THE ECological APPROACH TO VISUAL PERCEPTION

James J. Gibson
Cornell University
I suggested in the last chapter that a picture is a surface so treated as to make available an arrested optic array, of limited scope, with information about other things than the surface itself. What, then, is a motion picture? There is a treated surface, but the treatment has to consist of throwing shadows on the surface by projection instead of depositing traces or pigments on it. An optic array of limited scope is delivered, and it contains information about other things than just the surface itself. The main difference is that the array is not arrested. Its structure undergoes change, disturbance, or transformation. It is not frozen in time. And that is what we need to understand about it.

This definition of the motion picture is broad enough to include not only the kind made with photographic film that uses the stroboscopic principle in the projector and camera but also the kind made with a modulated scanning beam as in television, and the kind made with shadows projected on a translucent screen as in a shadow play, and the various kinds made with the optical gadgets now being tried out experimentally by kinetic artists.

The technology of cinema and television has reached the very highest level of applied science. The psychology of the awareness provided by a motion picture, however, is nonexistent, apart from an essay by J. Hochberg and V. Brooks (1978), to whom I am indebted for much good talk about the problems of the film. There are no experts on this form of perception. Muddles and misconceptions prevail. We are led to conceive a sort of apparatus inside the head that is similar to the apparatus for making a picture show outside the head. We have been taught that a picture is sent from the eye up to the brain, and so we conclude that a series of pictures can be sent up to the brain. We all know what a snapshot is, and we know that a film is a series of snapshots. If we are told that a movie presents us with a sequence of retinal snapshots joined by what is called the "persistence of vision," we believe it. But we are misled. Nevertheless, this is what we are told by movie commentators who have read physiological optics and believed it.

The motion picture camera and projector do not comprise the only method that can be used to produce a changing optic array. Nor is the stroboscopic principle the
only principle that can be applied. There are other ways of doing it, and the inventors of the nineteenth century tried out dozens of gadgets with names such as kinescopes and kinegraphs, vitascopes and vitagraphs, that have now been mostly forgotten. It is not always true that the motion in a changing array is “apparent” and not “real.” The important thing is not the apparatus devised for the motion picture but the information it provides for our vision.

THE CHANGING OPTIC ARRAY

Let us recall once again that the arrested optic array is an unusual case of the changing array; it is obtained in a frozen world by an observer who holds still and uses one eye. The eye continues to work, but it is not what the organ evolved for. Optical rest is a special case of optical motion, not the other way around. The eye developed to register change and transformation. The retinal image is seldom an arrested image in life. Accordingly, we ought to treat the motion picture as the basic form of depiction and the painting or photograph as a special form of it. What a strange ideal! It goes counter to all we have been told about optics. But it follows directly from ecological optics. Moviemakers are closer to life than picture makers.

THE PROGRESSIVE PICTURE

Unfortunately, we have no adequate term to describe what I will call the progressive picture as contrasted with the arrested picture. The term motion picture implies that motion has been added to a still picture. Cinematography, or cinema, is no better. The term photoplay is not right. Film sounds neutral, but live television does not use film. What gets depicted is a flow of events. What gets displayed are disturbances of structure in the array, with underlying invariants of structure. These are what the visual system picks up.

The progressive picture displays transformations and magnifications and nullifications and substitutions of structure along with deletions and accretions and slippage of texture. These are the “motions” of the motion picture, as I put it in Chapter 6. They are thoroughly saturated with meaning. They are lawful, even if not described by geometry. They can show people, animals, objects, places, and events with the utmost precision and elaboration. They need to be studied by experimental psychologists on the one hand and by experimental moviemakers on the other. They cannot be usefully studied by taking the parameters of “motion” and plodding through the systematic
variations of the traditional experiment. But on the other hand, neither can they be understood by playing around with aesthetic intuitions.

The progressive picture can also of course depict the movement of the observer himself in the environment as well as the motions of objects: it can arouse visual kinesthesis as well as visual event perception. This fact was pointed out in Chapter 10 and will be further discussed later in this chapter.

The progressive picture can turn into an arrested picture if a "stop-action" shot is inserted in the sequence. This is much used in film and television nowadays. The differences between the two kinds of depiction become clearly evident, along with the similarities between them.

THE ARRESTED PICTURE

If it is true that a drawing, painting, or photograph is actually an arrest of the normally changing array, we shall have to revise our thinking. The arrest has to be artificial, for no event can be stopped in midflight. It is an abrupt noncontinuation of the event, with a continuing nontransformation of the array. The picture is not, as we have supposed, the optic array at an instant, a single moment of time, but an unnatural stopping of the flow. The painter of a quiet landscape, to be sure, arrests only the very slowest of the changes and emphasizes what is invariant in the scene but nevertheless stops action.

We can now understand, I think, why painters stubbornly continue to insist that they can portray "motion" in a still picture. This is one of the paradoxes mentioned in the last chapter. Painters cannot display or represent motion, but they can certainly specify an event. The stopped event may contain the information for perceiving it. The wind in the trees can be depicted if the painter selects the right form in the transformation. So can the smile of a sitter for a portrait. The act of dancing can be conveyed by a photographer because the invariants are different from those of standing or walking. There are event-invariants as well as formless object-invariants.

The arrested picture can specify a progressive event. What the progressive picture can do is to specify it more completely. As for the formless invariants, they are stronger in a changing picture than they are in a changeless picture, but they are still present.

WHAT CAN THE MOVIES MAKE AVAILABLE?

The film, like the photograph and the painting, makes possible not only perceptual awareness but also several kinds of nonperceptual awareness. I refer to the edited film
made with a motion picture camera or its television equivalent. The perception or
imagination is vicarious, an awareness at second hand. Consider the possibilities.

A film can depict situations and problems that you will have to face at a later time,
we call this an educational film. It can depict vistas of distant scenic places to which
you may never go and the connections between vistas, a travel film. It may show events
that happened only yesterday, a news film. It may depict ways of life, histories,
adventures, encounters with wondrous persons, prophetic events, fictions, and fantasies;
we call these documentary films, historical films, adventure films, and wish-
fulfillment films. They are usually full of what their producers call "action." We are
addicted to them, all of us, children and adults. The beholder is apt to identify himself
with a protagonist to whom he feels sympathy, and this means he puts himself at the
point of observation of the protagonist in the way I have described. He then gets
perception, knowledge, imagination, and pleasure at second hand. He even gets re-
warded or punished at second hand. A very intense empathy is aroused in the film
viewer, an awareness of being in the place and situation depicted. But this awareness
is dual. The beholder is helpless to intervene. He can find out nothing for himself. He
feels himself moving and looking around in a certain fashion, attending now to this and
now to that, but at the will of the film maker. He has visual kinesthesis and visual self-
awareness, but it is passive, not active.

To behold a motion picture is thus similar in important ways to observing the
ordinary happenings of life. But it is also radically dissimilar in other ways that are just
as important. Both need to be understood. In the case of the film, one's movements
of approaching to scrutinize or retreating to get a fuller view are controlled by the film
maker. In the case of the real environment, one is free to move as one pleases, that is,
as one "wills." But note that the scanning of fine details in the array sample is free and
unconstrained in both cases. The film maker cannot interfere with your eye movements.
He can control only your head movements and your locomotion.

WHAT DOES A VERBAL NARRATION
MAKE AVAILABLE?

A narration or description can also of course give one the kinds of awareness at second
hand that the film can. And the reader is controlled by the writer as much as the film
viewer is controlled by the film maker. Neither can look for herself, or visualize for
herself, or imagine for herself. She is at the mercy of the artificer, the artist, the maker,
the one-who-shows. But let us not confuse the kind of information that has been put
into words with the kind that has been simply displayed. Film is not a language with
Figure 16.1
Cartoon by Peter Arno. (Drawing by Peter Arno; © 1946, 1974 The New Yorker Magazine, Inc.)
a grammar, as some film makers like to believe. A graphic depiction is not an explicit description and, similarly, a motion picture is not a verbal narration.

**A THEORY OF FILMING AND FILM-EDITING**

The ecological theory of perceiving advanced in this book has implications for filmmaking. Film-viewing, I said, is both similar and dissimilar to natural observing. Let us follow up this suggestion.

**THE COMPOSITION OF A FILM**

A motion picture is composed of virtual events joined together. One kind of junction is obtained by turning the camera from one event to another during a continuous run of film (panning) or rolling the camera stand from one location to another during a run of film (dollying). But another kind of junction is obtained by splicing together strips of film, each being the result of a single run called a *take* or a *shot*. These junctions, the transitions between the events displayed, are crucial for film-making. Events have to be nested in a coherent way if the superordinate events are to be intelligible.

Note that the cameraman or motion picture photographer is the person who moves the camera and changes the lenses, whereas the editor or cutter is supposed to be the one who puts together the shots, using either splices called *cuts* or other optical transitions called *fades*, *wipes*, *dissolves*, and the like that are made by photographic special effects. But both functions ought to be performed by the same person, or at least under the direction of the same person for, if I am right, the cameraman and the editor are doing the same thing. Camera movement and film-splicing are not separate kinds of composition.

**THE CAMERA AND THE HEAD OF THE VIEWER**

The motion picture camera occupies a point of observation in a studio set or on a real location, just as the head of an observer does in an ambient environment. The camera can turn, look up or down, and undergo locomotion, more or less as the head does. The field of view of the camera is analogous to the combined field of view of the eyes in the head in the sense that both fields are bounded by occluding edges, although the visual solid angle sampled by the camera is much smaller than the visual solid angle.
sampled by the head, which is nearly a hemisphere. Note that the light entering an eye and forming a retinal image is emphatically not analogous to the light entering a camera, as photographers assume.

In this analogy, the camera, film, projector, and screen are all components of the same device, a way of providing information to a seated viewer. The field of view of the camera becomes the optic array to the viewer, even when he is not placed so as to get the same-sized angular field of view that the camera got. The screen picture functions as a mobile window hiding most of the environment being filmed, and the edges of the picture can sweep over the ambient array of the environment in the way described in Chapter 12, with gain and loss at the leading and trailing edges of the picture. The seated viewer never actually turns his head, of course, but he gets the essential optical information for doing so. And thus, he becomes aware of a whole new world behind the magic window.

The window can turn sideways quite naturally. This is called panning, on the intuitive belief that the awareness becomes "panoramic" during such a shot. It does in a way, but not because the picture is panoramic. The window can also look up or down. Theoretically it could also tilt, although this is seldom done in practice. The edges of the window can also go forward or backward, or travel sideways, when the camera is moved on a dolly, truck, or crane. The dolly shot is a well-known way of moving close up so that an item of interest fills the window or far back so that a whole array of items is included. The use of a zoom lens that alters the field of view from wide angle to narrow angle or the reverse is now common as a substitute for the dolly shot. It too gives the feeling of approach or retreat, but the dolly shot is preferable when it is possible, for the zoom shot cannot display the deletion or accretion that occurs at occluding edges.

The modes of camera movement that are analogous to the natural movements of the head-body system are, in this theory, a first-order guide to the composing of a film. The moving camera, not just the movement in the picture, is the reason for the empathy that grips us in the cinema. We are onlookers in the situation, to be sure, not participants, but we are in it, we are oriented to it, and we can adopt points of observation within its space. The illusion of participation can be enhanced by having the camera occupy the point of observation of one of the protagonists in the story. This has been done commercially only once to my knowledge, in *The Lady in the Lake*, a Hollywood murder mystery in which Robert Montgomery played the hero but was scarcely visible since his acts of locomotion and exploration, his adventures, and his encounters with the woman and the villain were carried out by the camera. It was the camera that was punched in the face and kissed by the woman. The so-called subjective camera does not deserve the neglect in which it is now held by film producers.

Films for training and education can profit by having the camera occupy the point of observation of the learner. A student can be shown what it is like to land an airplane
or operate the controls of an earthmover or tie knots. But because a theory of visual kinesthesis and control has been lacking (Chapter 13), the method has not been exploited.

**THE PSYCHOLOGY OF FILM-SPLICING**

We have been talking about the filming of natural movements; what about the joining of shots? It seems plausible to me that the various kinds of cutting that a film editor can perform also have analogues in perception and that the insertion of fades, wipes, dissolves, and other special effects is at least an attempt to create transitions with psychological meaning. Composers of film are guided only by their feeling for what works. Some film theorists, as we shall note, try to take lessons from painters, but the lessons to be learned are not clear.

A cut represents a displacement of the camera between shots. The most intelligible cuts, I suggest, are those between shots that have some invariant structure in common. A displacement of the camera forward or backward yields a structure that is magnified or minified, and one sees the same layout afterward as before. Such a displacement is the same as a dolly shot or a zoom, except that it is discontinuous. The familiar sequence—long shot, medium shot, close-up—has a common structure at the center of the picture. A rotary displacement of the camera yields a shot that overlaps its predecessor unless the angle is greater than the camera's field of view. It is thus the same as a pan.

Next, there are cuts in which the viewer is displaced on a circular path around the event being filmed. He sees the lovers, say, from the north where the man's face is in sight and then from the south where the woman's face is in sight. Such differing vantage points, revealing different surfaces, seem to be called "camera angles," but it is not a good term. If underlying invariants are shared, the viewer will perceive the same two persons as before and be aware that he has been instantly transported to another viewpoint, not that the lovers are different persons or that they have been rotated.

Instant transportation of the onlooker can be attempted from one room to another in the same house or from one neighborhood to another in the same countryside—in short, from place to place. Intelligibility depends on whether the viewer has been previously oriented to the environment being portrayed, that is, whether the nesting of places has been established. This can be done with establishing shots, and it can also be done by connecting the major vistas of the environment with dolly shots. Orientation is crucial for comprehension.

Cutting back and forth between distant places, as in scenes of the heroine tied to the railroad tracks and the hero riding to the rescue, ought to suggest that the hero is getting closer, not farther away. The events are concurrent, but at what places? Chase
sequences have a similar problem. There is no overlap of structure between such alternating shots, but there must be some common invariants. What are they?

The split screen provides a way of depicting concurrent events at widely separated places without cutting back and forth. Instant transportation of the onlooker is avoided, but the ecological paradox of being in two places at once is introduced.

Instant transportation in time is attempted by the so-called flashback. Characters who have already been depicted in later events are depicted as involved in earlier events, often in the same place. But the jump in sequence, like the jump from place to place, must be made intelligible. Aristotle had a psychological point when he argued that drama should maintain the "unities" of time and space.

The cut is abrupt. Gradual transitions are possible, such as the fade-out, the fade-in, and the combining of the two in a dissolve that superimposes the structures of the two shots so that a perception of transparency is induced and one layout of surfaces is gradually converted into another layout by way of becoming insubstantial, by passing into "thin air" and out again. Another gradual transition is the wipe, where a line something like an occluding edge (but not optically the same) passes across the screen, concealing one vista and revealing another. The psychological meaning of these transitions has never been studied experimentally by either film editors or perceptionists, but the ecological approach to vision suggests how they might do so. Cuts, fades, dissolves, and wipes are not pure conventions the meaning of which can be arbitrarily decided by film makers and taught to us. The practice of jump-cutting in films and television seems to me ill founded.

THE THEORY OF MONTAGE

A quite different theory of the nature of filmic transitions seems to be widely accepted by directors and critics. This theory is based on the assumption that any juxtaposition of shots, however disparate, will form a unified "image" with a new meaning. The combination is more than the sum of the parts. The doctrine is identified with a book translated as The Film Sense (Eisenstein, 1942). The author, a famous Russian director, is celebrated for the boldness with which he combined shots of events that did not ordinarily occur together.

Montage, in this sense of the term, is related to collage. The latter was invented by painters who tried composing a work of art by pasting items on the canvas instead of painting forms on it. The associating of scraps, pieces, pictures, or forms not previously associated was thought to yield a fresh insight, or an unexpected gestalt. The word collage means a paste-up. The creation could be photographed and displayed. Similarly, strips of film could be spliced together. "The juxtaposition of two separate shots resembles not so much a simple sum of one shot plus another shot as it does a
DEPICTION BY FILM

If the foregoing approach is correct, there is such a thing as filmic depiction that is distinct from ordinary depiction. Its aim is to produce in the viewer the awareness of a train of events, and of the causal structure of these events. They are virtual instead of real events, to be sure, and no one is ever wholly deceived, as when having a hallucination, but the feeling of being present in the world behind the magic window is very strong.

This awareness of events is achieved by segmenting the flow of the pictorial optic array so that it specifies the same kinds of subordinate and superordinate happenings that are specified in a natural optic array. Persons, animals, places, objects, and substances are depicted along with the events. The segments of the optical flow are crucial, that is, the transients between parts as well as the parts themselves. Simply to call them “motions” is not to do justice to them.

Filmic depiction shares with verbal narration, storytelling, the capability of showing what happens if so-and-so happens, the predictable causal sequences of the world, along with the accidental happenings, the unpredictable sequences. But it shares with ordinary depiction, perspective pictures, the capability of putting the observer into the scene.

MOTION PICTURES AND VISUAL AWARENESS

301
SUMMARY

What we call the motion picture as distinguished from the still picture might better be called the progressive picture as distinguished from the arrested picture. It is not characterized by "motion" so much as by change of structure in the optic array. And the ordinary picture is not so much "still" as it is stopped.

The progressive picture yields something closer to natural visual perception than does the arrested picture. The nameless transformations that constitute it and that are so hard to describe are actually easier to perceive than the familiar frozen forms of the painting or photograph.

It provides a changing optic array of limited scope to a point of observation in front of the picture, an array that makes information available to a viewer at the point of observation. This delimited array is analogous to the temporary field of view of a human observer in a natural environment surrounding the observer.

The information in the display can specify the turning of one's head, the act of approaching or withdrawing, and the adopting of a new point of observation, although one is all the time aware of holding still and looking at a screen from a fixed position in a room. This is over and above the information in the display for an awareness of events and the places at which the events are happening, along with an awareness of the objects, persons, or creatures of the imagination to which the events are happening. The invariants to specify the places, objects, and persons emerge more clearly in the transforming array than they would in a frozen array.

The art of film-editing should be guided by knowledge of how events and the progress of events are naturally perceived. The composing of a film is not analogous to the composing of a painting. The sequential nesting of subordinate events into superordinate events is crucial. The transitions should be psychologically meaningful, and the sequential order of happenings should be intelligible. But the picture theory of vision and the stimulus sequence theory of perception are very poor guides to moviemaking. The theory of ecological perception, of perception while moving around and looking around the environment, is better. The various kinds of filmic transition—zoom, dolly, pan, cut, fade, wipe, dissolve, and split-screen shot—could usefully be evaluated in the light of ecological optics instead of the snapshot optics that is currently accepted.