

## Capitolo terzo

### Ritmi visivi, ritmi sonori, ritmi audiovisivi

#### *Suono e immagine, una saldatura inevitabile*

Il ruolo del suono nella costruzione del montaggio audiovisivo è stato oggetto di dibattito sin dall'introduzione del sonoro nel cinema muto verso la fine degli anni Venti.

In questo primo periodo, in cui sono particolarmente diffusi i paragoni tra cinema e musica, viene coniata l'espressione *contrappunto audiovisivo* per designare "la formula ideale in abstracto di cinema sonoro in cui i suoni e le immagini appartengono a due catene parallele e liberamente collegabili, senza alcun genere di dipendenza unilaterale" (Chion 1990, p. 37).

Nella terminologia della musica classica occidentale il termine *contrappunto* indica un tipo di scrittura che impone a differenti voci simultanee di essere eseguite ciascuna in uno svolgimento orizzontale, coordinato con l'esecuzione delle altre, ma al tempo stesso autonomo. Una forma di scrittura radicalmente differente è invece l'*armonia* in cui prevale una relazione che privilegia i rapporti verticali tra le note e le voci in funzione della realizzazione degli accordi.

L'utilizzo del termine *contrappunto* in ambito cinematografico è estremamente controverso. Mentre in musica questo tipo di relazione interviene a regolare i rapporti tra elementi che condividono un'unica sostanza dell'espressione (sonora), nel film i suoni e le immagini sono caratterizzati da materialità del tutto differenti. Un presunto *contrappunto audiovisivo* implicherebbe la presenza di una sorta di "voce sonora", percepita orizzontalmente come

coordinata alla colonna video, ma al tempo stessa del tutto distinta. Il cinema tuttavia, tende a escludere la possibilità di un simile funzionamento orizzontale e contrappuntistico, privilegiando decisamente i rapporti armonici e verticali tra i suoni e le immagini.

In ambito cinematografico la nozione stessa di colonna audio si rivela problematica. Infatti, se da un punto di vista strettamente tecnico è possibile isolare una pista sonora che attraversa il film interamente o in parte, è tuttavia impossibile ricondurre l'insieme dei suoni e dei rumori a una totalità unitaria.

La nozione di colonna audio, così come viene utilizzata, è in realtà un puro e semplice calco meccanico dell'idea di colonna immagine, la quale invece esiste, poiché essa deve il proprio essere e la propria unità alla presenza di un quadro, di un luogo di immagini investito dallo spettatore (p. 40).

L'ordine sequenziale dei suoni in un film non è una condizione sufficiente per dimostrare che esiste un complesso sonoro, un insieme chiuso, dotato di una qualche forma di coesione e quindi potenzialmente confrontabile con la colonna delle immagini. Piuttosto, ogni elemento sonoro (suoni/rumori) allaccia con gli elementi narrativi contenuti nell'immagine (personaggi/azioni) e con gli elementi visivi compositivi e scenografici, dei rapporti verticali simultanei più immediati, forti e significativi di quelli che questo stesso elemento sonoro può allacciare parallelamente con gli altri suoni che lo precedono o lo seguono nel flusso musicale.

L'esempio del suono fuori campo può aiutare a chiarire ulteriormente questo passaggio teorico. In un film i suoni fuori campo sono quelli che intrattengono con la colonna video il rapporto più semplice e vincolante, l'esclusione visiva della sorgente sonora è infatti la condizione sufficiente a separare questi suoni da tutti gli altri, a farli emergere dal continuum della pista audio. Eliminando l'immagine, tuttavia, diviene immediatamente evidente che nulla distingue

questo tipo di suoni da tutti gli altri, “la struttura globale crolla, e i suoni, insieme, ne ricostruiscono una del tutto nuova” (ib.).

Le forme di testualità audiovisiva dunque non sono caratterizzate dalla presenza separata di una colonna immagini e di una colonna audio, piuttosto si configurano come lo spazio della loro reciproca fusione, *un luogo di immagini e di suoni*.

Le ragioni di questo legame così vincolante devono essere ricercate nell'operazione del montaggio. Per quanto riguarda le immagini, è il montaggio che ha reso possibile la creazione di un'unità specifica del linguaggio cinematografico, il *piano*. Sia che lo si consideri come un'unità narrativa fortemente strutturata (Chion 1990), o esclusivamente come una porzione di pellicola impressionata dalla macchina da presa tra l'inizio e la fine di una ripresa (Vanoye, Lété 1992), il piano presenta il vantaggio innegabile di essere un'unità neutra, oggettivamente individuabile.

Al contrario, per quanto riguarda la componente audio, il montaggio sonoro non ha prodotto alcun tipo di unità specifica. In un film l'ascolto della successione dei suoni non permette infatti di isolare alcun genere di unità, soprattutto perché le giunte tra i singoli suoni non sono affatto evidenti, come invece avviene per gli stacchi che separano le inquadrature in una sequenza di montaggio. Ciò non significa che la componente sonora di un film costituisca per il suo ascoltatore un flusso del tutto privo di cesure; è infatti possibile individuare delle unità, come frammenti di dialogo, rumori, melodie, cellule ritmiche, ma la percezione della loro presenza dipende essenzialmente dalle caratteristiche peculiari dei suoni, dal grado di competenza musicale e dal livello di attenzione.

Se si tratta di rumori, realizziamo un *découpage* percettivo scomponendoli in eventi sonori, il che risulta più facile se si tratta di suoni isolati. In una musica isoliamo delle melodie, dei temi e delle cellule ritmiche, secondo il grado della nostra cultura musicale. In breve, ci comportiamo come sempre, e

abbiamo a che fare con unità che non sono specificamente cinematografiche, e dipendono completamente dal tipo di suono e dal livello di ascolto prescelto (Chion 1990, p. 44).

Il film non è concepibile dunque come il luogo della libera associazione di due diverse sostanze dell'espressione che rinviano a sistemi di segni paralleli e autonomi, ma esclusivamente come il prodotto del loro sincretismo, lo spazio in cui agiscono *combinazioni audiovisive*.

### *Il montaggio audiovisivo*

Le "origini" del montaggio si trovano nella composizione plastica.

Il "futuro" del montaggio sta nella composizione musicale (Ejzenštejn 1963-70, p. 6.).

Ciò che rende i videoclip delle forme originali di testualità audiovisiva non è unicamente la capacità di condensare in uno spazio breve un gran numero di inquadrature, complessi movimenti di macchina ed elaborati effetti speciali, ma anche l'attitudine a esplorare le potenzialità del sincretismo audiovisivo per spingersi al limite della soglia che separa una narratività forte dall'assemblaggio del tutto arbitrario di suoni e immagini.

In questo senso il montaggio non si riduce semplicemente a un'operazione tecnica necessaria a "chiudere il testo" o a mettere in forma in modo virtuosistico suoni e immagini, piuttosto costituisce il luogo stesso della realizzazione del senso, lo spazio in cui gli elementi appartenenti a linguaggi diversi (scritto, visivo, musicale) confluiscono in combinazioni audiovisive che sfruttano i meccanismi sinestesici per stimolare nello spettatore una reazione di tipo passionale.

Se la pratica del montaggio costituisce "l'arma segreta" dei videoclip, lo strumento privilegiato che ne assicura l'ef-

ficacia, per comprenderne il funzionamento e individuarne i dispositivi è necessario innanzitutto risalire all'origine della riflessione sull'unione di suoni e immagini, all'autore che più di ogni altro ha teorizzato le potenzialità espressive del montaggio audiovisivo, Sergej M. Ejzenštejn.

Come si tenterà di dimostrare in questo capitolo, le considerazioni avanzate da Ejzenštejn sulle forme della correlazione audiovisiva si rivelano infatti straordinariamente attuali e decisamente utili per affrontare il tema complesso dell'efficacia ritmica dei videoclip.

Nel 1937 il cineasta russo dedica al montaggio un'opera fondamentale, la *Teoria generale del montaggio*. In questo lavoro egli propone di distinguere innanzitutto tra *rappresentazione* (izobrazenie) e *immagine* (obraz).

Nella riflessione di Ejzenštejn il primo concetto esprime l'ordine delle cose percepito nella sua semplice "datità" (ci sono cose nel mondo che possono essere rappresentate tramite le parole, le fotografie o le immagini in movimento), il secondo, l'*immagine*, rinvia alla decostruzione di quest'ordine preesistente e alla sua riconfigurazione significante.

Mentre la rappresentazione costituirebbe un atto semplicemente riproduttivo, l'*immagine*, configurandosi come un procedimento di smontaggio e rimontaggio dei dati reali, si caricherebbe di virtù espressive o interpretative.

Nella riflessione di Ejzenštejn tuttavia, il termine *immagine* non rinvia esclusivamente all'insieme di procedure utili a intervenire sul dato reale per manipolarne l'ordine costitutivo, piuttosto esso esprime "l'orizzonte istitutivo della sensatezza in generale" (Montani 1999, p. 20). L'*immagine* infatti mira a ricondurre il soggetto fino alle radici profonde, alle precondizioni stesse di quell'attività originaria che rende possibile la comparsa del senso.

Se con questa attività originaria, in altri termini, noi stiamo rispondendo a un incontro che ci ha già presi, l'*immagine* mira a mostrare, al tempo stesso, questo esser-presi e questa risposta: immaginare è dunque un'attività che si riconosce in debito con un aver patito, con una passività, con un pathos (p. 20).

Tra *rappresentazione* e *immagine* non esiste uno scarto temporale, non è possibile, in altri termini, isolare una riproducibilità originaria, pura, soggetta in un secondo momento a un insieme di operazioni che ne rendono possibile una riconfigurazione assoluta. La stessa rappresentazione del mondo reale è già sempre un'attività istitutiva cui non è possibile contrapporre nessun genere di materia grezza, nessun elemento amorfo successivamente plasmabile. Ciò che si offre alla rappresentazione del soggetto, pur risultando da un'istituzione originaria, costituisce dunque sin dall'inizio la premessa della realizzazione di un senso. Poiché tuttavia questa attività di costituzione del reale sfugge a una percezione piena, consapevole, cessando di essere avvertita nella sua processualità, emerge per il soggetto la necessità di riaffermarne il principio tramite un intervento creativo, l'*immagine*.

Il testo estetico emerge dunque come il luogo di questo ritorno, lo spazio in cui diviene possibile esibire le potenzialità e i limiti di questa originaria formatività.

È precisamente a partire dalla consapevolezza che il rapporto tra rappresentazione e immagine è di natura processuale che nella riflessione di Ejzenštejn emerge il ruolo *fondante* del montaggio. Se, infatti, l'immagine è un intervento espressivo che rivela la natura costitutivamente creativa della rappresentazione, allora lo scopo dell'arte cinematografica sarà innanzitutto quello di rendere percepibile questo processo immaginativo, dandogli un tempo e una forma.

L'arte cinematografica, autentico territorio dell'immagine, non presenta mai risultati, ma sempre e solo processi. Proprio per questo, l'immagine richiede tempo e sviluppa tempo (p. 21).

Nella riflessione di Ejzenštejn il montaggio si configura dunque come l'insieme delle procedure necessarie a qualificare il tempo dell'immaginare. Significativamente egli distingue tre diversi modi del montaggio che rinviano ad altrettante manifestazioni della temporalità dell'immagine.

Il primo tipo di montaggio è chiamato *compositivo*. Visibile compiutamente nelle arti figurative, esso trova realiz-

zazione nel cinema nella composizione plastica delle singole inquadrature, dove "il tempo si fa presente nel percorso che l'occhio dello spettatore è tenuto a effettuare, per apprensioni successive, lungo il profilo della composizione: è questo percorso a costituire, per sintesi finale, l'unità dell'immagine" (ib.).

La seconda modalità è il *montaggio sequenziale* che rende percepibile la scansione delle sequenze marcando l'alternanza degli stacchi tra le inquadrature. Questo secondo tipo di montaggio sovrintende direttamente alla durata e alla distribuzione degli intervalli temporali, inscrivendo nel testo un'*ossatura ritmica* di base.

Diversamente dal montaggio compositivo, il montaggio sequenziale è tipicamente cinematografico, il senso infatti non emerge dalla composizione delle singole inquadrature ma si dispiega a partire dal ritmo della loro alternanza. Con il montaggio sequenziale il tempo si impone all'attenzione del soggetto come incoattività, come processo in piena fase di sviluppo, esso aspira a esibirsi come tempo configurante.

La terza modalità è definita da Ejzenštejn *montaggio audiovisivo* ed è caratterizzata innanzitutto dalla presenza del sonoro. Questo terzo tipo di montaggio si differenzia profondamente dal *montaggio compositivo* e dal *montaggio sequenziale* non solo per la presenza materiale di una nuova sostanza dell'espressione, quanto piuttosto per le inedite possibilità espressive che scaturiscono dall'associazione tra elementi visivi ed elementi sonori.

Come rileva Pietro Montani (pp. 23-24) con il montaggio audiovisivo la dimensione espressiva, l'intervento creativo, si spostano in uno "spazio-tempo intermedio che non dipende, come nel caso precedente, da un'alternanza (da un movimento sequenziale), bensì consiste (...) in una simultaneità internamente differita o ek-statica, in un gioco multiforme (perché occasionato da molti stimoli sensibili) di ritensioni e di protensioni".

Con la terza modalità di montaggio si apre dunque per il cinema la possibilità di spingersi oltre le forme consolidate della narrazione classica.

L'arte comincia propriamente solo a partire dal momento in cui l'associazione tra il suono e la rappresentazione visiva non è più semplicemente *registrata* secondo il rapporto esistente in natura, ma è *istituita* secondo il rapporto richiesto dai compiti espressivi dell'opera. Nelle forme più rudimentali si tratterà della sottomissione di ambedue gli elementi allo stesso *ritmo*, rispondente al contenuto della scena. È questo il caso più semplice, più accessibile e più frequente di montaggio audiovisivo: in cui i pezzi della rappresentazione visiva vengono tagliati e montati secondo il ritmo della musica che scorre parallelamente sulla colonna sonora. (...) A partire da questo caso più elementare – la semplice coincidenza “metrica” degli accenti nella “scansione” – è possibile ottenere un gran numero di combinazioni sincopate e un “contrappunto” puramente ritmico che consiste nel gioco calcolato della non coincidenza degli accenti, delle lunghezze, delle frequenze, delle ripetizioni ecc. (Ejzenštejn 1986, pp. 157-158).

Ne *Il montaggio verticale* (1986, p. 132) Ejzenštejn approfondisce ulteriormente la riflessione relativa alle innovazioni espressive rese possibili dall'introduzione del sonoro. Pur sostenendo che tra la concezione di un montaggio puramente visivo e quella di un montaggio che relaziona elementi di aree diverse – in particolare l'immagine visiva e l'immagine sonora – non ci sono differenze di fondo sulla linea della creazione di un'immagine audiovisiva unitaria e coerente, il cineasta russo insiste sulla necessità di ripensare in termini nuovi la correlazione tra elementi visivi ed elementi sonori.

Per chiarire la nuova concezione di montaggio, in cui prevale una correlazione di tipo verticale tra suoni e immagini, Ejzenštejn ricorre non a caso a un esempio tratto dall'ambito musicale:

a tutti è familiare l'aspetto esteriore di una partitura d'orchestra: una certa quantità di pentagrammi in ciascuno dei quali è scritta la parte di un determinato strumento. Ogni parte si sviluppa con un movimento progressivo in orizzontale. Ma la connessione verticale è un fattore non meno importante e decisivo: la correlazione musicale dei diversi elementi dell'orchestra in ogni unità di tempo. Così con il movimento

progressivo della verticale che coinvolge tutta l'orchestra e avanza orizzontalmente, si realizza il complesso e armonico movimento musicale dell'intera orchestra. Se ora passiamo dall'immagine di una partitura musicale a quella di una partitura audiovisiva, osserveremo che in questo nuovo stadio è come se alla partitura musicale si aggiungesse un ulteriore pentagramma: quello delle inquadrature che procedono l'una dopo l'altra conformandosi plasticamente al movimento della musica e viceversa (ib.).

Nel montaggio audiovisivo il rapporto tra suoni e immagini non privilegia la dimensione della sequenzialità, in altri termini non consiste in una semplice successione orizzontale secondo una logica del contrappunto, quanto piuttosto nella simultanea sovrapposizione verticale di elementi appartenenti a due sistemi semiotici indipendenti<sup>1</sup>.

Le risonanze audio-visive che scaturiscono da questa originale forma di sincretismo inscrivono nel testo delle complesse configurazioni ritmiche che valorizzano la dimensione estetica e passionale della significazione, aprendo a un'esperienza piena della fruizione.

koiné aisthesis, unità del "sentire", unità nelle differenze del sentire. Ecco con cosa intende lavorare la terza modalità del montaggio: non solo con un "io vedo" e un "io ascolto", ma, dice Ejzenštejn (1986, p. 59), con un "io sento", con l'unità ek-statica o "differente" (perché ottica, acustica, tattile...) di questo sentire (Montani 1999, p. 24).

Questa spinta a riformulare i canoni tradizionali del racconto, l'attitudine a "forzare" i confini dello spazio letterario, si estendono ben oltre i limiti del linguaggio cinematografico. La lezione di Ejzenštejn e delle avanguardie storiche nel corso di un secolo ha influenzato profondamente le diverse forme di testualità audiovisiva, innescando pratiche di sperimentazione videoartistica, delineando spazi di intervento del tutto inediti per il grande schermo, condizionando i tentativi successivi di restituire la forma e il ritmo della musica attraverso le immagini.

In questa prospettiva le soluzioni ritmiche delle forme brevi contemporanee, in particolare nell'ambito dei videoclip, possono essere rilette a partire dalla nozione di montaggio audiovisivo.

*Il videoclip. Oltre l'associazione di suono e immagine*

Nei casi più interessanti la manipolazione delle sonorità, la decostruzione/riconfigurazione del materiale visivo si sottraggono spesso in maniera decisa a qualsiasi tentativo di ricomposizione narrativa da parte dello spettatore.

Nei videoclip il montaggio,

ritornando sugli stessi motivi, e giocando ogni volta su quattro o cinque temi visivi di base, (...) è, più che un modo per far avanzare l'azione, un modo per far girare le facce del prisma, e per creare così, tramite la rapida successione dei piani, una sensazione di polifonia visiva e persino di simultaneità (Chion 1990, p. 140).

Il suono non si riduce a un semplice contrappunto delle immagini, piuttosto contribuisce in maniera esplicita alla sperimentazione di nuovi regimi della significazione, in cui la coincidenza marcata con le immagini costituisce solo uno dei momenti che compongono un gioco ritmico complesso fatto di scarti audio/visivi, corrispondenze mancate, fratture esibite.

Il videoclip soprattutto procede per frammentazione: le riprese convergono, a volte collegandosi, altre volte respingendosi, senza che nessun tema o oggetto si imponga alla vista. Le inquadrature magari hanno un "contenuto" perché sono stati filmati luoghi, scene, persone, ma il modo in cui sono state raggruppate nella continuità della pellicola impedisce che si scopra in esse un punto di riferimento comune. (...) La dissolvenza per il videoclip non è la forma incerta in cui si opera un passaggio, ha una sua durata, le immagini che la compongono crescono ciascuna su se stessa, e anche nella relazione

forzata imposta loro dalla simultaneità, si afferma quindi come un'immagine indipendente il cui fascino poggia sulla combinatoria che sviluppa (Sorlin 1997, pp. 235-236).

Un videoclip che esprime in modo esemplare queste potenzialità espressive del montaggio audiovisivo è senz'altro *Drop* del musicista giapponese Cornelius.

Il video si apre con una combinazione audiovisiva decisamente insolita: alcune gocce cadono da un rubinetto, si infrangono sulla superficie del lavandino a un ritmo costante producendo un lieve rumore.

Dopo alcuni secondi sullo sfondo ritmico prodotto da questo rumore persistente si insinua una sonorità elettronica estremamente dilatata. L'effetto di questa sovrapposizione tra musica e rumori rafforza ulteriormente la sensazione di una forte solidarietà audiovisiva.

Nelle sequenze successive un bambino si avvicina al lavandino, apre definitivamente il rubinetto e inizia a lavarsi meticolosamente le mani.

La sequenza dei gesti del bambino si ripete davanti a uno specchio, i suoni e le immagini sono perfettamente coincidenti, disegnano un flusso assolutamente lineare, compatto.

Due *microritmi* distinti (visivo e musicale) si sovrappongono fino a fondersi in un'unica combinazione audiovisiva, sancita dal gesto simbolico del bambino che con l'indice della mano destra interrompe a intervalli regolari il flusso dell'acqua, simulando in modo esplicito una forma decisamente originale di metronomo.



Fig. 1. Cornelius, *Drop*.



Fig. 2. Cornelius, *Drop*.

La durata delle immagini, la composizione delle inquadrature, il ritmo del montaggio inscrivono progressivamente nel video un andamento tensivo crescente.

I gesti ripetitivi del piccolo protagonista sembrano rinviare non tanto all'esecuzione quasi meccanica di una banale attività quotidiana come la pulizia personale, quanto piuttosto a un preciso rituale che si annuncia drammatico.

Il bambino ora immerge completamente il viso sott'acqua, il suo sguardo è assente, gli occhi sbarrati sono velati dal liquido che ormai ricopre il pavimento.



Fig. 3.  
Cornelius, *Drop*.

La musica si è fatta distante, durante gli interminabili secondi che trascorrono sott'acqua retrocede addirittura a elemento di sfondo, la sua percezione è filtrata dalla presenza del bambino, i suoni sono ovattati, il ritmo appena distinguibile.

Finalmente il bambino rialza la testa, tossisce, solleva lo sguardo sullo specchio, mentre la musica interviene di nuovo a rivendicare con forza il suo ruolo, agganciandosi vigorosamente alle immagini. Questa breve ma intensa immersione nell'acqua ha prodotto una frattura nell'andamento lineare del testo, una trasformazione radicale che investe sia il piano della narrazione sia il livello più superficiale del videoclip, lo stile di regia. L'effetto di forte realismo inscritto nel video a partire dalle prime inquadrature viene completamente rovesciato, la tosse del bambino dà avvio a una sequenza in cui viene figurativizzato un vero e proprio stato di allucinazione.

Immagini irregolari, mosse, sfocate, manipolano il corpo del bambino e l'ambiente circostante.

Fig. 4. Cornelius, *Drop*.Fig. 5. Cornelius, *Drop*.

Gli effetti di distorsione alterano linee, figure, colori, compromettono la coerenza e la leggibilità del mondo reale aprendo contemporaneamente un varco in una dimensione “aliena”.

Fig. 6. Cornelius, *Drop*.

Lo specchio del bagno assolve pienamente la sua funzione di dispositivo autoriflessivo, trasformandosi in una finestra che permette al bambino di proiettarsi in uno spazio aperto e incontaminato, un luogo onirico, un sogno a occhi aperti, a portata di chiunque e al tempo stesso distante.

Fig. 7. Cornelius, *Drop*.

Una dimensione “invisibile”, effimera, pronta a svelarsi o a svanire come un’epifania improvvisa, come una banale scritta su una t-shirt, come una semplice decorazione che per un istante si fa segno, traccia di un’intenzione e di uno stato d’animo. “DREAM”.

Nello spazio di poche inquadrature interno ed esterno si rovesciano costantemente l’uno nell’altro in un movimento irregolare dagli esiti assolutamente imprevedibili.

Ecco allora che in un angolo del bagno compare per pochi istanti Cornelius, l’autore di *Drop*, figura-rinvio che, come accade spesso nei videoclip, non resiste alla tentazione di legittimarsi nei confronti del pubblico confondendo realtà e finzione, esasperando il meccanismo dei rimandi, delle simulazioni.

Il rimedio all’allucinazione è a portata di mano. Il bambino afferra una piccola bottiglia verde e riversa lentamente alcune gocce negli occhi. Lo stato alterato della percezione scompare immediatamente.



Fig. 8. Cornelius, *Drop*.



Fig. 9. Cornelius, *Drop*.

Nel finale del video l’acqua è ancora protagonista. Un’inquadratura singolare colloca lo sguardo dello spettatore sul fondo del lavandino e permette di ammirare nel dettaglio i gesti del bambino che ora non si limita a interrompere con le dita il flusso dell’acqua ma interviene sulla sua superficie con movimenti ampi e controllati fino a tradurre il ritmo e la melodia della musica sotto forma di gocce, schizzi, increspature e piccole onde.

Apparentemente *Drop* è solo uno tra i tanti video che giornalmente vanno in rotazione nei palinsesti delle emittenti musicali, tuttavia è evidente come in questo piccolo testo siano condensati numerosi elementi di interesse.

L'utilizzo musicale dei rumori che rovescia significativamente la tradizionale distinzione tra musica diegetica ed extradiegetica, la costruzione di un ritmo audiovisivo composto di anticipazioni, conferme, scarti, che trasforma il racconto di un gesto quotidiano in una sequenza di azioni fortemente emozionali, la rielaborazione originale della soggettiva, una figura-cardine del linguaggio cinematografico, che viene impiegata inizialmente per esasperare il piano visivo, e in seguito per rilanciare strategicamente l'appeal melodico del brano musicale.

In questo video il montaggio non è semplice esibizione tecnica, puro assemblaggio finalizzato a restituire uno stile alla sequenza di suoni-immagine, ma risponde a un'esigenza (promuovere un album) con una scommessa: pur di catturare l'attenzione dello spettatore esso non esita a relegare sullo sfondo, seppur temporaneamente, il suo elemento più significativo e prezioso, la musica.

### *Figure del ritmo e strategie di montaggio*

Nei video musicali il sincretismo di suoni e immagini risponde a una logica che in parte si distingue da quella cinematografica. In questi micro-testi audiovisivi non c'è quasi mai una narrazione sostenuta dal dialogo e la musica costituisce una componente autonoma poiché è sempre preesistente rispetto alle immagini. Di conseguenza le sequenze visive sono parzialmente slegate dalla linearità imposta dal suono.

Nei videoclip le combinazioni tra suoni e immagini non sono il prodotto di una semplice giustapposizione di sostanze dell'espressione differenti, ma l'effetto di una *sincretisi*, termine che esprime la riunificazione di due movimenti diversi, il *sincronismo* e la *sintesi*. La sincretisi

è la saldatura inevitabile e spontanea che si realizza tra un fenomeno sonoro e un fenomeno visivo puntuale quando questi accadono contemporaneamente, indipendentemente da ogni logica di tipo razionale (Chion 1990).

La sincesi non è un processo del tutto automatico: dipende infatti dal senso complessivo di un testo, e si organizza secondo leggi gestaltiche. Per far emergere con chiarezza il processo della sincesi è sufficiente accostare in modo del tutto casuale elementi visivi ed elementi sonori, in alcuni punti i suoni e le immagini daranno vita a combinazioni audiovisive estremamente efficaci, in altri punti resteranno del tutto separati, inconciliabili.

La sequenza si “fraseggia” da sola, in ragione di fenomeni di rinforzamento, di “buona forma” che non obbediscono a una semplice legge. Talvolta questa logica è evidente: quando si tratta di un suono più potente degli altri, esso si coagula con l'immagine che gli è sincrona più degli altri che lo precedono o lo seguono. Questo può essere un fenomeno di senso o di ritmo (ib.).

In una sequenza audiovisiva i punti in cui si realizza compiutamente l'unione tra elementi sonori ed elementi visivi sono definiti *punti di sincronizzazione*.

Nei videoclip esiste un rapporto elementare tra colonna audio e colonna video che si riduce alla presenza puntuale ma estremamente efficace di punti di sincronizzazione in cui l'immagine interviene a “mimare” la produzione del suono. Al di fuori di questi momenti di coincidenza audiovisiva le immagini e i suoni rispondono a logiche di sviluppo indipendenti. Nei video il montaggio è fatto sulla musica. In un certo senso si può dire che la musica costituisca il punto di partenza e di arrivo di un video musicale.

La sincesi, coordinando i tempi separati del suono e dell'immagine in funzione della riunificazione di sostanze dell'espressione diverse, costruisce la struttura ritmica del testo audiovisivo, contribuendo in maniera decisiva a orientarne la fruizione.

Nelle forme brevi dell'audiovisivo i punti di sincronizzazione scandiscono i momenti di massima solidarietà tra suoni

e immagini. Questi intervalli in cui il sincretismo audiovisivo esibisce tutte le proprie potenzialità espressive inscrivono nei testi dei punti di forza, suscitando nello spettatore una specifica disposizione percettiva in cui lo sguardo e la vista si influenzano reciprocamente e contribuiscono a trasformarsi.

Nei videoclip un punto di sincronizzazione diffuso è costituito dai *raccordi a stacco del suono e dell'immagine*. In questi casi la sintesi audiovisiva si realizza come una doppia cesura sincrona.

Un esempio rappresentativo di questo legame tra suoni e immagini è senz'altro quello dei video hip hop. In questo senso sono esemplari i lavori del regista Hype Williams (*She's a bitch, No scrubs*) in cui il ritmo frenetico imposto alle immagini dai numerosi tagli di montaggio, si aggancia, rilanciandolo, al ritmo dei brani musicali e ai gesti sincopati dei performer.

Tra le figure della sincronizzazione audio-visiva è possibile individuare innanzitutto quelle fondate su una *manipolazione esibita della sostanza visiva*, realizzate prevalentemente nella fase di postproduzione.

Uno dei primi espedienti visivi utilizzato nei videoclip per marcare il ritmo del brano musicale è stato senz'altro l'*alternanza tra il bianco e nero e il colore*. In questo senso sono esemplari i video realizzati nella seconda metà degli anni Ottanta dal regista Derek Jarman per la band inglese degli Smiths. In particolare il video di *How soon is now* (1986) in cui le immagini colorate e fortemente astratte di alcuni frammenti di pellicola vengono utilizzate come efficaci inserti visivi per ribadire il ritmo della canzone scandito dalle percussioni.

Un tipo di manipolazione dell'immagine molto frequente nella realizzazione della sincretisi audiovisiva è lo *sfocato*.



Fig. 10. Nirvana,  
*Heart Shaped Box*.

Si pensi al videoclip di *Smells like teen spirit* realizzato da Anton Corbjin per i Nirvana. Nel video viene messa in scena la performance del gruppo musicale all'interno di uno spazio estremamente ristretto e scarsamente illuminato. L'uso alternato di immagini nitide e sfocate contribuisce a restituire allo spettatore una sensazione di forte claustrofobia, ma soprattutto si rivela uno strumento straordinariamente efficace di valorizzazione del suono. Le sonorità estremamente dilatate dei Nirvana vengono infatti prese in carico e rilanciate dai continui passaggi tra le immagini perfettamente a fuoco degli ambienti e i dettagli sfocati del leader Kurt Cobain.

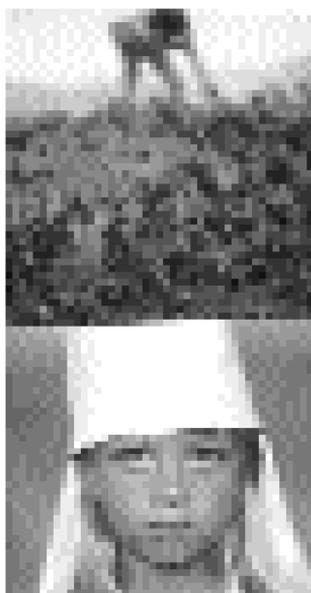


Fig. 11. Nirvana,  
*Heart Shaped Box*.

Nei videoclip l'uso dello sfocato si combina spesso con un altro tipo di manipolazione della sostanza visiva che consiste nell'intervenire direttamente sulla *grana delle immagini* alternando l'alta definizione alla bassa definizione.

Questo espediente diviene marca stilistica nei lavori realizzati da Floria Sigismondi, regista italo-canadese particolarmente sensibile alla sperimentazione tecnica ed espressiva sui regimi di visibilità.

In *Beautiful People*, videoclip realizzato per Marilyn Manson nel 1997, il ritmo veloce della canzone e le sonorità sporche si saldano sul piano visivo a brevi sequenze caratterizzate dalla rapida alternanza tra immagini nitide e sgranate.



Fig. 12. Tricky, *You make me wanna die*.

Nel video di *You make me wanna die* realizzato dalla regista per Tricky nello stesso anno, i suoni si sviluppano in modo uniforme e continuo scandendo in modo regolare il ritmo musicale. Sul piano visivo l'agganciamento con la musica si realizza tramite un movimento estremamente fluido di parziale defigurativizzazione: in sincrono con l'evoluzione regolare del flusso sonoro le immagini degli ambienti e del musicista vengono infatti private progressivamente di tutti i dettagli fino a esibire un livello prefigurativo in cui dominano linee confuse e macchie sfumate di colore.

Un intervento estremo di manipolazione visiva è il *morphing*, una figura molto usata nei video musicali che permette di "tradurre" in maniera efficace sul piano visivo la definizione del suono e la sua durata.

Nel pluripremiato video di *Black Hole Sun* realizzato da Howard Greenhalgh nel 1994 per i Soundgarden, i dettagli anatomici dei protagonisti vengono totalmente sfigurati in corrispondenza di suoni prolungati, regolari e ricchi di frequenze acute contribuendo in questo modo a drammatizzare fortemente l'ascolto del brano musicale.

Il repertorio di queste forme della sincronizzazione audiovisiva è estremamente vasto. Se le figure appena esaminate contribuiscono alla costruzione del ritmo agendo direttamente sulla qualità materiale delle immagini (grana, il-

luminazione, definizione), esiste tutta una serie di interventi che mirano a saldare i suoni e le immagini lavorando principalmente sulla riconfigurazione del tempo visivo.

Tra questi gli effetti di *velocizzazione e rallenti* che intervengono direttamente sulla dimensione ritmica del brano musicale. Un caso esemplare è il video di *Ray of light* realizzato da Jonas Akerlund nel 1998, in cui le immagini accelerate degli spazi metropolitani vengono utilizzate come “sfondo ritmico” per accompagnare la performance coreografica di Madonna. Un uso strategico della decelerazione caratterizza invece il video di *It's oh so quiet* realizzato da Spike Jonze per Bjork nel 1995. Il video sfrutta in maniera assolutamente originale la melodia del brano musicale mettendo in scena un vero e proprio micro-musical in cui il rallenti interviene a marcare su un piano visivo le pause ritmiche della canzone.

Due interventi ampiamenti utilizzati nella costruzione del ritmo audiovisivo sono il *congelamento dell'immagine* (*freeze frame*), puntuale intervento di sincronizzazione che consiste nel bloccare temporaneamente il movimento della macchina da presa o degli elementi profilmici per marcare una pausa significativa nel flusso sonoro, e il *frazionamento dello schermo in riquadri* (*split screen*) che permette di modulare in modo flessibile i tempi di ingresso delle immagini sui suoni.

Una figura emblematica del ritmo audiovisivo è il *loop*. La ripetizione della stessa sequenza di immagini è un espediente molto frequente soprattutto nei video realizzati per la musica dance in cui le ritmiche generate dagli strumenti elettronici (sequencer e drum machine) sono caratterizzate dalla ricorsività ossessiva di poche battute.

Nella costruzione del ritmo audiovisivo infine svolgono un ruolo fondamentale le *figure di avvicinamento*, come gli stacchi sull'asse e gli zoom.

Il ritmo del video di *Don't tell me* (2000), realizzato da J. B. Mondino per Madonna, ad esempio, è costruito proprio sull'uso combinato di zoom e stacchi sull'asse che marcano in modo estremamente rigido le linee evolutive della musica.



Fig. 13. Beck, *Mixed Bizness*.

Nello stesso anno anche il regista Stephan Sednaoui ha utilizzato in modo simile queste due figure della sincronizzazione in *Mixed Bizness*, di Beck. Nel video l'alternanza costante tra zoom in e zoom out viene impiegata per tradurre sul piano visivo l'andamento sincopato della canzone, un brano funky in cui prevalgono ritmi irregolari e scarti melodici.

### *Il ritmo contagioso delle forme brevi*

Il ritmo dei video musicali non è semplicemente il risultato di una coincidenza audiovisiva marcata, ma dipende in gran parte dal grado di prevedibilità dei punti di contatto tra elementi visivi e sonori.

Da un punto di vista orizzontale in una sequenza audiovisiva i suoni e le immagini non appaiono come elementi semplicemente allineati, distribuiti nel testo in modo tale da inscrivere al suo interno un andamento ritmico lineare e uniforme. Piuttosto, essi hanno delle tendenze, seguono leggi di evoluzione fatte di ripetizioni e scarti, che mirano a coinvolgere lo spettatore suscitando un senso di attesa e speranza, di saturazione da spezzare o di vuoto da riempire.

I punti di sincronizzazione stimolano lo spettatore a individuare consapevolmente o inconsciamente delle linee di sviluppo e a verificare in seguito se questa evoluzione innescata si realizza come previsto<sup>2</sup>.

Per il soggetto *sentire il tempo* non significa semplicemente ricondurre mentalmente i singoli elementi a una forma compiuta, ricomporre gli istanti in una durata. La sensazione del ritmo, la percezione della sua articolazione è innanzitutto un'esperienza timica, passionale.

Come afferma Jacques Geninasca (1992, p. 14) nei suoi studi dedicati all'analisi morfologica dei discorsi estetici "nell'esperienza stessa del ritmo il sentimento del tempo non si riduce alla semplice coscienza di un succedersi cronologico degli istanti ma si confonde con l'esperienza di una pienezza esistenziale".

Il ritmo, da un punto di vista teorico, si presenta dunque come "un sintagma organizzato di stati tensivi: attesa, distensione per l'attesa soddisfatta, sorpresa e disorientamento per un'attesa delusa, scoperta infine ed inversione euforica delle tensioni accumulate nella fase inventiva dello smarrimento" (ib.).

In tutte le forme testuali la successione di termini discreti (lessemi, figure, note musicali) a intervalli regolari è sufficiente a stabilire tra questi una relazione di dipendenza orientata. Questa ripetizione istituisce una *norma testuale* che dispone il fruitore in uno stato iniziale di attesa.

Ogni infrazione nei confronti di questa norma provoca sorpresa, suscita nel soggetto una delusione, non tanto per l'oggetto atteso che non si è manifestato, quanto piuttosto per l'errore nella previsione.

La sorpresa generata da un'aspettativa fallita viene vissuta in modo fortemente disforico, le sensazioni prevalenti sono il disagio e lo smarrimento, poiché "il mondo, il testo, quando non si conforma al contenuto dell'attesa, appare privo di coerenza e comunque non intelligibile" (p. 16).

Questo stadio intermedio costringe il soggetto a tornare sul testo, a insistere nella ricerca di un nuovo ordine, di una logica di sviluppo regolare cui ricondurre i singoli elementi.

La scoperta di una coerenza che per definizione si distingue dall'ordine ricercato inizialmente sfocia infine in un sentimento euforico. Per effetto di questa intelligibilità

riconquistata il soggetto riconfigura profondamente anche il suo rapporto con il tempo, “al senso di sgomento di chi è travolto dallo sgranarsi insignificante dei momenti succede quello di un tempo vivo e pieno per il senso di partecipazione all’ordine intelligibile che organizza il divenire” (ib.).

Geninasca definisce questo modello generale di organizzazione discorsiva sintagma seriale, proponendo di estenderne l’applicazione a “ogni enunciato compiuto e completo che abbia la forma di una serie finita di termini, qualsiasi sia la sua natura linguistica e la sua dimensione”.

Il sintagma seriale si presta a letture simultanee di natura diversa definite *prensioni* (*saisie*).

Nella riflessione di Geninasca il termine *prensione* rinvia al modo con cui il soggetto esperisce il mondo e i testi (che del mondo costituiscono una parte integrante). Le *prensioni*, pur disponendosi su un continuum graduato, possono essere ricondotte a tre forme distinte.

La *prensione molare* definisce un sapere di tipo associativo sociolettale o idiolettale. Questa prima forma di conoscenza permette al soggetto di individuare figure, configurazioni e percorsi figurativi di una semiotica del mondo naturale, così come concetti e insiemi concettuali. La *prensione semantica* deriva dalla comprensione della struttura semantica profonda dei testi. La *prensione ritmica o impressiva* infine esprime un modo del senso che si rivela compiutamente nell’esperienza di fruizione dei testi estetici. La *prensione ritmica* rinvia infatti a un rapporto pre-logico tra testo e soggetto in cui prevale il coinvolgimento sinestesico, in cui si colgono il rapporto sensibile, gli effetti passionali, il valore vissuto che il testo produce nel soggetto.

La nozione di sintagma seriale e la riflessione articolata sulle dinamiche di costruzione del ritmo audiovisivo, in particolare la descrizione dei modelli di agganciamento tra suoni e immagini, si rivelano estremamente utili per affrontare la questione dell’efficacia ritmica dei video musicali.

Uno degli espedienti maggiormente utilizzati nei video per suscitare quel senso di attesa che è indispensabile per

far presa ritmicamente sul soggetto è l'anticipazione controllata dei punti di sincronizzazione.

Così come in una canzone le prime note sono fondamentali per suscitare l'interesse dell'ascoltatore, in un video le combinazioni audiovisive iniziali ne condizionano fortemente l'efficacia. Non è un caso che l'*attacco audiovisivo* di queste forme brevi si imponga spesso all'attenzione dello spettatore esibendo una corrispondenza marcata tra suoni e immagini.

Si pensi all'attacco del video di *Drop* analizzato in precedenza, in cui l'immagine in dettaglio delle gocce e le sonorità regolari del brano musicale si riunificano sin dall'inizio in un'unica combinazione audiovisiva, inscrivendo all'interno del testo un punto di sincronizzazione che dà avvio al ritmo dell'intero videoclip, ne organizza l'evoluzione, momento incoativo che anticipa e prefigura un vero e proprio arco tensivo.

*Viorar vel til loftárása*, videoclip recente della band islandese dei Sigur Ross, è uno straordinario esempio di *solidarietà audiovisiva*.

Nel video viene messo in scena l'amore omosessuale tra due adolescenti in un paese di provincia tramite l'esibizione di un'originale sincretismo audio-visivo. Sin dalle prime inquadrature il video anticipa la propria strategia di sincronizzazione, su un tappeto sonoro rarefatto, in cui le sonorità del pianoforte e degli archi delineano un ritmo estremamente lento, si inseriscono in sincrono delle immagini fortemente rallentate: un ragazzo gioca con due bambole, il suo sguardo è rapito dai giocattoli che stringe al petto, alle sue spalle un secondo adolescente lo osserva incuriosito.

La sintesi tra la decelerazione delle immagini e il ritmo cadenzato del brano musicale drammatizza fortemente la prima sequenza del video, rendendone estremamente coinvolgente l'attacco. Questo sincronismo stretto tra musica e immagini viene ulteriormente rafforzato nel resto del video, in cui i risvolti drammatici della storia d'amore vengono narrati con una lentezza volutamente esasperante. *Vio-*

*rar vel til loftárása* rappresenta in modo esemplare la vitalità espressiva della forma breve-videoclip.

In assoluta controtendenza rispetto alle tradizionali forme di costruzione del ritmo audiovisivo, questo video dei Sigur Ross privilegia il rallenti all'accelerazione, le lunghe pause alle rapide sequenze di montaggio. In questo caso l'anticipazione controllata di punti di sincronizzazione stimola sin dall'inizio il soggetto a prevedere delle linee evolutive, ad anticipare i movimenti reciproci di suoni e immagini, precorrendo le cadenze di una sincesi marcata, apparentemente inevitabile.

Nell'ambito della produzione videomusicale *Viorar vel til loftárása* appare tuttavia come un caso isolato. Infatti, così come nella musica occidentale esistono le *cadenze evitate* (cadenze che inizialmente vengono anticipate tramite l'inflessione melodica e la progressione armonica, e poi improvvisamente evitate), nelle forme brevi dell'audiovisivo è molto frequente l'utilizzo di punti di sincronizzazione evitati, combinazioni tra suoni e immagini che si sottraggono in maniera del tutto imprevedibile a una coincidenza piena.

La mancata sincesi tra elementi sonori e visivi si rivela spesso efficace tanto quanto una completa riunificazione, poiché, intervenendo come uno scarto inaspettato nel flusso audiovisivo, modifica talvolta in maniera drastica il ritmo della sequenza, sorprendendo lo spettatore e mantenendone alto il livello di attenzione.

La desincronizzazione non si produce a partire da uno scollamento radicale di suoni e immagini; in altri termini, non ha origine da un montaggio video che "cade nel vuoto" ma è un effetto di senso che nasce dalla variazione improvvisa dei punti di contatto tra le inquadrature e le battute musicali. Ad esempio, se i tagli di montaggio si susseguono in corrispondenza del ritmo scandito da uno strumento come le percussioni, il cui suono si impone immediatamente all'attenzione dello spettatore, è sufficiente "staccare" le immagini su un suono differente o di sfondo, ritardando o anticipando l'aggancio audiovisivo, per produrre un effetto di desincronizzazione.

Contrariamente a quanto si pensa, nei videoclip la *desincronizzazione* è una componente fondamentale della costruzione ritmica. La sua diffusione, del tutto trasversale rispetto alle tradizionali distinzioni dei generi musicali, deve essere messa in relazione con l'evoluzione delle dinamiche di consumo del medium televisivo.

A questo proposito Chion sostiene che nei paesi in cui la televisione non ha diversificato la propria offerta in decine di canali attivi durante tutto l'arco della giornata, essa viene ancora percepita come un medium prevalentemente visivo, al contrario "laddove estende la propria programmazione a tutta la giornata e la propria presenza a diversi luoghi di lavoro e di abitazione, essa finisce inevitabilmente per assumere la propria natura radiofonica" (Chion 1990 p. 139).

In particolare, alcune fasce di puro flusso videomusicale si possono seguire "distrattamente", continuando a svolgere altre attività lavorative o ricreative, ma con la possibilità di gettarvi lo sguardo per cogliere qualche frammento audiovisivo.

In ciò che viene chiamato clip, vale a dire qualsiasi cosa di visivo messo su una canzone, si trova certamente di tutto, di tutti i budget e di tutte le qualità, ma talvolta anche cose superbe per vita e invenzione, in cui l'energia del disegno animato si combina con la presenza concreta della ripresa. In esso si inventa o si ritrova tutto un arsenale di procedimenti (ib.).

In questa complessa retorica dell'immagine la desincronizzazione colpisce direttamente lo spettatore, lo costringe a tornare allo schermo, a ricongiungersi con un ritmo imprevedibile che gioca sulla rottura dell'abitudine. Proprio questo è il paradosso della *televisione a immagine facoltativa* (ib.): libera gli occhi.

La televisione non è mai tanto visiva quanto in questi momenti di programmazione di clip, proprio nel momento in cui l'immagine va ad aggiungersi apertamente a una musica che bastava già a se stessa (p. 140).

Tra i numerosi interventi che permettono di tradurre sul piano visivo il ritmo del brano musicale la soluzione che si avvicina maggiormente alla simultaneità polifonica dei suoni e della musica è senza dubbio la *rapida successione delle immagini singole*. Questo espediente di montaggio stimola infatti la memoria dello spettatore a funzionare come un “perfetto mixer che è in grado di miscelare, molto meglio di una macchina, impressioni visive concatenate le une alle altre nel tempo” (ib.).

In questo, paradossalmente, il videoclip rivela una prosimità sorprendente con il cinema muto,

è proprio nella misura in cui vi è una musica alla base, e non vi è narrazione sostenuta da un dialogo, che l'immagine è totalmente slegata dalla linearità imposta dal suono. Se i clip funzionano, è sicuramente perché c'è un rapporto elementare tra colonna audio e colonna immagine, e le due non sono del tutto indipendenti (ib.).

I videoclip, facendo leva “su immagini e suoni puntuali che non hanno, alla lettera, nulla a che vedere gli uni con gli altri, e che formano nella percezione degli agglomerati mostruosi, ma irresistibili e inevitabili” (p. 59), rivelano dunque un'efficacia sorprendente.

Ciò che rende estremamente seducenti questi micro-testi è la capacità di tradurre in immagini il ritmo della musica, che per definizione è

un'arte passionale, poiché in essa il trattamento della disposizione del tempo – che tanto ha a che fare con la passione – è del tutto cruciale. (...) la passionalità è l'immediata forma del contenuto che ha come forma di espressione il ritmo della musica (Fabbri 1998, p. 46).

In questo senso si rivela del tutto infondata l'idea secondo cui la musica non sarebbe un linguaggio poiché priva di un contenuto. La musica piuttosto deve essere considerata come un sistema ritmico, o meglio, come una

“forma di organizzazione ritmica che ha come forma del contenuto certe determinate passioni, cioè i ritmi che sono costitutivi delle emozioni” (ib.).

Nei video, diversamente da quanto accade nel cinema, è al lavoro una prassi ritmica che mira a suscitare l'interesse dello spettatore a prescindere dal suo coinvolgimento cognitivo, rivendicando piuttosto una reazione corporea immediata. In questo senso i videoclip si offrono a una prensione impressiva prima ancora che semantica, un'esperienza del senso che si realizza nell'esplorazione del proprio sentire, nelle risonanze passionali che agiscono sul corpo del soggetto sottoponendolo a tensioni e distensioni.

I video sono complessi micro-testi in cui temi e figure vengono presi in carico e rilanciati da forme ritmiche complesse, risultato sorprendente di una pratica di montaggio in cui l'esplorazione delle potenzialità espressive del sincretismo audiovisivo prevale sulla costruzione di effetti di realtà.

Queste forme brevi rivelano dunque una straordinaria capacità di manipolare il soggetto, agendo in particolare sulla sua *sensomotricità* (Fontanille 1999).

Per effetto del montaggio verticale suoni e immagini danno vita a configurazioni plastiche talvolta estremamente complesse che agiscono sul soggetto, mettendo in moto il suo corpo, sollecitando la reazione della sua carne.

I video non si limitano a suscitare attenzione, reclamano piuttosto un coinvolgimento immediato, una partecipazione attiva, che si realizza compiutamente nel gesto quotidiano, ma denso di significato, di tenere il ritmo con i piedi o di battere il tempo con piccoli movimenti delle mani o della testa. L'efficacia seduttiva di queste forme brevi è tutta qui, nella capacità di costringere il soggetto a prendere in carico sul proprio corpo un ritmo travolgente, nell'armonia disarmante con cui i battiti e le pulsioni del nostro sentire individuale si accordano a brevi sequenze di suoni-immagine.

<sup>1</sup> Bisogna inoltre tenere presente che per sincronia non intendiamo una necessaria consonanza. Per noi è del tutto lecito qualsiasi gioco di coincidenze e non-coincidenze del “movimento”, alla sola condizione che il nesso sia stabilito e motivato da un punto di vista compositivo (p. 139).

<sup>2</sup> Spesso, evidentemente, è più interessante quando la tendenza innescata viene contrastata, talvolta quando tutto va come previsto, la dolcezza e la perfezione nella realizzazione dell'anticipazione bastano alla nostra emozione (Chion 1990, p. 44).