CARNEGIE MELLON UNIVERSITY

School of Architecture

College of Fine Arts

Thesis

Submitted in Partial Fulfillment of the requirements for the degree of

Masters of Science in Computational Design

TITLE:

Dance Videography Grammar:

A Shape Grammar for the Description, Analysis And Generation Of Dance Performance Videos

AUTHOR:

Tsz Wing Clover Chau

ACCEPTED BY ADVISORY COMMITTEE:

Vernelle Noel		May 7th 2025
Professor NAME	Principal Advisor	DATE
Eddy Man Kim		May 6th 2025
Professor NAME	Advisor	DATE
Daragh Byrne		May 6th 2025
Professor NAME	Advisor	DATE

Dance Videography Grammar:

A Shape Grammar for the Description, Analysis and Generation of Dance Performance Videos

Tsz Wing Clover Chau

School of Architecture
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213

Thesis Committee:
Vernelle Noel
Eddy Man Kim
Daragh Byrne

ABSTRACT

Before video recording became the most prevalent method of reproducing dance, written systems like Labanotation, Benesh Movement Notation and Eshkol-Wachman Movement Notation used symbols to record dance. These systems tackle the greater theme of movement literacy as posited by Brenda Farnell, and articulate medium-specific concepts for new paradigms of human movement.

The proliferation of mobile phone cameras has allowed anyone regardless of skill level to make videos. Thus, videography has emerged as an experimental field with no centralised pedagogy or institutions, where the tacit technical knowledge is obtained through experience and the paradigms are difficult to communicate and teach.

As a sub-genre of videography, dance videography encompasses the creation of 'dance films' or 'dance performance videos' through the recording of dance and corporeal movement using a video camera. It draws heavy influence from commercial music video techniques (Pajala-Assefa) as well as traditional film principles to highlight the choreography.

This thesis aims to develop a notation that describes how a dance video can be filmed. By breaking down and analysing a corpus of [*Performance37*] dance videos, this work aims to develop a grammar that can extend applications of existing videographic styles and develop new ones. To demonstrate the efficacy and generative capabilities of the proposed workflow, a dance video was produced and evaluated following the derived rules.

ACKNOWLEDGEMENTS

I would like to thank Vernelle Noel, Eddy Man Kim and Daragh Byrne for their invaluable support and advice as my advising committee. From the first year of my undergraduate degree until now, your guidance was crucial in enabling me to make it this far. I would also like to thank Pedro Veloso, Emek Erdolu, Ardavan Bidgoli, Jinmo Rhee and the MSCD PhDs who showed undergraduate Clover so much patience, and whose work inspired me to enter the MSCD program in the first place.

Thank you to Geriatrics Crew, for being amazing friends who gave me both emotional and academic support. From dancing to filming to lighting design, you support me in all my endeavors and give me the grace to be better on my bad days. Your competence and intelligence inspires me to keep pushing forward with my work.

Special thanks to abtech for providing such a supportive and healthy environment, that gave me a space to grow and take shelter throughout the semester. Your words enabled me to stand up against mistreatment and advocate for myself in multiple areas of my life, in a way I would never have the audacity to before.

I would also like to extend my gratitude to Steve, Elliott, Kenrick and the West Coast Collective for explaining the basics of parsing music and audio analysis. While I was not able to fully explore them here, they will serve as invaluable cornerstones for future work.

Lastly, shout out to my Studio children - Jiaxin, Eliora, Amika, Tiffany, Aastha, Russell - who made my first and last time teaching Architecture Studio so fulfilling. I am so proud of you and look forward to seeing the architects and designers you will become.

TABLE OF CONTENTS

ABSTRACT 1

ACKNOWLEDGEMENTS 2

TABLE OF CONTENTS 3

1 Introduction 6

1.1 Background 8

1.1.1 Videography 8

1.1.1.1 Film Industry Tools - Storyboard & Shot List 10

1.1.1.2 Film Industry Tools - Shot List 11

1.1.2 Movement Notation Systems 12

1.1.2.1 Alternative Notation Systems 16

1.1.3 Shape Grammars 17

1.1.4 Visual Analysis 18

1.1.5 Encoding Cinematography 18

1.2 Hypothesis 20

2 Methods 22

2.1 Examining Videographic 'Style' 22

2.2 Building a Corpus 22

2.3 Video Analysis 23

2.3.1 Visual Analysis 24

2.3.1.1 Shots, Cuts, Sequence 24

2.3.1.2 Shot Size 26

2.3.1.3 Shot Angle 27

2.3.1.4 Quantitative Analysis 30

2.3.2. Audio Analysis 37

2.3.2.1 Song Structure 37
2.3.2.2 Cut Distribution 39
2.3.3. Audio-Visual Analysis 40
2.3.3.1 Shot Size Distribution 40
2.4 Developing Vocabulary 42
2.4.1 Naming Convention 42
2.4.2 Shot Lexicon 43
2.5 Corpus Analysis 44
2.5.1 Visual Analysis 44
2.5.1.1 Shape Grammar Vocabulary 46
2.5.2. Audio Analysis 47
2.5.2.1 Song Structure 47
2.5.2.2 Cut Distribution 48
2.5.3. Audio-Visual Analysis 49
2.5.3.1 Shot Size Distribution 49
2.5.3.2 Section Starting & Ending Shot Size 51
2.6 Shape Grammar Ruleset 53
2.6.1 Dance Videography Shape Grammar for *[Performance37]* Style 53
3 Results 56
3.1 Video Production 57
3.2 Evaluation 58
4 Contributions 59
4.1 Limitations & Next Steps 59
5 Conclusion 61
6 Appendix 62
6.1 Transitions 62

6.1.1 Shot Types 62

6.1.2 Angles 66

6.2 Additional Audio Analysis 68

6.2.1 Music as Videographic Driver 68

6.2.2 Defining Beat Conditions 69

WORKS CITED 73

1 Introduction

In *Music Videos: The Look of the Sound*, Pat Aufderheide argues that "Videos are perhaps the most accessible form of that larger tendency known as postmodern art." Music videos are a popular genre of video, allowing artists to craft a visual aesthetic to complement their song that offer "a more engaging, enjoyable, and meaningful media experience than songs alone." (Schlue. 3.) These experiences are identified by Railton and Watson in *Music Video and the Politics of Representation*: pseudo-documentary, art music video, narrative video and staged performance.

There are two active components to the staged performance archetype: 1. Singing/ Rapping and 2. Dancing. A brief glance at Fig.1 shows they can be grouped into the following three categories based on who is performing these actions:

Artist(s) sing / rap and dance	Artist(s) sing / rap only, backup dancers dance	Backup dancers dance only
Tate Mcrae – Greedy 	Eric Nam - Potion 	Justin Bieber - Sorry 
Tate Mcrae – It's Ok I'm OK 	Jessi – Cold Blooded (with SWF) 	Iggy Azalea - Team 
BIBI – Vengeance 		

Fig. 1 - The Three Staged Performance Archetypes

Dancers will also film ‘performance videos’ or ‘concept videos’, where the camerawork and editing of the video is tailored to emphasize their movements. This work will refer to the production of these videos as 'dance videography'.

1.1 Background

This work sits at the intersection of dance, video and computational design. In order to provide a system with adequate specificity and flexibility for such intersectionality, this work reviews relevant literature on the field of videography, precedent movement notation systems, and applications of shape grammars, situating the work in a broader human pursuit of democratizing craft and design.

1.1.1 Videography

Dance videography is a subset of videography that focuses specifically on the recording and presentation of dance. Many of the hallmarks of contemporary dance videography draw from commercial music video techniques like rapid tempo editing and visual effects. (Pajala-Assefa.) Unlike traditional film, where film school offers a pedagogical pipeline to becoming a practitioner, videographers learn by working under an existing practitioner; or more commonly by producing a critical mass of work over a number of years.

Dayna Chun's description of 'video dance' provides a look at the experimental nature of this genre: "There are no rules in this type of art. There is an infinite amount of variables to play with including time, space, and energy… One aspect about video dance that really separates it from the other purposes of recording is its ability to manipulate the audience… With a video dance, the audience can only see and hear what the camera captures. The video points to exactly what the videographer wants the audience to see." (Chun 90)

This work will focus on the following three problems within the field of videography.

1. Tacit technical knowledge

In the film industry, "tacit knowledge is the most prominent form of knowledge… knowledge is bound within the worker, and is often difficult to formalise. Knowledge is subsequently rooted in the action of creating the project's product." (Alony et al. 6) This knowledge is built up over years of expertise, and varies from individual to individual. Acquiring and sharing this knowledge is both time and energy-intensive.

2. No standardized form of communication

Explicating tacit knowledge requires codifying it in a context-specific format. According to Alony et al., "there are major difficulties in the process of explicating knowledge. One of these results from the knowledge owner not being aware or conscious that they possess the knowledge. Another comes from the difficulty people have in communicating the knowledge in an articulate form." (Alony et al. 8)

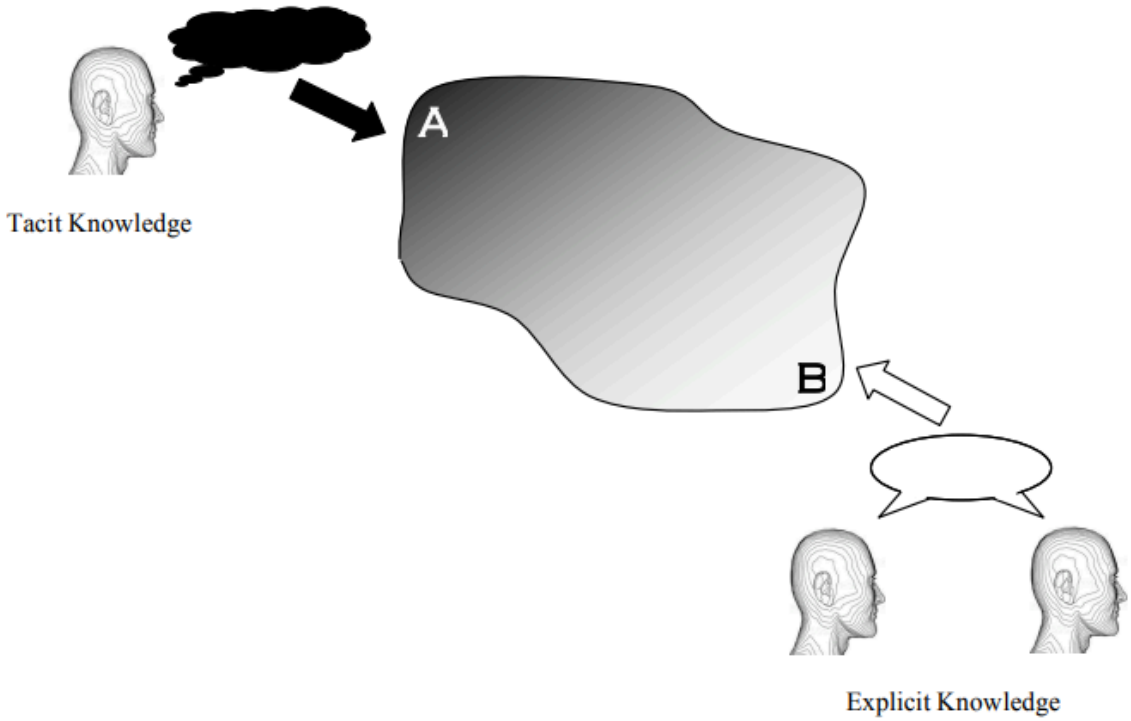


Fig. 2 - The explicability zone of knowledge (Alony et al. 9)

Effective communication between practitioners requires years of collaboration through which they learn how to better communicate with each other. Cinematographer Hyunwoo Nam's comments in *The Making of a K-Pop Music Video* reveal the effort such communication requires: "Besides being friends for a long time, Sungwook [the Director] and I always try to understand and be considerate of each other. It allows us to work together easily." Codifying videographic knowledge requires additional steps, for which there are no current universal systems. (“The Making Of A K-Pop Music Video.”)

On the other hand, communicating with clients requires a different approach. Storyboard artist Sylvain Despretz describes the need for tailoring knowledge for specific parties: “One of the reasons we don’t hear about technical jobs in the industry is because, in order to talk about them, we have to agree to bring the reader into the minutiae.” (Gorny.) When communicating with clients, artists should privilege the overall creative vision; when communicating with other practitioners, the conversation may expand to elaborate on niche concepts and technical details.

3. Access to experts

In *Skills Shortages, Gaps and Training Needs in the screen industries in Yorkshire and the Humber: Scoping Report*, Ozimek describes the link between tacit knowledge and the skill shortage in the screen industries (i.e. film, television, animation): "Tacit knowledge is acquired not only in specific work settings but also through engagement with professional networks and communities both online and offline…

barriers exist in accessing those settings and networks, and therefore tacit knowledge is not easily acquired for all." (Ozimek 24)

Access to industry experts is crucial in the acquisition of tacit knowledge, but this access is limited to most individuals."Contrary to other industries and professions, the screen industry does not require specific professional qualifications to work in the industry… much recruitment is based on reputation and personal contacts… The industry 'lacks any clear structure and pathways or routes from education to the industry.'" (Ozimek 24 - 25) My encounters with these issues in videography have informed the development of a system that codifies and explicates tacit videographic knowledge.

1.1.1.1 Film Industry Tools - Storyboard

Storyboards are visual representations of a video sequence, and show the progression of the story through shots. (Bzymek) They are both internal tools that help guide the development of narratives of films and video games, and external tools that communicate the intended vision to the client. In complex productions like music videos, a storyboard also specifies additional elements like the types of physical sets and computer graphics needed.



Fig. 3 - TripleS Hit the Floor MV Storyboard Panels by Park You Seung (Park)

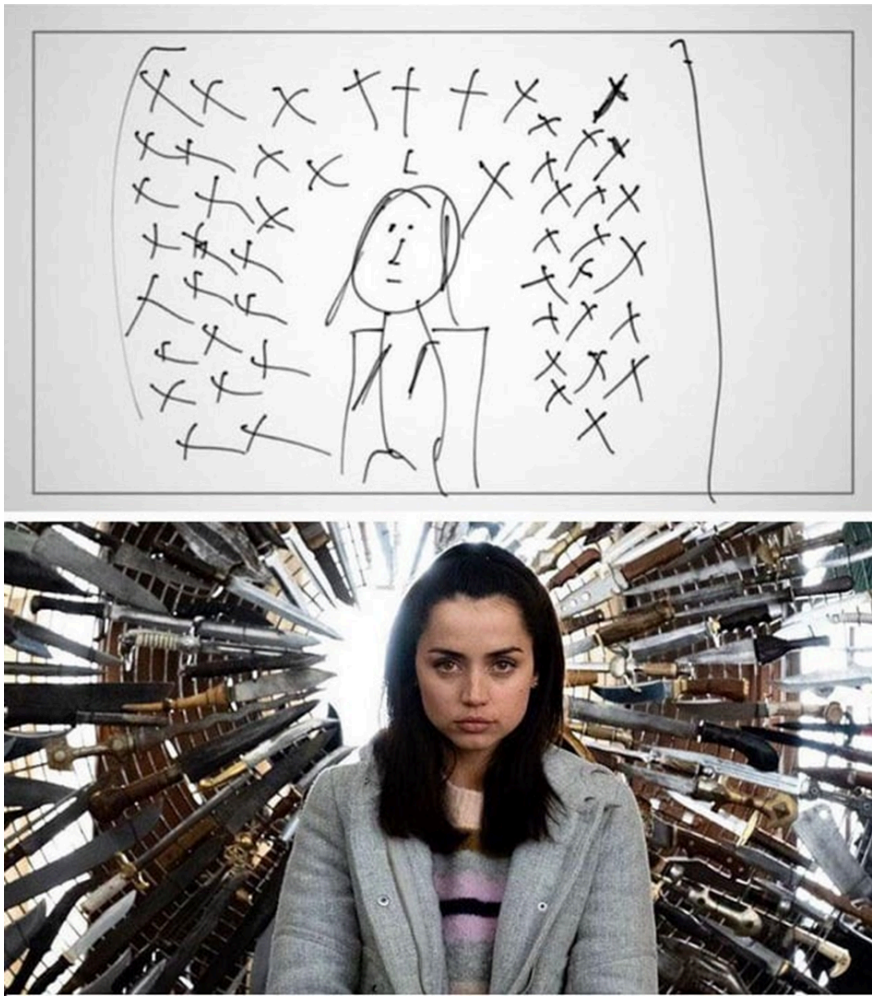


Fig. 4 – Knives Out Storyboard by Director Rian Johnson (Johnson)

1.1.1.2 Film Industry Tools - Shot List

Whereas storyboards are the visual representation of the desired final product, shot lists are a textual representation of the process. According to Simon, "a shot list is a document that maps out exactly what will occur and what will be used in that particular shot, or scene, of the video… [they] go hand-in-hand as part of writing the script, creating a storyboard, and the overall pre-production process." (Simon) It is both a list of mechanical instructions that specify the technical aspects of each shot, i.e. equipment, angle, shot length, composition, etc. and also an internal organisational tool to manage time, props, actors, equipment, etc.

The article 'Film 101: What Is a Shot List? How to Format and Create a Shot List' by Masterclass gives the following 12 elements of a shot list, each with their own column and description per shot[1]:

1. Shot number: the reference number assigned to each individual shot.

[1] Only the bolded elements are relevant within the scope of dance performance videos.

2. Shot description: a short description of the action and/or dialogue.
3. Shot size: how big or small the subject is in the frame.
4. Shot type: the camera angle, or how the camera frames the subject.
5. Movement: how the camera does (or doesn't) move within the shot.
6. Equipment: the type of camera that captures the shot.
7. Lens: the camera lens used to capture the shot.
8. Frame rate: the frequency at which the frames are captured.
9. Location: where the shot is captured.
10. Actors: the actors included in the shot.
11. Sound: how the sound and/or dialogue are captured.
12. Extra notes: remaining anecdotes the director wants to convey to the crew about the shot.

1.1.2 Movement Notation Systems

In *Writing for the Body*, Mark Franko positions video as a tool for the "transmission and reproduction of dances" to the general public (Franko 330). Before video became the medium of choice, "at least 87 movement writing systems have been used in Europe and North America since the fifteenth century", developed with the intention of communicating the movement of bodies in space over time (Farnell 1). Over the years, the following three notation systems have become the most prevalent:

1. Labanotation (Kinetography Laban)
2. Benesh Movement Notation (BMN)
3. Eshkol-Wachman Movement Notation (EWMN)

Of the three, Labanotation is the most well known. Proposed in 1926 by Austro-Hungarian choreographer and dancer Rudolf Laban, it uses a 9-staff system organised around a central staff representing the human spine. Parts of the body - arm, body, leg gesture, support - are indicated by lines on their respective side relative to the central staff. The notation is read from top to bottom with time denoted by the vertical axis.

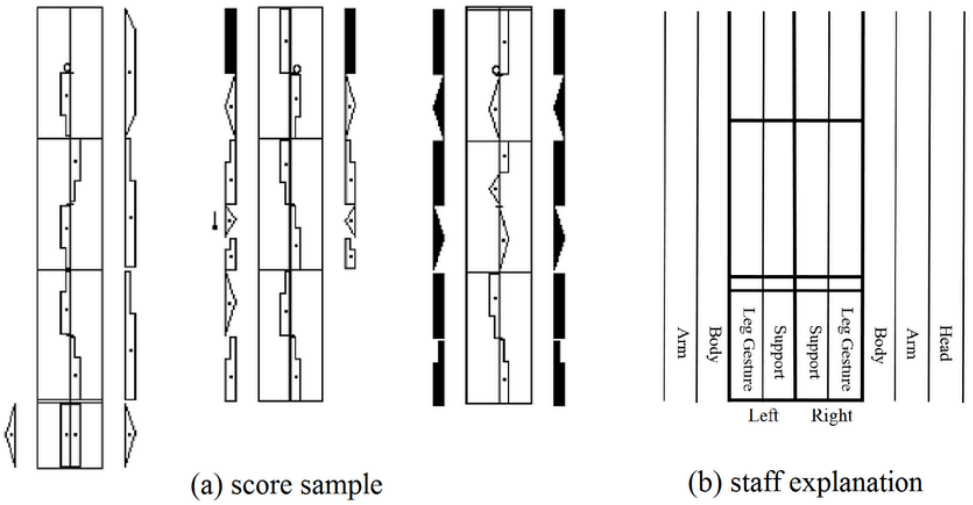


Fig. 5 - Labanotation Score Sample & Staff Explanation (Li, Min, et al.)

There are two main sets of symbols: ones indicating joint, limb, and limb surface; and ones indicating direction and height. Placement of these symbols along the 9 staffs indicate which body parts are moving where.

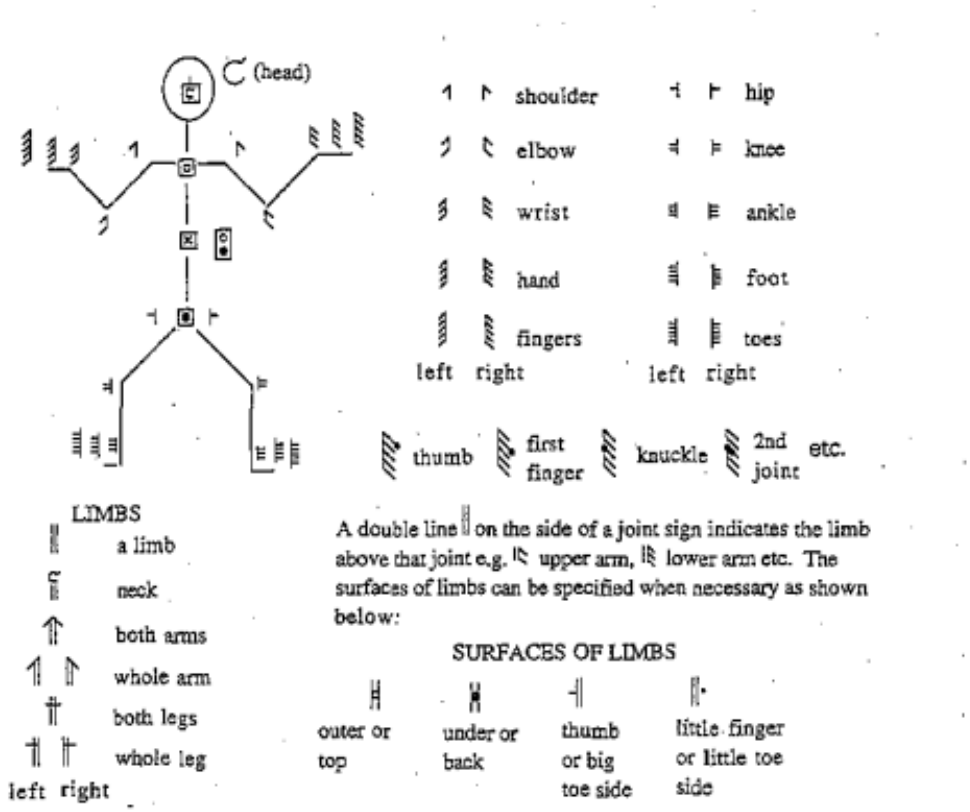


Fig. 6 - Labanotation System: Graphic signs for body parts and surfaces (Farnell 159)

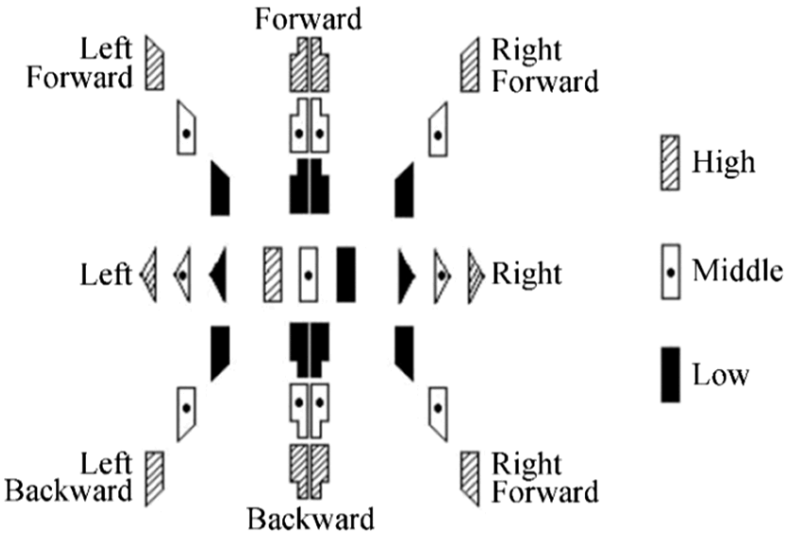
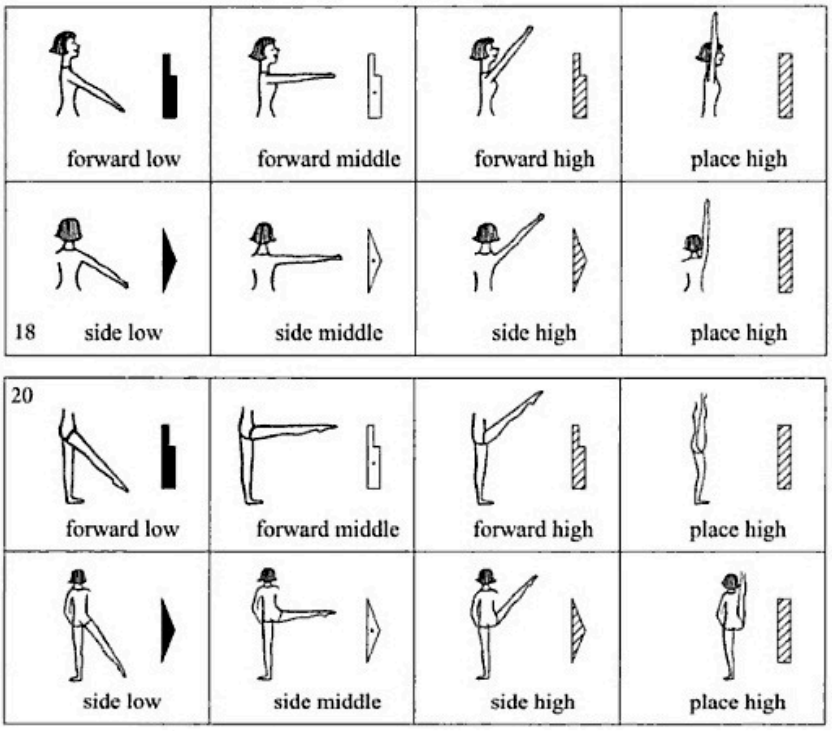


Fig. 7 - Basic Labanotation Symbols for Direction (Zhang, Xueyan, et al. 2)



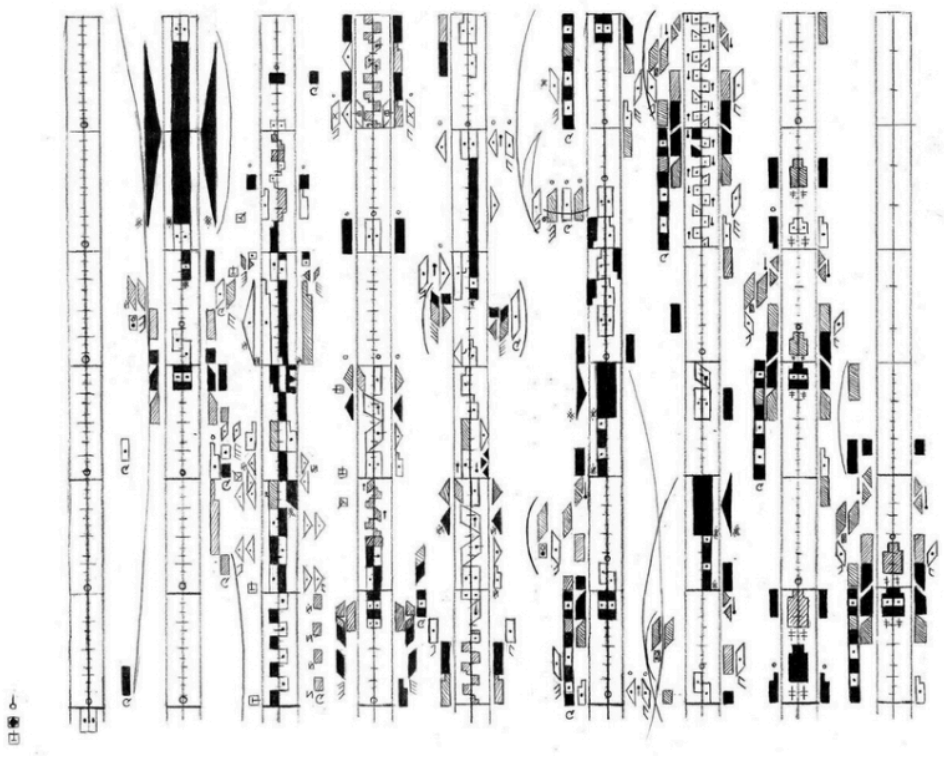


Fig. 8 - Labanotation Example Score (Robinson)

A major force behind the adoption of Labanotation is the NYC Dance Notation Bureau (DNB). Established in 1940, it advocates for Labanotation as a dance preservation and copyright tool. While originally developed to record European modern dance, Labanotation has been used to describe Western theatre choreography, Eastern European traditional folk dances, sociocultural anthropology, religious studies, Plains Indian sign language, and kinesiology. However, complexity and workflow issues have been cited as limitations to the system. Literacy requires training, transcription is time-intensive and each score requires external verification.

In general, the lack of an existing universal dance notation system belies the complex nature of dance notation. In their examination of Labanotation, Whitney E. Laemmli identifies three traits necessary for a movement notation system:

1. universality
2. economy
3. accuracy (Laemmli 12)

Likewise, Farnell posits: "A writing system creates concepts specific to the medium… A movement script offers much more than a new methodology for inquiry into human movement. The possibility of movement literacy opens up an important theoretical alternative to objectivist talk *about* the body or phenomenological-subjectivist talk *of the feeling of* body movement." (Farnell. 168.) Thus, a grammar that communicates videography-specific concepts could open up new paradigms for dance videography.

1.1.2.1 Alternative Notation Systems

In *The Future is Graphic: Animated notation for contemporary practice*, Hope presents animated notations as an alternate paradigm for time-based notations, with intended application for electronic music and electronic-acoustic hybrid music that feature elements not covered by standard musical notation, including micro-tonality, interactivity, non-linear structures, etc. (Hope)



Figure 1. Marta Tisenga 'Snæfellsnes' (2015).

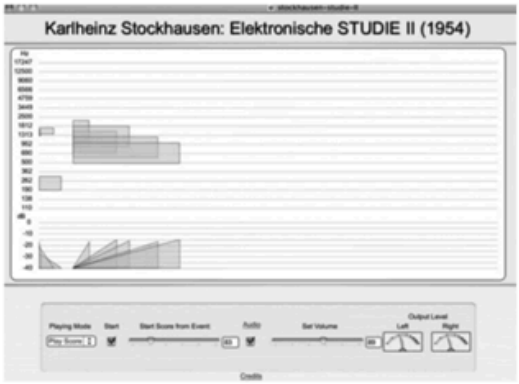


Figure 2. Georg Hajdu, electronic reading of Karlheinz Stockhausen's *Electronic Studie No. 2* (1954) created in MaxMSP (2012).

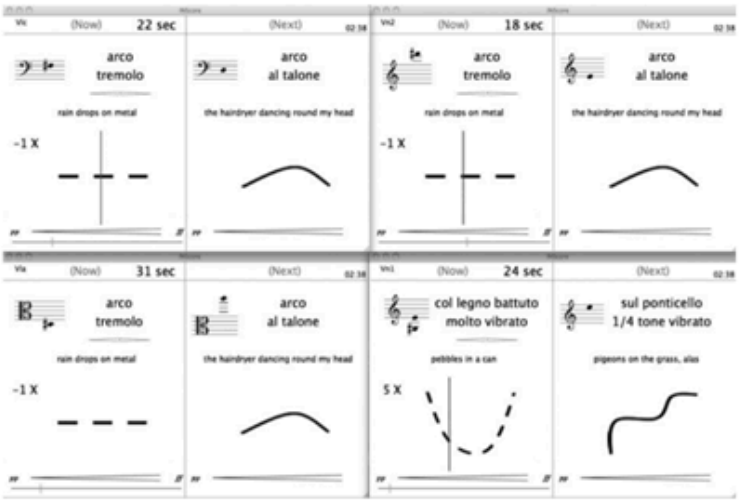


Figure 3. Sandeep Bhagwati, Monochrom, 2011, for string quartet (working version; note the playheads are red).

Fig. 9 - Alternative Notation Systems (Hope)

Similarly, the workflow developed by Park Sang-woo of South Korean video production company Digipedi uses manually colored cells in Excel to visualise music.

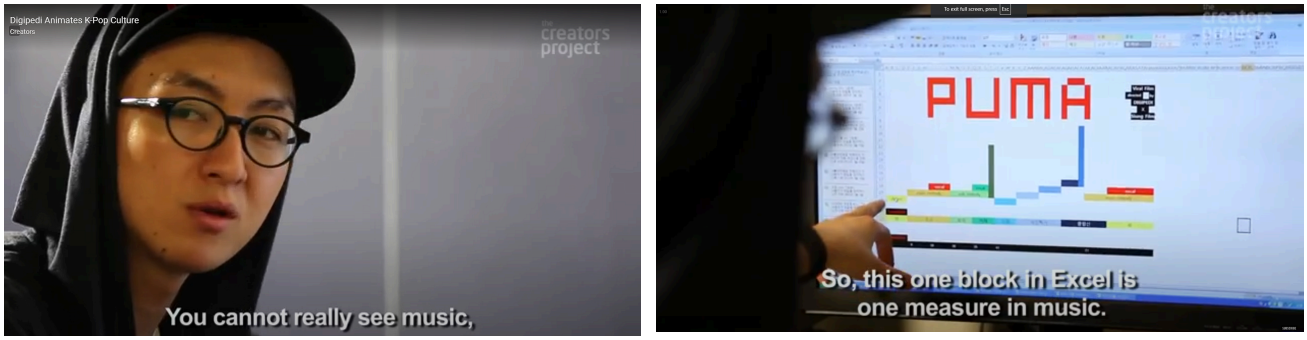


Fig. 10 - Park Sang-woo explains his music visualisation process (Creators)

These alternate systems demonstrate Sally Ness's observation on the emergence of notations: "Wherever codified techniques have evolved, wherever habits have stabilized enough to produce a continuity of style, general principles of conduct are operative in a discipline." (Franko 323)

1.1.3 Shape Grammars

Shape Grammars are a form of computation where “a set of shape rules apply in a step-by-step way to generate a set, or language, of designs” (Knight). Movement notation systems are visual languages composed of sets of shapes, and by definition are types of shape grammar. In *Introduction to Shapes and Shape Grammar*, Stiny defines shape grammar as an "algorithm [that] defines a language of shapes" and has four components:

1. S = {set of shapes}
2. L = {set of symbols}
3. R = {set of shape rules in the form $\alpha^2 \rightarrow \beta^3$}
4. I = initial labelled shape in $(S, L)^+$

(Stiny 346 - 347.)

Shape grammars provide a framework for codifying tacit knowledge as “the visual nature of shape grammars” makes them well suited at capturing “ambiguity and emergence” (Jowers). Thus, this work will leverage the shape grammar methodology to describe how one shot becomes another (i.e. $\alpha \rightarrow \beta$).

[2] α = labelled shape in $(S, L)^+$
[3] β = labelled shape in $(S, L)^*$

1.1.4 Visual Analysis

In *CineScale: A dataset of cinematic shot scale in movies,* Savardi et al. outlines a dataset of more than 792,000 image frames annotated according to the following shot scale categories:

1. Extreme Close Up (ECU)
2. Close Up (CU)
3. Medium Close Up (MCU)
4. Medium Shot (MS)
5. Medium Long Shot (MLS)
6. Long Shot (LS)
7. Extreme Long Shot (ELS)
8. Foreground Shot (FS)
9. Insert Shots (IS)

These frames were taken at 1 fps from a corpus of 124 movies by 6 famous directors and manually annotated by 3 human coders. Aside from establishing shot scale as a significant factor for shot type determination, the Cinescale dataset provides a precedent framework for the processing of videos via shot scales that can inform the visual analysis within this work.(Savardi et al. 8)

1.1.5 Encoding Cinematography

A way of encoding the ‘minutiae’ of tacit videographic knowledge is explored by Christianson et al via their proposed Declarative Camera Control Language (DCCL), which they claim "represent cinematic knowledge at the same abstraction level as expert directors and producers" (Christianson et al. 7). There are 4 primitives:

1. fragments
2. views
3. placements
4. movement endpoints

where fragments are singular camera positions or motions.

```
(AcFilmIdiom name Arijon-13-2
  :parameter (AcParamApproach :actor1 :actor2 :start :stop)
  :line (AcLineIdiom :primary ?actor1 :other ?actor2 :side left)
  (AcFilmShot name shot1
    (AcFragGoBy name frag1
      :time ?start :primary-moment beginning :entry-pos center :exit-pos out-left
      :placement (AcPlaceInternal :primary ?actor1 :other ?actor2 :range closeup :primary-side center)))
  (AcFilmShot name shot2
    (AcFragGoBy name frag2
      :time ?frag3.first-tick :primary-moment end :entry-pos on-right :exit-pos center
      :placement (AcPlaceExternal :near ?actor1 :far ?actor2
                                  :primary-subject near :range longshot :primary-side center)))
  (AcFilmShot name shot3
    (AcFragGoBy name frag3
      :time ?stop :primary-moment end :entry-pos out-right :exit-pos right12
      :placement (AcPlaceApex :primary ?actor1 :other ?actor2 :range mediumshot :primary-side right12))))
```

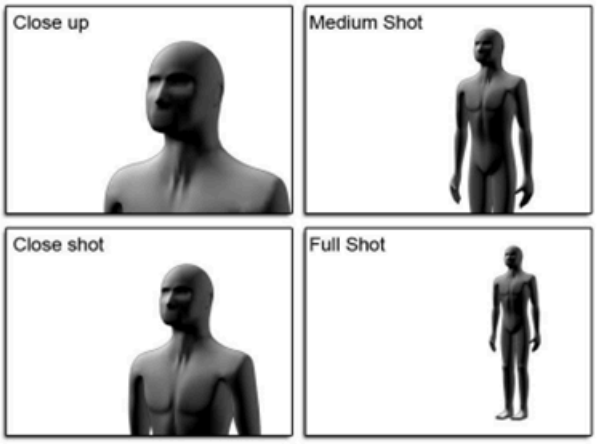
Figure 3: DCCL code corresponding to the idiom depicted in Figure 2.

The second part of the proposal concerns a 3 stage Camera Planning System (CPS) that produces an animation that follows the specified cinematographic idioms:

1. Sequence Planner
2. DCCL Compiler
3. Heuristic Evaluator

The Sequence Planner partitions and classifies each film sequence into 1 of 10 predefined activities (see Table 1). It generates scene nodes per partition, matching it to 1 of 16 predefined idioms (also shown in Table 1), resulting in a film tree with each branch populated by candidate idioms (see Figure 7).

DCCL is a generative system that describes videographic sequences via idioms with changeable parameters and outputs a 3-D visualization of the scene based on the ‘best idiom’. (Christianson et al.)



Close up	`Height(head, [40-60]%, t);` `InView(head, yes, t);` `Occlusion(head, no, t);`
Close Shot	`Height(headtorso,[60-80]%, t);` `InView(headtorso, yes, t);` `Occlusion(headtorso,no,t);`
Medium Shot	`Height(headknee,[80-90]%,t);` `InView(headknee, yes, t);` `Occlusion(headknee,no,t);`
Full Shot	`Height(body, [80-90]%, t);` `InView(body, yes, t);` `Occlusion(body, no, t);`

Figure 2: Conventional framings for human characters

Fig. 11 - Encoding Videographic Idioms with DCCL (Christianson et al.)

In *Specification of Cinematic Idioms*, Schofield et al. defines idioms within the context of cinematography, as representations of "idiosyncratic ways for conveying actions" (Schofield et al. 26). There are 4 main motivations for camera movement (Schofield et al. 23):

1. establishing context
2. establishing relations between targets
3. tracking targets
4. establishing relations between subject and context

Similarly, they make the argument for editing in cinematography: "Multiple camera idioms are used in cinematography because of their ability to arrange the narrative in space in a way that can be easily read by the viewer." (Schofield et al. 25) Making this multiple-idiom relationship clear is crucial, with the goal of building on the parameters and systems as defined by Schofield.

Regarding camera planning and control, the literature appears to show a convergence on the need for a method of quantifying desirable cinematography traits (e.g. cost function) and some evaluative function that sorts candidate sequences according to desirability. This requires a rigorous ruleset and is to be addressed once the core notation vocabulary is established.

1.2 Hypothesis

Following the shape grammar model, a generative notation for dance videography uses symbols to represent a camera's movement in three-dimensional space relative to a subject or subjects over a period of time.

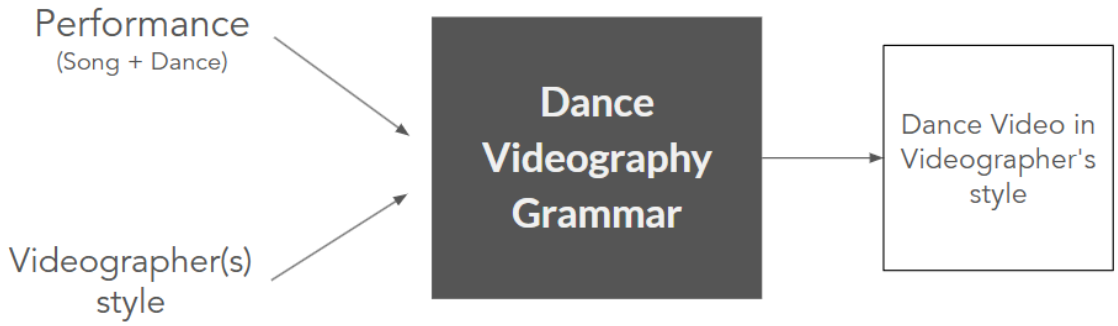


Fig. 12 - Mechanics of Proposed Dance Videography Grammar

The grammar will take two inputs - Performance (Song & Dance Choreography) and Specified Videographic Style - and allow the user to generate a dance video in the given style. There are three intended applications:

1. Replication
    a. Description of an existing Performance in an existing Style
2. Extension
    a. Description of a new Performance in an existing Style
3. Generation
    a. Description of any Performance in a new Style

This work hypothesizes that, given a grammar that exhibits Laemmeli's three essential traits - universality, economy, accuracy - a video of any dance performance in any style can be described (Laemmli 69). Capturing the tacit technical knowledge embedded in dance videography enables both lateral and vertical communication and makes accessible esoteric videographic concepts, positioning the grammar as a tool for movement literacy .

2 Methods

2.1 Examining Videographic 'Style'

There are numerous video production companies that produce dance performance videos. To differentiate their work, each company has developed a signature style.



Fig. 13 - Various Dance Performance video series viewable on Youtube

The thumbnails in Fig. 13 each represent a ‘unique’ viewing of the song and dance informed by the director’s approach to how dance ‘ought’ to be perceived. All draw from the same pool of videographic techniques, but are distinct in application and intended message. This relationship between technique and vision is called 'style'.

Defining a videographic ‘style’ is highly subjective; most viewers can tell that certain videos 'feel different', but are unable to identify why. Thus, by identifying and parameterising shared concepts in dance videography, this work intends to develop a notation system that qualitatively describes a given video 'style', allowing multiple styles to be compared and evaluated. The notation's two main generative capabilities are:

1. The production of new videos using existing styles;
2. The generation and description of completely new styles.[4]

2.2 Building a Corpus

In particular, this work will focus on production company M2 and their video series *[Performance37]*. Of the 229 videos listed on the *[Performance37]* playlist, a corpus of three videos have been selected for thorough analysis. To ensure fair representation of the *Performance37* style, these videos vary in artist, number of dancers, music style, dance style and BPM. Lee Hanhyung and Lee Juri are credited as Executive Producers and Director, respectively.

[4] Re-production of existing videos in existing styles is not included as it does not require this notation system.



Artist: LE SSERAFIM
No. of Members: 5
Song: *Eve, Psyche & The Bluebeard's Wife*
Music Genre: Funk, Electropop, Disco
Dance Genre: Girls Hip Hop, Vogue, Dance hall
BPM: 143

("[Performance37] LE SSERAFIM(르세라핌))



Artist: UNIS
No. of Members: 8
Song: Dopamine
Music Genre: Hyper pop, Dance
Dance Genre: Girl Hip Hop, Girls Jazz Funk
BPM: 170

("[Performance37] UNIS(유니스))



Artist: izna
No. of Members: 7
Song: IZNA
Music Genre: Hip Hop, R&B
Dance Genre: Girls Hip Hop
BPM: 106

("[Performance37] Izna(이즈나))

[Performance37] is only one of many video series produced by M2; other popular dance content IPs include *Relay Dance*, *Suit Dance, Studio Choom* and *Move to Performance* which appear to be filmed in the same location using similar equipment, involving the same personnel. Notable overlaps include Lee Hanhyung as Executive Producer and Kim Boram as Gaffer in both *Studio Choom* and *[Performance37]*, yet each series feels distinct in style.

2.3 Video Analysis

The first video in the corpus, *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* is a 3:23 min video posted by the M2 YouTube channel on May 23, 2023. Since then, it has reached 6M views and is the 3rd most viewed video in the Performance37 playlist.

To identify the hallmarks of the *[Performance37]* style, this work conducts analysis of corpus video to answer the following questions:

1. [Visual Analysis] What is occurring in the video?
2. [Audio Analysis] When are they happening?
3. [Audio - Visual Analysis] How are 1 and 2 related?

2.3.1 Visual Analysis

2.3.1.1 Shots, Cuts, Sequences

Videos operate with an established logic. In *Cinematography: Theory and Practice*, Blain Brown explains: "If we think of a language of cinema, these shots are the vocabulary; how we edit them together would be the syntax of this language." (Brown 121) Establishing and then following a coherent syntax allows viewers to understand the videographer's intent, regardless of whether they understand the language. Thus, a rule set that articulates this 'syntax' will capture the video's style.

Using terms as defined by the Film Terms Glossary on Filmsite.org, the *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* video can be described as a sequence of shots, separated by hard cuts. ("Film Terms Glossary.")

Shot	the basic building block or unit of film narrative; refers to a single, constant take made by a motion picture camera uninterrupted by editing, interruptions or cut; can also refer to a single film frame (such as a still image)
Cut	an abrupt or sudden change or jump in camera angle, location, placement, or time, from one shot to another[5]
Sequence	a scene, or connected series of related scenes that are edited together and comprise a single, unified event, setting, or story [within a film's narrative]

[5] A hard cut is a cut with no effect added in post-processing to obscure the transition.

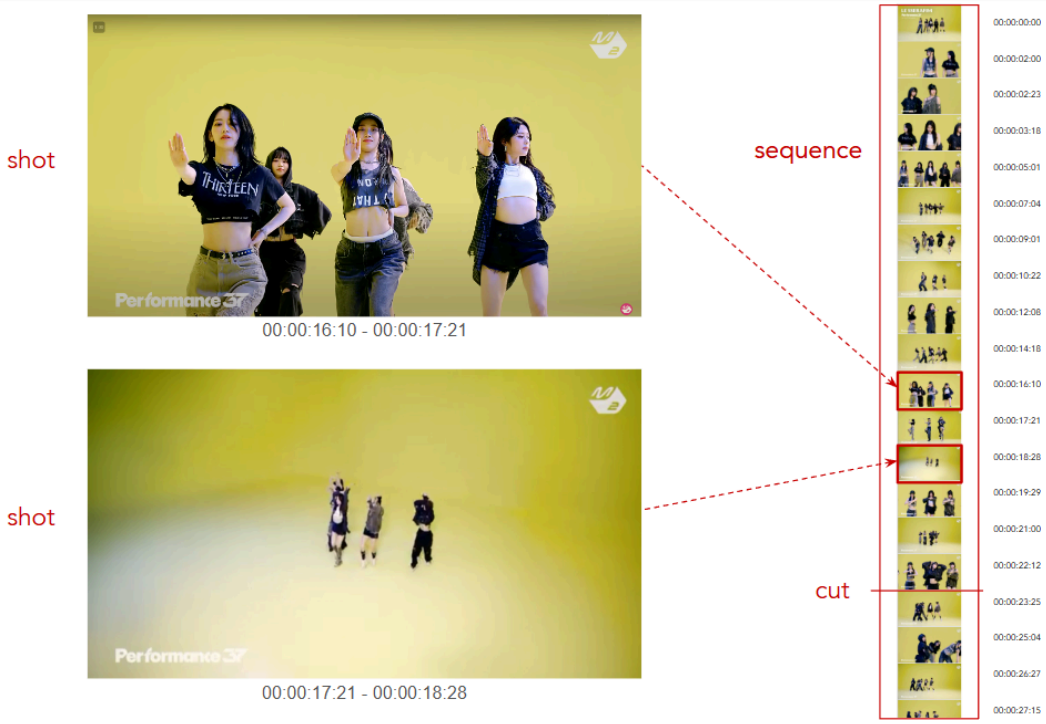


Fig. 14 - Shot, Cut & Sequence

There are 119 shots, and therefore 118 cuts in *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'*.

In *Moving in Filmic Spaces: Relating Camera Movements to Spatial Archetypes in architectural Animations*, Al-Saati proposes a framework for categorising shots based on camera movement, such as Point, Line or Circle. *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* is a sequence of 'Point'-type shots. (Al-Saati 632)

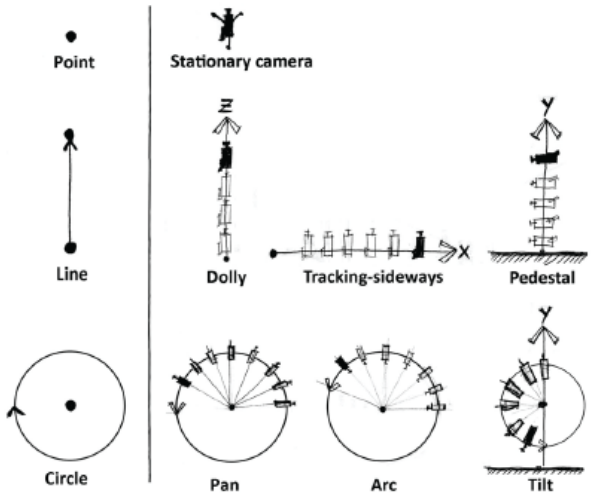


Fig. 15 - Point, Line & Circle Camera Movements

Due to the 'Point'-type nature of *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* , there only 2 factors that affect the type of shot, and therefore it's 'syntax':

1. Shot Size - The size of the subject relative to the frame
2. Angle - The orientation of the subject relative to the camera

2.3.1.2 Shot Size

Breaking down and sorting the video according to shot size reveals the six different shot size used:

1. Medium Close-Up (MC)
2. Medium Shot (M)
3. Medium Full Shot (MF)
4. Full Shot (F)
5. Long Shot (L)
6. Extreme Long Shot (EL)

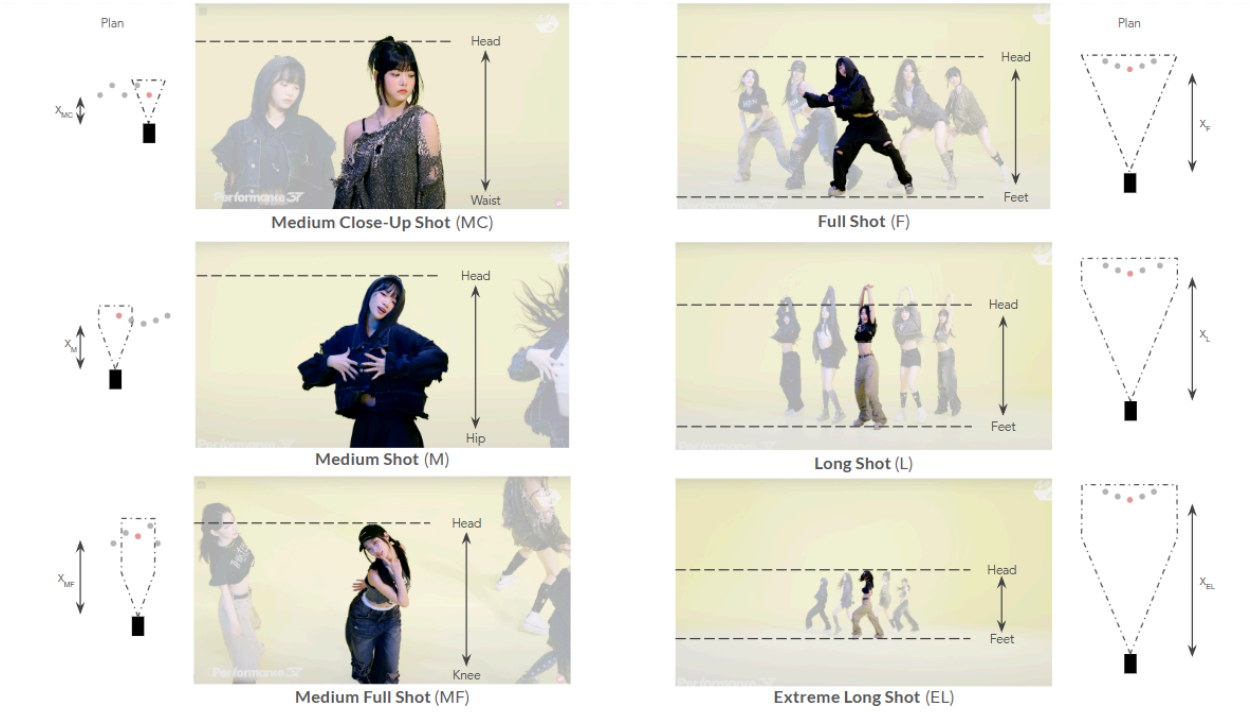


Fig. 16 - Shot Sizes and corresponding Example Video Frames

The size and amount of a dancer's body shown directly correlates to the perceived distance between subject and camera. In a Medium Close-Up shot, only the head to the waist of the dancer is shown. The camera is close to the dancer, showing greater detail. In contrast, the dancer's entire body is shown in a Full Shot. The camera is farther away, the focus is more on the dancers' movements. In an Extreme Long Shot, the dancer's entire body is still shown, but they take up less than one-third of the

screen. As the camera is quite far away, only large movements register, and the viewer better understands the dancers as bodies within space.

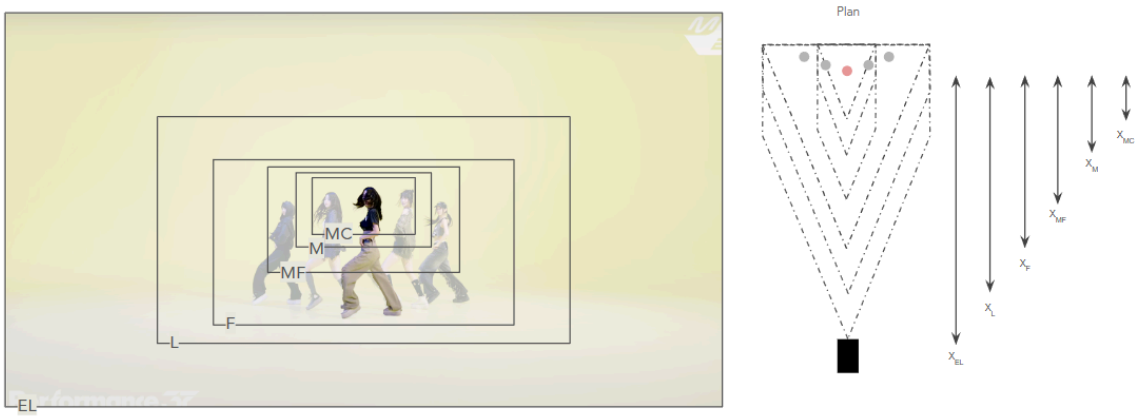


Fig. 17 - Shot Sizes and corresponding Subject-Camera-Distance

2.3.1.3 Shot Angle

According to Michael Chapman, cinematographer for Taxi Driver and Raging Bull, the role of angles is to elicit an subconscious emotional reaction: "Angles tell us emotional things in ways that are mysterious—emotional things that I am often unaware of." (Brown 122) Camera angles change the way viewers engage with the video content, as they change the relationship between subject and camera.

The camera can rotate around three axes: x (pitch), y (yaw) and z (roll). As there are no yaw rotations present in the video, only pitch and yaw will be described. Noticeably, this work will treat the main dancer of the choreography at the given period of time as the point of rotation. Additionally, shots will be categorised through general orientation to the subject(s) instead of specific rotation angle. Shots filmed with no camera rotation are considered 'Straight' shots.

Pitch

There are 3 pitch angles present in the video:

1. High
    a. Camera is higher than the subject's chest level, pointing down. Foreshortening causes the subject to appear small and diminutive.
2. Low
    a. Camera is lower than the subject's chest level, pointing up. Foreshortening causes the subject to appear large and dominating.
3. Top
    a. Camera is directly above the subject, oriented perpendicular to the floor. This is a naturally uncommon perspective, creating a sense of distance from the subject.

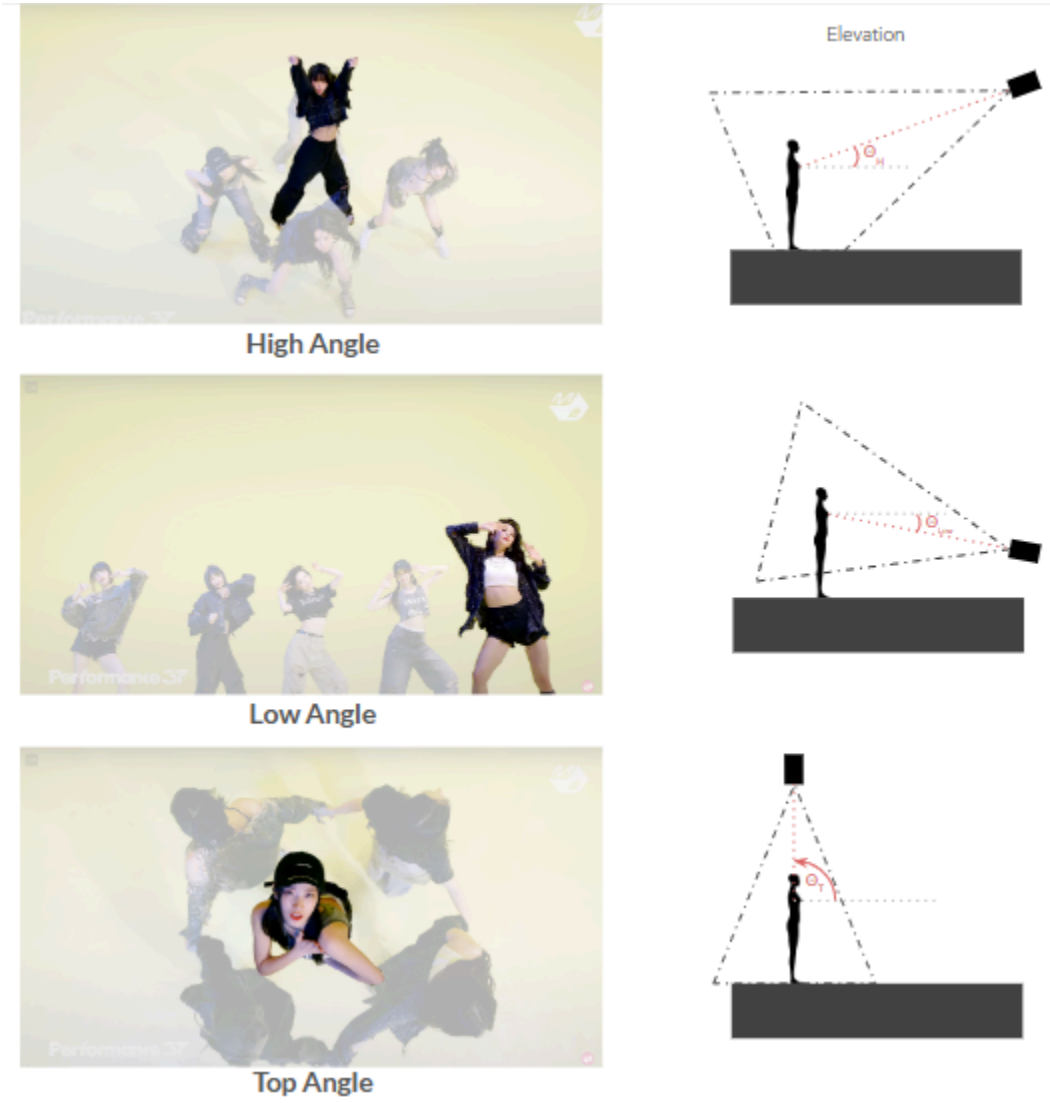


Fig. 18 - Pitch angles and corresponding example video frames

Yaw

There are 2 yaw angles present in the video:

1. Left
    a. Camera is rotated clockwise relative to the dancers, pointing to the right. Actions on the subject's right are obscured.
2. Right
    a. Camera is rotated anti-clockwise relative to the dancers, pointing to the left. Actions on the subject's left are obscured.

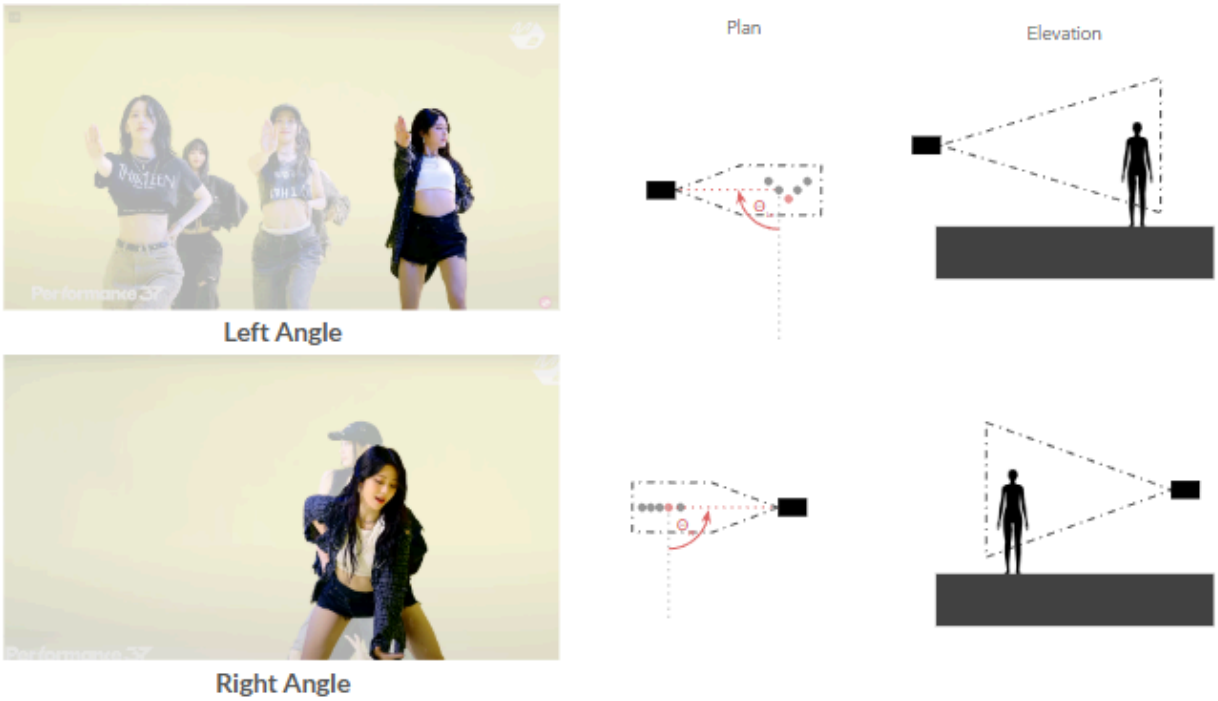


Fig. 19 - Yaw angles and associated example video frames

Pitch + Yaw

Identifying these angles allows for the generation of all possible camera angles with both pitch and yaw, as well as the identification of these angles within the video, as shown in Fig and Fig below.

	High	Low	Top
Left	Left High	Left Low	
Right	Right High	Right Low	
Left Back	Left Back High	Left Back Low	
Right Back	Right Back High	Right Back Low	

Fig. 20 - Pitch-Yaw Matrix showing all possible camera angles

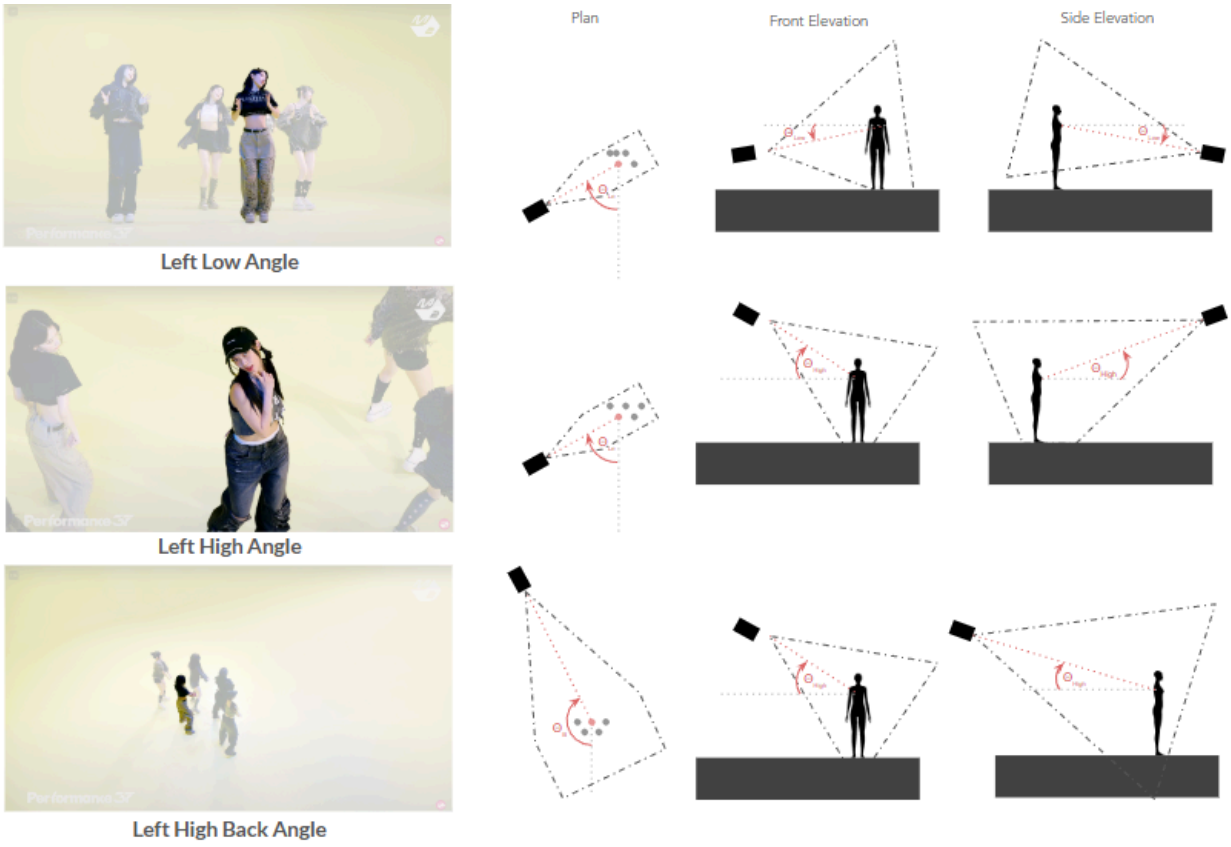


Fig. 21 - Left Pitch-Yaw angles and associated example video frames

2.3.1.4 Quantitative Analysis

Shots

Plotting all shots in the video in a matrix according to the identified parameters reveals that 25 out of 72 possible unique shot types are used. Additionally, this matrix shows that the range of shot sizes present in [*Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* skew long with no yaw angle shots (i.e. Left & Right) closer than Medium Full used.

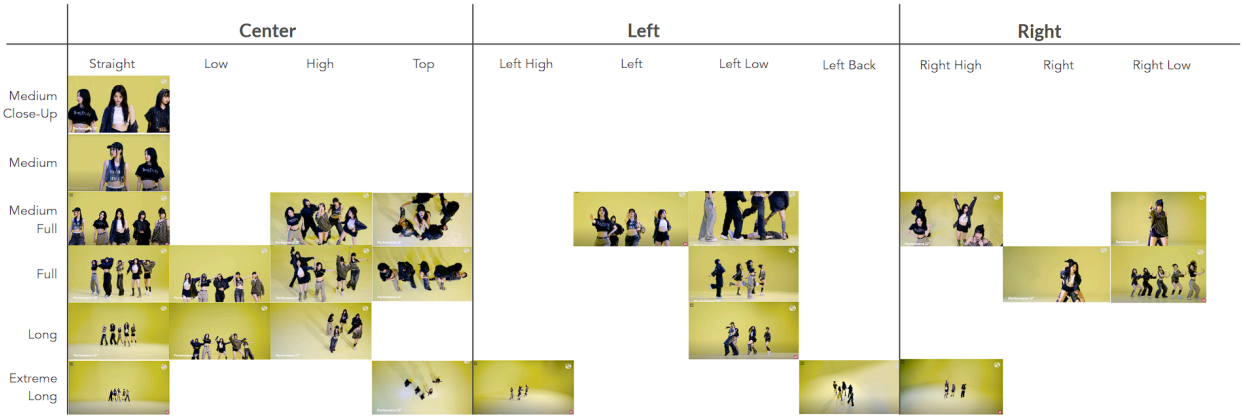
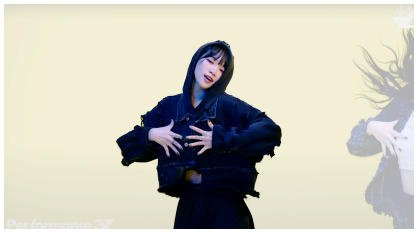


Fig. 22 - Shot Type Matrix for[*Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'*

Shot Size

Comparing the duration of each shot size provides another way of comparing the 'prevalence' of each shot type. All 119 shots are sorted according to shot size and counted, to calculate shot size duration as a percentage of the video's total run time.

Medium Shot Size (33.05%) and the Long Shot Size (24.97%) are the most and second-most prevalent shot size, respectively. Notably, the Medium Closeup Shot Size (2.33%) is the least prevalent.

Shot Size	Example Frame	Count	Time Elapsed (secs)	% of Total Run Time
Medium Closeup ($S_{MC-?}$)		3	4.4	2.33%
Medium ($S_{M-?}$)		39	62.47	33.05%
Medium Full ($S_{MF-?}$)		9	14.87	7.87%
Full ($S_{F-?}$)		21	37.07	19.61%

Long ($S_{L-?}$)		28	47.2	24.97%
Extreme Long ($S_{EL-?}$)		19	23	12.17%
Total		119	189	100.00%

Angles

The overwhelming majority of shots do not have a pitch or yaw angle - Straight (58.06%). Of the angled shots, the most prevalent angle types are the High angle (18.69%), Low angle (9.61%) and Left Low angle (4.59%), respectively.

Angle	Example Frame	Count	Time Elapsed (secs)	% of Total Run Time
Straight ($S_{?-S}$)		72	109.73	58.06%
High ($S_{?-H}$)		19	35.33	18.69%

Low ($S_{?\text{-}L}$)		8	18.17	9.61%
Top ($S_{?\text{-}T}$)		3	3.9	2.06%
Left ($S_{?\text{-}Le}$)		1	1.37	0.72%
Right ($S_{?\text{-}R}$)		2	1.67	0.88%
Left High ($S_{?\text{-}R}$)		3	3.7	1.96%
Left Low ($S_{?\text{-}LeLo}$)		6	8.67	4.59%

Right High ($S_{?\text{-}RH}$)		4	5.07	2.68%
Right Low ($S_{?\text{-}RLo}$)		1	1.4	0.74%
TOTAL		119	189	100.00%

This analysis reveals that *[Performance37]* uses different shot sizes to create visual contrast and communicate different streams of information. Not all shot types are shown for the same amount of time, implying that certain shot types are more desirable than others. The Medium shot size and Straight angle are shown the most.

Cuts

There are 118 cuts which can be distinguished based on the start and end shot type.

Shot Size 1 → Shot Size 2

Cuts are treated as instantaneous operations and are quantified by count instead of duration. The most prevalent transitions are the Long → Medium shot size (18 cuts) and the Medium → Long shot size (16 cuts). The same table with associated frames can be found in Appendix - 6.1.1 Shot Types.

Transition Type			Count	% of Total Cuts
	→	Medium Closeup ($S_{MC\text{-}?}$)	1	0.84%
	→	Medium ($S_{M\text{-}?}$)	0	0.00%
	→	Medium Full ($S_{MF\text{-}?}$)	1	0.84%
	→	Full ($S_{F\text{-}?}$)	1	0.84%

Medium Closeup

($S_{MC\text{-}}$)

	→	Long ($S_{L-?}$)	0	0.00%
	→	Extreme Long ($S_{EL-?}$)	0	0.00%
Medium ($S_{M-?}$)	→	Medium Closeup ($S_{MC-?}$)	2	1.68%
	→	Medium ($S_{M-?}$)	3	2.52%
	→	Medium Full ($S_{MF-?}$)	3	2.52%
	→	Full ($S_{F-?}$)	7	5.88%
	→	Long ($S_{L-?}$)	16	13.45%
	→	Extreme Long ($S_{EL-?}$)	8	6.72%
Medium Full ($S_{MF-?}$)	→	Medium Closeup ($S_{MC-?}$)	0	0.00%
	→	Medium ($S_{M-?}$)	4	3.36%
	→	Medium Full ($S_{MF-?}$)	0	0.00%
	→	Full ($S_{F-?}$)	2	1.68%
	→	Long ($S_{L-?}$)	2	1.68%
	→	Extreme Long ($S_{EL-?}$)	1	0.84%
Full ($S_{T-?}$)	→	Medium Closeup ($S_{MC-?}$)	0	0.00%

	→	Medium ($S_{M-?}$)	6	5.04%
	→	Medium Full ($S_{MF-?}$)	2	1.68%
	→	Full ($S_{F-?}$)	3	2.52%
	→	Long ($S_{L-?}$)	4	3.36%
	→	Extreme Long ($S_{EL-?}$)	5	4.20%
Long ($S_{L-?}$)	→	Medium Closeup ($S_{MC-?}$)	0	0.00%
	→	Medium ($S_{M-?}$)	18	15.13%
	→	Medium Full ($S_{MF-?}$)	2	1.68%
	→	Full ($S_{F?}$)	4	3.36%
	→	Long ($S_{L-?}$)	1	0.84%
	→	Extreme Long ($S_{EL-?}$)	3	2.52%
Extreme Long ($S_{EL-?}$)	→	Medium Closeup ($S_{MC-?}$)	0	0.00%
	→	Medium ($S_{M-?}$)	8	6.72%
	→	Medium Full ($S_{MF-?}$)	1	0.84%
	→	Full ($S_{F-?}$)	4	3.36%

	→	Long ($S_{L-?}$)	4	3.36%
	→	Extreme Long ($S_{EL-?}$)	2	1.68%
Total			118	100.00%

Given that the Medium and Long shot size were the most and second-most prevalent shot sizes, cuts between these shot types occur the most.

Angle 1 → Angle 2

The most cuts occur between Straight shots ($S_{?-S} \rightarrow S_{?-S}$: 41 cuts), the second-most cuts occur from Straight to Angled Shots ($S_{?-S} \rightarrow S_{?-A}$: 31 cuts). The same table including example frames can be found in Appendix - 6.1.2 Angles

Transition Type			Count	% of Total Cuts
Straight Shot ($S_{?-S}$)	→	Straight Shot ($S_{?-S}$)	41	34.75%
	→	Angled Shot ($S_{?-A}$)	31	26.27%
Angled Shot ($S_{?-A}$)	→	Angled Shot ($S_{?-A}$)	16	13.56%
	→	Straight Shot ($S_{?-S}$)	30	25.42%
Total			118	100.00%

The equal number of Straight to Angled shots and Angled to Straight shots show there is a mostly even distribution of angled shots amongst the straight shots. Consecutive Angled shots are comparatively rare, implying their use is special.

2.3.2 Audio Analysis

2.3.2.1 Song Structure

According to Genius, the song *'Eve, Psyche and the Bluebeard's Wife'* can be broken into the following 12 sections:

1. Instrumental / Intro
2. Refrain 1

3. Verse 1
4. Chorus 1
5. Post-Chorus 1
6. Refrain 2
7. Verse 2
8. Chorus 2
9. Post-Chorus 2
10. Bridge
11. Chorus 3
12. Outro

Each section can be analysed according to duration (secs), number of cuts and cut frequency (shots/ sec). Of the song sections, Chorus 3 and Chorus 2 have the highest and second-highest cut frequencies (0.76 and 0.77, respectively).

Song Section	Start	End	Duration (secs)	No. of Cuts (#)	Cut Frequency (shots/ sec)
Instrumental	00:00:00:00	00:00:07:04	7.13	5	0.70
Refrain 1	00:00:07:04	00:00:21:00	13.87	9	0.65
Verse 1	00:00:21:00	00:00:45:27	24.90	18	0.72
Chorus 1	00:00:45:27	00:01:01:04	15.23	9	0.59
Post-Chorus 1	00:01:01:04	00:01:14:14	13.33	8	0.60
Refrain 2	00:01:14:14	00:01:27:24	13.33	8	0.60
Verse 2	00:01:27:24	00:01:54:24	27.00	13	0.48
Chorus 2	00:01:54:24	00:02:08:01	13.23	10	0.76
Post-Chorus 2	00:02:08:01	00:02:21:22	13.70	8	0.58
Bridge	00:02:21:22	00:02:36:14	14.73	8	0.54
Chorus 3	00:02:36:14	00:02:49:15	13.03	10	0.77
Outro	00:02:49:15	00:03:09:00	19.50	13	0.67

Fig. 23 - Song Section Analysis

The song sections vary in duration, number of cuts and cut frequency. In particular, the cut frequency ranges from 0.48 to 0.77, indicating that cuts are not evenly spaced throughout the song.

2.3.2.2 Cut Distribution

Plotting cuts as lines with respect to time elapsed in the music shows the proximity of these cuts to each other, confirming the irregular cut distribution.

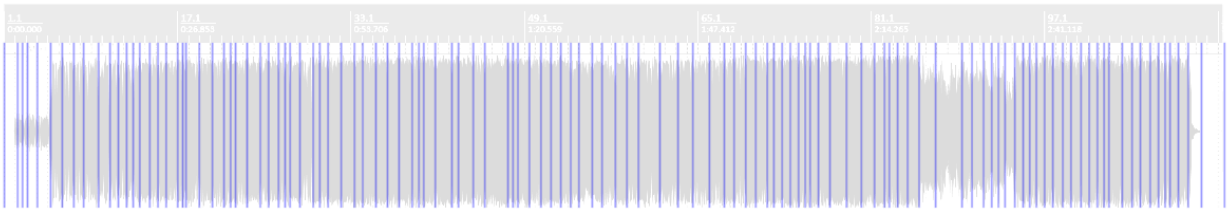


Fig. 24a - Cut Distribution with respect to Time

While it presents the intended information, Fig. 24a is highly illegible. With 118 cuts, it is difficult to discern whether cuts are 'related' or 'unrelated' to each other. Consequently, groups are 'clustered' or grouped together with any immediately preceding or following cut occurring within p[6] seconds. Each cut is displaced in the y-axis according to its index in its group, generating a line graph that shows the relative size and position of each cluster. Additionally, lines denoting discrete song sections are included to show position with respect to the music.

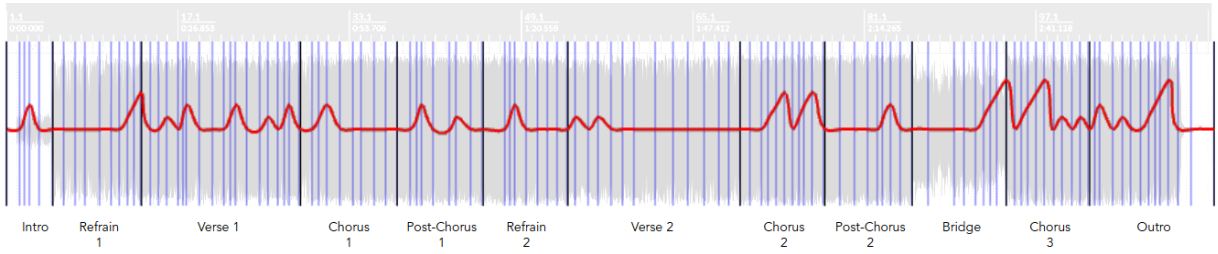


Fig. 24b - Cut Clusters with respect to Music

A cut introduces a visual change, forcing the viewer to re-engage. Placing cuts in quick succession provides the viewer with many streams of information at once, building a sense of momentum when matched to the rhythm and or beat of the song. Thus, the cut cluster line captures the subjective 'ebb' and 'flow' of the video, with an escalation to a climax in Chorus 3 and Outro. The data presented by Fig. 23 and Fig. 24b are complementary; Fig. 23 shows each section's general pacing while the peaks in Fig. 24b show how cut placement builds energy. Most notably, they corroborate the contrast between sections of high cut frequency (Chorus 3 and Chorus 2) and low cut frequency (Verse 2 and Bridge).

[6] where p is the seconds that result in the highest number of groups (i.e. g_{max})

2.3.3 Audio-Visual Analysis

2.3.3.1 Shot Size Distribution

The distribution of shot types can be visually represented in a similar way, where warm colors signify 'close' shot types (i.e. Medium Full, Medium, Medium Closeup) and cool colors signify 'far' shot types (i.e. Full, Long, Extreme Long). See Fig. 25 for reference.

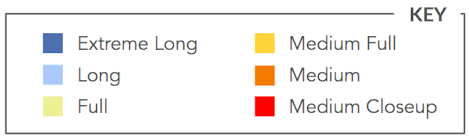


Fig. 25 - Shot Size Key

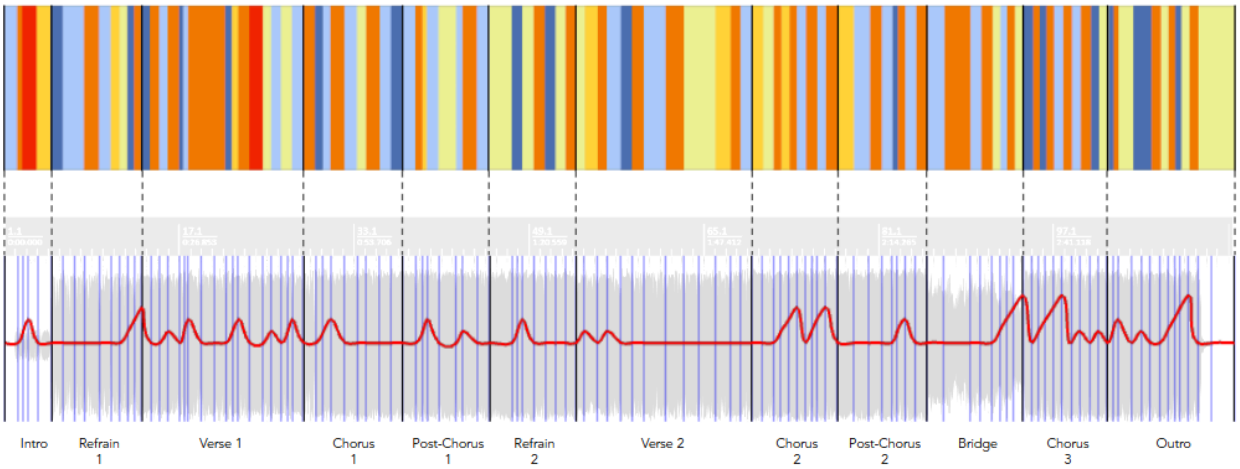


Fig. 26 - Shot Size Distribution for all Shots

A shot size cluster is defined as two or more consecutive shots of a similar shot size. Fig. 27a and 28a show the data separated via ‘close’ and ‘far’ shot size, Fig 27b and 28b isolate shot size clusters.

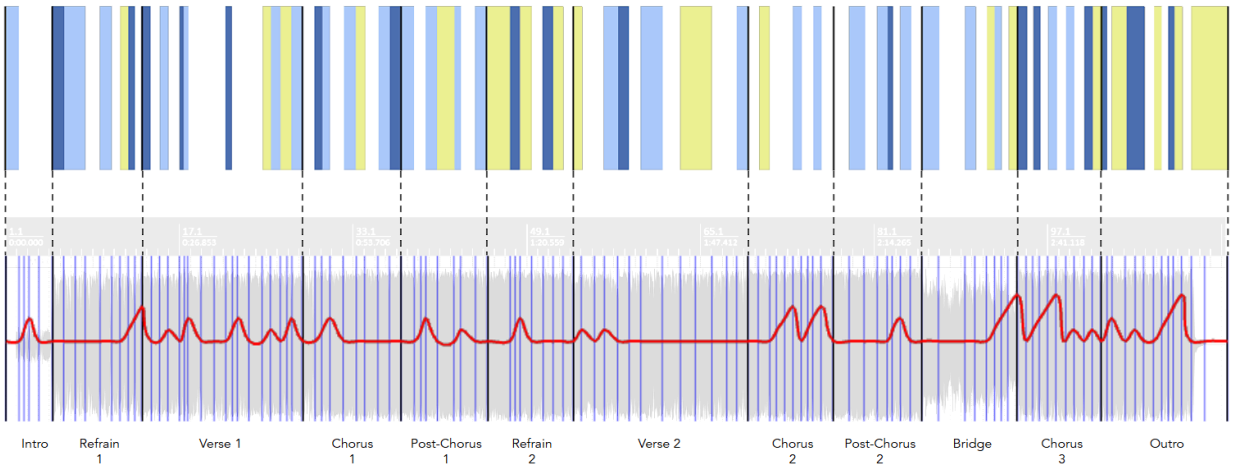


Fig. 27a - Shot Size Distribution for 'Far' Shots

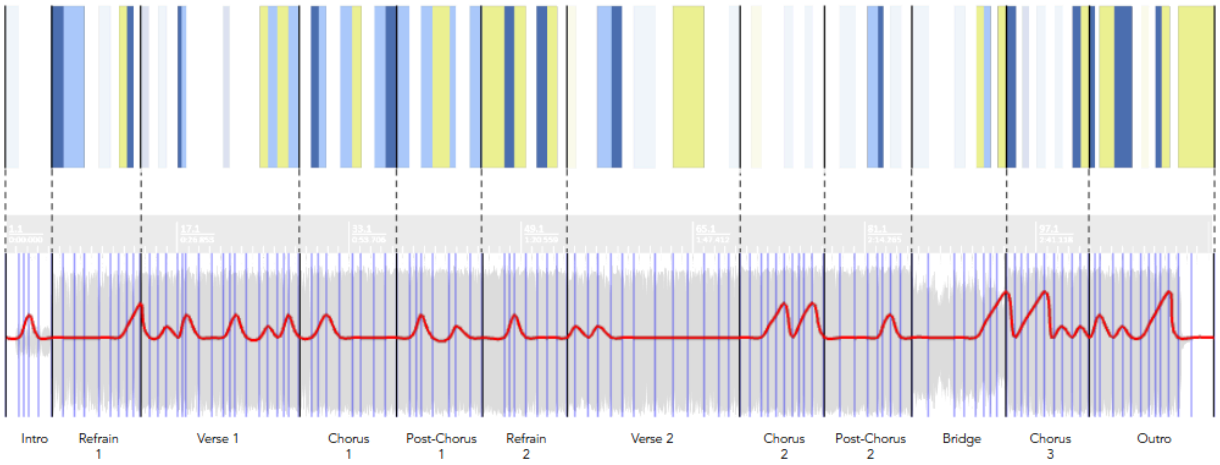


Fig. 27b - Shot Size Clusters for 'Far' Shots

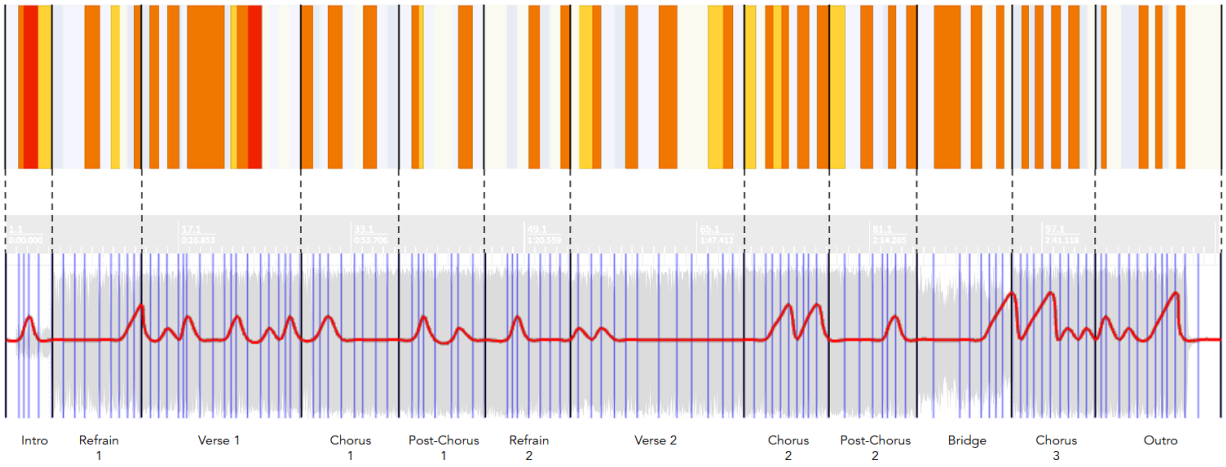


Fig. 28a - Shot Size Distribution for 'Close' Shots

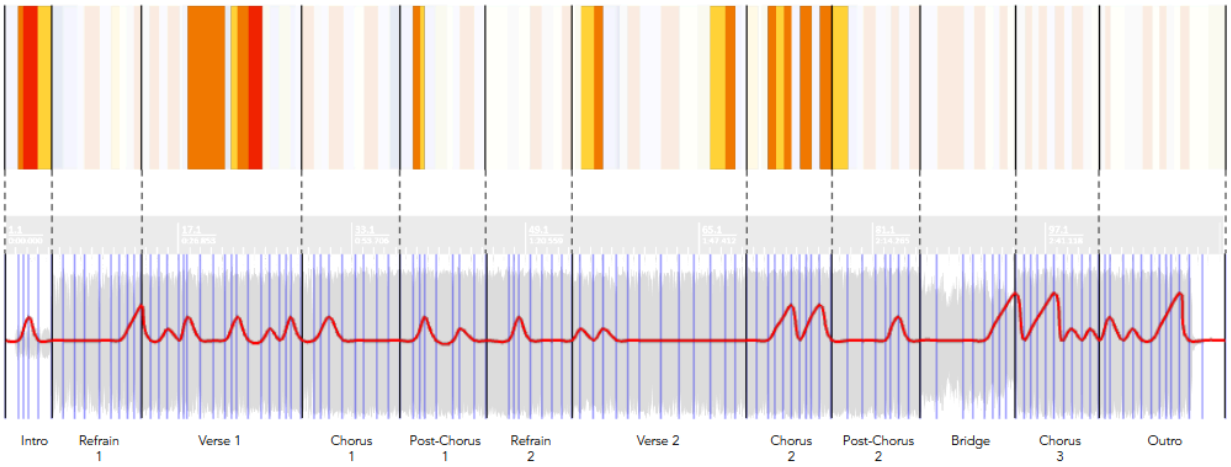


Fig. 28b - Shot Size Clusters for 'Close' Shots

When comparing Fig. 27b and Fig. 28b, it becomes apparent that not only are there more 'far' shot size clusters, but they are also more evenly distributed throughout the video. In contrast, there is one major 'close' shot size cluster in Verse 1 and several consecutive clusters in Chorus 2, with no clusters from then on.

2.4 Developing Vocabulary

2.4.1 Naming Convention

Textual labels and diagrams are useful when communicating spatial information about specific shots. However, a notational system offers a more streamlined workflow for parsing data at the corpus scale as well as the possibility of notation-based tools for novel dance video generation. Thus, this work proposes the use of the Dance Videography Grammar labeling convention to signify Shot Size and Angle, and can be applied to both the video frames and also the camera-subject diagrams.

$$\text{Shotsize}_{\text{(Yaw) Pitch Angle}}{}^{7}$$

Fig. 29a - Dance Videography Grammar Proposed Labelling Convention

Type	Notation	Meaning
Shot Size	MC	Medium Closeup
	M	Medium

[7] If no yaw angle, leave blank. Back angles are considered yaw angles and listed as such, i.e. Back Left.

	MF		Medium Full
	F		Full
	L		Long
	EL		Extreme Long
Angle	Yaw	R	Right
		Le	Left
	Pitch	H	High
		S	Straight
		Lo	Low
		T	Top
	Additional	B	Back

Fig. 29b - Dance Videography Grammar Proposed Labelling Convention Key[8]

2.4.2 Shot Lexicon

Labelling Fig. 22 using the proposed Dance Videography Grammar naming convention constructs a 'shot lexicon' that encompasses all shots present in the *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* video.

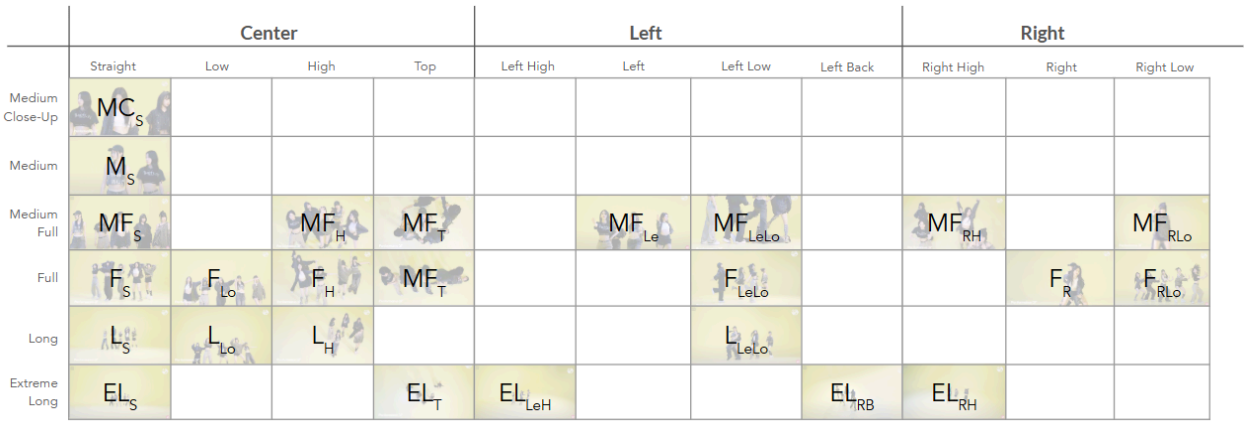


Fig. 30 - Shot Type Matrix with Labels for *[Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'*

[8] There are 3 parameters beginning with L - Long (L), Left (Le), Low (Lo) - which will be referred to with their abbreviations, respectively.

2.5 Corpus Analysis

2.5.1 Visual Analysis

Conducting the same shot type survey on the other two videos in the corpus generates the following shot type matrices which can be labelled according to the Dance Videography Grammar labelling convention.



Fig. 31 - Shot Type Matrix with Labels for *[Performance37] UNIS 'Dopamine'*



Fig. 32 - Shot Type Matrix with Labels for *[Performance37] izna 'IZNA'*

From Fig. 31 and Fig. 32, *[Performance37] UNIS 'Dopamine'* uses 29 shot types while *[Performance37] izna 'IZNA'* uses 26 shot types out of the 72 possible shot types. Overlaying these matrices reveals the spread - i.e. the pool of general candidate shot types - and the constants - i.e. the shot types that are present in all three videos.

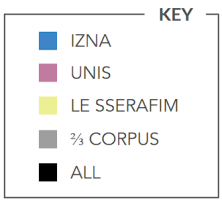


Fig. 33 - *[Performance37]* Video Corpus Key

	Center				Left			Right			
	Straight	Low	High	Top	Left High	Left	Left Low	Right High	Right	Right Low	Right Back
Medium Close-Up	XXX		X			X	X	XX	XX		
Medium	XXX		XXX		XX			X	X	X	
Medium Full	XXX	XX	XXX	X	XX	XX	X	XX		X	
Full	XXX	XXX	XXX	X	X		XXX	XX	X	X	
Long	XXX	XX	XXX	X	X		X	X		X	
Extreme Long	XXX		XX	X	XX			XX			X

Fig. 34a - Shot Type Matrix for *[Performance37]* Video Corpus

	Center				Left				Right		
	Straight	Low	High	Top	Left High	Left	Left Low	Left Back	Right High	Right	Right Low
Medium Close-Up	X		X			X	X		X	X	
Medium	X		X		X				X	X	X
Medium Full	X	X	X	X	X	X	X		X		X
Full	X	X	X	X	X		X		X	X	X
Long	X	X	X	X	X		X		X		X
Extreme Long	X		X	X	X			X	X		

Fig. 34b - Shot Types present in all *[Performance37]* Video Corpus videos

	Center				Left				Right		
	Straight	Low	High	Top	Left High	Left	Left Low	Left Back	Right High	Right	Right Low
Medium Close-Up	MC_S										
Medium	M_S		M_H								
Medium Full	MF_S		MF_H								
Full	F_S	F_{Lo}	F_H				F_{LeLo}				
Long	L_S		L_H								
Extreme Long	EL_S										

Fig. 34c - Shot Types present in all *[Performance37]* Video Corpus videos with Labels

Fig. 33a, 33b and 33c show a spread of 44 shot types, where 12 are present in all three corpus videos, and 24 present in at least two. In particular, the Straight and High angle and the Full shot size appear to be central to the *[Performance37]* style.

2.5.1.1 Shape Grammar Vocabulary

For each of the 12 requisite corpus shots, corresponding Camera-Subject diagrams can be produced showing the position, distance and angle of the camera relative to the subject. The number of diagram types (Plan, Front Elevation, Side Elevation) correspond to the 'complexity' of the shot type. Straight shot types are signified by only a Plan diagram, shot types with only pitch (e.g. High, Low) are signified by both Plan and Side Elevation diagrams, while shot types with both pitch and yaw are signified by Plan, Side Elevation and Front Elevation diagrams. These diagrams form the basis of the Dance Videography Grammar.

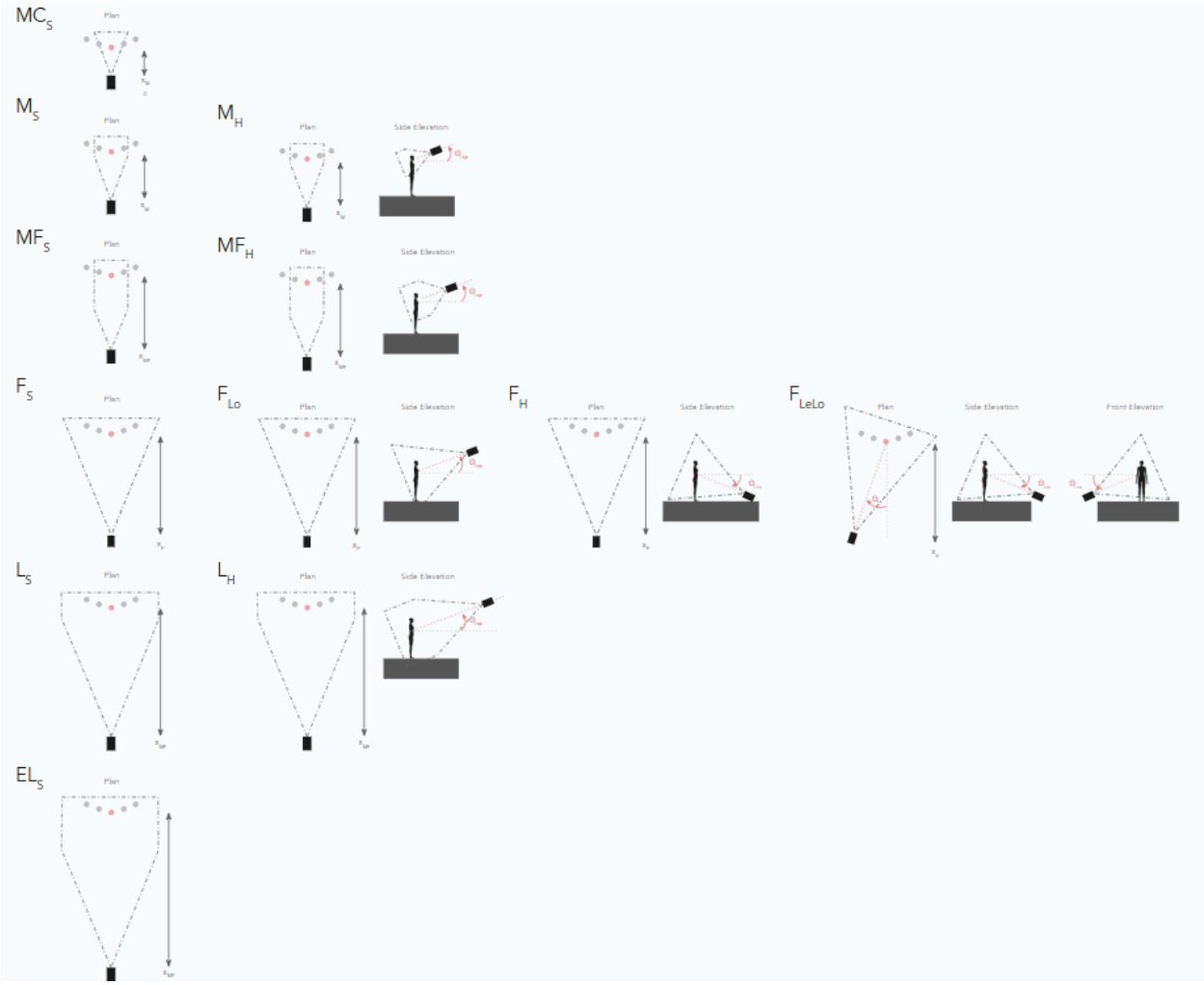


Fig. 35 - *[Performance37]* Vocabulary of Shots

2.5.2 Audio Analysis

2.5.2.1 Song Structure

The songs in the corpus differ in song structure. For instance, LE SSERAFIM'S '*Eve, Psyche & The Bluebeard's Wife'* has 12 sections, UNIS's *'Dopamine'* has 11 sections and IZNA's *'IZNA'* has 9 sections. Fig. 34 shows the number of cuts that occur across each song section, totalling to 119 shots, 160 shots and 128 shots for each video, respectively.

LE SSERAFIM			UNIS			IZNA		
Song Section	No. of shots	Duration (secs)	Song Section	No. of shots	Duration (secs)	Song Section	No. of shots	Duration (secs)
Intro	5	7.13	Intro	9	11.77	Intro	6	11.73
Refrain 1	9	13.87						

Verse 1	18	24.90	Verse 1	10	11.77	Verse 1	16	19.27
			Refrain 1	9	11.43			
			Pre-Chorus 1	18	21.67	Pre-Chorus 1	15	15.73
Chorus 1	9	15.23	Chorus 1	22	23.9	Chorus 1	18	20.03
Post-Chorus 1	8	13.33						
Refrain 2	8	13.33						
Verse 2	13	27.00	Verse 2	8	10.67	Verse 2	15	18.67
			Refrain 2	8	11.63			
			Pre-Chorus 2	15	21.6	Pre-Chorus 2	10	15.70
Chorus 2	10	13.23	Chorus 2	23	23.47	Chorus 2	15	20.90
Post-Chorus 2	8	13.70						
Bridge	8	14.73	Bridge	17	21.57	Bridge	9	15.07
Chorus 3	10	13.03	Chorus 3	21	27.97	Chorus 3	24	27.80
Outro	13	19.50						
Total	119	189.00		160	197.43		128	164.90

Fig. 36 - Song Sections and Associated Cuts for *[Performance37]* Video Corpus

2.5.2.2. Cut Distribution

Processing cut clusters for all corpus videos generates the graphs in Fig. 37a, where the biggest clusters generally occur in Chorus 3. Conversely, there are no or comparatively fewer clusters in the sections before Chorus 2 (Pre-Chorus 2 or Verse 2) and Bridge.

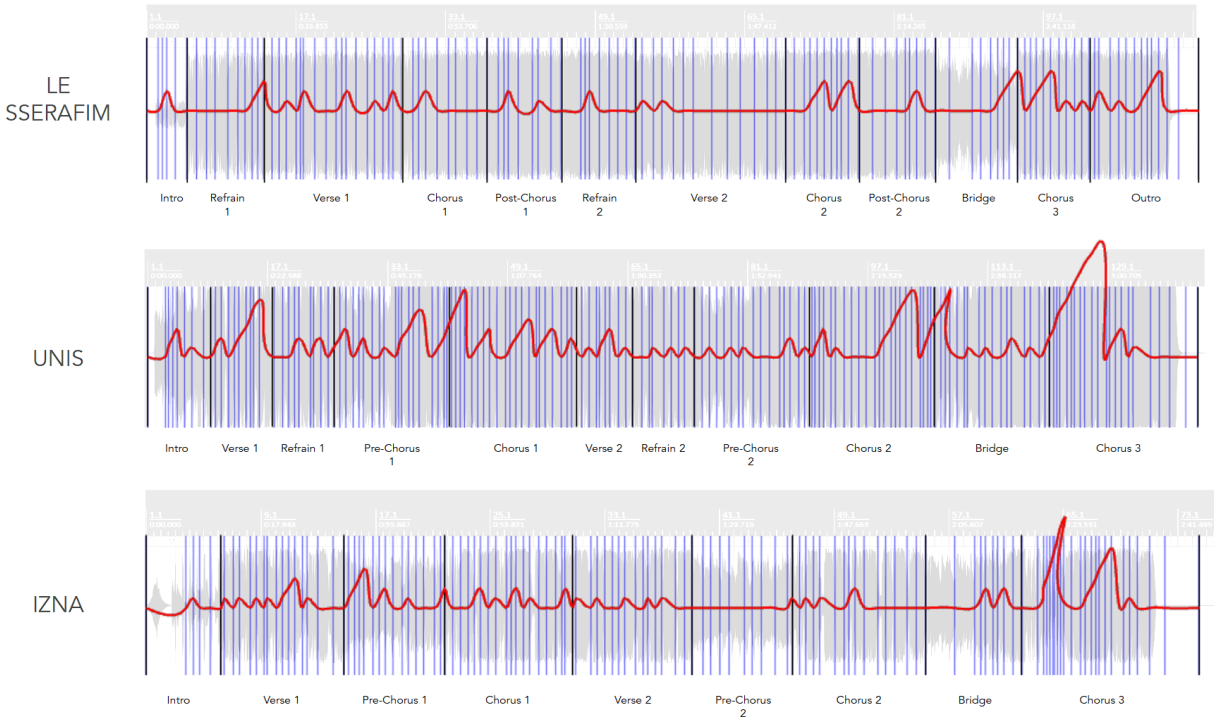


Fig. 37a - Cut Distribution and Associated Cut Cluster Lines for *[Performance37]* Video Corpus

When comparing the graphs, noticeably more cuts occur in *[Performance37] UNIS 'Dopamine'* than the other corpus videos. This aligns with the higher total shot count shown in Fig. 34 (160 shots). Not only are there more cuts, but the greater number of peaks show they are more closely distributed. A possible explanation considered for the increased clustering could be the BPM. UNIS '*Dopamine*' has a BPM of 170 while LE SSERAFIM's '*Eve, Psyche and the Bluebeard's Wife*' has a BPM of 143 and IZNA's '*IZNA*' has a BPM of 107. Re-examination of the Fig. under this context does not appear to suggest that there is a direct correlation between BPM, number of shots and cut clustering.

2.5.3 Audio-Visual Analysis

2.5.3.1 Shot Size Distribution

Fig 36. shows the shot size distribution graphs for all of the corpus videos. While the information in these graphs is not immediately legible, they effectively communicate the concept of distinct compositional differences between videos.

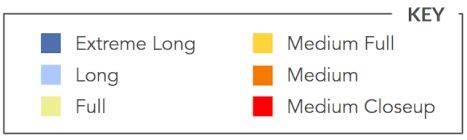


Fig. 25 - Shot Size Key

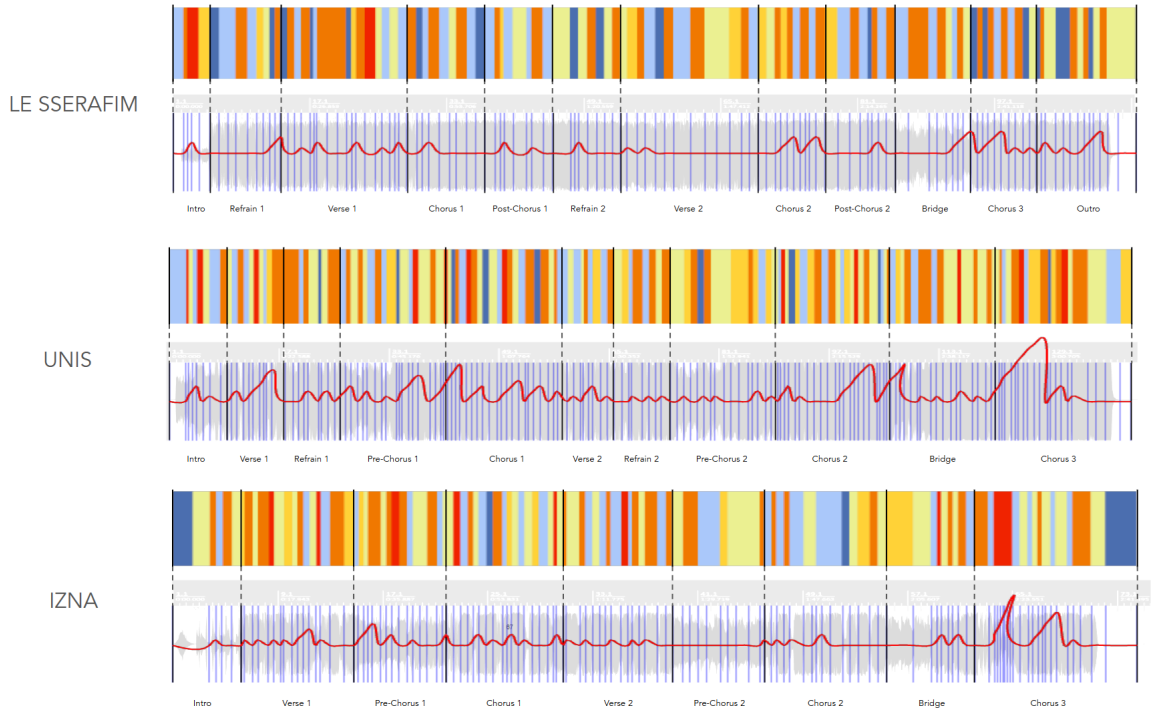


Fig. 38 - Shot Size Distribution for all Shots in [*Performance37*] Video Corpus

The variation in distribution is particularly clear when separated into 'close' and 'far' shot types as shown in Fig 39a. and Fig 39b.

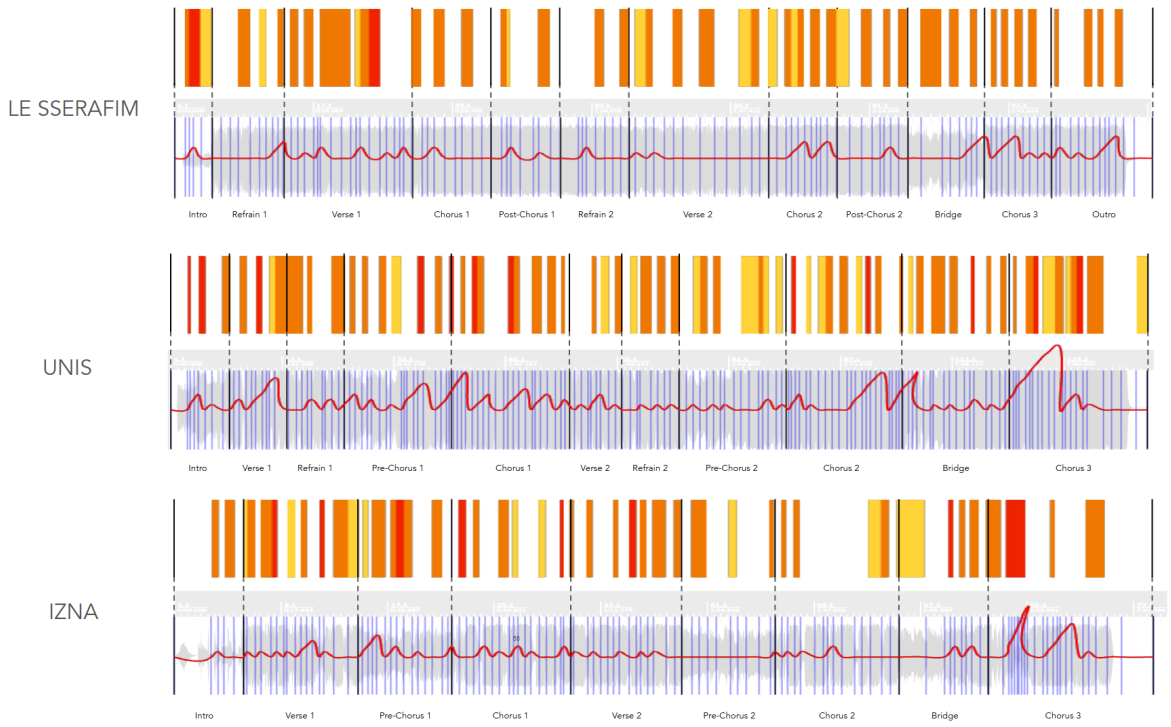


Fig 39a. Shot Size Distribution for Close Shots in [*Performance37*] Video Corpus

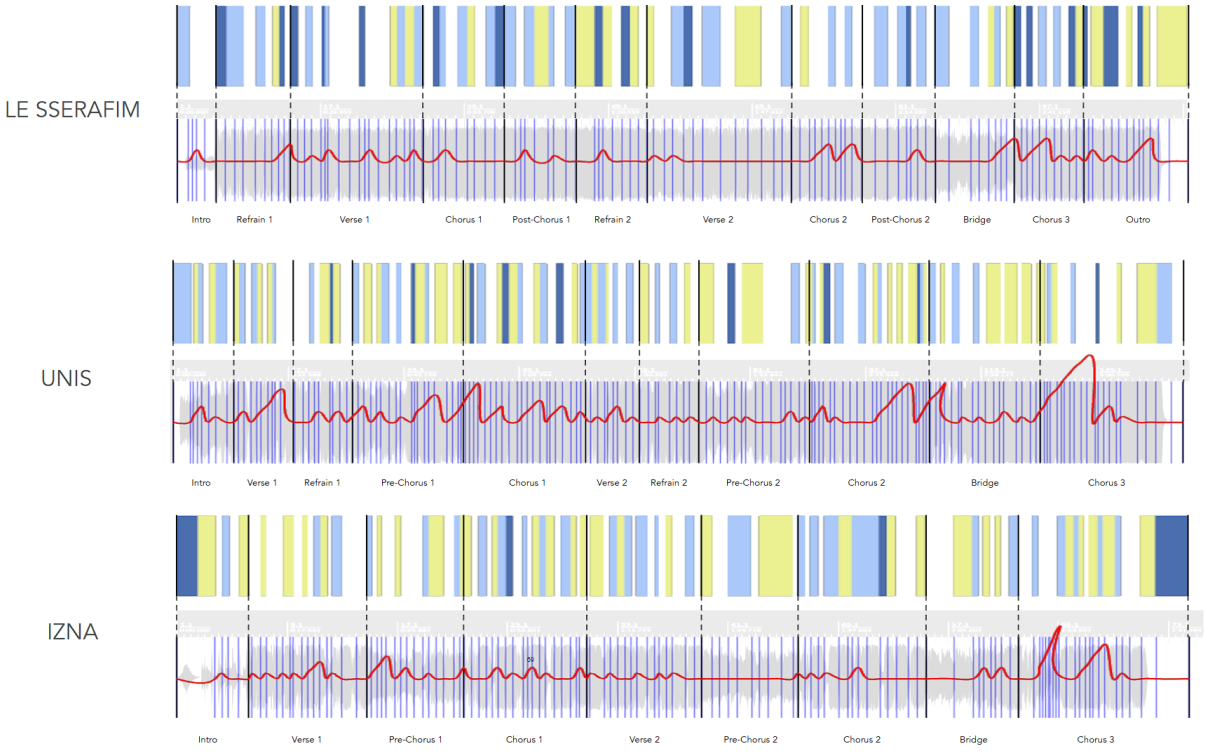
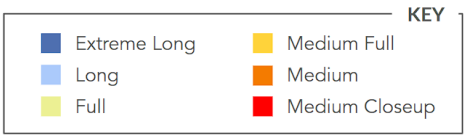


Fig 39b. Shot Size Distribution for Far Shots in [*Performance37*] Video Corpus

Cutting directly between the same shot size is a cinematographic technique called a jump cut. It is considered "abrupt, disorienting" and used "to create discontinuity for artistic effect". (“Film Terms Glossary.”) In a shot distribution graph such as Fig. 38, Fig. 39a and Fig. 39b, and is indicated by a section of substantial length dominated by the same color. Overall, the *[Performance37]* corpus does not exhibit this trait. The only exceptions are Pre-Chorus 2, Chorus 2 and Chorus 3 in *[Performance37] izna 'IZNA'*; Pre-Chorus 2 and Chorus 2 are not considered jump cuts due to differing angles not represented in this graph type, while Chorus 3 features a fade between each shot and is a deliberate use of match cuts to highlight each individual dancer.

2.5.3.2 Section Starting & Ending Shot Size

Fig 40a. shows each video broken down via musical section, with the first and last shot of each section labeled and color coded according to shot size.



LE SSERAFIM				UNIS			IZNA	
Song Section	Starting Shot	Ending Shot		Starting Shot	Ending Shot		Starting Shot	Ending Shot

Instrumental	L_S	MF_S		L_{Lo}	M_S		EL_S	F_S
Refrain 1	EL_S	M_S						
Verse 1	EL_S	L_{Lo}		F_{RH}	M_H		MF_S	MF_S
Refrain 1				M_S	M_S			
Pre-Chorus 1				L_S	MC_S		L_S	F_H
Chorus 1	$S_{M\text{-}S}$	$S_{EL\text{-}S}$		$S_{MC\text{-}S}$	$S_{F\text{-}Lo}$		$S_{F\text{-}S}$	$S_{F\text{-}Lo}$
Post-Chorus 1	$S_{L\text{-}Lo}$	$S_{L\text{-}S}$						
Refrain 2	$S_{F\text{-}H}$	$S_{M\text{-}S}$						
Verse 2	$S_{F\text{-}Lo}$	$S_{L\text{-}Lo}$		$S_{F\text{-}Lo}$	$S_{M\text{-}S}$		$S_{M\text{-}S}$	$S_{M\text{-}H}$
Refrain 2				$S_{F\text{-}Lo}$	$S_{M\text{-}S}$			
Pre-Chorus 2				$S_{F\text{-}Lo}$	$S_{F\text{-}Lo}$		$S_{F\text{-}H}$	$S_{M\text{-}S}$
Chorus 2	$S_{MF\text{-}RH}$	$S_{M\text{-}S}$		$S_{F\text{-}Lo}$	$S_{M\text{-}S}$		$S_{L\text{-}S}$	$S_{F\text{-}Lo}$
Post-Chorus 2	$S_{MF\text{-}T}$	$S_{M\text{-}S}$						
Bridge	$S_{L\text{-}S}$	$S_{F\text{-}S}$		$S_{L\text{-}S}$	$S_{F\text{-}RLo}$		$S_{MF\text{-}S}$	$S_{L\text{-}S}$
Chorus 3	$S_{EL\text{-}S}$	$S_{F\text{-}RLo}$		$S_{F\text{-}S}$	$S_{MF\text{-}S}$		$S_{M\text{-}S}$	$S_{EL\text{-}S}$
Outro	$S_{L\text{-}S}$	$S_{F\text{-}Lo}$						

Fig 40a. Section Start and End Shot Size in [*Performance37*] Corpus

There are three ways a section can progress:

1. Far[9] ⟶ Close[10]
2. Close ⟶ Far
3. Same ⟶ Same

Fig 40b. shows each section color coded according to the relationship between start and end shot size.

LE SSERAFIM				UNIS			IZNA	
Song Section	Starting Shot	Ending Shot		Starting Shot	Ending Shot		Starting Shot	Ending Shot
Intro	$S_{L\text{-}S}$	$S_{MF\text{-}S}$		$S_{L\text{-}Lo}$	$S_{M\text{-}S}$		$S_{EL\text{-}S}$	$S_{F\text{-}S}$
Refrain 1	$S_{EL\text{-}S}$	$S_{M\text{-}S}$						
Verse 1	$S_{EL\text{-}S}$	$S_{L\text{-}Lo}$		$S_{F\text{-}RH}$	$S_{M\text{-}H}$		$S_{MF\text{-}S}$	$S_{MF\text{-}S}$

[9] Starting shot distance > Ending shot distance
[10] Starting shot distance < Ending shot distance

Refrain 1				$S_{M\text{-}S}$	$S_{M\text{-}S}$			
Pre-Chorus 1				$S_{L\text{-}S}$	$S_{MC\text{-}S}$		$S_{L\text{-}S}$	$S_{F\text{-}H}$
Chorus 1	$S_{M\text{-}S}$	$S_{EL\text{-}S}$		$S_{MC\text{-}S}$	$S_{F\text{-}Lo}$		$S_{F\text{-}S}$	$S_{F\text{-}Lo}$
Post-Chorus 1	$S_{L\text{-}Lo}$	$S_{L\text{-}S}$						
Refrain 2	$S_{F\text{-}H}$	$S_{M\text{-}S}$						
Verse 2	$S_{F\text{-}Lo}$	$S_{L\text{-}Lo}$		$S_{F\text{-}Lo}$	$S_{M\text{-}S}$		$S_{M\text{-}S}$	$S_{M\text{-}H}$
Refrain 2				$S_{F\text{-}Lo}$	$S_{M\text{-}S}$			
Pre-Chorus 2				$S_{F\text{-}Lo}$	$S_{F\text{-}Lo}$		$S_{F\text{-}H}$	$S_{M\text{-}S}$
Chorus 2	$S_{MF\text{-}RH}$	$S_{M\text{-}S}$		$S_{F\text{-}Lo}$	$S_{M\text{-}S}$		$S_{L\text{-}S}$	$S_{F\text{-}Lo}$
Post-Chorus 2	$S_{MF\text{-}T}$	$S_{M\text{-}S}$						
Bridge	$S_{L\text{-}S}$	$S_{F\text{-}S}$		$S_{L\text{-}S}$	$S_{F\text{-}RLo}$		$S_{MF\text{-}S}$	$S_{L\text{-}S}$
Chorus 3	$S_{EL\text{-}S}$	$S_{F\text{-}RLo}$		$S_{F\text{-}S}$	$S_{MF\text{-}S}$		$S_{M\text{-}S}$	$S_{EL\text{-}S}$
Outro	$S_{L\text{-}S}$	$S_{F\text{-}Lo}$						

Fig 40b. Section Start and End Relative Shot Size in [*Performance37*] Corpus

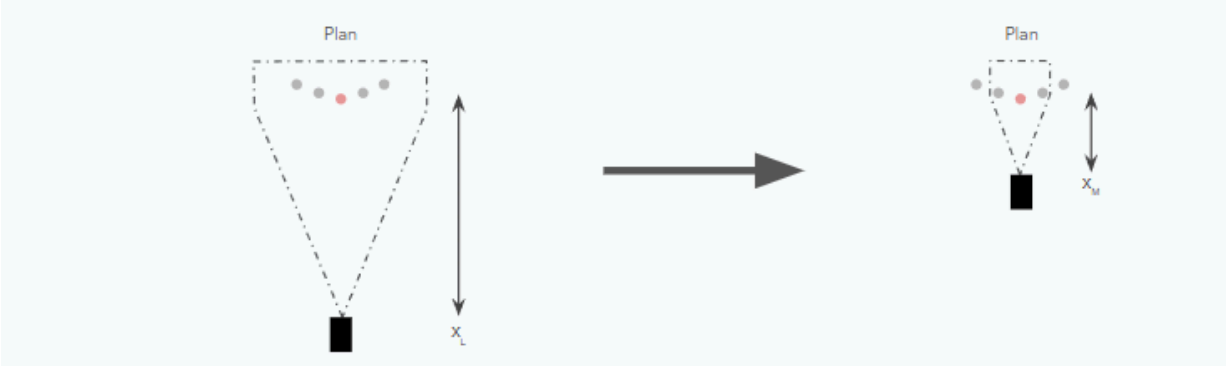
The overwhelming majority of sections display the 'Far ⟶ Close' relationship; Intro and Chorus 2 consistently display this progression in all corpus videos.

2.6 Shape Grammar Ruleset

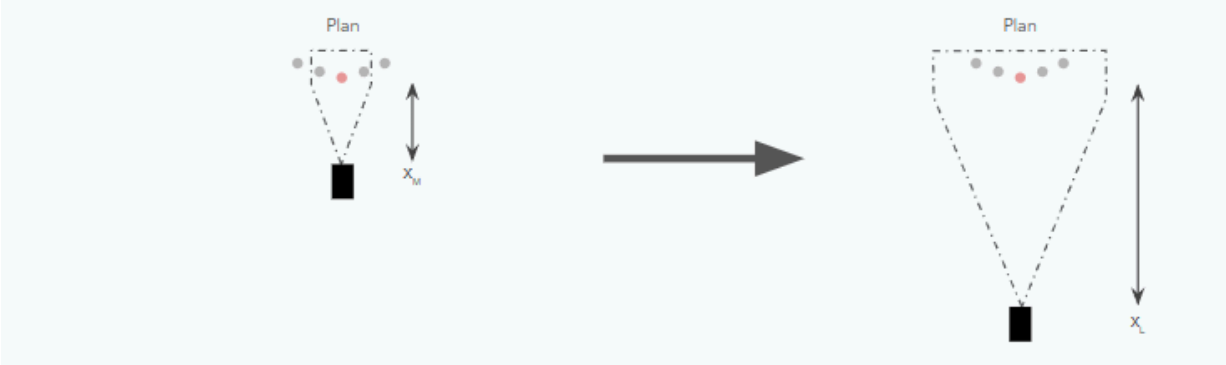
2.6.1 Dance Videography Shape Grammar for [Performance37] Style

From the corpus analysis, a ruleset describing the *[Performance37]* videography style can be developed. There are 4 rules defining the *[Performance37]* style, articulated below using the vocabulary produced with Camera-Subject diagrams.

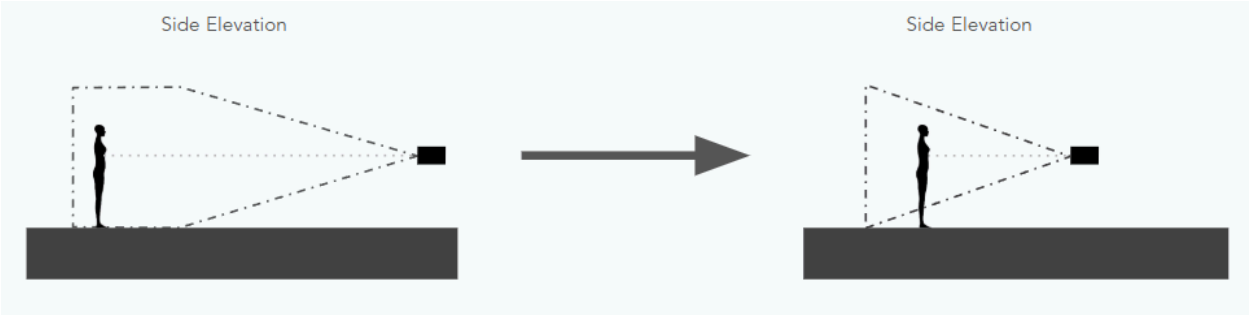
1. Shot Size Relations
    a. Starting shot (L) ⇒ (M)



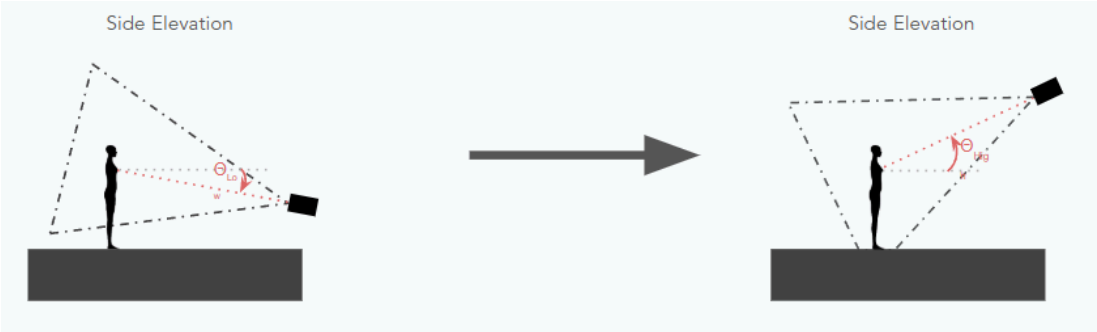
b. Starting shot (M) ⇒ (L)



2. Shot Angle Relations
   a. Starting Angle (S) ⇒ (S)

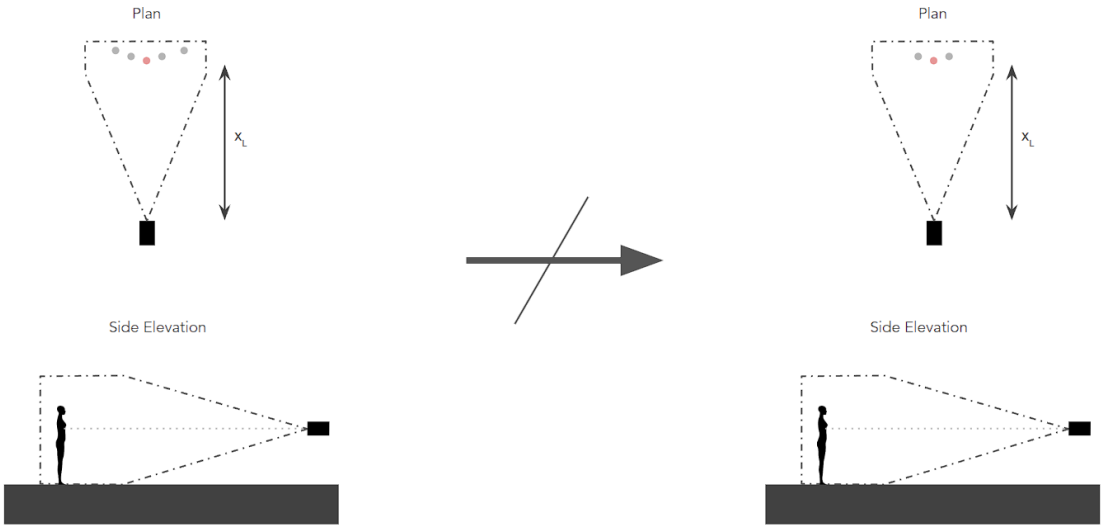


b. Starting Angle (Lo) ⇒ (H)

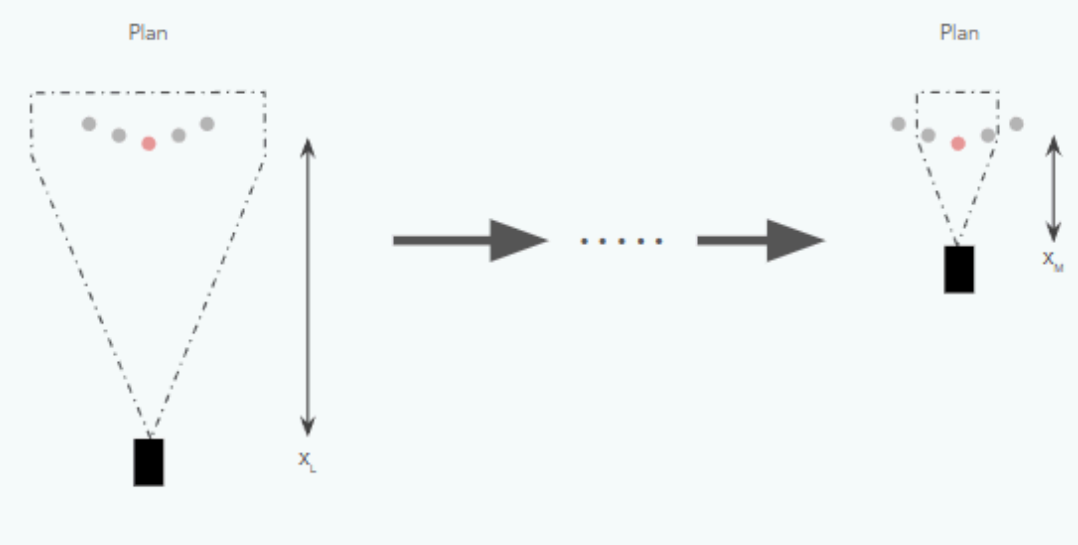


3. No Jump Cuts

a. Starting Shot (L_S) ⇏ (L_S)
    i. This rule can be intentionally broken for artistic effect, but only in specific circumstances.



4. Section Progression (i ⟶ f)
    a. Initial Shot (EL), (L), or (F) ⇒ … ⇒ Final Shot (MC), (M) or (MF)
        i. Most musical sections exhibit this relationship, only Chorus 1 does not.
        ii. Intro & Chorus 2 always obey this rule



Any video produced following these rules using the vocabulary of shots as detailed in Fig. 35 should exhibit the Performance37 style.

3 Results

A dance performance video for the song *Sad Song* by *P1Harmony* was created following the *[Performance37]* rules. This song and dance was chosen as it was not present within the greater *[Performance37]* corpus.

Links:
Demo Video: https://youtu.be/-TgaAZkK384
Demo Video with Annotations: https://youtu.be/Cx2uqtUnOSY

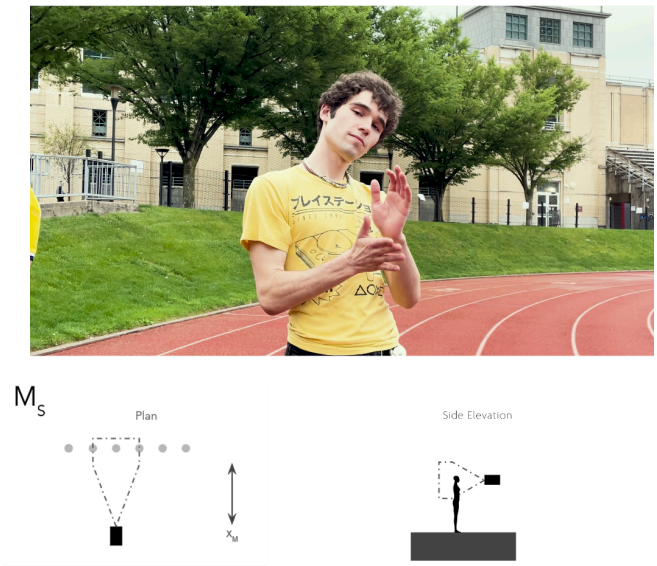


Fig 41a. Sample Frame from Generated [Performance37] Style Video for P1Harmony 'Sad Song' with Notation (M_S)



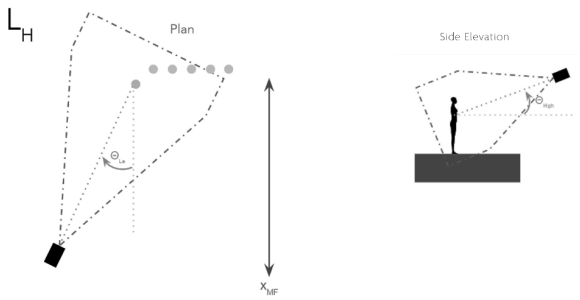


Fig 41b. Sample Frame from Generated [Performance37] Style Video for P1Harmony 'Sad Song' with Notation (L_H)

3.1 Video Production

Production of this video was planned and conducted using the Dance Videography Notation System via the Annotated Storyboard Tool. Reference frames taken from the *P1Harmony (*피원하모니*) 'SAD SONG' Dance Practice* video and the P1Harmony (피원하모니) - 'SAD SONG' MV were cropped and labelled according to desired shot type, with timestamps specifying duration of each shot. ("P1Harmony (피원하모니)- 'SAD SONG' MV.") ("P1Harmony (피원하모니) 'SAD SONG' Dance Practice.")

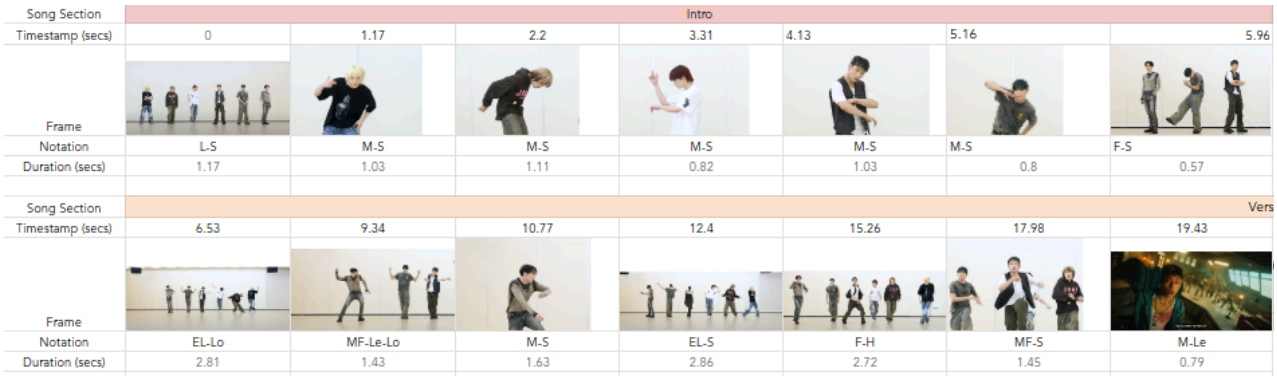


Fig 42. Annotated Storyboard for P1H Sad Song

This format makes the relationship between successive shots, allowing for easy application of [*Performance37*] rules.

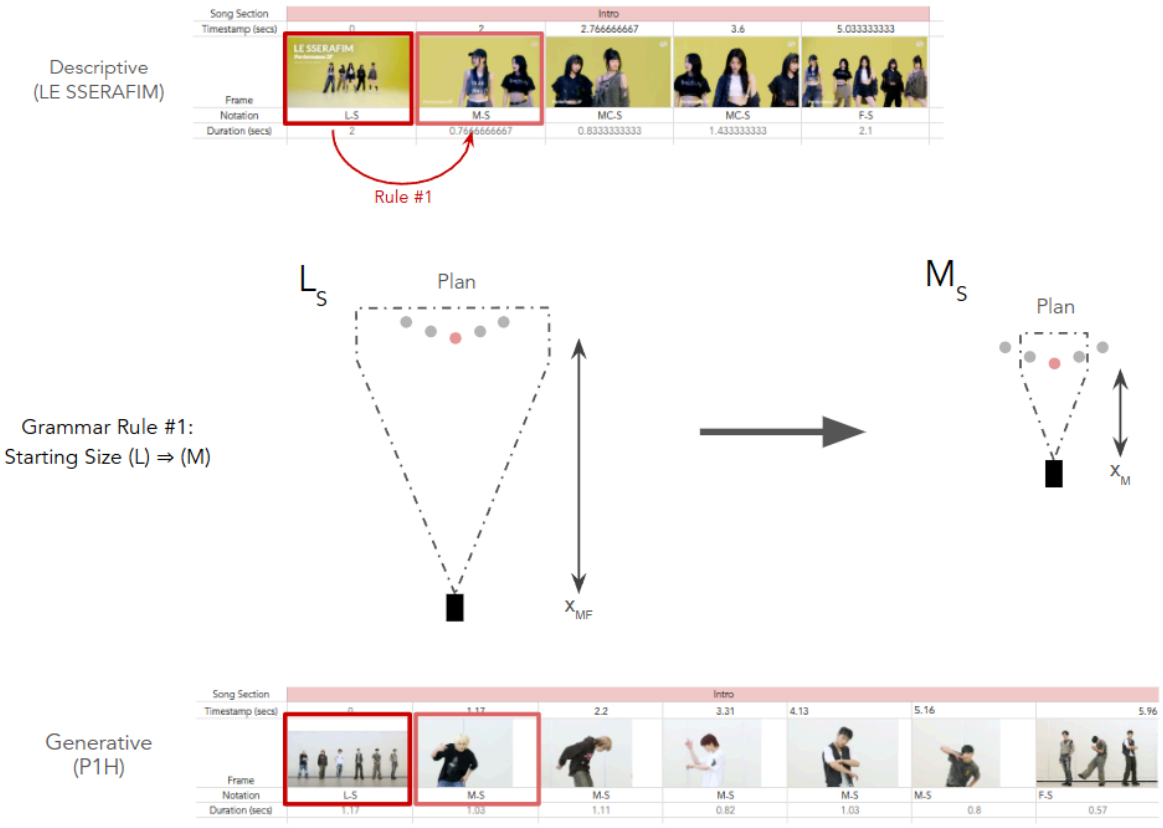


Fig 42a. Application of [*Performance37*] Rule #1 in pre-production of *Sad Song* by P1H

There were many benefits afforded by this workflow. As a videographer, I found it intuitive to read, easy to iterate upon, and quicker to produce than previous methods. Most importantly, I could grasp the structure of entire sections of the video at once, which allowed me to better tailor the overall flow of the video and plan for the shoot. This level of design agency in my work is unprecedented, as historically a clear understanding of the video only emerges after the shoot during post-production.

3.2 Video Evaluation

This video follows the [*Performance37*] videographic style. As a proof of concept, it illustrates the utility of the Dance Videography Grammar as well as the viability of the proposed tools and workflow. While this video may not fit seamlessly into the [*Performance37*] catalog due to production-level differences such as lighting, dance quality and camera quality, nevertheless, it feels like a [*Performance37]* video. Thus, this work demonstrates that shape grammars are a successful method of codifying and communicating the tacit technical knowledge embedded in videography.

4 Contributions

This work offers four main contributions:

1. *[Performance37]* Style Rules

The *[Performance37]* Style Rules explicate the tacit technical knowledge embedded in a *[Performance37]* video. In describing how these videos are made, they allow for novel applications of the style.

2. Dance Videography Grammar
    a. Shot Type Labelling & Notation System

The Dance Videography Grammar is a framework that allows for the description of any video, and thus any videographic style, composed of purely static shots. While this work applies the grammar on the *[Performance37]* corpus of videos to yield the [Performance37] Style Rules, application of this grammar on a different corpus will yield a set of different rules.

3. Annotated Storyboard Format

The Annotated Storyboard Format is a generative tool that can be leveraged to produce videos. The layout makes the applications of style rules easier than conventional methods, but they can also be used on their own as exploratory production tools.

4. Pipeline for Explicating Videographic Style

Overall, the Methodology outlined in this work provides a pipeline for explicating tacit videographic knowledge, and need not be limited to dance performance videos.

4.1 Limitations & Next Steps

Outside of *[Performance37]*, Lee Hanhyung is the lead producer for various other video series depicted in Fig. 13, such as *Studio Choom* and *Move to Performance*. In particular, *Studio Choom* is characterized by smooth sweeping camera movements and visually striking lighting design. Expanding the Dance Videography Grammar and Notation system to include additional factors like lighting, camera movements, editing and visual effects would not only allow for comparisons between these video series, but also the description of a greater 'Lee Hanhyung' videographic style.

Implementing a system for describing shots where the camera is moving (dynamic shots) versus the established vocabulary of static shots is a vital next step in increasing the use cases for the Dance Videography Grammar.

Future development of this work could enable the description and generation of a 'style DNA' where the shot distribution graph convention is leveraged to create visual representations of different videographic styles. This would not just be limited to the styles of different practitioners, but also as a

way of representing and comparing the differences within a single practitioner's work over the span of their career.

Additionally, this framework for categorizing shot types can be vectorized to provide a framework for Frame-by-Frame Analysis for Machine Learning Applications. Automation of the shot categorisation and labelling process appears to be a natural next step.

5 Conclusion

In conclusion, a notation system can allow for the analysis, design and generation of dance performance videos by encoding the tacit videographic knowledge in a system of shapes and shape rules. Following the shape grammar formalism allows the proposed Dance Videography grammar to be skill agostic, circumventing the need for additional training and years of technical expertise.

Traditional forms of communicating videographic paradigms use words to communicate images, which is labor and time-intensive to both produce and understand. Additionally, structural changes in the overall creative vision mean new versions of these documents must be re-created to match the new direction. As a visual language, the proposed Dance Videography Grammar is not only more efficient at communicating visual intent, but also provides a more streamlined method of updating production tools. Instead of specifying a whole new shot, replacement shots can be pulled from the existing shot type vocabulary pool, and adjustments visualised and made accordingly. Additionally, the proposed formats offer quality-of-life upgrades for practitioners. For instance, the Annotated Storyboard format allows for entire sub-sequences to be viewed at once, and automatically updates shot durations whenever timestamps are altered.

This work democratizes the craft of videography and expands existing ways of thinking by promoting camera movement literacy. It makes the work of experts accessible by providing a framework through which their work can be broken down, analysed and described. Through recreating, iterating and generating videographic styles without formal training or direct access to an expert, users have more agency to explore the boundaries of the medium.

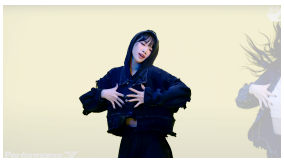
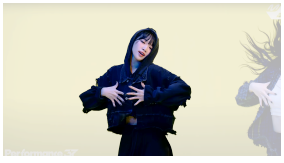
As a practitioner, I find this work exciting. There are many videographic styles to analyse, techniques to describe, dimensions to fold into the Dance Videography Grammar. This system was developed out of a desire to use this work to inform my growth as a videographer. Ultimately, my aspiration for this work is to one day use it to describe a 'Clover Chau' videographic style.

6 Appendix

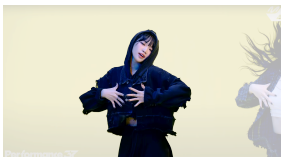
6.1 Transitions

6.1.1 Shot Types

Transition Type			S_1 Example Frame	S_2 Example Frame	Count	% of Total Cuts
Medium Closeup ($S_{MC-?}$)	→	Medium Closeup ($S_{MC-?}$)			1	0.84%
	→	Medium ($S_{M-?}$)			0	0.00%
	→	Medium Full ($S_{MF-?}$)			1	0.84%
	→	Full ($S_{T-?}$)			1	0.84%
	→	Long ($S_{L-?}$)			0	0.00%
	→	Extreme Long ($S_{EL-?}$)			0	0.00%

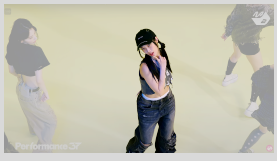
Medium ($S_{M-?}$)	→	Medium Closeup ($S_{MC-?}$)			2	1.68%
		Medium ($S_{M-?}$)			3	2.52%
		Medium Full ($S_{MF-?}$)			3	2.52%
		Full ($S_{T-?}$)			7	5.88%
		Long ($S_{L-?}$)			16	13.45%
		Extreme Long ($S_{EL-?}$)			8	6.72%
Medium Full ($S_{MF-?}$)	→	Medium Closeup ($S_{MC-?}$)			0	0.00%
		Medium ($S_{M-?}$)			4	3.36%

		Medium Full ($S_{MF-?}$)			0	0.00%
		Full ($S_{T-?}$)			2	1.68%
		Long ($S_{L-?}$)			2	1.68%
		Extreme Long ($S_{EL-?}$)			1	0.84%
	→	Medium Closeup ($S_{MC-?}$)			0	0.00%
		Medium ($S_{M-?}$)			6	5.04%
		Medium Full ($S_{MF-?}$)			2	1.68%
Full ($S_{T-?}$)		Full ($S_{T-?}$)			3	2.52%

		Long ($S_{L-?}$)			4	3.36%
		Extreme Long ($S_{EL-?}$)			5	4.20%
	→	Medium Closeup ($S_{MC-?}$)			0	0.00%
		Medium ($S_{M-?}$)			18	15.13%
		Medium Full ($S_{MF-?}$)			2	1.68%
		Full ($S_{T-?}$)			4	3.36%
		Long ($S_{L-?}$)			1	0.84%
Long ($S_{L-?}$)		Extreme Long ($S_{EL-?}$)			3	2.52%

	→	Medium Closeup ($S_{MC-?}$)			0	0.00%
		Medium ($S_{M-?}$)			8	6.72%
		Medium Full ($S_{MF-?}$)			1	0.84%
		Full ($S_{T-?}$)			4	3.36%
		Long ($S_{L-?}$)			4	3.36%
Extreme Long ($S_{EL-?}$)		Extreme Long ($S_{EL-?}$)			2	1.68%
Total					118	100.00%

6.1.2 Angles

Transition Type			S_1 Example Frame	S_2 Example Frame	Count	% of Total Cuts
Straight Shot ($S_{?-S}$)	→	Straight Shot ($S_{?-S}$)			41	34.75%

	→	Angled Shot ($S_{?-A}$)			31	26.27%
Angled Shot ($S_{?-A}$)	→	Angled Shot ($S_{?-A}$)			16	13.56%
	→	Straight Shot ($S_{?-S}$)			30	25.42%
Total					118	100.00%

6.2 Additional Audio Analysis

6.2.1 Music as Videographic Driver

Both Chorus 2 and Chorus 3 share the same number of cuts and a similar duration yet differ in cut clustering. Fig. 24b shows Chorus 2 with two equal peaks, and Chorus 3 with one large peak and two smaller peaks, suggesting a different factor is driving cut distribution.

According to David Bennett, 94% of pop music, including *'Eve, Psyche and the Bluebeard's Wife',* is in 4/4 time. Within a given measure or bar, there are four beats, each of which can be divided into half-beats (½), quarter-beats (¼), quavers (⅛), sixteenth-beats (1/16), even thirty-secondths (1/32).Similarly, the primary method of counting in dance is called an '8-count. Similarly, the primary method of counting dance is called an '8-count' that breaks two bars of 4 beats into quarter beats[11]. The following tables shows 4 counts of the systems aligned:

1								2								3								4						
1	-	e	-	and	-	a	-	2	-	e	-	and	-	a	-	3	-	e	-	and	-	a	-	4	-	e	-	and	-	a	-
1	⅛	¼	⅜	½	⅝	¾	⅞	2	⅛	¼	⅜	½	⅝	¾	⅞	3	⅛	¼	⅜	½	⅝	¾	⅞	4	⅛	¼	⅜	½	⅝	¾	⅞

Opening the music file in an audio processing software like Reaper, the placement of beats within the song can be visualised over time with regards to beats in a bar. The intro is the least sonically dense part of the song, with 13 beats within the first 2 bars or '8 count' - 7 in the first bar, 6 in the second. The beats present within the music are: 1, ³⁄₁₆, ⅜, ½, ⅝, ¹¹⁄₁₆, ⅞.

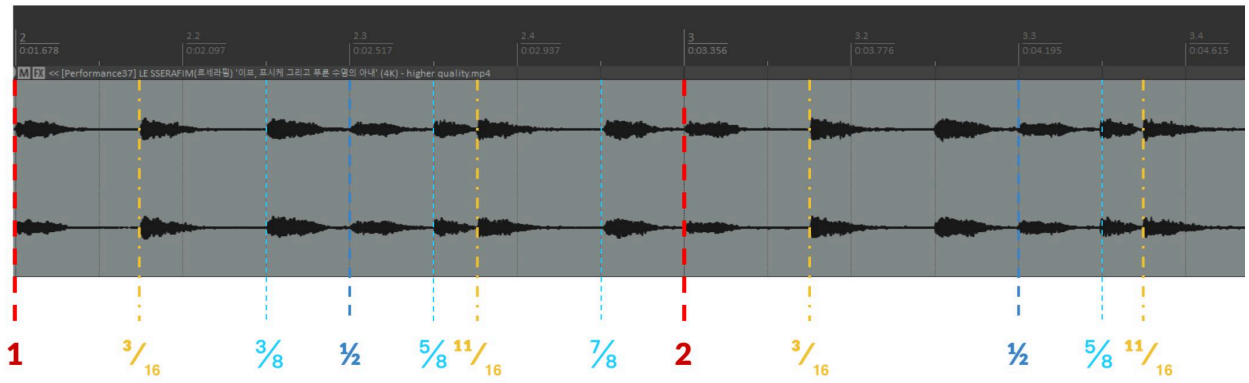


Fig. Audio Analysis of *'Eve, Psyche and the Bluebeard's Wife'* Intro

This pattern is carried through Refrain 1, with an additional ¾ beat added and the ⅝ beat removed. Thus, the complete list of beats present are: 1, ³⁄₁₆, ⅜, ½, ⅝, ¹¹⁄₁₆, ¾, ⅞.

[11] Musical counts go into smaller divisions like ¹⁄₃₂ while dance counts do not go farther than ¼ , due to the physical limitations of the time taken to execute a dance move.

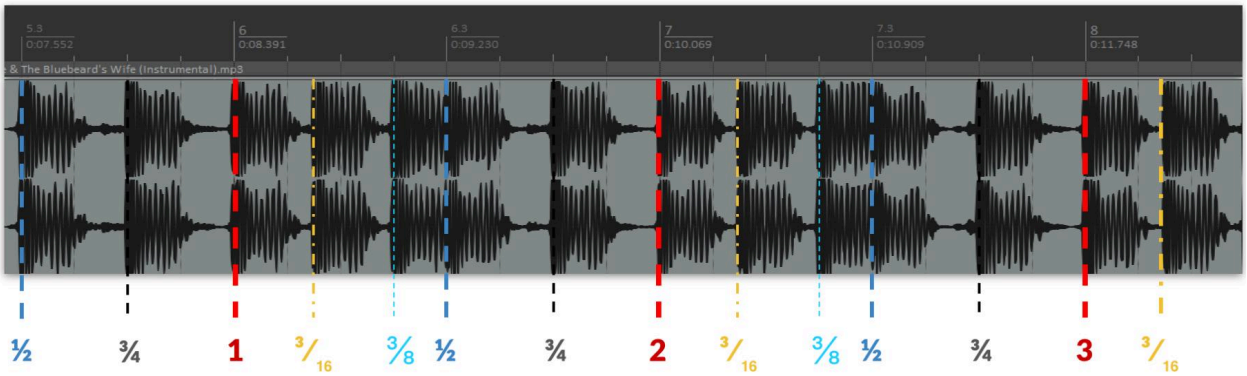


Fig. Audio Analysis of *'Eve, Psyche and the Bluebeard's Wife'* Refrain 1

6.2.2 Defining Beat Conditions

Discerning whether or not a cut occurs 'on beat' requires the following knowledge:

1. (Temporal) Position of the cut
2. (Temporal) Position of the nearest beats

Humans visually perceive cuts at a slower rate than they audibly perceive beats in the music. This phenomenon is explained in *Effects of visual, auditory, and combined cues on human movement and brain regions involved in perception action*: "Reaction time to auditory cues is faster by 20-50 milliseconds compared to visual or tactile cues." (Mate. 2022.) Consequently, if a cut and a beat occur at the same time, the cut will feel 'late'. To account for this, a cut will be considered 'on beat' if it occurs either 1 frame (0.033 secs)[12] before or on the same frame as the beat. Additionally, the absolute start of the music in dance performance videos may not align with the absolute start of the original track. Fortunately in the case of *'Eve, Psyche and the Bluebeard's Wife'*, the music is offset by one full beat exactly, meaning the timecodes do not need to be adjusted. For each cut, the following function can be used to determine if a given cut occurs 'on' a given beat:

If ($Beat_A$ - ⅓₀ <= (($Timecode_{Cut\ 1}$ - Offset) % 0.419) <= $Beat_A$) == True,

Cut_A = On $Beat_A$,

Else,

Cut_A != On $Beat_A$

A full table for showing the breakdown for each individual cut can be found in Appendix 3; Fig. summarizes these results, showing 68 of the 119 cuts (42.77%) are 'on beat'. These findings were corroborated by comparing the Appendix 3 results against the author's own perception of the [*Performance37] LE SSERAFIM 'Eve, Psyche & The Bluebeard's Wife'* video.

Beat Type	Count
1 beat	9

[12] The video is filmed in 30 frames per second, thus the duration of one frame is 1/30 or 0.033 secs).

Beat Type	Count
½ beat	5
¼ beat	7
⅛ beat	34
1/16 beat	13
Total No. of Cuts 'On Beat'	68
Total No. of Cuts Not 'On Beat'	50
Total	119

Focusing on the cuts within Chorus 2 and Chorus 3, six out of ten are 'on beat' in Chorus 2 compared to eight out of ten in Chorus 3.

Song Section	No. of Cuts	No. of Cuts On Beat	% of Cuts on Beat	Duration (secs)
Chorus 2	10	6	60.00%	13.23
Chorus 3	10	8	80.00%	13.033

These results can be visually represented by plotting these cuts along their respective waveforms.

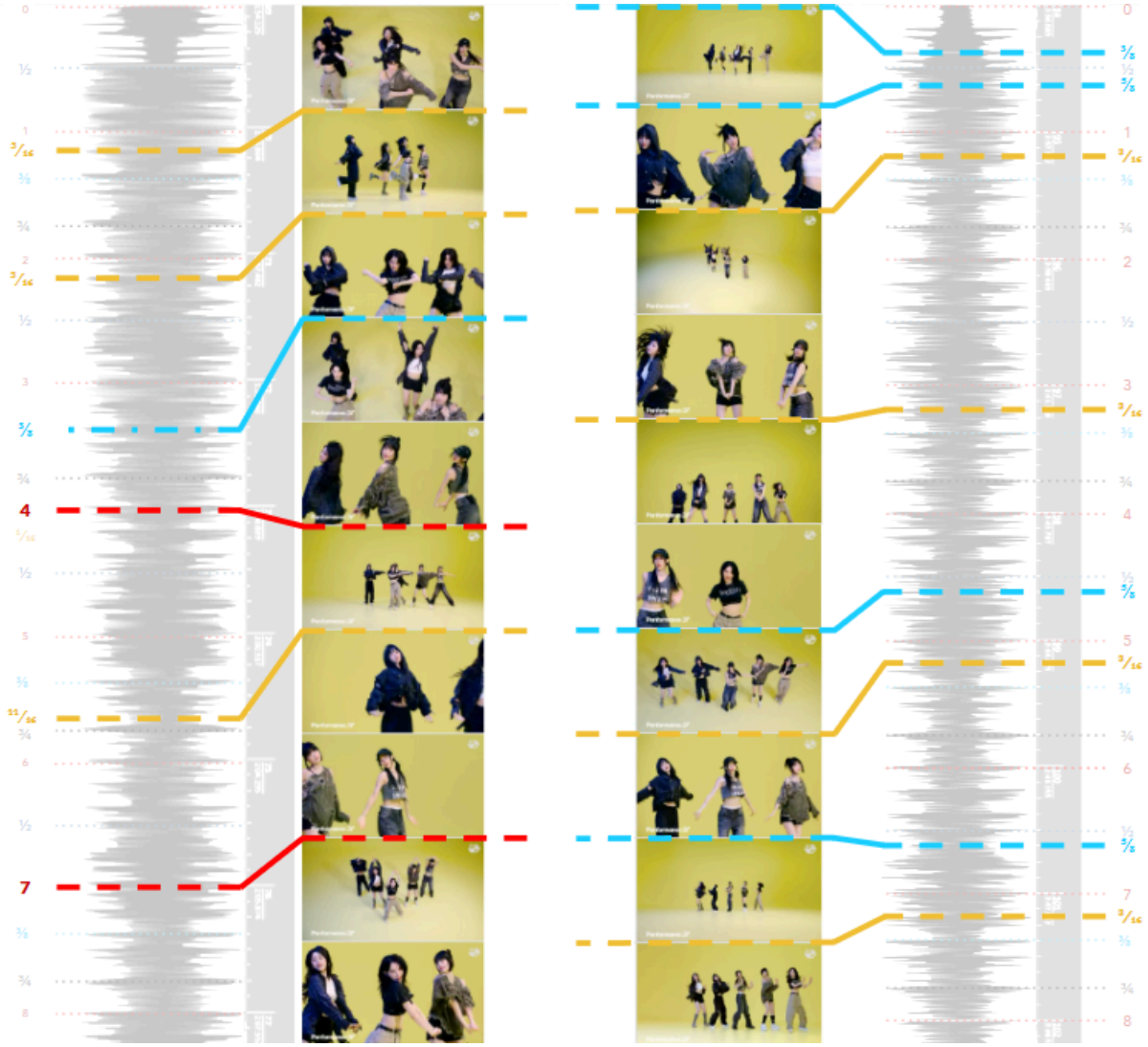


Fig. Position of Cuts in Music in Chorus 2 versus Chorus 3

	Chorus 2	Chorus 3
1 beat	2	0
⅛ beat	1	4
1/16 beat	3	4

Fig. Beat Types for On-Beat Cuts in Chorus 2 versus Chorus 3

While the duration of the sections are not exactly the same, both span two bars of 4, or one 8-count[13]. Given the same general conditions - number of beats in the music, moves in the choreography - the distribution, and crucially, the beat type of the cuts are different. Two of the cuts in Chorus 2 land on whole beats while the on-beat cuts in Chorus 3 land exclusively on ⅛ and 1/16 beats.

[13] The difference in duration is due to two lines of additional lyrics present in Chorus 3 compared to Chorus 2.

Works Cited

Alfaro-Contreras, María, et al. “Decoupling Music Notation to Improve End-to-End Optical Music Recognition.” *Pattern Recognition Letters*, vol. 158, June 2022, pp. 157–63. *DOI.org (Crossref)*, https://doi.org/10.1016/j.patrec.2022.04.032.

Alony, Irit, et al. “Sharing tacit knowledge: A case study in the australian film industry.” InSITE Conference, vol. 7, 2007, https://doi.org/10.28945/3119.

Al-Saati, Maha. Moving in Filmic Spaces: Relating Camera Movements to Spatial Archetypes in Architectural Animations. 2012.

Arijon, Daniel. *Grammar of the Film Language*. Silman-James Press, 1991.

Arthurs, Yoko, and Karin Petrini. “Musicians’ Views on the Role of Reading Music in Learning, Performance, and Understanding.” *Musicae Scientiae*, vol. 28, no. 1, Feb. 2023, pp. 3–17, https://doi.org/10.1177/10298649221149110.

Benini, Sergio, et al. “On the Influence of Shot Scale on Film Mood and Narrative Engagement in Film Viewers.” *IEEE Transactions on Affective Computing*, vol. 13, no. 2, Sept. 2019, pp. 592–603, https://doi.org/10.1109/TAFFC.2019.2939251.

Bernardet, Ulysses, et al. “Assessing the Reliability of the Laban Movement Analysis System.” *PLOS ONE*, edited by Emily S. Cross, vol. 14, no. 6, June 2019, p. e0218179. *DOI.org (Crossref)*, https://doi.org/10.1371/journal.pone.0218179.

Bzymek, Zbigniew. “COMD3523 Storyboard Concepts, Fall 2016.” *COMD3523 Storyboard Concepts FALL 2016*, New York City College of Technology | City University of New York, 7 Nov. 2016, openlab.citytech.cuny.edu/comd3523-fa2016/2016/11/07/origins-of-storyboarding/.

Brown, Blain. “Coverage & Continuity.” *Cinematography: Theory and Practice: Image Making for Cinematographers and Directors*, 3rd ed., Routledge, New York, NY, 2016, pp. 116–193.

Callaway, Charles, et al. “Automatic Cinematography and Multilingual NLG for Generating Video Documentaries.” *Artificial Intelligence*, vol. 165, no. 1, June 2005, pp. 57–89.

Christianson, David B., et al. “Declarative Camera Control for Automatic Cinematography.” *AAAI’96: Proceedings of the Thirteenth National Conference on Artificial Intelligence - Volume 1*, vol. 1, AAAI Press, 1996, pp. 148–55.

Chun, Dayna. “Essence of Dance and a Camera.” Mānoa Horizons, vol. 2, 2017, pp. 87–94, https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/568e2043-051b-4350-9d2e-36b5bcf10143/content.

"Craft | Definition in the Cambridge English Dictionary." Cambridge Dictionary, Cambridge University Press & Assessment, dictionary.cambridge.org/us/dictionary/english/craft. Accessed 22 Nov. 2024.

Drucker, Steven M., et al. "Cinema: A System for Procedural Camera Movements." *Proceedings of the 1992 Symposium on Interactive 3D Graphics*, ACM Press, 1992, pp. 67–70.

Farnell, B. M. Movement Notation Systems. P. T. Daniels, & W. Bright, *The World's Writing Systems*. Oxford University Press. 1996, pp. 855-879.

"Film Terms Glossary." Filmsite, Filmsite LLC, www.filmsite.org/filmterms.html. Accessed 26 Apr. 2025.

Franko, Mark. "Writing for the body." *Common Knowledge*, vol. 17, no. 2, 1 Apr. 2011, pp. 321–334, https://doi.org/10.1215/0961754x-1188004.

Gorny, Liz. "The Story of Storyboarding: Exploring the Hidden Art Form behind Films." It's Nice That, Bureau for Visual Affairs, 6 Oct. 2021, www.itsnicethat.com/features/the-story-of-storyboarding-animation-film-061021.

Guest, Ann Hutchinson. *Labanotation: Or, Kinetography Laban : The System of Analyzing and Recording Movement*. 3rd ed., Taylor & Francis, 1977.

He, Li-wei, et al. "The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing." *Seminal Graphics Papers: Pushing the Boundaries*, vol. 2, 2023, pp. 707–14.

Hope, Cat. "The Future Is Graphic: Animated Notation for Contemporary Practice." *Organised Sound*, vol. 25, no. 2, Aug. 2020, pp. 187–97, https://doi.org/10.1017/S1355771820000096.

Jenkins, Henry. "Henry Jenkins: Spreadable Media: Creating Meaning and Value in a Networked Culture." *Henry Jenkins: Spreadable Media: Creating Meaning and Value in a Networked Culture :: News & Events :: Swarthmore College*, Swathmore College, 8 May 2012, www.swarthmore.edu/news-events/henry-jenkins-spreadable-media-creating-meaning-and-value-a-networked-culture.

Johnson, Rian. 18 Jan. 2020. *X*, X Corp, https://x.com/rianjohnson/status/1218717738564087808. Accessed 21 Nov. 2024.

Jowers, Iestyn, et al. "Shapes, structures and shape grammar implementation." *Computer-Aided Design*, vol. 111, June 2019, pp. 80–92, https://doi.org/10.1016/j.cad.2019.02.001.

Kerr, Will, et al. “Unsupervised Scouting and Layout for Storyboarding in Movie Pre-Production.” *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences Workshops*, Association for Computing Machinery, 2023, pp. 86–93, https://doi.org/10.1145/3604321.3604372.

Kim, Hanseob, et al. “ASAP for Multi-Outputs: Auto-Generating Storyboard and Pre-Visualization with Virtual Actors Based on Screenplay.” *Multimedia Tools and Applications*, Aug. 2024. *DOI.org (Crossref)*, https://doi.org/10.1007/s11042-024-19904-3.

Knight, Terry. *Introduction*, Massachusetts Institute of Technology, 14 Sept. 2000, www.mit.edu/~tknight/IJDC/page_introduction.htm.

Laemmli, Whitney E. “Paper dancers: Art as information in Twentieth-century America.” Information & Culture: A Journal of History, vol. 52, no. 1, Jan. 2017, pp. 1–30, https://doi.org/10.1353/lac.2017.0000.

Lai, Pei-Chun, et al. “A Pattern-Based Tool for Creating Virtual Cinematography in Interactive Storytelling.” *Smart Graphics*, edited by Marc Christie and Tsai-Yen Li, vol. 8698, Springer International Publishing, 2014, pp. 121–32. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-319-11650-1_11.

Li, Min, et al. “Dance movement learning for labanotation generation based on motion-captured data.” IEEE Access, vol. 7, Nov. 2019, pp. 161561–161572, https://doi.org/10.1109/access.2019.2951588.

Lino, Christophe, et al. *Automated Camera Planner for Film Editing Using Key Shots*. ACM SIGGRAPH / Eurographics, 2011.

Mackendrick, Alexander. *On Film-Making: An Introduction to the Craft of the Director*. Edited by Paul Cronin, Faber & Faber, 2006, 2006.

Mascelli, Joseph V. *The Five C’s of Cinematography*. Grafic Publications Hollywood, 1965.

Mate, Kedar K. “Effects of visual, auditory, and combined cues on human movement and brain regions involved in perception action.” *McGill Journal of Medicine*, vol. 20, no. 2, 5 July 2022, https://doi.org/10.26443/mjm.v20i2.928.

Muender, Thomas, et al. “Analysis of Previsualization Tasks for Animation, Film and Theater.” *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 2019, pp. 1–6. *DOI.org (Crossref)*, https://doi.org/10.1145/3290607.3312953.

Ozimek, Anna Maria (2021) *Skills Shortages, Gaps and Training Needs in the screen industries in Yorkshire and the Humber:Scoping Report.* Research Report.

"P1Harmony (피원하모니) - 'SAD SONG' MV." *YouTube*, P1Harmony, 20 Sept. 2024, www.youtube.com/watch?v=uHJDposrTMw.

"P1Harmony (피원하모니) 'SAD SONG' Dance Practice." *YouTube*, P1Harmony, 22 Sept. 2024, www.youtube.com/watch?v=Fkp5pix-PlU.

Pajala-Assefa, Hanna. "Dance Film – an Alliance of Dance and Moving Image." *Dance Arts – Historical Perspectives and Contemporary Practices*, Theatre Academy University of the Arts Helsinki, disco.teak.fi/tanssin-historia/en/dance-film-an-alliance-of-dance-and-moving-image/. Accessed 5 May 2025.

*Park, You Seung. '*트리플에스 *Visionary Vison - Hit the floor'* 뮤비에 스토리보드로 참여하였습니다 아 *Baaaam-. 29 Oct. 2024. Instagram, Instagram, https://www.instagram.com/p/DBsvoNgPMBf/?hl=en&img_index=1. Accessed 21 Nov. 2024.*

Railton, Diane and Watson, Paul. "2. GENRE AND MUSIC VIDEO: CONFIGURATIONS AND FUNCTIONS". *Music Video and the Politics of Representation*, Edinburgh: Edinburgh University Press, 2011, pp. 41-65. https://doi.org/10.1515/9780748633241-004V CF

Robinson, Adelaide. "Documenting Dance: Keeping Score." The Digital Shoebox, WordPress.com, 22 Feb. 2017, adafrobinson.wordpress.com/2017/02/22/keeping-score/.

Rudy Hendrawan, I. Nyoman, et al. "Video Dataset of Balinese Dance Basic Movement for Action Recognition." *Data in Brief*, vol. 53, Apr. 2024, p. 110189. *DOI.org (Crossref)*, https://doi.org/10.1016/j.dib.2024.110189.

Savardi, Mattia, et al. "CineScale: A Dataset of Cinematic Shot Scale in Movies." Data in Brief, vol. 36, June 2021, p. 107002. DOI.org (Crossref), https://doi.org/10.1016/j.dib.2021.107002.

Schlue, Lea Margarete. *Music Videos as Meaningful Entertainment? Psychological Responses to Audiovisual Presentations of Song Narratives*. 2016. West Virginia University, Graduate Theses, Dissertations, and Problem Reports, https://researchrepository.wvu.edu/etd/6584/.

Schofield, Guy, et al. *Specification of Cinematic Idioms (Revised Version)*. FP7-ICT-231824, 15 June 2010, https://cordis.europa.eu/docs/projects/cnect/4/231824/080/deliverables/001-IRISFP7ICT231824Deliverable51v2.pdf.

Serrano, Anna, et al. “Movie Editing and Cognitive Event Segmentation in Virtual Reality Video.” *ACM Transactions on Graphics*, vol. 35, no. 4, June 217AD, pp. 1–12, https://doi.org/10.1145/3072959.3073668.

Simon, Justin. “How to Write a Shot List That Will Transform Your Video.” *The TechSmith Blog*, TechSmith Corporation, 15 Nov. 2024, www.techsmith.com/blog/shot-list/.

Stiny, George. “Introduction to shape and shape grammars.” *Environment and Planning B: Planning and Design*, vol. 7, no. 3, Sept. 1980, pp. 343–351, https://doi.org/10.1068/b070343.

Talbot, Thomas Brett, et al. “Storyboarding the Virtuality: Methods and Best Practices to Depict Scenes and Interactive Stories in Virtual and Mixed Reality.” *Advances in Usability, User Experience, Wearable and Assistive Technology*, edited by Tareq Ahram and Christianne Falcão, vol. 1217, Springer International Publishing, 2020, pp. 129–35. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-030-51828-8_17.

“The Making Of A K-Pop Music Video.” *YouTube*, Youtube Originals, youtu.be/VjTiEqLmklU?si=jI5YBHOOYMAfx83x. Accessed 6 May 2025.

Wang, Zixuan, et al. *DanceCamera3D: 3D Camera Movement Synthesis with Music and Dance*. IEEE, 2024, https://doi.org/10.1109/CVPR52733.2024.00754.

Zhang, Xueyan, et al. “Skeleton-based automatic generation of labanotation with neural networks.” Journal of Electronic Imaging, vol. 28, no. 02, 5 Apr. 2019, pp. 510–515, https://doi.org/10.1117/1.jei.28.2.023026.

“[Performance37] Izna(이즈나) ‘IZNA’ (4K).” *YouTube*, M2, 2 Dec. 2024, youtu.be/xvRTdbbXxj8?si=-i4aYRRkueM9Ixq_.

“[Performance37] LE SSERAFIM(르세라핌) ‘이브, 프시케 그리고 푸른 수염의 아내’ (4K).” *YouTube*, M2, 23 May 2023, www.youtube.com/watch?v=y1ApSB0FpAM.

“[Performance37] UNIS(유니스) ‘Dopamine’ (4K).” *YouTube*, M2, 30 Mar. 2024, youtu.be/7xHZpIxtSZ4?si=jKT8-oycWrcTld-4.