The aesthetics of music video: an analysis of Madonna’s ‘Cherish’

CAROL VERNALLIS

When we become engaged with a music video, what draws us in? What constitutes craft or artistry in the genre? Theorists of music video have usually addressed these questions from the perspective of sociology, film theory or popular cultural studies. Film theory, in particular, has had a tremendous influence on the analysis of music video, because of the two genres’ apparently similar structuring of sound and image. But by the criteria of film, music videos tend to come off as failed narratives; the genre’s effectiveness eludes explanation.

Much has been written about the ways that advertising, film, television sitcoms and newscasts have borrowed from the rhetoric of music video. However, there has yet to be a detailed analysis of any one video, an analysis that can describe how particular moments are set up and departed from and why some moments seem important and others less so.1 This absence of close readings results in part from the difficulties associated with analysing music, particularly popular music. Nor are there adequate theories of how music and image might work together to create a hybrid form. The need for such a theory has been emphasised by a number of theorists, including Alf Björnberg, John Fiske, Dick Hebdige, and Susan McClary. Andrew Goodwin is the most outspoken in his call for a reading that would reflect musical concerns:

In the study of music television a number of major lacunae are evident, but underlying many of them is the neglect of the music itself. This deafening silence in the corridors of the academy combines with an overestimation of the power of the visual to disfigure the study of music television. (1992, p. 2)

As Goodwin suggests, no one has attempted an analysis that takes musical codes, processes, and techniques as providing means by which video image can be structured.

This article attempts to accomplish such goals. It provides both a description of the ways that musical and visual codes operate in a music video, and an in-depth analysis that shows these operations at work in a temporal flow. These two modes, one largely taxonomical and the other more processual, work together to inform us about music video as an artistic practice and as an ideological apparatus. If we attend to the particular features of a single video, we can begin to understand how music video works as a distinct medium. It is by attending to these features – many of which would be called aesthetic features – that we can learn about music video’s modes of representing race, gender, and sexuality. The first section of this
article looks at aspects of ‘Cherish’ in order to develop the analytical tools for reading music video, drawing from music theory and popular music studies. The second section provides a chronological reading of the entire video, as well as close analyses of two particular sections. The final section takes up more fully the video’s representations of race, gender and sexuality.

Section one
The ‘Cherish’ video

As a clip on Madonna’s Immaculate Collection, ‘Cherish’ is one of the most widely available music videos. The video is set on the beach. Madonna lip-syncs from the shoreline, wearing a dark, wet, form-fitting dress. A group of mermen swim further out, in formations reminiscent of Busby Berkeley. A boy/merchild moves between Madonna and the mermen. ‘Cherish’ was seen by critics as designed in part to provide an opportunity for Madonna to display her newly muscle-bound physique; but it was at the same time undeniably arty, in the whimsical myth-making of the mermen, and through being shot in black and white.

The video is directed by photographer Herb Ritts, and reflects the same impulses as Ritts’ still photography; as Allen Ellenzweig (1992, p. 188) says about Ritts’ photographs, the video exaggerates the heroic statuesque, yet also gives a sense of weightlessness and transparency. Ritts might be described as a video-artist with a ‘classical’ impulse: in ‘Cherish’ all musical parameters are reflected in the image with a sense of clarity and balance. It can therefore serve as a model for describing the issues encountered in many music videos. Other videos place greater emphasis on tension and contrast; nevertheless, ‘Cherish’ makes a useful text for showing the nature of correspondence generally, its evenhandedness, notwithstanding. While other videos do not achieve, or even attempt, the balanced structure of ‘Cherish’, they use many of the same techniques.

‘Cherish’ is remarkable for the way that it reflects both local musical features and larger sections. One of the video’s most unusual aspects is the strong, clean contours that it traces across edited images. The clarity of these lines enables Ritts to make the video into a large form which responds to a viewer’s changing experience of the song. Ritts’ classicism can also be discerned in the grace and self-restraint with which he treats both the figure of Madonna and the imagery of gay desire.

Because a music video must – above all – sell the artist and a particular song, the degree of self-restraint demanded of its director can be considerable. A director must usually abandon hope of creating a traditional narrative, even one which the song’s lyrics relate. Moreover, he or she will often find that the pressure exerted by the song prevents the accurate representation of fixed objects: objects in music video will tend to shimmer, change continually, and threaten to fade away. Some directors, including Ritts, have developed strategies better suited to the conventional requirements of music video. What Ritts’ work on ‘Cherish’ suggests, and what is shown by other videos, is that music video image can relinquish qualities traditionally associated with vision and adopt those that resemble the experiential qualities of sound. Walter Ong’s characterisation of the differences between sonic and visual perception can provide a useful basis for comparison:
Sight isolates, sound incorporates. Whereas sight situates the observer outside what he views, at a distance, sound pours into the hearer. Vision dissects . . . When I hear, however, I gather sound simultaneously from every direction at once: I am at the centre of my auditory world, which envelops me, establishing me at a kind of core of sensation and existence . . . By contrast with vision, the dissecting sense, sound is thus a unifying sense . . . The auditory ideal, by contrast, is harmony, a putting together. (1985, p. 32)

Ong’s description of sound reveals perfectly the qualities of the image in a music video like ‘Cherish’. In ‘Cherish’, the image reflects sonic properties through its continuity of motion, most clearly in the imagery of the ocean. The permeability of the water’s surface and the force of the waves set the tone for the video, helping us to notice that the boundary between natural and mythic, or human and animal, has also become permeable in the figure of the boy/merchild. The dynamic nature of sound is also reflected through editing, as shots lose their sense of focus when they move towards the edit point. The inevitable decay of sounds shows in the way that figures move away from us, and in a natural process like the approach of dusk.

**Flow**

The muscular movement of the huge figures in slow motion, almost pulling themselves through the space, along with the waves rushing to and fro, gives the ‘Cherish’ video a particular feel, which might be called a capacity to carry the viewer through the video (Figure 1). This parallels the way that the propulsive elements in the music – the bass line, the rolling drum tracks, the harmonic motion – create and maintain the song’s momentum.

This sense of pull characterises the feel of many videos, and helps to distance the feel of music video from that of most narrative film. David Bordwell (1985, p. 54) argues that narrative films place viewers in a position of mastery. These films, he says, are edited in such a way as to create the illusion that the viewer owns a secure position in the space, from which they can judge the action objectively. In ‘Cherish’, as in many videos, the viewer is drawn through the space by the constant motion within the frame. The searching eye of the camera moves too much to provide the viewer with a stable position. This kind of camera movement exists in a give-and-take with the figures, as they lead us through the space, and with the waves, as they rush forward and back. The song’s groove – the rhythmic figures whose momentum continues across sectional divisions – works with the image’s continuous motion to encourage the viewer to give up their secure position and go along with the ride.

Another way that the image helps to pull us through the video is in the passage from shot to shot. First there is the edit, then a gradual establishment of motion, and then a lunge into a state of right proportion – the perfect photographic moment. Conventional film editing is designed to make the connections between images appear seamless except during periods of crosscutting and accelerated montage. Issues of repetition and variation within edits are played down. In ‘Cherish’, as in many music videos, the edit and the movement within the shot are highlighted for their ability to establish a characteristic rhythm – which in this instance can be bluntly stated as a three-part structure of catch, pull and hold.

There are several rhythmic strata of the music: the slower harmonic rhythm, the basic pulse in the drums, and the quicker tambourine articulations. We can also sense a similar rhythmic stratification in the image: the momentum of the waves;
the movement of the camera – hand-held by Ritts himself; the pace established through editing; the athletic movements of Madonna and the mermen; and the fine visual articulations created through reflected light off sand and surf, the spikiness of the figure’s hair, and the spray and foam from the ocean. These structures support one another, helping the viewer appreciate a level of detail that might well be overlooked in the song. Similarly, the rhythmic organisation of the song gives focus to the image.

Continuity

Traditional theory describes melodies as growing outward by preserving certain features and varying others. Part of the way that a sense of line is created is through
a quality of self-similarity among the materials of a video. Here, some parameters stay constant across a series of shots – for example, the side flank of a merman and a side flank of Madonna; the mermen pop up and Madonna pops up and puts her hands on her head; Madonna sashays back and the fish tail correspondingly slinks back into the water (Figure 2). In traditional Hollywood narrative, the editing techniques work to suggest the viewer’s mastery of the space (through shot/reverse shot, 180-degree rule, eyeline match and point-of-view (Bordwell 1985, pp. 55–7)). Music videos forego such mastery in order to create the sense of a continuous line. The editing attempts to keep the eye moving fluently through the space in a way that supports the directionality of the song.

Contour

The musical lines in a piece of music – the melody, the bass, the inner voices – have contours; composers often talk about these musical lines as visual shapes. In music video, the shape of the musical line can correlate to the shape of the visual image. There may be a few reasons for this. We have, perhaps, a culturally learned disposition to categorise movement by high–low relationships. Register often correlates to a sky and ground orientation. For an example, think of the cartoon figure of Wile E. Coyote falling, as the high pitch drops, and conversely of Orson Welles’ famous opera-house scene in Citizen Kane, in which the voice ‘soars’ higher and higher as the camera moves up towards, and then through, the ceiling of the opera house. (Imagine the opposite effect – the coyote falling and the pitch rising.) We also have
(Refer to caption on page 159)
a disposition to scan photos and paintings, as we read, from left to right. Many
music videos assist in our desire to scan the image by moving from left to right,
including ‘Cherish’, Metallica’s ‘Everywhere I Roam’, and U2’s ‘With or Without
You’. (One rare exception, Metallica’s ‘The Unforgiven’, moves from right to left,
but here movement signifies disintegration and disillusionment.) Spatial and aural
shapes can also correlate to emotional affects.

Musicologists have noted that, within a piece of music, melodic contours
relate closely to the affect we perceive in the music (Meyer 1989, pp. 128–9). Jagged lines produce music that seems anxious and intense. Lines with a narrow
ambitus seem more meditative. The contours of an image have qualities similarly
suggestive of certain affects: tall is courageous; flat or near the ground is safe;
off angle is unstable.3 We respond to imagery and music that work together to
reflect these spatial relationships. At the high points of phrases, hard rock and
heavy metal artists will jump toward the top of the screen and fireworks will
go off. Bon Jovi’s ‘Livin’ on a Prayer’ provides one good example. At the end
of musical sections, the image often seems to darken, to slow down and collapse
into itself. In the first verse of the ‘Cherish’ video, the highest pitch in the voice
(D4) is accompanied by the image of a cresting merman. The verse closes with
a shot of Madonna on the sand, which provides balance, followed by a fade to
black. The undulation of the melodic line is supported by the curved shapes
traced by the figures within the frame. Images of height, depth and balance
correspond, respectively, to high points, low points and moments of stability in
the vocal line (Figure 3). A more precise description of the musical and visual
contours is included in section II.

Form
The video traces many large-scale structures: a gradual shift from day to evening,
an implied maturation of the boy – being born, growing up, and separating from
Madonna; and a tracing of the human body – there are more shots of heads in the
beginning, more torsos and hips in the middle, and more legs and feet towards the
end of the video.

Music videos often sketch a large-scale formal design that matches the large
scale musical structure; for example, in Madonna’s ‘Like A Prayer’ and ‘Open Your
Heart’ videos, the sections of the music continually return in their original form,
and the space can be seen as built upon a spiral. ‘Cherish’ seems more continuous,
less repetitive,4 and the image’s sense of motion from left to right and of continual

Figure 3. Contours – in support of the contours of the primary contrapuntal lines, there are images
of height, depth and balance.
a) Madonna dips as the melody drops a tone and returns. Lyrics: ‘broken hearts’.
b) Madonna bends forward as the melody drops a major third. Lyrics: ‘before I start this dance’.
c) Madonna crouches as the melody drops a fifth at the end of the period. Lyrics: ‘more than just a
romance’.
d) Madonna stands erect as the melody centres itself at a fifth above the opening. Lyrics: ‘You are my
destiny’.
e) Madonna drops her hands to her hips – the listener may hear the drop of a fifth, and perhaps may
even pick out the falling line A–G–F♯–E–D from the instrumental texture. Lyrics: ‘Can’t you see?’
f) The merman crests underwater and the melody reaches its high point. Lyrics: ‘Cupid, please’.
g) The image of Madonna’s head following after the figure of the merman has a quality of balance.
The melody rests on A. Lyrics: ‘Take your aim at me’.
branching outward matches this aspect of the song. In the ‘Cherish’ video, both music and image create large sectional divisions. These sectional divisions can be seen in the shot-by-shot description of the video and the transcription of the song’s lyrics. I have included a transcription of the lyrics and a brief video ‘narrative’, as the analysis often refers to specific points in the song (Appendices A and B). Most importantly, over the course of the video, music and image shift from a state of close interdependence, to a greater degree of freedom, to a return to synchronisation in the closing section.5

Basic shape

Many music theorists argue that a primary musical motive changes continually throughout a piece, providing a key to the piece’s structure (Reti 1978, pp. 13-14). The most obvious example might be the first eight notes from Beethoven’s Fifth Symphony. This is the basic building-block that informs the whole of the first movement. Music video image often works similarly by presenting a recurrent shape. In ‘Cherish’, an arch shape occurs at several different formal/temporal levels6 (Figure 4). There is a secondary emphasis on spiral motion, which functions as a transformation of this basic shape (Figure 5).

There are arching contours in the song, but their perceptibility is hard to ascertain. The melody of the verse forms an arch as it sweeps up an octave and then drops a fourth. The bass line, too, has a wide ambitus. The song’s smallest identifiable group of pitches – a leap followed by a step in the opposite direction – forms an arch, and some of the transformations of this cell – through addition of another step – continue to retain the arch shape. Similarly, one can hear spiral shapes in some of the melodic materials. In any case, the idea of a basic shape remains a useful analytical concept for music video image, even in cases in which the song’s melodic materials do not reflect this shape. Pop songs, of course, do not get written according to the logic of motivic development, and hooks – a related analytical concept that is more useful here – can as easily be rhythmic, timbral, harmonic or verbal as they can be melodic. Connections between visual motives and the song’s hooks, therefore, will often come down to a question of affective resemblance across media.

Motive

In the ‘Cherish’ video, some of the song’s most prominent melodic shapes are linked to visual motives. In the chorus, the melodic gesture F#-G-A-D, which has a quick moving harmonic background, is often associated with Madonna’s assertively taking three steps forward. When, in the verse, the synthesizer rises above the voice to hit a high D, a wave or a hand crowns Madonna’s head. When she then sings a high D, we see her coming forward in a state of suspension, or with a merman swimming above us in an arc. In both the chorus and the verse, a shot of Madonna’s half-moon face is often associated with an ebb in the melodic line. These last examples help to delineate the phrase structure of the music. For example, Madonna’s round face carries an inertial force that slows down the visual material so that it is able to keep in sync with the
music. The repeated image of her half-moon face – most often occurring at a point two-thirds into the verse – becomes a marker that the verse will soon end.

**Phrase**

The image, in ‘Cherish’, forms small sections closely related to the phrases of the music. In the first verse and chorus, the image parallels the music by beginning with a strong articulation and ending with a movement toward a closing lilt that leads us into the next phrase. In ‘Cherish’, the most clearly articulated moments are those that serve to emphasise the beginnings and endings of musical sections.
At the opening of every section, the image begins with a sharp visual attack. As the energy of the musical section is used up, so is that of the image. Fade-outs help further to close sections (Figure 6).

**Lyrics**

Lyrics, in ‘Cherish’ as in most videos, provide only one among many kinds of material to attend to. As might be expected, the hook-line or -word in ‘Cherish’ is strongly underscored – the imagery is focused on expressing an imperative (‘cherish!’). The pseudo-high diction that marks much of the song (in phrases like ‘I’ll perish the thought’), along with the references to Cupid and to Romeo and
Figure 6. (a) The beginnings of phrases and sections (b) The close of a phrase and section (c) Internal articulation of a phrase

Juliet, contribute to both the nostalgic character of the music and the whimsical myth-making of the video. Lyrics that have a strongly performative or theatrical dimension in Madonna’s vocal performance (‘Who? You!’) receive confirmation of this quality in her deliberately stagey turn towards the camera followed by her and the merboy’s playful mouthing of the word ‘you’! One might also suggest that the line ‘can’t get away, I won’t let you’ might have encouraged the director to let the boy make a run for it.

Direct word painting – we hear the word ‘bell’ and see the image of a bell – is less common, however. More frequent is a linking of image, word and music that is more tenuous and enigmatic. For example, there may be a lag between the delivery of a lyric and the appearance of a corresponding image. There may also be a confusion over what the lyrics point to. For example, the phrase ‘Cupid, please take your aim at me’, in the context of the video’s opening – the shot of Madonna’s hips, followed by an undulating merman – might be read as a wish for closer connections among the figures, or even, in the larger context of the video, for impregnation. This reading might inspire the viewer to reconsider all the lyrics, the image, and the music that has come before.

**Timbre and texture**

Frequently, the sound world of the song provides stronger associations for the video-maker than do the lyrics, and can exert a greater influence upon the video. In ‘Cherish’, though no lyrics refer to mermen or to water, the snare and tambourine,
the voice, the multi-layered chorus accompanied by its glittery synthesized timbres, the sibilants ‘ch’ and ‘sh’ all have a white noise component that we might associate with foaming, rushing water and the prismatic light reflected off the ocean and the sand. Each character corresponds by association to one of the song’s countermelodies, and these correspondences – built upon a familiar connection between physical size and registral placement – reflect gender and power stereotypes. The mermen are associated with a synthesized saxophone in its lower register, Madonna with a sassy synthesized trumpet in the middle register, and the merboy with a slightly comic synthesizer patch in the high register. Later in the analysis, we will see that the video questions the primacy of Madonna relative to the mermen and the boy – on this hierarchical plane, all of the figures have an opportunity to move into the foreground.

The bass line, with the bass drum, gives the image its impetus for movement. In a successful music video, the video can trick the viewer into believing that the image and the music are so closely intertwined that the image is spurred on by the propulsiveness of the sound, or conversely, that the image sparks activity in the music. In ‘Cherish’, what animates the figures is mysterious. We almost never see the fins and feet as they animate the figures. Instead, it seems that it is the bass line, as it moves toward and then over the crest of its line, that propels the images onward. When the bass line takes on a different type of musical authority, we have a different type of motion. For example, in the bridge, when the bass arpeggiates two root-position triads, Madonna skips sideways through the water. When the bass line drops out – at the end of Bridge 1, there is a long, held chord without the bass – the figures seem to lack momentum: here, there is the sustained image of a dragging fish-tail. Similarly, at the end of Bridge 2, when there is a repeated chordal guitar riff, with a reduced accompaniment, Madonna slinks back into the water.

Other, more tenuous, sound-image connections include the imagery of the clinging wet cloth on Madonna’s body, the snugly fitting costumes on the mermen, and the water flowing against Madonna’s flesh and her dress, any or all of which might stand in for the synthesizer line that wraps around and closely tracks the voice. By contrast, the finer rhythmic articulations might be carried by the spikiness of Madonna’s hair and the fine detail of the foaming water.

The song is mixed in such a way that none of the backing tracks are brought to the fore. The prominent bass line commands more attention because it moves more actively, not because of any emphasis in the mix; the goal of the production must have been to create a wall of sound that places all the backing tracks on an equal plane. To complement this approach, the images of Madonna, the mermen, and the child are correspondingly huge and engulfing. Nevertheless, the arrangement varies as percussion, bass and backing vocals are added or drop out, and, more subtly, as the reverberation effects on Madonna’s voice change through the song. Similarly, the video makes an issue of how Madonna and the mermen move closer to or further away from us within our field of vision. For example, when she sings the words ‘who’ and ‘you’, her voice comes to the front of the mix and she advances towards the viewer. The beginning and end of the song are produced to sound distant, and it follows that the figures appear small and move away from us.

These local sound-image connections play a large part in shaping the video. What I can only call the nostalgic quality of the video originates in various aspects of the song. Besides the archaic touches of the lyrics, there is the swing rhythm
An analysis of Madonna’s ‘Cherish’

(Perhaps always nostalgic), the finger snaps, Madonna’s chirpy, child-like voice, the girl-group backing vocals, and the hook-line taken from the Association’s 1966 hit ‘Cherish’, all of which allude to earlier pop music. It is also important to recognise that the song’s arrangement scheme, production values and performance styles had been in place for at least five years before the release of ‘Cherish’ in the summer of 1989. These conservative qualities, along with the song’s pop acumen, help to bring its retrospective dimension to the fore.

Harmony

The song’s harmonic language emphasises smooth, sometimes almost elliptical motion, using a large number of first- and second-inversion chords, and moving through the bVII and the II as it wends its way to the subdominant or dominant. This aspect of the song, too, shows the balance and restraint that bring out what I have termed Ritts’ classicism. Yet within this narrow range, the video makes a subtle response to harmonic changes. For example, in the verses, the beginning and end rest on or near the tonic, D. Here we see stable images of Madonna, or the mermen and Madonna, as if these figures stood for arrival points, or places of stability. Within the verse, as the harmony begins to shift away from and then towards the tonic, Madonna starts to turn away from us and shift her focus towards the ocean.

The chorus is more active than the verses in this song: here, the harmonic rhythm moves at a faster rate than in the verse, the harmony oscillates between the subdominant and the dominant without much tonic, the orchestration opens slightly to include a clavinet and guitar, the melodic sequences become shorter, and the drums accentuate the pulse – many of these features contribute to a more march-like or anthemic quality to the chorus. (Although the song is in a shuffle groove, the chorus has a more ‘square’ feel, because of the crotchet harmonic rhythm.) The images in the chorus respond to these features with a more lively set of paired two-bar phrases. Unlike the verses, in which beginnings and ends of sections regularly feature Madonna, the chorus creates a sense of uncertainty as to whether Madonna or the mermen will appear first – sometimes Madonna takes up the first half of a sequence, and sometimes the mermen do.

At the bridge, the harmony implied by the bass line conflicts with that of the keyboard pad, with the bass line really defining the harmony. The motion of the bass line moves towards the dominant, unfolding a III–IV–V progression: F♯–A–C♯, G–B–D, A. In response to this, Madonna steps sideways for the first time in the video, as if drawing us toward something new. When the bridge returns, the bass line reiterates its movement towards V. This time, we do find ourselves in ‘another region’, with imagery of Madonna and the little boy beached on the sand.

Much of what we know about the disposition of the figures in ‘Cherish’ is defined through the harmony. For example, the child commonly appears at the same time as the subdominant in second inversion. This chord is pulled in two directions: it is often subsumed by the tonic, while it also resists the dominant. This harmonic pull might contribute to the way that the child is volleyed back and forth between the mermen and Madonna. The mermen tend to land on the relatively stable chords of I, IV and V, and the way they seem to float suspended in space, performing a slight acrobatic twist, suggests their freedom from the influence of the song’s harmonic motion. Madonna’s appearances, on the other hand, coincide with
passing chords; these correspondences make her seem somewhat hemmed-in by the harmony. In each section of the video, she forges a winding path towards us, yet monumental harmonic forces towards the end of the section pull her back towards the water’s edge. (The most pronounced instances occur in the first and second verses.) Later in the video, Madonna breaks free of these constraints, affecting the video as a whole.

Rhythm

‘Cherish’ has a strong triplet feel, but the drums count a standard pattern with the bass drum accentuating beats one and three, and the snare, beats two and four. At important points in the video, the editing and the broad physical gestures within the shot (like the shrugging of shoulders) fall squarely within the beat pattern, emphasising, for example, beats one and three, one and four, or simply one. This accentuation of the beat pulls us back in whenever we start to drift along with the flowing imagery. Perhaps to compensate for the absence of a live rhythm section, the mostly sequenced rhythm arrangement fills in nearly every quaver in a number of instrumental parts (Example 1). This rolling pattern might have helped to suggest the carriage of the characters, the setting on the ocean and the figure of the mermen. Against these more general rhythmic qualities, each section of the video displays its own particular pattern. In verse 2, the move towards the downbeat of the measure carries the metaphor of wave and flow into another domain.

Section two

A close reading of two sections

I will now look closely at two particular sections of the video. In music video, the images frequently divide into sections according to the song’s sectional divisions. The image can highlight differences among musical sections through shifts in colour, pacing, gesture, or topic. Often the image exhibits a character particular to
a section of the music. In ‘Cherish’, as in other videos, the image for the verse features constrained or restrained imagery, typically a solitary artist in a barren landscape. The imagery of the bridge often points away from that of the rest of the video. The chorus, on the other hand, can usually be characterised as communal, and will be set with imagery of freedom (running, jumping, flying), of fusion (between a couple or among a group), of plenitude (crowds, a large group of objects), of paradise (fields of grass).

I will first discuss the second verse. As I have mentioned, the chorus is more march-like, more public and extroverted, while the verse is more reflective. We have already seen the first verse. The image traced the winding and sinuous quality of the line, closely following local inflections, and the high D4 was matched by the image of a merman seeming to soar under water. Now, in verse two, the image encourages us to listen no longer to small articulations, but to hear one broad line.13

In verse 2, the image focuses on a larger form – the overall contour of the melodic line (Figures 7 and 8). This focus is established through a use of space similar to that of dance. Dance theorist Rudolf von Laban (1974, p. 10) speaks of a ‘kinosphere’ of the body: a centre of gravity and an implied larger sphere that the limbs can move into. In music video, one can speak of a larger kinosphere established through camera placement and disposition of the figures – a series of shots creates the illusion of a centre and defines a space around it. In verse 2, a point in the water is fixed as the centre. As we begin to move in different directions away from this centre, we feel the breadth of the melodic line. Yet, simultaneously, the image also plays with our sense of time. Since we have already heard the first verse,
we know the length of the second verse. In the second verse, the implied visual centre has been established, and soon, there seems to be nowhere to go. Yet the verse is not over. The voice continues searching for the higher register, which leaves us with a sense of constriction.

Our anxiety breaks as we see the image of Madonna crossing a boundary. She rises out of the water and moves toward us, like the first amphibian emerging from the ocean to walk upon land. As Madonna continues to approach, the music hangs on the tonic before finally moving to the dominant. Both image and music continue to linger, raising the question as to whether the music or the image will spark the next section of the video. Music video often works by scrambling effects and causes, making it hard to remember that the song is ontologically prior to the image. A characteristic power-relation is thus established, in which nothing is taken for granted, and in which each medium must work with and against the other. The possibility exists that either song or image will dominate, forcing the other’s hand.

The next example comes from the third verse, at the moment when the mermen swim underwater in formation. A long visual and musical slowing takes us into this section. This is the only moment in which the music departs from the structure of the conventional pop song. Interrupting the synthesizer solo and vocal ad-libs, the second half of this verse suddenly returns to the arrangement scheme and vocal melody of the previous verses, as if it had been a portion of another verse all along. Because this fragment seems both isolated and out of context, we will recognise the melody as familiar, but we may be unable to remember exactly where it comes from and why it has suddenly reappeared. Here the image assists in the process of reconstruction, as it draws from and condenses visual material from the earlier verses. The imagery in this sequence combines material from the culmination of verse 1 (the image of the merman swimming underwater) with the spatial arrangement at the close of verse 2 (the kinosphere established by Madonna and the mermen). Here, the imagery with the music recalls the earlier half of the video, and we experience a moment of very potent recollection (Figures 7 and 8).

Herb Ritts is known for his homoerotic photography, and Madonna for her close ties to the gay community. Hinted at throughout the video is the possibility that some desirable being resides in the depths of the water. We assume, as Madonna continually backs into the ocean, that it is she who is our object of desire. Yet, in this musically isolated fragment, veiled by the murky water, Ritts is able to express the possibility that our primary interest in the depths includes an erotic attraction to the mermen.

<table>
<thead>
<tr>
<th>verse 1</th>
<th>chorus</th>
<th>verse 2</th>
<th>chorus 2 – bridge 1</th>
<th>verse fragment</th>
<th>bridge 2 – outchorus</th>
</tr>
</thead>
<tbody>
<tr>
<td>merman</td>
<td>kinosphere:</td>
<td>figures</td>
<td>exchange</td>
<td>mermen</td>
<td>swimming</td>
</tr>
<tr>
<td>swimming</td>
<td>moving</td>
<td>within a</td>
<td>kinosphere</td>
<td>mermen</td>
<td>swimming</td>
</tr>
<tr>
<td>underwater</td>
<td>glances</td>
<td>visual material</td>
<td>from verse 1,</td>
<td>sense of space</td>
<td>from verse 2</td>
</tr>
<tr>
<td>visual material</td>
<td>sense of space</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 8.*
As I have suggested, music video suppresses some important facets of the music in order to draw our attention towards others. A beautiful moment in the song occurs eight bars earlier where Madonna steps back into the water, and the fish-tail is dragged along the water’s edge. The long-held IV chord serves as an extended suspension to the tonic. In the video, the held IV is subsumed under a long visual ritard that leads into the close of the third verse. This ritard is established through a descending contour in both the image and the music. Our attention sweeps across the musical texture: the image of the child reflects the high-pitched synthesizer melody; the image of Madonna holding the child connects with Madonna’s voice in the mid-range; and the image of the watery depths and the merman placed at the bottom of the frame is linked to the low register of the synthesized bass.

In the extended sequence which features the mermen, their visibility is obscured, and one might think that their power is therefore muted. Yet, in a large part of the gay community, the politics of concealment and revelation is a complicated, richly inflected phenomenon. Like Michelangelo’s *The Holy Family*, in which the artist may have chosen to paint a group of nude men into the recesses of the painting for his own desire and enjoyment, Ritts may have partially obscured his mermen for personally defined erotic reasons. Gay colleagues of mine talk about the titillation of seeing adventure or nature films of foreign lands, because one knows that one might see gorgeous naked men in the background – the recently filmed version of *Last of the Mohicans* is a case in point. For those who are willing to take time with the ‘Cherish’ video, the obscurity of the mermen works to good effect. Viewers who watch the video only a few times, and who have little contact with gay culture, may lose an opportunity.

I have discussed these two sections in some detail, before moving on to a chronological reading of the entire video, partly because they have an internal coherence that makes them worth studying independently. More important, these sections form the locus of certain themes that I will bring out further in the final third of this article, particularly themes of gay desire. I will return to the end of verse 3, which features underwater footage of the mermen, as it constitutes a nodal point in the video.

*A chronological reading*

A central claim of this article has been that music video image creates its meanings within the flow of the song. The clarity and stability of these meanings remain subject to the song’s temporal unfolding. I will give an abbreviated account of the overall flow of the video, picking up what has been left out thus far, and pointing to the aspects of the video that will be discussed in the sections of the article that follow. A chronological account of the video will prove helpful as a way of describing the relations among sections and the progress of many long-term processes. Because this chronological reading differs from most close analyses of film, it will be worth explaining quickly how it is organised. Close analyses of film usually use the shot or the scene as the fundamental unit of analysis, and close readings of music video have tended to do the same. Here, the method of detailed, sequential description is similar, but the fundamental unit is the musical section, rather than the scene or the shot. The use of the musical section as the fundamental unit places an emphasis upon varied repetition of materials over linear development of plot.
Treating the form of the song as the analytical ground for the video better reflects its semantic and formal structure.

Introduction

We are frequently led into the world of a music video across a threshold or through a liminal state. 'Cherish' achieves this effect through a shot of the child running along the shoreline, with sounds of the ocean, before the song begins. The somewhat dream-like character of this shot confirms its liminal function. In this shot, it is not clear whether the child is running away from or towards Madonna's breast, and whether he imagines Madonna, or she imagines him. Similarly, the sound of the ocean may seem to suggest a lack in the song, and vice versa. When the song begins, Madonna's repeated 'cherish' sounds like an echo or a call heard across a great distance, either in space or in consciousness.

In the opening of the video, the disposition of the figures seems to conflict with the contour of the bass line and the voice. The figures move laterally. The music here, on the other hand, is made up of falling gestures in the synthesizer and bass. The weight of this introductory musical material is not taken up by the image, but rather is deferred or carried over into the first verse. The image of Madonna's gradual descent towards the sand in verse 1 is suggested by the dramatic falling gestures of the song's introduction.

Verse 1

The outer voices move in contrary motion, first outward, and then inward. As a complement to this motion, the video presents images of increasing height (Madonna's head, the crest of the wave, and, as the high point, the merman cresting at the surface of the water) interspersed with images of increasing depth (Madonna crouching and rolling on the sand). The images that follow reflect the balance between the bass and the melodic lines above it, especially the voice (Madonna turning with her arms spread out, and the paired images of a merman soaring and Madonna's head on the sand).

Chorus 1

The chorus begins with one of the simplest and most functional relations in music video: the arrangement thickens as the frame fills with more figures. The music in the chorus more clearly articulates the crotchet pulse, and the metallic synthesizer doubles the voice, with the background vocals further emphasising the punctuated style of Madonna's vocal delivery. The melody in the chorus is made up of shorter, more regular phrases than is the verse. Chorus 1 loses some of its iterative, terse quality as it moves into verse 2.

Verse 2

As mentioned earlier, in verse 2 the imagery has a quality of line and extension.

Chorus 2

This section distills earlier aspects of the song. As verse 2 intensifies and simplifies the linear quality of verse 1, so does chorus 2 intensify and simplify the
clipped, march-like character of chorus 1. Images of upward motion are even more prevalent here than in chorus 1. The emphasis on the breadth of the figures and their high placement in the frame helps to create a visual arch, as it were, that spans the length of the tape. Chorus 2 contains imagery that foreshadows what we will see later in the video. For example, the image of the child held in the merman’s outstretched arms, and the mermen’s and Madonna’s broad shoulders prepare for Madonna’s movement from left to right in the first bridge. The image of Madonna placing a frond of seaweed on her shoulder (a makeshift feather boa), and of her reaching towards her ankles, prepares for the ‘vamp’ section in the second bridge. In retrospect, we will be able to look back upon the first bridge as a section in which energy is concentrated. This section returns to the questions of identity hinted at in Chorus 1 (who is it that we see, Madonna or the mermen?), but here these questions become more pointedly about difference (what do you have, fins or legs?). The issue of sexual or gender difference will be revisited in bridge 2, when Madonna and the merboy playfully flex their muscles. The imagery of three mermen moving into the frame after the end of the second chorus preceded by a lengthy fade to black can be seen as a culmination of a section.

Bridge 1

Musically, the bridge is quirky. It contains a new element, a varied form of the ‘Cherish’ motive from the introduction: G–F♯–D (bar 1) becomes F♯–G–B. It also combines musical fragments drawn from both the verse and the chorus. In response to this construction, the first half, like the chorus, reflects a celebratory ‘surf’s up’ attitude. A closer reading of this section shows that it goes far in creating the video’s clear sense of nostalgia. The lead synthesizer patch seems rather kitschy and outdated, and, in tandem with the image, may conjure up memories of old beach movies. The placement and movement of the figures point to early – and inexpensive – modes of visual entertainment, like the spinning top, the carousel, the zoetrope, and Disneyland’s (now dismantled) ‘Circle of Progress’: here, the mermen and Madonna move from left to right along the seashore, simultaneously but at different rates, and it seems as if someone were dragging a painted backdrop behind Madonna while she moves towards us and then away from us.

Verse 3 (with instrumental and vocal ad-libs)

This verse contains an interesting formal conceit. Its first half is an instrumental over the rhythm arrangement of the verse (which is typical for an instrumental), with the synthesizer taking the solo. This synthesizer line, however, is dominated by vocal ad-libs which Madonna sings over it (and which return counterpointed to vocal material from the chorus in the outchorus). More strikingly, the second half of this verse abruptly returns to the vocal melody of the previous verses, abandoning the synthesizer lead and the vocal ad-libs and proceeding as a ‘regular’ verse. The first half of verse 3 is composed of music and images that slow down gradually, and that move towards the second half of the verse (when the vocal line from the earlier verses returns). The visual material comes from the more restrained, tender imagery of the first two verses. The second half of this section features the
mermen swimming in formation. The underwater choreography echoes Busby Berkeley's famous waterfall sequence in the movie *Footlight Parade.*

**Bridge 2**

In the second bridge, a varied repetition of bridge 1, Madonna and the merboy loll on the beach. It is the only section that makes a feature of being edited off the beat. Because it is edited in this way, this section seems rhythmically out of step with the rest of the video. The fact that it takes place on the sand, unusual for this video, further helps to set it apart. The merboy shows off his muscles; Madonna picks up on this gesture in the break, when she adopts body-builders' poses. A two-shot of Madonna's feet and the merboy's tail foreshadows the set of oppositions that end the video. At the end of this bridge, Madonna backs into the water.

**Break**

Madonna vamps for the camera in a section that reminds one of old home movies. The break does more than create a nostalgic flavour, however. For all of Madonna's activity, this section has a strange sense of emptiness. The propulsive elements of the music – the bass and the snare and bass drums – have dropped out. The space behind Madonna is open and still. At first, it is not clear, in the quick editing, whether Madonna is moving to the music, or whether the music is dragging her around as if she were a marionette. The attempts at close synchronisation here sharply contrast with the rhythmic freedom of the first bridge. Throughout the video, Madonna has seemed constrained by the harmony. However, by the end of the break, she is depicted as no longer having to struggle with the harmonic structure, but, like Atlas himself, as being capable of upholding it.

The way that the harmony is realised in the 'Cherish' video relates to the way that liberal corporate culture tends to picture social obligations. We speak of the 'smaller, more intimate domestic sphere' and the 'glass ceiling of the workplace', and we can notice the way that the harmonic structure constrains Madonna. In this video, both harmony and social obligations are figured as reassuring, but also palpably constraining, sets of boundaries. Because they are treated similarly, one could be considered as a figure for the other. By the close of the break, however, the video reveals something new: we begin to know more about the mermen, the boy and Madonna, and what they aspire to.

In the break, Madonna adopts a triumphant stance. For the first time, she steps assertively into the emphatic 'ch!' of 'Cherish'. Her squared-off arms and torso condense the kinospheres of verse 2 and the second half of verse 3 into a single form. As she adopts a bodybuilder's pose, showcasing her biceps, a second vocal line, drawn from the first half of verse 3, enters in counterpoint to the vocal line of the chorus. This second line, made up of shorter phrases and fragments from earlier in the song, has an insistent quality that lends Madonna additional authority.

There is a shift across the video in the representations of the characters. For the first half of the video, the mermen move assertively, propelling themselves out of the water. In the second half of verse 3, however, they move with restraint and poise. Here the break, and the sections that follow, are dominated by Madonna's
compact, precise gestures. There are shifts in the arrangement which correspond to these changes of character: in the second half of verse three, the nearly static synthesizer pad becomes the most prominent instrumental voice, thereby underscoring the mermen’s timeless existence. In the break, the bright trumpet patch cuts through the texture, emphasising Madonna’s assertiveness. By giving the mermen and Madonna these actions at these particular points in the song, Ritts can make use of the cultural associations of the song’s timbres in order to play with traditional notions of gender.

For the first half of the break, the image emphasises episodic, rather than teleological, features of the song. A strength of this approach lies in its not conferring greater authority upon either Madonna or the mermen. In the second half of the break, the snare and bass drum return, and the backing vocals from the chorus enter behind Madonna’s ad libs, anticipating the outchorus.

**Outchorus**

By the time that the outchorus appears, the video has established a variety of connections between music and image. We can therefore hover over the image, waiting for a moment of engagement between it and the music. At this point, we may be waiting for the grounding images of feet to return, as well as for the completion of a spiral that is traced out across many edits. The second vocal line seems to match the prismatic late afternoon light on the ocean, through both its kaleidoscopic construction and its use of echoes from earlier in the song. The return of the synthesizer lead (the ‘Steve Winwood’ patch) works likewise to a similar effect. Typically for an outchorus, this section brings in all of the timbres we have heard in the song. The image, for its part, creates the sense of an ending through similarly conventional means: the mermen ride off into the sunset, Madonna and the boy share a final embrace before he runs away and she lies down on the sand.

The video closes with Madonna and the child facing off against each other – the image of the child seems, strangely, to be composed of adult legs and a child’s head. Popular culture theorist Philip Hayward (1991, p. 98) reads this image as a sign of patriarchal authority. He argues that Madonna’s freedom is reined in at the end of the video. The legs, torso and heads, however, function most importantly on a formal level. As a pair of bookends, the images serve to stop the flow of the video, to slow down the image so that it can close in sync with the music. The merman’s legs still remain topped by a child’s head (so that the phallic presence is not that impressive) and Madonna assumes a siren’s pose, a mythic image of great power. The boy’s legs and Madonna’s torso come at the moment that the song is over, and function as a frame for the song. Madonna lands on beat one, a position of power within the song.

The ending works through a compression of thematic oppositions. All of the video’s dualistic imagery – adult and child, female and male, human and animal, myth and camp – is squeezed into the closing shots. This kind of an ending is common in music video. It achieves its effect partly by providing a thematic payoff in the absence of any conclusion to the narrative. The enigmatic character of the final series of shots asks the viewer to return to the beginning and watch one more time, in order to see how the video could have arrived at this ending.
Section three

Representations of race, gender and sexuality

This article, thus far, has focused on music–image relations that would be called formal. I have argued that the logic of these relations establishes a ground for an investigation of race, gender and sexuality. Here, I will explore these issues, beginning with one example that shows the interdependence of formal and depictive modes. In the remainder of this final section, I will discuss other parameters – narrative and emblematic modes, affect and self reflexivity, and musical and visual space. This discussion forms the necessary complement to the explication of formal modes of continuity.

The video for ‘Cherish’ contains three kinds of figure: Madonna; a group of mermen; and a boy/merchild who moves between Madonna and the mermen. Madonna and the mermen are clearly European–American, while the boy is clearly not – although his specific ethnicity is unknown, he is pointedly ‘other’ to the rest of the figures in the video. One might well wonder about the boy’s ethnicity and its role in ‘Cherish’. The imagery of a young child of colour being passed between Madonna and the mermen is not innocent when one considers that the depiction of adult black men and women passing around a single white child is rarely seen in popular culture. In the current critical climate, it might be acceptable to note this as an instance of aestheticised racism, and terminate the analysis there.

Yet, on a more formal level, a darker figure as an object of exchange creates a strong degree of separation between the white figures of Madonna and the mermen. The video is shot entirely in black and white, and, because of the way that tones are developed gradually over the course of the video – dark black dress, black roots, a merman in the water with very black hair and white skin as opposed to Madonna’s blond hair and black dress, grey water, medium grey and white mermen, and a very white-skinned Madonna – the young boy of colour is almost argued for on formal grounds, as one shade within a grey scale. The young boy’s colour in this context is a transitional tone.

But we cannot stop here: we must acknowledge that this formal device reflects a specific racialist way of viewing skin colour. We could imagine a tape produced in a culture where such differences do not mean the same thing, such as in Brazil, Puerto Rico, or North Africa. The decision to emphasise these differences so strongly relies implicitly on the fact that skin colour, and race, mean a great deal to us as Europeans and North Americans.

Yet it is also important that the boy’s colour in this context is a transitional tone. Imagine the video with everyone of the same skin tone: the figures would lack separation. The image of the child of colour is, at one level, offensive, at another, progressive. The child’s skin-tone subverts the viewer’s projections concerning the nuclear family, allowing for both the imagery of gay desire and the imagery of Madonna’s independence.

This kind of interdependence of aesthetic and ideological aspects shows that, ultimately, we cannot look to any one place for an understanding of music video, but must rather deal with the relations among a video’s narrative, formal, and sociocultural aspects, particularly as they are complicated by the tensions between music and image. The ways that music and image combine cannot simply be taken as natural; styles of performance footage change from year to year, and the construction of a music video always requires effort. Videomakers have developed a
set of practices for putting image to music in which the image must give up its autonomy and abandon some of its representational modes. In exchange, the image gains in flexibility and play, as well as in polyvalence of meaning. Many of the meanings of music video lie in this give-and-take between sound and image, and in the relations among their various modes of continuity.

Questions of narrative

The first wave of academic writing to deal with video concerned itself with music video’s ability (or inability) to sustain a narrative. More recently, however, there have been attempts to put narrative in its proper place, as but one of several ways to establish continuity. Music videos suppress narrative direction for various reasons. The fact that the figures cannot speak and seem preternaturally animated by the music may work against narrative clarity. The brevity of the medium contributes as well. Pop songs, usually, are sectional forms (verse, chorus, bridge), and it is difficult for an image track to maintain a strong narrative drive against this sectional differentiation and repetition. But most important, it may be in a video’s interest to point only vaguely to a narrative. If the image were overly narrative in orientation, we might be drawn to the image as we are in a traditional film. The music for the video would most likely resemble film music – usually unacknowledged, almost unheard.

David Bordwell defines narrative as having 1) an agent with identifiable goals and distinct characteristics, and 2) obstacles to this agent’s success (1985, pp. 12-23). This definition shows immediately why music video might limit the role of narrative. If we were to engage with the figures in music videos as if they were people with clear traits and identifiable goals, who were approaching difficult or dangerous encounters, we might try to predict upcoming events. That kind of engagement would pull us outside of the here and now of the video, its moment-to-moment flow, and we might well lose the detail in both the music and the image. Videos draw heavily, but schematically, from traditional forms – for their familiarity as much as for their novelty.

‘Cherish’ is similar to many videos of a certain type, in which there are moments that suggest a story line. In such videos, these moments function more like ‘hooks’ than like parts of a story. We are carried between these narrative moments by the ultimately more important play of movement and texture. With this understanding of narrative in mind, one can suggest several possible plotlines for the ‘Cherish’ video:

(1) A community gives birth to a child, and turns him over to a woman for instruction about culture; the child is later returned to the nurturing community (this inverts the Lacanian notion of a society in which men teach the ‘law of the father’).
(2) Madonna and the mermen give birth to and raise a male child.
(3) Madonna imagines a love affair with a merman, and transforms one into a child whom she can love and nurture.
(4) The child is Madonna’s son and the mermen are merely imagined.
(5) The child is a member of a family that will not shield him from adult expressions of sexuality.
(6) Madonna steals both power and a child from the gay community.
(7) The child cannot fit into human society: ‘Cherish’ is a kind of ‘coming out’ story, and the mother is supportive.

For these narratives to exist, the viewer must infer or find them by
‘connecting the dots’ between particular charged moments in the video, and must pay close attention to the music and the image between these events. Because no parameter comes to the fore to the annihilation of another (although features become submerged or move to the background), multiple storylines can seem to exist simultaneously in the video. The viewer must consider all the visual gestures and all the musical codes in order to understand the connections among these moments. The vagueness of ‘Cherish’ may serve it well. It gives a committed viewer enough space to imagine what might be the relations among the figures.

Space
One way that music video substitutes for a lack in the narrative is through a focus on a set of relations. In music video, the disposition of figures and their movement on a ground often takes the place of plot and character development in a traditional sense. In ‘Cherish’, one simple way that this is played out is in the sense of scale between musical and visual space. The musical ‘field’ of the song seems smaller than that of the image for the video. As I have mentioned, the musical field for ‘Cherish’ seems narrow, even thin – it contains very few dramatic rhythmic, harmonic, melodic or timbral changes. By contrast, the visuals for ‘Cherish’ seem to occupy a broader realm. Though the visual field is filmed entirely in black and white, and is set exclusively on the beach, it has a quality of great expansiveness – the distance from sea to land, from the surface of the water to its depths, or from one point along the shoreline to another, seems quite wide. In this instance, the larger expanse of imagery, against a more narrowly constructed band of music, gives the viewer the sense that the figures are highly individuated yet only loosely bound to one another.

Of all of the figures, the mermen are particularly well defined through the use of space. They possess their own realm, and act as if they have always been and will forever be in the ocean. Some elements of the music, perhaps the sectional changes, seem to become real barriers – like the rock formations in the water – but the mermen seem to communicate across this distance through a secret code. The mermen seem comfortably placed in the musical and visual materials of the video.

The sense of space also helps to define Madonna. An approach that focuses on the imagery without sufficient attention to musical time and space might suggest a deeply sexualised text: Madonna is a siren who lures us into the video with a sexuality that moves from chastity to naughtiness to polymorphous perversity. The viewer may desire the child. Yet the appearance of the fish-tail dragged along the water’s edge provokes a moment of anxiety at the sight of the uncanny. The mermen remain remote, deaf to our interest. At the video’s close, the child realises his own desirability. Madonna drops her social graces and threatens to consume us.

Yet this story leaves out much of what is important in the video. What might seem frightening in the image is safely marked off as separate space – through the clear delineation of ocean and sky, and water and sand. Only when we are underwater do we deal with the presence or absence of male genitalia, or with male sexuality in general. When we are on land, both the little boy and Madonna seem quite comfortable. The membranes of separation – the surface of the water,
the shoreline – are very gently elided. The time and distance that separate the appearance of those visual hooks that carry narrative charge also provide a sense of security.

The music also works hard to repress what might be frightening about the images. Madonna is held within the harmony; similarly, the fish-tail is held within the long sweep of verse 3’s first half; the merboy and Madonna stand as a pair of balanced closing images. These images remain detached from one another. A reading that attends to both music and image will draw attention away from these moments and towards other features such as the process of the figures’ individualization and their relation with one another. Neither the song nor the video is committed to an ending. Both are non-teleological in nature.

This does not mean that the video is free of emotional complexity. Rather, those elements which are more sharply edged appear on another level and are safely submerged. The characters’ actions and expressions, which sometimes seem to reflect a sense of anger or hostility, provide one way that the figures in ‘Cherish’ are given voice. These moments do not determine the video’s tone; rather, they provide contrast in a circumscribed way. Since one knows that they will not overwhelm the video, they can function well as ornament and texture.22

One must not exaggerate, however, the extent to which the characters are in fact given voice. Of all of the figures in ‘Cherish’, the young boy is the most short-changed. More than any other figure in the video, he fills the role of a structural device – he is both a powerful figure of exchange and an agent of disruption. Through his attempts to both run out of the video’s bounds and to disrupt internal boundaries, he holds the integrity or the disintegration of the video within his frame. For dramatic charge, music videos often have a figure that threatens the video’s surface. Yet the video’s point of view can be said to be structured for him. Aspects of the ‘Cherish’ video are meant to evoke feelings associated with childhood. Large sections of the image engender feelings of being rocked, bounced on a knee, thrown up in the air, and twirled around. The sections of music that accompany this imagery match this physical movement. For example, the chorus has a bouncier, more anthemic quality, and there is a gentler, rocking motion to the verse. Much of the imagery revolves around the child’s expressions of fear, distance and pleasure, though the video works to suppress this fact.

Though all of the figures are limited by the conventions of music video, ‘Cherish’ remains committed to expressing the roles of the characters within a set of relations: Madonna is powerful enough to uphold the whole social structure; the mermen constitute a community of men who are playful, non-competitive, and free; the child becomes so autonomous that he almost seems like an adult. These relationships can be seen as a proposal for a new social order.

The star performer

Madonna appears throughout the video and her role as ‘star’ is clearly important. Ritts treats her as he treats Janet Jackson in the video ‘Love will Never Do’. The infectiousness of the singer’s smile functions as a hook that carries the viewer into the video. As the video progresses, this approach is modified in order to work more with spatial relations and with the body as form.

Most of what we know about Madonna derives from her facial expressions and physical gestures as she switches between coyness and exuberance. These
moments highlight musical structure. Madonna’s flirtatious, sweet invitation in verse 1 helps us to hear the contours of the melodic lines, partly because, with a tilt of her head, she continually beckons downward, and the camera follows the lines of her body. The flat, deadpan expressions in verse 2 (with mouths drawn into straight lines), help us to hear the breadth of that section. It is hard to argue, though easy to sense, that the sparkle in Madonna’s eyes and her joyous expression give a buoyant levity to the image which seems to match a playful element in the music. Yet the video is more complicated than this. It was shot in the winter and the water was extremely cold. To keep warm, Madonna was wrapped in blankets between takes. A trace of athletic stoicism may be present in the video that exists in tension, rather than harmony, with the song.23

Self-reflexivity

Against Madonna’s buoyancy, hard work and flirtation, the song and video sustain a critique of Madonna. The song has a naive, untutored quality because of Madonna’s vocal delivery and the thin orchestration of its two-voice structure. The impulsiveness of the jump from D1 to D2 in the second half of the verse’s vocal line seems to highlight this naïveté – it does not proceed according to traditional formal rules.24 However, if one listens to the two-bar synthesizer break, in the middle and at the end of verse 1, one might find it a bit too sarcastic or mocking, in its relation to the voice, since it is articulated as even, measured minims. When Madonna’s voice is chorused, the image shows her alone on the beach flexing her muscles. This image ironicises Madonna’s auto-eroticism, self-determination, and self-absorption. The image also comments upon itself and refers to its relation with the music. For example, the fluke is often used as a foil to the vocal line. At the bridge, the music starts off with an instrumental lick rather than with Madonna singing, and we see the mermen, not Madonna, propelled upwards by their flukes. The mermen with their flukes have gotten a taste for stardom, and later in the bridge, the flukes impertinently appear in small gaps between the singing.

A reading of the song with the image can broaden our understanding of Madonna’s work. Critics claim that Madonna merely effects a series of poses or masks, and that her work is therefore fragmentary – that it relies upon the most superficial of associations. Yet attention to the ‘Cherish’ video, in a way that is respectful of the song’s role in shaping it, can offer this vision: that the youthful, girl-next-door vocal delivery in the song is connected to, not detached from, the video’s more direct, sexualised modes of expression.

If I remember the song, it is for Madonna’s singing, which flows in tandem with, but is not given over to, a charismatic bass line and a minimally differentiated arrangement. I think of Herb Ritts’ supportive vision of Madonna, and of the long lines of his images which seem to radiate like spokes across the surface of the video. All of these features, through their attentiveness, clarity and directness, give a sense of integrity and commitment, qualities not acknowledged in the scholarly community, either for Madonna’s work or for Ritts’.

Homo-erotic configurations

The mermen are native to the place of the video and Madonna but a visitor. The video hints that they may be Madonna’s, and also our, real object of attention: it
was built around the mermen – they were shot separately, with a great deal of concern lavished on the mechanics of the tails. (Does this make the men in this video merely ‘pieces of tail’?) The second half of verse 3, which features these tails, is a critical moment in both song and video. If we look more closely, it becomes clear that the mermen reflect a homo-erotic perspective.

The mermen exist in a self-contained world, a world without women, and they procreate their own kind, not biologically, but socially. It is true that the mermen do not seem to possess genitalia, but the men in Herb Ritts’ other work – more clearly marked as homo-erotic – show a similar tendency towards becoming sculptural forms without genitalia. The prominent tails, however, call forth numerous associations, including sexual ones. The mermen’s flukes can evoke Christian symbolism, Hans Christian Andersen’s The Little Mermaid, the mythography of dolphins and their noble rescues of people, the birth of Venus, and the TV series Flipper, but also sperm and phalluses.

Philip Hayward (1991, p. 98) reminds us that images of mermen are quite rare – we are much more familiar with mermaids. If their origin is unclear, this unclarity might itself be read as a gay image – gays are sometimes called ‘fairies’, perhaps partly because we do not know how they come to be. The mermen’s mysteriousness and elusiveness play a crucial role in defining them. They never address the camera, and are often shown disappearing from view. Invisibility is a central theme in the gay experience. It is linked to oppression, but also to desire – to watch and not be seen; to be seen but not to acknowledge being seen.

From a homo-erotic angle, the elusiveness of the mermen makes them, if anything, more powerful. It creates a context in which their every appearance carries meaning. Some sequences wittily upstage Madonna, making the mermen into the real stars: the image of Madonna’s opening her blouse, possibly to reveal her breasts, interrupted by a shot of the merman’s bare chest instead; or, perhaps, the moment when Madonna sings ‘makes me feel so good’, and the merman rubs noses with the merboy.

A homo-erotic perspective allows us to sense the force of figures that would seem, at first, to add no more than a whimsical touch to a video dominated by the presence of Madonna. Like many figures in the backgrounds of videos, the mermen take on an enigmatic character partly because the conventions of music video do not allow them to speak. In ‘Cherish’, however, the mermen’s muteness seems, not merely conventional, but a matter of will or nature. Their silent presence and the way of life it suggests become integral to the video.

**Conclusion**

This article was written to provide a means for analysing music video. ‘Cherish’ shows that a video can reflect a multitude of musical parameters. It argues for the sheer complexity of the relation between music and image: it reveals that correspondences between music and image can range from the most strict to the most subtle or enigmatic – and that the most fragile may be the most engaging. The ‘Cherish’ video shimmers between the most traditional of texts – the topos of mother and child – and the most radical – a social order in which gays, women and children can live with independence.

I do not want to claim that connections between music and image are natural, or that we have an innate capacity to see these connections, or even that all or most
people do see them. More than proving what people do with music videos, I want rather to show what videos and videomakers can offer us. I might say, paraphrasing Wittgenstein, that learning the language of music video means learning a form of life (1968, p. 19).

Appendix A: ‘Cherish’ video narrative

Introduction
The boy runs along the shoreline.

Verse 1
Madonna sings while the waves crest behind her.

Chorus 1
Three mermen swim away from the camera. The section cuts rapidly among Madonna, the mermen and the boy, now a merboy.

Verse 2
Water runs up Madonna’s legs. Madonna and the mermen spy upon each other.

Chorus 2
Madonna spirals away from the camera. The figures of Madonna, the mermen and the merboy are placed high in the frame.

Bridge 1
With their arms reaching upward, three mermen fill the frame; Madonna moves from left to right at the water’s edge. Madonna remains at the shoreline while the mermen remain at sea.

Verse 3 (instrumental)
The camera travels a path out to sea, passing Madonna as she cradles the merboy. The mermen swim in formation under water.

Bridge 2
Madonna turns towards the camera. She frolics with the merboy on the sand.

Break
A merman dives backwards into the ocean, revealing his chest. Madonna flexes her muscles at the water’s edge.

Outchorus
A merman twists away from us, pulling himself out of the water. The boy, human again, runs towards the mermen as they swim out to sea. Madonna remains on the sand. The boy (now older?) returns to face her.
Appendix B: ‘Cherish’ lyrics

Introduction
Cherish, cherish.

Verse 1
So tired of broken hearts and losing at this game.
Before I start this dance I take a chance
in telling you I want more than just romance.
You are my destiny I can’t let go, baby, can’t you see?
Cupid, please take your aim at me.

Chorus 1
Cherish the thought of always having you here by my side.
Oh, baby, I cherish the joy, you keep bringing it into my life.
I’m always singing it.
Cherish the strength, you got the power to make me feel good.
And, baby, I’ll perish the thought of leaving. I never would.

Verse 2
I was never satisfied with casual encounters.
I can’t hide my need for two hearts that bleed
with burning love that’s the way it’s got to be.
Romeo and Juliet, they never felt this way I bet.
So, don’t underestimate my point of view.

Chorus 2
Cherish the thought of always having you here by my side.
Oh, baby, I cherish the joy, you keep bringing it into my life.
I’m always singing it.
Cherish the strength, you got the power to make me feel good.
And, baby, I’ll perish the thought of leaving. I never would.

Bridge 1
Who? You! Can’t get away, I won’t let you.
Who? You! I could never forget to.
Cherish is the word I use to remind me of your love.
Together, you’re giving it to me, boy.
Keep giving me all, all, all your joy.
Give me faith. I will always cherish you.
Romeo and Juliet, they never felt this way I bet.
So, don’t underestimate my point of view.

Verse 3 (instrumental)

Bridge 2
Who? You! Can’t get away, I won’t let you.
Who? You! I could never forget to.
Cherish is the word I use to remind me of your love.

Acknowledgements
I would like to thank Juli Brown for providing the drawings and Arun Bharali, Stevan Key, Mitchell Morris, Jann Pasler and especially Charles Kronengold for their insightful comments and careful reading of the text.
Endnotes

1. The attempts which come closest are Tricia Rose’s (1994) analysis of Public Enemy’s ‘Baseheads’, Andrew Goodwin’s (1992) shot-by-shot discussion of George Michael’s ‘Father Figure’ and Melanie Morton’s (1994) analysis of Madonna’s ‘Express Yourself’. None of these analyses, however, is sufficiently attentive to visual detail and to the flow of the music.

2. Composer and theorist Robert Erickson (1955, p. 24) states: ‘Any melody is analogous to a moving line. The line is strictly speaking never continuous; there is always a jump from tone to tone, and the visual analogy of a series of dots producing the effect of a line would be more accurate.’

Visual images can also be said to have fundamental structural contours. Leonardo Da Vinci (1961, p. 146) said something very beautiful about this: ‘The air is full of an infinite number of radiating straight lines which cross and weave together without ever coinciding; it is these which represent the true form of every object’s essence.’

3. Seurat (1945) thought of the horizontal line as passive, the vertical line as active; lines descending from the horizontal seemed lugubrious; lines ascending from the horizontal tended towards joy.

4. The song presents the chorus and the complete verse only twice each. This in itself is unusual for a pop song of this length at this tempo. The version of ‘Cherish’ which is used in the video is about 4 mins 30 secs.

5. I discuss the role of sectional divisions more fully in the second part of this article.

6. Though a relation, in music video, can exist through similarity or contrast, confluence may be more easily perceived than disjuncture. Similarly, an arch shape is simple to grasp. The ‘Cherish’ video’s accessibility derives in part from the video’s clear transformations of its basic shape, as well as from its continuity and flow.

7. Note the firm articulations, in both image and music, at the beginnings of phrases in verse 1. Note also how, throughout each section, the images gradually diminish in force. In the first chorus, the image punches through the section, from Madonna marching up to the camera, the little boy and the mermen surfacing, the mermen pulling through the water, to Madonna crunching her shoulder up and down. As the music in the chorus gradually winds down, the image too, softens; the residual ‘pop’ in the imagery of the older merman nudging the little boy on the nose and the last image of Madonna’s knee propped up reflects music that has lost its drive.

The image slavishly follows the phrase structure of the music in the first verse and chorus. During these two sections, the image supports the song’s regular phrase structure and is edited to emphasise meaningful word groupings and motivic cells. In verse 2, however, the image begins to play with the phrase structure – sometimes coming slightly before, sometimes lagging slightly behind the musical phrase. Since we can recognise both the rule and its variation, we still experience the phrasing of the music and the image in verse 2 as contrapuntally related.

8. Although probably produced by a digital synthesizer, this patch is designed to have the richness and grain associated with earlier, analogue synthesizers. It resembles a Minimoog ‘lead guitar’ timbre popularised by fusion keyboardists like Chick Corea, Jeff Lorber and Jan Hammer. On digital synthesizers, this timbre is sometimes referred to as the ‘Steve Winwood’ patch, because of its frequent presence on his albums from the early and mid-1980s.

9. The bass line is a very significant feature of the ‘Cherish’ song. Both it and the vocal line are equally charismatic. The bass line is interesting because it is one of the song’s few performerly aspects – there is some flexibility in the rhythmic feel, as well as some improvisatory flair in the choice of notes. The vocal line is, as one would expect, in the foreground. ‘Cherish’ is better thought of as a two-voice structure, and not as melody and accompaniment, because the bass line provides so much melodic interest, and seems to carry the harmony, even when it is contradicted by the keyboard pad (as in the bridge).

10. Synthesizer- and drum machine-based dance-pop songs with a human feel became a staple on the Rhythm & Blues charts starting in 1982. Examples from 1982-83 include ‘Sexual Healing’ by Marvin Gaye, ‘ Ain’t Nobody’ by Rufus and Chaka Khan, ‘ Just be Good to Me’ by The S.O.S. Band, and ‘Yah mo b there’ by James Ingram with Michael McDonald. Examples can be found on the Pop charts soon after: Prince’s ‘When Doves Cry’ (1984), which reached number one on both the Pop and R&B charts, Madonna’s own ‘Borderline’ (1984), and Howard Jones’ ‘ Things Can only Get Better’ (1985). Madonna’s sometime collaborator, Manhattan DJ and producer ‘Jellybean’ Benitez, was instrumental in popularising this sound and connecting it with certain strands in the British ‘New Romantic’ style. One might
say that this sound achieved hegemonic status on American radio from 1986–88.

11. Many aspects of the song are constructed to emphasise sameness over difference, restraint over assertion. The melodic materials of the chorus and bridge relate closely to those of the verse. The pitches G–F♯–A–B, which form the principal motive of the chorus, come out of the synthesizer’s cambiata figure in the verse, D–C♯–A–B. These connections help to make the similarities among the sections of the song as important for the videomaker as are the differences among them. The orchestration remains a continuous wall of sound. The only element that seems really to progress in the song is the synthesizer’s counter-melodies. For example, the synthesized trumpet is buried very deep in the mix in the first verse. It does not come to the fore until the second bridge. Later, it plays a crucial role in bringing the song to a close.

12. One way that the video draws attention to sectional divisions is through shifts in pacing. The imagery of the verse is slowed down, and the imagery for the chorus moves more quickly. The particular pacing of ‘Cherish’ derives from the fact that the chorus has a faster harmonic rhythm than the verse.

13. This moment illustrates a crucial point about music video: most often, the image can point to but a few musical parameters. At one moment, the image may comment upon the subtle nuances of the music, and at another, it may suggest the way that the music articulates larger sectional divisions. However, by the close of the video, our attention will have been drawn to almost all of the musical parameters. Music video works by saying, ‘first listen to this, then listen to this, now this’. Relations between music and image can exist within all degrees of concordance. Some correspondences may be literal one-to-one mappings, like those seen in cartoons (referred to as ‘mickey mousing’). An example in ‘Cherish’ would be the moment when Madonna raises her shoulders on beats 1 and 3. Yet the most successful connections may be ones that are obscure and enigmatic. Such connections may ask viewers to tease out the relation between music and image, and contribute to an important aspect of music videos – that they encourage us to watch them repeatedly. Some more enigmatic and subtle connections in ‘Cherish’ might include the relation between the sibilants in the voice and the percussion samples, and how they, in turn, relate to the prismatic light on sand and water; how the bass line seems to carry the figures forward; the echo of the theme song from the television series Flipper, and the way that the nostalgic elements in the music are paralleled by imagery drawn from beach films and old home movies, myth and fairytale.

14. This image of Madonna makes a vague joke, which would relate the men’s coming onto land (as in Hans Christian Andersen’s fairytale) to the first amphibian stalking forward without much grace. This kind of oblique association is common in music video. Much of the imagery in music video operates on jokes and associations which are not quite consciously acknowledged. We can, however, turn to the music and to the whole music video to fill in the context, in order to get the tone or flavour of such an association.

15. The expected form for a song of this genre and period would be (introduction)–verse–chorus–verse–chorus–bridge–verse (instrumental)–outchorus. ‘Cherish’ departs from the letter, if not the spirit, of this form by repeating the bridge and adding a break. The presence of a break is somewhat unexpected, since the break is a formal conceit associated more with dance music than with middle-of-the-road. The song’s most significant modification of formal conventions, however, lies in the internal construction of the third verse. The beginning of the third verse sounds more like an extension of the previous section (bridge 1) than like the downbeat of a new section. The voice elides this sectional division: it draws out the end of its phrase as the rhythm section begins the third verse. (Madonna sings, ‘love. Together, you’re giving it to me boy.’) We may think, for a moment, that we are still in the bridge. The first half of this verse functions as an instrumental with prominent vocal ad-libs. The ‘Steve Winwood’ synthesizer patch enters the texture with a solo line supported by the rhythm arrangement of the verse. Madonna, drawing attention away from the synthesizer solo, sings, not material from the verse, but new material that seems more appropriate to the bridge because of its improvisatory character and shorter phrase-lengths. If it were not for the voice, these eight bars would seem like part of a conventional instrumental. Interrupting the synthesizer solo and vocal ad-libs, the second half of this verse suddenly returns to the arrangement scheme and vocal melody of the previous verses, as if it had been just another verse all along. It only becomes clear that this is the third verse when, as in the second verse, Madonna sings, ‘you. Romeo and Juliet, they never felt this way I bet’.

Not only is verse 3 unusual for its internal organization – the conflicting tendencies of its first half, its surprising split down the middle, and its elided beginning – this section is pre-
ceded and followed by appearances of the bridge. The attempt to place the 'normal' portion of the verse demands some work on the part of the listener. The image assists in the task.

16. Another director might have emphasised the suspension, say through imagery of unrequited love. Ritts moves past this suspension and draws attention to the verse fragment, thereby underscoring the mermen's collective identity. (See pp. 179–80 for discussion of this identity and its construction within the video.)

17. The gesture of a broad sweeping arc which resolves into a balanced image of height and depth occurs frequently in 'Cherish'. Visual arcs imply phrasing that can be shaped either with or against that of the music. One can see an antecedent to this gesture in Ritts' still photography. The photograph 'Male Nude With Bubble' (reproduced in Ellenzweig 1992, p. 194) bears interesting similarities to 'Cherish': many of the qualities of balance in the photo seem, in the video, to be extended in time and placed in relation to the song.

18. The image highlights two falling gestures in the bass, and the rising line in the melody.

19. The Busby Berkeley sequence is much beloved, particularly in the gay community, for its spectacular opulence and willingness to dispense with narrative.

20. It is difficult to determine whether we are seeing feet that belong to a child or to an adult. The ambiguity thus created can be expressed as follows: if the legs are a child's, then the merboy meets Madonna as a child; if they are adult legs, he may meet her as a lover. In the latter case, more time will have elapsed, most likely in the company of the mermen.

21. Michael Jackson's 'Thriller' provides a good example of this phenomenon, as shown by Kobena Mercer (1994, pp. 93–109). Mercer points to the many thematic threads in the image and the song that might bear upon the video's ending.

22. The expressions on the mermen's faces in the group close-up in the first bridge could be read as an expression of anger at Madonna having stolen their child from the ocean. The boy's teeth chatter as Madonna reaches to hold him. Yet the imagery in this section also points to the notion that the boy comes ashore of his own accord — here, we see the fish-tail moving by its own power towards the shore: the merboy has beached himself. The imagery of Madonna in the second bridge — her shielded breasts, vulnerable belly, and heavy movements — seems campy. Her movements undercut her attempt to look beautiful. In addition, the relationship between the mermen and the boy looks warm and unconflicted, but the relationship between Madonna and him seems less so.

23. Another example of correspondence between affect and music can be drawn from Madonna's 'Open Your Heart' video. Her still movements, heavy mascara, brooding expression, and dark colours (emerald green, deep blue, black) match the severity of the music. Music video imagery can ignore, match, or supersede various parameters of the music. In 'Cherish', the affective character of the image is more extreme than that of the music. The mood of the image seems like a heightened version of the mood of the song — with images of people crashing and diving through frothy water, Madonna's rocking movements, and expressions of elation. At one level, music videos need to exceed the bounds of the music — it gives their existence a raison d'être. The claim of greater affective range is an easy way for the image to show that it is more authoritative than the music. This is an easy way, but not the only one. In U2's 'With or Without You', the image envelopes the music by moving at pulses both faster and slower than that of the music. In Peter Gabriel's 'Mercy St', a slow song that evokes a sense of solemnity, the slowest pulse is associated with the image.

24. The vocal line, here, embodies a non-traditional construction. Much of the song, until the high D4, can be heard as circling around D3. The octave leap up to D4 is not prepared for, in terms either of linear direction or harmonic flow. Its intrinsic appeal is questionable, but it does support the wide-eyed quality of the singer's subject position in this song.

25. I would suggest that Flipper may be the most concrete possibility, because its theme song resembles the opening of 'Cherish'.

26. There seems to be a very wide range of competence in the ability to view music video, with teenagers who have watched a great deal of television often displaying the greatest visual and musical acuity. My experiences, from speaking with videomakers and their assistants, and students who watch music video, suggest that their engagement with the medium is quite intense. The students in my undergraduate course on music video often write sustained and incisive analyses of videos. I discovered, further, that their fluency in the medium meant that music videos could be used as a tool for teaching music fundamentals. I hope to present examples of this application in a future paper.
References

von Laban, R. 1974. Laban’s Principles of Dance and Movement Notation (Boston)
Reti, R. 1978. The Thematic Process in Music (Westport)
Rose, T. 1994. Black Noise (Hanover)