

adapted from

The Rhetoric of Film

by

John Harrington

first published by Holt, Rinehart & Winston

© John Harrington

Chapter 1: BASIC PROPERTIES

By the time you reached fourteen, you witnessed 18,000 murders on the screen, along with over 350,000 commercials. By eighteen, you stockpiled nearly 17,000 hours of viewing experience. Even now, if you're a typical American, you watch at least 30 movies for every book you read. Eventually, the viewing experience will absorb over ten years of your life. What follows is an effort to provide some words--tools, really--to help you understand the way images affect you while you watch a film.

The Frame

A **frame** is a single photographic image printed on a length of film. A viewer can see a single frame only under certain artificial conditions: when a projector is stopped at "still" position, when a frame is excerpted and projected as a slide or printed on photographic paper, or when a freeze-frame appears on the screen. Like a single letter in a word, a frame is not a part of a viewer's perceptions until it is isolated. Even then, it seldom has meaning.

In normal usage a *frame* differs from a *still*. A still is a photograph taken with a still (versus motion) camera and printed on photographic paper. Most pictures displayed outside theaters or appearing in newspaper ads or magazine articles are taken with still cameras on a movie's set and are made to be photographs standing alone, even though an almost identical frame might appear in a film.

A **freeze-frame** is a single frame repeated for an extended time, making the image appear like a still photograph within a moving picture. It is produced mechanically in a laboratory by printing the same frame over and over until the image on the screen resembles a projected slide. Actually, the viewer does not see a single frame during a freeze-frame, but a repetition of the same picture even though the eye cannot detect any difference between a projected frame and a freeze-frame.

Although a single photographic frame cannot be discerned during actual viewing, it contributes to a larger unit and is understood in terms of that unit. During normal projection, twenty-four frames per second (approximately a foot and a half of 35mm film) pass through the projector's gate. Each image flashes on the screen, then the screen turns black and is followed by another frame. However, the human eye misses the period of blackout since the eye retains an image one-tenth of a second longer than the image exists. It is this physiological phenomenon (called **persistence of vision**) that allows motion pictures to be seen in continuous movement with no apparent jumps or single frames visible. (Take two frames out of a shot, however, and the eye can often detect a jump.) The average feature contains close to 130,000 separate frames.

The word "frame" also has another meaning in the filmmaker's jargon. The frame is the outer boundary of a projected image--the lines of the rectangle on the screen where an image ends and blackness begins. Because the frame serves as the boundary of an image, it is the starting point in the filmmaker's composition.

The **aspect ratio** of the frame is the relationship of width to height of the image you see on the screen. For silent films, the ratio was 1.33:1 (or 4:3, i.e., four measures wide to each measure high). That ratio became known as the standard "Academy aperture" (from standards set by the Academy of Motion Picture Arts and Sciences--the same folks who put on the Oscars). Experiments in the 1950s in response to the rise of television led to the development of CinemaScope with a 2.35:1 ratio. The standard ratio for theatrical films today is 1.85:1 (1.66:1 in Europe).

Composition is the arrangement of the elements within the frame--not only characters and props, but also patterns of lighting. Composition directs the eye by lines and masses as they are related to other lines and masses. Like the composition of a

painting, the composition within a frame can be viewed in terms of shape, balance, lines, texture, rhythm, color, and chiaroscuro (use of light and dark in an image). Lines provide a sense of direction. Vertical lines seem to soar, and the eye expectantly follows their uplifting energy. Horizontals imply stability and solidity. Diagonals usually convey feelings of confusion, division, or incompleteness. Mass (an occupied area, often shadowed) conveys heaviness or blocks out parts of the frame. Mass is often found in patterns, as in classical paintings. For example, balanced composition is evenly weighted on each side (hence it's often static), while geometric composition often uses circular or triangular patterns.

By controlling the way the viewer perceives an image within the frame, composition contributes both to meaning and to mood. Cinematic composition differs from that of painting and still photography in its direct ties to motion. Composition is almost always dynamic in film because either the camera or the subject changes position. Only by stopping a film can a viewer consider composition; usually, composition makes its point and is quickly replaced by an image with yet a different composition.

The frame has definite and fixed limits, and filmmakers can do little to change its shape. Still photographers, by contrast, can escape the fixed frame of their instrument and make a picture in almost any shape. If they want to emphasize the soaring quality of a Gothic cathedral, they will make a picture much taller than it is wide by cropping (cutting away) unwanted material. For all practical purposes, filmmakers cannot crop but must add as many secondary movements as possible to a horizontal rectangle. Of course, a wide screen helps establish the horizontal lines needed for panorama (the impression of an almost unlimited view in all directions along the horizon), thus giving a filmmaker some control over horizontal composition. But the wide screen has trouble filling in its left and right edges when not taking in vast horizontal expanses of territory.

One fortunate aspect of film composition is that a filmmaker can, by implication, extend the range of the frame. A frame of the bottom half of the Eiffel Tower seems to keep rising, and the viewer senses upward motion. An upward sweep of the camera further enhances a sense of rising movement. The feeling that "there is more there" has direct counterparts in prose. A **synecdoche**, for example, is a figure of speech that uses a part to

signify the whole ("ten hands" means ten workmen; a "set of wheels" means a car; and so on). The viewer of a film need only see a periscope to recognize "submarine" or a hand to be aware of a "person." The whole need not be seen to be realized. This reliance on parts allows a filmmaker psychologically to extend the frame. Viewers sense great lateral movement when a hand and arm cross the screen because they are aware that the hand and arm are attached to a person. Through viewers' ability to imagine--to read images not in a literal and dissected way--a film can transcend its physical limits.

Unlike the human eye, the camera sees indiscriminately. Everything in its view will be recorded. Consequently, the relationships between **foreground** (in front where the main action occurs) and **background** (behind the main action) are extremely important. Depending on how a filmmaker places objects, they can either divert or attract attention. If a man's face, for example, fills the right one-third of a screen in an action scene, the face either blocks out part of the action to lead the eye to what will happen, or the face provides a response to what is happening, becoming central for an instant. If the former is intended, the face will be either slightly or completely out of focus; if the latter is intended, the face will be completely in focus, while other parts of the action will be out of focus or in focus less sharply than the face. The director consciously leads the viewers' eyes to where they will perceive the most significant details.

Since film depends on motion, the movement of people and objects within the frame is as important an element of composition as the way people and objects line up originally. Unlike still photography or painting, composition within the frame of a film changes with each movement. Each change affects the viewer's perception of an image. Since film has boundaries on four sides but none in depth except that determined by a filmmaker, most effective movement occurs at least in part toward and away from the camera. Vertical and horizontal movements of characters remind the viewer of the limits of film's dimensions.

Movement toward and away from the eye, however, creates a sense of depth and of the viewer's involvement in what occurs. Movement toward the camera generally intensifies action and creates a sense that something significant is about to happen. Movement away from the camera appears wistful and serious. Cross-

screen movement always seems much more rapid than depth movement, creating a feeling of excitement, but frequently at the expense of a sense of direction and purpose. Depth movement, on the other hand, normally displays significant content, while cross-screen movement creates flurry and action.

The Shot

At a normal projection speed of twenty-four frames per second, a large number of frames make up the basic unit of film: the **shot**. A **shot** is a single uninterrupted action of a camera.¹ Some shots last only one or two frames, although such short shots appear rarely in commercial films. But anyone who has seen experimental films (such as Charles Braverman's *An American Time Capsule* or *The World of '68*) knows how rapidly shots can operate and how many shots the eye will accept in a small amount of time. Although longer shots are "standard," few last over thirty seconds. The exceptions, of course, run for as long as a filmmaker chooses to keep film running through his camera. The average shot runs from about two to thirty seconds.

Because it is the smallest functional unit of film and combines to form a larger statement, the shot syntactically parallels the word of spoken and written communication. The frame, on the other hand, resembles the single phoneme or letter of a word. Shots make up the vocabulary that film's visual grammar and syntax connect into statements with meaning. The vocabulary of film is primarily the vocabulary of a series of photographic images. But consider for a minute the notion "shot" in relation to the notion "word" in order to grasp the syntactical workings of the basic unit of cinematic composition.

¹A **take** is also a single uninterrupted action of a camera, but a take is the unedited footage and is seen from the point of view of the filmmaker rather than that of the viewer. A take will frequently be shortened at both ends, and perhaps another shot or two will be cut into the middle, creating three, four, five, or more shots from a single take. For instance, during an interview, two cameras might be trained on the two persons talking. Later, an editor will cut and splice in order to alternate between the two speakers, creating many separate shots from only two takes.

The shots of a film draw meaning from their context much as words derive significance almost exclusively from their linguistic context. When isolated, the meaning of either a word or a shot is imprecise. Consider the word "stand." Is it a verb (such as a command to assume a certain physical position, or a description of what someone is doing or did do), or is it a noun (such as an ideological position one takes, a structure to sit on, a courtroom place of witness, or a group of trees)? Without a context, one cannot ascertain meaning or function. Similarly, a single shot has meaning, but without a context a particular meaning is difficult to identify.

While analogies can be drawn between shot and word, the shot also resembles the written paragraph. A paragraph normally articulates an idea then offers supportive evidence or arguments. Similarly, a shot in context assumes a general idea or mood and also offers many equivalents of simple declarative and descriptive sentences, providing a viewer with supportive information. Imagine the elements of a hypothetical shot put into statement form: The woman sits in the kitchen. The baby is in the highchair. The baby is crying. The woman is holding baby food. The wall is yellow. On the wall stands a picture of a horse. There is a table in the foreground. The table is round. The table is dark. There are four chairs around the table. All this, and far more, a viewer perceives while watching a shot. A shot, like a paragraph, offers both detailed information and an idea or mood.

Any direct analogy between the shot and the paragraph, however, quickly breaks down. The elements of a paragraph are met with one at a time. They are linear. The content of a shot is, for all practical purposes, available all at once. (Some shots, though, do reveal some components of content linearly; for example, a moving camera presents different pieces of information in a defined order.) Ideas and details are not easily separated. Abstract ideas are seldom stated as such in film--and then usually in documentaries. Film argues almost entirely by evidence, inexorably forcing a viewer to supply appropriate abstract ideas. We are not told, for instance, that Mr. Jones loves his wife. We see him love her. Film is a visual medium, and expresses concepts visually.

Shot Length

Shots are categorized according to the apparent closeness of the camera to the person or object photographed. With the early single-focal-length lenses, distance literally became the factor determining the "length" of a shot. With the present variety of lenses, only the illusion of distance counts. If an object or person seems very far away, the result is normally called an **extreme long shot** (ELS), also called an **establishing shot** because it places objects in context and prepares a viewer for a closer look later. If only a small part of a person or object appears, the shot is called an **extreme close-up** (ECU). In between lie the **long shot** (LS), **medium long shot** (MLS), **medium shot** (MS), **medium close-up** (MCU), and **close-up** (CU).

The distinctions among shots by distance are relative, and no precise lines or measurements separate the various shots. Usually, the human figure provides the chief standard for measurement. In an **extreme long shot**, a person might be visible, but the setting clearly dominates. The same person fills a good part of the vertical line of the frame in a **long shot**, although the setting also receives strong emphasis. A **medium long shot** reveals about three-fourths of the subject, while a **medium shot** (also called a mid-shot) would show the subject only from the waist up, focusing a viewer's attention more on the subject than on the setting but maintaining a clear relationship between the two. A **medium close-up** shows a person from the shoulders up, and a **close-up** shows only the head. An **extreme close-up** reveals only a small part of the face, such as the nose or an eye.

What is a close-up in one situation, however, can be a mid- or long shot in another. The length of a shot depends on the subject of a film (or of an individual scene). The longer a shot is, the more it shows of the subject; the shorter a shot is, the more it emphasizes detail that is part of the subject. Hence, the length of a shot is relative and depends upon what the filmmaker has chosen for a subject. If, for example, a film is about cities of the world, a shot of a red double-decker bus in London would be a close-up. The same shot would be a long shot in a film about London's buses, while a shot of an instrument panel would be a close-up.

Shot Movement

The sense of the frame moving also contributes to the effect of a shot. Not only does the frame seem to move from shot to shot because of the mix among close-ups, mid-shots, and long shots, but the frame also appears to move during a shot.

A **pan** is used either to follow a horizontal action or to sweep across a stationary scene. A pan occurs when a stationary camera turns horizontally and the viewer sees new areas revealed in movement from right to left or left to right. A pan can also direct the audience's attention to an important action or convey a subjective sense of sharing the gaze of a person whose eyes are moving.

A **tilt** shot occurs when a stationary camera angles up or down, providing new views above or below the camera's level. It often lets viewers feel as if they are following a subject or looking through a character's eyes. The direction the camera moves is called "tilt up" or "tilt down."

A **dolly** is a platform on wheels serving as a camera mount capable of movement in any direction (forward, backward, circular, sideways). During a **dolly-in**, the camera moves toward the subject; for a **dolly-out**, the camera travels away from the subject. The dolly provides a sense of dynamic movement, with the same kinds of changes in perspective the human eye sees when a person moves.

A **tracking shot** includes any shot with a camera moving on tracks. Since a viewer can't tell whether a shot is a dolly shot or a tracking shot, the terms are used almost interchangeably to mean "moving shot." (Tracks allow more precise control over a camera's movement.)

A **zoom** shot is accomplished with a lens capable of smoothly and continuously changing focal lengths from wide-angle to telephoto. As the lens **zooms in**, the center of composition gradually fills the screen; the center diminishes as the lens **zooms out**.

While dolly (and tracking) shots and zoom shots resemble each other in their ability to move toward and away from a subject, they differ in effect. Once focus is set on a zoom lens, it need

not be reset as the focal length is increased or decreased since the camera-to-subject distance is constant. A dolly shot, on the other hand, changes the relationship between camera and subject, requiring a change in focus and also bringing about a change in the relationship between subject and background. The effect is the difference between walking closer and closer to someone versus looking at her from a distance through increasingly powerful field glasses. Through the glasses, only the size of the subject changes, and that change occurs at the same rate that the background increases or decreases in size. The dolly, on the other hand, provides changing perspectives and consequently a strong sense of movement through space. Usually a dolly involves a relatively small change in the size of the image, while a zoom moves freely from very close to very far away.

A **crane shot** is taken from a boom that can move vertically and horizontally at the same time. A crane is a large camera trolley with a moveable arm (or boom) with a platform for the camera. The crane can move like a dolly, but the boom also allows movement both vertically and horizontally. A crane shot often provides a dynamic sense of movement through space.

The Scene

A filmmaker puts shots together to make up a **scene**. While a shot comes from a single operation of the camera, a **scene** is a series of shots with a continuous action. A viewer perceives a series of shots as taken at the same location during a rather brief period of time. The classic car chase furnishes a good example of a scene. Some action prompts a chase, usually shot from a number of shot lengths and perspectives. Finally, one person catches the other, usually after one or more crashes, and the scene ends. Because the action of a chase is fast, a scene that lasts three minutes or less might have 50-100 separate shots. The time seems continuous (although the number of shots shown may cause the film version to last much longer than the actual time to cover a given distance), and the location is the same even though changing because the area of the chase gets treated as a single continuous area of travel. Action then resumes at another time or location.

Film scenes vary greatly in length. Sometimes a scene will be only a single shot long; in other cases a whole movie will have

but one scene. Usually, however, scenes last for several minutes.

The Sequence

A number of shots and scenes make up a **sequence**, forming a coherent dramatic action. A **sequence** is usually composed of a series of scenes that are related in location, time, generating action, point of view, or cast. Sequences are consequently much like chapters in a book, with beginnings, middles, and ends. At one time, sequences were clearly set off by strong "punctuation marks" such as the fade-out and fade-in. These strong forms of punctuation alerted a viewer that a major segment of a film had ended and that a new one was about to begin. Contemporary filmmakers, however, have abandoned such obvious punctuation marks, relying instead on straight cuts to subtly connect sequences.

Frame, shot, scene, and sequence are the most basic terms in the lexicon of film. They suggest the fundamental structures with which film operates, and they indicate the complex "cut-and-paste" nature of the medium. The rhetoric of any film grows from the way a filmmaker manipulates these basic structural units.