Figure/Ground Assignment in Natural Images

Xiaofeng Ren, Charless C. Fowlkes, and Jitendra Malik

Computer Science Division,
University of California at Berkeley,
Berkeley, CA 94720, USA

Abstract. Figure/ground assignment is a key step in perceptual organization which assigns contours to one of the two abutting regions, providing information about occlusion and allowing high-level processing to focus on non-accidental shapes of figural regions. In this paper, we develop a computational model for figure/ground assignment in complex natural scenes. We utilize a large dataset of images annotated with human-marked segmentations and figure/ground labels for training and quantitative evaluation.

We operationalize the concept of *familiar configuration* by constructing prototypical local shapes, i.e. *shapemes*, from image data. Shapemes automatically encode mid-level visual cues to figure/ground assignment such as convexity and parallelism. Based on the shapeme representation, we train a logistic classifier to locally predict figure/ground labels. We also consider a global model using a *conditional random field* (CRF) to enforce global figure/ground consistency at T-junctions. We use loopy belief propagation to perform approximate inference on this model and learn maximum likelihood parameters from ground-truth labels.

We find that the local shapeme model achieves an accuracy of 64% in predicting the correct figural assignment. This compares favorably to previous studies using classical figure/ground cues [1]. We evaluate the global model using either a set of contours extracted from a low-level edge detector or the set of contours given by human segmentations. The global CRF model significantly improves the performance over the local model, most notably when using human-marked boundaries (78%). These promising experimental results show that this is a feasible approach to bottom-up figure/ground assignment in natural images.

1 Introduction

Figure/ground organization, as pioneered by Edgar Rubin [2], is a step of perceptual organization which assigns a contour to one of the two abutting regions. It is commonly thought to follow region segmentation, it is an essential step in forming our perception of surfaces, shapes and objects, as vividly demonstrated by the pictures in Figure 1. These pictures are highly ambiguous and we may perceive either side as the figure and "see" its shape. We always perceive the ground side as being shapeless and extended behind the figure, never seeing both shapes simultaneously.

Figure/ground organization is a classical topic in Gestalt psychology, and over the years many factors have been discovered which play a role in determining

A. Leonardis, H. Bischof, and A. Prinz (Eds.): ECCV 2006, Part II, LNCS 3952, pp. 614–627, 2006.

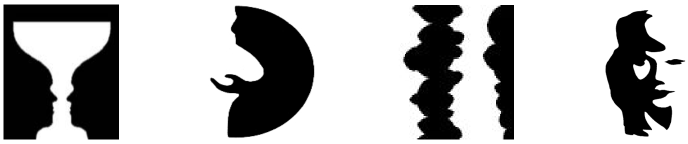


Fig. 1. The figure/ground assignment problem. We perceive that each boundary belongs to one, but not both, of the two abutting regions. The figure side has a "shape" and the ground side is "shapeless", extending behind the figure.

what regions are seen as figural [3]. The most important of these factors include *size*, *convexity*, *symmetry*, *parallelism*, *surroundedness* and *lower-region* as well as *familiar configuration*. Recent studies in psychophysics show that familiar configurations of contours provide a powerful cue for figure/ground [4], which often dominates more generic cues.

In computer vision, partly due to its lack of immediate applications, figure/ground organization has received little attention. Nevertheless, a few influential studies on figure/ground persist: many focusing on modeling and exploiting global structure such as T-junctions (e.g. [5, 6, 7, 8, 9]), others studying the use of local cues such as convexity (e.g. [10]). Typically such approaches have only been demonstrated on a limited set of images, mostly synthetic.

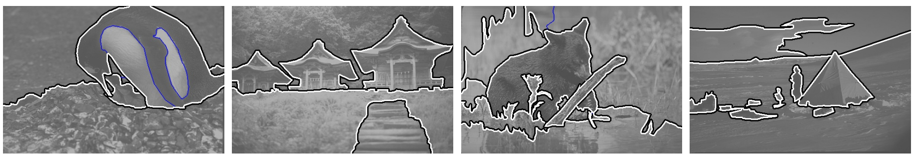


Fig. 2. Examples from the figure/ground dataset of natural scenes. Each image is first segmented by a human subject; then two human subjects assign figure/ground labels to each boundary in the segmentation. Here the white boundary indicates the figure side and black the ground side. Blue boundaries indicate contours labeled by subjects as not having a clear figure/ground assignment (e.g. surface markings).

In this work we utilize a large dataset of natural images where human subjects provide segmentations as well as figure/ground labels (the Berkeley Figure/Ground Dataset [1]). Figure 2 shows a few images from this dataset, each annotated with a segmentation and corresponding figure/ground labels. The purpose of this work is to address the challenges of figure/ground assignment in such complex natural scenes, in the presence of hierarchical object structure, arbitrary occlusion and texture as well as background clutter and imaging noise.

We propose a two-step approach: a local model using prototypical local shapes to represent context; and a global model using a random field to enforce consistency along contours and junctions. We train both models with human-marked groundtruth data and quantitatively evaluate their performance.

2 Figure/Ground Assignment in Natural Images

A standard view in perceptual organization is that images are first grouped into smooth contours, regions and junctions. Then each contour is assigned to one of the two abutting regions, after which shape analysis and object recognition happen. Recently this theory of sequential processing has been brought into question. Psychophysical experiments suggest that recognition of familiar contour configurations is a powerful figure/ground cue and may occur prior to figure/ground assignment [4]. On the other hand, studies from neurophysiology indicate that figure/ground organization may occur early in the visual pathway [11, 12], long before grouping is completed.

There is, of course, nothing contradictory between these findings and the traditional Gestalt emphasis on global processing. It could well be that informative cues (including familiar shape) are available in the local context of each contour independently extracted quite early. After this initial step, more global structure, such as T-junctions, may be constructed and used to enforce consistency between local figure/ground assignments.

This is the philosophy behind our approach, which is outlined in Figure 3. Starting from an image, first we compute its edge map using the **Pb** (Probability of Boundary) operator [13]. Then we use Geometric Blur [14], a local shape descriptor, to represent the local context around each image location. The representation is in terms of its similarity to set of prototypical local shapes, or **shapemes**, that we find in advance from clustering training data. These similarity terms are then combined using a linear classifier to predict the figure/ground label at each image location. We show that the shapeme-based classifier performs much better than a baseline model using size/convexity.

Next we develop a global figure/ground model which enforces labeling consistency at junctions. First we integrate local figure/ground cues over continuous

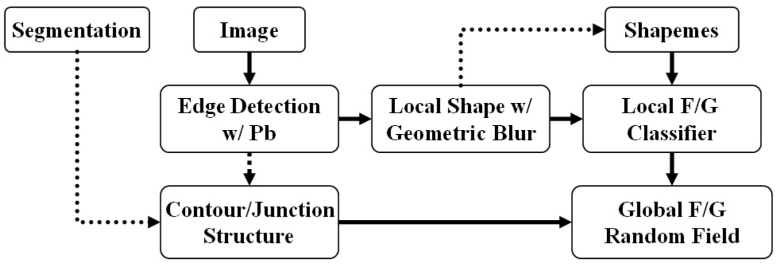


Fig. 3. Summary of our two-stage approach. First we use the **Pb** operator [13] to compute a soft edge map. The Pb map is used to compute local shape descriptors using Geometric Blur [14]. These shape descriptors are clustered into prototypical shapes, or shapemes, which encode rich mid-level visual information. Our local figure/ground model is a logistic classifier based on the shapeme representation. Our global figure/ground model uses a conditional random field to enforce global consistency by learning junction frequency and continuity. It operates either on a human-marked segmentation or thresholded Pb boundaries.

contour segments. We consider the following two cases separately: (1) if we assume that a segmentation is available, we obtain a contour/junction structure from a human-marked segmentation; or (2) if we don't assume to have a segmentation, a contour/junction structure is constructed from bottom-up based on thresholded Pb edges.

We use a *conditional random field* model [15, 16] to build a joint probabilistic model over the figure/ground labels on the complete set of contours segments. Empirical frequencies of junction types (such as valid or invalid T-junction labels), along with continuity of foreground contours, are exploited to correct locally ambiguous labelings. Inference is done with loopy belief propagation. We learn maximum likelihood model parameters with gradient descent.

We quantify the performance of our models by testing them against groundtruth labels. In the case of using human segmentations, each pixel on a human-marked boundary has a figure/ground label, and we count the percentage of figure/ground labels correctly predicted. In the case of bottom-up contour detection, we use the Canny's hysteresis to threshold Pb boundaries and apply a bipartite matching process to "assign" groundtruth labels to each pixel on the Pb boundaries. We then count the percentage of correct predictions of our models on these transferred labels.

3 Local Figure/Ground Model with Shapemes

Many of the classical figure/ground cues are *mid-level* cues. Unlike edge detection, which measures contrast at a point, visual cues such as convexity, parallelism and symmetry are about the relations between points or elements. On the other hand, these cues can still be estimated within a moderately sized neighborhood, without requiring a complete segmentation or recognition of objects.

Such mid-level cues are not trivial to operationalize. Parallelism, symmetry and convexity have precise mathematical definitions but models constructed from mathematical/geometric analysis are seldom flexible enough to cope with the variety of natural phenomenon including noise, texture and clutter. Another challenge with natural scenes is that they often contain multiple objects/parts and hence have a complex junction structure which is impossible to reliably detect using local operators[17].

3.1 Shapemes: Prototypical Shapes

Instead of seeking a mathematical definition for every local figure/ground cue, we take an empirical approach, using a generic shape descriptor to discover **shapemes**, or prototypical shapes, from data. This is in the spirit of Wertheimer's *familiar configuration* and Brunswik's *ecological theory* of Gestalt principles.

We use the Geometric Blur operator [14] to describe local shape. Let I be an input image and E an edge map. The *geometric blur* centered at location x, $GB_x(y)$, is a linear operator applied to E whose value is another image given by the "convolution" of E with a spatially varying Gaussian. GB_x has the property that points farther away from x are more blurred, making the descriptor robust

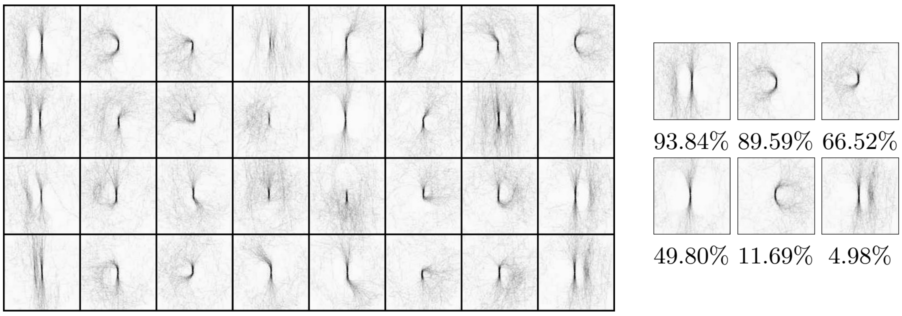


Fig. 4. Shapemes, or clusters of local shapes from a set of human-marked boundaries of baseball players. Shown here are the average shapes in each cluster. We find that shapemes encode rich contextual information, such as parallelism (row 1, col 1), convexity (row 1, col 2), sharp corners (row 2, col 3) or straight lines (row 2, col 5). On the right we show a few shapemes and the percentage of the shapes in each cluster that have the left side as figure. Empirical data confirm that mid-level cues such as parallelism or convexity are very useful for figure/ground assignment; figure/ground labelings are heavily biased in such cases.

to affine distortions. The value $GB_x(y)$ is the inner product of E with a Gaussian centered at y whose standard deviation is $\alpha|y-x|$. We rotate the blurred image GB_x so that the locally estimated contour orientation at x is always vertical. We choose $\alpha = 0.5$ and sample the blurred and rotated image GB_x at 4 different radii (increasing by a factor of $\sqrt{2}$) and 12 orientations, to obtain a feature vector of length 48.

We then cluster these Geometric Blur descriptors to find prototypical shapes, or shapemes. The use of shapemes was first introduced in [18] as a means to efficiently index and retrieve object specific shapes. Here we use shapemes in a rather different way, as a representation derived from data to capture mid-level cues. Our shapemes also differ in that they are orientation-independent, as we align them to local boundary orientations. This allows us to encode rich contextual information with a small set of shapemes.

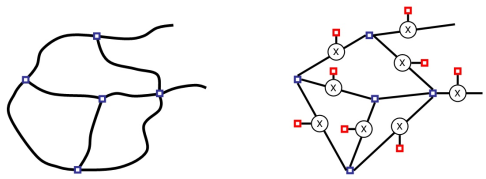
To illustrate the concept, Figure 4 visualizes 32 shapemes constructed (using k-means) from a simpler database containing silhouettes of baseball player photos [19]. We find that mid-level cues such as convexity and parallelism are implicitly captured in the shapemes, making it an appealing representation for figure/ground organization.

For experiments on the more complex Berkeley Figure/Ground Dataset, we use 64 shapemes constructed from Pb edge maps and modeled as a mixture of Gaussians. For each local shape, we use the mixture of Gaussian to obtain a feature vector f of dimension 64, which is the log posterior probability of each component mixture. We use these features to predict a binary label $Y \in \{-1, 1\}$ indicating which side is figure and which side is ground, a binary classification problem. A logistic classifier is fit to the human-marked labels using standard iteratively re-weighted least squares.

4 Global Figure/Ground Model with Conditional Random Fields

Although local shape is quite informative, figure/ground organization is not a local phenomenon. Contour form parts of object boundaries in the scene, and they interact through junctions and regions. One classical problem in the early days of computer vision is the labeling of line drawings. There, T-junctions are probably the most important cue in interpreting objects and scenes. Following this tradition, many previous studies focus on the global inference of figure/ground relations through junctions [5, 6, 7, 8, 9].

In the previous section we have shown that shapemes encode rich mid-level cues and can be used to construct a local model for figure/ground organization. To combine local cues and enforce global consistency, we assume that we have a discrete graph structure of the image, as shown in Figure 5(a), where edge pixels form contours and contours join to form junctions. This structure may either come from a human-marked segmentation or, as we will show in the next section, from a thresholded edge map.



(a) (b)

Fig. 5. Global inference of figure/ground assignments. Suppose we have a discrete contour/junction structure as in (a), which comes either from a human-marked segmentation or from thresholded Pb edge maps. We use a conditional random field to enforce global consistency of the figure/ground labels on individual edges. (b) shows the factor graph of our probabilistic model corresponding to the edge structure in (a). Edge potentials combine evidence from the local figure/ground model. Junction potentials ensure that the figure/ground labels are consistent with one another, forming valid junctions.

We use a *conditional random field* (CRF) model for global figure/ground inference on this discrete structure. Conditional random fields were first introduced by [15] for natural language processing, and have been shown to outperform traditional Markov Random Fields in many domains. It has been previously applied to image labeling [20, 21], as well as contour completion [16].

For every contour e in the image, the local model provides us with an estimate p_e, the probability that the "left" side of e is figure (averaged over all pixels on e). We associate with e a ternary variable X_e, where $X_e = 1$ if the "left" side of e is figure, $X_e = -1$ if the "right" side is figure, or $X_e = 0$ if neither (e.g. a surface marking). Let X_V be the collection of variables for all contours which join at a junction V in the graph. We consider an exponential family distribution over the collection of edges of the form

$$P(X|I,\Theta) = \frac{1}{Z(I,\Theta)} \exp\left\{\sum_e \phi(X_e|I,\Theta) + \sum_V \psi(X_V|I,\Theta)\right\} \quad (1)$$

where ϕ is a unary potential function on each contour e, ψ a potential function on each junction X_V, and Θ is the collection of model parameters. An example factor graph, showing the conditional independence structure of our CRF model, is illustrated in Figure 5(b).

The contour potential ϕ incorporates local figure/ground evidence, defined as $\phi(X_e) = \beta X_e log(\frac{p_e}{1-p_e})$, where p_e is the local estimate that the "left" side of e is figure.

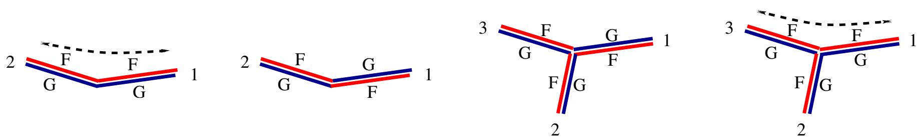
The junction potential ψ assigns a weight to each distinctive "junction type". Suppose a junction V contains k contours $\{e_1, \cdots, e_k\}$, sorted in a clock-wise way, with a figure/ground label assignment X_V (we do not consider any contour with a label $X_e = 0$). The type of the junction V can be represented by a vector of dimension k: $T(X_V) = \{X_{e_1}, \cdots, X_{e_k}\}$. We define

$$\psi(X_V) = \sum_{t \in T_a} \alpha_t \cdot \mathbf{1}_{\{T(X_V)=t\}} + \sum_{t \in T_c} \gamma \cdot \theta(X_V) \cdot \mathbf{1}_{\{T(X_V)=t\}} \quad (2)$$

where T_a is the set of all possible junction types, and T_c is a subset of junction types on which a continuity term θ may be defined (explained below).

Figure 6 shows a few examples of junction types. Intuitively, junction types (a) and (d) are sensible junction labelings, (a) being a continuation of contours, and (d) being a classical T-junction; while junction types (b) and (c) seem highly unlikely. We can count the empirical frequencies of these junction types; and indeed we find that type (a) and (d) are much more common than (b) and (c).

In order to analyze junctions, we need to know the geometric configuration in addition to its type. For example, in a T-junction, we need to know which two contours form the "top" of the "T". This is accomplished by defining a continuity term between a pair of contours. For junction type (d), we know from



(a): type (1,-1), frequency 0.126, weight 0.185 **(b)**: type (-1,-1), frequency 0.039, weight -0.611 **(c)**: type (-1,-1,-1), frequency 0.006, weight -0.857 **(d)**: type (1,-1,-1), frequency 0.086, weight 0.428

Fig. 6. A number of junction types, red indicating the figure side and blue the ground. Each type is represented by its set of figure/ground labels collected in a clockwise way. The empirical frequencies of these junction types confirm that type (a) and (d) are common junction labelings but (b) and (c) are uncommon. This is encoded into the CRF model parameters by maximum likelihood learning. For type (a) and (d), we may define a continuity term θ, which is the angle between the two contours that belong to the foreground.

the figure/ground labels that the contours 1 and 3 form the boundary of the foreground object, while the contour 2 is an occluded contour in the background. Therefore the continuity of this junction is the angle between the contours 1 and 3. We also use continuity in junction type (a) by measuring the angle between the two contours.

Because the contour/junction graph typically contains many loops, exact inference on the CRF model is intractable. We perform approximate inference using loopy belief propagation to estimate marginal posterior distributions of the figure/ground labels X_e. We then assign a binary figure/ground label to a contour e (hence all the pixels on e), the figure being on its "left" side if $P(X_e = 1) > P(X_e = -1)$, or otherwise the "right" side. In our experiments, we find loopy belief propagation converges quickly (< 10 iterations) to a reasonable solution.

We fit the model parameters $\Theta = \{\alpha, \beta, \gamma\}$ using maximum likelihood criterion. The partial derivatives of the log-likelihood take on a simple form as the difference between two expectations. For example,

$$\frac{\partial}{\partial \alpha_t} \log \left(\frac{1}{Z(I,\Theta)} \exp \left\{ \sum_e \phi(X_e|I,\Theta) + \sum_V \psi(X_V|I,\Theta) \right\} \right)$$
$$= \sum_V \mathbf{1}_{\{T(X_V)=t\}} - \left\langle \sum_V \mathbf{1}_{\{T(X_V)=t\}} \right\rangle_{P(X|I,\Theta)}$$

where the first term is the empirical frequency of junction type t, and the second term is the frequency of type t under the current parameter setting. Learning parameters with simple gradient descent converges quickly (< 500 iterations).

5 Figure/Ground Assignment Without Segmentation

The figure/ground models we have introduced are based on the assumption that figure/ground organization occurs after region grouping. Using human-marked segmentations, these models provide valuable insights into the figure/ground process, such as relative powers of the local and global cues involved. To utilize these cues in a practical algorithm, however, we need to compute a segmentation first. Unfortunately, segmentation is a hard problem itself and requires the use of all available visual information, potentially including figure/ground cues.

In this section we replace the human segmentation with a bottom-up grouping process based directly on edge detection. There are two main questions that need to be addressed:

1. The groundtruth labels in the dataset are all given on human-marked boundaries. How do we transfer these labels to a set of (potentially mislocalized) edges, so that we may train and test our models as before?
2. The global conditional random field model requires a discrete contour/ junction graph structure. How do we construct such a structure from the image?

To transfer groundtruth labels, we run a bipartite matching between pixels on human-marked boundaries and pixels on Pb edge maps, illustrated in Figure 7 (a). For each Pb edge pixel, we have an estimate of the local orientation at that location. We then look at the matched pixel on human-marked boundary, compute its figure/ground label at that particular orientation, and transfer it to the Pb pixel. Figure 7(b) shows a few examples from this matching process.

Fig. 7. Transferring groundtruth labels to Pb edge maps. (a) a bipartite matching establishes the correspondence between thresholded Pb edges and human-marked boundaries. This correspondence determines the figure/ground label on each Pb pixel, at its local orientation. (b) examples of the transferred groundtruth labels. White indicates that "left" is the figure side, black the ground side; blue pixels are either not matched, or matched to a boundary with no figure/ground labels (not used in evaluation).

To construct a discrete junction structure on which our conditional random field model can be applied, we use Canny's hysteresis thresholding to trace salient contours in Pb edge maps. Junctions are discovered during the process when two or more contours join. A heuristic is used which merges two vertices when they are sufficiently close to each other.

Pb edge maps have nice non-maximum suppression properties so a naive contour tracing approach suffices most of the time. Figure 8 shows an example of the resulting contour/junction graph, alongside the human-marked segmentation. Our conditional random field model can directly apply to either of these discrete structures.

Fig. 8. Constructing contour/junction structure (c) from thresholded Pb edge maps (b). Contours are marked in black and junctions in red. Such a discrete structure allows global inference on junction consistencies. However, bottom-up contours are much more fragmented and not nearly as clean (and useful) as the human-marked boundaries (d).

6 Experimental Results

The figure/ground dataset we use for both training and testing include 200 Corel images of size 321×481. Human subjects provide one segmentation for each image as well as two sets of figure/ground labels. We use 100 images for training and 100 for testing.

We test the performance of four models: a baseline size/convexity model, the local shapeme model, the local shapeme model averaged on continuous contours, and the global conditional random field model. Each model provides a binary figure/ground label on boundary pixels, and we count the percentage of correct predictions. Surface markings are excluded in these experiments. The models are evaluated in two cases, with or without using human-marked segmentations. Chance is 50%. Since each image is labeled by multiple human subjects, we can measure the labeling consistency between human subjects. For this dataset the self-consistency is 88%.

The baseline size/convexity model is constructed in the following way. Given a segmentation, suppose p is a pixel on a contour c between two segments S_1 and S_2. Let D be a disk around p with a radius r. We can measure the area of overlap $A_1 = |D \cap S_1|$ and $A_2 = |D \cap S_2|$: if the area $A_1 < A_2$, then we label S_1

Table 1. Performance evaluation based on human-marked segmentations. The baseline size/convexity model has difficulties around junctions. Its performance slowly increases with scale/radius, capped at 55.6%. The shapeme model, incorporating convexity, parallelism and textureness, performs much better than the baseline. Averaging local cues over human-marked boundaries proves to suppress noise and significantly increase the performance. The global CRF model, by enforcing labeling consistency at T-junctions, performs the best, achieving 78.3% accuracy.

Chance	Size/ Convexity	Local Shapeme	Averaging Shapemes on Contours	Global CRF	Dataset Consistency
50%	55.6%	64.8%	72.0%	78.3%	88%

Table 2. Performance evaluation based on Pb boundaries. Without the knowledge of a segmentation, the baseline size/convexity model cannot be applied. The local shapeme model, based solely on local image-based descriptors, performs as good as in the case with human-marked boundaries. The global models, however, are severely handicapped without a perfect segmentation. The contour/junction structure constructed from edge maps is useful at enforcing global consistencies (4% improvement); but as it is much more fragmented and noisy than the set of human-marked boundaries, the benefit of global integration is much smaller. Clearly, a more sophisticated contour/region grouping algorithm is needed here to produce better junction structures.

Chance	Size/ Convexity	Local Shapeme	Averaging Shapemes on Contours	Global CRF	Dataset Consistency
50%	n/a	64.9%	66.5%	68.9%	88%

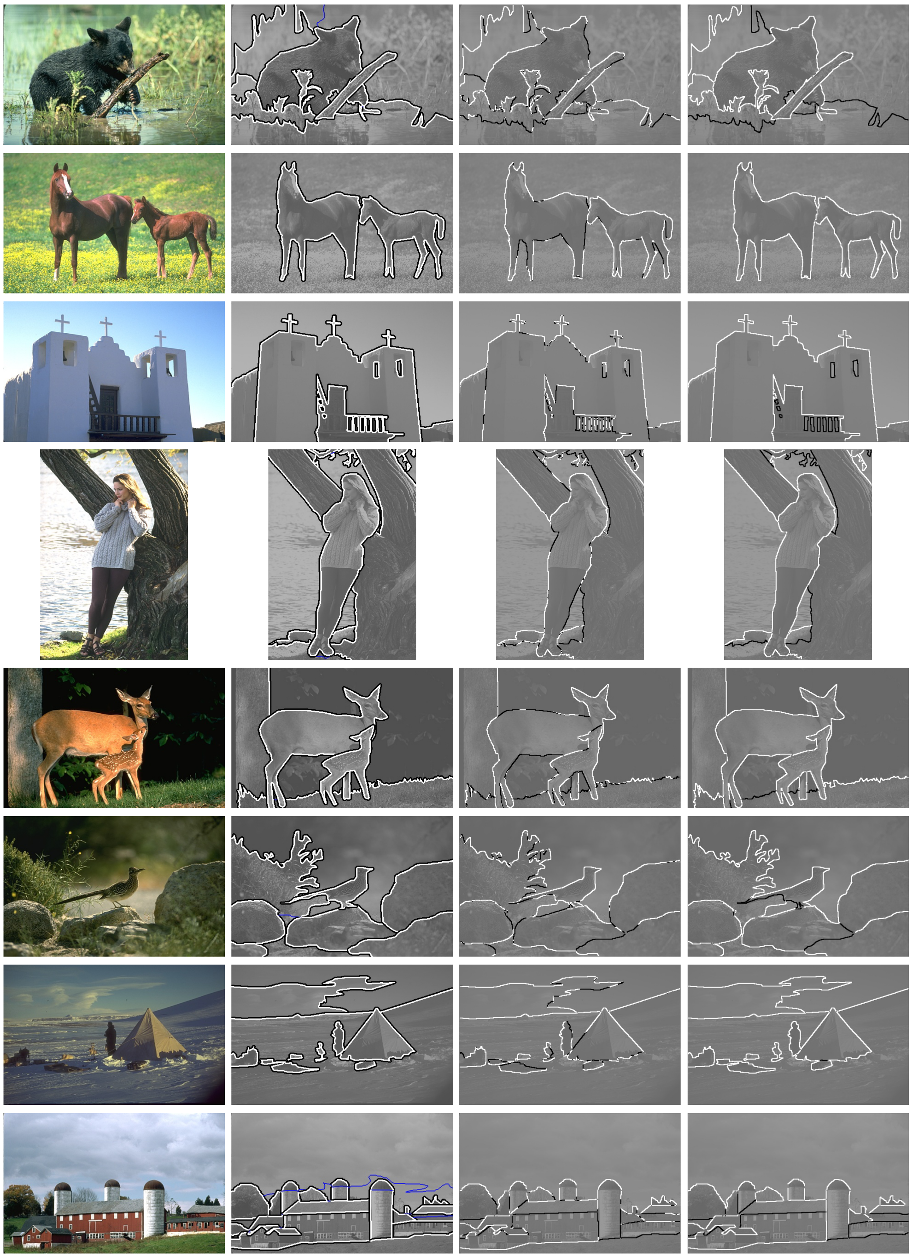


Fig. 9. Results based on human-marked boundaries. Shown here are the images, the groundtruth labels (white being the figure, black the ground, blue neither), figure/ground labels from the local shapeme model (white being correct, black incorrect; average accuracy 64.8%), and labels from the global CRF model (average accuracy 78.3%). The global model performs well in most cases, suggesting that figure/ground assignment in natural images is a feasible problem, if a good segmentation is available.

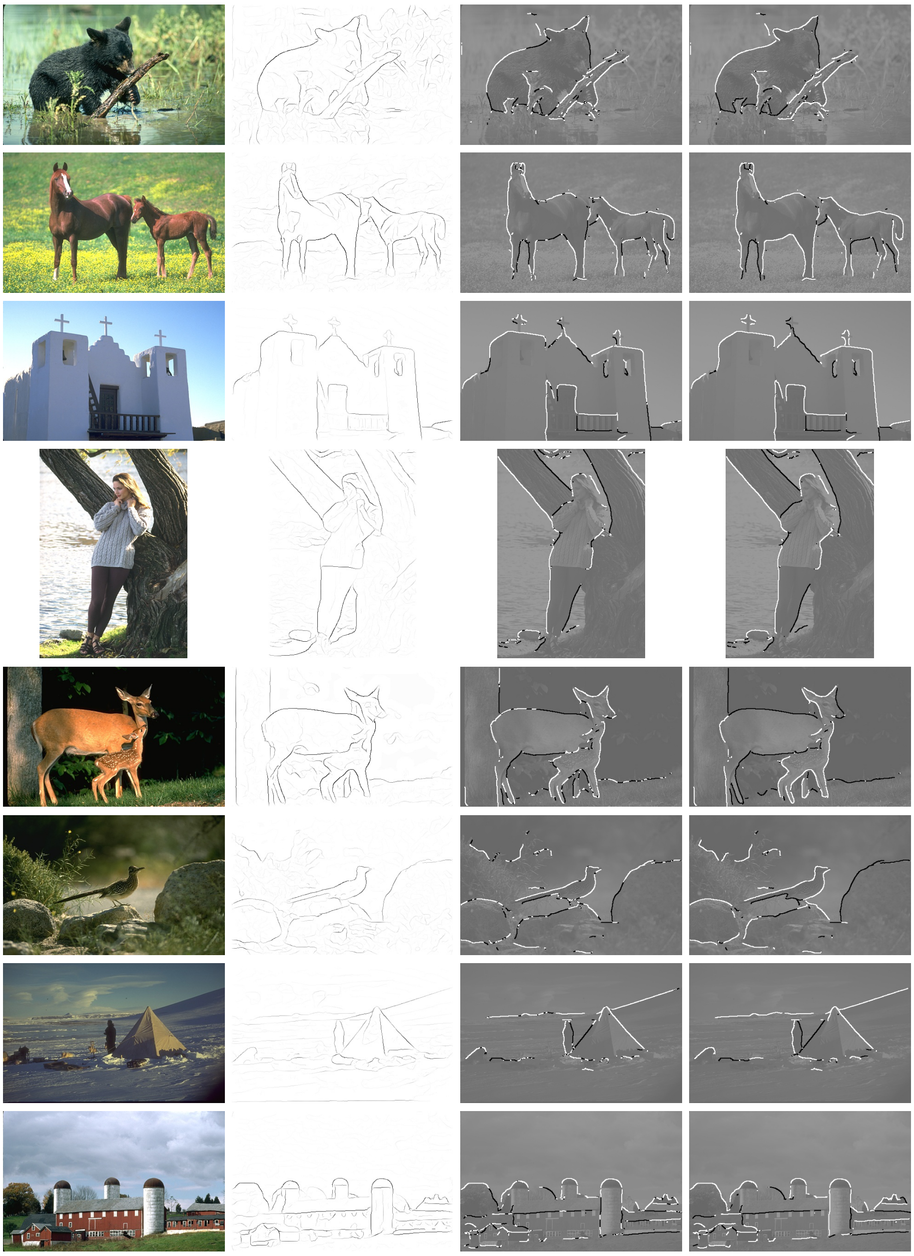


Fig. 10. Results based on Pb boundaries. Shown here are the images, the Pb edge map, figure/ground labels from the local shapeme model (average accuracy 64.9%), and labels from the global CRF model (average accuracy 68.9%). Without using human-marked segmentations, the results are more noisy and less consistent. Nevertheless the local shapeme model applies without any difficulty, and global inference on a bottom-up contour/junction structure still significantly improves performance.

as figure, and vice versa. This simple size cue is closely correlated with convexity (convex regions typically have a smaller size, if the boundary is smooth enough) and has been shown to perform well on this dataset without interference with junctions [1]. This cue, however, relies on the availability of a segmentation and performs poorly near junctions.

Table 1 lists the average labeling accuracy for the case of using human-marked segmentations. Table 2 lists the results in the case of bottom-up contour detection. We find that the local shapeme model performs well in both cases, achieving an accuracy of 64.8%, much higher than the baseline size/convexity model. Enforcing global consistency improves performance in both cases, most notably when using human-marked segmentations. Sample results can be found in Figure 9 and Figure 10.

7 Conclusion

In this work we have developed a model for figure/ground assignment in natural images using shapemes to represent context and a conditional random field to enforce labeling consistency at junctions. We train and test our models on a large dataset of natural images with human-marked groundtruth data, using either a high-quality segmentation or a bottom-up edge detector to determine junction structure.

The local figure/ground prediction based on shapemes performs well in both cases, comparing favorably to previous studies using classical Gestalt figure/ground cues. Shapemes automatically discover contextual cues such as parallelism or convexity, and are robust to complex variability in natural images. The global CRF model significantly improves the performance, most notably when using human-marked boundaries. Experimental results suggest that figure/ground assignment in natural images is a feasible problem and a good segmentation algorithm would greatly facilitate figure/ground organization.

References

1. Fowlkes, C., Martin, D., Malik, J.: On measuring the ecological validity of local figure/ground cues. In: ECVP. (2003)
2. Rubin, E.: Visuell wahrgenommene figuren. In: Kobenhaven: Glydenalske boghandel. (1921)
3. Palmer, S.: Vision Science: Photons to Phenomenology. MIT Press (1999)
4. Peterson, M.A., Gibson, B.S.: Must figure-ground organization precede object recognition? an assumption in peril. Psychological Science **5** (1994) 253–259
5. Kienker, P.K., Sejnowski, T.J., Hinton, G.E., Schumacher, L.E.: Separating figure from ground with a parallel network. Perception **15** (1986) 197–216
6. Heitger, F., von der Heydt, R.: A computational model of neural contour processing: figure-ground segregation and illusory contours. In: ICCV, Berlin, Germany (1993) 32–40
7. Geiger, D., Kumaran, K., Parida, L.: Visual organization for figure/ground separation. In: CVPR. (1996) 155–160

8. Saund, E.: Perceptual organization of occluding contours of opaque surfaces. CVIU Special Issue on Perceptual Organization (1999) 70–82
9. S. Yu, T.L., Kanade, T.: A hierarchical markov random field model for figure-ground segregation. In: EMM CVPR 01. (2001) 118–133
10. Pao, H.K., Geiger, D., Rubin, N.: Measuring convexity for figure/ground separation. In: ICCV. (1999) 948–955
11. Lamme, V.A.F.: The neurophysiology of figure-ground segregation in primary visual cortex. Journal of Neuroscience **15** (1995) 1605–1615
12. Zhou, H., Friedman, H.S., von der Heydt, R.: Coding border ownership in monkey visual cortex. Journal of Neuroscience **20** (2000) 6594–6611
13. Martin, D., Fowlkes, C., Malik, J.: Learning to detect natural image boundaries using brightness and texture. In: Advances in Neural Information Processing Systems 15. (2002)
14. Berg, A., Malik, J.: Geometric blur for template matching. In: CVPR. (2001)
15. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proc. 18th International Conf. on Machine Learning. (2001)
16. Ren, X., Fowlkes, C., Malik, J.: Scale-invariant contour completion using conditional random fields. In: ICCV. (2005)
17. McDermott, J.: Psychophysics with junctions in real images. Perception **33** (2004) 1101–1127
18. Mori, G., Belongie, S., Malik, J.: Shape contexts enable efficient retrieval of similar shapes. In: CVPR. Volume 1. (2001) 723–730
19. Mori, G., Ren, X., Efros, A., Malik, J.: Recovering human body configurations: Combining segmentation and recognition. In: CVPR. Volume 2. (2004) 326–333
20. Kumar, S., Hebert, M.: Discriminative random fields: A discriminative framework for contextual interaction in classification. In: ICCV. (2003) 1150–1159
21. He, X., Zemel, R., Carreira-Perpinan, M.: Multiscale conditional random fields for image labelling. In: CVPR. Volume 2. (2004) 695–702