Analysis of Ramp Discontinuity Model for Multiscale Image Segmentation

Himanshu Arora
University of Illinois
Urbana, IL61801, USA
harora1@uiuc.edu

Narendra Ahuja
University of Illinois
Urbana, IL61801, USA
n-ahuja@uiuc.edu

Abstract

This paper presents an algorithm for multiscale image segmentation. Towards this, it proposes a new region model, that of a homogenous region surrounded by ramp discontinuities (a scenario usually encountered in real images). This paper presents the analysis of this model, leading to a robust algorithm for detection of ramp discontinuities in the image, and finally segmentation of the image at different photometric scales. The algorithm is further specialized for detecting very thin regions. The final segmentation algorithm can detect regions in the image, with varying photometric scales, sizes and arbitrary geometric properties. The properties of the algorithm are experimentally verified on synthetic and real images.

1 Introduction

Low level segmentation attempts at partitioning the image on the basis of the gray level homogeneity. Since different regions might have different homogeneity levels, multiscale analysis is an essential part of image segmentation. Also, since region boundaries can be of arbitrary shape, no priors on the geometry of the regions should be used for segmentation. Most of the segmentation algorithms proposed in the past either neglect scale, or induce some priors for geometry of the edges. Traditional approaches such as region growing [9], watersheds [6], and rule based approaches [5] completely neglect scale. The energy minimization based approaches, e.g. Markov Random field (MRF) [3], and active contours [4] induce prior constraints on curvature and shape of segments. Normalized cuts based segmentation [7] requires the number of regions as an input, which has no intrinsic relation with the scale of spatial or photometric homogeneity. Mean-shift segmentation [2] clusters the image pixels in joint intensity and spatial domain, using a parzen-window density estimate of the image. Choosing a fixed bandwidth for the parzen-window estimate induces averaging, due to which steep corners can never be detected.

We draw our motivation from the work in [1, 8] which attempts at achieving multiscale segmentation of an image, with the scale explicitly representing the gray-level homogeneity and spatial extent of the region. In the above work, segmentation is achieved by using physically motivated attraction forces between pixels, according to their respective photometric and geometric scales. These forces do not use any prior information about the geometry of the boundaries. The authors however use a very simplistic region model, that of a homogeneous region surrounded by a step-discontinuity. This is also the case with all other segmentation methodologies in literature. In real images, due to discretization and blurring effects, an edge between regions appears as a continuously varying ramp discontinuity, which is usually perceived as multiple regions by all segmentation algorithms.

In this paper, we propose a new region model, that of a region surrounded by ramp discontinuities. We present an analysis of this model, leading to a robust algorithm for estimation of the ramp parameters. The analysis facilitates developing a strategy for detecting very thin elongated regions in the image. The segmentation at a photometric scale is achieved by evolution according to attraction forces between pixels proposed in [1, 8]. The segmentation has the property that regions at higher photometric scales are formed by strict merging of regions at lower photometric scales.

In Sec. 2, we briefly describe the attraction force fields originally introduced in [1]. In Sec. 3, we present the analysis of region model, and explain the algorithm for segmentation. In Sec. 4, we present results of the algorithm, verifying the properties of segmentation on synthetic and real images. Finally, we conclude the paper in Sec. 5.

2 The Transform

The force transform introduced in [1] maps an image I into an attraction force field $\boldsymbol{F}$. The vector $\boldsymbol{F_p}$ at image location $\boldsymbol{p}$ for a given spatial scale parameter $\sigma_s(\boldsymbol{p})$, and

photometric scale parameter $\sigma_g(\boldsymbol{p})$ is defined as

$$\boldsymbol{F_p} = \int_{\mathcal{R}} d_g(\Delta I, \sigma_g(\boldsymbol{p})) \cdot d_s(\boldsymbol{r}, \sigma_s(\boldsymbol{p})) \frac{\boldsymbol{r}}{||\boldsymbol{r}||} d\boldsymbol{q}, \quad (1)$$

where $\mathcal{R}$=domain$(I)\backslash\{\boldsymbol{p}\}$, ΔI=$|I(\boldsymbol{p})-I(\boldsymbol{q})|$, and $\boldsymbol{r} = \boldsymbol{q} - \boldsymbol{p}$. The functions $d_g(\cdot)$ and $d_s(\cdot)$ are positive, bounded kernel functions with symmetric, non increasing profiles. We use the boxcar function as the kernel, given by

$$\begin{array}{l} d_g(\Delta I, \sigma_g) \propto \mathrm{B}_{\Delta I}(\sigma_g), \\ d_s(\boldsymbol{r}, \sigma_s) \propto \mathrm{B}_{||\boldsymbol{r}||}(\sigma_s), \end{array} \quad B_x(c) = \begin{cases} 1, |x| \leq c, \\ 0, |x| > c, \end{cases} \quad (2)$$

Consider a region R surrounded by step discontinuities, with minimum contrast across the discontinuities being σ_g. For a point $\boldsymbol{p}$ inside R at a distance σ_s away from R's boundary, it was shown in [1] that the force $\boldsymbol{F_p}$ for parameters σ_g and σ_s is always directed towards the medial axis of R. Thus the force field always exhibits divergence at region boundaries. Also, for a given photometric scale σ_g, the spatial scale associated with each point p can be computed as the smallest value of σ_s with $\boldsymbol{F_p} > 0$. This simplistic model thus easily yields segmentation at different photometric scales by examination of forces for each point $\boldsymbol{p}$ in the image, given that the region boundaries are step discontinuities.

3 The Model and Algorithm

It is possible to easily extract regions using the force transform, given there exist only step discontinuities in the image. The transform however cannot handle the case when these discontinuities are spread over an area, forming a ramp discontinuity. In this section, we first introduce the ramp discontinuity based region model. We then present an analysis leading to an algorithm for estimation of width and contrast across the ramp. Finally, we present an algorithm to achieve final segmentation based on detected ramp pixels.

3.1 Region Model

A region R is modelled as a set of homogeneous contiguous pixels surrounded by a ramp discontinuity. Inside the ramp discontinuity, the intensity changes rapidly between homogenous intensity levels of its adjoining regions. This change in intensity is assumed to be monotonic. The ramp discontinuities usually occur due to blur introduced by image acquisition process (defocus and low resolution). Each ramp discontinuity has an associated width, which might vary as we go around the region. The contrast across the cross section of the ramp might also vary across the region. Note that a step discontinuity is a special case of a ramp with zero width. Every region R in the image has an associated photometric scale $\sigma_g(R)$, defined by the minimum contrast across the ramp discontinuity surrounding it. Thus, the lower the photometric scale of a region, the *fainter* are its edges. Photometric scale also reflects the level of homogeneity of the region, since the maximum variation in the neighborhood of any point in the interior of R will always be less than $\sigma_g(R)$. Segmentation at photometric scale σ_g is defined as the set of regions with photometric scale greater than σ_g. As we increase σ_g, there are fewer ramp discontinuities with contrast greater than σ_g. Hence, regions strictly merge to form bigger regions as the photometric scale is increased. Segmentation at scale σ_g can be achieved by estimating the pixels inside ramps of contrast greater than σ_g.

To analyze the model, let us first define the following function over contrast for each point $\boldsymbol{p}$ in the image.

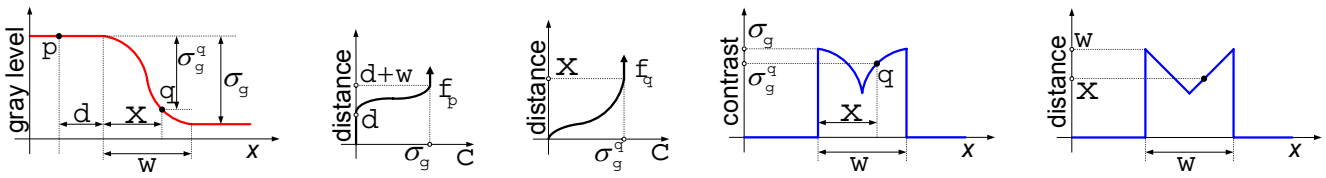
$$f_p(c) = \min\{||\boldsymbol{p} - \boldsymbol{q}|| : |I(\boldsymbol{p}) - I(\boldsymbol{q})| \geq c\}, \quad (3)$$

where, $||\boldsymbol{p} - \boldsymbol{q}||$ denotes the distance between points $\boldsymbol{p}$ and $\boldsymbol{q}$, $I(\boldsymbol{p})$ denotes the intensity at point $\boldsymbol{p}$. This function evaluates the smallest radius around point $\boldsymbol{p}$ within which at least one point with contrast greater than or equal to c appears. Consider a ramp cross section with width w and contrast σ_g (see Fig. 1a). Also assume that the regions surrounding the ramp are of constant intensities I_1 and I_2, with $I_2 - I_1 = \sigma_g$. The intensity inside the ramp increases monotonically from I_1 to I_2. For a point $\boldsymbol{p}$ in the interior of the region, at a distance d from the closest point in the ramp, we have $f_{\boldsymbol{p}}(0) = 0$ (see Fig. 1b). As we increment c by a small amount, $f_{\boldsymbol{p}}(c)$ instantly jumps to d, since all the points within a radius of d have zero contrast with respect to $\boldsymbol{p}$. Thereafter, $f_{\boldsymbol{p}}(c)$ monotonically increases from d to $d + w$ as c goes from 0 to σ_g, [1] and then jumps to a high value. We thus have discontinuities at $c = 0$ and $c = \sigma_g$. Now consider a point $\boldsymbol{q}$ inside the ramp, a distance x from the farthest point inside the ramp along the ramp cross section (say $\boldsymbol{r}$), and let the contrast between $\boldsymbol{q}$ and $\boldsymbol{r}$ be a particular value $\sigma_g^{\boldsymbol{q}}$. For this point (see Fig. 1c), $f_{\boldsymbol{q}}(c)$ monotonically increases from 0 to x as c increases from 0 to $\sigma_g^{\boldsymbol{q}}$, with a discontinuity at $c = \sigma_g^{\boldsymbol{q}}$. For the case when the regions surrounding the ramp are not of constant intensity, but have gradual shading, or noise, the function $f_{\boldsymbol{p}}(c)$ has similar profile, except that instead of discontinuities, we have very high gradients.

Detecting ramp pixels. According to the discussion above, for ramp pixels, $f_{\boldsymbol{q}}(c)$ has low gradient for small values of c. For the pixels in the interior of regions, $f_{\boldsymbol{p}}(c)$ increases rapidly for small values of c (discontinuity). We use this property to distinguish between ramp and non ramp pixels.

Estimating ramp parameters. To get the final segmentation, we need to estimate the contrast associated with each

[1] Actually the profile is the inverse of the profile of the ramp.



(a) ramp discontinuity (b) $f_p(.)$ (c) $f_q(.)$ (d) $\hat{\sigma}(.)$ (e) $\hat{w}(.)$

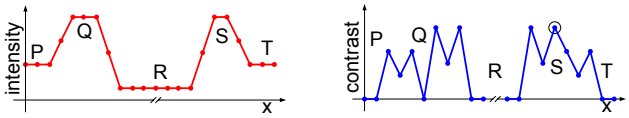
Figure 1. Region model (a) Intensity profile along the cross section of a ramp discontinuity of contrast σ_g and width w. Point $\boldsymbol{p}$ is outside the ramp, at a distance d from the closest ramp point. Point $\boldsymbol{q}$ is inside the ramp at a distance x and contrast σ_g^q from the farthest ramp point. **(b)** $f_{\boldsymbol{p}}(c)$ for point $\boldsymbol{p}$. **(c)** $f_{\boldsymbol{q}}(c)$ for point $\boldsymbol{q}$. **(d)** and **(e)** corresponding $\hat{\sigma_g}(.)$ and $\hat{w}(.)$ profiles respectively.

ramp, since segmentation at a photometric scale σ_g should only consider ramp pixels with contrast greater than σ_g. For each of the estimated ramp pixels, let us define the value of lowest value contrast at which a discontinuity in $f_{\boldsymbol{q}}(c)$ occurs as $\hat{\sigma}_g(\boldsymbol{q})$, and the corresponding function value as $\hat{w}(\boldsymbol{q})$. From the discussion above, $\hat{\sigma}_g(\boldsymbol{q})$ represents the contrast between the ramp pixel q and the farthest pixel along the ramp cross section, inside the ramp, denoted by $\boldsymbol{r}$. Similarly, $\hat{w}(\boldsymbol{q})$ represents the distance between the points $\boldsymbol{q}$ and $\boldsymbol{r}$. Assuming that these two functions take zero values in the interior of regions, the profiles of $\hat{\sigma}$ and $\hat{w}$ for the ramp in Fig. 1a are shown in Fig. 1d and Fig. 1e respectively. Thus, both profiles take maximum values at the boundary of the ramp, where the values are equal to the true contrast $\hat{\sigma}_g(\boldsymbol{q}) = \sigma_g$ and the true width $\hat{w}_g(\boldsymbol{q}) = w$. Starting from the boundary of the ramp, as we go closer to the center of the ramp, both functions decrease to a local minimum. All pixels inside the ramp should be assigned true contrast of the ramp σ_g and true width w. To estimate the true contrast, we first find the cross section of the ramp as the direction of the gradient in the original image. We then search for the closest local maximum of $\hat{\sigma}$ along the estimated cross section. We compute the ramp width in the same fashion.

We thus divide the pixels in image I into two sets: a) P, the pixels inside homogenous regions, and, b) Q, the pixels inside ramp discontinuities. For the pixels in set Q, we find the contrast Σ_g and width W of their respective ramp cross section. Next we describe the algorithm to achieve segmentation at a given photometric scale.

3.2 Algorithm

Segmentation at a photometric scale σ_g is a partition of an image into regions with contrast greater than or equal to σ_g. Thus all the ramps at region boundaries of the segmentation should have a contrast greater than σ_g. Let us define the set $Q(\sigma_g) = \{\boldsymbol{q} \in Q | \Sigma_g(\boldsymbol{q}) \geq \sigma_g\}$, i.e. the set of all ramp pixels with contrast greater than or equal to σ_g. This set can be divided into: a) Pixels inside ramps constituting the segmentation boundaries, and, b) Pixels inside ramps not on the segmentation boundaries. Pixels of type (b) do not have any isolated connected region inside



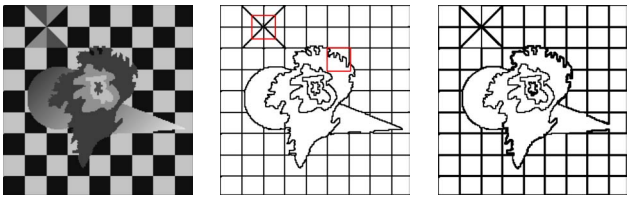
(a) intensity profile (b) $\hat{\sigma}_g(.)$ profile

Figure 2. Thin regions (a) Intensity along a cross section encompassing regions P, Q, R, S and T. Pixels are represented by dots along the profile. Region Q is three pixel wide, S is two pixel wide, and all other regions are wider than three pixels. **(b)** Estimated profile for $\hat{\sigma}_g(.)$ for the profile shown in Fig. 2a.

them, since if they do, then by definition of segmentation, this isolated connected region is a part of segmentation, implying that these pixels are on segmentation boundaries, i.e. of type (a). Thus, all isolated regions formed by connected components of the non-ramp pixels are part of the segmentation. We next discuss detection of thin regions, which are not identified in these connected components.

Thin Regions. Fig. 2a shows a sample intensity cross section passing through regions P, Q, R, S and T. Here, Q is three pixel wide, S is two pixel wide, and, P, R and T are wider than three pixels. It is assumed that the ramps between $P-Q$, $Q-R$, $R-S$, and $S-T$ have the same cross section, the one shown in the figure. Fig. 2b shows the corresponding profile for $\hat{\sigma}_g(.)$. The pixels at the periphery of the regions have non zeros values for $\hat{\sigma}_g$, and thus are labelled as ramp pixels. Regions P, Q, R, and T have a unique connected component associated with them due to the detected non ramp pixel. This is not true for region S, which does not have any non-ramp pixel inside it. Note however that $\hat{\sigma}_g$ has a local maximum inside region S which is not the case for any of the other wider regions. This fact can be used to detect thin regions. We put a seed for a new region at local maxima which completely lie inside the ramp area, i.e. the encircled point in S (Fig. 2b) is declared as a non ramp point.

Labelling of the connected components of non-ramp pixels thus achieved gives us seeds for all regions present in segmentation at photometric scale σ_g. These seeds are grown into the ramp pixels to achieve an initial segmen-



(a) (b) (c)

Figure 3. Segmentation results for a synthetic image with blurred edges. **(a)** Original image. **(b)** Result for $\sigma_g = 5$ obtained by our algorithm. Details for the regions marked with red squares are shown in Fig. 4. **(c)** Result of mean-shift algorithm for color bandwidth $\sigma_g = 5$.

tation. We next discuss the attainment of final segmentation by evolution of this intermediate segmentation using attraction forces.

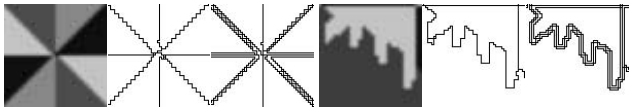
Evolution to achieve final segmentation. We want the final segmentation boundary to be at the position maximally separated from the regions surrounding it, i.e. associating a ramp pixel with a surrounding region with intensity closest to the intensity of the ramp pixel. From the analysis in Sec. 3.1, the contrast between a point $\boldsymbol{q}$ inside the ramp, and the closest intensity region is given by $\sigma_g = \sigma_g(\boldsymbol{q}) - \hat{\sigma}_g(\boldsymbol{q})$ and the distance is given by $\sigma_s = w(\boldsymbol{q}) - \hat{w}(\boldsymbol{q})$. The force $\boldsymbol{F_p}(\sigma_g, \sigma_s)$ thus will always be directed towards the region of closest intensity. We use this force to evolve the segmentation to get the final segmentation. Note that this force has to be computed just once for each pixel, irrespective of the scale at which segmentation is performed. The segmentation algorithm is summarized in Algorithm 1.

Algorithm 1: *Segmentation*

1. Estimate the ramp pixel set Q, respective contrasts Σ_g, and widths W;
2. Compute forces $\boldsymbol{F_q}$ at the ramp pixels $\boldsymbol{q}$;
3. **for** $\sigma_g = \sigma_g^{min} : \Delta\sigma_g : \sigma_g^{max}$ **do**
4. Compute the ramp pixels for contrast greater than σ_g, i.e., $Q(\sigma_g)$;
5. Detect seeds for thin regions;
6. Grow the connected components of non ramp pixels to get initial segmentation;
7. Evolve the initial segmentation according to the forces computed in Step 2 to achieve the final segmentation $F(\sigma_g)$;
8. **end**

4 Experiments and Results

In this section we test the following aspects of the algorithm: (1) Detection of complex topological boundaries; (2) Detection under blur; (3) Disambiguation between a shaded region and a blurred edge; and (4) Results on real images.



(a) (b)

Figure 4. Details of results in Fig. 3 for regions shown in red rectangles in Fig. 3b. Each of **(a)** and **(b)** contains the original image in the left, results of our algorithm in the center and result of the mean-shift algorithm on the right.

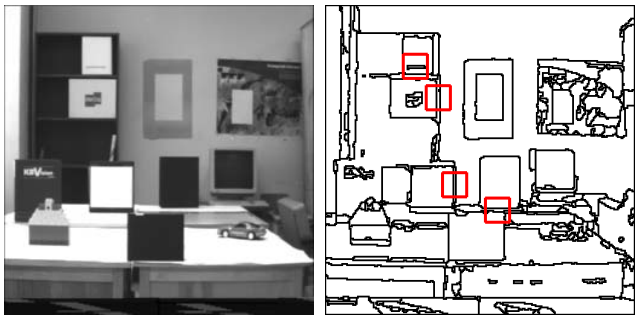


Figure 5. Detection of thin regions. **Left**: Original image. **Right**: segmentation for $\sigma_g = 10$. Note the detection of thin regions marked in red rectangles.

Fig. 3 shows the results on a synthetic image. The original image shown in the Fig. 3a consists of regions placed over a checkerboard pattern. The map shaped region in the middle has complex jagged boundaries. The cone shaped region extending from left to right has a constant gradient, emulating shading. All other regions are of constant gray level. Four checkerboard squares in the top left of the image are diagonally broken into eight regions, meeting at a single point, creating a complex topological corner. All the edges between regions are blurred using a 3×3 averaging filter, creating ramp discontinuities. The segmentation result of the algorithm discussed in Sec. 3.2 for $\sigma_g = 5$ is shown in Fig. 3b. Fig. 3c shows result of mean-shift segmentation [2] with intensity bandwidth for clustering $\sigma_g = 5$. For visualization, the edges are dilated to two pixel boundaries. For both algorithms, all the regions are recovered. Mean-shift however returns multiple regions inside the ramp discontinuities, due to which many edges seem wider. Fig. 4 shows detailed comparison between the two algorithms in regions shown in red in Fig. 3b. Fig. 4a (and similarly b) shows the original patch on the left, followed by the result of our algorithm, and mean-shift. As one can see, there is considerable blur on the diagonal edges. Since there is no explicit analysis of ramp discontinuities in the mean-shift algorithm, it perceives the blurred edges as multiple regions of constant intensity (Fig. 4a right), and hence multiple regions appear inside the ramp. Our algorithm (Fig. 4a center) however gives much better result, with no false regions inside the ramp. All edges are in the center of the ramps and the dif-

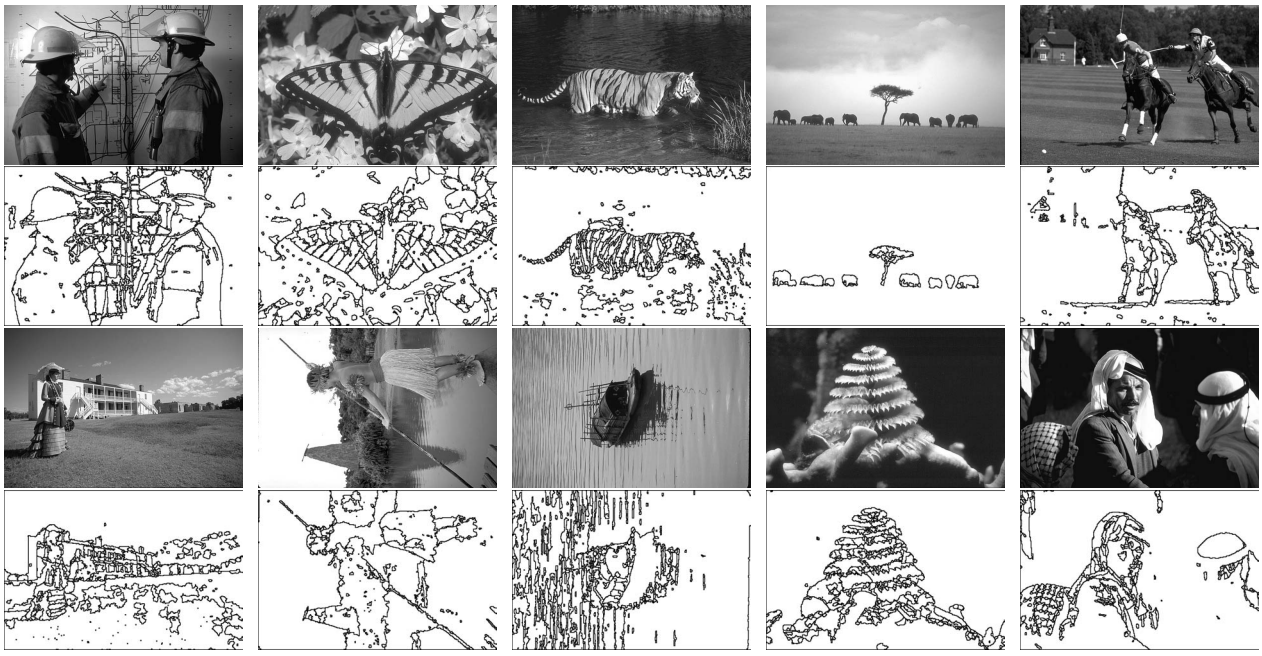


Figure 6. Segmentation results on **Berkeley images** [7]. Segmentation is obtained for photometric scale $\sigma_g = 20$.

ficult topological boundary is correctly recovered. Fig. 4b shows results for thin (single pixel or two pixel) protruding regions. The mean-shift algorithm again exhibits multiple regions inside the discontinuity, whereas our algorithm correctly identifies the regions.

Fig. 5 shows the detection of thin regions. The original image is shown on the left. The image on the right shows the segmentation for $\sigma_g = 10$. Note the thin regions, marked by red boxes are detected by our algorithm. All of these thin regions are less than 3 pixel in width.

Fig. 6 shows results on images from the Berkeley segmentation dataset [7]. The results are shown for $\sigma_g = 20$. Any edge with contrast lower than 20 is considered to be a noisy edge and ignored. Very thin (one or two pixel wide) regions like the roadmap layout in the left image, stripes on the tiger, and trunk of the tree in the elephant image, are correctly recovered. There are no multiple edges inside the ramps. The images have considerably blurred edges, as well as shading from which the region boundaries are correctly recovered.

5 Conclusions

In this paper, we presented an algorithm for image segmentation using a ramp discontinuity model. The algorithm has a multiscale nature in the sense that a segmentation for higher photometric scales is formed by strict merger of regions at lower scales, a property not experimentally verified in this work. This property suggests for a mechanism for constructing a segmentation tree of an image. In our future work, we intend to analyze an image in terms of such a segmentation tree.

Acknowledgment

The support of the Office of Naval Research under grant N00014-06-1-0101 is gratefully acknowledged.

References

[1] N. Ahuja. A transform for multiscale image segmentation by integrated edge and region detection. *IEEE TPAMI*, 18(12):1211–1235, 1996. 1, 2

[2] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. PAMI*, 24:603–619, 2002. 1, 4

[3] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. PAMI*, 6(11):721–741, 1984. 1

[4] D. Mumford and J. Shah. Boundary detection by minimizing functionals, i. In *CVPR*, pages 22–26, 1985. 1

[5] A. Nazif and M. Levine. Morphological segmentation. In *Proc. 2nd Int. Joint Conf. Pattern Recog.*, pages 555–577, 1974. 1

[6] P. Salembier. Morphological segmentation. *JVCIP*, 1(1):21–46, 1990. 1

[7] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. PAMI*, 22(8):888–905, 2000. 1, 5

[8] M. Tabb and N. Ahuja. Multiscale image segmentation by integrated edge and region detection. *IEEE Trans. Image Processing*, 6(5):642–655, 1997. 1

[9] S. Zucker. Survey region growing: Childhood and adolescence. *Comp. Vis. Graph. Image Proc.*, 5(4):382–399, 1976. 1