

# Bayesian inference of form and shape

Pascal Mamassian\*

CNRS UMR 8581, LPE Université Paris 5, 71 av. Edouard Valliant, 92100 Boulogne-Billancourt, France

**Abstract:** The ability to visually perceive two-dimensional (2D) form and three-dimensional (3D) shape is one of our most fundamental faculties. This ability relies on considerable prior knowledge about the way edge elements in an image are likely to be connected together into a contour as well as the way these 2D contours relate to 3D shapes. The interaction of prior knowledge with image information is well modeled within a Bayesian framework. We review here the experimental evidence of shape perception seen as a Bayesian inference problem.

**Keywords:** visual perception; Bayesian inference; form; shape; contour; three-dimensional perception

## Introduction

What are shapes good for? Imagine you are interested in identifying a tree in your local park. Neither its size nor its color is a reliable cue for its name because the tree's size is dependent on its age and its color on the current season. In contrast, a very robust cue is the shape of the tree, be it pyramidal, columnar, V-shaped, round, or oval. Thus for example, a pin oak will be easily distinguished from a white oak because the former is pyramidal and the latter is round in shape. Not only is the global shape important, but the local shapes also carry critical information. To pursue our example, the pin oak has leaves that contain lobes with bristle-tipped teeth whereas the leaves of the white oak have rounded lobes.

In addition to the difference in scale, the whole tree and the individual leaves also differ in another fundamental way. The tree is described as a 3D shape whereas the leaves are 2D. While the distinction between 2D and 3D shapes is clear in the world, it is often much more subtle for the visual system. Figure 1 illustrates how a simple line

drawing can appear 3D. Interestingly, once the 3D percept arises, it is difficult to simultaneously hold a planar interpretation.

In this chapter, we discuss how 2D and 3D object shape can be inferred from an image. It is well known that a single image is consistent with an infinite number of scenes and yet our perception does not reflect these ambiguities. Therefore, it seems that additional information is used to disambiguate the image. This extra information has

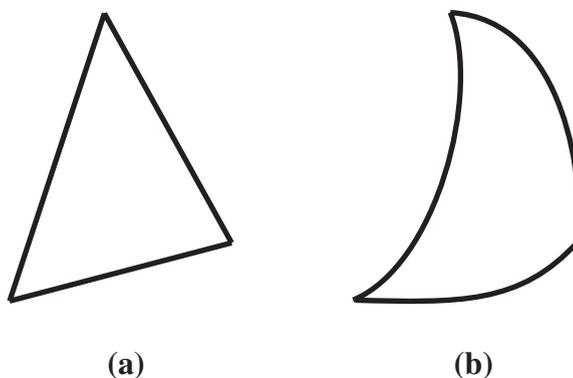


Fig. 1. Distinction between 2D and 3D shapes. (a) Three vertices connected by straight lines produce the impression of a flat triangle. (b) In contrast, when those three vertices are connected by curved lines, a curved 3D surface can be perceived.

\*Corresponding author. Tel.: +33-155-205-930,  
Fax: +33-155-205-854; E-mail: pascal.mamassian@univ-paris5.fr

been variously called a Gestalt principle, a rule, a law, or a property. In contemporary Bayesian models, this extra information is called prior knowledge. We start this chapter with a brief reminder of the way the Bayesian framework can be used in visual perception, and then describe how this framework has been used to understand contour groupings. We then review some studies on 2D and 3D shape from contours that can be cast within this Bayesian framework.

### Basic Bayes

The Bayesian framework has the merit to make a clear distinction between the information available at the level of the receptors and the information specific to the organism. The former is called the likelihood and the latter the prior knowledge. Priors are often seen as supplementary information brought in to disambiguate sensory information.

Imagine that you are interested in inferring the shape of a closed contour from an image such as the outline of a sculpture from Constantin Brancusi (Fig. 2). The sculpture is egg-shaped and sliced by a plane on its left side thus forming an elliptical section. The section is discontinuous in the image because of the particular illumination conditions when the image was taken but the elliptical contour does exist on the sculpture. How can we infer the closed contour in the world given

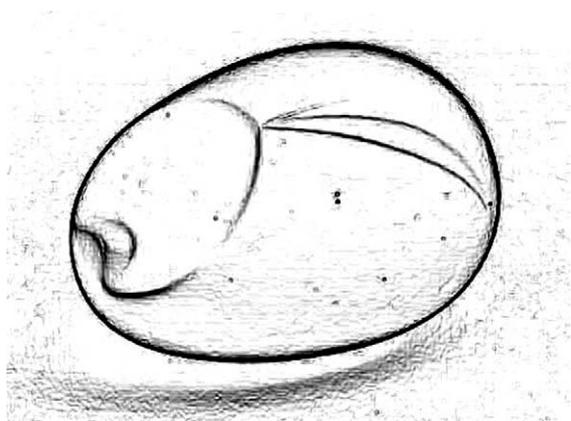


Fig. 2. Contour image of the sculpture “The Newborn” by Constantin Brancusi (1915). The original sculpture can be seen at the Philadelphia Museum of Art.

the broken contour in the image? We come back to this question in the next section; suffice to say at this stage that the image impinging on our retinas should be complemented by additional information if we want to conclude that the object does contain an elliptical cut.

The Bayesian framework offers the optimal way to combine multiple sources of information. The likelihood is combined with the prior knowledge thanks to Bayes’ theorem, thus producing a posterior distribution. The posterior represents the relative probability of all possible interpretations of the stimulus, taking into account all the information available. When one is interested in the choice of the organism for one particular interpretation, we also need to take into account a decision rule that is applied on the posterior distribution (Fig. 3). More details about the Bayesian approach to visual perception can be found in Mamassian et al. (2002) and Kersten et al. (2004).

### Bayesian models of contour grouping

Let us come back to our example of the broken elliptical contour in the picture of Brancusi’s sculpture. If we want to infer the closed contour, we need some prior knowledge about contours. This extra knowledge is likely to be similar to what Gestalt psychologists referred to as the principle of good continuation: two nearby line segments in an image can be thought to belong to the same contour in the world if there is a smooth curve of minimal curvature that joins them.

Gestalt principles have often been criticized for being qualitative descriptions of phenomena rather than proper explanations. Recent work in experimental psychology has however allowed us to quantify these principles. For instance, Kubovy and Wagemans (1995) have quantified the principle of proximity with the elegant use of dot lattice stimuli. This kind of experimental approach will allow us to measure the relative strength of Gestalt principles and thus to make quantitative predictions on how these principles can act as prior knowledge in the interpretation of contour shape.

If human observers use a prior assumption that contours are smooth, then we should be able to

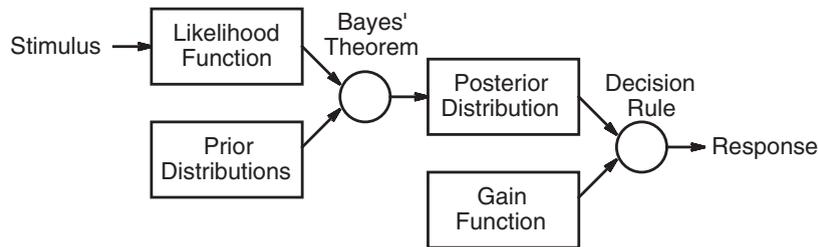


Fig. 3. Flowchart of the Bayesian framework. Reprinted from Mamassian et al. (2002), with permission from MIT Press.

predict when a series of dots are perceived as a single contour or as two contours connected at a corner. Feldman (2001) measured the probability to perceive one smooth contour versus a corner in a variety of conditions where one dot was more or less aligned with the dots on either side (see also Warren et al., 2002). He could account for the relative probability to perceive a corner thanks to prior probabilities on the angle between successive dots along a smooth contour.

The prior in Feldman's study had a general shape (it favored straight contours) and had a degree of freedom adjusted to the data (the belief that most contours are straight). Other studies have attempted to extract the prior probabilities from the statistics of natural scenes. Two groups of researchers have independently found that scenes in our environment do indeed contain a predominance of straight contours. Geisler et al. (2001) extracted the edges of 20 natural scenes and measured how nearby edges varied in orientation. They found a prevalence not only of parallel contours, but also of co-circular contours (edges that would fall on a circle). This latter property is a physical property that could be at the origin of the Gestalt principle of good continuation and the existence of association fields in the human visual cortex (Field et al., 1993). From their statistics, Geisler et al. (2001) extracted a likelihood ratio that was a unit-free estimate that two edges belonged to the same contour and found that this measure was a good descriptor of humans' ability to detect a contour in noise. In a similar vein, Elder and Goldberg (2002) extracted the edges of nine natural scenes and measured statistics on proximity, good continuation and luminance similarity. They found that proximity was by far the most reliable cue to group

elements along a contour and that human observers were in good agreement with a model based on the proximity cue. Finally, one should also note the effort of Howe and Purves (2005) to relate the statistics of oriented edges to contours in the 3D world instead of just the 2D retinal image. They also found a predominance of aligned contours in their analysis.

### From contours to 2D shape

We have seen in the previous section how a fragmented contour could be virtually reconnected into a continuous contour. Once the contour has been revealed, its shape can be inferred. Here again, at the level of the 2D shape of the contour, the Bayesian framework has proved useful. One elegant demonstration of the role of prior knowledge on shape perception is the work of Liu et al. (1999) on amodal completion. When the contour of an object is occluded by another object, the missing part of the contour can be interpolated via a process called amodal completion (Kellman and Shipley, 1991). For instance in Fig. 4, the upper parts of two objects are perceptually completed behind the occluder and linked to their respective lower parts. Liu et al. (1999) placed the occluder either in front or behind the object parts thanks to some binocular disparities and found that convex objects were completed more easily than concave ones. They reached this conclusion by placing the upper and lower parts at different depths and found higher depth discrimination thresholds when the object underwent a convex completion.

As we discussed in the introduction, there are a few objects in the world, like leaves, that are

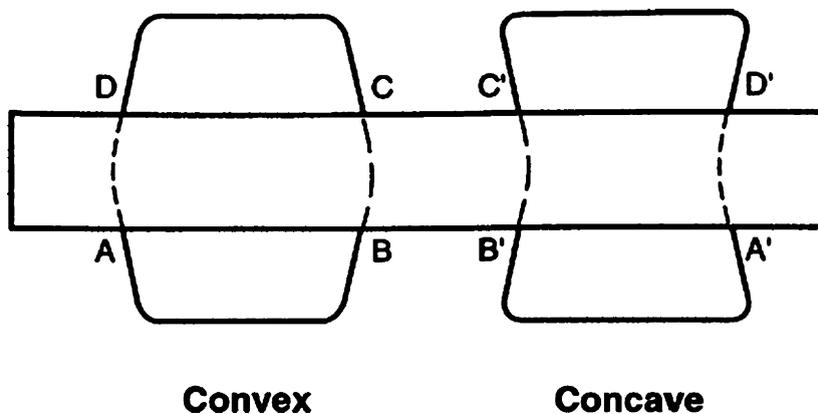


Fig. 4. Convex and concave amodal completion. Reprinted from Liu et al. (1999), with permission from Elsevier.

basically 2D. There are also quasi-objects that have a shape and that are 2D. The most common example is probably the shadow cast by an object on a remote surface. Interestingly, we are very tolerant of the relationship between the shape of the casting object and the shape of the cast shadow, so much so that we seem insensitive to impossible shadows (Mamassian, 2004). Other 2D quasi-objects are holes in a surface (Casati and Varzi, 1994; Bertamini & Croucher, 2003). The extent to which the shapes of cast shadows and holes are processed the same way as other physical objects is still an empirical question.

### From contours to 3D shape

Contours in an image are not only informative about the outline of an object, but also about the 3D shape of the object. Perhaps, the most celebrated demonstration of the 3D shape information available from contours is Attneave's cat shown in Fig. 5 (Attneave, 1954). Attneave argued that most of the information along a contour was concentrated in parts of high curvature. This argument has since received considerable support from both computational and psychophysical studies (Norman et al., 2001; Feldman and Singh, 2005). However, proving that information is concentrated at points of high curvature does not prove that the rest of the contour is uninformative. Koenderink (1984) showed that for a smooth surface, there is a one-to-one relationship between the



Fig. 5. Attneave's cat. Reproduced from Attneave (1954), with permission.

sign of curvature of the occluding contour and the sign of curvature of the surface. More specifically, a convex contour is the signature of a convex surface patch, and a concave contour indicates a hyperbolic (saddle-shaped) surface patch. Moreover, the curvature magnitude of the object can be inferred from the occluding contour when the observer is active (Mamassian and Bülthoff, 1996). Once the shape of the occluding contour is inferred, the 3D shape can be extrapolated thanks to a surface filling-in mechanism (Tse, 2002).

There are also other sources of information about 3D shape than those present on the occluding contour. For instance, parallel structure in a solid (like in marble) will create stripes on the surface of the object and the deformation of the resulting texture will be informative about the

shape of the surface (Li and Zaidi, 2004). Lines that appear painted on the surface on an object tend to be perceived as geodesics (contours of minimum curvature) and are therefore also informative about the 3D shape of the object (Knill, 1992). In addition, these contours tend to be interpreted in such a way that the surface that supports them is seen from above (rather than from below), thereby constraining further the 3D shape of the object (Mamassian and Landy, 1998).

Other contours are also informative about the 3D shape of the surface on which these contours appear to be painted. For instance, alternating dark and bright parallel contours tend to be perceived as shaded bevelled patterns on the surface (Mamassian and Goutcher, 2001). These contours are informative about the 3D shape of the surface only if you know the illumination position, namely above the observer. In summary, contours are informative about 3D shape provided a number of assumptions about the illumination position, the viewer position, or the alignment of the surface contours with geometric properties of the object. Without these assumptions, the contours are ambiguous about the 3D shape of the object. The fact that human observers are able to consistently perceive the same 3D shape from these images is consistent with a mechanism that integrates image information with prior knowledge. Bayesian models are precisely models of this type.

## Discussion

The state-of-the-art we have presented in this chapter on shape perception can be schematized by the diagram in Fig. 6. Once basic features such as oriented edges are extracted from an image, they can be grouped together to form continuous contours. These contours can then be used to infer the shape of an object in various ways. First, if the object is 2D, the shape of the contour will simply be related to the shape of the object. Second, if the contour is identified as the contour of the object against the background, there are mechanisms that will allow us to fill-in the surface from this occluding contour. Finally, if a contour is near other similar contours, these can be grouped together into a texture pattern, and a 3D shape can then be inferred from this texture. Even though we have drawn uni-directional arrows in this figure, it is clear that feedback mechanisms do influence early representations (Murray et al., 2004).

One critical aspect of this way to understand shape perception is that prior knowledge is brought in at all stages of visual processing. For instance, features are grouped into contours following preferences that contours are straight. Contours are grouped into texture patterns following some principles of similarity. 3D shape is inferred from texture following homogeneity assumptions on the texture. Each stage can be seen as a Bayesian inference process where information

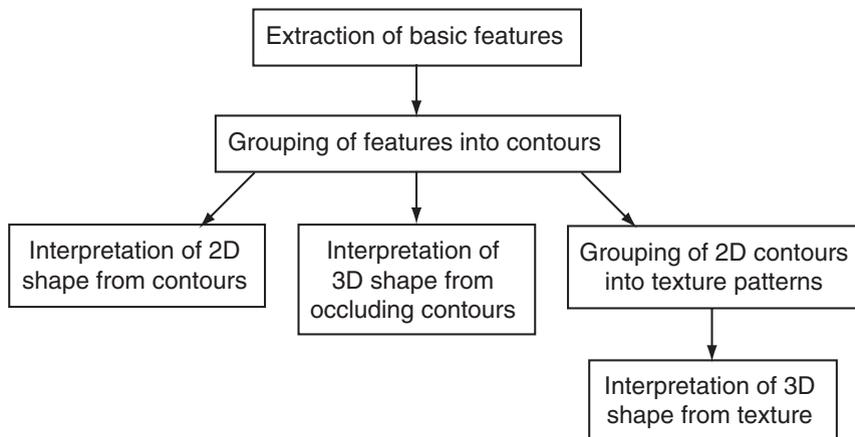


Fig. 6. A schematic diagram of the perception of 2D and 3D shape from an image.

from the earlier stage is combined with the relevant prior knowledge for the current stage. While there are a growing number of studies on the individual stages, little work has been done to integrate these stages into a hierarchical Bayesian model of shape perception.

We conclude this chapter with some outstanding questions. First, what are the Bayesian aspects that are equivalent to generic principles such as “simplicity” or “smoothness”? We have argued that such principles would intuitively translate as priors in a Bayesian model, but the precise nature of these priors and their origin remains a serious empirical question. Theoretical works showing the equivalence between simplicity and some statistical quantity are very important to explore this question (Chater, 1996).

A second issue is whether priors necessarily reflect regularities of the natural world. Some works within the Bayesian framework do indeed take this assumption as a premise, but some believe that this is not a requirement. For instance, we found that human observers behaved as if their assumption on the light source position was biased to the left (Mamassian and Goutcher, 2001), an assumption that is difficult to relate to statistics of the natural world. Future work should investigate the origin of such priors.

Finally, we may wonder what are the neural bases of Bayesian processing? The empirical determination of prior assumptions used by human observers when they perceive shape opens the door to fascinating investigations on the way these priors are implemented.

## References

- Atneave, F. (1954) Some informational aspects of visual perception. *Psychol. Rev.*, 61: 184–193.
- Bertamini, M. and Croucher, C.J. (2003) The shape of holes. *Cognition*, 87: 33–54.
- Casati, R. and Varzi, A.C. (1994) *Holes and Other Superficialities*. MIT Press, Cambridge, MA.
- Chater, N. (1996) Reconciling simplicity and likelihood principles in perceptual organization. *Psychol. Rev.*, 103: 566–581.
- Elder, J.H. and Goldberg, R.M. (2002) Ecological statistics of Gestalt laws for the perceptual organization of contours. *J. Vision*, 2: 324–353.
- Feldman, J. (2001) Bayesian contour integration. *Perception Psychophys.*, 63: 1171–1182.
- Feldman, J. and Singh, M. (2005) Information along contours and object boundaries. *Psychol. Rev.*, 112: 243–252.
- Field, D.J., Hayes, A. and Hess, R.F. (1993) Contour integration by the human visual system: evidence for a local ‘association field’. *Vision Res.*, 33: 173–193.
- Geisler, W.S., Perry, J.S., Super, B.J. and Gallogly, D.P. (2001) Edge co-occurrence in natural images predicts contour grouping performance. *Vision Res.*, 41: 711–724.
- Howe, C.Q. and Purves, D. (2005) Natural-scene geometry predicts the perception of angles and line orientation. *Proc. Nat. Acad. Sci., USA*, 102: 1228–1233.
- Kellman, P.J. and Shipley, T.F. (1991) A theory of visual interpolation in object perception. *Cogn Psychol.*, 23: 141–221.
- Kersten, D., Mamassian, P. and Yuille, A. (2004) Object perception as Bayesian inference. *Annu. Rev. Psychol.*, 55: 271–304.
- Knill, D.C. (1992) Perception of surface contours and surface shape: from computation to psychophysics. *J. Opt. Soc. Am. A*, 9: 1449–1464.
- Koenderink, J.J. (1984) What does the occluding contour tell us about solid shape? *Perception*, 13: 321–330.
- Kubovy, M. and Wagemans, J. (1995) Grouping by proximity and multistability in dot lattices: a quantitative Gestalt theory. *Psychol. Sci.*, 6: 225–234.
- Li, A. and Zaidi, Q. (2004) Three-dimensional shape from non-homogeneous textures: carved and stretched surfaces. *J. Vis.*, 4: 860–878.
- Liu, Z., Jacobs, D. and Basri, R. (1999) The role of convexity in perceptual completion: beyond good continuation. *Vision Res.*, 39: 4244–4257.
- Mamassian, P. (2004) Impossible shadows and the shadow correspondence problem. *Perception*, 33: 1279–1290.
- Mamassian, P. and Bühlhoff, H. H. (1996) Active kinetic depth effect. Technical Report No. 27. Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany.
- Mamassian, P. and Goutcher, R. (2001) Prior knowledge on the illumination position. *Cognition*, 81: B1–B9.
- Mamassian, P. and Landy, M.S. (1998) Observer biases in the 3D interpretation of line drawings. *Vision Res.*, 38: 2817–2832.
- Mamassian, P., Landy, M.S. and Maloney, L.T. (2002) Bayesian modelling of visual perception. In: Rao, R., Olshausen, B. and Lewicki, M. (Eds.), *Probabilistic Models of the Brain: Perception and Neural Function*. MIT Press, Cambridge, MA, pp. 13–36.
- Murray, S.O., Schrater, P. and Kersten, D. (2004) Perceptual grouping and the interactions between visual cortical areas. *Neural Netw.*, 17: 695–705.
- Norman, J.F., Phillips, F. and Ross, H.E. (2001) Information concentration along the boundary contours of naturally shaped solid objects. *Perception*, 30: 1285–1294.
- Tse, P.U. (2002) A contour propagation approach to surface filling-in and volume formation. *Psychol. Rev.*, 109: 91–115.
- Warren, P.E., Maloney, L.T. and Landy, M.S. (2002) Interpolating sampled contours in 3D: analyses of variability and bias. *Vision Res.*, 42: 2431–2446.