

# Influence of Camera Movement on Attention: Evidence from Eye Tracking

Hsien-Chih Chuang<sup>1</sup>, Han-Yi Tseng<sup>2</sup>, Da-Lun Tang<sup>3</sup>, and Yuh-Chang Wei<sup>4</sup>

<sup>1</sup>Department of Information Communications, Chinese Culture University, Taipei, Taiwan, zxz7@uivlve.pccu.edu.tw

<sup>2</sup>Department of Photography and Virtual Reality Design, Huaan University, New Taipei City, Taiwan, jumeitw@gm.hfu.edu.tw

<sup>3</sup>Department of Mass Communication, Tamkang University, New Taipei City, Taiwan, 136216@mail.tku.edu.tw

<sup>4</sup>Department of Information Communications, Chinese Culture University, Taipei, Taiwan, ycwei@faculty.pccu.edu.tw

## ABSTRACT

Numerous papers have discussed practical filming techniques and subjective criticism in cinematography, but few have assessed how camera movement influences viewing behavior. This gap is indicative of the limited knowledge of the movie-viewing process and the lack of objective evidence supporting the correlation between audience visual cognition and motion in cinematography. This research compared the sightline properties of four types of cinematography shots, with each type comprising five segments of 15-s film fragments, by measuring indicators such as fixation duration, number of fixations, and fixation dispersion to further understand the mechanisms with which audiences watch movies. Studies have indicated that the use of motion shots rather than static shots can influence the viewing process. Changing the position of the main actor using follow shots induces an increase in the number of fixations and subsequently maximizes fixation dispersion for the viewer, who wants to view the details within the various shots.

**Keywords:** Attention Evidence, Camera Language, Camera Movement, Eye Tracking.

## 1. Introduction

If the manner in which movies tell a story were analogous to a language, then each shot could be regarded as a word; the entire movie is composed of these words, and the understanding of a work of cinema thus begins with a grasp of the fundamental vocabulary. Similarly, the understanding of how a viewer watches a movie may also require knowledge of how individual shots influence the viewer. A movie is the product of a combination of visual and auditory components, and the syntax of movies comprises the camera angle, storyboarding, cutting, view finding, and image composition. A quality movie requires skilled acting, an engaging script, and a highly unified production team, with the actual filming forming a crucial part of the film production process. The movement of the camera, the medium for storytelling, can generate new language and bestow new meaning and logic to a movie, providing the audience with a novel type of imagery and increasing the joy of movie watching. Suboptimal camera movements, however, can dampen the flair of a story. Film, television, and video are ubiquitous, and viewers of these media generally have similar narrative experiences despite the complexity of the audiovisual stimuli and the considerable individual differences across viewers (Hutson,

Smith, Magliano, & Loschky, 2017). Although various forms of film depict something related to real life, such depictions can be rather conventional. Although directors are generally familiar with the rules, they may not be able to articulate them to their full potential. For example, camera movement, actor movement, lighting, and color all have specific techniques for attracting audience attention (Marchant et al., 2009).

The psychology of film is a subfield in applied psychology that focuses on movie image attributes, movie composition, and audience experience. In 1916, German psychologist Hugo Münsterberg published *The Photoplay: A Psychological Study* in which he stated that movies do not exist in film form or on the screen but instead exist in the viewer's mind. He stressed that movies are psychological constructs and performed preliminary empirical investigations. *Film as Art* by Arnheim (1957) was based on Gestalt psychology and systematically investigated the form, visual perception, and artistic characteristics of film. *The Aesthetics and Psychology of the Cinema* by French movie theorist Jean Mitry (1963) is another seminal work on film psychology analyzing subjective imagery and elaborating on psychological elements of film such as panorama; the first-person shot, deep view shot, and moving shot;

structure; color; and music. To comprehend a movie, the director must guide the viewer's sight and be aware of where and what draws the viewer's focus (Glebas, 2008). The psychological study of film watching analyzes the narrative within the imagery and the natural transitioning of viewers' attention. Studies of this nature have often been conducted through observation, questionnaires, and film analysis, but empirical studies of film aesthetics are rare. Eye tracking is a critical method of quantifying viewing behavior. The present study investigated viewing behavior and psychological elements such as the focal point of a light line and the scan path during movie watching to elucidate the relationship between movies and visual perception and to initiate an empirical study of movie narratives and film aesthetics.

### 1.1 Language of the Lens

The core of the art of filmmaking is photography, but the photographic techniques of filming, such as camera movement, filming angle, cutting, view finding, and image construction, create a language unique to films; this language of the lens has the ability to convey meaning through camera movement and plays a critical role in filmmaking as a delivery medium. Because imagery and camera movement can convey meaning in a story, camera movement is the basis of creating movie imagery (Block, 2001; Bordwell & Thompson, 2001; Hu, 2013; Monaco, 2000). Photography is also a key contributing aspect in filmmaking. Camera movement is not simply about moving the camera but also about the lens focus, which is used to draw the gaze of the audience, define the image, and shift the attention of the viewer (Hu, 2013). Without changing the story, character design, or background, differences in camera movement and motion segmentation can create differences in visual tension and offer the viewer a distinct emotional experience (Block, 2020). Effective camera movement can improve the visual tension and conveyance of a story. Camera movement is thus not only the first lesson in photography but is also the "make-or-break" factor for capturing the visual interest of the audience. If camera movement is utilized properly, telling stories through the lens, changing viewers' perspectives, guiding their gaze and focus, and immersing them in a movie can all be achieved to provide viewers with an enriched emotional experience.

Filming techniques have advanced with each generation of filmmaking and have resulted in the creation of various camera movement techniques, such as dolly, track, and stabilizer techniques, all of which instill the imagery with "freshness" and energy. Esteemed directors such as Steven

Spielberg and Martin Scorsese have created their own signature filming styles using some of the aforementioned techniques and have created numerous timeless film segments. Moving camera shots can be categorized into panning, tilting, swish panning, crane, tracking, dolly, pull-back, traveling, follow, zoom lens, static camera, and arc shots, among other variations. This research focused on the following commonly used camera movement types:

A follow shot can display the intricate details of the scenery. In this shot, both the camera and object being filmed move in tandem. This type of shot is typically employed in a long continuous take, whereby the frame follows the object or person being filmed while simultaneously featuring complex background elements, character movements, and various other details.

A whip pan shot is often used in filming as a means of transitioning between scenes. When the camera quickly "whips," a blurred image is generated to exit from the previous scene and to impart movement for added vibrancy, with the changing scenery or character positions connecting one scene to the next. This is an effective method of scene transition and conveying to the audience the transition between two spaces or characters. The imagery created has a strong visual impact and a surprising or refreshing effect.

A dolly shot can be captured by fixing a camera to a moving platform that allows for stable movement of the camera when filming. The platform moves the camera toward or away from the scenery, creating variation in the field of view and visual angles. The dolly shot is suitable for displaying characters' emotions to enhance the atmosphere in crucial scenes. The velocity of the dolly can be used to adjust the pace of the scene, which can heighten emotional intensity. Famed director Steven Spielberg has often applied this technique at different velocities when shooting characters. Rapid movement conveys a sense of urgency, whereas slow movement is suitable for expressions of deep thought. In *Close Encounters of the Third Kind*, the dolly shot pulls in from afar for a close-up to express a character's awe at seeing a descending flying saucer.

A dolly zoom, a combination of the dolly shot and zoom, is another movement often applied in movies, whereby the subject being filmed remains a fixed size while change is presented in the background. This technique is also often used by Spielberg to guide the focus of the audience and was one of Alfred Hitchcock's signature techniques for exaggerating tense emotions, such as anger, anxiety, and fear. When used to film large scenery, this technique can highlight the

spatial contrast between the character and background and, when used to capture the facial expression of characters, can impel the audience into an intended mental state. Mercado (2013) observed that the dolly zoom generates an unusual perspective change that can transform meaningful moments or scenes into images, direct the audience's focus onto something peculiar, and effectively display a character's strong emotional or internal state.

## 1.2 Eye Movement

Eisenstein (1943) may be the first film theorist who paid attention to visual guidance. He stated that the art of cinematography is about precisely guiding the audience's attention toward the creator's intended sequence of scenes and paths, which is synonymous with how a painting guides a viewer's eye across a canvas in the processing of each image segment. Anderson considered the observation of change in gaze in the structure of film theory to be a critical aspect. In practice, skilled film directors have established a set of movie syntactical elements as the means of expression, which include the use of camera movement and cutting to control the tempo of the movie, to guide the audience's viewing process. However, evidence of the effectiveness of such guidance is lacking (Tang & Lai, 2016).

Numerous experiments in psychological research have verified that eye tracking is a crucial method of effectively observing the thought process of a participant. It can naturally and instantaneously detect shifts in focus. Observing the human eye can thus provide an objective reflection of complex inner states (Chuang, Wei, Tang, & Tang, 2022; Duchowski, 2003; Henderson & Hollingworth, 1999; Ma & Chuang, 2015; Rayner, 1998; Zhou et al., 2016). The most direct method of measuring an individual's movie-viewing process is to record the movement of their eye during the entirety of a movie-watching session. Eye-tracking experiments have been applied under complex and dynamic conditions to analyze the viewing process (Rayner, 1998; Smith, 2012; Smith & Henderson, 2008; Smith & Mital, 2013; Tang & Lai, 2016; Treuting, 2006; Vig, Dorr, & Barth, 2009).

Henderson and Hollingworth (1998) revealed two crucial aspects of observing eye movement, namely the area where the fixation is centralized and the duration of the fixation on that location. Because eye movement can reveal the process of watching a movie, observing signals from eye movements can reveal where the viewer's attention is fixed and what catches their interest. Thus, at which point an audience is gazing and the space where attention is focused are strongly related (Antes, 1974; Cheng, 2015; Duchowski,

2003; Henderson & Hollingworth, 1998, 1999; Megaw & Richardson, 1979).

## 1.3 Research Variables

Number of fixations and saccades: Fixation, or gazing, is the state under which the eye moves slowly and has been critically studied in eye movement research. These two fundamental indicators can potentially reveal the complexity of processing external stimuli during gazing (Henderson, 2007; Ma & Chuang, 2015; Wang & Jian, 2022). The more complex or detailed the information presented is, the fewer times fixation occurs (Antes, 1974; Baker & Loeb, 1973; Mackworth & Morandi, 1967). Fixation can also reveal the level of immersion, whereby less gazing can indicate that the viewer is more immersed in the screen imagery (Salvucci & Anderson, 1998). When total viewing time is fixed, a lower number of fixations can indicate that each fixation lasted for a longer period of time. A saccade is the rapid movement from one gaze point to the next. Therefore, similar to fixations, the number of saccades should in theory reflect a viewer's mental state.

Spatial Dispersion Index (SDI) of fixation: Fixation dispersion is a less commonly used indicator that involves a topographical processing system producing coordinates and concentration data, which compensates for the lack of spatial information of the previous indicator. It reveals the level of dispersion of the gaze points on the screen, which provides an understanding of the level of concentration on the area of interest. Research has indicated that, compared with viewing still photos, watching films involves a significantly higher level of attentional synchrony from the viewer from one frame of the motion picture to the next. Fixation dispersion can be used to reveal this phenomenon (Ma & Chuang, 2017).

Smith (2013) reported that people have limited attention capacity and are unable to receive the entirety of visual information during movie watching; instead, the eye must continuously travel to extract the desired information within view. When watching a film, viewers typically move their eyes two to five times per s to extract information, and those eye movements are likely related to viewers' understanding of the film they are watching (Eisenstein & Leyda, 1948; Hutson et al., 2017; Jesionowski, 1989; Murch & Coppola, 2001; Smith, 2012). With information gathered through eye tracking, such as viewers' fixation dispersion, number of fixations, and fixation duration, the process through which information is processed internally can be extrapolated. Smith (2013) asserted that attention is the culmination of unconscious gazing and the

point at which the subject is looking, which indicates where the attention is focused. Studies have reported that the reason filming techniques and film editing can influence the level of enjoyment and comprehension of the viewer is because movie makers inadvertently apply the theory of attentional synchrony. The successful arrangement and control of camera movement, actor movement, editing, and other methods successfully capture the attention of the audience and allow them to effectively grasp the message and creative intention of the filmmaker at a particular moment in the narrative, achieving emotional harmony and logical resonance between the audience and director (Luan, 2018).

### 1.4 Film-Watching Research

Film is ubiquitous, but the processes that guide viewers' attention while viewing film narratives are poorly understood (Hutson et al., 2017). Film critics traditionally review films from the subjective perspective of the film as an art form (Dmytryk, 1984), and most studies have employed qualitative analysis or questionnaires. When directors make films, their main concern is how the work is presented and conveyed; however, their knowledge of how people view movies is limited due to the few experimental aesthetics studies on camera movement and viewing behavior. Münsterberg (1916) maintained that psychology provides a perspective for understanding how movies can affect audiences. Interdisciplinary film psychology study is applied to address an aspect ignored in film studies, namely the effect of the viewer's consciousness and preconscious (Bordwell, 1989). If the relationship between controllable storytelling factors that affect image saliency can be discerned, then this type of study represents a potentially illuminating path in film research (Dyer & Pink, 2015; Martinez-Conde et al., 2004; Parkhurst, Law, & Niebur, 2002; Tatler, 2014). Researchers have employed magnetic resonance imaging to study the viewing of film segments and have discovered that films can influence activity in the brain through the use of imagery, cutting, and filming methods (e.g., Hasson et al., 2008). This type of research into movie-viewing patterns applies empirical aesthetics and psychology to analyze the relationship between film elements and the viewer's experience (d'Ydewalle & Vanderbeeken, 1990; Flagg, 1978; Goldstein, Woods, & Peli, 2007; Hasson et al., 2008; Hochberg & Brooks, 1978).

The present research into film viewing combined quantitative and qualitative experimentation in a scientifically objective approach. The relationship between camera movement and the

viewing process and their ability to generate changes in experience are highly relevant for film producers and directors wanting to determine the effectiveness of filming techniques. This study focused on which camera movement techniques directors use to attract interest. If different camera movements can affect the viewing focal point and viewing pattern during movie watching, then the eye would be expected to exhibit different fixation states, saccade changes, and dispersion patterns, all of which reflect the viewer's internal experience. An appropriate understanding of filming methods can support the cameraman in improving storytelling through imagery to more effectively convey the artistic conception of the director. This can make the movie experience more enjoyable for the viewer and can benefit the development of modern film art as well as practical film theory. Eye-tracking experiments with film viewing generate a large quantity of data, which increases the complexity and time consumption of statistical calculations. Therefore, in this research, an independent algorithm was developed to streamline the process of obtaining crucial data of the scan path and to generate a dynamic heat map.

## 2. Methods

This research primarily used eye tracking to explore the effects of different camera movement methods on viewers' cognitive experience, viewers' subjective emotional experience, and the difference in viewers' sightline distribution to identify the key factors affecting how people view movies as well as particular viewing patterns that can be integrated into film making. The experimental structure of the study is presented in Figure 1. We proposed the following hypotheses:

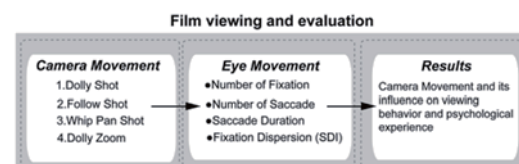


Figure 1. Research Framework

*H1:* Different camera movement segments in a movie can induce different viewing patterns, and each viewer's sightline focal point differs slightly.

*H2:* Dolly zoom segments in a movie result in fewer fixations and saccades, a shorter saccade duration, and an overall less even dispersion of sightline focal points.

### 2.1 Participants



A monetary reward was offered to encourage participation, and a total of 26 participants were recruited. Among them, 12 were men and 14 were women; they were aged between 18 and 24 years and watched a reported average of 1.76 movies per month. All participants were Mandarin-speaking college students. The experiment was designed to present stimulus material in random order. We applied a within-subject design, whereby all participants were subjected to all stimuli.

## 2.2 Stimuli

A film expert was asked to evaluate and select

segments that matched the four types of camera movement studied in this research (i.e., follow, dolly, whip pan, and dolly zoom shots). Segments with the four camera movement types were selected from five iconic movies for the experiment, for a total of 20 representative film segments. The experiment consisted of playing 20 segments of moving shots and 20 segments of still shots for a total of 40 film segments. Each segment was 15 s long, with segments displayed in random order. The segments had a resolution of  $1024 \times 768$  pixels played at 25 fps. Only audiovisual content was presented without subtitles. The segments are listed in Table 1.

**Table 1. Experimental Stimuli**

Camera movement type	Film title of segments				
Follow shot	The Wrestler (2008)	Boogie Nights (1997)	Pulp Fiction (1994)	Three Billboards Outside Ebbing, Missouri (2017)	007: Spectre (2015)
Dolly	Whiplash (2014)	Strays (1997)	20th Century Women (2016)	Indiana Jones and the Raiders of the Lost Ark (1981)	Jurassic Park (1993)
Whip pan	Magnolia (1999)	Shaun of the Dead (2004)	Hot Fuzz (2007)	Whiplash (2014)	La La Land (2016)
Dolly zoom	The Quick and the Dead (1995)	Poltergeist (1982)	Road to Perdition (2002)	Goodfellas (1990)	E.T. the Extra-Terrestrial (1982)

## 2.3 Variables

The experiment comprised the camera movement type (i.e., follow, dolly, whip pan, and dolly zoom shots) as the independent variable and eye movement indicators (i.e., number of fixations, number of saccades, saccade duration, and fixation dispersion) as four dependent variables.

## 2.4 Procedure

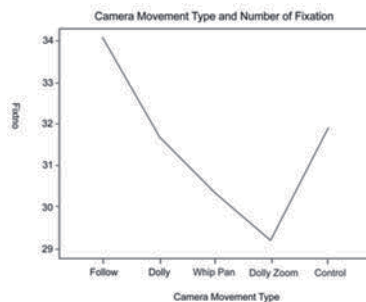
Before the experiment began, the participants were requested to sit 60 cm in front of a 21-in viewing screen centered and facing the observer. The participants first looked directly at an eye-tracking device (Tobii Pro Nano), with the sampling frequency set to 60 Hz to capture gaze

data, and then a 9-point calibration sequence was initiated. The participants practiced operating the testing apparatus and read the instructions before the experiment began. During the experiment, they were exposed to film segments comprising 20 15-s clips played in random order. To ascertain whether the viewers paid attention to the segments, after each trial, each participant was immediately asked about the contents of the clips they had viewed. The entire experiment lasted approximately 30 min.

## 3. Results

Different camera movement techniques have different effects on sightline. To elucidate the

sightline distribution and movement during film viewing, we incorporated all of the gaze trail data of the 15-s segments into the analysis in addition to a set of gaze trail data with static shot segments as a control. The data revealed that camera movement had a significant effect on the number of fixations (Mauchly's sphericity test ( $W = 0.75$ ,  $p = 0.081$ ),  $F(4, 100) = 3.667$ ,  $p < 0.01$ ), indicating that the motion shot method affects the overall sightline distribution. The segments with dolly zoom shots had the lowest number of fixations, and those with follow shots had highest. This result implied that the viewer's eyes are guided by the camera shots and rapidly move to track what is occurring on the screen, as illustrated in Figure 2.

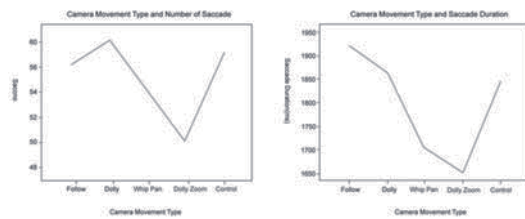


**Figure 2. Film Segments With Camera Movement Types. The dolly zoom shot had the lowest number of fixations**

The results of the ANOVA analysis indicated that the number of saccades (Mauchly's sphericity test ( $W = 0.81$ ,  $p = 0.054$ ),  $F(4, 100) = 2.743$ ,  $p < 0.05$ ) and saccade duration (Mauchly's sphericity test ( $W = 0.81$ ,  $p = 0.086$ ),  $F(4, 100) = 3.974$ ,  $p < 0.01$ ) were both significantly affected by the type of camera movement, indicating that camera movement affects the overall sightline distribution. The dolly zoom shot had the lowest number of fixations and shortest saccade duration. Because saccade amplitude reflects the complexity of the viewed information, this result indicates that dolly zoom segments are more complex and require more time to mentally process. The follow shot generated the highest number of fixations and saccades, indicating that this shot is more visually attractive and maintains the viewer's interest, as presented in Table 2 and Figure 3.

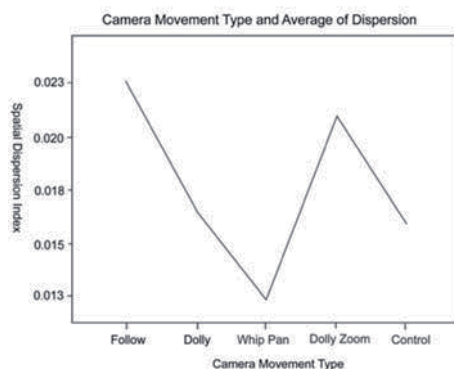
**Table 2. Characteristics of Gaze for Various Camera Movement Types**

Camera movement	Number of fixations (Mean, SD)	Number of saccades (Mean, SD)	Saccade Duration (Mean, SD)	Spatial Dispersion Index (SDI) (Mean, SD)
Follow shot	Most (34.0, 11.8)	High (56.6, 30.8)	Longest (1918.4, 820.0)	Most Dispersed (0.023, 0.006)
Dolly shot	Moderate (31.7, 12.9)	Highest (58.9, 37.6)	Long (1862.4, 885.9)	Moderate (0.016, 0.005)
Whip pan shot	Few (30.3, 10.4)	Moderate (53.3, 34.1)	Short (1700.1, 839.5)	Most Focused (0.012, 0.003)
Dolly zoom	Fewest (29.2, 9.8)	Fewest (49.4, 28.2)	Shortest (1651.4, 745.0)	Dispersed (0.021, 0.005)
Control (static shot)	Moderate (31.8, 10.4)	Moderate (56.5, 31.8)	Moderate (1842.3, 786.4)	Moderate (0.015, 0.004)



**Figure 3. Number and Duration of Saccades Among Camera Movement Types. The dolly zoom generated the lowest number and shortest duration of saccades.**

The results revealed that camera movement type has a significant effect on the dependent variable of dispersion (Mauchly's sphericity test ( $W = 0.617, p < 0.001$ ),  $F(4, 1496) = 303.784, p < 0.01$ ), indicating that different camera movements generate different fixation dispersion. The whip pan shot SDI dispersion was lower and more centralized, reflecting that the swift switch in scenery results in a concentrated fixation sightline. The follow shot SDI distribution was higher because this movement requires the viewer to continue observing changing surroundings, resulting in a dispersed sightline, as illustrated in Figure 4.

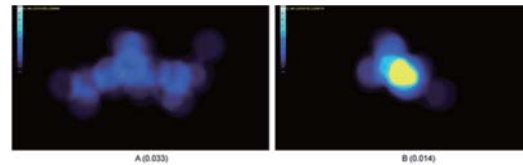


**Figure 4. Gaze Dispersion of Different Camera Movement Types. The dolly shot generated the most centralized fixation points, and the follow shot generated the greatest fixation dispersion.**

## 4. Discussion

We analyzed the camera movement types and created a dynamic heat map. Quantifying where and when the viewer is looking is difficult, as is determining which type of movie attracts the most attention. We merged the viewer gaze data with dynamic heat mapping, as depicted in Figure 6, to conduct a general exploratory comparison. Using an appropriate sightline change attribute value, we could identify the dynamics during a viewing session to construct a general viewing model. Quantified dynamic heat map analysis of each

camera movement type was performed to visualize the fixation dispersion area of greatest dispersion and highest density, after which we could layer the visualized gaze data as a dynamic heat map onto the actual movie frame.



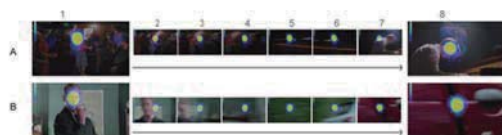
**Figure 5. Sightline Dispersion. (A) The lack of a yellow area indicates that the sightline is dispersed and not focused on a particular area. (B) Obvious focus of gaze in the center of the screen, which was significant, as indicated by the yellow area. The dispersal (0.033) in (A) is higher than that in (B) (0.014)**

In the heat map, we stacked the fixation positions from each frame, calculated the mean, and represented the frequency spectrum from highest to lowest with different levels of brightness to determine the fixation dispersion. Next, the calculated figures that were 5 standard deviations higher than the average were colored in yellow to represent the significant difference of these areas. We then individually analyzed every dynamic heat map and extracted dispersion information as a reference for model construction. The heat map properties and general comparison information are summarized in Table 2.

The aforementioned analysis revealed that movies with different camera movements affect the overall sightline distribution. The eye movement experiment demonstrated that the control group, being a static shot film segment, produced relatively moderate eye movement attributes. The dolly zoom segments differed from the other camera movement types (see Table 2), generating a significantly lower number of fixations and saccades; however, its sightline distribution was not the most dispersed. The follow shot generated the most dispersed sightline distribution as well as the highest number of fixations and saccades and the longest saccade duration. This result verifies that this camera movement type is more complicated and requires more time to mentally process. Human faces are the most noticed aspect of imagery. Psychologists have long known that when people look at static images, if faces are visible, they are the first aspect people notice (Crouzet, Holle, & Simon, 2010; Vassallo, Cooper, & Douglas, 2009; Yarbus, 1967). Treuting (2006) discovered that viewers direct a high degree of attention to faces when watching movies, which is consistent with our result; in the present research, which used dynamic film segments, the effect was the same

regardless of the distance of the faces in the shot. Most of the audience focuses their attention on the film protagonist, and other visual elements are unlikely to redirect the audience's attention. Only after the task of the main character is completed and a new character appears onscreen does the audience's attention shift to the new protagonist.

The whip pan shot often appears in films and has a particular meaning in the language of the lens and the narrative of film and television alike; namely, it is assistive for scene consolidation and postediting. The whip pan shot is used to shift the focus and attention from one character to another. Compared with changing scenes through cutting, the use of a whip pan to create subjective shots results in more dynamic imagery, enabling the audience to grasp the difference in distance between two spaces or characters. When the screen quickly switches between shots, this shot can smoothly connect elements and accelerate the narrative pace. The movie *La La Land* contained numerous sequences matching the whip pan to the rhythm of the soundtrack. In the scenes in which the shot switches back and forth between dancing and rapid playing of the piano, the study participants' sightline data demonstrated that their attention was highly focused. Even if the camera panned rapidly, generating a blurred image, the process of switching multiple scenes with no obvious subjects still created a high level of sightline consensus. The viewer was forced to continually adjust their visual focus, but, through the fluctuation in visual attention, the emotion being conveyed through the film was still felt (Figure 6A). The same sightline results were obtained using segments from *Shaun of the Dead*, where the director used the whip pan to string together various scene subjects. This method of quickly switching scenes maintains the pace of the movie but blurs the transition point, prompting the viewer to focus on the main character and not process other aspects in detail; this can enhance the viewer's interest and prevent boredom (Figure 6B).



**Figure 6. Fixation Dispersion of Whip Pan Shots.** The dispersion is more centralized, indicating the focus of visual attention.

According to the principle of vision, the moment a camera shakes is the time a viewer is most likely to become distracted. The use of the whip pan solves this problem, allowing for a scene switch that momentarily distracts yet retains the attention

of the viewer. This method was used frequently in *Birdman*. The director Alejandro González Iñárritu and cameraperson Emmanuel Lubezki used the whip pan to create editing points. Through the rapid horizontal movement of the camera, the transition between scenes occurred rapidly while the audience was momentarily distracted. The postproduction team cleverly joined the segments, creating a visual effect whereby the whole film resembled one long take. This research has thus revealed that unless a strong attraction to the character or subject exists, visual attention can be restored through the process of the whip pan. This result is consistent with that of Tang and Lai (2016), who determined that the director can control the rhythm of the film to attract attention and increase the viewer's sightline focus consensus, making more than 80% of the sightline more concentrated on a protagonist who only occupies approximately 20% of the film.

The follow shot is used to present complex details of a movie, such as the setting features and actor movements. This type of shot emphasizes using the camera to follow and film the object or character as it or they move and is typically used in a long take. The audience's attention moves with the target, with this technique adding a strong sense of immersion and participation for the viewer. The follow shot focuses on the filmed object or character and enables the audience to more deeply understand the relationship between the object or character and its environment. The heat map revealed that the viewer's gaze continually followed the camera focused on the main object or character but would also observe the surrounding scene; the sightline was thus more distributed, and the number of fixations increased. One shot conveys a single primary theme. Therefore, if the shot is too long, the attention on the object or character being filmed can be lost (Figure 7).

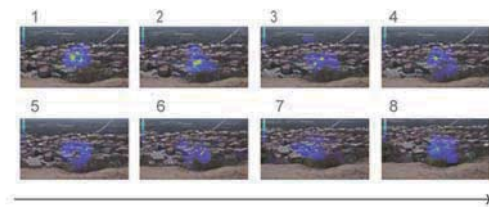


**Figure 7. Follow Shot.** This shot generated the greatest fixation dispersion, indicating a lack of visual focus.

In the dolly zoom, the camera is moved forward while simultaneously zooming in or out of focus to create a visual effect whereby the filmed target remains the same size but other objects in the shot



are perceived as shrinking or enlarging. This method effectively highlights the filmed target while presenting background changes and, if the image composition is enclosed, enables the viewer to easily discern changes between the background and foreground, guiding their attention toward the center of the screen. This commonly used technique in films displays changes in space and is typically applied to communicate characters' emotions, such as anger, anxiety, and fear, and generate an eerie and sometimes supernatural atmosphere. Therefore, this shot is the preferred camera movement of thriller and horror films. The heat map revealed that, with the dolly zoom shot, the audience invests attention and emotion into the film and lacks time to become distracted, which resulted in this shot having the lowest number of fixations and saccades as well as a more dispersed gaze focal point (Figure 8).



**Figure 8. Dolly Zoom Shot. This shot generated a dispersed sightline and the lowest number of fixations and saccades.**

## 5. Conclusion

Films are a form of modern art formed through the skilled combination of the film languages of image composition and sound. As the director conducts storytelling through images, each shot is equally crucial. Directors carefully consider the arrangement of the camera angle, the filming technique, lighting, color palette, and camera adjustments during filming to produce a quality movie that visually engages the audience. Considerations of the filming composition and camera movements are key to the success of a movie because they affect the atmosphere of story and the conveyance of the emotional upheavals and meanings within the plot. Effectively executing the conveyance of emotion can uplift viewers' spirits, stimulate their emotions, and arouse their thought processes.

This research applied quantitative methods based on empirical psychology and cognitive science to investigate camera movements in film making. We further quantified these features through eye tracking and classified viewing patterns to test the effectiveness of directing methods and to analyze changes in eye movement properties. The establishment of objective change indicators

broadens the understanding of the psychological experience generated through different camera movements and uncovers critical elements of film psychology.

This study revealed that different camera movement can generate different viewing patterns. Although no two participant's visual focus was identical, most of the data from each film segment indicated a high level of visual correlation, which supports H1. The whip pan shot switches from one scene to the next while the viewer's attention is diverted. The viewer's attention is thus focused strongly on the primary characters and objects rather than on various details, and the sightlines are thus more centralized, guiding the viewer to adjust their visual focus and attention to match the pace of the film and to feel the emotions conveyed. The dolly zoom segments are more complex, and viewers must therefore invest more attention and concentration into the film. Lacking the time for their attention to be diverted, the participants exhibited the lowest number of fixations and saccades, as well as more dispersed gaze focal points, for this type of shot, which is consistent with H2. The follow shot generated the greatest sightline dispersion and visual focus dispersion. The viewers fixated on the main character or object while continuing to look at the surroundings for details; hence, this shot had the highest number of fixations and saccades and the greatest sightline dispersion.

## References

- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, 103, 62–70. <https://doi.org/10.1037/h0036799>
- Arnheim, R. (1957). *Film as art*. California: University of California Press. ISBN 9780520248373
- Baker, M. A., & Loeb, M. (1973). Implications of measurement of eye fixations for a psychophysics of form perception. *Perception & Psychophysics*, 13, 185–192. <https://doi.org/10.3758/BF03214125>
- Block, B. (2001). *The visual story: Seeing the structure of film, tv and new media*. Oxford: Focal Press. ISBN 9780240804675
- Block, B. (2020). *The visual story: Creating the visual structure of film, tv and digital media*. Oxford: Focal Press. ISBN 9781138014152
- Bordwell, D. (1989). *Making meaning: Inference and rhetoric in the interpretation of cinema*. Cambridge: Harvard University Press. ISBN 9780674543355

- Bordwell, D., & Thompson, K. (2001). *Film art: An introduction*. New York: McGraw-Hill. ISBN 9780072317251
- Cheng, M. (2015). Cinematography research based on audience eye attention data. *Film Art*, 1, 108–114. <https://doi.org/10.29514/TJCD>
- Chuang, H. C., Wei, Y. C., Tang, D. L., & Tang, T. H., (2022). An eye-tracking investigation into the influence of filming method on the viewing behavior. *The Journal of Commercial Design*, 25, 3–11. <https://doi.org/10.29514/TJCD>
- Crouzet, S. M., Holle, K., & Simon, J. T. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, 10, 1–17. <https://doi.org/10.1167/10.4.16>
- Dmytryk, E. (1984). *On film editing: An introduction to the art of film construction*. Oxford: Focal Press. ISBN 9781138584327
- Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. New York: Springer. ISBN 9781852336660
- Dyer, A., & Pink, S. (2015). Movement, attention and movies: The possibilities and limitations of eye tracking? *Refractory: A Journal of Entertainment Media*, 25, 1–10. <http://www.axonjournal.com.au/issue-8-1/essaying-fabpod>
- d'Ydewalle, G., & Vanderbeeken, M. (1990). Perceptual and cognitive processing of editing rules in film. In R. Groner, G. d'Ydewalle, & R. Parham (Eds.), *Studies in visual information processing, vol. 1. from eye to mind: Information acquisition in perception* (pp. 129–139). Search and Reading. Amsterdam: North Holland. ISBN 9780444884350
- Eisenstein, S. (1943) *The film sense*. London: Faber and Faber. ISBN 9780156309356
- Eisenstein, S., & Leyda, J. (1948). *Film form: Essays in film theory*. New York, NY: Harcourt. ISBN 0156309203
- Flagg, B. N. (1978). Children and television: Effects of stimulus repetition on eye movement. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychology functions* (pp. 279–291). New York: Lawrence Erlbaum Associates. ISBN 9781138219816
- Glebas, F. (2008). *Directing the story: Professional storytelling and storyboarding techniques for live action and animation*. Oxford: Focal Press. ISBN 9780240810768
- Goldstein, R. B., Woods, R. L., & Peli, E. (2007). Where people look when watching movies: Do all viewers look at the same place? *Computers in Biology and Medicine*, 37, 957–964. <https://doi.org/10.1016/j.compbimed.2006.08.018>
- Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections: The Journal for Movies and Mind*, 2, 1–26. <https://doi.org/10.3167/proj.2008.020102>
- Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, 16, 219–222. <https://doi.org/10.1111/j.1467-8721.2007.00507.x>
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269–293). Amsterdam: Elsevier. ISBN 9780080433615
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243–271. <https://doi.org/10.1146/annurev.psych.50.1.243>
- Hochberg, J., & Brooks, V. (1978). *Film cutting and visual momentum*. Oxford: Oxford University Press. ISBN 9780195176919
- Hu, P. Y. (2013). *The influence of digital media on the reality of film space* [Unpublished doctoral dissertation]. University of National Chiao Tung. <http://140.113.39.130/cdrfb3/record/nctu/#GT079216821>
- Hutson, J. P., Smith, T. J., Magliano, J. P., & Loschky, L. C. (2017). What is the role of the film viewer? The effects of narrative comprehension and viewing task on gaze control in film. *Cognitive Research*, 2, 46. <https://doi.org/10.1186/s41235-017-0080-5>
- Jesionowski, J. E. (1989). *Thinking in pictures: Dramatic structure in D.W. Griffith's biograph films*. California: University of California Press. ISBN 9780520067929
- Luan, F. L. (2018). Film perceptive and cognitive development: Past, present, and future. *Journal of Beijing Film Academy*, 3, 98–104. <http://bjdyxyxb.izhengao.com>
- Ma, M. Y., & Chuang, H. C. (2015). How form and structure of Chinese characters affect eye movement control. *Journal of Eye Movement Research*, 8, 1–12. <https://doi.org/10.16910/jemr.8.3.3>
- Ma, M. Y., & Chuang, H. C. (2017). An exploratory study of the effect of enclosed structure on type design with fixation dispersion: Evidence from eye movements. *International Journal of Technology and Design Education*, 27, 149–164. <https://doi.org/10.1007/S10798-015-9342-Z>

- MackWorth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, 2, 547–552. <https://doi.org/10.3758/BF03210264>
- Marchant, P., Raybould, D., Renshaw, T., & Stevens, R. (2009). Are you seeing what I'm seeing? An eye tracking evaluation of dynamic scenes. *Digital Creativity*, 20, 153–163. <https://doi.org/10.1080/14626260903083611>
- Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5, 229–240. <https://doi.org/10.1038/nrn1348>
- Megaw, E. D., & Richardson, J. (1979). Eye movements and industrial inspection. *Applied Ergonomics*, 10, 145–154. [https://doi.org/10.1016/0003-6870\(79\)90138-8](https://doi.org/10.1016/0003-6870(79)90138-8)
- Mercado, G. (2013). *The filmmaker's eye: Learning (and breaking) the rules of cinematic composition*. Oxford: Focal Press. ISBN 9780240812175
- Monaco, J. (2000). *How to read a film: The world of movies, media, and multimedia: Language, history, theory*. Oxford: Oxford University Press. ISBN 9780195139815
- Münsterberg, H. (1916). *The photoplay: A psychological study*. Montana: Kessinger Publishing. ISBN 9781436581578
- Murch, W., & Coppola, F. F. (2001). *In the blink of an eye: A perspective on film editing*. Los Angeles: Silman-James Press. ISBN 9781879505629
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107–123. [https://doi.org/10.1016/s0042-6989\(01\)00250-4](https://doi.org/10.1016/s0042-6989(01)00250-4)
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Salvucci, D. D., & Anderson, J. R. (1998). *Tracing eye movement protocols with cognitive process models*. Hillsdale, NJ: Lawrence Erlbaum Associates. ISBN 9780805832310
- Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections*, 6, 1–27. <https://doi.org/10.3167/proj.2012.060102>
- Smith, T. J. (2013). Watching you watch movies: Using eye tracking to inform cognitive film theory. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies* (pp. 165–190). Oxford: Oxford University Press. ISBN 9780199862139
- Smith, T. J., & Henderson, J. M. (2008). Attentional synchrony in static and dynamic scenes. *Journal of Vision*, 8, 773. <https://doi.org/10.1167/8.6.773>
- Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behaviour in static and dynamic scenes. *Journal of Vision*, 13, 16. <https://doi.org/10.1167/13.8.16>
- Tang, D. L., & Lai, S. C. (2016). Exploring the impact of film's rhythm and viewer's construal level on the film viewing behavior: Evidence from an eye-tracking methodology. *Research in Arts Education*, 32, 69–101. <https://doi.org/10.6622/RAE.2016.32.03>
- Tatler, B. W. (2014). Eye movements from laboratory to life. In M. Horsley, M. Eliot, B. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research* (pp. 17–35). Cham: Springer. ISBN 9783319028675
- Treuting, J. (2006). Eye tracking and cinema: A study of film theory and visual perception. *SMPTE Motion Imaging Journal*, 115, 31–40. <https://doi.org/10.5594/J11499>
- Vassallo, S., Cooper, S. L., & Douglas, J. M. (2009). Visual scanning in the recognition of facial affect: Is there an observer sex difference? *Journal of Vision*, 9, 1–10. <https://doi.org/10.1167/9.3.11>
- Vig, E., Dorr, M., & Barth, E. (2009). Efficient visual coding and the predictability of eye movements on natural movies. *Spatial Vision*, 22, 397–408. <https://doi.org/10.1163/156856809789476065>
- Wang, T. N., & Jian, Y. C. (2022). A systematic review of eye-tracking studies on text-diagram science reading. *Bulletin of Educational Psychology*, 53, 773–799. [https://doi.org/10.6251/BEP.202206\\_53\(4\).0001](https://doi.org/10.6251/BEP.202206_53(4).0001)
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Springer. ISBN 9780306302985
- Zhou, L., Zhang, Y. Y., Wang, Z. J., Rao, L. L., Wang, W., Li, S., ... Liang, Z. Y. (2016). A scanpath analysis of the risky decision-making process. *Journal of Behavioral Decision Making*, 29, 169–182. <https://doi.org/10.1002/bdm.1943>

