EXCERPTED FROM

STEPHEN WOLFRAM

A NEW KIND OF SCIENCE

SECTION 10.7

Visual Perception

Visual Perception

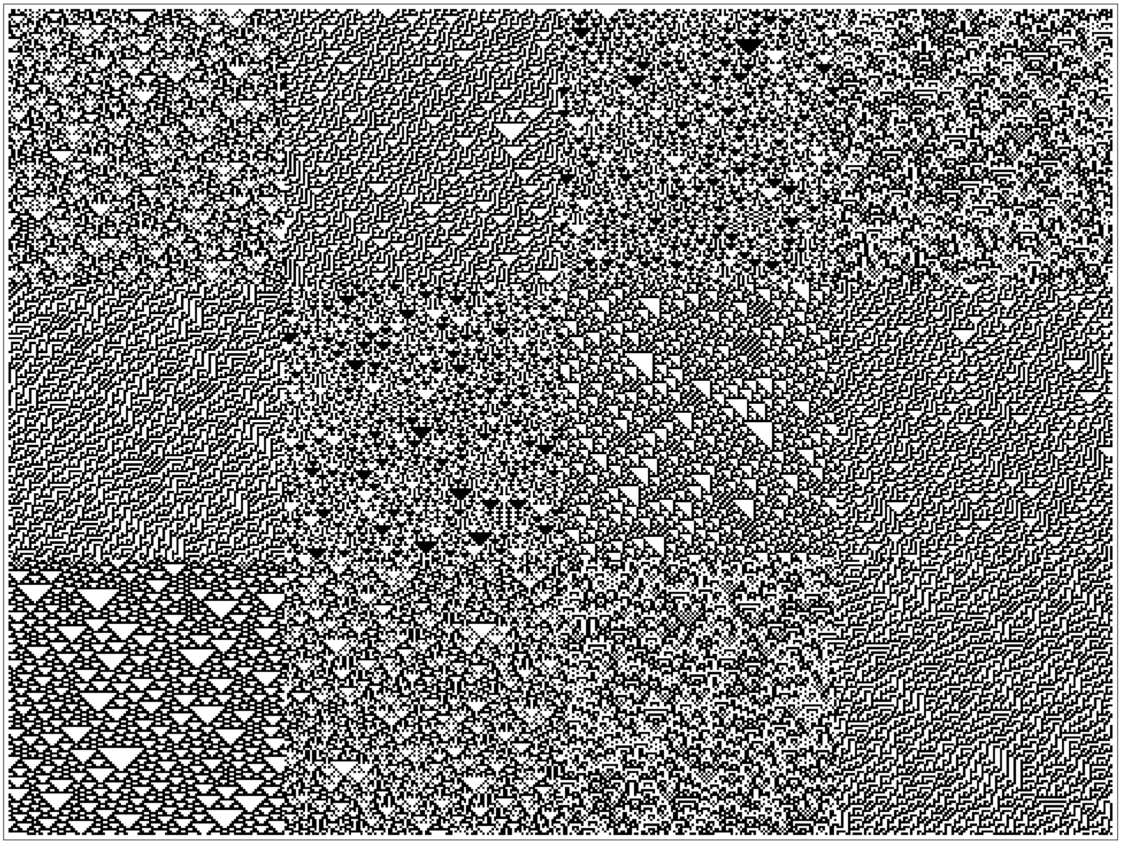
In modern times it has usually come to be considered quite unscientific to base very much just on how things look to our eyes. But the fact remains that despite all the various methods of mathematical and other analysis that have been developed, our visual system still represents one of the most powerful and reliable tools we have. And certainly in writing this book I have relied heavily on our ability to make all sorts of deductions on the basis of looking at visual representations.

So how does the human visual system actually work? And what are its limitations? There are many details yet to be resolved, but over the past couple of decades, it has begun to become fairly clear how at least the lowest levels of the system work. And it turns out—just as in so many other cases that we have seen in this book—that much of what goes on can be thought of in terms of remarkably simple programs.

In fact, across essentially every kind of human perception, the basic scheme that seems to be used over and over again is to have particular kinds of cells set up to respond to specific fixed features in the data, and then to ignore all other features.

Color perception provides a classic example. On the retina of our eye are three kinds of color-sensitive cells, with each kind responding essentially to the level of either red, green or blue. Light from an object typically involves a whole spectrum of wavelengths. But the fact that we have only three kinds of color-sensitive cells means that our eyes essentially sample only three features of this spectrum. And this is why, for example, we have the impression that mixtures of just three fixed colors can successfully reproduce all other colors.

So what about patterns and textures? Does our visual system also work by picking out specific features of these? Everyday experience suggests that indeed it does. For if we look, say, at the picture on the next page we do not immediately notice every detail. And instead what our visual system seems to do is just to pick out certain features which quickly make us see the picture as a collection of patches with definite textures.



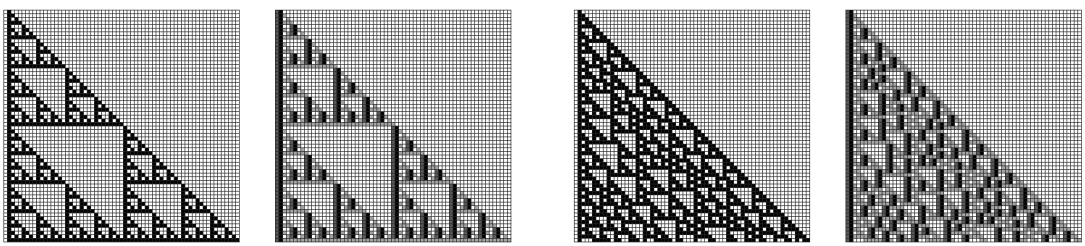
Patches generated by a variety of one-dimensional cellular automaton rules. Each patch is set up to have a roughly equal number of black and white squares. But despite this, our visual system quickly notices that different patches have different textures. And presumably this is because the visual system is automatically identifying particular features in each patch. Everyone appears immediately to be able to see some patches when shown this picture. But after looking at the picture for a while, the boundaries between the patches seem to get somewhat clearer.

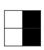
So how does this work? The basic answer seems to be that there are nerve cells in our eyes and brains which are set up to respond to particular local patterns in the image formed on the retina of our eye.

The way this comes about appears to be surprisingly direct. Behind the 100 million or so light-sensitive cells on our retina are a sequence of layers of nerve cells, first in the eye and then in the brain. The connections between these cells are set up so that a given cell in the visual cortex will typically receive inputs only from cells in a fairly small area on our retina. Some of these inputs will be positive if the image in a certain part of the area is, say, colored white, while others will be positive if it is colored black. And the cell in the visual cortex will then respond only if enough of its inputs are positive, corresponding to a specific pattern being present in the image.

In practice many details of this setup are quite complicated. But as a simple idealization, one can consider an array of squares on the retina, each colored either black or white. And one can then assume that in the visual cortex there is a corresponding array of cells, with each cell receiving input from, say, a 2×2 block of squares, and following the rule that it responds whenever the colors of these squares form some particular pattern.

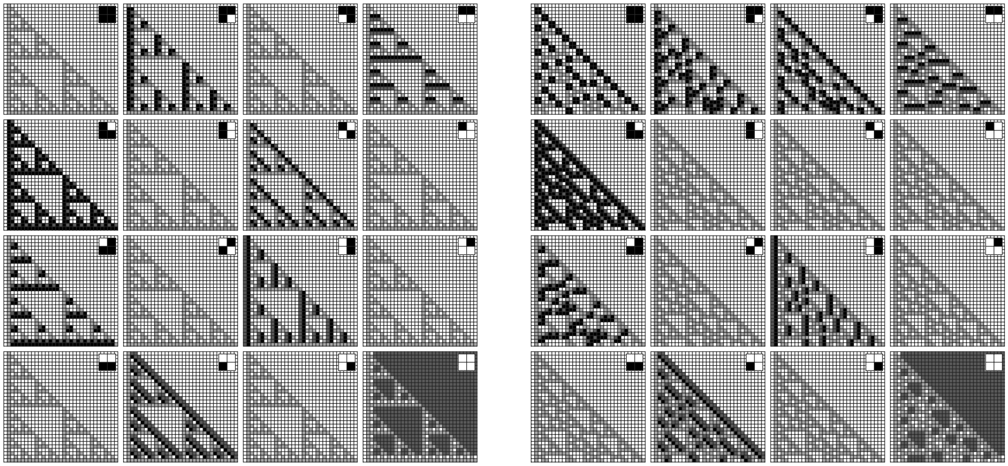
The pictures below show a simple example. In each case the first picture shows the image on the retina, while the second picture shows which cells respond to it. And with the specific choice of rule used here, what effectively happens is that the vertical black edges in the original image get picked out.





Responses to two sample images of cells sensitive to the 2×2 template shown on the left. The cells that respond are indicated by darker squares in the second picture in each pair. Such responses occur whenever the 2×2 template on the left appears, corresponding to the presence of a vertical black edge. The extraction of features by this kind of simple template matching appears to be a key element in human visual perception—as well as being common in technological image processing. The sample images used here are ones generated by the evolution of elementary one-dimensional cellular automata with rules 60 and 124 respectively.

Neurophysiological experiments suggest that cells in the visual cortex respond to a variety of specific kinds of patterns. And as a simple idealization, the pictures on the next page show what happens with cells that respond to each of the 16 possible 2×2 arrangements of black and white squares. In each case, one can think of the results as corresponding to picking out some specific local feature in the original image.

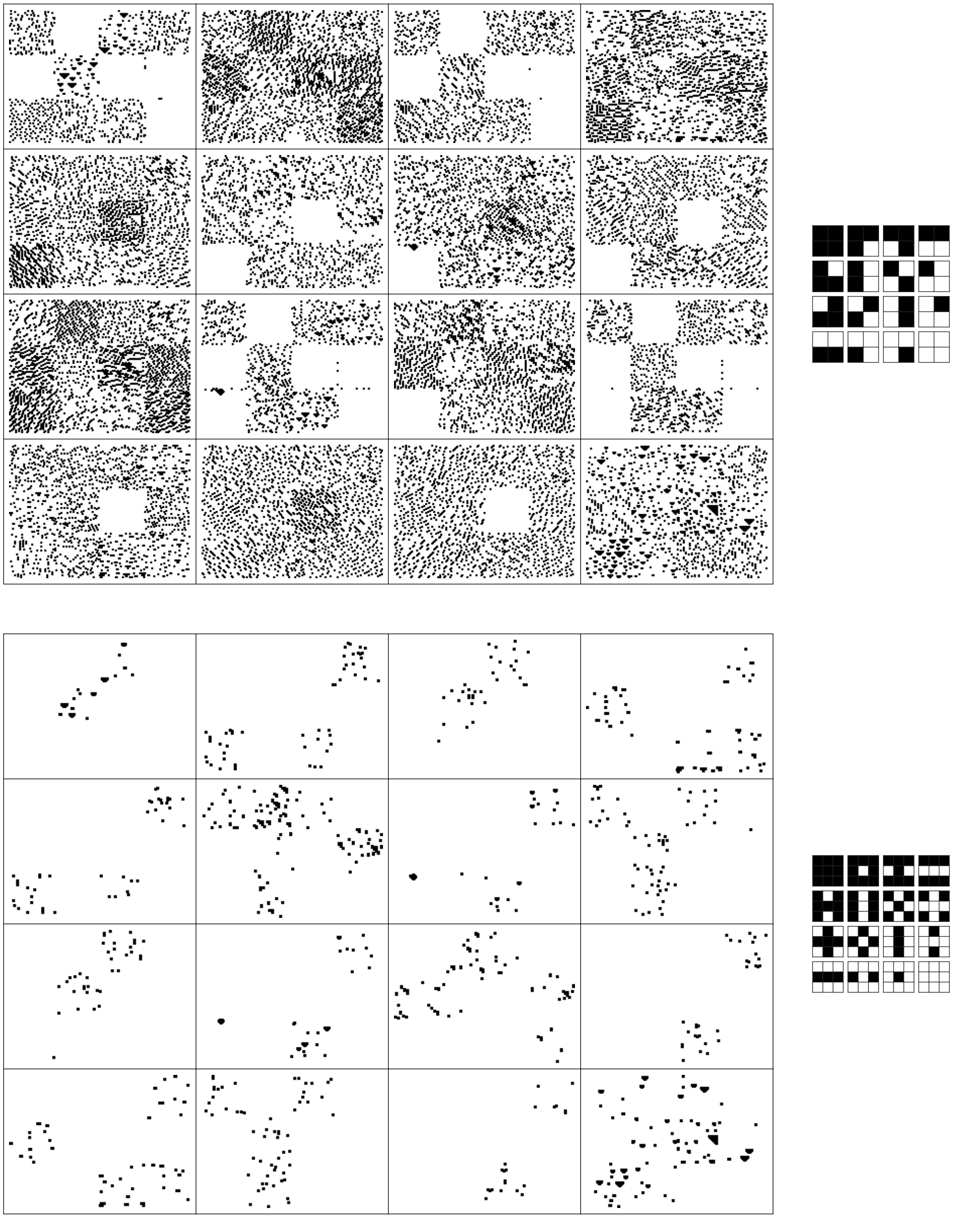


Responses to the sample images from the previous page by types of cells sensitive to each of the local arrangements of black and white squares shown. In each case, one can think of the resulting patterns as being filtered versions of the original images in which only parts that exhibit particular features are kept. The patterns can also be viewed as outputs from a single step in the evolution of two-dimensional block cellular automata in which the rules specify that a block becomes dark if it has the arrangement of cells shown, and becomes light otherwise. The comparative sparsity of dark blocks is a consequence of the fact that at any given position a dark block can occur in only one of the 16 cases shown. The absence of any dark blocks in many of the cases shown can be viewed as a reflection of constraints introduced by the construction of the images from one-dimensional cellular automaton rules.

So is this very simple kind of process really what underlies our seemingly sophisticated perception of patterns and textures? I strongly suspect that to a large extent it is. An important detail, however, is that there are cells in the visual cortex which in effect receive input from larger regions on the retina. But as a simple idealization one can assume that such cells in the end just respond to repeated versions of the basic 2×2 patterns.

So with this setup, the pictures on the facing page show what happens with an image like the one from page 578. The results are somewhat remarkable. For even though the average density of black and white squares is exactly the same across the whole image, what we see is that in different patches the features that end up being picked out have different densities. And it is this, I suspect, that makes us see different patches as having different textures.

For much as we distinguish colors by their densities of red, green and blue, so also it seems likely that we distinguish textures by their

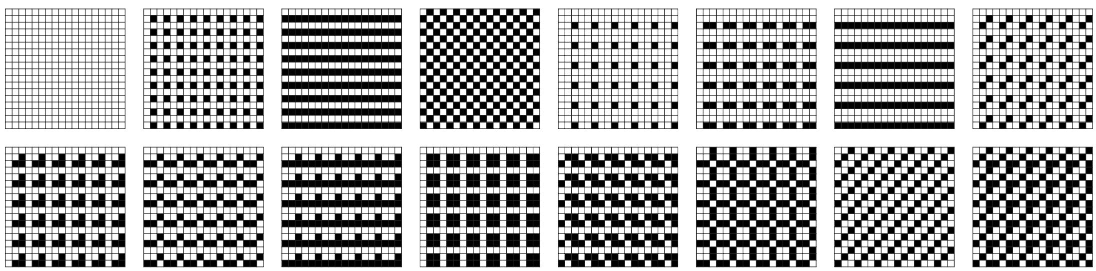


Responses to a smaller version of the image from page 578 by cells sensitive to all 16 possible 2 × 2 blocks, as well as their repetitive 3 × 3 extensions. Patches which appear to have different textures in the original image are seen to contain characteristically different densities of these various blocks. I strongly suspect that it is density differences such as these that allow our visual system to distinguish textures.

densities of certain local features. And the reason that this happens so quickly when we look at an image is no doubt that the procedure for picking out such features is a very simple one that can readily be carried out in parallel by large numbers of separate cells in our eyes and brains.

For patterns and textures, however, unlike for colors, we can always get beyond the immediate impression that our visual system provides. And so for example, by making a conscious effort, we can scan an image with our eyes, scrutinizing different parts in turn and comparing whatever details we want.

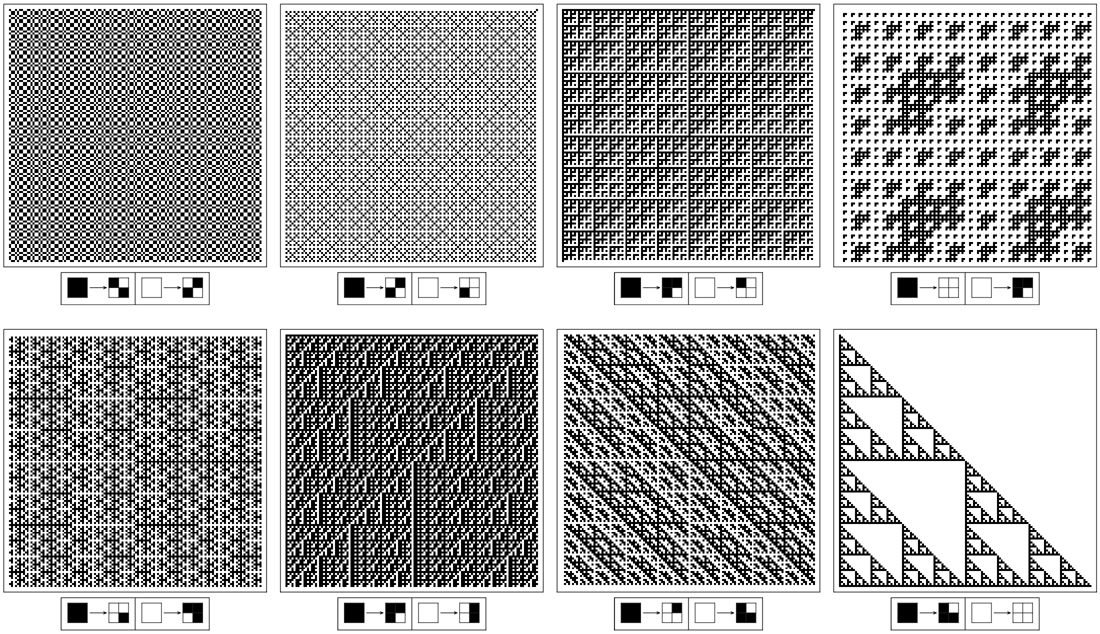
But what kinds of things can we expect to tell in this way? As the pictures below suggest, it is usually quite easy to see if an image is purely repetitive—even in cases where the block that repeats is fairly large.



Examples of all the distinct repetitive patterns that can be formed from arrays of 2 × 2 and 3 × 3 blocks. In every single case the presence of pure repetition is easy to recognize by eye. Note that in a pattern generated by repeating one particular block, there will normally be other blocks that occur with other alignments. Page 215 shows patterns obtained in systems based on constraints in which one effectively requires that only certain blocks or sets of blocks occur.

But with nesting the story is quite different. All eight pictures on the facing page were generated from the two-dimensional substitution systems shown, and thus correspond to purely nested patterns. But except for the last picture on each row—which happen to be dominated by large areas of essentially uniform color—it is remarkably difficult for us to tell that the patterns are nested. And this can be viewed as a clear example of a limitation in our powers of visual perception.

As we found two sections ago, many standard methods of data compression have the same limitation. But at the end of that section I showed that the fairly simple procedure of two-dimensional pointer



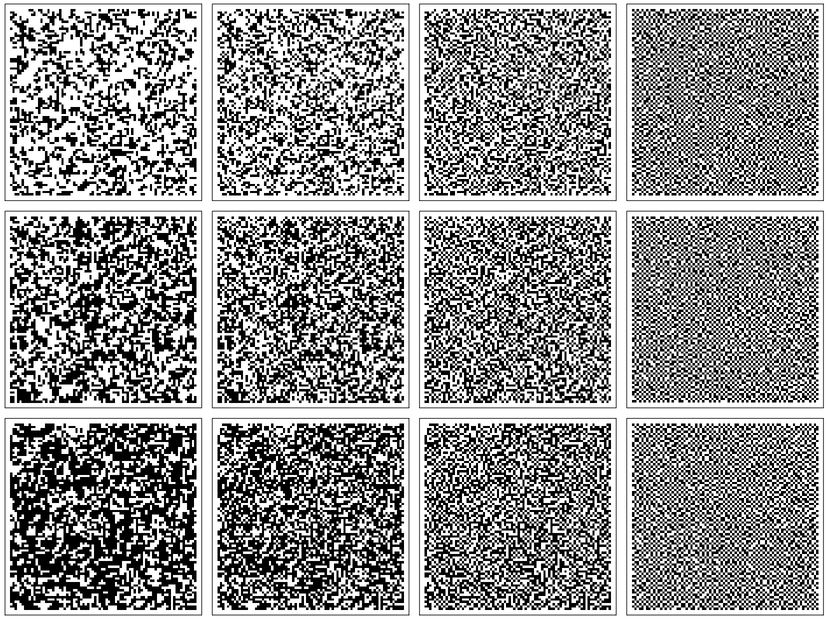
Examples of nested patterns created by following the two-dimensional substitution rules shown. Except for the last examples on each row, it is remarkably difficult to recognize the nested structure in these patterns by eye, even with quite careful scrutiny. The two-dimensional pointer-based encoding scheme from page 571 does however manage to recognize the structure in all cases.

encoding will succeed in recognizing nesting. So it is not that nesting is somehow fundamentally difficult to recognize; it is just that the particular processes that happen to occur in human visual perception do not in general manage to do it.

So what about randomness? The pictures on the next page show a few examples of images with various degrees of randomness. And just by looking at these images it is remarkably difficult to tell which of them is in fact the most random.

The basic problem is that our visual system makes us notice local features—such as clumps of black squares—even if their density is consistent with what it should be in a completely random array. And as a result, much as with constellations of stars, we tend to identify what seem to be regularities even in completely random patterns.

In principle it could be that there would be images in which our visual system would notice essentially no local features. And indeed in



Examples of images that approximate perfect randomness. The second image on each row has squares chosen independently to be black with probabilities 0.4, 0.5 and 0.6 respectively. In the other images various features are added or removed. In the first image on each row, if any square is surrounded by four squares with identical colors, then the square is forced to have the same color. In the third image, any clump of squares with the same color is broken up by reversing the color of the center square. And in the fourth image, the same is done with lines of squares of the same color.

the last two images on each row above all clumps of squares of the same color, and then all lines of squares of the same color, have explicitly been removed. At first glance, these images do in some respects look more random. But insofar as our visual system contains elements that respond to each of the possible local arrangements of squares, it is inevitable that we will identify features of some kind or another in absolutely any image.

In practice there are presumably some types of local patterns to which our visual system responds more strongly than others. And knowing such a hierarchy, one should be able to produce images that in a sense seem as random as possible to us. But inevitably such images would reflect much more the details of our process of visual perception than they would anything about actual underlying randomness.