

# Dead Rectangles as a Stimulus for Perceptual Organisation Research

Heiko H. Schütt (heiko.schuett@nyu.edu)

Center for Neural Science and Department of Psychology, New York University, 4 Washington Place  
New York, NY 10003 USA

Wei Ji Ma (weijima@nyu.edu)

Center for Neural Science and Department of Psychology, New York University, 4 Washington Place  
New York, NY 10003 USA

## Abstract

We describe dead rectangles, a new stimulus class for research on perceptual organisation. These stimuli are generated by randomly placing rectangles inside an image window, which allows for occlusion, surface properties and ambiguous stimuli. To show their utility we asked 14 observers to judge whether two points in a dead rectangle stimulus belong to the same rectangle or not. We find observers to perform around 70% correct and judge the points to belong to the same rectangle considerably more often than they do. Also, some pairs of points were consistently judged to belong to the same rectangle although they were not. The possible decompositions of one stimulus into rectangles form a tree, which can be used for formal analysis. These stimuli may allow research on perceptual organisation to move on to more natural conditions, while maintaining experimental control and a rigorous mathematical framework.

**Keywords:** dead rectangles; dead-leaves; stimulus; grouping; perceptual organisation

## Introduction

Perceptual organisation is the process by which humans segment their percept into objects by grouping smaller parts. This process is effortless and leads to highly accurate segmentation under natural conditions. Perceptual organisation is central to perception as it affects many other aspects of perception like attention (Soto & Blanco, 2004) and potentially even low-level processes like pooling and normalisation (Neri, 2011; Coen-Cagli, Kohn, & Schwartz, 2015). Thus, researchers study perceptual organisation since the Gestalt psychologists in the beginning of the 20th century (Koffka, 1935). These researchers asked subjects whether they would group simple stimuli on a grid rather in one or the other direction, which lead to a list of factors influencing grouping like proximity or similarity. However, most of these stimulus groups are not perceived as single objects and whether and how the grouping principles apply to natural stimuli and how multiple objects interact remain unclear.

We believe that one obstacle for perceptual organisation research is the lack of stimuli and an experimental paradigm which allow deeper insights into the problem. Most research on grouping and perceptual organisation is done using highly simplified stimuli like dots (Feldman & Singh, 2006) or Gabor patches (Field, Hayes, & Hess, 1993). These stimuli lack im-



Figure 1: Example stimulus used in our experiment, with a size exponent  $\alpha$  of 3 and 40 pixels distance between the queried points.

portant aspects of natural scenes like surfaces, textures and occlusions, which may be highly relevant for perceptual organisation. As a new artificial stimulus we here suggest dead rectangles, which consist of randomly placed stacked rectangles of various sizes. These stimuli are a special case of dead-leaves stimuli, which are generated by randomly placing simple shapes and named after foliage on the forest floor (Lee, Mumford, & Huang, 2001; Gousseau & Roueff, 2007; Pitkow, 2010; Zoran & Weiss, 2012). We display an example dead rectangles stimulus in Figure 1.

Dead rectangles are still artificial stimuli, but add important aspects, previously used stimuli do not possess: Dead rectangles have a surfaces and can thus occlude each other, which is possibly the most important visual interaction of objects. Dead rectangles stimuli contain many objects, mimicking natural conditions. Dead rectangles can mimic natural image statistics like edge frequencies as all dead-leaves stimuli can (Gousseau & Roueff, 2007; Pitkow, 2010). Finally, dead rectangles can create ambiguous situations without losing

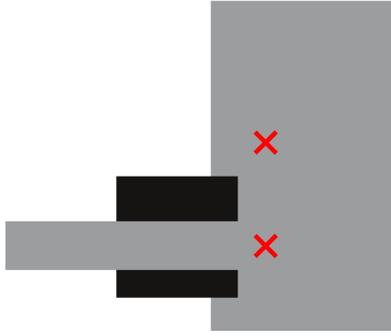


Figure 2: An ambiguous example stimulus. It is invisible how long the thin horizontal rectangle is. If it is long enough the two red dots are on different rectangles, otherwise they are on the same rectangle. Removing the black rectangle would generate another ambiguity which of the two gray rectangles is in front.

the percept of object-hood, as we show in Figure 2. This true ambiguity motivated us to use rectangles instead of the more commonly used ellipses and circles, for which a tiny part of the contour is sufficient to infer the whole shape.

As artificial stimuli dead rectangles also have a number of advantages over natural stimuli: We can control and vary any aspect of them. We know the true decomposition into objects which is unambiguously defined. There are no additional cues to object assignments like object classes or correlations among objects. Finally, subjects are less likely to have prior knowledge or expectations about dead rectangles influencing their decisions.

### Formal Description

To generate a dead rectangle stimulus we draw rectangles from a discrete set of possible rectangle sizes  $\Omega_R$  and place them uniformly randomly over the positions rendering any part of the rectangle visible. We sample rectangles until every pixel in the image is covered by a rectangle and then reverse the stack. This samples from the same distribution as visible on top of an infinitely deep stack of rectangles (Pitkow, 2010).

We draw the rectangle sizes  $s_x$  and  $s_y$  in the two dimensions independently from a range  $[s_{min}, s_{max}]$  with a stepsize  $s_{step}$  and probability  $p(s) \propto s^{-\alpha/2}$  with a size exponent  $\alpha$ . According to earlier analyses this leads to a size invariant distribution with an exponent  $\alpha = 3$  (Lee et al., 2001).

To adjust this distribution for the fact that we draw rectangle positions from a different area for different sizes of rectangles, but want the distribution of rectangles to be independent of the image size  $d$  we use a corrected distribution  $p'(s)$  for the size of a rectangle:

$$p'(s) \propto (s + d)s^{-\alpha/2} \quad (1)$$

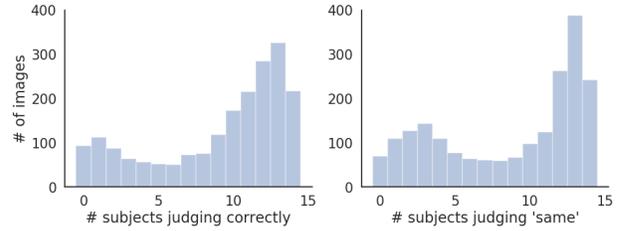


Figure 3: Histogram of stimuli showing how many pairs of points were judged correctly and how many were judged to be on the same rectangle by how many subjects.

## Behavioural Experiment

### Methods

In our behavioural experiment we asked 14 subjects whether two marked points in a dead rectangle stimulus belonged to the same rectangle or not. Subjects were recruited from the NYU campus community, were naive to the aim of the study and gave informed consent.

We pre-generated 2000  $300 \times 300$ px dead rectangles stimuli for this experiment, which were shown as up-sampled  $900 \times 900$  pixel images centrally on a screen with  $1280 \times 960$  pixels resolution until the subject responded for up to 5 seconds. We drew sizes from the range  $[5\text{px}, 400\text{px}]$  with a stepsize of  $s_{step} = 5\text{px}$ . We varied the exponent  $p$  of the size distribution from 1 to 5 and chose the queried points at five different distances symmetrically around the image centre with the points displaced horizontally, vertically (5, 10, 20, 40 or 80 px apart) or along either diagonal (4, 7, 14, 28 or 57 px apart) resulting in 4 orientation conditions. Thus, we had  $5 \times 5 \times 4 = 100$  conditions in total allowing for 20 different stimuli per condition.

We blocked the trials by distribution exponent to keep the image distribution constant within a block. Within a block trials were shown in a new random order for each subject. Each block was preceded by a 500 frame long movie showing the image generation with the distribution for that block starting from a blue background image and adding one rectangle every frame. We split the experiment into two sessions with 5 blocks for the 5 exponents of 200 trials each. At the beginning of each session we added a training session of 100 trials, which showed a single trial for each condition, which was not used in the main experiment, while keeping the separation into the 5 blocks and the movies.

### Results

Subjects respond reasonably consistently in our task as shown by clearly bi-modal distributions for how many subjects judge a pair of points to fall on the same rectangle and how many subjects are correct (see Figure 3).

Subjects all performed clearly above chance and their responses follow the big trends in the correct probabilities for two points to lie on the same rectangle. I.e., the true probability that two points belong to the same rectangle and subjects'

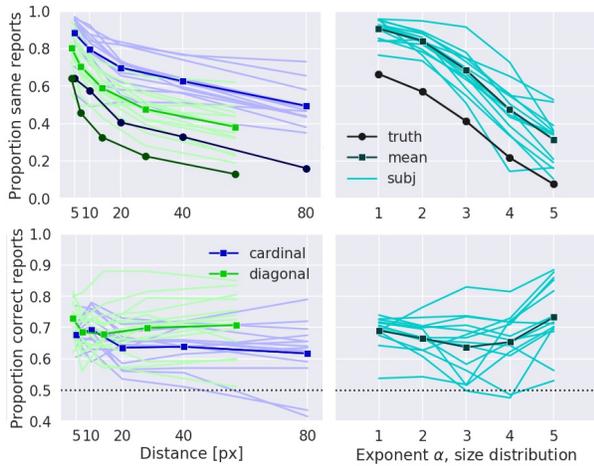


Figure 4: Experiment results. Top row shows the proportion of point pairs judged to be on the same rectangle. Bottom row shows the proportion of correct responses. Blue lines are for the horizontally and vertically separated point pairs, the green lines for the diagonally displaced point pairs. The darkest lines show the true proportion of point pairs on the same rectangle, the medium dark ones show the subject average and the light ones show individual subject data. For plotting against the size distribution exponent  $\alpha$ , we average all directions and plot in turquoise.

judgements decrease with increasing distance between the two points and with increasing size exponent (smaller rectangles), as displayed in Figure 4.

However, subjects generally report to see the two points belonging to same rectangle considerably more often than they do. This trend is consistent across all conditions. Interestingly, this effect is also seen in the consistent errors subjects make. In our dataset 93 stimuli were judged wrongly by all subjects and 112 were judged wrongly by 13 of 14 subjects. These 205 point pairs were all judged to belong to the same rectangle despite belonging to different ones.

## Discussion

Our subjects were able to do the task with little instruction and produced reasonably consistent results. Nonetheless they make a relatively large number of errors which makes this data interesting for modelling as simply predicting the correct response will not be the best model. Finally, they produce a peculiar tendency to make consistent errors only in one direction, which is an interesting qualitative pattern future models should explain.

## Graph Description & Tree Search

For formal descriptions, optimal inference, and models of human behaviour it is helpful to consider the set of all possible decompositions of a dead rectangle stimulus into rectangles as a graph tree, which starts from an empty image as the root

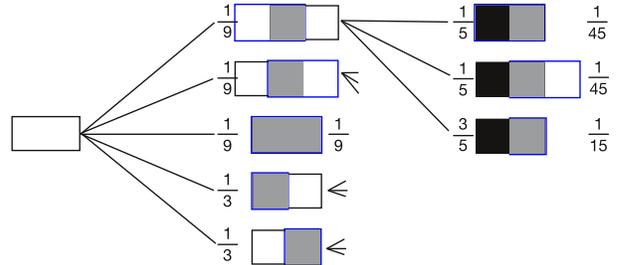


Figure 5: Example graph constructed for a two pixel image with only  $1 \times 1$  and  $2 \times 1$  rectangles with probabilities  $p'$  of  $\frac{2}{3}$  and  $\frac{1}{3}$  respectively. Gray areas mark the part of the image covered by the added rectangle. The black areas the area already covered. Blue outlines show the rectangle added.

and adds one rectangle at every step through the tree starting from the top. If we list all possible, visible rectangles as possibilities at each node each path from the root to a leaf of the tree represents one decomposition and the graph as a whole contains all possible decompositions. We can generate this graph recursively by first searching for all possible sizes and positions of the top rectangle in the stimulus. For the next rectangle we then allow everything for the area behind the first rectangle and find all possibilities for the second rectangle. As we remove the possible but invisible rectangle positions behind the already placed rectangles this recursion will eventually end.

From our construction we can also deduce the prior probability for each branch of the tree and thus for each decomposition into rectangles. This probability is the product of probabilities for each branch taken in the tree, where the probability for a branch is calculated by normalising the probability to place the rectangle at that position such that the probabilities for all visible rectangles (consistent and inconsistent) add to 1.

To compute the posterior over possible decompositions of a dead rectangle stimulus we need to find all decompositions consistent with the image and weigh them with their prior probability times the likelihood to produce the observed pixel colours. To find the consistent decompositions we can find which additional rectangles are consistent with the image at every node of the graph and can restrict our search for possible decompositions to these options. The likelihood to produce the colours we observe is simply  $\frac{1}{N^l}$ , where  $N$  is the number of colours and  $l$  is the number of rectangles in the decomposition.

This allows us to compute the optimal observer for the task in our behavioural experiment by summing up the posterior probabilities of all decompositions where the two points are on the same rectangle and the ones where the two are on different rectangles. These will be the posterior probabilities for the two cases.

We illustrate the tree of possible rectangle decompositions with the prior probabilities in Figure 5 for a minimal example of

a  $2 \times 1$ px image with only two possible rectangles ( $2 \times 1$  and  $1 \times 1$ ) with probabilities  $p' = \frac{1}{3}$  and  $\frac{2}{3}$ . First, we find all possible rectangles which could be on top of the stack, which in this case are 5: Three positions for the two pixel rectangle, which each have probability  $\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$  and 2 positions of the one pixel rectangle, which each have probability  $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$ . One of the positions of the 2 pixel rectangle fills the whole image and is thus a leaf. For all other positions there is one pixel left to be covered. For these situations we find that three of the five rectangle positions remain visible: two of the 2 pixel rectangle with probability  $\frac{1}{9}$  and one position of the 1 pixel rectangle with probability  $\frac{1}{3}$ . Re-normalising these values gives the  $\frac{1}{5}$  and  $\frac{3}{5}$  probabilities in the figure.

Thus, the prior probability for the two pixels to end up on the same rectangle is  $\frac{1}{9}$ , the probability of the single decomposition where both pixels are on the same rectangle. To calculate the posterior over the decompositions we need the additional information that there are  $N$  possible colours. For example, if the two pixels have the same colour, the probability that the one decomposition with both pixels on one rectangle creates the image is  $\frac{1}{N}$ , while the probability for all others is  $\frac{1}{N^2}$ . Thus, the posterior probability for the two to be on the same rectangle is:

$$P(\text{same}) = \frac{\frac{1}{9N}}{\frac{1}{9N} + \frac{8}{9N^2}} = \frac{N}{N+8} \quad (2)$$

As the search for possible rectangle positions can be implemented as a single convolution and thresholding for each rectangle shape and colour, paths through the graph can be computed relatively efficiently for moderately sized images.

For larger stimuli however, the number of decompositions consistent with any particular stimulus is extremely large. For the stimuli we used in our experiment there are typically hundreds of branches from each node and each decomposition consists of hundreds of rectangles. This extremely large space of possibilities renders the exact computation of the optimal observer computationally infeasible. Furthermore, it seems unlikely that our human observers considered all those possibilities in the 1-2 seconds they typically took to answer a trial.

There are, however, many possible approximations to the optimal solution. For example, one can try to search for high probability decompositions and base the decision on only few decompositions. Another possibility is to consider only a subset of the pixels in the image for the decision, which can drastically reduce the depth of the tree.

## Conclusion

We present dead rectangle stimuli as a stimulus for research on perceptual organisation. In contrast to previously used stimuli dead rectangle stimuli are proper images, allow occlusion among the objects in the stimulus and can be ambiguous. Nonetheless, we know the true solution for each stimulus and can compute the distribution over stimuli in contrast to natural images. Furthermore, subjects in our experiment could

quickly and consistently judge the stimuli, while making errors and consistently judging about 10 % of the point pairs to belong to the same rectangle although they did not. Thus, dead rectangles seem to be interesting stimuli for further experiments and modeling endeavours.

## Acknowledgements

This work was funded by the National Institutes of health (NIH) through grant R01EY020958 awarded to Wei Ji Ma and by the German Research Foundation (DFG) through grant SCHU 3351/1-1 to Heiko Schütt.

## References

- Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nature Neuroscience*, *18*(11), 1648–1655. doi: 10.1038/nn.4128
- Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, *103*(47), 18014–18019. doi: 10.1073/pnas.0608811103
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research*, *33*(2), 173–193. doi: 10.1016/0042-6989(93)90156-Q
- Gousseau, Y., & Roueff, F. (2007). Modeling Occlusion and Scaling in Natural Images. *Multiscale Modeling & Simulation*, *6*(1), 105–134. doi: 10.1137/060659041
- Koffka, K. (1935). *Principles of Gestalt psychology*. Oxford, England: Harcourt, Brace.
- Lee, A. B., Mumford, D., & Huang, J. (2001). Occlusion Models for Natural Images: A Statistical Study of a Scale-Invariant Dead Leaves Model. *International Journal of Computer Vision*, *41*(1-2), 35–59. doi: 10.1023/A:1011109015675
- Neri, P. (2011). Global Properties of Natural Scenes Shape Local Properties of Human Edge Detectors. *Frontiers in Psychology*, *2*. doi: 10.3389/fpsyg.2011.00172
- Pitkow, X. (2010). Exact feature probabilities in images with occlusion. *Journal of Vision*, *10*(14), 42:1–20. doi: 10.1167/10.14.42
- Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: a comparison within a single task. *Vision Research*, *44*(1), 69–81. doi: 10.1016/j.visres.2003.08.013
- Zoran, D., & Weiss, Y. (2012). Natural Images, Gaussian Mixtures and Dead Leaves. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp. 1736–1744). Curran Associates, Inc.