# Second-Order Visual Processing

**Andrew Derrington**

**To segregate the zebra from its background, does the visual system analyze the local spatial variation in contrast within visual channels, in the same way one might use grey-level to segregate a black horse, or color to segregate a piece of fruit in other images?**


1047-6938/01/01/0018/3-$0015.00 

It's not hard to spot a zebra, except perhaps among a herd of zebras like the one shown in the image on the facing page. The high-contrast stripes that are the zebra's most distinctive visual characteristic stand out from almost any background. One of the challenges for vision scientists is to understand the stages of human visual system processing that analyze spatial structures like the zebra. The zebra's key property in this respect is the fact that it stands out from the background by virtue of a second-order attribute: its stripes. This visual characteristic sets the zebra apart from objects distinguishable by virtue of their first-order attribute of color, such as being grey, white, black or red.

The zebra's stripes are rendered distinct from each other and from any but the most extreme background by primary stages of visual processing that, over the last 25 years, have been well-characterized in psychophysical and physiological experiments. When it comes to processing simple variations in grey-level we know both the overall performance of the human visual system and the physiological properties of relevant mechanisms at low levels of the primate visual system.

The challenge facing vision scientists is embodied in a simple question that springs from the distinct nature of each stripe, or in other words from the fact that it stands out both from neighboring stripes and from the background. Does the human visual system analyze the shape of the zebra by tracing the outline of each stripe and then assembling the stripes into the outline of the zebra? Or to segregate the zebra from its background, does the visual system use "stripedness," perhaps by analyzing the local spatial variation in contrast within visual channels, in the same way one might use grey-level to segregate a black horse, or color to segregate a piece of fruit in other images?

Yet after more than 25 years of research involving tasks in which observers are asked to analyze the spatial variations in texture patterns, there is no widely accepted answer. There are data to support both of the possibilities outlined above. For example, it is clear that when the patterns on the retina have very high contrast and very high spatial frequency, as can be the case when they are formed from laser interference fringes, very early in the visual pathway a non-linearity demodulates the pattern, making the contrast variation indistinguishable from a variation in luminance. On the other hand, when spatial frequency is much lower, it can require several seconds of inspection for an observer to distinguish between a perfect and an imperfect pattern of expanding motion made up of four independent patches (Fig. 1). Although this may not be quite long enough for the experimental observer to count every stripe in the stimulus, it is certainly more than an order of magnitude longer than the inspection time required to make the same discrimination using first-order stimuli (patches of sinusoidal grating).

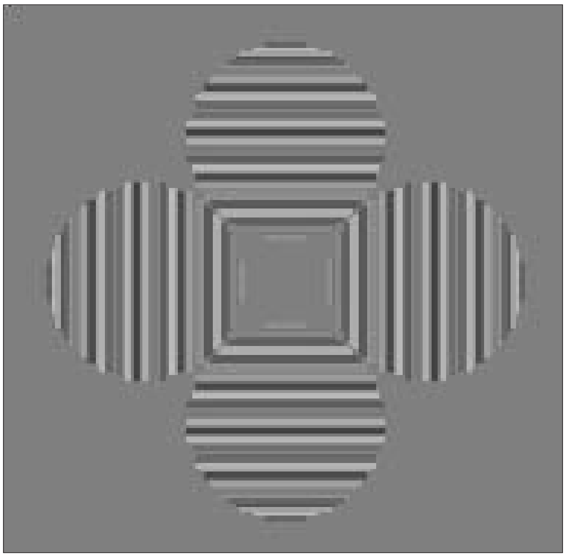

**Figure 1.** Diagram of expanding motion patches.[1] It takes observers 1.5 seconds or longer to distinguish a stimulus in which the contrast envelopes in all four patches move outwards from one in which three of the envelopes move outwards and one moves inwards.

It is appropriate to pose a more general form of the "zebra question" to encompass textures based on a range of higher-order attributes of image elements. These include direction and speed of motion, binocular disparity, flicker rate, contrast, orientation and spatial structure of texture elements. Posing the question in this way has enabled vision researchers to develop and test models of the relevant stages of visual processing.

A simple model can be applied to the processing of a very wide range of second-order stimuli. The model consists of two stages of filtering separated by a non-linearity. The first filter stage extracts a signal related to the texture. The non-linear stage (often a rectifier) makes the output of the first filtering stage available to the second stage of filtering which can extract spatial structure or direction of motion. By choosing appropriate textures and tasks it is possible to probe the nature of the filters and the nature of the non-linearity.

One central question is whether second-order analyses are hard-wired into the visual system or whether they depend on high-level cognitive processing. The complex motion task illustrated in Fig. 1 belongs almost certainly in the latter category. The observer has to distinguish whether an array of four patterns that surround the fixation point are all moving outwards, or whether three of the patches are moving outwards and one (selected at random) is moving inwards. Of course, if the visual field were tiled with appropriate motion analyzers, one might expect that the motion of all four patches would be analyzed concurrently in about the same time it takes to analyze the motion of a single patch. Indeed this is what happens when the patches are simple first-order luminance patterns (sinusoidal gratings). The viewing time required for successful performance is about 100 msecs.

However, the task becomes extremely difficult when the patches are second-order stimuli like those illustrated (beat patterns with a moving contrast envelope). Observers need to view the pattern for several seconds in order to perform better than 75% correct. They seem unable to analyze the patterns concurrently, and they give the impression that they are forced to search the four patches to look for one that is moving inwards. Yet it would be wrong to conclude from this experiment that second-order analyses all require conscious effort. Figure 2 shows a paradigm case and illustrates three separate properties of the visual mechanisms that signal the position of contrast modulations. First, these visual mechanisms operate concurrently in different retinal locations. Second, they can be adapted. Third, they pool contrast across orientations.

These properties can be illustrated by careful experimentation using Fig. 2 as follows. First, examine the three patches of contrast increments in panel (**a**) and of decrements in panel (**b**) and confirm that the patches appear to be aligned vertically. Then fixate for 10 or 20 seconds on the

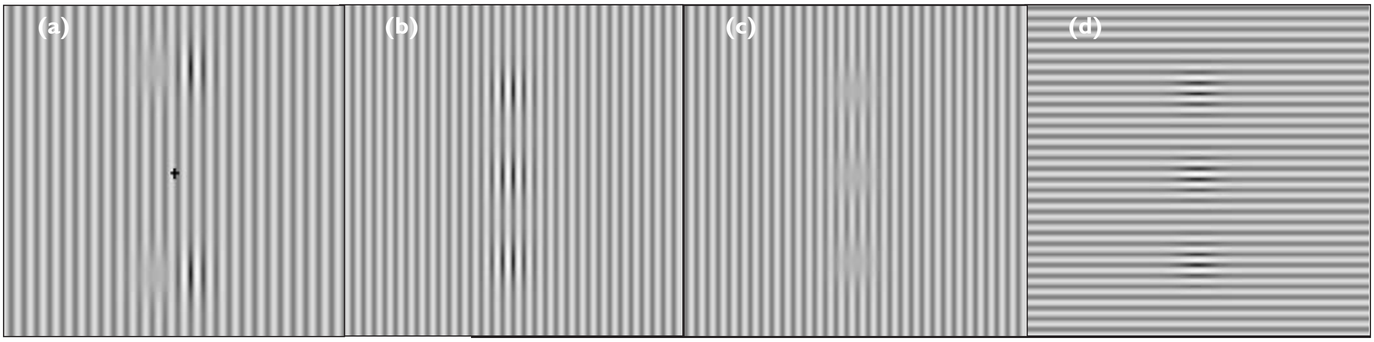


**Figure 2.** Spatial localization demonstration.[2]

cross in the center of panel (**a**) and re-examine (**b**) and (**c**). The patches no longer appear vertically aligned: they are shifted laterally in opposite directions. These lateral shifts are induced by adapting to the contrast gradients in the upper and lower parts of (**a**). Thus, panels (**a**), (**b**), and (**c**) demonstrate that the mechanisms that signal location of contrast modulations are susceptible to adaptation and that signals generated by mechanisms in different retinal locations can be examined concurrently—otherwise the judgments of co-linearity would be impossible.

Finally, by adapting to panel (**a**) of Fig. 2 and then inspecting panel (**d**), the reader will notice that the patches of high-contrast grating in (**d**) are also displaced in opposite directions. The mechanisms that signal their location are affected by adaptation to panel (**a**), even though the orientations of the first-order patterns (the carriers of the contrast modulation) in the two panels are 90° apart.

Of course, the reproduction of Fig. 2 may not be perfect, so the contrast variations might also contain unintended luminance variations. However, the demonstrations that can be carried out by adapting to this figure give qualitatively the same results as experiments carried out using contrast modulations presented on a calibrated monitor that contained no unintended luminance modulation.

These two sets of experiments merely scratch the surface of the problem of understanding second-order visual mechanisms. Later this year a special issue of JOSA A is to be dedicated to research on second-order visual processes.

## References

1. H. A. Allen and A. M. Derrington, Slow discrimination of contrast-defined expansion patterns, Vision Research, **40** 735-44 (2000).
2. P.V. McGraw, D. M. Levi, and D. Whitaker, Spatial characteristics of the second-order visual pathway revealed by positional adaptation, Nature Neuroscience **2**, 479-84 (1999).

**Andrew Derrington, School of Psychology, University Park, Nottingham, UK.**