where $\overline{r}_a$ and $\overline{r}_u$ are the average ratings for user $A$ and $u$ on all other movies, and $s_{ad}$ is the similarity between user $A$ and $u$. The summations are over all users $u \in U$ who have rated movie $M$. For item-based CF, the direct weighted average can be used (no differences in individual user scales need be taken into account).

$$P_{aM} = \frac{\sum_{i \in I} (r_{iM}) * s_{ai}}{\sum_{i \in I} |s_{ai}|}.$$  \hspace{1cm} (4)

The summation $i \in I$ is over all items (movies) $i$ that have been rated by active user $A$.

5

FILM THROUGH THE HUMAN VISUAL SYSTEM

FINDING PATTERNS AND LIMITS

Jordan E. DeLong, Kaitlin L. Brunick, and James E. Cutting

You would be hard-pressed to find someone who does not watch film. Cultures around the globe have embraced the art of the moving image and run with it, creating so many movies that no one person can hope to watch even a majority of them in his or her lifetime. Cinema has become such a fixture in our lives that the average American watches five films in theaters every year, as shown in Figure 5.1. Cinema's prominent place in society makes it easy to forget that film (in a form we would recognize) has only existed for roughly 100 years. Film has progressed from a technical curiosity to a large-scale form of entertainment that engages viewers from all walks and stages of life. Filmmakers have constantly changed and updated their craft, using trial and error to map out some of the "rules" needed to interface film effectively with the human mind. Several of these rules include matching action, eye gaze, and spatial layout between shots. Determining the bounds of what makes sense to viewers was only the beginning: knowing how to transition effectively between shots is a complex process under constant revision by a community of skilled filmmakers.

Although some people might describe today's films as "uniform" and "formulaic," films continue to evolve. This long-ranging and systematic reimagining of films can afford us insight into elements of the human visual system. In other words, movies have the potential to
within the loose confederation of the cognitive sciences, such as psychology, philosophy, computer science, and linguistics.

It is important to note that this chapter is written from a perspective even more radical than most cognitive film theorists would adopt. As researchers with a cognitive psychological viewpoint, we see film as a stimulus with a number of fascinating properties, many of which have not been examined quantitatively. The types of analysis presented in this chapter are agnostic to the types of interpretation found in most film studies; the data produced by our analysis are largely quantitative. Our methodology conspicuously ignores aspects of film like character development, setting, design, critical review, cultural relevance, director's intent, and most aspects of cinematography. Our data do not describe whether a single film is "good" or "bad," but instead track a number of low-level, slow-changing statistics of popular films. Other film researchers have also been interested in this type of data, extracting comparison statistics from films released throughout the last century.

For our sample, we chose films from every 5 years, starting in 1935 and ending in 2005. The films were selected based on a number of criteria such as box office gross, coarse genre type, and viewer rating in the Internet Movie Database (IMDB). Digital versions of these films were converted into a series of 256 x 256 grayscale images. This collection of movies makes up the dataset that we use throughout the different types of analysis in this chapter. A selection of films included in our database includes The 39 Steps (1935), Back to the Future (1985), and Star Wars Episode IV: A New Hope (1977). A complete list of the films can be found in the supplementary materials of Cutting, DeLong, and Noyelles (2010).

Our first analysis involved finding the boundaries between shots in the visual sequence. In film, a "shot" is continuous footage from the same camera. Shots are then pieced together using a number of different transitions, such as the straight cut (the vast majority of modern transitions), dissolve, fade, and wipe.

Detecting transitions between shots may appear to be a trivial task, but editors do their best to "hide" these discontinuities; in particular, some jump cuts (cuts that bind two shots with little perspective change) are regularly missed by human observers (Smith & Henderson, 2008). In addition, the rules for continuity editing have become so commonplace in popular film that viewers regularly miss cuts that follow these continuity rules (Smith & Henderson, 2008). Although many computer algorithms are somewhat adept at detecting straight cuts, slowly changing dissolves are difficult for them to detect. To raise accuracy in our analyses, human observers also viewed the films to supplement the results of our computerized analysis. After this process, we were left with a series of precise lengths for every shot within in the 150-film sample.
Changing Shot Lengths

The most popular type of quantitative film data to examine is average shot length (ASL), a metric of how long a shot is on-screen before transitioning to a new shot. David Bordwell noted that ASLs have been getting shorter than those during the "studio era" of Hollywood (Bordwell, 2002). This result may not be surprising if Bordwell was simply looking at the earliest of films, but data from more than 13,000 films have shown that ASL is still decreasing today (Salt, 2006). Our database of 150 films supports these findings, showing a decrease in shot length beginning at the end of the 1960s. An overview of these data is presented in Figure 5.2.

One common method for detecting ASL is to simply count how many cuts a film contains and then divide that number by the length of the film, a tedious enough task. However, cuts may frequently pass without the viewer noticing, requiring that researchers looking for these boundaries be either highly skilled at detecting subtle changes in real time or examine the film at an arduously slow pace. Our analysis utilized custom video-processing software to look for the statistical changes that accompany a transition in film, as well as confirmation by human observers (Cutting, Delong, & Nochelfer, 2010). After verifying the location of each cut throughout the film, we were able to compute the length of every shot. Despite the popularity of this metric, ASL may be inappropriate because the distribution of shot lengths does not follow a normal bell curve, but is rather a highly skewed, approximately log-normal distribution. Thus, although most shots are short, a small number of remarkably long shots inflate the mean. The large majority of shots in a film are actually below average in length, leading to systematic overestimation of an individual film's shot length. A better estimate of a film's median shot length, a metric that shows the same decrease in shot length over time but provides a better estimate of shot length, as shown in Figure 5.3. Regardless of metric used, however, it's clear that shot lengths in film have been decreasing over time.

The most common explanations for the decrease in shot length usually revolve around technology or cultural factors. The argument from technology claims that cheaper film and the rise of digital editing give directors and editors the ability to cut at a pace that earlier generations would have done if given the chance. Others explain decreasing shot length as an effect of culture, of the younger generation's lowered attention span or rises in attention-deficit hyperactivity disorder; in the 1980s, this change was blamed on the fast cutting and short duration of music videos that catered to the "MTV Generation" (Postman, 1983). Recent literature warns that the next generation's attention span is being damaged by video games (Swing, Gentile, Anderson, & Walsh, 2010), the internet (Carr, 2010), and Twitter (Ebert, 2010). This isn't the first

![Figure 5.2](image1.png)

**Figure 5.2** Average shot lengths (in seconds) increased with the advent of sound films in the late 1920s, but have been experiencing a steady decrease since 1960 (adapted from Cutting, DeLong, and Brunick, in press).

![Figure 5.3](image2.png)

**Figure 5.3** A plot of shot lengths in seconds by how common shots of those lengths are in the film *A Night at the Opera* (1935). Because the distribution has a heavy positive skew, median shot length can be seen as a more accurate description than average shot length for shots in a film.
time. Youth culture has been vilified, however; Frank Sinatra once claimed that the music of Elvis Presley “fosters almost totally negative and destructive reactions in young people” (Shapiro, 1981, p. 305).

Both arguments from culture and technology can be countered by a simple fact; Salt’s data, shown in Figure 5.2, show us that films in the late silent era (1920s–1930s) exhibited editing that was essentially as fast-paced as today. Critics of modern culture would be reticent to say that the pace of life in the late 1920s was equivalent to today. We can also rule out a purely technological explanation for the decrease in ASL because editing equipment in the 1920s would be considered primitive even by 1960s’ standards.

A more satisfying explanation for the equivalence between ASLs in the late 1920s and 1990s lies in the introduction of sound. Incorporating dialogue and a set soundtrack changed a number of ways that films were made. These changes promoted increased use of techniques such as shot/reverse-shot that are often the backbone of contemporary films. As the narratives within sound films became increasingly complex, shot length also increased. Near the conclusion of Hollywood’s studio era, many filmmakers felt that film needed to compete with television to combat falling viewership (shown in Figure 5.1). The resulting push created larger, event-centered films like *Tora! Tora! Tora!* (1970), a film considered to be a commercial failure at the time. Filmmakers were struck with a problem: How do you create complex storylines while keeping audiences interested?

A number of films from this era exhibited a different way of presenting a narrative, one that was inspired by foreign styles of editing, such as French New Wave; this was quickly adopted and modified by a new generation of filmmakers. One often-examined film from within this era is *Easy Rider*, a 1969 film directed by a violent and cocaine-addicted Dennis Hopper. The original cut for *Easy Rider* was more than 4.5 hours in length but was pared down to a palatable 90 minutes by adopting a number of quick cuts out of necessity, as well as for the sake of being stylistically different.

Bordwell (2002) highlighted a number of stylistic changes that have taken place since the 1960s that have led to more condensed and intense narratives. These films were made using a fast cutting pace and different lens types, including close-up shots in dialogue and free-ranging cameras that move around an otherwise static scene. It’s also worth pointing out that films from the 1960s weren’t just changing thematically, but also show increasingly different structure as well.

The quick-cutting style that has become more commonplace in film may also have benefits outside of simply compressing the narrative. In recent work, Pronin and Wegner (2006) found that quicker “thought speed” generates a more positive affect in an individual. The speed of thought can be induced by external sources, including the speed of shots in a film clip; people who were shown clips with a rapidly moving shot pace reported a more positive mood than did those shown similar clips with slower moving shots (Pronin & Jacobs, 2008).

Although the speed of cut sequences no doubt influences perceptual and emotional elements in the viewer, cut speed is not the only variable responsible for the perception of newer films as more “fast paced.” The increased prominence of the action genre has coupled quick cutting with increased motion (optical change resulting from objects in the environment) and movement (the camera itself changing position). We chose to conduct analysis on this other type of “speed” in film, the speed with which activity occurs on-screen.

Motion and Movement on Film

There is little doubt that the tools filmmakers use to shoot and edit films have changed dramatically since the 1930s. Cameras have continually become smaller, lighter, and higher in quality in nearly every decade (Salt, 2006). Regardless of these changes, Hollywood has practiced conservative camerawork from the beginning, when filmmakers feared that any amount of camera motion would confuse and disorient their viewers (Bottomore, 1990). These fears were eventually dampened; films today often have subtle camera motion that viewers don’t even notice. Today, we know that some degree of camera motion can be tolerated, but how much can we deal with?

A number of recent films have pushed the envelope of camera motion, leaving some viewers to question whether these “queasy-cam” films are hitting a limit (Ebert, 2007). One of these films is J. J. Abrams’s *Cloverfield* (2008), a romp through monster-ravaged New York City filmed from the perspective of a handheld camera. The deliberately unsteady camera work was so extreme that several theaters were forced to put up warnings so that they weren’t liable for any ill-effects related to induced seizures or motion sickness. Not all films feature the same level of continuous movement as *Cloverfield*; other action films, like *The Bourne Ultimatum* (2007) and *Quantum of Solace* (2008), feature sequences with very fast cuts and extreme camera movement as a means of giving the viewer a chaotic interpretation of events.

Moviegoers who watch these films walk away with an understanding that the films feature a different type of editing, but how can we quantify this change? How can we place the films of the 1940s on a scale to *Cloverfield*? The simplest way to investigate this relationship is to compare how much change occurs between two sequential still frames. This can be done by using a two-dimensional Pearson correlation, a technique that compares every pixel in
an image to that of a second image. For the films in our database this required comparing 65,536 pairs of pixels for each of the roughly 165,000 image pairs in a typical Hollywood film. At the turn of the millennium, this technique would have required considerable processing power, storage resources, and months of processing time. It can currently be accomplished within a feasible timeframe on a basic laptop computer. To make our results more intuitive, we calculated the effects of camera motion and scene movement into a single metric, the Visual Activity Index (VAI), which can be described as 1 minus the median interframe correlation.

It is clear that visual activity has been increasing over time, a trend shown in Figure 5.4, with action films leading the way (Cutting, DeLong, & Brunick, 2011). The motion and movement in film is becoming more pronounced, but where will this trend stop? Research in the area of visual perception has shown us that a series of images can be recognized even when they are presented every 100 milliseconds, a methodology known as rapid serial visual presentation (RSVP; Potter, 1976). It seems clear that our visual system limits how dissimilar frames can be in a feature film; interpreting a disconnected series of images is difficult for more than a couple seconds at a time. Average RSVP sequences have a VAI of roughly 0.80, but also depend on the images being displayed. Cloverfield’s VAI for the entire film is only 0.24. This places it well short of being a random sequence of images, but with a vastly higher amount of motion and movement than films like All About Eve (1950), which has a VAI of 0.012.

Motion and movement have developed a distinct relationship to stimulus duration in recent years. Figure 5.5 shows the relationship between the duration of presentation (either of a shot, series of shots, or image in a rapid serial visual presentation sequence) and visual activity. When considering a whole film, a relatively low level of visual activity is present, even in films considered distinct in their levels of visual activity (for example, The Bourne Ultimatum and Cloverfield). However, when sequences are extracted from a film, a trend emerges that suggests more activity can be tolerated as long as it occurs over a shorter period of time. For example, a segment of the escape from the burning hotel in Quantum of Solace lasting for about 10 seconds exhibits a higher overall VAI than a 1-minute segment of the tunnel car chase sequence at the beginning of the film.

It appears as if the amount of camera movement and object motion that will be tolerated in a film isn’t a constant rate, but rather a dynamic saturation point we reach when we’ve simply processed too much variation for too long. Highly noncorrelated sequences in film exhibit a high amount of visual activity, but can only persist with that level of activity for a certain amount of time before having

![Figure 5.4](image1.png)  
**Figure 5.4** Visual activity (described as 1 minus the median interframe correlation for each pair of frames in a film) has been slowly increasing with time (adapted from Cutting, DeLong, and Brunick, in press).

![Figure 5.5](image2.png)  
**Figure 5.5** The amount of visual activity in different types of media appear to fall on the same line. Rapid serial visual presentation sequences (images shown in quick succession) can be viewed for a few seconds at a time, but would not be tolerated at longer timescales (adapted from Cutting, DeLong, and Brunick, in press).
to back off and retreat to baseline. This finding makes sense intuitively; however, psychological experiments rarely take into account how our perceptual abilities may fluctuate over timescales longer than 3 seconds. This fluctuation not only has consequences for the use of visual activity, but also for how shots are distributed. It begs the question as to whether shot lengths are catered to these fluctuations in the same way visual activity is; however, these fluctuations are likely linked to attentional systems rather than perceptual systems.

Shot Structure: Evolving Patterns

Although ASL and visual activity give us a good metric of how films are changing over the decades, they aren't very descriptive about the structural components of an individual film. As sequences of shots in film become more standardized, the length of an individual shot isn't independent from its position in the film. To numerically determine the presence of these types of patterns, we borrowed a technique from David Gilden, an astrophysicist-turned-psychologist who has studied hidden structure within human reaction time data.

Many experiments in the field of psychology utilize reaction time, a metric in which participants are asked to respond quickly to a particular target stimulus and inhibit responses to the nontarget stimuli. Individuals performing this type of task perform individual trials at different speeds even when performing the same task repetitively. These variations are usually averaged out and "relegated to a kind of statistical purgatory" even though they may actually hold some kind of structure within them (Gilden, 2001, p. 33). Mathematical tools have uncovered similar patterns in other phenomena, such as natural scenes (Field, 1987), the presence of solar flares (Lu & Hamilton, 1991), the population of different cities across the US (Newman, 2005). Gilden was the first to use these techniques to characterize this structure within the patterns of human attention.

When performing cognitive tasks (such as deciding whether a string of letters makes up a word), human reaction times exhibit a type of temporal structure called a $1/f$ pattern (Gilden, 2001; Gilden & Hancock, 2007). This pattern is also known by a number of other names including "pink noise" or "fractal noise." Fractal noise is an especially apt description because the presence of this pattern suggests self-similarity at different scales. Mathematicians have known these patterns, often called "mathematical monsters," for centuries. They are characterized as being difficult to describe using Euclidian geometry, but pioneering work by Benoit Mandelbrot in the 1970s offered elegant explanations for these patterns.

Finding a $1/f$ pattern within the context of reaction times suggests that our bodies and brains have a number of different mechanisms that contribute to the completion of a reaction time task. The time in which these mechanisms complete the task isn't necessarily constant and varies based on whether these mechanisms are in sync. It is also important to note that the magnitude of the influence of these mechanisms varies proportionally with the amount of time it takes for these fluctuations to occur.

To test if the pattern of shot lengths in Hollywood films follows a similar pattern, we analyzed our previous data, using cut boundaries to calculate a series of shot lengths for an entire film. We then used Gilden's technique to calculate the power spectra for each of the 150 films, using the sequence of shot lengths as a time series. This technique allowed us to estimate the slope of the power spectra, a diagnostic metric of self-similarity. If the slope of the power spectrum is equal to 0 (known as a "flat spectrum"), then all frequencies are equally likely in the signal, meaning that there is no way to predict the next value, and no temporal structure exists within the signal. A slope of -2 means that the process can be modeled as a random walk, commonly described as a mathematical abstraction in which something moves by simply choosing a sequence of random steps. A slope of -1 lies directly between these types and is indicative of a $1/f$ pattern.

Our data show that after the 1960s (roughly the beginning of shot length's most recent decrease), films increasingly adhere to a $1/f$ pattern in their shot lengths (Cutting, DeLong, & Noltehelfer, 2010). There is no clear reason why this change would occur at the same time as decreasing shot lengths; computing the slope of the power spectra isn't affected by the average value, but rather by the relationships between values. Fluctuations in human attention that follow a $1/f$ pattern tend to mirror the same type of pattern found in the shot lengths of Hollywood films. These similarities lead us to believe that film may be evolving; the characteristics of films may have changed over time to better serve cognitive mechanisms such as attention (see Figure 5.6). Filmmakers clearly haven't been consciously attempting to replicate a $1/f$ pattern while editing the next big blockbuster, so why is this pattern becoming more common in Hollywood films? Filmmakers are trained to use their own intuition to use cuts, camera work, and various other techniques to create something that simply feels right. This process of trial and error started when early filmmakers like Georges Méliès had to invent strategies to string multiple events together in a single film. This system of trial and error is still alive today; a pioneering filmmaker introduces a new technique or style that is remembered, reused, and refined. It's not difficult to see how this process acts like a genetic algorithm in which the most successful techniques are remembered and copied, and the failed experiments languish in obscurity. Over time, filmmakers settle on methods that
simply look and feel right, emerging with an intuitive understanding of how the brain understands film.

The fact that Hollywood film is a good modulator of attention appears obvious: presenting a film distracts and pacifies both unruly children and airline passengers. Exciting research shows that this effect can also be seen in our neuroanatomy. Researchers have found that some films can exert considerable control over brain activity and eye movements when subjects were shown different types of film while undergoing a functional magnetic resonance imaging (fMRI scan) (Hasson et al., 2008, p. 1). The results of the scan showed that when a group of viewers watched a highly structured film like Sergio Leone’s *The Good, the Bad, and the Ugly* (1966), the number of brain regions that showed synchronous activity between subjects was ten times larger than when subjects watched unedited footage of a public park.

Film is not only becoming better at modulating our attention, but it can also provide researchers insight into the inner workings of the human visual system. Ongoing work aims to continue this exchange and extend our dialog beyond the mechanisms of attention to other important topics such as event segmentation, emotion, and narrative. Although our work has only scratched the surface, we hope to help introduce a fine-grained perspective into the cognitivist approach to film studies; not simply focusing on a philosophical interpretation of a single film, but also exploring how the smallest parts of a film (cuts, shots, movement) combine to construct meaning from moving pictures.

### Glossary

**1/f distribution** Representing 1 over frequency, this type of pattern can be characterized as the output of a system that is halfway between random and rigid. This distribution is found in many places in nature and is thought to signify self-similarity at multiple scales, as shown in fractal patterns.

**ASL** Average Shot Length. The measure of how often a filmmaker cuts during a movie. Over the past few decades, the rate of cutting has increased and ASL has decreased as a direct result.

**VAI** Visual Activity Index. A metric designed to approximate visual motion and movement present within part of a film. This metric can be characterized as a two-dimensional Pearson correlation that is subtracted from 1. Higher values denote more visual activity.

### Note

1. It’s worth noting, however, that 1/f patterns aren’t only found in the fluctuations of human attention but in a number of varied phenomena across the earth and in space. The patterns of change found when measuring the height of the Nile River, the diameter of asteroid impacts on the moon, and the size and position of leaves on branching plants all follow the same type of pattern, yet we wouldn’t dare make a claim about how they are related to Hollywood films. “Patterns of attention” merely seems like the best current explanation, but future research may show that film’s gradual movement toward a 1/f temporal pattern may be catering to something different entirely.

### References


Other Films


**Filmography**

**Database Films**


THE SOCIAL SCIENCE OF CINEMA

Edited by James C. Kaufman and Dean Keith Simonton