June 2016

What Makes a Movie

Richard L. Anderson
Visual Thinking Laboratory, University of North Texas, rich.anderson@gmail.com

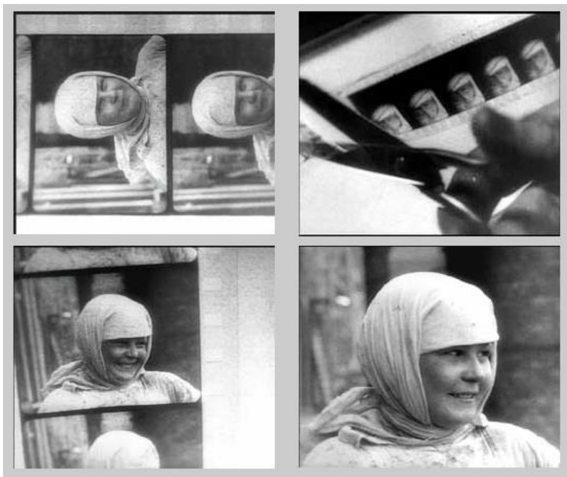
Brian C. O'Connor
Visual Thinking Laboratory, College of Information, University of North Texas, brian.oconnor@unt.edu

Movies do not move. Essentially all movie formats are made up of still images displayed rapidly. Each of the 16mm frames to the left is about the size of a fingernail. In projection, a frame is held motionless, a shutter opens and allows light to pass through and project an image onto a screen, the shutter closes, another frame is pulled into place, the shutter opens, … 24 times per second. The process of intermittent motion was the invention of the Lumiére brothers in 1895.

Electronic analog and digital formats, while they do not present still images observable by the naked eye, store data in single frame packets. The frame has been the addressable unit of the movie since the earliest of days. The frame is a still photograph, so a movie can be said to be a collection of still photographs.

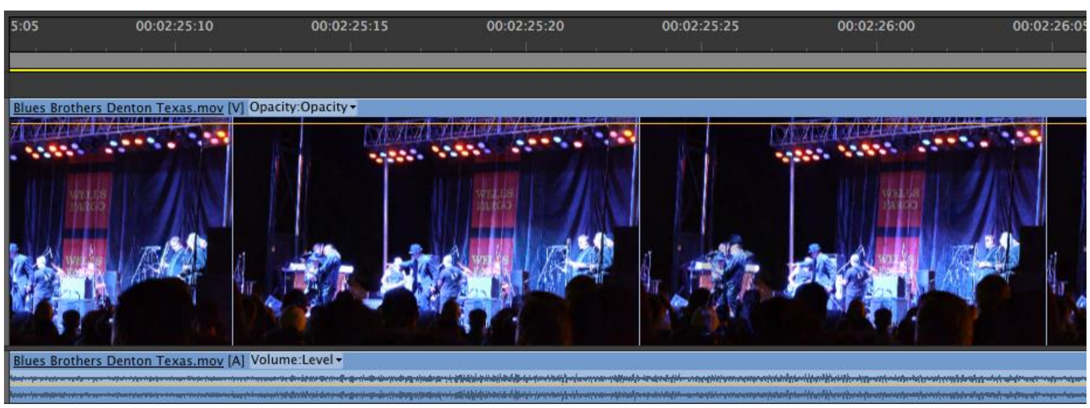
What makes a movie is something more than viewing a collection of still images.

The frame has been the fundamental unit of production of movies, enabling control of the viewing experience down to the fraction of a second. Johnson notes:



Montage—juxtaposing images by editing—is unique to film (and now video). During the 1920s, the pioneering Russian film directors and theorists Sergei Eisenstein and Dziga Vertov demonstrated the technical, aesthetic, and ideological potentials of montage. The 'new media' theorist Lev Manovich has pointed out how much these experiments of the 1920s underlie the aesthetics of contemporary video. Eisenstein believed that film montage could create ideas or have an impact beyond the individual images. Two or more images edited together create a "tertium quid" (third thing) that makes the whole greater than the sum of its individual parts.

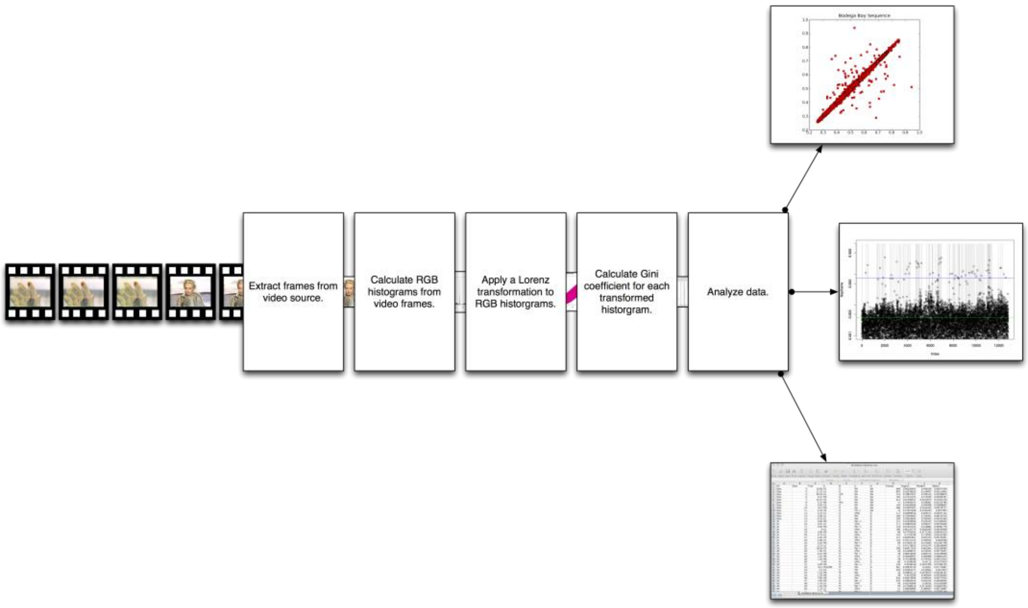
Eisenstein and Vertov (above) and most editors working in analog film made mechanical cuts at the frame lines; digital editors (below) work with pixels and timelines, but still cut at the frame level. The frame serves as a robust means of sampling the movie data stream and an explanation of what is a movie.



For some time we have been examining ways to describe filmic documents in unambiguous ways, to describe the structure of a movie, to compare structures of movies, and to engineer a robust model of moving image documents. We had made significant progress toward these goals combining the idea of seeing moving image documents as signal sets together with what might broadly be called a behavioral component. This behavioral component consisted in the well-established semiotic literature, particularly Metz, Bellour, and Augst; and the theories and practices of behavior analysis.

Our first step was to step away from the debates and failures inherent in seeing the "shot" as the unit of analysis. As Bonitzer notes, the definition of "shot" is: "endlessly bifurcated," essentially rendering the shot useless as a unit of analysis. We used changes in the Red, Green, and Blue components of every pixel in every frame of a film sequence to find points of discontinuity in a film. By itself, this approach is interesting but does not provide any particular way to find significant points of discontinuity. Bellour had wrestled for some time with the notions of how films generate meaning; he, too, looked to significant points of discontinuity in the signal set. In his work on the Bodega Bay sequence from Hitchcock's ***The Birds*** he used his highly regarded critical expertise to determine the significant points of discontinuity.

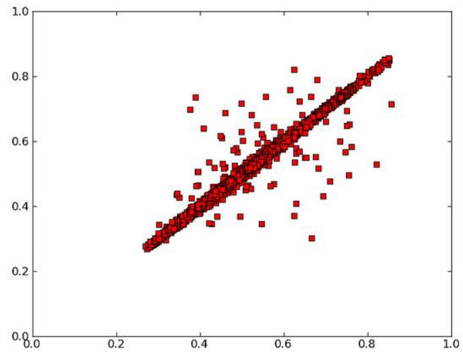
We used Bellour's approach to develop a computational heuristic for description of any film -we assumed he was engaging a signal set and characteristics of the signal made it possible for him/necessary for him to see points of discontinuity. Our efforts replicated Bellour's work very well and we validated the Bellourian heuristic with our analysis of Looney Tunes films by two different directors. The work with our heuristic met with enthusiasm from film theorists and documentalists (e.g. Buckland in Document (Re)turn: *Anderson, O'Connor and Kearns provide a striking example of combining radically different qualitative and quantitative analytical methods in their discussion of the [Bodega Bay] sequence of Hitchcock's* ***The Birds***. p. 319)



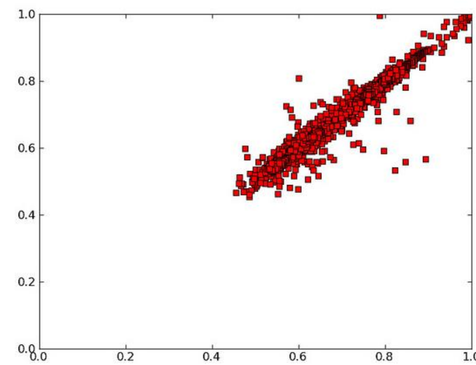
Still, a heuristic is of only limited value for defining “moving image document” and describing films in a manner useful for classification. Our current challenge is to engage more films and push beyond a heuristic. We currently have RGB signal data for the frames of 60 filmic documents – Hollywood titles, experimental of various sorts, TREC (Text Retrieval Conference) test documents, animations, TV shows, etc.

Briefly, we use the same sort of signal data acquisition as in our previous work, we simply use a different form of analysis. We derived RGB values for each frame (1800 frames per minute); posited an even distribution (as per Gini analysis); derived the area between the RGB histogram and the line of even distribution; for each and every pair of frames we subtracted the derived area for frame *n* from the derived area for frame *n+1*. Plotting the differences yielded a graphical representation of structure, particularly points of discontinuity.

A seemingly simple shift of perspective provides another way to look at the frame-to-frame change. If we plot the same data on a Cartesian plane with value for frame *n* as the X-coordinate and the value for frame *n+1* as the Y-coordinate, we have a system in which the unit of analysis is the CHANGE – this depends on the pixel level data stream (actually sub-pixel as R, G, B.)



Digraph of Birds

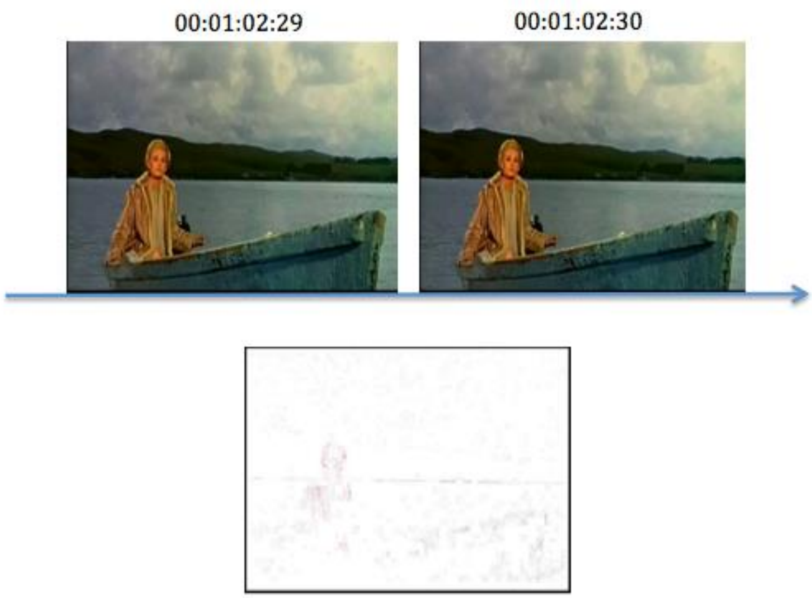


Digraph of Hyde and Go Tweet

Presenting our data in this digraphic way allows us to see a structural pattern within an entire film. The greater the deviation of a plotted point for any frame pair from the norm, the greater the probability that pair bounds a point of significant discontinuity. In examining data with digraph we see the same frame pairs data as in our previous method, but we see them more obviously. Also, we now have the means of constructing a formula for what constitutes a movie – most frames would have to lie along the line, some would have to lie off the line. The art and craft of movie making, and a way of characterizing filmic structure, lies in how many lie off the line and by how much.

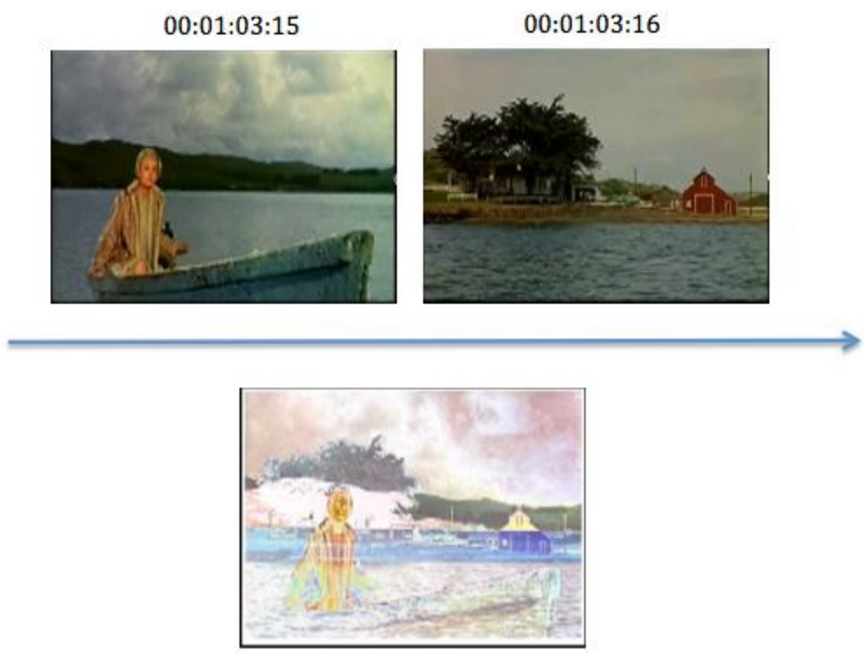
Significance of points of discontinuity can be presented and examined in two ways. With Bellour we have significance defined by a recognized expert in his expert subjective viewing. With empirical data derived from RGB values and shown to be consistent with Bellour's expert notion of consistency, we can define significance (on the whole and with some intriguing exceptions) to be any plotted point of change falling outside one standard deviation. With diagraphic presentation of RGB data and a much larger set of filmic documents, we have gone from heuristic to the algorithmic. We can take this same data and present it in a rather different form – *synthetic frames*. It is not too facile to say that each plotted dot in the digraph is roughly equivalent to a synthetic frame.

The data for just those pixels that are different between frame N_1 and frame N_2 can be used to generate a viewable image that is neither of the two frames nor is it made up of some regions of one and some regions of the other; in other words, it is synthetic. In most movies there are periods where most of the frames are similar, though not exactly alike; then there is some significant change. In our frames from *The Birds* we see Melanie in a boat for several



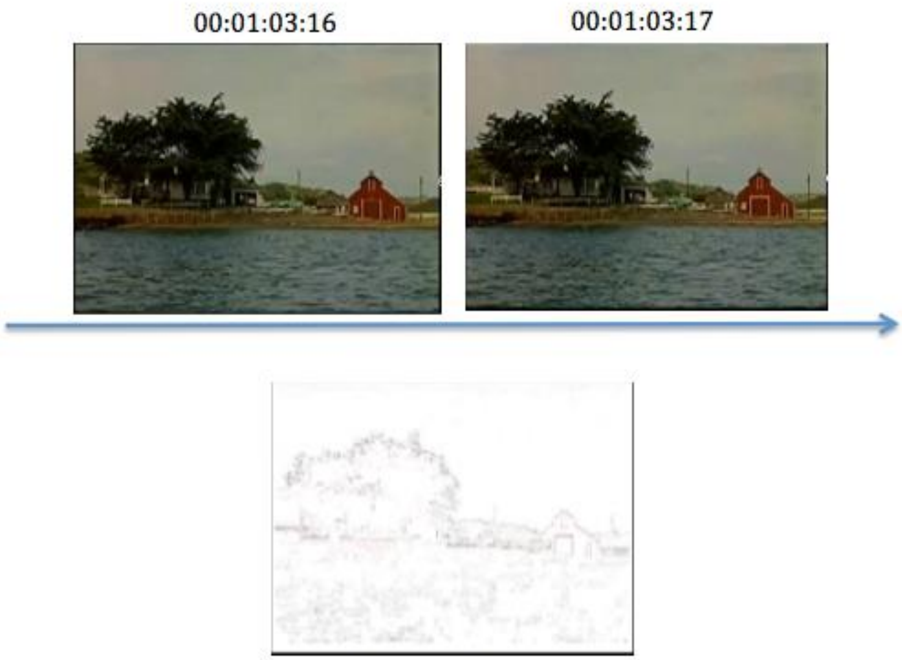
Synthetic frame

seconds, then we see the farmhouse she is approaching, then we see her in the boat again. In the theatrical release of the *The Birds* there were 24 frames for each second of viewing time, so in a sequence of four seconds length we would see 96 frames of Melanie in the boat. Not much changes from frame to frame, but there are some changes from frame to frame; the boat is in slightly choppy water, so the woman and the boat have slightly different distances from the frame edges. These small differences yield what almost looks like a pencil sketch of just the major outlines, since the watercolor remains the same, the boat color remains the same, the hair color remains the same, and the coat color remains the same – they just shift a bit from frame to frame. Timing is in standard format of hours: minutes: seconds: frames.



Synthetic frame

When we reach the point of change from Melanie in the boat to the farmhouse – frame X_{last} (00:01:03:15) and Y_{first} (00:01:03:16), as one might expect, there are many more points of difference so the synthetic frame shows many more points than the sketched outline. Then, once we are at the difference between frame Y_{first} (00:01:03:16) and Y_{second} (00:01:03:17) the synthetic frame is made up of only a few points of difference; though the camera has the point of view of the woman in the boat and the boat moves, so there are small shifts from frame to frame.



Synthetic frame

What is it then that distinguishes a movie from a static still photograph or a set of static still photographs, as in a slideshow? The narrow constraints that provide the viewer of the document the illusion of motion and a sense of narrative in the broadest sense make the distinction. There is a narrow window of entropy necessary for maintaining the illusion of motion; too much entropy and the document loses coherence, while too little entropy and the document no longer engages the viewer.

We need a little more though. The illusion of motion is normally brought about by the slight changes in data from frame to frame when played back at the intended or nominal speed of the medium. A viewer of a collection of random photographs could arrange a set of prints or digital files and allot a set time period for viewing each image and an order in which they would be viewed, but this would not necessarily present any perception of motion, nor would it necessarily be considered a representation of motion. It would be, essentially, a slide show; it might have thematic coherence, yet would not be a moving image document.

Here we might turn to a recent development in video to find a transitional case – the Ken Burns effect. Documentarian Ken Burns developed a technique by which “Action is given to still photographs by slowly zooming in on subjects of interest and panning from one subject to another.” The illusion of motion is generated by moving the camera (or software version of a camera) over the image, thus producing a set of frames that have the sort of difference between any two consecutive frames we discussed above. The image on the screen, the stimulus set to the eyes of a viewer, is changing at a standard rate; the illusion of motion though is motion of the still photograph rather than of the objects in front of the original camera. Here a sample of frames from two seconds of panning to the left across an image of a city street.

This is not necessarily a cheat in terms of message making or story telling and the effect does depend on the same persistence of vision that seems to account for what would normally be called a movie, yet there is no illusion of motion in the ordinary sense of some objects moving against a static backdrop and with regard to one another. We are speaking here of message making, of a filmmaker coding a message; as Hayes suggests, the filmmaker dances with the viewer, making assumptions about the viewer’s decoding abilities. Persistence of vision sets limits on coding practices; it frames the rate of change in the visual data stream at playback. Too little change from frame to frame and the viewer perceives no motion; too much change from frame to frame and the ability to merge the data is lost.

Any single pixel address within a frame is comprised of four values: Red, Green, Blue, and Opacity – RGBA or RGBα. For any pair of frames two additional values are added to the pixel address data: directionality and magnitude. These form a vector describing the amount of change over time; in a movie this time period is now ordinarily 1/30th of a second.

So what? Movies present movement. In order to analyze movies to understand how they are coded to generate meaning and, at the same time, to develop methods of categorizing movies based on their coding structures – what might be called fingerprinting – we need to be able to describe movement in rigorous terms. We need to be able to describe and compare sorts of motion without losing sight of the motion.

References

Anderson, Richard L., Brian C. O'Connor, and Jodi L. Kearns, "The Functional Ontology of Filmic documents" A Document (Re)turn: Contributions from a Research Field in Transition. Ed. Roswithe Skare, Niels Windfeld Lund, and Andreas Varheim. Frankfurt am Main, Germany: Peter Lanf GmbH, 2007, 345-363.

Anderson, Richard L., Brian C. O'Connor, Melody J. McCotter. Outside the Frame: Modeling Discontinuities in Video Stimulus Streams. iConference 2010 Proceedings, p. 508

Bateson, Gregory. Mind and Nature: A Necessary Unity. New York: E.P. Dutton, 1979.

Buckland, Michael K. "Northern Lights: Fresh Insights into Enduring Concerns." A Document (Re)turn: Contributions from a Research Field in Transition. Ed. Roswithe Skare, Niels Windfeld Lund, and Andreas Varheim. Frankfurt am Main, Germany: Peter Lanf GmbH, 2007, 319.

Hayes, Robert M. Measurement of Information, *Information Processing and Management*, v29 n1 p1-11 Jan-Feb 1993

Zacks, Jeffrey M. Flicker: Your Brain on Movies. New York : Oxford University Press, 2015.

DOI: 10.35492/docam/3/1/3