

**Faculty of Media, Society and Culture
Department of Media and Information**

**The Cinesthetic Montage of Music-Video:
Hearing the Image and Seeing the Sound**

Joachim Wichman Strand

**This thesis is presented for the Degree of
Master of Creative Arts
of
Curtin University of Technology**

July 2006

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material that has been accepted for the award of any other degree or diploma in any university.

Signature:

Joachim Strand

Date: 8 / 11 / 2006

The Cinesthetic Montage of Music-Video

Abstract

This thesis examines the interconnected relationship that exists between sound and moving-image in the music-video. The flow of images used in many music videos often carries no definite meaning. Rather, the viewer must perceive the physiological sensations of the video's audiovisual expression to make sense of it. Thus, both the expression and the perception of music-video is a cross-modal process. Using Vivian Sobchack's theory of *cinesthetics* as a framework, the thesis contends that the music-video produces an aural visuality in which sound can be cinesthetically expressed and perceived as image and the image perceived and expressed as sound.

Contents

Acknowledgements	1
General Introduction	3
Chapter One - Cinesthesia and Sound	
1.1 Introduction	10
1.2 Film Phenomenology	12
1.3 Audio and the Image	21
1.4 Conclusion	27
Chapter Two - The Music of the Image / The Image of the Music	
2.1 Introduction	28
2.2 The Music Video Context	30
2.3 The Music Video Framework	35
2.4 The Music-Video Montage	43
2.5 Conclusion	53
Chapter Three - Analysis of Chris Cunningham's "Only You"	
3.1 Introduction	54
3.2 The Music-Video: Audio-Form and Visual-Flow	58
3.3 Cinesthetic Music-Video Montage: "Only You" Dissected	67
3.4 Conclusion	75
Chapter Four - Making Music-Video: A critical reflection of the production process of <i>Moment</i>	
4.1 Introduction	77
4.2 Connecting Image and Music	79
4.3 The Cinesthetic Montage Process	82
4.4 Conclusion	86
General Conclusion	87
Bibliography	90
Films and Music Videos	95
List of Illustrations	96
Appendices:	
A: <i>Only You</i> Synopsis	97
B: <i>Only You</i> Lyrics	98
C: Audio Breakdown of <i>Only You</i>	99

Acknowledgements

I would like to thank my beautiful wife for love, patience and inspiration, my supervisors Antonio Traverso for patient application of his vast knowledge and Lezli-An Barrett for extra guidance, the tech staff at Curtin FTV for good friendship and great advice, my family and my friends. I would also like to acknowledge Vivian Sobchack and her work for being the catalyst for my thesis, Sergei Eisenstein for being 40 years ahead of his time, as well as Carol Vernallis, Andrew Goodwin and Michael Chion for providing works of great significance and inspiration. Big thanks also to Bruce Ridley at Curtin University and Sally Hanson at the Australian Copyrights Council for solid advice. I would also like to thank EMI, Warner Music Australia, and Universal Music Australia.

Finally, I would like to thank all the music video creators, particularly Chris Cunningham, Mark Romanek and Jonathan Glazer, for making great pieces of 'haiku for the eyes' that are an endless source of inspiration and cinesthetic pleasure.

- Joachim Wichman Strand, Perth 31.07.2006

“To remove the barriers between sight and sound,
between the seen world and the heard world!
To bring about a unity and a harmonious
relationship between these two opposite spheres.
What an absorbing task!
The Greeks and Diderot, Wagner and Scriabin –
who has not dreamt of this ideal?
Is there anyone who has made no attempt
to realize this dream?”

- Sergei Eisenstein (1970, p. 74).

General Introduction

While watching, and listening to, the opening sequence of Terrence Malick's *The New World* (2005 United States: New Line Home Entertainment) it struck me how the sound was completely integrated into the image, and the image onto the sound. The fluid motion of the colonist's ships juxtaposed with the graceful bodies of the swimming Native Americans was correlated perfectly with the slow undulations of Wagner's *Das Rheingold*. It was as if the montage of images had been composed *from* the music, and in such tight relationship to it that the two sensations could not be separated. The result was an audiovisual flow that moved as one unified poetic expression, generating an aural-visuality that reminded me of the music video. Just like a music video, the film's audiovisual flow poetically captured the mood and the 'feel' of the experience through careful and harmonic synchronisation of image and music that expands the moving image beyond the screen. And like the music video, it did not convey a definite meaning or follow a progressive narrative. Rather, the tightly connected audiovisual elements enabled the viewer to experience the feel of the water, the smell of the land and, in short, opened up an expansion of tactile sensations. Such an expansion relies on the cross-modality of the audiovisual elements, which enables them to flow into and onto each other seamlessly. This cross-modal flow generates one unified expression that can be felt by the viewers' bodies rather than seen by their eyes and heard by their ears.

The montage of frequently abstract images used in many music videos carries no apparent meaning. Rather the viewer/listener must participate in the creation of the meaning by 'listening' to and acknowledging the physiological sensations offered up by the music-video stream.¹ Furthermore, the music-video image is composed and represented so that it imitates the aural qualities of the song. Thus, the music-video makers ambition is to reconstruct the poetic atmosphere inherent in the song so that the image can expand the aural sensations. Consequently, the aural and visual elements flow and interconnect as one munificent stream that will have a multifaceted

¹ Viewer/listener: this expression points to the notion that the perception of the moving image involves not simply viewing but also simultaneous listening. Music-video: I use the hyphenated word music-video in this thesis to indicate the totally connected and dependent relationship that exists between the images and the audio in this moving image format, particularly in regards to the synchronicity and unified expression and perception of such an audiovisual flow.

physical impact on the viewer rather than just an intellectual one. This co-constructive and physical perception process of the audiovisual moving image carries great interest for filmmakers. By creating profuse streams of aural-visuality it is possible to produce moving images that makes sense through physical and psychological participation, consequently generating multiple and expanded levels of meaning that might not be available otherwise. The generation of such multifaceted and complex meanings is dependant upon the physical and psychological memories, perceptions, and expressions of music-video maker *and* viewer.

Such an understanding of the moving image's audiovisual possibilities has been explored by Vivian Sobchack in her theory of the cinesthetics of cinema. Sobchack's phenomenological theory of the cinesthetic subject establishes the perception and expression of the audiovisual moving image as a two-way synaesthetic gestalt.² Consequently, her focus is the relationship between the viewer's sensorium and the perception and expression of the audiovisual moving image. Using Sobchack's theory of cinesthetics as a framework, this thesis examines the specific relationship between sound and moving image in the music-video. In doing so, it focuses on how the application of an embodied cinematic apparatus enables the generation of moving images that illustrate audio in such a way that the sound transposes into and onto the images, and vice versa. Thus, this thesis contends that the music-video produces an aural-visuality in which sound can be cinesthetically expressed and perceived as image and the image perceived and expressed as sound. In order to demonstrate this argument it is first necessary to establish a working definition of cinesthetics, which is achieved by looking at Sobchack's phenomenological account of the audiovisual moving image.

According to Sobchack viewers make sense of the moving image because it is perceived through their embodied sensorium.³ This is possible because the embodied sensorium can recognise and comprehend the flow of audio and images because, as Sobchack says, the audiovisual moving image has its own "body" (2003, 165). In other words, the audiovisual moving image has its own set of senses that work in the same commutable and co-operative manner as the lived-body sensorium. In Chapter

² See V. Sobchack (1992) *The Address of the Eye: A phenomenology of film experience*.

³ Ibid

One I argue that the moving image functions in a manner analogous to that of the synaesthetic condition. Synaesthesia is a state in which sensations or perceptions from one sense modality is transposed onto a different sense modality, unifying the impressions from the different modalities into one munificent sensation. In this way, the images provide sensations that go beyond the mere visual and venture into the areas of touch, smell, and most importantly aural impressions. Chapter One explores how the cinematic apparatus is utilised to produce a flow of embodied images by examining the cinesthetic qualities of the moving image, as charted by Sobchack. Having looked at the first element of the music-video, namely the moving image, I then move on to the second element, the sound.

Sound has many functions as part of the audiovisual moving image. One of the most important is that it links disparate images together and expands the physical and psychological geography of the images beyond the screen. According to sound researcher Michael Chion, this expansion is possible because sound functions as an integral factor in the visual perception process.⁴ Thus, in Chapter One, I correlate Chion's notions to Sobchack's concept of cinesthetic perception and expression. This shows how sound is embodied, just like the moving image, and how they both provide sensory interconnections in order to generate audiovisual expressions and perceptions that go beyond simple seeing and hearing. In this way I illustrate how the cinesthetic instrumentality facilitates the transposition of sonic velocity into the image and visual shape onto the sound, generating a cinesthetically munificent aural-visibility that is analogous to the synaesthetic condition. Thus, as with the direct experiences perceived by the lived-body sensorium, each audiovisual sense the moving image expresses is met by a rush of other sensory associations. And so, having outlined the embodied characteristics of the audiovisual flow and the cinesthetic connections and transpositions it facilitates, I move on to show how these audiovisual connections function in the music-video.

Consequently, Chapter Two contends that the music-video expresses its aural-visibility through and by a *cinesthetic montage* of audio and moving image elements. Thus, through the concept of cinesthetic montage, I link Sobchack's notion of

⁴ See M. Chion (1994) *Audio-Vision: Sound on screen*.

cinesthetics to Sergei Eisenstein's theory of overtone montage. It is not my intention to collapse Sobchack's study of the cinesthetics of the moving image and Eisenstein's theory of the synchronisation of the senses through montage. Rather, what I seek to achieve is an underscoring of the fertile correlation between two concepts that both emphasise an embodied audiovisual synthesis that generates sonic reverberations in the image and visualisation of sound. In order to show how this correlation works, it is first necessary to provide a historical and technological context for the music video. Although the music video format draws conceptual techniques and technology from both cinema and television, it should be seen as different from these as it has its own way of arranging audio and visual elements through the exploration of their connections and cohesive temporal flow. Furthermore, the music-video format also draws on techniques for audiovisual synthesis that can be seen to stretch as far back as the Gesamtkunstwerk experimentations of Kandinsky and Scriabin and the theatrical sound and visual synthesis of Wagner.

The first section of Chapter Two outlines a chronology of audiovisual synthesis from the sound and visual experimentations of Kandinsky, Scriabin and Wagner; throughout the scopitone machines of the 1950s; the image projection systems of the 1960s and video scratching of the 1980s; through to the advent of music television in order to contextualise the music video format. Having provided this context, I move on to discuss the conceptual techniques that are deployed to facilitate the transposability of music and moving images, as put forward by Carol Vernallis.⁵ These techniques provide a visual framework for the music video in which the moving image elements are broken down into discernible components according to the musical sections of the song. This framework uses the musical section as the fundamental unit for analysis, rather than the shot that is used in the analysis of narrative cinema. This is done in order to emphasise the varied repetition of *materials* rather than *narrative* and the linear development of plot.

This kind of framework not only draws attention to the intimate connection that exists between sound and moving image, but also the way this connection flows *from* the music *to* the images. The direction of this influence will be particularly important for

⁵ See C. Vernallis (2004) *Experiencing Music Video: Aesthetics and cultural context*.

the conceptualisation process of the music video imagery discussed throughout this and the following chapter, as well as for the creative component discussed in Chapter Four. In fact, in these chapters it will be made clear that sound is the primary element of the music-video. While in the cinema sound and music are added *onto* the images to enhance and expand the visual experience; in the music-video it is the sound that provides the cues for the conceptualisation, production and fusing of the video imagery.

Thus, the last section of Chapter Two demonstrates how the visual elements exemplified by Vernallis's framework are cinesthetically merged with the audio through a process not unlike that described in Eisenstein's theory of *overtonal montage*. The music video synthesises audio and image in order to achieve a level of audiovisual commutability and transposability. Sound and moving images are joined to generate an aural-visibility that reverberates within the viewers. This audio-visual welding can be illustrated through an application of Eisenstein's theory of montage. Montage is first and foremost a conceptual technique to combine two images to create a third separate meaning that is greater than the sum total of the two images. With Eisenstein's subsequent examination of the commutability of the aural sense and the visual sense, he devised a model of overtone montage.⁶ This model is characterised by the intimate interconnection of the compositional effects within the shot, the form and flow of the images, the textures of the shot's surface, the timbre and harmonic development of the audio and the rhythm and synchresis of the audio-visual elements. In this way, the sound and the images are welded together, enabling the expression and perception of *all* the resonances contained within the interconnected sound and image elements. In order to illustrate how the cinesthetic montage of sound and image can create such connections, the next chapter comprises an in-depth analysis of a music-video.

Chapter Three, therefore, provides an analysis of the music-video *Only You*⁷ by one of the more visually creative music video makers of the past decade: Chris Cunningham. Working with major musical acts such as Björk, Madonna, and Portishead, as well as

⁶ See S. Eisenstein (1988b). An Unexpected Juncture. In R. Taylor (Trans. & Ed.), *Selected Works Vol. 1: Writings 1922-34* (pp. 115-122).

⁷ Made for the Portishead song *Only You* in 1998.

underground avant-garde acts such as Aphex Twin, Cunningham has developed a recognisable style of kinetic, tactile and fluid imagery. Thus, the chapter demonstrates how Cunningham's music-video cinesthetically expresses the music as image and the image as music through an application of cinesthetic montage. The first part of Chapter Three consists of a brief synopsis and description of the video, and a sectional breakdown of the song itself. Such a breakdown provides access to distinguishable musical sections that function as the fundamental unit of the analysis. The second part of the chapter examines *Only You* according to the conceptual techniques examined earlier in Chapter Two. This part makes up the most significant element of the analysis, as it illustrates the process of cinesthetic montage by which music is translated into image and the image is made to resound.

The final part of the chapter provides a close chronological reading of the video, expanding upon the different elements formulated previously in the chapter. The close chronological reading of *Only You* provides a more detailed illustration of how the sound and image elements are cinesthetically mounted together.⁸ It demonstrates how the imagery flows *from* the music and re-connects with it in a mutually reversible manner. The analysis shows how the video fuses the audio structure to the image, creating an aural-visuality in which all the audiovisual elements reverberate concurrently. This harmonisation of music-video generates an audiovisual expression that cinesthetically affects the viewer's embodied sensorium. In this way, the cinesthetic montage of the music-video facilitates the transposition of music onto and into the visuals, and the visuals into and onto the music, making it possible, so to speak, to hear the image and see the sound.

This thesis also contains a creative component, which consists of a 24-minute experimental music-video. It is a visualisation of an experimental piece of spectral music composed by Edith Cowan University Master of Creative Arts student Brett Mabury. The objective of the creative component was to apply the conceptual framework discussed and used in the previous chapters to the creation of an original piece. This application illustrates the claims made about music-video in general and Chris Cunningham's *Only You* in particular. Chapter Four therefore comprises a

⁸ I use this word as a verb indicating the act of executing the montage process.

critical reflection on the entire production process from conceptualisation, through shooting and to mounting the 24-minute music-video *Moment*. The chapter begins by exploring how the images were conceptualised using the music as a starting point, and how the music was used actively as a point of reference throughout the entire process. The second part of the chapter illustrates how the conceptual framework was applied to the shooting, and more importantly, the editing process. Thus, this chapter explicates how the conceptual framework is applicable to not only the analysis of music videos but also the creation and cinesthetic montage of music-video.

Chapter One – Cinesthesia and Sound

“Music is the short-
hand of emotion”
- Leo Tolstoy

1.1 Introduction

According to Vivian Sobchack, viewers make sense of the moving image experience because it is perceived through their embodied sensorium. This is possible, Sobchack claims, because the moving image has its own “body” with its own set of senses that work in the same commutable and co-operative manner as the lived-body sensorium (2003, 165). While Sobchack’s phenomenological investigation focuses on how the moving image is perceived and understood through and because of the viewer’s body, I turn my attention towards the body of the moving image itself. Using Sobchack’s notion of cinesthetics as a framework, I consider how the audio and the visuals of the moving image body exhibit certain characteristics similar to the condition known as synaesthesia. This condition transposes sensations or perceptions from one sense modality onto a different sense modality, unifying the impressions from the different modalities into one munificent sensation. Thus, the audio-visual of the moving image can both behave and be perceived in a synaesthetically analogous manner. This does not mean that everybody can literally taste sounds, or hear colours in the manner of the true synesthete, but that the direct flow of sound and visuals is perceived as one holistic perception of expression in which the modality of sound can be transposed onto and into the image and vice versa.

This chapter contends that the audio-visual arrangements of the moving image can be seen as a system of cross-modal associations analogous to the synaesthetic condition both in terms of its perception and expression. This means that the visuals can retain the qualities of and enhance the sound, and vice versa, creating an aural-visibility. The notion of aural-visibility implies that the sound and the visuals are commutable and transposable entities that can assume each others phenomenological qualities. In this way a merged cross-modal audiovisual form and flow is generated in which sound materialises like image and image reverberates like sound. But initially, I consider the

two main elements, or senses, of the moving image body: namely the visuals and the sound.

Accordingly, this chapter has a three-part structure. Firstly, I provide a brief overview of phenomenology and its application to the moving image in order to contextualise Vivian Sobchack's notion of the moving image body and its similarities to the human lived-body, as well as introduce the concept of synaesthesia. The second part of the chapter examines the synaesthetic analogy of the moving image sensorium through an examination of Sobchack's concept of cinesthetics. This notion highlights the commutability and transposability between the audio-visual elements of the moving image, and how this structure corresponds to the synaesthetic condition. In this way, I examine the commutability found in the human lived-body sensorium, and how this commutability can also be applicable to the embodied moving image and the cinematic apparatus through which it is expressed. To finish the chapter and prepare for the discussion of the aural-visuality of the music video in Chapter Two, I extend my discussion to the modality of sound. Specific attention is given to the ideas of sound researcher Michel Chion, in particular his approach to the conglomeration of the aural and visual elements in the moving image. Different sound techniques and applications of aural elements are discussed, emphasising Chion's notion of audiovision and the transposition of sound into image and image into sound.

1.2 Film Phenomenology

Phenomenology is a methodology for the study of concrete human existence that seeks to attain the central features of what and how we experience the world by using the intuitive experience of phenomena (what presents itself to us in conscious experience) as its starting point. It analyses and describes human consciousness as an intentional and embodied experience, in its search for a philosophy that accounts for space, time, and the world, just as we actively engage, experience, and live it. In the view of its founder, phenomenology is the study of the “knowledge of essences”, that is, the essential experiences that present the world to us (Husserl 1962, 42). In its post-Husserlian expressions, phenomenology considers the world already to be there before reflection, establishing the body as subjectively experienced by the individual and as a person’s vehicle of “being in the world” with conscious and intentional acts of expression and perception. As Merleau-Ponty puts it, “[c]onsciousness must be faced with its own unreflecting life in things and awakened to its own history, which it is forgetting” (Merleau-Ponty 1962, 31). Existential phenomenology rejects Husserl’s view on ‘essences,’ concentrating rather on historicized and ‘qualified’ descriptions as its goal. Existential phenomenology then is, as Don Ihde describes it, a “[p]hilosophical style that emphasizes a certain interpretation of human *experience* and that, in particular, concerns *perception* and *bodily activity*” (Ihde 1990, 29). These notions of perception and bodily activity have been applied to cinema by Vivian Sobchack who, through an application of Merleau-Ponty’s and Ihde’s existential phenomenologies, investigates the expression and perception of the moving image.

According to Sobchack, perception is a gestalt, an organising activity of embodied intentionality that connects to the world through a structured and structuring encounter (1992, 76). But perception is not just vision. Vision is only one of its modalities that work in conglomeration with all the other sense modalities of the human sensorium. The senses can therefore be said to be different openings to the world, functioning as a unified access system to experience. By being in the world all the functions of the sensorium, located in the body, are exercised and linked together in a continuously collaborative relationship that translates and transfers impressions across and within the entire sensorium. Thus, the entire body is implicated in this

process (Sobchack, 1992, 78). This is also true for the audio-visual stream experienced while watching and listening to the moving image.

The stream of sound and visuals is experienced by the body as both a physical reflex to multi-sensory inputs, as well as a reflection on how this reflexive process is mirrored by the actual audiovisual stream coming from the screen and by the representations on the screen. Thus, the viewer will experience the moving image while simultaneously being concerned about the process of perception and expression through a reflection on the workings of the cinematic apparatus, and the sensory experiences represented by the on-screen action. In this way, the perception of the moving image experience becomes an instantaneous dialogue between the body of the screen, the bodies on the screen and the body watching and hearing the screen.

The moving image body

The audiovisual imagery expressed by the moving image enables the spectator to engage with the film within a structure of meanings and metaphors in which relations between the subject and the object are “cooperative”, “co-constructive”, “dynamic” and “reversible” (Sobchack 1992, 138). When we watch film, we perceive the filmmaker’s expression of perception, and in turn reverse this structure by expressing our experience back onto the film. This does not, however, mean that we are living inside the film. Rather, we perceive the vibrant on-screen action, translate it using our co-operative sense modalities into a bodily experience, and act it out in our heads. We do this by dynamically co-constructing the film action using our lived-body experiences as raw material triggered and moulded by the perception and/or expression of the film. In this way, as Sobchack puts it, in the act of watching a film, viewers will not only:

spontaneously and invisibly perform these existential acts directly for and as ourselves in relation *to* the film before us, but these same acts are coterminously given to us *as* the film, as mediating acts of perception-cum-expression that we take up and *invisibly perform* by appropriating and incorporating them into our own existential performance; we watch them as a *visible performance* distinguishable from, yet included in, our own. (1992, 10)

Thus, the process of perception involves a circular route of transaction. We watch the film that provides vibrant action, this action is translated into and onto the body through the sensorium, and then this translation is put forward into and onto the film, and the process starts over again.

This reflexive two-way process would not be possible if the human body were not in some way acquainted with the elements found in the moving image, and able to recognise, acknowledge, and reflect upon them. According to Sobchack, this process is possible because the cinematic apparatus is an intentional technology just like the human-body (2003, 165). The perception and expression of the audiovisual moving image is therefore a completely sensuous experience informed by and comprehended through the sensory modes of the body. As previously explained, the lived-body provides access to and engages with the world through multi-sensory perception. This commutable and co-operative structure of the sense modalities can be said to be the technical methods of the lived-body's incarnate instrumentality, which are used to grasp the world. These notions can equally be applied to the moving image, since sense modalities are used here as well to express perception. As Sobchack says: "If the modes of intentional consciousness correlate and correspond to technical methods, then the film is to cinematic technology as human perception and its expression is to human physiology" (1992, 166). In this way, the moving image becomes a structure of interaction based on embodied perception, a notion that can equally be applied to the moving image because the film has its own body. That is not to say that the film is a lived-body with a concrete materiality. But insofar as it performs in ways similar to the lived-body, constituting the experience of consciousness, and more importantly the expression of perception, it can be said, as Sobchack puts it, to be "embodied" (1992, 162).

The moving image reveals a similar commutable system in its use of sense modalities; it is subjectively and objectively reversible, and it performs the double act of perception of expression and expression of perception that can be found in the material lived-body. In this way the moving image becomes: "an agent who autonomously, introceptively, and visibly perceives the *visual behaviour* of others" (Sobchack, 2003, 167). Thus, it is the multi-sensory system of access enabled by the human sensorium that makes it possible for the film medium to express sensations

beyond the sound-image elements. Through the audio-visual imagery created by the cinematic apparatus (editing, lighting, camera, sound, effects) it is possible to create a moving image that appeals to the entire sensorium simultaneously. Furthermore, this moving image can visualise the lines⁹ of the sound and make the image lines reverberate with them, in a manner resembling the synaesthetic condition.

Synaesthetic audiovisuals

Synaesthesia derives from the Greek *syn* (union/together) and *aisthesis* (sensation/perception), and entails a sensation or perception involving impressions from the whole sensorium that are unified and interconnected generating *one* munificent sensation. There have been several accounts of synesthesia through the 300 years it has been known to medicine. Marks (1978) points to a reported case from 1735 and a hearsay incident from 1690. Both Marks and Cytowic (1989) claim that the interest for synesthesia intensified in the nineteenth century, particularly in France, attracting serious attention within a number of different areas such as art, music, language, literature and natural philosophy.¹⁰ A demarcation between humanities and sciences in terms of what type of experiences can be nominated as synesthetic experiences was made in the 1980s. According to Cytowic (1995) synesthesia should no longer be considered a “cross-modal association of metaphoric language” or a “sensory fusion in art”, but a distinct neurologic condition. Within the field of neuropsychology it was made particularly clear that synesthesia is an involuntary physical experience that is clearly distinguished from metaphor, sound symbolism and other forms of artistic multi-sensory joinings. But, along the lines of Cretien Van Campen (1999) and Lawrence Marks (1978), it seems clear that synesthesia can be used to explain the multi-sensory experiences created by audio-visual art forms such as film and music video. To distinguish the two notions I will refer to the artistic convention as synaesthesia, taking my cue from Sobchack.

⁹ Eisenstein called the directions of movement located within an image or a piece of sound for its line. These lines could be the rising sound of a trumpet solo, or the movement a camera executes in a particular shot. The idea was to combine sound lines and image lines that matched. See Eisenstein (1970). *The Film Sense*.

¹⁰ See Brougher, K. (2005). Visual music: synaesthesia in art and music since 1900, Van Campen, C. (1997). Synesthesia and artistic experimentation. *Psyche*, 3(6), and Milicevic, M. (1997). *Experimental/Abstract Film and Synaesthetic Phenomena (1725-1970)*.

It was particularly within the literary Symbolist movement in the 1890s that synaesthesia was developed into a regularly utilized artistic convention. Synaesthetic imagery transported meaning between different sensory modalities and interconnected these different meanings into a whole in which the sum was greater than the individual parts. Literal statements consisting of correspondences between two or more senses were deployed to create fresh verbal images. In this way, two or more sensory elements were joined together, creating a 'new' heterogeneous element with subsequent opportunities for multiple systems of meanings and knowledge. This process of combining two elements to create a third munificent element can be correlated to Sergei Eisenstein's notion of overtonal montage.¹¹ Just like the expression of the moving image through montage, the synaesthetic expression created by the Symbolists goes beyond objective reason. It creates an equivalent to direct experience, enabled by the embodied sensory access to the world, through the joining of cross-modal sensory elements. This is possible because of the inherent and meaningful similarities that exist across the different sense modalities and because the bodies of the audience and the creator are intimately familiar with and perceptive of such elements. Consequentially, such cross-modal similarities can be grasped by virtually everyone. Lawrence Marks has called this principle the "unity of the senses" (1978, p. 212). This unity, or linking of separate sensory elements, also has implications beyond poetry and metaphor.

The painter Vassily Kandinsky, together with the composer Schoenberg, started experimenting with the conceptual joining of visuals (colour, lights and movement) and sound (music) in order to create a total art experience, or *Gesamtkunstwerk*, in the early 1900s (Cytowic, 1995, p. 22). With a piece like *Der Gelbe Klang* (The Yellow Sound, 1912) Kandinsky wanted to push the viewers past a purely analytical explanation and experience, moving them into a state more closely resembling the qualities of "direct experience" that synaesthesia typified (Cytowic, 1995, p. 22). This form of direct experience can be closely linked with the experience of the moving image, where sound and image are conceptually joined in order to interact and create

¹¹ Eisenstein's notion of overtonal montage will be discussed in Chapter Two.

a total and direct experience.¹² As we have seen, this is enabled by the embodied perception and expression of the moving image body. Sobchack claims that the moving image experience is presupposed on the intelligibility and intersubjectivity of our experience of the world, which is enabled, by the synoptic and synaesthetic sensory system of the body. A semiotic or hermeneutic understanding of the moving image must therefore be returned to a radical reflection on “cinematic communication” (Sobchack, 1992, 6). This communication will originate from, and be informed by, the practical structures of existential experience and embodied perception.

In this way, the embodied expression and perception of the moving image gives us a concrete and empirical insight into the workings of our own subjective vision. It lays bare the reversible, dialectical and social nature of the embodied vision. This unified sensory expression is informed by and grounded in the synaesthetic experience of embodied perception. Thus, the viewers will comprehend the audiovisual flow of the moving image with their entire bodily sensorium rather than just the eyes and the ears, thus feeling the sound, hearing the images, and seeing the textures. In this way the moving image body is experienced by the viewer’s bodies *through* their eyes and ears. As Sobchack writes: “We see and comprehend and feel films with our entire bodily being, informed by the full history and carnal knowledge of our acculturated sensorium” (2004c, 63). Sobchack has called this notion “cinesthetics”, highlighting the synaesthetically analogous commutability and transposability between the aural-visual elements of the moving image.

Cinesthetics

The neologism of cinesthetics denominates the cohesive relationship between the moving image and synaesthesia, foregrounding the complex and rich bodily experience of the world that grounds both the expression and perception of cinema. Furthermore, it points to “ways in which the cinema uses our dominant sense of vision and hearing to speak comprehensibly to our other senses” (Sobchack, 2004c, 67). Cinesthetics then consists of audiovisual imagery (sound, visuals, lights, colours)

¹² The combination of sound and visuals in terms of the audio-visual moving image will be discussed in Chapter 2.

that communicate intelligibly to all our senses through our eyes and ears, much like symbolist imagery and Kandinsky's *Gesamtkunstwerk*. In addition, as Sobchack claims, it makes us realise that cinesthesia is actually the rule, rather than the exception. Everyday equally available sensory impressions are continuously heightened and diminished in a unified manner regulated by the historical and cultural embodied knowledge. This is done without us even noticing because of habituation (Sobchack, 2004c, 69). Of course, this does not mean that the viewers can actually taste the fruit showed on the screen, or feel the wind they hear howling on the soundtrack, but neither does it mean that those embodied sensations are not available.

The sensations expressed by the audio-visual moving image are still present, viable, and capable of inducing cross-modal sensory occurrences, they have just been converted into sound and image. As Laura Marks says: "Once [the synaesthetic relationship to the world] is mediated through an image, the multisensory experience is condensed into visual form. It does not vanish but is translated into the image" (2000, 214). In this way, when bodies are seen floating around as if they were underwater in Chris Cunningham's music video *Only You*, the viewers can feel the floating sensation, brought on by the fluid qualities and structure of the sound and visuals, because they too have floated in water.¹³ The optical image is translated by the viewers through their application of sense memories located within the sensory circuits of their lived-bodies. This comes as a result of the embodied qualities inherent in the audio-visual elements comprising the moving image body. The viewers, as pointed out earlier, translate and reverse the embodied sense impressions located in the image through our cinesthetic system.

This dialogue with the on-screen action is not a conscious decision; rather it occurs without a thought, since the viewer's consciousness is actually directed towards the embodied world of the film, and not their own bodies. "Thus, 'on the rebound' from the screen - and without a reflective thought - I will reflexively turn toward my own carnal, sensual, and sensible being to touch myself touching, smell myself smelling, taste myself tasting, and, in sum, sense my own sensuality" (Sobchack, 2004c, 77). This means that the audio-visual senses create impressions that are commutable, and

¹³ Chris Cunningham's *Only You* will be analysed in Chapter Three.

working in a mutual and reversible relationship, with the other senses. Although the audience can only hear the wind on the soundtrack, their embodied sensorium transposes the audio onto the other senses, drawing on the accumulated memories of their sensorium to make them 'feel' the wind.

One sense modality is not more important than the other as they provide equally significant openings to experience and work as one unified system, making the perception of the moving image a direct and total experience. According to Sobchack, moving image viewers as cinesthetic subjects: "possess an embodied intelligence that opens our eyes far beyond their discrete capacity for vision, opens the film far beyond its visible containment by the screen, and opens language to a reflective knowledge of its carnal origins and limits" (2004c, 84). This foregrounds the reciprocity and reversibility of sensible matter and sensual meaning. The direct and total experience it creates is true for both sides of the screen. In an effort to produce this direct experience, much in the same way as the painter Vassily Kandinsky did with *Gesamtkunstwerk*, filmmakers apply the tools of the cinematic apparatus onto their embodied perceptions. Utilising their unified sensory system and the commutability between the visual and aural modality of the moving image body to create a moving image that talks to all our senses through a conglomerate system of commutable sound and moving image elements.

As pointed out by Sobchack and Laura Marks, and dependant upon Lawrence Marks notion of the "unity of the senses", the optical image and its aural companion is co-constructed by the viewers through their application of sense memories located within the sensory circuits of their lived-bodies, because the aural and visual flow is created through an application of the same sensory circuits. In this way the audio-visual moving image provides expressions of perception, grounded in and created from embodied vision and intelligible because of it, that is cinesthetically experienced by the spectators embodied sensorium, and cinesthetically expressed by the cinematic apparatus. Thus, the auditory and visual elements that make up the moving image are perceived by the human-lived body in a manner analogous to the synaesthetic condition.

The different sense impressions are experienced in a commutable process by the embodied sensorium that combines the sense inputs into one munificent sensation. This process is, as pointed out by Sobchack, reflected in and representative of the way the cinematic apparatus works. Accordingly, the synaesthetic analogy is present not only in the perceptive process of the human-lived body, but also in the expressive process of the moving image itself, enabling its aural and visual elements to create sensations far beyond temporal sound and visual movement. Thus, the image can appropriate the qualities of the sound, and vice versa, creating an aural-visuality. The notion of aural-visuality implies that the audio-visual arrangements of the moving image are not only commutable sense modalities enabling a merged sensory output/input, but also interchangeable modalities enabling the image to take on sonic qualities, and vice versa. The next section will investigate the element of sound in order to further examine the notion of aural-visuality. This will establish how the inherent qualities found in the sound and the moving images are interchangeable, so that the audio resounds the moving image, and the moving image visualises the audio.

1.3 Audio and the Image

Sound is the temporal element that ties the visual images together. It can stretch the visual space beyond the screen, creating an ambiguous and oscillating relationship between the image-space and the audio-space. Sound provides an expressive and informative richness to the moving image, dramatising the shots, creating a feeling of imminence and expectation, as well as one of immediacy. Music, the most important form of sound for this investigation, can add affect to the moving image, providing figurative, semantic or evocative values. But most importantly, sound can function as a visual element, and vice-versa, through the isomorphic and cross-modal system of audiovisual expression and perception.

Sound as linkage

According to Michael Chion, the most widespread functions of sound in the moving image medium is to bind the flow of images together, bridge the gap between visual breaks, and bring unity by establishing atmosphere: “as a framework that seems to contain the image, a ‘heard’ space in which the ‘seen’ bathes” (1994, 47). Furthermore, sound can be used to bring emphasis to the action or a specific part of the on-screen images, through punctuation. Film music functions in a similar way, even though it usually exists outside the world of the film. In this way, film music can enhance the effect of the on-screen action, providing evocative clues, or characterizing, even describing, a person or act in the film through the use of systematized and specifically assigned musical themes, also called leitmotifs. For example, in the beginning of Steven Spielberg’s *Saving Private Ryan*, a slow tracking shot in the American military cemetery in Normandy goes past several of the white marble crosses (see fig. 1.1). Each time a cross goes past the camera, a thump can be heard in the music, almost like a heartbeat. This little sound, part of a larger musical piece, signifies the short life lying underneath each cross, and expands the geographical place into psychological space. In this way, the music creates a cinesthetic element that has a direct impact on the viewer.



Fig. 1.1 – The camera moves from right to left. A thump is heard as each cross traverses the screen.

Another example can be drawn from the music video. In *Only You*, the men looking down on the main characters from surrounding windows always appear in a direct relationship with a specific musical horn sample. The horn sample becomes tied to the men in the windows and the creepy quality of the sample affects the viewers' perception of them, and as we shall see later, the eeriness of the men themselves affects the perception of the horn sample. Thus, the sound, whether musical or ambient, works directly and physiologically *on* the viewer by way of the entire embodied sensorium. As a result, the sound becomes a subtle means of affective and semantic manipulation that influences the perception of the image. It makes the viewer see more than what is actually present on the screen, or see it in a different way (Chion, 1994, 34). This correlates to Laura Mark's notion of the multisensory condensation through the mediated audio-visual image mentioned earlier (2000, 214) and the multi-sensory joinings of Kandinsky and the Symbolists. The sound affects the viewer's vision, and the vision affects the viewer's aural perception, creating one munificent sensation that is perceived by the viewer's entire body. This sensation is, as Sobchack points out, informed by the complete record and carnal awareness of the viewer's acculturated sensorium (2004c, 77). Consequently, the linked audio elements enhance and alter the feelings associated with the on-screen situation perceived visually, and as a result directly affect the visual lines creating an aural-visibility.

Synchresis: Audio-visual welding

The most important aspect of this aural-visibility is the synchronism or weld that Chion has called *synchresis*.¹⁴ This is a link that develops between a sound phenomenon and a visual phenomenon when they occur at the same time on and

¹⁴ A joining of the words synchronism and synthesis, implying the spontaneous and irresistible weld between a sound and an image occurring at the same time. See Chion, *Audiovision* p. 63.

through the screen. The synchresis relationship works independently of any rational logic, and enable sounds to enhance the image either realistically or poetically, forming meaningful or signifying agglomerations in our perception (Chion, 1994, 63). In *Only You*, the rising sound of strings and horns becomes the synchresis to a specific rising camera movement. As the sound line rises so does the visual line and the viewer feels that the rise of the music is the sound of the image moving, and vice versa. Similarly, in Jonathan Glazer's music video for Radiohead's *Karma Police*, the imagery of flames engulfing a car is connected to an undulating and piercing high-pitched synthesised sound (see fig. 1.2). The visual movement reverberates according to the sound, and the sound lines echo the visual movement. Thus, sound can add value to the moving image either as an on-screen element, floating on the surface of the images, or off-screen beyond the images, but always as an integral and reciprocal part of the moving image.



Fig. 1.2 – A high piercing, undulating synthesised sound is echoed by flames engulfing a car.

This sound-image relation is the expressive and informative value through which sound enriches the image. Furthermore, added value temporalises the images. It can render time in the moving image as exact, detailed, immediate and concrete or, on the contrary, it can render the moving image time as vague, fluctuating and broad. Synchronous sound imposes a sense of succession in the images, and orients these images towards a future or goal (Chion, 1994, 13). But most importantly, Chion goes on to state, added value creates the definite impression that what is heard comes from that which is seen, and is therefore already contained in these images creating a sense of immediacy (1994, 5). However, since the sound and the moving image are separate elements, often recorded, even created, at different times and in separate locations, it is not a given that the images and the sound actually belong to each other. But because of synchresis, sound and moving image become directly and intimately linked, creating a unified channel of sensory input that enables moving image expression and a synaesthetically analogous perception of this moving image.

This unified channel of sensory input implies that sound and the images exist and work in a transsensory and cinesthetic relationship. Following my proposed analogy, sound and moving image are therefore not perceived as separate elements, but rather as a singular sensory line in which the images visualise the production of sound, despite their often separate and incompatible origin and status. Thus, the sound and the moving image co-exist in a transsensory relationship in which one sense flows over and into the other; enhancing, altering and ultimately expanding its affect on the listening viewer (Chion, 1994, 167). Although Chion claims that the transsensory relationship between sound and image does not correspond to the intersensorial ideas of the symbolist movement, in which one sense could become the other, it is apparent that the two concepts express the same notion of cross-modality. They both highlight the notion that sensory inputs do not exist or are not perceived independently of each other, but rather that they function in a reciprocal, commutable and transposable system.

In this system one sense modality collaborates and is commutable with the other sense modalities, completely in line with Sobchack's concept of cinesthetics where the sense modalities are seen as one commutable and co-operative system of perception and expression. Through an agglomeration of the moving image and the sound in a reciprocal and interdependent relationship that crosses the boundary between the senses, this system generates "rhythmic, dynamic, temporal, tactile, and kinetic sensations that make use of both the auditory and the visual channels" (Chion, 1994, 152). Thus, as shown with the synchresis between the rising sound and the rising camera in *Only You*, Chion's audiovision can be said to be a cinesthetic system that saturates and short-circuits our visual perception. However, it is important to understand that this does not mean the eye and the ear work in the same way.

Audio-visual transposability

The ear works much faster than the eye. Audio lines are processed, synthesised, and analysed much faster than visual lines, producing a sound trajectory with a much clearer and more definite shape than the image. As Chion says, while the eye must simultaneously explore space and follow along in time, the ear isolates details

instantly and follows these details in time (1994, 10). Quick bursts of multilayered lines of sound can be easily registered and analysed by the ear, while rapid successions of complex visual lines causes the eye to struggle. Rapid movements that the eye struggles to analyse will be perceived by what Chion calls the “ear-that-is-in-the-eye”, an auditive eye that traces out the lines of rapid visual movements as a trajectory of a complex series of auditory phenomena, and subsequently etches them into our consciousness, quickly and clearly (Chion, 1994, 134). The camera that rises in *Only You* is combined with a rising sound that the auditive eye traces, aiding the eye in its analysis of this visual movement. Thus, as a result of the commutable and reciprocal system at work in the cinesthetic audiovision, the sound ‘spots’ the images. This provides an added value to the complex visual moment that leaves a strong audiovisual ‘memory’ that aids in the visual analysis of the complex visual line.

Such audiovisual assistance can be directly related to Eisenstein’s theory of overtone montage, as we shall see in Chapter Two. In this form of audio-visual montage, the sound and moving image elements are mounted together according to the similarities, or ‘feel’, of their inherent lines. Accordingly, a line of sound will be put together with a visual line that resembles or ‘feels’ similar or vice versa, generating an audio-visual flow in which the perception of the complex visual line is sustained by the reverberation of sound, and the audio is resounded by the moving image lines. Thus, the sound surrounds the viewers and reverberates within them, while constantly, and without drawing attention to the process, synergising with their vision. But the visible can also transpose a certain sonic velocity into the moving image.

Images containing fluttering and vibrating light, shifting or undulating smoke, create a luminous patterning and a visual volubility that behaves like sound; flowing, coming to the fore and fading away. An example of this can be seen in Mark Romanek’s video for Keith Richards’s *Wicked As It Seems* where light flashes, and in-camera flashes, take on the quality of fluttering sparkles of sound. These visible textures are what Chion calls microrhythms (1994, 134), and they mirror what Laura Marks calls intersensory links, by which a sound can evoke a texture or an image can evoke smell (2000, 213). The sound and the image are perceived simultaneously by the sensorium of the viewers as one unified expression of perception. “Each image is synthesised by a body that does not necessarily divide perceptions into different sense modalities”

(Marks, 2000, 222). Furthermore, the mechanism of the microrhythms can also be correlated to Sobchack's notion of cinesthetics. As pointed out earlier, when one sense modality is engaged, for instance through music, it will correspond to and enhance the experience of another sense modality, for instance vision. Thus, the auditory and visual sense modalities of the moving image body appeals to the lived-body of the viewer, bringing forth associations from the other sense modalities, as well as making one take on the role of the other, like a twinkling of light being perceived as a tinkling of sound, generating an aural-visuality.

Chion's concepts of synchresis and microrhythms can be seen as concrete expressions of the cinesthetic qualities of expression and perception laid out by Sobchack. Consequently, Chion's audiovision and its transposition of sound into moving image and moving image into sound, as well as the transsensory attributes of the images' aural-visual elements and the auditive eye, function as an embellishment of Sobchack's notions of the cinesthetic moving image. The sound and the image melt together into aural-visuality, experienced and expressed through a synaesthetically analogous system, that creates a unified and commutable sensory expression in which it is hard to determine whether it is the sound that make the images move and expand, or if it is the images that give the sound shape and visuality. Thus, the arrangements of sound and visuals in the moving image body generates an aural-visuality in which images and sound merge in such a way that sound can be manifested visually, and moving images resonate aurally.

1.4 Conclusion

The senses do not operate independently. They are part of a non-hierarchical and unified system in which one sense modality is commutable and transposable onto and into the other sense modalities of the lived-body sensorium as it perceives, experiences, and expresses the world. The process by which this commutable sensory system works is analogous to the synaesthetic condition. Furthermore, this interconnection of aural and visual elements is applicable to the moving image because it is embodied and synaesthetic just like the human lived-body. This is so because the moving image body is reversibly co-constructed by the lived-body of the viewers. Such a structure, explained by Sobchack's notion of cinesthetics, highlights the welding of sound and image into one channel of sensory input. In this way, as in the direct experiences perceived by the lived-body sensorium, each audiovisual sense the moving image expresses, cinesthetic to start with, is met by a rush of other sensory associations. This instrumentality of the cinesthetic moving image enables the transposition of sonic velocity into the image and visual shape onto the sound, generating a cinesthetically munificent aural-visuality that is analogous to the synaesthetic condition.

In the next chapter, I will explore the concept of aural-visuality in terms of the music video. This particular form of moving image assembles and mounts visuals according to musical patterns, interpreting sound through image by employing the embodied cinematic apparatus, overtone montage, and the notion of cinesthetics. The next chapter will show how the music video concretises the synaesthetic analogy and the notion of aural-visuality established in this chapter. In this way, the music video can be used to explicate how the sound can echo the image, and the image resound the audio.

Chapter Two – The Music of the Image / The Image of the Music

“Video is as close the cinema
has come to the impressionism
of the lyric poem”
-Frank McConnell

2.1 Introduction

The music-video is the moving image form that most clearly exemplifies both the cinesthetic analogy and the notion of aural-visibility established in the previous chapter. The music video is at the forefront of visual experimentation, especially in terms of the unification of audio and moving image. While drawing heavily on cinema, television and photography, it should be seen as distinctly different from these, as it has its own way of arranging audio and visual elements, through the exploration of their inter-connections and cohesive temporal flow. Accordingly, the music video may be seen as an experimental audiovisual genre that concretises the notion of aural-visibility. The purpose of this chapter is to conclude my discussion of cinesthetics through the critical application of this theory to an analysis of the music video and its elements. Thus, the chapter contends that the music video expresses its aural-visibility through a cinesthetic montage of audio and moving image elements. This cinesthetic montage produces an audiovisual synthesis of music and video that is expressed by and perceived through the embodied sensorium, thus facilitating the sonic reverberation of the moving image and the visualisation of the sound.

In order to achieve this, I firstly provide a historical context of the technology of the music video, particularly the development of the process of music and visual synthesis. This background traces a line from the sound-image experimentations of composers such as Wagner and Scriabin, throughout the scopitone machines of the 1950s, the image projection systems of the 1960s and video scratching of the 1980s, through to the advent of the music video. Secondly, I discuss the conceptual techniques put forth by Carol Vernallis that facilitate the transposability of music and moving image. These techniques provide a visual framework for the analysis of music videos, as well as for the production of a music-video. Vernallis’s concepts are discussed in a logical progression from the more general macrostructures of form and

flow, timbre and texture, to the more specific microstructures of contour, phrases, motifs and visual shapes. Finally, in the last section of this chapter I propose that the elements of Vernallis' visual framework can be understood as cinesthetically merged with the audio. In order to support this statement, I apply Eisenstein's theorisation of the process of overtone montage. This process generates an audiovisual ecology in which all sound and moving image elements are interconnected creating *one* amalgamated expression of music-video. The application of Eisenstein's theory does not collapse the former with Sobchack's notion of cinesthetics. Rather, it presents a fertile correlation between the cinesthetics of the moving image and the synchronisation of the senses as facilitated by the montage process. Consequently, the chapter follows a three-part logical progression from historical background, through conceptual techniques, to an exposition of the cinesthetics of the music-video montage.

2.2 The Music Video Context

This section will outline links from the early experiments of music-visual synthesis carried out by Kandinsky and Schoenberg, Scriabin, and Wagner, through the image projection systems of the 1960s and onto the scratch video experiments of the 1980s in order to provide a technological context for the music video as a form of music-image synthesis. The experimentations of the scratch-video artists will provide a connection to the cinematic apparatus of the music video format, while the links to the image projection systems will outline the relationship between Kandinsky's total art and the cinesthetic experience that can be provided by the aural-visibility of the music video.

Music and image synthesis

The synthesis of visuals and music is not a new notion. As long as both mediums have existed they have been mixed. In the 1900s, as pointed out in Chapter 1, composers such as Schoenberg collaborated with the visual artist Kandinsky to create total art (music and image together). Furthermore, composers like Rimsky-Korsakov and Scriabin, had long been interested in and experimented with the connection between music and coloured visuals. This can be seen as the musical expression of the cinesthetic literary experimentations of the symbolist movement. So-called colour organs, organs that displayed coloured lights connected to tones being played, were already a popular dance-hall act. Scriabin even wrote a separate part for a mute keyboard controlling coloured lights/shapes/beams for his piece *Prometheus: The Poem of Fire*.¹⁵ The German composer Richard Wagner had been experimenting with what he called 'mood setting' in his operas. He produced a mixture of music, theatre, lights and other visual elements in a darkened theatre in order to fully immerse the audience in a multi-sensory experience.

The conventions pioneered by Wagner and Scriabin were further developed and advanced through experimentation in the 1950s with events such as the Vortex concerts, where images were projected onto the dome of a planetarium accompanied

¹⁵ See Cytowic, R. E. (1989). *Synesthesia: A Union of the Senses* and Kandinsky, W. (1982). *On the spiritual in art*.

by music, and in the 1960s through the use of image projection systems in musical concerts (see fig. 2.1.). These systems were comprised of psychedelic imagery created through the use of lights, film and slide projections, and oil/water colour on glass slides to create organic ever changing “[...] multi-coloured, amoebic forms which dissolved and reformed in a dream-like fashion” (Walker, 1986, 103). The idea was to create a more complete sensory experience, just like Scriabin and Wagner wanted, and fully comparable with the textual cinesthetic experimentation of the Symbolists.



Fig. 2.1 – Liquid lights would move and morph according to the music, in an endless variety of shapes.

The English band Pink Floyd claimed they wanted to create a 360° experience through the synthesis of music and images. Although the music and the visuals were not synched in any exact fashion, a certain sense of synchresis was achieved through the application of imagery that thematically and stylistically complemented the music. The projections were also performed live and can therefore be said to be directly influenced by the music, much like the direct influence the mood of the sound and/or music has when performing an overtone montage. “For maximum effect the audio-visual stimulation had to be all encompassing, hence the efforts of musicians and projectionists to create a sound-image environment” (Walker, 1987, 102). Thus, such experiments can be seen as an early expression of the much more complex audio-visual fusion carried out by the music-video makers. With the introduction of the Video Cassette Recorder (VCR) and consumer oriented video-based recording and editing equipment in the late 1970s, image projection became easier and more economically viable, as well as changing the focus from liquid lights to the montage of televisual moving images.

The introduction of the VCR also created a new form of collage art, namely the scratch-video. Appearing in the US in the late 1970s and in the UK in the early 1980s,

scratch video is the moving image equivalent of photomontage and collage.¹⁶ Video artists would record broadcast images directly from the television and use these as raw material (Rush, 2005, 90). However, it is important to understand that scratch-video was not a music video, even though its techniques and use of juxtaposition and montage should be seen as the predecessor for the music-video apparatus. Since the objective was the creation of a strictly visual flow, the video scratch artists only used music, if they used it at all, as accompaniment, occasionally cutting to the beat of the song (Walker, 1986, 152). Consequentially, only an outline of the synchresis and image-sound synthesis we find in the music-video was achieved, as with the image projection systems. This is because, as we shall later see, the production of music-video implies that the sound takes centre-stage, turning the moving image into accompaniment.

Technology and the direct experience

In terms of synched sound and moving image experiences, there were 'experiments' done even before the advent of the sound film, since most silent films were accompanied by a piano, theatre organ or a full orchestra (Thompson & Bordwell, 1994, 13). With the advent of the sound film came *The Jazz Singer* and many other films that combined moving images with music, including several short musical clips to promote these films featuring acts such as Cab Calloway, Billie Holiday and Duke Ellington (Reiss & Feineman, 2000, 13). These were followed in the 1960s by The Beatles's *A Hard Day's Night* and The Monkees's *Head*. Both these films made use of a bold and visually engaging style, utilizing jump cuts and surrealistic imagery in a fashion that anticipated the music video by fifteen years (Reiss & Feineman, 2000, 15, and Mundy, 1999, 192).

With the introduction of television and popular music programs, the demand for international music acts, as well as an explosion of the number of acts available, made the promotional clip an economic necessity. These clips were made to satisfy the ever-demanding market without having to send the musical act around the world on

¹⁶ This form of audiovisual synchronisation is far from over. With the introduction of the VJ (video-jockey), and easily accessible and usable digital equipment, a new culture of audiovisual synthesis and live visual performances has emerged. See Spinrad, P. (2005). *The VJ Book: Inspirations and practical advice for live visuals performance*.

expensive and arduous trips. In 1965 The Beatles made a series of such clips in order to satisfy this growing demand for their appearance on broadcast music shows (Mundy, 1999, 207 and Howells, 2003, 236). Many of these early clips could be said to be conceptual clips rather than simple performance recordings, which broke the dominant performance aesthetic of the time. Thus, such clips paved the way for the expanded and experimental iconography and visual economy of the music-video we know today.

The production of music video clips grew rapidly throughout the 1960s and 1970s, gaining even more momentum as film and television technology advanced, particularly with the introduction of the broadcast video recorder. Not only did these advances in the television and cinematic apparatus offer new ways of visualising and representing the music, it also made the programs, and consequentially the clips themselves, into a commercial and cultural commodity.¹⁷ In 1975, a promotional clip was made for Queen's *Bohemian Rhapsody*, which proved to be an immensely successful marketing tool. This can be said to be the birth of the music video as we know it (Reiss & Feineman, 2000, 16). Following on from this, Music Television (MTV) and the continual broadcast of music videos was introduced five years later. This led to an explosive and continuous expansion of the experimentation and progression of form and technique throughout the 1980s and 1990s, and into the 2000s. According to Mundy, the development of the visual strategies found in the music video should not be seen as original innovation, but rather a continuation of the visual economy developed over decades by the film and television industry in their effort to visually represent music (1999, 216). But as I have shown above, this visual economy can actually be traced back to even before the start of the moving image industry.

The efforts of Wagner and Scriabin, the experimentations of the vortex group in the 1950s and the image projection systems of the 1960s outlined above, demonstrate that many of the concepts and devices of the music-image synthesis endeavour were already present. As the director of The Monkees film *Head*, Bob Rafelson said:

¹⁷ For a discussion on the shaping of the visual economy of the music video as a result of the commercialism of the broadcast medium see J. Mundy (1999) *Popular music on screen: from Hollywood musical to music video*.

“Almost all the effects you see in video today, the psychedelic solarization, the quick cutting, are things we were doing years ago” (quoted in Reiss & Feineman, 2000, 15). Furthermore, the devices utilised by the image projection movement and the vortex group, such as multiple overlays, multiple exposures, juxtaposition through montage, and painting and/or scratching directly on the film, had already been developed through the moving image experimentation of the surrealist movement and Man Ray, and experimental film-artists such as Len Lye.¹⁸ These techniques were then appropriated by video artists such as Nam June Paik and the scratch-video movement and modified to the VCR and non-linear editing equipment, in their efforts to produce experiences that would blur the division between fact and fiction, past and present, and reality and memory.¹⁹

Hence, one can see that although the video-scratch artists did not make music videos, the techniques they adopted from the experimental film movement and others, and customised for the VCR, now make up some of the most important tools deployed by music video makers today. In summary, I think that the experimental concepts that form the basis for the moving image techniques, as well as the notions of music and visual synthesis, can be traced all the way back to Schoenberg, Scriabin and Wagner, via the music-image experimentations of the 1950s and the image projection systems of the 1960s. The common denominator that connects all these concepts is the desire to create an all-encompassing, multi-sensory and directly experienced audio-visual environment. A current proponent of such an experience can be found in the music video.

¹⁸ See Milicevic (1997). *Experimental/Abstract Film and Cinesthetic Phenomena (1725-1970)*, Booklet from Len Lye video, *Avant Garde: Experimental cinema of the 1920's and '30s* (2005), and Matthews (1971) *Surrealism and Film*.

¹⁹ See Rush, M. (2005). *New Media in Art*.

2.3 The Music Video Framework

According to Chion, cinema is a place of images plus sound, where the sound is that which seeks its place within the space of the moving images (1994, 47). In the music video, on the contrary, it is sound that is the space and image that is the added element finding its place. The way in which the image locates this place is through an emulation of the form of the sound to which it is linked. To show this it will be fruitful to deploy a set of conceptual techniques outlined by Carol Vernallis.²⁰ This set of techniques provides a visual framework for the music video in which its moving image elements are broken down into exact and discernible components according to the musical sections of the video. The framework uses the musical section as the fundamental unit of the analysis, rather than the shot, which is used in the analysis of narrative cinema. This is done in order to place emphasis upon the varied repetition of materials rather than narrative and the linear development of plot. Furthermore, such a framework highlights the intimate connection that exists between sound and moving image in the music-video, a connection that will be examined in the video analysis in Chapter 3. The concepts will be discussed in logical order from the broad and all-encompassing concept of form and flow, followed by the more specific aural-visual connections of timbre and texture. Lastly, I will look at how contours, shapes, motifs, and phrases generate synchresis between exact aural and visual instances.

Form and flow

The form can be said to be the overall or large-scale structure of the music video. As pointed out by Vernallis (2004, 214) it outlines the broad connection between the sound and the image, and can therefore be said to be a thematic, as well as iconographic, breakdown of the music. In this way, it is a broad translation of the mood and feeling that pervades the sound. Thus, the form is a comprehensive set of visual elements, like colours, shapes, *chiaroscuro*, or sets, which is deployed to emulate the comprehensive mood and sensation of the sound. This indicates the most important difference between the way sound works in cinema and the way it is used in the music video. In cinema, the soundtrack, whether music or soundscape is used to

²⁰ See C. Vernallis (2004) *Experiencing Music Video: Aesthetics and cultural context*.

augment the image: it follows the image, trying to find its place within or around it. In the music video, it is the image that augments the sound; it is the foundation on or around which the visual elements are situated. Thus, one can construe that the visual elements are conceptualised *from* the music. The sound is the starting point from which the visual form is drawn, and to which the visual elements that make up the form are intimately connected and ultimately return.

It can therefore be said that the music video translates music into form, not into content, in a manner that reflects Roger Fry's comments that art is the expression of emotion communicated through formal arrangements, and not through mimesis or 'realistic' allusions (quoted in Howells, 2003, 35). Furthermore, the formal arrangement of the images reflects the way in which sound itself works. As indicated by Howells, and echoed by Chion and Ihde, music is not mimesis (2003, 36). It does not provide a copy of the world; rather it is an affective translation of an embodied emotional response to the world. Its affect on the listener is therefore achieved not through content, but through form expressed by the embodied cinematic apparatus. Formal arrangements of aural elements such as melody, pace, timbre and texture, arrangement, and acoustic space, reflect and generate moods and sensations that wash over the listeners and communicate directly with their embodied emotions and memories. The visual form of the music video functions in the same way.

Broad visual associations, constituting formal arrangements of colour, *chiaroscuro*, mise-en-scene and so forth, working either through metaphor or metonymy, are linked directly to the music in order to create a visual form that reflects and transposes the form of the sound (Goodwin, 1992, 55). In The video *Only You* the blue/green colour space and underwater mise-en-scene visualise the cold yearning emotion of the song and fluid qualities of the sound space. The large-scale structure of the claustrophobic setting of a backstreet surrounded by tall buildings, further augmented by the underwater mise-en-scene and the 'watchers' situated in the buildings, also reflects the claustrophobic mood of the music. In *Cold Beverage*, the lazy and laid-back summertime mood of the song is reflected through a set of slow-moving black and white images of people drinking lemonade in a hot city setting while the band is playing on a street corner (see fig. 2.2). The form is the first step in generating a moving image that will be transposable with the sound it visualises. By creating a

comprehensive and over-arching visual structure that can be related to the large-scale structure of the song's affective mood and acoustic space, the music video generates a foundation in which the sound and the image can work cinesthetically. The transposability of the sound and the image is further expanded through the flow of the moving image form.

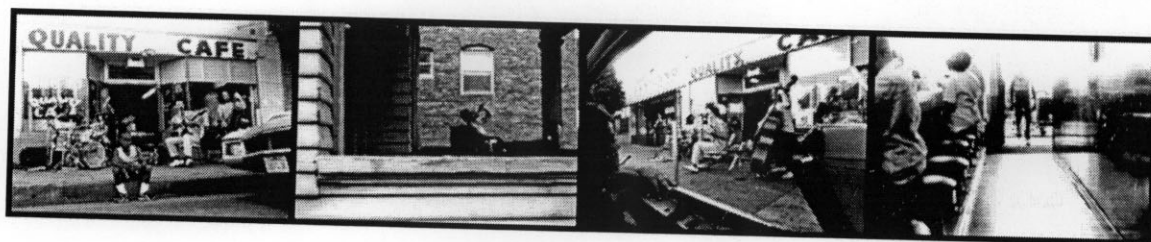


Fig. 2.2 – B & W images suits the low-fidelity blues of G. Love & Special Sauce in *Cold Beverage*.

The flow outlines how, and in which way, the form moves through editing pace, camera movements, and the movement located within the pro-filmic events. It is the next phase in rendering the sound into a visual shape. While form indicates the general concept and iconography of the song's evocative atmosphere, flow designates the general movement and timing of the form through a similar translation of the song's linear and sequential progression of aural instances. Thus, since the flow is the temporal action of the overall structure of the form, it is as intimately tied to the sound, as the form is tied to the sound. Movements in the shot, whether camera movements or the actions within the pro-filmic events, and the rhythm of the montage establish a characteristic rhythm of repetition and/or variation that reflect the number of beats, rhythm, tempo, and harmonic development of the audio (Vernallis, 2004, 212). Furthermore, the flow of the form also reflects the separation of the audio into the linear progressive elements of verse, chorus, bridge, and coda that constitute the song.

In this way, the repetition of certain parts of the mise-en-scene, and sequences, will reflect the repetition of the verse and chorus. Used as part of the montage process, Vernallis calls this procedure "visual rhyming" (2004, 160). In *Cold Beverage* the verse is visually represented by seemingly repetitive shots of the band playing, sometimes interlaced with a variety of disparate images that are connected through iconography or theme, while the choruses use shots of the lead singer coupled with

the same, or similar, disparate images deployed for the verse. Thus, we can see how the large-scale structures of the audio influence the form and flow of the moving images, and how this form and flow is audiovisually co-constructed. In addition to reflections of the temporal and linear flow of audio, the flow of images can also be determined by the timbre and texture of the sound.

Timbre and texture

According to Goodwin, timbre is the “colour of the tone” often described in terms of visual characteristics such as “brightness” and “sharpness” (1992, 57). This audio characteristic can be used to augment the visual form and flow associations and to establish a quality by which the two elements are linked together. Through the use of colour or *chiaroscuro*, the mood of the song, evoked through the timbre of the sound, can be reflected, as I have shown with *Only You*. By applying certain textural qualities to the image, for instance grain, tint, out-of-focus imagery or other forms of visual ‘noise’, the image can take on more of the timbral qualities of the sound it illustrates. These kinds of images are what Laura Marks calls “haptic images”, and used in combination with sound, montage and camera movement it achieves “sensuous effects” (2000, 172). Such ‘noisy’ images are also what Sobchack reflects upon in regards to cinema in her essay on the inaugural shots of the film *The Piano*.²¹ The “unrecognisable blur” of these out-of-focus shots, Sobchack claims, was understood by her body, and comprehended by her fingers that: “*grasped* it with a nearly imperceptible tingle of attention and anticipation and, offscreen, ‘felt themselves’ as a potentiality in the subjective and fleshy situation figured onscreen”, even before she refigured her carnal comprehension into conscious thought (2004c, 63). This sensuous perception is equivalent to the cinesthetic aural-visuality of the music video. The texture of the image becomes a visual translation of the acoustic space generated by the music. In this way, the sound can be cinesthetically represented through the use of the image’s visual surface.

²¹ See Sobchack (2004c). What My Fingers Knew: The cinesthetic subject, or vision in the flesh. In *Carnal Thoughts: Embodiment and moving image culture*.

The grainy black and white images used in *Cold Beverage* augment the old-style bluesy and low-fidelity acoustic space of the song, visually representing the song's timbre (see fig. 2.2). Furthermore, timbre can be seen to be the linear progression of the sound changing over time. If the timbral qualities of the audio changes over time, in other words creates an aural 'movement', it can also be made visual through the editing pace, or movement of the camera and/or the pro-filmic events. A change to a 'sharper' timbre can for instance be visualised by a distinct camera movement (for instance a quick tracking, or rising crane shot). Changes in timbre can equally be reflected by changes in the visual texture, by applying more grain or noise or changing the colour space or light-to-shadow ratio. In this way, the texture of the image becomes the audio timbre equivalent, and thus, the 'sound' of the sound is reflected by the visual quality of the visuals. The audio itself can also contain textural elements, like noise, scratching, electronically produced loops or a wide variety of other audio elements. These aural textures can be visualised in much the same way as the timbre. Textural elements are usually added in the production stage of the music, rather than the actual recording process. Thus, they become elements that are deposited onto the surface of the recorded audio in order to create a soundspace that more accurately and munificently expresses the song's emotion and sensation.

Given that these aural textures are located on the surface of the song, surface textures of the image can be equally deployed to cinesthetically translate the audio-textures. This will enable what Laura Marks calls "haptic looking", a concentration or focusing in on the surface of the image in order to discern its texture (2000, 162). These haptic images, she points out, are so incomplete that the viewers are forced to deploy their imagination and memories in order to complete them. In this way, such an image: "forces the viewer to contemplate the image itself instead of being pulled into narrative" (Marks, 2000, 163). This also echoes Sobchack's "carnal comprehension" of the initial shots of *The Piano*. Thus, the visual textures will intermingle and commute with the audio textures to augment the aural-visibility. As with timbre, certain textural elements can also be visualised through movement. In *Only You* the scratching, repetitive loops are visualised through ramped movements created by speeding up and slowing down the pro-filmic movements during the editing process. In this way the young boy in *Only You* seems to move in accordance with the textural scratching elements, or from a cinesthetic point of view, the audio elements are

moving in accordance with the boy. By connecting the visual quality of the image space to the aural quality of the acoustic space, the sound and image is further augmented into transposable elements that connect and associate cinesthetically through the aural-visual system.

The large-scale elements of the moving image form reflect and visualise the form of the sound. Equally, the surface textures of the sound are replicated by elements located on the surface of the image. In this way, the overall form and flow, as well as the surrounding surfaces of the form, can be said to enable sound and moving image transposability, creating a cinesthetic aural image. The next step then, is to create small-scale and specific instances of sound and image transposability by deploying Chion's concept of *synchresis*. While form and flow, timbre and texture, point to the macro-structure of sound and image transposability in the music video, the contours, shapes, motifs, and phrases highlight the microstructure and detailed syncretic connections of the music video's cinesthetic qualities.

Contour, phrases, motifs and the visual shape

Contours, motifs, phrases, and shapes, as listed by Vernallis, are elements that identify different aspects in which the sound and image transpose on a micro-level within the form and flow of the music video. Each aspect indicates one particular type of sound-image transposability. Originating from audio jargon, these aspects point to the close connection that exists between the aural and the visual elements that make up the music video. In addition, it provides an interesting reference to Goodwin's notion that the music must be seen as the starting point for all understanding of the visual elements of the music video (1992, 92). The music establishes the parameters and logic within which the visuals progress. Often, the music will be visualised through microrhythms, which were discussed in Chapter One, such as the twinkling of light or the undulating flow of flames. These microrhythms, or visual sound patterns, will often be directly linked to a particular sound element, replicating its phenomenological qualities through inventive use of the cinematic apparatus.

Contours can be said to be the particular shape of the musical lines that make up a piece of music as it rises and falls (Vernallis, 2004, 160). Each element of the song,

for instance the bass line or the guitar riffs have a contour. These musical 'shapes' are then visualised through image shapes that correlate in some way to the musical lines, for instance through similar direction of movement, twinkling qualities or direct visual hits. Furthermore, the contours can also be correlated to emotional affects. For example, jagged lines in the music will create sounds that seem anxious and intense (Vernallis, 2004, 213). Such audio qualities are then translated into image shapes that, through the use of microrhythms and/or the other visual forms outlined above, mimic the sound in such a way that the sound and image commute.

A rising sound, or contour, can be visualised through a rising camera. The camera traces the arc of the sound to which it is linked, emulating the shape of the sound, accordingly augmenting the audio-visual sense impressions generated by the music video. Phrases delineate the melodic variations and repetitions of certain instrumental sections that make up the form and flow of the audio, and are used to correlate the image and the sound within the overall form and flow of the music video. Particularly, the start and finish of a phrase often acquire visual attention, as this is usually where the contours of the sound elements are most prominent, reaching either a peak or opening up towards a significant development in the audio. To visualise this progression the camera can move according to the contour arc highlighted by the phrase, through instances such as a whip-pan, zoom, track or crane shot, or the change within the phrase can be emphasised through a quick change of shot, for example from Close-Up (CU) to Wide-Shot (WS).

At the smallest level of sound and image transposability one can find the motif. According to Vernallis this is a "recognizable melodic/rhythmic shape" that can gain meaning and distinctiveness through visual underscoring" (2004, 156). These aural elements are small, direct links between specific audio elements and precise visual moments achieved through synchresis, and are therefore directly comparable to Chion's visual microrhythms. The motif can for example be a small musical detail that is repeated or the shape of a recurring guitar section, and visual correlation can be achieved through synchresis or through the repetition of parallel visual shapes. This parallelism further highlights the cross-modal relationship between the sound and the images. As Vernallis points out: "The repetition of these visual and musical materials tells the viewer that an isomorphic relationship exists between them" (2004, 157). The

crash of a cymbal can be correlated to a flash of light; very effective indeed if the flash is created through the movement of the cymbal in front of a sharp light source, as the cymbal is hit. In *Wicked As It Seems* a rhythmically regular drum hit is correlated to flashes both within the image and on the surface of the image (see fig. 2.3). Thus, the shape of the sound, whether located in the contour, the phrase, or the motif is mirrored through visual shapes created by the mise-en-scene and the cinematic apparatus, creating an audio-visual commutation and transposability that creates the cinesthetics of the music video.



Fig. 2.3 – A drum hit correlates to flashes in the image and on the surface of the image.

The images of the music video contain shapes that, in order for the video to create an aural-visuality, must parallel and emulate the qualities of the sound elements. As shown above, this can be done on all levels of the music video image, whether it be through the large-scale structure of the video, the surface of the image, or the microstructures within the image. Thus, by using the song and the aural qualities inherent within its audio space as a starting point, visuals are created that correlate to the phenomenological qualities of the sound in such a way that the images become the sound, undulating and streaming around the viewers, pulsing and reverberating through them. When the visual elements emulate the quality of the sound structures, when the visual form is parallel to the aural flow, when the timbre is made into texture, and the motifs are generated into microrhythms, the image and sound commute and transpose into a cinesthetic aural-visuality. This aural-visuality surrounds and immerses the viewers in a cross-modal flow of music-video. In the next section I will look at how Eisenstein's method of overtone montage can be applied as a comprehensive process that cinesthetically connects and bonds together the sound and the moving image of the music-video.

2.4 The Music-Video Montage

The music video is a moving image form that constantly experiments with the synthesis of audio and image in order to achieve a level of commutability and transposability at which the video's audio-visuality creates a total experience. This experience can be seen as the post-modern extension of Kandinsky's total art, the 360° effect of the image projection systems, and the general intent to combine image and sound in such a way that the two elements transfer onto and into each other. The sensorium that informs the body does not separate the sensory inputs but makes them available to us as one munificent and cross-modal flow. In this way, the perception and expression of sound and moving image does not separate or give more weight to one element over another. The music-video works in the same way. Sound and moving image are joined together in the formation of an aural-visuality that surrounds the viewers and reverberates within them, without drawing attention to the process of synergising the image with the sound. But it is clear, as pointed out earlier, that in terms of the music video the sound is the primary element. In cinema, sound and music are added onto the images to enhance and expand the visual experience, whereas in the music video the images are added onto the sound to visualise its form and flow. It is the mood and 'feeling' of the sound that provide the cues for the atmosphere of the visual space. And it is the rhythm, timbre, and harmonic development of the sound that shape the form, flow, and quality of the images. This audio-visual bonding can be concretised through an application of Eisenstein's process of overtone montage.

One plus one equals three

Montage is first and foremost a theory of how to combine two images to create a third separate meaning that is greater than the sum total of the two images that are put together. This technique was developed by Russian filmmakers and theorists in the 1920s in order to create a form of cinema that was an aesthetic expansion of the political idea of dialectics (Sobchack, 1980, 69). Vertov, Pudovkin, and Eisenstein devised an editing system in which: "the meaning of a shot was not inherent in that shot but stemmed from its juxtaposition with other shots" (Sobchack, 1980, 66). In its most basic form, the montage theory equals the juxtaposition of image A with image

B, in order to create image C in the minds of the viewer, a totally new and synthetic image which does not really exist on the screen. The method was later developed and expanded, particularly by Eisenstein, in order to express a wider range of meanings beyond the simple verbal metaphor, through the clash or attraction between the images and the elements of composition, movement, and idea contained within them in order to give birth to a concept. This created a technique, as Sobchack points out, through which:

Screen time is expanded and contracted as actions are broken down and reassembled into separate rhythmic units; screen space in image after image becomes a dynamic field of motion which can then be juxtaposed with a contrasting parallel motion in the following image (1980, 67).

Although the invisible form of editing, pioneered by D. W. Griffith and perfected by the Hollywood system, is now recognised as the standard editing convention for mainstream cinema, the montage technique is still widely used, particularly in commercials and in music videos.

The basic concept of montage ($A + B = C$) was augmented by Eisenstein into a system that was far more elaborate in terms of its capabilities of producing meaning, and to elucidate sound and image synthesis. With the introduction of sound to the cinema, and Eisenstein's subsequent examination of the commutability of the aural sense and the visual sense²², the montage system came to regard the audio and the image as integrated and commutable elements of one cross modal flow rather than separate sensory inputs. Furthermore, Eisenstein pointed out that the only place from which one could draw the necessary rhythm and 'feeling' of how to express this cross-modal perception was from the richest source of experience, namely; "Man himself" (1970, 61). Thus, despite the differences in methodology, I feel this notion draw a parallel between Eisenstein's approach and Sobchack's cinesthetic approach.

²² See An Unexpected Juncture, in *Selected Works Vol. 1: Writings 1922-34* (pp. 115-122) for an overview of Eisenstein's consideration of the synaesthetic qualities of the kabuki theatre in which the transference of affective intention from one material to the other and the fusion of sound and image senses generate an audio-visual expression that creates non-differentiation of the audio-visual perceptions. Also see The Fourth Dimension of Cinema, in *The Eisenstein Reader* (pp. 111-123) and the Synchronization of the Senses in *The Film Sense* (pp. 60-92) for a consideration of the synaesthetic aspects of the cinema.

It draws attention not only to the embodied sensory process that serves as a foundation for cinematic perception and expression, but it also locates the embodied configuration of the cinematic apparatus itself, which in this instance is concretised by the montage process. Sound and moving image elements are mounted together according to the physiological qualities of the elements. This welding of moving image and sound generates a direct embodied experience that speaks comprehensibly to all the senses.

Montage enables the creation of meaning and emotion on a level that is not possible through continuity editing. By ‘crashing’ images together the viewers are forced to participate in the creation of meaning and emotion, drawing on their embodied knowledge to fill up the expression given to them so that their perception expands and goes beyond what is presented, both in terms of meaning and sense perceptions. Thus, the montage process enables the viewer to integrate and combine the audio and visual lines in their own heads (Eisenstein, 1970, 34). This means that when the sound works in parallel with the underlying resonances of the image it creates vibrations that are perceived “purely as physical parallaxes on the part of the perceiver” (Eisenstein, 1998b, 121). Such a process can be correlated to Sobchack’s notion of how the audio-visual flow speaking to the viewer from the screen, is dynamically co-constructed into a reversible audio-visual expression and/or experience by the viewer.²³ This creates more than what is present in the audiovisual flow, as if one plus one actually makes three, because the embodied sensorium will ‘fill in the gaps’ and trace out the lines generated by the montage process using the sensorium’s embodied memory, as pointed out by Laura Marks (2000, 145) and Sobchack (2004c, 63). Thus, montage enables, one can almost say equals, embodied viewing.

In order to achieve this physiological effect Eisenstein devised a system of what he called overtone montage, drawing an analogy from music, that would take into account *all* the audio-visual lines, stimulants, vibrations, and secondary resonances that made up the shot (1998b, 119). This would enable perception to be enhanced from: “a melodically emotional colouring to a direct physiological sensation” (120). Eisenstein’s augmented montage takes into account the complex interwoven

²³ See Sobchack *The Address of the Eye*, p. 138, and section 1.2 of this exegesis.

collection of secondary “stimulants” or “resonances” that make up the sum total of the effect of the audio-visual elements (1998b, 112). According to Eisenstein, the montage of music and image should be based on the “sense” of the sound, as indicated by Goodwin (1992, 55). Thus, the assembly of the moving image and the sound segments is done according to the *feeling* and *rhythm* of the music rather than the metre rule (1998c, 170). With the result that the viewer will not simply hear the sound and see the image, rather they will “feel” them (1998b, 115), both in the process of expression and perception.

The compositional effects within the shot, the form and flow of the images, the contours, motifs, phrases, and shapes, the textures of the shot’s surface, the timbre and harmonic development of the audio and the rhythm and synchresis of the audio-visual elements, in short *all* the audio-visual qualities inherent in the images and the sound, are joined and interconnected. In this way, the sound and the images are welded together to provide one complex, cross-modal audiovisual expression. This cinesthetic montage of the audiovisual moving image enables the expression and/or perception of all the resonances contained within the interconnected sound and image elements. Thus, as a consequence of the cinesthetic process of the cinematic apparatus, an aural-visibility in which the sound and the image reverberate into and onto each other, ultimately echoing within the viewers’ sensorium, can be generated.

Eisenstein claimed that montage was the building up of separate sensations into the image, an image that would evoke the associations and sensations of the experience (1970, 26). This process mirrors Sobchack’s view of cinema as a flow of separate sensations experienced through a cinesthetic sensorium. Accordingly, the montage of the audio-visual images also mirrors the way the lived-world, and consequently the audio-visual moving image, is perceived. The eyes wander over the scene that is expressed, continuously refocusing on and picking out new ‘images’ to look at, cinesthetically combining them with the flow of other sensory inputs perceived by the lived-body sensorium in order to ‘see’ what it is we are looking at. Cinesthetic montage combines different audio-visual images, as well as all the sonic and visual resonances inherent within them, to move the viewer’s perception beyond that which is actually expressed. It enables sonic qualities to be transferred into and onto the

visual form and flow, bonding the music to the video and welding the video to the music.

The cinesthetic montage of music-video

The music-video immerses the viewers in a flow of aural-visuality in which the optical movements merge with the sonic reverberations, creating a cinesthetic expression that surrounds and involves not only the eye and the ear, but the entire interplay of the sensorium. According to Kevin Williams, the musical visuality of the music video: “is an aesthetic structuring of perceptual experience – a response to consciousness and a technological apprehension of a communicative world” (2000, 104). Using the music as a starting point, the video is constructed to create a flow of visuals that not only mirror and match the form of the music, but also inject the image with sonic qualities, thus visually manifesting the audio. The mood, or atmosphere, of the song’s audio space, can be made visual by *chiaroscuro*, colour, or the movement of light. Harmonic development can be visualised by camera movement or editing pace and the timbre of the audio can be visually resonated by changing layers of textures located on the surface of the image.

The image, or the elements located within it, can flow, undulate, reverberate, twinkle or pop in a similar fashion to the audio. Thus, the cinematic apparatus enables cinesthetic connections to be made between the sound and the image, generated by meticulous overtone montage. These connections are perceived through, or expressed on the basis of, the embodied sensorium. As pointed out in Chapter One, the human sensorium perceives the lived world through a conglomerate system of commutable senses. All the sense modalities provide impressions that are united into one cross-modal flow. There is no hierarchy. Rather, there is a flat structure in which each sense modality is as important as the next and the sum total and interconnectedness of the combined sense modalities is what matters. Such a combined cross-modal flow can also be shaped by artistically joining two or more sensory elements together, as done by the Symbolists and the likes of Kandinsky. This joining of the cross-modal associations enabled by the united embodied sensorium provides the direct experience of the lived world.

For the moving image, sound and visuals are joined according to the expressive directions of the filmmakers, presupposed on the directors' direct experience of the world and enabled by the synoptic and cinesthetic sensory system of their lived-body. This cinesthetic process, inherent in the human sensorium, is reflected in and representative of the way the cinematic apparatus works. Consequently, the viewers can translate, or co-construct, the audio-visual moving image through their application of sense memories located in their lived-body sensorium, because the cinematic apparatus expresses the embodied qualities inherent in the audio-visual elements it comprises. In addition, the expressions generated by the cinematic apparatus come from the embodied perceptions of the filmmakers that are, as we have seen, cinesthetic. This is also true for the montage process. As specified above, the *monteur*²⁴ 'feels' the primary component of sound and thusly senses how to augment the image components in accordance to the audio lines and how to connect the images to each other and cinesthetically link these images to the audio. Thus, the entire cinematic process by which the music-video is conceptualised, generated and constructed is embodied. This makes it possible to fuse the sound and image elements by mounting them according to their inherent lines or tones, in such a way that each element *cinesthetically* connects to the form and flow of the other element.

This cinesthetic connection is initially conceptualised from the music. The flow of audio elements serves as the starting point for the construction of the form and flow of the video elements. The images are generated and mounted together to reflect the music on a textual as well as a formal level. The mood, timbre, contours, and harmonic developments of the sound are mirrored and matched by the flow of images through formal conventions and technical experimentation. This generates a correlation between the sound and the image, a *synchresis* as Chion says, in which the form and the flow of already recorded sound connects to a disparate, yet fitting, form and flow of cinesthetically constructed and connected images. Such connections can be found, as we have seen, on several levels. There are broad correlations between the feel of the sound and the large-scale structure of the video's form, and there are micro-structural links between specific elements of the sound and particular visual shapes or developments. And it is these correlations that ensure the qualities of sound

²⁴ The person performing the montage. Take from the French word for a person who assembles or puts something together.

and image elements are transferred onto and into each other. To aid this cinesthetic commutation, there is the syncretic practice of the viewer's embodied and cinesthetic vision. It connects and links sounds to images, and vision to audio, even if there is no such connection with which to begin. This is what makes the image projection systems work, and what connects these systems and their audio-visual synthesis to the music-video.

Arguably, the visually abstract and infinitely morphing undulations generated by the image projection systems discussed earlier could be said to be more cinesthetically connected to the sound than is evident in the music video. The colours, shapes, and flow of the projected visuals are generated live to the sound, or expressed *from* the perception of the audio, in an embodied process. And because of the abstraction of the images there are no easily identified meanings, 'forcing' the sound and the image to be connected on a more cerebral and direct level, directly plugging into the sensorium and bypassing objective reason. But, such imagery is also generated in cinema, as indicated by Laura Marks's notions of 'haptic looking' and 'haptic imagery', and explored by Sobchack in terms of the abstraction of the initial images from *The Piano*. However, these instances are only fleeting moments of cinesthetic-inducing abstractions, because with cinema there usually is a linear narrative with specific meanings that takes precedence.

The images of the music video do not necessarily carry any such specific meaning nor is it dependant on the narrative path. Although linking sound and image signifiers can garner some form of meaning, the sensory experience of the music video is not presupposed on an exposition of this meaning. In reality, analogous to the visual flow of the image projection systems, the direct experience of the music video is possible *because* of the lack of any direct meaningful understanding of its audiovisual elements. It is the sensual flow of sound and images, and the way these elements bond as an outcome of the syncretic, formal, and cinesthetic links that make the music-video a direct experience.

The beginning is the end is the beginning

The connections between sound and image in the music video exist on the textual and the formal level. The video generates a visual extension or concretisation of the music. This extension translates both the form and flow of the audio, meaning its sound progression and instrumental form, as well as the textual 'narrative' of the song. Of course, most songs do not present a directly linear narrative, it is rather a narrative that influence the listeners' emotive and sensory responses to the music and lyrics. This is because the music does not work through mimesis but through a poetic and emotive translation of the sensory responses to the world, as indicated by Howells (2003, 35). Thus, the music creates sensations, feelings and moods in the listeners, rather than a clearly structured and linear narrative. These sensations can then suggest a narrative, but this narrative will be *created* by the viewer, using their emotional responses triggered by the poetics of the music. This creation of an emotive narrative is how the music-video makes sense to the viewing listener. But more importantly, this creation is what forms the basis for the visualisation process that turns a song into music-video.

In order to visualise the song, the music-video creator would first catalogue the sensory and emotive responses the music creates in them through their embodied sensorium. This catalogue becomes the point of departure for the visual conceptualisation of the audio feelings. But in order to locate or generate the proper shots that would translate the sensations into visuals, one does not bluntly represent the feelings visually, as in loneliness equals lone person standing on a pier looking woefully into the sunset. Rather, the sensations generated by the audio will be conceptualised in such a way that the imagery becomes subtly suggestive of the subject matter, as it does in poetry. It would thereby trigger similar emotions as the song through a visually poetic input that works in the same non-mimetic manner as the song. The fully conceptualised shots will then generate similar emotive embodied responses from the viewers' co-construction of the affective 'narrative' as the audio does. Thus, the video becomes a visual extension of the song's moods and feelings. When these sensually conceptualised shots are combined into a visual form and flow through cinesthetic montage, the visual extensions of the audio are linked directly to the form and flow of the song. The montage process therefore assembles the direct

audiovisual connections, *and* enables the generation of a new level of cinesthetic combinations of music and video.

So, at one level, the shots are conceptualised through the emotive responses to the song and produced to transfer these emotions into visuals. This sensual conceptualisation process would use the audio breakdown of the song as its base structure, as indicated by Vernallis. The song is broken down into sections (intro, verse, chorus, etc.), which are subsequently further broken down into the working musical parts generated by the instruments, which make up each section. Each sensually conceptualised shot will then echo, both on the shot's macro- and micro-level, the audio elements, cinesthetically expanding and visualising the form and flow already established by the music. On another level the shots, and the macro- and micro-structural audiovisual connections inherent within them, are attached to the song by means of the montage process. Through the cinesthetic montage procedure the 'correct' image, or sequence of images, is found for its correlating sound or flow of sounds, and welded together. In this way, direct connections between the form and flow of the music and the form and flow of the video are created. This generates rhythmic reverberations that oscillate from the shot or the elements within it, to the audio and its inherent elements and back again. Thus, it is the cinesthetics of the montage process that connects all the sensory and sensual audiovisual elements, and enables the cinesthetic montage to generate an audiovisual form and flow that affects and is perceived by the entire embodied sensorium, in ways that are beyond simple hearing and seeing, through the cross-modal expressions of the music-video.

Through the detailed conceptualisation, construction and montage of the visual elements according to the form and flow of the music, the music-video generates a stream of audio-visual movement. This stream cinesthetically reverberates within the viewer/listener and rather than just impart various images for them to look at and various sounds to listen to, it *surrounds* them as they experience the aural-visuality of the music-video with their entire body. Through this embodied sensorium and its cinesthetic sensory system and process, the world and the moving image is perceived and expressed. Since the aural-visuality of the music video is conceptualised through the body, and the components are generated and connected through the use of an

embodied apparatus, the embodied perception of the viewers will be familiar with the structure of the components and its process of audio-visual transference.

The 'lack' of a cohesive narrative and distinctly clear meaning, ensures a polyphonic and flat structure that accentuates the cinesthetic sound-to-image connections with their constantly shifting audio-visual synchresis, rather than the narrative developments. Since the initial component of the synchresis is continuously shifting, the experience is equally in constant motion. It never settles or comes to an end, but evolves as it swells and recedes, coming to the fore only to be transformed and replaced. Thus, just as it ends it begins again. In summary, the cinesthetic apparatus of the video makers enables the merging of components originating from embodied experience through a careful process that takes into account and correlates the developing visual and sonic lines of the music and the moving images. It is this process of matching the lines that exist between the audio-visual components, lines generated by detailed conceptualisation and careful montage, which enables the music-video to be expressed and consequently perceived cinesthetically.

2.5 Conclusion

In this chapter I have argued that the cinesthetic audiovisual montage of sound and image instituted by the music video enables an aural-visibility in which the sound and the vision transfer *onto* and *into* each other. To demonstrate this contention I have shown how the embodied cinematic apparatus facilitates an image creation and montage process in which the lines traced out by the sound are translated and echoed by the flow of images. This music-image synthesis can be produced through a variety of conceptual techniques. Common to all is the way in which the music is used as the starting point for the conceptualisation of the sound-image synchronisation. Thus, the images follow the music, echoing and visualising the sound, and through montage are linked in such a way that they flow together and interweave. Through a meticulous cinesthetic montage of audio and moving images that correlates the audio-visual lines to each other, a flow of sound and images is created that expresses a direct experience to the listening viewers. The music-video is perceived as one *merged* form and flow of cross-modal expressions. As a result, the direct experience of the music-video is co-constructed by the viewers, combining mounted images with the lines of audio, generating an audio-visual form and flow in which the image becomes auditory and the sound is visualised. Thus, it is the union of the senses, perceiving the combined sensual flow of expressions and audio-visual transference enabled by the cinesthetic cinematic apparatus and the montage process, which facilitates the transposition of sound into image, and image onto sound.

In the next chapter I will put the concepts discussed thus far to use through an in-depth analysis of Chris Cunningham's music video *Only You*. The analysis will apply the concepts put forward by Carol Vernallis and Vivian Sobchack in order to demonstrate how the music-video's techniques work through cinesthetic montage to create a cross-modal flow of audiovisual moving images.

Chapter Three – Analysis of Chris Cunningham’s *Only You*

“All of a sudden it hit me – if there was such a thing as composing music, there could also be such a thing as composing motion. After all, there are melodic figures, why can’t there be figures of motion?”
- Len Lye

3.1 Introduction

One of the more interesting music video creators of the past decade is Chris Cunningham. Working with major musical figures such as Björk, Madonna, and Portishead, as well as underground avant-garde acts such as Aphex Twin, Cunningham has developed a recognisable style. The vision that pervades his work is personal, enabled by the fact that he chooses only to work with artists to which he feels a strong, personal connection. Combined with his background in sculpture and robotics, this quality makes Cunningham’s audio-visual artworks dark, twisted and very tactile. The starting point for his conceptualisation process is usually a memory or dream, and his videos often centre on or around the human body, epitomising the phenomenological credo that knowledge can only be found in the direct and embodied experience of the lived-world. This makes the work of Chris Cunningham incomparable material for the critical examination and analysis of the cinesthetic montage of music-video.

This final chapter comprises an in-depth analysis of Chris Cunningham’s video for the Portishead song *Only You*. The analysis seeks to demonstrate how the music-video cinesthetically makes the music into image and the image into music, as well as showing how the music video concretises the notion of aural-visuality through the application of cinesthetic montage. Firstly, I provide a brief synopsis and description of Cunningham’s video. This part of the chapter also includes a sectional breakdown of both the video and the song. As shown in Chapter Two, the song is the most important element of the music video and, therefore, all images and ideas flow from, are connected by, and ultimately return to, the song. The sectional breakdown dissects the song into segments based on an identification of the musical elements contained within them, the repetition of these elements, and the way in which each segment

connects to other segments. By segmenting the song in this manner the similarities, repetitions and differences between each section can be identified. This also provides access to easily distinguishable musical sections that function as the fundamental unit of analysis, rather than the shots, as in narrative cinema analysis. The musical section is used in order to place emphasis upon the varied repetition of materials rather than the linear development of plot. The framework indicated by the musical breakdown therefore functions as the backbone of the analysis, as suggested by Carol Vernallis (2004, 209). In the second part of this chapter, I analyse the video according to the conceptual techniques listed and examined in Chapter Two, demonstrating the application of cinesthetic montage. This part comprises the most significant element of the analysis, as it illustrates the process by which music is translated into image, the image is made to resound, and, finally, the cinesthetic audiovision arises. In the final part of the chapter, I provide a close chronological reading of the video, expanding upon the different elements formulated in the main section of the analysis.

Both Vernallis (2004) and Goodwin (1992) have claimed that the song lyric can be seen as one of many elements used to understand the music video's multimedia construction and its possible meanings. The connection between this meaning and the visual element, as well as the augmenting effect the lyric can have on the understanding of the music video image, can provide the viewer with clues on how to interpret the symbolic significance of the images. This sometimes also creates juxtaposition between what is being said and what is being shown. Despite the emphasis on the lyric as a central element of the song and the notion that some words can directly inform the content of the image, the lyric itself does not have any implication for the cinesthetic and audiovisual relationship between the music and the moving image. The audio of the song is a musical description of the lyric, capturing the mood and sentiment of the words. In addition, the video is a visual description of the mood described by the sound, and consequently the cinesthetic relationship between sound and image comes to the forefront. Given that it is the cinesthetic relationship between audio and moving image that is the focus of this analysis, the lyric element is not considered to any great extent. Thus, the analysis focuses on the audio-visual forms and flows, and the way in which these are connected through cinesthetic montage.

The video

In a dark and dirty back alley a young boy floats, as if underwater, into frame. His features are starkly lit by neon, and his movements are slow and undulating as he glides through the dreamlike space (See fig. 3.1). As he hits the ground he is surprised that he can walk. Taking a few tentative steps, he notices a man looking crossly down at him from the window of a neon-blue lit room in one of the surrounding buildings. A woman is also floating in the same alleyway. Oblivious to the boy, she sings her mournful song, watched over by more lone men in neon-blue rooms. But the men seem not to see her, as their attention is focused on the boy, almost controlling the ramped²⁵ and meaningless motions he performs. Seeing the boy, the woman smiles and reaches her hand out to him. Looking scared, the boy reluctantly reaches out and takes her hand. But the hands are immediately broken apart by the angry stares of the men in the surrounding windows. The boy starts to flip around and around in perpetual motion. As the music fades, he drifts back flipping and is swallowed by the dark shadows of the alleyway.²⁶



Fig. 3.1 – The boy floats into frame, surrounded by the muted blue/green colour space of the submerged back alley.

The video for Portishead's *Only You* is set in a dimly lit back alley. The dirty ambience and muted greens and blues of the image are reflected in the muffled, liquid tones of the song. The young boy has been filmed underwater and superimposed onto the back alley scene, but because of the flawless colour grading and lighting the two images blend seamlessly, making the boy float while on land. The boy's movements are ramped in time with the scratching of the samples used in the song. As the lyrics of the song begin we are introduced to a woman. She is Portishead's singer, and exists

²⁵ Ramping is an editing technique that involves manipulating the speed of a moving image, first speeding up, and then slowing down (or vice versa) the motion of a continuous movement making it jerk fluidly.

²⁶ See Appendix A for a detailed synopsis.

in the same space as the boy, but they are somehow separated. There are also several men, standing in neon blue lit windows in buildings surrounding the back alley. The video crosscuts between these three main elements, establishing a relationship between them at certain points and, at other points, isolating the elements. The slow motion, underwater movements of the boy and the woman, combined with an underwater colour spectrum dominated by greens and blues and a gritty texture to the image, reflect and echo the languid melancholy of the fluid audio.

Using the sectional breakdown as the foundation, the next section of the analysis will outline the different audio-visual elements that make up the *Only You* video. These elements will then be used to illustrate the conceptual techniques laid out by Vernallis, and demonstrate the process of cinesthetic montage that enables the cross-modal experience of the music video. The concepts that are utilised in the next section of the analysis were put forward by Carol Vernallis, and examined in Chapter 2. They will be explored in individual sections following a logical progression from general to specific. It is important to point out that although each of the conceptual techniques will be analysed separately, they are unified elements of an audiovisual ecology. This ecology warrants that the reverberations of both image and music will point to each other and be connected in a system that goes beyond simple image-to-sound synchronisation.

3.2 The Music-Video: Audio-Form and Visual-Flow

The audio-visual elements that make up *Only You* are mounted together according to the lines inherent within the sound and image elements. These lines are interwoven through careful montage, harmonising the sound to the image and the image to the sound. Thus, the audio-visual elements are coordinated to each other according to their innate qualities, and in such a way that the qualities of one element are transferred onto and into the other element through a cooperative, reversible and dynamic system. The first and overarching concept of this system that will be considered is the form and flow.

Form and flow

The large-scale formal design of *Only You*, with its back alley and surrounding industrial buildings, fluid underwater imagery and blue/green colour space are matched by the song's musical sections. According to Vernallis, music videos often outline a large-scale formal design that reflects the large-scale musical structures and the overall mood inherent in the music (2004, 214). The flow of the song, its timing, key musical moments, and audio loops are closely linked to the form and flow of the images. The slow-motion underwater movements match the slow drift of the audio. The ramped movements are linked with the scratching audio samples, and the forward momentum of the song is linked with the progression of the visuals from constrained images to moments of expanded visual geography, and ultimately with the conclusion of the song as the visual geography of the introduction is replicated in the last musical section. In this way, large-scale lines are drawn between the montage of the audio elements and the montage of the visual images, considering the overtone aspects of all the elements and mounting them together accordingly. This produces an aural-visibility in which the sense of movement in the music and the images, and the atmosphere inherent within them, interweaves the connecting lines between them.

The overtone montage, together with the additional harmonising lines created by the self-similarity between the video's elements, creates a sense of continuity in the otherwise non-narrative progress of the video. This sense of continuity is most easily understood by looking at the sectional divide of the song. *Only You* is divided into

nine main sections: three instrumental parts (including intro and solo/coda), two verses, three choruses and a bridge.²⁷ The verse consists of a montage of the boy and the woman performing repetitive actions and movements, cutting between them. The first instrumental section and the first chorus introduce the men in the windows, expanding the kinosphere,²⁸ as well as introducing new camera movements. In the second instrumental section, new actions and movements are introduced. The bridge further expands the kinosphere with the introduction of a new setting for the woman, and a new geographic relationship between the figures.



Fig. 3.2 – Various repeated CU's and MCU's of the boy and the woman.

This structure is similar to the one Vernallis puts forward as the regular large-scale structure for music videos (2004, 214). She claims that the verse will usually consist of constrained images using the CU or the Medium Close-Up (MCU) of the star combined with similar images of the other main figures in the clip. In the chorus, the kinosphere will expand to reveal more of the video's geography, while the bridge functions as a link between the verse and the chorus. Most of these aspects are also present in *Only You*, although the focus rests as much on the boy as on the star. The verses are made of constrained images in terms of kinosphere, consisting of a montage of CUs and MCUs of the boy and the woman in a repeated space (see fig. 3.2). The choruses expand the geography with the introduction of the men in the windows, filling the frame with more figures and a larger kinosphere in relation to the thickening of the musical arrangement and the lines of the swelling sound of the chorus. But there are important variations. While the bridge certainly links the first two thirds of the song to the last part in terms of the music, it is actually in this section

²⁷ See appendix C for a complete key of the songs main shapes and phrases, and a detailed audio breakdown of the song, including the recurrence rate of audio shapes/phrases.

²⁸ Taking the concept from dance theorist Rudolf von Laban, Vernallis defines the kinosphere as the centre of gravity and implied larger space that the figures and the camera move around and within. See Vernallis "Experiencing the Music Video" p. 220.

that the most radical expansion of the visual geography happens, in tight connection with the foremost expansion of the musical space.



Fig. 3.3 – The woman and the boy meet, join hands but are instantly torn apart.

When moving from the second instrumental section to the bridge, the boy releases a dove as part of a magic trick performed during the previous section. This dove flies up, in time with the vibraphone flutter, past the towering industrial building and up past the full moon in the darkened sky. This moment introduces the larger geography in which the woman is placed for the entire bridge section. She is now no longer in the back alley, but in some sort of square. She has the large building and the man in the window behind her, or perhaps this is the place she has always been in. Nevertheless, this space is only utilised for the bridge section of the video. In the second chorus the boy and the woman find themselves in the same space for the first time. Although this possibility is hinted at throughout the previous musical sections, it is only here that these two main figures definitely exist together in time and space. As the woman reaches her hand out, the boy does the same, and in the next shot we see their arms going towards each other, and their hands finally connecting (see fig. 3.3). But the connection is instantly broken at the start of the last part of the song, the third instrumental section. Again the kinosphere is expanded, more of the building is revealed along with the introduction of new floors of windows containing men. This line of movement is also aptly connected to the upward lines produced by the musical high point of the song.

The most prominent aspect of *Only You's* musical form and flow is the liquid and very organic soundscape constructed by Portishead's distinctive electronic music. The band utilises authentic instruments combined with samples from old jazz records, record scratching and a variety of other noise samples (crackling from old record players, white noise, snippets from radio/TV, etc.) as well as singer Beth Gibbons's distinctive voice. This montage of sound elements creates a very cinematic and

textured soundscape that instantly evoke images of smoke-filled jazz clubs, melancholy and lost love. The sense of nostalgia generated by the song is juxtaposed with its post-modern montage of sound elements and samples, producing an interesting amalgamation of old and new. The song produces a very fluid and muted overall contour, with a distinct nostalgic quality that connects seamlessly to the submarine visual space created in the video. This overall sense of fluid form, nostalgia, and longing found in the overall contour of the song is further supported by the choice of both musical and visual shapes used to construct the song and, by extension, the video image.

According to Vernallis, the visual elements of the music video can work with the music by taking on the phenomenological qualities of the sound. She continues by stating: “[...] these images, like sound, come to the fore and fade away, ‘stream’ around and surround us, and even reverberate within us, and mimic timbral qualities.” (2004, p. 177). This echoes the idea of how the audio functions as a flow in which the ‘seen’ bathes, proposed by Michel Chion, but with one significant difference. In the music video, it is the moving image body that creates a visual flow within which the ‘heard’ bathes. Thus, a reciprocal, intimate and embodied connection between the audio and the visuals is created by mounting images with certain harmonic lines that correspond to lines found in the audio segments, onto these audio segments. Thus, the moving-image body of this music video traces out a large-scale form that emulates the formal and harmonic qualities of the audio. The sombre sonic mood and dark atmospheric reverberations are reflected in the dark alleyway set-up, the muted and soft blue/green colour palette, and the slow, undulating movements reflect the fluid, percolating progress of the audio (see fig. 3.1 and 3.5). But, the overtone relations between the musical lines and the moving image body also exist on a smaller, more detailed scale.

Contours and shapes

Every song contains basic musical shapes that are repeated and mounted together to make up the song. In the case of Portishead’s *Only You*, there are five key shapes: the bass line and individual bass ‘Bops’ (B), the vibraphone ‘Bing’ and flutter (VB, VBF), the scratching of samples (S), the organ (O), and the sampled horns (H). These

elements are sometimes combined into key phrases, and sometimes appear on their own. There are three main phrases:

- The organ/bass/vibraphone-bing (O/B/VB),
- The organ/bass-bass/vibraphone-bing-bing (O/BB/VBB)
- Organ/scratch/vibraphone-flutter (O/S/VBF).

There is also a key scratching phrase that involves the scratching of a sample (SS) of the line “move it like that”²⁹ as either single or multiple scratches and often combined with breaks in the audio. Likewise there are visual shapes, or lines, that make up the video. There is the ramping of movements, slow-motion underwater movement including undulating hair, shots of specific body parts or other on-screen objects, camera zooms and tracks, cranes, tilts and pans, the rigid body of the boy, and the boy somersaulting. Eisenstein called these visual segments compositional units, and claimed that the organisation of such units create an “expressive whole” that was not a copy but an interpretation of the object itself. The montage of these compositional units, or shapes, would then convey or denote the perception of this object (1998a, 90). Similar to Sobchack’s notion of the expression and perception of the image, the montage of compositional units generate a reversible cinesthetic flow of sound and image in which each element is interconnected and communicates more than the sum total of its parts. A sum total that can be perceived and interpreted by the embodied sensorium of the listening viewer because the expression of the audiovisual flow is cinesthetically conceptualised, cinesthetically produced and ultimately, cinesthetically mounted.

As we have seen, the form and flow of the moving images that makes up a music video are directly connected to the form and flow of the audio. But the overtone lines that cinesthetically connects sound and image, are also evident on a more specific and detailed level. A specific musical shape will be directly connected to a specific visual shape, enhancing and extending the emotional effect of the audio shape. These visual shapes can be a colour, a camera movement or any other visual shape, that are interweaved into and onto the music through cinesthetic montage. In other words, the visual and aural shapes’ reflection and echoing of each other is what creates the connection. The process of reflecting and echoing the audiovisual elements is initiated

²⁹ See Appendix B for complete lyrics.

by the conceptualisation of the moving images. The echoes and reflections are then generated by the cinematic apparatus on the basis of the embodied conceptualisation process. The moving images are finally interwoven with the music from which it initially stems through a careful montage process. Vernallis describes musical contours as the visual shape of the sound or musical line (2004, 213).

And so, to be effective, the visual shapes must reproduce the contour of the musical shape. She further suggests that the shape of the contour, whether it is alleged as being jagged, narrow, or liquid, will closely affect the way the audience perceives the music when they listen to it. These 'visual' lines correlate directly to the emotional effect the song generates because the sound and the moving images closely echo and reflect each other. This interwoven association between audio and image is the most important element in the cinesthetic translation of music into moving image. The most interesting aspect of the musical and visual shapes is, however, the way the lines communicated by the audio-visual elements are connected and interwoven into and onto each other as motifs and phrases.

Motifs and phrases

In the music video, certain visual motifs are directly linked up to a specific musical shape according to the overtone aspects that connect the audio to the images. These visual motifs are used as repeated building blocks of the video in the same way that musical shapes are used as the building blocks of the song. Again we can see how the overtone montage design, in which the collective tone or mood of every single element as well as the total mood created by these elements combined, is present in both the montage of audio segments, visual elements, and the totality of the music-video. The motif is a specific image that will repeatedly occur in a direct link with a specific musical shape. The specific image can therefore be said to function as a symbol for the sound, as pointed out by Vernallis. More importantly however, are the ways in which a musical shape will be linked to images, usually including some sort of movement, in order to visually emulate the sound shape. These moments are referred to by Vernallis as phrases, and do not necessarily contain a link between a specific sound to a specific image (2004, 216). In the phrase, the image will visualise the musical shape, rather than symbolise it.

This visual representation is closely connected to the contour of the sound. Unlike the motifs, that are specific for each video, phrases are more general and are often specified by culturally informed dispositions, indicated by Laura Marks' notion of cinesthetic memories. Rising sound contours are often linked to a ground-to-sky oriented movement, while sharp, contained sounds are often associated with a distinct, short and sharp motion or moment, such as hitting something or releasing something. Extended or elongated sounds are often connected to continuous and progressively extensive movement, either by a figure, an object, or the camera. These connections are perceived by the embodied sensorium of the viewer. By enabling the sensorium to dynamically follow the audio-visual lines traced out by the embodied application of the cinematic apparatus, the music and image cinesthetically merge to create a sensory experience beyond simple hearing and seeing. In this aural-visibility the sound is bonded onto and into the image, generating audio-visual motifs and phrases that inform and enhance the form and flow of the video on multiple levels.



Fig. 3.4 – The men in the windows are always phrased with the horn sample.

In *Only You* there are several recurring motifs, but there is only two that come to the forefront as a central ingredient in the video. Firstly, there is the sound of the vibraphone either as a flutter or a straight Bing. The VBF is connected to the waving of hands, both in images of the boy and the woman, and the VB occurs both times the boy's feet hit the street. Then, there is the connection between the horn sample/VBF and the men standing in the neon blue windows that also comes to the forefront as a significant motif (see fig. 3.4). An extended horn sample utilised in the bridge is also connected to the same motif, and an extended horn sample foreshadows the musical high point of the song in the third instrumental section. The motif of the man in the window is utilised in connection with the VBF at the start of Chorus' One and Two, and the horn sample at the end of the same choruses. The camera movements connected to this motif are reversed within in each chorus section.



Fig. 3.5 – The boy's movement are often ramped in time to the scratching of samples.

Using the VBF as its phrase the camera moves towards the man in the window at the start of Chorus One, using a pan left/tilt up/zoom in combination. This is the same move that is used in the opening instrumental section. At the end of the chorus section the camera pans right/tilts down/zooms out, away from the man in the window, but this time using the horns as the phrase. These movements are repeated for Chorus Two, except that in this instance the camera-move combination does not include a zoom. The combination of phrasing with motif is used extensively in this video, but phrasing is also used separately as an important element of the video image. The most prominent use of phrasing is the linking of the scratching with the ramping of movements (see fig. 3.5). This element is introduced, along with the key motif of the men in the window, in the opening instrumental section and is only repeated in the second instrumental section. In these sections the movements of the boy will be ramped either in time (slow or fast motion) or in direction (backwards or forwards) in direct connection to the scratching. Several of the other key musical shapes are used as phrases for either figure movement; in the form of pure movement or ramping; camera movements; as mentioned above; and lastly to smooth the transition or cut from one visual section to another. Cinesthetic audio-visual connections can also be located on the resounding of the sound space in the surface of the images.

Timbre and texture

Timbre and texture are the last elements I will consider in this analysis. Both these concepts can be applied equally to the audio and visual elements of the music video. A song, or some musical shape within that song, can have a particular timbre or texture, often providing a sense of tactility. The timbre and texture subtly colours the listener's understanding of the musical space, much like a vague scent would colour a space in the lived-world. Vernallis claims that the use of texture helps to elicit a

visceral response from the audience (2004, p. 127). The use of visual textures in the music video image can also help create continuity between scenes that are removed in time and space. By using this temporal indeterminacy of texture, additional space is created for the audio to operate within. The texture of the image, like the timbre of the song, provides the moving image with a haptic surface creating an image with a sensible exterior that resounds with the musical timbre of the aural space.

The timbre and texture of a song's sound space can provide very strong visual associations, elicited from the visceral responses mentioned above. In *Only You* the most prominent timbre is the muted, underwater quality of the song's sound. This is directly related to the underwater imagery and the muted blue/green colour space that makes up the video, as we have seen in the examination of its form and flow. The song itself moves in slow motion, gently oozing forward, all the musical shapes combining to make a fluid sound-space that is old, nostalgic, and strangely familiar, like an odd dream. Even the film noise crackling throughout the whole clip supports this nostalgic, phantasmagorical feel of the song. This quality of the audio's texture is utilised in full by Cunningham, who creates a music video image that like a dream both in terms of movement and colour, but also in terms of the dramatic elements and 'narrative' the video represent. Again, we can see how the music is used as the conceptual point of departure for the visual flow. It functions as the initial element of the montage, but it is the connection of the overtone lines of *both* the image and the sound that enables the embodied image to echo the sound in such a way that the two elements flow into and onto each other. The texture and timbre of the audio-visual image are central elements in making the music video a viscerally embodied experience. It draws the attention to elements within the song, as Vernallis states (2004, 127). Because of this it floods our senses with cinesthetic imagery that elicits a visceral response, making the experience internally felt, rather than externally understood.

3.3 Cinesthetic Music-Video Montage: “Only You” Dissected

A variety of techniques have been deployed, invented and tested to generate a synaesthetic experience, as pointed out in Chapter Two. All these techniques have found their way into the ecology of the music-video language implying that one of the music-video’s main objectives is the creation of a cinesthetic audio-visual flow that makes contact with the viewer/listener on a level far beyond simple seeing and hearing. The shape, timbre and phrases of the sound generate harmonising lines that are resounded by the lines of the image’s form, motifs and texture. Thus, through an application of the embodied cinematic apparatus to the sound, a cinesthetic image is generated that infers the mood and feel of the music and visualises its progression. As we have seen in the analysis, the audio and image shapes and elements assume each of the other’s form and flow, resounding the progression and timbre of the music, through the use of movement, colour, montage and camera movement. In this way all the different audio-visual elements come together, relating, interacting and connecting with each other, to make up the music video’s cinesthetic structure. The audio-visual elements used to build *Only You* and their construction, composition and assembly, constitute a form of film language. According to Sobchack, cinema uses “modes of embodied existence”, meaning seeing, hearing, and physical and reflected movement, as the vehicle or substance of its language. Furthermore, it deploys “structures of direct experience” as the basis for the structure of this language (1992, 5).

The same is true for *Only You*. As shown in Chapter Two, the cinematic apparatus can generate a cinesthetic moving image body. In *Only You*, this embodied moving image is produced through a direct connection with the music it visualises. All the levels of the image flow from the music, connect with the music in a mutually reversible manner, and ultimately return it, a structure that is true and essential for the process of expression as well as that of perception. In this way, the sound and the image become interconnected and mutually dependent sensory expressions that are expressed as *one* cinesthetic flow. This flow is experienced by the viewers/listeners through their embodied sensory system and triggers sensory memories as a consequence of the audio-visual language’s source in embodied existence and direct experience. The following close chronological reading of *Only You*’s nine sections provides a more

detailed illustration of how the sound and image elements are cinesthetically mounted together according to their inherent lines. The close reading of specific sections will also reflect the flow from general features to particular moments in the video, as an attempt to try to reproduce the changing sense of engagement that the video creates with multiple viewings.

Instrumental One

The video opens on an empty, dark back alley. Phrased on the O/VB, a young boy floats into frame (see fig. 3.1). The underwater movements of the boy are reflected by the large-scale tone of the song, making the audio and image float into and become each other. Phrased with the VBF, the boy's head is ramped to establish an eyeline upwards. The horn sample provides a powerful thrust as we cut to a man in a window, using a pan left/tilt up/zoom in camera movement, and establishes the ongoing connection between the man and the horn sample motif. The scratching commences with the O/B/VB and the boy's movements are ramped with the scratching, establishing another audio-image phrase. On the O/BB/VBB, the boy's feet hit the street, and he seems surprised he can actually walk. Another section of scratching-ramped movement phrasing then commences. Each break between the scratches is emphasised with a ramped movement, making the viewer/listener float along as they watch and listen. Again the cinesthetic relationship between the audio and the image is underlined, making them inseparable elements rather than simply linked elements. On the O/BB/VBB, the video changes sections, using this audio element as a link for continuity between two sections.



Fig. 3.6 – The woman's hair and movements undulate, echoing the music.

Verse one

Phrased with the VBB, the woman floats into frame and starts singing. Again the underwater movement in the image is reflected by the overall structure and sound quality of the song. This section is made up of cuts between the woman and the boy, often phrased on either a B or a VB. A relationship between the two is established using eyeline shots, putting them both within the same kinosphere. Links between the figures are also established by connecting the lyrics and the images. On the word "illusion" we cut to a low angle shot of the boy floating in a proud stance with a determined smile. This tentative link might indicate some sort of dream relationship, further underlining the dreamlike quality of the audio. This section is reasonably straightforward in terms of cinesthetic qualities. It is more of an establishing section, introducing the main figures and their relationship. But there are phrases used here as well, to underscore the underwater cinesthetic union between sound and image. Throughout the entire section and for most of the video, the stretched organ waver and the VBF are resounded by the slow undulation of the woman's hair and movements (see fig. 3.6). This verse and the following chorus are linked by the organ stretching between the sections.

Chorus one

On the opening O/S/VBF, the boy is again ramped into frame, and an eyeline established. On the VBF we cut to a light flash, cinesthetically connecting the flash and the VBF, and pan left/tilt up/long zoom in to the man in the window, again establishing a relationship between the man and the boy that expands the kinosphere. This phrasing breaks the motif of the man and the horn sample, giving it additional weight. Again the image harmonises with, and according to, the sound, as the upward movement of the camera follows the upward line generated by the VBF. The video

cuts to the woman as she finishes singing, and then cut to the boy using the O/S/VBF to ramp his movements. Here the boy floats upwards with the VBF on the word “you”, finishing up in the proud stance seen in the previous section, timed on the words “turn my wooden heart”. Again we can see the connection between the two main figures as a result of the link between lyrics and, more importantly, image. As the horn sample kicks in again on the eyeline of the boy looking upwards, another light flash appears and the left/tilt up/long zoom in camera move is repeated, re-establishing the previously instituted motif and cinesthetic connection (see fig. 3.7). The jump to the next section is again linked by the O/B/VB, cutting from the man to the woman on the VB.



Fig. 3.7 – Light flash and the camera moves.

Verse two

This section is similar in its visual structure to the first verse, except that the timing of the shots is different. As we cut to the woman, we stay on her for most of the verse, which also contains a longer sentence than the previous verse, with the first O/BB/VB only cutting to the boy once. In that cut, the boy’s upward swimming motion is ramped with the O/B/VB synchronically timed with the end of the word “selfishly”. This again relates the two figures through the lyric, and cinesthetic audio-visuality. At the end of the verse the boy looks screen left, and on the VBF we again float into a camera movement that traces a line upwards to the man in the window. This move, as seen earlier, ties in with the O/S/VBF that also connects the verse to the following chorus.

Chorus two

In this section the word "you" is again connected to the proud stance of the boy. The singer is almost certainly singing about him. We then cut to a CU of the man in the window looking grimly at the boy, and us, and back to the boy who now looks despondent. As the woman turns away on the words "my wooden heart", the horn sample kicks in echoed by the camera panning right/tilting down/zooming out from the man in the window and into the light flash. As the last of the camera movement is ramped with the scratching sample of the next section we cut to another CU of the boy looking despondently upwards, screen right, towards the man in the window. Throughout the entire section the undulating hair and general movement of both the boy and the woman is phrased with the O and the VBF.

Instrumental two

As this section begins the boy's head is ramped downwards with the scratching, resounding the movement of the camera phrased with the VBF at the end of the previous verse. Then follows a long montage of crosscutting between the boy and woman, the boy's ramped motions phrased with the scratching, while the motions of the woman are more freely cinesthetically linked to the B or the O. The boy is being 'forced' to perform the motions by the man in the window, supported by the words of the scratch sample "move it like that". On the O/S/VBF we cut to the boy holding a handkerchief, and a complex set of ramped motions played backwards are phrased to the scratching of the "move it like that" sample (see fig. 3.8). Again the careful montage of the image and the scratch/break structure of the audio create a cinesthetic form. As the boy's other hand takes one end of the handkerchief, we cut to a CU of the boy smiling as the music cuts. The handkerchief is then pressed into his clenched fist as if in a magic trick and on the hit of a solo VB that rings out and reverberates, the boy opens his hands, releasing a dove that flies up towards the sky resounding the line of the reverberating VBF (see fig. 3.9). The VBF sound also links the chorus to the following bridge, smoothing out the leap from one section to the other. The bridge functions as a segue from the first two thirds of the song to the last part. It also indicates that the song is coming to an end.



Fig. 3.8 – The boy's movement are ramped to the scratching sample.



Fig. 3.9 – On the solo vibraphone Bing he releases a dove that flies up, phrased with reverberating vibraphone.

Bridge

Establishing an eyeline for the boy phrased with the VBF, the clip then cuts to the dove as it flies past a neon-blue lit window with a new man behind it. This camera movement is timed with the horn sample, including an emphasised double take using the same sample, continuing the horn/window-man motif established throughout the clip. The double take is used to phrase a cut from a wide shot to a CU of the man in the window. Beneath these images the woman has started singing again. Following the dove past the building and the moon, the music kicks in and we cut to a wide shot. The shot is dominated by the woman in CU filling the right side of the screen, and the building with the man in the window looming behind her filling the rest of the screen (see fig. 3.6). This radical expansion of kinosphere, directly related to the line of a thickening musical arrangement, is not in accordance with Vernallis' notion of large-scale form and flow. It is also the first time the video establishes a relationship between the woman and the man in the window, linking the opening and middle section of the song with the end. A new setting is introduced, and a new relationship between the figures of the video is established.

The actions of the boy also establish a more tangible relationship between the boy and the other figures. Is the magic trick to impress the woman, or is it to appease the man in the window? Both readings are credible. Either way, it shows that the boy is aware that he is not alone, and that he sees his actions as part of a cause and effect system. Phrasing is used extensively in this section, connecting both the VBF and the horn

sample with an upwards-rising line of both on-screen actions and the cinematic apparatus, something that will be embellished in the closing section of the song. As the bridge nears its end, we have a similar musical section found in the middle section of each verse. With an O/BB/VBB used to phrase a ramp movement of the boy's head, and the organ phrasing the undulating movement of the woman's hair, the last line of the bridge is sung, and we move into familiar territory again. The boy is somersaulting in the back alley. The woman is back there also. Again, the stretched reverberating shape of the organ line carries the transition between the two sections.

Chorus three

This extended chorus follows the structure of the other choruses, but the crosscutting between the boy and the woman is not as rapid and frantic as in the other choruses, and the image sections that are used do not change between cuts. The montage consists of the boy somersaulting slowly, and the woman singing with her head turned away from the camera. The cuts here are not done on any particular sound shape, as in the rest of the song, and there are no ramped movements, highlighting the slow and undulating flow of this section. Such an unhurried structure emphasises the languidly fluid large-scale structure of the song, bringing this element to the foreground and generating a sense of closure and return to the opening segment of the song and the video. At the second to last line of the chorus, we cut to a shot of the woman reaching her hand out as she smiles, and on the word "wooden", and the B, we cut to the woman's hand and the boy's hand moving towards each other. They connect on the word "heart", and we cut to the boy looking scared, timed with some scratches. A larger and extended horn sample then kicks in and we are moved into the final section of the song.

Coda

On the first three bam-bam parts of the extended horn sample we see the boy look scared and then despondently turn his head towards the camera. As a more forceful and rapid bambambam section of the horn sample kicks in there is a quick light flash (see fig. 3.10). The camera then quick-pans 180 degrees, and the last and final BAAAMMM is used to phrase a camera track towards a new man in the window,

standing with his arms ominously raised. Then a backwards organ line mixed with a rising string line dominates the sound-scape and phrases an upwards crane shot past



Fig.3.10 – The camera follows the music up past the men in the windows.



Fig.3.11 – The camera continues to move, resonating the upwards sound.

several floors with men in windows watching the boy, and us (see fig. 3.11). As the crane reaches its end, a tut-tut from the organ is used as we cut to a shot of the hand of the boy and the hand of the woman drifting apart. They then drift away from each other, as the organ starts its solo. For the rest of the video we crosscut between the woman and the boy, the woman's slowly undulating hair phrased with the organ as she twists and turns, and the boy swimming desperately. Another horn sample resounds with a zoom-in towards a man in a window, this time phrased with the rising line of the organ solo. On the O/B/VB we cut back to a slow outwards tracking shot of the boy somersaulting around and around. The song fades out as the outward tracking shot fades to black, each fading line and outward motion mounted together, generating the final cinesthetic connection that melts the audio into and onto the moving image as the image fades to black.

3.4 Conclusion

Only You is a sad lament for the loss of innocent young love and for a dream that could have been. The young boy symbolises the innocent dream, the woman his desire and the men in the window are the patriarchal 'big brothers', watching and ultimately crushing the boy's fantasy, as they force him into perpetual limbo. Both the lyrics and the music establish the boy as the subject the woman is singing about, while the main motifs and audio samples establish the men as the controlling society and the woman as the one that challenges this power relation, without success. The way several of the men look straight down the barrel of the camera directly at us creates a feeling that we are positioned from the boy's perspective as if we were the ones whose dream is being crushed. The nostalgic lament is for us, and for the way we have lost the innocence and dreams of childhood. In this way the video unites its large-scale audio structure with the image, creating a cinesthetic aural-visibility through the use of a fluid underwater sound scape that is visually resounded by muted blue/green colours, undulating images and a fluid use of the cinematic apparatus. All the mirroring elements are cinesthetically mounted together in order to generate direct connections between the sound and the visuals, in which the sound and the images transpose onto and into each other.

The embodied cinematic apparatus facilitates the creation of music-video in which lines traced out by its audio-visual segments, interconnected by cinesthetic montage, echo and harmonise with each other. These harmonic reverberations that connect the elements of the audiovisual flow can be found on both the macro-scale of form and flow and on the micro-scale in which specific musical moments are directly resounded by specific visual instances. The harmonic reverberations are generated by the visual conceptualisation process as well as by the method through which the audio-visual segments are mounted together. The process of cinesthetic montage mounts image-to-image, image-to-sound, and sound-to-image according to the feel of each element, connecting the lines inherent in each element with comparable lines in other elements. In this way, an interwoven expression of synchronised video and music is generated through an embodied cinematic apparatus. Because of this inherent embodied structure the viewers perceive the audiovisual flow through their embodied

sensorium as one *united* flow in a co-constructive, reversible and synaesthetically analogous process.

By relating the images step-by-step to the music through an identical motion that lies at the base of the sound and image movement and the movement of each element's harmonic lines, an organic bond is generated between audio and images. This audiovisual ecology exists on each level of the music-video, and within, as well as between, each audio-visual element of which it is composed. Therefore, the music video generates an aural-visuality in which sound and image reverberate concurrently; as if the image were the music and the music were the image. This generates an audio-visual flow in which the perception of the complex visual line is sustained by the reverberation of sound, while simultaneously the audio is echoed in the image lines. The resulting synchresis is present as a result of a conceptualisation, production, and montage process that is as cinesthetic as, and cinesthetic because of, the embodied perception of the music-video. Thus, the cinesthetic montage of audio-visual elements according to their inherent lines generates an entwined audiovisual expression that elicits a visceral response from the viewer/listener's embodied sensorium as they perceive the interwoven flow of music-video. Therefore, it is the union of the senses, when combined through the expression of the embodied cinematic apparatus and the perception of the human sensorium that facilitates the transposition of sound into image, and image onto sound.

The next and final chapter comprises a critical production reflection of the entire process of creating this thesis' practical component from conceptualisation, through shooting and to putting together. Chapter Four therefore examines how the images of *Moment* were conceptualised using the music as a starting point, and how the music was used actively in the entire process.

Chapter Four - Making Music-Video: A critical reflection of the production process of *Moment*

4.1 Introduction

The creative component of this thesis consists of a 24-minute experimental music-video. The music-video visualises an experimental piece of spectral music composed by Edith Cowan University Master of Creative Arts student, Brett Mabury. The objective of the creative component was to apply the conceptual framework discussed in the previous chapters to the creation of an original music-video. This application illustrates the claims made about music-video in general and Chris Cunningham's *Only You* in particular. This chapter therefore comprises a critical reflection of the production process from conceptualisation, to shooting, and to the cinesthetic montage. Firstly, the chapter examines how images were conceptualised using the music as a starting point, as well as the fundamental point of reference throughout the entire process. The second part of the chapter illustrates in detail how the conceptual framework was applied to the cinesthetic montage process. This section also contains a detailed examination of three excerpts from the *Moment* music-video. Thus, this chapter explicates how the conceptual framework is applicable to not only the analysis of music videos but also the creation of music-video.

The practical element for my thesis is a 24-minute visual exploration of an experimental musical piece. My video piece is therefore not a music video, if music video implies a piece of video made to accompany and illustrate a song for advertising and commercial purposes. Rather, it is a music-video, a visual interpretation of a piece of music, conceptualised, produced and put together so that the flow of images would resound and echo the music. I was asked by composer Brett Mabury to participate in an exploration of an audiovisual synthesis that was part of his thesis program. Mabury's idea was to record the 18 Swan Bells of the Perth Bell Tower and use these recordings as the basis for his exploration of spectral music. My assignment was to interpret the finished piece and produce a visual echo. The initial idea was that the video artist would make a visual piece that Mabury would illustrate with sound, but I suggested that it would be more constructive to do it the other way.

In this way the flow of sounds he generated would serve as the emotional, sensory and conceptual starting point for the visual exploration. The following section is a statement made by Brett Mabury that outlines his approach to the project and my task as the visual artist, as well as giving some detailed information about the musical process.

Composer's statement

“As a musician I have long been interested in working with artists from other idioms. In film the composer will often respond to what they are seeing on the screen. However, in ‘Moment’ I was encouraged to do the opposite by Joachim Strand. In this case the footage captured was a response to the music heard, a process I was excited to explore. The genre I have been looking at as part of my Master of Arts research is called ‘Spectral Music’, coming out of France and Germany in the 1970’s. These composers developed a new approach to music by uniquely bringing into account the physical structure of sound. Computers have been used to analyse different sound structures: everything from a trombone playing a low note to cascading water. The composer in a variety of musical settings can then use the frequency information gathered. Essentially these composers were more interested in exploring different colours in sound than creating a melody.

In this production I decided to investigate and manipulate the sound structure of the Swan Bells, composing a piece for a virtual orchestra that reflected the discovered resonance of the bells. Each of the 18 bells became an opportunity to create a different musical colour. Researching spectral music has certainly challenged my way of musical thought and my approach to composition in this case. Sets of bells with unique peals have not only signalled time, but also the significant events held by time. Consequently, I decided to make the piece 24 minutes in length, reflective of a 24-hour day. The visual artist was made aware of this approach, and asked to think about the visual moments in their day that capture them, especially those that we can easily miss. Joachim Strand has created his own piece in which the editing process allowed the vision he captured to be connected rhythmically with the music.”

– Brett Mabury

4.2 Connecting Image and Music

Mabury's music was supplied to me in pieces of varying lengths and differing levels of completion. The idea was that I would use these pieces to develop and conceptualise ideas and fragments of video that would then be connected into a whole as the audio-piece was completed. But it was difficult for me to generate any ideas when I did not have a sense of the complete form and flow of the piece. Although I did produce a few ideas and sketches using the various pieces, the visual conceptualisation did not fully commence until I listened to the song as one complete and continuous piece. This was necessary in order for me to record my sensory and emotional responses to it as a whole. In this way, my visual concept was constituted as a sensory interpretation of the sound (in a way not dissimilar to that formulated by Eisenstein) and, as a result of which, it would, as Goodwin states, flow directly from the music (1992, 92).

The initial responses were then embellished through repeated listening to the song. And so, as with Vernallis's method, I needed to go from the complete piece and work myself downwards through each section. But because Mabury's song was particularly intricate I did not do a complete musical breakdown of the *Moment* piece, as I have done for *Only You* in my thesis. Rather, I kept the breakdown procedure, and the methodical approach it enabled, in mind when conceptualising the visual flow. The musical piece was too complex, with its multiple sections and numerous instruments, to generate a usable breakdown. However, the approach I used followed the same steps without having the ease of tracking the map that the breakdown provides. Furthermore, as detailed below, I also saw this 'freedom' as a possibility for further testing of the cinesthetic montage process.

My first reaction was not simply a progression of sound, but rather a sense of moving towards something, travelling towards an as yet unknown destination. The song had a slow, dreamlike, yet oddly unremitting, structure that underlined its progressive flow towards some place or some space. Various bells, almost like a signal bringing the listener back into reality, interrupted these dream journey sequences at regular intervals. It was as if the song indicated that while you were dreaming, time was still moving and the clock was still ticking. Likewise, my visual excursion did not become

a concrete voyage to a specific location, rather it was a sub-conscious, sensory and dream-like journey based on my sensory impressions of the audio. To follow this dreamlike audio progression, I came up with the idea of a journey into a pear and out of an ear, images that comprise the opening and ending sections respectively. Coming out of an ear at the end was also a cinesthetic pun: as Mabury's musical journey entered the ears of the audience, my own visual trip, a journey that was made up of a visual flow coming out of my ear, entered their eyes.



Fig. 4.1 - 'Moment' invitation and poster.

I had very specific designs for some of the sections, for instance the road driving and the leaves dripping, as well as the colour drops that expand into water for the bell strikes, while for other sections I only had sketchy notions of what I wanted. This might seem like an anomaly in terms of the structured process outlined in the main thesis section, but I wanted to explore how effective the cinesthetic montage process would be if the images had not been accurately conceptualised, planned and generated. Therefore, I spent the next four days shooting. The first two days were spent filming the previously planned sections. One day was spent filming

experimentations with colour and oil drips in water, which I planned to use as textures or fill between the main sections. The fourth day I put the song on my iPod and listened to it while I filmed whatever came to mind as a direct sensuous response. The entire interlaced tree section, as well as the zoom/in out of the tree, and the pink sky section was filmed during this day.³⁰ The freedom of experimenting with the camera in direct sensory response to the music was an interesting way to work, but only approximately 10 minutes of the two hours of footage shot on that day were used in the final piece.

The second stage was then to identify the usable visual sections. Some sections were of course already planned, and in these instances it was only a matter of selecting out the best takes, while other sections had to be chosen based on their sensory 'feel' in correlation to the song. This is where the practice of finding parallel audiovisual lines becomes a significant method. Some visual sections seemed to jump out as I was watching them, as if they responded to the music echoing in my memory. The visual sections would either present themselves to me as the main element of an entire section, or as a minor element that would trigger a process of fitting this element with other elements to make up a section. Of course, more work was needed not only to mount each visual section together and provide appropriate textures and colour settings for each element, but also to mount each visual element to others within each section.

³⁰ Three of these sections will be discussed in detail below, and can be found as separate chapters on the *Moment* DVD.

4.3 The Cinesthetic Montage Process

As soon as I had selected the all visual sections I was to use, I started enhancing and mounting them together using Final Cut Pro, After Effects and Motion. Again, the process was one of listening to the song and then cutting, augmenting, and transforming the visuals so that they would correlate directly to the sound. The key audio instance was of course the bells that provided instant moments of auditory contours that could be correlated to visual shapes. Then there was the contour of rippling piano sounds spread throughout the song, as well as the stretching of undulating sounds that replicated the sound spectrum of the bells. The bells became visually represented drops of coloured ink dissolved in water. The tinkling of piano was correlated to two specific visual shapes, the first one being the twinkling of sunlight seen through branches, the second, raindrops dripping from leaves or hitting a surface of water. The stretched undulations of sound were correlated to shapes that stretched with the sound. This was done using Motion's real-time effects recorder. Thus, while listening to the elongated undulations of sound I was stretching sections of the image in real-time and recording the process. These audiovisual correlations would all be created by listening to the song and then finding the appropriate piece of image that would visually echo the feel of the particular instance of the song.

In this way my process followed the same procedure as the one outlined in the thesis. Filters and other textures were added to make sure the overall form echoed the song, while careful montage and superimposition would generate micro-scale instances of cinesthetic audiovisual links. Both methods also used the song as their starting point. The feel of any given audio instance directly influenced my montage choices in a direct and instantaneous cinesthetic relationship, consequently generating cinesthetic links between the audio and the images that would echo and resound in the finished music-video. Many of the visual sections were already conceptualised and shot according to and on the basis of certain qualities of the audio. But, during the montage process these visual sections were changed once more, again in a direct correlation to the audio flow, when I listened to the song and made my montage decisions based on the mood and feel it generated. Then, I refined and adjusted these selections, making sure visual instances synced properly to the audio moment and that textures surged and moved in tight accordance with the progression of the audio.

By replaying the sections over and over again, I was able to make minute adjustments as I progressed, always using the audio as a point of reference and direct cinesthetic influence. When I was satisfied with a section, I mounted it to another section at an appropriate audio instance or I devised a visual bridge, using the same cinesthetic montage method, which would link one section to the next. In this way each visual instance was generated as an outcome of the audio input, and the images and the audio flowed into and onto each other as a consequence of the cinesthetic montage. Thus, the montage process of music-video proved to be fluid. It was in constant flux as the selection, montage and augmentation of the visual sections came as a consequence of the mood and feel generated by the audio flow either during the conceptualisation, shooting, or actual mounting. In the following three sections I will discuss in detail three excerpts of the full piece. These excerpts can be viewed on the attached DVD along with the full video piece.

Excerpt One: Dripping leaves

This section was one of the first that I conceptualised. The idea came as a direct response to the sounds of tinkling piano notes, which reminded me of raindrops. Thus, I devised this section of water dripping off leaves, almost a Zen-moment of passing time, to be correlated to each piano sound. Shooting several leaves in various compositions with water dripping off them, I ensured that the drop and its release from the tip of the leaf would be the focus of the shot. I then slowed down several sections, and selected the ones that correlated to the feel of the audio instances to which I wanted to connect them. Several layers of dripping leaves were superimposed onto each other in order for the multiple piano contours to have a drop shape each. Some of the piano notes were lighter than the other ones, so I reversed the drop images, so that the drop of water would come in reverse from the bottom of the screen and connect back to the tip of the leaf. Thus, the different lines of piano notes were visualised using the same footage, but reversing the movement of the visual line to harmonise with the two distinctly different aural lines.

Then, I augmented the form of the image to create a colour space and a textured surface that would correlate to the overall form of this particular audio section. The

image of multiple raindrops hitting a surface of water was superimposed on top of the leaves section to add more texture to the image as well as generate an interesting link between the drops and the surface of the image in which they existed, generating their own synchresis with the tinkling of the piano. I also devised a water flash for the impact of the bell/horn sound that comes in the middle of this excerpt. It was then necessary to link this section with the previous section in order to create an uninterrupted visual flow. The piano tinkles had been visualised in the previous section through a superimposed shot of blinking fairy lights filmed out of focus. I prolonged that particular surface texture so that it would flow on top of the sun twinkles, creating a visual bridge. In this way a double set of light twinkles was used to cinesthetically visualise the tinkles of piano. I reused the fairy-light texture later in this section to generate a deeper surface texture and add more visual twinkles. The section ended as another bell struck, linked with a drop of red ink diluting in water carrying the drop metaphor into the next section.

Excerpt Two: Colour drop

For this section I needed a massive visual impact to correlate to the profound strikes of two separate bell instances. Furthermore, these strikes would elongate into stretched undulations of sound, something the image needed to echo. The overall form and flow of this section is rather dark and ominous with its warlike marching drums. Thus, I devised a dark and frenzied section of multiple flashing lights, fast-motion driving and other elements that would echo the portentous mood of the audio, while generating synchresis with the drums. The deeply textured and furious visual section was effective, but the bell strikes still needed their own specific visual 'thrusts'. Since these hits needed to stretch out, I used the ink drops that hit the water almost like inverted drops, which seep and surge into constantly changing shapes as they dilute. I chose one drop for each of the bells, but felt the second one needed some extra impact to echo the increase in audio intensity. The drop was therefore doubled so that it would move towards a copy of itself as it elongated across the image surface. Superimposed onto the frantic driving section, it generated a continuously shifting tapestry of coloured texture that elongated according to the undulations of the bell.

Excerpt Three: Tree zoom

This section contained a lot of elongated audio textures, a set of intense bells striking avidly, as well as a beautiful little piano element that seemed to open up like a flower blossoming. The audio textures were correlated to a visual surface of water slowly moving and changing colours as coloured ink spread through it, superimposed on top of a section of interlocking and interweaving tree branches. The water flash used in the dripping leaves section was reused, followed by a continuously shifting texture that would correlate to the flow of the audio as it stretched into the massive impact of the ringing bells. These bells were linked to a double set of sped-up zoom sections. As a result, the turbulent, but fluid, in-out movement of the zoom interacts directly with the bells. Simultaneously, the texture of the slowly moving and twinkling water on the image's surface, carried on from the previous cut, undulates to the contour of the elongated violin notes stretching through this section. As the piano begins to dominate and the audio space starts to expand, I took the image back to the interlocking branches but added a warm glow to the image that progressed according to the 'blooming' feel of the audio. This image then changes into a sky at sunset that intensifies as the piano reaches its peak, and the synchresis is complete. As the song opens up and expands, the image and its colour space develop accordingly and generate a cinesthetic link that works on both the micro- and the macro-level.

4.4 Conclusion

The cinesthetic montage process is a painstaking and complex process, but it generates a very effective result in terms of audiovisual synchresis. By adhering to the cinesthetic technique through every step of the process, including conceptualisation and shooting, it was possible for me to generate an audiovisual flow in which the music and the video surge into and onto each other. The image reflects the qualities of the song because the song, and the sensory responses it generates, were always the starting point and point of reference for each section of the process. It also enables immediate cinesthetic links to be generated, because the process is fluid enough to encompass augmentation based on the instantaneous sensory input of the mood and feel of the song at that particular moment. Every sequence of images was generated in a direct sensory response to a sound or flow of sounds. Such an embodied response took into account which visual elements echo which audio sections, and vice versa, consequently welding them together in terms of their inherent and resounding lines. Because this audiovisual expression process deploys the sensory system of the video artist, it facilitates an equal sensory impact when perceived by the audience. Thus, through the utilisation of this process it is possible to generate an audiovisual flow in which sound makes the images move and expand, while simultaneously images give the sound shape and visuality.

General Conclusion

Music is a poetic and emotive translation of the sensory responses to the world generating 'sensations', and 'moods' in the listeners, rather than a clear and structured narrative. The music-video adds a visual rendition, a mirror, to this array of sensuous mental states. It attempts to visually represent the emotional geography and physiological impact of the audio. Thus, to translate the music into images, the music-video maker visually conceptualise the 'feelings', 'moods' and 'sensations' provided by the music. The purpose of this process is to generate a form and flow of moving images to trigger similar emotions in a musical form but through a visual input. Such a sensual conceptualisation process starts with the audio breakdown of the song. The intro, verses, choruses, instrumental sections and codas are identified. Then, the working musical parts generated by the instruments, which make up each section, are categorised. These musical elements are then connected directly to each sensually conceptualised visual element that makes up the video. In this way, the images connect to the music, echoing and visualising the sound, linking up in such a way that they interweave and flow onto *and* into each other.

Therefore, each audiovisual sense the music-video expresses enables the transposition of sonic velocity into the image and visual shape onto the sound, providing sensations that go beyond the simple separation of hearing and seeing. This is possible because the senses of the embodied sensory system that experience the audiovisual flow do not operate independently. They are part of a non-hierarchical and cross-modal system in which one sense modality is commutable and transposable onto and into the other sense modalities of the lived-body sensorium. The process by which this commutable sensory system works is analogous to the synaesthetic condition and, consequently, is applicable to the moving image because it is embodied and synaesthetic just like the human lived-body. By meticulously interconnecting the images and the music through the embodied cinematic apparatus, it is possible to generate synaesthetically analogous links between the audiovisual elements. The lines inherent within each working musical part are correlated to resounding lines found within each visual element. Thus, by being expressed, as well as perceived, through the cinesthetic structure, the sound and image elements are welded together. The weld

generates a unified and commutable sensory expression, in which it is virtually impossible to determine whether it is the sound that makes the images move and expand or if it is the images that give the sound shape and visibility.

The audiovisual weld is generated through a process of cinesthetic montage that correlates the audio-visual lines to one another. Cinesthetic montage joins image-to-image, image-to-sound, and sound-to-image according to the 'feel' of each element, interconnecting the lines inherent in each element with comparable lines in other elements. By relating the images step-by-step to the music through the identical motion that lies at the base of the sound and image movement and the movement of each element's harmonic lines, an organic bond is generated between audio and images. This audiovisual bond exists within, as well as between, each audio-visual element from which the music-video is constructed. Thus, the harmonic reverberations that connect the elements of the audiovisual flow can be found on both the macro-scale of form and flow and on the micro-scale of phrases and contours in which specific musical moments directly resound in specific visual instances.

Consequently, on the one hand, the shots are conceptualised through the emotive responses to the song and produced to transfer these emotions into visuals. Each sensually conceptualised shot echoes in the audio elements, cinesthetically expanding and visualising the form and flow already established by the music. In contrast, both the shots and the macro- and micro-level audiovisual connections inherent within them are welded to the song by means of cinesthetic montage. Every image, or sequence of images, is determined by its correlated sound or flow of sounds, that is, by the interconnections between the form and flow of both the music and the video. Because the sound and the image flow, undulate, and develop in tight accord and through pre-determined connections with each other, the embodied perception of the cinesthetically expressed audiovisual form and flow works directly *on* and *in* the viewers. Thus, the initial ignition of this harmonic connectivity does not reside either in the images or the sound. The accord, which is clearly *there*, cannot be perceived as starting or ending at any particular time indicating its correspondence with the direct experience.

In the cinesthetic audiovisual flow the audio and visual elements harmonise and echo each other because they are conceptualised from and with each other, and subsequently welded together according to their inherent and resounding lines. Therefore, this cinesthetic audiovisual form expressed and generated by an embodied cinematic apparatus, in which sound can be cinesthetically expressed and perceived as image and the image perceived and expressed as sound, is perceived and co-constructed by the embodied sensory system of the viewing listener.

Bibliography

- Bjørnberg, A. (1999). Structural relationships of music and images in music video. *Popular Music*, 13(1), 51-74. Retrieved January 2, 2005 from JSTOR database.
- Brougher, K. (2005). *Visual music: synaesthesia in art and music since 1900*. New York: Thames & Hudson.
- Casebier, A. (1991). *Film and Phenomenology*. New York: Cambridge University Press.
- Chion, M. (1994). *Audio-Vision: Sound on screen*. (C. Gorbman, Trans. & Ed.). New York: University of Columbia Press. (Original work published in 1990).
- Cytowic, R. E. (1989). *Synesthesia: A union of the senses*. New York: Springer-Verlag.
- (1995). Synesthesia: Phenomenology and neuropsychology. *Psyche*, 2(10). Retrieved March 6, 2005, from <http://psyche.cs.monash.edu.au/v2/psyche-2-10-cytowic.html>
- Eikhenbaum, B. (1973). Literature and cinema. In S. Bann & J. E. Bowlt (Eds.), (T. L. Aman, Trans.), *Russian Formalism: A collection of articles and texts in translation* (pp. 122-127). Edinburgh: Scottish Academic Press. (Original work published 1926)
- Eisenstein, S. (1970). *The Film Sense*. (J. Leyda, Trans. & Ed.). London: Faber & Faber. (Original work published 1943)
- (1988a). The Montage of Film Attractions. In R. Taylor (Trans. & Ed.), *Selected Works Vol. 1: Writings 1922-34* (pp. 39-58). London: BFI Publishing. (Original work published in 1924)
- (1988b). An Unexpected Juncture. In R. Taylor (Trans. & Ed.), *Selected Works Vol. 1: Writings 1922-34* (pp. 115-122). London: BFI Publishing. (Original work published in 1928)
- (1988c). The Dramaturgy of Film Form (The dialectical approach to film form). In R. Taylor (Trans. & Ed.), *Selected Works Vol. 1: Writings 1922-34* (pp. 161-180). London: BFI Publishing. (Original work published in 1929)
- (1998a). Beyond the Shot. In R. Taylor (Trans. & Ed.), *The Eisenstein Reader* (pp. 82-92). London: BFI Publishing. (Original work published in 1929)
- (1998b). The Fourth Dimension. In R. Taylor (Trans. & Ed.), *The Eisenstein Reader* (pp. 111-123). London: BFI Publishing. (Original work published in 1929)

- (1998c). From lectures on Music and Colour in *Ivan the Terrible*. In R. Taylor (Ed.), (W. Powell, Trans.), *The Eisenstein Reader* (pp. 167-186). London: BFI Publishing. (Original work published in 1947)
- (1998d). Statement on Sound. In R. Taylor (Trans. & Ed.), *The Eisenstein Reader* (pp. 80-81). London: BFI Publishing. (Original work published in 1928)
- Eshun, K. (2000). The Kinematic Pneumacosm of Hype Williams: The rhythm of vision is a dancer. In P. Brophy (Ed.), *Cinesonic: Cinema and the sound of music* (pp. 51-58). Sydney: AFTRS.
- Flips 5: Music vs Motion*. (2001). London: Systems Design Limited.
- Goodwin, A. (1987). Music video in the (Post) modern world. *Screen*, 28(3), 36-55.
- (1992). *Dancing in the Distraction Factory: Music television and popular culture*. Minneapolis: University of Minnesota Press.
- Gow, J. (1992). Music video as communication: Popular formulas and emerging genres. *Journal of Popular Culture*, 26(2), 41-70.
- Grossberg, L. (1988). 'You [still] have to fight for your right to party': Music television as billboards of post-modern difference. *Popular Music*, 7(3), 315-332. Retrieved January 2, 2005 from JSTOR database.
- Hill, J. (2000). Film and postmodernism. In J. Hill & P. C. Gibson (Eds.), *Film Studies: critical approaches* (pp. 94-103). Oxford: Oxford University Press.
- Howells, R. (2003). *Visual Culture*. Cambridge: Polity Press.
- Husserl, E. (1962). *Ideas: general introduction to pure phenomenology*, trans. W. Boyce-Gibson. New York: Collier.
- Ihde, D. (1970). Auditory Imagination. In F. J. Smith (Ed.), *Phenomenology in Perspective* (pp. 202- 215). The Hague: Martinus Nijhoff.
- (1990). *Technology and the Lifeworld: From garden to earth*. Bloomington: Indiana University Press.
- Kandinsky, W. (1982a). On stage composition. In K. C. Lindsay & P. Vergo (Eds. & Trans.) *Kandinsky: Complete writings on art* (Vol. 1, pp. 257-266). London: Faber & Faber. (Original work published 1912)
- (1982b.) On the spiritual in art. In K. C. Lindsay & P. Vergo (Eds. & Trans.) *Kandinsky: Complete writings on art* (Vol. 1, pp. 114-220). London: Faber & Faber. (Original work published 1911)
- Kaplan, E. A. (1987). *Rocking Around the Clock: Music television, postmodernism, and consumer culture*. New York: Methuen.

- Kinder, M. (1984). Music video and the spectator: Television, ideology and dream. *Film Quarterly*, 38(1), 2-15. Retrieved January 2, 2005 from JSTOR database.
- Laing, D. (1985). Music video: Industrial product, cultural form. *Screen*, 26(2), 78-83.
- Lant, A. (1995). Haptical cinema. *October*, 74 (Autumn), 45-73. Retrieved April 24, 2005 from JSTOR database.
- Marks, L. E. (1978). *The Unity of the Senses: Interrelations among the modalities*. New York: Academic Press.
- Marks, L. U. (2000). *The Skin of the Film: Intercultural cinema, embodiment and the senses*. Durham, NC: Duke University Press.
- Matthews, J. H. (1971). *Surrealism and Film*. Ann Arbor, Michigan: University of Michigan Press.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*, trans. C. Smith. London: Routledge.
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
- Milicevic, M. (1997). *Experimental/Abstract Film and Synaesthetic Phenomena (1725-1970)*. Retrieved March 1, 2005, from <http://www.lmu.edu/acad/personal/faculty/mmilicevic/pers/exp-film.html>
- Mundy, J. (1999) *Popular music on screen: from Hollywood musical to music video*. Manchester: Manchester University Press.
- Pisters, P. (2003). *The Matrix of Visual Culture: Working with Deleuze in film theory*. Stanford, CA: Stanford University Press.
- Ramachandran, V. S., & Hubbard, E. M. (2003, April 15). Hearing colors, tasting shapes. *Scientific American*. Retrieved March 6, 2005, from http://www.sciam.com/print_version.cfm?articleID=0003014B-9D06-1E8F-8EA5809EC5880000
- Reiss, S. & Feineman, N. (2000). *Thirty Frames Per Second: The visionary art of the music video*. New York: Harry N. Abrams, Inc.
- Rush, M. (2005). *New Media in Art* (2nd ed.). New York: Thames & Hudson.
- Schapiro, M. (1978) *Modern Art: 19th & 20th century*. New York: George Brazillier.
- Shklovsky, V. (1991). *Theory of Prose* (B. Sher, Trans.). Elmwood Park, IL: Dalkey Archive Press.

- (1994). Poetry and prose in cinema. In J. Hill & P. C. Gibson (Eds.), *Film Studies: critical approaches* (pp. 63-64). Oxford: Oxford University Press.
- Sobchack, T., & Sobchack, V. (1980). *An Introduction to Film*. Boston: Little Brown
- Sobchack, V. (1990). The Active Eye: A phenomenology of cinematic vision. *Quarterly Review of Film and Video*, 12 (3), pp. 21-36.
- (1992). *The Address of the Eye: A phenomenology of film experience*. Princeton, NJ: Princeton University Press.
- (2004a). The Expanded Gaze in Contracted Space: Happenstance, hazard, and the flesh of the world. In *Carnal Thoughts: Embodiment and moving image culture* (pp. 85-108). Berkeley, CA: University of California Press.
- (2004b). The Scene of the Screen: Envisioning photographic, cinematic, and electronic “presence”. In *Carnal Thoughts: Embodiment and moving image culture* (pp. 135-162). Berkeley, CA: University of California Press.
- (2004c). What My Fingers Knew: The cinesthetic subject, or vision in the flesh. In *Carnal Thoughts: Embodiment and moving image culture* (pp. 53-84). Berkeley, CA: University of California Press.
- Spinrad, P. (2005). *The VJ Book: Inspirations and practical advice for live visuals performance*. Los Angeles, CA: Feral House.
- Thompson, K. & Bordwell, D. (1994). *Film History: An Introduction*. New York: McGraw-Hill.
- Van Campen, C. (1997). Synesthesia and artistic experimentation. *Psyche*, 3(6). Retrieved March 6, 2005, from <http://psyche.cs.monash.edu.au/v2/psyche-3-06-vancampen.html>
- (1999). *Artistic and Psychological Experiments with Synesthesia*. Retrieved March 6, 2005, from <http://home-1.tiscali.nl/~cretien/pub/syneng/htm>
- Van Leeuwen, T. (1999). *Speech, Music, Sound*. Houndmills, Basingstoke: MacMillan.
- Vernallis, C. (2001) The Kindest Cut: functions and meanings of music video editing. *Screen*, 42(1), 21-48.
- (2004). *Experiencing Music Video: Aesthetics and cultural context*. New York: Columbia University Press.
- Wahlberg, M. (2003). *Figures of Time: On the phenomenology of cinema and temporality*. Stockholm, Sweden: Stockholm University.
- Walker, J. A. (1987). *Cross-overs: Art into pop/pop into art*. London: Comedia/Methuen.

- Wallis, R. (1988). Push-pull for the video clip: A systems approach to the relationship between the Phonogram/Videogram industry and music television. *Popular Music*, 7(3), 267-284. Retrieved January 2, 2005 from JSTOR database.
- Williams, K. (2000). Music Video, Technology and the Reversal of Perspective. In R. Ascott (Ed.) *Art, Technology, Consciousness: mind@large* (pp. 103-108). Bristol, UK: Intellect.
- Wollen, P. (1986). Ways of thinking about music video (and post-modernism). *Critical Quarterly*, 28(1/2), 167-170.
- Youngblood, G. (1970). *Expanded Cinema*. New York: P. Dutton & Co., Inc.
Retrieved May 20, 2005 from
http://www.vasulka.org/Kitchen/PDF_ExpandedCinema/ExpandedCinema.html

Films and Music Videos

- Cunningham, C. (1998). Portishead – Only You. From Bangs, L. (2003). *The Work of Director Chris Cunningham*. USA: Palm Pictures.
- Glazer, J. (1997). Radiohead – Karma Police. From Bangs, L. (2005a). *The Work of Director Jonathan Glazer*. USA: Palm Pictures.
- Lye, L. (2001). *Rhythm*. France: Re:Voir Video.
- Malick, T. (2005). *The New World*. United States: New Line Home Entertainment.
- Man Ray (2005). *Avant Garde: Experimental cinema of the 1920s and '30s*. United States: Kino on Video.
- Romanek, M. (1992). Keith Richards – Wicked As It Seems. From Bangs, L. (2005b). *The Work of Director Mark Romanek*. USA: Palm Pictures.
- (1994). G. Love & Special Sauce – Cold Beverage. From Bangs, L. (2005b). *The Work of Director Mark Romanek*. USA: Palm Pictures.
- Spielberg, S. (1998). *Saving Private Ryan*. United States: Paramount Pictures.
- Strand, J. (2005). *Moment*. Australia: Curtin University, Department of Media Information.

List of Illustrations

Chapter One

- | | | | |
|----|----------|-----------------|---|
| 22 | fig. 1.1 | “Crosses” | from Spielberg – <i>Saving Private Ryan</i> |
| 23 | fig. 1.2 | “Flames on Car” | from Glazer – <i>Karmapolice</i> |

Chapter Two

- | | | | |
|----|----------|-----------------|--|
| 31 | fig. 2.1 | “Liquid Lights” | from http://www.nkp.com/wick1/imag1.htm |
| 37 | fig. 2.2 | “B & W Images” | from Romanek – <i>Cold Beverage</i> |
| 42 | fig. 2.3 | “Light Flashes” | from Romanek – <i>Wicked As It Seems</i> |

Chapter Three

- | | | | |
|----|------------------|---------------------|-----------------------------------|
| 56 | Fig. 3.1 | “Boy in Alleyway” | from Cunningham – <i>Only You</i> |
| 59 | Fig. 3.2 | “CU/MCU” | from Cunningham – <i>Only You</i> |
| 60 | Fig. 3.3 | “Hand Reaching Out” | from Cunningham – <i>Only You</i> |
| 64 | Fig. 3.4 | “Men in Window” | from Cunningham – <i>Only You</i> |
| 65 | Fig. 3.5 | “Ramped Movements” | from Cunningham – <i>Only You</i> |
| 69 | Fig. 3.6 | “Woman’s Hair” | from Cunningham – <i>Only You</i> |
| 70 | Fig. 3.7 | “Flash/Camera Move” | from Cunningham – <i>Only You</i> |
| 72 | Fig. 3.8 & 3.9 | “Dove” | from Cunningham – <i>Only You</i> |
| 74 | Fig. 3.10 & 3.11 | “BamBAAAMMM Move” | from Cunningham – <i>Only You</i> |

Chapter Four

- | | | | |
|----|----------|-------------------------|--|
| 80 | Fig. 4.1 | “ <i>Moment</i> Poster” | created and supplied by bunniboi.com |
|----|----------|-------------------------|--|

Appendix A

Only You Synopsis

In a dark and dirty back alley a young boy floats, as if underwater, into frame. His features are starkly lit by neon, and his movements are slow and undulating as he glides through the dreamlike submerged space. His movements are ramped and stilted in time with the music. As he hits the ground he is surprised that he can walk, taking a few tentative steps, he notices a man looking crossly down at him from the window of a neon-blue lit room in one of the surrounding buildings.

A woman is also floating in the same alleyway. Oblivious to the boy, she sings her mournful song, looked on by more lone men in neon-blue rooms. But the men seem to not see her, they are focused on the boy, almost controlling the ramped and stilted meaningless motions he has to perform. Using a handkerchief, the boy starts doing a magic trick involving elaborate backward-moving motions, ending with the handkerchief turning into a dove that the boy releases. The dove flies up past a large looming building and another lone man in neon-blue room, and finally towards the full, yellow moon dominating the black threatening sky. Seeing the boy, the woman smiles and reaches her hand out to him. Looking scared, the boy reluctantly takes her hand.

But the hands are immediately broken apart by the angry stares of the men in the surrounding windows, and the seething horn section high-point of the song. The camera floats towards one of the men striking a threatening pose, and swoops up past a building, revealing floor after floor of neon-blue rooms occupied by the overseers. As the camera plunge past the top of the building and out, the boy is forced into a set of repeated meaningless flips. The boy flips around and around in perpetual motion. As the music fades, he drifts back, flipping, and is swallowed by the dark shadows of the alleyway, and the image fades to black.

Appendix B

Only You Lyrics

We suffer everyday,
What is it for,
These crimes of illusion
Are fooling us all,
And now I am weary,
And I feel like I do.

It's only you,
Who can tell me apart,
And it's only you,
Who can turn my wooden heart.

The size of our fight,
It's just a dream,
We've crushed everything,
I can see, in this morning selfishly,
How we've failed,
And I feel like I do.

It's only you,
Who can tell me apart,
And it's only you,
Who can turn my wooden heart.

Now that we've chosen to take all we can,
This shade of autumn, a stale bitter end,
Years of frustration lay down side by side.

And it's only you,
Who can tell me apart,
And it's only you,
Who can turn my wooden heart.

It's only you,
Who can tell me apart,
And it's only you,
Who can turn my wooden heart.

Appendix C

Audio Breakdown of *Only You*

Symbol Key:

SHAPES: main repeated elements of the audio.

- **B:** individual bass beat.
- **VB:** vibraphone-bing.
- **VBF:** vibraphone flutter: vibraphone-bing that extends into an echoed flutter.
- **S:** scratching.
- **O:** organ. This shape is sometimes also extended into an extended flutter.
- **H:** horn-samples.

PHRASES: combination of shapes to make repeated musical motifs.

- **O/B/VB:** combination of organ, bass, and vibraphone-bing.
- **O/BB/VBB:** organ, double bass beat, double vibraphone-bing.
- **O/S/VBF:** organ, scratching, and vibraphone flutter.
- **SS:** scratching of sample line: “move it like that”.

Instrumental 1/Intro

Bass/VibraphoneBing Opens:

Organ —Stretches→

Organ/ScratcheS/ViBraphoneFlutter —Stretches→ Horns →

↪ (**Organ —Stretches→**)

O/B/VB: SS – (Break) “Move It Like That”

↪ (**Organ —Stretches→**)

O/B/VB: SS – (Break) “Move It Like That”

(Break) ↪ (**Organ —Stretches→**)

Organ/BassBass/VibraphoneBingBing: SS – “Move It Like That”

(Break)

O/S/VBF: Scratch→Break

B: Scratch→Break

B: SS – (Break) “Move It Like That”

O/BB/VBB

Verse 1 ↪

“We suffer everyday,”

(Break)

B “What is it for,”

(Break)

O/B/VB “These crimes of illusion”

(Break)

B “Are fooling us all,”

(Break)

O/BB/VBB “And now I am weary,”

Organ

↓

↓

“And I feel like I do.”

Chorus 1 ↪ **O/S/VBF** “It's only you,”

↓

“Who can tell me apart,”

O/S/VBF “And it's only you,”

↓

“Who can turn my wooden heart.”

B→(Break)→**H**

Horns

Verse 2 ↪ **O/B/VB** “The size of our fight,”

(Break)

B “It's just a dream,”

(Break)

O/BB/VBB “We've crushed everything,

Organ

↓

↓ “I can see, in this morning selfishly,” (LONGER SENTENCE)

↓
↓
↓
↓
↓

↓ "How we've failed,"
↓
↓ "And I feel like I do."

Chorus 2 ↪ **O/S/VBF** "It's only you,"

↓
"Who can tell me apart,"

O/S/VBF "And it's only you,"

↓
"Who can turn my wooden heart."

↑ ↑
B **B → Horns →**

Instrumental 2 (Break) **B/VB: SS** – "Move It Like That" (with Breaks)

(Break) ↪ (Organ —Stretches→)
O/B/VB: SS – "Move It Like That"

(Break) ↪ (Organ —Stretches→)
O/BB/VBB: SS – "Move It Like That"

O/S/VBF: Scratch ("Ove It")→ Scratch ("Like")
B: Scratch ("Ove It")→ Scratch ("Like")
B: Scratch ("Ove It")→ **SS** – "Move It Like That"
↳ (Music starts again)

(Break) Scratch: **VBF(solo)**

Bridge ↪ → Stretches →

"Now that we've chosen to take all we can,"

(No music: record player scratching sound) ↓
Horns

O/B/VB (Music starts again) "This shade of autumn, a stale bitter end,"

Organ ↓
↓ **O/BB/VBB** "Years of frustration lay down side by side."

Chorus 3 ↪ → **O/S/VBF** "And It's only you,"

↓
"Who can tell me apart,"

O/S/VBF "And it's only you,"



“Who can turn my wooden heart.”

O/S/VBF “It's only you,”



“Who can tell me apart,”

O/BB/VBB + S/VBF “And it's only you,”



“Who can turn my wooden heart.”



B



B → SS →

Instrumental 3/Solo/Outchorus

Horns:

Bam-Bam/Bam-Bam/Bam-Bam/BamBamBam → BAAAAMMMM! →

Rising Backwards **Organ + Strings** → **Organ: TutTut:** → **Organ solo**

Guitar: Baeowh (repeated) with **O/B/VBx5** (then once with **O/BB/VBB**)

SS/H/O/VBF + (Breaks) throughout

Song fades out