

UNIVERSITÀ DI PADOVA



FACOLTÀ DI INGEGNERIA

TESI DI LAUREA

**SISTEMI INFORMATICI PER L'ESECUZIONE
ESPRESSIVA DELLA MUSICA**

Corso di Laurea in Ingegneria dell'Informazione

Laureando: Dany Nichele

Relatore: Prof. Sergio Canazza

Anno Accademico 2012/2013

Sommario

Scopo di questa tesi è presentare alcune modalità per rendere possibile l'esecuzione automatica espressiva della musica.

Nel primo capitolo, si illustrano diverse metodologie che permetteranno di ricavare un modello matematico di un'esecuzione espressiva musicale; sono esposti i metodi di analisi per misura, per sintesi, *machine learning* e *case based reasoning*.

Nel secondo capitolo, si presentano varie tipologie di spazio che sono usate per la rappresentazione dell'espressività contenuta all'interno di un brano. Spazi espressivi di tipo discreto e continuo, con dimensione superiore a due; saranno, inoltre, riportate le caratteristiche, le limitazioni e le casistiche di utilizzo degli stessi.

La trattazione si conclude con il terzo capitolo, nel quale si cerca di trovare una relazione tra caratteristiche musicali ed emozioni; sono esposti, quindi, i parametri fondamentali su cui si deve agire per riuscire a trasmettere all'ascoltatore determinate emozioni.

Ringraziamenti

Desidero ringraziare tutti coloro che hanno permesso la realizzazione di questo elaborato, partendo dal Professor Sergio Canazza, per i preziosi insegnamenti, consigli e per il tempo dedicato alla mia tesi. Ringrazio i miei genitori, che mi sostengono da ventidue anni in tutto e per tutto, insieme a Igor, che per me non è soltanto un fratello. Proseguo ringraziando i miei amici, grandi compagni di feste e di vita; presenti anche nei momenti di bisogno. A questo proposito vorrei ringraziare, in maniera particolare, Andrea B. e Francesco F. Un ringraziamento anche agli amici del coro, Academia Ars Canendi, e della scuola di musica con cui ho passato momenti indimenticabili e grandi emozioni, in concerto e in compagnia. Infine, desidero ringraziare, i compagni di corso, con i quali ho condiviso la mia esperienza universitaria; ringrazio, inoltre, tutti coloro che mi hanno dato delle lezioni di vita, che hanno saputo ascoltarmi quando ne avevo più bisogno, che erano presenti regalandomi un sorriso e dandomi la forza di andare avanti; ringrazio anche quelli che hanno ostacolato la mia strada, rendendomi, così, ancora più forte.

Dany Nichele

Indice

Sommario	i
Ringraziamenti	iii
1 Sistemi automatici per l'esecuzione espressiva della musica	1
1.1 Modelli automatici per l'esecuzione	2
1.1.1 Metodo di analisi per misura	3
1.1.2 Metodo di analisi per sintesi	3
1.1.3 Machine Learning	4
1.1.4 Case-Based Reasoning	4
2 Modelli per la rappresentazione di emozioni ed espressività	7
2.1 Spazi espressivi	7
2.1.1 Spazio emozionale discreto	7
2.1.2 Spazio emozionale continuo	9
2.1.3 Spazi emozionali con dimensione superiore a due	13
3 Struttura musicale ed emozioni	17
3.1 Emozioni e caratteristiche musicali	17
3.2 Emozioni ed esecuzione	20
4 Considerazioni finali	23
Bibliografia	25

Elenco delle figure

1.1	Raffigurazione del processo Case-Based Reasoning	5
2.1	Gli otto <i>clusters</i> emozionali introdotte da Hevner	9
2.2	Aggiornamento della lista di Hevner Shubert (2003)	10
2.3	Rappresentazione stimoli musicali nel piano <i>valence-arousal</i> Vieillard et al. (2008)	12
2.4	Rappresentazione stimoli musicali nel piano <i>valence-arousal</i> (Russel) Yang and Chen (2012)	13
2.5	Rappresentazione dei 24 termini emozionali nello spazio quadri-dimensionale Fontaine et al. (2007).	14
3.1	Riassunto dell'analisi di regressione Eerola et al. (2012)	18
3.2	Dipendenza da <i>valence-energy</i> al variare dell'articolazione Eerola et al. (2012)	19
3.3	Comunicazione emozionale dell'esecutore.	20

Capitolo 1

Sistemi automatici per l'esecuzione espressiva della musica

Introduzione

Uno dei principali motivi per cui si prende parte ad attività musicali, componendo, eseguendo, o semplicemente ascoltando è dato dal fatto che la musica è capace di suscitare delle emozioni profonde e significative. La musica, però, non si ferma solo a trasmettere emozioni, la si trova, infatti, utilizzata in svariati campi e forme, ad esempio come mezzo didattico, di supporto in caso di disabilità; usata anche nell'ambito della sonificazione¹. Nella trattazione seguente ci si limiterà a un'analisi sul contenuto emozionale trasmesso con l'esecuzione musicale. Le emozioni, all'interno di un brano musicale, possono andare da un puro godimento estetico per un costruito sonoro, alla gioia o alla malinconia, che la musica a volte evoca o sostiene, al semplice sollievo dalla monotonia, dalla noia, dalla depressione, che le esperienze musicali quotidiane possono fornire.

La musica, inevitabilmente, ha la capacità di elevare il livello della vita emotiva. Ovviamente, individui e società fanno uso della musica anche per altri motivi, alcuni di essi già ricordati prima. Poiché molte attività musicali hanno anche un carattere sociale, possono acquisire molti significati anche su questo piano, offrendo così ricompense sociali a chi vi partecipa. Per esempio, la conoscenza di certi tipi di musica è un prerequisito per essere considerati membri a pieno titolo di numerose sub-culture.

Se i fattori emotivi sono fondamentali per l'esistenza della musica, diventa allora spontaneo chiedersi come la musica riesca a influire sulle persone. A prima vista, un evento musicale può sembrare solo una raccolta di suoni di varia altezza, durata, e altre qualità misurabili. Se si pensa però al "suono organizzato" di Varèse, che consiste in un procedimento compositivo attentissimo all'auscultazione dei fenomeni sonori più sottili, alle interferenze provocate da un determinato

¹L'idea alla base della sonificazione è quella di utilizzare dei procedimenti matematici per tradurre una successione di punti (ad esempio un grafico) o di colori (ad esempio una mappa) in una successione di note, in modo da generare un segnale percettibile con l'udito, dando così all'utilizzatore un'idea in tempo reale della situazione senza bisogno che questo interpreti un grafico, azione che comporterebbe un dispendio più tempo oltre che una capacità di lettura dello stesso.

incontro di timbri, agli agglomerati dei suoni, agli armonici superiori esaltati dall'intervento di un certo strumento piuttosto che un altro; si capisce che un'esecuzione musicale è un insieme di molti aspetti, non tutti così evidenti. La mente umana attribuisce a questi suoni un significato, si cercherà, quindi, di creare una corrispondenza biunivoca tra caratteristiche intrinseche del pezzo musicale e uno spazio, che meglio possa rappresentare il contenuto emozionale del pezzo; si analizzerà il rapporto che esiste tra musica ed emozioni. La struttura musicale fornita dalla partitura unita all'espressività data dall'esecutore diventano un elemento unico, che farà piangere o ridere, che piacerà o meno, che commuoverà o lascerà indifferente l'ascoltatore.

L'espressività di un brano musicale si trova, quindi, in diversi livelli di analisi; in prima battuta si trova il messaggio che il compositore intende trasmettere attraverso la scrittura dello spartito, con l'ausilio di segni espressivi riportati nella partitura. Si prosegue poi con l'interpretazione data dall'esecutore introducendo, intenzionalmente o meno, delle deviazioni rispetto alla partitura originale, saranno proprio queste deviazioni oggetto di studio. Si cercherà di costruire dei modelli computazionali che riescano a inglobare l'entità, il momento e il motivo dell'introduzione da parte dell'esecutore di queste variazioni; una volta costruito tale modello, sarà possibile effettuare un'esecuzione del brano automatica che tenga anche conto dei parametri espressivi. All'ascoltatore è lasciato il compito di interpretare il suono cercando di estrapolarne il contenuto emozionale.

1.1 Modelli automatici per l'esecuzione

La caratteristica che sta alla base di un sistema automatico per l'esecuzione espressiva della musica, è la capacità di convertire una partitura musicale in una performance che include deviazioni di tempo, intensità, timbro, e altre caratteristiche che non sono riportate esplicitamente all'interno della partitura. Sebbene lo spartito sia il mezzo per comunicare l'intenzione espressiva del compositore, questo non contiene una descrizione completa dell'espressività, infatti, sono riportati nella partitura solo informazioni riguardanti la struttura ritmica e melodica del pezzo; elementi, ad esempio come: tempo e timbro sono a discrezione dell'esecutore. Per riuscire, quindi, ad avere una performance automatica partendo da una partitura, si dovranno costruire dei modelli che rappresentino l'espressività da dare al brano, unita poi alle informazioni provenienti dallo spartito, renderà possibile l'esecuzione espressiva tramite l'ausilio di mezzi automatici come un calcolatore. Il controllo espressivo è diventato, così, un'area di grande interesse nel campo del *Sound and Music Computing*² e del *Music Information Retrieval*.³ Generalmente, le principali strategie per progettare un sistema di esecuzione automatica, sono il metodo di analisi per misu-

²Gruppo di ricerca in ambito scientifico, educativo e di diffusione di discipline legate all'applicazione di nuove tecnologie alla musica e al suono. Le attività del gruppo sono sempre state basate su un approccio interdisciplinare attraverso la collaborazione tra ricercatori e musicisti. Le attività di ricerca principali sono: sintesi sonora basata su modelli, rendering 3D, analisi e modellizzazione di contenuto emozionale ed espressivo in esecuzioni musicali e altri.

³MIR-Tecniche di Music Information Retrieval vengono usate al fine di estrarre caratteristiche significative direttamente dal segnale musicale e permettere all'utente di interagire con il sistema per mezzo di interfacce di alto livello come ad esempio selezionare tutti i contenuti multimediali aventi lo stesso tempo o altre proprietà desiderate.

ra e il metodo di analisi per sintesi. Di recente si sono utilizzate anche tecniche che fanno ausilio dell'intelligenza artificiale (*machine learning e case-based reasoning*).

1.1.1 Metodo di analisi per misura

Il primo metodo, analisi per misura, è basato sull'analisi delle deviazioni che contraddistinguono un'esecuzione fatta da un essere umano. Lo scopo è quello di cogliere le regolarità presenti all'interno di esecuzioni espressive e descriverle poi mediante modelli matematici come viene esposto in Repp (1992) e ripreso ad esempio in Yang and Chen (2011): viene usato un modello di regressione per ottenere una densità di probabilità che descrive l'espressività di un pezzo. Il metodo di analisi per misura inizia dalla selezione del tipo di performance. L'esecutore potrà eseguire il pezzo in maniera libera oppure gli verrà imposto un certo tipo di espressività da riprodurre; dipenderà dallo scopo dello studio il grado di libertà della performance. In esecuzione si misureranno le proprietà fisiche di ogni singola nota. I parametri del brano misurabili, e quindi soggetti a variazioni durante l'esecuzione, sono molteplici: durata, involuppo temporale, frequenza, vibrato ecc. La decisione su quante variabili considerare e la modalità di misurazione dipenderà dallo scopo dell'esperimento, dallo strumento utilizzato e dalle proprietà tecniche dello stesso. Una volta misurate le variabili fisiche della performance è necessario valutare affidabilità e consistenza dei dati misurati, classificando le esecuzioni in diverse categorie a seconda dei dati raccolti. Seguirà poi la scelta delle variabili più significative al fine della ricerca, le quali saranno usate per formulare un modello matematico. Il numero di variabili da considerare ai fini dello studio è sempre un punto delicato e ancora un problema aperto in quanto varia da caso a caso. Talvolta si esegue un'analisi multidimensionale, considerando quindi un maggior numero di variabili, con l'obiettivo di identificare dei pattern indipendenti, si rimanda questa trattazione al capitolo successivo, in cui verranno discusse le caratteristiche di vari spazi usati per la rappresentazione dell'espressività trasmessa con l'esecuzione di un brano.

1.1.2 Metodo di analisi per sintesi

Il metodo di analisi per sintesi prende in considerazione diverse versioni del brano in cui le variabili, scelte per la ricerca (intensità, durata, ecc.), variano in maniera sistematica. Si crea un modello matematico dell'intera performance, viene poi formulato un giudizio basato sulla versione modellizzata del brano in relazione agli aspetti selezionati; bisogna, ovviamente, conoscere le variabili in gioco e la loro scala di rappresentazione. Del giudizio espresso dall'ascoltatore, deve essere valutata l'affidabilità e, inoltre, i vari ascoltatori, devono essere classificati in gruppi. All'inizio di una ricerca, quando vengono reclutati degli ascoltatori per avere un riscontro riguardante espressività percepita, viene loro sottoposto, preliminarmente, un questionario nel quale ogni persona è tenuta a rispondere ad alcune domande (età, sesso, nazionalità, esperienze musicali, ecc.); riuscendo così a collocare ogni individuo in un determinato insieme di appartenenza. Come evidenziato in Rentfrow and Gosling (2003), la personalità incide sulle preferenze musicali e quindi anche sulla percezione e valutazione del contenuto emotivo. Il giudizio sull'espressività è influenzata, inoltre, dalla cultura di appartenenza; come trattato in un confronto transculturale in Balkwill et al. (2004). Il metodo di analisi per sintesi procede osservando la differenza presente

tra le variabili del modello matematico creato e i dati reali, nel caso non ci sia coincidenza dei risultati si esegue una modifica delle variabili considerate durante lo studio. La procedura con il metodo di sintesi continua, iterando il processo, fino a quando le variabili selezionate per lo studio convergono alla situazione reale. Il modello matematico così costruito fornirà una buona rappresentazione della situazione reale, proprio perché creato tramite aggiustamenti progressivi delle variabili da considerare ai fini della ricerca.

1.1.3 Machine Learning

Un modello basato sul riconoscimento automatico delle emozioni musicali tramite il metodo di *machine learning*, è sviluppato cercando di trovare una relazione tra caratteristiche musicali e valori emozionali percepiti dal brano. Si procede, pertanto, con il formulare delle ipotesi iniziali sugli aspetti della performance che si vogliono modellare; successivamente si cerca una corrispondenza del modello testandolo con dati provenienti da situazioni reali. Come si tratterà nel capitolo successivo, si è visto che, assegnando semplicemente un valore alle emozioni, la valutazione, sul contenuto emotivo, sarà diversa a seconda dell'individuo interpellato. Per risolvere questo tipo di problema viene introdotto da Yang and Chen (2011) un nuovo approccio che rappresenta le emozioni con una distribuzione di probabilità, in uno spazio che meglio possa rappresentare il contenuto emotivo del pezzo. Sviluppando, inoltre, una metodologia che ricava la distribuzione dell'emozione contenuta all'interno di una clip, tramite la stima da campioni di tipo discreto, utilizzando tecniche di tipo *machine learning*; un algoritmo di fusione integrerà poi diverse caratteristiche dell'espressività, trovandone una rappresentazione opportuna, potenziando, inoltre, il modellamento delle emozioni percepite. Lo scopo, delle tecniche di *machine learning*, è di scoprire complesse dipendenze in un insieme molto grande di dati, senza formulare ipotesi preliminari e, quindi, non esplicitando la natura dei dati, si ha quindi la possibilità di ottenere nuove informazioni evitando qualsiasi assunzione di carattere musicale.

1.1.4 Case-Based Reasoning

Il Case-Based Reasoning (CBR) è basato sul confronto tra casi che presentano tra loro delle affinità. Con questo tipo di approccio, quando si deve risolvere una problematica, si va alla ricerca all'interno di una banca dati di un problema simile già risolto, pensando che questo possa avere uguale soluzione o comunque molto simile. Effettuando una ricerca e indicando al sistema le caratteristiche fondamentali della problematica in questione, il CBR interviene confrontando le informazioni fornite dall'utente con le informazioni presenti nell'archivio (come mostrato in *Figura 1.1*); questi sistemi infatti, sono in grado, analizzando un elevato numero di casistiche, di generalizzarne le caratteristiche e individuare le similitudini tra casi diversi. In questo modo riescono a proporre come risultato della ricerca i contenuti che saranno utili per la risoluzione del problema. Si fa uso di questa metodologia, ad esempio nella ricerca Zeng et al. (2009), per il riconoscimento dell'espressività in maniera multimodale. Si incrociano cioè dati provenienti da espressioni spontanee del viso, movimenti della testa e movimenti del corpo. Vengono cercati, all'interno di un database, problemi simili a parità di dati raccolti al fine di riconoscere lo stato emotivo dell'individuo. Nel caso si riesca a trovare una soluzione non esistente di un determinato

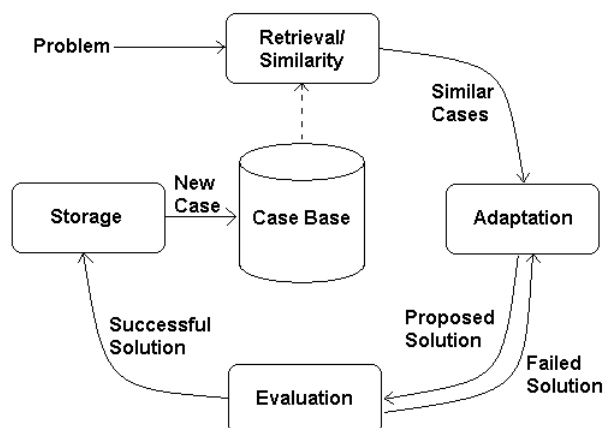


Figura 1.1: Rappresentazione del processo Case-Based Reasoning

problema, questa verrà inserita all'interno della base di dati; potendo, così, essere utile per la risoluzione di problematiche future. Il sistema CBR è utile quando si ha a disposizione un insieme considerevole di problemi già risolti, disporre però di un archivio di soluzioni in una certa quantità e oltretutto organizzati in maniera efficiente, in modo da favorirne la ricerca, risulta essere difficile.

Capitolo 2

Modelli per la rappresentazione di emozioni ed espressività

2.1 Spazi espressivi

Generalmente è impossibile comporre, eseguire o ascoltare un pezzo musicale senza avere un coinvolgimento di tipo affettivo. Si andrà prima di tutto ad analizzare vari metodi e approcci per riuscire a rappresentare le sensazioni provate. Un problema centrale del riconoscimento automatico delle emozioni riguarda la modellizzazione dello spazio espressivo, questo può avere una rappresentazione di tipo discreto oppure continuo.

Per quanto riguarda l'approccio discreto al problema, si procede con la divisione delle emozioni in un certo numero di categorie; tramite tecniche di *machine learning*¹ si attribuisce un'etichetta che caratterizzi ciascuna categoria. L'approccio di tipo continuo usa valori numerici plottati nello spazio bidimensionale, tridimensionale o in generale multidimensionale a seconda di quanta accuratezza si vuole dare al modello; lo spazio creato sarà comunque di tipo continuo. Tramite un modello di regressione sarà possibile rappresentare il contenuto emozionale in uno spazio che meglio lo rappresenti; di seguito sono esposte pertanto le varie tipologie di spazio emozionale, sottolineandone le rispettive proprietà.

2.1.1 Spazio emozionale discreto

Le esperienze che portano a delle emozioni, possono essere considerate come eventi distinti gli uni dagli altri; come affermato da: Barrett (1998); Yang and Chen (2012). Alla base di questo approccio c'è il concetto di emozione primaria, emozione cioè che non si può scomporre in altre emozioni più semplici. Esiste pertanto un certo numero limitato di emozioni universali considerate classi primarie come: *happiness, sadness, anger, fear, disgust* e *surprise*; secondo Ekman oppure secondo Izard: *happiness, sadness, hostility, guilt, surprise* e *interest*; tramite combinazione di emozioni primarie si possono ricavare tutte le altre emozioni definite secondarie. Esistono

¹Tramite l'apprendimento automatico si cerca di estrapolare nuovi elementi in comune e complesse dipendenze su insiemi molto grandi di dati; (si veda § 1.1.3)

molte varianti alle classi prima citate, proprio per la difficoltà di riuscire a capire quali siano le emozioni che stanno alla base e quali loro combinazioni. La rappresentazione discreta considera le emozioni come eventi che si manifestano in un individuo in modo consequenziale, almeno per un intervallo di tempo abbastanza piccolo. Come evidenziato da Barrett (1998), alcune esperienze riportano un alto grado di correlazione tra due diverse classi di emozioni considerate come primarie, questo denota che in determinate circostanze, una persona non è in grado di separare nettamente due diverse emozioni; si sottolinea, infatti, l'alto grado di correlazione tra lo stato d'animo *anxiety e depression*, appartenenti concettualmente a due classi di emozioni discrete distinte. Questo tipo di approccio è particolarmente adatto in soggetti che non si limitano a valutazioni di sola positività o negatività dell'evento, ma considerano anche il grado di attivazione della manifestazione. È però, evidente, il problema di granularità e ambiguità; lo scopo di un modello dovrebbe essere quello di rappresentare il più fedelmente possibile i dati reali, in questo caso, però, la soggettività dell'emozione e vincoli imposti dal linguaggio non permettono di avere una rappresentazione che rispecchi appieno la situazione reale. A parità di contenuto emozionale il giudizio da parte di individui diversi potrebbe non essere lo stesso.

L'approccio discreto al problema si presenta pertanto piuttosto limitativo, poteva però risultare comodo da un punto di vista computazionale: inserimento delle etichette, quindi stringhe rappresentanti le emozioni, in basi di dati facili da gestire. La modellizzazione discreta diventa ancora più limitativa se si pensa a una difficoltà di interpretazione degli aggettivi espressi in lingue diverse dalla propria, dando così sfumature semantiche ancora più ampie. È pertanto opportuno, analizzare altri tipi di approcci, che riescano a risolvere le problematiche evidenziate; prima di passare però all'analisi di tipo continuo sarà presentato un altro tipo di approccio, che tenta di eliminare le ambiguità semantiche, raggruppando in insiemi gli aggettivi che descrivono emozioni simili.

Clusters emozionali

Un altro tipo di approccio di tipo discreto è la *checklist* di aggettivi introdotta da Hevner (1936). Attraverso uno studio, Hevner, introdusse otto insiemi contenenti aggettivi con lo stesso contenuto emozionale, disponendoli in maniera circolare, come mostrato in *Figura 2.1*. Spostandosi da *cluster a cluster*, in maniera circolare, crescerà la differenza del contenuto emotivo tra i vari insiemi, fino a raggiungere il massimo della discrepanza nella posizione diametralmente opposta. La divisione di Hevner (1936) fu successivamente rivisitata da Farnsworth (1969) e recentemente cambiata con l'aggiunta di una nuvola in più per un totale di nove da Shubert (2003). Sedici aggettivi sono stati cancellati dalla lista originaria che ne conteneva 67, poiché non rispecchiavano appieno la semantica di nessun insieme. Farnsworth (1969) aveva introdotto dieci *clusters*, ma il decimo insieme, conteneva solo l'aggettivo *frustrated*: è stato pertanto eliminato nella successiva revisione. Gli *clusters* contengono, appunto, più aggettivi proprio per eliminare il più possibile l'ambiguità; la presenza di un insieme con solo un termine si discosta dalla struttura del metodo in divisione per *clusters*. Nella rivisitazione di Shubert (2003), viene chiesto ai partecipanti della ricerca di esprimere, tramite un indice numerico da 0 a 7, quanto gli aggettivi, presentati in ordine casuale, potessero descrivere il brano ascoltato. Gli aggettivi che ottenevano una valutazione media minore di 4 venivano scartati da quella che doveva essere la lista finale. Raccolti tutti i dati



Figura 2.1: Gli otto clusters emozionali introdotte da Hevner

ed eliminata, come detto in precedenza, il decimo *cluster*, quello che si ottiene è la lista mostrata in *Figura 2.2*. Si nota, osservando la tabella, che sono state invertite gli insiemi H e I, l'insieme E posizionato dopo il G, si possono pertanto disporre i vari insiemi, partendo dal primo quadrante, in modo circolare nella successione: A, B, C, D, F, G, E, I, H; il *cluster* H si ricongiunge con la A per chiudere il cerchio. Il *cluster* G, per problemi di simmetria tra i quadranti e per mantenere una forma circolare uniforme, è stata omessa dalla rappresentazione grafica. Sette aggettivi, che preventivamente erano stati tolti da Farnsworth, sono stati ripristinati. Uno dei problemi dell'approccio discreto, consiste nel numero delle classi delle emozioni primitive, che è di gran lunga inferiore in comparazione alla ricchezza delle emozioni musicali percepite da un individuo. Usare una granularità più alta non risolve il problema perché il linguaggio utilizzato, per descrivere le emozioni, è intrinsecamente ambiguo e varia da persona a persona. Diventa impensabile, per un'analisi della psicologia delle emozioni, aumentare il numero di classi di suddivisione. Si presenta ora un approccio di tipo continuo, permettendo, così, una caratterizzazione più puntuale delle emozioni.

2.1.2 Spazio emozionale continuo

Mentre l'approccio di tipo discreto si focalizza principalmente nel distinguere un'emozione da un'altra, l'approccio dimensionale continuo si focalizza sul posizionare le emozioni in un grafico cartesiano di dimensione bassa, solitamente due o tre dimensionale, con l'intento di rappresen-

UPDATED HEVNER ADJECTIVE LIST							
	M	SD	Source		M	SD	Source
Cluster A				Cluster F			
Bright	6.04	1.63	HF,6,a	Dark	5.84	1.95	HF,2,f
Cheerful	5.35	1.82	HF,6,a	Depressing	4.64	2.31	HF,2,f
Happy	5.29	1.97	HF,6,a	Gloomy	4.64	2.18	HF,2,f
Joyous	5.22	2.01	HF,6,a	Melancholy	5.36	2.17	H,2,f
Cluster B				Mournful	4.64	1.76	HF,2,f
Humorous	4.67	2.25	H,5,	Sad	5.53	2.02	HF,2,f
Light	6.02	1.73	HF,5,b	Solemn	5.09	2.11	HF,1,f
Lyrical	6.23	1.74	HF,4,c	Cluster G			
Merry	4.30	1.88	HF,6,a	Heavy	5.61	1.72	H,2,
Playful	5.17	2.01	HF,5,a	Majestic	5.81	1.87	HF,8,g
Cluster C				Sacred	4.92	2.35	HF,1,g
Calm	6.06	1.76	HRW,4,	Serious	5.08	2.15	HF,1,
Delicate	5.13	2.15	HF,5,c	Spiritual	5.30	2.06	HF,1,g
Graceful	6.02	1.71	HF,5,c	Vigorous	5.06	1.84	HF,8,g
Quiet	5.69	1.97	HF,4,c	Cluster E			
Relaxed	5.04	2.05	RW,,	Tragic	4.71	1.99	HF,2,f
Serene	4.66	2.15	HF,4,	Yearning	4.10	2.39	HF,3,c
Soothing	5.69	1.92	HF,4,c	Cluster I			
Tender	4.26	2.18	HF,3,c	Agitated	4.86	2.06	HF,7,I
Tranquil	5.29	1.97	HF,4,c	Angry	4.07	2.53	RW,,
Cluster D				Restless	4.63	2.23	H,7,
Dreamy	4.79	1.89	HF,3,d	Tense	5.71	1.77	RW,,
Sentimental	4.48	2.18	HF,3,d	Cluster H			
				Dramatic	6.54	1.52	HF,7,h
				Exciting	5.58	1.96	HF,7,h
				Exhilarated	4.27	2.08	HF,7,h
				Passionate	5.92	2.00	H,7,
				Sensational	4.83	2.31	H,7,
				Soaring	4.25	2.27	H,7,
				Triumphant	5.27	2.03	HF,7,r

Figura 2.2: Aggiornamento della lista di Hevner Shubert (2003)

Note- Le colonne mostrano: la media delle valutazioni date in una scala da 0 a 7 (M), la deviazione standard (SD), l'origine dell'aggettivo (H = lista aggettivi di Hevner, F = Farnsworth, R = Russel, W = Whissell), indicate poi il numero di originario di appartenenza alle nuvole di Hevner e la lettera originaria di appartenenza delle nuvole di Farnsworth. La sequenza delle nuvole è A, B, C, D, F, G, E, I, H per ritornare poi alla A.

tare le emozioni umane. La dimensione interna, dello spazio che andrà a rappresentare le varie emozioni, viene trovata analizzando la correlazione tra i termini affettivi. A questo fine viene svolta una ricerca e ai soggetti sottoposti alla stessa è richiesto di quantificare con un numero, per esempio utilizzando la scala Linkert a 7 punti (0 = non presente, 3 = moderatamente presente, 6 = certamente presente), l'ammontare di emozione descritta da un certo aggettivo contenuta nello stimolo musicale; questa procedura viene ripetuta per un numero abbastanza grande (60-80) di aggettivi, in modo da formare un quadro abbastanza completo. Tramite poi, tecniche di analisi, vengono ricavate un numero di dimensioni fondamentali, solitamente due o tre; successivamente si plotteranno i dati nel sistema cartesiano creato. Data la complessità nel rappresentare lo spazio espressivo, i modelli continui usati sono molteplici, caratterizzati da dimensioni, parametri misurati e approccio di misurazione diversi.

Valence and Arousal

Lo spazio espressivo continuo *valence-arousal* è uno spazio bidimensionale nel quale le componenti fanno riferimento a delle esperienze soggettive. Le due dimensioni sono considerate essenziali e indipendenti, oltre che ad essere il più possibile ortogonali. La prima componente cartesiana è *valence*, rappresenta quanto i sentimenti di una persona sono influenzati dalle valutazioni positive o negative di persone, cose o eventi; tipicamente posta sull'*asse x*, con valenza positiva verso destra. La seconda componente è *arousal*, rappresenta il grado di attivazione, cioè quanto una persona è disposta a compiere delle azioni o meno, quest'ultima posta sull'*asse y*, con attivazione positiva al crescere con il verso convenzionale dell'asse.

Viene analizzato da Barrett (1998) il tipo di approccio da adottare per rappresentare le emozioni; *valence-arousal* viene utilizzato in soggetti che prediligono la facilità di riconoscimento della prima delle due componenti, cioè il grado di positività. Viene evidenziata un'alta correlazione tra le emozioni soggettive di simile valore di *valence*, ciò potrebbe indicare che questi individui stiano riportando parecchi stati insieme, oppure potrebbe essere sintomo di un'incapacità nel riuscire a discernere due stati comunemente considerati distinti. Nella ricerca sulle emozioni fatta da Vieillard et al. (2008), viene usato un modello bidimensionale *valence-arousal* per rappresentare le varie emozioni. Vengono scelte tre emozioni definite di tipo primitivo, quali: *happy*, *sad*, *scary*; viene inserito inoltre l'aggettivo *peaceful*, anche se non considerata un'emozione base, viene fatta questa scelta per fornire un aggettivo opposto al termine *scary*, visto che *happy* e *sad* sono in contrasto. In realtà, nella lista delle emozioni base di Ekman, non era presente l'aggettivo *scary* ma bensì *fear*, viene fatto questo cambio perché quest'ultimo aggettivo viene considerato troppo ambiguo per la descrizione delle emozioni di uno stimolo musicale. Un esecutore può esprimere *fear* suonando delicatamente oppure suscitare la stessa emozione nell'ascoltatore suonando con veemenza. Per ogni pezzo musicale vengono calcolate le medie per i valori di *valence* e *arousal*, esprimendoli con una scala da 0 a 9, e riportati nel corrispondente spazio bidimensionale, come si può vedere in *Figura 2.3*. Gli aggettivi *happy* e *peaceful* vengono classificati come piacevoli mentre *scary* come non piacevole. L'emozione *sad* viene collocata dalla parte di piacere dell'asse di *valence*. Lungo l'asse *arousal*, cioè quella che rappresenta l'attivazione, *sad* e *peaceful* hanno ricevuto un valore basso; un valore alto, quindi grande attivazione, per *happy* e *scary*. Osservando i risultati della ricerca, si nota, nella dimensione *arousal*, un distacco sostanziale tra i due gruppi di stimoli musicali, cioè tra quelli con bassa e alta attivazione; per quanto riguarda la dimensione di *valence*, la distanza tra i vari campioni diventa molto meno significativa fino, nel caso di *scary* e *peaceful*, a trovare un'intersezione delle emozioni. L'emozione *happy* è stata riconosciuta nel 99% dei casi, seguono poi *sad* con l'84% e *scary* con l'82%; l'emozione meno riconosciuta risulta essere *peaceful* con il 67% e per un 12% viene valutata in modo ambivalente; con il 75% viene scambiata con il *scary*, che risulta avere valori simili in termini di *valence*. In precedenza, era già stato sottolineato da Barrett (1998), che esisteva una grande correlazione per quanto riguarda, emozioni con valore simile in *valence*; ne è testimonianza anche quest'ultimo esempio.

Il modello *valence-arousal* risulta essere un buon modello di rappresentazione delle varie emozioni, anche se si nota una dipendenza accentuata in aggettivi con valori in *valence* comparabili. L'emozione *peaceful*, nello studio precedente, non viene considerata come basilare, è

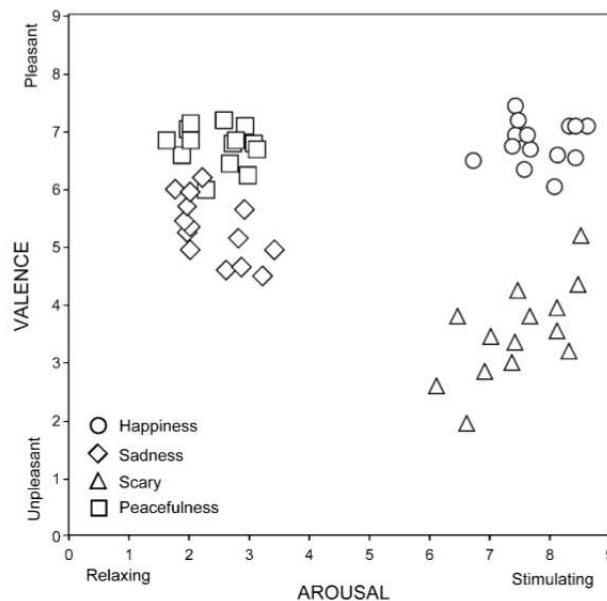


Figura 2.3: Rappresentazione stimoli musicali nel piano valence-arousal Vieillard et al. (2008)

Note: Media di valence-arousal su una scala di 10 punti. Ogni simbolo rappresenta uno stimolo musicale

pertanto naturale che questa abbia una certa correlazione con altre emozioni, in quanto, considerata proprio combinazione di emozioni base. Da tenere ben presente, la relazione che esiste tra l'esternazione dell'emozione e la risposta biologica, un'esperienza soggettiva non è altro che una traduzione di un fenomeno biologico in una rappresentazione cosciente dello stato. Giudizi uguali possono pertanto essere espressi per descrivere emozioni diverse in individui distinti. Non esiste, quindi, un modello universale di rappresentazione emozionale; verranno discusse in seguito circostanze che possono influire sull'emotività di una persona.

Modello circolare per Valence-Arousal

Si passa ora a esaminare un modello di rappresentazione che mette assieme due modelli già visti. La rappresentazione introdotta da Russel (1980), consiste in un modello bidimensionale *valence-arousal* con l'assegnazione in varie posizioni, dello spazio bidimensionale, di etichette rappresentanti varie emozioni, si viene a formare una struttura circolare, come si può vedere da *Figura 2.4*. Le considerazioni per questo modello sono del tutto analoghe a quelle fatte per i modelli precedentemente presentati, infatti, se il modello a *clusters* di Hevner (1936) (*Figura 2.1*) lo si inserisce opportunamente nel piano bidimensionale, considerando cioè la dipendenza con gli assi di *valence-arousal*, si ottiene praticamente il modello appena presentato. Altro esempio si ha in Vieillard et al. (2008), si vede da *Figura 2.3* il posizionamento di alcune emozioni proprio nel piano bidimensionale in questione. Viene riportato da Yang and Chen (2012) il pensiero tratto da un articolo di Sloboda and Juslin (2001), nel quale si afferma come da un punto di vista teorico la dimensione di attivazione, cioè *arousal*, è una delle maggiori variazioni percettibili nell'ambito

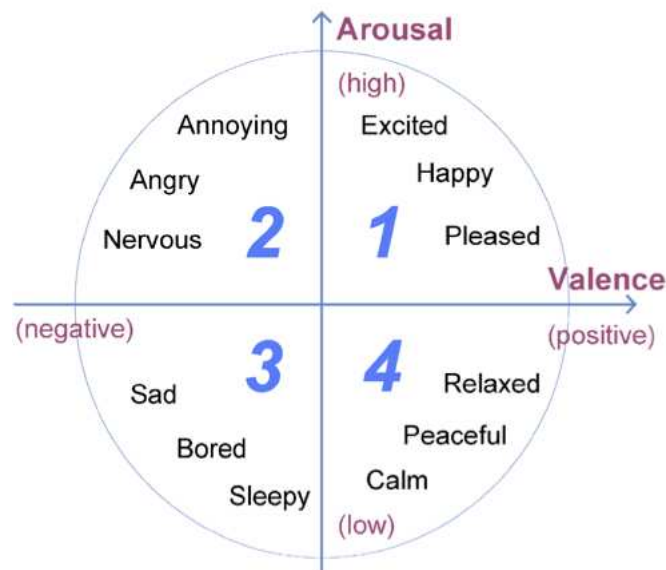


Figura 2.4: Rappresentazione stimoli musicali nel piano valence-arousal (Russel) Yang and Chen (2012)

Note: Le posizioni delle varie emozioni non sono esatte ma approssimative.

delle emozioni, insieme alla dimensione *valence*, la percezione cioè di qualcosa di piacevole o meno, ha una diretta correlazione su quello che è il comportamento di un individuo. Affermando, così, che il modello *valence-arousal* di Russel (1980) è il “cuore” dell’emozione, rappresenta cioè concretamente l’esperienza emozionale a un livello crudo e primitivo. Descrivere le emozioni usando un modello bidimensionale, però, non è libero dalle critiche del caso. Si sottolinea, infatti, come l’emozione *anger*, nello spazio *valence-arousal*, viene rappresentata nelle vicinanze immediate di *fear* (entrambe nel secondo quadrante), pur essendo molto differenti in termini di reazioni biologiche nell’organismo; un’emozione è di tipo attivo l’altra passivo. In risposta a questa mancanza, sembra opportuno introdurre una terza dimensione o più, per poter avere un quadro completo di rappresentazione dello spazio emozionale.

2.1.3 Spazi emozionali con dimensione superiore a due

Il problema su quante dimensioni usare, per poter rappresentare lo spazio emozionale, è tuttora aperto. Nella ricerca di Fontaine et al. (2007) si cerca di trovare una soluzione a questa grande problematica. Durante lo studio vengono osservati i cambiamenti attraverso sei attività: stima di eventi, cambiamenti psicologici (sensazioni corporali), espressioni motorie (faccia, voce, gesti), tendenza nell’azione, esperienze soggettive e regolazioni emozionali; nella maggior parte degli studi precedenti si sono osservati solo una o due di queste caratteristiche. Per ottenere, in modo evidente, quale sia la dimensione minima ottima a rappresentare lo spazio emozionale si usa un campione eterogeneo di persone, nello specifico: inglesi, francesi e tedeschi; chiedendo loro di

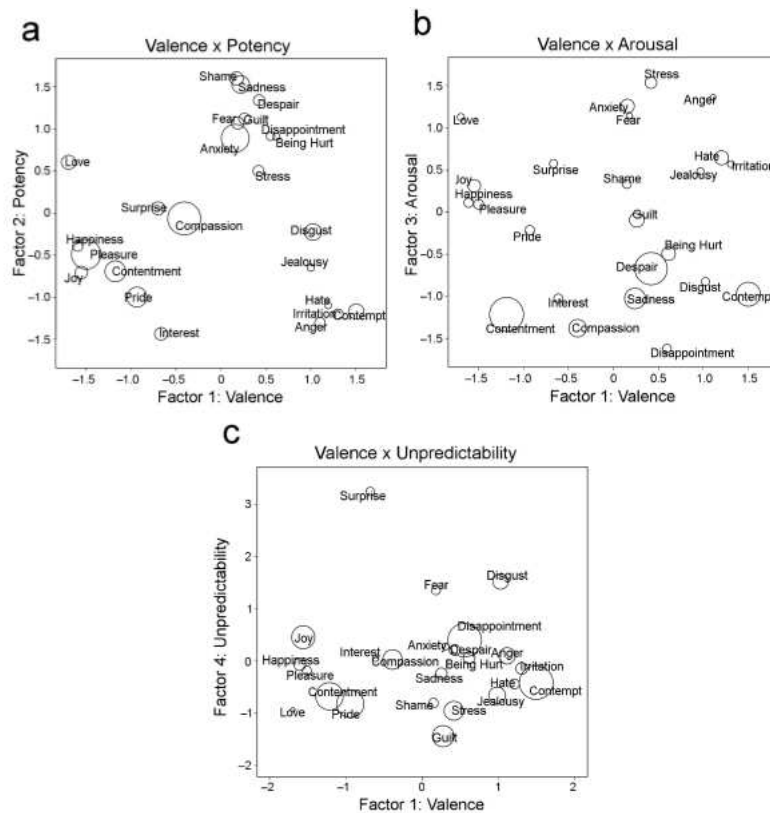


Figura 2.5: Rappresentazione dei 24 termini emozionali nello spazio quadri-dimensionale Fontaine et al. (2007).

Note: I punti medi dei vari cerchi rappresentano le coordinate medie lungo le tre lingue. Il diametro dei cerchi rappresenta la media della distanza euclidea tra le coordinate delle tre lingue; più piccolo è il cerchio, più simili sono i termini rispettivi attraverso le tre lingue. I tre grafici raffigurano i termini emozionali nelle coordinate: (a) Evaluation-Pleasantness x Potency-Control, (b) Evaluation-Pleasantness x Activation-Arousal, (c) Evaluation-Pleasantness x Unpredictability.

valutare 24 termini emozionali mettendoli in scala, considerando nella valutazione 144 possibili caratteristiche. Per ridurre la dimensione dello spazio emotivo, osservando le componenti lungo ogni dimensione, viene usato il PCA (*principal component analysis*), per trovare la minor varianza nei dati raccolti. L'elaborazione dei dati produce un modello composto da quattro dimensioni, quali: *evaluation-pleasantness*, *potency-control*, *activation-arousal*, *unpredictability*. L'interpretazione delle quattro dimensioni è basata sulle relazioni tra le 144 caratteristiche emozionali e sulle coordinate cartesiane dei 24 termini emozionali. In Figura 2.5 sono rappresentate le coordinate cartesiane, dei 24 termini emozionali, nel nuovo spazio di rappresentazione in quattro dimensioni. Le quattro dimensioni non sono rappresentabili graficamente contemporaneamente, vengono così presentati, dei grafici cartesiani divisi, ciascuno con ascissa *valence*. La prima dimensione, *evaluation-pleasantness*, può essere interpretata come tendenza all'azione contro l'assenza di movimento. La seconda dimensione, *potency-control*, è caratterizzata dal controllo, sensazione di potenza, dominazione o l'essere dominati. In questa dimensione emozioni come *pride*, *anger*, *contempt* sono in opposizione a *sadness*, *shame*, *despair* (Figura 2.5).

La terza dimensione, *activation-arousal*, è principalmente caratterizzata dall'attivazione, come il rapido battito del cuore oppure la prontezza nell'agire. L'ultima dimensione, *unpredictably*, è caratterizzata dalle cose nuove, dalla novità, sorpresa e tutto ciò quindi, che non può essere previsto. Si può affermare che le dimensioni più importanti sono le prime tre, l'ultima è degna di nota nel caso in cui si considerino le reazioni istantanee a stimoli nuovi, di sorpresa e pertanto non familiari.

Viene presentato da Schimmack and Reisenzenin (2002) un ulteriore spazio emozionale, questa volta composto da tre dimensioni, quali: *Energetic Arousal*, *Tense Arousal* e *Valence*; si ha pertanto una divisione di quello che prima era chiamato solo l'asse *arousal* in due assi che riportano il termine *arousal* nel finale del loro nome, si cercherà pertanto di provare, che *Energetic Arousal* e *Tense Arousal* non sono combinazione di *Valence* e *Activation*, ma bensì due distinte dimensioni; valide per essere incluse nella formazione di un nuovo spazio rappresentativo. Attraverso *energetic arousal* si va a rappresentare la sensazione di addormentamento oppure il grado di energia posseduta, per quanto riguarda *tense arousal*, invece, quantifica il grado di nervosismo o di calma di un individuo. La ricerca sottolinea, in prima battuta, che le due nuove dimensioni sono scatenate da cause differenti; per esempio *energetic arousal* è influenzato dal battito cardiaco che corrisponde al grado di attività delle celle del cervello che regolano l'organismo, cosa che non succede per *energetic arousal*. Si evidenzia, inoltre, che le due dimensioni possono crescere in direzioni opposte; per esempio, secondo uno studio del livello di zuccheri contenuti nel sangue, si osserva che *energetic arousal* decresce, mentre *tense arousal* cresce all'aumentare della glicemia. Terzo motivo, che sostiene la tesi di due dimensioni indipendenti è la diversità delle conseguenze; per esempio: numerosi studi dimostrano che *energetic arousal* rappresenta meglio le operazioni cognitive di *tense arousal*. Lo studio di Schimmack and Reisenzenin (2002) presenta, quindi, un nuovo approccio per risolvere la controversia riguardante la dimensione dell'attivazione. In un primo momento i ricercatori considerano *energetic arousal* e *tense arousal* due dimensioni separate, in quanto pensavano fossero causate da fattori differenti e basate su sistemi neurologici diversi. In contrasto a questa tesi c'è l'ipotesi di: *valence activation hypothesis*, la quale afferma che *tense* e *energetic arousal* siano combinazione della stessa dimensione di attivazione con la valenza. In linea con quest'ultima ipotesi si è trovato una correlazione positiva tra le due componenti e *valence* dovuta, però, soltanto alle correlazioni tra i residui, eliminando, quindi, la correlazione tra i residui, si trova un legame tra le due dimensioni prossimo allo zero. Il modello bidimensionale di attivazione, *Energetic Arousal*, *Tense Arousal*, non può essere preso da solo come rappresentativo, poiché non tutte le emozioni riescono a trovarne una rappresentazione, non si riesce nemmeno a trovare una corrispondenza biunivoca con *valence-activation*. Uno spazio completo potrebbe essere quello tridimensionale, come anticipato in precedenza, cioè *Energetic Arousal*, *Tense Arousal* e *Valence*, anche se, le dimensioni nuove di attivazione, presentano con *valence* un'alta correlazione, si ha pertanto un modello piuttosto ridondante.

Il numero ottimo di dimensioni, da introdurre in uno studio, dipende dalle domande che la ricerca si pone. Per ricerche che si concentrano soltanto sull'attivazione potrebbe essere sufficiente una sola dimensione (*arousal*). Per una ricerca che mira nella caratterizzazione di termini emozionali, con significato semantico molto simile, potrebbero non essere sufficienti neanche quattro dimensioni. Si può affermare che, in via generale, già quattro dimensioni riescono a dare una caratterizzazione molto buona dello spazio emozionale. I risultati di Fontaine et al.

(2007), portano ad affermare che un semplice modello bidimensionale, come *valence-arousal*, non riesce a caratterizzare in maniera opportuna emozioni come *anger e fear*, che sono invece nettamente separate nella dimensione *potency-control* e *unpredictability*; come si può vedere da *Figura 2.5*. Un modello come quello *valence-arousal*, bidimensionale, riesce a rappresentare degli stati emotivi da un punto di vista comprensivo, cioè, una volta che questi si manifestano nel nostro organismo si riesce a darne una caratterizzazione da un punto di vista logico. La dimensione *potency-control*, invece, riesce a dare una connotazione a emozioni, che si manifestano sia in maniera comprensiva, sia a emozioni che si manifestano all'interno dell'organismo come, ad esempio, sintomi gastrointestinali; delineandone l'impronta biologica. La quarta dimensione dello spazio, cioè *unpredictability*, svolge anch'essa un ruolo molto importante; un individuo di fronte a uno stimolo mai provato o che non si aspetta di provare avrà una reazione diversa, rispetto a una situazione già familiare. L'influenza dell'ambiente e le circostanze, in cui si trova una persona al momento dell'emozione, sono elementi che influiscono ampiamente sull'intensità e modo della manifestazione della stessa. Non esiste pertanto, tenendo in considerazione quanto discusso in precedenza, un numero di dimensioni ottimale da prendere in considerazione, la decisione della cardinalità dello spazio rappresentativo emozionale deve essere fatta di caso in caso per non incorrere ad un sovra o sotto dimensionamento dello stesso, avendo, nel primo caso delle coordinate dimensionali ridondanti tra loro, mentre, nel secondo, una carente caratterizzazione delle emozioni.

Capitolo 3

Struttura musicale ed emozioni

Ci sono diversi fattori che influenzano le emozioni percepite in musica. Emozioni differenti sono associate a caratteristiche musicali diverse. Per esempio, mentre *arousal* è legato al Tempo (veloce/lento), volume (alto/basso), e timbro (intenso/morbido), *valence* è legata al modo (maggiore/minore) e all'armonia (consonante/dissonante). Si nota, inoltre, che una percezione emozionale non è raramente legata a una sola caratteristica, ma a una combinazione di esse. L'emozione percepita è frutto di una combinazione tra vari fattori, quali l'esecuzione del performer, ma prima ancora alla scrittura della partitura e ovviamente dall'intenzione del compositore. Il ricercatore per valutare l'emozione, può avvalersi di diversi metodi: può chiedere una valutazione diretta all'ascoltatore riguardante le sue emozioni provate durante l'ascolto oppure procedendo con la variazione di determinati parametri intrinseci del pezzo (Tempo, ritmo, modo) uno alla volta, estrapolando un legame tra la struttura musicale ed il contenuto emozionale espresso.

3.1 Emozioni e caratteristiche musicali

Si cercherà ora di elencare le principali caratteristiche di un brano musicale cercando di legarle con le emozioni tipiche che si manifestano al variare delle stesse.

Tempo Come evidenziato da Balkwill et al. (2004), il Tempo veloce è associato a: *happiness, joy, anger, fear, activity*; quindi ci saranno alti valori di entrambe le componenti *valence-arousal*; con un Tempo lento, si vanno ad esprimere emozioni come: *sadness, disgust, boredom, solemnness*; caratterizzate per avere alta *valence* mentre bassa attivazione, quindi, un valore basso in *arousal*.

Modo Canazza et al. (2011) analizza la dipendenza delle emozioni al variare del modo del brano, caratteristica intrinseca dello stesso, si evince pertanto la caratterizzazione nei modi maggiori di emozioni con alta *valence* ma bassa *arousal*, emozioni come: *serene, happiness, joy*; il modo minore è caratterizzato da alto valore in *arousal* indipendentemente dal valore in *valence*, identificando emozioni come: *sadness, tension, disgust, anger*.

Intensità A un brano con molta intensità verranno associati aggettivi come: *intensity, tension, power, anger*; che hanno un alto valore in *arousal*; per brani con bassa intensità avviene

	Valence	Energy
	R^{2adj}	R^{2adj}
Prediction rate	.54	.74
	β	β
Attack Slope	-.17*	.32***
Envelope Centroid	.71***	-.42***
Ratio of HF-LF energy	-.52***	.71***
Spectral Skewness	-.30***	-.30**
Spectral Regularity	.03	-.13
Spectral Flux	.81***	-.09
Sub-Band No. 6 Flux	-.75***	.33**

* $p < .05$, ** $p < .01$, *** $p < .001$

Figura 3.1: Riassunto dell'analisi di regressione Eerola et al. (2012)

l'associazione di aggettivi come: *softness, tendeness, sadness*; caratterizzati, quindi, da valori bassi in *arousal*.

Altezza dei suoni La presenza di picchi elevati comporterà all'espressione di: *exciting, happy, potency, surprise, activity*; caratterizzati, quindi, da un alto valore della componente *arousal*, viceversa, la presenza di picchi di bassa entità denoterà emozioni del tipo: *boredom, pleasantness, sad*; la componente *arousal* sarà di valore basso.

Armonia Con l'armonia di tipo consonante si esprime: *serene, happy, relaxed*; mentre, con una di tipo dissonante aggettivi espressivi come: *disgust, anger, unpleasantness, tension*.

Tonalità Con armonia di tipo cromatico si trasmette: *angry, sad*; con armonia tonale: *joy, peaceful, happiness*; con armonia di tipo atonale si trasmette, ad esempio, emozioni come *angry*.

Timbro Assieme a Tempo e modo del brano, il timbro è un altro elemento caratterizzante l'espressività della performance, come sottolineato, per esempio, da Canazza et al. (2011). L'influenza sull'espressività, da parte del timbro, viene discussa da Eerola et al. (2012), conducendo una ricerca che cerca di capire la correlazione tra variazioni timbriche e variazioni espressive. Dalla ricerca, emerge una dipendenza del timbro principalmente da tre parametri, quali: tipo di attacco, brillantezza, che dipende a sua volta dall'energia in alta frequenza, e dal flusso spettrale o irregolarità nello spettro. Il timbro viene quindi caratterizzato da una varietà ampia di caratteristiche, che vanno dal tipo di strumento a modalità di esecuzione. In *Figura 3.1* sono riassunti alcuni parametri caratteristici di un'esecuzione e la loro correlazione con lo spazio bidimensionale *valence-arousal*, scelto per rappresentare le varie emozioni. Dai risultati dello studio si riescono a classificare le emozioni al variare delle caratteristiche prima citate; si identificano, ad esempio, in suoni con molte armoniche, emozioni come: *surprise, disgust, fear, potency*. In suoni con le armoniche alte amplificate si trovano emozioni come *anger*. Emozioni come: *sadness, pleasantness,*

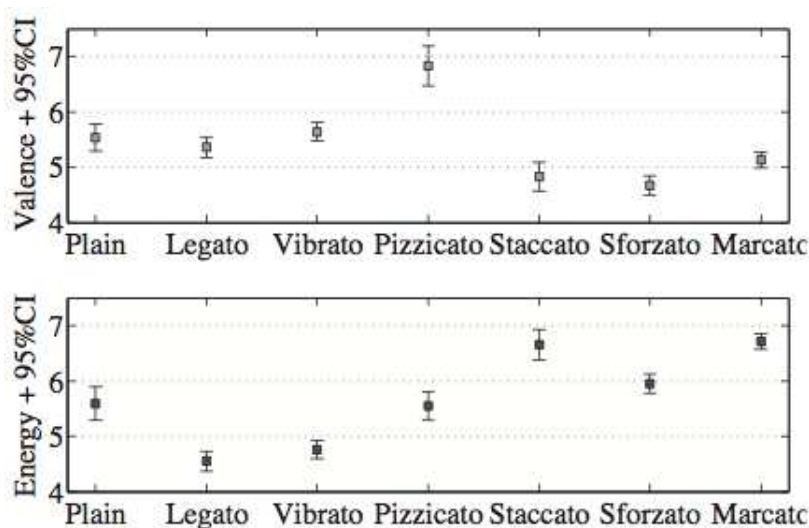


Figura 3.2: Dipendenza da valence-energy al variare dell'articolazione Eerola et al. (2012)

boredom; in suoni con poche armoniche, mentre, in quelli con l'energia distribuita in bassa frequenza si trovano: *tenderness, sadness*. Strumenti a strofinio delle corde si prestano a trasmettere emozioni, quali: *sadness, anger*; strumenti a fiato, come il flauto si trovano adatti a trasmettere una sensazione di pace.

Articolazione Da Eerola et al. (2012), si evince, che articolazioni con attacco veloce, come: sforzato, marcato e staccato; sono caratterizzate da un alto valore in energia; altro elemento caratterizzante per la distinzione emozionale sarà il tipo di inviluppo. Con articolazioni di tipo staccato si avrà la comunicazione di emozioni come: *anger, fear, activity, energy*; con articolazione di tipo legato si esprimerà: *softness, solemnity, tenderness, sadness*. Come detto in precedenza anche l'inviluppo gioca un ruolo importante dal punto di vista espressivo, si avrà pertanto: con un inviluppo di tipo quadrato, l'espressione di emozioni come: *activity, surprise, anger*; con un inviluppo tondeggiante, l'espressione di emozioni come: *boredom, disgust, fear, sadness*. L'articolazione è caratterizzata dalla variazione della componente *valence*, come si può vedere da *Figura 3.2*, risultati ottenuti dalla ricerca di Eerola et al. (2012). Si nota che i suoni pizzicati sono principalmente a più alto valore in *valence*. In contrasto, si vede come, la componente *energy* varia notevolmente da articolazioni di tipo sostenuto (legato, vibrato) ad articolazioni con inviluppo di tipo impulsivo (pizzicato, staccato, sforzato, marcato); caratterizzazione che non avviene per quanto riguarda la componente *valence*.

Come detto in precedenza, le intenzioni emozionali sono una combinazione di vari aspetti: indicazioni esplicitamente scritte nello spartito, tipo dello strumento ed esecuzione da parte del performer. I parametri sopracitati, essendo il mezzo per comunicare espressività, sono anch'essi combinazioni provenienti dallo spartito, strumento ed esecutore, la fusione di questi tre elementi da vita al suono; separare i fattori che influiscono sull'espressività è pertanto un'operazione non sempre facile.



Figura 3.3: Comunicazione emozionale dell'esecutore.

3.2 Emozioni ed esecuzione

Il messaggio musicale è parzialmente codificato nello spartito. Quando si ascolta un pezzo percepiamo il contributo sia dell'esecutore, sia del compositore. Il performer del brano cerca di trasmettere, prima di tutto, la struttura imposta dalla lettura dello spartito, unendola poi a una propria interpretazione del messaggio musicale. L'esecutore, variando parametri come Tempo, articolazione, intensità ecc. cerca di trasmettere all'ascoltatore le proprie emozioni. Il processo è schematizzato in *Figura 3.3*. L'ascoltatore combina le varie caratteristiche, che gli vengono trasmesse, cercando di interpretare il contenuto espressivo. I parametri, contenenti espressività, derivano sia da come il suono viene prodotto, quindi dallo strumento, sia da come l'esecutore si avvale delle caratteristiche sopracitate per trasmettere emozioni.

Si elencheranno di seguito alcune importanti emozioni legandole ai parametri musicali su cui deve agire un esecutore per riuscire a comunicarle.

Happyness: Tempo veloce, poche variazioni di Tempo, alta intensità, poca variazione di intensità, staccato, alta variabilità nell'articolazione, timbro chiaro.

Anger: Tempo veloce, poca variazione di Tempo, no ritardando, alta intensità, staccato, attacchi veloci, vibrato.

Sadness: Tempo lento, finale ritardando, bassa intensità, legato, poca variazione di articolazione, attacchi lenti, leggermente vibrato.

Fear: Tempo veloce, alta variabilità di Tempo, staccato, intensità molto bassa, alta variabilità di intensità, vibrato irregolare.

Se si vuole dare un'interpretazione dal punto di vista dello spazio emozionale bidimensionale *valence-arousal*, si nota che, ad esempio: Tempo, intensità e articolazione sono principalmente legati alla componente di attivazione, quindi *arousal*; mentre il timbro, inerentemente alla sua correlazione con l'intensità, è caratterizzato da una variazione della componente *valence*. A un'intensità media e all'energia ad alta frequenza sono associate emozioni positive, mentre per valori estremi di intensità, sia alti che bassi, sono associate emozioni di tipo negativo. Si evidenzia, per esempio da Eerola et al. (2012), che alcune caratteristiche come: Tempo, modo, timbro e articolazione sono importanti per quanto riguarda la distinzione del tipo di espressività trasmessa,

definendo, ad esempio, se l'emozione è di tipo positivo o negativo; gli altri parametri sopracitati svolgono un ruolo significativo solo in una fase finale dell'inquadramento dell'emozione, apportando solo delle lievi sfumature al contenuto emotivo.

L'analisi dell'espressività, nel contempo la trasmissione dell'emozione, non è un'azione semplice da svolgere, in quanto dipende da numerose sfaccettature, esposte nella trattazione; la soggettività dell'emozione di certo non aiuta in questo compito. Sono state comunque riportate le caratteristiche generali di inquadramento dell'emozione, fermo restando che, vista la delicatezza della questione, ogni caso va trattato singolarmente per averne una caratterizzazione completa e opportuna.

Capitolo 4

Considerazioni finali

In questa tesi si sono discusse alcune modalità per rendere possibile l'esecuzione automatica espressiva della musica; si sono cercate delle connessioni tra due diversi mondi che possono sembrare antitetici: le macchine e le emozioni.

Tramite lo studio dell'esecuzione di un essere umano, cercando di capire come la musica influisca sull'emozione, si è cercato di costruire dei modelli che riescano a rappresentare l'espressività del brano, permettendo a un calcolatore di eseguire il pezzo servendosi dei dati provenienti dallo spartito e dei dati sull'espressività, ricavati dal modello creato. I musicisti di rado suonano la musica come compare sullo spartito, ma la interpretano, aggiungendo la propria personalità. Opportune introduzioni di elementi espressivi in fase di esecuzione andranno ad arricchire il modello matematico.

Sono state, inoltre, presentati diverse tipologie di spazio con lo scopo di riuscire a rappresentare il contenuto emotivo; dalla discussione emerge, che lo spazio emozionale deve essere valutato di caso in caso, secondo lo scopo dell'esperimento, solo così si avrà una rappresentazione adatta al grado di accuratezza desiderato. Si è cercato, infine, di trovare un legame tra i vari parametri intrinseci di un brano o riguardanti il modo di esecuzione e la comunicazione dell'espressività; è pertanto emersa la grande importanza di parametri come: Tempo, modo, timbro e articolazione. Al variare di una di queste caratteristiche, si avrà una diversa sfumatura dell'emozione percepita.

L'espressività è frutto di numerose varianti tecniche, strutturali, di esecuzione ecc, che vanno a coinvolgere la psicologia dell'individuo, e quindi la soggettività, con limitazioni, inoltre, di tipo linguistico e culturale; è impossibile, quindi, riuscire a creare un modello universale di espressività; il modello che si andrà a creare, è fortemente legato al contesto per cui è stata pensata la realizzazione.

L'obiettivo dell'esecuzione automatica espressiva della musica non è di sottrarre agli esseri umani questo compito, bensì cercare di migliorare l'interazione con le macchine utilizzando i canali tipici della comunicazione uomo-uomo, che sono meno frustranti e più semplici da assimilare, in modo particolare per utenti diversamente abili o comunque per persone inesperte in ambito informatico.

Bibliografia

- Laura-Lee Balkwill, William Forde Thomson, and Rie Matsunaga. Recognition of emotion in Japanese, western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, 46(4):337–349, 2004.
- Lisa Feldman Barrett. Discrete emotion or dimension? the role of valence focus and arousal focus. *Psychology Press Ltd*, 12(4):579–599, 1998.
- Sergio Canazza, Giovanni De Poli, and Antonio Roda. Emotional response to major mode musical pieces: Score-dependent perceptual and acoustic analysis. Technical report, Università di Padova, 2011.
- Tuomas Eerola, Rafael Ferrer, and Vinoo Alluri. Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *An Interdisciplinary Journal*, 30(1):49–70, September 2012.
- P. R. Farnsworth. The social psychology of music. *State university press*, 1969.
- Johnny R.J. Fontaine, Klaus R. Scherer, Etienne B. Roesch, and Phoebe C. Ellsworth. The world of emotions is not two-dimensional. *Association for Psychological Science Journal*, 18(12): 1050–1057, 2007.
- Kate Hevner. Experimental studies of the elements of expression in music. *American Journal of Psychology*, 48:246–268, 1936.
- Peter J. Rentfrow and Samuel D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6):1236–1256, 2003.
- Bruno Haskins Repp. Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Journal of Acoustical Society of America (JASA)*, 44: 241–281, 1992.
- J. A. Russell. A circumplex model of affects. *Social Psychol*, 39(6):1161–1178, 1980.
- Ulrich Schimmack and Rainer Reisenzein. Experiencing activation: Energetic arousal and tense arousal are not mixtures of valence and activation. *American Psychological Association Journal*, 2(4):412–417, 2002.

- Emery Shubert. Update of the hevner adjective checklist. *Perceptual and Motor Skills*, 96: 1117–1122, 2003.
- J. A. Sloboda and P. N. Juslin. Psychological perspectives on music and emotion. *Oxford University Press*, 2001.
- Sandrine Vieillard, Isabelle Peretz, Nathalie Gosselin, Stéphanie Khalfa, Lise Gagnon, and Bernard Bouchard. Happy, sad, scary and peaceful musical excerpts for research on emotions. *Psychology Press Ltd*, 22(4):720–752, 2008.
- Yi Hsuan Yang and Homer H. Chen. Prediction of the distribution of perceived music emotions using discrete samples. *IEEE Computer Society*, 19(7):2184–2196, 2011.
- Yi Hsuan Yang and Homer H. Chen. Machine recognition of music emotion: A review. *ACM Transactions on Intelligent Systems and Technology*, 3(3):40:1–40:30, 2012.
- Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Computer Society*, 31(1): 39–58, 2009.