Chapter 2

Inevitability of Illusions

This chapter primarily concerns a very general constraint on brains: that they take time to compute things. This simple fact has profound consequences for the brain, and vision in particular. I will put forth evidence that it is the visual system's attempting to deal with this computing delay that explains why we experience the classical geometrical illusions. Figure 2.1 shows a sample such illusion; basically, the illusions are those found in any introductory Psychology course. I will also, along the way, briefly discuss a general approach to modeling brain computation: that approach is decision theory, wherein the brain, or some portion of it, is modeled as an ideal rational agent acting to maximize its expected utility on the basis of probabilities concerning the nature of the uncertain world. This is referred to as the Bayesian framework for visual perception, and with it researchers have made some important breakthroughs. We will need to understand it, and its shortcomings, to understand how the visual system copes with the time it takes to compute a percept. I also discuss the difficulties of one of the older and more established inference-based theories of the geometrical illusions. Before proceeding, it is important to understand why there may be computing delays in perception.

Computation is sometimes slow, sometimes fast, but never instantaneous. Computation takes time. Running software on your computer takes time. For example, it takes about one second to start Microsoft Word on my lap top, over two seconds to start Adobe Acrobat, and over half a minute to run LaTex with this book as the input. Despite the orders of magnitude increase in computation speed over the last twenty years since the advent of the personal computer, there seems to always be significant delays for contemporary software. This is presumably because software producers have figured out the time delays

M. A. Changizi, *The Brain from 25,000 Feet*

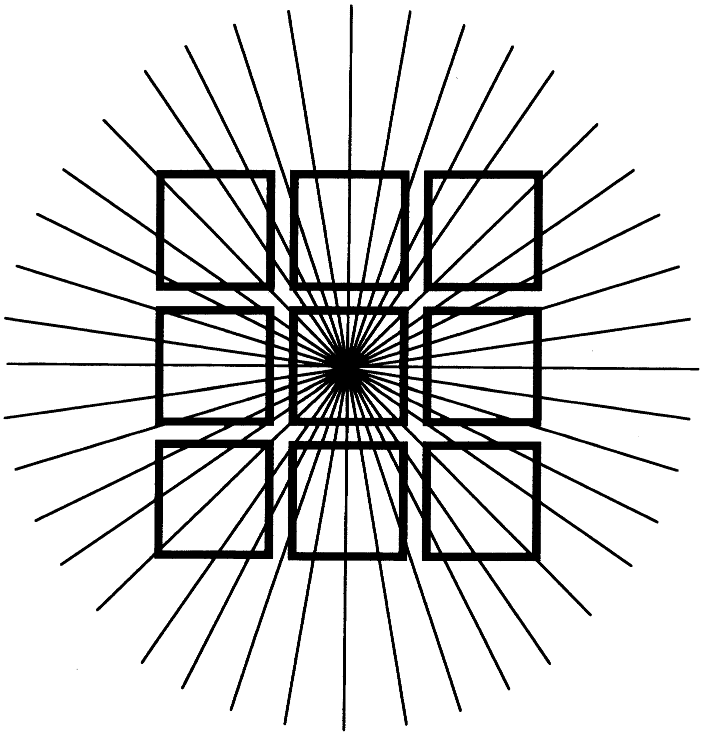

Figure 2.1: Nine perfect, identical squares on a radial display induce an illusion, which is a version of the Orbison illusion.

consumers are willing to put up with and can use this time to carry out more sophisticated computations for the consumer.

Brain computation takes time as well. In addition to the computation delays due to simply traveling through individual dendrite and axon arbors, and to the time it takes signals to traverse synapses, computation delays are also due to the complex time course and pattern of neural firings that actually implement the computation. How much time can the brain afford to take in carrying out its computations? To answer this, consider the brain (and evolution) as the software producer, and the animal (and his genes) as the consumer. The brain will presumably have figured out the time delays the animal is willing to put up with—i.e., delays that the animal is able to deal with without compromising survival too much—so as to be able to use this time to compute more powerful functions of use to the animal. More exactly, the brain and evolution presumably will have discovered how to optimally trade off computation time with computational power. How much time is given to computations in this optimal trade-off will depend on the details of the animal's ecology, but it seems *a priori* unlikely to be exceedingly long—e.g., 10 second delays—or microscopically short—e.g., 0.001 seconds. Because the world changes too much and too unpredictably during a long, say 10 second, interval, long delays will lead to computational solutions that are moot by the time they are computed. Nearly instantaneous computations would avoid this problem, but would leave the brain with too little time to compute much of interest to the animal. Somewhere in between these extremes will be an optimal middle ground, allowing sufficient time for powerful computations, but the time is short enough that the computations are still applicable to the changing world. These considerations are relevant for any brain—Earthly or not—having to deal with an uncertain and dynamic world, so long as they are not literally infinite in computational speed.

One effective possible strategy for a brain to use in its attempt to increase its computation time is to attempt to *correct* for the computation delay (De Valois and De Valois, 1991; Nijhawan, 1994, 1997, 2001; Berry et al., 1999; Sheth et al., 2000; Schlag et al., 2000; Khurana et al., 2000; Changizi, 2001). That is, suppose it would be advantageous to have a time interval Δt to carry out some useful computation, but suppose that Δt is long enough that the world typically has changed to some degree during this time, making the computation moot. What if, to deal with this, the brain took a different tact? Rather than trying to compute something that is useful for dealing with the world the way it *was* when the computation started, the brain might try, instead, to compute a function that *will* be useful for dealing with the world as it probably will be by the time the computation is finished. Such a strategy might be called *latency correction.* To the extent that latency correction is possible, the brain will extend its computation duration to derive more powerful functionality. At some point the computation interval will be so long that latency correction algorithms will no longer reliably work, but such a strategy will buy the brain more time to provide neater software for the animal, thereby increasing the animal's prospects.

Vision is one kind of brain computation that is needed swiftly and is difficult to compute. The visual system computes from the retinal stimulus a perception of the way the world out there is, and since the world is typically in flux either because it is itself changing or because the observer is himself moving, the percept must be computed in a timely manner lest the information from the retinal stimulus be irrelevant. Visual perception is also difficult: it is a classic example of an underdetermined problem, as there is no unique solution to it, there being (infinitely) many possible ways the world could be that would lead to the information on the retina (see also Chapter 3). Our own artificial computer algorithms for vision, despite a few decades of progress, still fall far short of success, where success is defined as the recognition of or navigation within scenes under a wide variety of circumstances. Because vision is difficult, to do a good job at it the visual system would like to have as much time as it reasonably can. In fact, the visual system in mammals *does* take a significant, but not exceedingly long, period of time: there is a latency on the order of magnitude of 100 msec (Lennie, 1981; De Valois and De Valois, 1991; Maunsell and Gibson, 1992; Schmolesky et al., 1998). This is ecologically significant because a lot can happen in 100 msec, or a tenth of a second. Even walking at just one meter per second means that the positions of objects change by 10 cm during that time. If the visual system generated a percept of the way the world probably was when the information was picked up at the retina, the percept would be about the way the world probably was 100 msec in the past. At one m/sec, objects perceived by an observer to be within 10 cm of being passed would, in fact, already have passed the observer... or the observer will have bumped into them. Catching a ball and the other complex activities we engage in obviously worsen this problem.

Latency correction is thus a beneficial strategy, if the visual system can carry it off. That is, the strategy is this: rather than computing a percept of the scene that probably caused the retinal stimulus—a percept that would need to be generated nearly instantaneously to be of much use to the animal—the visual system can, instead, compute a percept of the scene that will probably be out there by the time the computation is finished and the percept is elicited. That is, the visual system attempts to perceive not the past, but, instead, to "perceive the present." In this way the visual system can generate percepts that are typically coincident with reality, but it can also secure itself some elbow room for solving the tough problem of vision.

If a visual system were able to implement latency correction, what kind of algorithm might we expect it to employ? To answer this, let us consider what a latency correction algorithm would have to do. In order to reliably generate a percept at time t of what is out there at time t on the basis of retinal information from $t - 100$msec, the visual system would need to solve the following two conceptually distinct problems.

1. The visual system must figure out what the scene at time $t - 100$msec probably was (e.g., a 10 meter flag pole 5 meters away), and

2. the visual system must determine what scene that scene will probably become by time t (e.g., a 10 meter flag pole 4.5 meters away).

[Note that a *scene* consists of the properties of the objects in the vicinity of the observer, including the observer's viewpoint. Thus, a room viewed from a different position would make for a different scene.]

Each of these problems is an *inference problem*, as it is underdetermined by any information the observer may have. The visual system must infer what might be out there at time $t - 100$msec (the time of the retinal stimulus), even though there are infinitely many scenes that can, in principle, have led to the same information on the retina. And the visual system must also infer how the scene will probably change, even though there are infinitely many ways that the scene might, in fact, change. [I am not claiming that a brain must actually make this distinction between 1 and 2. A brain could solve the latency correction "all at once," but it still would have conceptually dealt with both problems.]

2.1 Visual inferences

Therefore, if the visual system could carry out latency correction, it would have to be good at making inferences. But making inferences is something the visual system actually *is* good at, as has been noticed at least since Helmholtz (1962), and has been taken up by many since (e.g., Gregory, 1997). The visual system appears to act like a scientist, using the evidence present in the retinal stimulus to make a reasoned choice. The visual system also acts like a scientist in that it can learn from past experience. Finally, the visual system is even like a scientist in that it is also simply born with certain biases, or preconceptions, toward some perceptual hypotheses over others. (In fact, it *must* be born with such biases; see Chapter 3.) In this section I discuss two research paradigms within this inference tradition.

Traditional visual inference

My main task for this chapter is to show how the classical geometrical illusions are consequences of the visual system implementing a latency correction strategy. Since, as we discussed earlier, latency correction is something we might expect from any brain with finite computing speed, we also expect any such brain to perceive illusions. But there already exist many theories of the visual illusions; what are wrong with them? First, note that I am only interested here in considering theories of visual perception that concern the purported function computed by the visual system, and also the general kind of algorithm used. I am not interested here in theories about the implementation-level mechanisms found in the visual system (e.g., lateral inhibition, or some neural network). One of the most venerable and most well-entrenched such (functional) theories of the geometrical illusions is what I will call the *traditional inference approach* (Gregory, 1963, 1997; Gillam, 1980, 1998; Rock, 1975, 1983, 1984; Nundy et al., 2000).

Before stating what the general form of this kind of theory is, it is useful to present a sample stimulus with which I will introduce the theory. Consider Figure 2.2, where observers perceive the bold vertical line on the right to have greater angular size than the bold vertical line on the left; this is the illusion. Note that observers *also* perceive the *linear* size of the line on the right to be greater; that is, they perceive that it is a taller object in the depicted scene, when measured by a ruler in, say, meters; and they also perceive that it is farther away. But this latter perception of linear size is not what is illusory about the figure: no one is surprised to learn that observers perceive that the line on the right has greater linear size in the depicted scene. What is illusory is that observers perceive the line on the right to have greater *angular size*—to fill more of the visual field—than the line on the left, despite their angular sizes being identical.

The traditional inference explanation for this illusion states that the line on the right is perceived to be longer because the cues suggest that it probably *is* longer. Describers of the theory will usually also say that such a perception is

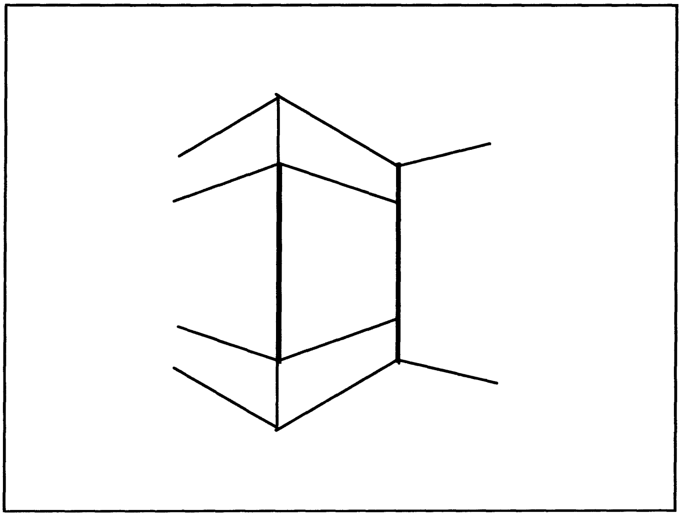

Figure 2.2: An illusion which is a variant of the Müller-Lyer. The two bold vertical lines are the same angular size, but the right one appears to have greater angular size. One of the most commonly accepted functional explanations for this illusion is an inappropriate inference explanation which says that the line on the right is perceived to be bigger because the cues suggest that it is *a bigger line out there. The cues suggest this as follows: the right line is nearer to the vanishing point of the converging lines and thus is probably farther away, and since it has the same angular size as the other line, it follows that it must be bigger. It is "inappropriate" because, in this case, the lines are at the same distance, namely they are both on the page. The deep problem with this explanation is that it equivocates between two notions of perceived size: perception of angular size, and perception of linear (or objective) size. Because the right line probably is bigger in linear size, we should perceive it to be bigger in linear size. Fine. But observers also perceive the right line to be bigger in angular size, and its probably being bigger in linear size does* not *imply that it is bigger in angular size. They are, in fact, probably the* same *angular size, since they project identically onto the retina. The traditional inference explanation therefore cannot explain the illusion.*

useful for us in the real-world scene version of Figure 2.2—i.e., when you are standing in front of a real hallway—but when the stimulus is from a piece of paper as it *actually* is in this figure, this perceptual strategy is said to become "inappropriate." There is, however, a deep conceptual problem with this explanation. To start, let us look again at the main statement, which is along the lines of

> The line on the right is perceived to be longer because the cues suggest that it probably *is* longer.

What does the statement mean by 'longer'?

The *first possibility* is that it means 'greater linear size.' That is, the statement would be,

> The line on the right is perceived to have greater linear size (e.g., in meters) because the cues suggest that it probably *is* greater in linear size.

The statement in this case would be fine, as far as it goes, since it is certainly useful to perceive the linear size to be what it probably is. For example, if the line on the right is probably three meters high, then it is appropriate to perceive it to be three meters high. However, this interpretation is no longer relevant to the illusion, since the illusion concerns the misperception of their angular sizes.

The *second possible interpretation* is that 'longer' means 'greater angular size,' in which case the statement becomes,

> The line on the right is perceived to have greater angular size (measured in degrees) because the cues suggest that it probably *is* greater in angular size.

This, however, is no good because the cues do *not* suggest that the line on the right has greater angular size. The lines have, in fact, identical angular size, and the visual system "knows" this since equal angular sizes are unambiguously projected onto the retina. And it is a fallacious argument to say that the angular size of the line on the right is probably greater because it's linear size is probably greater; linearly huge objects very often have tiny angular size (e.g., the moon), and linearly tiny objects often have tremendous angular size (e.g., hold your finger up near your eye).

So far, the traditional inference explanation statement is either irrelevant (the first interpretation) or false because the cues do *not* suggest that the line on the right has greater angular size (the second interpretation).

The *third and final possible interpretation* I will consider is that the *first* occurrence of 'longer' is interpreted as 'greater angular size' and the *second* occurrence of 'longer' is interpreted as 'greater linear size.' That is, in this possibility the statement is equivocating between two meanings of 'longer.' The statement is now,

> The line on the right is perceived to have greater angular size (measured in degrees) because the cues suggest that it probably is greater in linear size (bigger in meters).

This appears to be the interpretation that people actually have when they utter this view. It is sometimes even phrased as something along the lines of, "the perception of the projective properties of the lines are biased toward the probable objective properties of the lines." The statement is not irrelevant as in the first interpretation; this is because the claim concerns the perception of angular size, which is what the illusion is about. The statement also does not err as in the second interpretation by virtue of claiming that the line on the right probably has greater angular size. One preliminary problem concerns what it could possibly mean to bias a projective property toward an objective property; how can something measured in degrees get pushed toward something that is measured in, say, meters? Another issue concerns *how much* the angular size should be increased in the probably-linearly-longer line; the explanation gives us no apparatus by which it is possible to say. I will focus on another problem, which concerns the supposed *usefulness* of such a strategy for vision: of what possible use is it to perceive a greater angular size merely because the linear size is probably greater? The visual system's goal according to these traditional inference approaches is to generate useful percepts, and, in particular, to generate percepts that closely represent reality (because this will tend to be useful). To accurately represent the angular sizes in Figure 2.2 would be to perceive them as being identical in angular size. The visual system would *also* want to perceive them as having different linear sizes, but there is no reason—at least none that this traditional inference explanation gives—for the visual system to misperceive the angular sizes.

It is sometimes said that the illusion is only an illusion because Figure 2.2 is just on a piece of paper. The inferential strategy of increasing the perceived angular size of the line on the right because it is probably linearly longer is inappropriate in this case because, it is said, the figure is just a figure on a page, where the lines in fact have the same linear size. If, the argument continues, the proximal stimulus were, instead, due to a real live scene, then the strategy would be appropriate. Unfortunately, the strategy would be inappropriate in this latter scenario too. To see this, let us imagine that the stimulus is not the one in Figure 2.2, but, instead, you are actually standing in a hallway of the kind depicted, and your eye position is placed in just such a manner that the line on the right has the same angular size as the one on the left. Is there anything "appropriate" about perceiving the line on the right to have greater angular size merely because its linear size is probably greater? It is not clear what would be useful about it, given that its angular size is the same as that of the line on the left, and perceiving their angular sizes to be equal does not preclude perceiving their linear sizes to differ. (E.g., hold your finger out until it fills just as much of your visual field as a tree off in the distance. You now perceive their angular sizes to be identical, but you also perceive the tree to be linearly larger.)

Some may think I have constructed a straw man position for the traditional inference explanation, and that the authors behind such explanations have more sophisticated positions. Perhaps this is so, although I do not think so; I have no interest, however, in whether or not this is *really* the explanation they intended. What is important is that the idea as I stated it is what the "average psychologist and neurobiologist on the street" appear to understand the explanations to be. For example, pulling out the nearest undergraduate perception textbook to me, the cogent author describes the traditional inference explanation for a figure essentially just like Figure 2.2 as follows.

> ...the converging lines are unconsciously interpreted as parallel lines receding into the distance...and the [vertical] lines as lying in the same receding...plane as the converging lines.... The unconscious perception of differential depth leads to the conscious perception of differential size: The [right] line would have to be longer because it...connects the receding parallel lines, whereas the lower one is not even close [Palmer, 1999, p. 324].

Now, as we will see later in this chapter, there *is* a good reason to perceive the angular size of the line on the right to be greater: namely because its angular size probably *will* be greater by the time the percept is actually generated (due to the observer's probable forward movement toward the vanishing point). However, the traditional inference explanation of the geometrical illusions provides no such reason, and is thus, at best, an explanation that provides no real explanation for why the visual system would generate the illusions.

In addition to the above conceptual difficulties, the traditional inference approach has more run-of-the-mill difficulties in explaining the illusions. As one example, consider the Orbison illusion where the square is directly below the vanishing point (see Figure 2.1). The square in this case appears to project as a trapezoid, with its longer edge on top. To explain this in the traditional inference manner, one needs to argue that the top edge of the projection is actually due to a real world line that is bigger in meters than the bottom edge of the square. For this to be the case, the projected square would have to be due to a real world trapezoid with its top edge tilted backward. The difficulty is: Why would such a tilted trapezoid be the probable source of a square projection? This is a highly coincidental, or non-generic (Freeman, 1994), projection for such an object. It seems obviously much more probable that the source of the perfectly square projection is a square in the observer's fronto-parallel plane and near the vanishing point. But in this case, the top and bottom of the object are identical in length, and so the traditional inference approach predicts no illusion. Other explanations by the traditional approach require similarly improbable sources. For example, in the Hering illusion (Figure 2.19), the probable source of two vertical lines on either side of the vertical meridian cannot be that they are two vertical lines, for then the distance in meters between each line would be the same and the traditional inference account would predict no illusion. Instead, for traditional inference to work here, the probable source would have to be that the two lines bend away from the observer, and as they bend away, they also get farther apart in meters; and all this in just such a manner that they happen to project perfectly straight. With this strange source, the lines are farther apart in meters when nearer to the vanishing point, which is why they are perceived to bow out according to the traditional inference approach. However, it seems much more plausible that the probable source of the two lines is that they are two vertical lines.

2.1.1 The standard Bayesian approach

In recent years the visual-system-as-inference-engine approach has been reinvigorated by a *Bayesian approach* to inference. There are many ways of modeling inference, but the Bayesian framework is a particularly good one. I will not discuss it in detail here, but will only try to communicate what is so good about it. [See Chapter 3 for an introduction to the Bayesian framework.]

The basic idea is that an agent has a numerical degree of confidence in each of the perceptual hypotheses, the hypotheses which are mutually exclusive. These degrees of confidences are modeled as probabilities, where each hypothesis has a probability in the interval from 0 to 1, the sum of the probabilities over all the hypotheses equals 1, and the probability of no hypothesis being true is 0. A probability of 1 for a hypothesis means that the agent has complete confidence in the hypothesis. A probability of 0 means the agent has complete confidence that the hypothesis is *not* true. The Bayesian framework tells us how these probabilities should be altered when evidence, or retinal information, is accumulated. This approach is, in a certain sense, optimal, because if you do *not* follow this approach, then others can dupe you out of all your money; I am here intimating an important result called the *Dutch Book Theorem*, or the Ramsey-de Finetti Theorem (Ramsey, 1931; de Finetti, 1974; see also Howson and Urbach, 1989, pp. 75–89 and 99–105, for discussion).

It is not only a nice framework because of this kind of optimality argument, it is also nice because it makes certain conceptual distinctions that allow us, the scientists, to make better sense of the inferential process. In particular, the framework distinguishes between

- *prior probabilities*, which are the probabilities in the hypotheses before seeing the evidence,
- *likelihoods*, which are the probabilities that the evidence would occur given that a hypothesis were true, and
- *posterior probabilities*, which are the probabilities in the hypotheses after seeing the evidence.

The reader is invited to read the introduction to the Bayesian framework in Chapter 3, but it is not necessary to cover it in any detail here.

The main idea to get across is that the inference-engine idea appears to apply well to the human visual system, and has been taken up during the 1990s within the Bayesian framework (Knill and Richards, 1996), where considerable success has been made: e.g., the perception of 3D shape (Freeman, 1994), binocular depth (Nakayama and Shimojo, 1992; Anderson, 1999), motion (Kitazaki and Shimojo, 1996), lightness (Knill and Kersten, 1992) and surface color (Brainard and Freeman, 1997).

In fact, if the visual system truly can be described within a probabilistic framework, then the proper treatment is a decision theoretic one, where the brain is treated as attempting to maximize its expected utility. That is, perception is an act, and an agent cannot decide how to act purely on the basis of the probabilities of hypotheses. For example, suppose there are two main possibilities concerning the scene that caused the retinal stimulus: the first is that there is a bed of flowers, and the second is that there is a tiger. Even if a flower bed is more probable than the tiger, the costs are so high for not recognizing a tiger that the perception that maximizes your expected utility may be the tiger perception. We would therefore expect that visual perception should be modifiable by modifying only the utilities of the observer, and evidence exists that even appetitive states such as thirst can modulate low-level perceptions such as transparency (Changizi and Hall, 2001).

Although the Bayesian approach has allowed significant advances in understanding visual perception, there is a difficulty with the way in which it is typically conceived. It is always assumed, either explicitly or implicitly, that the visual system is attempting to generate a percept of the scene that probably caused the retinal stimulus. That is, the "standard Bayesian perception approach" is to assume that the perceptual hypotheses are about the various possible scenes that are consistent with the retinal stimulus actually received. So, for example, when we say that a stimulus is bistable (such as the Necker cube, which is just a line drawing of a wire cube), we mean that the visual system jumps back and forth between two percepts of scenes that are consistent with the stimulus. The possible percepts are confined to percepts of scenes that could have caused the retinal stimulus. It is not, then, possible within the standard Bayesian perception approach to have percepts of scenes that are not even consistent with the retinal stimulus. The standard Bayesian approach can only accommodate *consistent perception*. Note that misperception can be consistent perception, since the perception could be of something that is not actually there, but is nevertheless consistent with the retinal stimulus. Many of our perceptions are consistent with the retinal stimulus, and the standard Bayesian approach is fine in such cases. For example, the examples of Bayesian successes I mentioned earlier—perception of 3D shape, binocular depth, motion, lightness and surface color—appear to be consistent perceptions. E.g., for the motion aperture phenomenon there are many different possible motions consistent with a line moving behind a circular aperture; it is a case of consistent perception since subjects appear to perceive one of the possibilities consistent with the stimulus.

The difficulty for the standard Bayesian perception approach lies in the fact that there are many perceptual phenomena where the observer perceives a scene that could *not* have caused the retinal stimulus. That is, there are cases of *inconsistent perception*. For example, the geometrical illusion from the earlier Figure 2.1 is an example of inconsistent perception. The angles of the squares in the figure project toward your eye at nearly 90°, supposing you are looking straight at it and are not too close. Yet many or all of the projected angles are perceived to be significantly different from 90°. Why is this a case of inconsistent perception? Because the actual stimulus does *not* project (much) differently than 90°.

[Note that if proximal stimuli possessed significant errors, it *would* be possible for the standard Bayesian perception approach to handle inconsistent perception. For example, suppose that an object projects toward an observer with an angular size of θ, but that the retina records this angular size with error according to some normal distribution. Then an ideal probabilistic engine would realize that projected lines in the world can have angular sizes markedly different from the angular size measured by the retina, and could sometimes generate perceptual hypotheses inconsistent with the proximal stimulus, but hopefully consistent with the true angular size. However, this does *not* appear to be relevant for the retina and visual system; at least, any error for angular sizes (and projected angles) are negligible.]

To help drive home the point, consider Figure 2.3. There is an object X in the lower half of the figure, and whatever it may be, it is projecting toward your eye as a perfect square. No, you certainly are not perceiving it to project as a perfect square, but we'll get to that in a moment. First, let us ask about what the three-dimensional shape and orientation of object X are. Well, there are infinitely many possible three-dimensional shapes and orientations for X that would allow it to project toward you as a square. For example, it could simply be a square in your fronto-parallel plane; or it could, instead, be a trapezoid with its longer edge on top and tilted away from you until it projects as a perfect square. And so on. So long as the three-dimensional shape and orientation you perceive is consistent with its projecting toward you as a square, then you are having a consistent perception. Now, however, let us ask about what the *projected* shape of object X is. Despite the fact that X may be infinitely many different things, all those things would *still* project as a square, so the projected shape of X is, in fact, unambiguously a square. To perceive the projected shape in a manner consistent with the stimulus, you must perceive X to project toward you as a square. The problem is that we *don't* perceive X to project toward us as a square, despite the fact that it does project toward us as a square. Instead, we perceive X to project toward us as a trapezoid. This is inconsistent perception.

In fact, all the classical geometrical illusions are cases of inconsistent perception: in each case, observers perceive a projected angle, an angular size, or an angular separation to be inconsistent with the information in the proximal stimulus. Note that all these are cases of perception of projective properties, and projective properties are more likely to change quickly in time, and thus ripe for latency correction. In Figure 2.3, observers misperceive all these three

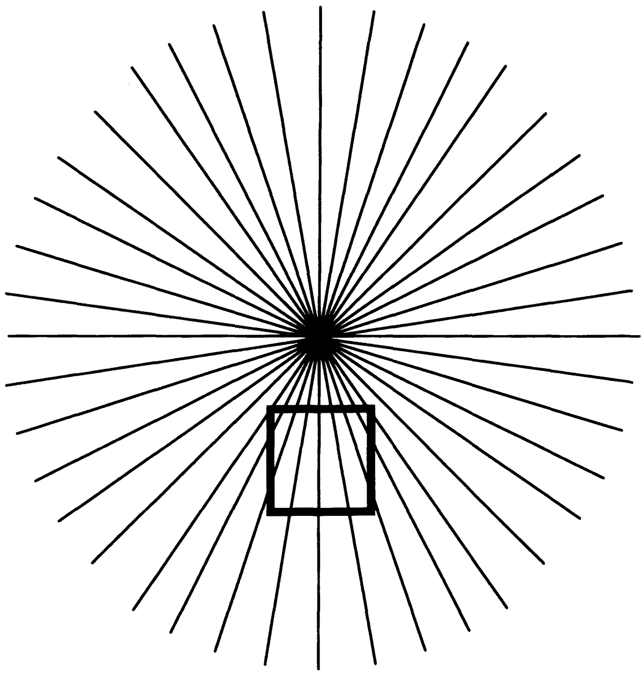

Figure 2.3: One perfect square on a radial display induces an illusion. There are many possible three-dimensional orientations and shapes for the square-like object that are consistent with the stimulus; i.e., where the object would project toward the observer as a perfect square. Perceiving any one of these possible objective orientations and shapes would be a case of consistent perception. However, since the square-like object in the figure actually projects as a perfect square, it is not *consistent with the stimulus that it projects in any other way. Nevertheless, we perceive it to project not as a perfect square, but as a trapezoid. This, then, is a case of inconsistent perception.*

projective properties. (1) It has misperception of projected angle because observers perceive the top two angles to project differently (namely, smaller) than the lower two angles, when they in fact all project identically. (2) It has misperception of angular size because observers perceive the top side of the box to have longer angular size than the bottom, but they have the same angular size. (3) And it has misperception of angular separation because observers perceive the higher parts of the sides to have greater angular separation than the lower parts of the sides, but the angular separations are in fact identical. [(2) and (3) are essentially the same kind of misperception, but I have distinguished them here because in some classical illusions it is more natural to think in terms of one over the other.]

It was only after understanding that certain kinds of illusions are cases of inconsistent perception that I both realized the inadequacy of the standard Bayesian approach to perception, and was propelled toward a nonstandard Bayesian approach to perception: latency correction.

The principal feature making the standard Bayesian approach "standard" is that, as mentioned, it presumes that the visual system is trying to choose among hypotheses concerning the scene out there at the time the retinal stimulus occurred. What if, however, the visual system is not trying to use the evidence to figure out what *was* out there when the stimulus occurred, but, instead, is trying to use the evidence to determine what is *going* to be out there by the time the percept actually occurs? That is, what if the visual system is implementing latency correction? For latency correction, the perceptual hypotheses the visual system is picking from are not hypotheses about what was out there when the retinal stimulus occurred, but hypotheses about what will be out there when the perceptual computations are completed.

With this alternative Bayesian approach for perception, it becomes possible for inconsistent perception to occur. Why? Because now it is quite possible that the scene probably out there at the time t the percept is generated is different than any possible scene that could have caused the retinal stimulus (which occurred at $t - 100$msec). That is, it is entirely possible that the probable scene out there at time t is causing a new retinal stimulus that is different from the one at time $t - 100$msec. For example, in Figure 2.3, imagine that the object X at the bottom actually is a perfect square in your fronto-parallel plane, but a little below the horizon. Furthermore, suppose you are moving toward the center point of the radial display. How would the projection of the square change as you move forward? Well, the top would project larger—i.e., have greater angular size—than the bottom because you are closer to the top than the bottom.

That is, object X would project trapezoidally in the next moment. If, upon being presented with Figure 2.3 as a retinal stimulus, your visual system infers that you are moving toward the center point, then the elicited percept will, if a latency correction strategy is being employed, be of object X projecting trapezoidally. This is, in fact, the central idea behind my latency correction theory of the classical geometrical illusions, which we take up in detail in the next section.

The latency correction (nonstandard Bayesian) approach to perception does not predict *only* inconsistent perceptions; consistent perceptions are still possible. Consistent perceptions *better* be possible, since many of our perceptions *are* (or at least appear to be) consistent. In what circumstances would latency correction lead to consistent perception? That's easy: any time the probable scene properties causing the stimulus are probably unchanging. What kinds of properties do not typically change much in the short term? Although *projective properties*—how objects project toward the observer, either geometrically or spectrally—change very quickly through time since they depend on the observer's position relative to the objects, *objective properties*—the properties of objects independent of their relationship to other things—do not typically change much through time. For example, the angular size of a flag pole is a projective property, as it depends on how far you are from it. Accordingly, it often changes in the next moment as you move, projecting either larger or smaller. The linear size of a flag pole, however, is an objective property, as it is, say, 10 meters high independent of where you stand with respect to it. In the next moment the linear size is very unlikely to change. Accordingly, we expect that perception of the flag pole's linear size will be a consistent perception because latency correction will generate a percept of a 10 meter pole, which is still consistent with the retinal stimulus. In fact, the cases where the standard Bayesian approach has mainly excelled are in applications to the perception of objective properties, like surface color and object recognition.

To sum up some of our discussion, a latency correction approach to vision can explain the existence of inconsistent perceptions; the standard Bayesian approach cannot. This latency correction approach is a nonstandard Bayesian approach, which means (i) it is a Bayesian approach, acquiring all of its powers and benefits, but (ii) it has a slightly different view concerning the kind of perceptual hypotheses the visual system is looking for... namely, it is looking for perceptual hypotheses about the way the world *is*, not about the way the world *was*. Furthermore, in the special case of perception of objective properties this alternative, latency correction, Bayesian approach collapses to the standard Bayesian approach, thereby squaring with the Bayesian approach's many successes.

2.2 A simple latency correction model

In this section I describe my model for how latency correction leads to the classical geometrical illusions (Changizi, 2001; Changizi and Widders, 2002). The following section applies the model to the illusions.

Recall that the latency correction hypothesis is, in my statement of it, as follows:

> On the basis of the retinal information the visual system generates a percept representative of the scene that will probably be present at the time of the percept.

The 'probably' that appears in the statement means that the statement is a probabilistic hypothesis, a Bayesian one in particular (but not a standard Bayesian one where the percept would represent the scene probably causing the proximal stimulus). And as mentioned earlier, we may conceptually distinguish two problems the visual system will have to solve.

1. First, the visual system must figure out what scene probably caused the proximal stimulus.

2. And, second, the visual system must figure out how that scene will change by the time the percept is elicited.

Again, this does not mean that the visual system's algorithm or lower-level mechanisms must distinguish these things, only that whatever function the visual system is computing, it would, in effect, have to solve both of these problems. This conceptual distinction is helpful for us scientists who wish to make predictions from the latency correction hypothesis: to predict what percept a visual system will generate given some proximal stimulus, we can subdivide our task into two smaller tasks. Namely, we must, for the geometrical illusions, try to determine what the probable scene is that would cause the geometrical figure, and then try to determine how the observer will typically move in the next moment (i.e., by the time the percept occurs).

We first need a way of deciding what the probable scene is for simple geometrical figures like those in the classical geometrical stimuli. That is, we need a way of figuring out what a figure probably depicts. Before I describe a

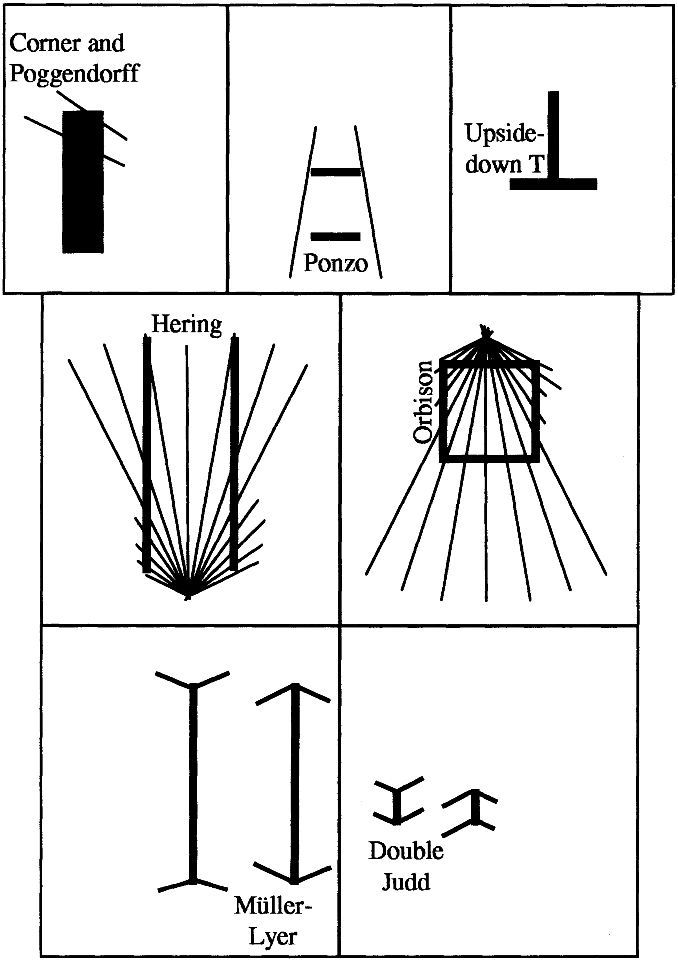

Figure 2.4: Eight classical geometrical illusions. Corner Poggendorff: the line through the corner of the rectangle appears to be bent. Poggendorff: the line through the rectangle appears to be two, parallel, non-collinear lines. Hering (also a variant of the Zöllner stimulus): the two parallel lines appear to be farther apart as one looks lower. Upside-down 'T': the horizontal bar appears to be shorter than the same-length vertical bar resting on top of it. Orbison: the right angles near the top appear to be acute, and the right angles at the bottom appear to be obtuse. Ponzo: the higher horizontal line appears to be longer than the same-length lower one. Double Judd: the vertical shaft of the left figure appears higher than the same-height one on the right. Müller-Lyer: the vertical shaft on the left appears longer than the same-length one on the right. See Coren and Girgus (1978) for references; see Greene (1988) for the corner Poggendorff.

model for helping us do this, let us see some of the classical figures, as shown in Figure 2.4.

The first feature to notice is that these most famous classical geometrical illusions consist entirely of straight lines. The second thing to notice is that, in addition to many oblique lines, there are also many horizontal and vertical lines, many more than we would expect if lines were thrown onto the page with random orientations. Finally, we can see that for many of the illusions there is a subset of the obliques that seem to all point toward the same point. All these features suggest that there may be simple rules for determining what the figures depict. That is, that there may be simple rules for determining what kind of real world line is the source of any given projected line.

Three principal kinds of line

The question I ask now is, For each kind of projected line in a figure, what kind of real world line or contour probably projected it? To help us answer this, let us look at a geometrical figure, namely Figure 2.5, where there are so many cues that it is obvious what the source lines of the projected lines are. The figure clearly depicts a room or hallway. It is my hypothesis that the projected lines in the geometrical stimuli are typically caused by lines and contours in "carpentered" environments like rooms and hallways. Furthermore, I hypothesize that observers typically move down hallways and rooms; they do not tend to zigzag wildly, nor do they tend to move vertically. The focus of expansion—the point of the forward-moving observer's visual field from which objects are expanding radially outward—is thus the vanishing point.

There are three kinds of line in the scene depicted in Figure 2.5: x lines, y lines and z lines.

- x lines are the lines that lie parallel with the ground, and perpendicular to the observer's direction of motion.
- y lines are the lines that lie perpendicular with the ground, and are also perpendicular to the observer's direction of motion.
- z lines are the lines that lie parallel with the ground, and are parallel to the observer's direction of motion.

Note that these kinds of line are defined in terms of the observer's probable direction of motion, which, again, is toward the vanishing point. In my simple model, I will assume that these are the *only* kinds of line in the world; I call them the *principal lines*. All we really need to assume, however, is that these

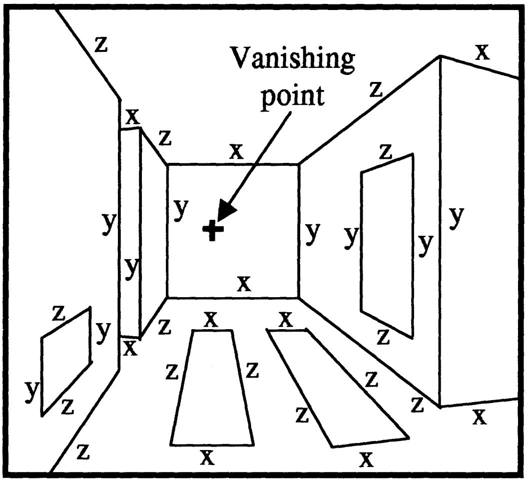

Figure 2.5: A sample geometrical figure showing the probable kind of source line for each line segment in the stimulus. The assumed observer direction of motion in such a stimulus is toward the vanishing point. The classical geometrical figures will be interpreted in this fashion.

three kinds of line are sufficiently more frequent in our experiences than other kinds of line that, in simple geometrical stimuli, one of these kinds of line is always the probable source.

Given that there are just three kinds of line in the world, we can ask of each of them, How do they typically project toward the observer? Once we have learned how each kind of line typically projects, we can work backward and ask, Given the projection, which kind of line probably caused it?

How do principal lines project?

x lines typically project horizontally in figures, as one can see in Figure 2.5. In particular, they project from the left straight to the right when they are near the *vertical meridian*, which is the vertical line drawn through the vanishing point in the figure. When an x line is off to either the left or right side, however, x lines begin to project more and more obliquely, as can again be seen in Figure 2.5. In fact, at the right side, the projections of x lines begin to point toward a vanishing point way off to the observer's right side; and, similarly, on the left side x line projections begin to converge toward a vanishing point on that side. We can understand how x lines project more clearly by considering a *projection sphere*. A projection sphere allows us to visualize the way things in the world project toward an observer. Projections are, by definition, devoid of depth information; they only possess information about the direction from which the stimulus was received. The set of all possible such directions from the outside world toward the observer's eye can be encapsulated as a sphere with the observer's eye at its center; each point on the sphere stands for a different projection direction from the outside world. Figure 2.6 shows how the three kinds of line may project toward the observer within our simple model. The "x line" sphere in Figure 2.6 shows how x lines project. Every x line segment lies along some great circle extending from the left pole to the right pole. The contour on the sphere that goes through the focus of expansion (the cross) is the way the horizon, for example, projects toward the observer, and is called the *horizontal meridian.* If the observer is about to cross over railroad tracks, the tracks project like the contours on the lower half of the sphere; as the observer nears the tracks, they project progressively lower and lower, eventually projecting along the very bottom of the projection sphere. As long as x line segments are near the *vertical meridian* (which is the contour extending from directly overhead, through the focus of expansion, and down to directly below the observer), they project horizontally onto the projection sphere, and parallel to one another. However, in the left and right peripheral parts of the sphere, x lines begin to converge toward the left and right poles; they no longer project horizontally, and they are no longer parallel to one another. We must be careful, however, because x lines *do* project horizontally in the periphery *if* they happen to lie along the horizontal meridian.

How do y lines project? From Figure 2.5 one may see that y lines typically project straight from the bottom to the top (i.e., non-obliquely), and that they are parallel to one another. Although it is not all that common in our experience, if a y line segment is very high above or very low below an observer, they begin to project obliquely, are no longer parallel to one another, and begin to converge toward the top or bottom pole, respectively. We can make this more precise by looking at the y line projection sphere in Figure 2.6. Every y line segment lies along some great circle extending from the top or North pole to the bottom or South pole. Suppose you are floating in front of a pole that goes infinitely far above you and infinitely far below you. If you are moving directly toward it, then it projects as the contour on the sphere that goes through the focus of expansion. Suppose now that you are going to pass the pole on your right. As you near it, the pole will project progressively more and more along the right side of the projection sphere (which is on the left in the figure). As long as the y segments are relatively near the horizontal meridian, they project nearly purely up and down, and are parallel to one another, as they all are in Figure 2.5. When the y line segments are in the upper or lower periphery, however, they begin to converge toward a pole of the sphere, and are no longer parallel to one another. y lines in the periphery can still project non-obliquely if they happen to lie along the vertical meridian.

Finally, how do z lines project? From Figure 2.5 we can see that z lines project obliquely, and that they share a vanishing point, namely at the focus of expansion. The z line projection sphere in Figure 2.6 encapsulates how z lines project. Each z line segment lies along a great circle from the focus of expansion all the way to the focus of contraction (which is directly behind the observer). For example, if you are walking on a sidewalk, the sides of the sidewalk project on the lower left and lower right of the projection sphere. z lines typically project obliquely, but a z line can project horizontally if it happens to lie along the horizontal meridian, and it can project vertically if it happens to lie along the vertical meridian.

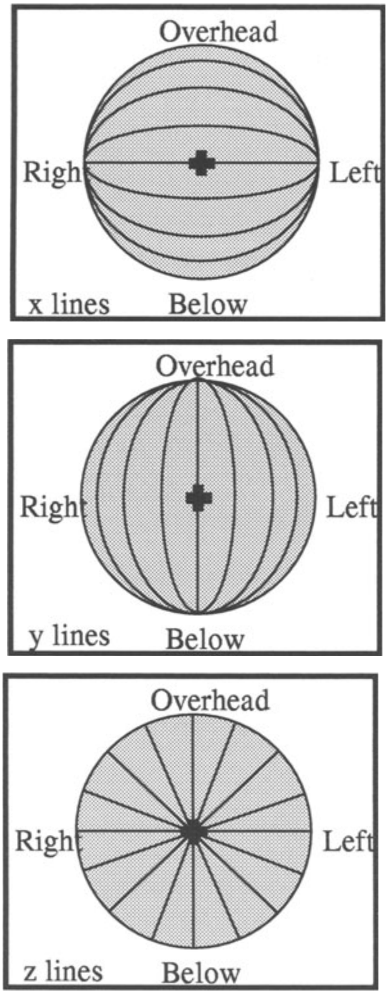

Figure 2.6: Three projection spheres showing, respectively, how x lines, y lines and z lines project toward an observer. The focus of expansion is shown as the cross. Note that each of these figures depicts a convex sphere (even the "z line" one), and the contours are on the near side.

An aside on the non-Euclidean visual field

Before using these insights on how principal lines project to determine the probable source of a projected line, there is a neat observation we may make from the discussion thus far. Suppose you are about to cross railroad tracks. The projection of each of the two tracks is a straight line in your visual field (each follows a great circle on the projection sphere). Furthermore, these two projected lines are parallel to one another when at the vertical meridian of your visual field. However, the two projected lines become non-parallel to one another in the periphery of your visual field, and eventually even intersect. How is it possible that two straight lines which are parallel at the vertical meridian can intersect one another? Can this really be?

It *can* really be, and it is possible because of the non-Euclidean nature of the geometry of the visual field. The geometry that is appropriate for the visual field is the surface of a projection sphere, and the surface of a sphere is not Euclidean, but, well, spherical. There are three main kinds of geometry for space: elliptical, Euclidean (or flat), and hyperbolic. Spherical geometries are a special case of the elliptical geometries. In Euclidean geometry, the sum of the angles in a four-sided figure (a quadrilateral) is $360°$; in elliptical it is more, and in hyperbolic it is less. Let us ask, then, what the sum of the angles in a four-sided figure in the visual field is. A four-sided such figure is built out of four segments of great circles. Figure 2.7 shows an example four-sided figure on a projection sphere. In particular, it is a square. It is a square because (i) it has four sides, (ii) each side is a straight line (being part of a great circle), (iii) the lines are of (roughly) the same length, and (iv) the angles are (roughly) the same. Notice that each angle of this square is bigger than $90°$, and thus the square has a sum of angles greater than $360°$. The visual field is therefore elliptical.

One does not need to examine projection spheres to grasp this. If you are inside a rectangular room at this moment, look up at the ceiling. The ceiling projects toward you as a four-sided figure. Namely, you perceive its four edges to project as straight lines. Now, ask yourself what each of its projected angles is. Each of its angles projects toward you greater than $90°$; a corner would only project as exactly $90°$ if you stood directly under it. Thus, you are perceiving a figure with four straight sides, and where the sum of the angles is greater than $360°$. The perception I am referring to is your perception of the projection, not your perception of the objective properties. That is, you will *also* perceive the ceiling to objectively be a rectangle, each angle having $90°$. Your perception of

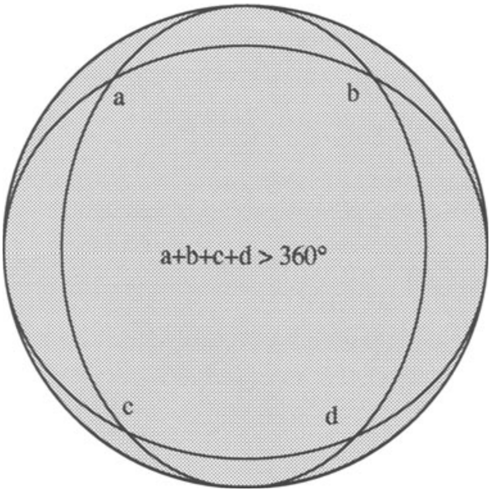

Figure 2.7: Four great circles on a sphere (or on a visual field). In this case they make a square on the sphere; that is, the figure is four-sided, each side is straight and of equal length, and the angles are equal. Each angle of the square is greater than 90°*, however. [To see this, the reader must judge the angle on the sphere, not the angle on this page.] Thus, the sum of the angles of the square is greater than* 360°*, which means the space must have elliptical geometry. In particular, it is spherical.*

the objective properties of the ceiling is Euclidean, or at least approximately so. Your perception of the way the ceiling projects, however, conforms to elliptical geometry. [There is a literature which attempts to discover the geometry of our perception of three dimensional space, and it is argued to be hyperbolic. This is an entirely different issue than the one we are discussing, as we are focusing just on the perception of projective properties (without depth information).]

It is often said that non-Euclidean geometry, the kind needed to understand general relativity, is beyond our everyday experience, since we think of the world in a Euclidean manner. While we may think in a Euclidean manner for our perception of the objective lines and angles, our perception of projective properties is manifestly non-Euclidean, namely spherical. We therefore *do* have tremendous experience with non-Euclidean geometry, it is just that we have not consciously noticed it. But once one consciously notices it, it is possible to pay more attention to it, and one then sees examples of non-Euclidean geometry at one's every glance.

Given a projection, which principal line is source?

We have seen earlier the way that the three principal kinds of line—x, y and z lines—project toward an observer. Now we wish to utilize this knowledge to ask the "inverse" question: Given some projected line in a proximal stimulus, which of the three kinds of line is the probable source?

Observers typically are looking forward as they move, and it is therefore reasonable to assume that, unless there are cues to the contrary, a projected line is probably not in the extreme peripheral regions of the visual field. This fact is useful because examination of the regions relatively near the focus of expansion (i.e., not in the periphery) of the projection spheres in Figure 2.6 reveals some simple regularities. The only kind of line projecting obliquely in this regime of the projection sphere is the z line, and all z lines converge to the same vanishing point (which is also the focus of expansion since observers are assumed to move parallel to the z axis). As a consequence of this, we may state the following rule.

> Rule 1: *If there is a single set of oblique projected lines sharing a vanishing point, then they are probably due to z lines.*

The only kind of line projecting horizontally in this regime, and *not* lying on the horizontal meridian, is the x line. Therefore...

> Rule 2: *A horizontal projected line that does not lie on the horizontal meridian is probably due to an x line.*

Both x and z lines can cause horizontal projected lines lying on the horizontal meridian, and thus the following rule applies.

> Rule 3: *A horizontal projected line that* does *lie on the horizontal meridian may be due to either an x line* or *a z line.*

The only kind of line projecting vertically in the relatively-near-the-focus-of-expansion regime is the y line, and the following rule therefore applies.

> Rule 4: *A vertical projected line that does not lie on the vertical meridian is probably due to a y line.*

Analogously with Rule 3, we also have...

> Rule 5: *A vertical projected line that* does *lie on the vertical meridian may be due to either a y line* or *a z line.*

One kind of proximal stimulus we will want to decipher is one where there are two sets of converging projected lines with distinct vanishing points. Because observers tend to look where they are going, one of these sets probably consists of projected z lines, and it will probably be the set with lines for which it is most salient that they converge to a vanishing point. The other set of converging lines consists of either projected y lines (which would point toward a vanishing point above or below the focus of expansion) or projected x lines (which would point toward a vanishing point to the left or the right of the focus of expansion). This is recorded as the following rule.

> Rule 6: *When there are two sets of projected lines with different vanishing points, the set with the more salient vanishing point probably consists of projections of z lines, and the other of either x or y lines, depending on where they point.*

These rules are all consequences of the simple model of three kinds of principal lines and forward motion parallel to the z axis.

Recall that, for us scientists to make predictions using the latency correction hypothesis, we must determine the probable scene causing the proximal stimulus, and we must determine how that scene will probably change in the next moment. Although we now have machinery enabling us to infer the probable scene, we have not yet addressed the latter. How a scene will change in the next moment depends on where the observer is moving toward, and how fast. Where the observer is moving can be determined by the vanishing point of the z lines; since the observer moves along the z axis, the vanishing point of the projected z lines is also the focus of expansion, or the direction of motion. Therefore, once we have discovered which projected lines in the proximal stimulus are due to z lines, we have also discovered the direction of motion. Because of the importance of this consequence, I record it as a final rule.

> Rule 7: *The probable location of the focus of expansion is the vanishing point of the projected z lines.*

The observer's speed can be set to some reasonable value; I typically set it to 1 m/sec. I also often assume in simulations a latency of 50 msec, which is an underestimate.

One important aspect of the probable scenes that this simple model does not accommodate is distance from the observer. If all the probable sources were as in the model, but were probably a mile away, then we can expect no change in the nature of the projections in the next moment. It is reasonable to assume that latency correction will be primarily tuned to nearby objects, objects that we can actually reach, or that we might actually run into. Accordingly, it is plausible that the visual system interprets these geometrical stimuli as scenes having a distance that is on the order of magnitude of meters away (rather than millimeters or hundreds of meters).

How general is this model?

The model I have proposed above requires that the ecological environment of the observer have an abundance of x, y and z lines, where the observer moves parallel to the z lines. I have called this a "carpentered world assumption" (Changizi, 2001c), but how much does my explanation depend on this assumption?

First consider z lines. One of the principal roles they will play in the explanation of the geometrical illusions is that they provide the cue as to the location in the visual field of the focus of expansion. That is, the visual system figures out where the observer is probably going on the basis of where the vanishing point is for the z lines. However, there need not actually be any z lines in the world for there to be, for the moving observer, projections which are like the projections of z lines. If an observer is moving forward in an unstructured environment, the optic flow itself will cause "optical blur" projected lines, and these will converge to the focus of expansion (Gibson, 1986). Thus, my explanation does not require that the ecological environment actually possess a propensity for z lines. The projected z lines may be due not to z lines at all, but to optic flow; radial lines mimic optic flow, and may trick the visual system into believing it is probably moving forward.

For x and y lines, all that my model really requires is that the probable source of a horizontal projected line (not on the horizontal meridian) is an x line, and that the probable source of a vertical projected line (not on the vertical meridian) is a y line. It could, for example, be the case that x lines and y lines are relatively infrequent, but that they are still the most probable source of horizontal and vertical projected lines, respectively.

It is also worth noting that a propensity for y lines does not require a carpentered world assumption. Forests, for example, have a propensity for y lines; gravitation makes the y axis unique, and any gravitational ecology will probably have a propensity of y lines. x lines, on the other hand, really do seem to require a carpentered world assumption; e.g., although the forest will have a propensity for there to be lines parallel to the ground, which is half the definition of an x line, it will not have a propensity to lie perpendicular to the observer's direction of motion. The model therefore *does* depend, in this regard, on the carpentered world assumption. Humans raised in non-carpentered environments would, then, be expected to have a different repertoire of geometrical illusions, which appears to be the case (Segall et al., 1966).

2.3 Explaining the geometrical illusions

In this section I explain how the latency correction hypothesis explains the classical geometrical illusions. The first subsection answers the question, What is the probable scene underlying the geometrical stimuli? This includes determining what the lines are and where they are with respect to the observer's direction of motion. The next subsection tackles the geometrical illusions that are misperceptions of projected angle, which includes the corner, Poggendorff, Hering and Orbison. The following subsection explains the illusions of angular size or angular distance, which includes the double Judd, Müller-Lyer, Hering, Orbison and the upside-down 'T'. The final subsection tests a psychophysical prediction of the latency correction hypothesis, providing further confirmation.

2.3.1 The probable source and focus of expansion

The rules developed in the previous section can now be applied to the illusions from Figure 2.4, both in determining what are the probable sources of the stimuli, and in determining what is the probable direction of motion for the observer. This is our task in this subsection.

The probable sources

Figure 2.8 shows the same illusions as in Figure 2.4, but each projected line has been labeled with the probable kind of source line via the earlier rules. The explanations for the probable sources are as follows.

- No vertical line in any of the illusory figures has cues suggesting it lies along the vertical meridian, and thus each is probably due to a y line.
- Of all the horizontal lines, only the one in the upside-down 'T' illusion possesses a cue that it might lie along the horizontal meridian. The cue is that there is a 'T' junction, and such junctions are typically due to three-dimensional corners (i.e., an x-y-z corner). The horizontal segment of the 'T' junction is probably, then, due to two distinct segments, one the projection of an x line, and one the projection of a z line. That is, it is probably a corner that is being viewed "from the side." I have arbitrarily chosen the left segment to be the projection of an x line, but the cues in the upside-down 'T' illusion (which consists of just the upside-down 'T') do not distinguish which is which.
- All the remaining horizontal projected lines are parts of stimuli without any cues that they lie along the horizontal meridian, and so are thus due to x lines.
- All that is left are the obliques. In the Hering, Orbison, Ponzo, Corner and Poggendorff there exists just one set of converging obliques, and they are thus probably due to z lines.
- In each of the Müller-Lyer and the Double Judd there are two sets of converging projected lines: one set consists of the four inner obliques (the ones *in between* the two vertical lines), and the other set consists of the four outer obliques (the ones not in between the two vertical lines). The four inner obliques are more salient and clustered, and appear to share a vanishing point more clearly than do the outer ones. The inner obliques are therefore probably due to z lines. Since the outer obliques have a vanishing point horizontally displaced from the vanishing point for the inner obliques, the outer obliques must be due to x lines. [While this serves as an adequate first approximation, greater analysis in fact reveals that the outer obliques probably do not share a vanishing point at all (and thus they cannot all be principal lines). Consider just the Müller-Lyer Figure for specificity. Lines in the world project more obliquely as they near their vanishing point (see Figure 2.6). The two outer obliques on the left are far in the visual field from the two outer obliques on the right; if they were projections of the same kind of line in the world, then they would *not* project parallel to one another, one pair being considerably closer to the vanishing point (for that kind of line) than the other. But the outer obliques on the left *are* parallel to the outer ones on the right, and thus they cannot be projections

of the same kind of line, and they do not point to a single vanishing point. Only the four inner obliques are approximately consistent with a single vanishing point.]

The probable focus of expansion

Now that we know what the probable sources are for the eight illusory proximal stimuli, we can use the information about the projected z lines to determine the focus of expansion. That is, the z line vanishing point is the focus of expansion (see Figure 2.8). Figure 2.9 shows the earlier figures, but where the illusions now share the same focus of expansion, and Figure 2.10 shows the key features of each illusion embedded in a display which provides a strong cue as to the focus of expansion.

- For the Hering, Ponzo, Orbison and Müller-Lyer stimuli there is exactly one focus of expansion determined by the projections of the z lines, and Figure 2.9 shows them embedded in a radial display at the appropriate location with respect to the focus of expansion. Notice that for the Müller-Lyer the fins act as cues as to the location of the focus of expansion, and that in Figure 2.10, where the radial display does the cueing work, the fins are no longer necessary for the illusion.
- The projected z lines for the double Judd are so similar in orientation that they may converge either up and to the right of the figure, or down and to the left; that is, the focus of expansion may be in one of these two spots. I have placed the fin-less version of the double Judd in Figure 2.10 into these two positions with respect to the focus of expansion. Note that the illusion is qualitatively identical in each case to the earlier one (since the cues to the focus of expansion are provided by the radial display rather than the fins).
- The corner and Poggendorff illusions could be placed anywhere in the display so long as the projected z line points to the probable focus of expansion; I have chosen one spot arbitrarily. Any conclusions I draw later will not depend on this choice.
- The upside-down 'T' illusion could be placed on either side of the vertical meridian, so long as the horizontal segments lie along the horizontal meridian. I have arbitrarily chosen one spot. Any conclusions I draw later will not depend on this choice.

At this point in this section I have used the model to determine the probable source and focus of expansion given a sufficiently simple geometrical proximal stimulus. The model has concluded that each of the eight classical geometrical illusions I have been discussing are probably caused by a particular kind of source in the world, and are probably located in a certain position in the visual field with respect to the focus of expansion. These conclusions were summarized in Figure 2.8.

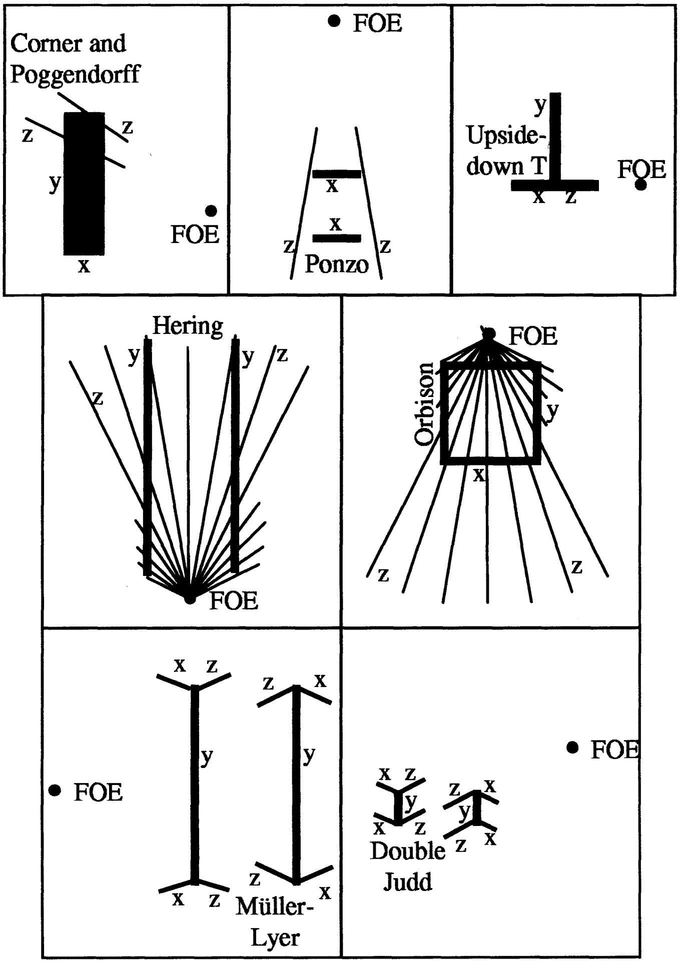

Figure 2.8: Eight classical geometrical illusions, now showing for each projected line the probable kind of source line, x, y or z. The probable focus of expansion is also shown in each case.

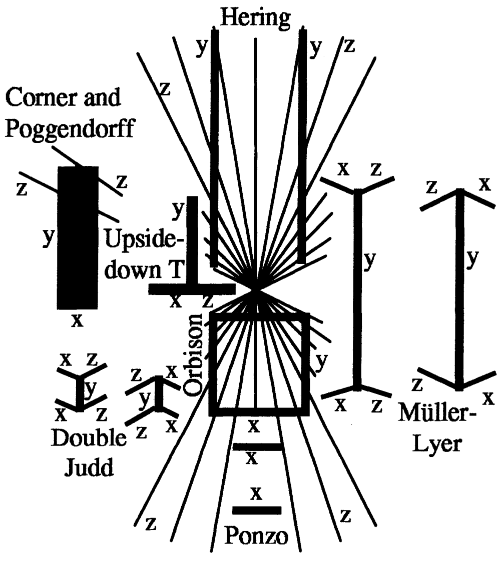

Figure 2.9: The same eight classical geometrical illusions, showing, as in an earlier figure, for each projected line the probable kind of source line, x, y or z. They have been placed such that their focus of expansion is the same.

We still have not explained the illusions, however. Recall that, under the latency correction hypothesis, in addition to determining the probable scene causing the proximal stimulus—which is what we have done thus far—we must also figure out how that scene will probably change by the time the percept occurs. Well, since we know the probable scene, and we know which direction the observer is probably moving, all we have to do is to determine how the sources will project when the observer is moved forward a small amount.

2.3.2 Projected angle misperception

One kind of illusion concerns misperception of projected angle. First, let me be clear about what I mean by perception of projected angle. If you look up at a corner of the room you are in, you will notice that you perceive there to be three right angles; this perception is the perception of the objective angles. You simultaneously perceive there to be three obtuse angles summing to $360°$; this perception is the perception of the projected angles. It is the latter that is relevant for the geometrical illusions.

The corner, Poggendorff, Hering and Orbison can be treated as misperceptions of projected angle. In the corner and the Poggendorff the angles appear to be nearer to 90° than they actually are. The same is true for the angle between the vertical line and the obliques in the Hering illusion. In the Orbison illusion, the right angles appear to be bent away from $90°$. How do we make sense of these projected angle illusions? And why are some misperceived towards 90° and some away from it?

First, let us distinguish between two kinds of projected angle. Since there are just three kinds of line in my model, the only kinds of angle are those that result from all the possible ways of intersecting these kinds of line. They are x-y, x-z and y-z angles; these are the *principal angles*. That is, x-y angles are any angles built from an x line and a y line, and so on. The latter two kinds of angle are actually similar in that, having a z arm, the plane of these angles lies parallel to the observer's direction of motion. I call x-z and y-z angles xy-z angles. x-y angles, on the other hand, lie in a plane perpendicular to the observer's direction of motion, and must be treated differently.

xy-z projected angles

Note that the corner, Poggendorff and Hering illusions have angle misperceptions where the angles are xy-z angles, and the misperception is that observers perceive the projected angles to be nearer to 90° than they actually are. Why is

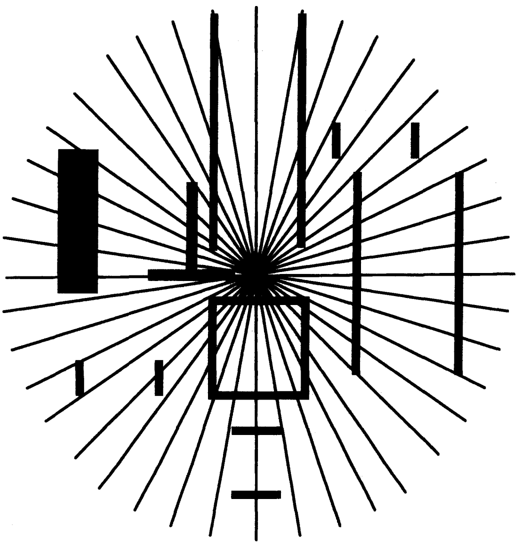

Figure 2.10: Each illusion from Figure 2.9 is "transferred" into a stimulus with strong cues as to the location of the focus of expansion. The same kinds of illusion occur, suggesting that it is cues to the location of the focus of expansion that is of primary importance in the illusions. In the case of the double Judd and Müller-Lyer figures, the probable location of the focus of expansion is the same as in Figure 2.9, but now its location is due to the radial lines rather than the fins; because the double Judd is also consistent with being in the upper right quadrant, it has been transferred there as well as in the bottom left quadrant. The corner and Poggendorff stimuli could be placed anywhere in the radial display so long as radial lines traverse them in the appropriate fashion.

this? The latency correction hypothesis says it is because in the next moment the angles will project nearer to 90°, and thus the misperception is typically a more veridical percept. [It is inappropriate in the case of a static stimulus.] But *do* xy-z angles actually project more like 90° in the next moment? Yes, and there are a number of ways to understand why.

The most obvious way to see this is to hold something rectangular, maybe an index card, out in front of you, below, and to the right of your eyes. Hold the card out flat (i.e., parallel with the ground), and orient it so that the near edge is perpendicular to your line of sight. This is depicted in Figure 2.11 (A). Observe how the four right angles of the card project toward you. The projected angles nearest and farthest from the vertical meridian—i.e., angles a and d—are both acute, and the other two are obtuse. What happens to these projected angles if you move your head and eyes forward as if you are going to pass the card? If you move so far that the card is directly below and to the right of your eyes—i.e., you are just passing it—you will see that the four angles all project as 90°. Thus, as you move forward, each of these x-z projected angles changes toward 90°, eventually becoming exactly 90° when you pass it. The same observation applies any time an observer moves forward in the vicinity of x-z angles (e.g., a rug, or the ceiling). The same observations also apply for the projection of y-z angles, one case which is depicted in Figure 2.11 (B). A real world example is when you walk past a window: all the projected angles begin either very acute or very obtuse, but as you near the window, they progressively project more and more as 90°.

Another way of comprehending how xy-z angle projections change is to examine projection spheres upon which xy-z angles have been projected. Figure 2.11 (C) and (D) show such projection spheres; (C) consists of the intersections of x and z line projections, and (D) of y and z line projections. Recall the nature of optic flow on the projection sphere: flow begins at the focus of expansion and moves radially outward toward the periphery of the sphere. Thus, projected angles nearer to the focus of expansion are the way they project when they are farther away from the observer, and projected angles in the periphery of the projection sphere are the way they project when the observer is nearer to passing the angle. For example, in Figure 2.11 (C) the reader may see asterisks at four projected angles along the same radial line. The asterisk nearest the focus of expansion is roughly 60°, the next one a bit bigger, the next still bigger, and the last one is 90°. [Recall that to judge these angles the reader must judge the angle on the sphere, not on the page.] We could have, instead, put our asterisks on the other side of the projected z line, and we would have had

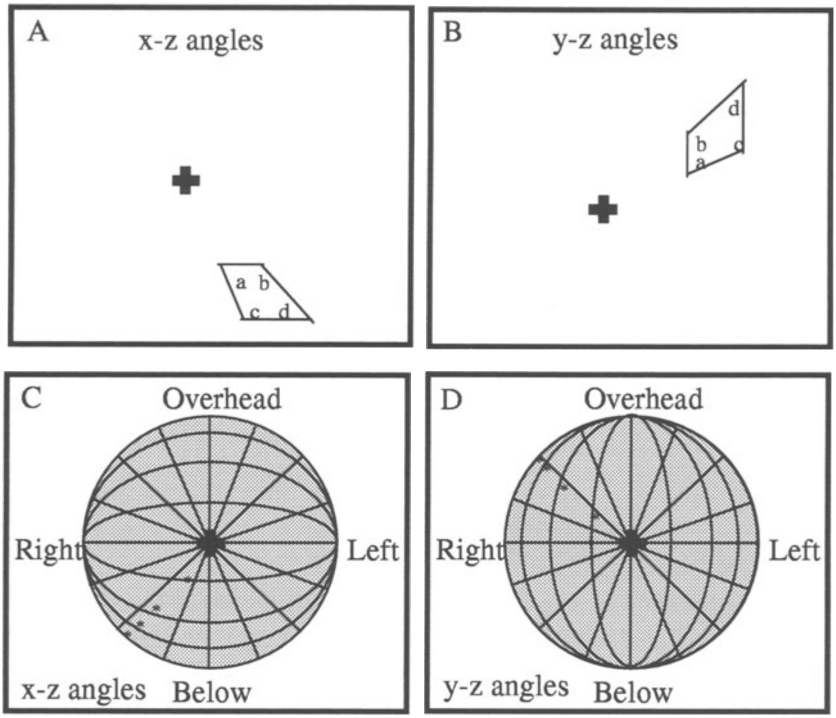

Figure 2.11: (A) Depiction of the view of a rectangular card below and to the right of an observer's direction of motion (represented by the cross). The card is lying parallel to the ground, with one axis parallel to the observer's direction of motion, and the other perpendicular to the direction of motion. The card's angles are thus x-z angles. (B) This depicts the analogous card as in (A), but now the angles are y-z. (C) A projection sphere upon which x and z lines are projected; their intersections are x-z angle projections. Notice how, along any radial line, the angles of intersection between x and z lines become more like $90°$ in the periphery (see the asterisks); that is how they change in the next moment, since the angles move toward the periphery as the observer moves forward. (D) A projection sphere upon which y and z lines are projected; their intersections are y-z angle projections. Notice how, along any radial line, the angles of intersection between y and z lines become more like $90°$ in the periphery (see the asterisks); this is how they change in the next moment.

the projected angles starting from around 120° and falling to 90°. A similar account applies to the projections of y-z angles as shown in Figure 2.11 (D).

In short, xy-z angles project more like 90° as observers move forward. If a proximal stimulus has cues suggesting that a projected angle is due to an xy-z angle, then latency correction predicts that observers will perceive the angle to be more like 90° than it actually is. That is, people should perceive the projected angle to be "regressed" toward 90° (Thouless, 1931a). The corner, Poggendorff, and Hering illusions each had projected xy-z angles, and each is perceived to be nearer to 90° than it actually is. These illusions are, therefore, consistent with latency correction.

The Poggendorff has another salient illusory feature in addition to the projected angles being perceived nearer to 90° than they are: the two oblique lines are collinear, but do not appear to be. Each oblique line appears to, intuitively, undershoot the other. Latency correction explains this illusory feature as follows. Suppose that a single z line lies above you and to your left along the wall (perhaps the intersection between the wall and the ceiling). Now also suppose that there is a black rectangle on your upper left, but lying in your fronto-parallel plane. That is, the rectangle is made of x and y lines. Suppose finally that the rectangle is lying in front of the z line. The projection of these objects will be roughly as shown by the Poggendorff illusion in Figure 2.9. I say "roughly" because the projection will not, in fact, be as in this figure. Consider first the projected angle the z line will make with the right side of the rectangle. Suppose it is 60°; that is, the (smaller) y-z angle on the right side of the rectangle is 60°. What will be the projected angle between that same z line and the other vertical side of the rectangle? The part of the z line on the other vertical side of the rectangle is farther away from the focus of expansion and more in your periphery. Thus, this more peripheral y-z angle will be nearer to 90°; let us say 63° for specificity. That is, when the *same* z line crosses through or behind a rectangle as constructed, the projected angles will *not* be the same on either side. Now, the two projected angles in the Poggendorff figure *are* the same on either side, and thus the projected lines on either side *cannot* be due to one and the same z line. Instead, the more peripheral y-z projected angle, being farther from 90° than it would were it to be the projected angle made with the z line from the other side, must actually be a line that is physically higher along the wall. The visual system therefore expects that, in the next moment (i.e., by the time the percept is generated), the oblique projected line on the left should appear a little higher in the visual field compared to the extension of the oblique line on the right (since differences in visual field position are accentuated as an observer moves forward).

x-y projected angles

The Orbison illusion primarily concerns the misperception of the four projected angles, each which is 90°, but which observers perceive to be greater or lower than 90°. The squares in the Orbison illusion are composed of x and y lines (Figure 2.8, and see also Figure 2.1), and we must ask how the projections of x-y angles change as observers move toward the focus of expansion (which is the vanishing point of the projected z lines in the Orbison figure).

The most straightforward way to understand how x-y angles change is to hold up a rectangular surface like an index card in your fronto-parallel plane, with one axis vertical and the other horizontal, and move your head forward. When the card is sufficiently far out in front of you, each of its four angles projects nearly as 90°. As you move your head forward as if to pass it, the angles begin to project more and more *away* from 90°. Some angles begin to project more acutely, and some more obtusely. When you are just about to pass the card, some of its angles will project closer and closer to 0°, and the others will project closer and closer to 180°. If the card is in your lower right quadrant, as is depicted in Figure 2.12 (A), two angle projections fall toward 0°—*b* and *c*—and two rise toward 180°—*a* and *d*. If, instead, it is directly below you, the top two angle projections fall to zero and the bottom two rise to 180°. If it is directly to your right, then the near two go to zero and the far two straighten out. If the card is directly in front of you—i.e., each of its four angles is in each of the four quadrants of your visual field—then each angle gets more and more obtuse as you move forward. For example, as you walk through a doorway, each of the corners of the door projects more and more obtuse, so that when you are just inside the doorway each corner projects as 180° (and the doorway now projects as a single line all the way around you).

We may also comprehend how x-y angles change in the next moment by examining a projection sphere on which x and y lines have been projected, as shown in Figure 2.12 (B). If the reader follows the asterisks from the focus of expansion outward, it is clear that these x-y angles begin projecting at approximately 90° and as the observer moves forward and the angle moves peripherally the projections become greater and greater. Following the '#'s shows the same thing, except that the projected angles get smaller and smaller as the observer moves forward.

In sum, x-y angles project further away from 90° in the next moment; they

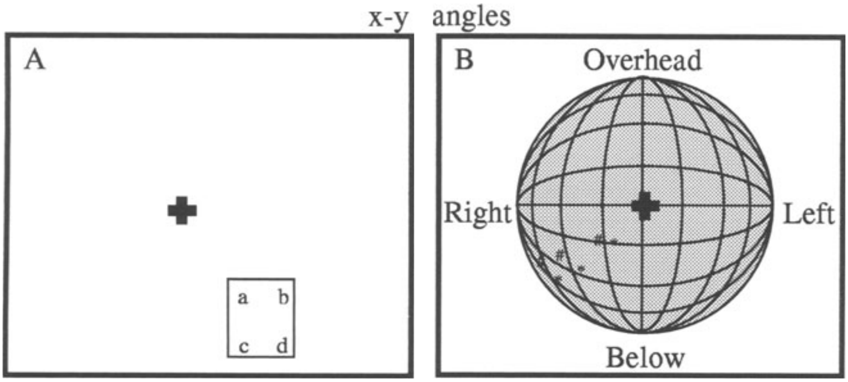

Figure 2.12: (A) Depiction of the view of a rectangular card below and to the right of an observer's direction of motion (represented by the cross). The card is lying upright and in the observer's fronto-parallel plane. The card's angles are thus x-y angles. (B) A projection sphere upon which x and y lines are projected; their intersections are x-y angle projections. Notice how, along any radial line, the angles of intersection between x and y lines become less like 90° in the periphery (see the asterisks and '#' signs); that is how they change in the next moment since the angles move toward the periphery as the observer moves forward.

are "repulsed" away from 90° instead of regressed toward as in xy-z projected angles. Which direction a projected x-y angle will get pushed away from 90° depends on the angle's orientation and its position relative to the focus of expansion. Figure 2.13 shows how one kind of x-y line changes its projection in the next moment, and one can see that it depends on the quadrant. The figure also shows the magnitude of the projected angle change as a function of position along the x and y axes on a plane one meter in front of the observer. For x-y angles of other orientations the plot looks similar, except that it may be rotated by 90°. Figure 2.14 (A) summarizes the directions which x-y projected angles change in the next moment as a function of position in the visual field with respect to the focus of expansion.

The latency correction hypothesis predicts that if cues suggest that a projected angle is due to an x-y angle, then observers will misperceive the angle to be whatever it will probably be in the next moment (by the time the percept is elicited). Figure 2.14 (B) shows the same squares as in (A), but now embedded in a radial display which provides strong cues as to the location of the focus of expansion. In every case, observers misperceive the right angles in (B) in the direction predicted by latency correction in (A). This is just a special case

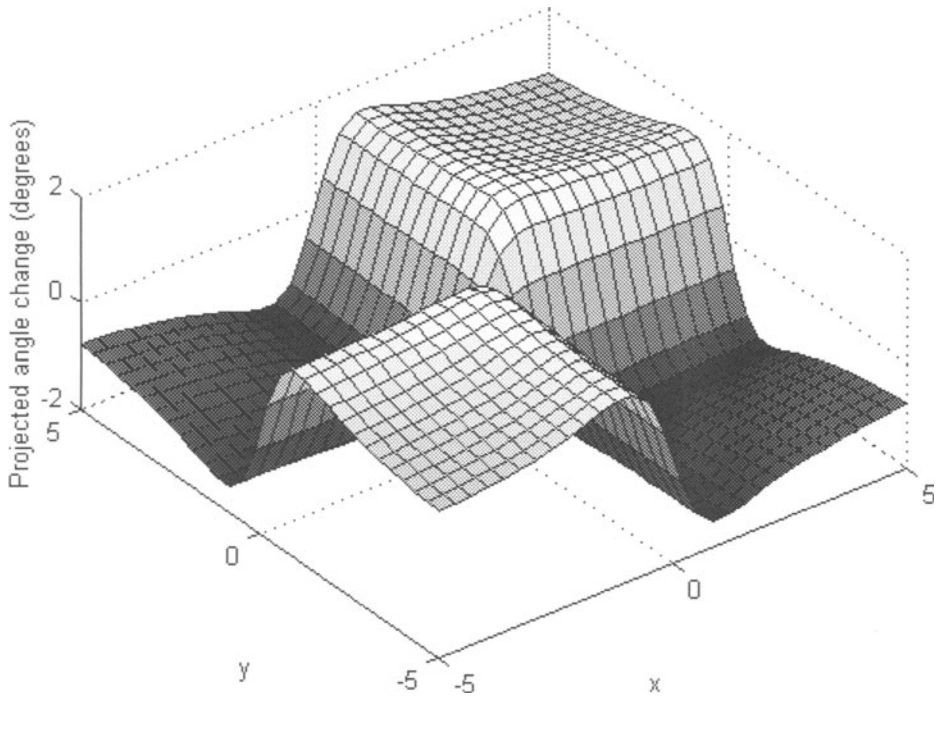

Figure 2.13: *The change in projected angle as a function of position with respect to the focus of expansion on a plane one meter ahead and perpendicular to the direction of motion, for an x-y angle with one arm pointing up and another arm pointing right. This assumes the observer moves forward 5 cm (or that the latency is 50 msec and that the observer is moving at 1 m/sec.)*

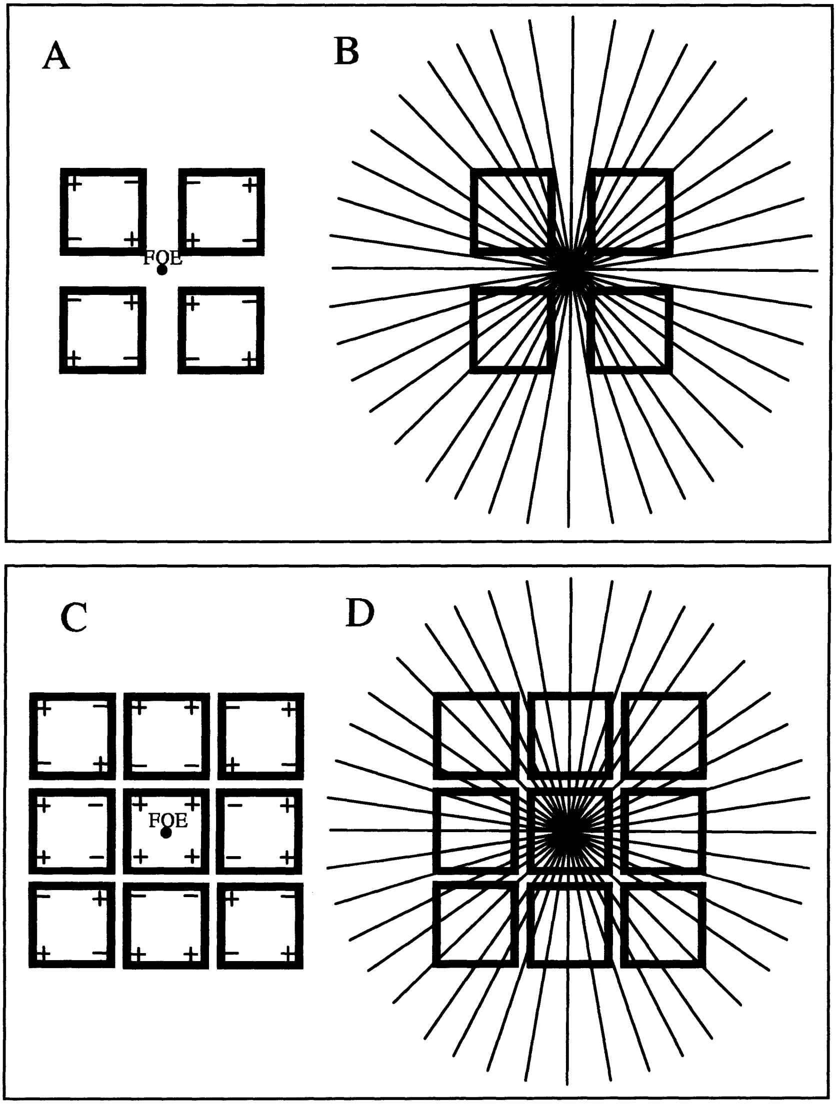

Figure 2.14: (A) Four perfect squares. If an observer were moving toward the cross at the center, the projected angles for each angle would change away from 90°*, and the pluses and minuses show the direction of change. (B) The same squares embedded in a stimulus with strong cues that the focus of expansion is at the center. Subjects perceive the angles to be different than* 90° *as predicted by the directions shown in (A). (C) and (D) are the same, but with more squares.*

of the Orbison illusion, and thus the Orbison illusion is consistent with latency correction.

Before we leave the Orbison illusion, I should note that there is another illusion also called by the same name (and also discovered by Orbison (1939)), and shown on the left in Figure 2.15. The illusion is primarily that the two projected angles nearer to the center of the circles appear to be a little obtuse, and the other two a little acute. Recall that when the square is on the right side of the other Orbison display with the radial lines (on the right in Figure 2.15) the two near-the-center angles are perceived to be a little acute, and the two farther ones a little obtuse. That is, the concentric-circle version of the Orbison leads to qualitatively the opposite illusion of the radial-line version. The square in the concentric-circle Orbison looks more like the square on the opposite side of the radial-line Orbison (see Figure 2.15), although it is less dramatic. My model has no apparatus at this point to accommodate concentric circles, but it is unclear what kind of ecologically valid scenario would lead to such a stimulus. My hunch at the moment is that concentric circles have no strong association with some probable source, and that, instead, the visual system is noticing that in the vicinity of the square there are many oblique lines, and they are all pointing to the right, and pointing toward the horizontal meridian. The oblique lines are not straight, and they do not point to a single vanishing point, but to the extent that the obliques all point to the right and toward the horizontal meridian, the visual system may guess that the focus of expansion is more probably somewhere on the right side of the square. This would explain why observers misperceive this square in a qualitatively similar manner as the square on the left in the radial-line Orbison.

Ambiguous projected angles

We have to this point showed how the latency correction hypothesis can explain the misperception of projected angles with cues as to whether they are due to xy-z angles or x-y angles. When cues suggest a projected angle is due to an xy-z angle, observers perceive the projected angle to be nearer to 90°, just as predicted by latency correction. And when cues suggest a projected angle is due to an x-y angle, observers perceive the projected angle to be farther from 90°, also as predicted by latency correction.

But even ambiguous projected angles—i.e., projected angles for which there are no cues as to what kind of angle caused it—are misperceived toward 90° (Fisher, 1969; Bouma and Andriessen, 1970; Carpenter and Blakemore,

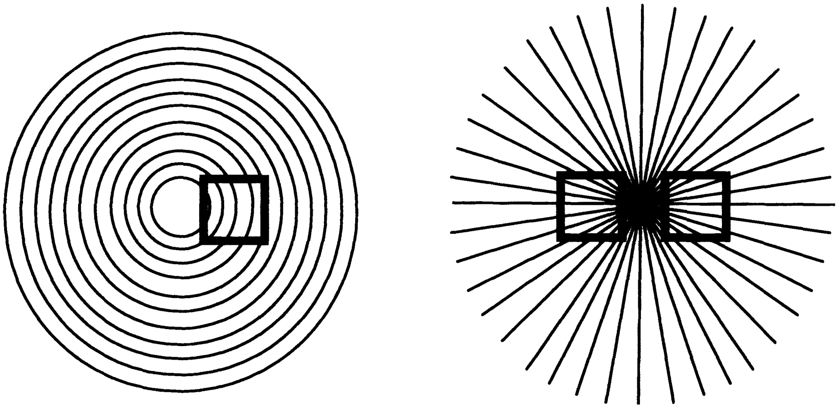

Figure 2.15: The concentric-circle version of the Orbison illusion, and the radial-line version.

1973; Nundy et al., 2000). The illusion magnitude can be as great as a couple degrees or so. For example, the projected angle in a 'less than' symbol—'<'—is perceived nearer to 90° by observers. It is as if the visual system decides that ambiguous projected angles are probably xy-z angles, since it is misperceiving them like it misperceives projected angles with cues they are xy-z angles. Why should this be?

Let us ask, what *is* the probable source of an ambiguous projected angle? There are no cues, but it may still be the case that in the absence of cues one kind of angle is much more probable. If you are in a rectangular room, look over at one of the walls and suppose you are about to walk straight toward the wall. Consider one of the corners of the wall you are approaching. It consists of three right angles, an x-y angle, an x-z angle, and a y-z angle. Two of these three angles are therefore xy-z angles, and therefore xy-z angles tend to be around twice as frequent in our experience. If a projected angle is ambiguous, then without knowing any more information about it we should guess that it is an xy-z projected angle. Another difference between x-y angles and xy-z angles is that the former project nearly as 90° when they are in front of an observer, only projecting much different from 90° when the observer is about to pass the angle (Figure 2.16). xy-z angles, on the other hand, project in all ways when out in front of an observer (Figure 2.16). This may be made more precise by placing an observer at random positions within a room with the three

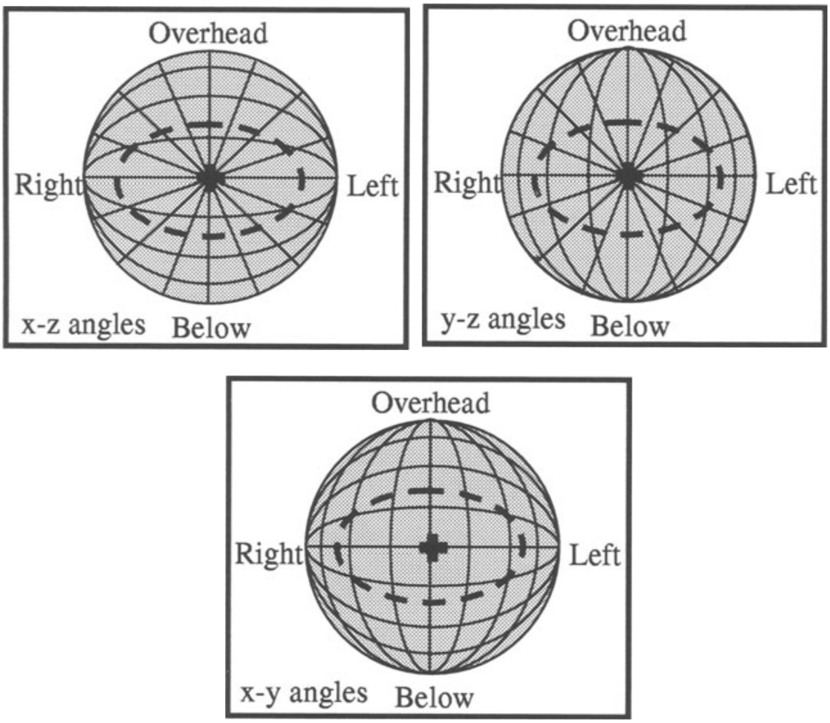

Figure 2.16: Projection spheres for the three kinds of angle. The top two are the two kinds of xy-z projected angles, and the bottom is for x-y projected angles. The dashed ellipse in each identifies the region of the projection sphere observers typically view (i.e., observers tend to look in the direction they are going, and look left and right more than up and down). The reader may observe that within the ellipse xy-z angles project in all sizes, acute to obtuse, but that x-y angles project only very near 90°.

angle types stuck in one spot, and seeing how often each projects a given angle. Figure 2.17 shows this, and we can see that when a projected angle is near 90° it is probably due to an x-y angle, but otherwise it is probably due to an xy-z angle.

In sum, then, ambiguous projected angles are probably xy-z angles if they are acute or obtuse. We therefore expect the perception of ambiguous projected angles to be like the perception of projected angles with cues that the projected angle is due to an xy-z angle, and, as mentioned earlier, this is indeed what observers perceive. In particular, the latency correction hypothesis can predict how much the projected angle should be misperceived depending on its angle. Figure 2.18 shows how, on average, the projected angle changes in the next moment as a function of the starting projected angle. We see that there is no projected angle change when the angle is very close to 0°, 90° or 180°; the projected angle change is maximally positive somewhere in between 0° and 90°, and maximally negative somewhere in between 90° and 180°. The latency correction hypothesis therefore predicts that this will be the shape of the psychophysical function for observers of ambiguous projected angles. The inset of Figure 2.18 shows how observers misperceive projected angles as a function of the angle, and the psychophysical function is indeed very similar to that predicted by the latency correction hypothesis.

2.3.3 Angular size misperception

We have now seen that the illusions of projected angle—the corner, Poggendorff, Hering and Orbison illusions—are just what we should expect if the visual system engages in latency correction. We have not, however, touched upon the double Judd, the Müller-Lyer or the upside-down ‘T’ illusion. Each of these illusions involves the misperception of an angular distance or an angular size. Even the Hering can be treated as a misperception of angular distance, since the angular distance between the two lines appears to be greater nearer the vanishing point (see Figure 2.19 for two versions of the “full” Hering illusions). The Orbison, too, can be classified as a misperception of angular size since the sides of the squares are not all perceived to be the same. In this subsection I describe how latency correction explains these angular size illusions.

Projected x and y lines

How do the angular sizes of projected x and y lines change as an observer moves forward? Let us focus on how x projections change, and what we learn

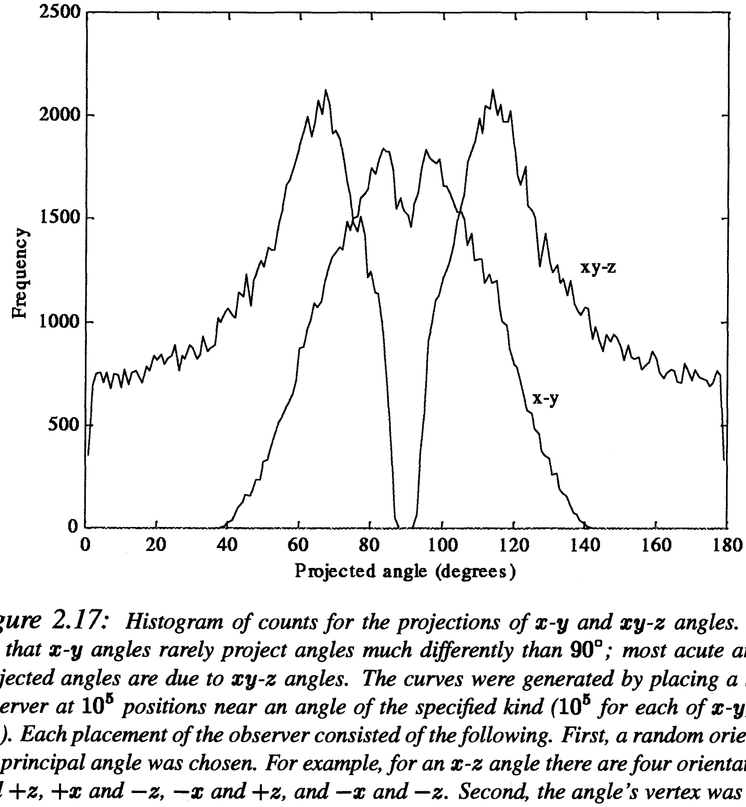

Figure 2.17: Histogram of counts for the projections of x-y and xy-z angles. One can see that x-y angles rarely project angles much differently than 90°; most acute and obtuse projected angles are due to xy-z angles. The curves were generated by placing a simulated observer at 10^5 positions near an angle of the specified kind (10^5 for each of x-y, x-z and y-z). Each placement of the observer consisted of the following. First, a random orientation of the principal angle was chosen. For example, for an x-z angle there are four orientations: $+x$ and $+z$, $+x$ and $-z$, $-x$ and $+z$, and $-x$ and $-z$. Second, the angle's vertex was placed at the origin. Third, a position for the simulated observer was determined by randomly choosing values for x uniformly between 0.1 m and 1 m to one side of the angle, values for y uniformly between 1 m above and below the angle, and values for z uniformly between 0.5 m and 1 m in front of the angle. The simulation was confined to these relatively nearby positions since one might expect that veridical perception of nearby objects matters more in survival than veridical perception of objects far away. The nature of my conclusions do not crucially depend on the particular values used in the simulation.

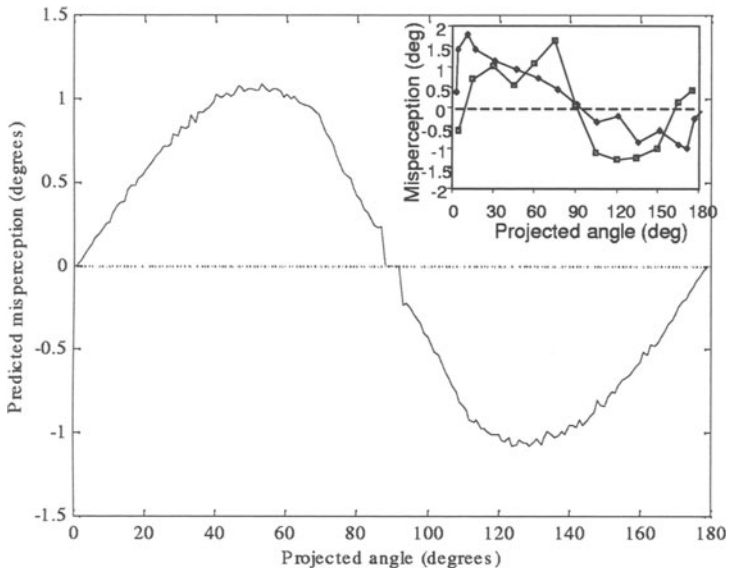

Figure 2.18: Average projected angle change as a function of the pre-move projected angle, for principal right angles lying in a plane parallel to the direction of motion (xy-z angles). One can see that the latency correction hypothesis predicts that, for projected angles that are probably due to xy-z angles, acute projected angles are overestimated and obtuse projected angles are underestimated. The graph was generated from the same simulation described in Figure 2.17, except that for each placement of the observer, the observer was then moved along the z-axis toward the angle (i.e., z got smaller) at a speed of 1 meter/sec for (a latency time of) 0.05 sec. The particular position of the peak is not important, as it depends on the allowed range of pre-move positions in the simulation. Inset shows two plots of actual misperceptions for subjects. Diamonds are averages from one representative non-naive subject (RHSC) from Carpenter and Blakemore (1973, Figure 3), and squares are averages from six naive subjects from Nundy et al. (2000, Figure 5).

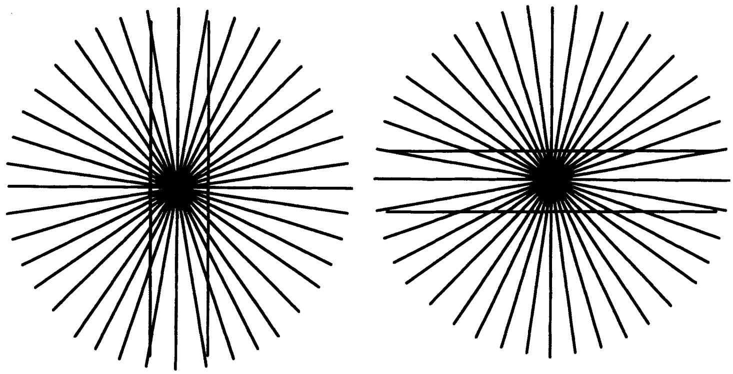

Figure 2.19: Two versions of the Hering illusion. The perceived angular distance between the two lines is greater near the middle than near the edges.

will immediately apply to y line projections as well.

Consider the angular distance between the sides of a doorway at eye level. As you approach the doorway, the angular distance between the sides increases. When you are just inside the doorway the angular distance is at its maximum of 180°. Consider how the angular distance between the sides of the doorway a little above eye level changes as you move forward. As before the angular distance increases, but it now does more slowly, and when you are just inside the doorway, the angular distance reaches its maximum at a value below 180°. The farther above or below eye level you look, the slower do the sides of the doorway expand as you approach. The angular distance between the sides of a doorway is really the length of a projected x line, namely an imaginary line extending between the two sides. The same is true for projected y lines: the angular distance between the top and bottom of the doorway increases as you approach, and does so most quickly for the angular distance between the part directly above and below your eye.

This may also be understood via examining projection spheres, as shown in Figure 2.20. In each sphere of the figure there are three pairs of squares and circles, the inner-most, the next-farther-out, and the outer-most pairs. They represent three snapshots of the horizontal (A) or vertical (B) angular distance between the points. Focusing just on (A), one can see that although the inner-most pair of squares and circles have about the same horizontal angular distance

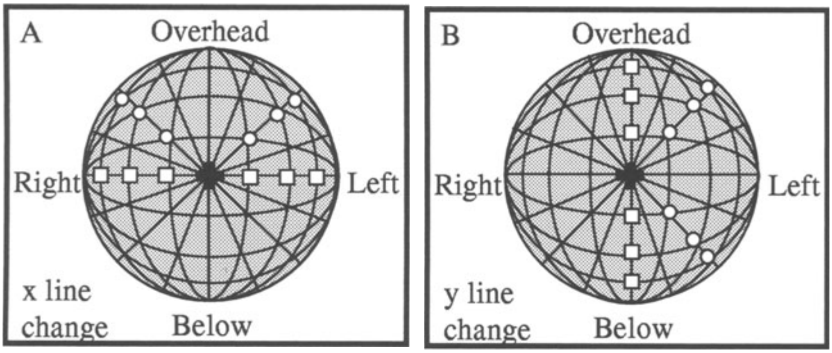

Figure 2.20: Projection spheres with x, y and z line projections. (A) This aids us in understanding how angular sizes of x line projections change as an observer moves forward. The inner-most pair of squares and circles depict the sides of a doorway that is far in front of an observer, the squares are at eye level (i.e., lying on the horizontal meridian) and the circles above eye level. The angular distance between the two squares is about the same as that between the two circles. But as an observer moves forward, in the next moment the sides of the door expand, the sides at eye level project as the next-farther-out pair of squares, and the sides above eye level project as the next-farther-out pair of circles. The horizontal angular distance between the squares is now greater than that between the circles. Similarly, in the next moment the sides are depicted by the next pair of squares and circles. (B) Identical to (A) but shows how vertical angular distances grow most quickly when they lie along the vertical meridian.

between them—this corresponds to the sides of a doorway far in front of an observer, the squares at eye level and the circles above eye level—by the time the observer approaches, the horizontal angular distance between the squares has grown considerably more than the horizontal angular distance between the circles.

There is one major summary conclusion we can make concerning how projected x lines change as observers move forward:

> *The angular distance between any point and the vertical meridian increases as observers move forward. Furthermore, this angular distance increase is maximal for points lying along the horizontal meridian, and falls off as the point gets farther away from the horizontal meridian.*

This statement is just another way of saying that as you approach a doorway, its sides bow out most quickly at eye level (and less and less quickly the further it is from eye level). The analogous conclusion holds for y lines.

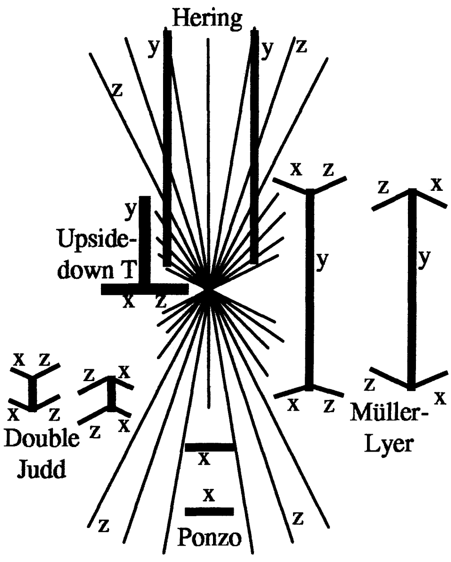

Figure 2.21: The geometrical illusions which rely on misperception of angular distance are shown again here for convenience.

> *The angular distance between any point and the horizontal meridian increases as observers move forward. Furthermore, this angular distance increase is maximal for points lying along the vertical meridian, and falls off as the point gets farther away from the vertical meridian.*

These conclusions are sufficient to explain the angular size illusions shown in Figure 2.21, except for the upside-down 'T' illusion (which I take up in the next subsubsection). I will explain each in turn.

- *Double Judd*: The double Judd illusion consists of two projected y line segments, projections which do not cross the horizontal meridian (see Figure 2.21). It suffices to treat each segment as if it were a point. We are interested in the angular distance between each segment and the horizontal meridian. They are, in fact, the same in the figure. However, the conclusion above states that in the next moment the segment nearer to the vertical meridian—i.e., the inner segment—will have a greater distance from the horizontal meridian than the other segment. The latency correction hypothesis therefore predicts that observers will perceive the segment that is nearer to the vertical meridian to have greater angular separation from the horizontal meridian. And this just is the illusion for the double Judd illusion: the inner segment of each pair in Figure 2.21 appears to be farther away from the horizontal meridian. [A similar explanation would work if the double Judd stimulus were rotated 90°.]

- *Müller-Lyer*: The Müller-Lyer illusion consists of two projected y line segments, projections which *do* cross the horizontal meridian. Consider just the tops of each projected y line. The top of the projected y line on the left in Figure 2.21 is nearer to the vertical meridian than the top of the other projected y line, and so it will move upward more quickly in the next moment. Thus, the angular distance between the top of the left projected y line and the horizontal meridian should appear to observers as greater than that for the right projected y line. The same also holds for the lower halves of each projected line, and thus the total angular distance from the top to the bottom of the left projected line will be longer in the next moment than that of the right projected line, and thus should be perceived in that way if latency correction applies. And, of course, this is the illusion in the case of the Müller-Lyer. The same explanation holds for the variants of the Müller-Lyer in Figure 2.22.

- *Ponzo*: The explanation for the Ponzo illusion follows immediately from the argument for the Müller-Lyer illusion, except that it concerns the distance from points to the vertical meridian.

- *Hering*: In the Hering illusion in Figure 2.21, there are two projected y lines on either side of the vertical meridian. The angular distance between the lines is perceived to depend on how high one is looking above or below the horizontal meridian. At the horizontal meridian the perceived angular distance between the two projected y lines is greatest, and it falls as one looks up or down. The conclusion concerning x lines above explains this: points on one of the Hering lines nearer to the horizontal meridian will, in the next moment, move away from the vertical meridian more quickly. [A similar explanation would hold if the Hering had been presented as two projected x lines lying on either side of the horizontal meridian.]

We see, then, that one simple latency correction rule underlies these three, seemingly distinct, classical geometrical illusions.

Projected z lines

The angular size and distance illusions discussed above concerned the angular sizes for x and y lines. What about the angular size of z lines? Consider how projected z line segments change as an observer moves forward. When the segment is very far away, it projects small, and as you near it it projects larger. This is no different from the behavior of x and y lines. Consider, though, how a z line projection changes when you are already relatively nearby. It still projects larger in the next moment. This is partly because it is closer, but also partly because it begins to project more perpendicularly toward the observer. Consider, as a contrast, how an x line segment lying on the horizontal meridian and to one side of the vertical meridian projects as an observer near it moves

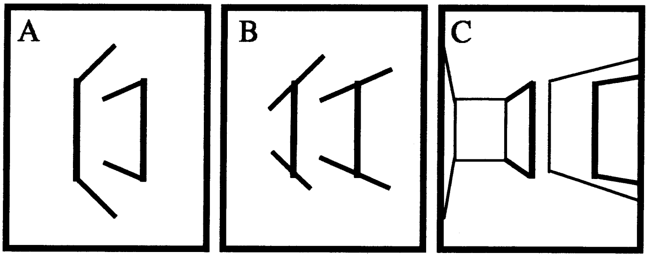

Figure 2.22: The angular size of the vertical bold lines are the same in each figure, but the left one appears larger because the cues suggest that the focus of expansion is to the left, and thus the left one will grow more quickly in the next moment. Note that in (C) the illusion is the opposite of the standard Müller-Lyer: the fins-in line appears longer than the fins-out line.

forward. Eventually, the x line begins to project less perpendicularly toward the observer—i.e., less of the line is facing the observer. When the observer passes the x line, its angular size will have fallen to zero. For the z line segment, however, when the observer passes it, its angular size will be at its maximum.

With this under our belts we can ask and answer the question of how the probable source of the upside-down 'T' illusion will change in the next moment. Recall that the source of the 'T' is a corner made of an x, y and z line, whose point lies on the horizontal meridian, and thus so does the x and z line. The probable focus of expansion is somewhere on the same side as the z arm, but past the tip of the projected z arm (e.g., see Figure 2.21). The angular size of the horizontal bar is due to the sum of the angular sizes of the x line and the z line, these lines being at right angles to one another in the world. Suppose each line has a length of L meters. It's angular size could then be mimicked by a single straight real world line (it is not a principal line) going from the tips of each line that is the square root of $L^2 + L^2$, or $1.414L$. The y line must, then, be approximately $1.414L$ meters long as well, since it projects the same angular size and is approximately the same distance away. Consider now what happens when the observer is about to pass the corner. Since the x line is to one side of the vertical meridian, its angular size has fallen to $0°$. The angular size of the z arm is at its maximum, however. The bottom of the y arm rests on the horizontal meridian, and it will therefore not get smaller in the last moments before passing it, but, instead, will increase to its maximum. Since the z line is of length L and the y arm length $1.414L$, and since each is about the same distance from the observer, the angular size of the y arm will be about 1.41 times as large as the angular size of the z arm. This is how the corner will project when the observer is just passing it, but the more general conclusion is, then, that the total angular size of the bottom of the 'T' grows less quickly than does the angular size of the y line. Latency correction therefore predicts that observers will perceive the vertical line to have greater angular size, as is the case.

A new illusion

In the explanation of the upside-down 'T' illusion, we learned that, when relatively nearby, x line segments lying on the horizontal meridian and on one side of the vertical meridian—like the one in the upside-down 'T' illusion—increase their angular size more slowly than do z line segments lying in the same part of the visual field. We can use this observation to build a novel illusion. Figure 2.23 shows two identical horizontal lines lying on the horizontal meridian, one on each side of the vertical meridian. The one on the left has cues suggesting it is due to an x line, and the one on the right has cues that it is due to a z line. Although they are at equal distances from the vertical meridian, the z line appears to have greater angular size, as latency correction predicts. (The bold vertical lines are also identical in angular size to the horizontal lines.)

2.3.4 Psychophysical confirmation

It is possible to summarize the explanation for all those illusions that did not rely on misperception of the angular size of z lines; i.e., all the illusions except for the upside-down 'T' and the new illusion just discussed above. Figure 2.24 shows how much a point in an observer's visual field moves away from the horizontal meridian in the next moment. The figure for movement away from the vertical meridian is identical, but rotated 90°.

I will describe how this one plot explains most of the illusions discussed thus far.

- *xy-z projected angles, including corner, Poggendorff, Hering, and ambiguous angle perception*: The two white dots in Figure 2.24 can be thought of as the endpoints of an x line extending between them. The figure indicates that the dot nearer to the vertical meridian will move up more than the other dot in the next moment. Consider the angle this projected x line segment makes with a projected z line (which goes radially outward from the focus of expansion in the figure). The projected x line will, after the move, make an angle with the projected z line that is more near 90° than it originally was. This

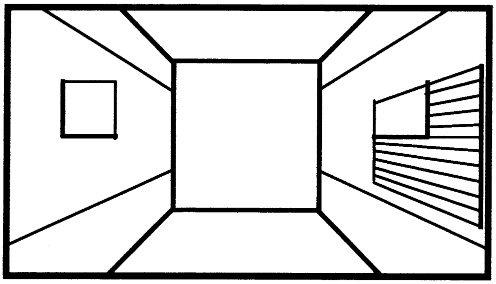

Figure 2.23: Two predicted illusions. First, the left horizontal line appears to have smaller angular size than the right one, but they are identical. The reason is that the right one is probably due to a z line (being part of the flag on the wall), whose angular size will increase in the next moment more than that of the x line on the left. Second, and for the same reason, the horizontal line on the right appears to have greater angular size than the adjacent vertical line, but the two lines on the left appear roughly identical (and, the predicted perception on the left is that the vertical line should be a little larger than the horizontal line).

captures the regression to 90° phenomenon we discussed earlier. The explanation for y-z projected angle illusions is similar, but relies on the plot that is rotated by 90°.

- *x-y projected angles, and the Orbison illusion*: We just learned that the projected x line extending between the two points in Figure 2.24 will "lift up" on its left side (i.e., its left side will acquire greater angular distance from the horizontal meridian than the right side). Consider the projected angle the x line makes in the next moment with a y line. The projected angle begins at 90°, but gets pushed away from 90° in the next moment. Projected y lines also change, and change so as to accentuate this projected angle change.

- *x and y angular distances, including the double Judd, Müller-Lyer, Ponzo and Hering illusions*: When we just consider the two dots in Figure 2.24, we have the raw material of the double Judd illusion, and the plot states that the one nearer to the vertical meridian moves away from the horizontal meridian more in the next moment, which agrees with perception. Not only does the dot on the left appear higher, the angular distance between it and the horizontal meridian appears greater, which is essentially the Müller-Lyer, Hering and Ponzo illusion.

Since Figure 2.24 encapsulates most of the predictions my model of latency correction has made, it would be nice if we could test observers to see if their perceptions of the angular distance between each dot and the horizontal merid-

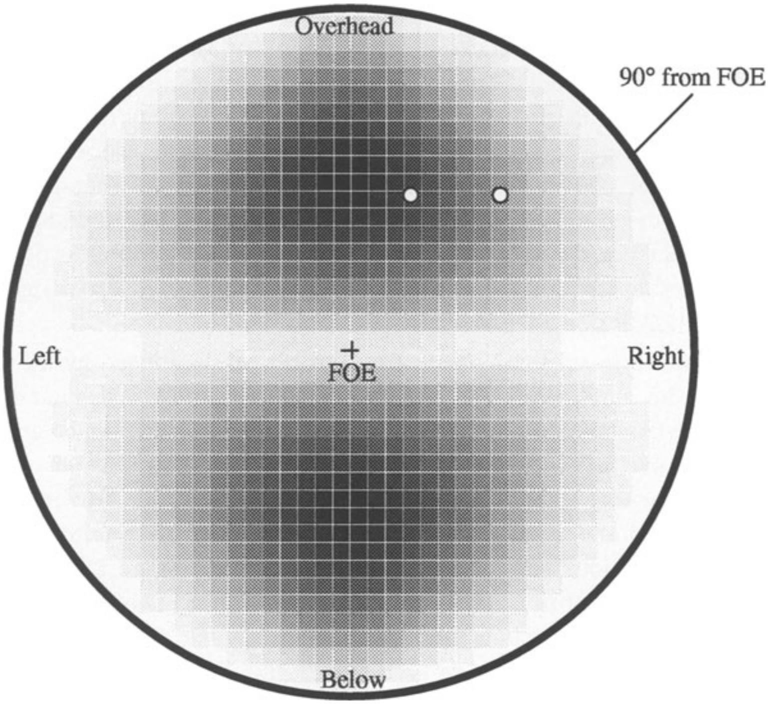

Figure 2.24: Change in angular distance from the horizontal meridian as a function of position within the visual field. Rim of circle is 90° from the focus of expansion. Plot uses a linear gray scale, with white representing zero degrees angular distance change, and black representing approximately two degrees. The two dots are props referred to in the text. By rotating the plot by 90°, one obtains the plot for the change in angular distance from the vertical meridian as a function of position within the visual field.

ian fits this composite prediction. This is what an undergraduate student and myself did, with intriguing and encouraging results (Changizi and Widders, 2002).

Using a computer, two dots were placed on a radial display of black lines, the whole display was 20 cm in diameter, and subjects typically sat about one to two feet from the screen (this was not experimentally controlled). The dots were kept horizontally separated by about 2 cm, and were red to be easily distinguished from the radial lines. They were moved as a pair to each of 300 different positions in an 18 by 18 grid in the radial display (six positions at the extremity of each quadrant lie outside the radial display and were not measured). For each position, the subject was asked to move the outer dot (the one farther from the vertical meridian) up or down until its perceived angular distance from the horizontal meridian was the same as that for the less peripheral dot. The resolution was roughly a third of a millimeter. (See Changizi and Widders, 2002, for detailed methods.)

The data from subjects is not of a form directly predicted by the plot in Figure 2.24 because the subjects were judging the *difference* in angular distance from the horizontal meridian, whereas the plot measures how much any given point will move upward in the next moment. Instead, the predictive plot we want is the one that records, for each point in the visual field, how much more the less peripheral dot will move away from the horizontal meridian than the more peripheral dot. This plot can be obtained from Figure 2.24 by simply taking, for each point in the visual field, the next-moment angular distance of the less peripheral dot minus the next-moment angular distance of the more peripheral dot. This is shown in Figure 2.26; this figure shows the predicted strength of the vertical angular distance illusion as a function of position in the visual field. This one plot encapsulates the predicted illusion magnitude for nearly all the illusions discussed in this section. If the visual system follows a latency correction strategy, then we expect it to approximate the predicted plot, at least to first order; this plot is the fingerprint of latency correction. The predicted plot assumes that all points are equidistant from the observer, whereas in reality it may be that points at different positions in the visual field have different probable distances from the observer. However, the basic "bauplan" of the predicted plot is expected to be followed, even if not the particulars.

Figure 2.27 shows averages from the above described experiment for myself, David Widders, and one naive undergraduate (NG), along with the average of our averages. In each case, the psychophysical results have the latency correction fingerprint, which provides further strong confirming evidence for

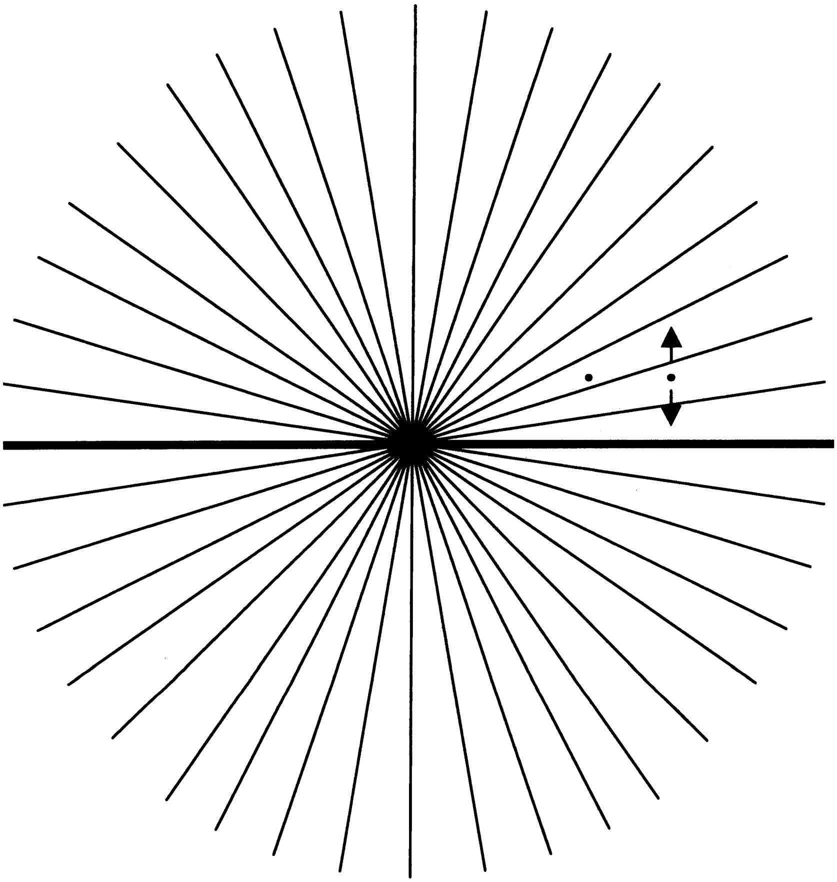

Figure 2.25: An example of the stimulus used in the psychophysical test of the latency correction hypothesis. The arrows indicate that the more peripheral dot could be moved up and down by the subject.

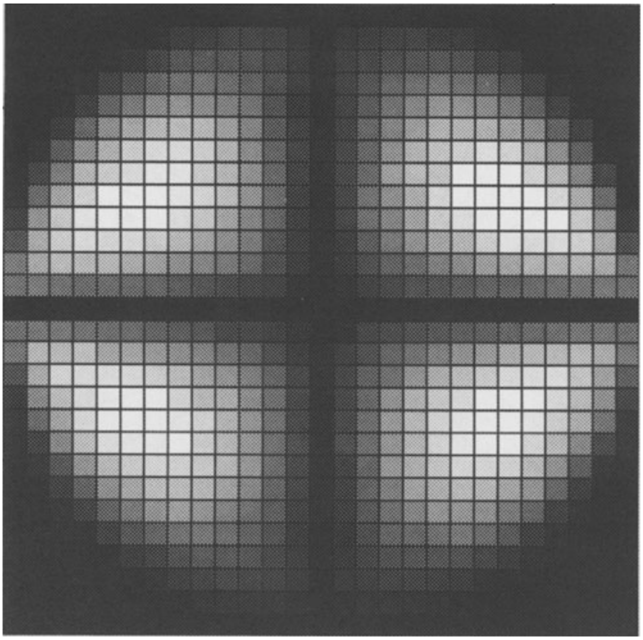

Figure 2.26: This plot shows the predicted misperception for the vertical angular distances from the horizontal meridian for two horizontally displaced dots, as a function of position of the points in the visual field. This plot is the predicted fingerprint of a latency correction strategy for vision. The plot is generated by assuming that all dots are at the same distance from the observer. Whiter here means greater predicted illusion.

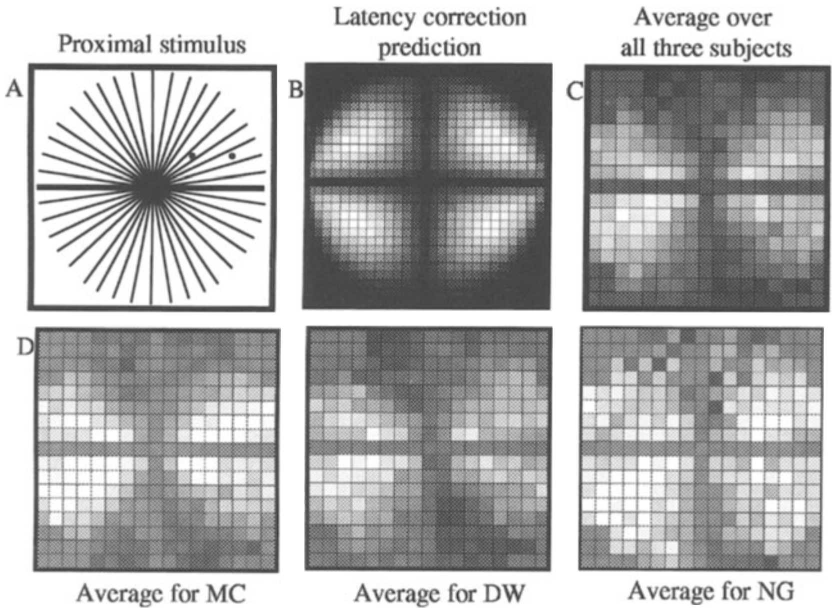

Figure 2.27: (A) The general kind of stimulus used in the experiment is repeated here for convenience, as is (B) the predicted plot for the latency correction hypothesis. (C) The average of the average results over the three subjects. (D) Three experimental plots for three subjects individually, the first two (non-naive) averaged over four experiments, and the last (naive) averaged over two runs. The range of misperceptions for the three subjects are approximately, in centimeters: MC $[-0.07, 0.13]$, *DW* $[-0.04, 0.14]$, *NG* $[-0.03, 0.11]$, *and average of the averages* $[-0.04, 0.11]$. *Experiments for angular-distance-from-vertical-meridian perceptions were similar. Note that the predicted plot ranges over the entire visual field, whereas the experimental results are for some subset of it. Whiter means greater illusion. For each plot, zero misperception is represented by whatever gray level lies along the horizontal and vertical meridians (where subjects experience no illusion). [I thank Nirupa Goel for being the naive subject here.]*

latency correction. Not only do observers experience illusion gradients in the areas predicted, but the illusion magnitude tends to be more clustered, in any given quadrant, nearer to the horizontal meridian, which is also a property found in the predicted plot. Our experimental results are qualitatively identical for the perception of angular distance from the vertical meridian. We have also noticed a tendency for greater illusion magnitudes in the bottom half of the visual field, which may be due to the fact that, on average, objects tend to be nearer to observers in the lower half of their visual field, and they consequently move more quickly in the visual field in the next moment.

2.4 Further directions for latency correction

In this chapter I have introduced a basic strategy for vision, a strategy so useful that we might expect any kind of computational system to utilize it. That strategy is latency correction: rather than carrying out computations whose intent is to provide a solution relevant to the problem that initiated the computation, the intent is, instead, to provide a solution that will be relevant when the computing is finally finished. This strategy is useful because it allows an optimal tradeoff between fast computation and powerful computation. More powerful computations can be carried out if the system has more time, and the system can buy itself more time for computing if it can correct for this computing time, or latency. I have concentrated only on vision in this chapter, and provided evidence that the visual system utilizes a latency correction strategy. The evidence thus far has concerned the perception of classical geometrical illusions; we have seen that observers perceive the projected angles and angular sizes of scenes not as they actually project, but as they probably will project in the next moment, i.e., at the time the percept is actually elicited.

The explanatory value of the latency correction hypothesis is, I believe, much greater than just explaining the classical geometrical illusions or cases such as the flash-lag effect (which I will mention below). I believe that a considerable fraction of all visual illusions may be due to latency correction; in particular, I believe that all inconsistent perceptions are due to latency correction. There is much work ahead of us in understanding the consequences of latency correction, and before leaving this chapter I will discuss preliminary ideas and research in progress.

Motion-induced illusions

Evidence for latency correction in the literature has, except for my own work, concentrated on motion-induced illusions. (The stimuli in the work I have described here are all static.) The most famous effect is called the flash-lag effect, where an unchanging object is flashed in line with a continuously moving object such that, at the time of the flash both objects are identical (MacKay, 1958; Nijhawan, 1994, 1997, 2001; Schlag et al., 2000; Sheth et al., 2000). [There has also been a fireworks-like debate about this interpretation (Baldo and Klein, 1995; Khurana and Nijhawan, 1995; Whitney and Murakami, 1998; Purushothaman et al., 1998; Lappe and Krekelberg, 1998; Krekelberg and Lappe, 1999; Whitney et al., 2000; Eagleman and Sejnowski, 2000; Brenner and Smeets, 2000; Khurana et al., 2000).] The continuously moving object appears to be "past" the flashed object, even though they are identical. In the first flash-lag effect, the continuously moving object is a rotating bar, and the flashed object is a light that flashes in line with the moving bar; observers perceive the flashed light to lag behind the moving bar. [Some of this extrapolation may even be carried out by retinal ganglion cells (Berry et al., 1999).] Sheth et al. (2000) showed that the effect holds for other modalities besides perceived position. The continuously changing stimulus may be in the same position, but changing in luminance from dim to bright, and the flashed stimulus has, at the time of its appearance, the same luminance as the other stimulus; in this case observers perceive the changing stimulus to be brighter than the flashed one. It also works for hue and other modalities. Other evidence for latency correction can be found in Thorson et al. (1969) who have shown that when two very nearby points are consecutively flashed, motion is perceived to extend beyond the second flashed point. Also, Anstis (1989) and DeValois and DeValois (1991) have shown that stationary, boundaryless figures with internal texture moving in a direction induce a perceived figure that is substantially displaced in the same direction (see also Nishida and Johnston, 1999; and Whitney and Cavanagh, 2000).

One difficulty with these illusions is that they cannot be shown in a book; one needs a computer display or real live moving objects to see them. There are, however, two illusions from the literature that are motion-induced and *are* able to be displayed here. Furthermore, although neither illusion was introduced by the authors for the purposes of latency correction, there is a relatively straightforward latency correction explanation for both.

The first is due to Foster and Altschuler (2001) and is called the Bulging

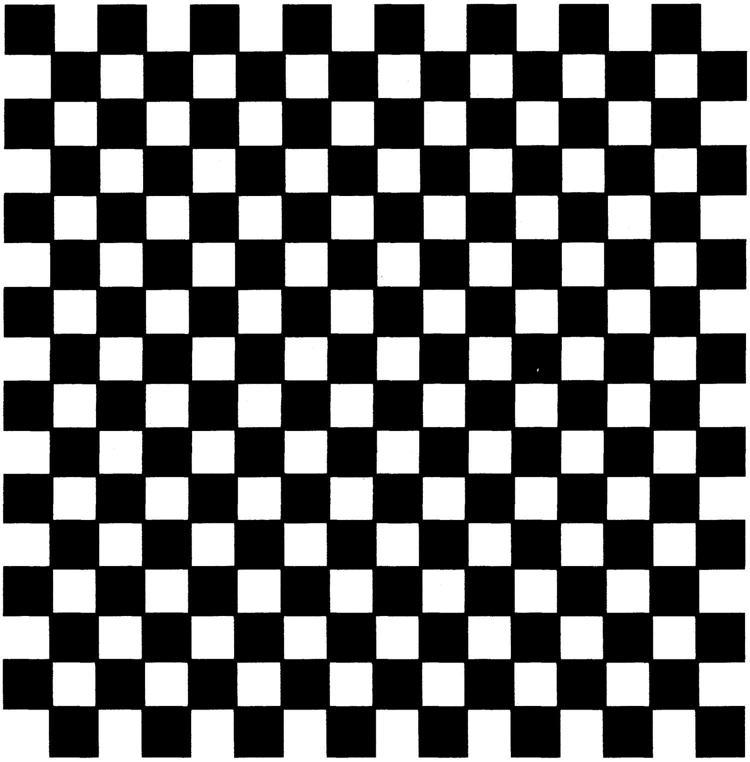

Figure 2.28: The Bulging Grid illusion (Foster and Altschuler, 2001) occurs when you move your head quickly toward the checkerboard. In addition to the perception of a bulge, the projected angles and angular sizes change just as in the Orbison illusion. The Orbison illusion does not require observer motion because the radial lines provide the cue as to the focus of expansion.

Grid (Figure 2.28). Before I tell you what the illusory aspects of it are, note that it is essentially a bunch of squares, or projections of x and y lines. The Orbison illusion (see Figure 2.1), recall, was when the cues suggest that the probable source consists of x and y lines, and the radial lines in the Orbison acted as cues to the observer's direction of motion—the vanishing point was the probable focus of expansion. The Bulging Grid figure is, in a sense, then, like the Orbison illusion, except that it does not possess any cues as to the probable focus of expansion. Well, there is one obvious way to create a probable focus of expansion: move your head toward the image. There is arguably no better cue to a focus of expansion than optical flow emanating radially from some point. We should predict that if an observer moves his head toward it, he should experience an Orbison-like illusion. Indeed this is what occurs. Try it. Forget for the moment about the bulge, and focus just on the perception of the projected angles and the angular sizes. The angles change away from 90° as in the Orbison illusion, and in the same ways. [What about the bulge? I have no good answer to this as of yet, although I can make two observations. First, if one overlays the bulging grid (or any grid of squares) with a radial display, one also perceives a bulge (albeit smaller). This suggests that the radial lines are indeed serving to cue direction of motion. Second, a bulge *is* consistent with the misperceived projected angles, although it is inconsistent with the actual projected angles. That is, the actual projected angles are best explained by a flat grid in front of the observer, and so I would expect that they would be perceived as such. Instead, the perception of a bulge suggests that it is as if the visual system determines the perceived projected angles according to latency correction, and *then* uses these angles to compute the probable depths. At any rate, more thinking on this is needed.]

The next motion-induced illusion worth bringing up is one by Pinna and Brelstaff (2000), and is displayed in Figure 2.29. In my opinion is it is *the* most striking illusion ever; plus it requires no complex computer display or stereoscopic glasses, etc. It was invented by them without latency correction in mind, but it has a relatively straightforward latency correction explanation. When you move toward the point at the center that point becomes the probable direction of motion. What kind of scene probably generated this stimulus? One intuitive conjecture is that the observer is walking down a circular tube or tunnel. Consider just the inner ring. The oblique lines pointing roughly toward the center do not actually point at the center. If they did, then the lines would probably be the projections of z lines. Instead, they point inward and a little counter-clockwise. What kind of line would project this way? Answer:

A line that was painted on the inside wall of the tube, but was spiraling around it and going down the tube simultaneously. That is, if there were lines on the inside wall of the tube that wrapped counter-clockwise around the tube once every, say, ten meters, the nearest segments of those lines would project just like the nearly-radial lines of the inner ring. We must also suppose that, for whatever reason, the observer is only able to see the nearby parts of these spiraling lines. Now that we have an idea of what kind of scene might cause such a stimulus as the inner ring, we can ask how that scene will project as the observer moves forward. In the next moment, the spiraling lines nearest the observer will no longer be at the same positions, but will, instead, have rotated or spiraled counter-clockwise a little. That is, as the observer moves forward, the spirals will move counter-clockwise around him. And this is exactly the illusion we experience here. The illusion may, then, be due to the visual system attempting to engage in latency correction; but it is inappropriate here since there is no tube. The explanation is similar for the outer ring, and is similar for when you move your head away from the stimulus rather than toward it. Much work is needed to examine in detail such images and whether latency correction really is the explanation. At this time, it is just highly suggestive and encouraging.

Brightness and color

There is considerable evidence since Helmholtz that, when cues make it probable that a surface has a certain reflectance, brightness and color judgements—to be distinguished from lightness and surface color judgements—are influenced away from the actual luminance and chromaticity in the proximal stimulus and partially towards the probable reflectance, or partially towards the "typical" or "generic" luminance and chromaticity emitted by the probable surface (Arend and Reeves, 1986; Arend and Goldstein, 1990; Arend et al., 1991; Arend and Spehar, 1993; Adelson, 1993; Kingdom et al., 1997). There has been little success, however, in explaining this "regression toward the 'real' object" (Thouless, 1931a, 1931b) phenomenon. The reason it has been difficult to explain these brightness and color illusions is that they are cases of inconsistent perception (see Subsection 2.1.1). For example, in brightness contrast (Figure 2.30) two identical gray patches are surrounded by, respectively, dark and light surrounds. The patch in the dark surround is perceived to be lighter *and* brighter than the other patch.

That the patch is perceived to be lighter—i.e., perceived to have greater

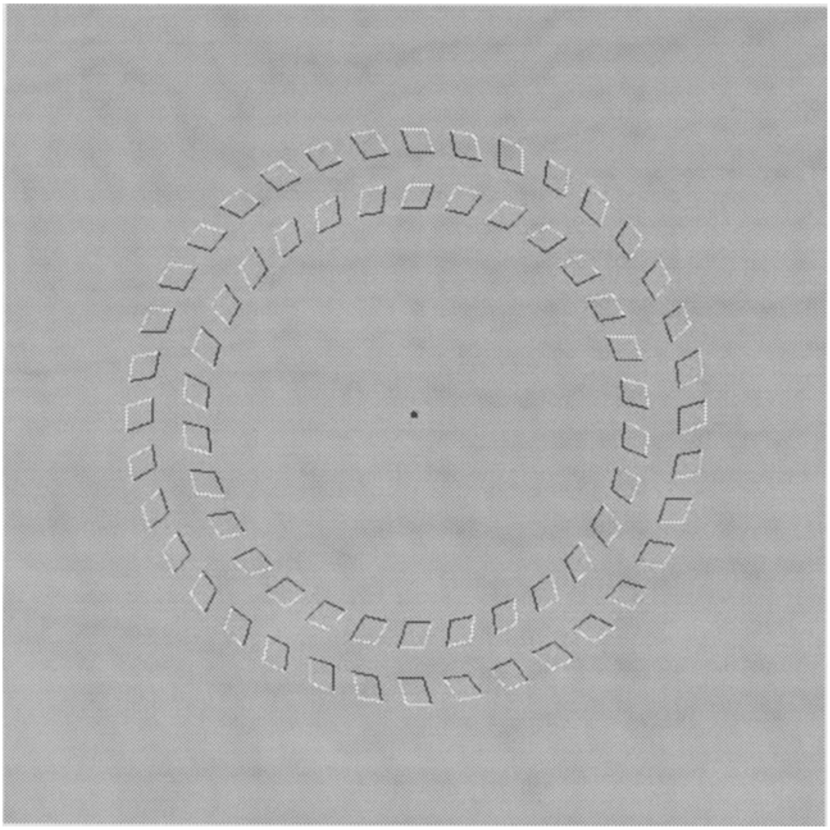

Figure 2.29: In this striking illusion from Pinna and Brelstaff (2000), you should move your head either toward or away from the figure while focusing on the point at the center.

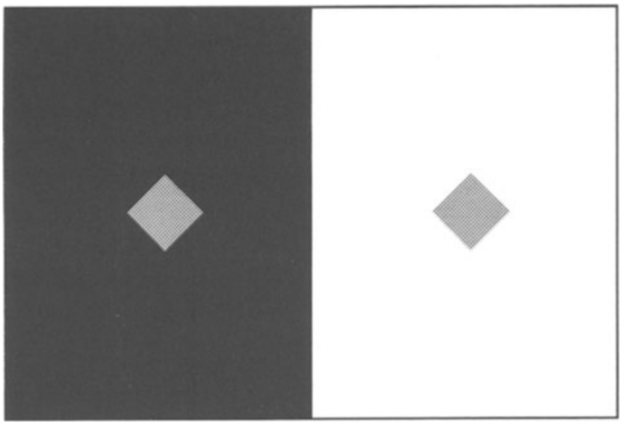

Figure 2.30: Demonstration of lightness and brightness contrast. The gray patches are identical, but the one in the dark surround appears to have greater reflectance than the one in light surround. This is called lightness contrast, and is easily accommodated by an inference approach (namely, a "subtracting the illuminant" account going back to Helmholtz). The display also demonstrates brightness contrast, where the patch in dark surround appears to send more light to the eye (i.e., appears to have greater luminance) than the patch in light surround.

reflectance, or greater ability to reflect more light—is easily explained by an inference or Bayesian framework. The explanation is that the dark surround suggests that its patch is under low illumination, and the light surround suggests that its patch is under high illumination. Since the patches have the same luminance—i.e., they send the same amount of light to the eye—the patch in the dark surround must be a more reflective object. This is sometimes referred to as the "subtracting the illuminant" explanation. The same idea applies for perception of surface color, which refers to the perception of the reflectance of the object, where now we care about the full spectral reflectance properties, not just the amount of light reflected.

However, the explanation for why the patch in the dark surround is perceived to be *brighter*—i.e., perceived to have greater luminance—is *not* explainable by the traditional inference or Bayesian account. The reason is that the luminances of the two patches are probably identical; the retina "knows" this. Yet the brain generates a percept of the patch in dark surround having greater luminance than the patch in light surround. The brain therefore generates a percept that is inconsistent with the proximal stimulus. As we discussed earlier in this chapter, inconsistent perception *can*, in principle, be accommodated within a latency correction approach. This observation led me to look for latency correction explanations for brightness illusions. Similar arguments lead us to the same conclusion for the perception of color—perception of the chromatic quality of the light sent to the eye, or perception of the chromaticity—as opposed to the perception of surface color.

Latency correction is, indeed, suggestive of an explanation for brightness (and color) contrast illusions. As an observer walks through the world, the luminance and chromaticity received from any given surface can change radically as a function of the surface's angle with respect to the observer. It is reasonable to assume that the following is true:

> If a surface currently has a luminance/chromaticity that is atypical for it, then the luminance/chromaticity is probably going to become more typical in the next moment, not less.

For example, if a surface with high (low) reflectance currently has low (high) luminance, then it is more probably going to have higher (lower) luminance in the next moment. Similarly, if the chromaticity from a red surface is currently yellowish (because of a yellow illuminant), the chromaticity is more probably going to become less yellowish and more reddish in the next moment. Latency correction accordingly predicts that if cues suggest that the actual lumi-

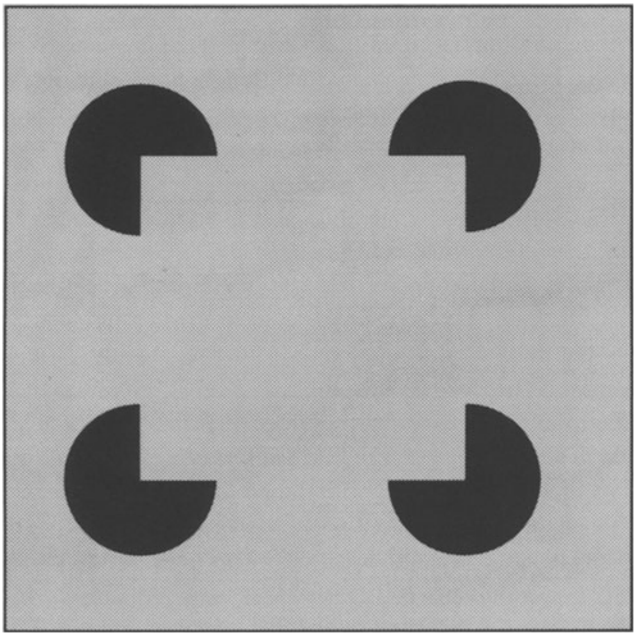

Figure 2.31: The Kanizsa square. An illusory square is perceived, along with a illusory luminance contours at its top and bottom.

nance/chromaticity is atypical for the probable source, an observer will perceive a brightness/color representative of a luminance/chromaticity more toward the typical luminance/chromaticity of the probable source, as that is more probably what will be present at the time the percept is elicited. This prediction is qualitatively consistent with actual psychophysical trends in the perception of brightness and color, as cited above.

Even the illusory contour phenomenon is a case of inconsistent perception, as one perceives luminance contours despite the proximal stimulus being inconsistent with luminance contours. Illusory contours are perceived along the edges of objects that are probably there, like in the Kanizsa square (Figure 2.31). This may be expected within latency correction, however, since although there is no luminance discontinuity at the time of the stimulus, since there probably *is* a surface discontinuity, it is probable that in the next moment there *will* be a luminance discontinuity. That is, a surface discontinuity without a luminance discontinuity is a rare situation, and it is much more likely to have a luminance discontinuity by the time the percept occurs, so the visual system includes one.

At best, though, this is all just encouraging; it does not provide any strong evidence that latency correction explains brightness and color illusions. These illusions are merely roughly what one might, *prima facie*, expect if latency correction were true. What I need are more detailed models akin to what I have put forth for geometrical stimuli. Such theoretical work is in progress.

I leave this subsection with a very exciting illusion that strongly suggests that brightness illusions *will* fall to latency correction explanations. It is a motion-induced illusion like the Bulging Grid and the spiral discussed earlier—each seemingly explainable by latency correction. This illusion, however, is a brightness illusion, also with a latency correction explanation. The illusion is due to David Widders, an undergraduate student of mine, and is shown in Figure 2.32. Move your head towards the center of the figure, and you should perceive the middle to become brighter and the dark edges to become brighter. The brightness appears to flow radially outward. If the probable scene causing such a stimulus is a tunnel with an illumination gradient along it (due to, say, some light at the end), then latency correction would predict such an outflowing brightness illusion, since that is how the luminances would be in the next moment. If you move your head backward the effect is the opposite. Even more interestingly, the illusion works for hue and other gradients as well, and is especially stunning on a computer or a glossy printout. We are further examining motion-based illusions of this kind within a latency correction framework.

Representational momentum

There exists a literature, possibly of great relevance to latency correction, called "representational momentum" (see, e.g., Freyd, 1983a, 1983b; Freyd and Finke, 1984, 1985; Hubbard and Ruppel, 1999, and references therein). The phenomenon is as follows. Freyd (1983b) showed subjects images taken from a scene possessing motion. The images possessed ample cues as to the motion in the scene, and two images were chosen from successive moments in the scene, so that one image, A, obviously just preceded the next, B. Subjects were presented two images in succession, and asked to say whether the two were the same or different. When the images were presented in the order they actually

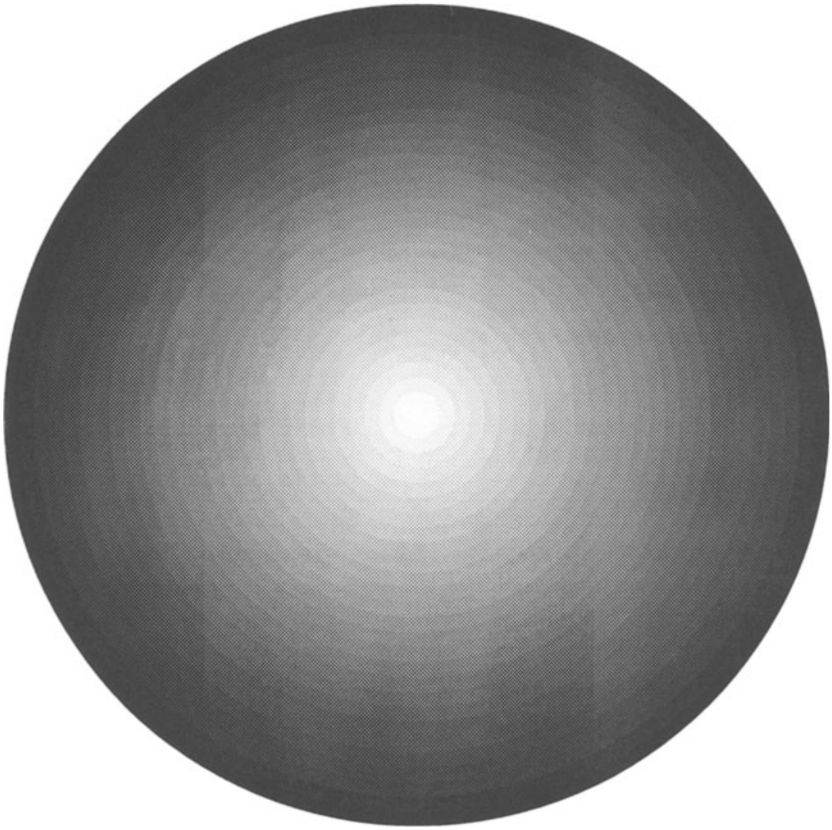

Figure 2.32: Move your head toward the center of this figure, and you should perceive the brightness to "flow" outward towards the edges. It works best on either a computer screen or glossy paper. If the stimulus is probably due to a tunnel with an illumination gradient, then as an observer moves forward the brightness will, indeed, "flow" outward. Thus, the illusion is consistent with latency correction. The illusion was invented by David Widders.

occurred—A followed by B—subjects took more time to respond that they were different than when the images were presented in the opposite order. It is as if subjects, upon seeing image A, forward it a little, so that by the time B is displayed, their memory of A is already depicted in B, and they have difficulty noticing any difference. I introduce the connection here only to note that representational momentum may be another long-known effect that, like the classical illusions, may be due to latency correction.

Other cues to the focus of expansion

Thus far, the cues to the location of the focus of expansion have been projected z lines, which, as we have discussed, may be due either to real live z line contours, or may be due to optic flow itself (that is, the radial lines may mimic optic blur). Optic flow is not actually necessary, however, to perceive forward movement (Schrater et al., 2001); all that is necessary is that the overall size of image features increases through time. Accordingly, we might expect that we can create a static image such that there is a size gradient, with larger image features (i.e., larger spatial frequency) near the periphery and smaller image features near the center. Such an image would suggest that the peripheral parts of the image are nearby, that the parts nearer to the center are farther away, and that the "flow" is that the smaller image features near the center are becoming the bigger images features on the sides. The probable focus of expansion therefore is the center, and we expect to find the same kinds of illusions as in the radial display stimuli from earlier.

In this light, consider moving down a tubular cave with constant-sized "rocks" along the inside wall. At any given radial angular distance from the focus of expansion, all the rocks at that angular distance will project at the same size. Thus, arguing backward, *if two features in an image have the same size (i.e., the objects causing them are projecting the same size), then they are probably due to rocks at the same angular distance from the focus of expansion.* Consider Figure 2.33 (A) which shows a bunch of similar-sized projected shapes, and the probable location of the focus of expansion is shown at the center. Since this is the probable focus of expansion, the vertical dotted line on the left should increase in angular size more in the next moment than the equal angular size vertical dotted line on the right. And this is what observers perceive. Actually, this is now just like a standard class of variants of the Müller-Lyer illusion, one which is shown in Figure 2.33 (B). These variants may, then, be explained by the fact that the probable focus of expansion for them is a point

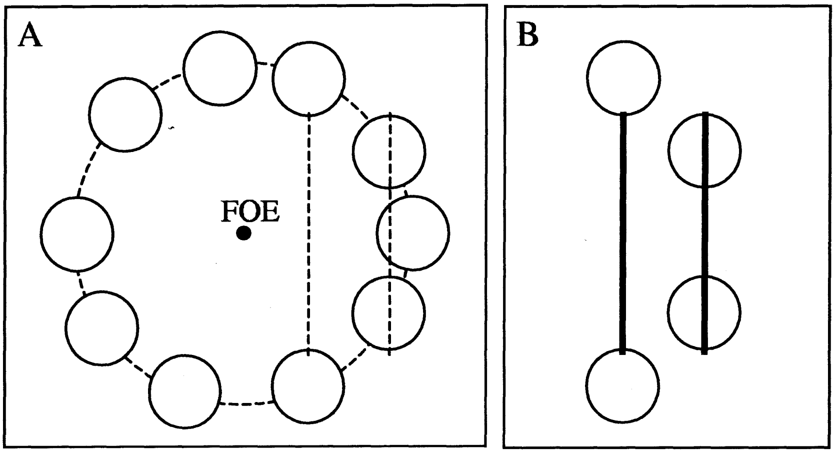

Figure 2.33: (A) When there is a set of projected shapes all of the same size and projecting roughly equal angular distances from a single point in the visual field, this point is probably the focus of expansion. (This is because the projected shapes are probably due to similar objects at similar positions relative to the observer.) Since the observer is probably, then, moving toward this point, the angular distance of the dotted line on the left is expected, if latency correction is true, to be perceived to have greater angular size than the one on the right. This is consistent with what observers, in fact, perceive. (B) Furthermore, this explains a class of variants of the Müller-Lyer illusion, where there is some object—in this case a circle—outside and inside the vertical lines. The vertical line with the object outside is always perceived to have greater angular size, no matter the object's shape. This is because the probable focus of expansion is the point that is equi-angular-distant from the four objects, and this point must be nearer to the left vertical line, as depicted in (A).

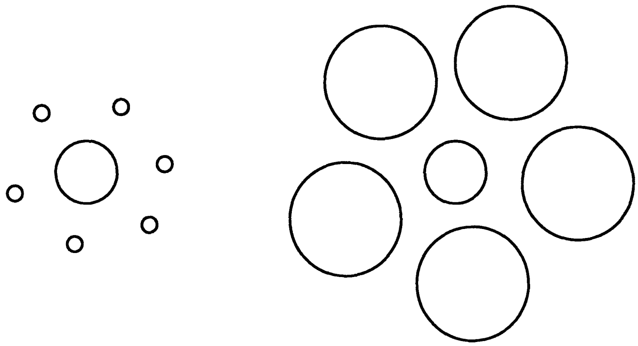

Figure 2.34: The Ebbinghaus illusion: the middle circles on the left and right are identical, but the one on the left appears to have greater angular size. Since the circle on the left is surrounded by smaller circles, it is probably nearer to the focus of expansion, and should increase in angular size in the next moment (supposing it is not too different in distance from the observer than the other circle).

nearer to the line with the objects on the outside of the line.

The same observations concerning the tubular cave above can allow us to make another qualitative conclusion. The objects nearer to the focus of expansion project smaller, being farther away. Thus, all things equal, *the part of the visual field with smaller image features is more probably nearer the focus of expansion.* Consider now Figure 2.34 showing a classical illusion called the Ebbinghaus. The circle in the middle on the left and in the middle on the right are identical, and yet the one on the left appears to have greater angular size. This is readily explained by the conclusion just made: since the circle on the left is surrounded by smaller projections, it is probably nearer to the focus of expansion than is the circle on the right. Supposing that the probable distances from the observer for each of the two middle circles is roughly the same, the one on the left will increase in angular size more quickly in the next moment. Thus, latency correction predicts that the left middle circle will be perceived to be larger, which is consistent with actual perception.

These ideas are preliminary at this point, but very promising. They suggest that my latency correction model may be easily extendable beyond x, y and z lines, so that it may be applied to stimuli with many other kinds of projections.

Further discussion of these ideas will have to await further research on my part.