Learning a Classification Model for Segmentation

Xiaofeng Ren and Jitendra Malik
Computer Science Division
University of California at Berkeley, Berkeley, CA 94720
{xren,malik}@cs.berkeley.edu

Abstract

We propose a two-class classification model for grouping. Human segmented natural images are used as positive examples. Negative examples of grouping are constructed by randomly matching human segmentations and images. In a preprocessing stage an image is oversegmented into superpixels. We define a variety of features derived from the classical Gestalt cues, including contour, texture, brightness and good continuation. Information-theoretic analysis is applied to evaluate the power of these grouping cues. We train a linear classifier to combine these features. To demonstrate the power of the classification model, a simple algorithm is used to randomly search for good segmentations. Results are shown on a wide range of images.

1. Introduction

Perceptual grouping can be formulated as an optimization problem in a number of different frameworks, such as graph partitioning [22, 16, 8, 5] or variational approaches[13]. The objective function being optimized is typically driven by the designer's intuition or computational convenience. The theme of this paper is to derive the "right" optimization criterion. This is done in a learning approach using a database of human segmented images.

We formulate the computational problem of segmentation as a classification between "good" segmentations and "bad" segmentations. Figure 1 illustrates our basic approach. Figure 1 (a) is an image from the Corel Imagebase, (b) shows the image superimposed with a human marked segmentation, and (c) is the same image with a "wrong" segmentation. Our intuition tells us that the segmentation in (b) is "good" and the one in (c) is "bad".

How do we distinguish good segmentations from bad segmentations? Classical Gestalt theory has developed various principles of grouping [25, 14] such as proximity, similarity and good continuation. The principle of good continuation states that a good segmentation should have smooth boundaries. The principle of similarity is twofold:

(a) (b) (c)

Figure 1. We formulate segmentation as classification between good segmentations (b) and bad segmentations (c). We use Gestalt grouping cues as features and train a classifier. Human segmented images are used as examples of good segmentations. Bad segmentations are constructed by randomly matching a human segmentation to a different image.

1. *intra-region similarity*: the elements in a region are similar. This includes similar brightness, similar texture, and low contour energy inside the region;
2. *inter-region (dis)similarity*: the elements in different regions are dissimilar. This in turn includes dissimilar brightness, dissimilar texture, and high contour energy on region boundaries.

These classical principles of grouping have inspired many previous approaches to segmentation. However, the Gestalt principles are *ceteris paribus rules*, which means that they distinguish competing segmentations only when everything else is equal. Many of the previous works have made ad-hoc decisions for using and combining these cues.

In this work, we learn a classification model for segmentation from these Gestalt cues. A database of human marked segmentations has been established [10]. We use the human segmented images in this database as positive examples. For negative examples, we randomly match a human segmentation to a different image, an example of which has been given in Figure 1 (c).

The outline of the paper is as follows. In § 2 we introduce a preprocessing stage which organizes an image into "superpixels". In § 3, we define a set of features for segments, including Gestalt cues of contour, texture, brightness and good continuation. These features are evaluated using information-theoretic measures. From these features we train a logistic regression classifier. Based on this model for segments, in § 4 we formulate segmentation as an optimization problem of a linear objective function over the space of segmentations. To demonstrate the power of our classification model, we design a simple algorithm to randomly search for good segmentations. The experimental results are shown in § 5 . § 6 discusses related works and concludes the paper.

2. Oversegmentation as Preprocessing

In this section we present a preprocessing stage to group pixels into "superpixels". The motivations of this preliminary grouping are: (1) pixels are not natural entities; they are merely a consequence of the discrete representation of images; and (2) the number of pixels is high even at moderate resolutions; this makes optimization on the level of pixels intractable. We would like to work with "superpixels' which are local, coherent, and which preserve most of the structure necessary for segmentation at the scale of interest.

We apply the Normalized Cuts algorithm [22, 8] to produce the superpixel map. Both contour and texture cues are used. The affinity matrix has local connections only. Figure 2 shows an example of the oversegmentation with the number of superpixels $k = 200$. We observe from this example that the superpixels are roughly homogeneous in size and shape; this fact simplifies the computation in later stages. Some structures in the human segmentation are lost, but they are usually minor details, much smaller in scale than the objects we are interested in. The reconstructed segmentation is a good approximation of the original one.

To quantify the quality of this approximation, we use a contour-based measure to verify the superpixel maps against the human segmentations. In particular, we compute the percentage of the human marked boundaries being "covered by" (within a fixed small distance of) the superpixel boundaries. This is the *recall* rate used in [9]. Figure 3 shows the results for images of size 240-by-160. As expected, the recall rates increase with k. In our experiments we have found that $k = 200$ is sufficient.

3. What is a Good Segment?

A segmentation is a collection of segments. To answer the question "What is a good segmentation?", we need to answer "What is a good segment?" first. In this section, we will define a set of features for segments, evaluate the usefulness of these features, and train a classifier from them.

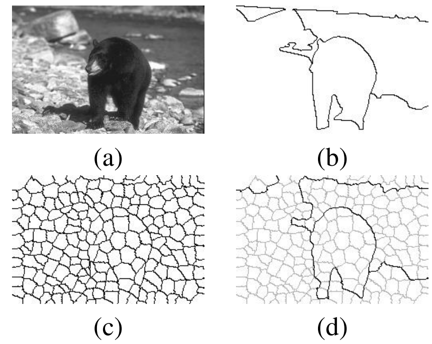


Figure 2. An example of superpixel maps. (a) is the original image; (b) is a human marked segmentation; (c) is a superpixel map with $k = 200$; (d) shows a reconstruction of the human segmentation from the superpixels: we assign each superpixel to a segment in (b) with the maximum overlapping area and extract the superpixel boundaries.

3.1. Features for grouping

For static images, the classical Gestalt principles of grouping include *proximity*, *similarity*, *good continuation* (curvilinear continuity), *closure* as well as *symmetry* and *parallelism*. In our model, for a segment S we define the following features $\mathbf{F}_j$:

1. inter-region texture similarity $\mathbf{T}_{ext}(S)$;
2. intra-region texture similarity $\mathbf{T}_{int}(S)$;
3. inter-region brightness similarity $\mathbf{B}_{ext}(S)$;
4. intra-region brightness similarity $\mathbf{B}_{int}(S)$;
5. inter-region contour energy $\mathbf{E}_{ext}(S)$;
6. intra-region contour energy $\mathbf{E}_{int}(S)$;
7. curvilinear continuity $\mathbf{C}(S)$.

Texture Similarity

For texture cues we follow the discriminative framework of texton analysis (e.g., [8]). The image is first convolved with a bank of filters of multiple orientations. Based on a vector quantization of the filter outputs, the pixels are clustered into a number of texton channels. This gives us a descriptor for each region, namely the distribution of textons inside its support. The texture difference of two regions is then measured as the χ^2 distance between two histograms. We make one modification here in the filtering stage: when we apply the filterbank to the image, we restrict the support of the filters to be within a single superpixel (normalized convolution, e.g., [6]).

In the next step, we convert the χ^2 distance into a log likelihood ratio: let Ω_{same} denote the set of all superpixels

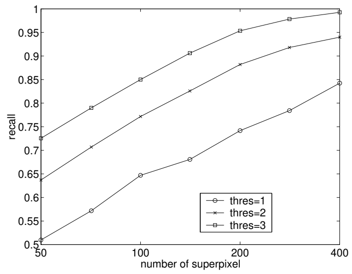
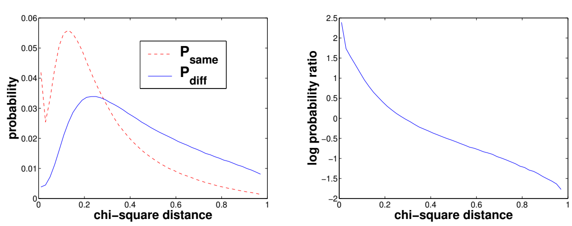


Figure 3. The percentage of human marked boundaries covered by the superpixel maps. The number of superpixels k varies from 50 to 400. Distance tolerance is set at 1, 2 and 3 pixels respectively, for images of size 240-by-160. For $k = 200$ and a tolerance of 2 pixels, approximately 90% of the human marked boundaries are covered.

pairs (q_1, q_2) such that they appear in the same segment of a human segmentation, and let Ω_{diff} denote the set of all superpixel pairs such that they appear in different segments in a human segmentation. We compute the χ^2 distance d_T for all pairs of superpixels in Ω_{same}, and denote the distribution of these χ^2 distances as P_{same}. Similarly we collect P_{diff}, the distribution of χ^2 distance on the set Ω_{diff} (see Figure 4 for the empirical distributions). Let $d_T(q, S)$ be the χ^2 distance between the texture histogram of a superpixel q and a segment S. The texture similarity between q and S is defined as:

$$T(q, S) = \log \frac{P_{\text{same}}\,(d_T(q, S))}{P_{\text{diff}}\,(d_T(q, S))}$$



(a) (b)

Figure 4. The empirical distributions of χ^2 distance of texture histograms. (a) P_{same} is between a pair of superpixels in the same human marked segment; P_{diff} is between a pair of superpixels in different segments. (b) shows the log likelihood ratio $\log\left(P_{same}/P_{diff}\right)$.

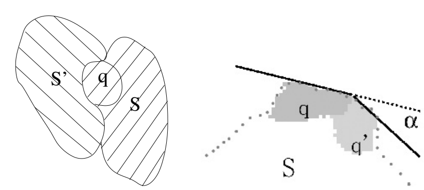
The log likelihood ratio measures the significance of the χ^2 value. We use this basic texture similarity measure $T(q, S)$ to define two texture features for a segment S, *intra-region* texture similarity and *inter-region* texture similarity. Figure 4 (a) illustrates the definition. The intra-region texture similarity sums over all the superpixels in the region:

$$\mathbf{T}_{int}(S) = \sum_{q \in S} T(q, S)$$

and the inter-region similarity sums over all the superpixels on ∂S, the boundary superpixels of S:

$$\mathbf{T}_{ext}(S) = \sum_{q \in \partial S} T(q, S'(q))$$

where $S'(q)$ is the segment adjacent to q. If there are multiple segments adjacent, we take the average of similarity values.



(a) (b)

Figure 5. (a) The intra-region similarity compares the descriptor of a superpixel q to the segment S containing it. The inter-region similarity compares the descriptor of a superpixel q on the boundary of S to the adjacent segment S'. (b) Curvilinear continuity of S is measured by the tangent changes at superpixel junctions along the boundary of S.

Brightness similarity

The intra-region brightness similarity $B_{int}(S)$ and inter-region brightness similarity $B_{ext}(S)$ are defined in an identical way. The brightness descriptor for each region is a histogram of brightness values. We compute the χ^2 distance of histograms and use empirical data to convert the χ^2 distance into a log likelihood ratio. This basic similarity measure is incorporated into the intra- and inter-region similarity cues.

Contour energy

Contour cues are computed at the level of pixels. We first compute the orientation energy OE [12, 8] at each pixel. The orientation energy is converted to a soft "contourness", p_{con}, by a non-linear transform [8, 18]. The inter-region contour energy $\mathbf{E}_{ext}$ is the summation of $p_{con}(x)$ over all the *pixels* on the boundary of S, and the intra-region contour energy $\mathbf{E}_{int}$ is the summation of $p_{con}(x)$ over all the pixels on the superpixel boundaries inside S.

Good continuation

Curvilinear continuity is measured as follows: for each adjacent pair of superpixels q and q' on the boundary of a segment S, there is a change of tangent at the junction (see Figure 4 (b)). This measures the first-order smoothness of the boundary: the larger this angle α, the less smooth the boundary of S. From the boundaries in the human segmentations, we collect the distribution $P_{\text{tangent}}(\alpha)$ of tangent changes. Let $J(S)$ be the set of all superpixel junctions on the boundary of S, the curvilinear continuity $\mathbf{C}(S)$ is defined as

$$\mathbf{C}(S) = \sum_{(q,q') \in J(S)} \log P_{\text{tangent}} \left(\alpha(q, q') \right)$$

Normalizing the features

The features we have defined are unnormalized and cannot be directly compared with one another. To normalize them, we notice that all the features we have defined are summations of basic quantities. For example, consider the intra-region texture similarity $\mathbf{T}_{int}(S) = \sum_{q \in S} T(q, S)$. We assume that the $T(q, S)$'s are random variables with the same mean μ and variance σ^2 for all pairs (q, S) such that $q \in S$. If there are m superpixels in S, we normalize $T_{int}(S)$ as $\left(T_{int}(S) - m\mu\right) / \sqrt{m}\sigma$. The maximum likelihood estimates of μ and σ are used. Other features are normalized in the same way.

3.2. Power of the Gestalt cues

Before we train a classifier of segments from the features above, one interesting question is to ask how useful the grouping cues are. We conduct an information-theoretic analysis to measure the power of the cues in a model- and algorithm-independent way.

Each segment S is associated with a set of features $\{\mathbf{F_i}\}$ and a class label $\mathbf{h}$: if S is a segment from a good segmentation, $\mathbf{h} = \mathbf{1}$; if S is from a bad segmentation, $\mathbf{h} = \mathbf{0}$. From the datasets we collect the joint distribution of $\mathbf{h}$ and the features. For any feature $\mathbf{F_i}$, we compute the mutual information $I(\mathbf{h}, \mathbf{F_i})$. This is the amount of information contained in $\mathbf{F_i}$ about the classification $\mathbf{h}$. The distributions are normalized and the marginal entropy of $\mathbf{h}$ is 1.0 (bits). The first column of Table 1 (a) shows the results for individual features. We also combine each pair of inter- and intra-features together to evaluate the overall power of contour, texture, and brightness cues. These results are listed in the first column of Table 1 (b).

From this analysis of mutual information we find that the presence of boundary contours is the most informative grouping cue. The texture cues and brightness cues are approximately equally informative. The intra-region cues by themselves are usually non-informative. Combined with inter-region cues, however, they make significant contributions. Curvilinear continuity turns out to be a powerful cue in our analysis. The power of continuity is revealed due to the way we construct the dataset of bad segmentations (see the introduction). Because a randomly assigned segmentation often disagrees with the superpixel map, which is constructed from the image, the resulting boundaries are jig-jaggy, poor in continuity.

One interesting question to ask is whether the normalized convolution based on the superpixel map is helpful for grouping. We repeat the texture analysis with standard filtering, i.e., not making use of the superpixel masks. Information-theoretic analysis shows that the joint information of inter-region texture and inter-region brightness cues, $I(\{T_{ext}, B_{ext}\}; \mathbf{h})$, increases from 0.170 to 0.188. This result suggests that a reasonable support mask of image regions does help with texture analysis, if texture and brightness are used simultaneously.

3.3. Training the classifier

We have formulated the problem of segmentation as a two-class classification. This is one of the most well-studied problems in the statistical learning and we have a variety of techniques at our disposal. We use a simple logistic regression classifier, which linearly combines the features $\mathbf{F}_j$:

$$\mathbf{G}(S) = \sum_j c_j \mathbf{F}_j(S) - \theta \tag{1}$$

The higher the value of G is, the more likely S is a good segment. The weights c_j are easily learned by maximizing the likelihood on training data with the standard *iterative reweighted least squares* algorithm [4]. 160345 segments are used as training data and 80178 segments as test data. The initialization is random and the convergence is fast. For intra-region features, the weights are negative.

To gain more insights into this combination of features, we examine the empirical distributions more closely. We collect the joint distributions of a pair of features, both on the positive examples and the negative examples, to see if they are linearly separable. Figure 6 shows contour plots of two examples of the empirical density functions. We have found that the normalized features are roughly Gaussian distributed and a linear classifier fits the data well. (The authors of [9] also reported that for vision data of low dimension and poor separability, logistic regression performs as well as other sophisticated techniques.)

One way to evaluate our model is to look at the final segmentation results, which we present in Section 5. Information-theoretic analysis again provides us an alternative way of evaluation. Ideally, we would like our model to capture all the information contained in the grouping cues.

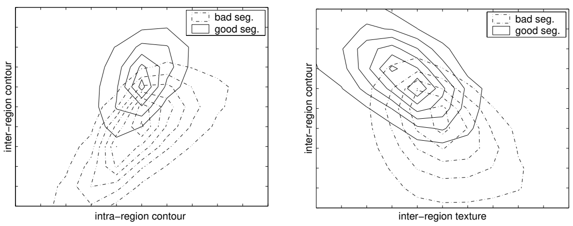


Figure 6. Iso-probability contour plots of empirical distributions for a pair of features. These plots suggest that: (1) the normalized features are well-behaved; for both classes a Gaussian model would be a reasonable approximation. And (2) a linear classifier would perform well.

That is, the label $\mathbf{h}$ and the features $\mathbf{F}$ would be conditionally independent given the output of the model $\mathbf{G}$. The residual information is measured by the mutual information of $\mathbf{h}$ and $\mathbf{F}$ conditioned on $\mathbf{G}$. The results have been listed in the second columns of Table 1. We observe that there is little residual information left in the features, which indicates that the linear classifier fits the data well.

To further evaluate our classification model, we use the precision-recall framework [19, 2]. *Precision* is the fraction of detections which are true positives. *Recall* is the fraction of true positives which are detected. Figure 7 shows the precision-recall curves for three cases; the results are almost identical. This suggests that for this problem (1) the logistic regression model generalizes well; and (2) sophisticated classification techniques may not outperform the simple linear model.

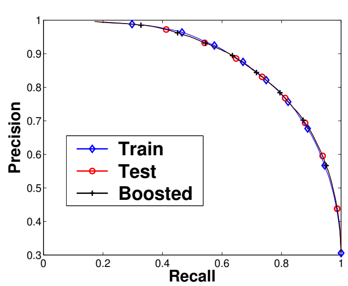


Figure 7. Precision-recall analysis for (1) simple logistic classifier on training data; (2) simple logistic classifier on test data; and (3) boosted logistic classifier on test data. The three curves are almost identical.

4. Finding Good Segmentations

"What is a good segmentation?" We make the simplifying assumption that the "goodness" of the segments S in a segmentation $\mathcal{S}$ are independent. This leads us to the following criterion:

Feature		Information	Residual Info.
Contour:	inter-	0.387	0.010
	intra-	0.012	0.010
Texture:	inter-	0.137	0.005
	intra-	0.030	0.008
Brightness:	inter-	0.112	0.005
	intra-	0.049	0.007
Continuity:		0.198	0.002

(a)

Combined Feature	Information	Residual Info.
Contour	0.510	0.024
Texture	0.220	0.026
Brightness	0.232	0.025

(b)

Table 1. Information-theoretic analysis of grouping cues. (a) shows the results for individual features. (b) shows the results when pairs of intra- and inter- region cues are combined. The first column is the amount of information these features contain about the class label. The second column is the amount of residual information these features retain when conditioned on the model output. The marginal entropy of the class label is 1.0 (bits).

$$f(\mathcal{S}) = \sum_{S \in \mathcal{S}} \left(\sum_j c_j \mathbf{F}_j(S) - \theta \right) \tag{2}$$

which sums the classifier function in Eqn (1) over the segments. The problem of finding the best segmentation becomes the optimization of f in the space of all segmentations.

The objective f is simple in form but the search space of all segmentations is large. Following the Markov Chain Monte Carlo paradigm [3, 24], we adopt a simple strategy of random search based on simulated annealing.

The dynamics in this random search involves three basic moves: (1) *shift*: a superpixel is shifted from its segment to an adjacent segment; (2) *merge*: two adjacent segments are merged into one; and (3) *split*: a segment is split into two. The first two moves are straightforward. For splitting a segment, we use a simple method by clustering the superpixels in the segment based on location and mean intensity. This clustering is also used to initialize the search.

At each step, the algorithm randomly picks one of the moves above and construct a new segmentation $\mathcal{S}'$. If

$f(\mathcal{S}') > f(\mathcal{S})$, we accept the move. Otherwise, we accept with probability $\exp\left(\frac{f(\mathcal{S}')-f(\mathcal{S})}{T}\right)$ where T is the temperature, decreasing linearly over time.

The algorithm is naive; nevertheless it demonstrates the power of our classification model. There exist many other possible ways of exploring the space of segmentations.

5. Experimental Results

Figure 9 shows some results of our algorithm on images from the Corel Imagebase. The images are all gray-scale and of size 240-by-160. In our current implementation, the random search itself takes about 15 to 30 minutes on a Pentium III 550 Hz processor.

We have found that the segmentations are biased toward small regions. One reason is that our objective function f is a simple sum over individual segments and is not normalized w.r.t. the number of segments. To segment an image into approximately equally sized regions, and also to provide a degree of user control over the scale of the segmentation, we introduce a prior distribution on segment size $|S|$. Figure 8 shows the empirical distribution of $|S|$ in the human segmented images. We approximate this prior with a log-normal distribution, and the objective function becomes:

$$\tilde{f}(\mathcal{S}) = f(\mathcal{S}) + \sum_{S \in \mathcal{S}} \left(-\frac{(\log |S| - \mu_s)^2}{2\sigma_s^2} \right) \qquad (3)$$

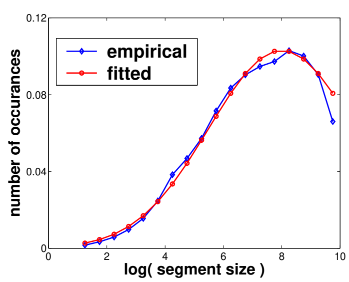


Figure 8. The empirical distribution of (the logarithm of) segment size $|S|$. Both extremely small and large segments are uncommon. We approximate this prior on $|S|$ with a log-normal distribution.

Our naive algorithm does occasionally suffer from the problem of local maxima. One reason lies in the asymmetry of the simple dynamics: merging is trivial but splitting is hard. To obtain the final results we present in Figure 10, the algorithm is allowed to run for three times and the solution with the best $\tilde{f}$ value is picked. A weak prior with $\mu_s = \log 20$ (20 superpixels) and $\sigma_s = 2.0$ is used.

6. Discussions

In this paper we have presented a discriminative framework for segmentation as the classification of "good" segmentations and "bad" segmentations. The Gestalt grouping cues are combined in a principled way and we have empirically measured the power of these cues. A linear classifier and a simple random search algorithm have produced promising results on a variety of natural images.

This work is motivated in part by both the Normalized Cuts [22, 8] and the DDMCMC work [24]. Our approach is in the discriminative paradigm as the Normalized Cuts. The basic contour and texture cues in our model is similar to those in [8]. However, our approach differs from [22] in two important aspects. First, we have constructed our model on the level of segments. This has enabled us to (1) define relationships between the parts and the whole; (2) to easily incorporate mid-level cues such as good continuation, and, (3) instead of relying on intervening contours [7], we use the contour cues in a straightforward way. Second, we have formulated grouping as a two-class classification problem. This framework connects the segmentation as a computational problem with the ecological statistics of natural images [9, 2] and the rich theory of learning.

The Normalized Cuts criterion is driven by computational convenience, and does not have a clear statistical interpretation. The random walk formulation of Normalized Cuts [11] defines segmentation as a one-class problem and has only been applied to special classes of images. The Normalized Cuts criterion does lead to a computationally tractable problem of spectral clustering. Our framework, on the other hand, has to solve a difficult optimization in the space of all segmentations.

The DDMCMC [24] is a generative approach which builds explicit models of image regions. The DDMCMC framework faces a computational challenge similar to ours. The main difference in philosophy is discriminative vs. generative. Solving an easier problem of discrimination, we are able to succeed with a linear classifier and a naive search algorithm. As we have found out, boundary contour is the most informative grouping cue, and it is in essence discriminative. Such contour cues are used indirectly in [24].

Acknowledgments. This research was supported by NSF through a Digital Library Grant IRI-9411334.

References

[1] E. Borenstein and S. Ullman. Class-specific, top-down segmentation. In *ECCV '02*, volume 2, pages 109–124, 2002.

[2] C. Fowlkes, D. Martin, and J. Malik. Learning affinity functions for image segmentation: combining patch-based and gradient-based approaches. In *CVPR '03*, volume 2, pages 54–61, 2003.

Figure 9. Segmentation results with no prior.

[3] S. Geman and D. Geman. Stochastic relaxation, gibbs distribution, and the bayesian retoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–41, Nov. 1984.

[4] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer-Verlag, 2001.

[5] I. Jermyn and H. Ishikawa. Globally optimal regions and boundaries. In *ICCV '99*, pages 904–910, 1999.

[6] H. Knutsson and C.-F. Westin. Normalized and differential convolution: methods for interpolation and filtering of incomplete and uncertain data. In *CVPR '93*, pages 515–523, 1993.

[7] T. Leung and J. Malik. Contour continuity in region-based image segmentation. In *ECCV '98*, volume 1, pages 544–59, 1998.

[8] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *Int'l. Journal of Computer Vision*, 43(1):7–27, June 2001.

[9] D. Martin, C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using brightness and texture. In *NIPS '02*, 2002.

[10] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV '01*, volume 2, pages 416–423, 2001.

[11] M. Meila and J. Shi. Learning segmentation by random walks. In *NIPS '00*, pages 873–879, 2000.

[12] M. Morrone and R. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–13, 1987.

[13] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions, and associated variational problems. *Comm. Pure Math.*, pages 577–684, 1989.

[14] S. Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, 1999.

[15] P. Parent and S. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(8):823–39, Aug. 1989.

[16] P. Perona and W. T. Freeman. A factorization approach to grouping. In *ECCV '98*, pages 655–670, 1998.

[17] J. Puzicha, T. Hofmann, and J. Buhmann. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. In *CVPR '97*, pages 267–272, 1997.

[18] X. Ren and J. Malik. A probabilistic multi-scale model for contour completion based on image statistics. In *ECCV '02*, volume 1, pages 312–327, 2002.

[19] C. V. Rijsbergen. *Information Retrieval, 2nd ed.* Dept. of Comp. Sci., Univ. of Glasgow, 1979.

[20] S. Sarkar and K. Boyer. Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors. In *CVPR '96*, pages 478–483, 1996.

[21] R. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.

[22] J. Shi and J. Malik. Normalized cuts and image segmentation. In *CVPR '97*, pages 731–7, 1997.

[23] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):888–905, August 2000.

[24] Z. Tu, S. Zhu, and H. Shum. Image segmentation by data driven markov chain monte carlo. In *ICCV '01*, volume 2, pages 131–138, 2001.

[25] M. Wertheimer. Laws of organization in perceptual forms (partial translation). In W. Ellis, editor, *A sourcebook of Gestalt Psychology*, pages 71–88. Harcourt Brace and Company, 1938.

[26] L. Williams and D. Jacobs. Stochastic completion fields: a neural model of illusory contour shape and salience. In *ICCV '95*, pages 408–15, 1995.

[27] S. Yu, R. Gross, and J. Shi. Concurrent object segmentation and recognition with graph partitioning. In *NIPS '02*, 2002.

[28] S. C. Zhu and A. L. Yuille. Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.

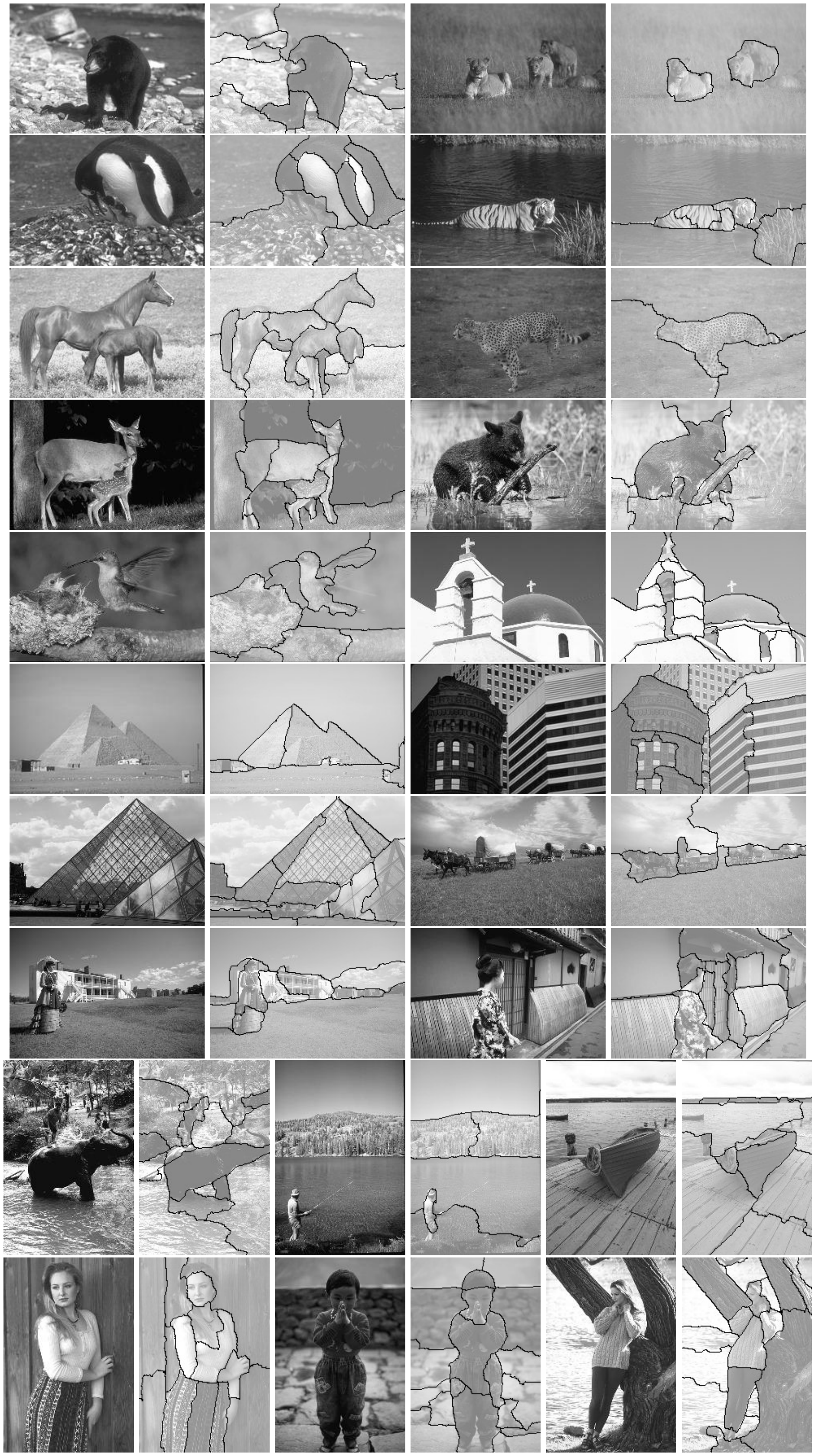


Figure 10. Segmentation results with the log normal prior on segment size.