Eurographics Workshop on Intelligent Cinematography and Editing (2017)
W. Bares, V. Gandhi, Q. Galvane, and R. Ronfard (Editors)

# Analyzing Elements of Style in Annotated Film Clips

Hui-Yin Wu[1], Quentin Galvane[2], Christophe Lino[1], Marc Christie[1]

[1] Inria, Université de Rennes 1, France
[2] Technicolor Rennes, France

**Abstract**

*This paper presents an open database of annotated film clips together with an analysis of elements of film style related to how the shots are composed, how the transitions are performed between shots and how the shots are sequenced to compose a film unit. The purpose is to initiate a shared repository pertaining to elements of film style which can be used by computer scientists and film analysts alike. Though both research communities rely strongly on the availability of such information to foster their findings, current databases are either limited to low-level features (such as shots lengths, color and luminance information), contain noisy data, or are not available to the communities. The data and analysis we provide open exciting perspectives as to how computational approaches can rely more thoroughly on information and knowledge extracted from existing movies, and also provide a better understanding of how elements of style are arranged to construct a consistent message.*

## 1. Introduction

In the context of virtual cinematography – an attempt to adapt elements of cinematography from real movies into virtual environments – many approaches rely on the formalization of empirical pieces of knowledge listed as recipes in filmmaking textbooks. However, practical evidence shows that to capture and reproduce elements of cinematographic style requires more than arranging rules and conventions. In this perspective, extracting and adapting quantified elements of film style from existing movies appears both as a challenge, and as a mean to provide more expressive approaches to virtual cinematography. However, despite a strong interest from research communities in film analysis, it appears there are limited reliable sources on which to perform data extraction and analysis.

In this paper, we propose an open database of annotated film clips together with an analysis of elements of style related to shot composition, shot transition, and shot arrangements. The database currently encompasses 22 clips (from 18 popular movies, released from 1966 to 2012, and covering well-known styles such as comedy, drama, war, heroic-fantasy, science-fiction and western) for a total of more than 1000 shots. The database reports shot information relative to the number of characters appearing on the screen, their on-screen positions and head orientations, the type of camera shot (ranging from extreme long shot to extreme close-up shot), the type of camera angle (high, medium, low), the type of camera motion, by relying on features defined in the Patterns language proposed in [WC16].

Using our film database, we here report three analyses we have conducted. First, we analyzed the staging and framing properties in the shots, i.e. how filmmakers are staging actors (spatial arrangement in the scene), how they are composing their shots (arrangement of actors on the screen), and how the staging can influence their choices in term of framing. Second, we analyzed transition properties between two consecutive shots (*cuts*), i.e. how they make choices with regards to the next shot to cut to, depending on the current shot properties. Third, we analyzed transition properties along sequences of shots (*editing patterns*), i.e. how composition properties within shots evolve over a sequence of shots.

The paper is organized as follows. After reviewing related works (Section 2), we detail how our dataset has been collected from film clips (Section 3). We then report the results of our three analyses: staging and framing properties (Section 4), transition properties (Section 5), and editing pattern properties (Section 6). We finally conclude by discussing our results and perspectives (Section 7).

## 2. Related Work

The analysis of film style is a central concern in film literature through the objective of understanding the complex relationship between *appearance* – how technical elements of film are arranged – and *meaning* – what is the impact, emotional and cognitive, of a film sequence. Film literature offers a great source of knowledge. Text books from David Bordwell for example, offer insights on elements of film style through time [Bor98] and on staging [Bor05].

The precise annotation of technical elements such as frame composition, shot size and shot lengths offers a complementary an quantified point of view, for example in studying how such parameters evolve over the history of movies [Sal83, Cut15], and help



DOI: 10.2312/wiced.20171068

in identifying long-term trends in style such as *intensified continuity* [Bor02]. More detailed elements of film composition have been reported by studies of James Cutting. Based on manual annotation of characters head positions, the authors report the distribution of these positions with relation to shot size, and number of characters on the screen [Cut15]. Most of the annotated data is however not released, or not formatted in ways that can easily be exploited. Some tools, such as Cinemetrics (www.cinemetrics.lv/), provide means to collect and share the data. The process however makes it difficult to exploit the data directly due to non-guided annotations without an agreed terminology, for example on elements such as shot size.

While multiple tools dedicated to media annotations exist (see Elan (tla.mpi.nl/tools/tla-tools/elan/) and Anvil (www.anvil-software.org/), these are not designed in mind to annotate elements of camerawork, staging or lighting. The "Ligne de Temps" tool (www.iri.centrepompidou.fr/outils/lignes-de-temps/) offers a specific film annotation process, but currently lacks the ability to extend the tool with dedicated annotation components.

With the quick evolution in computer vision techniques, a number of tasks performed manually in the past can be semi or fully automated. Relying on such results, approaches such as [CBL13] can classify shot sizes, but also detect specific shot angles such as over-the-shoulder shots [SBA*15].

Relying on this wealth of information, approaches to virtual cinematography have used learning techniques to reproduce elements of film editing [MCB15]. Interestingly languages have been proposed to support tasks like annotation, and also generation of film sequences, encompassing elements of shot description [RGB13], and properties of shot transtion [WC15].

## 3. Dataset structure and annotation

Our database currently encompasses 22 film clips extracted from 18 popular movies (detailed in Table 1). The period in which these films were released spans from 1966 (*The Good, the Bad, and the Ugly*) to 2012 (*The Hunger Games*). The clips were chosen from famous scenes that users uploaded to youtube, which indicates the clips' popularity and interest to the general public.

We manually annotated all film clips by relying on the *Insight* annotation tool introduced in [MWS*15]. We annotated each shot of the clips, and each key framing inside shots. Information gathered on each framing includes:

- the names of actors appearing on the screen;
- the position of their head centers, their scale (relative to the screen height), and their orientation (in azimuth relative to the camera, and in elevation when available);
- the positions of their eyes in the frame (when available);
- the properties over their feet, hands, or any other important element of the actors when their head and eyes are not on-screen;
- other important inanimate objects appearing on-screen, that are crucial to the plot;
- the shot size on a scale of 9 sizes (estimated);
- the vertical shot angle on a scale of 7 angles (estimated);

The annotations are represented in an XML schema in which each video clip has its own associated file. Each file contains header information concerning the characters in the clip and video information. Then the annotation is split into framings (a single still frame) and shots (a continuous sequence of framings). For each shot, at least one framing was selected to be the "keyframe" that is annotated with the information detailed above.

The full database is openly available on Github: https://github.com/husky-helen/FilmAnnotation.

| Title | Year | Dur | No.Shots | Avg.Shot | Type |
|---|---|---|---|---|---|
| Good, Bad, Ugly | 1955 | 6m16s | 98 | 3.83 | Action |
| Godfather | 1972 | 1m19s | 30 | 2.63 | Action |
| The Shining | 1980 | 3m09s | 5 | 37.8 | Dialogue |
| American History X | 1988 | 3m57s | 57 | 4.16 | Dialogue |
| Forrest Gump | 1994 | 2m20s | 17 | 7.72 | Mixed |
| Pulp Fiction | 1994 | 2m07s | 28 | 4.34 | Dialogue |
| Shawshank(1) | 1994 | 3m32 | 48 | 4.41 | Action |
| Shawshank(2) | 1994 | 3m09s | 26 | 7.26 | Action |
| Gattaca(1) | 1997 | 1m52s | 19 | 5.89 | Dialogue |
| Gattaca(2) | 1997 | 1m59s | 19 | 6.21 | Dialogue |
| Armageddon | 1998 | 3m10s | 100 | 1.88 | Action |
| Lord of the Rings | 2001 | 3m33s | 56 | 3.67 | Mixed |
| Infernal Affairs | 2002 | 3m05s | 47 | 3.91 | Dialogue |
| Big Fish | 2003 | 2m10s | 47 | 2.77 | Mixed |
| Constant Gardener | 2005 | 3m33s | 64 | 3.33 | Mixed |
| Benjamin Button | 2008 | 3m10s | 78 | 2.44 | Action |
| Departures | 2008 | 5m28s | 62 | 5.29 | Action |
| Invictus | 2009 | 1m29s | 22 | 4.04 | Mixed |
| The Help(1) | 2010 | 2m50s | 28 | 6.07 | Dialogue |
| The Help(2) | 2010 | 3m39 | 55 | 3.96 | Dialogue |
| Hunger Games(1) | 2012 | 3m59s | 101 | 2.37 | Mixed |
| Hunger Games(2) | 2012 | 22s | 11 | 2 | Action |

**Table 1:** *Overview of annotated datasets. The average shot duration is represented in seconds.*

The cumulative duration of all our clips exceeds one hour. There are a total of 1184 different framings from 1018 shots. Shots with more than a single framing are those where we have annotated the framing at different key time-frames due to some camera motions inside the shot. Table 1 summarizes the details of each film clip: title, release year, clip duration, number of shots, average shot duration (in seconds), and a categorization for the scene type (either action, dialogue, or mixed). The films are sorted by increasing order of release year.

Our *Insight* annotation tool (Figure 1) provides the 3D model of a human head that can be manually rotated, moved, and scaled to each actor in the annotated framing, and the position, scale, and

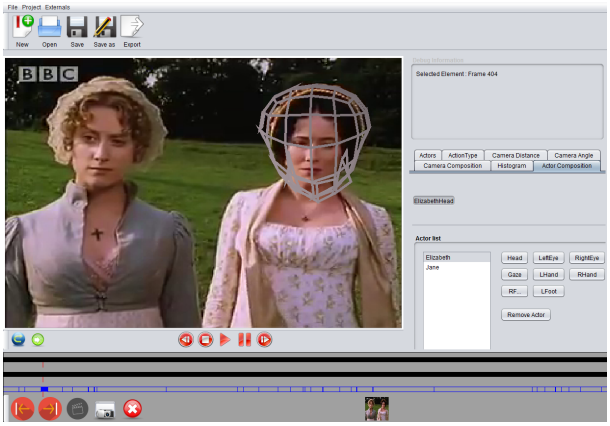
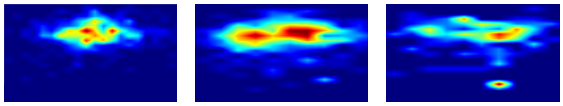


**Figure 1:** *A screen capture of the Insight annotation tool showing the on-screen head tool to estimate actor head positions, orientations, and sizes. The clip shown is from Pride and Prejudice*



**(a)** *One actor* **(b)** *Two actors* **(c)** *Groups of actors*

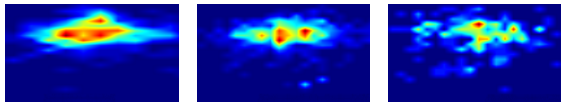
**Figure 2:** *Density of head positions depending on the number of filmed actors.*

rotation parameters are automatically calculated and output. On-screen positions of other body parts or objects are achieved through cursor clicks on the framing for each target. The shot size and angle are estimated based on the scale introduced in [WC16], though shot size can also be assigned automatically based on the relative head size on the screen. Each shot is then annotated with a shot movement feature, such as zoom, pan, dolly, or tilt. Most shots in the database are static shots, which reflects general film practice.

## 4. Staging and Framing

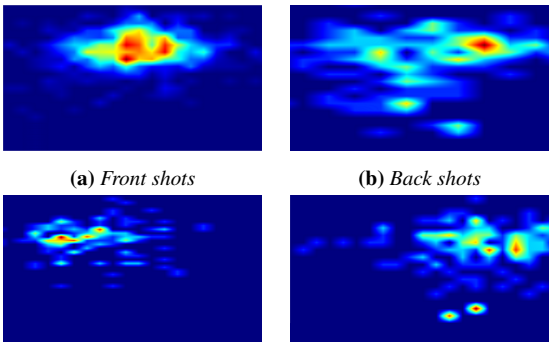
Decisions regarding the staging and framing of actors can be both motivated by narrative goals related to the overall story and/or by aesthetic reasons. Film directing is a combination of directing actors so that they play certain actions at precise locations in a scene, and of directing (i.e. placing and moving) cameras relative to both these actors and the scene to obtain a visual layout (or framing) that best convey the director's vision of the story. In other words, all together, what appears on the screen and how it appears in terms of size, position, and angle have a strong influence on how the audience will perceive the scene.

In this section, we present results of our analysis on both the staging and framing in shots containing one, two, or groups of actors. Our analysis focuses on the on-screen head location and orientation of actors, as well as the shot size.



**(a)** *Close Shots* **(b)** *Medium Shots* **(c)** *Long Shots*

**Figure 3:** *Density of head positions depending on the shot size.*



**(a)** *Front shots* **(b)** *Back shots*

**(c)** *Right profiles* **(d)** *Left profiles*

**Figure 4:** *Density of head positions depending on the actor's head orientation (or* profile*).*

### 4.1. Head locations

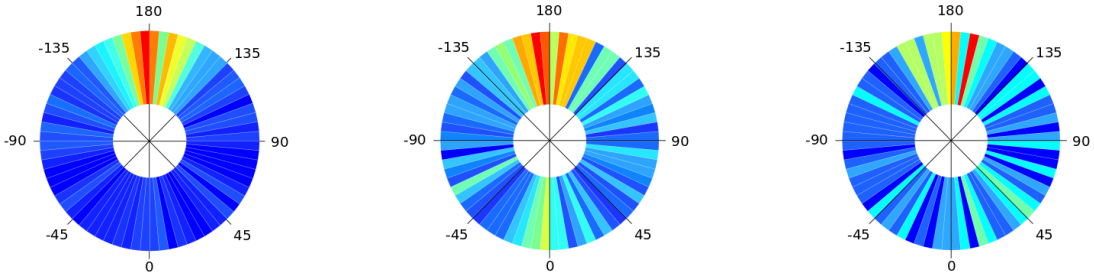
In a shot, the audience first notices the head of the actors before observing other elements of the scene presented. Thus, the on-screen position where the head appears has its own storytelling and emotional function.

We analyze the position of actors' heads in our dataset, with regards to three criteria: (i) the number of filmed actors, (ii) the shot size and (iii) the actor's head orientation (or *profile*). Our results are illustrated in Figures 2, 3 and 4.

On average, we find that actors are placed in the top third of the frame, and actors' positions are balanced horizontally: a single actor is placed in the middle, two actors are placed at each third, and for three actors or more the common practice seems to balance them horizontally along the top third. Looking at our results for three or more actors, our assumption is also that the main actors are often placed in the right, either on the top or bottom third of the frame (Figure 2). These results correspond to common framing practice in filmmaking known as rule of thirds for optimal framing.

When not considering the number of actors, the shot size does not seem to be a dominant factor on placing actors. In average, actors are still placed in the top third of the frame, with a Gaussian distribution around the middle on the horizontal axis (Figure 3).

The orientation of actors seem to have a higher influence on the distribution of actors in the frame. When an actor is viewed from front, she is centered in the top of the frame. When she is framed from back, she is still placed on the top of the frame, but on either side horizontally, following the rule of thirds. When she is viewed from the side, she is placed in the side opposite to her gaze direction, which follows a very common framing rule (*gaze-room*) and is also a common practice used to create an opposition between




**(a)** *One actor* **(b)** *Two actors* **(c)** *Group of actors*

**Figure 5:** *Distribution of characters orientations relative to the camera*

two actors in a dialogue scene – shooting both actors either from over-the-shoulder of one of them or from an apex view, or cutting between shots of each actor separately – (Figure 4).

### 4.2. Head orientations

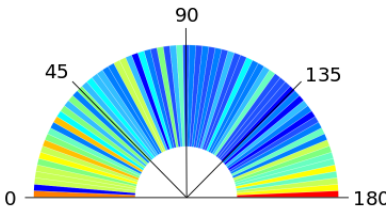



**Figure 6:** *Distribution of differences in actors orientations*

The orientations and gaze directions of filmed actors strongly drives the viewers' reading of a shot content and her expectations on the content of the next shot. For instance, if an actor is looking at another element inside a shot, the viewer will naturally follow the actor's gaze to look at this element. If the actor is looking at an element located outside the screen, the viewer will also often expect this element to appear in the next shot.

By observing how actors are staged, through their relative orientation in the shot (Figure 6), we find that they are most often placed so that they look in the same direction or in opposite directions. In the latter, though we cannot distinguish cases where they are back-to-back, following common film practices in dialogues (which represent the majority of shots with two actors) we can here make the assumption that they might be most often facing each other.

Figure 5 details our findings regarding the actors' orientation depending on how many of them appear on the screen. When filming only one actor, the actor is most often filmed from a front, or sometimes three-quarter front view, but more rarely with a side or back view (Figure 5a). This corresponds well to static shots (which represent a large percentage of our dataset), where neither the actor nor the camera is moving. This is often motivated by the need to film an actor when he is speaking or reacting to another event, and where focusing on facial expressions is thus a crucial element of the story.

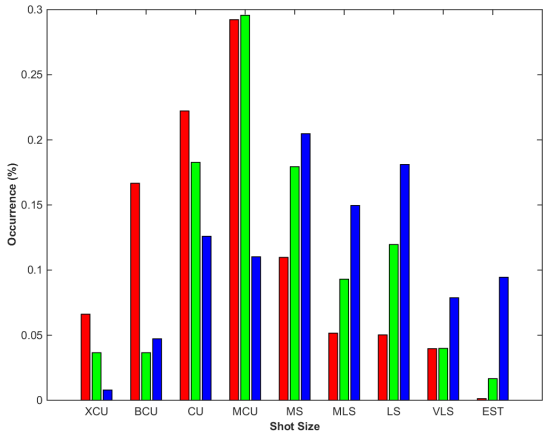
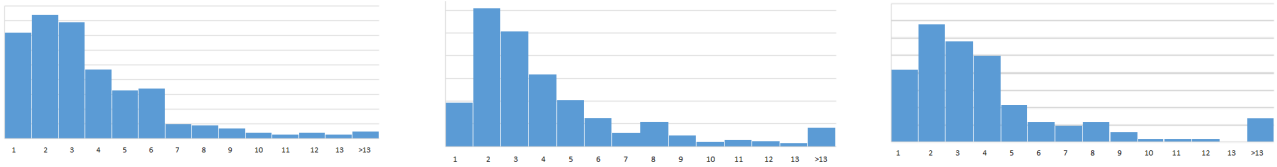



**Figure 7:** *Distribution of shot sizes for shots with one actor (red), two actors (green), and three or more actors (blue), ranging from extreme closeups (XCU) to very long shots (VLS) and establishing shots (EST).*

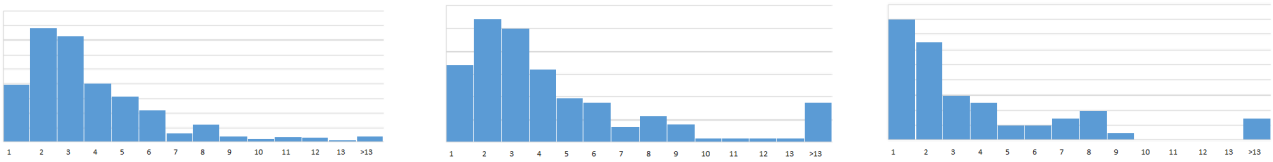
When filming two actors, there are a large number of shots filmed from a front, three-quarter front, or back view, and fewer filmed from a side view (Figure 5b). This corresponds well to common practice in filming dialogues through over-the-shoulder shots (often combined with shot-reverse-shot edits) in a wide proportion, and apex shots (perpendicular to the actors' line of action) in a more limited proportion.

When filming groups of actors, the common practice seems to show that most actors should be looking toward the camera, i.e. viewed from either front or three-quarter front, and some of them in other directions it, i.e. viewed from any orientation (Figure 5c). Our assumption is that the face of protagonists should be visible, while secondary actors might be viewed from any angle, as they are only providing some knowledge about the context in which the action occurs.



**(a)** *Close shots* **(b)** *Medium shots* **(c)** *Long shots*

**Figure 8:** *Distribution of shot durations for various shot sizes*



**(a)** *One Actor* **(b)** *Two Actors* **(c)** *Group of Actors*

**Figure 9:** *Distribution of shot durations for various number of actors*

### 4.3. Shot size

The size of a shot (i.e. how much the filmed actors are filling the image) is an important feature in the viewers' understanding of the unfolding actions. Filming an actor from a close shot (where only her head appears on the screen) will enable conveying details such as subtle facial expressions or hand gestures, while filming from a long shot (containing the whole body of the actors and their surrounding) will enable conveying the context in which an action occurs. Filming an action from intermediate medium shots will enable to make a balance, to not only conveying some (not subtle) details on the action but also its (close) context. Our dataset contains annotations on a scale of 9 shot sizes. The three sizes extreme close-up (XCU), big close-up (BCU), and close-up (CU) make up the category of *Close* shots; medium close-up (MCU), medium shot (MS), and medium long shot (MLS) make up the *Medium* category; finally, long shot (LS), very long shot (VLS), and establishing shot (EST) make up the *Long* shot category.

We analyze the distribution of shot sizes in function of how many actors appear on the screen. As illustrated in figure 7, this distribution seems to follow a Gaussian or log-normal law, where the mean shot size is either a MCU or a MS. Further, as expected, there is an increase in the proportion of medium and long shots as the number of actors increases. Symmetrically, there are also more often close shots when filming only one actor.

## 5. Transitions

From single shots, we move onto the analysis of shot transitions by studying the stylistic choices directors and film editors make to put shots into a coherent storyline. In our analysis, we start by only considering transitions between two shots (i.e. cuts), which are another key component influencing the viewers' perception of the story. Indeed, performing a cut from a shot to another shot provides directors with means to introduce explicit or implicit links between the contents of these two shots. Two additional stylistic choices offered to directors are the consecutive shot sizes and the amount of time spent in each shot, which will have a strong impact on the amount of information viewers will be able to learn in these shots, and on the general emotion conveyed throughout the film. We here analyze two stylistic choices: the cutting rhythm (in terms of the mean shot duration), and the cutting preferences (in term of the shot sizes of two consecutive shots).

### 5.1. Cutting rhythm

The duration of the shot affects the rhythm of the scene. Quick successive shots create a fast-paced and tense emotion (which director are using more often in action scenes), whereas shots that take more than a few seconds allows the audience to slow down and observe the scene (e.g., directors are often making use of such slow-paced shots to convey love stories). The decision on how long a shot should be can also be strongly related to the amount of information in the shot. We here analyze the duration of shots along two criteria: (i) the number of filmed actors and (ii) the shot size.

From Figures 8 and 9, we can observe that both the shot size and the number of filmed actors have an impact on the distribution of shots duration (in seconds).

For medium and long shots, the overall distribution seems to generally follow a log-normal law, with some shots of less than 2 seconds, a maximum number of shots that last for around 2 or 3 seconds, and then a slow decrease in the number of shots with a few shots of more than 10 seconds. The distribution of long shots also span over more time, probably because viewers need more time to browse their (more complex) content. In proportion, there are also more long shots of one second or less. Our assumption is that these ones are essentially establishing shots, only providing the overall context of the scene but not much details on the actions. For close shots, the shots are often shorter, and their distribution is closer to a Gaussian with a moderate standard deviation. Most shots last between 1 and 3 seconds, less shots last between 4 and 6 seconds, and few of them last more than 6 seconds.

For shots with only one or two filmed actors, the distribution

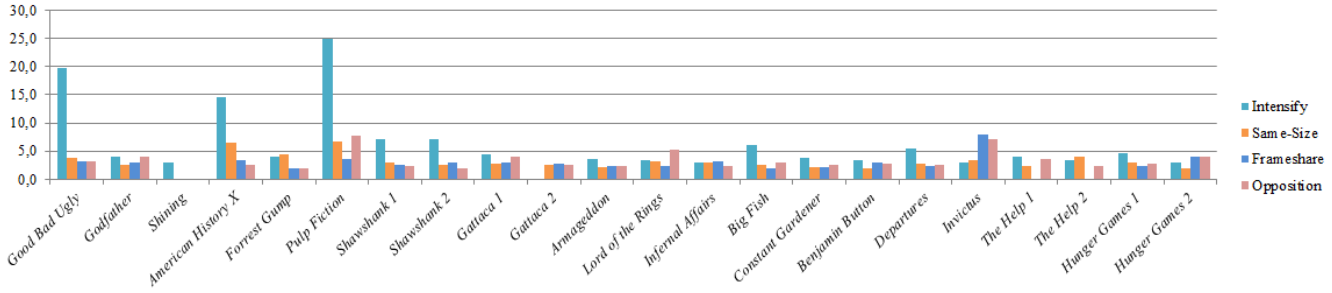



**Figure 10:** *Average length (in seconds) for all ECPs (embedded constraint pattern) used through shot sequences in the film clips.*

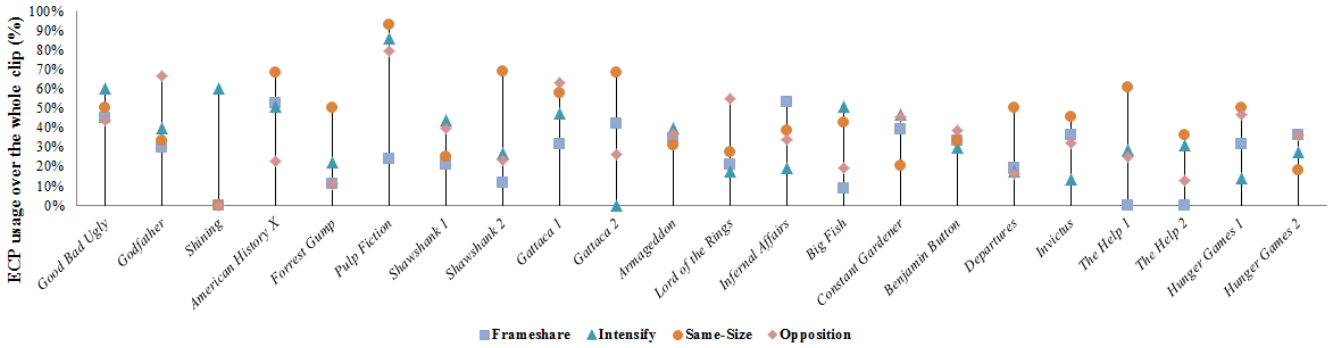



**Figure 11:** *Coverage of each ECP (embedded constraint pattern) in the films clips. Each dot on the histograms represents the percentage of shots that a pattern is using throughout a given film clip.*

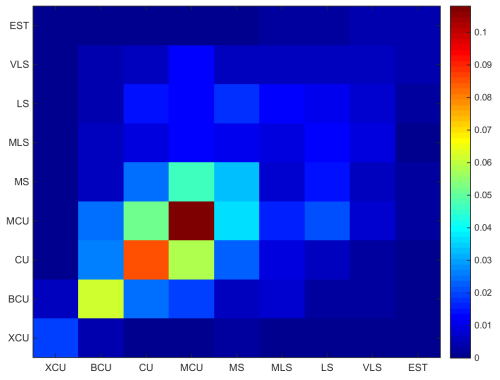



**Figure 12:** *Transition matrix showing the distribution of cuts from a given shot size (abscissa) to another (ordinate).*

follows a log-normal distribution with a mean duration around 2 seconds. Conversely, for groups of three or more actors, shots are most often very short, there is a very fast decrease in the number of shots between 2 and 4 seconds, a few shots last between 5 and 9 seconds. This could indicate that very often groups shots are very similar to long shots in the way they are used to convey the context of an action. Further, both two-shots and group shots rarely last more than 10 seconds, unless they are much longer (more than 13 seconds), which we could also interpret as a way to convey a set of consecutive actions through a single shot.

### 5.2. Size transitions

As we have seen from the analysis of staging, there is much insight to be gained through the analysis of the size of the actor on-screen. Shot size is also a element frequently used to express story details not just in single shots, but also across long sequences.

Figure 12 shows the distribution of shot size transitions between any two shots. From Figure 7 we find that there were the most medium close-up shots (MCU) as compared to any other shot type. It is thus understandable, that we find more instances of two consecutive shots with at least one of them being a MCU. From the transition matrix, we can further conclude that transitions are more often performed between shots of the same size or within the distance of two units of size. For example,it is very uncommon to cut from a close shot to a long shot. Instead, directors tend to first make a transition to a close or medium shot. In the same way, it is uncommon to directly cut from an establishing or long shot to a close shot. As shown previously, transitioning through a medium shot seems more common. As well, after establishing the scene, it seems uncommon to re-establish it (i.e. cuts to an establishing shot are very rare in our data).

## 6. Patterns

In actual film practice, visual elements within a single shot, and from one shot to another, are stylistic aids directors often use to express space-time logic over a whole action, event, or scene. The arrangement of these visual elements – shot sizes, angles, and regions – in a recurring and repetitive manner across multiple shots is also another stylistic technique directors and film editors use.

In this section, we use ECPs as a means to analyze film style. An ECP, embedded constraint pattern, is defined in [WC16] as a way of defining constraints on recurring elements of style over a long sequence of shots, such as sequence of shots where the shot size does not change, or a sequence of 1-actor shots where all the actors appear in the center of the framing. We selected four ECPs introduced in the same paper, and analyzed the database for occurences of these ECPs. We analyzed four ECPs, described as follows:

- **Same-Size**: all shot sizes in the sequence are the same;
- **Intensify**: shot sizes in the sequence become closer and closer;
- **Frameshare**: all shots are filming a single actor, where all successive actors appear in the same horizontal region (either left or right) throughout the whole sequence;
- **Opposition**: similar to Frameshare, but where actors appear in opposite horizontal regions (left vs. right) in every new shot;

The usage of ECPs in film clips allows us to make quantitative observations such as on the evolution of average ECP length (in terms of the number of shots in the sequence) over these film clips such as in Figure 10 or the usage of these techniques over the entire dataset, as shown in Figure 11. This is different from the case study of *Lord of the Rings* from [WC16] in that we can compare how ECPs are used.

From the data, we can observe that shot sizes are a common tool used in long sequences as a way to convey a consistent emotion throughout the clip. On average, intensify and same-size sequences can be much longer than any other ECP, applied to more than 5 shots in the sequence, as seen in *The Good, the Bad, and the Ugly*, *American History X*, and *Pulp Fiction*. In the clip from *Pulp Fiction*, three types of sequences (intensify, same-size, and opposition) also cover more than 80% of the shots in the clip, which clearly shows that all three techniques can be used in parallel throughout a whole scene.

## 7. Discussions

The benefit of analysis of films through large amounts of data can be seen from many different aspects. First, and foremost, observing empirical data is a great way to verify concepts and experience introduced in film educational material. This can particularly be seen from our analysis of head locations of actors in the framing, which reveals the popular use of the rule-of thirds as a framing guideline across the wide variety of genres and director styles. From this aspects, film analysis has an application in education, opening possibilities to learn not just from theory or past experience, but also from easily accessible and search-able data.

The second benefit lies in applications to online video streaming services. At the core of the development of accurate video recommendation systems is large amounts of data that is annotated and analyzed using machine learning techniques to discover the preference of users. Apart from the story or actor fame itself, visual style, lighting, and music have long been popular elements used to categorize movies into genres, and recommend suitable films and videos to users.

Finally, there is the potential of data in use for creative applications, such as in film pre-visualization tools, or editing of scenes and cinematography in virtual environments, which is a topic that has gained continued interest in the fields of graphics and animation.

In this paper, we have limited our analysis to the specific features of shot sizes, angles, actor orientations, and positions in single shots and across shot sequences. However, with more developed tools that can observe elements of color, lighting, motion, or combined with exterior data such as movie scripts or viewer reviews, we see a strong potential in future tools and algorithms for film analysis that can, for example, more accurately identify genres, directorial styles, or applied to generative contexts, select suitable framing and camera movement styles to express a specific story context.

## References

[Bor98] BORDWELL D.: *On the History of Film Style*. Harvard University Press, January 1998. 1

[Bor02] BORDWELL D.: Intensified continuity: Visual style in contemporary american film. 2

[Bor05] BORDWELL D.: *Figures Traced in Light: On Cinematic Staging*. 2005. 1

[CBL13] CANINI L., BENINI S., LEONARDI R.: Classifying cinematographic shot types. *Multimed Tools Appl 62*, 1 (2013). 2

[Cut15] CUTTING J.: The framing of characters in popular movies. *Art & Perception 3*, 2 (2015). 1, 2

[LC12] LINO C., CHRISTIE M.: Efficient composition for virtual camera control. In *Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation* (2012).

[LC15] LINO C., CHRISTIE M.: Intuitive and Efficient Camera Control with the Toric Space. *Transactions on Graphics 34*, 4 (2015).

[MCB15] MERABTI B., CHRISTIE M., BOUATOUCH K.: A Virtual Director Using Hidden Markov Models. *Computer Graphics Forum* (2015). 2

[MWS*15] MERABTI B., WU H.-Y., SANOKHO C. B., GALVANE Q., LINO C., CHRISTIE M.: Insight : An annotation tool and format for film analysis. In *Eurographics Workshop on Intelligent Cinematography and Editing, May 2015, Zurich, Switzerland* (2015), p. 1. 2

[RGB13] RONFARD R., GANDHI V., BOIRON L.: The Prose Storyboard Language: A Tool for Annotating and Directing Movies. In *FDG Workshop on Intelligent Cinematography and Editing* (2013). 2

[Sal83] SALT B.: *Film Style and Technology: History and Analysis*. Performing Arts, 1983. 1

[SBA*15] SVANERA M., BENINI S., ADAMI N., LEONARDI R., KOVÃĄCS A. B.: Over-the-shoulder shot detection in art films. In *CBMI* (2015), IEEE, pp. 1–6. 2

[WC15] WU H.-Y., CHRISTIE M.: Stylistic Patterns for Generating Cinematographic Sequences. In *Eurographics Workshop on Intelligent Cinematography and Editing* (2015). 2

[WC16] WU H.-Y., CHRISTIE M.: Analysing Cinematography with Embedded Constrained Patterns. In *Eurographics Workshop on Intelligent Cinematography and Editing* (2016). 1, 3, 7