UNIVERSITY OF SOUTHAMPTON

FACULTY OF ARTS AND HUMANITIES

Film

Film Buddhology

by

Ao Chen

Thesis for the degree of Doctor of Philosophy

June 2022

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ARTS AND HUMANITIES

Film

Thesis for the degree of Doctor of Philosophy

FILM BUDDHOLOGY

Ao Chen

This thesis proposes Buddhology as a methodology for discussing film ontology. The mind in Buddhology provides a way to reconsider the essence of all phenomena. Considering that film is a palpable phenomenon, I suggest that the mind engages with film at a Buddhist ontological level. My overarching argument is that film ontology is nominal. Nominal film arises from my act of combining film ontology with the Buddhist idea of nominal existence, referring to an intermediate state between film (film exists) and "no film" (film does not exist). On the one hand, my thesis argues that film does not have an independent and inherent quality to enable itself to exist in its own right, but on the other hand, because audiences' sense organs and sense consciousness together make the manifestation of film possible, it follows that film is not absolute nothingness. The framework of film Buddhology that I outline in my thesis is threefold. First, I will summarise some traditional approaches to and concepts in film ontology, such as phenomenology, realism, and cognitivism, in order to consider how other theorists analyse the nature of film. Second, I will put these theories into dialogue with Buddhist discussions concerning the association between phenomena and the mind in the context of film ontology, and examine the relation between film and the viewer's act of experiencing film. Moreover, I will explain how the mind and its several sub-ideas are selected from among many Buddhist sources. Third, I apply my ideas to specific case studies where each chosen film indicates a specific type of filmmaking. The selected examples are Michelangelo Antonioni's art film *Blow-up* (1966), Hollis Frampton's experimental film *Zorns Lemma* (1970), Abbas Kiarostami's docufiction *Through the Olive Trees* (1994), and William Wyler's classical Hollywood film *The Best Years of Our Lives* (1946). A Buddhist perspective on these films will trigger discussions about nominal images, nominal montage, nominal voices, and nominal deep-focus shots—manifold paths to the applicability of nominal film beyond any types of films.

Table of Contents

Table of Tables

Table of Figures

Table of Figures

Academic Thesis: Declaration of Authorship

I, Ao Chen, declare that this thesis entitled “Film Buddhology” and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signed:

Date: 28/06/2022

Acknowledgements

I am greatly indebted to my PhD supervisors Dr Malcolm Cook, Dr Beth Carroll, and Professor Tim Bergfelder for their guidance and constructive feedback on my PhD thesis *Film Buddhology* (2021). I also would like to thank them for offering helpful advice on my conference paper "Nominal Film Frame: The Ending of *Blow-up*" (2020).

I would like to express my gratitude to the Faculty of Arts and Humanities Graduate School at the University of Southampton for financing my conference presentation at the 30th International Screen Studies Conference in the UK and my research trip to the 43rd Annecy International Animation Film Festival in France for attending the world premiere of my animation *Red* (2019).

In addition, I am exceptionally grateful for the teaching opportunity given to me by the Graduate School and the Film Department.

Thanks also to my MFA thesis advisor Professor Michael O'Rourke and Chairperson Professor Peter Patchen from the Pratt Institute in the USA for providing references for my PhD application.

Last but not least, I want to thank my parents Professor Fei Chen and Qiaomei Jiang for their financial, intellectual, and spiritual support.

Notes on Transliteration and Translation

A pinyin transliteration system for Chinese words is employed throughout this thesis. I choose to write Chinese words in the form of simplified Chinese characters. In terms of Chinese language, according to the *Xinhua zidian* 新华字典 (translated into English as the *New China Dictionary*), the pronunciations of Chinese words are dependent on a Mandarin pronunciation system; pinyin and Bopomofo (also known as *Zhuyin* 注音, translated into English as Mandarin Phonetic Symbols) are used for the phonetic transcription of Chinese characters.[1] Those names under other transliteration systems such as Wade–Giles and Gwoyeu Romatzyh are uniformly converted into pinyin. In addition, this thesis engages with the International Alphabet of Sanskrit Transliteration from time to time.

Some Buddhist sūtras have several Chinese translations. In order to get as close as possible to the original Sanskrit meanings of Buddhist ideas, I select the Chinese translations that have been widely recognised as the standard reference versions. For example, with respect to the *Vajracchedikā Prajñāpāramitā Sūtra* (also known as the *Diamond Sūtra*, translated into Chinese circa 401 as the *Jingang jing* 金刚经 by Kumārajīva, 344–413), Lu Xuanchang emphasises that of those six translations, Kumārajīva's translation is regarded as the standard "because Kumārajīva has been the master translator for the past seven Buddhas. The sūtras he translated are the same as the Buddhas' heart, so everyone likes to read and recite them."[2] Moreover, as Lu suggests, Kumārajīva says that "I personally don't know if there are mistakes in the sūtras I have translated, but if there are none, when I am cremated my tongue will not burn."[3] History has it that his tongue remains untouched by the fire, which suggests that the sūtras that he has translated are the standard reference versions.[4]

---

[1] Zhongguo shehui kexueyuan yuyan yanjiusuo 中国社会科学院语言研究所 [The Institute of Linguistics at the Chinese Academy of Social Sciences], *Xinhua zidian* 新华字典 [The New China Dictionary] (Beijing: Shangwu yinshuguan 北京：商务印书馆 [Beijing: The Commercial Press], 2004), p. 6. My translation (Note: Since Arabic page numbers are used throughout this dictionary, in order to avoid confusion, I would like to explain that page 6 refers to page 6 of *Fanli* 凡例 [Explanatory Notes]).

[2] Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007); in the English translated version *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra*, the original text is accompanied by the commentary provided by Xuanhua (the pinyin form of Hsüan Hua [Hsüan Hua is a Wade–Giles-based name]), 1918–1995, trans. Heng Zhi (the pinyin form of Heng Chih [Heng Chih is a Wade–Giles-based name]) (San Francisco: Sino-American Buddhist Association, Inc., 1974), p. 41 (Note: Lu Xuanchang is the pinyin form of Lu Hsüan Ch'ang [Lu Hsüan Ch'ang is a Wade–Giles-based name]).

[3] Ānanda, *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra* (English), p. 41.

[4] Ānanda, *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra* (English), p. 41.

Another example is the *Avataṃsaka Sūtra* (also known as the *Flower Ornament Scripture*, translated into the first Chinese version circa 420 as the *Huayan jing* 华严经 by Buddhabhadra, 359–429, and into the second circa 699 by Śikṣānanda, 652–710). As for its Chinese translations, the English translator Thomas Cleary suggests:

> The first comprehensive translation of the *Flower Ornament Scripture* was done under the direction of an Indian monk named Buddhabhadra (359–429); the second, under the direction of a Khotanese monk named Śikṣānanda (652–710). The latter version, from which the present English translation is made, was based on a more complete text imported from Khotan at the request of the empress of China; it is somewhat more than ten percent longer than Buddhabhadra's translation.[5]

In addition, there are some Buddhist works directly written in English, such as the introductory chapter of *The Gelug/Kagyü Tradition of Mahāmudrā* by Alexander Berzin and *What Makes You Not a Buddhist* by Khyentse Norbu.[6] Furthermore, some Buddhist ideas in Chinese are translated into English by myself where needed.

[5] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), p. 2.
[6] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997); Khyentse Norbu, *What Makes You Not a Buddhist* (Boston and London: Shambhala Publications, Inc., 2007) (Note: Even if Dzongsar Jamyang Khyentse, also known as Khyentse, is the name on the front cover of the book, Khyentse is his first name. His surname is Norbu).

Chapter 1 An Introduction to Film Buddhology

1.1 Nominal Film

Throughout this thesis, my overarching argument is that film ontology is nominal. Nominal film arises from my act of combining film ontology with the Buddhist idea of *huanyou* 幻有 (nominal existence),[7] referring to an intermediate state between film (film exists) and "no film" (film does not exist). I will amplify nominal film and clarify how I use the term "intermediate state" in the latter part of this section. In order to apply nominal existence to film ontology, I will engage with the mind in Buddhology. It provides a way to reconsider the essence of all phenomena. Since film is a palpable phenomenon, I argue that the mind engages with film at a Buddhist ontological level. That is, I adopt a Buddhist approach to revisiting the ontological link between film and audiences. Before explaining the preceding argument, I would like to, in a concise manner, first put it into a wider context of film theory and examine how my perspective is different from other theorists' accounts of film ontology.

A dominant example of examinations of film ontology is André Bazin's change mummified in *What Is Cinema?*.[8] This idea is concerned with cinema's capability to trace reality by preserving a moment of temporal change captured from the flow of daily life. For Bazin, film ontology is a suspended state between life and death, presence and absence. Yet, death does not only refer to the absent moment, but also includes other connotations. For instance, if death is placed into the context of Paolo Cherchi Usai's discussion of film stock in *The Death of Cinema: History, Cultural Memory, and the Digital Dark Age*, it signifies that film stock as the image carrier decays.[9] In addition, if film is put into dialogue with other media such as drawing, television, and video, as Noël Carroll suggests in *Theorizing the Moving Image*, it turns out that film has no essence because it cannot establish itself as a distinctive medium due to its stylistic resemblance to other media.[10]

Although change mummified involves some kind of intermediate state, it is not what I call nominal film. Change mummified deals with the existential issue of the world simulated on the screen – or, in other words, the existential issue of the representation of film, whilst nominal film

---

[7] Dharmapāla, *Cheng weishi lun* 成唯识论 [The Discourse on the Perfection of Consciousness-only] (It was completed by Dharmapāla in 550, translated into Chinese in 659 as the *Cheng weishi lun* by Xuanzang, circa 602–664) (Taipei: Laogu wenhua shiye gufen youxian gongsi 台北：老古文化事业股份有限公司 [Taipei: Laoku Culture Foundation Inc.], 2014), vol. 2, p. 46. My translation.

[8] André Bazin, "The Ontology of the Photographic Image" (1958), in *What Is Cinema?*, ed. and trans. Hugh Gray (Berkeley, Los Angeles, and London: University of California Press, 1967), vol. 1, p. 15.

[9] Paolo Cherchi Usai, *The Death of Cinema: History, Cultural Memory, and the Digital Dark Age* (London: British Film Institute, 2001), p. 17.

[10] Noël Carroll, "Defining the Moving Image," in *Theorizing the Moving Image*, ed. Noël Carroll (Cambridge, New York, and Melbourne: Cambridge University Press, 1996), pp. 50–52.

engages with the existential issue of film itself. Instead of concentrating on the issue of the absent moment replayed in the present, nominal film is a mind-related idea, scrutinising how audiences' minds construct and deconstruct the entity of film itself. "Film does not exist," one of the key claims about nominal film, means neither that film stock decays, nor that film cannot establish itself as a unique medium. Rather, from a Buddhist perspective, "film does not exist" or "no film" implies that all filmic elements such as images, montage, voices, and deep-focus shots do not have independent and inherent qualities to allow themselves to exist in their own right, since their entities depend on the viewer's act of perceiving them. First, if filmic elements existed in an independent and inherent way, then even if the viewer did not perceive them, they should still be able to manifest themselves to the viewer in their own right. Second, filmic elements cannot retain fixed qualities, as their audio-visual appearances can be altered by the viewer's mind, which indicates that there are no inherently findable elements establishing their solid entities from their own side. "Film exists," another crucial statement regarding nominal film, suggests that the viewer's sense organs and sense consciousness together make the manifestation of film possible, it follows that film is not entire nothingness. Being devoid of existing independently and permanently does not mean to deny film. Instead, the viewer can feel the sensation of film. If film were total nothingness, then even if the perceiving functioned, film should never be manifested to the viewer. Since the arising of this sensation cannot be negated, film still has a relative entity in some ways. Nominal film is an intermediate state between film and "no film," suggesting that film is neither existent nor non-existent—a conversion process between the existence and non-existence of film.

Such intermediacy is not identical to *zhongdao* 中道 (the Middle Way), but the latter is helpful in understanding nominal existence. In the *Mūlamadhyamakakārikā* (also known as the *Root Verses on the Middle Way*, written by Nāgārjuna circa the third century, translated into Chinese in 409 as the *Zhong lun* 中论 by Kumārajīva, 344–413), Nāgārjuna points out that "*you ruo bucheng zhe*, *wu yunhe kecheng*, *yin you youfa gu*, *you huai mingwei wu*. […] *Ruo fa you dingxing*, *feiwu zeshi chang*, *xian you er jin wu*, *shizewei duanmie*." "有若不成者，无云何可成，因有有法故，有坏名为无。【……】若法有定性，非无则是常，先有而今无，是则为断灭。" (when the existent is not established, the non-existent is also not established. It is, indeed, the change of the existent that people generally call the non-existent. […] "Whatever that exists in terms of self-nature, that is not non-existent" implies eternalism. "It does not exist now, but existed before" implies annihilation).[11]

[11] Nāgārjuna, *Zhong lun* 中论 [The Root Verses on the Middle Way] (It was completed by Nāgārjuna circa the third century. The explanation was completed by Vimalākṣa circa the fourth century, translated into Chinese in 409 as the *Zhong lun* by Kumārajīva, 344–413), ed. Yan Yongkui (Chengdu: Zhongguo chuantong wenhua yanjiusuo 成都：中国传统文化研究所 [Chengdu: The Institute of Chinese Traditional Culture], 1995), pp. 113, 115; in the English translated version *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way*, the original text is accompanied by the annotation provided by David Kalupahana, 1936–2014, trans. David Kalupahana (Delhi: Motilal Banarsidass, 1999), pp. 230, 234.

On the one hand, *chang* 常 (eternalism) means that things exist through their essence. Yet, as things are always in a changing state, eternalism is an extreme idea that needs to be avoided. On the other hand, *duanmie* 断灭 (annihilation) means that things existed before, but because of the decision to resist temptations, one starts to force themselves to regard things as not existent. Yet, this situation will induce one to forsake everything experienced—another extreme view. It seems like the way to enter the Middle Way is to reach a compromise between eternalism and annihilation. That is to say, the entities of things should not be emptied too much. One should step back and acknowledge their existence somewhat. However, such intermediacy is not what Nāgārjuna means by the Middle Way.

In order to examine the Middle Way, I would like to associate the preceding statements quoted from the *Mūlamadhyamakakārikā* with the *Śūnyatāsaptatikārikā* (also known as the *Seventy Verses on Emptiness*, written by Nāgārjuna circa the third century, translated into Chinese in 1939 as the *Qishi kongxing lun* 七十空性论 by Fazun, 1902–1980). In the latter Buddhist work, Nāgārjuna states that "*you youxing ying chang*, *wu zhe dingcheng duan*, *youxing duo er shi*, *shigu buying xu*." "有有性应常，无者定成断，有性堕二失，是故不应许。" (if a phenomenon were to exist inherently it should be permanent. If a phenomenon were to disintegrate completely then you must accept the annihilationist view. If a phenomenon were to exist inherently it would either exist permanently or else undergo complete disintegration: it cannot occur in a way which is different than these two. Therefore one should not assert that a phenomenon has inherent existence).[12] In other words, both eternalism (the existent) and annihilation (the non-existent) are based on the premise that there is a phenomenon (the existent) in the first place. Whilst one has noticed the ever-changing, one is still likely to posit that there is "something" in a state of flux. Yet, the Middle Way has swept this premise away, considering that the premise is just a conceptual label that the observer adds to the perceived appearance. A conceptual premise does not serve as the phenomenon itself. Regarding Nāgārjuna's Buddhist view that "when the existent is not established, the non-existent is also not established,"[13] "the existent" refers specifically to the premise that there is a pregiven thing in the first place. When this premise has been removed, there is no place for both its eternalism and annihilation to happen.

---

[12] Nāgārjuna, *Qishi kongxing lun* 七十空性论 [The Seventy Verses on Emptiness] (It was completed by Nāgārjuna circa the third century. The first version of the explanation was completed by Nāgārjuna circa the third century, the second by Candrakīrti in the early seventh century, and the third by Pārahitā in the late eleventh century, translated into Chinese in 1939 as the *Qishi kongxing lun* by Fazun, 1902–1980. The recent version of the explanation was completed by Khenpo Shenga (also known as Shenpen Nangwa), 1871–1927, translated into Chinese in 2019 as the *Qishi kongxing lun* by Khenpo Sodargye), in the *Zhongguan zhuangyan lun shi* 中观庄严论释 [An Explanation of a Filigree of the Middle Way], eds. Guo xiaofei and Tang Ao (Lhasa: Xizang zangwen guji chubanshe 拉萨：西藏藏文古籍出版社 [Lhasa: Tibetan Ancient Books Publishing House], 2019), p. 306; in the English translated version *Nāgārjuna's "Seventy Stanzas": A Buddhist Psychology of Emptiness*, the original text is accompanied by the commentary provided by Sonam Rinchen, 1933–2013, trans. Tenzin Dorjee and David Komito (Ithaca: Snow Lion Publications, 1987), pp. 128–29.

[13] See p. 2.

We can take the fire as an example to illustrate the above point. If a proposition is that the fire is lit, its inverse is that the fire is extinguished. Strictly speaking, this inverse is viewed as a denial of a proposition with a remainder of clinging. A remainder of clinging suggests that whether the fire is lit or extinguished, the observer still assumes that the fire serves as a pre-established carrier making the processes of lighting and extinguishing itself in operation. However, in the context of the Middle Way, the pre-established "fire"—a conceptual premise—has been swept away, which indicates that the experiencer makes a denial of a proposition without a remainder of clinging. A denial of the fire lit is not an acknowledgement of the fire extinguished, as nothing has ever undergone the processes of lighting and extinguishing itself. In this light, this denial points neither to the fire lit nor to the fire extinguished. Calling it an "intermediate state" between them is also not precise because otherwise it looks as if one acknowledges that there are eternalism and annihilation that keep one in between.

However, there is a reason why I still want to utilise the term "intermediate state" throughout this thesis. In order to discuss this reason, I would like to link the above issue to two truths doctrine: *shisu di* 世俗谛 (conventional, provisional, or nominal reality) and *diyiyi di* 第一义谛 or *shengyi di* 胜义谛 (ultimate reality).[14] First, *shisu di* signifies that things arise and cease only in a relative sense. Their arising and ceasing are based on interdependent relationships. That is to say, if a phenomenon experiences its initiation and extinction, it must have hinged on the conditions for undergoing these processes. Thus, it is temporary and false, functioning but without a truly independent and inherent substance. Because of the variability of the conditions, the arising can convert into the ceasing, and vice versa. Nominal existence serves as an intermediate state between existence and non-existence at a relative level, indicating that both sides are not able to maintain a fixed state. Nominal existence is not exactly what Nāgārjuna means by the Middle Way in view of the fact that "nominal existence" or an "intermediate state" itself is still a conceptual premise or linguistic convention. Yet, Nāgārjuna further suggests that "*ruo buyi su di*, *bude diyiyi*, *bude diyiyi*, *ze bude niepan*." "若不依俗谛，不得第一义，不得第一义，则不得涅槃。" (without relying upon convention, the ultimate fruit is not taught. Without understanding the ultimate fruit, freedom is not attained).[15] Likewise, my work still relies on concepts and language to engage with Nāgārjuna's contention about sweeping conceptual premises away. If words were not applied to my research, I would not be able to express and explain anything, which is beyond conventional, provisional, or nominal existence. I do not intend to change my analysis into a blank paper because otherwise readers cannot receive and learn any information.

[14] Nāgārjuna, *Zhong lun* (Chinese), p. 174; *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way* (English), p. 331.
[15] Nāgārjuna, *Zhong lun* (Chinese), p. 174; *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way* (English), p. 333.

Based on the above quote, in order to scrutinise *diyiyi di* (the ultimate)—the second aspect of two truths doctrine—I want to look at the English annotations for this facet, as they outline multiple views on *diyiyi di*. David Kalupahana encapsulates the difference between the Buddha – Siddhārtha Gautama – and Vasubandhu's attitudes to ultimate reality, with the latter philosopher being one of the interpreters of the Abhidharma (a set of ancient [the third century BCE and later] Buddhist texts which contain a detailed scholastic reworking of doctrinal material appearing in the sūtras).[16] In this comparative analysis, as Kalupahana summarises, on the one hand, the Buddha does not see a view as the ultimate. Because of dependent origination, there is no place for the ultimate, which is similar to Nāgārjuna's concept sweeping. On the other hand, Vasubandhu's metaphysical viewpoint is seen as a form of epistemological idealism, reducing matter to consciousness or the knowledge of form, which signifies that the manifestation of subjectivity is affirmed. However, it is not purely subjective idealism because of the activity of "cause-effect" consciousness.[17] Although there are distinct views on the ultimate, they all imply that there is no pure being and no pure nothing.

I point out that nominal film—an intermediate state between the existence and non-existence of film—serves as a becoming or a conversion process between them, but simultaneously, this idea is subject to the potential of being swept away. That is to say, though I use nominal film throughout this thesis, it should not be regarded as a pure and inherent thing. Instead, it can be understood as a provisional name or label which at best denotes nominal film.

Before employing a Buddhist methodology to examine nominal film in more detail, I would like to further map several sources related to film ontology. Although those film theorists' contentions that I have mentioned previously serve as a foil for the general understanding of nominal film, it is necessary to conduct further examinations of the relevant sources about film ontology, which helps us reflect on film ontology from multiple perspectives. I will then scrutinise a Buddhist methodology and utilise Buddhist ideas to engage with others' works. The next section is composed of five realms: film phenomenology, film realism, cognitive film theory, auteur theory, and film theorizing and film medium specificity, with each route having its own theoretical approach to the essence of film. This does not mean that discussions of film ontology are only linked to these five fields. Rather, they are picked as influential examples related to this topic. I will encapsulate their key concepts and look at differences among these fields. They will provide a broader context for my subsequent examination of a Buddhist approach to film ontology.

---

[16] Nāgārjuna, *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way* (English), pp. 331–33.
[17] Nāgārjuna, *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way* (English), pp. 331–33.

1.2 A Broader Context of Film Ontology

1.2.1 Film Phenomenology

Film phenomenology is concerned with the association between film and audiences, engaging with the essence of film at the level of sensory fields. In *The Address of the Eye: A Phenomenology of Film Experience*, Vivian Sobchack conducts an analysis of the act of being with one's own eyes. She argues that "without an act of viewing and a subject who knows itself reflexively as the locus and origin of viewing *as* an act, there could be no film and no 'film experience.'"[18] She further asserts that there are two perceptual modes of access to a film: an anonymous mode of being aware of some "thing" seen and a situated mode of being aware of the perception of the self as capable of seeing.[19] The former emphasises that seeing a film with audiences' eyes is a prerequisite for manifesting a film, whilst the latter stresses that seeing a film with audiences' reflexive and reflective consciousness makes audiences realise their own activities of seeing a film. When these two parts come together, both film itself and film experience can be in a state of being determined. She also points out that "it is this homology that lays the philosophical and existential grounds for further consideration of the figures of self and other, seer and seen, as co-emergent, co-operative, and co-perceptive."[20] This idea signifies the interdependence of film itself and film experience.

Although Sobchack highlights the nexus of film and film experience and mentions "no film" and "no film experience," she does not scrutinise a transition zone between film and "no film." As stated previously, I argue that nominal film deals with a conversion process between being and non-being, an intermediate state where film is a part of "no film" and "no film" is a part of film. Nominal film enables film ontology to be in a variable situation, rather than to stagnate in a state either of film or of "no film." Although Sobchack suggests that without a reflection on the act of perceiving, there could be "no film," I add that it is also crucial to note that at a Buddhist ontological level, "no film" is not pure "no film." "No film" still embraces film. To put it another way, "no film" still encompasses a filmic potentiality that has yet to be manifested. The reason why "no film" contains film is because once "no film" is subject to audiences' experiences, it will not be able to retain its non-existent state and will become film. "No film" is not total nothingness.

The conversion process between film and "no film" rarely gets attention, whilst discussions of the manifestation of film are pervasive in different works. In addition to Sobchack's work, Laura Marks' tactile approach to film provides an alternative way to delve into the link between film and

[18] Vivian Sobchack, *The Address of the Eye: A Phenomenology of Film Experience* (Princeton and Oxford: Princeton University Press, 1992), p. 51 (Note: The word *as* in this quotation is Sobchack's italic emphasis).
[19] Sobchack, *The Address of the Eye: A Phenomenology of Film Experience*, pp. 51–54.
[20] Sobchack, *The Address of the Eye: A Phenomenology of Film Experience*, p. 55.

audiences. In *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*, Marks propounds the idea of haptic visuality and points out that "the eyes themselves function like organs of touch."[21] This suggests that the relationship between film and audiences can be embodied and intersubjective. She further discusses the distinction between haptic and optical visuality. Haptic visuality puts emphasis on the material presence of the image, whilst optical visuality adds weight to the representational power of the image.[22] More specifically, the former triggers "a look that moves on the surface plane of the screen for some time before the viewer realises what she or he is beholding."[23] The latter drives audiences to gaze into objects and to identify illusory figures in space. Instead of being enthralled by spatial depth and narrative, haptic visuality motivates audiences to draw on a repository for their own memories and imaginations, bringing them into contact with the object's material significance.[24]

Even though Marks' haptic approach is different from Sobchack's visual emphasis, her general ideas are still concerned with the manifestation of film (the explicit aspect of film ontology). She does not focus on "no film" (the implicit aspect of film ontology) and nominal film. I will carry out a detailed analysis of these ideas in sub-section 1.4.1.

There are some other sources exploring film ontology but not typically through experience-oriented methods. In the next sub-section, I will discuss the notion of realism. There are two reasons for focussing on this aspect in my thesis. First, realism also engages with film ontology, which is connected to my object of study. Second, even though the primary focus of realism is the world as represented and even though it deals with the mechanical function of the movie camera, rather than being directly associated with film experience, realism can still be reconsidered in an experience-oriented way when it is put into a Buddhist context. Before employing a Buddhist perspective to rethink realism, I would like to first look at how theorists like Siegfried Kracauer and Béla Balázs investigate realism.

1.2.2 Film Realism

Realism's view on film ontology is based on observations of the relation between the world and the movie camera. The world can be examined from the physical and mental aspects of daily reality. In *Theory of Film: The Redemption of Physical Reality*, Kracauer focuses on the technical fact that the mechanical function of the movie camera enables the provisional and unexpected moments of life

[21] Laura Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses* (Durham and London: Duke University Press, 2000), p. 162.
[22] Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*, p. 163.
[23] Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*, pp. 162–63.
[24] Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*, pp. 162–63.

to be stored, such as people on the street and leaves fluttering in the air. He argues that "now this reality includes many phenomena which would hardly be perceived were it not for the motion picture camera's ability to catch them on the wing."[25] The movie camera preserves the temporary and unpredictable appearances of physical reality. Unlike Sobchack and Marks' ideas, Kracauer's perspective does not stress the senses. Rather, he emphasises the movie camera's superior ability to keenly capture and present fleeting physical phenomena that may be ignored by us in daily life.

In addition to contingent physical phenomena, in *Theory of the Film: Character and Growth of a New Art*, Balázs canvasses how the movie camera makes facial expressions visible and how it shows the character's subtle mental activities behind the physical face. From Balázs' perspective, the close-up not only enables audiences to get close to the face and the movements of facial muscles, but also allows us to experience emotions and thoughts represented by facial expressions transcending physical space:

> When we look into the eyes in a close-up, we no longer think of that wide space, because the expression and significance of the face has no relation to space and no connection with it. Facing an isolated face takes us out of space, our consciousness of space is cut out and we find ourselves in another dimension: that of physiognomy. The fact that the features of the face can be seen side by side, i.e. in space—that the eyes are at the top, the ears at the sides and the mouth lower down—loses all reference to space when we see, not a figure of flesh and bone, but an expression or in other words when we see emotions, moods, intentions and thoughts, things which although our eyes can see them, are not in space. For feelings, emotions, moods, intentions, thoughts are not themselves things pertaining to space, even if they are rendered visible by means which are.[26]

This statement emphasises how the close-up technique of the movie camera is utilised to display the character's lifelike emotional states—mental phenomena that do not depend on physical space. This analysis is distinct from a phenomenological examination of how filmic physical and mental phenomena as the holistic object of the viewer's vision and the act of viewing construct each other's existential qualities.

Yet, if we rethink Kracauer and Balázs' discussions of the function of the movie camera, we realise that their ideas can be dissected at the level of audiences' senses. I will later put their ideas into a Buddhist framework and look at their potential relationship to sensory experience. This reconsideration intends to regard the movie camera as a mind assistant extending the spectator's perceiving process and to situate Kracauer and Balázs' ideas in my Buddhist analysis of nominal film.

---

[25] Siegfried Kracauer, *Theory of Film: The Redemption of Physical Reality* (New York: Oxford University Press, 1960), p. ix.
[26] Béla Balázs, *Theory of the Film: Character and Growth of a New Art*, ed. Herbert Marshall, trans. Edith Bone (London: Dennis Dobson Ltd, 1952), p. 61.

Apart from Kracauer and Balázs' analyses of moment and movement, Gilles Deleuze uses Henri Bergson's philosophy of movement and time as a theoretical basis to reflect on filmic elements such as the frame, the shot, and montage.[27] Regarding Deleuze's theoretical approach to film, one of the most noteworthy terms is the intermediate image in *Cinema 1: The Movement-image*,[28] considering that this concept is different from my idea of an intermediate state between film and "no film." The intermediate image signifies that cinema does not exhibit immobile sections to which movement is added, but rather immediately gives us a mobile section – or, in other words, movement established between images.[29] From Deleuze's point of view, distinct from the ancient conception of movement which recomposes movement with privileged instants, eternal poses, or immobile sections, cinema reproduces movement by virtue of linking it to any-instant-whatevers, with movement serving as a function of equidistant instants in order to create an impression of continuity.[30] Movement belongs to the intermediate image as immediate given, connecting objects to the duration of a whole which changes—in the sense of the qualitative change of a whole.[31] Montage, the mobile camera, and the emancipation of the viewpoint enhance the operation of the movement-image.[32] However, my idea of intermediacy (nominality) is not movement established between images, but rather a conversion process between the arising and void of all cinematic elements. I will dissect this issue in detail later.

In addition to ontological examinations of the movement-image, in *Cinema 2: The Time-image*, Deleuze draws a parallel between the screen and the brain or the body, which suggests that cinema expresses world thinking at a universal level.[33] Moreover, for Deleuze, "the purely optical and sound situation (description) is an actual image, but one which, instead of extending into movement, links up with a virtual image and forms a circuit with it."[34] Our bodies enable us to connect the image to another in the past through our memory. Although these two kinds of images are different in terms of their respective contexts, their interweaving presents temporal and spiritual recollection-images. He goes on to suggest that recollection-images trigger attentive recognition, stimulating our mental responses and producing, in conjunction with the actual image, a closed circuit which makes us pass from one world to another, from the present to the past, then brings us back to the present. During this process, images are endowed with an atmosphere of the world, transformed from the flat view into space which transcends an ordinary sensory-motor situation.[35]

---

[27] Gilles Deleuze, *Cinema 1: The Movement-image* (1983), trans. Hugh Tomlinson and Barbara Habberjam (Minneapolis: University of Minnesota Press, 1997).
[28] Deleuze, *Cinema 1: The Movement-image*, p. 2.
[29] Deleuze, *Cinema 1: The Movement-image*, pp. 2, 11.
[30] Deleuze, *Cinema 1: The Movement-image*, pp. 5–6.
[31] Deleuze, *Cinema 1: The Movement-image*, pp. 2, 11.
[32] Deleuze, *Cinema 1: The Movement-image*, p. 3.
[33] Gilles Deleuze, *Cinema 2: The Time-image* (1985), trans. Hugh Tomlinson and Robert Galeta (Minneapolis: University of Minnesota Press, 1997).
[34] Deleuze, *Cinema 2: The Time-image*, p. 47.
[35] Deleuze, *Cinema 2: The Time-image*, pp. 47–48, 62.

1.2.3 Cognitive Film Theory

Another field that needs to be mentioned is cognitive film theory, addressing the viewer's inference making. In *Narration in the Fiction Film*, David Bordwell summarises two facets of inference making: bottom-up and top-down processes. The former depends on the perceptual input, for instance the impressions of steady light, apparent motion, and colour, focussing on illusions created by artworks, whilst the latter is based on the viewer's prior knowledge including their engagement with daily life and multiple works.[36] Bordwell asserts that in perceptual-cognitive activities, schemata—organised clusters of knowledge guiding the spectator's hypothesis making—play a significant role in narrative understanding, with this contention indicating the active participation of the viewer.[37] For example, the spectator draws on schemata to construct hypotheses about and hold expectations of artworks. Moreover, the arrangement of events in linear sequence is a key consideration in film viewing. Once events challenge linear sequence or a "canonical" story format, the viewer will change the way they engage with the material and structure of the film in order to make and test new assumptions about what is shown.[38] Therefore, for Bordwell, film viewing is an active and variable perceptual-cognitive process shaping the viewer's construction of an intelligible story, rather than a passive experience.[39]

Moreover, Bordwell puts analyses of perceptual-cognitive activities into the context of Russian Formalist narrative theory, analysing the relationship among *fabula* (story), *syuzhet* (plot), and style. Briefly speaking, *fabula* is "a pattern which perceivers of narratives create through assumptions and inferences. It is the developing result of picking up narrative cues, applying schemata, framing and testing hypotheses."[40] Distinct from *fabula*, *syuzhet* is "the actual arrangement and presentation of the fabula in the film."[41] The material and structure of the film itself are displayed in a *syuzhet* which prompts the spectator to harness schemata of causality, time, and space to create a *fabula*. Style is "the film's systematic use of cinematic devices."[42] It supports the construction of *fabula* and *syuzhet*.

Apart from the spectator's logical inferences about canonical story formats, it is also significant to consider the cognitive engagement with non-canonical story formats. In "Is a Cognitive Approach to the Avant-garde Cinema Perverse?," James Peterson chooses avant-garde cinema as an example to emphasise that though avant-garde cinema generally expresses chaotic, confusing, and irrational events, a cognitive perspective on the viewing of avant-garde cinema is not perverse. Instead, it still

[36] David Bordwell, *Narration in the Fiction Film* (Madison: University of Wisconsin Press, 1985), pp. 31–33.
[37] Bordwell, *Narration in the Fiction Film*, pp. 31–33.
[38] Bordwell, *Narration in the Fiction Film*, pp. 33–36.
[39] Bordwell, *Narration in the Fiction Film*, pp. 31, 35–36, 38–39.
[40] Bordwell, *Narration in the Fiction Film*, p. 49.
[41] Bordwell, *Narration in the Fiction Film*, p. 50.
[42] Bordwell, *Narration in the Fiction Film*, p. 50.

relates to the basic comprehension of canonical story formats.[43] Peterson further points out that in order to understand avant-garde cinema, the spectator needs to acquire specialised knowledge of art history, film history, and the lives of filmmakers and use heuristics to discern and solve problems posed by the film. For example, the spectator can engage with the film by focussing on the elements over which the filmmaker has most control or associating the images and sounds with the concerns of the filmmaker.[44] For Peterson, the continuous learning of these viewing strategies indicates that comprehension can be treated as the discovery of structures. By reading and collecting information about films and filmmakers, the spectator is able to establish coherence among the elements of the film and to study the film within its social context.[45] Although different spectators may use different ways to understand the same film, they all utilise knowledge and heuristics to try to make sense of avant-garde cinema.[46] From Peterson's perspective, this inference-based model of communication encourages the rational spectator to learn the irrationality and the beauty of avant-garde cinema.[47]

Both Bordwell and Peterson's cognitive approaches reveal the roles of the spectator's ordinary experience in daily life and specialised knowledge of art and film in comprehending cinematic story formats. Bordwell uses a perceptual-cognitive perspective to provide an account of how the viewer perceives stimuli and makes and tests hypotheses about previous and next events. However, it does not address other facets of viewing, such as emotion, raw sensation, and supermundane experience which can be complemented by film Buddhology. In addition, though Peterson chooses avant-garde cinema as an example to put emphasis on the viewer's intellectual engagement with the unusual, I suggest that from a Buddhist perspective, the unusual is not necessarily attributed to non-canonical story formats but instead to the variable operation of the viewer's mind. Whether a film is canonical or non-canonical, the viewer can still enter a state in which their experience becomes different from a typical sensation, which will be detailed later. Furthermore, though Buddhology also engages with the word "cognition," it is not just concerned with how the viewer's brain works when they perceive objects. It regards all sensory systems and objects as a whole and reflects on their dependent arising.

1.2.4 Auteur Theory

Different from film phenomenology, film realism, and cognitive film theory, auteur theory examines film by linking it to the director. Alexandre Astruc and François Truffaut adopt an authorial approach

---

[43] James Peterson, "Is a Cognitive Approach to the Avant-garde Cinema Perverse?" (1996), in *Post-theory: Reconstructing Film Studies*, eds. David Bordwell and Noël Carroll (Madison and London: University of Wisconsin Press, 1996), pp. 108, 115–16, 127.
[44] Peterson, "Is a Cognitive Approach to the Avant-garde Cinema Perverse?," pp. 110–11, 117, 122.
[45] Peterson, "Is a Cognitive Approach to the Avant-garde Cinema Perverse?," pp. 111, 116, 120.
[46] Peterson, "Is a Cognitive Approach to the Avant-garde Cinema Perverse?," p. 120.
[47] Peterson, "Is a Cognitive Approach to the Avant-garde Cinema Perverse?," pp. 112, 118, 123.

and relate film to literature. In "The Birth of a New Avant-garde: La Caméra-stylo," Astruc proposes *caméra-stylo* (camera-pen), with this idea suggesting that film can be seen as a form used to express an artist's own thoughts, just like a writer using a pen to write an essay or novel.[48] He further makes a contrast between film and adaptation. As for these two realms, he maintains that expressing ideas with film form does not mean to illustrate text with pictures. Rather, film can convey personal views unique to the director, with the director's perspective reflecting the ideas of a written material.[49]

Similar to Astruc's contention, in "A Certain Tendency of the French Cinema," Truffaut reflects on how to prevent film from becoming pictorial illustrations for understanding a script and to make film become film. He criticises the French cinema for being controlled by the "Tradition of Quality" (i.e. the ossified framework of the adaptations of classical French literary works), endeavouring to fight against an intense current of traditional and commercial French cinema.[50] The director can be transformed from a craftsman who only puts the screenwriter's written material onto the cinematic screen into an auteur who has their own voices throughout the entire film. Truffaut points out that "I do not believe in the peaceful coexistence of the 'Tradition of Quality' and an '*auteur's* cinema.'"[51] The purpose of raising this idea is to promote the director's status and therefore to make it just as important as a literary author's status.

In "On the Politique des Auteurs," Bazin does not only consider auteurs' roles, but also reminds us that it would be in danger if we overly praised auteurs by rejecting some admirable and estimable films made by budding talent.[52] Bazin believes that films can surpass their auteurs. Instead of blindly using auteurs as a criterion of judgement, we should consider social, historical, and technical factors in order to carry out proper evaluations of films.[53] He also points out that directors' ageing does not necessarily weaken their talents. However, some auteurs' later works could be eclipsed not because of their ageing, but because of "a disharmony between the subjective inspiration of the creator and the objective juncture of the cinema which is involved."[54] Bazin recognises auteurs' roles, but at the same time, he juxtaposes auteurs with budding talent through a reflection on the aesthetic qualities of films themselves.

---

[48] Alexandre Astruc, "The Birth of a New Avant-garde: La Caméra-stylo" (1948), in *The New Wave: Critical Landmarks*, ed. and trans. Peter Graham (New York: Doubleday, 1968), p. 18.
[49] Astruc, "The Birth of a New Avant-garde: La Caméra-stylo," pp. 18–19.
[50] François Truffaut, "A Certain Tendency of the French Cinema" (1954), in *Cahiers du Cinéma in English*, ed. Andrew Sarris, trans. Rose Kaplin, G. Hersberger, and John K. E. Hitchcock (New York: Cahiers Publishing Company, 1966), no. 1, pp. 30–31.
[51] Truffaut, "A Certain Tendency of the French Cinema," p. 37 (Note: *Auteur's* in this quotation is Truffaut's italic emphasis).
[52] André Bazin, "On the Politique des Auteurs" (1957), in *Cahiers du Cinéma in English*, ed. Andrew Sarris, trans. Rose Kaplin, G. Hersberger, and John K. E. Hitchcock (New York: Cahiers Publishing Company, 1966), no. 1, p. 18.
[53] Bazin, "On the Politique des Auteurs," pp. 8, 12.
[54] Bazin, "On the Politique des Auteurs," p. 14.

In addition to Bazin's view, the juxtaposition of auteurs' works and budding talent's works has also been revisited in the context of film archives. In *Paris Hollywood: Writings on Film*, Peter Wollen maintains that auteurism is explicit in its canon-construction.[55] Not only does this viewpoint involve the issue of film preferences and thus have an impact on public taste and canon formation, but also reveals that the archival path can put auteurs' cinema and other cinema together, which makes the film canon diverse. Wollen engages with three criteria for the "360° Pan" list of "Treasures from the National Film Archive" in London. First, the National Film Archive list challenges but also keeps some "classics" in the cinema. Second, it selects the world-recognised auteurs' masterpieces. Third, even if it values the famous auteurs' works, it also promotes less well-known, but more deserving, films.[56] Instead of only featuring auteurs' works, the National Film Archive list embraces "classics" and little-known films as parts of its collection, thereby causing the aesthetic foundation of film to be variable.

Auteur theory constitutes a particular method for reflecting on film. However, considering that my research focus throughout this thesis is to use nominal film—a conversion process between film and "no film"—as a mind-oriented perspective to engage with film ontology, which is quite different from an authorial perspective, I will not conduct a close examination of auteur theory in subsequent sections. Nevertheless, this does not mean to ignore the function of auteur theory. Although auteur theory calls attention to the relationship between film and the filmmaker, rather than between film and the viewer's mind, I suggest that film Buddhology can still complement this issue by making the transition from the presence of filmmaking to the operation of the viewer's mind. This point will be discussed in sub-section 2.4.1.

1.2.5 Film Theorizing and Film Medium Specificity

Since "no film" is a part of my discussions, I want to touch upon Carroll's idea of "film has no essence" in *Theorizing the Moving Image*.[57] For Carroll, the essence of film can be scrutinised at the levels of theorizing and medium specificity. Theorizing is an ongoing process of examining theories, revealing that there is no perfect theory, as all theories are open-ended and can be revisited.[58] The nature of film is changeable, as it hinges on theorists' stylistic preferences, for example realist and montage styles.[59] As Carroll suggests, theorizing transforms "What is cinema?" into "What can be cinema?"

---

[55] Peter Wollen, *Paris Hollywood: Writings on Film* (London and New York: Verso, 2002), p. 224.
[56] Wollen, *Paris Hollywood: Writings on Film*, pp. 219–20.
[57] Noël Carroll, "From Real to Reel: Entangled in Nonfiction Film," in *Theorizing the Moving Image*, ed. Noël Carroll (Cambridge, New York, and Melbourne: Cambridge University Press, 1996), p. 245.
[58] Noël Carroll, "Introduction," in *Theorizing the Moving Image*, ed. Noël Carroll (Cambridge, New York, and Melbourne: Cambridge University Press, 1996), p. xiii.
[59] Noël Carroll, "Medium Specificity Arguments and the Self-consciously Invented Arts: Film, Video, and Photography," in *Theorizing the Moving Image*, ed. Noël Carroll (Cambridge, New York, and Melbourne: Cambridge University Press, 1996), p. 4.

and emphasises that it is the process of thinking, rather than a finished product of thinking, providing infinite possibilities for analysing the essence of film. A particular way of thinking initiates a particular framework for film ontology. When a thinking process alters, this framework is also reconfigured. The so-called perfect theory about film ontology has no fixed entity.[60] In addition, he considers that if the essence of film is analysed through a medium-specificity perspective, in other words, if the nature of film is put into dialogue with other media such as drawing, television, and video, it cannot be seen as a unique medium, as it shares some art forms with other media.[61]

Carroll does not mean to suggest that we must not give answers to the question of the essence of film. As he argues, "even if the doctrine of medium specificity and the sort of essentialism it espouses are false, it still may be the case that cinema has an essence."[62] He still proposes the five conditions for the essence of film. First, film is an alienated, disembodied, and detached display. Second, films, even static films, are part of the class of things from which the impression of movement is possible. Third, film performance is generated from a template such as a film print, a videotape, a laser disk, and a computer programme. Fourth, film performance is not artistic. Fifth, film is two-dimensional.[63] He further asserts that his "definition will not have stylistic ramifications for what film artists should and should not do."[64] This position implies that these five conditions are not eternal. They are changeable, depending on the particular contexts produced by film artists.

A question needs to be reconsidered: how are Carroll's five conditions of the essence of film related to his seemingly contradictory contention that "film has no essence?" A Buddhist perspective on this ontological issue will lead to nominality making the being and non-being of film become one. However, this does not signify that the being and non-being of film are completely identical. Rather, their oneness means that there is a conversion process between them, which makes them become each other's accompaniments.

In the next part, I will scrutinise film Buddhology—a philosophical methodology for reflecting on the question of what film ontology can be—and adopt it to revisit other theorists' works. Before utilising film Buddhology to fill in the gaps, I would like to first examine the distinction between film Buddhology and Buddhism-in-film, as film Buddhology could be easily misread as Buddhism-related films. I will also reiterate why the mind in Buddhology is used as a vital facet of my thesis and discuss how its several sub-ideas are selected from among many Buddhist sources. In order to organise and clarify these sub-ideas, I will draw in the following section a tree diagram of the mind in Buddhology

[60] Carroll, "Introduction," p. xiii.
[61] Carroll, "Defining the Moving Image," pp. 50–52.
[62] Noël Carroll, "Questioning Media," in *Theorizing the Moving Image*, ed. Noël Carroll (Cambridge, New York, and Melbourne: Cambridge University Press, 1996), p. 2.
[63] Carroll, "Defining the Moving Image," pp. 62–71.
[64] Carroll, "Defining the Moving Image," pp. 54–55.

applied to film ontology; the purpose of this diagram is to provide a general Buddhist context of the mind. Hence, some more detailed Buddhist notions are not listed in this diagram. Yet, those unlisted ideas still belong to the mind's sub-ideas and will be analysed in the subsequent case study chapters.

1.3 Film Buddhology and Buddhism-in-film

My theoretical approach is film Buddhology. This term should not be misconstrued as Buddhism-in-film. Buddhism as a religion has its own religious form, habit, and behaviour, including becoming a monk, building a temple, begging for alms, and conducting a ritual, whilst studying Buddhist theories belongs to Buddhology. This contrast suggests that Buddhism-in-film refers to Buddhist culture-related films and films with Buddhist preaching (Buddhist iconography and contemplation in films as a particular belief system or as a practical way of living), whilst film Buddhology is a philosophical methodology for reconsidering the question of what film ontology can be. Therefore, film Buddhology can be applied to any types of films, rather than merely films having salient references to Buddhist ritual and contemplative practice.

Victor Fan provides a Buddhist analysis of film ontology in *Cinema Approaching Reality: Locating Chinese Film Theory*. From Fan's perspective, "to the question 'What is cinema?' a Buddhist philosopher would probably answer, '*rushi*' (like/is; as is)."[65] (An account of *rushi* 如是 will be provided in sub-section 1.4.4.) Furthermore, according to Fan's analysis of film ontology, though he uses a Buddhist methodology to discuss film ontology, rather than Buddhism-in-film, the term that he uses throughout his work is still "Buddhism."[66] I argue that Buddhism as a very religious word that he applies to his analysis could be easily misconstrued in the same way as an analysis of Buddhism-in-film. Thus, I want to use the term "Buddhology" to highlight its theoretical quality.

I will employ the mind in Buddhology to explore the nominality of film ontology. To be specific, the mind in Buddhology refers to experience such as the processes of seeing, hearing, considering, and emotionally feeling objects.[67] In this context, the mind is not *yishi* 意识 (mental consciousness), but mental consciousness is an aspect of the mind. In other words, not just the ideation of thoughts, any types of cognition, for example the seeing of sights and the hearing of sounds, are facets of the mind. It embraces a broad scope of experience. The reason why I use the mind is because it provides

[65] Victor Fan, *Cinema Approaching Reality: Locating Chinese Film Theory* (Minneapolis: University of Minnesota Press, 2015), p. 14.
[66] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 77–81.
[67] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997), pp. 61, 64.

a way to rethink the essence of all phenomena.[68] I argue that it can be used to discuss how the mind of film—film experience—engages with film. The mind as an influential idea is pervasive in a variety of Buddhist traditions including *Theravāda* (the School of the Elders), *Mahāyāna* (the Great Vehicle), and *Vajrayāna* (the Diamond Vehicle). Yet, I will not carry out a comparative study of these Buddhist traditions. Rather, I will use the general meaning of the mind to re-examine film ontology. Therefore, choosing the mind as a part of my thesis is not dependent on the choice of a particular tradition. As film ontology is a theoretical topic, my application of the mind is predicated on its theoretical aspect, rather than practical. I will not draw on the mind to address practical problems in daily life, such as how to reduce greed, aversion, fear, anxiety, and crisis and how to sing Buddhist songs for praying.

I will put film ontology into the context of three Buddhist mind-based sub-ideas: *rulai* 如来 (the Buddha), *weishi sifen* 唯识四分 (the four aspects of cognition) (*weishi* is also known as *xinjing* 心镜 [the mind mirror]), and *weishi sanliang* 唯识三量 (the three principal means of knowledge).[69] As these three sub-ideas are concerned with the interconnection between phenomena and the mind, I suggest that they can also be used to examine the ontological relationship between film and audiences' minds. In order to map and clarify these sub-ideas, I provide a tree diagram of the mind in Buddhology applied to film ontology (table 1) in which some sub-ideas encompass several branches. The Buddha and the four aspects of cognition will be analysed in the next section, providing a broad picture of a Buddhist view on film ontology. The three principal means of knowledge will be placed into the case study chapters, for they are detailed sensory branches of the mind helpful in discussing specific audio-visual issues in films. However, I would like to reiterate that throughout this thesis, though some other specific terms are not listed in the diagram below since it means to show a general Buddhist context of the mind, those unlisted terms are still linked to this Buddhist mind strategy. They will be scrutinised in the subsequent case study chapters.

---

[68] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), pp. 34–35, 1529.

[69] •In *Cinema Approaching Reality: Locating Chinese Film Theory*, on page 220, Fan argues that a Buddha is also termed *rulai* 如来 (like-approaching or as though [it were] approaching).
•In "Liang Shuming and His Confucianized Version of Yogācāra" edited in *Transforming Consciousness: Yogācāra Thought in Modern China*, on page 224, Thierry Meynard translates *weishi sifen* 唯识四分 into the four aspects of cognition; on page 222, Meynard translates each aspect into English. I will expand on them in sub-section 1.4.1.
•In *Zongmi on Chan*, on page 313, Jeffrey Broughton translates *xinjing* 心镜 into the mind mirror. The mind mirror is a figurative paraphrase of cognition (*weishi* 唯识).
•In the *Encyclopedia of World Religions* (page number unknown), Bruno Becchio translates *pramāṇa* (*liang* 量) into the means of knowledge. As *liang* covers three aspects, they are known as *weishi sanliang* 唯识三量 (the three principal means of knowledge). The specific branches of perception and inference will be translated by myself. As for the specific facets of specious perception and specious inference, in "Lorig: Ways of Knowing" (page number unknown), Alexander Berzin translates them into English. I will expand on them in sub-section 2.4.2.

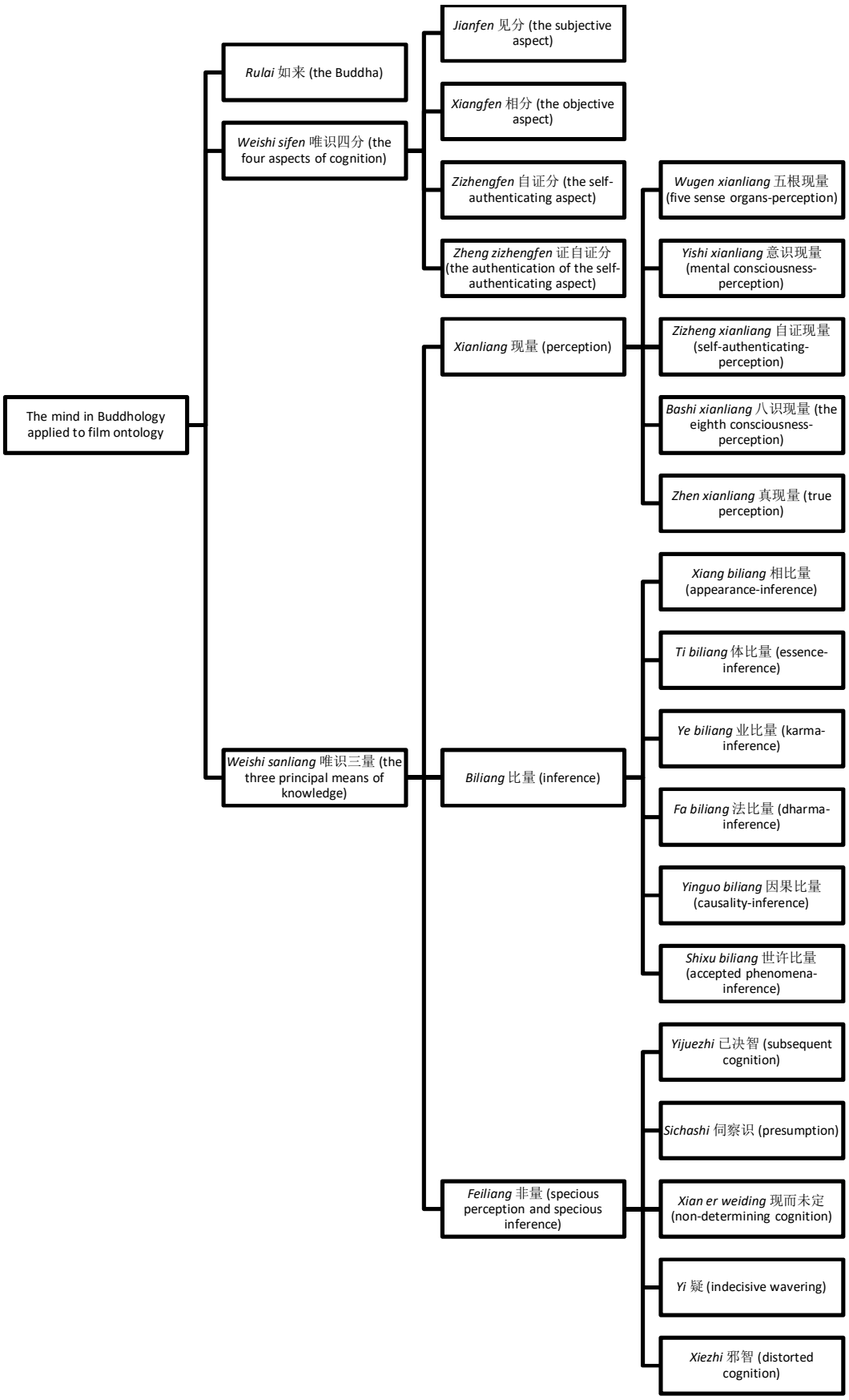

Table 1 A tree diagram of the mind in Buddhology applied to film ontology

1.4 A Buddhist View on Film Ontology

1.4.1 Nominal Film and *Weishi sifen* 唯识四分 (the Four Aspects of Cognition)

Previously, I examined the general Buddhist meaning of the mind which refers to the experience of objects.[70] According to the *Avataṃsaka Sūtra* (also known as the *Flower Ornament Scripture*, translated into the first Chinese version circa 420 as the *Huayan jing* 华严经 by Buddhabhadra, 359–429, and into the second circa 699 by Śikṣānanda, 652–710), "reality is […] a projection of the mind."[71] This idea implies that all phenomena are illusions since they are devoid of existing in an independent and inherent manner. All of them are compounds and therefore possess no self-essence.[72] This contention also points out that due to the mind of attachment regarding the object as able to manifest itself in its own right, it seems as if the object itself existed independently and inherently. However, the manifestation of the object needs to depend on the mind perceiving it. Without the mind, the object cannot be manifested to the perceiver, but rather serves as a potentiality that is not yet able to appear.[73] The object's dependence on the mind further suggests that the object is subject to the variable operation of the mind. When the perceptual mode is reconfigured, the perceived object will also be transformed into multiple appearances, which signifies that it has no fixed entity.[74]

It is my contention that nominal film can be further investigated from the perspective of *weishi sifen* 唯识四分 (the four aspects of cognition). This notion consists of the following portions: *jianfen* 见分 (*darśana-bhāga*/the subjective aspect), *xiangfen* 相分 (*nimitta-bhāga*/the objective aspect), *zizhengfen* 自证分 (*svasaṃvitti-bhāga*/the self-authenticating aspect), and *zheng zizhengfen* 证自证分 (*svasaṃvitti-saṃvitti-bhāga*/the authentication of the self-authenticating aspect).[75] First, the subjective aspect signifies the perceiver or the perceiving part, revealing that the mind has an ability to discern and understand differences among objects. Second, the objective aspect refers to objects or the perceived part which can be further subdivided into two kinds: *yingxiang xiangfen* 影像相分

[70] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, pp. 61, 64.
[71] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 1529.
[72] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 34.
[73] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 35.
[74] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 1529.
[75] Yongming Yanshou, *Zongjing lu* 宗镜录 [The Records of the Source Mirror] (961) (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2015), vol. 60, p. 1176 (Note: Even if Shi Yanshou is the name on the front cover of the book, "Shi" is a unified title of Buddhists. His name is Yongming Yanshou).

(*pratibimba*/the noema qua image) and *benzhi xiangfen* 本质相分 (*bimba*/the noema qua original stuff). The noema qua image borrows a metaphor from the image in the water, emphasising that it is similar but not identical to the noema qua original stuff. It serves as the content of cognition, with its experienced form arising as a result of the interdependence of external stimuli and the functions of sense organs and sense consciousness. It can be directly perceived by the mind, whilst the noema qua original stuff cannot be directly perceived. The act of identifying the noema qua image catalyses the formation of a conceptual construct overlaid by the conceptual mind on the noema qua original stuff. This conceptual construct is not the original stuff in itself. Third, the self-authenticating aspect is also known as self-cognition, suggesting that the perceiver's mind can confirm their own activities of experiencing objects. The subjective and objective aspects arise simultaneously through the self-authenticating aspect. To put it another way, the perceiving part is verified at the very moment that the perceived part is verified. Neither is possible without the self-authenticating aspect. Fourth, the authentication of the self-authenticating aspect means the cognition of self-cognition, implying that the mind can consider why things are the way the perceiver senses them.[76]

To make the above four aspects figurative, in the *Zongjing lu* 宗镜录 (also known as the *Xinjing lu* 心镜录, written by Yongming Yanshou in 961, translated into English as the *Records of the Source Mirror* or the *Records of the Mind Mirror*), Yanshou uses the metaphor of a mirror to further explain the four aspects of cognition.[77] For Yanshou, the subjective aspect symbolises the mirror's function to reflect objects. The objective aspect signifies the appearances of objects on the surface of the mirror. The self-authenticating aspect represents the surface of the mirror. This surface shows the appearances of objects. The authentication of the self-authenticating aspect denotes the back of the mirror. The reason why things are the way we sense them is because we merely pay attention to the front of the mirror and ignore its back. This consideration makes us aware that when the mirror is turned over, the appearances on the front are cut off from the perceiving process and thus vanish.[78]

During film viewing, we may believe that there is a fixed self to which we cling: this is my mind; that is a film outside my mind. We are usually controlled by the surface of the mind mirror, as there are a variety of filmic elements reflected in the mind mirror that attract our attention. We have not considered turning the mind mirror over and examining its back because there seems to be nothing to note on its back. The neglect of its back causes us to consider that film has an independent and

[76] Yongming Yanshou, *Zongjing lu* 宗镜录 [The Records of the Source Mirror] (961) (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2015), vol. 55, pp. 1083–85, vol. 60, p. 1176 (Note: Even if Shi Yanshou is the name on the front cover of the book, "Shi" is a unified title of Buddhists. His name is Yongming Yanshou).
[77] Yanshou, *Zongjing lu*, pp. 1176, 1182.
[78] Yanshou, *Zongjing lu*, pp. 1176, 1182.

inherent quality and to ignore our oneness with film. Yet, when the mind mirror is constantly turned over, it is presented neither as the static front nor as the static back. Rather, it becomes a dynamic and revolving mind mirror as a whole, namely nominal film juxtaposing film with "no film." Film's dependence on the viewer's mind implies that the audio-visual appearances of film vary with the operation of the mind. Film arises, but cannot maintain its existential state due to its impermanence.

1.4.2 The Movie Camera as a Mind Assistant

Although both Kracauer and Balázs discuss film ontology in terms of the technical ability of the movie camera, their ideas can still be put into a Buddhist context and re-examined at a mind level. This sub-section aims to make the transition from an analysis of the movie camera to an examination of the viewer's mind. The movie camera is not only regarded as a mechanical device, but also a mind assistant. By a mind assistant I mean that the movie camera's ability to record phenomena in detail can help the viewer enhance their potential ability to observe objects carefully.

Kracauer mentions that without the movie camera's ability to capture unobtrusive and transitory moments, it would be quite difficult for us to experience them.[79] From a Buddhist perspective, this also suggests that these moments cannot manifest themselves to the viewer in their own right. The movie camera's technical ability assists the viewer to extend their act of seeing. Because of the extension of their visual experience, imperceptible instants can be brought into a state of being. Without the enhancement of visual experience, these transient instants merely serve as potentialities that are not yet able to appear. At the Buddhist level of compounds or non-self, inconspicuous and fleeting instants are devoid of existing in an independent and inherent manner, considering that their manifestations need to depend on perceptual enhancement. However, this is not to negate the perceived moments. Rather, due to perceptual extension, the viewer can feel the sensation of moments. If moments were true nothingness, then even if visual experience were extended, they should never be manifested to the spectator. Because the arising of this sensation cannot be negated, the perceived moments still have relative entities.

Likewise, I argue that Balázs' physiognomy[80] is related to the Buddhist idea of compounds or non-self. A subtle facial expression is not an independent and inherent phenomenon. It cannot manifest itself to the spectator in its own right. The close-up technique of the movie camera encourages the viewer to develop their ability to observe objects closely. Perceptual extension makes the subtle motion of facial muscles become explicit. Without close observation, this implicit visual detail simply acts as a potentiality that is not yet able to arise. Thus, a subtle facial expression

[79] Kracauer, *Theory of Film: The Redemption of Physical Reality*, p. ix.
[80] Balázs, *Theory of the Film: Character and Growth of a New Art*, p. 61.

has no self-essence. Yet, this does not mean to deny the perceived expression. The reinforcement of observation allows the viewer to feel the sensation of the expression. As with moments, the existential quality of the expression is also in a nominal state.

1.4.3 Film Theorizing and *Rulai* 如来 (the Buddha)

Not only is nominality connected to cinematic physical and mental phenomena, but also to Carroll's theorizing.[81] I would like to clarify that despite the fact that the idea of theorizing does not focus on the relationship between film and audiences' minds, the Buddhist idea of nominal existence can be applied to theorizing because it can engage with the relationship between film and theorists' minds, which, I argue, implies a potential resemblance between these two relationships. Likewise, there is no inherently findable film validating its fixed entity from its own side because it hinges on theorists' stylistic preferences. It seems as if theorists acquired the nature of film, but its existential quality is illusory. This theorizing-oriented ontological issue can be further analysed in a Buddhist framework.

In the *Vajracchedikā Prajñāpāramitā Sūtra* (also known as the *Diamond Sūtra*, translated into Chinese circa 401 as the *Jingang jing* 金刚经 by Kumārajīva, 344–413), the Buddha – Siddhārtha Gautama – tells Subhūti that "*fan suoyou xiang*, *jieshi xuwang*. *Ruo jian zhuxiang feixiang*, *ze jian rulai*." "凡所有相，皆是虚妄。若见诸相非相，则见如来。" (all with marks is empty and false. If you can see all marks as no marks then you see the Tathāgata).[82] This Buddhist view can be understood in two ways. First, *rulai* 如来 or *Tathāgata* refers to Gautama Buddha. This term can be further divided into two words: *ru* 如 and *lai* 来. *Ru* 如 means like or as if; *lai* 来 means approaching or coming. Fan points out that "a Buddha is often referred to as *rulai* (like approaching or as though [it were] approaching). What it signifies is an appearance that *is* there, but has yet to be fully actualised."[83] This Buddhist idea is predicated on *rulai zang* 如来藏 or *Tathāgatagarbha* (the womb or embryo of the thus-gone). *Zang* 藏 or *garbha* is a storehouse of *rulai* or *Tathāgata*. Every sentient being carries a *rulai* or *Tathāgata* within them in seed, with this potentiality being able to grow into Buddhahood. If one engages with Buddhahood, then when a form appears, it becomes the unity of

---

[81] Carroll, *Theorizing the Moving Image*.

[82] Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007), p. 35; in the English translated version *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra*, the original text is accompanied by the commentary provided by Xuanhua (the pinyin form of Hsüan Hua [Hsüan Hua is a Wade–Giles-based name]), 1918–1995, trans. Heng Zhi (the pinyin form of Heng Chih [Heng Chih is a Wade–Giles-based name]) (San Francisco: Sino-American Buddhist Association, Inc., 1974), p. 95.

[83] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, p. 220 (Note: The word *is* in this quotation is Fan's italic emphasis).

appearance and emptiness.[84] Although Fan employs *rulai* to dissect cinematographic reality, which is quite different from the discussion of Carroll's theorizing, I suggest that *rulai* can also engage with theorizing in some ways. On the one hand, the essence of film can be specified at a relative level (e.g. realist or montage style), but on the other hand, its essence has yet to be fully actualised. Because its essence is dependent on theorists' minds and thus changed by theorists' distinct stylistic preferences, it has no independent and permanent existential quality. In this light, film ontology is situated at a threshold between initiation and extinction. Like *rulai*, once we have seen through and transcended the boundaries between film and theorists' minds, between being and non-being, we will realise that they are all one. As Subhūti tells Gautama Buddha, "*shi shixiang zhe zeshi feixiang*, *shigu rulai shuo ming shixiang*." "是实相者则是非相，是故如来说名实相。" (the real mark is no mark, therefore the Tathāgata calls it the real mark).[85] The being of film is non-being, and vice versa.

Second, what Gautama Buddha imparts to Subhūti is not meant to encourage us to deny all phenomena and to be attached to their lack of essence. As Khyentse Norbu emphasises, if we try to deny something, we "have to acknowledge that there is something to negate in the first place."[86] Likewise, this is not to say that film does not exist somewhat. Rather, it still exists relatively. As Gautama Buddha points out, "without attachment to it, it is empty. Without attachment to emptiness it is false. Abiding in the emptiness and falseness without attachment is the Middle Way."[87] As stated earlier, though nominality is different from the Middle Way, the former serves as the conceptual or linguistic embodiment of the latter, which reveals that nominality as a provisional name or label is subject to concept sweeping. The above Buddhist mindset suggests why Carroll still proposes the five conditions for film even if he states that it has no essence. Carroll does not cling to non-essentialism. Rather, he regards non-essentialism as a part of the process of theorizing film.

1.4.4 Cinematographic Reality and *Rushi* 如是 (Like/Is; As Is)

I mentioned in section 1.3 that in *Cinema Approaching Reality: Locating Chinese Film Theory*, Fan's Buddhist answer to the philosophical question of film ontology is *rushi* 如是 (like/is; as is). He argues

---

[84] Buddhabhadra, *A Buddha within: The Tathāgatagarbha Sūtra: The Earliest Exposition of the Buddha-nature Teaching in India* [Da fangdeng rulai zang jing 大方等如来藏经] (The first Chinese translated version was completed by Buddhabhadra in 420, and the second by Amoghavajra in the middle of the eighth century. The canonical Tibetan translation was completed circa 800, and the paracanonical Tibetan translation by Bathang), trans. Michael Zimmermann (Hachiōji: The International Research Institute for Advanced Buddhology at the Sōka University, 2002), pp. 44, 62–63.

[85] Ānanda, *Jingang jing* (Chinese), p. 65; *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra* (English), p. 159.

[86] Khyentse Norbu, *What Makes You Not a Buddhist* (Boston and London: Shambhala Publications, Inc., 2007), p. 66 (Note: Even if Dzongsar Jamyang Khyentse, also known as Khyentse, is the name on the front cover of the book, Khyentse is his first name. His surname is Norbu).

[87] Ānanda, *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra* (English), p. 147.

that "the phrase *rushi* preserves an indeterminacy between likeness and being, and is first and foremost understood as 'the way it is.'"[88] From Fan's Buddhist perspective, "cinematographic reality is at a threshold between here and there, present and past, presence and absence. It is neither virtual nor actual, being or becoming. It is neither a state nor a nonstate. It neither proliferates nor distinguishes."[89] Not only is Fan's discussion of *rushi* linked to the moments of life, but also to the process of observation-reflection. For Fan, the cinematographic image has an ability to put audiences in direct perceptual contact with life in all its presence and concreteness. He also notes that what audiences observe is not only the representation of life as a trace of the passage of time-passed preserved in film, but also the here-and-now object invoking in audiences' minds a reflection of themselves.[90] He goes on to suggest that the image serves as a reflection of audiences' subjectivity, which enables them to become one with the object. This oneness reveals that audiences regard life as actualised. Hence, for Fan, what audiences observe is considered an intermediate area between like and is.[91] However, I suggest that my idea of nominal film is not limited to a Buddhist account of a cinematographic threshold between the passage of time-passed and life in all its presence and concreteness. In the subsequent case study chapters, I intend to analyse how the viewer's mind affects, alters, voids, and retains cinematic elements.

1.5 A Rationale for the Choice of the Films as Case Studies

I will scrutinise the following four films: the art film *Blow-up* (Michelangelo Antonioni, 1966), the experimental film *Zorns Lemma* (Hollis Frampton, 1970), the docufiction *Through the Olive Trees* (Abbas Kiarostami, 1994), and the classical Hollywood film *The Best Years of Our Lives* (William Wyler, 1946). Considering that the dominant examples of discussions of film ontology are Kracauer and Bazin's accounts of the cinematographic image and that one of the primary traits of *Blow-up* is the existential power of images, I select it as my first case study. In film studies, Kracauer and Bazin's preference for realism is often contrasted with montage. I intend to examine *Zorns Lemma* because it features the contrast between montage and the long take. Its long-take continuity is accompanied by voices which will be regarded as one of the clues as to a fragmentary potentiality within the long take. Although the question of the vocal ontology is not my primary focus of the discussion of *Zorns Lemma*, it will be scrutinised in the subsequent case study *Through the Olive Trees* since it highlights an aesthetic process of how voices alter what audiences sense. As one of the Buddhist ideas about sounds can also be applied to and open up my Buddhist examination of deep focus, I plan to choose *The Best Years of Our Lives* as my following case study due to its influential use of deep-focus shots.

---

[88] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, p. 220.
[89] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 221–22.
[90] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 93–105.
[91] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 93–105.

These four examples constitute different types of films, which will open up an opportunity to engage with the universality of film Buddhology. Moreover, they are neither Buddhist culture-related films nor films with Buddhist preaching, which suggests that film Buddhology is not limited to Buddhist films. In addition, since these case studies have already received considerable attention in multiple theorists' works, they have been widely regarded as canonical examples. For instance, *Blow-up* as a famous art film has been examined in Peter Brunette's *The Films of Michelangelo Antonioni*, Seymour Chatman's *Antonioni: Or, the Surface of the World*, and Bordwell's "The Art Cinema as a Mode of Film Practice."[92] Likewise, numerous theoretical works have engaged with *Zorns Lemma*, which enables this film to earn an influential position in the realm of experimental cinema. Examples include P. Adams Sitney's *Visionary Film: The American Avant-garde, 1943–2000*, Maureen Turim's *Abstraction in Avant-garde Films*, Scott MacDonald's *Avant-garde Film: Motion Studies*, William Verrone's *The Avant-garde Feature Film: A Critical History*, and Allen Weiss' "Frampton's Lemma, Zorn's Dilemma."[93] As for *Through the Olive Trees*, it has been acknowledged as a renowned docufiction. The prime examples of examinations of this film cover Laura Mulvey's *Death 24x a Second: Stillness and the Moving Image* and Akram Zaatari and Abbas Kiarostami's "Abbas Kiarostami."[94] With respect to *The Best Years of Our Lives*, it has been widely recognised as one of the greatest classical Hollywood films ever made. The prominent examples of the research on this case study embrace Bordwell, Janet Staiger, and Kristin Thompson's *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960* and Bazin's "William Wyler, or the Jansenist of Directing."[95] Throughout this thesis, I suggest that these chosen case studies are all connected to a state of being nominal at a Buddhist ontological level, which further suggests the applicability of nominality beyond any types of films. Each film will be dissected from a specific nominal perspective. These nominal dimensions are but not limited to nominal images, nominal montage, nominal voices,

---

[92] Peter Brunette, *The Films of Michelangelo Antonioni* (Cambridge, New York, and Melbourne: Cambridge University Press, 1998); Seymour Chatman, *Antonioni: Or, the Surface of the World* (Berkeley, Los Angeles, and London: University of California Press, 1985); David Bordwell, "The Art Cinema as a Mode of Film Practice" (1979), in *The European Cinema Reader*, ed. Catherine Fowler (London and New York: Routledge, 2002), pp. 94–102.

[93] P. Adams Sitney, *Visionary Film: The American Avant-garde, 1943–2000* (New York: Oxford University Press, 2002); Maureen Turim, *Abstraction in Avant-garde Films* (Ann Arbor: UMI Research Press, 1985); Scott MacDonald, *Avant-garde Film: Motion Studies* (Cambridge and New York: Cambridge University Press, 1993); William Verrone, *The Avant-garde Feature Film: A Critical History* (Jefferson and London: McFarland and Company, Inc., 2012); Allen Weiss, "Frampton's Lemma, Zorn's Dilemma" (1985), in *October*, eds. Rosalind Krauss and Annette Michelson (Cambridge: MIT Press, 1985), vol. 32, pp. 118–28.

[94] Laura Mulvey, *Death 24x a Second: Stillness and the Moving Image* (London: Reaktion Books Ltd, 2006); Akram Zaatari and Abbas Kiarostami, "Abbas Kiarostami" (1995), in *Bomb*, eds. Betsy Sussler, Andrew Bourne, and Sabine Russ (New York: New Art Publications, Inc., 1995), no. 50, pp. 12–14.

[95] David Bordwell, Janet Staiger, and Kristin Thompson, *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960* (London: Routledge, 1985); André Bazin, "William Wyler, or the Jansenist of Directing" (1948), in *Bazin at Work: Major Essays and Reviews from the Forties and Fifties*, ed. Bert Cardullo, trans. Alain Piette and Bert Cardullo (New York and London: Routledge, 1997), pp. 1–22.

and nominal deep-focus shots, serving as manifold Buddhist avenues towards the ontological issues of nominal film.

1.6 Chapter Plan

The second chapter is concerned with the art film *Blow-up*. I will scrutinise this film in terms of nominal images. My argument in this chapter is that images do not exist, but also exist. First, I will outline the general features of art cinema and use Buddhist mind theory to delve closely into the difference between the representation of reality and reality itself. By analysing their difference and the operation of audiences' minds, I will initiate a nominal perspective on the entity of the image. Second, I will employ the nominal realm of the image to re-examine art cinema's oneness with Hollywood cinema and other types of cinema. Third, the nominal image will be utilised to readdress Antonioni's doubt about the revealed image. Fourth, I plan to draw on the various branches of *xianliang* 现量 (perception), *biliang* 比量 (inference), and *feiliang* 非量 (specious perception and specious inference) to rethink the ontological state of the images in *Blow-up*. I argue that the images in *Blow-up* cannot manifest themselves as pure images, for the spectator also sees the product of their own perception and inference coupled to images. It seems as if visual objects such as props, characters, and scenery had the appearance of being separate entities. However, because all of their manifestations are affected and configured by the viewer's mind, they are just compounds devoid of existing in an independent and inherent manner.

The third chapter examines the experimental film *Zorns Lemma*. I will engage with this film at the level of nominal montage. My argument in this chapter is that there is no fundamental division between montage and the long take because at a mind level, the long take carries the potential of a cut. First, I will adopt *yinguo biliang* 因果比量 (causality-inference) to set up an ontological link between montage and the long take. Second, I will scrutinise the nominal issues of the styles of experimental cinema. Third, I will pick *xiang biliang* 相比量 (appearance-inference) as a way of thinking to analyse the montage and long takes in *Zorns Lemma*. In the first place, at the levels of interdependence and impermanence, what the viewer perceives is not the long-take object, but rather a compound of the long-take object and other accompanying objects such as voices. If any component of this compound alters, the quality of the entire compound changes. It is the variation of the relation among objects giving rise to fragmentary traces. Furthermore, I will link appearance-inference to the blink of an eye to discover the uncertainty of the number of cuts. In addition, I will synthesise appearance-inference and the meaning-generating process to discuss the fragmentary meanings of the long take. Apart from a fragmentary state within the long take, I will draw on the Buddhist idea of unity to revisit a continuous rhythmic or conceptual flow within a montage system.

The fourth chapter engages with the docufiction *Through the Olive Trees*. I will delve into nominal voices in this film. Although one of the prominent facets of this film is an expression of how voices alter what audiences sense, I want to reverse the relation between voices and audiences and dissect how audiences' minds reconfigure the appearances of voices. Furthermore, I would like to emphasise that my Buddhist examination of voices will centre around their acoustic qualities, rather than their linguistic content. My argument in this chapter is that there is no thorough boundary between voices and "no voices," for "no voices," at a mind level, have the potential of voices. I will explain this statement from two aspects. One is that even if a film has no vocal tracks, vocal interventions are still present because of audiences' interpretive voices. Another is that the voice has no independent and permanent nature making itself fundamentally distinct from other voices, but there are still relative distinctions among them. In this chapter, first, I will encapsulate the general features of docufictions. Second, I will use *yinguo biliang* 因果比量 (causality-inference) to rethink the vocal aspect of docufictions and to conduct analyses of the interventions of voices from audiences' experiences, the hetero-homogeneity of sounds, and vocal nominality in six documentary modes. Third, multiple Buddhist ideas can be employed in conjunction with causality-inference, opening up a variety of perspectives on the vocal ontology in *Through the Olive Trees*. Specifically, I will use the Buddhist experiential process of vocal experience, the void, Buddhist echo-like phantoms, and the seventh and eighth consciousness to illuminate the nominal qualities of the opening voices, off-screen voices, vocal echoes, and ending voices.

The fifth chapter deals with the classical Hollywood film *The Best Years of Our Lives*. I will revisit this film in terms of nominal deep-focus shots. Although my argument in this chapter is that deep-focus shots do not exist, but also exist, this view does not merely engage with the nominal existence of deep focus itself, but also the nominal entities of its affiliated forms such as composition in depth and long-take continuity. First, I will use a Buddhist account of sensory consciousness to deconstruct deep focus. Second, I will summarise the overall characteristics of classical Hollywood cinema. Third, I will employ a Buddhist visual aspect (eyes and eye consciousness) of five sense organs-perception (*wugen xianliang* 五根现量) to analyse deep focus. Specifically, I will engage with *zixiang* 自相 (an individual and primordial appearance, which by definition transcends linguistic representation) and a Buddhist view on the retina to rethink selective focus and utilise knowledge-associated intuitive perception to canvass the flux of composition in depth. Fourth, regarding the deep-focus existential issue in *The Best Years of Our Lives*, I will draw on the Buddhist ideas of conceptual thoughts, mirrors, particles, the heavenly eye, and distraction to scrutinise deep focus' relations to the frames, mirror images, continuity, sizes, occlusion, and attention. These aspects, at a holistic level, will help readers consider the question of whether a wide range of sharpness exists in a truly inherent manner.

Chapter 2 Nominal Images in Art Cinema

2.1 Nominal Images

This chapter provides an analysis of nominal images in art cinema. My argument is that images do not exist, but also exist. In order to elucidate "images do not exist" or "no images," I would like to first engage with Alexander Berzin's Buddhist discussion of the association between phenomena and the mind. He argues that if the perceiver wants to know how the manifestations of things work, they can only do so by using their mind.[96] More specifically, he employs two principal Buddhist perspectives to discuss why phenomena have no inherent and findable self-essence:

> The first presents what exists in terms of phenomena being either mind or objects of mind – in other words, experience or the contents of experience. Phenomena, including minds, exist merely by virtue of the fact that mind can simply give rise to an appearance or occurrence of them as an object of cognition. We can establish that our children and love for them exist simply because we can know and experience them. The other major approach discusses what exists in terms of mental labelling, which means things exist as what they are simply in relation to words and what words refer to or signify. Phenomena exist as what they are by virtue of being simply the meaning of the words, mental labels, or conceptual formulations of them. We can establish that our children and love exist simply because we can give them names that refer to them.[97]

Likewise, this Buddhist statement can be used as a theoretical basis to examine the existential quality of the image. First, the spectator is able to confirm that the image exists because they can see it. Second, the spectator is able to use a mental label—the word "image"—to render the object seen the image. We can find out that in neither case is the existence of the image initiated from its own side by virtue of an inherently findable self-entity or a defining characteristic mark rendering it truly the image existing independently of the viewer's mind. The former situation reveals that the manifestation of the image is the play of a visual sensation, whilst the latter reveals that the manifestation of the image is empowered by a concept. This further implies that even the entity of the mind itself also depends on being labelled as the "mind." As with "no image," there is also no self-entity ultimately in evidence on the side of the mind that establishes the identity of the mind itself—in the sense of "no mind." Neither the image nor the spectator's mind can be conceived of as existing on its own.

---

[96] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997), p. 20.

[97] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, pp. 20–21.

The above analysis signifies that the image is empty of self-nature – or, in other words, the void of the image. By this I do not mean empty space represented in the image. Instead, the image's empty quality emphasises that the image itself, rather than the content of the image, is devoid of existing in an independent and inherent way. Even if the content of the image fills the entire space, an empty quality can still be applied to the image itself at a Buddhist ontological level, considering that the image does not exist independently of the spectator's mind. However, this does not mean to say that the image is the absolute void. Since the spectator is able to feel the sensation of the image, we cannot negate the arising of this sensation. If the image were the absolute void, then even if the spectator performed the acts of seeing and labelling it, it should never be manifested to the spectator. Yet, instead of seeing nothing, they are able to recognise that they see the image. In this light, the void of the image still encompasses its relative manifestation, which suggests that the image and "no image" are ontologically woven into a seamless whole.

This seamless whole is what I call the nominal image—an intermediate state where we neither cling to the image nor to "no image," neither being detached from the image nor from "no image." At every moment, being detached from the image is also a way of engaging with it cognitively. Therefore, the image and "no image" do not function one after the other, but rather function simultaneously. In this sense, when we look from the point of view of nominal images, we find out that there are no fundamental distinctions between the previous image and the next image, between the character and the scenery, between light and shadow, since none of them exist independently of eye consciousness and mental labelling and therefore all of them exist relatively.

2.2 Images and the General Features of Art Cinema

Before using a Buddhist methodology to delve further into nominal images, I would like to first look at the overall traits of art cinema. In "The Art Cinema as a Mode of Film Practice," David Bordwell defines art cinema as realistic cinema attenuating the classical narrative mode and making cause-effect logic fragile.[98] Unlike classical Hollywood cinema which usually displays clear goals and maintains the coherence of time and space, art cinema emphasises ambiguity and subjective reality and leads audiences to narrative enigmas and emotional mystery. Bordwell argues that art cinema is characterised by two foundations: realism and authorial expressivity.[99] Realism encompasses multiple aspects such as accidental and unforeseeable daily reality, real location shooting, real social conditions, and realistic or psychologically complex characters. Authorial expressivity puts emphasis on stylistic signatures in the narration – or, in other words, the presence of filmmaking.

[98] David Bordwell, "The Art Cinema as a Mode of Film Practice" (1979), in *The European Cinema Reader*, ed. Catherine Fowler (London and New York: Routledge, 2002), p. 95.
[99] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 95.

Realism encourages the spectator to explore those moments that are beyond their ability to perceive. As Ian Aitken suggests, both Siegfried Kracauer and André Bazin call our attention to our perceiving process. If our sense consciousness is extended, the concrete and indeterminate aspects of reality can be noticed.[100] I consider that though their ideas reveal that cinematographic reality is expressed by looseness, contingency, provisionality, and uncertainty, they provide less accounts of reality itself than the representational trace of reality. I will use Buddhology to re-examine art cinema and cinematographic reality in detail later.

András Bálint Kovács and Fei Mu employ a subjective-reality perspective to canvass the mental facet of reality. Kovács argues that European modern art cinema usually conveys confusing and enigmatic plots in which solutions are frequently inhibited by a labyrinthine psychological structure disengaging from the logic and unity of action and reaction, with those ambiguous and puzzling thoughts, dreams, memories, and phantasms confounding audiences' expectations of solutions.[101] In a similar way, Fei coins a term called *xuanxiang* 悬想 (suspension-imagination) and uses *kongqi* 空气 (air) as a form of suspension-imagination to examine the mental situation of the aimless life-process.[102] The function of this ontological idea is to shift the focus of the spectator's attention from cause-effect logic to inconspicuous and indeterminate mental reality. As Victor Fan clarifies, "what [Fei] would prefer to do is put together dramatic events and objects around the main theme so that the absence of the main theme would be crystallised as an air that lingers in the cinematographic image."[103] For Fan, Fei's discussion of air can be interpreted as atmosphere, referring specifically to the presence of desire's absence. Drama merely acts as a foil for the presence of desire's absence.[104] Fan also suggests that by showing ordinary daily life in all its simplicity and representing subtle facial expressions and vague actions, Fei discourages the spectator from immersing themselves in cause-effect logic, but rather encourages them to experience the potential traces of characters' unfulfilled desires suggested in images.[105] This account is linked to Laura Marks' phenomenological contention

---

[100] Ian Aitken, *European Film Theory and Cinema: A Critical Introduction* (Bloomington and Indianapolis: Indiana University Press, 2001), pp. 175–76, 183–84.

[101] András Bálint Kovács, *Screening Modernism: European Art Cinema, 1950–1980* (Chicago and London: University of Chicago Press, 2007), pp. 42, 99, 129.

[102] Fei Mu, "*Xiangxuehai* zhong de yige xiao wenti: 'Daoxu fa' yu 'xuanxiang' zuoyong" "《香雪海》中的一个小问题：'倒叙法'与'悬想'作用" ["A Small Problem in *Sea of Fragrant Snow*: 'Inverse Narration' and the Functions of 'Suspension-imagination'"] (1934), in *Yingmi zhoubao* 影迷周报 [Film Fan Weekly News], eds. Xin Xin and Zhou You (Shanghai: Qianqiu chubanshe 上海：千秋出版社 [Shanghai: Qianqiu Publications], 1934), vol. 1, no. 5, p. 82; Fei Mu, "Lüe tan 'kongqi'" "略谈'空气'" ["Let's Talk about 'Air' Briefly"] (1934), in *Shidai dianying* 时代电影 [Modern Cinema], eds. Zong Weigeng and Gong Tianyi (Shanghai: Shidai tushu gongsi 上海：时代图书公司 [Shanghai: Modern Publications], 1934), vol. 1, no. 6, p. 22.

[103] Victor Fan, *Cinema Approaching Reality: Locating Chinese Film Theory* (Minneapolis: University of Minnesota Press, 2015), p. 134 (Note: Although the surname mentioned in the original version of this quotation is "Fey," I decide to use "Fei," as "Fei" is the pinyin form of "Fey". "Fey" is a Gwoyeu Romatzyh-based surname).

[104] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 132–34.

[105] Fan, *Cinema Approaching Reality: Locating Chinese Film Theory*, pp. 122, 141.

about the haptic image which is capable of expressing visceral and emotional intensity in those any-space-whatevers transcending action.[106] Ephemeral physical moments are based on movement and the texture of light, whilst innermost feelings and desires are predicated on the psychological aspect of the presence of absence.

The above ideas about mental reality are similar to Bordwell's account of psychologically complex characters who may behave for unclear reasons and hence whose mental qualities give rise to uncertainty and correspond to life's untidiness.[107] The preceding views on mental reality bear a resemblance to Bordwell's art-cinema question "Is life just leaving loose ends?"[108] and to his contention that "when in doubt, read for maximum ambiguity."[109]

Yet, I argue that whether what the viewer perceives are contingent moments or secret desires, they are just the traces of reality, rather than reality itself. To borrow a vivid metaphor from a diamond, if reality were a diamond, all traces would act as its flat sides, rather than the diamond itself. Therefore, in addition to these physical and mental traces, it is also crucial to conduct a further analysis of cinematographic reality itself. My idea of nominal images is neither to depict an indexical imprint of unpredictable movement nor the character's hidden feelings, but rather focuses on the issue of how the viewer's mind affects, alters, voids, and retains images. Before utilising a Buddhist methodology to reconsider art cinema and cinematographic reality, I would like to mention that in film studies, the existential question of reality itself, rather than the traces of reality, still gets some attention. One of the prime examples is Aldo Tassone's interview with Michelangelo Antonioni in which Antonioni reflects on his art film *Blow-up* (1966) and reality itself:

> I wouldn't say that the appearance of reality equals reality, because there can be more than one appearance. There can also be more than one reality, but I don't know it, and I don't believe it. Maybe reality is a relation [rapport].[110] […] I don't know what reality is like. Reality escapes us, it changes continually. When we think we've reached it, the situation is already something else. I always doubt what I see, what an image shows me, because I "imagine" what's beyond that; and what's behind an image is unknown.[111]

---

[106] Laura Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses* (Durham and London: Duke University Press, 2000), p. 163.
[107] Bordwell, "The Art Cinema as a Mode of Film Practice," pp. 96, 98.
[108] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 98.
[109] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 99.
[110] Aldo Tassone, "Le cinéma italien parle: Michelangelo Antonioni" (1992), in *Michelangelo Antonioni: Entretiens et inédits 1950–1985*, eds. Carlo Di Carlo and Giorgio Tinazzi (Rome: Cinecittà International, 1992), p. 108 (Note: Peter Brunette translates Antonioni's statement quoted by Tassone into English, requoting it in his work *The Films of Michelangelo Antonioni* [Cambridge, New York, and Melbourne: Cambridge University Press, 1998], p. 121).
[111] Cesare Biarese and Aldo Tassone, *I film di Michelangelo Antonioni* (Rome: Gremese Editore, 1985), p. 136 (Note: Brunette translates Antonioni's statement quoted by Biarese and Tassone into English, requoting it in his work *The Films of Michelangelo Antonioni*, p. 125).

Although Antonioni does not sidestep the existential question of reality itself by saying he does not know, not knowing should not be considered the last word on reality itself. Rather, it is also necessary to embrace a variety of potential answers to this question. Moreover, even if Antonioni suggests that reality may be a "relation," he does not explain what kind of relation reality is. I will employ nominal images to rethink this problem at a Buddhist ontological level step by step. I would like to stress that though this chapter is concerned with nominal images in art cinema, my Buddhist view on a "relation" will also engage with images in other types of cinema. By utilising a macroscopic perspective to dissect the cinematographic ontology, we will get a better understanding of the universality of film Buddhology.

2.3 A Buddhist View on Images in Art Cinema

Instead of saying "I do not know," Buddhology gives a more specific answer to the question of reality. As suggested in the *Avataṃsaka Sūtra*, "reality is […] a projection of the mind."[112] This Buddhist idea means that all phenomena are phantoms, considering that they are devoid of existing in an independent and inherent manner. Without the operation of the mind, objects cannot be manifested to the perceiver. Hence, there are no inherently findable objects establishing their solid entities from their own side. They possess no self-nature to enable themselves to arise in their own right. Instead of merely saying an abstract word "relation," Buddhology situates reality itself in the relation (nominality) between the arising of the sensation of objects and a universal lack of independent and inherent existence.

One of the functions of nominal images is to dissolve the imaginary boundaries between art cinema and Hollywood cinema, between art cinema and other types of cinema. The image as a visual form can be reconfigured in a Buddhist framework of the *Prajñāpāramitā Hṛdaya* (also known as the *Heart Sūtra*, translated into Chinese in 649 as the *Bore boluomiduo xin jing* 般若波罗蜜多心经 by Xuanzang, circa 602–664):

> O, Śāriputra, form does not differ from voidness, and voidness does not differ from form. Form is voidness and voidness is form; the same is true for feeling, conception, volition and consciousness.
>
> Śāriputra, the characteristics of the voidness of all dharmas are non-arising, non-ceasing, non-defiled, non-pure, non-increasing, non-decreasing.
>
> Therefore, in the void there is no form, feeling, conception, volition or consciousness;

[112] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), p. 1529.

> No eye, ear, nose, tongue, body or mind; no form, sound, smell, taste, touch, mind-object, or eye realm, until we come to no realm of consciousness;
>
> No ignorance and also no ending of ignorance, until we come to no old age and death and no ending of old age and death.[113]

If audiences engage with the image and "no image" without attachment, they will maintain a state of "creating form based on the formless."[114] The image in art cinema shares this state with any types of cinema, arising from the void, dwelling during some time in audiences' minds, and simultaneously dissolving back into emptiness. My idea of nominal images can be used to re-examine Bordwell's comparative study of art cinema and Hollywood cinema. For Bordwell, the reason why art cinema is distinct from Hollywood cinema is because in general, the former has ambiguous goals, whilst the latter has clear objectives.[115] Yet, if we revisit this contrast in a Buddhist framework of the *Prajñāpāramitā Hṛdaya*, we will reach an alternative situation, which can be canvassed from the following two interrelated aspects of investigation.

In the nominal field of the image, neither the image nor "no image" is in an actual state. First, whether what the spectator experiences is the image in art cinema or the image in Hollywood cinema, none of them are established as existing from their own side, considering that it is the spectator's visual sensation on which the arising of both types of images are founded. Without the act of seeing, none of them can exist, but rather serve as potentialities that are not yet able to arise. Second, there is no ambiguous or goal-oriented appearance standing on its own feet. Rather, it is the mind of judgement labelling the image as "ambiguous" (having art cinema's style) or as "goal-oriented" (having Hollywood cinema's style). Since both the arising of the image in art cinema and the image in Hollywood cinema are upheld by mental labelling or conceptual formulations, they do not have independent entities to allow themselves to be fundamentally different from each other.

Although different types of cinema engage with the common state of "no image," this does not suggest that we should treat this pervasive "no image" as the ultimate level of cinematographic reality. If we call "no image" the "ultimate level," it will give the impression that it looks as if there were a real phenomenon named "no image" that could be grabbed as inherently existent, which runs counter to the void. "No image" does not mean that the orthodox difference between images

---

[113] Xuanzang (the pinyin form of Hsüan-tsang [Hsüan-tsang is a Wade–Giles-based name]), *Heart Sūtra* [Bore boluomiduo xin jing 般若波罗蜜多心经] (It is ascribed to Avalokiteśvara and Śāriputra, circa 568–circa 484 BCE, translated into Chinese in 649 as the *Bore boluomiduo xin jing* by Xuanzang, circa 602–664. The commentary was provided by Tanxu [the pinyin form of T'an Hsü {T'an Hsü is a Wade–Giles-based name}], 1875–1963), trans. Lok To (New York: Young Men's Buddhist Association of America, 2000), pp. 71, 84, 92, 93, 102.

[114] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 294.

[115] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 95.

in art cinema and Hollywood cinema should be forsaken. Despite no fundamental distinction, there is still a relative distinction between them. Relative differences among phenomena are the basis of communication and function within the limits of common-sense reality. Likewise, audiences' visual sensations and conceptual formulations are the principal causes of the idea of different types of images, which allows them to learn, know, and understand their respective characteristics. Hence, the state of no difference still embraces the possibility of exhibiting a variety of imaginable images.

It is noteworthy that though Bordwell focuses on their difference, he also takes their oneness into account. Yet, his approach to their oneness is distinct from mine. His contention is that there is an intricately stylistic interpenetration between these two types of cinema; they adopt and assimilate each other's expressive forms in new directions. More specifically, he chooses François Truffaut's art film *The 400 Blows* (1959) and George Roy Hill's Hollywood film *Butch Cassidy and the Sundance Kid* (1969) as examples to argue that both their endings use the stylistic form of the freeze image.[116] However, in my view, Bordwell's examination of their oneness has its limitations since the premise of his idea has to be that they use similar forms of images, narratives, and so forth. In other words, if art cinema and Hollywood cinema did not borrow and use each other's stylistic forms as references, a complex transformation putting them together would not arise and thus the boundary between them would be exaggerated. If we revisit this boundary at the Buddhist ontological level of nominal images, we realise that even if there were no stylistic interpenetration among images in different types of cinema, images would still become one because all of them are devoid of existing independently and inherently, relying on audiences' visual sensations and conceptual formulations.

Nominal images can also be utilised to reconsider Antonioni's doubt about the revealed image. Apart from Antonioni's reference to a "relation" mentioned previously, he also provides an account of the distinction between the revealed image and the one under the revealed image:

> We know that under the revealed image there is another one which is more faithful to reality, and under this one there is yet another, and again another under this last one, down to the true image of that absolute, mysterious reality that nobody will ever see. Or perhaps, not until the decomposition of every image, of every reality.[117]

In my view, this statement, combined with Antonioni's contention that "I always doubt what I see,"[118] is inclined to put more emphasis on the result of rational thinking than on intuitive

---

[116] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 100.
[117] Michelangelo Antonioni, *Sei film: Le amiche, Il grido, L'avventura, La notte, L'eclisse, Deserto rosso* (Turin: Giulio Einaudi Editore, 1964), p. xiv (Note: Allison Cooper translates this Antonioni's statement into English in Antonioni's "Preface to *Six Films*" [1964], in *The Architecture of Vision: Writings and Interviews on Cinema*, eds. Carlo Di Carlo, Giorgio Tinazzi, and Marga Cottino-Jones, trans. Allison Cooper, Dana Renga, and Andrew Taylor [New York: Marsilio Publishers, 1996], p. 63).
[118] See p. 30.

perception. In other words, it seems like doubts are more reliable than intuitions because doubts can provide access to the hidden image which is more real than (substantially different from) the one that we see directly. From a Buddhist perspective, Antonioni's doubt originates from the discriminating mind, which gives rise to the duality of the revealed image and the one under the revealed image. However, such duality-based cognition has been deconstructed in Buddhology. The *Avataṃsaka Sūtra* proposes that "any who see the enlightened free from discrimination in body and mind will have no more doubts or hesitation in regard to anything."[119] At the level of non-dual universality, both awareness of the revealed image and of the hidden one rely on the actions of the mind. None of them can show themselves to the perceiver in their own right. The arising of the revealed image hinges on our visual intuitions. The arising of the hidden one hinges on our doubts. Whether what we experience is the former or the latter, neither one is more faithful to reality than the other since nothing can separate from the mind and exist in a truly independent and inherent way. There is nothing to doubt because in order to doubt something, we need to assume at least something else as true. Since there is no true image that can be believed, there is no place for doubts to exist. Without the discriminating mind, doubts cease and therefore duality falls apart. There is no fundamental difference between the revealed image and the one under the revealed image. Both are in a state of a lack of self-existence and serve as "no images."

This does not mean that the revealed image and the one under the revealed image are completely identical. Doubts caused by the discriminating mind enable us to examine how multiple cinematographic appearances work during film viewing and remind us of the variability of the existential state of the perceived image. Such variability, in some ways, links doubts to no doubts since maintaining images in all their diversity (revealed, hidden, etc.) instead demonstrates their lack of permanent and inherent existence. To some degree, doubting something signifies that nothing out there can be doubted, and vice versa. If we cling to no doubts, then it seems as if there were an actual thing called "no image" that could be held tightly, which is in direct contradiction to the emptiness of all phenomena. By being in an intermediate situation where the difference between the revealed image and the hidden one and no difference between them are not held either too tightly or too loosely, the existential quality of the image can be initiated and extinguished at the same time. Antonioni's view of the hidden image being more faithful to reality than the revealed image can be considered one aspect of the nominal image, whilst "no image" can be treated as another aspect. Within a Buddhist context, they together constitute the seamless whole of the nominal image in which intuitions, doubts, and no doubts are one but not one.

[119] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 292.

2.4 *Blow-up* and Buddhology

2.4.1 The Transition from the Presence of Filmmaking to the Operation of the Viewer's Mind

Before employing Buddhology to engage with *Blow-up*, I would like to mention that in film studies, discussions of the images in *Blow-up* are commonly framed through an authorial perspective that assigns meaning to the filmmaker, which, I argue, underplays the spectator's role. For example, in Peter Brunette's view, "the 'true target' of Antonioni's work is a reflection on precisely what he is doing when he looks, that is, when he makes a film."[120] In *Blow-up*, the protagonist is a male fashion photographer named Thomas (played by David Hemmings) who uses his camera to photograph the surrounding things that capture his attention and carefully observes these photographs by blowing them up. During the filmmaking process, Antonioni reflects on an ontological resemblance between himself as a filmmaker and Thomas as a photographer: "I came to know reality by photographing it, when I began taking it with the movie camera, a little like in *Blow-up*; in this sense, I think that it's my most autobiographical film."[121] In order to illustrate this comment, Brunette selects Thomas' studio scene (figure 1) as an example where Thomas attempts to use his photographs to recreate a crime scene, with the photographs arranged in the form of a cinematic montage. This studio scene also recalls the filmmaker's engagement with the shaping of visual narrative.[122] Moreover, Brunette suggests that the sound editor artificially adds the sound of the wind rustling through the trees to these silent photographs in order to feature the presence of the creative process of editing a film.[123]

Figure 1 The storyboard-like photographs taken by Thomas in *Blow-up*

Furthermore, Bordwell considers the opening credits for *Blow-up* to be the embodiment of the filmmaker's ability to control what the spectator sees. By making the starlet's posture and the grass into a collage, the artistic intervention of the filmmaker is underlined.[124] Another typical example is

[120] Brunette, *The Films of Michelangelo Antonioni*, p. 125.
[121] Biarese and Tassone, *I film di Michelangelo Antonioni*, p. 138 (Note: Brunette translates Antonioni's statement quoted by Biarese and Tassone into English, requoting it in his work *The Films of Michelangelo Antonioni*, p. 111).
[122] Brunette, *The Films of Michelangelo Antonioni*, pp. 119–20.
[123] Brunette, *The Films of Michelangelo Antonioni*, p. 124.
[124] Bordwell, "The Art Cinema as a Mode of Film Practice," p. 98.

the final scene in the park (figure 2) where Thomas squats on the grass and then raises his head to see something above him. As Brunette argues:

> The camera, shooting down on him, shows him suddenly looking up at something. We cut to the familiar Antonioni shot of the tops of the moving trees, ostensibly from the photographer's POV. But then, the camera pans back down, in the same shot, back on *him*, now standing up, indicating that he was *not* the source of the look (since he cannot be looking at himself), not the motive of this particular shot, as we had thought. One could also add that this doubling of the look is the only real resemblance between the film and the Cortázar short story on which it is based, where a similar result is accomplished by an explicit and regular alternation between first- and third-person narration.[125]

Figure 2 The doubling of the look in *Blow-up*

In addition, as Seymour Chatman suggests, Antonioni sees Thomas' own disappearance at the very end as his "autograph."[126] This vanishing represents the post-production process of decreasing the opacity of the image. For Brunette, "we are also forcefully made aware of the presence of the director, of a controlling hand, an Other, outside the confines of the story and the film itself."[127]

In my subsequent discussion of *Blow-up*, I will regard the presence of filmmaking as one of the accompanying appearances experienced by the viewer and concentrate on the nexus of images and the viewer. The viewer does not only see images, but also sees the product of their own perception

[125] Brunette, *The Films of Michelangelo Antonioni*, p. 123 (Note: The words *him* and *not* in this quotation are Brunette's italic emphases).
[126] Seymour Chatman, *Antonioni: Or, the Surface of the World* (Berkeley, Los Angeles, and London: University of California Press, 1985), p. 145.
[127] Brunette, *The Films of Michelangelo Antonioni*, p. 118.

and inference coupled to images. A question needs to be considered: when seeing visual objects—the photographs arranged in the form of a cinematic montage, for example—how does the viewer know that this specific visual expression serves as a reminder of the presence of filmmaking? From a Buddhist perspective, as discussed previously, Berzin suggests that if the perceiver wants to know how phenomena arise, they can only do so by using their mind.[128] Likewise, I argue that if the viewer attempts to understand how these montage-oriented photographs exist as an authorial mark, they can only do so through an inferential construct of their mind. This issue is associated with Maitreya's account of *xiang biliang* 相比量 (appearance-inference). For Maitreya, "*weisui suoyou xiangzhuang xiangshu*. […] *You jianyan gu bizhi youhuo*." "谓随所有相状相属。【……】由见烟故比知有火。" (we infer what we experience from their appearances. […] We know fire due to our inferences from smoke).[129] If we link this Buddhist statement to the experience of the photographs, we find out that it is the viewer's mind giving them an inferential ability to use the knowledge of the appearance of a storyboard to associate the photographs with the mark of the presence of filmmaking. This means that the montage-oriented photographs do not possess a defining internal force empowering their existence as an authorial mark in the first place. Without the viewer's inference, these photographs disengage from an authorial mark. By being in a situation where there is no addition of inferential activities, the viewer, to some degree, can accept the photographs as the way they are. The relation between the photographs and an authorial mark can be thought of as a nominal construction in the viewer's mind. Depending on whether their mind engages in inferential activities, the state of these photographs moves between being accompanied by an authorial mark and not being accompanied by it. The variable operation of the viewer's mind changes the way the photographs appear. In other words, there is no inherently solid authorial mark rendering these photographs truly what they are.

Not only does the mind affect the state of the visible image, but also the invisible – or, in other words, the presence of the absent image. To illustrate, in the scene where Thomas participates in a mimed tennis match (figure 3), he picks up the imaginary ball and throws it back to the two players. It seems that the sound of the imaginary ball being played back and forth established its own absent visual presence. However, without the viewer's hearing, there is no basis for verifying its sound and therefore the idea of the presence of its absence falls apart. When sensing the presence of the ball's absence, the viewer also senses their own experience constructing a state of being present for this absent visual object. In order to know how this ball exists as the presence of absence, they can only

[128] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 20.

[129] Maitreya, *Yujia shidi lun* 瑜伽师地论 [The Treatise on the Foundation for Yoga Practitioners] (It was completed by Maitreya, circa 270–350, translated into Chinese in 648 as the *Yujia shidi lun* by Xuanzang, circa 602–664), ed. Li Li'an (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2005), vol. 15, pp. 209–10. My translation (Note: Within Western film theoretical traditions, fire-and-smoke has been dissected in terms of Charles Peirce's notion of index. However, index is not my concern. Although the fire-and-smoke example could trigger the discussion of indexicality, I do not discuss it).

do so in relation to an aurally perceptual construct of their mind. Yet, when their mind decides not to be involved with audio identification, there is no place for this invisible ball to be rendered as the presence of absence. Images, whether visible or invisible, are devoid of existing in an independent and inherent manner, initiated and extinguished by the changeable operation of the viewer's mind.

Figure 3 Thomas picking up the imaginary ball and throwing it back to the two players in *Blow-up*

2.4.2 Images and *Weishi sanliang* 唯识三量 (the Three Principal Means of Knowledge)

During film viewing, the viewer's mind is the key to the existential state of images, controlling the way images appear. In some ways, the viewer is creating what they see. In order to expand on this issue, I would like to examine a Buddhist system of cognition. In conjunction with Maitreya, Dignāga, and Jampa Gyatso's discussions of cognition, Chen Bing suggests that cognition engages with *weishi sanliang* 唯识三量 (the three principal means of knowledge): *xianliang* 现量 (perception), *biliang* 比量 (inference), and *feiliang* 非量 (specious perception and specious inference).[130] Perception has the following six aspects: *wugen xianliang* 五根现量 (five sense organs-perception), *yishi xianliang* 意识现量 (mental consciousness-perception), *zizheng xianliang* 自证现量 (self-authenticating-perception), *bashi xianliang* 八识现量 (the eighth consciousness-perception), *dingzhong xianliang* 定中现量 (yoga/Buddhist meditation-perception), and *zhen xianliang* 真现量 (true perception).[131] Among these aspects, I leave the discussion of the eighth consciousness-perception until chapter 4 and do not take yoga/Buddhist meditation-perception into consideration because it focuses on how to use Buddhism to address practical problems in daily life.

---

[130] Maitreya, *Yujia shidi lun*, pp. 209–11; Dignāga, *Jilianglun lüejie* 集量论略解 [The Compendium of Validities with Brief Explanation] (The text of the *Jilianglun* was completed by Dignāga, circa 480–circa 540, edited and translated into Chinese in 1980 as the *Jilianglun* by Fazun, 1902–1980. The *lüejie* [brief explanation] was completed by Fazun) (Beijing: Zhongguo shehui kexue chubanshe 北京：中国社会科学出版社 [Beijing: China Social Sciences Press], 1982), vol. 1, pp. 3–6 (Note: Since Arabic page numbers are used throughout this book, in order to avoid confusion, I would like to explain that pages 3–6 refer to pages 3–6 of *Juan yi: Xianliang pin* 卷一：现量品 [Volume 1: Perception Chapter]); Jampa Gyatso, *Yinmingxue qimeng* 因明学启蒙 [The Enlightenment of Buddhist Logico-epistemology] (It was completed by Jampa Gyatso, 1825–1901), trans. Yang Huaqun (Beijing: Zhongguo shehui kexueyuan shijie zongjiao yanjiusuo 北京：中国社会科学院世界宗教研究所 [Beijing: The Institute of World Religions at the Chinese Academy of Social Sciences], 1983), vol. 4, pp. 137–42; Chen Bing, *Fojiao xinlixue* 佛教心理学 [Buddhist Psychology] (Xi'an: Shaanxi shifan daxue chubanzongshe 西安：陕西师范大学出版总社 [Xi'an: Shaanxi Normal University General Publishing House], 2018), vol. 1, pp. 135–39. My translation.

[131] Chen, *Fojiao xinlixue*, p. 136.

Eyes and eye consciousness—a visual facet of five sense organs-perception—enable the image to be manifested to the viewer. However, the word "image" is used for ease of expression. Strictly speaking, in Buddhology, visual perception serves as the immediate or bare sensory impression of the visual object, not predicated on the applications of concepts, language, and knowledge.[132] For example, when the viewer engages in a moment where there is a direct and unmediated perceptual encounter with the park's image, they do not view it as the "park's image," as there is no conceptual recognition of the sense object intervening in this perception. Prior to recognising "this is the park's image," there must be an experiential moment where the viewer's perception engages directly with the "park's image." Without the interventions of thoughts and reflections, the viewer enters a non-discriminating state where nothing can be defined as an individual. The seeing of the "park's image" is merely the seeing of another object's image. By maintaining this state, all objects' images become one. Yet, when the viewer does not maintain their seeing, but rather uses their seeing as a basis for their inference and identification, these images can be distinguished and determined as individuals.

Mental consciousness-perception, also known as the sixth consciousness-perception, gives the observer an ability to know objects directly without putting much thought into their appearances.[133] Visual perception focuses on a non-conceptual state prior to knowing, whilst mental consciousness-perception takes knowing into account and treats it as a non-conceptual sensory form under certain conditions, even though in a general sense, knowing itself is closely linked to concepts and thoughts. For example, after noticing Thomas for the first time, the viewer can instantly identify this character in the following scenes without much thinking about his visual appearance. Once the viewer is quite familiar with what they recognise and accepts it as normal, conceptual recognition will be converted into intuition. The recognition of Thomas' appearance can be internalised and habitually performed in intuitive perception. Considering that Thomas' appearance undergoes a change from the product of conceptual recognition to intuition, there is no inherently findable visual appearance establishing its solid entity from its own side. The viewer sees the figure, but nothing out there can be truly seen.

As discussed in chapter 1, the self-authenticating aspect means that the perceiver verifies their own activities of experiencing objects.[134] With the aid of self-authenticating-perception, the viewer can confirm that they are capable of viewing the image. Regarding the opening credits for *Blow-up*, though Bordwell sees them as the mark of the filmmaker's ability to control what the viewer sees,[135] I consider that the viewer's self-authenticating-perception makes them notice that they are able to use their seeing to determine whether the visual objects in the opening credits can be manifested.

[132] Chen, *Fojiao xinlixue*, p. 136.
[133] Chen, *Fojiao xinlixue*, p. 136.
[134] See pp. 18–19.
[135] See p. 35.

True perception involves gaining a deep insight into the essence of reality itself.[136] This idea is also called *wuxin* 无心 ("no mind"), referring to *wupanyuanxin* 无攀缘心 (the non-clinging mind).[137] True perception does not mean that typical or standard perception can grant credibility to the seen. Rather, it motivates the observer to perceive the thusness or suchness of objects. Considering that both the involved objects of typical and distorted perceptions are determined by the viewer's mind, neither of them establishes its independent entity from its own side. Hence, the perceiver sees the thusness or suchness of objects as nothing but interrelationship itself. Likewise, the image becomes not the image of the park, but of interrelationship. There is no such thing as the "park's image" to which the viewer can cling. Instead, their seeing and the park's image are one simultaneous arising.

Inference embraces seven facets: *xiang biliang* 相比量 (appearance-inference), *ti biliang* 体比量 (essence-inference), *ye biliang* 业比量 (karma-inference), *fa biliang* 法比量 (dharma-inference), *yinguo biliang* 因果比量 (causality-inference), *shixu biliang* 世许比量 (accepted phenomena-inference), and *xinjie biliang* 信解比量 (religious belief-inference).[138] I put religious belief-inference aside, for it is primarily concerned with the practical issues of good and evil in the sense of Buddhism.

As discussed previously, in the context of appearance-inference, there is no inherently findable authorial mark rendering the montage-oriented photographs truly what they are, since the viewer's mind is able to decide whether they want to utilise the knowledge of a storyboard's appearance to make a hypothetical connection between the photographs and an authorial mark. Not governed by an authorial mark, the existential state of these photographs is affected by the viewer's decision.[139]

Essence-inference refers to making deductions about the object's past or future essence based on its present essence.[140] Not only does this term engage with appearances, but also involves time. To illustrate, when the viewer sees a blurry blow-up of a gunman hiding in the bushes, they tend to infer that a clear gunman's image will be disclosed in the later portions of the film, despite the fact that his visual appearance is never clearly displayed. I argue that the gunman's image does not exist not because the film never shows his distinct face. Even though his image were explicitly expressed afterwards, there would be no gunman's image-in-itself lurking behind his image, because once the viewer verified his visual appearance, it would already be affected by their mind and thus lacking in self-essence. Neither the blurry image appearing in the present nor the sharp image assumed to be occurring in the future has access to the image-in-itself. Yet, at a relative level, the viewer still needs to hypothesise the idea of the gunman's image-in-itself in order to know and explain what they see.

---

[136] Chen, *Fojiao xinlixue*, p. 136.
[137] Chen, *Fojiao xinlixue*, vol. 2, p. 608. My translation.
[138] Chen, *Fojiao xinlixue*, p. 138.
[139] See p. 37.
[140] Chen, *Fojiao xinlixue*, p. 138.

Karma-inference is not just linked to appearances, but also concerned with drawing inferences from actions.[141] A salient attribute of *Blow-up* is distracting actions. When seeing Thomas having an orgy with two girls (played by Jane Birkin and Gillian Hills), the viewer assumes that the orgy's image will be interrupted by another action's image since they have been familiar with multiple distracting actions before the orgy's image is presented. For example, Thomas is prevented from having lunch, as he notices a man (played by Dyson Lovell) spying on him from outside the restaurant. Likewise, Thomas' telephone conversation with Ron (played by Peter Bowles) is interrupted by the arrival of two girls. These distracting actions initiate a distracting field in which the viewer has much less hope of seeing uninterrupted things, motivating them to infer that the orgy's image is a part of distracting actions. This inference is based on a set of loosely connected actions performed by the characters. However, if the viewer shifts the focus of their attention from the characters' actions to the action of seeing, the distracting entities of images will be extinguished. From the point of view of the action of seeing, there is no disjointed state among the orgy and other actions since they are all manifested as a visual form directly encountered by the viewer's perception. For eye consciousness, the seeing of the orgy is just the seeing of another action's image. Therefore, there is no place for a distracting entity to appear. Depending on what kind of action the viewer concentrates on, the existential state of the orgy's image can be modified, engaging with but also disengaging from other actions' images.

Dharma-inference drives the observer to use a certain phenomenon to extrapolate its adjacent phenomenon. In Buddhology, a widely used paradigm of dharma-inference is deducing death from life.[142] I suggest that within the context of images, dharma-inference can be used to examine images in adjacent shots. For example, in respect of the final park scene where Thomas squats on the grass, when the viewer notices in the prior shot Thomas raising his head to see something above him, due to the impact of an understanding of continuity editing, the viewer infers that in the next shot, they will see objects from Thomas' point of view. Yet, I argue that there are no such things as "images in an adjacent shot" about which the viewer can make deductions. Although this idea seems to oppose dharma-inference, there is a reason behind this contention. To discuss this issue, I would like to first mention a life-and-death situation. As Yongming Yanshou argues, "*ji sheng ji si*." "即生即死。" (life is death).[143] This idea signifies that as life begins, so does death. Thus, life is death. It seems that life and death were two distinct states next to each other. However, they are one. In fact, this idea does not only refer to a living being's life cycle at a practical level, but also the start and end of the process

[141] Chen, *Fojiao xinlixue*, p. 138.
[142] Chen, *Fojiao xinlixue*, p. 138.
[143] Yongming Yanshou, *Zongjing lu* 宗镜录 [The Records of the Source Mirror] (961) (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2015), vol. 73, p. 1411. My translation (Note: Even if Shi Yanshou is the name on the front cover of the book, "Shi" is a unified title of Buddhists. His name is Yongming Yanshou).

of any phenomenon at a theoretical level. Likewise, regarding the final park scene, we can treat the playing of images in each shot as a lifetime. As the playing of images in the prior shot begins (I would like to call this process x_1), so does its end (x_2). Its end (x_2) indicates the start of the playing of images in the next shot (y_1). In this light, we are aware that x_1 is x_2 and x_2 is y_1, therefore x_1 is y_1. This equality divulges that there are no images in a shot carrying independent and inherent existence to establish themselves as fundamentally different from the ones in another shot. In some ways, a state of being adjacent—a state where two seemingly different processes, x_1 and y_1, are next to each other—does not exist because they have already become one. This suggests that deductions about images in an adjacent shot also fall apart because there is no such thing as a "state of being adjacent" existing in the first place.

Causality-inference allows the experiencer to infer an effect from a cause.[144] Because of seeing Thomas getting into and starting his car, the viewer deduces that driving will arise. This inference is based on the viewer's common sense that the function of a car is to transport people and their stuff from place to place. Apart from common sense, in Buddhology, there is a paradoxical issue in terms of cause and effect. Nāgārjuna argues that "*yinguo ruo yi zhe*, *shishi yi buran*. […] *Ruo yinguo shi yi*, *yin ze tong feiyin*." "因果若异者，是事亦不然。【……】若因果是异，因则同非因。" (if cause and effect were different, everything would be invalid. […] If cause and effect were different, cause would be the same as non-cause).[145] Non-cause refers to no specific cause. It can be understood as any cause. To illustrate, it appears as if driving (effect) occurs only after Thomas starts his car (cause), which means that the cause itself does not contain the effect prior to the manifestation of the effect. However, if driving does not exist prior to its manifestation, that is, if driving is not contained within Thomas' act of starting his car, then it can be produced by any cause—someone unknown starting this car, for instance—considering that driving itself is not associated with anything. Yet, this specific driving is produced by a specific cause (Thomas' act of starting this car), rather than by any arbitrary person. Therefore, driving exists prior to its manifestation, contained within Thomas' act of starting his car. Driving is an implicit state of starting, with the flow of time transforming starting into driving. In some sense, starting and driving are one. Cause-effect images are organised through non-duality. Yet, Nāgārjuna also suggests that "*yinguo shi yi zhe*, *shishi zhong buran*. […] *Ruo yinguo shi yi*, *sheng ji suosheng yi*." "因果是一者，是事终不然。【……】若因果是一，生及所生一。" (if cause and effect were one, everything would be invalid. […] If cause and effect were one, the producing and

[144] Chen, *Fojiao xinlixue*, p. 138.
[145] Nāgārjuna, *Zhong lun* 中论 [The Root Verses on the Middle Way] (It was completed by Nāgārjuna circa the third century. The explanation was completed by Vimalākṣa circa the fourth century, translated into Chinese in 409 as the *Zhong lun* by Kumārajīva, 344–413), ed. Yan Yongkui (Chengdu: Zhongguo chuantong wenhua yanjiusuo 成都：中国传统文化研究所 [Chengdu: The Institute of Chinese Traditional Culture], 1995), p. 148. My translation.

produced objects would be homogeneous).[146] Based on this idea, I suggest that if driving exists prior to its manifestation, then whatever can be done by driving can also be done by starting. Driving can make the car move, which indicates that Thomas' act of starting the car can also make it move. Yet, this does not happen because at the moment of starting the car, its wheels have not yet turned and hence it is still in the same place. Consequently, driving does not exist prior to its manifestation, not contained within Thomas' act of starting his car. In this light, we notice that cause-effect images are constituted by relative duality. Starting and driving are not one. When the viewer's inference moves between the non-duality and relative duality of cause and effect, nominal cause-effect images arise.

Accepted phenomena-inference deals with the act of making deductions based on phenomena accepted by the public.[147] For instance, when seeing the park's image, as most viewers acknowledge that it is full of green, an individual can further confirm their own understanding of the knowledge of a green appearance and judge that the park's image is full of green. However, according to Tian Maozhi's discussion of the Buddhist dependent-origination world view, different norms established for the type of experience transform the perceived object into multiple appearances. A colour-blind individual's experience of colours is distinct from the one with typical vision.[148] Likewise, even if the park's image is widely recognised as green, for the viewer with deuteranopia, it becomes yellow. If this green appearance existed in a truly inherent way, then it should maintain the same appearance all the time. Thus, there is no inherently findable green appearance establishing its fixed entity from its own side. Despite the fact that the viewer with typical vision is capable of recognising green, "no green" still serves as a hidden flow on the bottom of green, deconstructing green in an implicit way.

In addition to my Buddhist examinations of how perception and inference engage with images, it is also significant to put the ontological research on images into dialogue with the third means of knowledge called specious perception and specious inference. This means consists of five branches: *yijuezhi* 已决智 (subsequent cognition), *sichashi* 伺察识 (presumption), *xian er weiding* 现而未定 (non-determining cognition), *yi* 疑 (indecisive wavering), and *xiezhi* 邪智 (distorted cognition).[149]

In order to discuss subsequent cognition, we need to take into account another Buddhist term called *xinzhi* 新知 (fresh cognition).[150] The first microsecond of cognition is fresh, as it synchronises

---

[146] Nāgārjuna, *Zhong lun*, p. 148.
[147] Chen, *Fojiao xinlixue*, p. 138.
[148] This Buddhist discussion comes from Tian Maozhi's preface to the *Jingang jing* 金刚经 [The Diamond Sūtra]. Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007), p. 7. My translation.
[149] Chen, *Fojiao xinlixue*, pp. 138–39.
[150] Dignāga, *Jilianglun lüejie*, p. 3.

the act of seeing with the object seen and hence engages with the here-and-now—in the sense of valid cognition. This initial moment is followed by a sequence of subsequent cognition during which the experience of the object is no longer fresh because subsequent cognition is just a reflection on the existential state of the object previously experienced, which does not reflect the here-and-now. Subsequent cognition uses the first moment of cognition as a condition for its own intelligibility and thus is devoid of its own freshness. Moreover, subsequent cognition entails conceptual recognition, which suggests that the object seen has already been affected by subjective judgement. In this light, subsequent cognition no longer corresponds to fresh and primordial cognition, regarded as invalid. Take the man outside the restaurant, for example. At the first instant of cognition, the viewer does not see this character as a man, but rather as a certain figure. At this initial stage, the viewer's act of seeing and the figure seen synchronise. The cognition of the figure is in the process of here-and-now, being fresh and valid. Subsequently, the viewer reflects on this primordial figure, recognising that "this is a man." Then, the viewer discerns his visual details such as his facial feature, hair colour, and suit and examines his furtive action. Based on this visual information, the viewer starts guessing his identity. He could be the gunman, which motivates the viewer to associate him with Jane (played by Vanessa Redgrave), the woman in the park, because there is a possibility that the gunman might collude with Jane. It seems that the viewer's cognition is getting closer to the understanding of the man's image, but on the other hand, it is moving further away from the primordial figure sensed in the first microsecond of cognition. The subsequent cognition of the figure is no longer fresh, treated as invalid. In this context, invalidity does not mean that the viewer's cognition is wrong. It is obvious that this figure is a man. However, due to the impacts of reflections and concepts, the viewer usually considers that the object has a defining internal force empowering its existence as the man's image in the first place and rendering it truly the man's image existing independently of the viewer's mind. In some sense, the man's image is a conceptual product resembling the primordial figure previously experienced. All of the visual details of the man's image depend on the initial cognition of the figure. Without the first microsecond of cognition of the figure engaging with the here-and-now, the man's image cannot be crystallised. This does not suggest that the primordial figure is the embodiment of the ultimate level of cinematographic reality. Considering that the primordial figure has acquired a different form (the man's image), its ontological essence has already been in a changeable situation.

Presumption suggests that the perceiver does arrive at a tenable conclusion, but without really knowing why it is tenable.[151] To illustrate, as for the park's image, when the viewer just undoubtedly accepts the non-existence of a green appearance without conducting an analysis of it to understand why it has no fixed entity, their cognition is deceived. Such a deceptive quality causes the viewer to blindly treat green and other appearances as entire nothingness. By carrying out an examination of

[151] Chen, *Fojiao xinlixue*, p. 138.

the non-existence of a green appearance, the viewer realises that non-existence is a part of nominal nature. Learning how "no green" functions persuades the viewer to accept green's relative entity.[152]

Non-determining cognition means that the experiencer engages directly with objects, but fails to gain a clear understanding of them.[153] It is noteworthy that there is a subtle distinction between visual perception and non-determining cognition. For Lobsang Choephel, though visual perception also immediately encounters objects without the intervention of conceptual recognition, it prompts the perceiver to concentrate on the process of seeing objects. Even if the perceiver does not employ concepts to recognise objects, they will not forget what they have seen. However, non-determining cognition suggests that when a certain object appears to the perceiver, due to the situation where they are absorbed in another object, they cannot focus on the process of seeing it even though they see it. Therefore, after a few days, the perceiver will not be able to recall seeing this object.[154] Take the Yardbirds rock music club, for example. Although the viewer sees a crowd of people scrambling for a broken piece of a guitar, what the viewer cares about is whether the protagonist Thomas can grab that broken piece. Even if other characters also appear on the screen, the viewer does not pay particular attention to their visual appearances. Hence, the viewer cannot get a clear understanding of what they look like and will not be able to remember seeing their appearances. Other characters' visual features are in an implicit state until the viewer is preoccupied with them. The viewer's mind is capable of altering the way the crowd of people appear, making their images implicit and explicit. In other words, there are no inherently solid characters established as existing from their own side.

Indecisive wavering signifies that the perceiver dithers between two conclusions concerning a certain phenomenon.[155] For instance, as dissected earlier, there are two causality-based deductions about driving. First, driving exists prior to its manifestation, contained within Thomas' act of starting the car. Otherwise, driving can be produced by any arbitrary person because of not being associated with anything. This specific driving and its specific cause (Thomas' act of starting the car) are merged into one. Second, driving does not exist prior to its manifestation, not contained within Thomas' act of starting the car. Otherwise, whatever can be done by driving can also be done by starting due to their oneness, in which case this does not happen.[156] This raises an issue: which statement is viable? When the viewer vacillates between these two judgements, their cognition is invalid since it fails to establish decisiveness about the causal ontology of driving as a visual phenomenon. This wavering

---

[152] See p. 43.
[153] Chen, *Fojiao xinlixue*, p. 139.
[154] Lobsang Choephel and Zilkar Rinpoche, *Xinlei xue* 心类学 [The Types of Awareness] (It is from the lectures delivered by Choephel at the Atisha Buddhist Centre in Taichung, Taiwan, translated into Chinese as the *Xinlei xue* by Faju), ed. Jiandi (https://www.lamrimworld.org/treasure_view.php?news_view_sn=69&sn=7, 2006), pp. 94–95, 102–3. My translation.
[155] Chen, *Fojiao xinlixue*, p. 139.
[156] See pp. 42–43.

can be inclined towards one of these two judgements or evenly balanced between the two. I would like to clarify that this wavering is not the same as the cognition of nominal cause-effect images, as the latter is based on the viewer's decisive awareness of the Buddhist dependent-origination world view. Driving's oneness with starting arises from the viewer's conceptual reflection on a one-to-one correspondence between cause and effect, whilst driving's distinction from starting originates from the viewer's attention to different functions of cause and effect. When the viewer realises that it is the conversion process between these two judgements making each other sensible, in other words, when the viewer notices that these two conclusions together constitute their holistic understanding of why driving as a visual phenomenon is the way they experience it, they no longer get involved in wavering between them. Instead, they are decisively aware of nominal cause-effect images in which the viewer neither clings to their oneness nor to their distinction, neither being detached from their oneness nor from their distinction. There is no further indecisiveness interrupting awareness of this nominal situation, with such awareness arising simultaneously with each moment of image viewing.

Distorted cognition signifies that the observer's cognition is deceived due to apprehending the object as existing independently of their experience or because of their own sensory impairment.[157] For example, the grass in the film seems to be a pregiven visual object because it exists before being released. However, if the viewer wants to know how the manifestation of the grass works, they can only do so through the constructs of perception and inference. In this sense, the viewer's mind and the grass are one simultaneous arising. By raising awareness of this simultaneous arising, the viewer is able to engage with the non-self of the grass, rather than clings to it as the pregiven. Furthermore, as examined earlier, when the viewer's cognition is affected by a visual impairment—deuteranopia, for instance—green is distorted into yellow; the perceived yellow is deceptive, as it does not fit with the colour of the grass of typical perception. However, this does not mean that green is the ultimate existence of the colour of the grass. Based on the idea of true perception discussed previously, both green and yellow are determined by the viewer's eyes and eye consciousness and thus do not carry independent entities. In terms of true perception, the word "true" should not be misunderstood as typical or standard, but rather as interrelated. This suggests that both clinging to green (typical) and clinging to yellow (impaired or distorted) are regarded as distorted because of the ignorance of the interrelationship between the viewer's mind and the grass. Their mind causes the experienced grass to be in a state of flux. What is considered typical can be viewed differently. My Buddhist discussion of the grass does not only view it as a visual object devoid of existing in an independent and inherent way, but also as a figurative construct of an empty state of image viewing which is connected to the existential state of all visual objects. In the next sub-section, instead of looking at the tranquillity of the grass, I will utilise the grass as metaphorical material to investigate the image-viewing reduction.

[157] Chen, *Fojiao xinlixue*, p. 139.

2.4.3 The Image-viewing Reduction

At the very end of the film, Thomas disappears (figure 4). I call this phenomenon "representational vanishing," as it implies the post-production process of decreasing the opacity of Thomas. However, instead of focussing on representational vanishing, I would like to examine the ontological vanishing of images at the Buddhist level of the mind. In order to engage with this issue in a figurative manner, I regard the empty grass as an empty state of image viewing. When the viewer's mind is reduced to an empty state, that is, when the viewer's mind does not cling to the apprehension of independent and inherent existence, all perceived visual objects fall apart, which makes manifest the sphere of the void. It seems that Thomas, Jane, the gunman, a tree, a car, and a broken piece of a guitar had the appearance of being separate entities. However, all of them are compound appearances devoid of existing in an independent and inherent way. Once being cut off from the viewer's act of seeing, their visual appearances will not be able to manifest themselves to the viewer. Although the empty field of the viewer's mind itself does not express visual objects, the viewer's act of seeing causes an infinite variety of visual objects to be placed into this empty field. It can be understood as a field of potential from which every object seen arises.

Figure 4 The grass in *Blow-up*

In order to further analyse the above ontological issue, I use the grass as metaphorical material to present a dynamic Buddhist theoretical model of image viewing (figure 5). This model pictures a nominal state of image viewing as a turbulent sensory realm. When the viewer's mind moves in this sensory ocean, that is, when the sensory waves rise and fall, the peaks are from which visual objects arise. The mind usually pays attention to the peaks since they are more prominent than the rest of the sensory ocean. The rest suggests that visual objects have no independent and inherent nature—an implicit facet that the mind ignores. Although the rest is out of range of tangible cognition, it still serves as the empty potential of visual objects. To put it another way, every visual object arises and returns from this turbulent sensory field. Just as waves come from the ocean, rise, and go back into the ocean, visual objects come from the sensory realm, retain their relative manifestations for some time, and change into the void. Both their initiation and extinction are interrelated attributes of the sensory realm, with the conversion process between these two facets giving rise to nominal images.

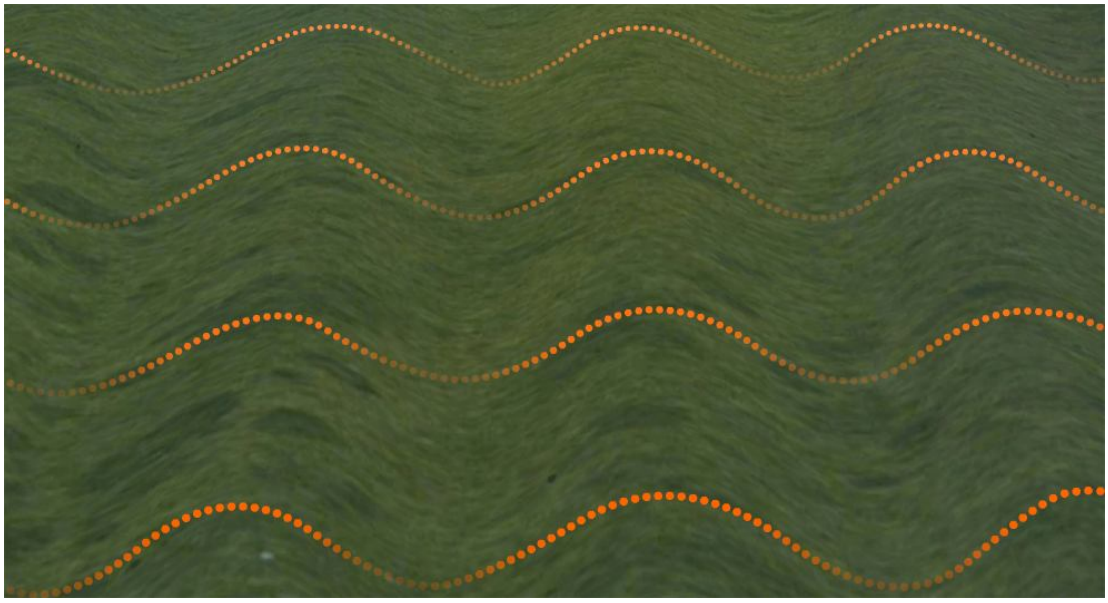

Figure 5 A dynamic Buddhist theoretical model of image viewing

When the mind is moving, the viewer understands images by virtue of the concepts of images, which tends to increase their clinging to images seen and drives them to regard images as the visible and knowable and thus to determine the entities of images. However, when the mind is maintained in a settled state, that is, when the viewer quiets the mind of the conceptual formulations of images, they gain access to a non-conceptual process of sensory consciousness. In this experiential process, the viewer focuses on the mere arising and engaging with what they see occurring at each moment without judging that what they experience are images, considering that there are no such things as independent and inherent images. Both the moving and settled minds are the interwoven fabric of image viewing. Being in a flowing state, images and “no images” are able to convert into each other. At the level of images, the perceived visual objects including props, characters, and scenery appear as different entities. At the level of “no images,” there are no fundamental distinctions among these objects. Considering that all of their manifestations depend on the viewer’s act of seeing them, they are devoid of existing in an independent and inherent way. Hence, every object seen can be reduced to the void. Building on these two aspects, we realise that a Buddhist ontological account of nominal images is not limited to the specific styles of images. Instead, it can be applied to any types of images.

Chapter 3 Nominal Montage in Experimental Cinema

3.1 Nominal Montage

In the second chapter, I discussed realism in art cinema and demonstrated why there is no essential partition between art cinema and Hollywood cinema at a Buddhist ontological level. Although David Bordwell has also argued that the boundary between them is relative, his approach is based on the representational transformation for these two types of cinema. This complex interchange suggests that because art cinema and Hollywood cinema can adopt the expressive forms of each other, there is no absolute boundary between them.[158] However, my Buddhist ontological view on their oneness is based on the idea that both are devoid of existing in an independent and inherent way since once they are separated from audiences' visual sensations and conceptual formulations, they cannot be manifested. Therefore, even if they did not adopt the expressive forms of each other, their common insubstantial existence would still bring them together.

In addition to putting realism in art cinema into dialogue with Hollywood cinema, realism also engages with formalist montage within the context of experimental or avant-garde cinema. Before analysing my argument in this chapter, I would like to first explain how I use the term "montage" in my subsequent discussion and provide a broad picture of why the long take—an aspect of realism—and montage are one within Buddhology. Montage is a specific historical term related to a particular type of editing—Sergei Eisenstein and Dziga Vertov's Soviet montage, for example. However, in this chapter, I use montage as an aesthetic and mind-related element, rather than recount its historical context. As for realism, I focus on its long-take aspect, not daily-life aspect, associating its continuity with montage's fragmentation in a Buddhist context of the mind.

My argument in this chapter is that there is no fundamental division between montage and long-take realism (abbreviated to the "long take") because at a mind level, the long take carries the potential of a cut. Although in film studies, Peter Wollen also dissects the relation between montage and realism, his view on their unity is distinct from mine since his account of realism deals with daily life. The next sub-section will outline how Wollen examines their oneness in an avant-garde context.

3.1.1 An Orthodox Account of the Oneness of Montage and Realism

In Wollen's 1975 article "The Two Avant-gardes," he chooses Soviet cinema as a prime example to interpret the interconnection between montage and realism, maintaining that "the Soviet directors

[158] David Bordwell, "The Art Cinema as a Mode of Film Practice" (1979), in *The European Cinema Reader*, ed. Catherine Fowler (London and New York: Routledge, 2002), p. 100.

of the twenties, though they saw themselves in some sense as avant-garde, were also preoccupied with the problem of realism."[159] These Soviet directors include Eisenstein and Vertov. Even though Eisenstein's films are somewhat distinct from Vertov's, both use montage as formalist intervention in the recorded flow of daily reality. Montage is an alternative system of editing, serving as a conflict (a dialectical collision between shots) from which new concepts arise. Eisenstein's *Strike* (1925) and Vertov's *Man with a Movie Camera* (1929) display a conflict between shots and social reality-related footage, producing powerful emotional effects and generating multiple intellectual meanings.[160]

Wollen's viewpoint on avant-gardes can be complemented by Laura Marks' phenomenological approach to experimental cinema. Marks stresses that Vertov as one of the influential experimental filmmakers values embodied spectatorship—the relation between perception and embodiment.[161] As Marks suggests, in early and experimental cinema, instead of promoting the illusion that permits distanced identification with action in deep space, it features a haptic quality evoking an immediate bodily response to the screen.[162]

Wollen's examination of the nexus of montage and realism depends on the premise that a film must feature a stylistic interpenetration between montage and the representation of social reality. This is similar to Bordwell's approach to the representational oneness of art cinema and Hollywood cinema. Despite the fact that Wollen and Bordwell's research foci are not identical, both their ideas suggest that the interweaving of different forms of expression presented in the same film serves as a basis for unifying multiple styles. I consider that this interweaving is not the only approach to the oneness of montage and realism. Wollen's analysis emphasises their visibly on-screen interweaving, whilst in the following sub-section, I will use a Buddhist ontological view to discuss how the viewer's mind reconfigures and unifies the existential qualities of montage and realism in an implicit manner.

3.1.2 The Buddhist Ontological Oneness of Montage and Realism

Instead of focussing on a synthesis of montage and daily-life realism, I examine the interconnection between montage and long-take realism. Readers may ask: "Is their ontological oneness attributed to the fact that the components of the long take are still photographs which have cut the long take into multiple static pieces—in the sense of the fragmentary equality between montage and the long take?" Although this mechanical function-related idea can serve as an interpretation of the oneness

[159] Peter Wollen, "The Two Avant-gardes" (1975), in *Studio International: Avant-garde Film in England and Europe*, eds. Richard Cork, David Curtis, Peter Gidal, Malcolm Le Grice, and Peter Wollen (London: Studio International Journal Ltd, 1975), vol. 190, no. 978, pp. 171–75.
[160] Wollen, "The Two Avant-gardes," pp. 171–75.
[161] Laura Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses* (Durham and London: Duke University Press, 2000), pp. 171, 215.
[162] Marks, *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*, pp. 170–71.

of montage and the long take, my approach to their oneness is not based on filmic technology, but rather based on the effect of the viewer's mind. I do not describe the visibly on-screen interweaving of montage and the long take. Rather, I consider that even if an experimental film like *Lemon* (Hollis Frampton, 1969) presents a lemon in a single long take and even if the long take is understood as a continuity of movement, rather than a succession of separate images, the long take itself still carries the potential of a cut. I call this latent fragmentary state "mind montage" since it is produced by an inferential construct of the viewer's mind. In order to expand on this issue at a Buddhist ontological level, I utilise *yinguo biliang* 因果比量 (causality-inference) to provide the first clue as to how mind montage within the long take operates during film viewing. In terms of this Buddhist idea, Maitreya points out that "*weiyi yinguo zhanzhuan xiangbi*." "谓以因果展转相比。" (we infer an effect from a cause, and vice versa).[163] It is significant to note that causality-inference emphasises the variability of interrelationship. That is, if a cause changes, its effect changes, and vice versa. The reason why I bring up causality-inference is because the variability of causality can be applied to the changeable cognition of fragmentation, which makes the perceived fragmentation transcend tangible montage and prompts the viewer to notice those implicit discontinuous traces within the perceived long take and to re-examine the question of whether both montage and the long take possess fragmentation.

For example, when the viewer perceives a lemon presented in a single long take, not only their seeing but also their thinking engages with it. In this shot, light and shadow constantly move around the surface of the lemon. During this entire process, they see a continuity of movement—a common cause giving rise to its contrast with montage's fragmentation, but simultaneously, they experience a fragmentary potentiality of the meaning of this long take. It does not carry a single meaning lasting forever because it varies with its context, namely the shapes of light and shadow. As Maureen Turim argues, the changing chiaroscuro lighting makes the shot of the lemon bear disjointed fragments of meaning, such as the male/female (sexual) and day/night (cosmic) metaphors.[164] By putting Turim's view into dialogue with causality, I suggest that even if there are no preceding and succeeding shots interrupting the continuous state of this single long take, the viewer's variable thought on this shot has cut its meaning into separate pieces and made them into a mind collage—a cause producing its fragmentary equality with montage. A turning point between a meaning and another one serves as one facet of mind montage within the long take. It marks the viewer's symbolic thinking during film viewing. Both montage and the long take possess fragmentary traces.

---

[163] Maitreya, *Yujia shidi lun* 瑜伽师地论 [The Treatise on the Foundation for Yoga Practitioners] (It was completed by Maitreya, circa 270–350, translated into Chinese in 648 as the *Yujia shidi lun* by Xuanzang, circa 602–664), ed. Li Li'an (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2005), vol. 15, pp. 210–11. My translation.

[164] Maureen Turim, *Abstraction in Avant-garde Films* (Ann Arbor: UMI Research Press, 1985), p. 25.

From the perspective of causality-inference, the viewer's reflection on disjointed fragments of the long take's meaning causes the cognition of the non-distinction (fragmentary equality) between montage and the long take. This does not mean to cling to their oneness as an ultimately inferential result. In Buddhology, the idea of causality-inference calls attention to the changeable operation of *yuan* 缘 (condition) which gives rise to the variability of cause and effect. In terms of this ontological issue, Bodhiruci states that "*yinyuan huiyu shi*, *guobao huanzi shou*." "因缘会遇时，果报还自受。" (when a cause and a condition come together, one accepts an effect).[165] This Buddhist idea suggests that if a cause can produce an effect, there must be a condition making this cause catalyse an effect. When a condition changes, aspects of deduction dependent on this condition also change. Likewise, the non-distinction between montage and the long take is not the only inferential result. When the condition alters, that is, when the viewer shifts the focus of their attention from symbolic cognition to the cognition of retinal images, the perception of a continuity of movement within the long take becomes a dominant cause which supersedes a reflection on disjointed fragments of the long take's meaning. Under this particular condition, the viewer recognises that its continuous state is different from montage's fragmentation and therefore their oneness is not initiated, even though disjointed fragments of meaning still act as potentialities on the bottom of the long take seen. As the condition can change between the intellectual attention to retinal images and symbols, neither the distinction nor non-distinction between montage and the long take is a fixed inferential result. A nominal realm between distinction and non-distinction suggests that a continuity of movement and mind montage within the long take interrelate. They are reflected in and enter into each other simultaneously. The manner in which they interpenetrate signifies that fragmentation does but also does not just appear at cut points between shots. Owing to a threshold between retinal images and symbols, one accepts fragmentation occurring at cut points between shots as embracing fragmentation happening at any point in time.

Although Frampton's *Lemon* is an example for the discussion of mind montage within the long take, I want to choose his another film *Zorns Lemma* (1970) as my primary case study in this chapter. Considering that the latter presents both montage and the long take, it can be utilised as a paradigm for reflections not only on mind montage within the long take, but also on the mind long take within montage. Before dissecting *Zorns Lemma*, I would like to first revisit a wider context of experimental cinema. The next section will demonstrate that all of the styles of experimental cinema engage with nominality at a macroscopic level. Within a Buddhist context, these avant-garde styles become one.

[165] Bodhiruci, *Dabaoji jing* 大宝积经 [The Mahāratnakūṭa Sūtra] (It was translated into Chinese circa the sixth century as the *Dabaoji jing* by Bodhiruci, floruit 508–535), ed. Bodhiruci (Shanghai: Shanghai foxue shuju 上海：上海佛学书局 [Shanghai: Shanghai Buddhist Books], 2004), vol. 57, p. 450. My translation.

3.2 A Buddhist View on Montage in Experimental Cinema

Before examining experimental or avant-garde styles, I want to first mention A. L. Rees' theoretical work *A History of Experimental Film and Video* in which he uses Clement Greenberg's interpretation of the meaning of avant-garde as an opening preface statement: "You don't define it; you recognise it as a historical phenomenon."[166] The undefined nature of avant-garde should not be misconstrued as the total nothingness of avant-garde. Instead, such unclear essence opens up infinite possibilities for studying avant-garde, thereby not restricting discussions of avant-garde to a one-point and fixed perspective. Rees adopts a historical approach to analysing experimental cinema, whilst some other theorists employ a formalist/stylistic perspective. In *Film: A Critical Introduction*, Maria Pramaggiore and Tom Wallis divide experimental cinema into five specific segments: surrealist film, abstract film, the city symphony film, structural film, and the compilation film.[167] All of them involve the frequent overthrow of continuity editing, challenging traditional methods for organising images, shots, and sounds. I assert that even though the specific functions of these experimental styles can be seen as different from each other, there is still a nominal thread linking them at a Buddhist ontological level.

With regard to surrealist film, Pramaggiore and Wallis suggest that it is generally expressed by irrational, unconscious, and absurd imagery. Fragmentary dream-like shots are edited together to achieve incoherence. Chronological time and conventional modes of narrative logic are weakened. Surrealist film is full of incongruous images that abruptly change in time and space.[168] Pramaggiore and Wallis choose some renowned films as case studies such as *Un Chien Andalou* (Luis Buñuel, 1929) and *Meshes of the Afternoon* (Maya Deren and Alexander Hammid, 1943) to illustrate the above characteristics of surrealist film.[169] However, Michael Richardson considers that a particular style of imagery cannot be used to define surrealist film itself because a style can only be utilised to delineate the representation of surrealist film, rather than to engage with its essence. As Richardson argues:

> Misunderstandings [of surrealist film] are founded in the fact that [critics] seek to reduce surrealism to a style or a thing in itself rather than being prepared to see it as an activity with broadening horizons. Many critics fail to recognise the distinctive qualities that make up the surrealist attitude. They seek something – a theme, a particular type of imagery, certain concepts

---

[166] Edward Lucie-Smith and Clement Greenberg, "Interview Conducted by Edward Lucie-Smith" (1968), in *Clement Greenberg: The Collected Essays and Criticism, Volume 4: Modernism with a Vengeance, 1957–1969*, ed. John O'Brian (Chicago and London: University of Chicago Press, 1993), p. 281 (Note: A. L. Rees quotes this Greenberg's statement in his work *A History of Experimental Film and Video* [London: Palgrave Macmillan, 2011], p. 11).

[167] Maria Pramaggiore and Tom Wallis, *Film: A Critical Introduction* (London: Laurence King Publishing Ltd, 2011), p. 292.

[168] Pramaggiore and Wallis, *Film: A Critical Introduction*, p. 293.

[169] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 293–94, 303–5.

> – they can identify as "surrealist" in order to provide a criterion of judgement by which a film or art work can be appraised. The problem is that this goes against the very essence of surrealism, which refuses to be *here* but is always *elsewhere*. It is not a thing but a relation between things and therefore needs to be treated as a whole.[170]

This statement gives weight to the difference between a surrealist style and a relation between things formed by an activity with broadening horizons, rather than just describes a certain surrealist style as Pramaggiore and Wallis do. Richardson's view on a relation seems parallel to Buddhabhadra and Śikṣānanda's idea of the interdependence of phenomena. This Buddhist idea suggests that "one depends on all and all depend on one, whether immediately or remotely; therefore, the existence of all is considered an intrinsic part of the existence of one, and vice versa."[171] Everything originates from a potentially interconnected source. In this sense, likewise, Richardson's analysis of a relation between things corresponds to the essence of surrealist film.

Before utilising a Buddhist perspective to further revisit Richardson's account of surrealist film, I argue that though I will not use the term "montage" in my subsequent discussion of surrealist film, montage has been included in a surrealist style. I want to interpret a surrealist style in a macroscopic way in order to extend the surrealist montage issue to a broad examination of the dialogue between a surrealist style and a surrealist relation. An extended discussion reveals that any formalist/stylistic elements, not just montage, engage with this dialogue which involves a Buddhist nominal thread. I want to use the Buddhist idea of nondualism to reflect on the ontology of surrealist film since it can be employed to complement Richardson's duality-led statement regarding the difference between a surrealist style and a surrealist relation. Nondualism can be traced back to the *Vimalakīrti Nirdeśa Sūtra* (also known as the *Vimalakīrti's Instructions*, translated into Chinese in 406 as the *Weimojie suoshuo jing* 维摩诘所说经 by Kumārajīva, 344–413). This Buddhist writing suggests:

> Darkness and light form a dualism; if there is no darkness and no light, there will be no dualism. Why? It is like entering the meditation that wipes out perception and conception, where there is neither darkness nor light. And the forms of all dharmas are the same as this. If one enters this state and views all with equality, one may thereby enter the gate of nondualism.[172]

---

[170] Michael Richardson, *Surrealism and Cinema* (Oxford and New York: Berg, 2006), p. 3 (Note: The words *here* and *elsewhere* in this quotation are Richardson's italic emphases).

[171] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), p. 1530.

[172] Kumārajīva, *The Vimalakīrti Nirdeśa Sūtra* [Weimojie suoshuo jing 维摩诘所说经] (Note: Even if the title of the book is *The Vimalakīrti Sūtra*, *The Vimalakīrti Nirdeśa Sūtra* is the full title of this sūtra. It was completed circa 100. The earliest Chinese translated version, now lost, was done in 188. The influential Chinese translated version is the sixth version, done in 406 by Kumārajīva, 344–413), ed. Alex Wayman, trans. Burton Watson (New York: Columbia University Press, 1997), pp. 109–10.

If a surrealist relation is placed into a Buddhist context, we discover that it is not fundamentally different from a surrealist style. Surrealist film exists merely by virtue of the fact that the perceiver's mind of conceptual formulations gives the essence of surrealist film a mental label such as a "style" or a "relation" that refers to it, and thus makes it become a tangible object of cognition that can be learnt, known, and understood. This implies that the nature of surrealist film is not established from its own side by virtue of an inherently findable self-entity or a defining characteristic mark rendering it truly surrealist film existing independently of the perceiver's mind. Both a "style" and a "relation" appear like illusions at a Buddhist ontological level, considering that they are simply names mentally labelled onto the essence of surrealist film without actually being that essence. Its existential quality cannot truly and inherently arise through its own power. For Richardson, if we recognise a surrealist relation, we can hold on to the essence of surrealist film, and if we fail to grasp this relation, we are inclined to mistake a surrealist style for its essence.[173] However, my Buddhist concern is not about whether we fail to identify a relation, given that there is no such thing as a "relation" existing as a substantial entity of surrealist film in the first place. Even though Richardson regards surrealist film as a relation between things formed by an activity with broadening horizons, which is analogous to interdependence in Buddhology, his contention treats a relation itself—the appearing object of his thought—as something actual. Both a style and a relation as conceptual names mentally fabricate an appearance of the essence of surrealist film. When we realise that both are merely illusions, we no longer consider that a style breaches the essence of surrealist film and that a relation conforms to it, as there is nothing as a true and inherent object going against or not going against this essence.

However, my Buddhist discussion of surrealist film does not mean to cling to the void of a style and of a relation as an ultimately cognitive approach to the nature of surrealist film. Khyentse Norbu suggests that "the path itself must eventually be abandoned, just as you abandon a boat when you reach the other shore. You must disembark once you have arrived."[174] Likewise, the empty qualities of a style and of a relation—"no style" and "no relation"—lack stable foundations, considering that they are also illusory labels at a Buddhist ontological level. By not being attached to "no style" and "no relation," we can avoid an extreme situation where we blindly see a style and a relation as total hollowness. Otherwise, our cognition will be bound by the concept of the void; we will learn nothing about surrealist film. This instead suggests that Richardson's statement on the distinction between a style and a relation can be viewed as tenable at a relative level. The formation of an idea relies on the process of the mind fixating upon the object. When Richardson puts emphasis on a relation, his mind drives the associated mental states towards this relation, which makes him reach a conclusion

[173] Richardson, *Surrealism and Cinema*, p. 3.
[174] Khyentse Norbu, *What Makes You Not a Buddhist* (Boston and London: Shambhala Publications, Inc., 2007), p. 77 (Note: Even if Dzongsar Jamyang Khyentse, also known as Khyentse, is the name on the front cover of the book, Khyentse is his first name. His surname is Norbu).

that this relation fits more with the essence of surrealist film. When other critics place emphasis on a style, they bring their minds towards this style and hence neglect a relation, which motivates them to arrive at a conclusion that reducing surrealist film to a particular style corresponds more with its nature. From a holistic perspective, a style and a relation are different appearances of the entity of surrealist film. The variability of the focus of attention makes the nature of surrealist film appear as a style or a relation. Instead of seeing a style-oriented approach as an obstacle to our understanding of the nature of surrealist film, a style can be regarded as one facet of its nature. By moving between a style/relation and "no style"/"no relation," we engage with a nominal style and a nominal relation. In this nominal state, a style and a relation exist but also do not exist as the nature of surrealist film.

Although Pramaggiore and Wallis draw a distinction between surrealist film and abstract film through different formalist/stylistic categories, I argue that they should be considered alike because abstract film also carries a nominal situation. This nominality will be dissected after a brief summary of pre-existing accounts of abstract film. Pramaggiore and Wallis discuss how animated geometrical patterns are rhythmically edited in movement.[175] They select Walter Ruttmann's *Lichtspiel: Opus I–IV* (1921–1925) and Hans Richter's *Rhythmus 21, 23, and 25* (1921–1925) as case studies to examine the relationship between graphic form and rhythmic editing. They also conduct an in-depth analysis of Stan Brakhage's *The Dark Tower* (1999) to reflect on how editing and the slices of light and colour display conflicting forces in movement.[176] In addition to hand-painted geometrical animation, there is an alternative way to express abstraction. Bordwell and Kristin Thompson point out that abstract film does not necessarily engage with animated geometrical images. Abstract elements like shapes, colours, and rhythmic motion can be extracted from concrete objects in daily reality, so that we can perceive abstract forms in everyday life more keenly.[177] They choose J. J. Murphy's *Print Generation* (1974) and Richard Serra's *Railroad Turnbridge* (1976) as instances to depict daily-life abstraction.[178]

Apart from geometrical animation and daily-life abstraction, abstract film gets involved in vivid metaphors. As analysed previously, Turim provides an account of abstract metaphors in Frampton's *Lemon*, examining the link between the abstract expression of a lemon and the viewer's association. Although a lemon is a concrete object, the shapes of light and shadow endow it with abstract forms. The viewer associates it with various symbolic fragments such as the male/female and day/night.[179] In addition to metaphors, Turim incorporates Jean-François Lyotard's view on avant-garde film into her research on the shot of the lemon. As Turim suggests, the lack-of-motion shot of the lemon can be understood as a pole of a polarity breaking a restricted performance in classic fiction films, whilst

[175] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 294–96.
[176] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 294–96.
[177] David Bordwell and Kristin Thompson, *Film Art: An Introduction* (New York: McGraw-Hill, 2008), p. 357.
[178] Bordwell and Thompson, *Film Art: An Introduction*, p. 357.
[179] Turim, *Abstraction in Avant-garde Films*, p. 25.

some other experimental films carry another pole, namely excess-of-motion shots (multidirectional motion, excessive shakes, and frame-by-frame cuts) which also subvert a restricted performance in classic fiction films.[180]

If we use the Buddhist idea of nondualism to re-examine Turim's account of a polarity between lack and excess-of-motion shots, we realise that "polarity" is not an appropriate term since it refers to a situation where two poles are completely different or opposite to each other. A lack-of-motion shot does not suggest that this shot has no motion. First, at an on-screen level, despite slowness, it still engages with a state of motion and therefore is not fundamentally at variance with a movement in excess-of-motion shots. Second, at a mind-montage level, the change (motion) of thought in time operates in conjunction with the perceived lack-of-motion shot, generating disjointed fragments of the meaning of the shot and hence sharing fragmentary traces with frame-by-frame cuts in excess-of-motion shots. From a Buddhist perspective, lack and excess-of-motion shots are different, but at the same time, they are one because of the universality of movement and of fragmentation—in the sense of a nominal domain between lack and excess-of-motion shots, which can be applied not only to geometrical animation, but also to daily-life abstraction.

Another detail of abstract film studies that needs to be noticed is that theorists P. Adams Sitney and Gene Youngblood discuss the relationship between abstract film and Buddhism. However, their contentions are quite different from my Buddhist analysis of abstract film. Broadly speaking, Sitney selects multiple paradigms—Harry Smith's *No. 10: Mirror Animations* (1956–1957), Jordan Belson's *Samadhi* (1967), and Paul Sharits' *N:O:T:H:I:N:G* (1968)—to discover how abstract patterns in these films signify Buddhist iconography and meditation.[181] Likewise, Youngblood explores how Belson is inspired by Buddhism and how he uses it as a practical form to depict abstract patterns.[182] Although Youngblood mentions the term "mind," his idea is concerned with how invisible mind meditation is visualised by abstract objects,[183] which, I argue, is not film Buddhology.

Not only can nominal entities be applied to surrealist film and abstract film, but also to the city symphony film. As for the latter, Pramaggiore and Wallis summarise that it uses montage to express the movements of modern urban life and visual comparisons. Examples embrace Ruttmann's *Berlin: Symphony of a Great City* (1927) and Vertov's *Man with a Movie Camera*.[184] Apart from daily reality

[180] Turim, *Abstraction in Avant-garde Films*, p. 26.
[181] P. Adams Sitney, *Visionary Film: The American Avant-garde, 1943–2000* (New York: Oxford University Press, 2002), pp. 246–48, 258–67, 360–61.
[182] Gene Youngblood, *Expanded Cinema* (New York: E. P. Dutton and Company, Inc., 1970), pp. 169–73.
[183] Youngblood, *Expanded Cinema*, pp. 171–73.
[184] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 295–97.

and graphic resemblances, they go on to point out that films like Vertov's *Man with a Movie Camera* and Shirley Clarke's *Bridges-Go-Round* (1958) reflect the self-reflexive presence of the filmmaker.[185]

Some other theorists express a similar view. In "Introduction: The City Symphony Phenomenon 1920–40," Steven Jacobs, Anthony Kinik, and Eva Hielscher argue that the general characteristics of the city symphony film are musical documentary footage of the modern city, the general public that makes up modernity, one-day-in-the-life of a city, fast, rhythmic, and associative montage, and self-reflexivity.[186] They provide accounts of the history and development of the city symphony film and scrutinise its different sub-types.

I do not want to repeat their analyses of the city symphony film's historical background and its sub-types. Rather, I intend to re-examine the ontological quality of the city symphony film. Although Jacobs, Kinik, and Hielscher raise a question of what the city symphony film is and think of an answer to this question: a combination of documentary film and experimental film, which is concerned with the musical rhythm of a city,[187] their point of view is merely based on its representation, rather than ontology. Within a Buddhist context, mind-related considerations can be added to their contention, raising an alternative interpretation. In *The Gelug/Kagyü Tradition of Mahāmudrā*, Alexander Berzin suggests that contents (phenomena/objects) are not pure contents, but the contents of experience or the objects of the mind, as their entities arise from the mind.[188] Likewise, in terms of the ontology of the city symphony film, I argue that Jacobs, Kinik, and Hielscher as perceivers need to utilise their eyes and eye consciousness in the first place in order to recognise urban-industrial iconography and avant-garde montage intertwined on the screen. Their observations enable them to confirm that a synthesis of documentary film and experimental film is the ontology of the city symphony film. This suggests that their perceptual process also serves as an indispensable component of the formation of its ontology. From a Buddhist perspective, their visual experiences have already transformed the attributes of the other two components—documentary film and experimental film—into two mind-related objects: documentary film of experience and experimental film of experience. Although for ease of expression, sometimes we can omit "of experience," the role of "of experience" is to remind us that all objects are dependent on our sensory fields.

---

[185] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 297–98.
[186] Steven Jacobs, Anthony Kinik, and Eva Hielscher, "Introduction: The City Symphony Phenomenon 1920–40" (2019), in *The City Symphony Phenomenon: Cinema, Art, and Urban Modernity between the Wars*, eds. Steven Jacobs, Anthony Kinik, and Eva Hielscher (New York and Abingdon: Routledge, 2019), pp. 3–42.
[187] Jacobs, Kinik, and Hielscher, "Introduction: The City Symphony Phenomenon 1920–40," p. 10.
[188] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997), pp. 20–21.

By using the above mind-related aspects to reconfigure the ontology of the city symphony film, I maintain that both documentary film and experimental film as the components of its ontology are nominal. On the one hand, they exist since urban-industrial iconography and avant-garde montage can be identified through visual cognition. On the other hand, they cannot truly and inherently exist since once they are cut off from visual cognition, they cannot be manifested and hence cannot form the existential quality of the city symphony film in their own right. By retaining a threshold between these two states, not only can the modern city-related images and montage be seen and confirmed, but also experiencers can realise their own activities of seeing and confirming these visual elements.

In addition to surrealist film, abstract film, and the city symphony film, nominality also engages with structural film. Pramaggiore and Wallis note that this experimental style features the following two parts: a reflection on the cinematic apparatus (strips of film, sound waves, cameras, and lenses) and the viewer's variable spatiotemporal experience. They use Michael Snow's *Wavelength* (1967) as an example to canvass these two facets.[189] In addition, Peter Gidal has argued that structural film stresses the materiality of film, such as flatness, grain, light, and movement, and that structural film encourages audiences to notice its material relation to themselves and to reconsider how structural film deconstructs cinematic illusions.[190] The aesthetic process of dissolving cinematic illusions is also raised by Sitney, as he suggests:

> The structural film insists on its shape, and what content it has is minimal and subsidiary to the outline. Four characteristics of the structural film are its fixed camera position (fixed frame from the viewer's perspective), the flicker effect, loop printing, and rephotography off the screen.[191]

By focussing on cinematic form itself and the change of the perception of time and space, the viewer realises their own intervention in a film. Gidal points out that "each film is not only structural but also structuring."[192] The existence of structural film is relative, as it depends on the viewer's act of structuring it. The reason why structural film pursues a more simplified expression is because its role is to discourage audiences from immersing themselves in a complex narrative and to offer them more space for reflecting on the process of structuring it.

Another key attribute that needs to be noted is the predetermined structure of structural film. The duration and frequency of images and sounds have been organised by a determinate system in advance. In discussing Gidal's *Room Film 1973* (1973), Malcolm Le Grice argues that it engages with the process of cutting the original sequence into multiple segments of equal length and each one

---

[189] Pramaggiore and Wallis, *Film: A Critical Introduction*, p. 299.
[190] Peter Gidal, "Theory and Definition of Structural/Materialist Film" (1975), in *Structural Film Anthology*, ed. Peter Gidal (London: British Film Institute, 1978), p. 1.
[191] Sitney, *Visionary Film: The American Avant-garde, 1943–2000*, p. 348.
[192] Gidal, "Theory and Definition of Structural/Materialist Film," p. 3.

presented twice, with the function of cuts being to show a salient predetermined structure by which the sequence is controlled.[193] Le Grice goes on to suggest that a predetermined structure does not necessarily mean that each element of structural film is controlled by a mechanical formula. Rather, it can be affected by subjective engagement including the viewer's reflexive attention to the object, different products of recognition and conception, and extended processes for conscious attention and structuring).[194] To put it another way, though a predetermined structure remains objective and mathematical, Le Grice calls attention to "more complex notions of procedural determinants which may not be mathematical, mechanistic or strictly predeterminate."[195] Whilst this viewpoint appears to make the ontological quality of the system of structural film become less predetermined, his idea still regards a totally predetermined structure as one possibility. That is to say, some films are totally affected by predetermined structures; some largely; some partly. Therefore, despite the fact that a predetermined structure is a dominant trait of structural film, we cannot always presume structural film to be a predetermined thing totally. From this perspective, Le Grice's contention seems to make a predetermined structure become relative and flexible, but he still treats this predetermination as a universal strategy in structural film.

However, if structural film is put into dialogue with Berzin's Buddhist discourse on the contents of experience discussed earlier, we realise that a predetermined structure is not justifiable because whether procedural determinants in structural film are strictly or partially predeterminate, there is no pregiven phenomenon existing independently of the viewer's experience. If the viewer wants to know how this structure exists and operates throughout the entire film, they can only do so through seeing and determining it. Otherwise, a structure cannot be manifested. In other words, there is no predetermined structure, as the viewer's mind and a structure are one simultaneous arising. I argue that instead of using the term "predetermined," we can simply call it a structure. However, this idea causes another problem. Any types of films, not just structural film, have structures, engaging with the viewer's mind. How to distinguish the structure of structural film from others? This is the reason why "predetermined" cannot be completely ignored. "Predetermined" can be seen as a conceptual mark indicating that this structure corresponds directly to structural film, enabling us to convey and communicate the idea of the unique feature of this particular structure. In a word, "predetermined" should be removed and retained at the same time—in the sense of the nominality of structural film.

The compilation film, just like experimental styles discussed previously, also involves a nominal thread. Pramaggiore and Wallis argue that this type of film restructures previously released footage

[193] Malcolm Le Grice, "Abstract Film and Beyond" (1977), in *Structural Film Anthology*, ed. Peter Gidal (London: British Film Institute, 1978), p. 23.
[194] Le Grice, "Abstract Film and Beyond," p. 23.
[195] Le Grice, "Abstract Film and Beyond," pp. 22–23.

in a new context in order to create new concepts. They choose Bruce Conner's *A Movie* (1958) as a case study to illustrate this recycling feature.[196] William Wees also summarises that the compilation film engages with the process of recycling and reordering pre-existing footage in a different context, with this process giving rise to new ideas.[197] He analyses the conversion process of the referentiality of the cinematographic image in the compilation film.

I want to use the Buddhist idea of impermanence to revisit Wees' discussion of the compilation film, considering that his view involves the change of pictorial meaning. Impermanence reveals that "all created things do not remain as one wishes and are not under one's control."[198] A meaning can be modified in an alternative context by virtue of re-editing, which suggests that this meaning has already possessed immanent variability. If a meaning were absolutely fixed, then even if the footage were re-edited in a new context, the meaning should still keep itself the same. This reveals that the reason why a meaning is variable is not because a film eschews teleology or finality in the first place, but because there is a quality of change operating within the meaning itself. That is to say, even if a film is goal-oriented and not put into a different context, which seems to anchor a meaning, it still has an innate tendency to change, considering that a truly invariable signification would have to be permanent when it is placed into any context. On the one hand, a meaning exists because it can be sensed. On the other hand, it does not exist because it has immanent variability. A nominal meaning is present throughout the compilation film. Even without re-editing, such nominality is still working.

After the discussion of several aspects of the nominality of experimental styles, I would like to refocus on the nominal issues of montage and the long take and select Frampton's *Zorns Lemma* as a case study. Moreover, I would like to emphasise that the previously analysed disjointed fragments of the meaning of the shot is not the only factor in producing mind montage. In the next section, I will dissect nominal montage—a threshold between mind montage and a continuous quality—from multiple facets.

3.3 *Zorns Lemma* and Buddhology

3.3.1 Previous Interpretations of the Montage Text

Before providing a close analysis of nominal montage in *Zorns Lemma* in a Buddhist context, I would like to first encapsulate how other theorists interpret this film. In Sitney's reading, Frampton's *Zorns*

[196] Pramaggiore and Wallis, *Film: A Critical Introduction*, pp. 299–300.
[197] William Wees, *Recycled Images: The Art and Politics of Found Footage Films* (New York: Anthology Film Archives, 1993), p. 35.
[198] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 1110.

Lemma is influenced by Jorge Luis Borges' labyrinthine structure of literature and contention about constructing and deconstructing the self.[199] Sitney further points out that what Frampton does is to present "montage as a logical function and cinematic construction in general as a system of thought in his film *Zorns Lemma*."[200] In this context, a logical function does not denote linear logic, but rather non-linear logic or the logic fitting with the framework of uncertainty. In addition to the labyrinthine construction, Sitney stresses that *Zorns Lemma*—the title of the film—is derived from Zorn's lemma which is a proposition of set theory in mathematics. Frampton expresses the alphabet, alphabetised signage, and wordless footage, making each shot of these symbols one second long throughout the entire film. These visual elements can be seen as ordered sets. Audiences are encouraged to engage with the process of ordering.[201] Being predetermined and simplified, the process of ordering in the form of the one-second pulse of the cuts from shot to shot is a prominent characteristic of this film.

For Turim, though the middle part of *Zorns Lemma* is established by a calculating machine-like structure, a one-second-pulse metronome, it is interwoven with Frampton's artistic intervention.[202] This interweaving is attributed to two aspects. First, Turim argues that the structure of *Zorns Lemma* is highly complex. Not only does it produce audio-visual dynamic expressivity, but it also encourages audiences to notice its incongruity, to participate in ordering structuration, and to train and change their perceptions of voices.[203] Second, Turim suggests that Frampton's subjective activity cannot be avoided. She takes Wanda Bershen's idea into account, noting that even if Bershen emphasises that this mathematical structure automatically determines how *Zorns Lemma* is presented, which makes the filmmaker absent, Bershen inadvertently (paradoxically) accepts the filmmaker's presence since Bershen takes the knowledge of Frampton's life experience into consideration.[204]

Scott MacDonald scrutinises *Zorns Lemma*'s cinematic schooling, which is analogous to William Verrone's notion of instructional avant-garde.[205] Both suggest that the film uses a contrast between montage and the long take to underline the process of the instruction in how to watch avant-garde cinema, to highlight the construction of *Zorns Lemma*, and to evoke alternative film experience. For MacDonald, the changes from the sound of reading a primer with the empty screen to daily textual symbols with the image-per-second rhythm, and from rapid montage to the sound of reading a text with a long-take snowy field reveal how a child gets to know intellectual verbal/linguistic tools and

[199] Sitney, *Visionary Film: The American Avant-garde, 1943–2000*, p. 369.
[200] Sitney, *Visionary Film: The American Avant-garde, 1943–2000*, p. 306.
[201] Sitney, *Visionary Film: The American Avant-garde, 1943–2000*, p. 369.
[202] Turim, *Abstraction in Avant-garde Films*, pp. 88–91.
[203] Turim, *Abstraction in Avant-garde Films*, pp. 88–91.
[204] Turim, *Abstraction in Avant-garde Films*, pp. 88–89; Wanda Bershen, "Zorns Lemma" (1971), in *Artforum*, ed. Annette Michelson (New York: Anthology Film Archives, 1971), vol. 10, no. 1, pp. 44–45.
[205] Scott MacDonald, *Avant-garde Film: Motion Studies* (Cambridge and New York: Cambridge University Press, 1993), p. 72; William Verrone, *The Avant-garde Feature Film: A Critical History* (Jefferson and London: McFarland and Company, Inc., 2012), pp. 88–91.

uses alphabetised words to learn and explore the world. In MacDonald's view, *Zorns Lemma* can be considered an intellectual journey from innocence and ignorance to maturity and enlightenment.[206]

Allen Weiss examines how the content of images in *Zorns Lemma* reflects religious cosmology. Overall, Weiss suggests that the theological emblems in the film include God, words and images as the representations of God, the ibis as a spiritual entity, divine light, and life and death, treating the film as the embodiment of practices of a religion.[207] I would like to stress that though I engage with a Buddhist methodology, unlike Weiss' discussion, my focus is not to link it to a religious system for living. In the following sub-sections, I will adopt Buddhology as a theoretical approach to rethinking how implicit disjointed traces arise within the long take, reflecting on their relationship to on-screen montage's fragmentation, and, in reverse, opening up discussions of what I mean by the mind long take within montage and of how such a potentially continuous quality functions during film viewing.

3.3.2 The Discontinuity of the Compound

I utilise *xiang biliang* 相比量 (appearance-inference) to examine nominal montage in *Zorns Lemma*. In sub-section 3.1.2, I argued that montage and the long take are one but not one. This idea includes two propositions. First, because of the symbolic cognition of disjointed fragments of the long take's meaning, both montage and the long take possess fragmentary traces. This common ground implies that they do not have independent and inherent existence to allow themselves to be fundamentally different from each other. Second, at a relative level, montage and the long take still retain different on-screen appearances, namely fragmentation and a continuity of movement. A threshold between the above two situations is what I call nominal montage and the nominal long take. For the sake of brevity, I shorten this term to nominal montage. There are two reasons for abbreviating it. A general consideration is that when engaging with nominal montage, the nominal long take has already been involved in the discussion since they are one. Another pertinent point is that most of *Zorns Lemma*'s contents—the forty-five-minute middle segment of the film—are displayed in the form of montage.

The first part of *Zorns Lemma* comprises the two-minute black screen and the sound of reading *The Bay State Primer* (figure 6). If this shot is examined through appearance-inference, its disjointed traces will be unveiled. As discussed previously, Maitreya suggests that "*weisui suoyou xiangzhuang xiangshu*. […] *You jianyan gu bizhi youhuo*." "谓随所有相状相属。【……】由见烟故比知有火。" (we infer what we experience from their appearances. […] We know fire due to our inferences from

[206] MacDonald, *Avant-garde Film: Motion Studies*, pp. 65, 72, 77.
[207] Allen Weiss, "Frampton's Lemma, Zorn's Dilemma" (1985), in *October*, eds. Rosalind Krauss and Annette Michelson (Cambridge: MIT Press, 1985), vol. 32, pp. 123–28.

smoke).[208] Similarly, from the appearances of the film's title *Zorns Lemma* and the sound of reading, the viewer infers that the film has begun and thus starts focussing on the black screen. Even though the black screen seems like an empty field which has no visible characters and scenery that need to be noticed and even though it lasts for two minutes, the viewer is likely to be eager to pay attention to it. This is because under the impact of the audio appearance of reading as a reminder of the start of the film, the viewer deduces that there will be subsequent visual changes. This inference suggests that what the viewer experiences is not the black screen itself, but rather a composite black-screen appearance coupled with the viewer's own expectations of visual variations. In this respect, it is not so much the empty black screen as an engaging long-take black symbol filling the entire screen. This black symbol encourages the viewer to see it as a crucial component of the film. In conjunction with the sound of reading, the black symbol serves as a potential force in the form of the long take raising the viewer's expectations of upcoming visual variations, rather than as total emptiness. However, I argue that this shot is not the absolute long take since the fitful sound of reading cuts it into multiple pieces. I will employ a Buddhist perspective to explain how the visual long take is sliced by the sound.

Figure 6 The black symbol in *Zorns Lemma*

I would like to clarify that even if sound is regarded as one of the clues as to fragmentary traces within the long take, I will not look into sound significantly in *Zorns Lemma*. Sound will be scrutinised in chapter 4. The term "sound" in my subsequent discussion is not concerned with Eisenstein's view on sound/image combination as a form of montage. Rather, at a Buddhist ontological level, I discuss how sound is utilised as one of the mind-related factors in the arising of the potential of a cut in the long take.

Appearance-inference relates to the Buddhist idea of the interdependence of phenomena. The latter suggests that any appearance that we experience is a compound. Additionally, in *What Makes*

[208] Maitreya, *Yujia shidi lun*, pp. 209–10. My translation (Note: Within Western film theoretical traditions, fire-and-smoke has been dissected in terms of Charles Peirce's notion of index. However, index is not my concern. Although the fire-and-smoke example could trigger the discussion of indexicality, I do not discuss it).

You Not a Buddhist, in discussing an assembled phenomenon, Norbu links it to its own impermanent nature.[209] For Norbu, "anything that changes in relation to another thing: even the slightest shift, is subject to the laws of impermanence."[210] If any component of a compound changes, the ontological quality of the entire compound alters. Because of impermanence, a compound lacks any underlying foundation to form a coherent, enduring, and unitary pattern. Instead, it has been transformed into discrete traces, understood as fragile and temporary collocations of its basic elements. A compound is a flux of momentary manifestations, namely a fragmentary and discontinuous string of moments which cannot hold the moments of existence together as a stable and continuous entity. Owing to the implicit discreteness of momentary being, a compound is merely a macroscopic approximation to an imaginary stable and continuous state.

Regarding the first part of *Zorns Lemma*, apart from the black symbol and the sound of reading, audiences also experience film grains and film-grain noise. From the point of view of an audio-visual compound, audiences sense an assembled phenomenon consisting of the above four elements. For the sake of brevity, I would like to call this phenomenon the "compound$_1$." Following this, owing to the fitful disappearance of the sound of reading, audiences note that the compound$_1$ has erratically changed into a compound of the remaining three elements—the black symbol, film grains, and film-grain noise. I would like to call the latter the "compound$_2$." It seems as if the perceived phenomenon were stable and continuous. Yet, at the Buddhist level of a compound, audiences are able to deduce that it has undergone two discrete traces which therefore form a latent split between them. I regard the transition from the compound$_1$ to the compound$_2$ as a mind cut, considering that this cut arises from audiences' inferential cognition of the alteration in the compound. The long-take black symbol is not the only thing that they sense. They experience a compound of which the black symbol is just a component. They can never experience the long-take black symbol as an independent entity. Even though it simulates a continuous appearance, it is still affected by a mind montage of the compound. In addition, I would like to add that even if the sound of reading were not fitful, a mind montage of the compound would still serve as a potentiality because if audiences interrupted their own hearing by covering their ears intermittently, the sound of reading could still vanish fitfully, thereby altering the state of the perceived compound. If the compound were truly stable and continuous, then even if audiences fully covered their ears, the sound of reading should still arise and hence, in conjunction with the black symbol, film grains, and film-grain noise, stabilise the state of the compound of four components. In this sense, the perceived long take always has fragmentary potential at a mind level.

[209] Norbu, *What Makes You Not a Buddhist*, p. 30.
[210] Norbu, *What Makes You Not a Buddhist*, p. 16.

The above discussion of fragmentation is not identical to the one examined in sub-section 3.1.2. The previously examined fragmentation refers to disjointed fragments of the long take's meaning—one aspect of mind montage within the long take, whilst the fragmentation canvassed in the current sub-section signifies disjointed fragments of the compound—another facet of mind montage within the long take. Even if these two approaches are somewhat different, both are fragmentary potential lurking beneath the ostensibly continuous long take.

The discontinuity of the compound is not directly visible. If audiences attempt to seek discrete traces by focussing only on the visibly long-take black symbol, rather than by thinking of the change of the compound, discrete traces cannot arise because the manifestation of the black symbol is not interrupted at an on-screen level. However, if audiences re-examine the perceived phenomenon at the Buddhist level of the compound, their inferential reasoning puts the black symbol in relation to other elements of the compound, making them realise the quantitative fluctuations in collocations of the elements. This awareness gives access to a fragmentary and discontinuous string of moments of compound existence, with the compound being devoid of coherent, enduring, and unitary quality. On the one hand, on-screen montage is manifested as fragmentary and discontinuous traces, which makes it different from the stable and continuous long take. On the other hand, there is no essential division between montage and the long take, as they share fragmentation with each other at a mind level. Depending on which side audiences engage with, long-take fragmentation can cease and arise.

The above discussion can be referred to as nominal montage. Montage does not exist, but also exists within the long take. There are some other theorists scrutinising the interconnection between montage and the long take, but their methods are different from mine. One of the most noteworthy examples is Lin Niantong's discourse on a synthesis of montage and the long take in Chinese cinema. For Lin, a prominent feature of Chinese cinema is "montage within the long take."[211] Lin argues that Chinese cinema utilises montage's conflicting representations as aesthetic bases including conflicts between two forces, thoughts, and intentions, and uses the long take to convey these contradictory things.[212] Although Chinese cinema is not my primary object of study, I emphasise that Lin's account of montage within the long take is distinct from my viewpoint. The premise of Lin's approach to the oneness of montage and the long take is to choose a film that must carry conflicting representations in the long take, whilst my perspective is not restricted by this premise. Even if the first part of *Zorns*

[211] Lin Niantong, *Jingyou* 镜游 [The Roaming Lens] (Hong Kong: Suye chubanshe 香港：素叶出版社 [Hong Kong: Suye Publications], 1985), pp. 2–11; in "In Search of Chinese Film Style(s) and Technique(s)" edited in *A Companion to Chinese Cinema*, on page 266, James Udden translates Lin's main point of the general feature of Chinese cinema into "montage within the long take."
[212] Lin, *Jingyou*, pp. 2–11. Based on Lin's detailed examination of "montage within the long take" and Udden's English translation, I provide a further English translation for Lin's contention.

Lemma does not exhibit antithetical objects, there is still fragmentary potential within the long take, which is ascribed to deductions about variable collocations of elements of the perceived compound.

In addition to disjointed fragments of the long take's meaning and disjointed fragments of the compound, there are other facets of mind montage. In the next sub-section, I will dissect the second part of the film—twenty-four letters of the alphabet displayed on the black symbol—and argue that though mind montage is different from on-screen montage, it can still be put into dialogue with on-screen montage. I will carry out a close examination of how the viewer's ability to draw appearance-inferences changes the ontological state of the perceived on-screen cuts between every two letters.

3.3.3 The Uncertainty of the Number of Cuts

In terms of twenty-four letters of the alphabet ("J" and "U" are missing), I do not discuss the process of ordering and the cognitive process of using letters to engage with all phenomena. Rather, I intend to analyse the number of cuts of twenty-four letters. At an on-screen level, the letters are presented one by one. Among them, "T" and "W" are respectively repeated once. There is a gap between "W" and "Z." "X" is repeated twice. From these visual arrangements, the spectator is able to deduce that the number of cuts of twenty-four letters is twenty-eight. Because everything shown in this finished film no longer changes, the number of cuts seems fixed. However, from a Buddhist perspective, the number of cuts is subject to uncertainty. In this regard, readers may ask: "Is this because the viewer can download and re-edit the film to change the number of cuts?" Although this can be a possibility, I want to reflect on a situation where even if the viewer does not re-edit the film at a technical level, their visual cognition can still be able to change the number of cuts at a mind level, interrupting the continuity of the long take. Moreover, these concealed cuts can appear at any point on the timeline.

In order to amplify the above situation, I synthesise the Buddhist idea of appearance-inference and Walter Murch's analysis of the blink of an eye. Murch asserts that even when people are awake, their blinks break the apparent stability and continuity of daily phenomena seen, and the signals of the blinks can be seen as reference points where cuts could have happened.[213] His view on the blink of an eye originally serves as a practice-based metaphor for film editing. He explains that the reason why audiences can quickly adapt to film editing is because they have daily experiences of dreaming and blinking. Regarding the former, images in dreams are fragmentary, thereby suggesting that film editing resembles the way images are organised in people's dreams. As for the latter, though people seem to sense the continuous world when they are awake, the world's appearance is discontinuous,

[213] Walter Murch, *In the Blink of an Eye: A Perspective on Film Editing* (Beverly Hills: Silman-James Press, 2001), pp. 59–60, 62–63.

for people's eye blinking cuts the flow of the perceived world.[214] Murch provides a further argument for the blink of an eye when he suggests:

> One of the tools to identify exactly where these cut points, these "branches," may be is to compare them to our patterns of blinking, which have been underscoring the rhythm of our thoughts for tens of thousands, perhaps millions, of years of human history. Where you feel comfortable blinking—if you are really listening to what is being said—is where the cut will feel right. […] I believe the sequence of thoughts—that is to say, the rhythm and rate of cutting—should be appropriate to whatever the audience is watching at the moment. […] In fact, statistically the two rates—of real-life blinking and of film cutting—are close enough for comparison. […] I certainly don't expect the audience to blink at every cut—the cut point should be a *potential* blink point. In a sense, by cutting, by this sudden displacement of the visual field, you are blinking *for* the audience: You achieve the immediate juxtaposition of two concepts for them—what they achieve in the real world by blinking.[215]

This contention takes into consideration a psychologically complex situation where the viewer may anticipate that the first idea should be brought to an end and that the second idea should start. The rhythm of eye blinking can be utilised to punctuate these ideas and sort them out, enabling the viewer's film experience to correspond directly to their own discontinuous visual experience in daily life. Moreover, the blink of an eye can be employed to place emphasis on the rhythm of the viewer's feelings and thoughts.

I argue that there is room to extend Murch's contention. If his analysis is put into dialogue with the Buddhist idea of appearance-inference, one could argue that the blink of an eye is not only used as a practical strategy for film editing, but also regarded as a theoretical factor in reshaping the way the viewer deduces the number of cuts. I focus on the latter and link it back to the discussion of the number of cuts of twenty-four letters of the alphabet. Previously, I calculated that this number was twenty-eight. However, this is not invariable. From a Buddhist perspective, as what the viewer sees is an assembled phenomenon linked to their visual cognition, one should not only count the number of on-screen cuts, but also consider any mind-related aspects that give rise to a break in the flow of the perceived phenomenon, even though these considerations can be subtle. An eye blinking-based cut in the experienced continuity can be considered another path to no essential boundary between montage and the long take.

Initially, the spectator sees "A" (figure 7). Although it only lasts for one second, this one-second shot, at a relative level, can be regarded as a very short long take. It carries a one-second continuous

---

[214] Murch, *In the Blink of an Eye: A Perspective on Film Editing*, pp. 58, 59–60, 62–63.
[215] Murch, *In the Blink of an Eye: A Perspective on Film Editing*, pp. 68–69 (Note: The word *potential* in this quotation is Murch's italic emphasis).

state. It is stable until it is displaced by “B.” However, if the viewer blinks before “A” is displaced by “B,” even though the fleeting flow of “A” is not cut at an on-screen level, its experienced momentary continuity can still be broken by the viewer’s eye blinking at a mind level. From the point of view of a compound, eye blinking is a part of visual experience, affecting the appearances of things that the viewer sees. From the moment of shutting their eyes, the viewer can note that the flow of “A” stops in the middle. They suddenly recognise a closed-eye dark field inside their eyelids. Because this dark field happens in a flash, it is easily neglected. Nevertheless, such extremely subtle discontinuity still forms a latent cut. From the moment of opening their eyes, the viewer can discover that the ignored instant of a closed-eye dark field ceases and that the flow of “A” starts again—the arising of another latent cut. Owing to a blink, the flow of “A” has already been cut twice before it is displaced by “B.” Thus, from the above experienced traces, the viewer can deduce that the number of cuts of twenty-four letters is thirty, rather than twenty-eight. A blink-based factor reveals the impermanent nature of the number of cuts.

Furthermore, the rate of eye blinking varies from person to person. In the process of perceiving twenty-four letters, some viewers may keep their eyes open for five seconds and some ten seconds, with many variations in the rate of eye blinking. Therefore, the viewer can infer that the number of cuts of twenty-four letters is uncertain. It is not necessarily thirty. The change of the number of cuts arises from awareness of the compound of on-screen cuts and the viewer’s eye blinking (mind cuts). Even though the number of cuts appears fixed, its experienced state can be affected and altered by the viewer’s eye blinking. When on-screen cuts and eye blinking are regarded as a whole, the viewer realises the prevalence of mind montage. In other words, cuts can arise at any point on the timeline. This is the reason why the viewer’s inferential cognition of the number of cuts of twenty-four letters can still be reshaped even if they do not re-edit the film. A reflection on the compound reminds the viewer to notice their own closed-eye dark field that breaks the flow of the perceived phenomenon.

Figure 7 The alphabet in *Zorns Lemma*

From a Buddhist point of view, the spectator's inferential cognition of a compound appearance enables themselves to notice that the blink of an eye is an unstable component of the formation of the perceived shot and therefore to concentrate on the impermanent nature of the number of cuts. This implies that because the continuous state of the perceived long take can be fragmented by eye blinking and because montage carries fragmentation, there is no absolute boundary between them.

However, in terms of the above discussion of the number of cuts, I would like to further suggest that just because Buddhology emphasises impermanence and interdependence does not mean that the viewer must abandon the cognition of twenty-eight cuts and cling to uncertainty as the ultimate level of the number of cuts. This number depends on how the viewer's inferential cognition engages with fragmentation. When the viewer only considers on-screen fragmentation, twenty-eight arises as an explicit number, whilst uncertainty is neglected and therefore serves as an implicit potentiality on the bottom of the number of cuts seen. When the viewer shifts the focus of their attention from on-screen fragmentation to compound fragmentation, that is, when the viewer connects on-screen fragmentation to their bodily function of eye blinking, twenty-eight is concealed, whilst uncertainty becomes explicit. Because the focus of attention can change between on-screen fragmentation and compound fragmentation, neither twenty-eight nor uncertainty can be seen as a firmly established inferential outcome. A Buddhist perspective is more concerned with an intermediate state between twenty-eight and uncertainty than inferential results themselves. A nominal domain between these two cognitive outcomes indicates that the number of cuts is in a flowing state where its determinate and indeterminate aspects convert into each other. Not only is this conversion process pertinent to impermanence, but also it is important to realise that the nominality of the number of cuts can only happen when both twenty-eight cuts and uncertain cuts are taken into consideration and regarded as interconnected, which reflects interdependence. This ontological interchange signifies that there is no essential distinction between twenty-eight cuts and uncertain cuts. On the one hand, montage has relative existence because the viewer can see and confirm changes between shots. On the other hand, montage does not have independent and inherent existence, as eye blinking can intervene in the perceived appearance to reorganise the fragmentary state. These two aspects respectively arise from specific foci of attention, but at a holistic level, both are the processes of inferential cognition.

Although blink-based compound fragmentation can be applied to an analysis of the next forty-five-minute montage of the alphabetised signage and the wordless footage, I do not want to repeat my observations. Rather, I will associate the Buddhist idea of appearance-inference with Wollen's interpretation of Jean-Luc Godard's view on the meaning-generating process to scrutinise this forty-five-minute montage. The reason why I bring the meaning-generating process into the subsequent discussion is because it deals with the disjunction between signifier and signified, which is linked to

an examination of signs. In the following sub-section, I will examine how appearance-inference adds an alternative view to sign-based montage.

3.3.4 Disjointed Meanings and Roughly-the-same Experiences

The forty-five-minute montage has two facets: the alphabetised signage and the wordless footage. Unlike the sudden change between the black symbol and the alphabet, there is a gradual transition from the alphabetised signage to the wordless footage—a step-by-step process of substituting the wordless footage for the alphabetised signage. The alphabetised signage is composed of words that appear on various urban signs in everyday life, such as advertising, traffic, and shop signs (figure 8). Each shot contains one word. The length of each shot is set to be one second—twenty-four frames.

Figure 8 The alphabetised signage in *Zorns Lemma*

I view each one-second shot as a very short long take and revisit how its apparently continuous state engages with a fragmentary potentiality. Specifically, I carry out a re-examination of disjointed fragments of the meaning of the shot of a sign. Moreover, I point out that even if a continuous short long take has only one dominant meaning, its signification is not totally fixed, stable, and unitary at a Buddhist ontological level, which therefore adds a deeper fragmentary state to the perceived shot of a sign. This expanded fragmentary state will be dissected after re-examining disjointed meanings.

If the spectator only pays attention to a particular shot of a sign, it can serve as an indicator of a certain place. However, when their inferential cognition puts the visual appearance of this shot in relation to on-screen montage, it can be transformed into a component of, for example, the process of ordering or an intellectual journey. Their variable thought on this shot has divided its signification into fragments and made them into a collage of meaning making. This perceived shot does not have to convey a certain location in the city all the time. Due to the change of symbolic thinking, this shot does not carry a coherent, enduring, and unitary meaning. Apart from a place, it can metamorphose into an element of a mathematical or linguistic skeleton, which gives rise to the disjunction between signifier and signified. This reveals that at a meaning-making level, the one-second continuous state of this shot has contained within itself discrete traces—the first clue as to a fragmentary potentiality within a very short long take.

I would like to re-examine Wollen's "The Two Avant-gardes" in which he discusses how Godard reflects on the disjunction between signifier and signified. Specifically, Wollen suggests that Godard addresses the process of deconstructing a persuasive appearance of the world. Such deconstruction arises from a conflict between signifier and signified, between distinct types of codes.[216] For Wollen, Godard calls attention to the whole process of signification and explores different possible paths to the meaning of a film. That is, Godard attempts to free the viewer from the constraints of ostensible naturalness in order to get involved in indeterminate activities of deconstructing and reconstructing meaning.[217] I argue that when experiencing the shot of a sign, just like a dialectical collision between shots through juxtaposition, disjointed fragments of meaning contained within this single shot itself can also be assembled in the viewer's mind, thereby endowing its apparently one-second continuity with mind montage.

I point out that another vital aspect of Wollen's interpretation of Godard's meaning-generating process that is helpful in engaging with *Zorns Lemma* is that Godard's contention does not mean to downplay a particular meaning.[218] The juxtaposition of numerous significations does not encourage the spectator to interpret a film arbitrarily. It also does not suggest that a certain meaning is useless. Rather, this juxtaposition opens up a threshold between fixed meaning and free meaning.[219] On the one hand, a particular meaning can help the viewer further comprehend a film. It is not a completely unmanageable code. On the other hand, it does not possess a dominant position. It is a temporarily fixed meaning subject to being replaced by another meaning.

Likewise, the shot of a sign should not be seen as something that cannot express a certain idea. A "place in the city" as a fragment of meaning does not have to be forsaken since it is still a facet of the viewer's understanding of a sign. However, a "place in the city" is not the ultimate code, as it is superimposed on other significations such as a "mathematical unit" and a "linguistic unit." It follows that a "place in the city" is a relative code. Such superimposition reflects the nominal fragmentation of the shot of a sign. On the one hand, a code arises in this one-second shot, maintaining its symbolic consistency at a relative level. At this very moment, there are no other codes interrupting it, thereby not yet creating a conflict between codes. That is, symbolic fragments have not yet happened within this single shot. On the other hand, a code lacks any basis for establishing a coherent, enduring, and unitary quality, as other codes break its symbolic consistency and create various symbolic fragments within this single shot. An intermediate state between these two aspects reflects nominal montage. This one-second shot itself is continuous, but simultaneously, it contains a fragmentary potentiality.

[216] Wollen, "The Two Avant-gardes," pp. 171–75.
[217] Wollen, "The Two Avant-gardes," pp. 171–75.
[218] Wollen, "The Two Avant-gardes," pp. 171–75.
[219] Wollen, "The Two Avant-gardes," pp. 171–75.

However, I argue that the arising of symbolic fragments is not necessarily based on the premise that a film deliberately suspends meaning. It seems that the reason why the shot of a sign possesses a fragmentary potentiality is because *Zorns Lemma* invites the viewer to interpret it differently. Yet, if *Zorns Lemma* had only one dominant code, if the film did not avoid teleology or finality, or if there were no flexible context for producing multiple possible paths to meaning, would meaning still have the potentiality of being freed from any constraint and initiating symbolic fragments within the shot of a sign? In a Buddhist realm, Tian Maozhi emphasises that "even among homogeneous life forms, the worlds that they experience can only be 'roughly the same,' rather than being exactly the same. 'Roughly-the-same' experiences reveal that the perceivers' 'minds' are similar."[220] If Tian's Buddhist contention is applied to the experience of a hypothetically dominant code, we discover that even if we choose to hypothesise that *Zorns Lemma* were a certainty-oriented film, the shot of a sign would still get involved in symbolic fragments. At first, when one sees and deciphers the shot of a sign, the process of visual cognition guides one towards a corresponding code: a "place in the city." However, even though this code were dominant, it is crucial to note that this inferential outcome only belongs to this specific viewer since such inferential reasoning arises from this particular viewer's perceptual system, not other viewers' sense organs and sense consciousness. When other viewers see this shot, multiple viewers' inferential results are produced. Even though all these results pointed to the same code, the cause of this code would still be freed from any constraint since its cause is not limited to one viewer's perceptual system. These inferential results can be viewed as multiple similar symbolic fragments on the bottom of an apparently dominant code. They are not identical and hence are not able to form a unitary state. These roughly-the-same inferential results are spliced as a mind collage within the shot of a sign.

Although the shot of a sign has obviously carried multiple codes such as a "place in the city," a "mathematical unit," and a "linguistic unit" discussed earlier, I reduce signification to one possibility in order to call attention to the pervasive freedom of meaning of experience. The arising of symbolic fragments is not absolutely ascribed to an uncertainty-related film. It can be caused by roughly-the-same experiences suggested in Buddhology, which makes the unitary state of a single code fall apart in an implicit way. The relative existence of a code and roughly-the-same experiences together form nominal montage within the shot of a sign. A code can retain its symbolic consistency relatively, but it can be understood as various similar appearances rooted in multiple viewers' perceptual systems.

[220] This Buddhist discussion comes from Tian Maozhi's preface to the *Jingang jing* 金刚经 [The Diamond Sūtra]. Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007), p. 11. My translation.

3.3.5 Unity as an Unfragmented Impression

As analysed previously, the long take or even the one-second shot (a very short long take) possesses a fragmentary potentiality or mind montage. This idea brings up another question: is there a reverse situation where montage carries the potential of the long take? With regard to sign-based montage, readers may ask: "Is this long-take potentiality attributed to the fact that each shot of a sign creates a very short but still continuous illusion, even though each of them is affiliated to the overall context of sign-based montage?" Although this is a consideration, I plan to analyse a continuous potentiality different from a visible one-second continuous illusion.

In Buddhology, multiple perceived appearances, even if they can be different at a relative level, are understood as a holistic unity. Buddhabhadra and Śikṣānanda suggest that the Great Vehicle of Buddhahood needs to be dependent on the unity of being and comprehensiveness of approach and engages with the purification of holistic knowledge, manifested as a potentially undefiled quality of reality continuous and unfragmented throughout past, present, and future.[221] In this sense, I assert that when unity is applied to the inferential cognition of sign-based montage, it draws attention to a continuous and unfragmented rhythmic or conceptual stream of experience of on-screen cuts. I would like to call this stream of experience the "mind long take." This continuous and unfragmented state of the mind is not based on long-take realism, but rather based on a holistic impression of on-screen cuts. Even if there are many on-screen cuts, the viewer can still recognise that these cuts are skilfully unified in a mathematical or linguistic system. From a holistic perspective, this system keeps its continuous engagement with a one-second-pulse metronome and the alphabetised signage. The state of the viewer's mind is not broken by the intrusion of other cinematic elements such as linear narratives. From this holistic consideration, the viewer can infer that sign-based montage possesses a continuous ordered-set or semantic thread conveying an unfragmented impression. This rhythmic or conceptual stream of experience transforms the ontological quality of sign-based montage from the discrete into the unified and unbroken, presenting in the viewer's mind an evolutionary flow of an ordering or intellectual journey.

The mind long take cannot be directly seen since its manifestation needs to rely on the viewer's holistic inference about cuts. What they see are not cuts as independent phenomena, but rather as dependently originated phenomena associated with an unfragmented impression. At the Buddhist level of unity, the viewer perceives a fragmentary appearance which is not fragmentary. It functions as a flow of a rhythmic or conceptual construct. Without the viewer's holistic appearance-inference, the mind long take cannot occur. On-screen cuts and the mind long take give rise to a nominal realm

[221] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, pp. 696, 1510, 1525.

between fragmentation and continuity. My contention that sign-based montage is fragmentary and continuous is connected to my previous idea that the long take is continuous and fragmentary. Both have transcended the imaginary boundary between fragmentation and continuity and become one.

3.3.6 Different Degrees of Fragmentation

The fourth part of the film is closely intertwined with the third one, with the wordless footage being gradually substituted for the alphabetised signage within the structure of the one-second pulse of the cuts from shot to shot. The wordless shots, for example an egg being cooked, a man walking on the street, and a wall being painted, are used as alternative symbols to represent the alphabetised signage (figure 9). The third part continues to cycle through the signs, but they are different in every cycle, whilst in the fourth part, when a certain sign is replaced, it has the same wordless substitution continuing for one second each time. Originally, the wordless replacements are long-take materials. However, they are cut into multiple pieces by the filmmaker at the stage of post-production in order to make them fit with a one-second-pulse metronome. At first, the viewer recognises a one-second substitution. After multiple repetitions, they are able to learn that the first frame of this substitution in the next cycle is connected to the last frame of this substitution in the previous cycle.

Figure 9 The wordless footage in *Zorns Lemma*

At the inferential level of the referent of the shot, the alphabetised shot and the wordless shot are distinct. The former presents abstract letters, whilst the latter shows concrete actions. However, if one shifts their inferential orientation from the referent of the shot to the shot length, they realise that both the shots of abstract letters and concrete actions reveal the same extent of fragmentation, as both are arranged in a one-second-pulse metronome, which reveals that there is no fundamental difference between them. When these two levels are both taken into consideration, a nominal state between difference and “no difference” arises. Based on the Buddhist idea of nondualism discussed earlier, I argue that all objects and inferential results, for example the difference and “no difference” between the alphabetised shot and the wordless shot, have equality, since without the mind linking them to their referents and lengths, both their difference and “no difference” cannot be established. This suggests that both their difference and “no difference” do not possess a defining internal force empowering their own existence in the first place. A nominal state between them is ascribed to the

variability of the experience of the shot. Depending on whether the mind connects the alphabetised shot and the wordless shot to their referents or to their lengths, they can be either heterogeneous or homogeneous. The changeable operation of the perceiver's mind alters the way the shot appears.

Although I suggest that at the deductive level of the shot length, the alphabetised shot and the wordless shot reveal the same degree of fragmentation, the degree of fragmentation is still variable if the experiencer continues to change their inferential orientation. For example, if the alphabetised shot and the wordless shot are experienced at the level of the original form of the shot, they do not necessarily show the same extent of fragmentation. From the direct link between the first frame of a wordless shot in the next cycle and its last frame in the previous cycle, the viewer can deduce that the original form of this wordless shot should be a long take. Due to the intention to organise a one-second-pulse metronome, the filmmaker cuts this wordless shot (a continuous concrete action) into multiple slices at the stage of post-production. When the viewer connects these pieces in their mind, they recognise the gradual development of an action. Some actions are still ongoing, such as turning the pages of a book, playing on a swing, and sending out smoke, whilst other actions are completed at the end of the cycles, such as a wall painted, a shoelace tied, and a tyre changed. Each reveals its own flow of daily life. When watching these wordless slices, the viewer becomes curious about how these actions develop throughout the film. Curiosity motivates them to wait for answers, that is, to expect more long take-originated wordless pieces which will be uncovered in the subsequent cycles. The viewer is persuaded to focus on the slices of an action, make the direct connection among them, and restore them to an originally complete long take. A desire to make these wordless pieces return to a long-take form makes them become less fragmentary. The wordless shot contains a continuous potentiality, with its degree of fragmentation being reduced due to the attention to its original form.

However, when the viewer experiences the alphabetised shot, they learn that there is no direct clue to support the idea that its original form is a long take. This is because in the next cycle, instead of keeping presenting this same sign, it is replaced by a different sign with a different word, thereby showing more extent of fragmentation than the wordless shot. Even if this alphabetised shot has a one-second continuous state, its substitution is not derived from the previous sign, which therefore lowers the viewer's expectations of linking it to a long-take form. Not only is the difference between the alphabetised shot and the wordless shot attributed to the viewer's reflection on their respective referents (abstract letters and concrete actions), but also to the viewer's low and high expectations of associating them with a long-take form. On the one hand, from the perspective of the shot length, the alphabetised shot and the wordless shot reveal the same extent of fragmentation, as both serve as a one-second flow. On the other hand, from the perspective of the original form of the shot, the alphabetised shot is more fragmentary than the wordless shot. In other words, the variability of the viewer's deductive perspective endows the extent of fragmentation with a great deal of uncertainty.

I want to clear up a problem lurking beneath my discourse on the alphabetised shot. Previously, I used unity—a Buddhist idea considering different perceived appearances a continuous whole—to point out that there is a continuous and unfragmented rhythmic or conceptual stream of experience of on-screen cuts. These cuts are skilfully unified in a mathematical or linguistic system which keeps its continuous engagement with a one-second-pulse metronome and the alphabetised signage. This continuous ordered-set or semantic thread leaves an unfragmented impression on the viewer. Thus, this suggests that the alphabetised shot is not the embodiment of fragmentation, which appears to conflict with the idea analysed in the last paragraph that the alphabetised shot shows a great extent of fragmentation. However, this instead indicates that the alphabetised shot is devoid of existing in an independent and permanent way. When the viewer shifts their inferential orientation from unity to the original form of the shot, the ontological quality of the alphabetised shot is transformed from the unified and unbroken into the discrete. Based on nondualism, I suggest that instead of regarding these two inferential results as completely different, both its continuous and fragmentary qualities are the creations of the viewer's mind. On the one hand, the ontological quality of the alphabetised shot is uncertain until the viewer's mind picks a particular inferential orientation. On the other hand, when perception and inference are wiped out, both its continuous and fragmentary qualities vanish, as both their manifestations need to lean on the viewer's acts of observing and considering the shot.

3.3.7 Revisiting the Discontinuity of the Compound

The fifth part of the film exhibits a long-take snowy landscape lasting for eleven minutes (figure 10). Two people and a dog walk through the snow and into the distant woods, with six voices alternating in reading the words of a passage from Robert Grosseteste's treatise *On Light, or the Ingression of Forms*. The text is read at a rate of one word per second, which implies that the fifth part of the film has greater frequency of intermittency than the first part because at the beginning of the film, some sentences read are longer than one second. As stated previously, from a Buddhist perspective, what the viewer experiences serves as a compound. Likewise, regarding the ending of the film, the viewer experiences a compound of five facets—the characters, the snowy landscape, film grains, film-grain noise, and the sound of reading.[222] Owing to a one-second-pulse metronome produced by six voices, the perceived compound has undergone the constant changes from "the compound with the sound of reading" to "without it" to "with it…" From the variation of the relation among the elements, the viewer is able to deduce that the quantitative fluctuations in collocations of the above facets divulge a fragmentary and discontinuous string of moments of compound existence. Moreover, I would like

[222] Apart from these five facets, some beat sounds are also synchronised with six voices. Yet, sometimes when a voice pauses for a little longer, a beat sound still continues. For the sake of brevity, I would like to group the beat sound into the sound of reading, considering that the beat sound, like the tempo-based sound of reading, can also be counted as a tempo-based facet.

to explain that even if the sound of reading were not punctuated by a one-second-pulse metronome, the discreteness of the compound would still act as a potentiality because if the viewer interrupted their own hearing by covering their ears one second at a time, the sound of reading could still vanish fitfully. Thus, a sense of discontinuity can still be created. Although it seems like the fifth part of the film were more fragmentary than the first part because the former's intermittency is more frequent than the latter, I assert that there are no essentially high and low degrees of intermittency at a mind level. If the viewer covered and opened their ears in a rapid and repeated way whilst perceiving the first part, the perceived compound would undergo more discrete traces and thus show more extent of fragmentation. Depending on how the viewer senses the compound, the extent of fragmentation is indeterminate.

Figure 10 The snowy landscape in *Zorns Lemma*

3.3.8 A Buddhist Bridge between Film Ontology and Film Representation

Concluding this chapter, I would like to point out that even if my research focus is to use Buddhology as a philosophical methodology to revisit film ontology, rather than to dissect Buddhist iconography and contemplation in film representation, it is feasible for my contentions to affect the way we look at Buddhism in film representation. In terms of the link between the first and fifth parts, MacDonald considers it an intellectual journey from innocence and ignorance to maturity and enlightenment.[223] He regards the black frame at the very beginning as childhood innocence and ignorance and regards the white frame at the very end as mature enlightenment—a straight unidirectional intellectual trip from the former to the latter. Yet, I regard this journey as circle-line-towards-enlightened ignorance. A circular cognitive mode can be traced back to the *Zhiyue lu* 指月录 (also known as the *Shuiyuezhai zhiyue lu* 水月斋指月录, edited by Qu Ruji in 1602, translated into English as the *Records of Pointing to the Moon*). Qingyuan Xingsi, one of the Buddhist scholars mentioned in this Buddhist work, states:

[223] MacDonald, *Avant-garde Film: Motion Studies*, pp. 65, 72, 77.

> Laoseng sanshinian qian, wei canchan shi, jian shan shi shan, jian shui shi shui. Jizhi houlai qinjian zhishi, you ge ru chu. Jian shan bushi shan, jian shui bushi shui. Erjin de ge xiuxie chu, yiqian jian shan zhishi shan, jian shui zhishi shui. Dazhong, zhe sanban jianjie, shi tong shi bie? Youren zisu dechu. Xu ru qinjian laoseng. 老僧三十年前，未参禅时，见山是山，见水是水。及至后来亲见知识，有个入处。见山不是山，见水不是水。而今得个休歇处，依前见山只是山，见水只是水。大众，这三般见解，是同是别？有人缁素得出。许汝亲见老僧。 (Thirty years ago, when I did not perform Chan meditation, I saw mountains as mountains, and water as water until I had acquired knowledge and found access to a situation where I no longer saw mountains as mountains, and no longer water as water. But now that I have got its very substance I am at rest. For it is just that I see mountains once again as mountains, and water once again as water. Everyone, are the foregoing three cognitive stages same or different? If anyone can clearly discuss this matter, I will invite this person to meet me and give this person a further sense of direction).[224]

The first cognitive stage refers to ignorance. Due to ignorance, the perceiver believes that both mountains and water have independent and permanent qualities to establish their inherent entities. The second cognitive stage signifies Buddhist knowledge. By virtue of acquiring Buddhist knowledge, the perceiver realises that what they perceive are simply compound and variable appearances. They are interdependent and impermanent and thus are devoid of existing in a truly inherent way. As for the third cognitive stage, it is not ultimately distinct from the first one. Once again seeing mountains as mountains and water as water can be regarded as returning to the first one, a circumnavigation as it were. Although this intellectual journey indicates a return to the same point, it does not suggest a return to the same point with exactly the same cognition. The perceiver still sees phenomena that they had seen during the stage of their ignorance, but now their seeing is connected to (enlightened) ignorance. In other words, just because mountains and water do not inherently exist does not mean that they are absolute nothingness. Rather, they exist in a relative condition. The perceiver still sees that mountains and water are there, but simultaneously, they realise that mountains and water are not there. This intellectual journey displays a nominal state between mountains and "no mountains," between water and "no water." Such self-knowing ignorance releases the perceiver from the prison of ignorance, but at the same time, it encourages the perceiver to re-examine objects that they had seen during the stage of their ignorance and to recognise objects at a relative level.

Likewise, in terms of *Zorns Lemma*, I maintain that the first part's black symbol (ignorance) can be seen as the start of a cognitive journey, whilst the fifth part's white landscape does not represent enlightenment, but rather signifies knowledge (access to enlightenment). Rather than MacDonald's one-way journey from ignorance to enlightenment, within a Buddhist context, it is a round trip from

---

[224] Qu Ruji, *Zhiyue lu* 指月录 [The Records of Pointing to the Moon] (1602) (Chengdu: Sichuan chuban jituan bashu shushe 成都：四川出版集团巴蜀书社 [Chengdu: Sichuan Bashu Publishing House Ltd], 2005), vol. 28, p. 814. My translation.

ignorance to (enlightened) ignorance. The white landscape is not a terminus, but rather knowledge leading us to (enlightened) ignorance. This circular journey gives access to a nominal state between "what we see is what we see" and "what we see is not what we see." This intermediate state neither accepts nor rejects these two parts, integrating them into a seamless whole. Whether the object of study is film ontology or film representation, from a Buddhist perspective, both relate to nominality.

Chapter 4 Nominal Voices in Docufictions

4.1 Nominal Voices

This chapter deals with the nominality of voices in docufictions, by which I refer to hybrid forms between fiction and documentary film.[225] My argument in this chapter is that there is no thorough boundary between voices and "no voices," for "no voices," at a mind level, have the potential of voices. This statement involves two parts. First, even if a film has no vocal tracks, vocal interventions can still be present. Second, the voice has no independent and permanent essence that allows itself to be truly different from other voices, but there are still relative differences among them. My Buddhist examination of voices will centre upon their acoustic qualities, rather than their linguistic meanings. Nevertheless, I would like to argue that this is not meant to refute their linguistic meanings. Rather, linguistic meanings make voices become one of the most widely scrutinised elements in docufiction-related theoretical works, though the titles of these works do not exclusively deal with docufiction but more generally with documentary film.[226] Examples include Stella Bruzzi's *New Documentary*, Bill Nichols' *Introduction to Documentary*, and Annabelle Honess Roe and Maria Pramaggiore's edited work *Vocal Projections: Voices in Documentary*.[227] All of them raise a point that voices play an influential role in recounting the history of an event and in explaining images.

4.2 Voices and the General Features of Docufictions

Before discussing voices' acoustic qualities in docufictions, I want to first outline the general aesthetic characteristics of docufictions. This aesthetic information can be illustrated by Soviet cinema. I want to suggest that though Soviet cinema is not my primary object of study, some of its conspicuous features should still be taken into consideration. In the third chapter, I summarised Peter Wollen's discussion of how montage in Soviet avant-garde films engages with daily-life realism. However, daily-life realism does not necessarily completely rely on recorded things, but also on fictional qualities. Soviet cinema has been treated as the first embodiment of docufictions. In *A New History of Documentary Film*, Betsy McLane argues:

---

[225] The term "docufictions" is proposed by Gary Rhodes and John Springer. Gary Rhodes and John Springer, "Introduction" (2006), in *Docufictions: Essays on the Intersection of Documentary and Fictional Filmmaking*, eds. Gary Rhodes and John Springer (Jefferson and London: McFarland and Company, Inc., 2006), p. 5.

[226] Within Western film theoretical traditions, documentary is not a radically independent object of study, as it has been put into dialogue with other filmic types; take fiction, for instance.

[227] Stella Bruzzi, *New Documentary* (London and New York: Routledge, 2006), pp. 47–50; Bill Nichols, *Introduction to Documentary* (Bloomington and Indianapolis: Indiana University Press, 2001), pp. 99–138; Annabelle Honess Roe and Maria Pramaggiore, eds. *Vocal Projections: Voices in Documentary* (New York and London: Bloomsbury Publishing Plc, 2019).

> It is sometimes impossible to separate documentary from fiction; hybrids have existed since the inception of cinema and continue to fascinate documentarians in the twenty-first century. […] The first recognised group of such films that fused fiction-documentary work are the Soviet silent features.[228]

Sergei Eisenstein's films *Strike* (1925) and *Battleship Potemkin* (1925) are notable examples of docufictions. On the one hand, McLane suggests that Eisenstein engages with real social events (e.g. a labour protest and the 1905 Russian Revolution), real location shooting, and non-professional performers—in other words, documentary-like characteristics. On the other hand, McLane suggests that Eisenstein shapes performances through well-made scripts—in other words, fiction-like characteristics.[229]

McLane's argument regarding docufictional features is picture-centred. In the subsequent sections, I will focus on voices in docufictions by using the Buddhist idea of *yinguo biliang* 因果比量 (causality-inference). This notion puts emphasis on the variability of interrelationship, revealing that "if a cause changes, its effect changes, and vice versa."[230] Causality-inference was mentioned in sub-section 3.1.2, but it can be further employed to explore the variability of causality within a vocal context. The variability of causality will be significant access to a nominal realm between voices and "no voices."

At the level of the Buddhist idea of causality-inference, I will draw on nominal voices to examine the following voice-related ontological problems. First, I reconsider Bruzzi's discussion of whether voices violate the visual quality of film.[231] Her analysis depends on her focussing on the soundtrack of voices *from film*. However, I point out that if her focus converts into voices *from film and audiences' experiences*, a different effect will appear. A film experience-based consideration includes a situation where there is still the potential of voices even if a film has no vocal tracks. These implicit voices do not result from a live narrator who provides commentary or characters' voices, but from audiences' experiences. I argue that these covert voices can be manifested if film and audiences' experiences are considered a seamless whole. I will use a Buddhist perspective to expand on voices from audiences' experiences, and how these potential voices help us reflect on the question of whether voices vitiate the visual quality of film.

Second, in addition to voices from audiences' experiences, another consideration that also has an effect on audio experience is physical space. I will use a nominal interpretation to engage with

---

[228] Betsy McLane, *A New History of Documentary Film* (New York and London: Continuum International Publishing Group, 2012), pp. 51–52.
[229] McLane, *A New History of Documentary Film*, pp. 53–54.
[230] See p. 51.
[231] Bruzzi, *New Documentary*, pp. 47–50.

Rick Altman's idea of the material heterogeneity of recorded sound.[232] This particular acoustic matter conflates audiences' experiences and the surrounding environment that transmits sound. Altman asserts that sound in cinema not only relates to invariable abstract and conceptual elements such as loudness, pitch, and timbre, but also to the changeable nearby space. It is mutable three-dimensional environment that causes the "same" sound to be heterogeneous.[233] In other words, because of the additional effects of different locations, sound is not independent and its quality can change. It cannot keep its permanent existence. From my perspective, its heterogeneity can be seen as one-sided, caused by Altman's attention to the changeable surrounding space. If the object of our attention turns to the above fixed aural concepts, the effect will be that a sound is still "that sound" at a relative level. The variability of causality thus makes the "same" sound heterogeneous and homogeneous. When its homogeneity is taken into account, a nominal phenomenon between heterogeneity and homogeneity of recorded sound will arise.

Third, though the above information offers readers the general facets of vocal nominality, these aspects do not involve specific types of voices. I will further dissect how nominality engages with these types. Nichols proposes six documentary modes, with each of their vocal issues producing a distinctive effect. They are subdivided into the poetic mode, the expository mode, the observational mode, the participatory mode, the reflexive mode, and the performative mode.[234] I would suggest, however, that the vocal issues of these different modes can be considered alike based on the fact that they all engage with nominality. I argue that this does not mean to say that Nichols' separate delineation obstructs nominality. His discrete delineation has its own role, helping us better understand what the specific characteristic of each mode's vocal issue is. I point out that each mode's vocal issue has two distinct situations which will be unified by a particular factor—a nominal state between difference and oneness. This intermediate region helps us reconsider a voice's ontological relationship to another voice. I will expand on how this intermediate state functions within the context of each mode's vocal issue.

4.3 A Buddhist View on Voices in Docufictions

4.3.1 The Interventions of Voices from Audiences' Experiences

In *New Documentary*, Bruzzi asserts that "voice-over, in both documentaries and fiction films, is an extra-diegetic soundtrack that has been added to a film."[235] Her examination of voice-over is

[232] Rick Altman, "The Material Heterogeneity of Recorded Sound" (1992), in *Sound Theory, Sound Practice*, ed. Rick Altman (New York and London: Routledge, 1992), pp. 15–31.
[233] Altman, "The Material Heterogeneity of Recorded Sound," p. 16.
[234] Nichols, *Introduction to Documentary*, pp. 99–138.
[235] Bruzzi, *New Documentary*, p. 47.

composed of its representational and ontological facets. In terms of the former, she suggests that voice-over possesses the function of narration that provides audiences with a deep insight into what happens in a film, emitted by a disembodied and omniscient narrator.[236] In addition, she deals with the ontological issue of whether voice-over spoils the visual nature of film.[237] In order to discuss this question from multiple perspectives, Bruzzi puts different filmmakers and scholars' ideas into dialogue, such as Paul Rotha, Robert Drew, and Sarah Kozloff. I intend to look deeply into this ontological aspect.

Bruzzi considers that in Rotha's opinion, voice is greatly hazardous to the visual characteristic of film since Rotha regards film as an exclusively visual expression. Although Bruzzi does not ignore Rotha's exceptional consideration of the topical newsreel as the only justifiable form of the dialogue film, she suggests that Rotha's general attitude to voice is that voice violates the visual characteristic of film.[238] In addition, Bruzzi engages with a similar view expressed by Drew that only non-voice-over documentaries function appropriately, as voice-over is just a didactic form preventing documentaries from being beyond exposition.[239] Furthermore, Bruzzi puts Kozloff's thought into the discussion. Kozloff suggests in *Invisible Storytellers: Voice-over Narration in American Fiction Film* that the reason why voice-over is usually treated as an unimaginative and incompetent tool is because film is mostly regarded as a visual medium, thereby undermining voice-over's aptitude for relaying expositional information.[240] Broadly speaking, all of the preceding ideas, from my perspective, reveal that if there are no vocal tracks, the visual quality of film will certainly not be subject to vocal interventions. However, the issue of whether "no vocal tracks" truly maintains a film's visual purity needs to be revisited, as I declare that a vocal potentiality cannot be fully avoided.

I want to employ causality-inference to reconsider vocal interventions. In the field of Buddhology, Bodhiruci states that "*yinyuan huiyu shi*, *guobao huanzi shou*." "因缘会遇时，果报还自受。" (when a cause and a condition come together, one accepts an effect).[241] This idea was brought up in sub-section 3.1.2, suggesting that "if a cause can produce an effect, there must be a

---

[236] Bruzzi, *New Documentary*, p. 47.
[237] Bruzzi, *New Documentary*, pp. 47–50.
[238] Bruzzi, *New Documentary*, pp. 47–48; Paul Rotha, *The Film till Now: A Survey of World Cinema* (New York: Twayne Publishers, Inc., 1949), pp. 406–10.
[239] Bruzzi, *New Documentary*, p. 48; Robert Drew, "Narration Can Be a Killer" (1983), in *Imagining Reality: The Faber Book of Documentary*, eds. Mark Cousins and Kevin Macdonald (London: Faber and Faber Ltd, 1996), pp. 271–73.
[240] Bruzzi, *New Documentary*, p. 48; Sarah Kozloff, *Invisible Storytellers: Voice-over Narration in American Fiction Film* (Berkeley, Los Angeles, and London: University of California Press, 1988), pp. 21–22.
[241] Bodhiruci, *Dabaoji jing* 大宝积经 [The Mahāratnakūṭa Sūtra] (It was translated into Chinese circa the sixth century as the *Dabaoji jing* by Bodhiruci, floruit 508–535), ed. Bodhiruci (Shanghai: Shanghai foxue shuju 上海：上海佛学书局 [Shanghai: Shanghai Buddhist Books], 2004), vol. 57, p. 450. My translation.

condition making this cause catalyse an effect. When a condition changes, aspects of deduction dependent on this condition also change."[242]

I point out that the condition of the foregoing similar attitudes to voices is that they all focus on the issue of voices from film; this attention-oriented condition makes the causality regarding the relation between voices and images appear as the following situation where vocal interventions in images are ascribed to voices from film. However, if this condition is shifted to the act of focussing on the oneness of film and audiences' experiences, vocal interventions in images will not be absolutely owing to voices from film. This implies that even if a film has no vocal tracks, voices from audiences' experiences still affect visual forms. Even film without vocal tracks cannot entirely separate from voices—the first clue as to the nominal region between voices and "no voices." In my subsequent analysis, I will explain what voices from audiences' experiences mean, and illustrate how these potential voices affect our way to understand vocal interventions. Within a Buddhist context, voice-related discussions can be traced back to the *Avataṃsaka Sūtra*. This Buddhist source brings up an argument that all verbal expressions can be considered a unity. More specifically, Buddhabhadra and Śikṣānanda declare:

> Just as the sunlight when it appears illumines all forms so that all with eyes can see them clearly, in the same way do great enlightening beings with knowledge understanding all speech enter into the masses of all verbal expressions and cause the intelligent to gain thorough understanding of all verbalisation.[243]

I argue that if this Buddhist idea is used to engage with voices in cinema, it can suggest that within the holistic context of film and audiences' experiences, all voices need to be regarded as a whole, rather than just notice voices from film. When the viewer is interpreting a film during viewing, they are making expository voices. Even if the viewer's comments are not spoken out, these latent expository voices still "exist" as inner voices in the viewer's mind. These potential voices are implicit until the viewer says them out loud. If the condition of the attitude to voices is based on the attention to voices from film and audiences' experiences, not just from film, we find out that when audiences sense a film, voices from the film are not the only vocal element that they sense, for they also perceive their own feelings and thoughts which can be conveyed in the form of interpretive voices. Even if these potential voices are not from the film, it is still a component of film experience.

---

[242] See p. 52.

[243] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), p. 866.

The potential of a voice cannot be completely avoided even if a voice is never actualised. This further suggests that there is no substantial dividing line between the voice and *no voice*. A nominal field between them is dependent on the changeable way of making inferences about a film without vocal tracks. If the viewer's interpretive voice is noticed, regardless of whether such a voice is spoken out, the effect will be that there is still a potentially vocal intervention in images. If the viewer's interpretive voice is neglected, the effect will be that there is no vocal intervention in images. Voice is there but also is not there. From a Buddhist perspective, the voice and "no voice" become a seamless whole, which makes the issue of a vocal intervention in images become even more complex.

Nominal voices do not mean to reject interventions made in Bruzzi, Rotha, Drew, and Kozloff's discussions of the distinction between film with vocal tracks and film without them, as their relative difference can still be recognised. The Buddhist view on vocal interventions in images has transcended their difference. In the realm of nominal voices, instead of just saying voices intervene in images, I argue that based on the compound context of film and audiences' experiences, both voices and "no voices" intervene in images, thereby making film experience not purely visual. However, if this holistic context is effaced, the difference between film with vocal tracks and film without them will become explicit, whilst nominal voice will become implicit.

4.3.2 The Hetero-homogeneity of Sounds

In addition to voices from audiences' experiences, an important consideration that influences audio perception is the surrounding environment that transmits sound. In "The Material Heterogeneity of Recorded Sound," Altman argues that the relation between audio perception and space reflects the changeability of recorded sound.[244] More specifically, Altman considers that though all film sounds relate to the criteria of loudness, pitch, and timbre, these abstract and conceptual criteria cannot make a particular sound maintain the same state always, since it is also affected and altered by the surrounding space. Once the "same" sound is emitted in different spaces, distinct vibrations, molecules, and pressures make the "same" sound become heterogeneous, which causes this sound to have multiple versions. For Altman, a certain sound is not a single, identical, and unidimensional acoustic representation, but a changeable appearance exposed to different spatial aspects.

From a Buddhist viewpoint, Altman's idea of the material heterogeneity of recorded sound suggests that sound cannot be truly grasped because it is subject to impermanence. Impermanence

[244] Altman, "The Material Heterogeneity of Recorded Sound," pp. 16–17.

reveals that "all created things do not remain as one wishes and are not under one's control."[245] Even if a sound can be recorded, it cannot be permanently maintained in a constant state because of the indeterminate effect of changeable space. As sound has no permanent nature, it has no inherent existence. Although the particular loudness, pitch, and timbre are able to produce a particular sound, it is not real. Its essence constantly flows, varying with distinct external factors. For example, if the viewer attends three screenings of the "same" sound film, one at the cinema, another at home, and a third outdoors, its acoustic attribute has been reconstructed three times. Each version is an audio appearance affected by space, rather than an independent and permanent object. Its quality is not inherent. What the viewer senses is not a pure sound, but a spatialised sound.

The idea of the material heterogeneity of recorded sound can also be used to re-examine the previously analysed voices from audiences' experiences. The material heterogeneity of voices from audiences' experiences signifies that even though the viewer's interpretive voice is recorded and preserved, its acoustic nature is still able to be reconfigured multiple times when it is replayed in distinct places. Even if an interpretive voice's content can be repeated in different spaces, an acoustic quality of this "same" voice is subject to impermanence. Uncertain spatial acoustic properties also cause the viewer's elucidatory voice to be heterogeneous.

However, under the Buddhist nominal system, the material heterogeneity and homogeneity of recorded sound can be regarded as one. Instead of moving away from sound as a single and homogeneous appearance, its homogeneous aspect, at a relative level, suggests that even if sound has no independent and permanent quality, this is not to say that sound does not exist somewhat, as the perceiver's mind can still produce the imaginary existence of sound, rather than nothingness. If the condition of the attitude to sound is based on the attention to its homogeneity, the viewer notices that they can still identify a certain sound even after it is replayed in distinct locations. This is because it can leave a general impression on the viewer. It is the rough impression that makes them familiar with this sound, which reveals that being replayed in distinct spaces does not mean that the sound has been fully distorted. It is not wholly changed into something else with which the viewer is entirely unfamiliar. Hence, instead of just recognising the heterogeneity of sound, it is also important to note that sound can retain a single and unidimensional state at a relative level.

When the condition of the attitude to sound alters back and forth between the attention to its heterogeneity and the attention to its homogeneity, a nominal field between them arises. Because of this nominal state, the "same" sound is not able to maintain a constant state but also is

[245] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 1110.

able to maintain it. It is neither completely homogeneous nor completely heterogeneous. This nominal state changes its homogeneity into heterogeneity, and vice versa. As for this conversion process, it is noteworthy that when Altman explores the heterogeneity of sound, he also engages with a certain conversion process, but his engagement with a conversion process is based on a physical situation where different spaces change the acoustic quality of the "same" sound. In this light, it appears that a conversion process were limited to heterogeneity. Yet, from the perspective of Buddhology, a process of change can be further applied beyond the heterogeneity of sound. My approach to conversion is based on the act of changing the focus of attention back and forth between its heterogeneity and homogeneity. I assert that my discussion of its homogeneity does not aim to reject its heterogeneity. Rather, a conversion process between its heterogeneity and homogeneity can arise only if both are considered. Just like heterogeneity, homogeneity also engages with conversion. Conversion does not necessarily signify the changeable impact of uncertain spatial acoustic properties. It can also connote the flowing process of changing back and forth between the heterogeneity and homogeneity of sound. If its homogeneity is not taken into account, even though its heterogeneity presents a conversion process where the "same" sound has been changed several times, its heterogeneity, from a wider perspective, will still be in a relatively immobile state. Without its homogeneous involvement, there is no varying process between them.

Instead of following Altman's dichotomous perspective on sound (i.e. the heterogeneity of sound pertains to change, whilst the homogeneity of sound does not), I concentrate on the hetero-homogeneous mixture of sound. With respect to this mixture, from a Buddhist perspective, it implies that its heterogeneity does not have sole access to conversion, as both its heterogeneity and homogeneity, to some extent, get involved in conversion. They are not isolated from each other.

A nominal state between the material heterogeneity and homogeneity of recorded sound can also be employed to engage with voices from audiences' experiences. To be more specific, when the viewer's interpretive voice is recorded and then replayed in different spaces, its acoustic attribute can be transformed by uncertain spatial acoustic properties multiple times. Nevertheless, if the condition of the perspective on the viewer's interpretive voice is dependent on the attention to its homogeneity, we become aware that the recorded interpretive voice can leave us with an overall vocal impression that makes us familiar with it and recognise it. The effect suggests that the state of the viewer's expository voice can be changed from self-evident variability into relative stability. Although uncertain spatial acoustic properties break the independent and permanent existence of the viewer's expository voice, it still has the relative existence. Its relative existence is attributed to the fact that even though it is replayed in different spaces, it can still be identified, rather than being totally disintegrated. When moving back and forth between its heterogeneity and

homogeneity, the viewer's interpretive voice engages with a nominal field where it is not the "same" voice but also is the "same" voice.

4.3.3 Vocal Nominality in Six Documentary Modes

Apart from the above general Buddhist system of voices, I intend to carry out a close analysis of specific types of voices. In *Introduction to Documentary*, Nichols discusses the vocal issues of six documentary modes: the poetic mode, the expository mode, the observational mode, the participatory mode, the reflexive mode, and the performative mode.[246] Although Nichols explains how each of documentary modes' vocal facets plays its own distinctive role, these facets can be considered alike, for all of them are related to nominality. Nominality within the context of voices will be discussed after a brief summary of Nichols' vocal examination of six documentary modes. Although his methodology helps readers understand the distinct aesthetic traits of voices, I argue that their shared nominal trait also needs to be considered.

The poetic mode does not give prominence to continuity editing and coherent narrative. Rather, it uses a series of spatiotemporal fragments to express rhythm, subjectivity, and looseness. The role of the voice is to resonate with these spatiotemporal fragments. Nichols selects Péter Forgács' *Free Fall* (1998) and *The Danube Exodus* (1998) as examples to argue that the voices in these films are used to recite diary entries in order to create emotional tempo, rather than merely describe the social situation in World War II. These voices possess lyric emotion.[247]

The expository mode typically engages with a voice of God/voice of authority—a voice that addresses audiences directly in a didactic way. Instead of stressing an artistically poetic quality, the voice in the expository mode emphasises logic, puts forward arguments, recounts history, and conveys the information behind moving images—a way to deal with objectivity and omniscience. In terms of the difference between voice of God and voice of authority, Nichols uses several case studies to clarify this issue. Films like *The City* (Ralph Steiner and Willard Van Dyke, 1939) and *Dead Birds* (Robert Gardner, 1963) employ a voice of God technique—a voice emitted by an invisible narrator. Films like *16 in Webster Groves* (Arthur Barron, 1966) and *The Selling of the Pentagon* (Peter Davis, 1971) are related to voice of authority—a voice emitted by a visible narrator.[248]

The observational mode focuses on the act of observing behaviour. Instead of using voice-over commentary and narration to provide further information for audiences, the observational mode aims to keep recorded appearances the way they are. In order not to control audiences' thoughts

[246] Nichols, *Introduction to Documentary*, pp. 99–138.
[247] Nichols, *Introduction to Documentary*, pp. 102–5.
[248] Nichols, *Introduction to Documentary*, pp. 105–9.

on what they perceive, voice-over commentary and narration needs to be weakened. To illustrate this issue, Nichols refers to examples including *Les raquetteurs* (Michel Brault and Gilles Groulx, 1958) and *High School* (Frederick Wiseman, 1968) which aim to tempt audiences to actively make their own inferences about the behaviour that they witness in the films, rather than make use of authoritative voice-over to manipulate audiences.[249]

The participatory mode accentuates the voice of the filmmaker. Audiences are able to hear the voice of the filmmaker within film. The participatory mode evades anonymous voice-over, stressing the filmmaker's presence. Nichols states that films such as *Chronicle of a Summer* (Jean Rouch and Edgar Morin, 1961) and *The Sorrow and the Pity* (Marcel Ophuls, 1969) involve the filmmaker's vocal impact.[250]

The reflexive mode calls into question cinema's ability to show the truth. With regard to the vocal aspect, Nichols chooses Michelle Citron's *Daughter Rite* (1980) as an example and argues that this film seems to deal with two sisters in real life, but in fact, they are played by two professional actresses. When they recount their relationship with their mother, their discussion is not based on their own personal experiences, but rather based on the insights derived from multiple prepared interviews with other people. That is to say, Citron utilises the vocal contents of other interviewees as materials and makes these contents to be recounted by two actresses on camera, so that two actresses possess synthetic experiences that enable them to represent other interviewees. The film appears to truly present the personal life of these two particular sisters and their mother, but it puts the interviewees' vocal information into a re-enactment-oriented context.[251]

The performative mode is interwoven with personal and evocative expression. As for the voice in the performative mode, Nichols selects Alain Resnais' *Night and Fog* (1956) as a case study and argues that the evocative tone of the voice-over commentary in the film is more striking than the historical account of the context of Nazi concentration camps. The haunting voice-over enables audiences to suffer the painful experience that the narrator recalls. This vocal power has transcended the historical evidence.[252]

I argue that from a Buddhist perspective, even if the above vocal aspects are different from each other, all of them engage with nominality. As for the voice in the poetic mode, Nichols takes Forgács' vocal expression as an example to argue that the reminiscent voice that recites diary entries is distinct from the descriptive narration of history, as the former is more emotional than

[249] Nichols, *Introduction to Documentary*, pp. 109–15.
[250] Nichols, *Introduction to Documentary*, pp. 115–24.
[251] Nichols, *Introduction to Documentary*, pp. 125–30.
[252] Nichols, *Introduction to Documentary*, pp. 130–37.

the latter. However, I declare that they are alike because both engage with *benxing xin* 本性心 (the luminous mind). This Buddhist idea can be traced back to the *Abhidharma-nyāyānusāra-śāstra* (also known as the *Conformity to Correct Principle*, written by Saṃghabhadra in the fifth century, translated into Chinese in 654 as the *Apidamo shunzhengli lun* 阿毗达磨顺正理论 by Xuanzang, circa 602–664). On the whole, the luminous mind, according to Saṃghabhadra's research, can be defined as follows:

> *Benxing xin zhe*, *wei wuji xin*, *fei qi fei xin renyun zhuanwei*, *zhuyou qing lei duo zhu ci xin*, *yiqie wei zhong jie rongyou gu*, *ci xin bi jing*, *fei ranwu gu*. 本性心者，谓无记心，非戚非欣任运转位，诸有情类多住此心，一切位中皆容有故，此心必净，非染污故。 (The luminous mind, also known as the indeterminate mind, is a condition where there is no taint of emotions such as joy and sorrow. Most of the time, our minds are not tainted [affected] by emotions, being in a condition of purification).[253]

A critical detail of this definition needs to be clarified further. The luminous mind is not meant to exclude emotions. Rather, it shows a primordial condition of emotions – or, in other words, a state that comes before emotions. Saṃghabhadra suggests that before a certain emotion appears, the mind is in a state of purification (luminescence)—a state of not being affected by emotion. When an emotion gets into the luminous mind, the emotion transpires, thereby defiling the luminous mind and making its luminescence implicit. However, whether or not there are emotional interventions, the luminous mind—a precondition for emotions—always functions. Even if an emotion taints the mind's purification, it does not completely eradicate the luminous mind itself. It has been coupled to this pre-emotional state in the first place. It merely forms an emotional layer over the luminous mind, rather than totally destroys the luminous mind. It is critical to note that the luminous mind is always present even if emotions cover it.

The luminous mind can also be employed to delve into the ontological relationship between the emotional voice that reads diary entries and the expository voice that narrates history. The distinction between these two voices, from a Buddhist perspective, is based on different states of the operation of the mind – or, in other words, no taint of emotions and a taint of emotions. However, both engage with the luminous mind as a pre-emotional state. On the one hand, the elucidatory voice is directly related to the luminous mind since this kind of voice has no emotions. On the other hand, the lyric voice that recites diary entries has emotions, but it needs to go through

[253] Saṃghabhadra, *Apidamo shunzhengli lun* 阿毗达磨顺正理论 [The Conformity to Correct Principle] (It was completed by Saṃghabhadra in the fifth century, translated into Chinese in 654 as the *Apidamo shunzhengli lun* by Xuanzang, circa 602–664), edited by Cangjing shuyuan 藏经书院 [Buddhist Texts Academy] in the *Wan zhengzang jing* 卍正藏经 [The Wan Tripiṭaka] (Taipei: Xinwenfeng chuban gufen youxian gongsi 台北：新文丰出版股份有限公司 [Taipei: Xinwenfeng Publishing Company], 1980), vol. 72, p. 1011. My translation.

the luminous mind in the first place. Without an originally purified condition that comes before emotions, emotions cannot emerge. This reveals that the poetic voice is in fact indirectly related to the luminous mind. In summary, these two voices become one because of the shared luminous mind, but their difference can still be maintained at a relative level at the same time. An intermediate state between their oneness and distinction brings about nominality.

There are alternative approaches to engaging with a nominal domain between vocal oneness and difference. For example, in terms of the voice in the expository mode discussed earlier, Nichols examines the difference between the voice of God (a voice emitted by an invisible narrator) and the voice of authority (a voice emitted by a visible narrator). I believe that if we want to excavate within the expository voice a nominal region between vocal oneness and difference, we also need to be aware of the oneness between the voice of God and the voice of authority.

In terms of their oneness, in film studies, reference to Christian Metz's account of aural objects is useful because he suggests that there is no fundamental dividing line between an off-screen voice and an on-screen voice.[254] More specifically, Metz argues:

> In a film a sound is considered "off" (literally off the screen) when in fact it is the sound's source that is off the screen, therefore an "off-screen voice" is defined as one which belongs to a character who does not appear (visually) on the screen. We tend to forget that a sound in itself is never "off": either it is audible or it doesn't exist. When it exists, it could not possibly be situated within the interior of the rectangle or outside of it, since the nature of sounds is to diffuse themselves more or less into the entire surrounding space: sound is simultaneously "in" the screen, in front, behind, around, and throughout the entire movie theatre.[255]

I suggest that if this statement is put into dialogue with Nichols' analysis of the voice in the expository mode, the distinction between the voice of God and the voice of authority becomes uncertain. For the same reason, neither voice is anchored in a particular location. They spread throughout the entire space. Their oneness is not interrupted by the frame of the screen. Thus, "of God" and "of authority" can be bracketed out. Both are commentary-related voices. I argue that one of the vocal ideas raised in the *Avataṃsaka Sūtra* is similar to Metz's account, but also resonant with Nichols' stance to some degree:

> Voice does not emerge from inside, outside, or both, yet it is possible to produce skilful expressions to accomplish explanation. Like echoes in a valley, it is produced by conditions and is not at

[254] Although Christian Metz's view could trigger an analysis of film semiotics, as he is best known for opening up film semiotics, I do not put voice into the context of film semiotics.
[255] Christian Metz, "Aural Objects" (1975), trans. Georgia Gurrieri, in *Cinema/Sound, Yale French Studies*, ed. Rick Altman (New Haven: Yale French Studies, 1980), no. 60, pp. 28–29.

> variance with the nature of things. It enables sentient beings each to understand according to kind and to be able to practice this learning.[256]

I contend that "voice does not emerge from inside, outside, or both" is akin to Metz's account, as Metz suggests that voice is not anchored in on-screen or off-screen space, but is pervasive in the whole space—no partition between voices from inside and outside. Voices have transcended the man-made ideas of inside and outside. The rest of the above Buddhist contention suggests that if one intends to express and explain voices' transcendence of the boundary between inside and outside, the terms "inside" and "outside" need to be mentioned in the first place—the oblique acknowledgement of the relative existence of inside and outside. That is to say, if the boundary between inside and outside is thorough nothingness, we are not able to understand what is transcended by voices. I argue that this Buddhist aspect, in part, corresponds to Nichols' idea of the boundary between the voice of God (off-screen narration) and the voice of authority (on-screen narration). From a holistic perspective, Metz's oneness-based perspective and Nichols' difference-based perspective can be considered two sides of the same coin. When the oneness and difference between the voice of God and the voice of authority are placed on the same equalising plane, nominality unfolds. Their oneness is ascribed to the attention to voice's diffusion-based trait, whilst their distinction is ascribed to the attention to voices' vision-based sources. By moving between these two states, we reach their oneness and distinction at the same time—a nominal (superimposing) effect.

In addition to nominality within the context of the expository voice, the vocal issue of the observational mode also gets involved in a nominal thread. Nichols states that by avoiding voice-over, audiences are able to use their own abilities to observe and understand recorded behaviour, and to come to their own conclusions without needing additional information given by voice-over. I suggest that even if there is no voice-over to manipulate audiences' observation, their film experiences are still exposed to a form of control. This control arises from a situation where different audiences' interpretive voices, to some extent, can still affect each other's observation and thinking.

Previously, I discussed voices from audiences' experiences, an idea built on the Buddhist statement that "all speech enter into the masses of all verbal expressions."[257] This Buddhist perspective emphasises that all vocal elements should be considered a whole. In the same way, regarding voices in cinema, all voices can be regarded as a unity. When audiences observe a film,

---

[256] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 872.
[257] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 866.

even without vocal interventions from film, audiences' interpretive voices can still create interventions. I argue that this consideration can be used to re-examine the vocal issue of the observational mode.

If the viewer does not understand the behaviour presented in a film, they may want to hear how other viewers observe and interpret such behaviour. Just because there is no voice-over to manipulate the viewer's observation does not mean that the viewer can totally use their own ability to engage with the behaviour. The viewer's perception can be built on other viewers' interpretive voices and thus can be partly controlled by others, even without being controlled by the voice-over from the film.

On the one hand, since voice-over is capable of addressing audiences directly, their observation is controlled by voice-over. By avoiding voice-over, audiences can draw their own conclusions without depending heavily on the information conveyed by voice-over. On the other hand, since multiple audiences' interpretive voices are still able to affect a certain viewer's perception, one's observation can be controlled by other viewers, even though a film has no voice-over. When the cause alters, its effect alters—an inferential conversion from uncontrolled observation to controlled observation. Nominality between these two observational states occurs.

Apart from the observational mode, a nominal thread can also be applied to the voice in the participatory mode. As discussed earlier, for Nichols, the voice of the filmmaker within film is distinct from anonymous voice-over, for the former is able to remind audiences that the filmmaker actively participates in, rather than unobtrusively observes the pro-filmic world. However, the reason why the voice of the filmmaker can be affirmed is because audiences can see the filmmaker appear and talk on camera. In other words, audiences identify it not by virtue of the voice itself, but by virtue of the filmmaker (the voice's visual source). Alternatively, even if the filmmaker is off the screen, the vocal content can also imply that the voice is made by the filmmaker. For example, the invisible filmmaker can say something like "I am the filmmaker of this film…" to remind audiences of the covert presence of the filmmaker in film. The underlying problem is that if audiences focus just on the voice itself without focussing on its visual source and content and if they are not familiar with the filmmaker's voice, it may be hard for audiences to differentiate it from anonymous voice-over. In other words, they may not know whether the voice that they hear is made by the filmmaker. This further suggests that at the level of the voice itself, the difference between the voice of the filmmaker and anonymous voice-over is not absolute.[258]

[258] The word "filmmaker" is not deployed to examine auteur theory. Rather, I want to point out that due to the influence of audiences' experiences, the voice of the filmmaker can be transformed into anonymous voice-over. Its substantial existence has been deconstructed.

The above discussion does not repudiate Nichols' statement on the difference between the voice of the filmmaker and anonymous voice-over. It is crucial to note that at the Buddhist level of the variability of causality, though not all filmmakers' voices can be recognised at a solely vocal level, anonymous voice-overs as they were, those unidentified filmmakers' voices can be changed into the identified if the condition of audiences' experiences changes. When audiences not only concentrate on the voice itself but also shift part of their attention to the voice's visual source or content, the implicit filmmaker becomes explicit. As there is an explicit correspondence between the voice and the filmmaker, audiences are able to deduce that this particular voice is obviously made by the filmmaker, rather than by an unknown narrator. When audiences focus exclusively on the unfamiliar filmmaker's voice itself, the filmmaker can only be in an implicit state. Since there is no direct clue as to the connection between the voice and the unfamiliar filmmaker, it makes no difference to audiences whether it is the filmmaker's voice or anonymous voice-over. Owing to the variability of causality, the voice of the filmmaker is distinct but also not distinct from anonymous voice-over—alternative access to nominality.

With respect to the voice in the reflexive mode, as summarised previously, Nichols selects Citron's *Daughter Rite* as an example to suggest that the expression of two performing interviewees is based on the vocal contents of prepared interviews with other real-life interviewees, rather than based on their own personal experiences. This film seems to record two real-life sisters, but they are professional actresses playing sisters. However, I consider that the impression of access to reality is not completely deconstructed even if these two characters are not actual interviewees. Their vocal expression still reflects other people's real-life experiences. Therefore, to some extent, their staged performance is still linked to the real. The real life-based vocal intervention gives rise to an intermediate state between the real and the virtual—nominality that arises.

In terms of the voice in the performative mode (personal and evocative voice), I argue that a nominal perspective on the poetic voice can be directly applied to the performative voice because as dissected earlier, the performative voice also deals with emotional issues. Similarly, if we use the previously discussed Buddhist notion of the luminous mind to rethink the performative voice, we will arrive at a similar conclusion that there is no inherent partition between the performative voice that creates a haunting atmosphere and the unemotional voice that narrates historical evidence. This conclusion is based on the fact that both engage with the luminous mind as a pre-emotional state.[259] By moving between their oneness and distinction, nominality arises. Because all of the six documentary modes' vocal facets get involved in nominality, they are one. Vocal oneness does not conflict with Nichols' vocal difference-based perspective. Instead, they buttress each other.

[259] See pp. 91–92.

4.4 *Through the Olive Trees* and Buddhology

4.4.1 Conventional Perspectives on Vocal Echoes and the Cube of Sensation

Before dissecting nominal voices in *Through the Olive Trees* (1994) at the Buddhist level of the mind, I would like to first summarise some sources that include voice-related discussions of this film. Laura Mulvey offers a brief description of the echoes of the voices in a slope scene where Mohammad-Ali Keshavarz (played by himself) and Farhad Kheradmand (played by himself) talk and walk together (figure 11). In this scene, Kheradmand shouts to an opposite hill several times, thereby creating an echo. However, Mulvey suggests that the vocal echoes, at a deeper level, are not simply delayed aural phenomena, but haunted "voices from the past echoing across time and the boundary between the living and the dead."[260] From my perspective, though Mulvey touches upon the voices, her analysis is based on the content of the voices. The reason why Mulvey asserts that the vocal echoes have haunted forms is not because the vocal echoes themselves have haunted forms, but because Keshavarz tells Kheradmand about "the echo that comes from the ruined village and the legend that its former inhabitants mysteriously answer a shouted greeting."[261] This reveals that without this linguistic context, Mulvey cannot know that the vocal echoes themselves possess ghostly entities. I will rethink the voices and echoes without relying on the vocal content.

Figure 11 The voices' echoes in *Through the Olive Trees*

Another noteworthy example is the interview-based article "Abbas Kiarostami" in which Kiarostami offers a concise account of how he understands sounds in his own films.[262] Although this article does not engage with the word "voice," Kiarostami's discussion of sounds can still be applied to voices. He argues:

[260] Laura Mulvey, *Death 24x a Second: Stillness and the Moving Image* (London: Reaktion Books Ltd, 2006), p. 132.
[261] Mulvey, *Death 24x a Second: Stillness and the Moving Image*, pp. 131–32.
[262] Akram Zaatari and Abbas Kiarostami, "Abbas Kiarostami" (1995), in *Bomb*, eds. Betsy Sussler, Andrew Bourne, and Sabine Russ (New York: New Art Publications, Inc., 1995), no. 50, p. 14.

> For some directors the significance of sound is more important than the visual. When we go out to shoot, sometimes people ask the crew where they're going and they say, "We're just going to record some sound, but we're taking a cinematographer with us, just in case." If you just concentrate on the visual, you would be dealing with only one side of the cube. Sometimes we put so much emphasis on our shot, it's as if we're telling the world, "Shut up, the picture is so important!" But if you look at you and me sitting here talking, there are all these noises around us. That's an important part of reality.[263]

The cube reveals that what audiences perceive is not just the two-dimensional screen, but the screen that implies the three-dimensional filmic world. In *Through the Olive Trees*, there are abundant off-screen voices. Just as noises around us are still present in the other sides of the cube, its other sides suggest the presence of off-screen voices. The picture on the front face of the cube is just a part of reality, rather than the whole of reality. The idea of the cube stresses that what audiences sense is beyond the screen. Aural phenomena can extend the screen's dimension.

Although the metaphor of the cube, to some extent, can be used to reveal the extension of the screen's dimension, the edges of the cube still keep off-screen and on-screen voices separate from each other. They are not present in the same side. However, based on the previously mentioned Metz's idea of audio phenomena, there is no partition between an off-screen voice and an on-screen voice because the audio is not interrupted by the frame of the screen. I argue that if Metz's contention is put into dialogue with Kiarostami's idea of the cube, we realise that what Metz does is to unfold the cube into the plane—the plane of oneness. All voices coexist in this integrated plane. No edges distinguish them. However, I suggest that not being interrupted by the frame of the screen is not the only consideration that merges off-screen and on-screen voices into one. I will adopt a Buddhist approach to putting forward an alternative path to their oneness in *Through the Olive Trees*.

4.4.2 The Buddhist Experiential Process of Vocal Experience

I apply causality-inference to the discussion of nominal voices in *Through the Olive Trees*. My analysis includes four parts: the opening voices, off-screen voices, vocal echoes, and ending voices. All voices can be connected to a nominal system. That is to say, voices do not exist, but also exist.

In the opening scene (figure 12) of this film, a man (played by Mohammad-Ali Keshavarz) speaks straight to the movie camera. Suddenly, a woman (played by Zarifeh Shiva) emerges from a crowd of people in the background, walks to Keshavarz hurriedly, and interrupts him. He stops talking and then starts listening to what Shiva says. The acoustic qualities of these two voices are

[263] Zaatari and Kiarostami, "Abbas Kiarostami," p. 14.

apparently different since audiences are able to recognise that one is Keshavarz's voice and another is Shiva's voice. In other words, these two different voices seem to have their respective independent entities.

Figure 12 The opening voices in *Through the Olive Trees*

However, from a Buddhist perspective, these two voices can be considered alike. Before discussing why they can become one at a Buddhist ontological level, I would like to point out that Buddhology has no sole access to aural oneness. Within Western film theoretical traditions, Altman has mentioned why all recorded sounds, to some degree, can be considered alike. His approach to aural oneness is based on the level of physical space, whilst a Buddhist approach to aural oneness is based on the level of the viewer's mind. Altman asserts:

> For what the record contains is not the sound event as such but a record of a particular hearing, a specific version of the story of the sound event. Every recording is thus signed, as it were, with the mark of the particular circumstances in which it was heard. […] Every recording carries the elements of this *spatial signature*, carried in the audible signs of each hearing's particularities. Even when those signs are contradictory or have been tampered with, even when they seem not to match the visual data provided with the sound record, they still carry information that is narrative and spatial in nature.[264]

This idea implies that all recorded sounds are documentary, as all of them are coupled to the spatial conditions of the recording. Even if each spatial situation of the recording has a unique material quality, all recorded sounds interconnect with physical space as a whole—commonality that activates the oneness of sounds.

Although Buddhology also provides access to the oneness of sounds, it is not based on the level of the spatial conditions of the recording. Rather, it illuminates aural oneness at the level of the viewer's mind. A prime example of a Buddhist analysis of aural oneness is Alexander Berzin's

[264] Altman, "The Material Heterogeneity of Recorded Sound," p. 24 (Note: *Spatial signature* in this quotation is Altman's italic emphasis).

reconsideration of Śāntideva's guidance about Buddhist meditation. Before scrutinising how Berzin employs Śāntideva's guidance to discuss the aural, I would like to first look at Śāntideva's primordial intention to give meditation guidance. Originally, Śāntideva's guidance does not centre around the aural but happiness, which can be traced back to the *Bodhisattvacaryāvatāra* (also known as the *Engaging in Bodhisattva Behaviour*, written by Śāntideva circa 700, translated into Chinese in 985 as the *Puti xing jing* 菩提行经 by Devaśāntika, floruit 980–1000, and re-translated into Chinese in 1997 as the *Ru pusa xing lun* 入菩萨行论 by Rushi). Śāntideva argues:

> Ru tong zhu xile, suo wei zhong shanye, xin ying ji danzhe, le bi wu yanzu. Shiren qin qiule, chengfou you weiding. Erli neng dele, buxing le heyou? 如童逐戏乐，所为众善业，心应极耽著，乐彼无餍足。世人勤求乐，成否犹未定。二利能得乐，不行乐何有？ (Like someone wishing for happiness as the result of play, any [positive] actions [a Bodhisattva's] engaged in. He clings to those actions and delights in those actions, never having enough. Although people do actions for the sake of happiness, it's not clear that they'll become happy or not. But for [a Bodhisattva] whose actions in fact bring happiness, how can he be happy without doing those actions).[265]

As Śāntideva suggests in this statement, the reason why a Bodhisattva intends to change the focus of his attention from happiness as a reward or as a benefit to the action as the cognitive process is because a Bodhisattva considers that the action itself has already made him cheerful. Just as children immerse themselves in play, asking nothing in return, a Bodhisattva also enjoys the action (process) of helping others. It is the action itself that makes him happy, rather than a reward. If we are desperate to obtain a reward, we will be disappointed if there is nothing in return.

Berzin reuses Śāntideva's statement to discuss Buddhist meditation within the context of hearing. Specifically, Berzin uses the sound of traffic and the sound of the ocean as examples to discuss the ontological relationship between them. They are very different on the surface. The loud noise of traffic is annoying, making people upset, whilst the sound of the ocean makes people feel relaxed. In Berzin's view, we can reuse Śāntideva's meditation advice (focus on the action [the cognitive process]) to dispel upset caused by the loud noise of traffic. Berzin argues:

> By shifting the focus of our attention from dwelling morbidly on the noise itself, to the cognitive process that is occurring of merely the arising of a sound and the hearing of it, we realise that the arising of the noise of traffic is the arising of just another sound, and the hearing of it is just another experience of hearing. There is nothing more. With such shift of focus, we subjectively experience the same event of hearing the traffic in a totally different qualitative manner. Our experience of

---

[265] Śāntideva, *Engaging in Bodhisattva Behaviour* [Ru pusa xing lun 入菩萨行论] (It was completed by Śāntideva circa 700, translated into Chinese in 985 as the *Puti xing jing* 菩提行经 by Devaśāntika, floruit 980–1000, and re-translated into Chinese in 1997 as the *Ru pusa xing lun* by Rushi), trans. Alexander Berzin (http://www.buddhistische-gesellschaft-berlin.de/downloads/bca.pdf, 2005), p. 65.

> hearing the noise can now be accompanied with indifference, peace of mind or even happiness, instead of anger, unhappiness and self-pity.[266]

Despite the fact that this sound-related statement eventually serves to train us to avoid getting caught in unhappiness (to practise Buddhist meditation), I plan to focus just on the act of hearing and to use this action to engage with Keshavarz and Shiva's voices. When the cause is that audiences try to identify and recognise what they hear, the effect will be that these two voices are different. One is Keshavarz's voice and another is Shiva's voice. However, when the cause changes, when audiences shift the focus of their attention from these two voices to the act of hearing these two voices, the effect changes from their difference to their oneness. From the point of view of hearing, it is just hearing. Both hearing Keshavarz's voice and hearing Shiva's voice suggest hearing as an action. That is to say, both pertain to the experiential side of vocal experience. The Buddhist perspective on audio experience reveals that audiences can remain with the process of "here-and-now"—a situation where there is just the experiential process of the mere arising and engaging with what they hear occurring at each instant, without recognising whether the voice is made by Keshavarz or Shiva. From the point of view of hearing, there is no radical distinction between hearing Keshavarz's voice and hearing Shiva's voice. To put it another way, they have no absolutely independent entities that enable themselves to be truly different from each other. The shift of audiences' attention can change the way they perceive Keshavarz and Shiva's voices. These two voices can become one at the Buddhist level of hearing.

Regarding examinations of how the shift of focus adjusts audiences' aural perceptions, I would like to note that though in film studies, Michel Chion also puts forward a similar idea, his approach is distinct from a Buddhist examination of aural perceptions. Chion advocates the masking method, which means that we can "screen a given sequence several times, sometimes watching sound and image together, sometimes masking the image, sometimes cutting out the sound."[267] He suggests that different masking procedures can lead to different foci of attention, thereby transforming the viewer's audio sensation. However, a Buddhist perspective on the variations of audio experience is not based on a technical measure to alter the way the same sequence is screened, but rather based on the viewer whose focus has changed in the first place. Changing the way the opening scene of *Through the Olive Trees* is screened is not necessarily a precondition for changing the way the

[266] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997), p. 70.
[267] Michel Chion, *Audio-vision: Sound on Screen* (1990), ed. and trans. Claudia Gorbman (New York: Columbia University Press, 1994), p. 187.

viewer perceives Keshavarz and Shiva's voices. Even if the masking procedure is not used to intervene in the screening of the opening scene, the perception of their voices can still change.

Berzin further asserts that staying with the process of experiencing phenomena does not mean to be dissociated absolutely from each phenomenon's particularity because otherwise we completely ignore everything.[268] Aural oneness does not always maintain a permanent state. Regarding Keshavarz and Shiva's voices, when audiences do not plan to stay with the process of hearing these two voices but plan to discern these two voices, their difference will be emphasised. Berzin believes that though staying with the process of perceiving phenomena is an embodiment of Buddhist insight, it is merely a facet of the focus of attention, not being a fixed situation. If another aspect—the unique qualities of phenomena—is entirely disregarded, then Buddhist cognition will be incomplete. I point out that when shifting back and forth between their voices and the act of hearing their voices, nominal voices arise. This suggests that these two voices have relatively independent entities that allow them to be distinct from each other, but simultaneously, they do not have inherently independent entities that allow them to be fundamentally distinct from each other. An intermediate state between these two facets reveals that depending on the variable quality of the focus of audiences' attention, Keshavarz and Shiva's voices are one but not one.

4.4.3 Voices and the Void

Throughout *Through the Olive Trees*, one of the most conspicuous features is the use of off-screen voices. A noteworthy example of the use of off-screen voices is an acting setting where Tahereh Ladanian (played by herself) and Hossein Rezai (played by himself) converse in a damaged adobe house (figure 13). Ladanian is upstairs (off-screen), whilst Rezai is downstairs (on-screen). Although Rezai sometimes stays in the upstairs room, most of the time he is downstairs, emitting an on-screen voice, and Ladanian is upstairs, making an off-screen voice.

Figure 13 The off-screen voices in *Through the Olive Trees*

[268] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 68.

I previously described how Kiarostami analyses sounds in his own films. He borrows a metaphor from the cube to suggest that the screen divulges the three-dimensional filmic world of which the front side is not the totality of reality. Its other sides suggest the presence of off-screen sounds, which is also a vital component of reality. In my view, Kiarostami's idea about a resemblance between the filmic world and the cube, to some degree, if his notion is applied to Ladanian's off-screen voice and Rezai's on-screen voice, divides these two voices since the cube has edges. That is to say, these two voices are not manifested on the same side of the cube. However, based on Metz's perspective on audio phenomena discussed earlier, if Metz's idea is put into dialogue with Kiarostami's account of the cube, one could argue that the cube can be unfolded into the plane of oneness where there is no partition between these two voices, as the vocal transmission is not blocked by the visible edges of the cube. There is no barrier between their voices. They coexist in the entire surrounding space.

However, diffusing into the entire surrounding space is not the only consideration that integrates Ladanian's off-screen voice and Rezai's on-screen voice into one. Buddhology provides an alternative way to merge them into one. It is worth noting that though Buddhology also puts audio phenomena into dialogue with space, the Buddhist use of space is not identical to Metz's. I intend to examine the Buddhist spatial relationship between Ladanian's off-screen voice and Rezai's on-screen voice.

In some contexts, the Buddhist engagement with the term "space" divulges its comparable relationship with a Buddhist idea called "void" or "emptiness," with a notable instance of the usage of this concept being in the *Saṃyukta Āgama* (also known as the *Mixed Āgama*, translated into Chinese sometime between 435–443 as the *Za ahan jing* 杂阿含经 by Guṇabhadra, 394–468). Regarding the void, this Buddhist source suggests that "[…] *wuyu xi xukong*, […] *feiyi hehe zhe*." "【……】 五欲悉虛空， 【……】 非义和合者。" ([…] all five desires [desires caused by five physical forms—sight, sound, smell, taste, and touch] are the void, […] without holding on to composite illusions [composite illusions imply that all phenomena are actually not inherent because they have no independent and permanent existence]).[269] In order to further illustrate the void in a more figurative sense, Buddhist scholars such as the Dalai Lama and Chen Bing compare the void with space, though not the same as space in physics. The Dalai Lama argues:

[269] Guṇabhadra, *Za ahan jing* 杂阿含经 [The Mixed Āgama] (It was translated into Chinese sometime between 435–443 as the *Za ahan jing* by Guṇabhadra, 394–468) (Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], 1995), vol. 48, pp. 314–15. My translation.

> Just as space is the lack of anything tangible or obstructive on the side of some material object that could impede its physical existence in three dimensions, voidness is the total lack of impossible modes of existence that could impede the conventional existence of any phenomenon.[270]

In addition, according to Chen's reinterpretation of "*xukong shi se*, *suiru sezhong*" "虚空是色，随入色种" (the void is form, being in form), Chen also employs space to further illustrate this statement.[271] Before elaborating on how Chen engages with the ontological connection between audio phenomena and space, I would like to first discuss the meaning of *se* 色 in this statement.[272] *Se* 色 has two meanings. First, it refers specifically to *sechen* 色尘 (the physical form of sight). Second, it refers broadly to *seyun* 色蕴 (the aggregation of the physical forms including sight, sound, smell, taste, and touch). In this particular statement, *se* 色 fits with its second meaning. Furthermore, Chen suggests that the void can be considered unimpeded and unobstructed space. Each form in space is a composite illusion. If illusions are erased, void space will be showcased, not obscured by so-called substantial phenomena. Although the void does not possess a material quality, it can be viewed as a state of material form because it permeates material form.[273] I will focus on the physical form of sound and consider how Chen's discussion of void space engages with Ladanian's off-screen voice and Rezai's on-screen voice.

Before putting their voices into dialogue with Buddhist void space, I would like to mention Chion's discussion of the disintegration of the quality of an off-screen voice. He analyses de-acousmatising a voice—a process making the existence of an acousmatic voice disintegrate (a voice that is heard without its cause or source being seen; in a film, an off-screen voice is acousmatic). In his analysis, the reason why the existence of an off-screen voice is subject to the possibility of vanishing is because a character who emits an off-screen voice might appear in the visual field at any moment. The identity of the character might be revealed, thereby converting an off-screen voice into an on-screen voice, which causes the entity of an off-screen voice to vanish.[274] However, based on Buddhist void space, I argue that even if there is no process of de-acousmatising a voice, the entity of an off-screen voice can still fall apart, and the same goes for an on-screen voice.

---

[270] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 325.
[271] Chen Bing, *Fojiao xinlixue* 佛教心理学 [Buddhist Psychology] (Xi'an: Shaanxi shifan daxue chubanzongshe 西安：陕西师范大学出版总社 [Xi'an: Shaanxi Normal University General Publishing House], 2018), vol. 1, p. 228. My translation; Guṇabhadra, *Lengqie jing* 楞伽经 [The Scripture of the Descent into Laṅkā] (It was translated into Chinese in 443 as the *Lengqie jing* by Guṇabhadra, 394–468. The commentary was completed in 1196 by Zhengshou, 1145–1208), ed. Puming (Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], 2018), vol. 1, p. 42. My translation.
[272] Although Chen does not scrutinise sound in his interpretation of Guṇabhadra's ideas of the void and form, Chen's review still relates to sound, for it is one of the physical forms.
[273] Chen, *Fojiao xinlixue*, p. 228.
[274] Michel Chion, *The Voice in Cinema* (1982), ed. and trans. Claudia Gorbman (New York: Columbia University Press, 1999), pp. 18, 22, 27–29.

In my analysis, Both Ladanian's off-screen voice and Rezai's on-screen voice are devoid of existing in an independent and inherent way. The existence of these two voices needs to depend on some components: Ladanian and Rezai as the acoustic sources that make them and on the surrounding environment as a physical medium that transmits them. If their voices possessed independent and inherent entities, then even without these components, their voices should still arise. Therefore, neither has absolute existence. Based on Chen's analysis of the void and form, I suggest that the composite illusions of their voices impede and obstruct the appearance of Buddhist void space. Even if their voices seem not to block and occupy space due to their invisibly physical properties, their relative existence reveals that they are still tangible because they can vibrate audiences' eardrums. If audiences do not want to shift the focus of their attention from these two voices to a state where the composite vocal illusions are devoid of existing in a truly inherent manner, then the voices will still "exist" as invisibly tangible forms, transmitted throughout space, and taking up space.

I maintain that when audiences focus on the relatively independent entities of Ladanian and Rezai's voices, their respective vocal attributes will be highlighted. For instance, their vocal differences in loudness, pitch, and timbre will be manifested, which produces an effect that each of their voices can produce a particular vocal fluctuation that fills space. Although Metz's previously discussed perspective on aural phenomena reaches vocal oneness by virtue of dissolving the boundary between off-screen and on-screen voices, this oneness is still not completely unified because different vocal fluctuations can still take up the entire space, if the focus of attention is voices' relatively independent entities.

When the viewer shifts the focus of their attention from these two voices to their lack of truly independent and inherent existence, the effect converts into no substantial vocal fluctuations that impede and obstruct space. These so-called different vocal fluctuations can be considered composite vocal illusions. It is the common insubstantiality that causes off-screen and on-screen voices to become a seamless whole. Not only is this oneness not broken by the frame of the screen, but also not vitiated by differences in their vocal fluctuations. Moreover, this oneness is not limited to the process of de-acousmatising a voice. It can also be attributed to the void. It is noteworthy that though both de-acousmatisation and the void as two different causes make the existential quality of an off-screen voice fall apart (the apparently same effect), the effect produced by the former cause and the effect produced by the latter cause are not totally the same. The former is representational dissolution (transforming the representation of a speaking character from the off-screen into the on-screen), whilst the latter is ontological dissolution (shifting the focus of attention from the composite illusions of off-screen and on-screen voices to their lack of truly independent and inherent existence).

I would like to point out that Buddhist void space does not mean to nullify vocal transmission in space. In accordance with the previously discussed Chen's clarification of the notions of the void and form, the void permeates form, rather than negates it. If his interpretation is put into the context of the vocal issue, we realise that the void's permeability discloses that though the void is present in every part of off-screen and on-screen voices, the appearances of these voices can still be maintained at a relative level. Void space does not suggest that there are no soundtracks of Ladanian's off-screen voice and Rezai's on-screen voice. Their voices are transmitted throughout the entire space, but at the same time, there is no truly inherent vocal transmission. Hence, the nominality of their voices arises. Within this nominal state, their voices are transmitted but also not transmitted, filling space and leaving space void at the same time. Depending on how the viewer's focus of attention engages with their voices, their voices can become one or relatively separate.

4.4.4 Vocal Echoes and Buddhist Echo-like Phantoms

Sometimes when a voice is transmitted across the entire space, its echo in the distance can be detected. In *Through the Olive Trees*, I already referred to the hillside scene where Kheradmand shouts to an opposite hill several times. The echoes of Kheradmand's voices appear as the form of the environmental sounds. Previously, I mentioned Mulvey's analysis of these vocal echoes. In Mulvey's view, these vocal echoes divulge a time travel dialogue between Kheradmand and the souls of the formerly local inhabitants. Mulvey's examination of the vocal echoes, from my perspective, is related to R. Murray Schafer's idea of the soundscape since Schafer also puts an analysis of audio sense into the context of an ecological environment inhabited by local residents.[275] From my perspective, though the local condition of the recording documented in the sound is not my focus, Schafer's research on the distinction between hi-fi and lo-fi soundscapes can be put into dialogue with a Buddhist perspective on the echoes of Kheradmand's voices. I want to first look at how Schafer elucidates hi-fi and lo-fi soundscapes. Specifically, Schafer argues:

> A hi-fi system is one possessing a favourable signal-to-noise ratio. The hi-fi soundscape is one in which discrete sounds can be heard clearly because of the low ambient noise level. The country is generally more hi-fi than the city; night more than day; ancient times more than modern. In the hi-fi soundscape, sounds overlap less frequently; there is perspective—foreground and background: "… the sound of a pail on the lip of a well, and the crack of a whip in the distance"—the image is Alain-Fournier's to describe the economic acoustics of the French countryside. […] In a lo-fi soundscape individual acoustic signals are obscured in an overdense population of sounds. The pellucid sound—a footstep in the snow, a church bell across the valley or an animal scurrying in

[275] R. Murray Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World* (Rochester: Destiny Books, 1994), pp. 3–12.

> the brush—is masked by broad-band noise. Perspective is lost. On a downtown street corner of the modern city there is no distance; there is only presence.[276]

In summary, in a hi-fi soundscape, sounds, even subtle sounds, can be clearly differentiated. In a lo-fi soundscape, sounds are mixed into unitary chaos because of noise pollution. Regarding the echoes of Kheradmand's voice, I argue that from a Buddhist perspective, even though the scene displays a village-based (hi-fi) soundscape, there is no essential distinction between the foreground voice and the background echo; they can be unified. I suggest that this particular oneness is neither due to the oneness through lo-fi chaos, nor due to a situation where the voice and the echo originate from the same person. Rather, it is attributed to a Buddhist idea suggested in the *Avataṃsaka Sūtra* that "all sounds are the same as echoes."[277] Within the context of this Buddhist statement, "echo" does not simply refer to a delayed audio phenomenon at a physical level. Rather, there is a deeper Buddhist philosophical and metaphorical meaning behind the word "echo."

In the *Avataṃsaka Sūtra*, echo is compared with shadow: "[enlightening beings] know that […] sounds and voices are like echoes, that sentient beings are like shadows, that all things are like phantoms."[278] Likewise, Asaṅga juxtaposes echo with shadow by making a statement that "*yunhe yingzhi yitaqizixing*? *Yingzhi piru huan*, *yan*, *meng*, *xiang*, *guangying*, *guxiang*, *shuiyue*, *bianhua*." "云何应知依他起自性？应知譬如幻、炎、梦、像、光影、谷响、水月、变化。" (how to get the point of forms arisen from interdependence? For example, these forms can be regarded as illusions, flames, dreams, images in the mirror, shadows, echoes in the valley, the moon reflected in water, and metamorphosis).[279] These Buddhist scholars mean to stress in a more figurative way that just like shadow as an insubstantial projection of the object blocking the light, and just like echo as an insubstantial reflection of the original sound, the original sound itself is an insubstantial projection of the mind.

---

[276] Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World*, p. 43.

[277] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 872.

[278] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 621.

[279] Asaṅga, *She dacheng lun shiqin shi jizhu* 摄大乘论世亲释集注 [Vasubandhu's Explanation of the Mahāyāna Compendium/Summary] (The text of the *She dacheng lun* [The Mahāyāna Compendium/Summary] was completed by Asaṅga, circa 310–390. The explanation was completed by Vasubandhu, circa 330–410, translated into Chinese between 648–649 as the *She dacheng lun shiqin shi* by Xuanzang, circa 602–664. The *jizhu* [commentary] was completed by Zhimin, 1927–2017), ed. Li Mingquan (Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], 2004), vol. 5, p. 1. My translation (Note: In terms of the holistic system of page numbers of this book, there are two aspects that should be taken into consideration. First, each section and sub-section start with page 1. Second, each sheet of paper is one page number; that is to say, each of front and back pages constitute one page number. In order to avoid confusion, I would like to explain that page 1 refers to right-hand page 1 (the second half of page 1) of *Juan wu: Suozhi xiangfen disan zhi er* 卷五：所知相分第三之二 [Volume 5: The Objective Aspect of Cognition – Part 3.2]).

Regarding sound as an insubstantial projection of the mind, Tian Maozhi uses sound-related examples from daily life to illustrate this Buddhist ontological idea in a concrete way. For Tian, if one suffers from hearing loss, sounds cannot appear for this person. However, for those who have the ability to hear sounds, sounds arise. If one's perception is subject to auditory hallucination, one may hear a sound that others do not hear.[280] Apart from these sound-related situations, Tian also discusses a sight-based example that I think is related to sound. Tian suggests that a painter is able to see more colours than ordinary people even though they see the "same" sight.[281] I argue that this instance can be applied to a sound-related situation where a musician can hear more aural subtleties than ordinary people even if they hear the "same" sound. Tian goes on to suggest that these different experienced results indicate that the appearance of the (sonic) world is insubstantial because it can be transformed by our own (aural) perceptions. Consequently, there are no permanent and inherent existence.[282]

In the slope scene of *Through the Olive Trees*, at the physical level of a hi-fi soundscape, each of the echoes in the background arrives at audiences' ears with a short delay after each of the foreground direct voices, with the voices and the echoes being able to be clearly distinguished because the ambient noise level is quite low. However, at the Buddhist level of an echo-like phantom, from the very first moment he shouts, his voice has already "existed" as an echo; to put it another way, audiences do not need to wait for his voice to become an echo physically across the hills in order to experience an echo. An echo-like quality has already been in his voice, suggesting that the so-called existence of the voice is an insubstantial projection of the mind, as a vocal appearance can be changed by different aural sensations, and the same goes for a physical echo. In brief, because both the foreground direct voice and the background echo are not invariable, there is no fundamental difference between them.

Film studies scholar Chris Cagle has also suggested that there is no clear-cut distinction between the foreground voice and the background sound. His analysis of their oneness is based on a technical situation where digital sound technology renegotiates the foreground voice and the background sound.[283] He suggests that in the era of analogue sound design, documentarians try to

---

[280] This Buddhist discussion comes from Tian Maozhi's preface to the *Jingang jing* 金刚经 [The Diamond Sūtra]. Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007), pp. 9–10. My translation.
[281] This Buddhist discussion comes from Tian's preface to the *Jingang jing*. Ānanda, *Jingang jing*, p. 8.
[282] This Buddhist discussion comes from Tian's preface to the *Jingang jing*. Ānanda, *Jingang jing*, pp. 9–10.
[283] Chris Cagle, "The Voice in Documentary Sound Design: A Digital Revolution" (2019), in *Vocal Projections: Voices in Documentary*, eds. Annabelle Honess Roe and Maria Pramaggiore (New York and London: Bloomsbury Publishing Plc, 2019), pp. 187–202.

attenuate the background sound and prioritise the foreground voice since sounds recorded on location mix with lots of noises which, to a large degree, obscure the primary speaking voice. In the era of digital sound design, digital sound technology has allowed documentarians to embrace the background sound as an aesthetic tool, without being interfered with by unwanted noises. Cagle notes that digital sound mixers mix multiple soundtracks to make characters' talking harmonise with surroundings, so that the verisimilitude of each can be juxtaposed with each other, rather than separated.[284]

I do not intend to examine the specific facets of digital sound technology, but what I want to emphasise is that from a Buddhist viewpoint, even if there is no digital sound technology to reconcile the foreground voice with the background sound, they can still become one, considering that both possess echo-like qualities, being insubstantial projections of the mind. This oneness cannot be nullified by noises.

However, this does not mean to reject the relative distinction between the foreground voice and the background echo. Otherwise, it looks as if there were an inherent quality called "echo-like quality" that could be grasped as real. I suggest that the shift of audiences' attention can change the way they experience these two aural appearances. If audiences concentrate on an echo-like quality, the effect will be that they are one. If the focus of attention is their physical characteristics, the effect will be that one is an original direct sound and another is a reflection of sound. Once vocal oneness and difference intersect, a nominal state between them arises.

4.4.5 Vocal Identification and the Seventh and Eighth Consciousness

A Buddhist examination of the insubstantial (echo-like) projection of the mind emphasises the variable relationship between the act of perceiving and the perceived object. This interconnection also involves another perceptual aspect: identification. In the following paragraphs, I will analyse the ending of the film in order to dissect the process of vocal identification at a Buddhist ontological level. I consider that a Buddhist perspective on the process of vocal identification will provide an alternative path to nominality. There is a three-minute sequence of alternating close-up shots of two characters: Ladanian walking ahead and Rezai following her. He endeavours to talk to her even though she ignores him. Sometimes when a shot only displays Ladanian walking, audiences are still able to hear someone talking off-screen (figure 14). How to verify that this off-screen voice is Rezai's voice and how does the Buddhist identification with this off-screen voice function?

[284] Cagle, "The Voice in Documentary Sound Design: A Digital Revolution," pp. 187–94.

Figure 14 The ending voices in *Through the Olive Trees*

In "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," Victor Fan uses the Buddhist ideas of the seventh and eighth consciousness to examine the intersubjectivity of identification.[285] I would like to suggest that though his focus is pictorial identification, his standpoint can also be applied to vocal identification.

Before employing a Buddhist system to discuss the identification with Rezai's voice, I give a brief summary of how Fan puts cinematic identification into the context of Buddhist philosophies. I will then use his method to engage with Rezai's voice and provide a reinterpretation of his method. Instead of immediately engaging with a Buddhist structure of cinematic identification, Fan outlines how Western philosophers look at the process of identification. He makes a comparison between Maurice Merleau-Ponty and Jean-Paul Sartre's views on how the identification with the image operates. As Fan suggests, Merleau-Ponty's view reveals an intersubjective relationship between the spectator and the image, whilst Sartre's contention lays emphasis on subjectivity, namely a temporally and existentially permanent sensory-perceptual body that tempts the spectator to make present the image.[286] Based on their respective approaches, Fan considers that there is an aporia in which the intersubjectivity of the identification with the image is not totally intersubjective but also is subjective since "it is extremely hard to theorise identification without imagining a *je ne sais quoi* from which a temporally and existentially permanent self is constituted."[287] In order to further examine this tangled issue, Fan uses the Buddhist ideas of the seventh and eighth consciousness to expand on how the intersubjectivity of cinematic identification is but also is not intersubjective.[288]

---

[285] Victor Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies" (2019), in *The Structures of the Film Experience by Jean-Pierre Meunier: Historical Assessments and Phenomenological Expansions*, eds. Julian Hanich and Daniel Fairfax, trans. Daniel Fairfax (Amsterdam: Amsterdam University Press, 2019), pp. 250–54.

[286] Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," p. 246; Maurice Merleau-Ponty, *Phenomenology of Perception* (1945), trans. Colin Smith (London and New York: Routledge, 2002); Jean-Paul Sartre, *The Imaginary: A Phenomenological Psychology of the Imagination* (1940), trans. Jonathan Webber (London and New York: Routledge, 2004), pp. 17–20.

[287] Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," p. 246 (Note: The phrase *je ne sais quoi* in this quotation is Fan's italic emphasis).

[288] Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," p. 246.

In a Buddhist mechanism, the seventh consciousness (i.e. *mānas-vijñāna* [self-consciousness]) carries obstinate power of attaching itself to the perceived results of seeing, hearing, smelling, tasting, touching, and thinking; the latter six aspects are subordinate to the great power of self-attachment which induces the perceiver to distinguish, judge, decide, and further surmise that there is inherent existence. The eighth consciousness (i.e. *ālaya-vijñāna* [storehouse consciousness]) provides a repository for storing the first seven consciousness' information as seeds that carry with them the cause-effect chain and psychological imprints or potentiality from the past; the latent seeds remain dormant until their conditions and the self-consciousness allow them to ripen and crystallise them as future consequences.[289]

Fan proposes four Buddhist steps towards the identification with the image. These steps are used to delve into how the intersubjectivity of the identification with the image is not totally intersubjective. First, the operation of the first six consciousness causes the image to be experienced by the perceiver and thus to be manifested. If perception is not working, the image cannot emerge for the perceiver. A pictorial manifestation also in turn suggests that perception is in operation. This intersubjective and interdependent relationship is accompanied by the seventh consciousness. They jointly produce seeds which will be preserved in the eighth consciousness, with these latent seeds carrying memories. Second, these lurking seeds are activated. To put it another way, the perceiver's memories and entrenched attachment to the self and to the phenomenon continuously leave the perceiver with an impression of the unity and integrity of the perceived object in time. This attachment (subjectivity)-based perceptual process makes the perceiver want the image to remain in time and to recognise what it represents. Third, the perceiver recognises the figure in the image. Finally, the figure is "presentified" and "affectivised" by the perceiver. The reason why the perceiver can identify is because the self-attachment cannot put aside a yearning for identifying.[290] Fan paraphrases Yinshun's Buddhist view and suggests that "we can understand perfectly that identification is an intersubjective process. Yet, when we explain how we identify, we are always tempted to assume that a permanent self is in operation."[291]

I argue that the foregoing Buddhist process towards the identification with the image can also be employed to examine the identification with Rezai's voice. At the ending of the film, there is a

[289] Maitreya, *Yujia shidi lun* 瑜伽师地论 [The Treatise on the Foundation for Yoga Practitioners] (It was completed by Maitreya, circa 270–350, translated into Chinese in 648 as the *Yujia shidi lun* by Xuanzang, circa 602–664), ed. Li Li'an (Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], 2005), vol. 63, p. 976. My translation.

[290] Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," pp. 253–54.

[291] Fan, "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies," p. 254; Yinshun, *Zhongguan lunsong jiangji* 中观论颂讲记 [Lectures on the Fundamental Verses on the Middle Way] (Beijing: Zhonghua shuju 北京：中华书局 [Beijing: Zhonghua Book Company], 2017), p. 66.

sequence of alternating close-up shots in which he tries hard to keep talking to Ladanian in order to get her attention. The shots of Rezai and the shots of Ladanian alternate. When audiences see the shots of Ladanian, they can still hear an off-screen voice. At the moment of seeing Ladanian only, how to confirm that the off-screen voice is made by Rezai and how does the Buddhist identification with this off-screen voice work?

First, audiences' hearing allows an off-screen voice to be perceived and thus to be manifested, with this audio manifestation in turn suggesting that audiences' auditory sense is in operation—an intersubjective and interdependent relationship between audiences of hearing and the voice for hearing. This complex association is supervised by the self-attachment. They together generate seeds that carry with them the mental imprints from the past; these mental imprints refer to audiences' memories of the particular tone of Rezai's voice. Before the ending of the film, Rezai has talked and been present in multiple scenes. The distinctive tone of his voice has been imprinted on audiences' memories, which makes them familiar with the above-mentioned ending off-screen voice. Second, the hidden seeds are activated, which signifies that audiences' memories of Rezai's voice and their own attachment to the self and to the phenomenon jointly endow this experienced off-screen voice with the qualities of presence and constancy. This attachment (subjectivity)-based sensory process makes audiences associate Rezai's vocal tone with this off-screen voice and pursue the persistence of this voice in time. Third, audiences can confirm that this particular off-screen voice is Rezai's voice. Finally, audiences believe that Rezai is emitting this off-screen voice in the present even if the voice that they experience is simply an elapsed audio moment replayed by the electronic speakers. Based on the presentified status of the voice, if audiences are also able to understand his Persian language, they will get a better understanding of the plot of *Through the Olive Trees* and immerse themselves in it, which allows his abstract and presentified voice to be even further concretised and affectivised.

The above four Buddhist steps exhibit the process of presentifying an extinct and lifelike off-screen aural instant replayed by the mechanical acoustic apparatus as a lived voice that is being emitted in the present by Rezai as a figure of flesh and bone. Based on the above analysis of the perception of Rezai's voice, the process of identification seems intersubjective. However, holding on to the self and to the phenomenon makes the process of identification seem subjectivity-centred.

In this respect, a nominal state between the intersubjectivity of vocal identification and the subjectivity of vocal identification emerges. However, if we reconsider its intersubjectivity, we realise that intersubjectivity seems only present when identification has not yet been in operation (i.e. the foregoing first Buddhist step). This first step is directly related to Berzin's reconsideration of Śāntideva's meditation advice – or, in other words, an analysis of the experiential process of the mere arising of a sound and the hearing of it. This process reveals that when we disengage from

acoustic identification, we will not be bothered by the loud noise and therefore will be able to eliminate unhappiness caused by the loud noise. Our hearing will be accompanied by "here-and-now." We can remain with the cognitive process of the mere arising and engaging with what we hear occurring at each moment, without the self-conscious agent judging what it really is and clinging to it. In other words, this cognitive process gives emphasis to non-identification (identification has not yet been in operation) and accentuates an intersubjective and interdependent linkage between the perceiver and a sound. By contrast, the last three Buddhist steps suggest that when we identify or seek to identify, the self-consciousness is in operation. In other words, we are consciously willing a sound to become our focus of attention and making the sound appear as an "inherent" phenomenon by thinking about it very hard. This self-led role makes us ignore the fact that phenomena cannot emerge from their own side. Therefore, it appears that identification is a subjective process, whilst non-identification is an intersubjective process.

In this regard, Fan still embraces intersubjectivity as a part of identification even though identification centres around subjectivity. Self-consciousness does not nullify the intersubjective relationship between the perceiver and the perceived object. It just creates a subjective layer over intersubjectivity, rather than thoroughly dismantles intersubjectivity. A subjective revelation suggests that our ignorance of intersubjectivity deludes ourselves into gripping the self and the phenomenon and thus masks our awareness of intersubjectivity. Veiled intersubjectivity acts as a potentiality, being implicit, rather than entirely lacking. When the cause is that audiences identify and hold on to Rezai's voice, the effect will be that the intersubjectivity of the identification with his voice becomes implicit. When audiences identify but do not hold on to his voice, the effect will be that the intersubjectivity becomes explicit—variable causality that puts vocal identification into a nominal state between subjectivity and intersubjectivity. A Buddhist analysis of the identification with the voice shuns both extremes, either too clinging to it or being too detached from it.

Chapter 5 Nominal Deep-focus Shots in Classical Hollywood Cinema

5.1 Nominal Deep-focus Shots

This chapter examines nominal deep-focus shots in classical Hollywood cinema. Before considering this particular nominality, I would like to first look at the basic idea of deep-focus shots, which refers to shots that are in focus all the way from the front to the back. Apart from this trait, there are several significant expressive variations of deep-focus shots that need to be taken into consideration. Specifically, a deep-focus shot is usually expressed in the forms of composition in depth and long-take continuity. Thus, I would like to provide clarification that even if my argument in this chapter is that deep-focus shots do not exist, but also exist—in the sense of deep-focus nominality, my argument is not simply concerned with the nominal existence of deep focus itself, but also the nominal entities of its affiliated forms. From a Buddhist perspective, all these nominal situations are attributed to the changeable operation of the viewer's visual experience. Throughout this chapter, in addition to the Buddhist discussion of the process of how deep focus is constructed and deconstructed at the level of the perception of focus itself, I will also use Buddhology to dissect this process at the level of the perception of its linked forms. Deep focus and its variations constitute the holistic quality of deep focus. I will analyse the existential issue of this holistic quality.

The visual effects of deep-focus cinematography can be produced by different kinds of technical methods such as small aperture, wide-angle lenses, and fast film, designed to increase depth of field, namely the distance between the nearest and farthest planes that are in acceptably sharp focus.[292] Yet, an ontological account of deep-focus shots still needs to be re-examined, considering that the coexistence of the sharp definition of all planes of a shot can be maintained or nullified, depending on how the spectator's experience operates. James Naremore suggests:

> Human vision is exactly the opposite of depth photography, because humans are incapable of keeping both the extreme foreground and the extreme distance in focus at the same time. The crucial difference between a camera and the human eye is that the camera is non-selective. Even when we look at the deep-focus compositions in [*Citizen*] *Kane* [directed by Orson Welles in 1941], we do not see everything in the frame at once; we are aware of an overall composition which exists simultaneously, but, as Bazin has noted, we need to make certain choices, scanning the various objects in the picture selectively.[293]

[292] The planes refer specifically to the layers of a shot, rather than aeroplanes. Although the case study of this chapter—*The Best Years of Our Lives* (1946)—has an aeroplane graveyard scene presenting a great number of aeroplanes' engines, they should not be confused with the planes as the layers of a shot.
[293] James Naremore, "Style and Meaning in *Citizen Kane*" (1978), in *Orson Welles's Citizen Kane: A Casebook*, ed. James Naremore (New York: Oxford University Press, 2004), p. 130.

I argue that this perception-related idea can be regarded as the first clue as to the uncertainty of the existence of deep focus. From a Buddhist perspective, there are alternative approaches to its uncertainty. For example, one of the Buddhist contentions that I discussed in the fourth chapter is that different voices are one at the level of hearing—a view developing from Alexander Berzin's re-examination of Śāntideva's meditation advice. After illuminating hearing voices, we can follow a similar procedure with seeing sights (e.g. the planes of a shot). By the same token, this procedure shows the experiential process of the mere arising of the planes and the seeing of them. When focus identification is not manifested, our seeing will be tied to "here-and-now" without judging what the planes' focus levels really are and clinging to the extents to which these planes are sharp. By maintaining this cognitive process, seeing the foreground, middle-ground, and background planes have no substantial distinctions, for they can all be deemed to be seeing sights appearing at every moment. This oneness is not ascribed to their common sharpness, but rather to the common seeing of them. Because focus identification has not yet functioned in this experiential process, the problem of whether a shot is deep-focus has not yet arisen, which suspends the entity of deep focus. Despite the understanding of it at a technical level laying emphasis on the common sharp state of all planes, this sharpness can still be subject to being insubstantial if the mind remains with the process of seeing them, rather than discerning them.

However, this process does not aim to nullify the entity of a deep-focus shot. When we start to recognise the planes' focus levels one by one, each of their sharp features will be in a state of being noticed, thereby making present a deep-focus shot. Yet, this brings up another issue of visual experience. For the spectator who is used to films that have soft or shallow focus, they will especially notice the shot style itself when it becomes different (e.g. deep focus), since it is almost impossible for this spectator not to notice this huge difference. Moreover, when seeing a classical Hollywood film, the spectator normally gives more attention to the object seen than the act of seeing, for the spectator is usually attracted by and emotionally involved in what they see. Even if some films may not be appealing, sometimes the viewer still expects to find something that would interest them during the film screening. Because of these factors, the spectator appears easily to concentrate on what is seen (e.g. the plane) and to get to know its further identified information including its focus level and the content that it carries. Hence, it is quite difficult for the viewer to maintain the experiential process of the mere arising of the planes and the seeing of them. It seems as if in terms of the above situation, focus identification has a propensity to function and therefore makes the viewer pay attention to the sharp features of the planes seen.

It is worth noting that in the cinema, before a film screening—if there is no pre-show advertising—audiences first see an empty screen. Although they see it, they normally do not give attention to it and do not identify it as the represented space of the filmic world, since they know

that the screening has not yet started. This state can be seen as the Buddhist experiential state of the mere arising of sight/object and the seeing of it—the decentralisation of sight/object itself. During the showing of the film with deep-focus shots, even though some viewers, due to being accustomed to soft or shallow focus, will be strongly drawn to this difference and caught up in deep focus as a distinct sight, a sight-decentred experience has already happened at the very beginning. The occurrence of this primordial perceptual step cannot be denied. It has the potential for changing the display of sight. Berzin suggests that such experiential operation forms sensory consciousness, being the key to the insubstantial showing of sight. More specifically, he asserts:

> Our focus simply on sensory consciousness itself causes our sensory cognition to be an inattentive perception. In other words, the sensory consciousness still gives rise to its object, for instance a sight, but because that sensory consciousness is primarily the object upon which our meditating mental consciousness is focussing, it does not decisively apprehend its object, the sight. It is inattentive of it, and thus our meditating mental consciousness does not give rise to a clear appearance of the sight. Eventually, as our single-pointed placement of mind on mind becomes perfected, our meditating mental consciousness gives rise only to the mere arising and engaging that constitute the sensory consciousness upon which it is focussed. It does not give rise at all to any appearance of the object of that sensory consciousness.[294]

Although focussing on sensory consciousness, rather than on sight itself, opens up a perceptual approach to the deconstruction of the manifestation of sight, it is normally extremely rare for audiences to maintain such an experiential process. When the surrounding lights are turned off, audiences are reminded that the screening of a film is about to start. The light source is primarily concentrated on the screen, driving audiences to immerse themselves in it and to give attention to a variety of cinematic sights such as deep-focus shots, lighting, characters, and sets. Yet, this does not signify that the experiential process of the mere arising of sight and the seeing of it gives way to sight itself, for instance deep focus. Instead, this experiential state can be regarded as a hidden flow on the bottom of deep focus seen, loosening this sight in an implicit way. In a cinematic context, a Buddhist perspective is more concerned with a holistic visual system than deep focus seen. This holistic visual system gives weight to an explicit-implicit conversion process between deep focus viewed and the act of viewing.

When the film is being presented, though the central preoccupation with deep focus—a cinematic sight—seems to marginalise awareness of the act of seeing, this awareness is never in a state of total powerlessness. As Shuman Chen contends, “the inherent Buddha-nature as

---

[294] Alexander Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā* (The first part was completed by Berzin in 1996; the second by the Panchen Lama, 1570–1662; the third by the Dalai Lama in 1978; and the fourth by the Dalai Lama in 1982), trans. Alexander Berzin (Ithaca: Snow Lion Publications, 1997), p. 72.

metaphysical principle is *present but concealed within* each sentient being."[295] Clinging to sight produces an "existential" layer over awareness of the act of seeing, which makes this seeing hidden and forgotten but not entirely lacking. By not being extremely tight or insistent on the tangible presentation of deep focus, audiences as sentient beings can still unmask awareness of the act of seeing and unlock the potential of maintaining this awareness in which no entity of deep focus unfolds. This evokes mindfulness of the state of audiences' minds and alertness to bring the focus back to "here-and-now"—the mere arising and engaging with what they sense happening at each moment, without identifying the object as deep focus.

Mindfulness does not entirely lose deep focus even though it makes the existential quality of deep focus fall apart. Being mindful of deep focus initiated and extinguished out of a holistic visual system divulges dependent arising. When achieving the juxtaposition of its initiation and extinction, the observer's mind, as Berzin states, becomes "as acute and constant as that of meditators."[296] This state leads to a nominal deep-focus shot—an intermediate area between the existence and *non-existence* of deep focus. Its entity arrives and goes, like the flow of water.

5.2 Deep-focus Shots and the General Features of Classical Hollywood Cinema

Before using Buddhology to further look at nominal deep-focus shots, I want to first analyse the broad features of Hollywood cinema. In *Film Art: An Introduction*, David Bordwell and Kristin Thompson provide a brief summary of the classical Hollywood cinema and the New Hollywood. The former usually shows the clear and complete cause-effect chain of the narrative, and the character who wants to reach the goal.[297] In addition to the coherent story, some techniques also play an influential role in it, affecting the cinematic action of expressing the story. In *The Way Hollywood Tells It: Story and Style in Modern Movies*, Bordwell argues that a variety of cinematic techniques including chiaroscuro lighting, steep angles, and deep-focus photography of the 1940s change Hollywood storytelling and film experience.[298] Orson Welles' *Citizen Kane*, for example, turns deep focus into a major aspect of classical Hollywood style in the 1940s; a deep-focus approach to visual narration presents an innovation that is distinct from the soft-focus style of the 1920s and 1930s, with the latter featuring a blurring of edges and textures.[299] Bordwell and Thompson further argue

[295] Shuman Chen, "Buddha-nature of Insentient Beings" (2014), in *Encyclopaedia of Psychology and Religion*, ed. David Leeming (Cham: Springer, 2020), p. 273 (Note: *Present but concealed within* in this quotation is Chen's italic emphasis).
[296] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 132.
[297] David Bordwell and Kristin Thompson, *Film Art: An Introduction* (New York: McGraw-Hill, 2008), pp. 94–96, 463–68.
[298] David Bordwell, *The Way Hollywood Tells It: Story and Style in Modern Movies* (Berkeley, Los Angeles, and London: University of California Press, 2006), p. 119.
[299] Bordwell and Thompson, *Film Art: An Introduction*, pp. 458–59.

that from a general perspective, despite the fact that deep-focus cinematography makes an outstanding impact on visual storytelling, the narrative itself does not fundamentally alter. It is the technical process of conveying the story that causes a revolutionary change.[300]

André Bazin raises a similar point. The storytelling of Orson Welles and William Wyler engages with the shot in depth and the unity of image in space and time, which challenges the dramatic impact of a montage approach to narrative form.[301] One of the most noteworthy examples that strongly engages with deep focus is Wyler's *The Best Years of Our Lives* (1946). Wyler uses numerous deep-focus shots to present the daily lives of the characters. I will carry out a detailed analysis of deep focus in this film in section 5.4.

In addition to engaging with the period of the classical Hollywood cinema, the preceding technical effects are further promoted in the era of the New Hollywood, refashioning the formalist/stylistic conventions of the classical Hollywood cinema. As David Bordwell and Kristin Thompson mention, some New Hollywood directors enrich Hollywood storytelling conventions with deep-focus techniques reminiscent of *Citizen Kane*.[302] For example, they argue that "during the 1970s and 1980s, deep-focus cinematography was revived in Steven Spielberg's work, notably *Jaws* [1975] and *Close Encounters of the Third Kind* [1977]."[303] These revived deep-focus effects can be used to reinforce the spatial and dramatic impacts of narrative constructions.

5.3 A Buddhist View on Deep-focus Shots in Classical Hollywood Cinema

5.3.1 The *Sight* of Deep Focus and *Zixiang* 自相 (an Individual and Primordial Appearance)

In this part, I will employ a Buddhist visual aspect (eyes and eye consciousness) of five sense organs-perception (*wugen xianliang* 五根现量) to re-examine deep-focus shots. This visual aspect implies seeing a phenomenon in a primordial way without the act of making explicable this perceived phenomenon. Chen Bing uses the blue sky as an instance to illustrate this visual facet. He states that a primordial visual perception of a phenomenon can be considered a non-conceptual or pre-reflective sensation where, for example, when we look out the window, we see the *sight* of the blue

[300] Bordwell and Thompson, *Film Art: An Introduction*, p. 459.
[301] André Bazin, "The Evolution of the Language of Cinema" (1958), in *What Is Cinema?*, ed. and trans. Hugh Gray (Berkeley, Los Angeles, and London: University of California Press, 1967), vol. 1, pp. 33–35.
[302] Bordwell and Thompson, *Film Art: An Introduction*, p. 466.
[303] Bordwell and Thompson, *Film Art: An Introduction*, p. 173.

sky—a spacious and transparent appearance tied to visual cognition—without saying or thinking "this blue sky is extremely bright and engaging."[304]

The word "*sight*" used in this particular example is slightly distinct from the "sight" used in section 5.1.[305] With regard to Chen's account of *sight*, it refers specifically to *zixiang* 自相 (an individual and primordial appearance, which by definition transcends linguistic representation); it can be further understood as the object of immediate or raw sensation without higher intellectual processing.[306] At the first moment of cognition, the object is viewed as something indefinite (e.g. a spacious and transparent appearance [a *sight* or the object of primordial and raw sensation]), and at the subsequent moment, its quality is expressed in the verbal and conceptual form of a detailed judgement (e.g. this blue sky […]). A primordial and raw perception and a conceptual judgement are different modes of cognition.

Although the blue sky, to some extent, can be seen as a sight, it is not a *sight* of a primordial visual perception. Instead, in some ways, it is seen as a conceptual label that we add to that spacious and transparent appearance (*sight*), in order to make intelligible and describe that instantly perceived appearance (*sight*). Thus, there is a complex ontological relation between the blue sky and the *sight* of the blue sky. In the following pages, I will put the Buddhist visual aspect into dialogue with film theorists' discussions of deep focus, in order to examine the ontological association between deep focus and the *sight* of deep focus. A Buddhist approach will open up the nominal field of deep focus. It will provide an alternative discourse on the initiated and extinguished phases of a deep-focus shot, distinct from the conversion process analysed in section 5.1.

As I have mentioned previously, André Bazin takes the films of William Wyler and Orson Welles as examples to argue that deep-focus shots and long-take realism together provide audiences with enough time to select their foci of interest based on their viewing preferences and to reconfigure interpretations within a spatiotemporal unity. He contends:

> Depth of focus brings the spectator into a relation with the image closer to that which he enjoys with reality. Therefore it is correct to say that, independently of the contents of the image, its structure is more realistic. [Depth of focus] implies, consequently, both a more active mental attitude on the part of the spectator and a more positive contribution on his part to the action in progress. Whilst analytical montage only calls for him to follow his guide, to let his attention follow

[304] Chen Bing, *Fojiao xinlixue* 佛教心理学 [Buddhist Psychology] (Xi'an: Shaanxi shifan daxue chubanzongshe 西安：陕西师范大学出版总社 [Xi'an: Shaanxi Normal University General Publishing House], 2018), vol. 1, p. 136. My translation.
[305] Throughout sub-section 5.3.1, in order to articulate this nuance in a Buddhist context, I would like to write the term "*sight*"—the object of primordial and raw sensation—in italics.
[306] Chen, *Fojiao xinlixue*, p. 136.

> along smoothly with that of the director who will choose what he should see, here he is called upon to exercise at least a minimum of personal choice. It is from his attention and his will that the meaning of the image in part derives.[307]

In the above quote, though the term that the English translator Hugh Gray uses is "depth of focus," I decide to replace it with "deep focus." The use of "depth of focus" in this context is not appropriate in view of the fact that "depth of focus" is a lens optics term that measures the tolerance of the film's displacement within the camera in relation to the lens, which is not the same as "deep focus." To put it another way, "depth of focus" is the region *behind* the lens in which the film plane or the sensor plane is put to produce an in-focus image, whilst "deep focus" is the area *in front of* the lens in which the nearest and farthest planes are in acceptably sharp focus. Based on Bazin's general views on the shot in depth in William Wyler and Orson Welles' films, I suggest that Gray mistranslates the term as "depth of focus." This terminological imprecision can be further demonstrated by some academic scholarship such as Laura Mulvey's study of *Citizen Kane* and Patrick Ogle's "Technological and Aesthetic Influences on the Development of Deep-focus Cinematography in the United States" in which both utilise the term "deep focus" to fine-tune the preceding Gray's English translation of Bazin's writing.[308] This terminological clarification helps us revisit the potential relationship between a deep-focus shot and visual perception. For Bazin, deep focus gives audiences the freedom to pick which part of the frame they want to pay attention to, which affects their interpretations of what they see and enhances the ambiguity of an entire film.

Similar to Bazin's contention about the freedom to scan the visual field, in *Film Phenomenology and Adaptation: Sensuous Elaboration*, David Richard links deep-focus shots to a phenomenological reflection. He points out that deep-focus shots provoke embodied ethical evaluation permitting the viewer's eyes to scan the frame and giving the viewer the freedom to choose to either pay attention to the character's emotions or shift their attention to other objects.[309]

Even though *depending on audiences' foci of attention and wills* is related to a Buddhist framework of the mind in view of the fact that Buddhology also suggests that the variable operation of the mind can change the experienced appearance of the world, Bazin's idea of audiences' perceptions does not entirely conform to this Buddhist system. Although deep focus allows the viewer to scan the planes selectively—which causes them not to be in-focus simultaneously—there

---

[307] Bazin, "The Evolution of the Language of Cinema," pp. 35–36.
[308] Laura Mulvey, *Citizen Kane* (London and New York: Bloomsbury Publishing Plc, 1992), p. 30; Patrick Ogle, "Technological and Aesthetic Influences on the Development of Deep-focus Cinematography in the United States" (1972), in *Movies and Methods: An Anthology*, ed. Bill Nichols (Berkeley, Los Angeles, and London: University of California Press, 1985), vol. 2, p. 60.
[309] David Richard, *Film Phenomenology and Adaptation: Sensuous Elaboration* (Amsterdam: Amsterdam University Press, 2021), pp. 216–17.

is always a plane of interest in-focus during plane viewing. In some ways, this suggests that there is *an inherently minimum in-focus quality* built into a deep-focus shot. In my view, Bazin's idea treats this in-focus quality as pregiven. It simply allows that this pregiven in-focus nature can be added to any single plane, depending on the spectator's viewing preference. However, from a Buddhist perspective, there are no pregiven inherent phenomena that can be viewed from distinct vantage points.

I would like to relate the ontological issue of deep focus to Berzin's Buddhist interpretation of the difference between traffic and the *sight* of traffic, in order to disclose why there is no inherently minimum in-focus characteristic, even though there appears to be a plane of interest during plane viewing. Berzin points out that what we see is not traffic, but rather the *sight* of traffic – that is, a traffic-like object transmitted to the retinas of our eyes in connection with visual cognition, whilst traffic itself is the focal or objective condition for the experience of traffic.[310] Similar to Chen's Buddhist discourse on the *sight* of the blue sky, though the *sight* of traffic is close to traffic, it is merely the instantaneously visual imaging of the eyes, the object of primordial and raw sensation, rather than being identical to traffic itself. It is something indeterminate until we utilise a verbal and conceptual form (e.g. the word "traffic") to explain it. The act of defining the *sight* prompts us to treat "traffic" itself as the so-called sight or object in a broad sense. Yet, this so-called and relative sight is not the primordial and raw *sight*. The latter cannot be stabilised by linguistic representation, as it varies with the operation of visual cognition. Berzin uses the *sight* of traffic to emphasise that what we view is something connected to our visual sensations, affected by the retinas of our eyes, rather than an independent object. Not being a pregiven, it is a being-given. When we see the *sight* without using the linguistic label "traffic" to make intelligible and describe that *sight*, our visual sensations are primordial. By going through this primordial *sight*-based viewing procedure, there is no fixed conception that drives us to consider that this is traffic. The change of the visual imaging of the eyes can alter the appearance of a perceived thing. What we see is an impermanent mirage.

It is my contention that the above Buddhist analysis of *sight* can also be utilised to analyse the ontological relationship between a deep-focus shot and the *sight* of a deep-focus shot. What the viewer sees is the *sight* of a deep-focus shot transmitted to the retina of the viewer's eyes. Once ocular imaging alters, the perceived deep-focus shot will also be exposed to transformation. For instance, a severely myopic individual's visual experience is affected by an eye disorder where light focuses in front of, instead of on, the retina. Even if this individual is given a chance to make a

[310] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 32 (Note: Despite the fact that Berzin mentions the term "retina," he does not plan to shift attention from Buddhist philosophy to neuroscience and cognitive studies. Rather, he wants to emphasise that the visual imaging of the eyes is not equal to what the experience is aimed at, which further gives rise to the deconstruction of the perceived phenomenon. This corresponds to insubstantiality in a Buddhist framework).

selection during plane viewing, there is no single plane sharp for this individual, for the visual information about the plane's focus level sent to the retina has been changed into an entirely blurry state. This implies that a minimum in-focus quality perishes at a *sight* level. When the primordial and raw *sight* becomes out of focus due to severe myopia, what this person sees conforms neither to the basic idea of deep focus—a sharp range covering a large area front-to-back—nor to perception-directed selective focus—selecting a portion of the image that the perceiver wants to be in focus, whilst blurring out the rest. Rather, such a *sight*-oriented (seriously short-*sighted*) sensation dismantles the overall sharp quality of a deep-focus shot. Regardless of which portion of the image is chosen, it is still in an indistinct state.

A Buddhist examination of the *sight* of visual phenomena suggests a hidden conversion process lurking beneath the experience of a deep-focus shot. For the perceiver having typical vision acuity, even if the unselected areas are out of focus, the one selected as the focus of attention is still sharp, thus maintaining a minimum in-focus quality. Such sharpness can be experienced and identified by this individual—an act of acknowledging the relative entity of a minimum in-focus quality. For a severely myopic experiencer, however, the *sight* is manifested as a blurry appearance in which each of the planes of a shot is out of focus, thus dismantling an in-focus potentiality. The conversion process—nominality—between a minimum in-focus quality and *no in-focus quality* implies that the selected area is and is not different from the unselected, which is distinguished from Bazin's selective focus. From a Buddhist perspective, Berzin's reference to the role of the retina reminds us of the effect of variable *sight*. An in-focus quality is what the viewer's mind looks for, but not what absolutely appears to the perception experiencing it.

It is noteworthy that some scholars have engaged with the deconstruction of the existence of deep focus but in an alternative way. For example, David Bordwell argues that many deep-focus shots in *Citizen Kane* are not produced by the camera, but by matte painting and optical printer work.[311] James Naremore also suggests that even if André Bazin stresses that the use of deep-focus shots in *Citizen Kane* reinforces the ambiguity in the structure of the image, Bazin underemphasises "invisible montage, a combining of two or more images in a complicated optical printing process that creates the illusion of a single shot."[312]

In addition to the disintegration caused by matte painting and optical printing process, Bordwell examines the deconstruction of the long-take quality of deep focus. Before looking at this

[311] David Bordwell, Janet Staiger, and Kristin Thompson, *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960* (London: Routledge, 1985), pp. 589–90.
[312] Naremore, "Style and Meaning in *Citizen Kane*," p. 128 (Note: In the third chapter, I also utilised "invisible montage," but by this I mean mind-oriented implicit discontinuities within the long take, which should not be confused with Naremore's "invisible montage").

idea, I would like to reiterate that a long-take form is a significant component of the *holistic* characteristic of deep focus since it gives the viewer sufficient time to view the details of all planes' focus levels and finally to verify that there is a wide range of sharpness. The deconstruction of long-take continuity, however, interrupts the cognitive process of confirming such front-to-back sharpness to some degree, and hence changes the way a deep-focus shot is manifested. Its interrelated forms (e.g. long-take continuity and great depth) need to be taken into consideration because all of the discussions of these related forms will lead to the existential issue of the *holistic* quality of deep focus.

Bordwell highlights a visual difference between *Citizen Kane* and *The Best Years of Our Lives* by pointing out that the conventional long-take quality of deep focus has undergone a certain transformation. This deconstructs the rigidity of long-take tableaux, making volatile the manifestation of deep focus. According to Bordwell, this issue can be further illustrated if we take notice of a difference between the long-take forms of deep focus in these two films.[313] One of the prime examples in *The Best Years of Our Lives* is a bar scene where Al Stephenson (played by Fredric March) looks from Homer Parrish (played by Harold Russell) and Butch Engle (played by Hoagy Carmichael) playing the piano together in the nearest plane to Fred Derry (played by Dana Andrews) in the farthest phone booth (figure 15). This deep-focus scene first presents all of them in a long take and then cuts in closer, so that the screen is only able to frame Stephenson and Derry; the scene then cuts back to all characters, exhibiting them in a long take, and cuts in closer again to display Stephenson and Derry on-screen only. By utilising editing to interrupt the continuity of a deep-focus shot repeatedly, deep focus gets rid of permanent power (i.e. the rigid law of long-take expression) and becomes something in a flexible flow. This deep-focus paradigm retains a wide range of sharpness and treats every plane equally. However, it uses editing to turn the phone booth into a potentially central area of the image and therefore destabilises the equal weight of the planes. Yet, initially, deep focus is usually linked to an uninterrupted long-take form. For example, in *Citizen Kane*, an uninterrupted long take is utilised to display the deep-focus shot in a boarding house that simultaneously shows Mary Kane (played by Agnes Moorehead), Walter Parks Thatcher (played by George Coulouris), Jim Kane (played by Harry Shannon), and the eight-year-old Charles Foster Kane (played by Buddy Swan) (figure 16). An aesthetic change of a long-take model from *Citizen Kane* to *The Best Years of Our Lives* stresses that a deep-focus entity does not function through a fixed criterion.

[313] Bordwell, Staiger, and Thompson, *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*, p. 592.

The deconstruction of deep focus is not only ascribed to the deconstruction of focus itself (a perceptual change from a state of being sharp to indistinct), but also to counterfeit deep-focus techniques, the deconstruction of the long-take tradition of deep focus, and the fluidity of composition in depth in deep focus (this facet will be discussed in detail later). Instead of simply looking at focus itself, the feature of deep focus can be regarded as *composite*, as something tied to its variations. These distinct factors deconstruct the existence of the *holistic* characteristic of deep focus.

Figure 15 The bar in *The Best Years of Our Lives*

Figure 16 Mrs Kane's boarding house in *Citizen Kane*

Both Bordwell and Naremore's approaches to the overthrow of deep focus result from technical factors (e.g. matte painting, optical printer work, and editing). Nonetheless, I mentioned that Naremore also touches upon the deconstruction of deep focus at a perceptual level by unfolding the contrast between the selective human eye and the non-selective camera-eye.[314] Naremore's perceptual account is similar to Bazin's view on selective focus. Although a deep-focus shot can be dismantled at a technical level, Naremore and Bazin's stances on selective focus suggest

[314] See p. 113.

that deep focus still carries a minimum in-focus quality and therefore, I believe, moderates the process of deconstruction.

Building on Berzin's Buddhist discussion of primordial and raw *sight*, even if a deep-focus shot is not produced by matte painting and optical printer work and not interrupted by editing and even if selective visual experience keeps a certain plane sharp, this sharp state can still be dissolved at a *sight* level. When audiences pay attention to one plane and if the short-*sighted* visual imaging of the eyes is also considered, a minimum in-focus quality undergoes a process of deconstruction from the sharpening to blurring of this chosen plane and becomes insubstantial. However, if this ocular factor is not taken into account, what is seen conforms more to selective focus than to the *sight* of deep focus, thereby keeping a minimum in-focus quality. When these two visual approaches are superimposed, an intermediate field between a minimum in-focus quality and *no in-focus quality* emerges. In-focus nominality makes us engage in non-dualising the ontology of deep focus, keeping calling to mind the ever-changing manifestations of the perceived planes.

5.3.2 A Nominal Wide Range of Sharpness and Knowledge-associated Spatial Intuition

Another cinematic phenomenon that has a certain connection with deep-focus cinematography is deep-space mise-en-scène. In *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*, David Bordwell argues that we need to build awareness of the significant distinction between deep-focus cinematography and deep-space mise-en-scène. The former creates spatial depth by engaging with different technical devices such as small aperture, wide-angle lenses, fast film, matte painting, and optical printer work, with every plane of a shot being in sharp focus. The latter creates spatial depth by composing primary action on two planes or by making use of some spatially divided elements such as doors and windows to frame distant action; however, deep-space staging usually has only one plane that is expressed in a sharp manner.[315] Despite the fact that Bordwell lays emphasis on their distinction, I argue that he still engages with the oneness of these two aesthetic systems, as he points out that both represent spatial depth. That is to say, Bordwell has already touched upon an intermediate region between their distinction and oneness. I would like to point out that composition in depth is not exclusive to deep-space mise-en-scène. It also serves as an important form of deep-focus cinematography. Spatial depth in a deep-focus shot, from my perspective, is subject to fluidity. Depending on audiences' perceptions, it can become shortened or lengthened, thereby raising a question of whether there is inherently great depth for

[315] Bordwell, Staiger, and Thompson, *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*, p. 581.

front-to-back sharpness to cover – or, in other words, a question of whether the *composite* feature of a deep-focus shot, a wide range of sharpness, exists inherently.

Before re-examining spatial depth, I would like to further consider the idea of perception—primordial perception, to be precise—in Buddhist thought. In addition to having a link with *sight*, Chen points out that primordial perception is also related to intuition.[316] He further declares that though primordial perception is not inference, an inferential thing can become an intuitive thing when the experiencer is quite familiar with a piece of inferential knowledge and accepts it as normal. To some extent, the act of restoring an inference to an intuition can be understood as an updated and evolutionary primordial perception. For instance, someone who has learnt the principles of physics is naturally aware that the falling of an apple is attributed to gravity without much thinking and hesitation.[317] This suggests that cumulative knowledge can be internalised and habitually performed in intuitive perception.

Although gravity as a physical law implies a conventionalist/historical basis for an experience, which seems to contradict a primordial *sight*-based experience as a non-conceptual or pre-reflective sensation, the qualitative transformation of gravity from a product of higher intellectual processing to an immediately perceived object suggests that this law becomes the product of natural perception. To some degree, the operation of primordial perception does not contradict knowledge or concepts, but rather embraces them. Buddhology contends that the act of perceiving things from a changeable perspective can alter things perceived. In other words, based on this wider orientation, when the familiarised type of knowledge becomes different, the experiencer's intuitive perception can be altered in multiple ways, thereby making the perceived object have impermanent appearances. This conversion process will be used as a nominal basis for an analysis of the spatial perception of deep focus.

For example, the experiencer's knowledge of painting has an impact on their intuitive perception of the distance represented on a two-dimensional surface, and the distance that this observer experiences can be different from the distance experienced by the other who is more familiar with the knowledge of cinema than painting. Great depth is an important component of the *holistic* characteristic of deep focus. Without a considerable distance, a wide range of sharpness cannot be brought to the viewer. Different types of knowledge change the distance seen and hence affect the way we look at deep-focus cinematography. This factor will open up, in my subsequent discussion, a Buddhist nominal situation where there is great depth but no great depth, which further leads to a Buddhist nominal state between a wide range of sharpness and *no wide range of*

[316] Chen, *Fojiao xinlixue*, p. 136.
[317] Chen, *Fojiao xinlixue*, p. 137.

sharpness. The reason why I pick painting as a part of my discussion is because the canvas, a plane like the screen, also engages with the issue of spatial depth and this comparability will be conducive to a Buddhist study of the nominal status of spatial expression. I will use Western paintings as instances in view of the fact that Eastern paintings remind us of a Buddhism-in-art representation, which is easily confused with Buddhology. This discussion will not look closely into the field of painting. Rather, I will engage with its several aspects that have relevance to deep-focus cinematography.

When a deep-focus shot is observed from a painterly perspective, it decreases the distance between the nearest and farthest planes, considering that in order to present a longer distance, a painter needs to utilise aerial perspective, a technique of producing an illusion of great depth by depicting distant planes as blurrier than near planes, which rather resembles shallow focus. One of the prominent examples is Leonardo da Vinci's Renaissance painting *Mona Lisa* (circa 1503–circa 1517) where a visual contrast between the neutral greyish-blue sky and dark greyish-blue trees in the background is attenuated (the trees are hazy, covered with the transparent layers of atmosphere), whilst the foreground figure is sharp (figure 17). These aerial layers desaturate the colours and blur the contours of the objects perceived at a distance to increase spatial depth. Building on Chen's Buddhist notion of knowledge-associated intuitive perception, the viewer familiar with the principle of the artistic expression of the molecules of air can be naturally and instantly aware that *Mona Lisa* emphasises considerable depth. However, if all planes, just like deep focus, are sharp, aerial perspective will be reduced. The common sharpness makes all planes appear to be on the same layer, which makes spatial depth become flatter, a low relief as it were. For example, in Paul Cézanne's post-impressionist painting *Madame Cézanne Leaning on a Table* (circa 1873), Cézanne uses a sharp contrast between the light warmish-grey sky and dark greyish-brown balcony bars to make the background stand out and thus narrow the distance between the farthest area and nearest Madame Cézanne seated at a table (figure 18). Furthermore, Cézanne employs an eye-catching contrast between the light blue wall and dark brown rear edge of the table to sharpen and highlight its rear edge, making its rear and front edges become equally important—a stylistic approach to restoring a lifelike three-dimensional space to a two-dimensional surface of the canvas. At the Buddhist level of intuition, the spectator familiar with the layers of air can immediately realise that this deep focus-like (air-reduced) painting compresses the distance, which deviates from great depth as a significant element of the *holistic* characteristic of deep focus. A cumulative study of painting makes the observer build heightened awareness of the aesthetic operation of aerial perspective. When seeing a surface that shows all things in a stand-out way, this observer can promptly perceive that this is attributed to the weakening of the layers of atmosphere, to spatial compression, which means that the distance is short—an experienced result automatically accepted as normal.

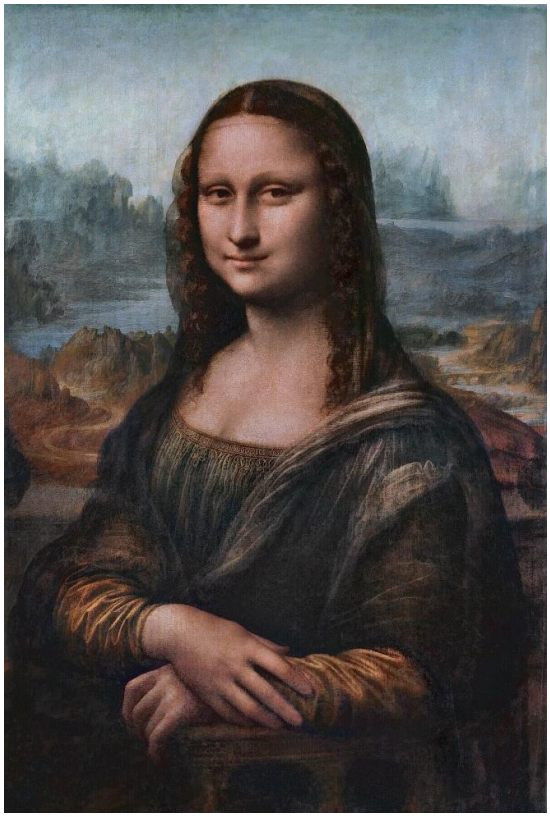

Figure 17 *Mona Lisa*

Figure 18 *Madame Cézanne Leaning on a Table*

In a cinematic context, one of the functions of a deep-focus shot is to highlight composition in depth. If there is no great depth, even shooting at the maximum focus, the sharp range can only cover this entire short distance. In order to achieve a wide range of sharpness, great depth is

indispensable, providing more space for the extension of the sharp range. A wide range of sharpness—a *composite* feature of deep focus—is constructed by considerable depth and a front-to-back in-focus state. In *The Best Years of Our Lives*, there is a scene where Fred Derry walks in an aeroplane graveyard, with a wide range of sharpness being present between the nearest and farthest aeroplanes' engines (figure 19). Instead of considering the art principle of aerial perspective, perception at this point is more concerned with the cinematic issue of depth of field. A great distance and deep-focus techniques together constitute a wide range of sharpness. The spectator perceives that there is considerable depth brought into focus. The aeroplanes' engines are not on the same layer. The graveyard is spacious. At the Buddhist level of intuition, the cumulative learning of depth of field in the domain of cinema tempts the viewer to be unhesitatingly aware that the wide sharpness of all of these aeroplanes' engines is assembled by great depth and front-to-back sharpness, and to take this perceived outcome for granted. Yet, if this scene is put into a painterly context, the distance sharp from front to back restricts aerial perspective and therefore decreases itself.

Figure 19 The aeroplane graveyard in *The Best Years of Our Lives*

In terms of a wide range of sharpness, seeing its depth as something shortened and as something lengthened seem contradictory. The first situation is to bring the farthest area closer by reducing aerial perspective and therefore to decrease depth. When it is viewed as something shortened, a sharp state does not truly cover a wide range, but simply covers a compressed range that appears spacious. The second is to stress that the reason why a wide range of sharpness can be perceived is because great depth indeed provides more space for the extension of the sharp range. Yet, these two situations can be integrated into a holistic process of visual perception at a Buddhist level. Different familiarised types of knowledge affect the spectator's intuitive perception of spatial depth differently, making a perceived distance have distinct appearances. It seems that a deep-focus shot engaged with composition in depth. However, from a Buddhist perspective, there is no inherent depth displayed. Its insubstantial state is not attributed to a commonplace view that a represented distance is an illusory spatial appearance simulated on a two-dimensional surface,

but attributed to a variable factor that a certain "established" distance can be seen as an elasticated potentiality (flow) that can be either lengthened or shortened. Its appearance is not predetermined, depending on what kind of knowledge engages with intuitive perception, which suggests that depth has no permanent essence in its own right. Elasticated depth provides but also does not provide more space for the extension of the sharp range, which leads to a nominal situation where a wide range of sharpness is initiated and extinguished, maintaining and dismantling the *holistic* characteristic of a deep-focus shot. The volatile state of spatial depth turns deep-focus cinematography into something in a flexible flow.

I would like to provide a further exposition of the aforementioned distinction between the commonplace view (a simulated distance) and the variable factor (an elasticated distance). Even if both suggest that distance is an illusion, they are not completely the same. The distance represented in the image as a simulation of the distance in the three-dimensional world simply means that the representation of the image is illusory, rather than reveals that the image itself is illusory. In other words, even though what we view is not a physical life-based spatial distance, we still experience something – that is, the rectangular image with concrete length and width. We see an image-based represented distance. The aim of an elasticated distance is to reveal that the image itself—an image-based represented distance, to be specific—is also insubstantial. Even if this represented distance seems fixed since everything displayed in a finished film no longer changes, it is not absolutely fixed. At the Buddhist level of intuition, when we internalise a piece of particular knowledge, it will become a part of our intuitive perceptions and have a particular effect on a perceived image-based represented distance. It can be regarded as an elasticated distance that can be either increased or decreased at a mind level. The variability of spatial depth divulges that what we perceive is not an inherent image-based represented distance since there is nothing that has permanent existence. It is neither a physical life-based spatial distance nor an image-based represented distance, but rather a fluid illusion devoid of existing in a truly inherent manner. A Buddhist examination of spatial depth emphasises its ontological engagement with knowledge-related natural intuition and with a certain degree of ambiguity, rather than seeing depth as independent. It seems that great depth and a front-to-back in-focus quality constituted a deep-focus shot, but when spatial intuition changes, sharpness does not necessarily occupy a wide space.

The non-inherent entity of distance does not mean that distance is nothingness. Distance has relative existence. Instead of experiencing nothing, audiences still see objects when the act of seeing is working. They are still able to recognise a distance manifested in the frame, to immerse themselves in the process of how characters engage with the represented space, and sometimes even to identify with characters in emotional surroundings. These perceived appearances, to some degree, cannot be absolutely negated because they happen, but happen in an insubstantial manner.

Their insubstantiality is not attributed to their simulations of the three-dimensional world, but rather to the variable knowledge-associated intuitive perception of an image-based represented distance. On the one hand, depth can be perceived as something shortened and lengthened because of different knowledge-related spatial intuitions. On the other hand, such fluidity makes depth have no independent and permanent entity. Great depth seems to be covered by sharpness. Yet, as depth has no inherent quality, there is no inherent wide space for a front-to-back sharp state to truly cover. This spatial entanglement in deep focus leads to a nominal wide range of sharpness.

5.4 *The Best Years of Our Lives* and Buddhology

5.4.1 Traditional Approaches to the Experience of Deep-focus Shots

Before utilising Buddhology to conduct a further analysis of deep-focus nominality in *The Best Years of Our Lives*, I would like to first provide an outline of how other theorists carry out ontological examinations of deep-focus elements of this film. Bazin selects the bar (figure 15) and wedding scenes (figure 20) as examples to explore the ontological process of how the spectator's attention is directed even though the common sharp quality of all planes, to a great degree, discourages the manipulation of the direction of attention. As for the shot in the bar, the foreground piano performance and background telephone call are given focal emphasis equally, presented in a long take. However, in order to prevent the spectator from forgetting the inconspicuous but important action in the phone booth, this long take is cut twice by close-ups of Al Stephenson peeking at the phone booth, which persuades the spectator to experience his anxiety.[318] Yet, I argue that even if editing interrupts this continuity, there is still a continuous state throughout this deep-focus scene—in the sense of continuous nominality. Continuous nominality will be utilised as the key to the deconstruction of spatial depth, which will further call into question the idea of a wide range of sharpness. This facet will be detailed in sub-section 5.4.4.

In addition, Bazin suggests that though the deep-focus wedding makes both sides of the shot carry equal sharpness and hence gives audiences multiple choices of what to see, they still follow the director's intention.[319] The viewer's attention will be directed towards the left area of the shot in which Fred Derry and Peggy Stephenson (played by Teresa Wright) look at each other lovingly. Since the marriage on the right has been confirmed, the viewer becomes even more curious to know whether two would-be lovers on the left can be together forever. Moreover, their lines of

[318] André Bazin, "William Wyler, or the Jansenist of Directing" (1948), in *Bazin at Work: Major Essays and Reviews from the Forties and Fifties*, ed. Bert Cardullo, trans. Alain Piette and Bert Cardullo (New York and London: Routledge, 1997), pp. 14–16.
[319] Bazin, "William Wyler, or the Jansenist of Directing," pp. 16–17.

sight can also direct the viewer's attention towards them. Yet, at a Buddhist level, every object in this shot can become out of focus if the viewer's distraction is considered. This factor will give rise to the fluidity of the common sharp quality of a deep-focus shot. I will expand on this issue in sub-section 5.4.7.

Figure 20 The wedding in *The Best Years of Our Lives*

Bordwell provides an ontological account of the deep-focus spatiality of *The Best Years of Our Lives* in comparison with *Citizen Kane* and *The Little Foxes* (Wyler, 1941). He considers that the compositions in the former are more spacious than the latter two. Although *Citizen Kane* and *The Little Foxes* emphasise the sense of depth, they usually jam many faces into the frame and make the nearest face get close to the camera. Therefore, the spatial expression of these two films looks less open than *The Best Years of Our Lives*. Bordwell picks several deep-focus shots as typical examples to specify this visual distinction. He contrasts the shot in the store where Fred Derry and Peggy Stephenson have an intimate chat at a perfume and cosmetics counter in the foreground area and the store manager in the background area (figure 21), with the shot in Mrs Kane's boarding house (figure 16). In addition, he contrasts the shot in the Parrishs' living room where the two families have a discreet conversation (figure 22), with the shot of the stairwell (figure 23).[320] Bordwell argues that the different arrangements of the characters' positions in the shots can make a spacious quality present to different degrees. Yet, I consider that even if the perceived object is the same composition, the entire space can still be subject to variations—being increased and decreased. Such fluidity originates from the variability of the viewer's mind. I will use Buddhology to conduct an analysis of how the shift of the viewer's mind reshapes the perceived space, which will lead to the issue regarding the nominal essence of a wide range of sharpness of deep focus.

[320] Bordwell, Staiger, and Thompson, *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*, p. 592.

Figure 21 The store in *The Best Years of Our Lives*

Figure 22 The living room in *The Best Years of Our Lives*

Figure 23 The stairwell in *The Little Foxes*

5.4.2 Deep-focus Shots and Conceptual Thoughts

Previously, I carried out a Buddhist examination of an elasticated distance in a deep-focus shot. This idea suggests that as depth has no inherent quality due to different knowledge-associated spatial intuitions, there is no inherent wide space for a front-to-back sharp state to truly cover. On the one hand, there is a considerable distance brought into focus. This deep-focus feature can be naturally

and instantly perceived by the viewer. On the other hand, heightened awareness of the potential connection between the common sharpness and the attenuation of aerial perspective promptly gives rise to an alternative perceptual situation where a sharp quality does not truly cover a large area, but merely covers a low relief-like area that seems spacious. The overlapping of these two experiences generates a nominal wide range of sharpness.

One of the ways to represent spatial depth is to place characters into a spatially divided composition. To illustrate, in *The Best Years of Our Lives*, there is a deep-focus shot in which Al Stephenson gets home and hugs Milly Stephenson (played by Myrna Loy) (figure 24). The entire screen is divided by the frames into several spatial segments, such as the door frames, the window frame, and the frame formed by a combination of the beam and the walls on both sides. The characters walk in this room, engaging with their multi-layered surroundings. Yet, from a Buddhist perspective, the frames used for spatial division in this shot do not necessarily promote the representation of depth, which brings up an ontological question of whether there is an inherently large area truly covered by a front-to-back sharp state. I assert that the subsequent Buddhist discussion of the nominal field between great depth and *no great depth* will be based on the mind's acts of tightening and loosening a conceptual thought on the characters within the frames. When being in direct perceptual contact with space without the intervention of a conceptual thought on the characters in space, what we experience can be a narrow distance that appears large brought into focus.

Figure 24 Al Stephenson being back home in *The Best Years of Our Lives*

A Buddhist study of how to deal with conceptual thoughts is provided in *Saraha's Dohākośa-gīti* (also known as the *Saraha's Treasury of Songs: People Dohā* or the *Saraha's Three Cycles of Dohā: People Dohā*, translated into Chinese in 2013 as the *Daoge baozang·shumin zhi ge* 道歌宝藏•庶民之歌 by Shi Xinhui). Even if *Saraha's Dohākośa-gīti* does not dissect the issue of characters within frames, its discussion of conceptual thoughts can still be applied to this issue. One of the statements in this Buddhist source vividly explains the operation of reducing conceptual thoughts.

"He who does not enjoy the senses purified, and practises only the Void, is like a bird that flies up from a ship and then wheels round and lands back there again."[321] Regarding this bird-and-ship Buddhist formula, the Dalai Lama offers a supplementary interpretation:

> In ancient times, before there were compasses, ships voyaging on the ocean would carry a few birds. If the navigator saw in the distance what seemed like a large wave, but thought that perhaps it might be dry land instead, he would release one of them. […] If it were only a wave, then since they were in the middle of the sea, there was no place for the bird to land other than back on the ship. No matter how far or for how long the bird flew, in the end it would have to return. The same is true with any conceptual thought let loose by the mind. It takes off from the here-and-now, non-contriving mind and, no matter how far it goes, can only return or disappear into the here-and-now, non-contriving mind once more.[322]

From the Dalai Lama's perspective, the here-and-now, non-contriving mind is seen as the coarse primordial mind. It reveals a state where we are not taken hold of by any conceptual thoughts. By maintaining this state, we are able to "see a vivid, non-obstructing bare absence and clarity as well – a pristine clarity that cannot be touched by any form of matter and which neither obstructs nor can be obstructed by anything."[323] I argue that a Buddhist analysis of the mind's acts of tightening and loosening a conceptual thought on the characters within the frames can change the way we see the Stephensons' apartment as well. Considerable depth and front-to-back sharpness construct a wide range of sharpness within this deep-focus shot. Yet, the existential quality of this *composite* feature still needs to be revisited.

The reason why audiences recognise great depth in this deep-focus shot is because their minds give rise to a conceptual thought on the characters within the frames. Usually, the character is regarded as a potential unit of measurement. When at least one of the characters appears in a spatially divided composition, there is a concretely proportional relationship between the sizes of the characters and of the surrounding frames. Being taken hold of by a conceptual thought on the characters allows audiences to measure other objects in space and to raise their awareness of great depth in a three-dimensional appearance. In addition to a conceptual thought on their sizes, thoughts on the character's movement and line of sight also enhance audiences' perceptions of depth in a solid representation. For example, by thinking about Al Stephenson walking along the

---

[321] Saraha, *Saraha's Treasury of Songs: People Dohā* [Daoge baozang·shumin zhi ge 道歌宝藏•庶民之歌] (The *Saraha's Treasury of Songs: People Dohā* is ascribed to Saraha, circa the first century, written by one of the chiefs of the eighty-four great perfected ones [the Siddhas of perfection] circa the ninth century. The explanation was completed by Khenpo Tsultrim Namdak, translated into Chinese in 2013 as the *Daoge baozang·shumin zhi ge* by Shi Xinhui), in *Buddhist Texts: Through the Ages*, ed. Edward Conze, trans. David Snellgrove (Oxford: Bruno Cassirer, 1954), p. 233.
[322] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, pp. 287–88.
[323] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 287.

corridor, audiences are able to identify the depth of the corridor based on his average pace length. Furthermore, the children's lines of sight can direct audiences' attention towards the back view of his walking, which further underlines considerable depth in a three-dimensional representation.

If a bird-and-ship Buddhist formula for the here-and-now, non-contriving mind is applied to this deep-focus example, we realise that a conceptual thought on the characters, like a bird, will land back on the non-contriving ship as well. When our minds let loose this thought and investigate it flapping its wings, we become aware that this thought is powerless. It cannot go anywhere but can only fly back to the non-contriving ship. It cannot demonstrate that the surrounding frames themselves can retain great depth in their own right. If we are in direct perceptual contact with the surroundings without having a thought on the characters, the surroundings can also have the potential of being flattened even if this deep-focus shot features great depth with front-to-back sharpness.

The reason for the likelihood of being flattened is not due to the commonplace fact that what we see is a three-dimensional illusion projected onto the two-dimensional screen, but rather due to the great uncertainty of the whole size of the surroundings seen. Once our minds are not interfered with by a thought on the characters, the holistic size of the surroundings will tend to be tremendously indeterminate. We can assume that the surroundings are either huge – like a spacious apartment, or tiny – like a doll's house with the same proportion, which weakens our perceptions of concrete depth and therefore calls into question the idea of a wide range of sharpness of this deep-focus shot. By being in a situation where there is no thought on the characters, we can sense space of fluidity in which there is no inherently great depth, which destabilises a wide range of sharpness as a *composite* feature of deep focus. This shot can be either manifested as a spacious appearance and provides a large area for a front-to-back sharp state to cover, or as a narrow appearance in which the sharp range becomes restricted but appears wide.

5.4.3 Deep-focus Shots and Buddhist Mirror Theory

Another method for engaging with a wide range of sharpness is the application of a mirror. By placing a mirror and displaying the mirror image sharply, audiences can recognise considerable off-screen depth reflected sharply in the on-screen field. More specifically, in *The Best Years of Our Lives*, there is a deep-focus example that displays Peggy Stephenson and Marie Derry (played by Virginia Mayo) having a chat in the ladies' room where there is a make-up mirror reflecting both Stephenson and an attendant (played by Louise Franklin) (figure 25). If the mirror were replaced by an empty wall, the sharp range would be blocked by the wall, as there would be no reflected off-screen depth for the sharp range to extend itself. The use of the mirror buttresses a wide range of sharpness. Yet, I suggest that the mirror images do not necessarily increase spatial depth. My idea

is not based on the commonplace fact that everything displayed in the mirror is evidently a two-dimensional reflection disengaging from depth, but rather based on the Buddhist theory of the connection between the mirror images and "*wusuo conglai*, *wusuo congqu*" "无所从来，无所从去" (*no coming*, *no going*). In the *Avataṃsaka Sūtra*, Buddhabhadra and Śikṣānanda offer a Buddhist analysis of how the mirror images engage with *no coming*, *no going*:

> Suppose there is a glass mirror called "reflector," clearly reflecting, equal in size to ten worlds, with the images of all the mountains and rivers, and all the sentient beings, including in the hells, the animals and ghosts, all reflected therein—what do you think, can you say those images enter and exit the mirror? The godlings answered, "No." All actions are also like this; though they can produce resulting consequences, there is no place they come or go. Just as a magician deceives people's eyes by illusion, so do actions also do the same.[324]

Figure 25 The mirror in *The Best Years of Our Lives*

Before applying *no coming*, *no going* to the mirror images, I want to scrutinise the meaning of *no coming*, *no going*. Some Buddhist writings provide accounts related to *no coming*, *no going* and will subsequently help us have a better understanding of "no place [the mirror images] come or go." A notable source engaging with this aspect is the *Vajracchedikā Prajñāpāramitā Sūtra* where the Buddha – Siddhārtha Gautama – informs Subhūti that "*rulai zhe*, *wusuo conglai*, *yi wusuo qu*, *guming rulai*." "如来者，无所从来，亦无所去，故名如来。" (the Tathāgata does not come from anywhere, nor does he go anywhere. Therefore he is called the Tathāgata).[325] With regard to

[324] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经] (The first Chinese translated version was completed by Buddhabhadra circa 420, and the second by Śikṣānanda circa 699), trans. Thomas Cleary (Boston: Shambhala Publications, Inc., 1993), p. 949.

[325] Ānanda, *Jingang jing* 金刚经 [The Diamond Sūtra] (The *Jingang jing* ascribes its discourses and dialogues to the Buddha – Siddhārtha Gautama, circa 563–circa 483 BCE, and Subhūti. The text was edited by Ānanda circa 494 BCE, translated into Chinese circa 401 as the *Jingang jing* by Kumārajīva, 344–413), in Tian Maozhi, an annotated edition of the *Jingang jing* (Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007), pp. 111–12; in the English translated version *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra*, the original text is

this statement, Tian Maozhi adds a note to it and explains that the Buddhist concept of *wusuo conglai*, *wusuo congqu* 无所从来，无所从去 (*no coming*, *no going*) is also known as *wusheng wumie* 无生无灭 (*non-arising*, *non-perishing*) – or, to be accurate, the *non-arising*, *non-perishing* of the Buddha's Dharma Body.[326]

To understand this alternative term, we need to first look at its connection to causality. One of the most significant Buddhist sources that engages directly with these two facets in an interweaving way is the *Karunikarāja Prajñāpāramitā Sūtra* (also known as the *Perfection of Wisdom Sūtra for Humane Kings Protecting Their Countries*, translated into Chinese in 765 as the *Renwang huguo bore boluomiduo jing* 仁王护国般若波罗蜜多经 by Amoghavajra, 705–774). This Buddhist text offers an analysis of why all things are subject to *non-arising*, *non-perishing* (*no coming*, *no going*): "*Yiqie zhongsheng xing wushengmie*, *you zhufa ji huanhua er you*, *yun chu jie xiang wuhe wusan*, *fa tong faxing jiran kong gu*. *Yiqie zhongsheng zixing qingjing*." "一切众生性无生灭，由诸法集幻化而有，蕴处界相无合无散，法同法性寂然空故。一切众生自性清净。" (All beings neither arise nor perish in their own right. Not possessing independent entities, all phenomena stem from interdependence [conditions], appearing thus as compound phantoms devoid of existing in a truly inherent manner. The aggregates of material and mental formations, sense bases, and worlds have never autonomously come or gone, which accords with dharmatā – or, in other words, the Buddhist principle of "thusness" or "suchness." All beings have abilities to quiet their minds of attachment to compound appearances, reach an empty state, and achieve the wondrous pure radiant body of enlightenment).[327] If a thing experiences its arising (coming) or perishing (going), it must have relied on the condition(s) for undergoing this process. For example, the arising (coming) of ice needs to hinge on water and a sub-zero environment; without these factors, ice cannot arise (come). Likewise, the perishing (going/melting) of ice depends on a temperature above freezing point; without this condition, ice cannot perish (go). In other words, nothing arises (comes) and perishes (goes) in its own right. This is the key to what Gautama Buddha means by *no coming*, *no going*.

---

accompanied by the commentary provided by Xuanhua (the pinyin form of Hsüan Hua [Hsüan Hua is a Wade–Giles-based name]), 1918–1995, trans. Heng Zhi (the pinyin form of Heng Chih [Heng Chih is a Wade–Giles-based name]) (San Francisco: Sino-American Buddhist Association, Inc., 1974), p. 235.

[326] This Buddhist explanation comes from Tian's annotations for the *Jingang jing*. Ānanda, *Jingang jing*, p. 112. My translation.

[327] Amoghavajra, *Renwang huguo bore boluomiduo jing* 仁王护国般若波罗蜜多经 [The Perfection of Wisdom Sūtra for Humane Kings Protecting Their Countries] (It was translated into Chinese in 765 as the *Renwang huguo bore boluomiduo jing* by Amoghavajra, 705–774), edited by Dazheng xinxiu dazang jing kanxinghui 大正新修大藏经刊行会 [Society for the Publication of the Taishō Tripiṭaka] in the *Dazheng xinxiu dazang jing* 大正新修大藏经 [The Taishō Tripiṭaka] (Tainan: Heyu chubanshe 台南：和裕出版社 [Tainan: Heyu Publications], 2007), vol. 8, no. 246, p. 39. My translation.

Based on the preceding discussion, we realise that what Buddhabhadra and Śikṣānanda mean by "no place [the mirror images] come or go" is that there are no independent mirror images entering and exiting the mirror. In *The Best Years of Our Lives*, even though it seems that Peggy Stephenson and the attendant's images entered the mirror and provided more space for a front-to-back sharp state to cover, there are no reflections in the mirror that arise in an independent and inherent manner. My idea draws on Khyentse Norbu's Buddhist discussion of the mirror image: "You could say that your reflection in the mirror does not truly exist because it depends on your standing in front of it. If it were independent, then even without your face, there should be a reflection."[328] In the same way, the entities of their mirror images also lean on Stephenson and the attendant sitting in front of the mirror; without this condition, their mirror images cannot arise. As for this Buddhist account, the *non-existence* of the mirror images is not ascribed to the commonplace fact that the mirror-image representations are simply simulations, but ascribed to the idea that the mirror images themselves have no independent qualities. Since there are no mirror images that appear in their own right, no reflected depth appears in the mirror inherently. Even if the application of the mirror in deep focus seems to be able to uphold a wide range of sharpness, the non-independent arising of reflected depth destabilises the entity of the wide range.

Even if the object of study is not the mirror image, Buddhist mirror theory can still be applied to it. In accordance with Buddhabhadra and Śikṣānanda's analysis of the mirror's link with *no coming*, *no going*, it not only refers to the material mirror at a physical level, but also refers to the mind mirror at a philosophical and metaphorical level. The mind mirror reveals that all things, just like the mirror images, are compound mirages, since all of them cannot come or go in an independent and inherent manner. It is the moving mind—the mind of attachment to compound mirages—that makes objects seem to have independent entities. Yet, when the mind is settled—that is, when the mind is not disturbed by the concept of independent existence—we can be aware that appearances that we sense, whether they are the mirror images or not, are all void of inherent essence.

For example, Peggy Stephenson (by Stephenson in this context I mean the character, not her mirror image) enters and sits in the ladies' room. Yet, one can argue that she is not in this room. The entering and sitting of this picture-based character cannot arise in their own right, as their arising leans on the acting of the actress of flesh and bone and the recording of the movie camera. When the actress enters the room, sits in front of, and is captured by the movie camera, her corporeal appearance can be transformed into a picture-based appearance presented on the

[328] Khyentse Norbu, *What Makes You Not a Buddhist* (Boston and London: Shambhala Publications, Inc., 2007), pp. 62–63 (Note: Even if Dzongsar Jamyang Khyentse, also known as Khyentse, is the name on the front cover of the book, Khyentse is his first name. His surname is Norbu).

screen. Without these conditions, the entering and sitting of this picture-based character cannot happen. Moreover, these conditions also relate to the subsequent moment where this character seems to exit this room. I argue that she cannot leave this room in an independent and inherent way, considering that such picture-based leaving hinges on the corporeal actress performing an action of leaving and on the movie camera converting this leaving into the picture-based leaving. Without these conditions, there is no exiting of this picture-based character. Not just the character, no objects in this shot can come or go in their own right since all of their picture-based appearances need physical and filming conditions. Although this discussion could trigger an analysis of indexicality, I would like to use it as a foil for a Buddhist examination of dependent arising itself. What the viewer sees is dependent arising where there are no characters, surroundings, and depth existing in a truly inherent manner, let alone a fusion of composition in depth and a front-to-back sharp quality. When the viewer undergoes this Buddhist perceptual process, they can experience a basis that exists in the sense of dependent arising but without giving rise to any inherent appearances of that basis. Due to this void quality—a non-obstructing bare absence—there is no wide range of sharpness inherently findable on its side. This discussion is still related to the Buddhist principle of *no entering and exiting of the mirror images* even if I do not focus on the question of the character's mirror image.

Yet, as argued before, this does not signify that this deep-focus shot shows absolute nothingness. My discussion does not seek to completely obliterate every object and plane. Rather, *nothing comes and goes in its own right* is merely one facet of cognition. A void state is ascribed to the quieting of the mind of the concept of independent and inherent existence, allowing the viewer to engage serenely with a non-conceptual process of sensory consciousness. However, once the viewer does not sustain the quieting of the mind, every object and plane in this deep-focus shot will surface with the moving of the mind. The moving mind suggests understanding things through the ideas of things, strengthening the viewer's attachment to what they see (e.g. composition in depth and front-to-back sharpness). Such attachment regards the idea of a wide range of sharpness as a visible and knowable deep-focus feature that has relative existence. If the void were absolute, then even if the mind moves, there should be a void all the time. It is the conversion process between the moving and settled minds transforming this *composite* deep-focus feature into something and nothing, rather than nothing always. Buddhist examinations of the mirror images and the previously discussed frames in deep focus both lead to the nominal state between the construction and deconstruction of a wide range of sharpness.

5.4.4 Deep-focus Shots and the Continuity of Nominality

Another attribute of deep focus is long-take temporality. Because a deep-focus shot gives audiences more freedom to choose what to look at, it is usually connected to the long take in order to offer audiences enough time to view all of these sharp details. Yet, based on the previously outlined analyses by Bazin and Bordwell of the bar scene (figure 15) in *The Best Years of Our Lives*, the long-take tradition of deep focus is broken to some degree, as its continuity is cut twice by close-ups of Al Stephenson peeking at the phone booth. Despite these two cuts, there is still a continuous potentiality throughout this deep-focus scene, and even throughout the entire film. This continuous potentiality does not refer to long-take continuity giving audiences ample time to observe sharp objects, but rather the continuity of nominality. Its operation is not obstructed by an editing-based fragmentary manifestation, regardless of how many cuts intervene in this bar scene. Instead of examining the issue of time, the continuity of nominality will be adopted as an alternative Buddhist approach to deconstructing spatial depth, which will further call into question the idea of a wide range of sharpness. A Buddhist analysis of the continuity of nominality is distinct from my previous research on the continuity within montage carried out in sub-section 3.3.5.

In order to dissect the continuity of nominality in deep focus, I first look at Tian Maozhi's Buddhist examination of the surface of the table. Originally, Tian utilises the surface of the table as an example to explain the Buddhist dependent-origination world view. This explanation reveals how different viewing approaches make the perceived object have varied appearances and thus make it insubstantial:

> If we roughly observe the surface of the table, we will believe that it is a complete plane. If we closely look at it, we will be aware that it is not a plane since some of its parts are slightly higher, whilst some lower. If we see it through a microscope, we will perceive that the so-called surface of the table turns out to be a hilly-gully region. If we use a microscope with a higher magnification, we will recognise that the surface of the table becomes a heap of atoms.[329]

This statement brings about an ontological question: which one is the surface of the table? The one seen by the flesh eye at a distance, the one seen closely, the one under a microscope, or the one with a higher magnification?[330] In terms of the fourth one, I point out that just because a microscope with a higher magnification allows us to see what our eyes cannot see does not mean that the true essence of the surface of the table is a heap of atoms. From a Buddhist perspective, if a heap of atoms were its nature, it should retain atomic appearances always. That is to say, if we see it with our eyes from a distance, it should still be manifested in the form of a heap of atoms,

[329] This Buddhist discussion comes from Tian's preface to the *Jingang jing*. Ānanda, *Jingang jing*, p. 7.
[330] This Buddhist discussion comes from Tian's preface to the *Jingang jing*. Ānanda, *Jingang jing*, p. 7.

rather than a plane. The surface of the table does not possess an ultimate attribute, as distinct viewing approaches make the experienced object have multiple appearances. The above four viewing approaches serve as dependent origination from which a plane, slight undulation, a hilly-gully region, and atoms arise respectively. The surface of the table does not carry self-nature. Rather, it is just a compound illusion of dependent origination, whilst a plane, slight undulation, a hilly-gully region, and atoms are all conceptual labels added to that illusion.

The above discourse can be further applied to all phenomena. Each of them is in the state of non-self – or, in other words, *no unchangeable and permanent self*. Depending on what kinds of visual approaches engage with a phenomenon, it can be transformed into various appearances. However, this does not mean that every phenomenon is total nothingness. It still makes us experience a concrete sensation. The surface of the table, for example, can still make us feel the specific sensation of a plane, slight undulation, a hilly-gully region, or atoms. This sensation arises and we are not able to negate this arising. Consequently, from a Buddhist perspective, the surface of the table does not exist, but also exists, which, I argue, divulges its nominal quality. This nominality is continuous, regardless of whether we notice it. Nominality pervades all phenomena ceaselessly, and at no time can any of them be dissociated from it. Even if we can choose to ignore it, we cannot entirely erase it. Just because we can imagine a certain phenomenon as something carrying self-nature does not mean that its nominal nature can be interrupted and displaced by self-nature. When we believe in its self-nature, its nominal state is still present but in an implicit manner in view of the fact that we have not yet got insights into this nominal potentiality. When we juxtapose its existence with *non-existence*, nominal existence becomes conspicuous. Whether this nominal state is implicit or explicit, it does not cease. Yet, this does not signify that nominality is permanent. Rather, it suggests a conversion process between existence and *non-existence*. Therefore, nominality has already engaged with impermanence, throughout all phenomena.

In the above-mentioned bar scene of *The Best Years of Our Lives*, the viewer first sees a long shot, just like seeing the surface of the table with the flesh eye at a distance, in which Al Stephenson pretends to listen to the piano performance, but his attention is on Fred Derry. Afterwards, this long shot is followed by a close-up of Stephenson, which helps the viewer get an even closer look at Stephenson, just like seeing the surface of the table under a microscope. It seems that the viewer saw the character Stephenson. However, if he were truly on the screen, he should be able to maintain the same appearance all the time. Regardless of being seen in a long shot or a close-up, he should always be in an unchangeable and permanent state. Yet, the problem is that the close-up of Stephenson displays more details of the textures of his hair, ear, skin, clothes, whilst in the long shot, those extremely subtle textures become blurry or even disappear. This suggests that the character Stephenson cannot retain his permanent entity on the screen. Therefore, the viewer does

not truly see him. In the same way, though Derry still looks tiny since he is in the background, compared to the long shot, the close-up still brings him closer to the viewer and presents more details of him. In addition to Stephenson and Derry, other objects in these shots and everything throughout the entire film are subject to such impermanence. That is to say, what the viewer sees are devoid of existing in a truly inherent manner. Multiple viewing strategies (e.g. seeing through a long shot and seeing through a close-up) transform the experienced thing into distinct appearances, which makes it devoid of an unchangeable and permanent attribute.

Nevertheless, at a relative level, the viewer can still recognise these characters. Even though they do not have unchangeable and permanent existence, the viewer can still identify that the one standing beside the piano is Al Stephenson, and that the one in the phone booth is Fred Derry. They are not completely transformed into something else. I contend that when their relative existence and void overlap, we become aware of their nominality. In other words, the characters are themselves but not themselves, just like seeing the surface of the table as itself but not itself. Whether seeing through a long shot or a close-up, the viewer, at all times, neither truly sees things nor sees total nothingness. This nominally cinematic phenomenon is present throughout the bar scene, being continuous throughout the entire film. This continuity does not mean the spatiotemporal unity of the image in movement, but rather means the unity of the nominal existence of all phenomena in every shot. Moreover, even though the long shot were not cut to the close-up, nominality would still be in operation. If we have already situated ourselves in a nominality-oriented perceptual habit, we do not need the following close-up to further remind us of the impermanence (e.g. textural impermanence) of the objects. Heightened awareness of nominality, to a great degree, can motivate us to unhesitatingly realise that all things in the previous long shot are themselves but not themselves even if we have not yet seen the subsequent close-up presenting the accompanying micro-textural changes of the objects. Nominality does not just arise at the turning point between the previous long shot and the next close-up. Rather, it is throughout both.

Composing Al Stephenson on the close plane and composing Fred Derry on the distant one provide great depth for front-to-back sharpness to cover. Yet, based on the Buddhist idea of the continuity of nominality, since the existential qualities of the characters and surroundings seen, just like the surface of the table, fall apart, there are no inherently pictorial representations supporting composition in depth and therefore there is no place a wide range of sharpness comes. However, when we have not yet gained insights into the continuity of nominality, every object in this shot appears in the way it is usually seen. A wide range of sharpness is thus initiated as a visible and knowable deep-focus feature. Yet, when the continuity of nominality is taken into account, a wide

range of sharpness becomes involved in a superimposing state where its existence can be regarded as tangible and intangible.

5.4.5 Deep-focus Shots and Buddhist Dust Theory

My Buddhist analysis of the bar scene is not just concerned with the nexus of the continuity of nominality and textural change. I want to revisit "atom" taken from Tian Maozhi's statement on the surface of the table, as it offers alternative access to nominality in the bar scene. It will open up the discussion of the nominal size of the object and nominal spatial depth, which further calls into question the idea of a wide range of sharpness. I point out that the subsequent discussion will be quite distinct from my previous analysis of the size and depth in the family reunion scene.

Although the concept of the atom is relevant to the field of quantum mechanics, it also relates to—after it is divided into subatomic particles at a more microscopic level and approaches the state of an elementary particle—a Buddhist term called *wufen weichen* 无分微尘 (extremely tiny and indivisible dust).[331] In physics, an elementary particle refers to a subatomic particle without substructure. In other words, it is not made of other particles, often treated as the smallest building block of matter. Likewise, in Buddhology, especially for the Buddhist schools of the Small Vehicle, extremely tiny and indivisible dust is acknowledged as the tiniest grain – that is, the basis of matter. Moreover, its indivisible quality implies that it is a zero-volume unit taking up no space. Yet, in Sakya Paṇḍita Künga Gyeltsen's view, this Buddhist term exposes a paradox that dust cannot reach the state of being indivisible.[332] I will reconsider how this paradox occurs and associate it with the deep-focus bar scene in order to explain why there is no fundamental difference between the sizes of the objects and how a wide range of sharpness is disintegrated.

Before applying the dust paradox to the bar scene, a reference of Francisca Cho's lecture "Ritual Apparitions and a Buddhist Theory of Film" is helpful, as its content is pertinent to the discussion of dust and the sizes of the objects. Cho proposes that Buddhist ritual and contemplative practice can be directly analogised to cinematic images. She suggests that "the abstract concept of

---

[331] Sakya Paṇḍita Künga Gyeltsen, *Liangli baozang lun shi* 量理宝藏论释 [An Explanation of the Treasure of a Buddhist Theory of Standard Measurement] (It was completed by Sakya Paṇḍita Künga Gyeltsen, 1182–1251), trans. Khenpo Sodargye (Beijing: Zhongguo wenshi chubanshe 北京：中国文史出版社 [Beijing: Chinese Literature and History Press], 2014), vol. 1, pp. 102–8. My translation (Note: Despite the fact that the Buddhist term "extremely tiny and indivisible dust" could trigger the discussion of quantum Buddhology, I do not dissect this interdisciplinary field).

[332] Sakya Paṇḍita, *Liangli baozang lun shi*, pp. 102–8.

emptiness is transformed into an image of non-obstruction in which one can see the entire universe even in a particle of dust."[333] She then adds:

> The macrocosm within the microcosm of the Buddha-fields is limitless whilst simultaneously occupying a limited space. Because film is something that is seen, it exists in space. It takes up space. But unlike ordinary objects, cinematic images exist within the bounded space of a screen. […] Within the two-dimensional frame of a movie, moreover, there are limitless possibilities for what can be seen. We can observe things and go places that are otherwise inaccessible to us. And we can also see things constructed from the pure imagination. […] [The two-dimensional screen] can play with space and time. It can dilate it. Or it can shorten it, both time as well as our relation to space. So cinematic screens provide the macrocosm within the microcosm wherein the past, present, and future can be imaged in endless and ever-new ways.[334]

Yet, in my view, there are two problems with Cho's analogical study of the limited screen and the macrocosm within the microcosm of the Buddha-fields. First, strictly speaking, a visual analogy to the macrocosm within the microcosm is not to use a Buddhist theory to engage with film ontology, but rather to use the cinematic screen as a form of symbolic representation to delineate Buddhist ritual and contemplative practice in a figurative way. Second, Cho argues that a vividly visual parallel between Maitreya's tower containing within itself the totality of that landscape of towers and the restricted frame of the screen containing within itself the landscape shot in real life implies that the landscape can be freely lessened or expanded and thus has no fixed size.[335] Yet, I assert that there is no comparability between the size of the photographic object and the size of the real-life object, as they do not have common attributes. By this I mean that if we apply the macrocosm within the microcosm to the object, the premise of this application is that its micro- and macro-appearances need to be either both picture-based (two-dimensional) or both real life-based (three-dimensional). Based on equal attributes, we can then continue to dissect why the landscape, for example, has no fixed size, rather than compare this landscape's micro- (cinematographic) appearance to its macro- (real-life) appearance and accordingly think that the landscape possesses no fixed size.

In other words, if we look at the bar scene of *The Best Years of Our Lives*, we cannot simply state that the reason why both Al Stephenson and Fred Derry have no fixed sizes is because the two life-size actors can be proportionally lessened and contained within the small screen, or expanded and contained within the large screen. When applying the macrocosm within the microcosm to

---

[333] Francisca Cho, "Ritual Apparitions and a Buddhist Theory of Film" (https://www.youtube.com/watch?v=ueKIAtLfr4g This lecture was delivered at the Harvard Divinity School in Cambridge, Massachusetts, United States of America on 2nd April 2018).
[334] Cho, "Ritual Apparitions and a Buddhist Theory of Film."
[335] Cho, "Ritual Apparitions and a Buddhist Theory of Film."

these two picture-based characters, we should rather demonstrate why the picture-based characters themselves have no fixed sizes, rather than mix the picture-based characters with the real life-based actors. In this regard, readers may ask: "Is this because of linear perspective? For example, Derry appears tiny, as he is in the background. If he were in the foreground, he would become larger and therefore not necessarily be smaller than Stephenson." Yet, I argue that this still puts a three-dimensional concept into the two-dimensional figures. I utilise Sakya Paṇḍita's Buddhist discourse on dust to examine why the areas of these two picture-based characters are equal on the screen. This size-related analysis will lead to the void of spatial depth. De-spatialisation suggests that there is no depth for front-to-back sharpness to cover—the deconstruction of a wide range of sharpness. My approach to size will be quite distinct from Cho's argument and the principle of linear perspective.

Before examining Buddhist dust theory, I want to mention the Mandelbrot set—a mathematical idea similar to the macrocosm within the microcosm of the Buddha-fields. It dissects the dynamic sizes of patterns merely on the complex plane, rather than mixes them with the sizes of 3D figures in solid geometry.[336] This idea is also related to Buddhabrot since if we rotate the image of the Mandelbrot set 90 degrees clockwise, it resembles a sitting Buddha. Yet, I do not plan to delve into this concept because these geometrical patterns are akin to the meditative patterns of Buddhism in Jordan Belson's abstract animations. I will explore size at the level of Buddhology.

Sakya Paṇḍita considers that the object is made of dust particles and that each dust particle can be further divided into many smaller dust particles. He then raises a question: after multiple divisions, can we eventually get extremely tiny and indivisible dust? Assuming that it is tenable, we notice that there is a paradox. Regardless of how many extremely tiny and indivisible dust particles gather together, they are not capable of forming the physical object because their indivisible qualities have already divulged that they have no volume. When zero-volume units merge, the assembled object should also dissociate itself from volume. In other words, this compound should possess an indivisible quality as well, which contradicts the fact that the object can be divided into smaller dust particles. In this light, dust is infinitely divisible.[337]

By the same token, in the preceding bar scene, Al Stephenson and Fred Derry as two picture-based characters are also produced by tiny pictorial fragments – or, in other words, pixels if they are placed into a digital grid system.[338] These tiny squares, just like dust particles, are infinitely

[336] The complex plane is not the layer of a shot. Rather, in mathematics, the complex plane is the plane associated with complex coordinate system, formed or established by the real axis and the perpendicular imaginary axis.
[337] Sakya Paṇḍita, *Liangli baozang lun shi*, p. 104.
[338] I do not delve into the technical issue of digital imaging, but rather make a comparison between dust and pixel.

divisible because otherwise a similar paradox would arise that if indivisible pixels integrated, nothing could be formed on the screen, as their indivisibility suggested that they were zero-dimensional points taking up no area and thus the assembled character should also disengage from the area. In order to constitute the visible character, pixels need to have infinitely divisible qualities. Yet, there is a problem about a zero-dimensional point that needs to be clarified. We may tend to consider that at least three points that are not on the same line determine a plane figure, which seems to go against "nothing can be formed on the screen."[339] However, a precondition for the formation of a plane figure is that these points themselves must have sizes. This means that they are concrete, tangible, and divisible point-like plane figures. On the contrary, the word "point" that I utilise within this context refers to the point without a size. Thus, it cannot be divided. It can be viewed as an abstract and intangible concept. Regardless of how many of them integrate, they cannot constitute a tangible area.

Sakya Paṇḍita points out that though the infinitely divisible quality of dust supports the construction of the tangible and divisible object, it also produces an alternative contradictory situation where the small object can have the same size as the large object. For example, though a mountain looks much larger than a blade of grass, considering that both their numbers of dust particles are infinity, their volumes should be equivalent; they should occupy the same amount of space. This implies that though a blade of grass appears tiny, it can be as large as a mountain.[340] Likewise, in view of the fact that a pixel can be infinitely divided, an infinite number of them need to be taken in order to form the picture-based Al Stephenson, and the same goes for the picture-based Fred Derry. Even if the area of Stephenson in the bar scene looks much larger than the area of Derry, they both remain equal in the number of the pixels, which means that their areas should also be equivalent. Derry seems tiny, but he can be as large as Stephenson. By applying Buddhist dust theory to the image, I investigate the link between the image and the macrocosm within the microcosm without getting caught in the description of the real-life actors and the three-dimensional concept derived from linear perspective.

One of the approaches to the filmic expression of spatial depth is to place the characters on different planes. In the bar scene, the salient difference between the sizes of the characters (great depth) and front-to-back sharpness together reinforce a wide range of sharpness as a *composite* feature of deep focus. Yet, based on the above Buddhist view on the sizes of the characters, since they can take up the same amount of the perceiver's visual field, this suggests that they are composed on one plane to some degree, thereby dismantling the distance between them. As there

[339] The plane figure refers specifically to the geometrical plane in the field of mathematics, rather than the layer of a shot.
[340] Sakya Paṇḍita, *Liangli baozang lun shi*, p. 105.

is no great depth for front-to-back sharpness to cover, the existential quality of a wide range of sharpness is destabilised. However, Sakya Paṇḍita argues that the idea that all objects' sizes are equal is just one way to perceive sizes. We should not cling to the equality of the number of dust particles because otherwise we will be caught in an extreme situation in which every object's relative size is denied.[341] The identification of difference between sizes should also be considered a part of experience. At the finite level of division—not including indivisibility—the sizes of the characters are hugely different, which suggests that there is considerable depth for front-to-back sharpness to cover, thereby stabilising the existential quality of a wide range of sharpness. Depending on how we look at the divisible attribute of dust, this *composite* deep-focus feature can arise and cease.

5.4.6 Deep-focus Shots and *Tianyan* 天眼 (the Heavenly Eye)

Another approach to the nominality of the *holistic* characteristic of deep focus develops from the nominality of occlusion. By occlusion I refer to the fact that some objects are partially hidden by closer objects. In *The Best Years of Our Lives*, the aeroplane graveyard scene (figure 19) provides a good example. The occluding engines seem closer, whilst the ones occluded seem more distant. A dense placement of these engines suggests that the spatial volume of this aeroplane graveyard is extremely large and therefore provides more space for front-to-back sharpness to cover.

However, I consider that if the Buddhist perceptual process of *tianyan* 天眼 (the heavenly eye) engages with this deep-focus shot, there will be, to some degree, no inherent and substantial occlusion among the front, middle, and back engines. The perceptual ability of the heavenly eye is to make those supposedly obstructed parts of the engines visible and knowable. I would like to clarify that despite the fact that the flesh eye can utilise the knowledge of three-dimensional shapes to infer those occluded appearances that the flesh eye cannot see, the heavenly eye can see through all of these engines' full appearances without drawing a three-dimensional inference. The heavenly eye does not perceive an occluded appearance to be behind something or farther away, but rather perceives it to be exposed and thus perceives all faces to be on the same plane, X-ray vision as it were, but sees more clearly than an X-ray machine. This perception is beyond the physical limits of obstruction and thus nullifies spatial depth. I will carry out a further analysis of the operation of the heavenly eye and examine its ontological relationship to the occlusion in this deep-focus shot. This discussion will lead to the nominality of a wide range of sharpness.

[341] Sakya Paṇḍita, *Liangli baozang lun shi*, p. 105.

One of the most important Buddhist theoretical sources that engages with the heavenly eye is the *Mahāprajñāpāramitopadeśa* (also known as the *Treatise on the Great Prajñāpāramitā*, written by Nāgārjuna circa the third century, translated into Chinese in 405 as the *Da zhidu lun* 大智度论 by Kumārajīva, 344–413). With regard to the heavenly eye, Nāgārjuna performs an in-depth analysis of it by comparing it with the flesh eye. The flesh eye allows the observer to see the front of the object, whilst the perception of its back is impeded. In order to see its back, the observer needs to rotate the object a little bit or move around it. However, Nāgārjuna suggests that the heavenly eye goes beyond obstruction, expanding the perceiver's powers of observation. More specifically, "*de shi tianyan*, *yuanjin jie jian*, *qianhou neiwai*, *zhouye*, *shangxia*, *xijie wuai*." "得是天眼，远近皆见，前后内外，昼夜、上下，悉皆无碍。" (with the heavenly eye, the far is seen as the near; it allows the experiencer to directly see without hindrance the front from the back, the back from the front, see the inside from the outside, the outside from the inside, see day at night, night in the daytime, see the top from the bottom, the bottom from the top).[342] By this he does not suggest that the perceiver sees something made of a transparent material. In fact, even if the object is transparent, the flesh eye still cannot directly see its back since there is a transparent barrier between the flesh eye and the back of the object; the viewing of its back is no more than indirect in this situation. Yet, for the heavenly eye, even if the object is entirely opaque, it can still directly view its back without rotating or moving around it and without putting a mirror behind it to expose its back. In addition, with regard to the inside and outside, Nāgārjuna does not mean that the perceiver inside/outside the room sees the outside/inside through a window or hole. Even though it is a totally enclosed room, the heavenly eye still enables the perceiver to directly see the outside/inside of the room without obstruction, putting the inside and outside on the same plane.

Nāgārjuna's statement on the heavenly eye implies the shaping of the dimension of time. Usually, if the perceiver wants to see the different external and internal faces of an array of engines in the three-dimensional world, they need to spend some time moving around these engines and disassembling them in order to make direct observations. In other words, their different faces are manifested at different observational moments. However, the heavenly eye is able to shape time and thus to juxtapose multiple observational moments in time with each other, so that all of their different external and internal faces appear simultaneously even if there are no reflectors or assistant devices such as multiple mirrors or cameras to expose those occluded parts at the same time. For the flesh eye, it is quite difficult to reach the Buddhist state of time shaping. We are able to shape a plane figure by drawing or shape space by 3D modelling. Yet, in terms of time (not the

[342] Nāgārjuna, *Da zhidu lun* 大智度论 [The Treatise on the Great Prajñāpāramitā] (It was completed by Nāgārjuna circa the third century, translated into Chinese in 405 as the *Da zhidu lun* by Kumārajīva, 344–413) (Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], 1991), vol. 33, p. 228. My translation.

timeline on an editing platform), we usually obey it. Although we can feel its flowing state, we cannot touch it and thus cannot shape it.[343] The heavenly eye's perceptual process can be considered a collage of time. The different faces manifested at different observational moments are placed together on one *exposed* plane simultaneously.

In a similar way, when the heavenly eye sees the two-dimensional shot of an array of the aeroplanes' engines, it can also see all of the spatiotemporal details contained within this shot. As Nāgārjuna maintains, "*shi tianyan jian hehe yinyuan sheng jiaming zhi wu*." "是天眼见和合因缘生假名之物。" (The heavenly eye endows the perceiver with powers to penetrate [see through] the entire dependent-origination process of the object).[344] For instance, the perceiver is able to penetrate the past and to see the filming location *at that time* in which this shot was captured. This past filming location provided a holistic environment for the production of this shot. This environment contained multiple potential positions for the camera to be put at. The shot of these engines is just one observational moment extracted from that environment in time. Yet, with the heavenly eye, the perception of this shot is not restricted by this selected moment. It allows the perceiver to have a penetrating view of the dependent origination of this shot – that is, the past filming location in which there were countless possible observation points to be chosen. Each of them indicated a particular observational moment. These could-be viewpoints (positions) would be able to show the distinct faces of the engines. The time-shaping capability of the heavenly eye enables the perceiver to penetrate all of the could-be viewpoints involved in this past filming location, to make the could-be viewpoints into a collage, and thus to place the different faces of the engines on the same plane. When all of their occluded faces are exposed under the system of the heavenly eye, there is no occlusion among the front, middle, and back engines; without the occlusion, spatial depth is nullified, which further reveals that there is no wide range of sharpness in this deep-focus shot, as there is no great depth for front-to-back sharpness to cover.

Yet, the above discussion does not mean to deny a wide range of sharpness in this deep-focus shot. Rather, this examination suggests that the existence of this *composite* deep-focus feature does not have a permanent state. Because of distinct perceptual systems, a wide range of sharpness is subject to impermanence. On the one hand, the flesh eye cannot see through the engines. The occlusion among them provides great depth for front-to-back sharpness to cover. On the other hand, the heavenly eye is able to shape time and to penetrate the engines, which deconstructs great depth and hence destabilises a wide range of sharpness. By shifting back and forth between

[343] It appears that sound could not be touched as well, but the reason why we are still able to shape a work of sound art or music is because its vibrations can be touched by eardrums. Regarding time, any parts of the physical body cannot touch it and hence cannot shape it.
[344] Nāgārjuna, *Da zhidu lun*, p. 228.

the flesh eye and the heavenly eye, a wide range of sharpness in this deep-focus shot is initiated and extinguished. The change of perception causes the fluidity of spatial depth, which further generates a nominal wide range of sharpness.

5.4.7 Deep-focus Shots and Distraction

In addition to composition in depth, another important aspect of the discussion of deep focus is audiences' focus of attention. In sub-section 5.4.1, I summarised Bazin's analysis of the perceiving of the wedding scene (figure 20) in *The Best Years of Our Lives*. As we saw, Bazin presumes that the viewer is more curious about Fred Derry and Peggy Stephenson than a group of people on the right, considering that the love between the former two is still uncertain, even though the common sharp state of all planes gives equal weight to every object. I argue that Bazin's interpretation assumes that the viewer is always in an undisturbed environment for perceiving this shot. Yet, I suggest that sometimes the viewer can be distracted by various reasons. Although two would-be lovers appear to grab the viewer's attention and underline their sharpness, their in-focus nature can still be dissolved if the viewer's personal distractive factors are taken into account. Despite the fact that these factors are not related to the content of this shot, they are able to change the way the experienced objects are manifested.

In terms of distraction, the Dalai Lama suggests that "distraction is an interruption to placement of mind on an object."[345] He further uses the terms "flightiness" and "mental wandering" to paraphrase "distraction." In fact, he does not use these alternative expressions in a completely negative way. Rather, they are deployed to emphasise the weak state of the mind's mental hold on its object of focus. Such a weak state is caused by the mind of attachment to something else more attractive than its original object of focus.[346] A desire to engage with other things or actions leads to an interruption to concentration, which further transforms the experienced appearance of the original object of focus. For example, when seeing the shot of the wedding ceremony, the viewer may suddenly have some other personal matters weighing on the mind that need to be addressed or feel sleepy in a moment due to constant fatigue. Once these additional conditions happen, for the viewer, not only will a group of people on the right be out of focus, but also two would-be lovers on the left will be out of focus. With distraction, the supposedly common sharpness is subject to dissolution. Yet, this does not necessarily suggest that the flightiness of the mind makes the experienced shot entirely lacking in focus. From the Dalai Lama's perspective, the flightiness of the mind has distinct degrees. Specifically, he argues:

[345] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 135.
[346] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, pp. 135–36.

> There are two general levels of flightiness. The gross level is when mind loses the hold of its mindfulness on its focal object completely. The subtle level occurs when mind maintains its hold, but either has a subtle level of wandering beneath the surface, like the slow flow of water on the bottom of a frozen river, or is more steady but seemingly "itchy" to move.[347]

On the one hand, when seeing the shot of the wedding ceremony, if some personal matters occur to the viewer and they are extremely urgent, the viewer will desire to tackle them immediately, thereby diverting their attention from this deep-focus shot and causing its sharp quality to be in an implicit state. In addition, if the viewer feels exhausted, they will desire to rest, to close the eyes, and to stop seeing this shot. As the perception of this shot ceases, any objects at this moment cannot be manifested, let alone sharpness. Both situations imply the loss of the hold of the viewer's mindfulness on this shot. On the other hand, if the matters weighing on the mind are not urgent, even though they can produce distraction at the moment of springing to the mind, the viewer's attention will then soon be directed towards the shot of the wedding ceremony, as those non-urgent matters coming into the mind can be handled later. The re-attention to this shot motivates the viewer to notice Fred Derry and Peggy Stephenson because the uncertainty of their love makes the viewer curious. In addition, if the viewer feels a little drowsy, even though this feeling can weaken the viewer's attention to these two would-be lovers, the mind does not totally lose its hold on them, considering that the viewer's eyes are not completely closed. Both situations suggest the subtle loss of the hold of the viewer's mindfulness on this shot. It is the placement of the attention on objects giving rise to differences between explicit and implicit appearances. However, Buddhabhadra and Śikṣānanda also consider a situation where there is no placement of the attention. They suggest:

> By transcendence of all perceptions of form, by disappearance of all perceptions of objects, by not placing the attention on various perceptions, [enlightening beings] attain to and abide in […] [gradually] the realm of nothingness, aware of the absence of anything at all. By totally transcending the realm of nothingness, they attain to and abide in the realm of neither perception nor non-perception.[348]

When audiences do not perform the action of attention, neither the shot of the wedding ceremony nor other matters weighing on the mind appears as the form of a focal potentiality—the common out-of-focus quality that makes them become one. However, when being situated in the Buddhist state of nominal attention—an intermediate area between perception and *non-perception*—audiences do not abandon the action of attention, but also do not cling to any

[347] Berzin and the Dalai Lama, *The Gelug/Kagyü Tradition of Mahāmudrā*, p. 281.
[348] Buddhabhadra and Śikṣānanda, *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra*, p. 724.

perceived objects simultaneously. That is to say, by focussing on the object but not becoming attached to it, audiences not only maintain forms in all their diversity, but also transcend the realm of differences.

Chapter 6 Conclusion

6.1 Nominal Film and the Mind in Buddhology

A possible Buddhist answer to the question of what film ontology can be is nominal film. This term originates from my act of combining film ontology with the Buddhist idea of nominal existence, referring to an intermediate state between film and "no film." In order to recapitulate nominal film, I would like to first reiterate how I use a Buddhist methodology. To be precise, my theoretical approach is film Buddhology, which is distinct from what I term Buddhism-in-film. Film Buddhology means adopting Buddhist mind theory as a philosophical strategy to engage with the question of what film ontology can be, whilst Buddhism-in-film signifies Buddhist iconography and meditation represented in films as a particular belief system or as a practical way of living.

The mind is the key element of film Buddhology. Within a Buddhist context, the mind means experience such as the processes of seeing, hearing, considering, and emotionally feeling objects. Broadly speaking, Buddhist mind theory maintains that reality is a projection of the mind. This idea reveals that all phenomena are illusions since they are devoid of existing in an independent and inherent manner. The mind of attachment regards the object as able to manifest itself in its own right. However, the manifestation of the object needs to depend on the mind perceiving it. Without a perceiving process, the object cannot be manifested to the perceiver, but rather serves as a potentiality that is not yet able to appear.

The mind in Buddhology provides a particular way to reconsider the essence of all phenomena. Since film is a palpable phenomenon, I declare that the mind engages with the existential question of film at a Buddhist ontological level. In terms of the aspects of the definition of nominal film, "film does not exist" should not be misconstrued as "the representation of film—the world simulated on the screen—does not exist." In addition, this idea means neither that film stock decays, nor that film cannot establish itself as a unique medium because of its similarity to other media such as drawing, television, and video. Rather, within my Buddhist theoretical context, "film does not exist" is treated as a mind-related idea, indicating that all cinematic elements such as images, montage, voices, and deep-focus shots are devoid of existing in an independent and inherent manner because their existence needs to rely on the viewer's act of perceiving them. First, if cinematic elements existed in a truly inherent way, then even if the viewer did not see and hear them, they should still be able to manifest themselves to the viewer in their own right. Second, cinematic elements cannot retain a permanently existential state since their audio-visual appearances vary with the operation of the viewer's mind. The change of the mind makes appearances in a flowing state. There are no

inherently findable cinematic elements establishing their solid entities from their own side. At a Buddhist ontological level, the existence of film is extinguished.

"Film exists" signifies that the viewer's sense organs and sense consciousness together make the manifestation of film possible, it follows that film is not complete nothingness. To put it another way, being devoid of existing in an independent and permanent manner does not indicate that we deny what we perceive. Instead, the viewer is able to feel the sensation of film and therefore cannot negate the arising of this sensation. To some degree, film still has a relative entity. The more the viewer's mind is attached to what they perceive, the more they tend to apprehend the appearance of film to exist as an initiated object of cognition.

By moving between "film exists" and "film does not exist" (for the sake of brevity, they have been abbreviated to film and "no film"), we reach an intermediate state called nominal film. By retaining this threshold, on the one hand, the viewer apprehends cinematic elements to exist in the manner in which they appear to exist. On the other hand, there are no essential distinctions among these cinematic elements, as all of them are compound and variable appearances devoid of existing independently and inherently. Every object perceived is thus reduced to the void. The initiation and extinction of an infinite variety of filmic elements convert into each other. I examined nominal film from, but not limited to, the following four branches: nominal images, nominal montage, nominal voices, and nominal deep-focus shots, with each of them accompanied by a particular type of film. Not only can nominal film be viewed from different aspects, but also choosing different types of films as case studies opens up an opportunity to engage with the universality of film Buddhology.

I put the ontological issue of images into a Buddhist context of the mind. First, images do not have independent existence, for they need to depend on the viewer's act of seeing them. Whether the experienced images come from art cinema or other types of cinema—Hollywood cinema, for example—without the operation of visual experience, their visual appearances will not be able to manifest themselves to the viewer. Second, the stylistic characteristics of images do not have self-nature. It is the viewer's mind of judgement labelling what they see as "art cinema's images" or "Hollywood cinema's images." Once images dissociate themselves from the mind of judgement, their stylistic qualities will be reduced to an empty state. The foregoing two facets suggest that even if art cinema and Hollywood cinema did not use each other's visual features as references, there would still be no fundamental difference between art cinema's images and Hollywood cinema's images because both their images and even other types of images are all devoid of existing in an independent and inherent manner. Yet, this does not mean that images should be denied. Rather, the viewer is capable of having the sensation of images and thus cannot repudiate the arising of this sensation, which reveals that images still possess relative entities. In order to engage with the ontological issue of images, I picked Michelangelo Antonioni's art film *Blow-up* as an example, using

the branches of *xianliang* 现量 (perception), *biliang* 比量 (inference), and *feiliang* 非量 (specious perception and specious inference) to dissect the ontological relationship between images and the viewer's mind. It appears as though visual phenomena including props, characters, and scenery had the appearance of being separate entities. Yet, it turns out that all of them are merely compound appearances devoid of existing in an independent and inherent manner. I also proposed a dynamic Buddhist theoretical model of image viewing. This model pictures a nominal state of image viewing as a turbulent sensory field in which images and "no images" become one. Although the idea of "no images" is out of range of tangible cognition, it still acts as the empty potential of tangible images.

Regarding montage, I suggest that it is not fundamentally different from the long take because at a mind level, the long take carries the potential of a cut. This idea should not be misunderstood as a technical view that the components of the long take are stills which have already cut it into multiple static pieces. Rather, I point out that even if the long take is considered a continuous trace, audiences' minds can still cut it into slices. First, I used *yinguo biliang* 因果比量 (causality-inference) and the changeable operation of the mind to argue that on the one hand, the long take keeps its continuous state, but on the other hand, audiences' variable thoughts on the long take trigger the fragmentary meanings of the long take. This factor serves as one aspect of mind montage. Second, I employed the Buddhist notions of nondualism, the contents of experience, and impermanence to stress that different styles of experimental cinema—surrealist film, abstract film, the city symphony film, structural film, and the compilation film—are one, as they all possess nominality. Subsequently, I refocussed on the nominality of montage and the long take. The fragmentary meanings of the long take are not the only factor causing mind-oriented discontinuity within a continuous state. In order to analyse other factors, I selected Hollis Frampton's experimental film *Zorns Lemma* as a case study, using *xiang biliang* 相比量 (appearance-inference) and impermanent compounds to complement an account of mind montage (invisible montage). As for the first two-minute long-take clip of *Zorns Lemma*, I discussed its discontinuous potentiality in the context of a cinematic compound. By this I mean that what audiences experience is a compound of the black symbol and other accompanying cinematic objects and that the fitful disappearance and reappearance of the accompanying sound of reading cause the entire perceived compound to be subject to the changes from "the compound with the sound of reading" to "without it" to "with it…" The black symbol is not the only object that audiences sense in the first portion. Rather, it is the variation of the relation among objects creating a sense of discontinuity. In addition, I synthesised appearance-inference and the blink of an eye to examine the uncertainty of the number of cuts, combining appearance-inference with the meaning-generating process to dissect the disjointed meanings of the long take. I drew on the Buddhist view on unity to explain why there is a continuous flow (the mind long take) within a montage system. I also engaged with the ontological link between the sign-based shots and the wordless shots at the Buddhist level of inferential expectations. Finally, I recapitulated the discontinuity of the compound.

In terms of voices, first, even if a film has no vocal tracks, vocal interventions are still present due to audiences' interpretive voices. Second, the voice has no independent and permanent nature making itself fundamentally distinct from other voices, but there are still relative distinctions among them. I used *yinguo biliang* 因果比量 (causality-inference) to delve into the interventions of voices from audiences' experiences and the hetero-homogeneity of sounds. Then, I utilised the luminous mind, the Buddhist view on voices, and the variability of causality to emphasise that the vocal facets of six documentary modes—the poetic mode, the expository mode, the observational mode, the participatory mode, the reflexive mode, and the performative mode—can be considered alike, as they are all related to nominality. Subsequently, I chose Abbas Kiarostami's docufiction *Through the Olive Trees* as an instance, drawing on the Buddhist experiential process of vocal experience, the void, Buddhist echo-like phantoms, and the seventh and eighth consciousness to engage with the nominal qualities of the opening voices, off-screen voices, vocal echoes, and ending voices.

As for the idea of nominal deep-focus shots, it means that deep-focus shots do not exist, but also exist. However, I not only engaged with the nominal entity of deep focus itself, but also the nominal entities of its affiliated forms such as composition in depth and long-take continuity. Apart from a Buddhist analysis of the process of how deep focus is constructed and deconstructed at the level of the perception of focus itself, I also employed a Buddhist strategy to discuss this process at the level of the perception of its linked forms. Deep focus and its variations together form the holistic quality of deep focus. I used Buddhology to canvass the existential issue of this quality. Throughout the chapter of nominal deep-focus shots, first, I examined a Buddhist account of sensory consciousness and deconstructed the manifestation of deep focus by concentrating on the experiential process of the mere arising of deep focus and the seeing of it happening at each moment, without identifying what the viewer sees as deep focus. Second, I utilised a Buddhist visual aspect (eyes and eye consciousness) of five sense organs-perception (*wugen xianliang* 五根现量), *zixiang* 自相 (an individual and primordial appearance), and a Buddhist statement on the retina as theoretical bases to argue that even if selective focus seems to retain an inherently minimum in-focus quality, it will still be subject to a blurry state if a severely myopic individual's visual experience is taken into account. Third, I used the Buddhist idea of knowledge-associated intuitive perception to examine the fluidity of composition in depth and thus to raise a question of whether there is inherently great depth for front-to-back sharpness to cover – or, in other words, a question of whether the composite feature of a deep-focus shot, a wide range of sharpness, exists inherently. In order to engage with the existential question of a wide range of sharpness, I used William Wyler's classical Hollywood film *The Best Years of Our Lives* as an example and employed the Buddhist views on conceptual thoughts, mirrors, particles, the heavenly eye, and distraction to scrutinise deep focus' relations to the frames, mirror images, continuity, sizes, occlusion, and attention. All of these Buddhist aspects give rise to the nominal essence of a wide range of sharpness of deep focus.

6.2 Nominal Average Shot Lengths in Contemporary Hollywood Cinema

Although I analysed in the fifth chapter classical Hollywood cinema and (to a lesser extent) New Hollywood, it is also interesting in this context to consider contemporary Hollywood cinema. In my subsequent discussion, I will extend my Buddhist approach by briefly examining how nominality engages with contemporary Hollywood cinema. In "Intensified Continuity: Visual Style in Contemporary American Film," David Bordwell points out that "the faster cutting rate, the bipolar extremes of lens lengths, and the reliance on tight singles are the most pervasive features of intensified continuity: virtually every contemporary mainstream film will exhibit them."[349] There is a general trend towards a decrease in the average shot length (ASL) in contemporary Hollywood cinema, considering that this type of cinema is usually accompanied by an increase in the speed of action. A typical example is *Fast and Furious Presents: Hobbs and Shaw* (David Leitch, 2019), with the film's title speaking for itself. Especially in terms of its car chase and fight scenes, the application of a shorter ASL creates a fast-paced, action-packed, and tense atmosphere.

Yet, I argue that there is no decrease in the perceived ASL in contemporary Hollywood cinema. I would like to emphasise that my perspective is not based on the commonplace fact that some of this type of cinema do not fit with the trend towards a decrease in the ASL,[350] but rather based on a philosophical point where a shorter ASL can be equal to a longer ASL. Before analysing this equivalence, I would like to first look at a Buddhist term called *wufen chana* 无分刹那 (an extremely short and indivisible instant) since it is a significant Buddhist idea relating to time.[351] It is quite similar to *wufen weichen* 无分微尘 (extremely tiny and indivisible dust) discussed in the fifth chapter, except that the latter deals with the amount of space/area, rather than the amount of time. Likewise, especially for the Buddhist schools of the Small Vehicle, an extremely short and indivisible instant is treated as the shortest unit of time, with its indivisible quality signifying that it is zero duration occupying no time. However, in Sakya Paṇḍita Künga Gyeltsen's view, this term produces a paradoxical effect in which an instant is not capable of truly attaining an indivisible state.[352] The length of time is composed of a myriad of instants and each instant can be further split

[349] David Bordwell, "Intensified Continuity: Visual Style in Contemporary American Film" (2002), in *Film Quarterly*, ed. Ann Martin (Berkeley: University of California Press, 2002), vol. 55, no. 3, p. 20.
[350] Examples include *Gravity* (Alfonso Cuarón, 2013) and *Birdman* (Alejandro González Iñárritu, 2014) notable for their long-take approach. I would like to clarify that even if these two films have implicit references to Buddhism, I still want to mention them since they are typical examples of an increase in the ASL in contemporary Hollywood cinema. Yet, my idea of "no decrease in the perceived ASL" is not ascribed to the advent of films like *Gravity* and *Birdman*. In my subsequent discourse, I will explain the reason for this idea at the level of Buddhology.
[351] Sakya Paṇḍita Künga Gyeltsen, *Liangli baozang lun shi* 量理宝藏论释 [An Explanation of the Treasure of a Buddhist Theory of Standard Measurement] (It was completed by Sakya Paṇḍita Künga Gyeltsen, 1182–1251), trans. Khenpo Sodargye (Beijing: Zhongguo wenshi chubanshe 北京：中国文史出版社 [Beijing: Chinese Literature and History Press], 2014), vol. 1, pp. 130–38. My translation.
[352] Sakya Paṇḍita, *Liangli baozang lun shi*, pp. 130–38.

into numerous shorter instants. Based on this issue, he raises a question: after multiple divisions, can we eventually get an extremely short and indivisible instant? Assuming that it is tenable, we become aware that regardless of how many extremely short and indivisible instants are added together, they are not able to constitute a longer length of time, as their indivisible qualities have already suggested that they occupy no time. When zero-duration units aggregate, the assembled length of time should also dissociate itself from duration. That is to say, this assembled length of time should have an indivisible quality as well, which is antithetical to the fact that the length of time can be split into shorter instants. Accordingly, from this Buddhist perspective, an instant is infinitely divisible.[353]

Similar to an account of the infinite divisibility of a dust particle, which I discussed previously, Sakya Paṇḍita argues that despite the fact that the infinite divisibility of an instant buttresses the formation of the palpable and divisible length of time, it also causes an additionally paradoxical effect in which short duration can have the same length as long duration. Considering that both short duration and long duration are made of infinite units of time, their lengths should be equivalent; they should occupy the same amount of time. This signifies that whilst an instant seems short, it can be as long as an aeon.[354] Likewise, with respect to the ontological relationship between a shorter ASL and a longer ASL, since an instant can be infinitely divided, an infinite number of them need to be taken in order to form a palpably shorter ASL, and the same goes for a longer ASL. Even if Bordwell stresses that there is a general trend towards a decrease in the ASL in contemporary Hollywood cinema, I argue that it can also be regarded as an increase in the ASL, as both remain equal in the number of the units of time. Therefore, from this Buddhist perspective, there is no truly inherent decrease in the ASL making itself fundamentally different from an increase in the ASL, and vice versa. By applying Buddhist time theory to the ASL in contemporary Hollywood cinema, I deconstruct the concept of a decrease in the ASL. It is also noteworthy that at the Buddhist level of the ASL, a general trend towards a decrease in the ASL in contemporary Hollywood films are not at variance with a long-take form used in other types of films such as Alfred Hitchcock's thriller *Rope* (1948) and Alexander Sokurov's historical drama film *Russian Ark* (2002).

Yet, Sakya Paṇḍita further suggests that the idea that all lengths of time are equivalent is merely one perspective to experience time. We should not be attached to the uniformity of the number of the units of time because otherwise we will be caught in an extreme situation in which every period's relative length is negated.[355] The recognition of difference between lengths should also be viewed as an aspect of experience. At the finite level of division—not including

[353] Sakya Paṇḍita, *Liangli baozang lun shi*, p. 135.
[354] Sakya Paṇḍita, *Liangli baozang lun shi*, pp. 135–36.
[355] Sakya Paṇḍita, *Liangli baozang lun shi*, p. 136.

indivisibility—a shorter ASL and a longer ASL are significantly different, which suggests that Bordwell's view on a general trend towards a decrease in the ASL as a salient feature of contemporary Hollywood cinema should not be rejected. Depending on how we engage with the divisible nature of an instant, this general trend can appear and disappear, which makes itself distinct and not distinct from an increase in the ASL—in the sense of the nominal ASL.

Building on the above discourse, I would like to suggest that nominal ASLs also affect the way we understand nominal images, nominal montage, nominal voices, and nominal deep-focus shots. At the level of nominal images, on the one hand, as a short ASL is identical to a long ASL, each shot's images that appear to pass quickly should be regarded as generally slowed down in order to match a long ASL, but on the other hand, since they are unequal, the mind keeps image-movements the way they usually look. The variable operation of the mind changes the perceived images into slow and fast appearances. Images themselves hence have no constant entities, but at the same time, the experience of images in a usual sense is not totally denied.

Likewise, as for nominal montage, due to the equivalence between a short ASL and a long ASL, shots as the units of montage can be viewed as generally extended infinitely and hence seen as extremely long takes, and extremely long takes can be viewed as generally reduced infinitely and hence seen as the units of montage. Yet, considering that the viewer feels the sensation of difference between a short ASL and a long ASL, the arising of this sensation cannot be negated. This suggests that when seeing the units of montage, discontinuity is the way the viewer senses it. We do not destroy the presence of discontinuity, but at the same time, we are aware that the equivalence between a short ASL and a long ASL makes discontinuity have the potential of continuity, which implies that there are no inherently findable units of montage establishing their solid entities from their own side. Rather, the viewer's identification of the presence of discontinuity merely constructs the relative entities of the units of montage, thus concealing their potential of being extended infinitely.

I contend that when the above Buddhist idea of the length of time is applied to the existential issue of the vocal waveform length (VWL), we find out that an infinite number of the units of the VWL need to be taken in order to form a short VWL, and the same goes for a long VWL. Due to the equality between a short VWL and a long VWL, when the spectator hears a loud and clear voice in a very short time, this voice should be regarded as slowed down and as low-pitched and nasal in order to match a long VWL. Conversely, when the spectator hears a loud and clear voice continuing for a long time, this voice should be regarded as sped up and as high-pitched and squeaky in order to match a short VWL. It seems that the spectator's hearing caught a voice, but there is no inherently detectable voice establishing its permanent entity from its own side, since it can be understood as either lowered or raised due to the equality between a short VWL and a long VWL.

Yet, readers may ask: “Considering that at a technical level, pitch correction on an editing platform allows a VWL to be extended or reduced without tonal distortion, is the vocal tone truly deconstructed in the sense of the equality between a short VWL and a long VWL?” I argue that even if pitch correction brings back the original tone to some degree, the corrected audio quality is disjointed because it needs to add gaps or cut parts of the audio spectrum in order to match a long or short VWL. However, being devoid of existing in a truly permanent manner does not mean that voices are negated. The spectator still hears voices, recognising that a short VWL is still somewhat incompatible with a long VWL.

When the temporal equivalence is applied to a deep-focus shot with a long-take form, we notice that because the long take is identical to the extremely short take and because the extremely short take cannot provide audiences with enough time to see the sharp details of all objects, a wide range of sharpness should not be able to be perceived. Even if this feature seems to be what audiences experience, there should be no sufficient time for them to confirm the entity of this feature. Yet, at a relative level, the long take is different from the extremely short take. The long take is not totally viewed as sped up. Hence, there is still ample time for audiences to sense the sharp details of all planes and therefore to recognise and verify a wide range of sharpness.

As I have argued in my thesis, film Buddhology as a theoretical approach to film ontology transcends any types or genre of film, with the idea of nominal film being able to engage with a wide variety of filmic elements. Each of my case studies not only highlights a particular branch of nominal film, but, as a whole, also contributes to the discussion of an intermediate state between film (film exists) and “no film” (film does not exist). Furthermore, my thesis requires openness to further theoretical reflection crucial to increasing our engagement with a Buddhist methodology. It is still an ongoing process for me to explore the Buddhist applicability to other types of films and cinematic elements.

Bibliography

Aitken, Ian. *European Film Theory and Cinema: A Critical Introduction*. Bloomington and Indianapolis: Indiana University Press, 2001.

Altman, Rick. "The Material Heterogeneity of Recorded Sound" (1992). In *Sound Theory, Sound Practice*, edited by Rick Altman, pp. 15–31. New York and London: Routledge, 1992.

Amoghavajra. *Renwang huguo bore boluomiduo jing* 仁王护国般若波罗蜜多经 [The Perfection of Wisdom Sūtra for Humane Kings Protecting Their Countries]. In the *Dazheng xinxiu dazang jing* 大正新修大藏经 [The Taishō Tripiṭaka], edited by Dazheng xinxiu dazang jing kanxinghui 大正新修大藏经刊行会 [Society for the Publication of the Taishō Tripiṭaka]. Tainan: Heyu chubanshe 台南：和裕出版社 [Tainan: Heyu Publications], vol. 8, no. 246 (2007).

Ānanda. *Jingang jing* 金刚经 [The Diamond Sūtra] (494 BCE). In an annotated edition of the *Jingang jing*, edited by Tian Maozhi. Zhengzhou: Zhongzhou guji chubanshe 郑州：中州古籍出版社 [Zhengzhou: Zhongzhou Ancient Books Publishing House], 2007; *The Diamond Sūtra: A General Explanation of the Vajracchedikā Prajñāpāramitā Sūtra*. Commented on by Xuanhua (the pinyin form of Hsüan Hua [Hsüan Hua is a Wade–Giles-based name]). Translated by Heng Zhi (the pinyin form of Heng Chih [Heng Chih is a Wade–Giles-based name]). San Francisco: Sino-American Buddhist Association, Inc., 1974.

Antonioni, Michelangelo. *Sei film: Le amiche, Il grido, L'avventura, La notte, L'eclisse, Deserto rosso*. Turin: Giulio Einaudi Editore, 1964.

Antonioni, Michelangelo. "Preface to *Six Films*" (1964). In *The Architecture of Vision: Writings and Interviews on Cinema*, edited by Carlo Di Carlo, Giorgio Tinazzi, and Marga Cottino-Jones. Translated by Allison Cooper, Dana Renga, and Andrew Taylor, pp. 57–68. New York: Marsilio Publishers, 1996.

Asaṅga. *She dacheng lun shiqin shi jizhu* 摄大乘论世亲释集注 [Vasubandhu's Explanation of the Mahāyāna Compendium/Summary] (circa 310–390). Commented on by Zhimin. Edited by Li Mingquan. Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], vol. 5 (2004).

Astruc, Alexandre. "The Birth of a New Avant-garde: La Caméra-stylo" (1948). In *The New Wave: Critical Landmarks*, edited and translated by Peter Graham, pp. 17–23. New York: Doubleday, 1968.

Balázs, Béla. *Theory of the Film: Character and Growth of a New Art*. Edited by Herbert Marshall. Translated by Edith Bone. London: Dennis Dobson Ltd, 1952.

Bazin, André. *Bazin at Work: Major Essays and Reviews from the Forties and Fifties*. Edited by Bert Cardullo. Translated by Alain Piette and Bert Cardullo. New York and London: Routledge, 1997.

Bazin, André. *What Is Cinema?*. Edited and translated by Hugh Gray. Berkeley, Los Angeles, and London: University of California Press, vol. 1 (1967).

Bazin, André. "On the Politique des Auteurs" (1957). In *Cahiers du Cinéma in English*, edited by Andrew Sarris. Translated by Rose Kaplin, G. Hersberger, and John K. E. Hitchcock. New York: Cahiers Publishing Company, no. 1 (1966): pp. 8–18.

Becchio, Bruno. *Encyclopedia of World Religions*. Edited by Johannes Schadé. Yardville: Foreign Media Group, 2006.

Bershen, Wanda. "Zorns Lemma" (1971). In *Artforum*, edited by Annette Michelson. New York: Anthology Film Archives, vol. 10, no. 1 (1971): pp. 41–45.

Berzin, Alexander. "Lorig: Ways of Knowing." In Study Buddhism: A Project of Berzin Archives e. V., founded by Alexander Berzin, page number unknown. https://studybuddhism.com/en/advanced-studies/science-of-mind/ways-of-knowing/lorig-ways-of-knowing.

Berzin, Alexander and the Dalai Lama. *The Gelug/Kagyü Tradition of Mahāmudrā*. Translated by Alexander Berzin. Ithaca: Snow Lion Publications, 1997.

Biarese, Cesare and Aldo Tassone. *I film di Michelangelo Antonioni*. Rome: Gremese Editore, 1985.

Bodhiruci (ed.). *Dabaoji jing* 大宝积经 [The Mahāratnakūṭa Sūtra]. Shanghai: Shanghai foxue shuju 上海：上海佛学书局 [Shanghai: Shanghai Buddhist Books], vol. 57 (2004).

Bordwell, David. *Narration in the Fiction Film*. Madison: University of Wisconsin Press, 1985.

Bordwell, David. *The Way Hollywood Tells It: Story and Style in Modern Movies*. Berkeley, Los Angeles, and London: University of California Press, 2006.

Bordwell, David. "Intensified Continuity: Visual Style in Contemporary American Film" (2002). In *Film Quarterly*, edited by Ann Martin. Berkeley: University of California Press, vol. 55, no. 3 (2002): pp. 16–28.

Bordwell, David. "The Art Cinema as a Mode of Film Practice" (1979). In *The European Cinema Reader*, edited by Catherine Fowler, pp. 94–102. London and New York: Routledge, 2002.

Bordwell, David and Kristin Thompson. *Film Art: An Introduction*. New York: McGraw-Hill, 2008.

Bordwell, David, Janet Staiger, and Kristin Thompson. *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*. London: Routledge, 1985.

Broughton, Jeffrey. *Zongmi on Chan*. New York: Columbia University Press, 2009.

Brunette, Peter. *The Films of Michelangelo Antonioni*. Cambridge, New York, and Melbourne: Cambridge University Press, 1998.

Bruzzi, Stella. *New Documentary*. London and New York: Routledge, 2006.

Buddhabhadra. *A Buddha within: The Tathāgatagarbha Sūtra: The Earliest Exposition of the Buddha-nature Teaching in India* [Da fangdeng rulai zang jing 大方等如来藏经]. Translated by Michael Zimmermann. Hachiōji: The International Research Institute for Advanced Buddhology at the Sōka University, 2002.

Buddhabhadra and Śikṣānanda. *The Flower Ornament Scripture: A Translation of the Avataṃsaka Sūtra* [Huayan jing 华严经]. Translated by Thomas Cleary. Boston: Shambhala Publications, Inc., 1993.

Cagle, Chris. "The Voice in Documentary Sound Design: A Digital Revolution" (2019). In *Vocal Projections: Voices in Documentary*, edited by Annabelle Honess Roe and Maria Pramaggiore, pp. 187–202. New York and London: Bloomsbury Publishing Plc, 2019.

Carroll, Noël (ed.). *Theorizing the Moving Image*. Cambridge, New York, and Melbourne: Cambridge University Press, 1996.

Chatman, Seymour. *Antonioni: Or, the Surface of the World*. Berkeley, Los Angeles, and London: University of California Press, 1985.

Chen, Bing. *Fojiao xinlixue* 佛教心理学 [Buddhist Psychology]. Xi'an: Shaanxi shifan daxue chubanzongshe 西安：陕西师范大学出版总社 [Xi'an: Shaanxi Normal University General Publishing House], vol. 1 (2018), vol. 2 (2018).

Chen, Shuman. "Buddha-nature of Insentient Beings" (2014). In *Encyclopaedia of Psychology and Religion*, edited by David Leeming, pp. 273–77. Cham: Springer, 2020.

Cherchi Usai, Paolo. *The Death of Cinema: History, Cultural Memory, and the Digital Dark Age*. London: British Film Institute, 2001.

Chion, Michel. *Audio-vision: Sound on Screen* (1990). Edited and translated by Claudia Gorbman. New York: Columbia University Press, 1994.

Chion, Michel. *The Voice in Cinema* (1982). Edited and translated by Claudia Gorbman. New York: Columbia University Press, 1999.

Cho, Francisca. "Ritual Apparitions and a Buddhist Theory of Film." This lecture was delivered at the Harvard Divinity School in Cambridge, Massachusetts, United States of America on 2nd April 2018. https://www.youtube.com/watch?v=ueKIAtLfr4g.

Choephel, Lobsang and Zilkar Rinpoche. *Xinlei xue* 心类学 [The Types of Awareness]. Edited by Jiandi. https://www.lamrimworld.org/treasure_view.php?news_view_sn=69&sn=7, 2006.

Deleuze, Gilles. *Cinema 1: The Movement-image* (1983). Translated by Hugh Tomlinson and Barbara Habberjam. Minneapolis: University of Minnesota Press, 1997.

Deleuze, Gilles. *Cinema 2: The Time-image* (1985). Translated by Hugh Tomlinson and Robert Galeta. Minneapolis: University of Minnesota Press, 1997.

Dharmapāla. *Cheng weishi lun* 成唯识论 [The Discourse on the Perfection of Consciousness-only] (550). Taipei: Laogu wenhua shiye gufen youxian gongsi 台北：老古文化事业股份有限公司 [Taipei: Laoku Culture Foundation Inc.], vol. 2 (2014).

Dignāga. *Jilianglun lüejie* 集量论略解 [The Compendium of Validities with Brief Explanation] (circa 480–circa 540). Commented on by Fazun. Beijing: Zhongguo shehui kexue chubanshe 北京：中国社会科学出版社 [Beijing: China Social Sciences Press], vol. 1 (1982).

Drew, Robert. "Narration Can Be a Killer" (1983). In *Imagining Reality: The Faber Book of Documentary*, edited by Mark Cousins and Kevin Macdonald, pp. 271–73. London: Faber and Faber Ltd, 1996.

Fan, Victor. *Cinema Approaching Reality: Locating Chinese Film Theory*. Minneapolis: University of Minnesota Press, 2015.

Fan, Victor. "Illuminating Reality: Cinematic Identification Revisited in the Eyes of Buddhist Philosophies" (2019). In *The Structures of the Film Experience by Jean-Pierre Meunier: Historical Assessments and Phenomenological Expansions*, edited by Julian Hanich and Daniel Fairfax. Translated by Daniel Fairfax, pp. 245–58. Amsterdam: Amsterdam University Press, 2019.

Fei, Mu. "Lüe tan 'kongqi'" "略谈'空气'" ["Let's Talk about 'Air' Briefly"] (1934). In *Shidai dianying* 时代电影 [Modern Cinema], edited by Zong Weigeng and Gong Tianyi. Shanghai: Shidai tushu gongsi 上海：时代图书公司 [Shanghai: Modern Publications], vol. 1, no. 6 (1934): p. 22.

Fei, Mu. "*Xiangxuehai* zhong de yige xiao wenti: 'Daoxu fa' yu 'xuanxiang' zuoyong" "《香雪海》中的一个小问题：'倒叙法'与'悬想'作用" ["A Small Problem in *Sea of Fragrant Snow*: 'Inverse Narration' and the Functions of 'Suspension-imagination'"] (1934). In *Yingmi zhoubao* 影迷周报 [Film Fan Weekly News], edited by Xin Xin and Zhou You. Shanghai: Qianqiu chubanshe 上海：千秋出版社 [Shanghai: Qianqiu Publications], vol. 1, no. 5 (1934): p. 82.

Gidal, Peter. "Theory and Definition of Structural/Materialist Film" (1975). In *Structural Film Anthology*, edited by Peter Gidal, pp. 1–21. London: British Film Institute, 1978.

Guṇabhadra. *Lengqie jing* 楞伽经 [The Scripture of the Descent into Laṅkā]. Commented on by Zhengshou. Edited by Puming. Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], vol. 1 (2018).

Guṇabhadra. *Za ahan jing* 杂阿含经 [The Mixed Āgama]. Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], vol. 48 (1995).

Gyatso, Jampa. *Yinmingxue qimeng* 因明学启蒙 [The Enlightenment of Buddhist Logico-epistemology] (1825–1901). Beijing: Zhongguo shehui kexueyuan shijie zongjiao yanjiusuo 北京：中国社会科学院世界宗教研究所 [Beijing: The Institute of World Religions at the Chinese Academy of Social Sciences], vol. 4 (1983).

Jacobs, Steven, Anthony Kinik, and Eva Hielscher. "Introduction: The City Symphony Phenomenon 1920–40" (2019). In *The City Symphony Phenomenon: Cinema, Art, and Urban Modernity between the Wars*, edited by Steven Jacobs, Anthony Kinik, and Eva Hielscher, pp. 3–42. New York and Abingdon: Routledge, 2019.

Kovács, András Bálint. *Screening Modernism: European Art Cinema, 1950–1980*. Chicago and London: University of Chicago Press, 2007.

Kozloff, Sarah. *Invisible Storytellers: Voice-over Narration in American Fiction Film*. Berkeley, Los Angeles, and London: University of California Press, 1988.

Kracauer, Siegfried. *Theory of Film: The Redemption of Physical Reality*. New York: Oxford University Press, 1960.

Kumārajīva. *The Vimalakīrti Nirdeśa Sūtra* [Weimojie suoshuo jing 维摩诘所说经] (Note: Even if the title of the book is *The Vimalakīrti Sūtra*, *The Vimalakīrti Nirdeśa Sūtra* is the full title of this sūtra). Edited by Alex Wayman. Translated by Burton Watson. New York: Columbia University Press, 1997.

Le Grice, Malcolm. “Abstract Film and Beyond” (1977). In *Structural Film Anthology*, edited by Peter Gidal, pp. 22–27. London: British Film Institute, 1978.

Lin, Niantong. *Jingyou* 镜游 [The Roaming Lens]. Hong Kong: Suye chubanshe 香港：素叶出版社 [Hong Kong: Suye Publications], 1985.

Lucie-Smith, Edward and Clement Greenberg. “Interview Conducted by Edward Lucie-Smith” (1968). In *Clement Greenberg: The Collected Essays and Criticism, Volume 4: Modernism with a Vengeance, 1957–1969*, edited by John O’Brian, pp. 277–82. Chicago and London: University of Chicago Press, 1993.

MacDonald, Scott. *Avant-garde Film: Motion Studies*. Cambridge and New York: Cambridge University Press, 1993.

Maitreya. *Yujia shidi lun* 瑜伽师地论 [The Treatise on the Foundation for Yoga Practitioners] (circa 270–350). Edited by Li Li’an. Xi’an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi’an: Northwest University Press], vol. 15 (2005), vol. 63 (2005).

Marks, Laura. *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*. Durham and London: Duke University Press, 2000.

McLane, Betsy. *A New History of Documentary Film*. New York and London: Continuum International Publishing Group, 2012.

Merleau-Ponty, Maurice. *Phenomenology of Perception* (1945). Translated by Colin Smith. London and New York: Routledge, 2002.

Metz, Christian. “Aural Objects” (1975). Translated by Georgia Gurrieri. In *Cinema/Sound, Yale French Studies*, edited by Rick Altman. New Haven: Yale French Studies, no. 60 (1980): pp. 24–32.

Meynard, Thierry. “Liang Shuming and His Confucianized Version of Yogācāra” (2014). In *Transforming Consciousness: Yogācāra Thought in Modern China*, edited by John Makeham, pp. 201–41. Oxford and New York: Oxford University Press, 2014.

Mulvey, Laura. *Citizen Kane*. London and New York: Bloomsbury Publishing Plc, 1992.

Mulvey, Laura. *Death 24x a Second: Stillness and the Moving Image*. London: Reaktion Books Ltd, 2006.

Murch, Walter. *In the Blink of an Eye: A Perspective on Film Editing*. Beverly Hills: Silman-James Press, 2001.

Nāgārjuna. *Da zhidu lun* 大智度论 [The Treatise on the Great Prajñāpāramitā] (circa the third century). Shanghai: Shanghai guji chubanshe 上海：上海古籍出版社 [Shanghai: Shanghai Ancient Books Publishing House], vol. 33 (1991).

Nāgārjuna. *Zhong lun* 中论 [The Root Verses on the Middle Way] (circa the third century). Commented on by Vimalākṣa. Edited by Yan Yongkui. Chengdu: Zhongguo chuantong wenhua yanjiusuo 成都：中国传统文化研究所 [Chengdu: The Institute of Chinese Traditional Culture], 1995; *Mūlamadhyamakakārikā of Nāgārjuna: The Philosophy of the Middle Way*. Translated by David Kalupahana. Delhi: Motilal Banarsidass, 1999.

Nāgārjuna. *Qishi kongxing lun* 七十空性论 [The Seventy Verses on Emptiness] (circa the third century). Commented on by Nāgārjuna. In the *Zhongguan zhuangyan lun shi* 中观庄严论释 [An Explanation of a Filigree of the Middle Way], edited by Guo xiaofei and Tang Ao. Lhasa: Xizang zangwen guji chubanshe 拉萨：西藏藏文古籍出版社 [Lhasa: Tibetan Ancient Books Publishing House], 2019; *Nāgārjuna's "Seventy Stanzas": A Buddhist Psychology of Emptiness*. Commented on by Sonam Rinchen. Translated by Tenzin Dorjee and David Komito. Ithaca: Snow Lion Publications, 1987.

Naremore, James. "Style and Meaning in *Citizen Kane*" (1978). In *Orson Welles's Citizen Kane: A Casebook*, edited by James Naremore, pp. 123–59. New York: Oxford University Press, 2004.

Nichols, Bill. *Introduction to Documentary*. Bloomington and Indianapolis: Indiana University Press, 2001.

Norbu, Khyentse. *What Makes You Not a Buddhist*. Boston and London: Shambhala Publications, Inc., 2007 (Note: Even if Dzongsar Jamyang Khyentse, also known as Khyentse, is the name on the front cover of the book, Khyentse is his first name. His surname is Norbu).

Ogle, Patrick. "Technological and Aesthetic Influences on the Development of Deep-focus Cinematography in the United States" (1972). In *Movies and Methods: An Anthology*, edited by Bill Nichols. Berkeley, Los Angeles, and London: University of California Press, vol. 2 (1985): pp. 58–83.

Peterson, James. "Is a Cognitive Approach to the Avant-garde Cinema Perverse?" (1996). In *Post-theory: Reconstructing Film Studies*, edited by David Bordwell and Noël Carroll, pp. 108–29. Madison and London: University of Wisconsin Press, 1996.

Pramaggiore, Maria and Tom Wallis. *Film: A Critical Introduction*. London: Laurence King Publishing Ltd, 2011.

Qu, Ruji. *Zhiyue lu* 指月录 [The Records of Pointing to the Moon] (1602). Chengdu: Sichuan chuban jituan bashu shushe 成都：四川出版集团巴蜀书社 [Chengdu: Sichuan Bashu Publishing House Ltd], vol. 28 (2005).

Rees, A. L.. *A History of Experimental Film and Video*. London: Palgrave Macmillan, 2011.

Rhodes, Gary and John Springer. "Introduction" (2006). In *Docufictions: Essays on the Intersection of Documentary and Fictional Filmmaking*, edited by Gary Rhodes and John Springer, pp. 1–9. Jefferson and London: McFarland and Company, Inc., 2006.

Richard, David. *Film Phenomenology and Adaptation: Sensuous Elaboration*. Amsterdam: Amsterdam University Press, 2021.

Richardson, Michael. *Surrealism and Cinema*. Oxford and New York: Berg, 2006.

Rotha, Paul. *The Film till Now: A Survey of World Cinema*. New York: Twayne Publishers, Inc., 1949.

Sakya Paṇḍita, Künga Gyeltsen. *Liangli baozang lun shi* 量理宝藏论释 [An Explanation of the Treasure of a Buddhist Theory of Standard Measurement] (1182–1251). Beijing: Zhongguo wenshi chubanshe 北京：中国文史出版社 [Beijing: Chinese Literature and History Press], vol. 1 (2014).

Saṃghabhadra. *Apidamo shunzhengli lun* 阿毗达磨顺正理论 [The Conformity to Correct Principle] (the fifth century). In the *Wan zhengzang jing* 卍正藏经 [The Wan Tripiṭaka], edited by Cangjing shuyuan 藏经书院 [Buddhist Texts Academy]. Taipei: Xinwenfeng chuban gufen youxian gongsi 台北：新文丰出版股份有限公司 [Taipei: Xinwenfeng Publishing Company], vol. 72 (1980).

Śāntideva. *Engaging in Bodhisattva Behaviour* [Ru pusa xing lun 入菩萨行论] (circa 700). Translated by Alexander Berzin. http://www.buddhistische-gesellschaft-berlin.de/downloads/bca.pdf, 2005.

Saraha. *Saraha's Treasury of Songs: People Dohā* [Daoge baozang·shumin zhi ge 道歌宝藏•庶民之歌] (circa the first century). Commented on by Khenpo Tsultrim Namdak. In *Buddhist Texts: Through the Ages*, edited by Edward Conze. Translated by David Snellgrove. Oxford: Bruno Cassirer, 1954.

Sartre, Jean-Paul. *The Imaginary: A Phenomenological Psychology of the Imagination* (1940). Translated by Jonathan Webber. London and New York: Routledge, 2004.

Schafer, R. Murray. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books, 1994.

Sitney, P. Adams. *Visionary Film: The American Avant-garde, 1943–2000*. New York: Oxford University Press, 2002.

Sobchack, Vivian. *The Address of the Eye: A Phenomenology of Film Experience*. Princeton and Oxford: Princeton University Press, 1992.

Tassone, Aldo. "Le cinéma italien parle: Michelangelo Antonioni" (1992). In *Michelangelo Antonioni: Entretiens et inédits 1950–1985*, edited by Carlo Di Carlo and Giorgio Tinazzi, p. 108. Rome: Cinecittà International, 1992.

Truffaut, François. "A Certain Tendency of the French Cinema" (1954). In *Cahiers du Cinéma in English*, edited by Andrew Sarris. Translated by Rose Kaplin, G. Hersberger, and John K. E. Hitchcock. New York: Cahiers Publishing Company, no. 1 (1966): pp. 30–41.

Turim, Maureen. *Abstraction in Avant-garde Films*. Ann Arbor: UMI Research Press, 1985.

Udden, James. "In Search of Chinese Film Style(s) and Technique(s)" (2012). In *A Companion to Chinese Cinema*, edited by Yingjin Zhang, pp. 265–83. Chichester, Malden, and Oxford: Wiley-Blackwell Publishing Ltd, 2012.

Verrone, William. *The Avant-garde Feature Film: A Critical History*. Jefferson and London: McFarland and Company, Inc., 2012.

Wees, William. *Recycled Images: The Art and Politics of Found Footage Films*. New York: Anthology Film Archives, 1993.

Weiss, Allen. "Frampton's Lemma, Zorn's Dilemma" (1985). In *October*, edited by Rosalind Krauss and Annette Michelson. Cambridge: MIT press, vol. 32 (1985): pp. 118–28.

Wollen, Peter. *Paris Hollywood: Writings on Film*. London and New York: Verso, 2002.

Wollen, Peter. "The Two Avant-gardes" (1975). In *Studio International: Avant-garde Film in England and Europe*, edited by Richard Cork, David Curtis, Peter Gidal, Malcolm Le Grice, and Peter Wollen. London: Studio International Journal Ltd, vol. 190, no. 978 (1975): pp. 171–75.

Xuanzang (the pinyin form of Hsüan-tsang [Hsüan-tsang is a Wade–Giles-based name]). *Heart Sūtra* [Bore boluomiduo xin jing 般若波罗蜜多心经]. Commented on by Tanxu (the pinyin form of T'an Hsü [T'an Hsü is a Wade–Giles-based name]). Translated by Lok To. New York: Young Men's Buddhist Association of America, 2000.

Yanshou, Yongming. *Zongjing lu* 宗镜录 [The Records of the Source Mirror] (961). Xi'an: Xibei daxue chubanshe 西安：西北大学出版社 [Xi'an: Northwest University Press], vol. 55 (2015), vol. 60 (2015), vol. 73 (2015) (Note: Even if Shi Yanshou is the name on the front cover of the book, "Shi" is a unified title of Buddhists. His name is Yongming Yanshou).

Yinshun. *Zhongguan lunsong jiangji* 中观论颂讲记 [Lectures on the Fundamental Verses on the Middle Way]. Beijing: Zhonghua shuju 北京：中华书局 [Beijing: Zhonghua Book Company], 2017.

Youngblood, Gene. *Expanded Cinema*. New York: E. P. Dutton and Company, Inc., 1970.

Zaatari, Akram and Abbas Kiarostami. “Abbas Kiarostami” (1995). In *Bomb*, edited by Betsy Sussler, Andrew Bourne, and Sabine Russ. New York: New Art Publications, Inc., no. 50 (1995): pp. 12–14.

Zhongguo shehui kexueyuan yuyan yanjiusuo 中国社会科学院语言研究所 [The Institute of Linguistics at the Chinese Academy of Social Sciences]. *Xinhua zidian* 新华字典 [The New China Dictionary]. Beijing: Shangwu yinshuguan 北京：商务印书馆 [Beijing: The Commercial Press], 2004.

Filmography

16 in Webster Groves. Directed by Arthur Barron. CBS, 1966. 16 mm. 45 minutes.

A Movie. Directed by Bruce Conner. Bruce Conner Productions, 1958. 16 mm. 12 minutes.

Battleship Potemkin. Directed by Sergei Eisenstein. Mosfilm, 1925. 35 mm. 75 minutes.

Berlin: Symphony of a Great City. Directed by Walter Ruttmann. Les Productions Fox Europa and Deutsche Vereins-Film, 1927. 35 mm. 65 minutes.

Birdman. Directed by Alejandro González Iñárritu. New Regency Productions, M Productions, Le Grisbi Productions, TSG Entertainment, and Worldview Entertainment, 2014. Codex. 119 minutes.

Blow-up. Directed by Michelangelo Antonioni. Metro-Goldwyn-Mayer, Premier Productions, Carlo Ponti Productions, and Bridge Films, 1966. 35 mm. 111 minutes.

Bridges-Go-Round. Directed by Shirley Clarke. Shirley Clarke Productions, 1958. 16 mm. 8 minutes.

Butch Cassidy and the Sundance Kid. Directed by George Roy Hill. Campanile Productions, George Roy Hill-Paul Monash Production, Newman-Foreman Company, and Estudios Churubusco Azteca S.A., 1969. 35 mm. 110 minutes.

Chronicle of a Summer. Directed by Jean Rouch and Edgar Morin. Argos Films, 1961. 16 and 35 mm. 85 minutes.

Citizen Kane. Directed by Orson Welles. RKO Radio Pictures and Mercury Productions, 1941. 35 mm. 119 minutes.

Close Encounters of the Third Kind. Directed by Steven Spielberg. EMI Films, Columbia Pictures, and Julia Phillips and Michael Phillips Productions, 1977. 35 and 65 mm. 138 minutes.

Daughter Rite. Directed by Michelle Citron. Iris Films, 1980. 35 mm. 49 minutes.

Dead Birds. Directed by Robert Gardner. Peabody Museum, 1963. 16 mm. 85 minutes.

Fast and Furious Presents: Hobbs and Shaw. Directed by David Leitch. Chris Morgan Productions, Dentsu, Seven Bucks Productions, and Universal Pictures, 2019. Codex. 137 minutes.

Free Fall. Directed by Péter Forgács. Balázs Béla Stúdió and Magyar Televízió Fiatal Müvészek Stúdiója, 1998. 8 mm. 75 minutes.

Gravity. Directed by Alfonso Cuarón. Warner Bros., Esperanto Filmoj, and Heyday Films, 2013. 65 mm and Codex. 91 minutes.

High School. Directed by Frederick Wiseman. Osti Productions, 1968. 16 mm. 75 minutes.

Jaws. Directed by Steven Spielberg. Zanuck/Brown Company and Universal Pictures, 1975. 35 mm. 124 minutes.

Lemon. Directed by Hollis Frampton. Hollis Frampton Productions, 1969. 16 mm. 7 minutes.

Les raquetteurs. Directed by Michel Brault and Gilles Groulx. National Film Board of Canada, 1958. 35 mm. 15 minutes.

Lichtspiel: Opus I–IV. Directed by Walter Ruttmann. Ruttmann-Film, 1921–1925. 35 mm. 13, 4, 3, and 4 minutes.

Man with a Movie Camera. Directed by Dziga Vertov. Vse-Ukrainske Foto Kino Upravlinnia, 1929. 35 mm. 68 minutes.

Meshes of the Afternoon. Directed by Maya Deren and Alexander Hammid. Maya Deren Productions, 1943. 16 mm. 14 minutes.

Night and Fog. Directed by Alain Resnais. Argos Films, 1956. 16 mm. 32 minutes.

No. 10: Mirror Animations. Directed by Harry Smith. Harry Smith Productions, 1956–1957. 16 mm. 4 minutes.

N:O:T:H:I:N:G. Directed by Paul Sharits. Paul Sharits Productions, 1968. 16 mm. 36 minutes.

Print Generation. Directed by J. J. Murphy. J. J. Murphy Productions, 1974. 16 mm. 50 minutes.

Railroad Turnbridge. Directed by Richard Serra. Richard Serra Productions, 1976. 16 mm. 19 minutes.

Rhythmus 21, 23, and 25. Directed by Hans Richter. Hans Richter Productions, 1921–1925. 35 mm. 3, 3, and 4 minutes.

Room Film 1973. Directed by Peter Gidal. Peter Gidal Productions, 1973. 16 mm. 55 minutes.

Rope. Directed by Alfred Hitchcock. Warner Bros. and Transatlantic Pictures, 1948. 35 mm. 80 minutes.

Russian Ark. Directed by Alexander Sokurov. Hermitage Museum, Hermitage Bridge Studio, Egoli Tossell Pictures, Ministry of Culture of the Russian Federation, Mitteldeutsche Medienförderung, BKM Kulturelle Filmförderung des Bundes, Filmbüro NW, WDR/ARTE, Fora Film, Koppmedia, NHK, Seville Pictures, Yle TV1, DR, AST Studio, Mariinsky Theatre, Filmboard Berlin-Brandenburg, Filmförderung Hamburg, and Kulturelle Filmförderung Sachsen, 2002. Digital (HDTV). 99 minutes.

Samadhi. Directed by Jordan Belson. Jordan Belson Productions, 1967. 16 mm. 6 minutes.

Strike. Directed by Sergei Eisenstein. 1-ya Goskino Fabrika, Goskino, and Proletkult, 1925. 35 mm. 82 minutes.

The 400 Blows. Directed by François Truffaut. Les Films du Carrosse and Sédif Productions, 1959. 35 mm. 99 minutes.

The Best Years of Our Lives. Directed by William Wyler. Samuel Goldwyn Productions, 1946. 35 mm. 172 minutes.

The City. Directed by Ralph Steiner and Willard Van Dyke. American Documentary Films Inc. and American Institute of Planners, 1939. 35 mm. 43 minutes.

The Danube Exodus. Directed by Péter Forgács. Lumen Film, 1998. 8 mm. 60 minutes.

The Dark Tower. Directed by Stan Brakhage. Jaren Vine Productions, 1999. 16 mm. 3 minutes.

The Little Foxes. Directed by William Wyler. Samuel Goldwyn Productions, 1941. 35 mm. 116 minutes.

The Selling of the Pentagon. Directed by Peter Davis. CBS, 1971. 16 mm. 60 minutes.

The Sorrow and the Pity. Directed by Marcel Ophuls. Télévision Rencontre, Société suisse de radiodiffusion et télévision, and Norddeutscher Rundfunk, 1969. 16 mm. 251 minutes.

Through the Olive Trees. Directed by Abbas Kiarostami. Abbas Kiarostami Productions, Ciby 2000, and Farabi Cinema Foundation, 1994. 35 mm. 103 minutes.

Un Chien Andalou. Directed by Luis Buñuel. Luis Buñuel and Pierre Braunberger Productions, 1929. 35 mm. 16 minutes.

Wavelength. Directed by Michael Snow. Michael Snow Productions, 1967. 16 mm. 45 minutes.

Zorns Lemma. Directed by Hollis Frampton. Hollis Frampton Productions, 1970. 16 mm. 60 minutes.