelism, symmetry, T-junctions, corners and U-structures [12]. Line segments are gradually grouped into predefined shapes, termed *collated features*, that are thought to be omni-present in images.

The major problem with these techniques is that they lack a crucial stage, which is to check whether the selected bridges are consistent with the evidence already present and the structure of the raw image, and thus may produce wrong junctions. Additional imperfections can also be identified:

- When using line-segments rather than boundary fragments the obtained boundaries lack smoothness.
- Predefined shapes used in [12] limit the sphere of applicability of such techniques to man made objects.
- Thresholding on angles as in [6] has the undesirable effect of eliminating simple L-junctions, which are so common in everyday scenes.
- The measurements presented for the co-curvilinearity in [11, 6] do not distinguish between genuine and false co-linearity, as shown in **Figure 1**. Genuine co-curvilinearity involves two disconnected boundary fragments (segments) of the same physical boundary of a narrow band. The very nature of conventional operators and tessellation effects result in the eventual displacement of sections of the boundary by a few pixels.

Another critical problem associated with previously outlined approaches is ensuring the right choice of size for the neighbourhood. Small windows lead to wrong bridges, and large windows bring into consideration other segments that may interfere with the correct bridges.

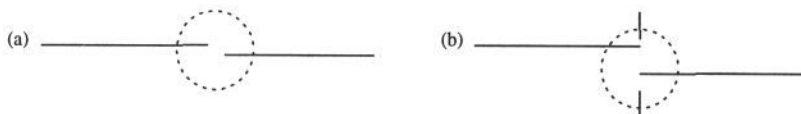


Figure 1. (a) shows a genuine colinearity. (b) false colinearity. These two cases need to be distinguished.

3 Our approach

We shall present an approach to junction detection and gap bridging which offers solutions to the problems identified above. Our approach uses curves rather than line segments, we combine perceptual groupings with evidence from existing boundaries and the initial raw image to connect boundary fragments. Co-curvilinearity, proximity, contrast and continuity are used to predict potential joins between boundary fragments which terminate in the same neighbourhood. A suitable band, based on the prediction stage, is then chosen for each potential join. Structural quantities are computed from intensities in this band to verify the plausibility of the predicted joins with image structure. This approach has some similarities with Canny's technique for edge detection [1]. Parts of the contour around the point of interest (*i.e.* evidence) are examined to decide on the entire segment using hysteresis thresholding (*i.e.* verification). Finally, a local non-iterative process based on ordering and partitioning selects the plausible joins and junctions.

3.1 Input used

The initial boundary fragments are extracted from Spacek's *first difference magnitude surface* [15] (see **Figure 2**), which combines both edge strength and boundary thinning. Spacek computes gradient components $\partial f/\partial x$ and $\partial f/\partial y$ by convolving the initial image $f(x, y)$ with two linear filters defined by: $abs(x).x/(x^2 + y^2)$ and $abs(y).y/(x^2 + y^2)$. Ridge points are obtained by checking for a local maximum in the direction of the gradient vector. A non-maximum suppression technique is applied to eliminate false points.

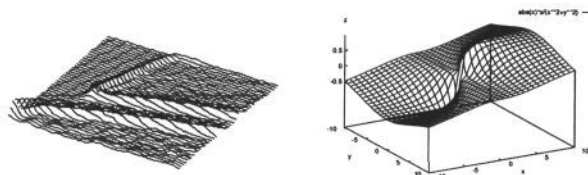


Figure 2. Left: first difference magnitude surface. Connected boundaries are encoded along ridges (not shown here) using an 8-connected grid. Right: Spacek's finite difference template for $\partial f/\partial x$.

3.2 Potential joins

Each boundary fragment, and its end points, are used to build the *adjacency graph*, in which co-terminations form the nodes and boundary fragments form the arcs (dual links) between them.

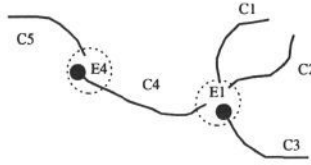


Figure 3. The nodes (dots) created at E_1 and E_4 capture the connectivity.

A node captures the neighbouring co-terminations (see **Figure 3** for illustration). A single node is created at E_1 . This node *attracts* neighbouring co-terminations at a distance *radius* from it; *radius* is a global parameter (see results section). The choice of E_1 is purely arbitrary.

The term *join* henceforth is used to refer to an abstract entity which denotes the perceptual and structural characteristics of a link between two boundary fragments.

Potential joins are searched for in the *adjacency graph* between neighbouring fragments. For every pair of neighbouring fragments, the following entities are computed:

3.2.1 Perceptual coherence

- The difference in contrast dc of the two boundaries involved in the join.
- The co-curvilinearity cl of the section PP' (defined later).
- End point proximity embodied in $d = \overline{QQ'}$.
- The eventual boundary which the join would give rise to, and its total length l .

Let P and P' be two points belonging respectively to B and B' such that $\overline{PQ} = \overline{Q'P'} = b$, where b is a global control parameter. Since the vectors \vec{PQ} and $\vec{Q'P'}$ could be assimilated to the tangents to B and B' at Q and Q' respectively, the co-curvilinearity is defined as follows (see **Figure 4**):

$$cl(j) = \begin{cases} 0 & \text{if } PQ \cap Q'P' = \infty \\ & \text{(straight join)} \\ \overline{PQ} \overline{Q'P'} & \text{if } PQ \cap P'Q' = I \\ & \text{(angled join)} \\ 0 & \text{if } PQ \cap P'Q' = \emptyset \text{ and } W \leq 2 \\ & \text{(tessellation effect)} \\ \pi & \text{if } PQ \cap P'Q' = \emptyset \text{ and } W > 2 \\ & \text{(wrong co-curvilinearity)} \end{cases}$$

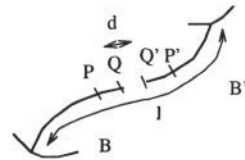


Figure 4. Computing perceptual characteristics of a join.

where I is the intersection point (when it exists) between lines PQ and $Q'P'$, and $W = d(PQ, P'Q')$ the distance between the lines PQ and $Q'P'$ when parallel. \emptyset is the empty set. Note that $cl \in [0, \pi]$. This definition distinguishes between

false and true co-curvilinearity by construction. Joins with small values of cl are preferred over joins with larger values. $cl = \pi$ indicates that there is no perceptual evidence at all for the join.

3.2.2 Structural coherence

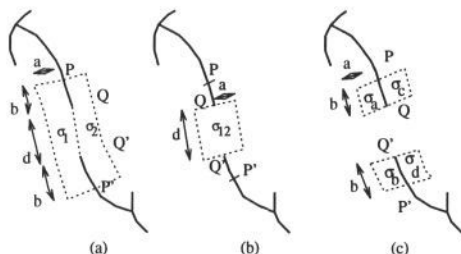


Figure 5. Explanative figure for computing the consistency of a join with image structure.

The differences in image intensities at the location of a gap cannot be used directly. Edge detectors would not fail at these locations in the first place, if these differences were relevant. However, we still can measure the uniformity of the intensities on each side of a potential join. This will give hints on whether the gap is a genuine one or not. To do so, we compute σ_1, σ_2 and σ_{12} the standard deviations in image intensity along and across a potential join (see **Figure 5** (a) & (b)).

The physical reason behind these measurements is that image intensities are uniform on both sides of a genuine join, *i.e.*, σ_1 and σ_2 are less than a certain threshold and less than σ_{12} . In case of a 3-way junction though, only one side is uniform and there is no uniformity of intensity across the potential join.

3.3 Selected joins

Potential joins are ordered in a list with respect to the following criteria of comparison: let j be a join, cl_j, d_j, l_j and dc_j respectively its co-curvilinearity, length, total length of the eventual new boundary, and the difference in contrast between the boundaries involved. The ordering is done by the following set of rules:

- 1) $|cl_j - cl_{j'}| > \epsilon_{cl}$ & $cl_j < cl_{j'} \Rightarrow j < j'$ 2) $|cl_j - cl_{j'}| > \epsilon_{cl}$ & $cl_{j'} < cl_j \Rightarrow j' < j$
- 3) $\frac{d_j}{d_{j'}} < \epsilon_d \Rightarrow j < j'$ 4) $\frac{d_{j'}}{d_j} < \epsilon_d \Rightarrow j' < j$ 5) $\frac{l_j}{l_{j'}} < \epsilon_l \Rightarrow j' < j$
- 6) $\frac{l_{j'}}{l_j} < \epsilon_l \Rightarrow j < j'$ 7) $TRUE \Rightarrow j' = j$

These rules are applied in order (*i.e.* 1,2,3,4,5,6 and 7). The first one to have its left hand side conditions satisfied decides on the ordering between j and j' . Rules 1 and 2 reflect the preference for joins that will result in the best co-curvilinearity first. If no decision can be taken on the basis of co-curvilinearity, *i.e.* there is no

strong evidence that one join is more co-curvilinear than the other, the proximity factor is used to decide on the ordering. This is embodied in rules 3 and 4. A very short join is much preferred over a long one. If still no decision can be taken on the basis of the proximity factor, *i.e.* there is no strong evidence that a join is much shorter than the other, preference is given to joins that will result in longer boundaries. Rule 7, when reached, indicates that joins j and j' are not discernable on the basis of the above grouping factors. This level of ambiguity hardly ever occurs. The constants ε_{cl} , ε_d and ε_l are positive thresholds in [0..1].

However, a join j is genuine only when the following structural conditions hold: $\sigma_1(j) \leq T_1$, $\sigma_2(j) \leq T_2$, $T'_{12} \leq \sigma_{12}(j) \leq T_{12}$, the bridge QQ' does not cross any other boundary, and the resulting boundary does not cross itself. Where T_1 , T_2 and T_{12} are local thresholds. The use of thresholds seems unavoidable in this context.

Automatic thresholding techniques can be roughly classified into two categories: one based on statistical concepts, the other on preserving the geometric structure of images [16]. We adopt the former strategy in computing local thresholds: Let σ_a , σ_b , σ_c and σ_d be the standard deviations in intensity levels in the specified locations in **Figure 5** (c). The sigmas represent evidence near strong edges found by the edge detector. We define local thresholds: $T_1 = \max(\sigma_a(j), \sigma_b(j))$, $T_2 = \max(\sigma_c(j), \sigma_d(j))$, $T'_{12} = \min(\sigma_{ac}(j), \sigma_{bd}(j))$, $T_{12} = \max(\sigma_{ac}(j), \sigma_{bd}(j))$.

The last condition is equivalent to Canny's hysteresis thresholding technique [1]. Joins which do not satisfy these conditions are marked as candidates for a 3-way junction (see next section).

Similarly, a junction is genuine (see **Figure 6**), only if the following structural conditions hold: $\sigma_1(j) \leq \sigma_a(j)$, $\sigma_2(j) \leq \sigma_b(j)$, $\sigma_{12}(j) \leq \sigma_{ab}(j)$, and the bridge $Q''J$ does not cross any other boundary.

The junction point J is the nearest point on PP' to Q'' (see **Figure 6**).

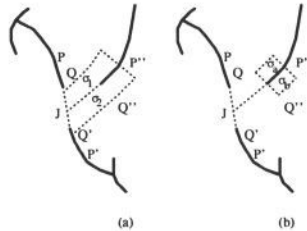


Figure 6. *Explanative figure for computing the structural coherence for 3-way junctions.*

3.4 Bridging strategy

The general idea lies in partitioning the ordered set of potential joins at a boundary end-point into two sets. One set consists of simple joins between pairs of boundaries. When fused, each will give rise to a single boundary. The other set comprises joins to be each combined with an additional boundary not involved in the first set in order to produce 3-way junctions. Joins of each set satisfy the

structural conditions. Moreover, they are consistent¹ with each other and with joins of the other set. It is only after computing these sets that bridging and junction reconstruction are performed.

Fused boundaries are deleted from the global list, and the adjacency graph is updated accordingly. Attributes of newly created boundaries, such as contrast and length, are also updated. Final boundaries are collected with the topological structure and junction points.

Line segment's points and corner's points are generated using an adaptation of midpoint line scan-conversion algorithm [5] to produce chain codes.

At this stage, we draw the attention to the following points:

- Contrary to previously described techniques, our approach offers a solution to the neighbourhood problem discussed earlier
- No assumption or domain limitation is needed
- The complexity of the approach is governed by the number of end-points $2n$ (where n is the number of boundaries) and the local computations at each end-point. These local computations are proportional to the radius r , as will be empirically shown in the results section. It follows that the overall complexity is $O(2nr)$.

4 The results

The parameters that govern our bridging process are l_{min} and $radius$ for the global control, ϵ_{cl} , ϵ_d and ϵ_l for local perceptual grouping, and a , b for computing the standard deviations and co-curvilinearity. Most of these parameters were fixed for all experiments. Typically, $l_{min} = 3$, $(\epsilon_{cl}, \epsilon_d, \epsilon_l) = (0.175, 0.50, 0.175)$, $a = 3$ pixels and $b = 5$ pixels.

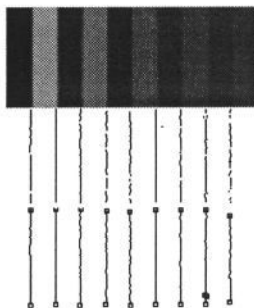


Figure 7. *From top to bottom: A 250×100 image of step edges at various contrasts, boundary fragments detected and boundaries obtained using our bridging algorithm. Panel intensities are 64, 190, 96, 158, 112, 142, 120, 134, 124 and 130. Edge finders responses are rounded to the nearest pixel position, hence the zigzags.*

The performance has been tested on a variety of synthetic and real images.

¹Two joins are said consistent if they do not have a boundary in common and do not cross each other. A binary image is used locally to compute intersections.

The first synthetic image is a reproduction of the data used by Fleck [4] to test the output of various edge detectors. This image represents step edges at various contrasts (**Figure 7**). The panels were first smoothed with a Gaussian of standard deviation 1 cell and then corrupted with Gaussian noise of standard deviation 1 intensity unit to simulate real imaging conditions.

It is apparent that the structural conditions for detecting gaps work well, even in cases of very low contrast. In this experiment only one gap out of 26 was not bridged. This miss is actually due to the fact that the co-curvilinearity of this join is π , and thus is not eligible for bridging on the basis of perceptual grouping.

The second synthetic experiment involves noisy step edges. The panels were first smoothed with Gaussian of standard deviation 1 cell and then corrupted with Gaussian noise of increasing standard deviation (**Figure 8**).

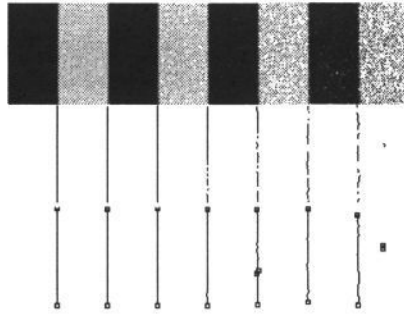


Figure 8. From top to bottom: A 400×100 image of noisy step edges, boundary fragments detected, and boundaries obtained using our bridging algorithm. Panel intensities are 27 and 227, and the boundaries occur between two cells. Odd-numbered panels are corrupted with Gaussian noise of standard deviation 8.0, 16.0, 32.0 and 64.0 units. Even-numbered panel boundaries are not corrupted.

The first real image shows a pair of scissors on a mouse pad (**Figure 9**).

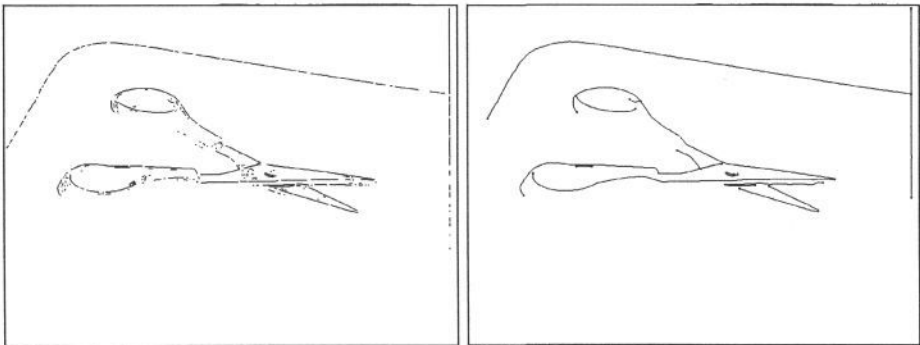


Figure 9. Left: boundaries extracted from 1st difference magnitude surface of a pair of scissors. Right: final boundaries obtained ($r = 40$). Junctions are marked with dots.

The table below shows the performance obtained with a *SPARC station ELC* when processing **Figure 9**.

r(pixels)	time(s)	% reduction
20	5.6	69
25	6.7	70
30	8.0	70
35	8.4	70
40	10.0	70
45	10.6	70
50	11.7	70
55	12.9	70
60	14.3	70
65	15.7	70

Table 1: *Recorded measurements from processing the scissors.*

The reduction factor (Table 1, 3rd column) is: $f = 1 - \frac{\text{final boundaries}}{\text{initial boundaries}}$. The times recorded are for the bridging process to run, including the construction of the two graphs and I/O operations. For this particular image, $r_{opt} = 25$ is the optimum radius. Radii $r > r_{opt}$, produce the same boundaries as $r = r_{opt}$. r_{opt} depends strictly on the quality of the input. The graph (r, time) confirms our claim that the complexity of the approach is $O(2nr)$.

For the above values of r , on average 83% of the joins were selected on the basis of a strong co-curvilinearity, 14% on the basis of proximity and 3% on the basis of the total length. The order in which co-curvilinearity and proximity were applied was not important.

The second experiment is a (768×576) picture of various objects (see **Figure 10**). Final boundaries for this image (see **Figure 10**) were obtained using the radius $r = 11$. Out of 160, only 44 boundaries were retained, which gives a reduction factor of 72%. All 3-way junctions were successfully solved. 38% of the joins were based on co-curvilinearity, 10% on proximity, 51% on the total length of the newly produced boundary, and only 1 join was ambiguous.

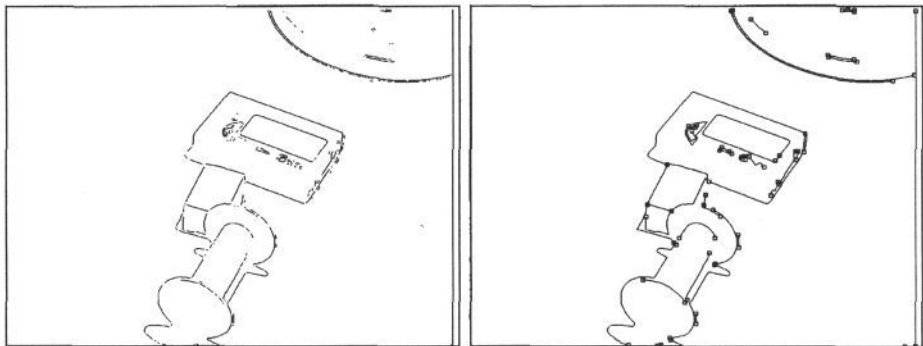


Figure 10. *Left: Input to our system: boundaries extracted from first difference magnitude surface of an image of a few objects. Right: final boundaries obtained ($r = 11$). Detected junctions and end-points are marked with squares.*

5 Conclusion

We have presented an approach to gap bridging which constructs coherent boundaries in images using both perceptual groupings and image structure criteria. The underlying characteristic of the approach is consistency with the initial image, which makes it suitable for general-purpose computer vision systems. The authors successfully used these boundaries for correspondence computation [14].

References

- [1] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:678–698, 1986.
- [2] Martin C Cooper. Interpretation of line drawings of complex objects. *Image and Vision Computing*, 11(2):82–90, 1993.
- [3] M. A. Fischler and R. C. Bolles. Perceptual organization and curve partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8:100–105, 1986.
- [4] Margaret M. Fleck. Multiple widths yield reliable finite differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(4):412–429, 1992.
- [5] J. D. Foley, A. VanDam, S. K. Feiner, and J. F. Hughes. *Computer Graphics: Principles and Practice*. Addison Wesley Publishing Company, Inc., 1987.
- [6] Radu Horaud, Françoise Veillon, and Thomas Skordas. Finding geometric and relational structures in an image. In *European Conference in Computer Vision*, pages 373–384, 1990.
- [7] Martin D. Levine and Robert Bergevin. Extraction of line drawing features for object recognition. In *IEEE International Conference on Pattern Recognition*, pages 496–501, 1990.
- [8] D. G. Lowe. *Perceptual Organization and Visual Recognition*. Hingham, MA: Academic, 1985.
- [9] D. Marr. Analysis of occluding contour. In *Royal Society*, volume B197, pages 441–475, 1977.
- [10] J. D. McCafferty. *Human and Machine Vision: Computing Perceptual Organisation*. Ellis Horwood Ltd, 1990.
- [11] Rakesh Mohan and Ramakant Nevatia. Segmentation and description of scenes based on perceptual organization. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 333–341, 1989.
- [12] Rakesh Mohan and Ramakant Nevatia. Using perceptual organization to extract 3-d structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1121–1139, 1989.
- [13] S. E. Palmer. The psychology of perceptual: A transformational approach. In *Human and Machine Vision*, pages 269–239. Beck, Hope and Rosenfield, Eds., New York: Academic, 1983.
- [14] T. Rachidi and L. Spacek. Boundary-based correspondence computation using the topology constraint. In *Proceedings of the 5th British Machine Vision Conference*, York, September 1994.
- [15] L. Spacek. Thinning image boundaries. CSM 187, Dept of Computer Science, University of Essex, Colchester C04 3SQ UK., 1993.
- [16] Din-Chang Tseng and Mao-Yu Huang. Automatic thresholding based on human visual perception. *Image and Vision Computing*, 11(9):539–548, 1993.
- [17] S. W. Zucker. Computational and psychophysical experiments in grouping: Early orientation selection. In *Human and Machine Vision*, pages 545–567. Beck, Hope and Rosenfield, Eds., New York: Academic, 1983.