

The kindest cut: functions and meanings of music video editing

CAROL VERNALLIS

When critics of film and television say that something is ‘cut like music video’ or refer to ‘MTV-style editing’, what do they mean? They might mention quick cutting or editing on the beat, and indeed one can see that the edits in music video come far more frequently than in film, that many stand out as disjunctive, and that the editing seems to have a rhythmic basis closely connected to the song. These last two features of music video editing – that it is sometimes meant to be noticed and that it brings out aspects of the song – suggest at once that it does something different, and perhaps something more, than does the editing in film. Music video editing bears a far greater responsibility for many elements than does classic Hollywood film editing. Not only does the editing in a music video direct the flow of the narrative, it can underscore non-narrative visual structures and form such structures on its own. Like film editing it can colour our understanding of characters, but it has also assimilated and extended the iconography of the pop star.

Much of the particularity of music video editing lies in its responsiveness to the music. It can elucidate aspects of the song, such as rhythmic and timbral features, particular phrases in the lyrics, and especially the song’s sectional divisions. Because it can establish its own rhythmic profile, the editing can provide a counterpoint to the song’s rhythmic structures. More subtly and also most importantly, the editing in a music video works hard to ensure that no single element (the narrative, the setting, the performance, the star, the lyrics, the song) gains the upper hand. Music video directors rely on the editing to maintain a sense of openness, a sense that any

1 Theorists like Roman Jakobson have noted that in classic Hollywood film all elements – lighting editing music – become subsumed in the narrative which functions as the dominant through-line. See Roman Jakobson *Readings in Russian Poetics* (Cambridge MA MIT Press 1971), pp 82–7. In music video on the other hand the narrative is only one element among many any parameter can come to the fore, grab our attention and then quickly recede from view. We do notice the editing in Hollywood films but only rarely, for example during a frenetic action sequence or at the heightened moment when the camera peers down a gun barrel.

I refer to classical Hollywood editing in its most generic Bordwellian sense see David Bordwell Janet Staiger and Kristin Thomson *The Classical Hollywood Cinema* (London Routledge & Kegan Paul 1985). There are of course, a number of directors who highlight editing as a stylistic feature, such as Leone or Peckinpah and those who break the rules, like Dreyer and Ozu. Music video editing has influenced contemporary filmmaking to the extent that the border between media has become less distinct.

I do not intend to minimize the ways that editing contributes to a narrative film's organization and its effect on spectators. See Ayako Saito Hitchcock's trilogy a logic of mise en scene in Janet Bergstrom *Endless Night* (Berkeley CA University of California Press 1999) pp 200–249.

Editing tends to be an undertheorized area of film studies. The most significant texts on film editing are Karel Reisz *The Technique of Film Editing* (Boston MA Focal Press 1988) and Ken Dancyger *The Technique of Film and Video Editing* (Boston MA Focal Press 1996).

2 David Bordwell notes that graphics rhythm space and time then are at the service of the filmmaker through the

element can come to the fore at any time. The editing does so in part simply through being noticed. By demanding attention, it prevents powerful images from acquiring too much weight and stopping the flow of information. The editing thus preserves the video's momentum and keeps us in the present. A striking edit can allow one to move past a number of strange or disturbing images while neither worrying about them nor forgetting them completely.¹

Music video's complexity stems not only from the sheer number of functions it serves, but also from the way that it moves unpredictably among these functions. It may be helpful to picture the succession of images in a video, and the edits that join them, more as a necklace of variously coloured and sized beads than as a chain. This picture not only emphasizes the heterogeneity of shots in music video, but also suggests the materiality of the edit itself. Indeed, sometimes the edit seems to function as a part of the image and sometimes as a gap.

This essay is divided into five sections: (1) an introduction to the grammar of shots and edits, and a discussion of how their functions differ between film and music video, (2) an analysis of the role of editing in making meaning, creating narrative, and establishing other forms of continuity, (3) an explanation of how music video image adapts to the processual nature of sound; (4) an examination of the ways that music videos treat closeups of the star; and (5) a presentation of the means by which editing can reflect musical features. Because it deciphers types of shots and edits, this essay functions as a grammar for music video editing; however, it also contains a theoretical component. I argue that the edits in music video mean something different – and create meaning differently – than do their filmic counterparts. Even when a shot and edit in a music video remind one of classic narrative film, their function may have undergone a change of valence. Music video editing, like camera movement and camera placement, enables relations between the song and the image.

Shots and edits

When constructing a taxonomy of shots and edits in music video, one can begin with traditional narrative film practices. The continuity system forms the basis of film editing but is much less common in music video. Common continuity edits in film include the 180-degree rule, which preserves screen direction, as well as the thirty-degree rule, which prevents a jump-cut between two shots, and also shot/reverse-shots, over-the-shoulder shots and matches on actions.² Such edits attempt to naturalize the movement from shot to shot and render the break as seamless as possible. Continuity editing seeks to preserve the flow of time and the coherence of spaces. The ultimate

technique of editing. Yet most films we see make use of a very narrow set of editing possibilities – so narrow indeed that we can speak of a dominant editing style throughout Western film history. The purpose of the system [continuity editing] was to tell a story coherently and clearly, to map out the chain of characters' actions in an undistracting way. David Bordwell, *Film Art* (New York: McGraw-Hill, 1987) pp. 284–5.

3 Music videos seldom present a clear path through their structure. Sustained sequences of pure cross-cutting are rare (the classic example is a shot of a man on a horseback racing to the train, then the speeding train, then back to the man on horseback) as is a figure chasing another in a single shot. More levels of activity may be necessary to underscore the heterogeneous quality of the song. At first glance, the Clash's 'Rock the Casbah' seems like a simple chase – Egyptians chase Israelis. A closer examination of the video reveals that the armadillo that runs at the bottom of the frame complicates the relations among the figures and the music.

4 Videos establish a sense of continuity on the surface partly by using dissolves. Smoother than cuts, dissolves provide a fuzzy articulation, rather than a sharp one, against the song. They do not therefore require a strong rhythmic commitment to a single musical feature.

Director David Fincher's background is in graphic art. In an interview, he told me that when he shot music videos he deliberately avoided learning the information commonly taught in film schools. He believed he would make more interesting work by not sharing a common background with other filmmakers.

goal of continuity editing is to create a single, clear path through a film's world. Because music videos seem to benefit from providing a multiplicity of incomplete, sometimes obscure paths, continuity editing will serve different functions and govern only isolated sections of a video.³ Perhaps music videos avoid continuity editing because such techniques would give the visual track too strong a forward trajectory: the image might seem to overtake the song. A music video's aim is to spark a listener's interest in the song, to teach her enough about it that she first remembers the song and second, purchases it. Music video's disjunctive editing keeps us within the ever-changing surface of the song. Though such edits may momentarily create a sense of disequilibrium, they force the viewer to focus on musical and visual cues, allowing her to regain a sense of orientation. In addition, the dense, oblique quality of a string of imagery can serve to showcase the star. The viewer may experience a jolt of accomplishment and pleasure as she passes through a thicket of imagery to come upon a clearing where she finds herself alone with a closeup of the performer. One of the most narrative music videos, Aerosmith's 'Janie's Got a Gun', comes the closest to following the rules of traditional Hollywood continuity, yet it also extends and breaks these rules. The video concerns incest and, at one point, the father stands at the threshold of his daughter's room while the mother watches. In this sequence, sightlines do not match. Consequently, it will take a while for the viewer to notice that, based upon the position of the characters, the mother is watching the father, yet he does not see her. In addition, the thirty-degree rule is violated between the medium and closeup shots of the father. (The camera angle between shots is narrower than thirty degrees, and the object in the frame appears to jump.) Music videos avoid matches on action, often extending or abbreviating a shot to give the sense of a cut in the 'wrong place'. This effect blunts narrative progress and creates a rhythmic emphasis on the moment when the edit occurs. (Figure 1)⁴

It will be helpful to widen our consideration of editing techniques to include not only those of classic Hollywood films, but also those in the Russian formalist film tradition. Karl Reisz's description of Eisenstein's *October* works as a characterization of music video editing.

Indeed, as a piece of narrative, [*October*] is extremely unsatisfactory. The incidents are loosely constructed and do not follow each other with the dramatic inevitability which a well-told story demands. We are not, for instance, shown Kerensky's character through a series of dramatically motivated episodes but through a number of random incidents, each suggesting a further aspect of Kerensky's vanity or incompetence. The time relationship between consecutive shots and scenes is left undefined and no

Downloaded from screen.oxfordjournals.org at University of Oregon Libraries on June 30, 2011

Figure 1
Aerosmith, 'Janie's Got a Gun'



5 Reisz, *The Technique of Film and Video Editing*, p 35

sense of continuous development emerges: the cut from [shots] 108 to 109, for example, takes us – without reason or explanation – from the Czar's study to a staircase somewhere in the palace. No attempt is made to explain or to conceal the time lapse between the shots, as could easily have been done with a dissolve.⁵

It seems intuitive that the Russian film formalists (precursors to the experimental filmmaking tradition) should share a lot with music video directors, but some of the reasons why this is so may not be immediately clear. We should remember that the early Russian film directors Kuleshov, Pudovkin and Eisenstein worked with minimal resources (film was so valuable that it was recycled for its silver) and they used almost no intertitles because they were making films for a largely illiterate public. Even though music video directors can command great resources in the era of late capitalist production, they too struggle with limitations. These limitations may seem trivial by comparison: music video is a short form; the music and lyrics may be banal; the singer must lip sync while the rest of the figures remain silent; much time must be spent showcasing the star. Like the early Russian filmmakers, they have to make the most of the brute materials of film, and to make frames and cuts as expressive as possible.

One kind of edit – graphic match – appears frequently in music video and Russian formalist films, but much less often in traditional narrative film. Such an edit joins two shots through shared compositional elements such as colour or shape, irrespective of content. In film, graphic matches are normally used to join scenes. Strikingly formal, they appear unnatural in most other contexts. While we might see one or two graphic matches in the whole of a Hollywood narrative film, we often find two or three in a five-minute music video. Music video can use graphic matches so freely because the genre has reason to draw attention to its materials and production methods: the viewer can revel in an interesting edit, in a nice shape shared by two images, and in the cleverness of the director's and the

6 In a music video a graphic match will produce a momentary sense of surprise, a shock of recognition. This edit may take the viewer outside of the tape yet without 'skipping a beat', he or she will quickly become immersed again in the video's flow. The effect is similar to the small shocks that are associated with a sudden shift in colour or a lyric that creates a punning or disorienting effect. Narrative films use graphic matches much more sparingly, because they aim to enmesh the viewer in the constant unfolding of the narrative. To achieve this effect, editing should be as invisible as possible.

Figure 2
Peter Gabriel, 'Mercy St.'



7 The most elementary requirement of a smooth continuity is that the actions of two consecutive shots of a single scene should match – if a scene is shot from more than one angle, the background and positions of the players remain the same in each take. A more difficult aspect of the same problem is to keep the action and movement shown in consecutive shots accurately continuous. If an actor starts a movement – say he is half way through opening a door – in one shot, then that movement must be continued in the next from the precise moment it was left. Reisz, *The Technique of Film and Video Editing*, pp. 216–17.

Figure 3
Janet Jackson, 'Love Will Never Do (Without You)'

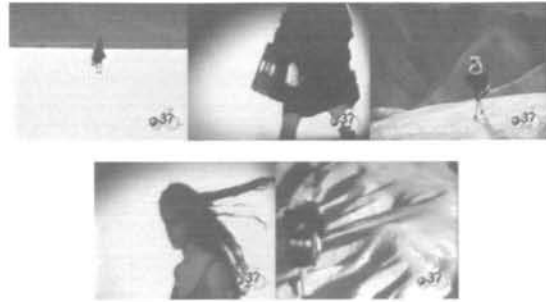


editor's work, any of which might draw us away from the narrative of a Hollywood movie.⁶ The graphic match can highlight elements of a popular song. As just mentioned, in the graphic match most visual parameters remain the same while one changes. In a pop song, a melody might be sung by the voice, and be picked up by a flute. In this case, the one feature that has changed is timbre. Both sound and image share qualities of transformation and continuity. (Figure 2) The concentration of imagery in music video has led to an expanded role for the visual principles of graphic match. In particular, videos will contain groups of non-adjacent shots that share bold compositional elements. I will discuss the nature and structural role of these groups of shots later in this paper. Film editing generally seeks to avoid placing a series of shots that contain motion against

one another; films will interpose a static shot when possible. When films cut between shots with motion, the editing tries to keep the viewer's eye in the same location to minimize the disjunctive effect.⁷ Only in action sequences or heightened dramatic scenes does the editing compel the viewer to shift focus rapidly from one edge of the frame to another. Music videos place movement against movement much more promiscuously, juxtaposing motion within the frame to camera movement, and mixing speeds, directions and durations. (Figure 3)

Unlike films, music videos frequently employ intentionally disjunctive edits. A jump-cut, which is generally unacceptable in film, can be avoided by shifting the camera at least thirty degrees between two adjacent shots that contain a change in scale. The jump-cut makes us feel that we are lurching forward or back. This kind of edit has been much discussed since its occasional deliberate use in films of the French new wave. (Figure 4) Music videos use jump-cuts liberally, along with a variety of other disjunctive edits. There are brusque edits that demand attention through a drastic shift in scale, colour or content. Sometimes we see an even stronger edit, one so clearly aestheticized that it separates itself from the flow of

Figure 4
Alanis Morissette, 'You Oughta Know'



the video. This edit might contain a roughened edge (say, a bit of film leader and a white flash sandwiched between the shots) that makes it work like a jump-cut.

It is also important that videos can create confusion about what is an edit and what is not. There are many ways to produce a meaningful articulation, both in camera and through post-production: the lens can continually change focal planes, or an element of the frame can pop forward, almost as though there had been an edit that affected only part of the frame. When other effects help to do the work of the editing in a video (defocusing, fading to black, strobing) the editing can perform other roles, such as creating an aesthetically pleasing visual line, or drawing the viewer's attention to the music. (Figure 5)

Figure 5
REM, 'Losin' My Religion'



Just as editing changes in the shift from film to music video, so does the function of shots. Music videos and Hollywood films share a basic premiss: that visual information can best be communicated by cutting between three kinds of shot – long, medium and closeup. One can describe these shots according to the relation each establishes between the figure and the space around him. In long shots, the space obtains a greater prominence than the figure; in medium shots, the relation is roughly equivalent; and in closeups, the figure dominates the space. Hollywood film has virtually standardized the cropping of the figures in these shots. Most textbooks recommend that the proportion of the figure to the space in the frame falls within set guidelines to achieve a sense of balance: if too much or too little space surrounds the figure, a shot is said to look awkward. Further, the camera should not frame the body in such a

8 See Gerald Millerson *The Technique of Television Production* (Boston, MA: Focal Press, 1999) p. 99

way that the frame's edge passes through a joint of the body, like the neck, elbows, knees or ankles.⁸ Music videos do not follow these rules. Not only is the relation between figure and space frequently off kilter, but the camera bisects the figure in places that would be unacceptable for classic Hollywood film. In film, the framing helps to draw attention to the content of a shot, rather than its composition, and to render the editing process invisible. The framing in music video makes us as aware of the edge of the frame (and of what we cannot see) as of the figure itself. This kind of framing can give a shot a precarious quality that the succeeding shot cannot always put right. In this way, the image moves forward, matching the momentum of the music. (Figure 6)



Figure 6

Viewers still learn the grammar of traditional shots from watching films and television. Music videos make use of these shots, but give them different functions and meanings. Our knowledge of film and television practices still provides a reference point, and can lend excitement to a shot in a music video that violates the rules of these practices. In Hollywood films, the extreme long-shot frequently serves to set the context for a new scene or to adopt a character's point of view. Such a shot might appear at any point in a music video as a way of exposing a space otherwise revealed only in fragments or of creating visual contrast. It may even help us to listen past the song's foreground elements to acknowledge the totality of its sound space. The absence or oblique presentation of master shots in music video means that the viewer does not own or know fully the space, but is taken through it. The closeup, in classic Hollywood film, will disclose something intimate about a character. In music video, the closeup can work similarly to showcase the star, but just as often it serves to underscore a lyric, a musical hook or the peak of a phrase. It may even be chosen simply to fit a pattern of shots already established within the edited sequence. In general, Hollywood cinematography's language of confrontation plays a greatly

diminished role in music video. Over-the-shoulder shots, separation shots and the 180-degree rule tend to make the relations between figures clear and specific in a way that would be inappropriate for most videos.

The use of camera angles can tell us much about the visual language of music video. Low-angle shots are used more extensively in music videos, partly because they reproduce the relations among audience, performer and stage. Such shots confer authority upon performers and assert their sexual charisma, often crudely, by highlighting the erogenous zones of performers. High-angle shots in music video, as in film, give the viewer a sense of power and mobility. (Figure 7) These shots perform other functions in videos. Sometimes an overhead or extremely high-angled shot is edited in to create a rhythmic unison with a key moment in the music, like the crest of a melody. Classic Hollywood film employs high-angle and low-angle shots sparingly; the camera quickly returns to a level perspective. A music video, by contrast, may contain a long series of high-angle or low-angle shots. When high- and low-angle shots are mixed together to form a series, the video will lack a sense of ground. The viewer turns to the music for additional spatial-temporal cues.



Figure 7
Guns N' Roses, 'November'

Traditionally, in forms such as Hollywood film, opera, the stage and oratory, the singer is placed in the centre and on a level field as a means of establishing centrality, stability, importance and clarity. When the singer is placed off centre (through framing, and so on) we might assume a different experience of the song. It is important to note that many music videos have parodied or deconstructed proscenium framing, for example in Nirvana's 'In Bloom', the band performs before an Arabian Nights backdrop for an Ed Sullivan Show. The scene is particularly humorous because the set was shot with low-resolution black and white video and the men, wearing skirts, quickly proceed to destroy it. Music videos often contain a long series of low- or high-angle shots which create a different relation between listener, music and image. For example, when there

is a long series of medium, low-angle shots (so that the performer appears from the waist up) as well as with an image that lacks stability – for example, with both camera movement and movement within the frame – one may have the sense that the song buoys the performer. In music video, there is no clear causality concerning which came first, music, lyrics or image, and at any moment, any one medium can be seen to influence the other. In this case, the performer is no longer the unambiguous source of the song. It can seem almost as if the image floats above and is carried by the music, literally as if sonic waves passed along the bottom of the screen, and the image bobbed up and down accordingly.

Camera movement in music video also differs from that of film. Most music videos make such extensive use of the dolly that a static shot seems anomalous. The dolly shot keeps the video moving; it starts almost invariably as soon as a video begins, and only ceases towards the end. It provides a simple way for a video to match and sustain a song's momentum. Director Marcus Nispel says that his work derives its musicality from a clever deployment of the dolly. He employs the dolly to create what he calls 'moves within moves' – the simultaneous use of tracking and panning.⁹ Nispel sometimes uses this scheme while a figure turns in the frame. (Figure 8) He thereby interposes three types of motion into one shot. A more common scheme places the artist at the centre of a circular track. The camera, often positioned at a low angle, moves back and forth along the track at various speeds. This scheme can create the sense of a performance space in almost any setting, while the low-angle and varied speeds give the camera a responsive, even performerly character. Videos can present a number of dolly shots edited together in order to build towards key moments in the video.

One type of music video camera movement that contrasts with the continuity editing system is the tracking shot. It is a camera

⁹ Interview with director, summer 1997



Figure 8
Amy Grant, 'House of Love'

movement whose meaning shifts when used in music video rather than in film, and is often used for special emphasis, frequently dominating a segment towards the middle of the song and punctuated by a few dissolves. Tracking shots play a crucial role in music video because they provide relief from a typically shallow sense of space. (In videos, we almost never pierce the background or stray far from the star.) The movement of the camera provides a change in point-of-view: instead of experiencing the music from a stationary position, as it rushes past, the viewer can get the sense of running alongside the soundstream. The tracking shot embodies perfectly the music video's attempt to match the energy of the song, to approach the song's rhythmic drive, even if the music remains just out of reach. The tracking shot can also constitute a distinct rhythmic stratum which will go in and out of synchronization with the song's other rhythmic strata. (Figure 9) Other kinds of camera movement function similarly. Cranes, pans, tilts and dramatic reframings are usually done by hand and can achieve the intimate effects associated with handheld camera work. These shots provide possibilities for textural detail and subtle expressive nuance. They can also mimic the ways that sound approaches and fades away.

Figure 9
Michael and Janet Jackson,
'Scream'



Different types of shots and edits can be mixed to create variety, and although one gets the sense that any shot can follow any edit, some shots and edits are particularly complementary. An edit or camera move can anticipate a gesture in the shot that follows. A tracking shot can complement a subsequent shot of a strutting figure. A crane shot that starts low and rises through the space seems to match a figure reaching outward. A brusque edit works well preceding a shot that contains a series of sharp rhythmic gestures performed by the dancer or musician. Videos can create this kind of play simply between shots and edits, almost irrespective of what a shot contains. A dissolve can pair nicely with a tracking shot, and the effect of a jump-cut can be extended by an unbalanced shot. This isomorphism or exchange of gestures and shapes teaches the viewer to move fluently from parameter to parameter while watching a video. Such movement can occur across many parameters, leading the viewer directly into the structure of the song. In pop music, materials are commonly shared among different domains, for example, a melodic line in the voice will be taken up by the guitar, though the rhythmic values may be expanded. The drums will be

performed 'out of the pocket' (off the beat) to showcase breaks in a singer's voice.

The syntax of shots in music video, taken as a whole, is less conventionalized than that of shots in film. In the traditional Hollywood film scene, the camera begins at a distance and gradually moves in – from long or master-shots to two-shots and medium-shots, to closeups like separation and over-the-shoulder shots. We seem to learn more about the film's world and the characters' inner lives as we narrow our focus in this way. The music video camera shifts more freely among types of shots. Since a shot decision is made partly according to the form of the music and the pattern established by preceding shots, the search for knowledge about people and places takes second place. One cannot construct a typical shot order for music video. One might thrill at a twirling overhead-shot that appears two-thirds into the video. In retrospect, the viewer will realize that this is a good choice within the structure of the video; however, while viewing the tape, he would not have been able to predict the appearance of the shot. Unlike films, music videos do not divide neatly into scenes. The song's sectional divisions provide a stronger basis for parsing a video. If one had to generalize about the syntax of music video image, one might take the musical phrase as the most significant unit. Music videos typically present segments of six to nine shots that last roughly the length of a musical phrase. While the beginnings and ends of these segments do not always align with those of musical phrases, they can be recognizable to a viewer because they contain internal repetition and often possess a kind of symmetry. (Figure 10)



Figure 10
Metallica, 'Enter Sandman'

Meaning, narrativity and continuity

How does music video image create meaning and in what ways does editing contribute to that creation? The meanings of music videos have been thought to present a puzzle. Most often, music video image is relatively discontinuous. Time unfolds unpredictably and without clear reference points. Space is revealed slowly and

incompletely. A video will hint at a character's personality, mood, goals or desires but will never fully disclose them. We seldom see an action completed – a figure's movement is often cut off by the edit. Stories are suggested but not given in full. Nor can the lyrics tell us what we need to know – they may be banal or purely conventional. A famous performer can also pull at the video's meaning – we cannot tell beforehand how or to what extent our knowledge of a star is intended to come into play in a given video.

Music video editing plays an interesting role in producing this effect of discontinuity. The editing in Hollywood film seeks to fill the gaps in our knowledge, to stabilize the meaning of an image. In music video, the editing seems rather to help create the discontinuity and sense of lack. If, as I have suggested, editing constitutes a distinct visual parameter of music video, we should expect that it can contribute to the creation of discontinuity. Since it, too, reveals things incompletely, makes promises it does not keep, it should be understood as but one of the elements fighting for attention in a video. And the case becomes more complicated: edits happen between images, they are not part of the image. Edits can literalize the discontinuity by making us aware of the space between images.

Scott McCloud, in *Understanding Comics*, writes about the pleasure of reading into incomplete images.¹⁰ He celebrates the interpretive work needed to transform black and white lines or spots of colour into meaningful characters. Music video images can provide the same pleasure. We know very little about the figures we see, but we still attempt to make sense of them, based on how they look and what they are doing, as well as the setting, the lyrics and the music. We must decide whether a figure functions as a character or merely as part of a tableau vivant. Extending the notion of the reader's share in the interpretation of the image, McCloud discusses the gap between panels of the comic. In this gap, the reader calculates the amount of time elapsed, the distance traversed, and any change in the figures. The edits in music videos work similarly. Partly by attending to the song, the viewer decides what has happened in the cut from one shot to another. The disjunctive force of the edit compels this decision. How do these two shots relate? On what basis does the edit link them together? And what is the net effect of these disjunctions on the video as a whole?

A video like Marilyn Manson's 'Beautiful People' highlights how difficult it can be to make sense of music video editing. Marilyn Manson first became known for a version of Eurythmics' spare synth-pop hit 'Sweet Dreams'. The Marilyn Manson version recast the song as a gothic metal dirge with a video that placed the androgynous singer in a decayed warehouse, wearing a variety of abject and incongruous costumes. 'Beautiful People', the band's next video, extended the gothic punk aesthetic of the first to encompass a much broader range of imagery and more serious themes. Where

10 Scott McCloud, *Understanding Comics* (Northampton MA: Kitchen Sink Press, 1994).

11 Christian Metz has noted that when we see a closeup of a gun on top of a bureau the image can be read as a sentence. Look this is a gun. David Bordwell similarly attentive to closeups remarks that when the camera focuses on a prop, we expect the film to make further use of it – in the case of the gun, for example a shooting. Unlike classical Hollywood cinema music video possesses few ways to make an image stand out from the rest of the visual flow. We may reach a peak in a music video, but feel unsure about what has led us to this moment. As we pass through an assortment of images, it can be hard to tell which serve to foreshadow an event and which perform a more purely decorative function. Consequently, our commitment to thinking about future events becomes attenuated.

12 The worm, as it dances at the edge of a precipice, seems highly marked, yet also, through its placement tangential – a thrown-away object. Only after reflecting upon the tape as a whole, might the viewer notice how key the image is. The worm stands for the spinal chords that Manson steals from his victims.

'Sweet Dreams' played with and against the singer's rock iconography. 'Beautiful People' remakes him as a Faustian figure who, we find out, experiments on human and animal subjects in order to gain control of the masses. The video's theme emerges quickly and its plot can ultimately be pieced together. 'Beautiful People' tells its story at an unpredictable rate, however, complicating the narrative with imagery whose origin and function are difficult to determine. The opening shots of the video demonstrate the modus operandi of music video's particular narrativity – the suggestion of a narrative along with a clear indication that this narrative will proceed elliptically and be rendered only in fragments.

The video opens with fifteen rapid static shots, many of which show parts of a human figure, prostheses or medical appliances. (Figure 11) Beakers and electronic devices suggest a laboratory. A worm, in closeup, dangles from the edge of a shelf. The video's theme of Faustian mad science is clear enough, but the elements are rendered with such detail that they begin to suggest a narrative. But which elements will be elaborated narratively and which simply provide colour? The video does not let on. Videos generally seem unable to mark images as important or unimportant in the ways that film can.¹¹ The laboratory shots, which one might expect to return, are not really taken up. The worm never reappears, but it is echoed by images of boot laces and metal cables twisted around a microphone stand.¹² The remainder of the video never seems to take stock of this opening, but rather moves forward to present three



Figure 11
Marilyn Manson,
'Beautiful People'

13 Often in music video each visual strand develops in isolation. It may be difficult to gauge when a particular strand will reappear or the degree to which it has changed during its absence. Each strand may contain clues that shed light on the performer, the supporting characters, or the general context. By piecing these clues together the viewer will gradually build a composite, if self-contradictory, image of who the characters are and what they tend to do.

The visual strands in music video can bear some similarity to those in popular music. Popular music contains distinct sections, each of which develops in a manner different from that of classical Hollywood narrative. Towards the end of each song section the music may thicken intensify, or simply be used up – one cannot really say that a particular song section possesses a teleology. Though a verse chorus or bridge can be highly differentiated, it can also share materials and one might be tempted to say that some cross-fertilization has taken place as the song unfolds. Similarly each visual strand in a music video remains distinct from the other changes over time (though rarely in a narrative fashion) and becomes affected to some degree by the other strands. It is common in music video for one strand to take on a particular patch of colour, prop or disposition from another. At some point both strands may return to their original identity yet by this point the viewer will have the sense that some sort of process has taken place and that the video is ready to draw to a close.

types of material: shots of the singer; images of dancers on stilts; and shots of the other bandmembers, a crowd, or people with prostheses, all of whom are shown to serve as experimental subjects. These strands proceed unpredictably – we do not know which will appear when, nor whether a given appearance will provide any new information. The viewer will realize at the video's close that 'Beautiful People' concerns process: Manson grows creatures out of stolen body parts. Yet, the clues are scattered across different domains and embedded in a variety of locations.¹³

Claims that videos lack coherence centre on wildly disparate juxtapositions and abrupt changes of style or production values. As 'Beautiful People' suggests, however, the connection between shots is sometimes clear, sometimes obscure, and many of the most interesting juxtapositions lie in between: we can have a vague understanding of a connection, but be unable to specify its nature. Metallica's 'Unforgiven' provides another example of a video that borders on incoherence precisely because, at some moments, the shooting and editing work as they do in film. It takes place in an abstract space with richly textured surfaces of sooty black and grey. A little boy and an old man perform repetitive tasks which seem impossible to complete. The band is set off from these characters, although the tonality of their surroundings remains consistent with the rest of the video's settings. Separation shots of a traditional sort imply a relationship between figures, but the video provides no way to determine which is the best of several possibilities – are they grandfather and grandson, allegorical figures of youth and old age, or do they represent one subject as child and adult? The matter is complicated further by separation shots of the singer and the old man. (Figure 12) In a film, these shots would presuppose a relationship between the two figures. We would not expect such a connection between musician and non-musician in a video that isolates these figures spatially, but the style of the shooting and editing almost demands that we imagine one. What is remarkable about these two shots and the edit between them is that each exists within a separate discourse. The shot of the lead guitarist belongs to the language of documentary, the shot of the old man to allegorical painting, while the edit between them derives from the realm of narrative filmmaking. When one recognizes these three elements, one becomes aware of an unbreachable rift among different modes of expression.

That 'Unforgiven' compels a viewer to pose these questions explicitly already marks it as different from Hollywood Film. Very few videos allow these conventions to perform their traditional functions unnoticed, and seldom present two adjacent shots that resemble paired shots in Hollywood cinema: we are unlikely to see two characters gaze at each other so that the sight lines match, each character takes up the same amount of space in the frame, and we

Figure 12
Metallica, 'Unforgiven'



can identify both and understand what their gestures mean. When this happens, the viewer may feel a shock of recognition. She may think, 'This feels like a film!'. The same is true for paired shots that carry clear narrative implications in film. Imagine a suspenseful sequence in which one shot shows a protagonist approaching a door and the following shot shows the door from the protagonist's perspective. If we were to see this sequence in a music video, we would know to feel suspense, but we would be so relieved to see something familiar, that we might well experience a sense of increasing rather than decreasing certainty. These moments often work as a pastiche of cinema.

In both 'Beautiful People' and 'Unforgiven', the editing verges on inscrutability. A typical video contains a broad range of connections, with the clear and egregiously unclear connections appearing unpredictably. The particular quality of videos may derive from those juxtapositions in which there is obviously a connection, but from which something is left out. Such juxtapositions represent the middle of the continuum. We see successive shots of people, whom we can identify by type, in a single space, but the people do not acknowledge one another and we cannot determine their relation. Or we might see people in different places and be unable to tell whether they are meant to relate at all. We must often extrapolate from what the shots provide if we are to give meaning to a juxtaposition. The early Russian filmmakers understood that this kind of extrapolation was crucial to cinema and argued forcefully that the editing could actually create a meaning in situations where the shots could not themselves provide one. Kuleshov performed an experiment in which he paired shots of an actor and coffin, an actor and bowl, and an actor and children. The results showed that the meaning created by placing one shot next to another could be that of a proper emotion directed towards an object: these pairs seemed to signify, respectively, mourning the loss of a beloved, yearning for food, and enjoying children at play. In music video, adjacent shots often relate but loosely; when separated by dramatic edits, each image will seem enclosed within its own semantic realm. Even paired shots of figures often withhold something. Such pairs of shots can resemble the images in the Kuleshov experiment. In these cases, the affect of the song provides the context for the image. The music cannot define the meaning of objects, but it can surely suggest the animating desire that characters bear towards objects or others. We read emotions into

the image before us, and, with the help of the song, make connections between this image and others in the video. (Figure 13)



Figure 13

The Tony Rich Project, 'Nobody Knows'

- 14 Andre Bazin argues in *What is Cinema?* (Berkeley, CA: University of California Press, 1971), p. 34, that the use of the graphic match atrophied once sound became used in film. He notes that the shots of women talking and then chickens clucking in Fritz Lang's *Fury* (1936) is a hold-over from silent filmmaking.

Eisenstein gave the word montage a special sense, to signify the way in which two shots edited together could create a new meaning that could not inhere in either shot alone. Eisensteinian montage, like the Kuleshov experiment, is predicated on an absence or incompleteness of meaning, but it establishes connections based on conceptual relationships. In a famous sequence from Eisenstein's *Battleship Potemkin* (1925), a ship's cannon fires, accompanied by shots of a stone lion reclining, crouching and standing upright, signifying the uprising of the proletariat against the Czar. Bazin notes that montage disappeared from cinema when sound arrived, to be replaced by the seamless editing we now take for granted.¹⁴ One can see how the silent image track of music video might lend itself to montage. Montage occurs with some frequency in videos, but the collision rarely creates more than a mildly humorous or clever effect. (If montaged images possessed the force that Eisenstein expected of them, they might detract from the song.) In Don Henley's 'The End of the Innocence', two girls, about sixteen and seven, sit alone in a movie theatre. We can see the projectionist's light behind them, and on the screen we see a shot of a train with people sitting as passengers and the cropping of the shot makes the image look like a strip of film. (Figure 14a) In Nine Inch Nails's 'Closer', the eel stands in for a phallus. (Figure 14b)

Figure 14a

Don Henley, 'The End of the Innocence'

Figure 14b

Nine Inch Nails, 'Closer'



The precarious relation of shot to shot, and the varied bases for this relation, affects a video's larger structures of meaning. The polysemic image track creates expectations it frequently leaves unfulfilled; we do not always know where – or how – a video is going. Films teach us to assume that we gain information as the narrative progresses, that we move steadily closer to revelation. Music videos work within this assumption but play against it by progressing haltingly and unpredictably, and by contradicting what has already been shown. Videos also draw us away from the

narrative by foregrounding other structures (especially formal and musical ones) and fulfilling other responsibilities (as to the star). We will therefore be unable to guess, as the video unfolds, whether a given shot will bear heavily, somewhat or not at all on a video's narrative. If narrative fulfillment does come, it will be at a time and in a form that cannot be predicted. This kind of expectation and interrogation of individual shots can suggest the way that videos build larger structures. The relationships between shots have more varied bases in music video than in film: they can relate not only because they present the same character, object or location but also because they share projection values, a lighting and colour scheme, a sense of scale or the use of a camera position. A single shot gives only an incomplete representation of that feature which makes it stand out, whether a character or a colour scheme. Successive appearances of this feature – even if not contiguous – therefore form a structure of partial revelation in which some questions are answered while new ones are asked. Indeed, the relation between non-contiguous shots linked by a single feature is unexpectedly potent in music video, sometimes closer than the connection between adjacent shots. Larger structures made of half-a-dozen or so shots, irregularly spaced and connected by the way they treat some visual parameter, play an important role in creating continuity. A music video can interlace several of these structures, which, like adjacent shots, may be unified by disposition of the figure, shape, colour, setting, theme and so on. It is important that they are not flat and featureless. They manifest changes of intensity as they unfold. A value may increase and decrease, in the case of colour, light, size, speed, height or depth; if we are considering plot, location or character development, it is the amount and nature of revelation that will vary. Some of these structures contain a shot that can function as a high or low point. The high points sometimes form tears in the musical-visual texture (by jutting out above an established level of intensity). This effect is momentary, and afterwards a viewer may quickly invest attention in some aspect of the song's texture in an attempt to prolong the moment of intensity. This prolongation takes place in the instant after the high point has been reached but before the cut to the next shot. The song can provide something to latch onto – a melodic high point, a sectional division, or the entrance of a new timbre. This musical feature takes on new meaning, in this heightened moment, which it carries into its subsequent occurrences. One may consider the feature's history, and in retrospect, invest it with a special meaning. The structures formed according to principles of graphic match are particularly interesting. One might be tempted to call these irregularly spaced graphic matches 'visual rhymes', since a match interrupted by a group of shots that do not share the feature held in common by the two matched shots, carries a chime of recognition and creates a momentary sense of completion.

Nirvana's 'Come As You Are' connects shots according to graphic similarities. Each of these shots contains an oblong shape with appendages, and, in tandem with the music, suggests a stunting of potential. (Figure 15)

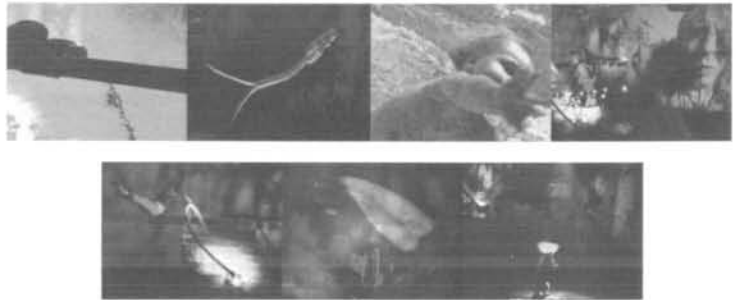


Figure 15
Nirvana, 'Come As You Are'

Editing and the experience of image and sound

I suggested earlier that music video editing exceeds the functions of film editing largely through its responsiveness to musical features – rhythmic, timbral, melodic and formal. I would like now to expand on this suggestion. When early discussions of music video mentioned music-image connections, they tended to notice simple rhythmic correspondences: 'cutting on the beat'. In order to show that the connections between the editing and the music of music videos can be more subtle and more various than this, it will be helpful to take a step back and consider those elements of music video image – crucial, in my view – that reflect the experiential properties of sound. This deep connection between image and song in music video allows for the responsiveness of editing and other visual parameters to musical features.

Theorists such as Edward Branigan, Michel Chion and Walter Ong have reminded us that sound and image possess different properties. Sounds ebb, flow and surround us.¹⁵ The cinematographic features and mise-en-scene of music video – extreme high, low and canted angles, long tracking-shots, unusual camera-pans and tilts, and the lively features within the frame, glittering surfaces, rippling light – can mimic sonic processes. The types of shots used in videos do not just reflect sonic processes, they also suggest a listening subject as much as a viewing one. We actually see figures turning, as if to listen, towards people and objects in the space. The camera's perspective often suffices to imply a listening subject. In order best to see something we might want to be placed squarely before it. If we want to listen attentively to a sound, however, a frontal position is unnecessary. Many positions may be satisfactory – above, below, off to the side. In fact, turning an ear towards the object will take

15 Edward Branigan, 'Sound and epistemology in film', *The Journal of Aesthetics and Art Criticism*, no. 47 (Fall 1989) pp. 312–24. Michel Chion, *Audiovision: Sound on Screen* (New York: Columbia University Press, 1990). Walter Ong, *Orality and Literacy: the Technology of the Word* (London: Methuen, 1985).

our eyes away from it. One of the most common camera positions in music video – below the subject and to one side – may privilege listening over viewing and grant greater authority to the soundtrack than to the image. The camera in music video also seems to mimic the ways that we direct our attention in a sonic space. We can throw our attention to focus on a sound that interests us. When we shift our focus among various sound sources in our environment, we experience a greater sense of mobility than viewing offers. The kinds of shots and editing that we see in music video – jumping from one location to another even before an image catches our eye – resemble what we do when we listen.

The camera normally takes time to explore the extent of a video's setting, so that a setting is only partly revealed in any single shot. How does this practice influence our hearing of a song? A pop song creates a sense of a space through arrangement, production, mixing and mastering. The acoustical properties of this constructed musical space seldom seem to match that of the video's setting. This lack of fit creates some confusion and some interest. How could the song's soundworld inhabit the space of the video? The camera, as it explores the space, suggests possible ways that parts of the arrangement might be distributed within this environment. Many music videos exploit our curiosity about how a song might sound in the actual space of the music video: walls, floors and ceilings are placed at odd angles and covered with materials that imply specific acoustical properties; objects that resemble speakers and baffles may be distributed throughout the space. Despite the fact that the camera never quite reaches the sides or the back of the setting, these videos encourage us to imagine the sound waves rolling into the walls and bouncing off them, much like dye moving through water. (Figure 16)



Figure 16
Usher, 'Make Me Wanna'

If the very walls and furniture can seem to respond to the music, what of the figures we see interacting with it through dance and other, more subtle kinds of movement? Sounds can seem to come through or from them. But how? All gestures in music video – the flick of a wrist, the flickering of light, or the fluttering of fabric – become like dance. We use sound to register the interiority of objects, whether hollow or dense. The way that the camera in music video hovers over the figures, slowly taking in their bodies, may look pornographic, but it might also be a way to register the sounds emanating from these bodily sources. If we think of a singer's voice

as reflecting the rhythms of her body, and the instruments as extending the voice, then the camera can be thought of as creating a fantasy of what lies inside the body – the spring of the muscles, the heartbeat, the flow of blood.

To sell the band and the song, a video most often places the singer front and centre. Some of the time, however, figures begin to turn away from us and show us only crowns of heads, crooks of necks, and elbows. These parts of the body seem to carry as much authority as does the face, and any part that is turned towards us can seem to lead us into the music. Indeed, the figures do not look at each other so much as they turn receptively towards one another as if to listen. The singer remains perpetually in motion, turning sometimes to address us, sometimes towards the figures in the background. The supporting figures may continue to turn towards and away from us, helping thereby to continue the image's rotation. This continuity of motion works to maintain the flow of the image against that of the music.

The description that I have laid out thus far suggests that music video creates an experience more like listening than viewing. As such, it encourages some of the receptiveness and sense of connection that sound creates. Music videos draw us into a playful space where attractive objects are distributed across the visual field. In the absence of a strong narrative, videos have other means to maintain a viewer's engagement. The figures, the camera and the edits each find ways to participate, but they do not always work together harmoniously to achieve this goal: the three often fight amongst themselves for attention, with the song's formal and rhythmic structures as the stable ground. It does not matter, in a sense, whether an elbow comes forward, an edit occurs, a camera tracks, or a figure walks – all are felt as articulations against the music. The bodies of the figures are often the first element to engage us. Music video reveals the body as an enormous but incomplete surface. We may feel tempted to extrapolate beyond the edge of the frame in order to fill in the missing arms and legs. At the same time, the intense focus on a fragment of the body invests it with a special expressive weight. We can imagine feelings and desires – a thinking subject – by watching the rate of release in the shoulders or the spring of the hips. As the video unfolds, we piece together what we have seen to make the body whole. We might remember a longer shot of the body, perhaps torso to feet, a closeup of the head and neck, and a high-angle shot that captures the figure from above. The image also creates associations with the song, matching sections or other musical features with particular visual materials. We see the body bob up and down during the third verse of the song, say, and we might recall the way it moved during the previous verse. As the video progresses, features of the song become associated with elements of the image: a rhythmic motive with the swivel of the

performer's hips, a lilting instrumental melody with a character in the background. By recalling what we have seen and heard, we imagine a phantasmagorical body. (Figure 17) If a video gives us enough material to create a picture of one body, we can attribute moods to other characters in the space who have been rendered more partially, and who often have been chosen because their carriage and gestures are so different from those of the lead performers. As a video progresses, we participate kinaesthetically in the video. We compensate imaginatively for what we do not immediately see in the frame.



Figure 17
Mariah Carey, 'Honey'

The camera functions similarly to create the sense of a consciousness. It is silent and invisible, yet it moves so concertedly – searching, jogging back and forth – that we imagine these movements adding up to something like a narrative voice. As the song unfolds, we can try to guess what the camera is hearing and what it will follow next. The edits balance the camera's movement, keeping things on track. Almost like a downbeat, the edit creates a new beginning. The edits form patterns that the viewer can project into the future. Such patterns are formed by the camera and the body as well. The body, the camera and the editing thus build a kind of momentum that can carry the viewer through a video.

Along with those features that attract and hold our attention, music videos have several ways to keep us at bay. The moving camera and the patterns formed by edits are among the techniques that engage us. We are also engaged by the ever-changing surface of music video, in which a lyric might come to the fore at one point, then a closeup, followed by a striking edit, and then some hook in the music. The song's unfolding and the performers' movements may draw us in. Though music videos rarely contain fully-wrought stories,

they can interest us in a narrative by inciting curiosity about who the characters are and what they might do. Videos distance us in a variety of ways. The borders of the television screen block our entry into the visual space. The figures move obliquely against the music and do not speak; their gestures are abruptly cut off, so one never knows what the next shot will be, nor how the rest of the video will evolve. We may get the sense that the figures are not quite human but not fully emblematic. What animates them seems strange. It is almost like looking at an aquarium. The mechanisms that draw us forward and keep us at a distance exist in constant tension in music video. One may have an urge to follow the unfolding of the image and the music, to enter the space of the music video. One can also feel as though one is locked forever outside, looking in. The body seems restrained somehow, glued to the chair.

Showcasing the star

A focus on editing can help us to understand the relation between music video's star-making dimension and its modes of continuity and signification. Closeups of the star, and the ways they are edited into the flow of a video, provide useful cases to study. The music video closeup possesses its own rhetoric. It has developed a unique look, revealing each wrinkle of the brow and blink of the eyelid as if to capture every emotion crossing the face. Music videos break down visual, lyrical and musical elements to their smallest constituent parts: a prop, a colour, a gesture, a few words, an intriguing riff. In this light, the closeup can be understood to serve specific structural functions. Closeups can leave a viewer with just a face and a moment of the song, unlike actors in narrative film, who bear a past and future that press in on them as we view them in closeup, the music video performer stands in a kind of temporal isolation. As the face fills the frame, it is subjected to so much visual analysis that it seems to move very slowly, almost to suggest the song's slowest rhythmic stratum. This rhythmic effect can serve a grounding function. The closeup of the singer's face is often shot and edited in such a way as to leave us with a single gesture. In its abbreviated simplicity, this gesture suggests a way of grasping hold of some musical element, which might be the main hook or a small detail.

Music videos often present a flow of images that are too rich and materials that seem to dissolve too quickly. The closeup gives us something to commit to memory. The music seems to set certain faces in amber, preserved and just out of reach. The face becomes a mask, drawn into contortions we associate with the most hyperbolic silent screen acting – more an archetype than an expression of the performer himself. This intense isolation keeps the viewer in the present, blocking access to the past or the future as the music rushes

16 Rudolph Arnheim *The Power of the Center* (Berkeley CA: University of California Press, 1988 pp 53–5)

by. The compositional features of the closeup, particularly the relation of the figure to the edges of the screen, contribute to this sense of the figure's being held in isolation. Rudolf Arnheim has taught us to recognize the force of the frame on the composition of a painting. He defines the balancing centre as the point around which the composition organizes itself. It is created by the configuration of vectors issuing from an enclosure such as the frame of a picture.¹⁶ One can liken this force or pressure from the frame upon the picture proper to that of the song upon the closeup of the star. Here the song drives toward the downbeat, the beginnings of phrases and sections, or the tonic chord. When a performer is shown in closeup moving a bit with the music, the music seems to buffet the figure, like rip tides pushing and pulling in different directions. One wonders whether the figure will hold position or be sucked into the centre.

This effect of a push and pull within an unyielding frame makes the closeup precarious. The moments of stability that closeups provide become high points of the video. The video brings us towards these peaks, holds us against them, and then releases us. Only a few moments of the video will provide this much pleasure, and as I, the viewer, reach for them they will be gone. As I watch a video and follow the song, I casually study the performer's body, just as I do when I look at models in magazines. I admire the lines of the jaw, the look in the eye, the light. Suddenly the performer's head turns towards me, the eyes gaze into mine, the singing voice demands my attention, and I am struck. Music can transgress both physical space and the borders of the body, changing our sense of time and of these boundaries themselves. At this moment, the performer crosses the limits of the screen and addresses me as a person, and I can no longer view this face and body as an object. Just as quickly, the head turns, the rhythm changes, the soul has gone, and again I am simply watching a blank human form.¹⁷

17 Earlier I noted that music can seem to leap across frames and link two shots together thereby connecting one character to another. In this example, the music seems to pierce the viewing plane and pour forth to envelop the viewer.

In the absence of a strong narrative, music video creates tension by varying basic visual materials, such as shots and edits. Much of music video editing consists in finding new relationships within space and amongst persons. When a video presents an alternation between shots that display a body and shots that emphasize the space around it, the body becomes the video's ground. One of the most sustained discussions in film theory concerns how a viewer is sutured into the diegesis of film through editing. A series of sightlines and shot/reverse-shots, most commonly, place the viewer in the position of the protagonist or the privileged onlooker. Music video, it seems to me, is much freer in terms of viewer identification and perspective. In the Backstreet Boys' 'Show Me the Meaning of Being Lonely', the viewer's empathy switches from figure to figure simply because someone is within the frame and/or lip synching. Identification occurs quickly – within one or two shots. In the most

- 18 At certain moments in some music videos a parameter's first order function is to highlight musical structure and second to represent if it serves as an articulation it does not matter if we are flashed something illicitly sexual (a breast or a crotch), or something that is not (an elbow or an ankle) When elbow breast, flash of light, hip, edit or buttock accentuate the beat or a timbre in the voice distinctions start to blur as well as the boundary between gender and sexuality Some of music video's intense pleasure may stem from a slide into a pre-Oedipal state one of abundance repetition and polymorphous perversity
- 19 Music video directors often covet particular editors Marcus Nispel admires his editor's special sensitivity to the ways an edit falls on or off the beat perhaps gained from years as a professional drummer Music video editing demands skills not taught in film school
- 20 While some types of shot and edit appear more frequently in some genres than others – a slow-mo low-angle long tracking-shot followed by a dissolve appears most frequently in rap – the language of shots and edits does not differ greatly from one musical style to another Even in country videos (such as Shania Twain's 'You're Still the One'), the rhetoric remains within the same language Neither has editing changed noticeably in music video's short broadcast history While a number of early 1980s videos such as INX's 'What You Want' may not be as densely articulated as some videos of today many of the same editing techniques and strategies are present A study of editing based on genre period or director would be fruitful My analysis of Madonna's 'Cherish' provides a close reading of the editing in a single video Carol Vernallis 'The aesthetics of music video: the relation of music and image in Madonna's "Cherish"' *Popular Music*, vol 17 no 2 (1998) pp 153–87

extreme examples in music video, the viewer's empathy and identification moves between multiple elements in the frame. A series of shots can all contain movement on the beat – a bob of a head, a slap of a wrist, a raising of a glass, a throwing of a ball – and the viewer's attention will seem to skip across the surface like a skimmed stone, following the movements and feelings of everything that moves within the frame¹⁸

Editing and the music of music video

Through its varied roles, editing loosens the representational functions filmed images traditionally perform, opening them up to a sense of polyvalent play The editing thereby places the video's images and the song's formal features in close relation.¹⁹ I doubt the numerous ways that music and image can be put into one-to-one relations would surprise musicians or pop music scholars. Obviously, editing can reflect the basic beat pattern of the song, but it can also be responsive to all of the song's other parameters. For example, long dissolves can complement arrangements that include smooth timbres and long-held tones. A video can use different visual material to offset an important hook or a different cutting rhythm at the beginnings and ends of phrases. And, of course, these effects can switch from one-to-one relationships to something that is more contrapuntal.²⁰

Tempo is one feature readily taken up by music video editing Music videos tend to underscore the most arresting features of a song, and if the song is striking for its sprightly rhythmic feel or its languorous, plodding tempo, the image will often unfold especially quickly or slowly – the image will seem actually to exceed the song's extreme speed Greenday's 'Jaded/Brain Stew' is really two short, connected songs During the slower first part, the video shows a tractor dragging a couch across a landfill in slow motion, along with shots of a dead horse and a sullen old man The performers' lack of engagement enhances the sense of lassitude When the music changes to a faster tempo for the second part of the song, the camera starts whipping around and the pace of the editing increases. The concentration of energy also derives from squeezing the performers into a small room and from using lurid, overheated colours. (Figure 18)

The editing can draw our attention to the general contours of the song's phrase structure Long takes underscore broad melodic phrases, while quick cutting is used to keep us focused on the beat of songs that emphasize smaller rhythmic elements Maxwell's high, pure falsetto floats over the arrangement in his 'Ascension' His singing suggests that he can extend the melodic line for measures on end The video unfolds in a performance setting typical of music



Figure 18
Greenday 'Jaded/ Brain Stew'

video – a stark space with an enormous winged backdrop. The phrasing is reflected in two principal ways. Editing occurs very infrequently, especially while Maxwell sings. The breadth of the melodic line is also matched by the long strides of models in close-fitting metallic suits who walk resolutely towards the camera. (Figure 19a) The vocal hook for Tag Team's 'Whoomp, There It Is', on the other hand, is constructed of short, rhythmic vocal interjections. The camera correspondingly adopts a high-angle point-of-view over a crowd of dancers who vigorously bob up and down. The camera darts in and out over the dancers, while the editing serves to break up the camera movement. (Figure 19b) When the editing diverges from the rhythm of the song, the departure can serve a number of functions. In most pop songs, the beat pattern is omnipresent and easy to follow. When the editing moves from coincidence with the beat pattern to divergence, or vice versa, the effect can be keenly felt. Occasionally, editing off the beat can create a rhythmic counterpoint to the song's beat pattern. Prince's 'Gett Off' presents a high level of rhythmic complexity. The video contains cutting before and behind the beat, which establishes another rhythmic voice and brings out the cross-rhythms created by the figures' movements.

Editing can, of course, carry on two roles simultaneously, like reflecting musical features and shaping the meaning of the video. In Madonna's 'Oh, Father', the verse is sedate, and the editing occurs regularly, separated by long intervals. On the other hand, the chorus, which narrates the story of a child being tormented by her father, is much more tumultuous. The rapid editing occurs sporadically and off the beat, while Madonna's voice cracks, and the drums are 'out of the pocket'. Since the image, alone, in music video cannot narrate a story (figures cannot speak, the form is short, and time and place is rendered incompletely), the parameters must do the work of telling the tale. In this instance, the editing bears much of the brunt of

Figure 19a
Maxwell, 'Ascension'



Figure 19b
Tag Team, 'Whoomp,
There It Is'



describing Madonna's distress, and it also functions musically, underscoring both the jagged quality of her voice and the rhythm arrangement.

In the last example, the editing reflects musical structure and at the same time conveys meaning. But editing can perform even more sophisticated functions. First, I would like to describe the ways that, by emphasizing certain sounds and images, a filmmaker can provide a path through the image. In one scene from Jacques Tati's *Monsieur Hulot's Holiday* (1953), vacationers relax at a resort hotel. In the foreground of the shot, some guests quietly play cards, while in the background, Hulot plays ping-pong. Early in the scene, the guests in the foreground are murmuring softly and Hulot's ping-pong game is louder. The sound encourages us to watch Hulot. As the guests become louder and we hear less of Hulot's ping-pong ball, our attention shifts to the front of the set. Let me now give a similar example from music video. In a video, our attention to the song shapes the way we perceive the image. but, to an equal extent, what we attend to in the image helps to determine how we hear the music. When a star jams his face in front of the camera, or when a hand or foot threatens to break through the viewing plane, we suddenly hear the music in a different way. We become aware that we should pay attention right now. If the same moment in the song were accompanied by a less assertive image – say, a long shot – we would more likely attend to the overall arrangement of the song than focus on any particular element. This experiment can work in reverse, with the music influencing our attention to the image. Imagine a scenario with two types of music. The first contains a city scene, shot in slow motion, with people walking down a busy street:

a medium-shot in slow motion is cropped so that we see the people from their knees to just above their eyebrows. Let us say that the song contains a pounding jungle beat and short synthesizer flurries. We might notice the intensity of the pedestrians' faces or the muscular armature of one or two people. On the other hand, if we hear a flowing synthesizer pad with a minimal rhythm arrangement, perhaps some innocuous 'CD jazz', we might attend instead to the spring and sweep of the bodies in motion, and to the flow of the crowd as a whole. Music videos frequently crop images like the example above, breaking bodies at the joint or rendering them partially, so that more of the context must be supplied by the music than by the image.

In music video, the musical or visual element with the sharpest profile tends to claim the viewer's attention. As a video unfolds, our attention shifts continually among music, image and lyrics, as each provides novelty at some point and then recedes into the background. A deployment of mixed shot sizes, some with very clear content, and some cropped so that they are vague or unspecific, can thus establish a path through the formal and timbral space of the song. The editing can even complicate the matter further by controlling the deployment of shots. Thus, as we move towards a moment of culmination in the song, the editing can tease us with the possibility of spoiling the peak moment's arrival, or feign disinterest by drawing attention to other features of the song. By anticipating what the song will do next, the image can create a sense of expectation. A change of shot sizes can also allow us to circulate within a musical parameter like rhythm or the arrangement. A viewer might first notice the music's smallest rhythmic value and then jump down one level to the basic quarter-note pulse. If one sees a long shot of performers in the background against an ornate curtain or a waterfall, one might attend to the microrhythms of the music. Imagine that the video cuts next to a medium shot in which the singer's face and chest are foregrounded and her head moves side to side, while she crosses and uncrosses her arms as if clocking to the music. The two shots together might encourage such a leap. The image can then serve as a guide to teach us about salient features of the song.

I hope I have provided a glimpse into the world of music video editing. As the video unfolds, the editing can shift rapidly in function, foregrounding musical structure, showcasing the star, reflecting experiential features of the sound, conveying meaning, and even constructing aesthetically pleasing visual strands in its own right. It is helpful, as one watches video, to be attentive to the way that the editing sometimes plays an equal role with other elements – such as colour, narrative, and the treatment of the star – that sometimes vie for attention and sometimes recede into the

background, and also to how editing will play a uniquely superordinate role, functioning as a switcher. Editing controls as it has traditionally served to control the order and duration of shots, and therefore helps to determine when and for how long another parameter will come to the fore.

The investigation of music video editing should be understood as one example of a kind of study that might be performed on any element. Such a study would acknowledge the field on which all elements interact, without forgetting that each has its own cultural history within and beyond music video, its own set of functions (traditional and non-traditional), and its own technical means. A group of these studies would allow us to appreciate music video as a discursive form without imposing a false unity or unjustly privileging one element over the others. When we follow the changing surface of a video, we can try to remember that a momentary effect that claims our attention is part of a structure that traverses the whole of the video, and that this effect is created within the context of that structure: it may mark the high point of some value or constitute a departure from a traditional role. If a video seems discontinuous, it is not because the image track consists of autonomous shots that do not relate to one another, it is because the video interlaces a number of such structures in an unpredictable way. The sheer density of this interlacing provides one of music video's greatest pleasures.

Thanks to the following companies for use of video stills

Arista Records for The Tony Rich Project, 'Nobody Knows', Usher, 'Make Me Wanna'

Interscope for Guns 'N' Roses, 'November Rain', Marilyn Manson, 'Beautiful People'

Life/Bellmark for Tag Team, 'Whoomp, There It Is'

Sony for Aerosmith 'Janie's Got a Gun', Mariah Carey, 'Gimme Your Love' and 'Honey', Michael and Janet Jackson, 'Scream', Maxwell, 'Ascension'

Universal for Peter Gabriel, 'Mercy St', Amy Grant, 'House of Love', Janet Jackson, 'Love Will Never Do (Without You)', Nine Inch Nails, 'Closer', Nirvana, 'Teen Spirit' and 'Come As You Are'

Warner Bros for Green Day, Jaded/Brain Stew', Guns 'N' Roses, 'November Rain', Don Henley, 'The End of the Innocence', Marilyn Manson, 'Beautiful People', Metallica, 'Enter Sandman' and 'Unforgiven', Alanis Morissette, 'You Oughta Know', REM, 'Losing My Religion'