



UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Dipartimento di Informatica
Corso di Laurea Triennale in Informatica Musicale

Tesi di Laurea

ANALISI QUALITATIVA E QUANTITATIVA DELLA PERCEZIONE DELL'AUDIO ATTRAVERSO INTERFACCIA BCI

Relatore:
Prof. LUCA A. LUDOVICO

Laureando:
SIMONE CASTELLANI

Correlatore:
Dott. MATTIA G. BERGOMI

ANNO ACCADEMICO 2012-2013

*A Italia
e Imer*

Indice

Introduzione	I
Elenco delle figure	III
Elenco delle tabelle	IV
1 Brain-Computer Interface	1
1.1 Cenni di anatomia e fisiologia del sistema nervoso umano	1
1.1.1 Sistema nervoso centrale - SNC	1
1.1.2 Neuroni e sinapsi	1
1.2 Elettroencefalografia (EEG)	3
1.2.1 Ritmi	4
1.2.2 Attività mentali e potenziali evocati	5
1.3 Classificazione delle BCI	6
1.4 Architettura di una BCI	7
1.4.1 Disposizione degli elettrodi	8
1.5 Artefatti e rimozione	9
2 Emozioni e spazi affettivi	11
2.1 Emozioni e cervello	11
2.1.1 Studi sulla percezione delle emozioni	12
2.2 Spazi affettivi	12
2.2.1 Modelli degli spazi affettivi	12
2.2.2 Spazi affettivi ed età	14
2.3 Affective computing	15
3 Indagine sperimentale	17
3.1 Stimolazione affettiva	17
3.1.1 Database affettivi: IAPS e IADS	17
3.1.2 Selezione degli stimoli affettivi	19
3.2 Acquisizione dati	21
3.2.1 Interfaccia	21
3.2.2 Applicativo	22
3.2.3 Protocollo sperimentale	23
3.3 Processamento	24
3.3.1 Preprocessing	24
Independent Component Analysis (ICA)	24
Filtraggio passa-banda con filtro di Čebyšëv	27

3.3.2	Scelta delle feature	27
	Recurrence Plot	28
	Constant Q transform	29
4	Analisi dei risultati	30
4.1	Recurrence Plot	30
4.1.1	Limitatezza del segnale	32
4.2	Risultati Q-transform	33
5	Conclusioni e sviluppi futuri	36
	Bibliografia	39
A	Codice Implementativo	42
A.1	Ricerca e selezione del materiale	42
A.1.1	importValoriDb.m	42
A.1.2	SelezioneAudioImmaginiExp.m	43
A.1.3	getValueFromID.m	43
A.1.4	getImgMinFromAudio.m	44
A.1.5	getAudioMaxFromImg.m	45
A.2	Estrazione delle feature ed analisi dei dati	46
A.2.1	preprocessingFilt.m	46
A.2.2	EEGtoERP.m	47
A.2.3	RP.m	48

Elenco delle figure

1.1	Rappresentazione schematica del sistema nervoso centrale e lobi cerebrali . . .	2
1.2	Anatomia e morfologia del neurone	3
1.3	Ritmi cerebrali	4
1.4	BCI invasive e BCI non invasive	7
1.5	Architettura di una BCI	8
1.6	International 10-20 System	9
1.7	10-10 Extension System e 10-5 System	10
2.1	Modello circonflesso delle emozioni	13
2.2	Spazio affettivo in funzione dell'età e del genere	14
2.3	Framework di ricerca per l'interazione basata sulle emozioni	16
3.1	Distribuzione spaziale IAPS e IADS	18
3.2	Emotiv Epoc	21
3.3	BCI2000	22
3.4	Protocollo sperimentale e set di stimolazione	23
3.5	Procedura di preprocessing del segnale EEG per ogni soggetto	24
3.6	Esempio di applicazione ICA	25
3.7	Risposte in frequenza e fase dei filtri HP e LP	27
4.1	Serie e i grafici delle ricorrenze	31
4.2	Serie delle differenze <i>difAcImg</i> e <i>difAsImg</i>	32
4.3	QT Set 1	34
4.4	QT Set 2	34
4.5	QT Set 3	34
4.6	QT Set 4	35
4.7	QT Set 5	35
4.8	QT Set 6	35
4.9	QT Set 7	35

Elenco delle tabelle

2.1	Categorie Affective Computing	15
3.1	Numero di file per classi di valenza e arousal	18
3.2	Medie e deviazione standard IAPS e IADS	19
3.3	Valori di valenza e arousal per ogni stimolo	21
4.1	Percentuali di omogeneità delle risposte inter-soggetto basate su RP	31
4.2	Valori delle differenze <i>difAcImg</i> e <i>difAsImg</i>	33
5.1	Andamento nei ritmi μ e β nelle diverse combinazioni di stimolazione	36

Introduzione

Obiettivo primario di questo elaborato è quello di indagare le variazioni di percezione della componente emotiva di una serie di stimoli audio/visivi.

L'obiettivo è stato perseguito allestendo una raccolta dati sperimentale e proseguendo con l'analisi di essi mediante tecniche matematiche. La percezione è misurata in termini di attivazione dei ritmi corticali μ e β , misurata utilizzando dispositivi hardware e software appositamente progettati.

Il primo capitolo presenta alcuni aspetti di anatomia e fisiologia del cervello umano, le basi della elettroencefalografia, la descrizione dei ritmi cerebrali e i problemi relativi agli artefatti introdotti nei segnali EEG. Successivamente presentata le modalità con cui queste teorie sono applicate all'informatica, analizzando l'architettura e il funzionamento delle Brain Computer Interface (BCI).

Il secondo capitolo è dedicato all'analisi della natura psicofisiologica delle emozioni e dei differenti spazi affettivi, tra cui il modello dimensionale Valenza/Arousal [1]. Questo modello è uno spazio bidimensionale, in grado di fornire una stima della positività o negatività (valenza) e la capacità di attivazione (arousal) di un determinato stimolo, attraverso due coordinate cartesiane. Il capitolo termina con una descrizione delle applicazioni che gli studi sulle emozioni hanno avuto attraverso lo sviluppo dell'affective computing [2].

L'idea, basata sul modello sopra descritto e sul concetto che la percezione di uno stimolo emotigeno genera una serie di risposte complesse da parte dell'organismo, è quella di misurare le risposte cerebrali di un soggetto, monitorando gli stati di attivazione dei ritmi μ e β (generalmente associati ad uno stato di attenzione e coinvolgimento) in risposta ad un protocollo di stimolazione che generi uno scollamento emotivo sempre maggiore tra gli stimoli.

Il protocollo è ampiamente descritto nel terzo capitolo, ma è così riassumibile: vengono proiettati su uno schermo sette set di stimoli visivi con valori di valenza sempre minore (i primi tre) e poi sempre maggiore (i restanti quattro), ogni set contiene una immagine che è presentata per tre volte, la prima senza audio in sottofondo, la seconda insieme ad un audio con valori di valenza simili, la terza insieme ad un audio con valori di valenza totalmente opposti.

Per scollamento è inteso il fatto che in tutti gli stimoli in terza posizione, le immagini e gli audio, dovrebbero suscitare emozioni totalmente opposte e la loro sovrapposizione potrebbe non essere coerente ne con uno ne con l'altro stimolo.

Come indicatore di queste reazioni complesse è stata considerata la variazione di risposta neurale, calcolata come differenza di ampiezza delle risposte medie dei soggetti durante le stimolazioni con la sola immagine, considerato livello zero (ideale), e le ampiezze medie dei ritmi presi in esame durante la stimolazione con audio correlato e scorrelato.

Nel terzo capitolo sono elencati i materiali e i metodi implementati; questo è suddiviso in tre blocchi fondamentali corrispondenti alle fasi percorse durante lo studio:

- l'analisi e la scelta del materiale: vengono presentati i database IADS e IAPS, che forniscono un insieme di campioni audio e immagini, catalogati secondo la valutazione statistica della loro posizione nello spazio affettivo Valenza/Arousal. La scelta degli stimoli da utilizzare è stata condotta in modo algoritmico al fine di mantenere oggettivi i criteri applicati. Sono stati scelti sette campioni audio da cui partire, posizionati agli estremi della coordinata valenza, e cercate le immagini che minimizzano la scorrelazione, misurando la minima distanza euclidea dalla proiezione del punto dell'audio sullo spazio affettivo delle immagini. Successivamente l'algoritmo esegue la ricerca degli audio che massimizzano la scorrelazione, misurando la massima distanza euclidea dalla proiezione dei punti di ogni immagine nello spazio degli audio;
- la raccolta dati: vengono descritte le tecnologie hardware e software utilizzate, l'interfaccia BCI Emotiv Epoc, tramite cui le onde cerebrali dei soggetti sono state campionate in sei posizioni differenti in accordo con lo standard internazionale per la disposizione degli elettrodi sullo scalpo definito sistema internazionale 10-20 [3]. Viene presentato il software BCI2000, utilizzato per condurre la stimolazione e la raccolta dai dati;
- l'analisi dei dati: vengono presentati i metodi matematici messi in atto per preprocessare i segnali, per rimuovere i possibili artefatti (Independent Component Analysis e filtraggio passa-banda) ed estrarre le features utili all'analisi dei dati (Recurrence Plot e Constant Q transform).

Il quarto capitolo presenta i risultati ottenuti dall'applicazione delle tecniche matematiche e nel quinto vengono presentate alcune considerazioni su quest'ultimi e i possibili sviluppi.

Capitolo 1

Brain-Computer Interface

1.1 Cenni di anatomia e fisiologia del sistema nervoso umano

Il sistema nervoso umano è da considerare come corpo unico le cui parti sono tra di loro in reciproca connessione, sia morfologica che funzionale[4].

Per comodità di spiegazione, però, è generalmente suddiviso in *sistema nervoso centrale* (SNC), *sistema nervoso periferico* (SNP) e *sistema nervoso vegetativo* (SNV).

Il tessuto nervoso, altamente specializzato nel ricevere, trasmettere ed elaborare gli stimoli interni ed esterni del corpo, è suddiviso in *sostanza grigia* e *sostanza bianca*.

La prima è formata dai corpi delle cellule nervose. La seconda è costituita da particolari prolungamenti, denominati *neuroni*, i quali sono emanati dai corpi cellulari stessi e sono adibiti al trasporto degli stimoli.

1.1.1 Sistema nervoso centrale - SNC

Nel SNC, i neuroni che svolgono funzioni simili si aggregano in modo ben definito a formare *nuclei*, *lamine*, *corteccie*, *colonne* e *centri*, mentre all'esterno del SNC i corpi cellulari si raggruppano per formare *gangli*.

In figura 1.1) sono illustrate le sezioni che compongono il SNC, tra cui: midollo spinale, ponte, cervelletto, diencefalo ed emisferi cerebrali.

Il diencefalo è costituito da *talamo*, *epitalamo* e *ipotalamo*. Il talamo rappresenta un nodo fondamentale per la trasmissione verso le aree della corteccia cerebrale specifiche degli impulsi sensitivi, i quali assumono carattere di percezione.

Sono presenti due emisferi cerebrali ricoperti dalla corteccia, il sinistro e il destro, i quali non sono simmetrici, sia per quanto concerne l'anatomia sia per le funzionalità. Ciascun emisfero elabora informazioni controlaterali.

La corteccia cerebrale è suddivisa in quattro lobi: *frontale*, *parietale*, *temporale* e *occipitale*. Le convoluzioni cerebrali hanno la funzione di aumentarne la superficie. Sono state individuate 52 diverse aree della corteccia, per molte delle quali è stata successivamente provata una specializzazione funzionale [5].

1.1.2 Neuroni e sinapsi

I neuroni sono cellule specializzate per la ricezione e per la generazione di segnali.

Ogni neurone è composto da:

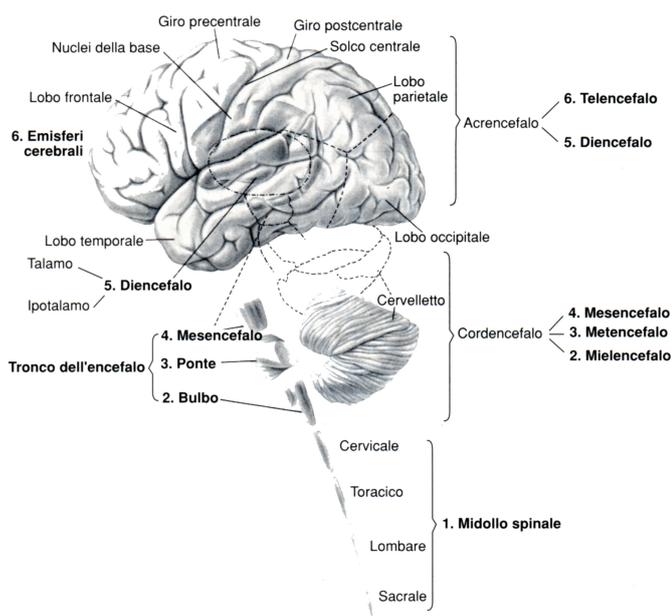


Figura 1.1: Rappresentazione schematica del sistema nervoso centrale e lobi cerebrali

- un *corpo cellulare* (o *soma*), il quale contiene il nucleo e svolge le principali attività metaboliche e di sintesi;
- i *dendriti*, i quali sono ramificazioni che originano dal soma e svolgono azione di interfaccia tra neuroni diversi;
- l'*assone*, ha origine dal soma ma a differenza dei dendriti si allontana anche di molto (fino ad un metro nei motoneuroni), in alcuni tipi di neuroni è ricoperto da mielina che svolge un'importante funzione isolante;
- la *sinapsi* (o bottone sinaptico), è la parte terminale dell'assone ed è deputata alla liberazione del neurotrasmettitore che consente la propagazione del segnale da una cellula all'altra.

La figura 1.2 mostra una rappresentazione dell'anatomia del neurone e le morfologie che essi possono assumere in base alla loro differenziazione.

I neuroni scambiano informazioni mediante l'invio e la ricezione di *segnali elettrici neurali*, i quali possono essere *potenziali graduati* o *potenziali d'azione*[5].

I potenziali graduati sono oscillazioni del potenziale di membrana, la loro ampiezza è proporzionale all'intensità dello stimolo e decresce con la distanza dal punto di origine; possono essere eccitatori, se determinano la depolarizzazione della membrana, oppure inibitori, se determinano iperpolarizzazione di membrana.

Questi potenziali sono soggetti a *sommazione spaziale* e *sommazione temporale*. La prima si basa sulla convergenza anatomica di più fibre sullo stesso neurone, la cui somma algebrica dei potenziali può determinare il raggiungimento della soglia e indurre la partenza del potenziale d'azione. La seconda dipende dal fatto che ogni fibra è percorsa da più potenziali d'azione, ciascuna dei quali è responsabile della formazione di potenziali graduati sul neurone innervato

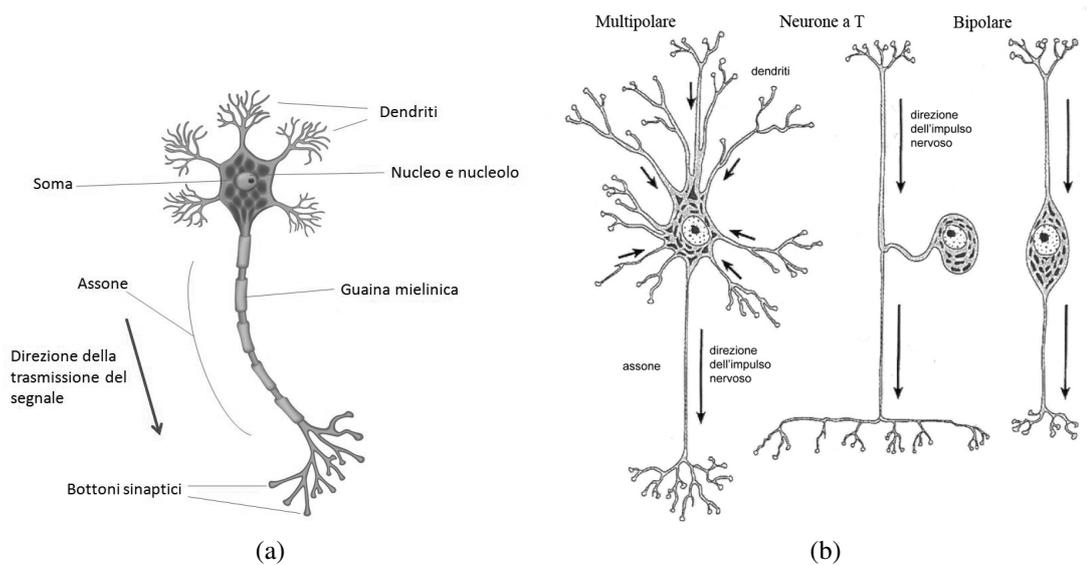


Figura 1.2: Rappresentazione schematica dell'anatomia (a) e delle differenti morfologie (b) di un neurone

dalla fibra considerata. Se la frequenza dei potenziali è opportuna si potranno sommare nel tempo.

I potenziali d'azione sono generati sempre alla stessa ampiezza indipendentemente dall'entità dello stimolo, il quale deve superare un valore di soglia per innescare l'attivazione.

Il potenziale d'azione dura nella cellula nervosa circa 1 *ms*, è generato nell'assone e procede fino alla sinapsi.

Le sinapsi si distinguono in elettriche e chimiche; nelle prime esiste un contatto diretto tra le cellule, in cui il passaggio del segnale avviene rapidamente, in maniera indistinta nelle due direzioni. Nell'uomo sono localizzate nella retina, nell'ippocampo e nel talamo. Nelle seconde il segnale elettrico media il rilascio a livello pre-sinaptico di un mediatore chimico, sono unidirezionali e molto più lente.

1.2 Elettroencefalografia (EEG)

L'elettroencefalografia (EEG) è la registrazione dell'attività elettrica dell'encefalo. La tecnica è stata inventata nel 1929 da Hans Berger, il quale scoprì che vi era una differenza di potenziale elettrico tra aghi infissi nello scalpo oppure tra due piccoli dischi di metallo (elettrodi) quando essi sono posti a contatto sulla cute sgrassata del cuoio capelluto. L'EEG misura la differenza di potenziale elettrico tra un elettrodo attivo ed un elettrodo indifferente, posto ad una certa distanza dal primo. Il contributo elettrico di ciascun neurone è molto piccolo e il segnale deve attraversare diversi strati, tra cui le meningi e la scatola cranica, prima di raggiungere gli elettrodi. In una determinata area corticale sono necessari migliaia di neuroni che scarichino il potenziale elettrico in sincronia per poter rilevare l'attività elettrica attraverso l'EEG.

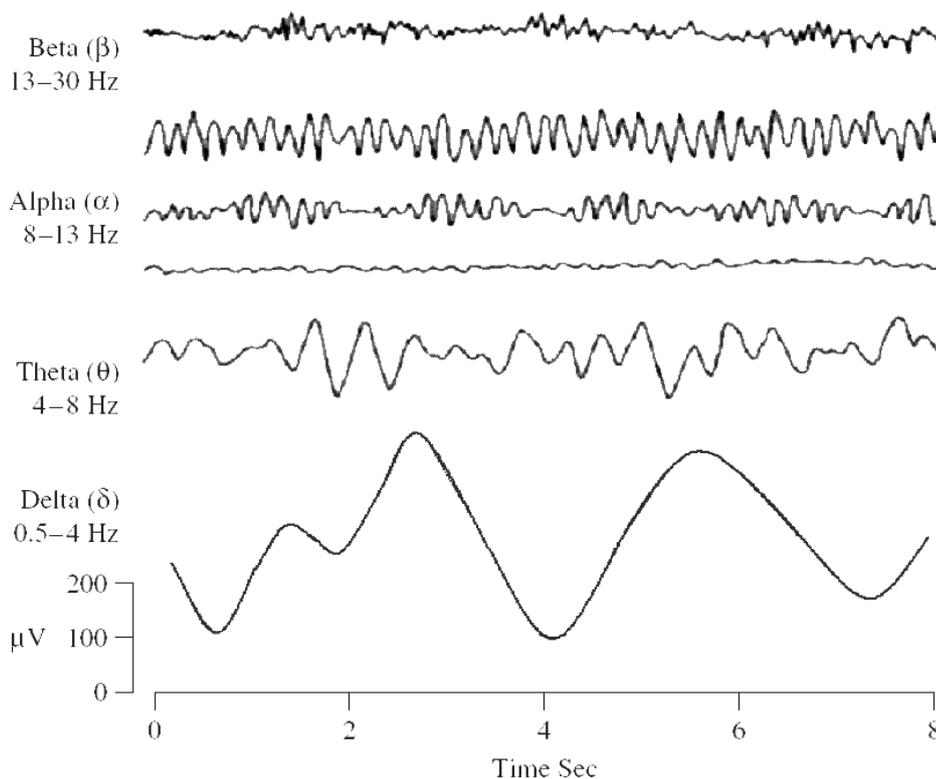


Figura 1.3: Ritmi cerebrali

1.2.1 Ritmi

Il segnale dipende in gran parte da quanto è sincronizzata l'attività neurale: se i neuroni di una zona ricevono la stessa quantità di stimolazione, ma in tempi diversi, la risposta di ciascuno non può sommarsi spazialmente alle altre e ne risulterà un segnale più debole; se la stimolazione raggiunge i neuroni nello stesso istante, le risposte saranno sincronizzate e sommabili, e l'attività registrata assumerà un'ampiezza maggiore. Questa attività si registra nel tracciato EEG come variazioni dei *ritmi cerebrali* in termini di ampiezza, misurata in microvolt (μV), e di frequenza, misurata in hertz (Hz). Al variare di questi parametri si rilevano onde appartenenti a diversi ritmi i quali sono associabili differenti stati psicofisiologici di un soggetto: attività, sonno, concentrazione, stimolazioni sensoriali ed emotive.

In un tracciato EEG si possono individuare cinque ritmi, figura 1.3:

- ritmi Alfa (α), sono ritmi caratterizzati da una frequenza tra 8 e 13 Hz, una ampiezza variabile tra 10 e 200 μV , registrabili durante stati di veglia ma a riposo mentale, tipicamente ad occhi chiusi. Si manifestano soprattutto nella zona occipitale e parietale;
- ritmi Beta (β), sono ritmi caratterizzati da una frequenza tra 13 e 30 Hz, con una ampiezza variabile tra 8 e 30 μV . Sono ritmi dominanti in stati di veglia ad occhi aperti e durante una qualsiasi attività cerebrale, quasi continuo negli stati di allerta (detta fase di arousal). Tipicamente, ad attività neurale più intensa, sono associati segnali EEG di minore ampiezza e maggiore frequenza, ciò è spiegabile con il concetto di desincronizzazione anticipato nel paragrafo precedente, l'elevata attività neurale a cui è associato questo rit-

mo rende difficile la sommazione di segnali provenienti da diverse sorgenti, limitandone così l'ampiezza;

- ritmi Mu (μ), sono ritmi caratterizzati da frequenza simili alle onde Alfa, ma si rilevano principalmente nella parte di cervello che controlla il movimento volontario. Maggiormente riscontrate quando il corpo è fisicamente a riposo, una persona sopprime queste onde quando è impegnata in una qualche attività motoria, nell'immaginazione di un movimento o nella visione di un altro soggetto che compia una azione motoria.
- ritmi Theta (θ), sono ritmi predominanti nei neonati e presente in molte patologie cerebrali dell'adulto. Caratterizzato da una frequenza tra 4 e 8 Hz ed una ampiezza media di $75 \mu V$;
- ritmi Delta (δ), sono ritmi caratterizzati da una frequenza molto bassa, tra 0.5 e 4 Hz, e ampiezza tra 20 e $200 \mu V$, presenti soprattutto nell'infanzia e nella fase di sonno profondo. Analogamente ai ritmi Beta, ma in maniera opposta, i neuroni implicati in questi ritmi svolgono una attività più lenta e sincronizzata, per questo motivo la sommazione dei segnali è più elevata e l'ampiezza registrata degli elettrodi è maggiore.

1.2.2 Attività mentali e potenziali evocati

Per attuare il controllo su una BCI l'utente è tenuto a svolgere alcune attività mentali cui conferiscono particolari conformazioni ai tracciati EEG [6]. Le tipologie di attività mentale possono essere:

- Attività sensomotoria. Questa attività interessa i ritmi Mu e Beta, originati nella corteccia motoria primaria e sono maggiormente presenti quando il soggetto non è impegnato in processi di input sensoriali/motori o la produzione di output motori. Un movimento volontario è riscontrato da una de-sincronizzazione in queste bande di frequenza la quale è denominata event-related desynchronization (ERD). L'immaginazione del movimento provoca ugualmente questo fenomeno; per questo motivo l'immaginazione del movimento di un arto è usato di sovente come input nelle BCI le quali processano il segnale e identificano la tipologia di movimento (destra/sinistra, mano/piede) mediante algoritmi di classificazione che utilizzano come feature i valori che l'elettrodo rileva nelle suddette bande;
- P300. Quando si intervallano a stimoli frequenti degli stimoli infrequenti o particolari, siano esse sonori, visivi o somatosensoriali, si rilevano, nell'EEG misurata nella zona parietale della corteccia, dei picchi positivi a circa 300 millisecondi dopo la presentazione dello stimolo. Questi picchi sono definiti P300. Le BCI che operano mediante le P300 utilizzano una serie di possibilità (solitamente organizzate in matrici) che vengono illuminate in maniera casuale. L'utente è tenuto a concentrarsi su una di queste possibilità è nell'istante in cui la scelta effettuata viene illuminata, l'evento è inteso come infrequente. Il rilevamento delle P300 in concomitanza con l'accensione di una particolare possibilità determina la scelta del soggetto ed esegue la relativa azione. L'algoritmo utilizza come feature i campioni di segnale campionati nella zona parietale;
- Steady State Visual Evoked Potential (SSVEPs). Quando ad un soggetto viene sottoposto uno stimolo ripetitivo ad una determinata frequenza >5 Hz, una risposta in continua oscillazione alla frequenza di stimolazione e/o armonica è riscontrata nella corteccia visiva.

La determinazione è effettuata con operazioni di filtraggio spaziale (tra gli elettrodi) e temporale (filtraggio centrato approssimativamente nella frequenza di stimolo);

- Potenziali corticali lenti. Questi sono lenti cambiamenti di potenziali senza movimento volontariamente generati dal soggetto. Si riflettono cambiamenti nella polarizzazione corticale dell'EEG della durata variabile tra 300ms ed alcuni secondi. Sono fortemente condizionati dalla capacità del soggetto;
- Risposte a compiti mentali. Si assume che diversi lavori mentali producano distinte e specifiche distribuzioni dei pattern frequenziali dell'EEG sullo scalpo.

1.3 Classificazione delle BCI

A seguito dello sviluppo tecnologico, e alla recente diminuzione dei costi dei dispositivi per eseguire le misurazioni, l'elettroencefalografia (EEG), mediante le brain-computer interface (BCI), si è affermata sempre più come strumento utilizzato, in ambito di ricerca, per il riconoscimento e la classificazione automatica degli stati emotivi.

Una Brain-Computer Interface (BCI) è un sistema di comunicazione in cui i messaggi o i comandi utilizzati da un individuo per interagire con il mondo esterno non dipendono dalle normali vie di output cerebrali, costituite da fasci di nervi e muscoli, ma vengono codificati e utilizzati da un computer e da altri dispositivi ad esso collegato[7].

L'uso combinato di queste tecnologie trova applicazione in ambiti che spaziano dall'assistenza ai portatori di disabilità fisiche, ad esempio sedie a rotelle il cui spostamento è effettuato tramite traduzione e attuazione di comandi "codificati" nelle onde cerebrali [8], alle applicazioni video-ludiche e multimediali[9].

Una prima classificazione suddivide le BCI in *dipendenti* e *indipendenti*:

- le BCI dipendenti sono interfacce che non utilizzano le normali vie di comunicazione cerebrali ma vie alternative. Mediante queste BCI non vi è quindi l'esclusione del tratto muscolare periferico ma sostituiscono una via interrotta o deteriorata affinché si possa generare l'attività cerebrale che codifica il messaggio da inviare al computer. In questo tipo di BCI si fa riferimento a segnali fisiologici generati per movimento muscolare, ad esempio i potenziali generati dallo spostamento degli occhi di una persona inferma;
- le BCI indipendenti sono interfacce che non dipendono in alcun modo dalle vie neuromuscolari. Non è necessario l'intervento di queste vie per generare l'attività cerebrale necessaria per l'invio dell'informazione codificata al computer. I potenziali sono quindi generati dal "pensiero" di svolgere una azione il quale, con un buon training, si manifesta in pattern più o meno specifici in grado di attivare la BCI.

Mentre una seconda classificazione, più classica, suddivide le BCI in *invasive* e *non invasive*:

- le BCI invasive, figura 1.4a, sono interfacce direttamente impiantate nella materia grigia mediante operazioni chirurgiche. Producono segnali con la massima qualità ma possono portare alla formazioni di cicatrici nella materia grigia che potrebbero diminuire la potenza del segnale, oltre ad avere possibili problematiche di rigetto da parte del corpo umano ad un corpo riconosciuto come estraneo. Nonostante l'invasività di questa tecnologia, la ricerca che utilizza questo tipo di BCI ha centrato i propri sforzi sull'assistenza a persone

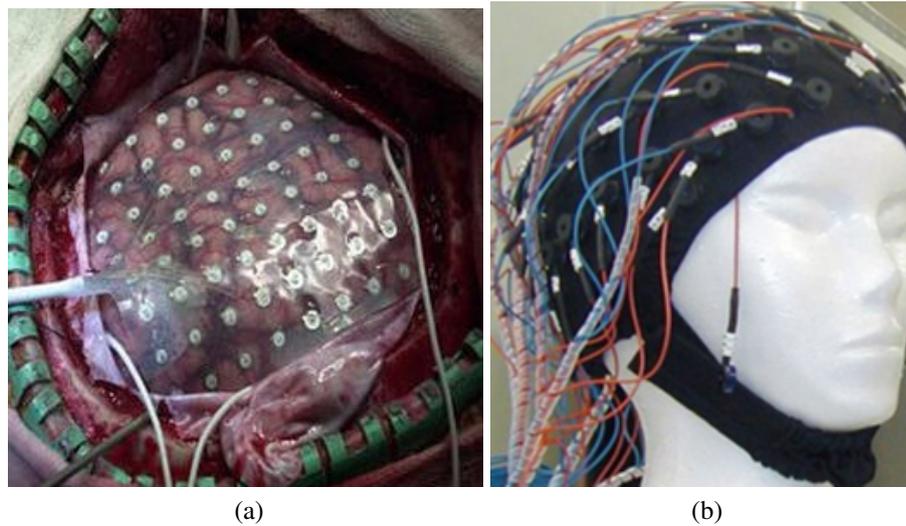


Figura 1.4: BCI invasive (a) e BCI non invasive (b)

con deficit motori, paralisi e alla riacquisizione di funzionalità sensoriali come la vista [10],[11].

- le BCI non invasive, figura 1.4b, sono interfacce che utilizzano le teorie fornite dall'elaborazione dei segnali e dalla neuropsicologia per il controllo e la comunicazione tra cervello e macchina, cercando di interpretare le intenzioni del soggetto, fino alle funzioni cognitive e di pianificazione. Utilizza segnali detti di neurofeedback, i quali sono segnali di condizionamento del sistema nervoso centrale (SNC), con cui si possono apprendere funzioni motorie senza la mediazione del sistema somatico-muscolare, ovvero senza la volontà del soggetto. I neurofeedback si presentano come pattern nell'attività cerebrale che possono essere identificati ed associati ad uno specifico messaggio o comando da inviare al computer. La facilità di installazione e di utilizzo di queste interfacce, rispetto a quelle invasive, tende a favorirle nell'impiego clinico, fornendo uno strumento per l'ausilio a soggetti affetti da patologie che portino sia a condizione di locked-in state, patologie in cui le facoltà motorie sono limitate ad un solo muscolo, ad esempio la SLA o la sclerosi multipla, sia alla completa infermità o completely locked-in state [12].

1.4 Architettura di una BCI

Quando si progetta una BCI bisogna tenere in considerazione diversi aspetti preliminari, tra questi: le finalità per cui è progettata, le tipologie di segnali fisiologici che si intendono acquisire, le tecniche di sollecitazione degli stimoli e i dispositivi che verranno utilizzati al termine della catena. In figura 1.5 è rappresentata schematicamente l'architettura basilare di tutte BCI, composta da diversi blocchi funzionali:

- l'acquisizione del segnale è la fase in cui i segnali dal cervello vengono acquisiti attraverso degli elettrodi posti sullo scalpo, sulla corteccia o direttamente innestati nel cervello, vengono amplificati e digitalizzati. Bisogna fare distinzione tra segnali di risposte evocate e segnali spontanei. Le risposte evocate sono reazioni del cervello a determinati stimoli

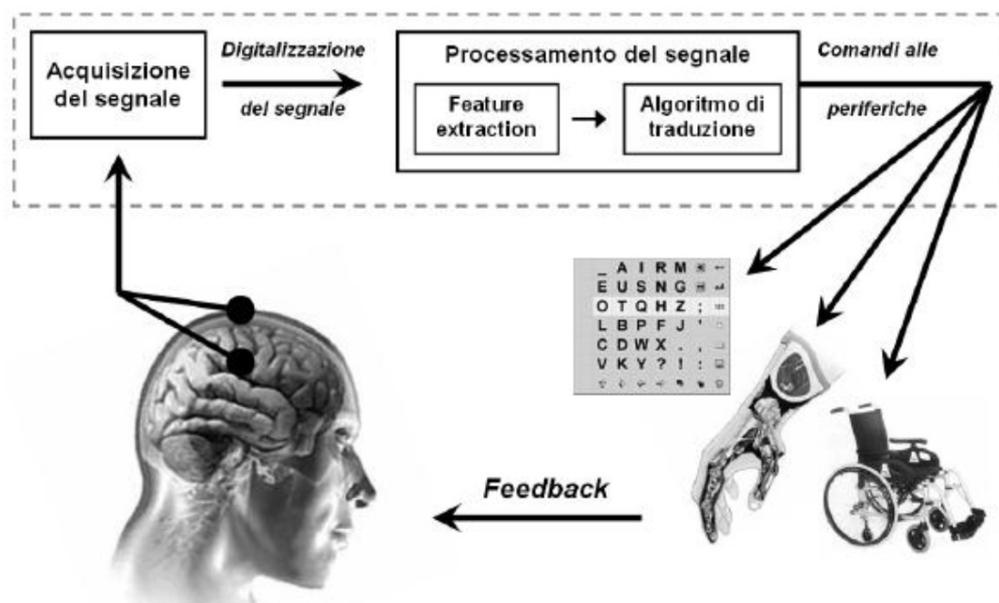


Figura 1.5: Architettura di una BCI

provenienti dall'esterno, a differenza dei segnali spontanei i quali non prevedono nessuna sollecitazione;

- processamento del segnale, fase in cui il segnale viene pulito dagli artefatti generati da segnali non strettamente correlati alle onde cerebrali, ad esempio EOG oppure EMG, e successivamente vengono estratte le features specifiche del segnale, come ad esempio l'ampiezza di determinati potenziali evocati o ritmi della corteccia motoria, i quali rispecchiano l'intento dell'utente. Le features possono essere nel dominio del tempo, delle frequenze oppure in entrambi i domini, le quali sembrano avere maggiore capacità discriminativa [13],[14];
- traduzione del segnale, nella quale le features sono tradotte attraverso algoritmi specifici in comandi da inviare alle periferiche, sia software che hardware, attraverso algoritmi che possono essere lineari, ad esempio i metodi statistici, oppure non lineari, come le reti neurali o le support vector machine (SVM) [15],[16]. L'efficienza dell'algoritmo si può valutare in base alla capacità di adattamento alle diverse risposte cerebrali dei diversi soggetti.
- i dispositivi terminali sono quelli che, progettati ad hoc, rispondono ai comandi inviati successivamente alla traduzione delle feature e possono essere di vario tipo: la possibilità di muovere un cursore su uno schermo, selezionare icone o muovere sedie a rotelle e protesi. Questi dispositivi, se provvisti della tecnologia adatta, inviano all'utente un segnale di feedback, il quale svolge una duplice azione di risposta e controllo.

1.4.1 Disposizione degli elettrodi

Nel 1947, al primo congresso internazionale riguardo l'EEG, si è realizzata la necessità di definire uno standard per la disposizione degli elettrodi sullo scalpo. Questo diviene utile per effettuare misurazioni comparabili tra soggetti e ricerche diverse. Per una accurata ricostruzione

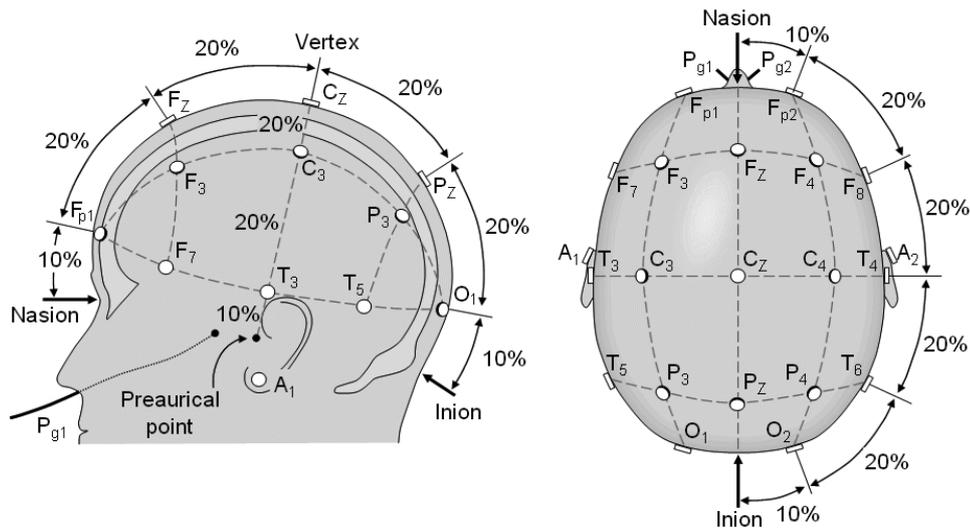


Figura 1.6: International 10-20 System

della attività corticale gli elettrodi sono applicati sullo scalpo secondo la disposizione standard denominata: International 10-20 System [3]. I numeri all'interno del nome si riferiscono alla percentuali di distanza tra due punti di riferimento cranici: inion, la prominenza alla base dell'osso occipitale, e nasion, l'attaccatura superiore del naso. Questa distanza varia tra le persone da 30 a 36 cm. In figura 1.6 è mostrata schematicamente la disposizione sullo scalpo degli elettrodi, il pedice Z indica quelli lungo la linea centrale del cranio, mentre i numeri pari e dispari indicano gli elettrodi disposti, rispettivamente, a destra e sinistra della stessa. A_1 e A_2 sono usati generalmente come riferimento. Occorre fare una precisazione riguardo le tecniche di riferimento, esistono tre metodi:

- unipolare, i potenziali di ogni elettrodo sono misurati comparandoli con quelli di un elettrodo di riferimento, generalmente posti in A_1 e/o A_2 , lo svantaggio di tale tecnica è una possibile distorsione del segnale dovuta alla eventuale presenza di attività vicino all'elettrodo di riferimento;
- bipolare, i potenziali sono misurati come differenza di tensione tra due diversi elettrodi, ciò permette di avere una risoluzione spaziale maggiore;
- riferito alla media, i potenziali di ogni elettrodo sono misurati rispetto alla media di tutti gli elettrodi, ciò risolve il problema riscontrabile con il metodo unipolare ma introduce inevitabilmente un'influenza reciproca nei pattern spaziali.

È possibile aumentare la risoluzione spaziale, e la quantità di informazione fornite dall'EEG, utilizzando un numero maggiore di elettrodi, diminuendone i rapporti di distanza. In figura 1.7 sono riportati i sistemi 10-10 System e 10-5 System [17].

1.5 Artefatti e rimozione

Se si vuole analizzare un tracciato EEG per calcolarne le features specifiche da utilizzare in sistemi BCI, bisogna tenere in considerazione il fatto che, seppur l'EEG possieda una buona risoluzione temporale ed è potenzialmente un ottimo strumento per analizzare l'attività cerebrale

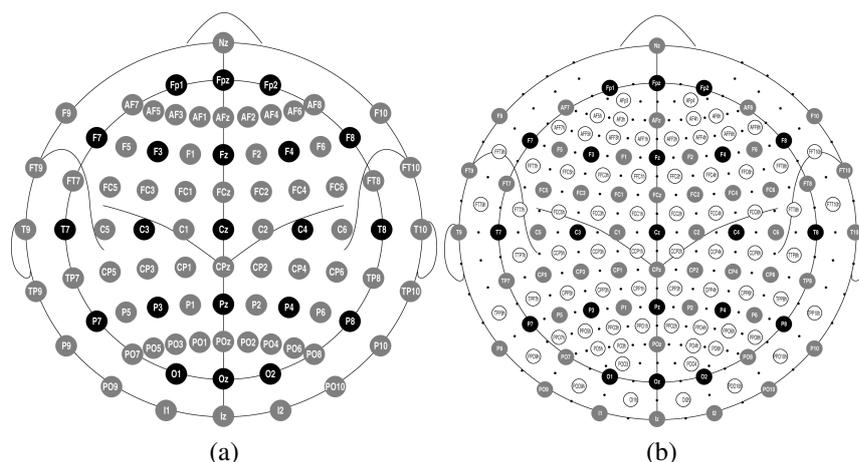


Figura 1.7: 10-10 Extension System (a) e 10-5 System (b)

e la ricognizione dello stato emotivo di un soggetto [18], è spesso modificato da elementi di disturbo e artefatti che sono relativamente correlati all'attività neurale e di origine extracorticale. Tra i più frequenti si possono riscontrare:

- interferenze di linea, le quali si manifestano come interferenze nelle frequenze 50 o 60 Hz, in funzione della linea elettrica utilizzata.
- attività cardiaca, misurabile attraverso ECG, la quale si manifesta come picchi anomali all'interno del segnale EEG che si ripetono ritmicamente in contemporanea al battito cardiaco oppure come onde lente ritmiche se l'elettrodo è posizionato sopra una arteria.
- movimenti dei muscoli oculari, rilevabili attraverso EOG, i quali sono dovuti principalmente all'azione di apertura/chiusura, denominata blink, e il movimento in generale degli occhi. Sono composti tipicamente da basse frequenze.

Per ovviare a questa problematica, e quindi aumentare il rapporto segnale/rumore del segnale EEG, sono state sviluppate diverse tecniche che sfruttano le teorie fornite dalla statistica, dall'elaborazione numerica dei segnali e dal buon senso. Prevenire la creazione di questi artefatti è spesso la soluzione migliore: dipendentemente dagli scopi della ricerca, disporre i soggetti in posizioni comode e rilassate, o fornire giustificate raccomandazioni, aiuta a prevenire artefatti dovuti, ad esempio, all'attività motoria, allo spostamento della testa o al serramento delle mascelle [19]. Artefatti dovuti ad una attività puramente fisiologica sono però difficili da prevenire, si può ricorrere quindi a diverse tecniche che operano nel dominio del tempo e nel dominio delle frequenze, tra cui: filtraggio del segnale con filtri adattivi in cascata [20], filtraggio e sottrazione con filtri di smoothing [21], algoritmi che utilizzino la Wavelet Transform, la Principal Component Analysis (PCA) o la Independent Component Analysis (ICA) [21],[22].

Capitolo 2

Emozioni e spazi affettivi

Le emozioni sono stati mentali e fisiologici associati a modificazioni psicofisiologiche, a stimoli interni o esterni, siano esse naturali o meno.

In termini evolutivisti, le emozioni rivestono il ruolo fondamentale nella gestione delle reazioni di un soggetto, in cui si rende necessaria una risposta immediata a determinate situazioni che minano la sopravvivenza.

2.1 Emozioni e cervello

La percezione di uno stimolo emotivo genera una serie di risposte complesse da parte dell'organismo. Le risposte interessano diverse strutture cerebrali:

- risposte ormonali, attraverso l'attivazione dell'ipotalamo;
- risposte esperienziali, attraverso la corteccia del cingolo anteriore;
- risposte sociali, attraverso il lobo frontale.

I primi studi effettuati permisero la formulazione della teoria per cui l'elaborazione degli stimoli emotivi era riservata al sistema limbico. Questo sistema era un circuito incentrato soprattutto sull'ippocampo e sulle sue connessioni con l'ipotalamo. L'evoluzione dei mezzi di studio, e le conoscenze acquisite sperimentalmente, hanno provato che il sistema non poteva essere composto solo da quegli elementi: si inclusero l'amigdala e la corteccia prefrontale[23].

Ciascuna area cerebrale risulta svolgere un ruolo specifico ed apporta un contenuto proprio all'elaborazione dello stimolo.

Le aree particolarmente attive nel processo sono [24]:

- ipotalamo, il quale è responsabile delle risposte fisiologiche dell'organismo, come la variazione della frequenza cardiaca, della temperatura corporea o della sudorazione;
- amigdala, la quale è raggiunta da fibre provenienti da tutti i lobi e dalle vie sensoriali, e svolge una azione principale nell'elaborazione delle emozioni;
- corteccia prefrontale, la quale è implicata nei processi emotivi ed affettivi e nella capacità di prendere decisioni finalizzate ad uno scopo. Da evidenze sperimentali si è dedotto che possa essere una plausibile zona di interazione tra processi cognitivi ed affettivi;

- corteccia cingolata anteriore, la quale è coinvolta nell'apprendimento emozionale condizionato, nella valutazione dei contenuti motivazionali e nella capacità di attribuire valore emotivo ai vari eventi;
- insula, la quale svolge un ruolo fondamentale nell'elaborazione dell'esperienze soggettive e nella valutazione di emozioni generate internamente.

2.1.1 Studi sulla percezione delle emozioni

Fin dal XIX secolo si è cercato di definire quali siano i meccanismi per la quale determinati stimoli provochino determinate reazioni emotive.

La teoria presentata in [25] propone la presenza di una retroazione sensoriale delle risposte emotive, diverse per ogni emozione, che ne determinino la percezione.

Con lo sviluppo delle scienze cognitive, nella metà del Novecento, l'interesse si è spostato dallo studio delle relazioni tra funzioni psicologiche e meccanismi neurali all'analisi di processi inerenti la percezione e la memoria (con un paradigma assimilabile alla computazione elettronica)[26].

La teoria dell'eccitazione cognitiva [27] afferma che le emozioni nascono dalla spiegazione che si attribuisce a certi stati fisici, emotivamente ambigui, sulla base di interpretazioni cognitive delle probabili cause, interne ed esterne.

Negli stessi anni, si fa strada la teoria della valutazione [28] secondo la quale il cervello deve prima valutare l'importanza di uno stimolo emotigeno, cui porterà alla formazione di certe tendenze all'azione (coscienti o meno) che determinano la formazione di emozioni.

Le neuroscienze cognitive si sono sviluppate insieme all'ingegneria informatica, la quale mediante le reti neurali, ha fornito la base virtualizzata del sistema di elaborazione delle attività cognitive umane. Esse trascurano lo studio delle emozioni, ma possono fornire una strategia che permette di studiarle indipendentemente dei caratteri soggettivi delle esperienze emotive [26].

2.2 Spazi affettivi

Sono stati creati alcuni modelli che mirano a fornire la rappresentazione di ciò che è lo *spazio affettivo*: il dominio entro la quale le emozioni possono essere classificate.

2.2.1 Modelli degli spazi affettivi

Modelli discreti

Esistono *modelli discreti* in cui ciò che si definisce è un insieme di emozioni basilari, innate e universali. Paul Ekman, si occupò dello studio delle emozioni e della loro classificazione e afferma che ogni stato emotivo è caratterizzato da agiti, manifestazioni neurofisiologiche e corrispettivi psicologici specifici [29].

Sulla base di queste affermazioni propose sei differenti emozioni basilari, identificate dall'espressione facciale: rabbia, disgusto, paura, felicità, tristezza e sorpresa. Queste emozioni sono denominate primarie e concorrono, sommandosi, alla formazione delle emozioni secondarie: il disprezzo come somma di rabbia e disgusto, oppure il disappunto come somma di sorpresa e tristezza.

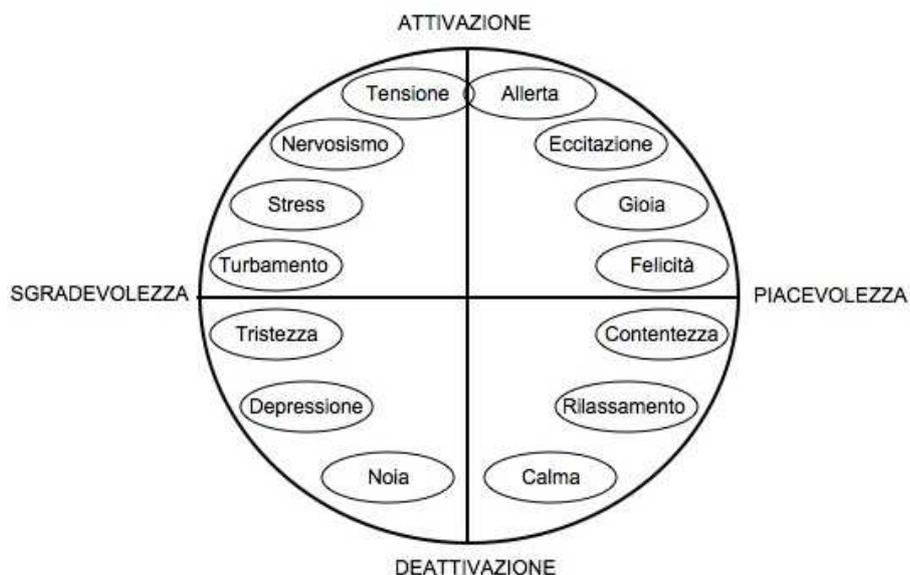


Figura 2.1: Modello circonflesso delle emozioni

Questo modello si dimostra insufficientemente adatto a descrivere alcune emozioni complesse. Inoltre, l'espressione di un'emozione basilare può variare in funzione delle differenze culturali ed interculturali, oltre che per differenze fisiche (ad esempio il genere sessuale)[30].

Modelli a valutazione o a stima

I *modelli a valutazione o modelli a stima* si basano sulla valutazione o stima di circostanze, siano esse correnti, ricordate o immaginate.

Introdotta da Magda B. Arnold [28], questo modello si basa sul concetto per cui ciò che si prova dipende strettamente da ciò che si ritiene importante, e quindi, la valutazione è legata in prima battuta a ciò che si vuole e successivamente a ciò che si prova. Ad esempio, la frustrazione è determinata dal fatto di non riuscire a raggiungere un obiettivo.

Modelli dinamici

I *modelli dinamici* considerano le emozioni come processi dinamici.

Secondo questo modello le emozioni sono caratterizzate in termini di andamento della risposta, e possono essere regolate attraverso cinque "tappe" che ne determinano la creazione: selezione della situazione, modificazione della situazione, sviluppo dell'attenzione, cambiamento della cognizione e modulazione della risposta.

Modelli dimensionali

I *modelli dimensionali* considerano, a differenza dei modelli discreti, uno spazio multidimensionale continuo in cui ogni dimensione è considerata come una proprietà fondamentale appartenente a ciascuna emozione.

Il modello a cui si riferirà l'elaborato appartiene a quest'ultima classe ed è definito Circumplex Model of Affects (CMA) (figura 2.1). In questo modello, introdotto da Russel [1], le cause

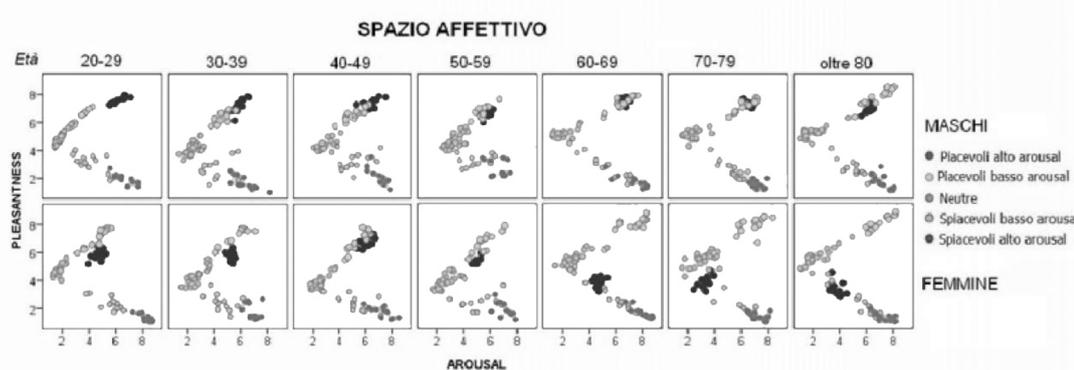


Figura 2.2: Rappresentazione grafica dello spazio affettivo in funzione dell'età e del genere

alla base dei meccanismi emotivi sono interpretate come uno spazio continuo di stati molto interconnessi tra loro e molto spesso ambigui.

Tra le dimensioni proposte, lo spazio affettivo è rappresentato ponendo:

- sull'asse delle ascisse la *valenza edonica*, la quale rappresenta quanto una determinata emozione è stata percepita come positiva o negativa;
- sull'asse delle ordinate l'*arousal*, la quale determina la forza con cui l'emozione si manifesta e quanto l'evento scatenante è rilevante.

Questo approccio facilita l'identificazione e la caratterizzazione di una emozione grazie alla possibilità di utilizzare solo i due parametri valenza e arousal.

La rappresentazione e l'analisi qualitativa delle valutazioni dei soggetti sottoposti a stimolazione diviene quindi più semplice.

2.2.2 Spazi affettivi ed età

Come anticipato all'inizio di questo capitolo, le emozioni sono associate a modificazioni psicofisiologiche dei vari soggetti. Le modificazioni psicofisiologiche sono dovute ai processi di risposta, introdotti nella prima sezione, i quali assumono carattere sia psicologico (come le risposte dovute a seguito della valutazione di una situazione specifica) sia fisiologico (come la produzione di ormoni). Ci si potrebbe domandare se queste risposte possano differenziarsi in funzione dell'età; se l'esperienza pregressa di un soggetto adulto influisca sulle valutazioni, oppure se la differente produzione di ormoni nella pubertà e in età senile abbia influenza sulla percezione.

In [31] si utilizzano immagini come stimolo e i risultati ottenuti (rappresentati in figura 2.2) mostrano che:

- per quanto concerne la valenza, stimoli con arousal alto (prevalentemente a contenuto erotico), subiscono una maggiore differenziazione nei due sessi all'avanzare dell'età. I soggetti maschi tendono a valutare questo tipo di immagine più positivamente rispetto ai soggetti di sesso femminile. Gli stimoli a basso arousal (famiglie e bambini) risultano valutate più positivamente dalle donne rispetto agli uomini. Per le immagini a bassa valenza, ovvero spiacevoli, non si sono riscontrate differenze rilevanti;

Computer	Incapaci di esprimere emozioni	Capaci di esprimere emozioni
Incapaci di riconoscere emozioni	I.	II.
Capaci di riconoscere emozioni	III.	IV.

Tabella 2.1: Categorie di affective computing focalizzate sul riconoscimento e sull'espressione delle emozioni

- in termini di arousal, nei soggetti di sesso maschile la tendenza a valutare l'intensità di attivazione di una determinata immagine rimane coerente con l'aumentare del livello di arousal. Nelle femmine le immagini ad alto e basso arousal vengono valutate come egualmente attivanti fino all'intervallo di età riferito ai 50-59 anni. In età più avanzata si riscontra un'inversione della piacevolezza e dell'intensità delle immagini ad alta valenza ed alto arousal (a contenuto erotico).

2.3 Affective computing

I risultati ottenuti dagli studi inerenti la generazione, l'elaborazione e la percezione delle emozioni, insieme alla contemporanea evoluzione della tecnologia a disposizione della ricerca, hanno posto le basi per l'avvio dell'*affective computing*. Quest'ultimo è lo studio e lo sviluppo di sistemi e dispositivi in grado di riconoscere, interpretare, processare e simulare emozioni umane [32].

Numerosi studi sono stati condotti per determinare metodi i quali, mediante l'interazione tra uomo e macchina, potessero essere utili per la riduzione della frustrazione dell'utente, per lo sviluppo di infrastrutture atte all'acquisizione delle informazioni emotive, per la progettazione di strumenti che possano migliorare le abilità socio-emotive e per l'apprendimento dei comportamenti umani [33],[34].

L'affective computing è una materia interdisciplinare che include e fonde concetti provenienti dalla psicologia, dalle scienze cognitive, dalla fisiologia e dall'informatica.

Nella tabella 2.1 viene proposta una classificazione delle tipologie di affective computing in cui[2]:

- I. Identifica la maggior parte dei computer, non vi è personalizzazione.
- II. Identifica la categoria in cui i computer assumono un comportamento naturale, delle tre categorie che includono le emozioni è la più avanzata tecnologicamente.
- III. Identifica i computer che sono in grado di riconoscere lo stato emotivo dell'utente ed adattare dinamicamente il proprio funzionamento, senza avere la possibilità di esprimere a propria volta uno stato emotivo.
- IV. Identifica la categoria che massimizza la comunicazione tra umano e macchina, potenzialmente offre una vera esperienza "personal" e non implica che il computer sia guidato dalle emozioni umane.

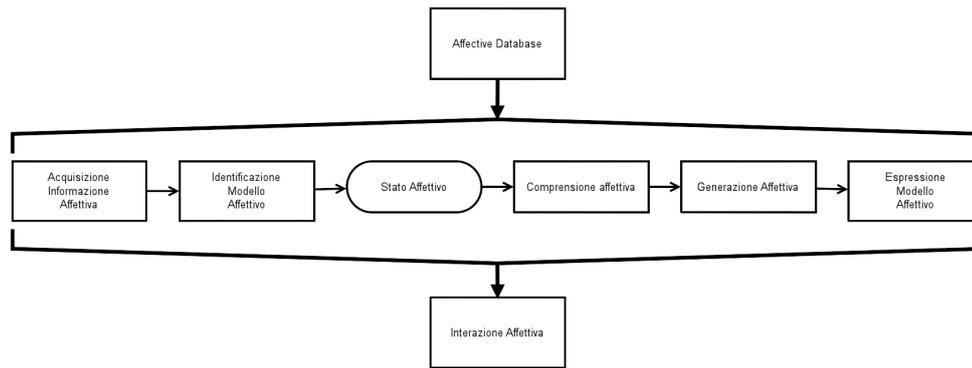


Figura 2.3: Framework di ricerca per l'interazione basata sulle emozioni

La figura 2.3 illustra la procedura standard per l'interazione basata sulle emozioni [32], prevede che vi sia una fase di acquisizione delle informazioni, la successiva modellazione, l'apprendimento dello stato emotivo e la sua espressione.

La fase di acquisizione dell'informazione affettiva rappresenta una parte cruciale per l'avvio dell'interazione. Dopo un'accurata tecnica di sollecitazione emotiva vengono considerati i diversi modi espressivi delle emozioni umane come espressioni facciali, movimenti del corpo, il comportamento della voce e altri segnali fisiologici. Acquisiti i segnali contenenti l'informazione emotiva, vi è la necessità di elaborare tali segnali al fine di ricavarne le features, ovvero alcune caratteristiche specifiche, inerenti ad un particolare stato emotivo di un soggetto per poter infine definire un modello utilizzabile nelle fasi seguenti[32]. Numerosi segnali fisiologici sono stati studiati come possibili indicatori degli stati emotivi, tra cui: elettrocardiogramma (ECG), risposta elettrodermica (EDR), reazione galvanica cutanea (GSR), elettromiogramma (EMG), battito cardiaco (heart rate, HR), variazioni di respirazione, temperatura cutanea, diametro pupillare[30],[35],[36],[37].

Tuttavia il riconoscimento di uno stato emotivo associato ad un soggetto non è di facile realizzazione; le problematiche che si riscontrano sono di varia natura e principalmente sono dovute alla non linearità delle risposte emotive alle sollecitazioni: le tecniche utilizzate possono talvolta dare luogo a risposte differenti.

Capitolo 3

Indagine sperimentale

Al fine di raggiungere l'obiettivo di questo elaborato è stata condotta una indagine sperimentale, la quale ha permesso di raccogliere i dati necessari per una primaria valutazione.

Basato sulla procedura standard dell'interazione emotiva, illustrata in figura 2.3 nel Capitolo 2, l'approccio ripercorre i primi step utilizzando coppie di immagini e suoni per generare una risposta emotiva nel soggetto. I ritmi cerebrali sono registrati attraverso una BCI e successivamente processati in ambiente MatLab per valutarne l'andamento nei diversi set di stimolazione.

3.1 Stimolazione affettiva

La sollecitazione di una emozione non è banale e pone problemi di natura pratica (e.g., procurare la stessa sensazione in soggetti diversi) ed etica (e.g., suscitare emozioni negative). Vi sono diverse tipologie di sollecitazione [6]:

- uso di set standard composti da film, clip, immagini, suoni... ;
- tecniche di immaginazione, le quali consistono nell'immaginazione e richiamo di elementi autobiografici per sollecitare emozioni già provate. È chiesto al soggetto di ricordare una particolare situazione oppure è guidato attraverso una storia che qualche volta include aiuti audio visivi;
- interazioni sociali prestabilite, in cui vengono utilizzate manipolazioni di forte impatto per ricreare situazioni realistiche le quali provocano emozioni simili a quelle provabili in situazioni del mondo reale.

In questo elaborato è stato utilizzato il primo metodo di stimolazione, attingendo materiale fotografico e audio da due database: *International Affective Picture System*, o IAPS [38], e *International Affective Digitalized Sounds*, o IADS [39].

3.1.1 Database affettivi: IAPS e IADS

Entrambi i database, IAPS e IADS, sono stati progettati e creati al *NIMH Center for the Study of Emotion and Attention*, University of Florida. L'uso improprio del materiale, compresa la divulgazione e la riproduzione delle immagini sui testi scientifici, è vietata; ciò per non comprometterne l'affidabilità di stimolazione.

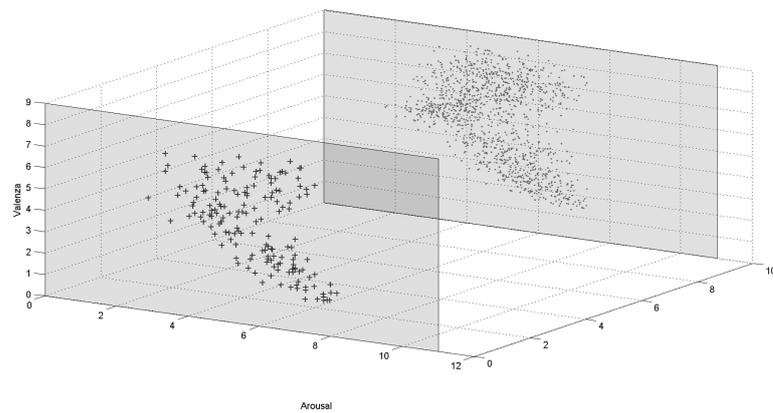


Figura 3.1: Distribuzione nello spazio valenza/arousal degli elementi contenuti in IAPS e IADS

IADS		Valenza						
		1	2	3	4	5	6	7
Arousal	1							
	2					1		
	3				1	3	2	1
	4			1	14	13	6	4
	5		3	5	9	7	12	3
	6		11	19	2	6	7	8
	7	8	11	1			5	3
	8		1					

(a)

IAPS		Valenza							
		1	2	3	4	5	6	7	8
Arousal	1				2				
	2				33	24	8	2	
	3		1	12	56	98	42	29	1
	4	1	25	45	44	66	82	57	12
	5	17	75	65	28	25	70	51	3
	6	32	57	20	9	12	38	22	1
	7	9	2				2	4	
	8								

(b)

Tabella 3.1: Distribuzione dei file contenuti in IADS (a) e IAPS (b) nelle diverse classi di valenza e arousal unitari ($v1 = \text{valenza} \cdot 1.00 \div 2.00$)

La creazione è stata eseguita sottoponendo ogni singolo stimolo ad un vasto numero di soggetti, chiedendo di dare una personale valutazione in termini di valenza e arousal. Ogni immagine e audio presente è quindi catalogata secondo questi valori, riportati poi in tabelle suddivise per sesso dei soggetti: valutazioni femminili, valutazioni maschili e valutazioni generali. Nell'elaborato sono stati considerati i valori della classificazione generalista poiché i soggetti dell'esperimento sono di entrambi i sessi.

Definendo con A e I l'insieme degli audio e delle immagini, la figura 3.1 rappresenta la distribuzione nello spazio valenza/arousal dei campioni audio $a_i(\text{val}, \text{aro}) \in A \subseteq \mathbb{R}^2$ e delle immagini $i_j(\text{val}, \text{aro}) \in I \subseteq \mathbb{R}^2$.

Mediante script Matlab (vedi A.1.1) è stato effettuato l'import e la suddivisione in classi unitarie di valenza e arousal di entrambi i database. Nelle tabelle 3.1a e 3.1b è riportato il numero di file per ogni classe di valenza e arousal.

La Tabella 3.2 riassume le medie, con la relativa deviazione standard, dei valori di valenza e arousal per ogni classe di valenza dei database.

Il modello dimensionale che descrive lo spazio affettivo definito in termini di valenza e arousal (figura 2.1) è tipicamente centrato nell'origine degli assi e associa quindi valori di valenza sempre più positivi per emozioni positive e sempre più negativi per emozioni negative. Il sistema adottato da IAPS e IADS invece utilizza esclusivamente valori positivi, ne consegue che per valori tanto più vicini a zero sono associate emozioni negative, mentre per valori $\gtrsim 5.5$

	Classe di valenza	n° elementi	Medie			
			Valenza	dev.	Arousal	dev.
IADS	v1	8	1.7413	0.1505	7.6675	0.2610
	v2	26	2.4938	0.3406	6.9088	0.6843
	v3	26	3.4373	0.2953	6.2527	0.5904
	v4	26	4.5292	0.2905	4.9958	0.7282
	v5	30	5.4870	0.3216	5.0453	0.9970
	v6	32	6.5072	0.3268	5.6072	1.1040
	v7	19	7.3816	0.2557	5.8674	1.0855
	Totale	167				
IAPS	v1	61	1.7389	0.1711	6.4023	0.6173
	v2	160	2.5210	0.2714	5.6957	0.7274
	v3	144	3.5101	0.2699	5.1917	0.7703
	v4	174	4.5757	0.3121	4.0074	1.1102
	v5	225	5.4692	0.2837	4.0678	0.9564
	v6	244	6.4843	0.2777	4.8290	1.0412
	v7	168	7.3569	0.2577	4.9347	0.9981
	v8	17	8.1294	0.1011	4.7459	0.6924
	Totale	1193				

Tabella 3.2: Medie e deviazione standard dei valori di valenza e arousal per ogni intervallo unitario di valenza

gli stimoli assumono carattere positivo. Considerazioni analoghe valgono anche per i valori di arousal.

3.1.2 Selezione degli stimoli affettivi

La selezione del materiale da utilizzare come stimolo è stata eseguita cercando di massimizzare lo scollamento tra i valori di valenza. Si sono perciò considerati solamente gli elementi appartenenti agli intervalli evidenziati nelle tabelle 3.1a e 3.1b. Si è preferito un approccio oggettivo, in cui l'unica selezione non automatica ha riguardato gli audio iniziali, formalizzabile come segue:

1. creazione di griglie per $A, I \subseteq \mathbb{R}^2$:

$$\mathcal{G}(A) = \{x = k, y = j\} : k \in \{1, \dots, 7\}, j \in \{1, \dots, 8\} \quad (3.1)$$

$$\mathcal{G}(I) = \{x = l, y = m\} : l \in \{1, \dots, 7\}, m \in \{1, \dots, 7\} \quad (3.2)$$

2. uniformazione degli spazi

$$\mathcal{G}'(A) = \{x = k, y = j\} : k \in \{1, \dots, 7\}, j \in \{1, \dots, 7\} \quad (3.3)$$

3. selezione di sette audio iniziali $\{a_1, \dots, a_7\} \in A : a_i = (k_i, j_i)$ con $i \in \{1, \dots, 7\}$;
4. possiamo immaginare i rettangoli A e I come giacenti su piani paralleli nello spazio. (Vedi figura 3.1). Sebbene sarebbe opportuno usare una nuova notazione come per esempio Π_A e Π_I per indicare i due rettangoli immersi in \mathbb{R}^3 , con un abuso di notazione continueremo a chiamarli A e I .

5. per la definizione 3.3, $\mathcal{G}(A)$ e $\mathcal{G}(I)$ sono isomorfi dunque definiamo

$$a \in A \text{ se } a = (x, y, z_A)$$

$$i \in I \text{ se } i = (x, y, z_I)$$

definiamo la proiezione di a su I come

$$\pi_A : A \mapsto \pi(A) \subseteq I$$

$$a = (x_a, y_a, z_A) \mapsto (x_a, y_a, z_I)$$

e cerchiamo l'immagine $\bar{i} \in \Pi_I$ tale che minimizzi la distanza dalla proiezione π_A

$$\bar{i} = \min_{i \in I} \{d(\pi_A(a), i)\} \text{ per } a \in A$$

6. ricerca audio scorrelato: sia $\bar{i} \in I$ l'immagine correlata all'audio $\bar{a} \in A$, sia

$$\pi_I : I \mapsto \pi(I) \subset A$$

$$i = (x_i, y_i, z_A) \mapsto (x_i, y_i, z_I)$$

la proiezione dell'immagine nello spazio degli audio e cerchiamo l'elemento che massimizzi la distanza da $\pi(I)$ cercando nella cella che ha valori opposti di valenza e arousal. Poichè IADS ha dati sparsi (si veda tabella 3.1a) la massimizzazione porta ad una serie di risultati identici, pertanto si è proceduto come segue (vedere A.1.5):

- (a) se I è l'immagine, settiamo i valori di ricerca della valenza v_r e arousal a_r come $v_r = \text{floor}(\text{val}(I))$ e $a_r = \text{floor}(\text{aro}(I))$;
- (b) la ricerca viene effettuata inizialmente tra gli audio A contenuti nella cella con valenza $v(A) = 7 - (v_r - 1)$ e arousal $a(A) = 7 - (a_r - 1)$;
- (c) se nella cella non è presente alcun audio allora $a(A) = a(A) + 1$ e si itera nuovamente la ricerca con i nuovi parametri;
- (d) la ricerca continua fino a che non sia disponibile un set di possibili audio corrispondenti ai primi 5 che massimizzino la distanza tra I e A ;
- (e) poiché è possibile avere un risultato uguale per più audio, l'associazione di A ad I viene ottenuta considerando il primo audio non attribuito ad una qualsiasi immagine.

Osservazione. Matematicamente questo procedimento corrisponde a centrare il rettangolo A nell'origine piegandolo lungo la sua diagonale che passa per il secondo e il quarto quadrante, tornando in questo caso a minimizzare le distanze tra la proiezione dell'immagine e gli elementi audio.

Mediante la funzione nativa $\text{pdist2}()$ di MatLab, sono state calcolate: distanza Euclidea, distanza di Minkovski, distanza di Hamming e distanza City Block. Dopo diverse prove, le differenze riscontrabili nei campioni selezionati utilizzando le diverse distanze non è stato rilevante; si è proceduto pertanto considerando la sola distanza Euclidea.

In tabella 3.3 sono riportati i valori di valenza e arousal per ogni stimolo.

SET	Immagine		Audio			
	Valenza	Arousal	Correlato		Scorrelato	
			Valenza	Arousal	Valenza	Arousal
1	2.96	6.82	2.89	6.91	6.62	3.36
2	2.37	7.35	2.78	7.54	6.84	4.46
3	1.59	7.34	1.68	7.88	7.44	3.36
4	6.44	7.07	6.47	7.32	2.80	5.39
5	7.21	7.31	6.94	7.54	1.63	7.79
6	7.40	6.58	7.32	6.56	1.68	7.95
7	7.63	6.07	7.64	6.03	1.57	7.27

Tabella 3.3: Valori di valenza e arousal per ogni stimolo, SET(1,2,3) assumono carattere sempre più negativo mentre SET(4,5,6,7) assumono carattere sempre più positivo. La scorrelazione è massima per SET(3) e SET(7).

3.2 Acquisizione dati

La stimolazione è stata condotta mediante l'utilizzo di un laptop in cui è installato il software BCI2000 [12] e della interfaccia BCI Emotiv Epoc [40].

3.2.1 Interfaccia

L'interfaccia utilizzata, figura 3.2a, dispone di 14 sensori più 2 canali di riferimento e giroscopio. I potenziali corticali sono campionati ad una frequenza di campionamento di 128 Hz e l'invio dei pacchetti dati verso il pc avviene wireless attraverso un dongle USB. La posizione degli elettrodi sullo scalpo rispetta le posizioni standard e in figura 3.2b è riportata schematicamente la disposizione rispetto al sistema 10-5. Gli elettrodi di rame presenti sulla interfaccia sono smontabili e il contatto con lo scalpo avviene attraverso dei tamponcini umettati con soluzione salina. L'utilizzo di soluzione fa sì che migliori la conduttività ma introduce il problema dell'ossidazione degli elettrodi, come nel caso dell'interfaccia utilizzata, i quali devono essere

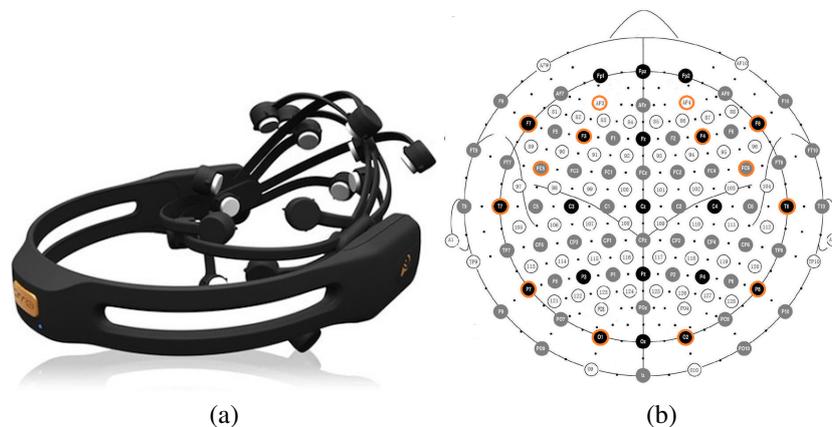
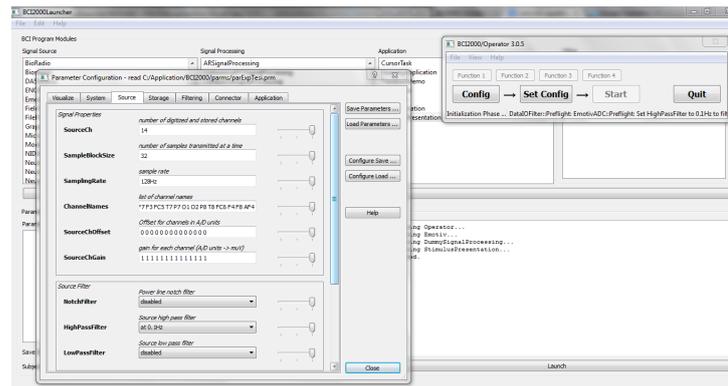
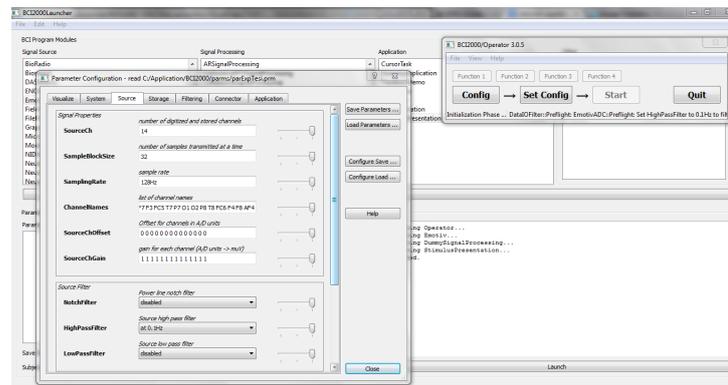


Figura 3.2: Emotiv Epoc



(a)



(b)

Figura 3.3: BCI2000: launcher per la selezione dei moduli e delle impostazioni (a) e fine di una sessione di prova con simulatore di segnale (b)

puliti e privati dell'ossido. Per l'esperimento sono stati utilizzati 6 canali nelle posizioni: AF3, AF4, F3, F4, O1 e O2, oltre ai canali di riferimento posizionati in P3 e P4.

3.2.2 Applicativo

BCI2000 è un software general purpose per la ricerca nel campo delle interfacce cervello-computer. Può anche essere usato per l'acquisizione dei dati, la presentazione di stimoli e applicazioni di monitoraggio dell'attività cerebrale. L'architettura di BCI2000 è basata sul modello illustrato nella figura 1.5 e consiste in quattro moduli [12]:

1. *Source*, digitalizza e memorizza i segnali derivanti da una interfaccia hardware e li passa al modulo successivo senza alcun tipo di preprocessing. Il formato file (.dat) consiste in un header ASCII, seguito dai valori binari dei campionamenti.
2. *Signal Processing*, converte i segnali cerebrali in segnali di controllo per dispositivi esterni elaborandoli in due fasi, l'estrazione delle feature e la traduzione di quest'ultime.
3. *User Application*, riceve i segnali di controllo dal modulo precedente e li usa per guidare un'applicazione.
4. *Operator*, definisce i parametri del sistema, come la durata degli stimoli o il valore di una specifica variabile, l'inizio e la fine di una operazione.

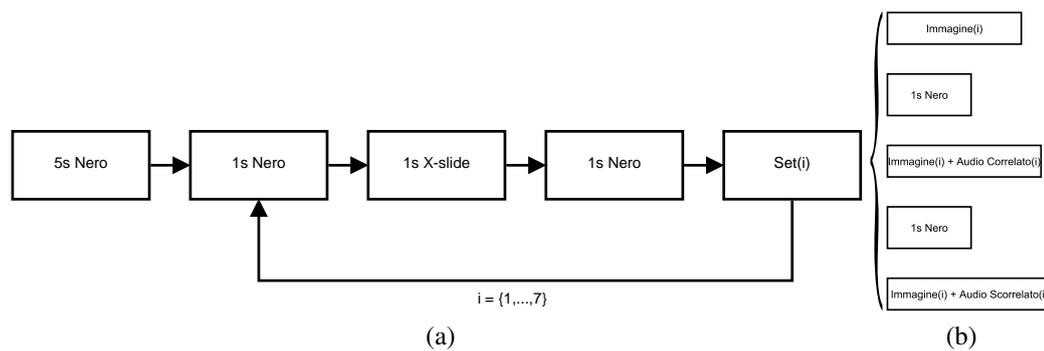


Figura 3.4: Protocollo sperimentale (a) e struttura generica di un set di stimolazione (b)

I moduli comunicano tra loro attraverso un protocollo di rete basato su TCP/IP, questo rende possibile sviluppare in qualsiasi linguaggio di programmazione ed eseguire l'applicazione su ogni macchina connessa in rete.

La figura 3.3a mostra l'interfaccia con la quale è possibile selezionare i moduli e gestire le impostazioni per l'esecuzione dell'esperimento:

- “Emotiv” è selezionato come modulo *Source*, il quale rende possibile l'acquisizione dei dati provenienti dall'interfaccia utilizzata;
- “DummySignalProcessing” è selezionato come modulo *Signal Processing*, è un modulo semplice che non esegue nessun processing del segnale. L'interesse è quello di acquisire i dati raw provenienti dall'interfaccia.
- “StimulusPresentation” è selezionato come modulo *User Application*, mediante cui è possibile presentare stimoli visivi e audio.

La figura 3.3b mostra il termine di una sessione di prova in cui è utilizzato un generatore di segnale come sorgente e la presentazione di stimoli audio/visivi come applicazione.

3.2.3 Protocollo sperimentale

La stimolazione è avvenuta posizionando i soggetti davanti allo schermo, con cuffie auricolari e interfaccia BCI posizionata sullo scalpo. La popolazione è composta da 31 soggetti (22 maschi e 9 femmine) volontari di età compresa tra 20 e 30 anni. L'intero esperimento si compone di sette stimolazioni in cui, per natura del metodo utilizzato nella selezione del materiale emotigeno, i valori di valenza e arousal aumentano gradualmente. Il protocollo sperimentale è illustrato in figura 3.4a. Ogni set sperimentale è composto da una sequenza di tre stimoli, come in figura 3.4b: l'immagine, l'immagine con audio correlato e l'immagine con audio scorrelato. Ogni stimolo dura 6 secondi e sono intervallati da 1s di schermo nero. Ogni set è anticipato dalla proiezione di una “X” (1s) per attirare la visione del soggetto sullo schermo [41]. Le immagini dei primi tre set sono a carattere negativo mentre le restanti sono a carattere positivo, tale ordine è stato attuato per avere il massimo scostamento al termine di ogni gruppo (positivo e negativo) ed evitare possibili effetti di “mascheramento emotivo” (una possibile emozione più intensa che possa predominare su un'altra più lieve [41]).

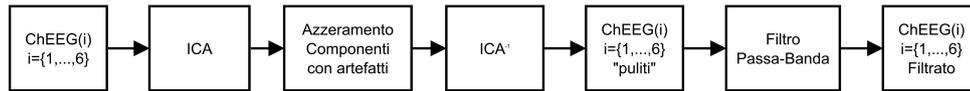


Figura 3.5: Procedura di preprocessing del segnale EEG per ogni soggetto

3.3 Processamento

I segnali corticali campionati dall'interfaccia sono contenuti in un file con estensione .dat che necessita di una conversione in un formato importabile all'interno dell'ambiente MatLab; a tale scopo è stato utilizzato un convertitore presente nel pacchetto dati di BCI2000 al fine di generare un file con estensione .CSV. Tramite script sono stati importati i file relativi a tutti i soggetti conservando solo i segnali rispettivi ai canali di interesse: AF3, AF4, F3, F4, O1 e O2. Si è proceduto poi preprocessando il segnale per l'eliminazione dei possibili artefatti e l'eliminazione delle frequenze non considerate nello studio. Successivamente i segnali sono stati segmentati in accordo con la durata degli stimoli per creare gli Event Related Potential (ERP) e si è eseguita l'estrazione delle feature per l'analisi dei dati (vedere A.2).

3.3.1 Preprocessing

I ritmi cerebrali e il rumore condividono talvolta porzioni di spettro, ad esempio i ritmi alfa sono spesso mixati con i movimenti oculari EOG; la rimozione completa degli artefatti comporta anche la perdita di informazioni utili dei segnali EEG ed è impossibile rimuoverli interamente in fase di acquisizione [13],[20].

Per poter analizzare meglio i segnali ed formulare delle considerazioni, il preprocessing è stato eseguito in due fasi: applicazione dell'algorithmo ICA per la ricerca e l'isolamento degli artefatti e successivamente filtraggio passa-banda con filtro di Čebyšëv. La figura 3.5 riassume schematicamente la procedura di preprocessing.

Independent Component Analysis (ICA)

L'ICA è nata per risolvere il più generale problema di separazione ed individuazione di sorgenti nascoste (Blind Source Separation BSS). Il problema consiste nel determinare i segnali originali, le sorgenti, dopo che questi sono stati sottoposti ad un'operazione di mixing.

Ora è un metodo largamente utilizzato per la rimozione degli artefatti nel processing dei segnali elettrofisiologici. L'algorithmo richiede dati di molteplici canali EEG e si basa sull'assunzione che il segnale registrato sia una combinazione lineare di sorgenti indipendenti all'interno del cervello. L'obiettivo è di trovare una trasformazione lineare dei segnali misurati dai sensori tale che ogni componente risulti il più indipendente possibile. Dopo il calcolo delle componenti indipendenti, si può ottenere un segnale EEG pulito mediante l'eliminazione delle componenti corrispondenti agli artefatti [21].

Definizione. Siano rispettivamente $\bar{x} = (x_1, \dots, x_m)^T$ e $\bar{s} = (s_1, \dots, s_n)^T$ i vettori delle osservazioni e dei segnali sorgenti. Il sistema allora si può scrivere come:

$$\bar{x} = A \cdot \bar{s} \quad (3.4)$$

dove con A si indica la matrice di mixing.

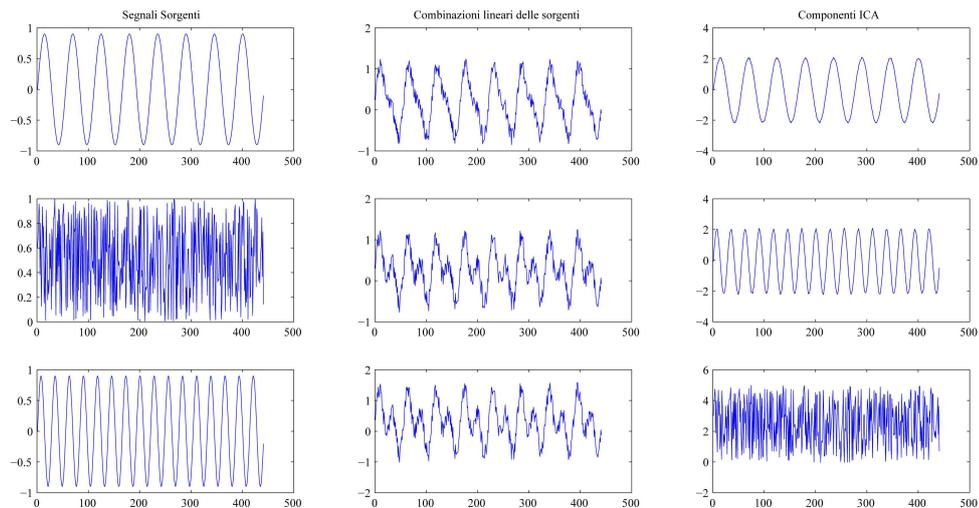


Figura 3.6: Esempio di applicazione ICA

Il problema è di stimare A e ricavare \bar{s} conoscendo solo il segnale misurato \bar{x} . Il problema ha soluzione se si riesce a trovare una matrice di trasformazione W tale per cui:

$$\bar{u} = W \cdot \bar{x} \quad (3.5)$$

dove \bar{u} è la componente indipendente, un'approssimazione della sorgente originaria \bar{s} . La matrice W è detta di unmixing ed è una stima della inversa di A , tale che:

$$\bar{u} = W \cdot \bar{x} = W \cdot A \cdot \bar{s} \quad (3.6)$$

e trovando W tale per cui $W \cdot A = I$, cioè $W = (A)^{-1}$ otteniamo $\bar{u} = \bar{s}$.

Vi sono alcune ambiguità sull'applicazione di questa tecnica: poichè sia A che \bar{s} sono incognite, e ogni scalare che moltiplica una componente può essere cancellato dal suo reciproco, non è possibile determinare le energie delle componenti indipendenti.

L'algoritmo di scomposizione in componenti indipendenti può essere utilizzato nel rispetto di alcune ipotesi sul segnale:

- le sorgenti sono assunte essere statisticamente indipendenti, ovvero la loro densità di probabilità congiunta $p^{x,y}$ può essere fattorizzata nel prodotto delle rispettive densità marginali p^x, p^y . Si ha quindi indipendenza statistica tra n variabili aleatorie se vale

$$p_{x_1, \dots, x_n}(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i);$$
- le sorgenti non possono avere una distribuzione gaussiana;
- il numero delle sorgenti è uguale al numero delle osservazioni.

In figura 3.6 è riportato un semplice esempio del funzionamento dell'algoritmo applicato su segnali generati sinteticamente in ambiente MatLab ed il relativo codice implementativo.

```

% Generazione Segnali

sr = 441;
fr = 8;
dur = 1;
N = dur * sr;
ts = 1/sr;

sine = .9*sin((0:N-1)*2*pi*fr*ts);
noise = rand(1,N);
sin2 = .9*sin((0:N-1)*2*pi*fr*2*ts);

% Combinazione lineare dei segnali
for n=1:length(sine)
    comb1(n) = 0.8 * sine(n) + (0.4*noise(n)) + (0.3*sin2(n));
    comb2(n) = 0.5 * sine(n) + (0.5*noise(n)) + (0.5*sin2(n));
    comb3(n) = 0.6 * sine(n) + (0.6*noise(n)) + (0.7*sin2(n));
end

% APPLICAZIONE ICA
sig = [comb1;comb2;comb3];
[weights,sphere] = runica(sig, 'extended', 1);
IC = sphere * sig;
IC = weights * IC;

% plot dei segnali
figure;

subplot(331),plot(sine),title('Segnali Sorgenti');
subplot(334),plot(noise);
subplot(337),plot(sin2);

subplot(332),plot(comb1),title('Combinazioni lineari delle sorgenti');
subplot(335),plot(comb2);
subplot(338),plot(comb3);

subplot(333),plot(IC(1,:)),title('Componenti ICA');
subplot(336),plot(IC(2,:));
subplot(339),plot(IC(3,:));

```

Ai segnali EEG è stato applicato lo stesso algoritmo dell'esempio, denominato RunICA [42]. La ricerca e l'eliminazione delle componenti indipendenti relative agli artefatti è stata eseguita manualmente osservando gli andamenti dei tracciati delle componenti. Le componenti contenenti gli artefatti sono state poste a zero e la ricombinazione lineare per ricavare i segnali EEG puliti è stata eseguita moltiplicando le componenti rimanenti per l'inverso della matrice di pesi fornito dall'algoritmo.

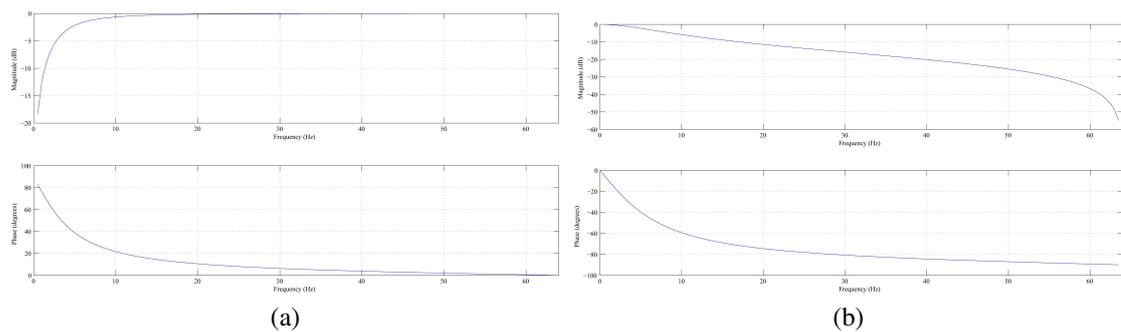


Figura 3.7: Risposte in frequenza e fase dei filtri passa alto (a) e passa basso (b) applicati ai segnali EEG

In questa fase sono stati scartati 5 soggetti per cui non è stato possibile identificare con chiarezza le componenti associabili agli artefatti.

Filtraggio passa-banda con filtro di Čebyšëv

Successivamente alla rimozione degli artefatti nei segnali EEG a disposizione, i sei canali di ogni soggetto sono stati filtrati con due filtri posti in serie, progettati e realizzati con il toolbox MatLab *fdatool*, per conservare le frequenze dei ritmi μ e β .

Il primo filtro è un passa alto con frequenza di taglio posta a 11 Hz mentre il secondo è un filtro passa basso con frequenza di taglio posta a 35 Hz. La figura 3.7 mostra le risposte in frequenza e fase di entrambi i filtri.

3.3.2 Scelta delle feature

In questo elaborato non viene effettuata una classificazione automatica degli stati emotivi dei diversi soggetti, piuttosto si vuole osservare qualitativamente e quantitativamente l'andamento dei ritmi μ e β nei diversi set sperimentali, per definire in qualche misura come la stimolazione mediante correlazione e scorrelazione emotiva influisca sull'attivazione cerebrale. L'osservazione di questi ritmi può fornire uno strumento di indagine per quantificare l'attivazione causata da uno stimolo proprio in virtù del fatto che la presenza di questi ritmi indica una maggiore attività corticale (vedere la descrizione dei ritmi cerebrali nella sezione 1.2.1).

La procedura di preprocessing applicata fornisce segnali il cui contenuto frequenziale è ristretto ai soli ritmi di interesse.

Per l'estrazione delle features è possibile implementare differenti metodi: Murugappan et al. classificano le emozioni umane basandosi su feature statistiche quali la deviazione standard, la misura dell'ampiezza del segnale EEG, l'entropia e la varianza ottenendo risultati con un grado di accuratezza fino al 90% [13]. Tseng et al. utilizzano un indice calcolato mediante combinazione lineare delle distanze di Mahalanobis tra differenti ritmi per identificare variazioni dello stato cognitivo un soggetto [9]. Huang et al. propongono un algoritmo di estrazione delle feature chiamato *asymmetrical spatial filter* il quale massimizza la differenza nella varianza di segnali EEG filtrati tra i due emisferi corticali [18]. Petrantonakis et al. propongono un algoritmo basato sull'asimmetria spaziale, combinata con l'approccio *multidimensional directed information*, per l'identificazione e la segmentazione di segnali EEG rilevanti ai fine della

classificazione emotiva [41]. Ryu et al. utilizzano una rete neurale la cui regola di classificazione si basa sulla pendenza di un indice di asimmetria calcolato come funzione nel tempo nei ritmi α e β [14]. Valenza et al., seppur non applicati direttamente a segnali EEG ma ad altri segnali fisiologici, propongono una serie di feature suddivise in standard e metodi non lineari: le prime raggruppano le feature calcolabili statisticamente nei diversi domini (temporale, frequenziale e tempo-frequenza), le seconde usano metodi non lineari per analizzare, ad esempio, il caos deterministico e la ricorrenza delle condizioni di un sistema (Recurrence Plot) [30]. Bin Yunus et al. comparano l'efficienza di diverse tecniche di classificazione emotiva, per valutare quali metodi di rimozione del rumore siano più efficaci, utilizzando l'analisi tempo-frequenza nei diversi ritmi per determinare l'effetto di uno stimolo emotivo nei diversi ERP [22].

Consigliati dai metodi e i risultati riscontrati in letteratura, sono state estratte le seguenti feature:

- massimi, minimi, media aritmetica e RMS dei valori di ogni canale, per ogni ERP;
- media dei valori RMS dei 6 canali per ogni stimolazione, sia $M(x)$ la media del vettore $x = \{x_1, \dots, x_6\}$ contenente i sei valori RMS di ogni canale EEG:
 - $img = M(RMS(immagine))$;
 - $imgAc = M(RMS(immagine + a.correlato))$;
 - $imgAs = M(RMS(immagine + a.scorrelato))$;
- le differenze delle medie delle diverse stimolazioni:
 - $difAcImg = imgAc - img$;
 - $difAsImg = imgAs - img$;
 - $difAcAs = imgAc - imgAs$;
- media delle differenza precedenti;
- Recurrence Plot delle differenze $difAcImg$, $difAsImg$ tra tutti i soggetti;
- Constant Q transform delle medie img , $imgAc$, $imgAs$ nei sette differenti set di stimolazione.

Recurrence Plot

Recurrence Plot (RP) è un grafico che mostra tutte le volte in cui uno stato di un sistema dinamico ricorre nel tempo. Eckmann et al. hanno introdotto uno strumento che può visualizzare la ricorrenza di uno stato x_i nello spazio delle fasi [43]. Questo strumento permette di osservare traiettorie nello spazio delle fasi m-dimensionali attraverso una rappresentazione delle sue ricorrenze in uno spazio bidimensionale. Quando uno stato al tempo i ricorre anche al tempo j , l'elemento (i, j) di una matrice quadrata N^2 è settato ad 1, o settato a 0 altrimenti. La rappresentazione grafica è chiamato RP, cui può essere espresso matematicamente come:

$$R_{i,j} = \Theta(\epsilon_i - \|x_i - x_j\|),$$

dove $x_i \in \mathbb{R}^m$, $i, j = 1, \dots, N$; N è il numero degli stati considerati x_i , ε_i è una soglia di distanza, $\|\cdot\|$ è una normalizzazione e $\Theta(\cdot)$ è la funzione Heaviside definita come segue:

$$H(z) = \begin{cases} 1, & \text{se } z \geq 0 \\ 0, & \text{se } z < 0. \end{cases}$$

Constant Q transform

In matematica e nell'elaborazione dei segnali, la Constant Q transform (QT) trasforma una serie di dati dal dominio del tempo al dominio delle frequenze, ed è relazionata alla trasformata di Fourier [44]. A differenza di quest'ultima, sia nelle sue varianti continue che discrete, in cui le componenti in frequenza sono calcolate con una costante spaziale e una risoluzione frequenziale fissa, la QT mantiene costante il rapporto tra la frequenza centrale e la risoluzione frequenziale. La QT può essere pensata come una serie di filtri logicamente spaziati, la cui ampiezza frequenziale del k -esimo filtro è un certo multiplo di quella del filtro precedente:

$$\delta f_k = \left(2^{\frac{1}{n}}\right)^k * \delta f_{min},$$

dove n è il numero di filtri per ottava, f_{min} è la frequenza centrale del primo filtro utilizzato.

Il "fattore di qualità" Q è definito come segue:

$$Q = \frac{f_k}{\delta f_k},$$

mentre la finestra del k -esimo filtro è definita come:

$$N[k] = \left(\frac{f_s}{\delta f_k}\right) = \frac{S_r * Q}{f_k}$$

e la lunghezza della finestra è quindi funzione del numero del filtro. Per compensare un decremento della potenza sulle alte frequenze si normalizza per $N[k]$. Qualsiasi funzione di finestra sarà funzione della lunghezza della finestra.

In conclusione, la QT è calcolata come segue:

$$X[k] = \frac{1}{N[k]} \sum_{n=0}^{N[k]-1} W[k, n] x[k] e^{\frac{-j2\pi Qn}{N[k]}}.$$

Capitolo 4

Analisi dei risultati

Nel corso del capitolo vengono illustrati i risultati ottenuti mediante i Recurrence Plot delle differenze di attivazione media dei soggetti nelle due tipologie di stimolo (immagine con audio correlato e immagine con audio scorrelato) al variare della soglia di distanza ϵ ; successivamente sono presentati gli spettrogrammi ottenuti mediante QT delle attivazioni medie nei diversi set sperimentali.

4.1 Recurrence Plot

Considerando le risposte agli stimoli come l'evoluzione nel tempo di un sistema dinamico, la rappresentazione delle ricorrenze può essere utilizzata per osservare e quantificare l'omogeneità delle reazioni tra i diversi soggetti. Per eseguire questa quantificazione sono state create sette serie (una per ogni set) in modo che ogni serie rappresenti l'insieme delle risposte di tutti i soggetti per ogni set sperimentale:

$$Set(i) = \begin{cases} img_i = \{sog1 + sog2 + \dots + sog26\} \\ imgAc_i = \{sog1 + sog2 + \dots + sog26\} \\ imgAs_i = \{sog1 + sog2 + \dots + sog26\} \end{cases}$$

con $i = \{1, \dots, 7\}$ e $sog(n) = \{x_1, \dots, x_{12}\}$ sono i valori corrispondenti alle medie ogni 64 sample dei segnali del n -esimo soggetto nell' i -esimo set sperimentale.

Di ogni serie è stato calcolato il grafico delle ricorrenze che offre un rappresentazione approssimata di quanto le risposte inter-soggetto siano simili tra loro, in figura 4.1 sono mostrate le serie e i grafici delle ricorrenze ottenuti con due differenti soglie di distanza ϵ , nel secondo set sperimentale.

Nella tabella 4.1 sono riportate le percentuali di punti di ricorrenza (Recurrence Rate, RR) tra le risposte, calcolate mediante l'equazione:

$$RR = 100 * \left(\frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \right),$$

dove N è il numero di punti nello spazio delle fasi e graficamente corrisponde al numero di punti sugli assi.

I risultati ottenuti da questa analisi riflettono l'eterogeneità delle risposte dei soggetti sottoposti alla sperimentazione, inoltre sono concordi con l'affermazione per cui i segnali EEG sono normalmente non stazionari e conseguentemente anche i valori statistici derivabili da essi [22].

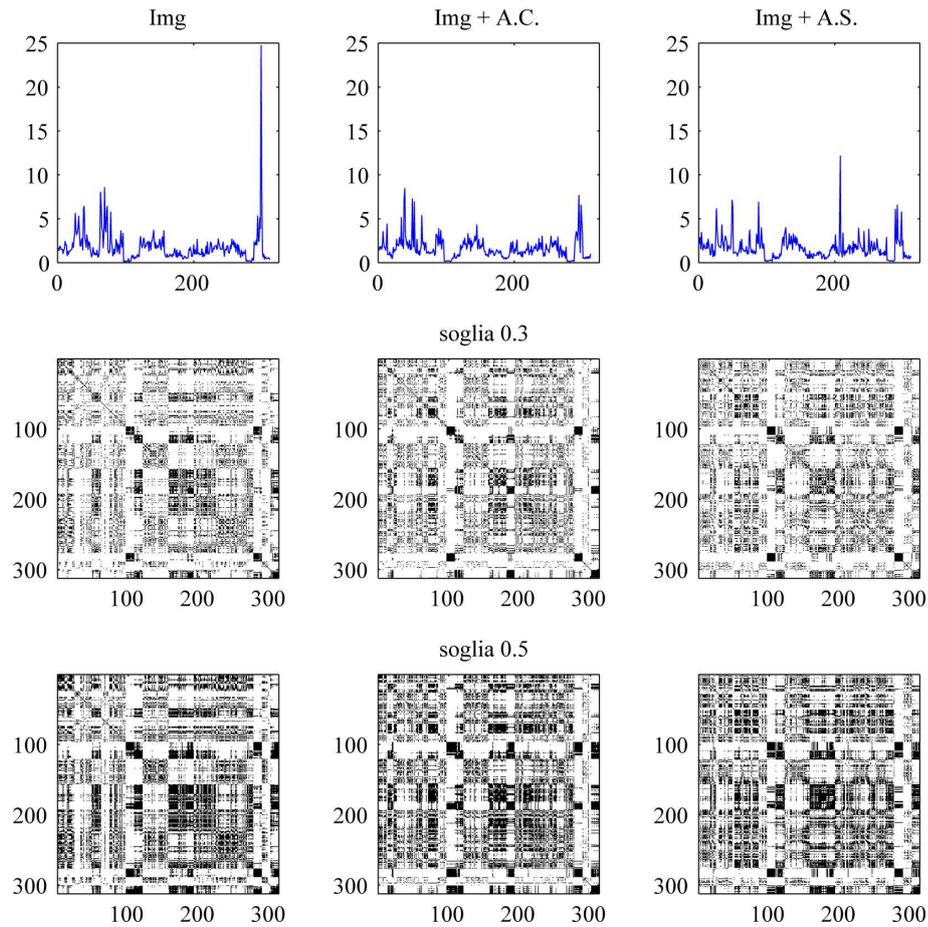


Figura 4.1: Serie e i grafici delle ricorrenze ottenuti con due differenti soglie di distanza ε nel secondo set sperimentale

	Set	Img	Img + A.C.	Img + A.S.
$\varepsilon = 0.3$	1	21.3819	20.6977	17.8357
	2	18.3987	19.0233	19.0417
	3	20.8600	21.6120	20.0999
	4	19.3438	18.1377	17.6775
	5	18.4377	19.4321	22.6866
	6	19.5924	21.1580	18.4788
	7	18.2220	18.8651	20.8107
$\varepsilon = 0.5$	1	34.4305	33.4874	29.1132
	2	29.8714	31.1904	31.0651
	3	33.8942	34.7798	32.4499
	4	31.4472	29.1831	28.5832
	5	30.0563	31.7041	36.5631
	6	31.7205	34.1757	30.1446
	7	29.9659	31.5288	33.8798

Tabella 4.1: Percentuali di omogeneità delle risposte medie dei soggetti nei diversi set sperimentali, calcolate come RR al variare della soglia di distanza ε

4.1.1 Limitatezza del segnale

Nelle tabelle 4.2a sono riportati i valori delle differenze di attivazione durante la stimolazione con audio correlato rispetto all'immagine e in tabella 4.2b le differenze tra stimolazione con audio scorrelato rispetto all'immagine.

Successivamente sono state create due serie costituite dalla concatenazione delle differenze di attivazione calcolate per ogni set $difAcImg$ e $difAsImg$; ciò rappresenta l'andamento delle variazioni di attivazione dei ritmi μ e β di tutti i soggetti nei differenti set. Può essere osservato che, a meno di alcune variazioni elevate ma isolate, le differenze dei valori RMS dei segnali sono limitate tra 0,5 e 1,5. La figura 4.2a mostra le differenze tra i valori medi durante la stimolazione con audio correlato, mentre in figura 4.2b sono mostrate le differenza tra i valori durante la stimolazione con audio scorrelato.

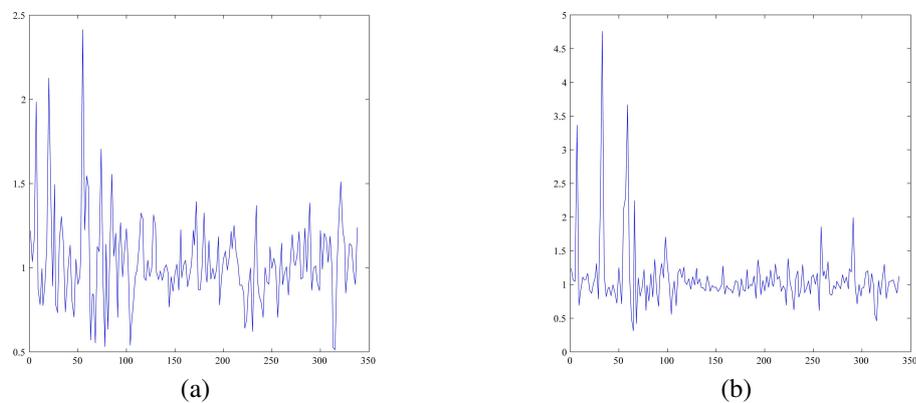


Figura 4.2: Plot delle serie contenenti i valori delle differenze $difAcImg$ (a) e $difAsImg$ (b)

Soggetto	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Soggetto	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7
1	1.2191	1.0358	1.1710	1.9841	0.8731	0.7811	0.9945	1	1.2395	1.0665	1.0459	3.3635	0.6928	0.9335	1.1104
2	0.7746	0.9016	1.1317	2.1263	1.4704	0.8918	1.4930	2	1.0748	1.0729	1.1645	0.9161	0.8706	1.0191	1.1220
3	0.7851	0.7331	1.1640	1.3038	1.1532	0.7383	0.8992	3	1.3102	0.7856	2.1804	4.7563	1.1413	0.8211	0.9427
4	1.0034	1.1327	0.8009	0.7073	1.0500	0.9024	0.9429	4	0.9599	0.8333	0.9958	0.8753	0.7236	1.2479	0.8888
5	1.0721	2.4136	1.2245	1.5452	1.4788	0.5712	0.8443	5	0.7142	2.1230	2.2975	3.6680	1.1586	0.5019	0.3138
6	0.8382	0.5536	1.1253	1.0956	1.7044	0.9142	0.5318	6	2.2448	0.4226	1.0970	0.8254	0.9582	1.2190	0.6165
7	1.1390	0.6346	1.1170	1.5548	1.0687	1.2048	0.7062	7	0.9864	0.7563	1.1572	0.8161	1.3690	0.8912	0.6757
8	1.0740	1.2673	0.9467	1.0931	1.2318	0.9920	0.5399	8	1.0912	1.3098	1.1000	1.7038	1.2851	1.0005	0.5584
9	0.6531	0.7768	0.9468	0.9865	1.1285	1.3240	1.2894	9	0.8349	1.0506	0.6845	1.1772	1.2291	1.1039	1.2559
10	0.9457	0.9243	1.0436	0.9481	0.9972	1.3124	1.2410	10	1.0409	1.0017	1.0838	0.9314	1.1134	0.9938	1.2401
11	0.9745	0.9296	0.9803	0.9248	0.9924	1.0173	0.9668	11	1.0089	1.0822	0.9514	0.9499	0.9060	1.1261	0.9894
12	0.7663	0.9453	0.8606	0.9666	1.0205	0.8666	1.2254	12	0.8980	0.9828	0.9582	0.9598	0.8971	0.9358	1.0104
13	0.9406	1.0174	1.0438	0.8870	0.9480	1.0195	1.2222	13	1.2738	0.8586	0.9814	0.9321	0.9277	0.8744	0.9883
14	1.1345	1.3931	0.8714	0.8678	1.0082	1.3247	0.9545	14	1.0566	1.0350	0.8191	1.0851	0.9189	0.9035	1.2197
15	0.9140	1.1603	0.9320	0.9968	0.9325	0.9958	1.1841	15	0.9362	0.9972	0.9838	1.1247	0.7933	1.3660	1.1371
16	0.7795	0.9383	1.0511	1.0982	0.9864	1.0632	1.2158	16	0.8485	1.0580	0.9050	1.1223	0.9575	1.2128	0.9988
17	1.1126	1.2514	1.0959	1.0233	0.8951	0.8969	0.8487	17	0.9733	1.3014	1.0786	1.1196	0.8998	0.9784	0.9233
18	0.6414	0.6891	0.9001	0.9980	0.6219	1.0663	1.3688	18	0.6948	1.3819	1.0252	0.9056	0.6243	1.0940	1.2020
19	0.9319	0.8307	0.7924	0.7052	1.0016	0.9143	0.9033	19	0.8124	0.9119	1.2933	0.8792	0.9471	1.0527	0.8592
20	1.1229	0.9967	1.0553	0.9999	0.7061	0.9546	1.1458	20	1.1079	1.1420	1.0033	1.1663	0.6153	1.8551	1.1250
21	0.8984	0.9660	1.0061	0.8378	1.0315	1.1959	1.0303	21	1.1984	1.0816	1.3358	0.8582	0.8426	0.9839	0.9367
22	1.0130	1.0730	1.2139	0.9328	0.9430	1.2340	0.9765	22	1.0477	0.9939	0.9358	1.1441	1.0217	1.1086	0.9350
23	1.1121	1.3844	0.8668	0.9961	1.0107	0.9068	0.8635	23	1.2329	1.1857	1.9916	0.9837	0.7125	1.0001	0.8271
24	1.2205	0.9916	1.2038	1.1697	1.0318	1.1861	0.9805	24	0.9426	0.9565	1.1785	1.2015	0.8710	1.1618	0.9743
25	0.5260	0.5129	0.9877	1.2994	1.5106	1.2232	1.1436	25	0.5722	0.4611	1.0542	0.8458	1.0829	1.2969	0.7925
26	0.8502	0.9880	1.1425	1.1312	0.9797	0.9010	1.2396	26	0.8859	1.0454	1.0487	1.0677	0.9731	0.8739	1.1199
MEAN	0.9401	1.0170	1.0260	1.1223	1.0683	1.0153	1.0289	MEAN	1.0380	1.0345	1.1673	1.3607	0.9436	1.0598	0.9524

(a)

(b)

Tabella 4.2: Valori delle differenze $difAcImg$ e $difAsImg$

4.2 Risultati Q-transform

Il calcolo della QT è stato effettuato sulle medie delle risposte dei soggetti. Il risultato di questa analisi in tempo-frequenza mostra l'andamento dei ritmi μ e β nel corso dell'esposizione allo stimolo, attraverso i diversi set sperimentali.

Di seguito sono riportati gli spettrogrammi calcolati per ogni gruppo di stimoli:

- QT1, QT2, QT3: questi set prevedono immagini la cui valenza e arousal diminuisce progressivamente, per cui l'immagine assume carattere sempre più negativo. Mediante la correlazione con audio avente gli stessi valori in termini di valenza e arousal si è registrata un'attenuazione dell'attivazione dei ritmi β ; mentre nella la fase di scorrelazione vi è un aumento di attività.
- QT4, QT5, QT6, QT7: in questi set la valenza e l'arousal delle immagini proiettate aumenta progressivamente, esse assumono perciò carattere sempre più positivo. Rafforzando la stimolazione con campioni audio aventi valenza e arousal correlati si registra un

sensibile aumento dell'attività nella bande prese in esame. La correlazione con stimoli aventi valori di valenza e arousal opposti invece non rivela un andamento costante in quanto alterna fasi in cui presenta aumenti di attività a fasi di diminuzione.

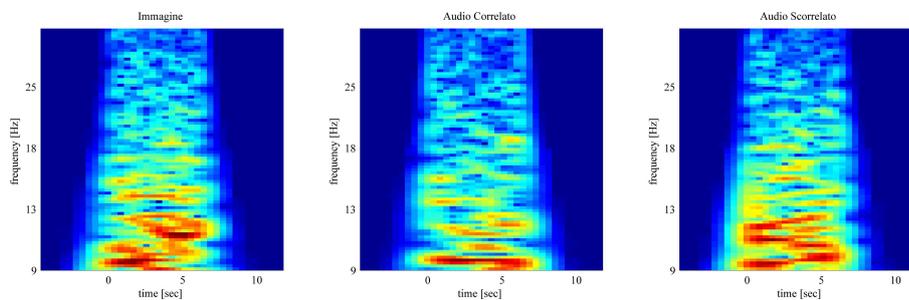


Figura 4.3: QT Set 1

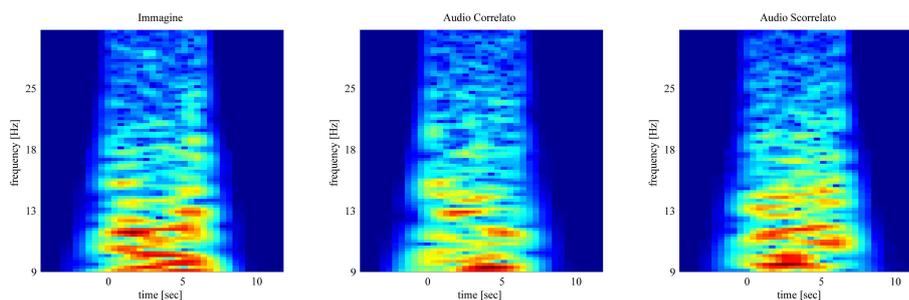


Figura 4.4: QT Set 2

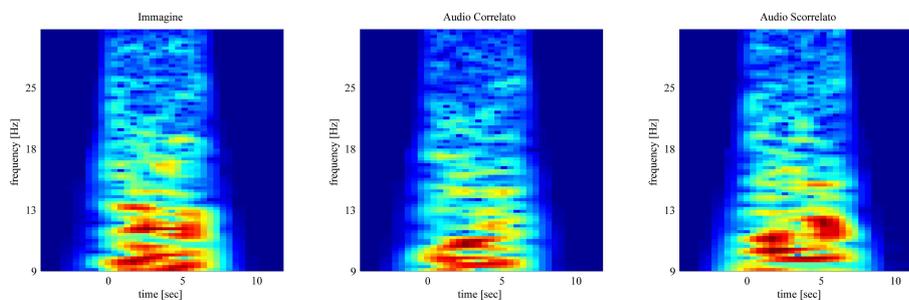


Figura 4.5: QT Set 3

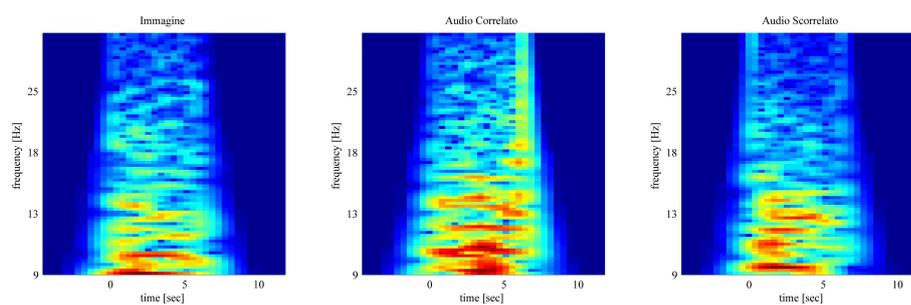


Figura 4.6: QT Set 4

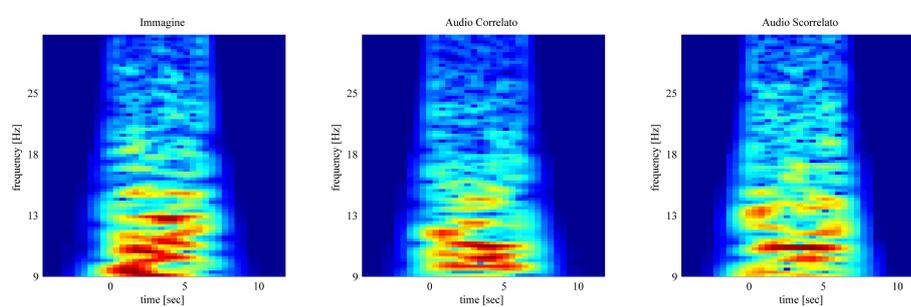


Figura 4.7: QT Set 5

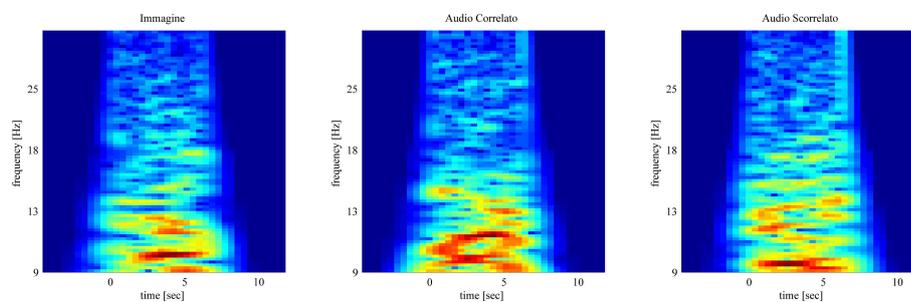


Figura 4.8: QT Set 6

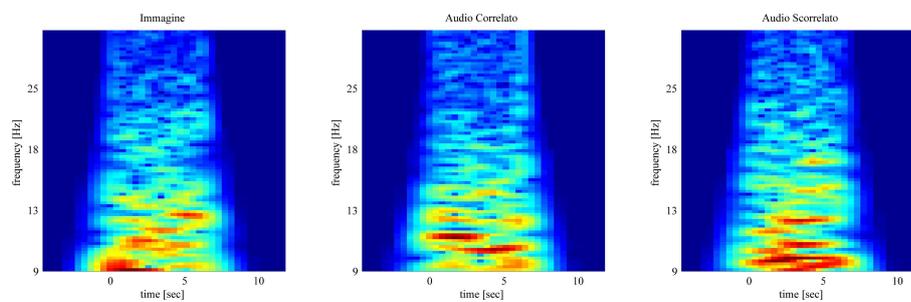


Figura 4.9: QT Set 7

Capitolo 5

Conclusioni e sviluppi futuri

Il risultato principale a cui siamo giunti è riassunto schematicamente in tabella 5.1 e mostra come l'attività dei ritmi ricercati sia variabile nelle differenti combinazioni di

Considerando che i ritmi β sono generalmente associati ad uno stato di attivazione neurale, di coinvolgimento del soggetto (vedi sottosezione 1.2.1), si può osservare che durante la stimolazione con audio positivo l'attivazione dei ritmi β aumenta in maniera indipendente dal fatto che esso sia associato ad una immagine con carattere emotivo negativo o positivo.

Questo potrebbe rivelare come, rispetto all'immagine, l'audio assuma un peso preponderante nella percezione della componente emotiva. Questo risultato è in linea con l'aspettativa iniziale per cui un audio positivo generi un'attivazione anche se contrapposto ad immagini negative, di forte impatto, come spesso è utilizzato in tecniche cinematografiche e pubblicitarie.

Le stimolazioni mediante audio negativo associato ad immagini negative, forniscono risultati che mostrano una diminuzione dell'attività neurale; ciò potrebbe essere dovuto al fatto che un'immagine negativa generi una forte attività, e di conseguenza uno stato d'allerta che l'audio non riesca ad alterare. Oltretutto gli stimoli visivi utilizzati sono, per loro natura, statici e questo rende difficile la creazione di quelle dinamiche di tensione/rilassamento tipiche delle clip video.

Nell'ultima tipologia di stimolazione, che utilizza audio negativo contrapposto ad immagini positive, ci si aspettava a priori un forte scollamento e un'attivazione tanto maggiore quanto maggiore è la distanza nello spazio Valenza/Arousal dei due stimoli. Questo non è accaduto e si sono registrati invece andamenti significativi sia in aumento che in diminuzione, in maniera totalmente dipendente dal soggetto, in accordo con risultati ottenuti tramite i grafici delle ricorrenze (vedi tabella 4.1) che forniscono un'idea dell'eterogeneità delle risposte dei soggetti. La differenza nelle risposte inter-soggetto potrebbe trovare spiegazione nell'osservazione della distribuzione spaziale dei campioni audio/visivi nei rispettivi spazi affettivi (vedi figura 3.1): si osserva che la distribuzione statistica dei campioni positivi è più sparsa di quella dei campioni negativi e quindi che uno di quest'ultimi genera una attivazione a differenza dei primi che sono percepiti talvolta come attivanti, talvolta meno.

Nell'ultima tipologia di stimolazione, lo scollamento dell'attivazione dei ritmi β dalla posizione "zero", fornita dalla proiezione della sola immagine, è ambigua ed è un'eccezione sulla quale potrebbe concentrarsi un possibile sviluppo del lavoro; le motivazioni di tale andamento

	Immagine Negativa	Immagine Positiva
Audio Positivo	Aumento di attività	Aumento di attività
Audio Negativo	Diminuzione di attività	Non vi è andamento costante

Tabella 5.1: Andamento nei ritmi μ e β nelle diverse combinazioni di stimolazione

non possono essere spiegate a priori e potrebbero dipendere da fattori diversi, quali ad esempio lo stato dell'interfaccia utilizzata o dalla fase di preprocessing e pulizia del segnale.

Di seguito sono riportate alcune idee che potrebbero permettere di affinare i risultati appena descritti:

- scegliere ed utilizzare una interfaccia più performante e aumentare il numero dei canali campionati; ciò fornisce una risoluzione spaziale maggiore indispensabile per integrare la ricerca con la teoria dell'asimmetria spaziale (l'area frontale sinistra è implicata nell'esperienza di emozioni positive mentre quella destra è associata ad emozioni negative [41]);
- cercare di ridurre l'eterogeneità delle risposte effettuando una selezione dei soggetti partecipanti, selezionandoli per classi di età e posizione sociale;
- utilizzare tutte le classi di valenza e arousal all'interno dei database, selezionando il materiale emotigeno con di una valutazione statistica da parte di una popolazione di soggetti non partecipante alla raccolta dati;
- utilizzare tecniche di machine learning (ad esempio, reti neurali o Support Vector Machine) per una classificazione più accurata dei segnali ottenuti.

I possibili sviluppi futuri di questo lavoro possono iniziare dalla definizione di un nuovo spazio affettivo, basato sul modello dimensionale valenza/arousal, che tenga in considerazione le variazioni di percezione dovute alla combinazione dei singoli stimoli, cercando quindi regole matematiche di predizione dello spostamento in tale spazio. Parallelamente si può procedere ad una generalizzazione dello studio a campioni audio e immagini non catalogati e proporre, successivamente, le regole per una clusterizzazione. Un ulteriore sviluppo potrebbe riguardare la dinamicità degli stimoli nel tempo, mediante l'utilizzo di clip video in sostituzione degli stimoli statici proposti in questo elaborato. Questo tipo di approccio è più facilmente riconducibile agli stimoli che l'essere umano vive tutti i giorni in maniera continua e dinamica. La durata dei segnali cerebrali campionati è in funzione della durata delle clip video che per far emergere la componente emotiva possono impiegare anche diversi minuti. Le tecniche presentate in questo lavoro potrebbero non essere efficienti in quanto la media di segnali così lunghi influisce drasticamente sulla risoluzione spaziale e frequenziale delle analisi. Una soluzione al problema potrebbe essere l'utilizzo di tecniche di estrazione delle feature mediante finestrazione del segnale e classificazione di esse con algoritmi di machine learning precedute da una accurata fase di training.

Ringraziamenti

Bibliografia

- [1] J. A. Russell, “A circumplex model of affect,” *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.
- [2] R. W. Picard, *Affective computing*. MIT press, 2000.
- [3] R. W. Homan, J. Herman, and P. Purdy, “Cerebral location of international 10–20 system electrode placement,” *Electroencephalography and clinical neurophysiology*, vol. 66, no. 4, pp. 376–382, 1987.
- [4] V. Mezzogiorno and A. Mezzogiorno, *Compendio di anatomia umana*. Piccin, 1994.
- [5] G. E. Caimmi, Philippe P., *Dispense di fisiologia*. Aracne, 2009.
- [6] G. G. Molina, T. Tsoneva, and A. Nijholt, “Emotional brain-computer interfaces,” pp. 1–9, 2009.
- [7] K. Crowley, A. Sliney, I. Pitt, and D. Murphy, “Evaluating a brain-computer interface to categorise human emotional response,” pp. 276–278, 2010.
- [8] F. Galán, M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, and J. d. R. Millán, “A brain-actuated wheelchair: asynchronous and non-invasive brain-computer interfaces for continuous control of robots,” *Clinical Neurophysiology*, vol. 119, no. 9, pp. 2159–2169, 2008.
- [9] K. C. Tseng, Y.-T. Wang, B.-S. Lin, and P. H. Hsieh, “Brain computer interface-based multimedia controller,” pp. 277–280, 2012.
- [10] P. R. Kennedy and R. A. Bakay, “Restoration of neural output from a paralyzed patient by a direct brain connection,” *Neuroreport*, vol. 9, no. 8, pp. 1707–1711, 1998.
- [11] J. Naumann, *Search for Paradise: A Patient’s Account of the Artificial Vision Experiment*. Xlibris Corporation, 2012.
- [12] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “Bci2000: a general-purpose brain-computer interface (bci) system,” *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [13] M. Murugappan, R. Nagarajan, and S. Yaacob, “Appraising human emotions using time frequency analysis based EEG alpha band features,” pp. 70–75, 2009.
- [14] C. S. Ryu, Y. Song, S. H. Kim, I. Yi, J.-E. Kim, and J.-H. Sohn, “A time-frequency analysis of the EEG evoked by negative and positive visual stimuli,” vol. 4, pp. 2012–2015, 1998.

- [15] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 11, no. 2, pp. 141–144, 2003.
- [16] Z. Zhu and K. He, "A novel approach of emotion recognition based on selective ensemble," vol. 1, pp. 695–698, 2008.
- [17] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution eeg and erp measurements," *Clinical Neurophysiology*, vol. 112, no. 4, pp. 713–719, 2001.
- [18] D. Huang, H. Zhang, K. Ang, C. Guan, Y. Pan, C. Wang, and J. Yu, "Fast emotion detection from EEG using asymmetric spatial filtering," pp. 589–592, 2012.
- [19] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based human emotion recognition and visualization," pp. 262–269, 2010.
- [20] A. G. Correa, E. Laciari, H. Patiño, and M. Valentinuzzi, "Artifact removal from eeg signals using adaptive filters in cascade," in *Journal of Physics: Conference Series*, vol. 90, no. 1. IOP Publishing, 2007, p. 012081.
- [21] D. Szibbo, A. Luo, and T. J. Sullivan, "Removal of blink artifacts in single channel EEG," pp. 3511–3514, 2012.
- [22] J. bin Yunus *et al.*, "The effect of noise removing on emotional classification," in *Computer & Information Science (ICCIS), 2012 International Conference on*, vol. 1. IEEE, 2012, pp. 485–489.
- [23] T. Dalgleish, "The emotional brain," *Nature Reviews Neuroscience*, vol. 5, no. 7, pp. 583–589, 2004.
- [24] K. L. Phan, T. D. Wager, S. F. Taylor, I. Liberzon *et al.*, "Functional neuroimaging studies of human emotions," *CNS spectrums*, vol. 9, pp. 258–266, 2004.
- [25] W. James, "What is an emotion?" *Mind*, no. 34, pp. 188–205, 1884.
- [26] J. E. LeDoux, "Emotion circuits in the brain," *Annual review of neuroscience*, vol. 23, no. 1, pp. 155–184, 2000.
- [27] S. Schachter, J. E. Singer *et al.*, "Cognitive, social, and physiological determinants of emotional state," *Psychological review*, vol. 69, no. 5, pp. 379–399, 1962.
- [28] M. B. Arnold, "Emotion and personality," 1960.
- [29] P. Ekman, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, pp. 45–60, 1999.
- [30] G. Valenza, A. Lanata, and E. P. Scilingo, "The role of nonlinear dynamics in affective valence and arousal recognition," *Affective Computing, IEEE Transactions On*, vol. 3, no. 2, pp. 237–249, 2012.
- [31] F. Battisti, *Modificazioni dello spazio affettivo nel ciclo di vita*, 2012.

- [32] J. Tao and T. Tan, "Affective computing: A review," in *Affective Computing and Intelligent Interaction*. Springer, 2005, pp. 981–995.
- [33] R. Picard, "Affective computing for hci," *Human-Computer Interaction: Ergonomics and User Interfaces*, vol. 1, pp. 829–833, 1999.
- [34] C. Kleine-Cosack, "Recognition and simulation of emotions," 2006.
- [35] F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals using wireless sensors for presence technologies," *Cognition, Technology & Work*, vol. 6, no. 1, pp. 4–14, 2004.
- [36] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 12, pp. 2067–2083, 2008.
- [37] C. Maaoui and A. Pruski, "Emotion recognition through physiological signals for human-machine communication," *Cutting Edge Robotics*, 2010.
- [38] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (iaps): Affective ratings of pictures and instruction manual," *Technical Report A-8*, 2008.
- [39] M. M. Bradley and P. J. Lang, "The international affective digitized sounds (iads-2): Affective ratings of sounds and instruction manual," *University of Florida, Gainesville, FL, Tech. Rep. B-3*, 2007.
- [40] E. Emotiv, "www.emotiv.com," 2013.
- [41] P. C. Petrantonakis and L. J. Hadjileontiadis, "Adaptive emotional information retrieval from EEG signals in the time-frequency domain," *Signal Processing, IEEE Transactions on*, vol. 60, no. 5, pp. 2604–2616, 2012.
- [42] S. Makeig, A. J. Bell, T.-P. Jung, T. J. Sejnowski *et al.*, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems*, pp. 145–151, 1996.
- [43] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhys. Lett*, vol. 4, no. 9, pp. 973–977, 1987.
- [44] J. C. Brown, "Calculation of a constant q spectral transform," *The Journal of the Acoustical Society of America*, vol. 89, p. 425, 1991.

Appendice A

Codice Implementativo

A.1 Ricerca e selezione del materiale

A.1.1 importValoriDb.m

Questa parte di codice suddivide la matrice contenente i dati relativi ai DB, copiando il valore all'interno della variabile coerente con la classe di valenza appartenente per ogni valore: se la valenza è compresa tra *classe* e *classe + 1* allora inserisce il valore nella variabile riferita alla classe di appartenenza e aumenta il contatore. Se *n* è uguale alla lunghezza dell'array, allora è stato controllato l'ultimo elemento della matrice, riporta *count = 1* e passa alla classe successiva.

```
count = 1;
for classe = 1:8
    for n=1:length(fetch_dati_ordinati_IAPS)
        if (fetch_dati_ordinati_IAPS(n,2) >= classe && ...
            fetch_dati_ordinati_IAPS(n,2) < classe + 1)

            IAPS.(var_name(classe,:))(count,[1, 2, 3]) ...
                = fetch_dati_ordinati_IAPS(n,[1, 2, 3]);
            count = count + 1;
        end
        if (n == length(fetch_dati_ordinati_IAPS))
            count = 1;
        end
    end
end
for classe = 1:8
    for n=1:length(fetch_dati_ordinati_IADS)
        if (fetch_dati_ordinati_IADS(n,2) >= classe && ...
            fetch_dati_ordinati_IADS(n,2) < classe + 1)

            IADS.(var_name(classe,:))(count,[1, 2, 3]) ...
                = fetch_dati_ordinati_IADS(n,[1, 2, 3]);
            count = count + 1;
        end
        if (n == length(fetch_dati_ordinati_IADS))
            count = 1;
        end
    end
end
```

```

        end
    end
end

```

A.1.2 SelezioneAudioImmaginiExp.m

Script per la selezione dei campioni audio e delle immagini: *id_audio* contiene gli interi identificatori dei campioni audio scelti inizialmente le altre linee di codice sono chiamate a funzione per l'estrazione delle immagini minimizzanti e gli audio massimizzanti a partire da queste ultime.

```

% Array contenente gli ID dei campioni audio selezionati
id_audio = [278 286 279 255 719 600 709 355 ...
            365 215 360 110 366 717 311 352 817];

% Recupero dei valori di valenza e arousal dei campioni selezionati
valori_audio = getValueFromID(id_audio,'audio');

immagini_minimiz= getImgMinFromAudio(valori_audio);
valori_immagini.dist_euclidea= immagini_minimiz.dist_euclidea(:,[4 5 6]);

audio_massimiz= getAudioMaxFromImg(valori_immagini.dist_euclidea);

```

A.1.3 getValueFromID.m

Questa funzione restituisce una variabile array in cui sono riportati i valori di valenza e arousal degli audio o delle immagini passate come argomento.

Input: *arrayID* = array contenente gli ID dei campioni di cui si vogliono i valori di valenza e arousal. *media* = specifica se audio o immagini.

Output: *valori* = matrice contenente *n* righe, con *n* = lunghezza dell'array in input e tre colonne: | ID | valenza | arousal.

```

function [ valori ] = getValueFromID( arrayID, media )

load('variabili_per_funzioni.mat')
valori = [];
if(strcmpi(media, 'audio'))
    for n = 1:length(arrayID)
        [r,c] = find (fetch_dati_ordinati_IADS==arrayID(n));
        valori = vertcat(valori, fetch_dati_ordinati_IADS(r,:));
    end
elseif (strcmpi(media, 'immagini'))
    for n = 1:length(arrayID)
        [r,c] = find (fetch_dati_ordinati_IAPS==arrayID(n));
        valori = vertcat(valori, fetch_dati_ordinati_IAPS(r,:));
    end
else
    error('Secondo parametro della funzione errato. ...
        Il parametro può essere "audio" o "immagini"');
end

```

```
end
```

```
end
```

A.1.4 getImgMinFromAudio.m

Questa funzione cerca e seleziona le immagini minimizzanti a partire dai valori dei campioni audio forniti come input.

Input: *audiolist* = variabile contenente i valori degli audio di cui si vuole cercare le relative immagini minimizzanti. La variabile deve essere composta di una lista di n righe e tre colonne:

| numero id audio | valenza audio | arousal audio |

es: 120|4.52|6.35

Output: *img_min* = variabile struttura contenente i valori delle immagini minimizzanti i campioni audio in input. La struttura contiene due campi:

- EUCLIDEA: minimo calcolato con la distanza euclidea
- MINKOVSKI: minimo calcolato con la distanza di Minkovski

```
function [ img_Min ] = getImgMinFromAudio( audioList )
    load('variabili_per_funzioni.mat')
    % recupero dati valenza/arousal IAPS
    val_aro_IAPS = fetch_dati_ordinati_IAPS(:,[2 3]);

    % DISTANZA EUCLIDEA %
    % calcolo della distanza
    [distance_min.euclidea, index_min.euclidea] = ...
        pdist2(val_aro_IAPS, audioList(:,[2 3]), ...
            'euclidean', 'Smallest', 1);

    % %
    % creazione matrice contenente gli audio
    % con le rispettive immagini che distano la
    % minore distanza euclidea.
    % %
    for n = 1:length(audioList)
        img_Min.dist_euclidea(n,:) = ...
            [audioList(n,:), ...
            fetch_dati_ordinati_IAPS(index_min.euclidea(n),:)];
    end

    % DISTANZA MINKOWSKI
    % con p = 8

    [distance_min.minkowski, index_min.minkowski] = ...
        pdist2(val_aro_IAPS, ...
            audioList(:, [2 3]), 'minkowski', 8 , 'Smallest', 1);

    for n = 1:length(audioList)
        img_Min.dist_minkowski(n,:) = ...
            [audioList(n,:), ...
```

```

        fetch_dati_ordinati_IAPS(index_min.minkowski(n),:)]];
    end
end

```

A.1.5 getAudioMaxFromImg.m

Questa funzione cerca e seleziona l'audio massimizzante la distanza da una serie di immagini. La funzione controlla la valenza e l'arousal di un'immagine e cerca il campione audio situato alla massima distanza euclidea. La ricerca viene effettuata calcolando le distanze dall'immagine ad ogni elemento contenuto nell'insieme degli audio che abbiano valori di valenza e arousal opposti a quello dell'immagine. Nel caso in cui non fosse presente alcun campione audio nello specifico intervallo di valori, viene ripetuta la ricerca nella classe di arousal superiore, mentre resta invariata la classe di valenza.

Input: *imglist* = variabile contenente i valori delle immagini di cui si vuole cercare i relativi audio massimizzanti. La variabile deve essere composta di una lista di n righe e tre colonne:

| numero id immagine | valenza immagine | arousal immagine |

es: 1000 | 4.52 | 6.35

Output: *audio_max* = variabile struttura contenente i valori dei campioni audio massimizzanti le immagini in input. La struttura contiene due campi:

- EUCLIDEA: massimo calcolato con la distanza euclidea

- MINKOVSKI: massimo calcolato con la distanza di Minkovski (non riportata nel codice)

```

function [audio_max] = getAudioMaxFromImg(imgList)
    load variabili_per_funzioni.mat;
    % %
    % Script per la ricerca dell'audio massimizzante nella sola casella
    % reciproca ai valori di valenza e arousal dell'immagine minimizzante
    % %
    audio_max.euclidea = [];
    for n=1:length(imgList)
        % ricavo le classi di valenza e arousal dell'immagine presa in
        % considerazione
        val_imm = floor(imgList(n,2));
        aro_imm = floor(imgList(n,3));
        % calcolo dei valori di valenza e arousal in cui cercare l'audio
        % massimizzante
        search_val = 8 - val_imm;
        search_aro = 8 - aro_imm;

        while_exit = 0;
        set_audio = [];

        while (while_exit == 0)
            candidati = IADS.(strcat('v',num2str(search_val)));
            for i=1:length(candidati);
                % if (arousal audio == search_arousal)
                % then aggiorno set_audio
                if (floor(candidati(i,3)) == search_aro)

```

```

        set_audio = vertcat(set_audio, candidati(i,:));
    end
end
% fine del for -> incremento per passare alla classe di
% arousal successiva. Questo potrebbe rivelarsi inutile
% se set_dati contiene almeno un elemento
search_aro = search_aro + 1 ;
% se vi è più di un valore all'interno di set_audio allora
% while_exit = 1 e termina il ciclo while
if( size(set_audio, 1) > 5)
    set_audio = set_audio(1:5,:);
    while_exit = 1;
end
end
% fine del while -> set_audio contiene sicuramente
% almeno un campione audio.

% calcolo dell'indice indicante la posizione
% dell'audio massimizzante la
% distanza euclidea nella matrice candidati
[dist_guidata_MAX.euclidea, index_guidata_MAX.euclidea] ...
    = pdist2(set_audio(:,[2 3]),imgList(n,[2 3]), ...
        'euclidean','Largest',5);
nome = strcat('img ', int2str(imgList(n)));
audio_max.euclidea.(nome) = [];
for i = 1:length(set_audio)
    audio_max.euclidea.(nome) = ...
        vertcat(audio_max.euclidea.(nome),set_audio(i,:));
end
audio_max.euclidea.(nome) = ...
    horzcat(audio_max.euclidea.(nome), ...
        dist_guidata_MAX.euclidea);
end
end
end

```

A.2 Estrazione delle feature ed analisi dei dati

A.2.1 preprocessingFilt.m

Questo script esegue il preprocessing del segnale applicando un filtro passa-banda con banda passante tra 11-35 HZ (ritmi beta).

SignalClean viene filtrato con con due filtri Chebyshev in serie: - Low-Pass Filter: prende in ingresso il segnale e filtra con una frequenza di cut-off uguale a 35Hz; - Hi-Pass Filter: prende in ingresso il segnale filtrato dal filtro passa basso e filtra con frequenza di cut-off uguale ad 11Hz;

```

% Parametri di inizializzazione dei filtri
order = 1;
sampleRate = 128;

```

```

cutOffFreqLowPass = 35;
cutOffFreqHiPass = 11;
ripple_mag = 5;

% DEFINIZIONE PARAMETRI DEI FILTRI PASSA-BASSO e PASSA-ALTO
[chebyLP_B, chebyLP_A] = cheby2(order, ripple_mag, ...
                                cutOffFreqLowPass/(sampleRate*2), 'low');
[chebyHP_B, chebyHP_A] = cheby1(order, ripple_mag, ...
                                cutOffFreqHiPass/(sampleRate*2), 'high');

% FILTRAGGIO DEI SEGNALI
for sogCount = 1 : 31
    s = signalClean.(strcat('s',int2str(sogCount)));
    signalFiltered.(strcat('s',int2str(sogCount))) = [];
    for ch = 1 : 6
        sog = strcat('s',int2str(sogCount));
        signalToFilt = s(ch, :);
        chebyFiltSignLP = filter(chebyLP_B, chebyLP_A, signalToFilt);
        chebyFiltSignHP = filter(chebyHP_B, chebyHP_A, chebyFiltSignLP);
        signalFiltered.(sog) = vertcat (signalFiltered.(sog), chebyFiltSignHP);
    end
end
end

```

A.2.2 EEGtoERP.m

Script per la suddivisione in ERP del segnale EEG campionato.

```

% Definizione variabili comuni
sr = 128;
durataStimolo = sr * 6;
s = signalFiltered.s1(1,:);

% Vettori contenenti i valori di sample iniziale per ogni stimolo suddiviso
% nelle 3 tipologie di stimolo
stImg = (1024:2816:length(s) - 3*durataStimolo + 2*sr);
stImgAudCorrelato = (1920:2816:length(s) - 2*durataStimolo + sr);
stImgAudScorrelato = (2816:2816:length(s) - durataStimolo);

% SUDDIVISIONE %

for sogCount = 1 : 31
    sog = strcat('s', int2str(sogCount));
    indSt = 0;
    for setCount = 1:7
        set = strcat('set', int2str(setCount));
        indSt = indSt + 1;
        sampleInizialeImg = stImg(indSt);
        sampleInizialeCorr = stImgAudCorrelato(indSt);
        sampleInizialeScorr = stImgAudScorrelato(indSt);
        sampleFinaleImg = sampleInizialeImg + durataStimolo;
    end
end

```

```

sampleFinaleCorr = sampleInizialeCorr + durataStimolo;
sampleFinaleScorr = sampleInizialeScorr + durataStimolo;
tImg = sampleInizialeImg:sampleFinaleImg;
tCorr = sampleInizialeCorr:sampleFinaleCorr;
tScorr = sampleInizialeScorr:sampleFinaleScorr;
% Solo immagine
ERP.(sogg).(set).img(:, :) = signalFiltered.(sogg)(:, tImg);
% Immagine + audio correlato
ERP.(sogg).(set).corr(:, :) = signalFiltered.(sogg)(:, tCorr);
% Immagine + audio scorrelato
ERP.(sogg).(set).scorr(:, :) = signalFiltered.(sogg)(:, tScorr);
end
end

```

A.2.3 RP.m

Questo script calcola il Recurrence Rate (RR) delle risposte in ogni singolo stimolo considerando i segnali dei 26 soggetti, l'algoritmo esegue:

1. costruzione di 3 serie composte dalla successione, sottocampionata a 64 sample, delle risposte medie di ogni soggetto in ogni di stimolazione. Le strutture elencate di seguito contengono le serie suddivise per set di stimolazione:

- dataImg : stim. con sola immagine, così composto:
 - |
 - | - st1 -< sog1 + sog2 + ... + sog26
 - | - st2 -< sog1 + sog2 + ... + sog26
 - | - ...
 - | - st7 -< sog1 + sog2 + ... + sog26
 - dataAcImg : stim. con immagine + Audio Correlato
 - dataAsImg : stim. con immagine + Audio Scorrelato
2. calcolo dei rispettivi RP per ogni set sperimentale.
3. calcolo del RR
- RRIImg : RR stim. con sola immagine
- 7 valori indicanti la ricorrenza in: Set1 | Set2 | ... | Set7
- RRAcImg : RR stim. con immagine + Audio Correlato
 - RRAsImg : RR stim. con immagine + Audio Scorrelato

Il passo di sottocampionamento può essere modificato cambiando il valore della variabile passoSamp. La soglia per il calcolo delle ricorrenze può essere modificata cambiando il valore della variabile soglia.

```

passoSamp = 64;
t = (768 / passoSamp) - 1;

for stCount = 1:7
    stimolo = strcat('st', int2str(stCount));

    dataImg.(stimolo) = [];
    dataAcImg.(stimolo) = [];
    dataAsImg.(stimolo) = [];

```

```

for sogCount = 1:26
    for sec = 0:t
        sample = sec * passoSamp;
        sampIni = 1 + sample;
        sampEnd = passoSamp + sample;
        m1(1,sec+1) = rms(ERPStImg.(stimolo)(sogCount,(sampIni: sampEnd)));
        m2(1,sec+1) = rms(ERPStImgCo.(stimolo)(sogCount,(sampIni: sampEnd)));
        m3(1,sec+1) = rms(ERPStImgSc.(stimolo)(sogCount,(sampIni: sampEnd)));
    end
    dataImg.(stimolo) = horzcat(dataImg.(stimolo),m1);
    dataAcImg.(stimolo) = horzcat(dataAcImg.(stimolo),m2);
    dataAsImg.(stimolo) = horzcat(dataAsImg.(stimolo),m3);
end
end

% Plot delle ricorrenze %
for stCount = 1:26
    stimolo = strcat('st', int2str(stCount));
    N = (length(dataImg.(stimolo)(:,:)))^2;

    figure;
    hold on;
    subplot(3,3,1);
    plot(dataImg.(stimolo));
    title('Img');
    axis square;
    subplot(3,3,2);
    plot(dataAcImg.(stimolo));
    title('Img + A.C. ');
    axis square;
    subplot(3,3,3);
    plot(dataAsImg.(stimolo));
    title('Img + A.S. ');
    axis square;

    soglia = 0.3;

    subplot(3,3,4);
    imagesc (heaviside(soglia - abs(dist(dataImg.(stimolo))))));
    axis square;
    subplot(3,3,5);
    imagesc (heaviside(soglia - abs(dist(dataAcImg.(stimolo)))));
    title('soglia 0.3');
    axis square;
    subplot(3,3,6);
    imagesc (heaviside(soglia - abs(dist(dataAsImg.(stimolo)))));
    colormap([1 1 1;0 0 0]);
    axis square;
    hold off;

```

```
soglia = 0.5;

subplot(3,3,7);
imagesc (heaviside(soglia - abs(dist(dataImg.(stimolo)))));
axis square;
subplot(3,3,8);
imagesc (heaviside(soglia - abs(dist(dataAcImg.(stimolo)))));
title('soglia 0.5');
axis square;
subplot(3,3,9);
imagesc (heaviside(soglia - abs(dist(dataAsImg.(stimolo)))));
colormap([1 1 1;0 0 0]);
axis square;
hold off;
end

% Conteggio Punti di ricorrenza d = 0.5 %
for stCount = 1:7
    stimolo = strcat('st', int2str(stCount));
    N = (length(dataImg.(stimolo)(:,:)))^2;
    soglia = 0.5;
    RPIImg = heaviside(soglia - abs(dist(dataImg.(stimolo)))));
    RRIImg(stCount) = (nnz(RPIImg == 1)*100/N);
    soglia = 0.5;
    RPACImg = heaviside(soglia - abs(dist(dataAcImg.(stimolo)))));
    RRAcImg(stCount) = (nnz(RPACImg == 1)*100/N);
    soglia = 0.5;
    RPASImg = heaviside(soglia - abs(dist(dataAsImg.(stimolo)))));
    RRASImg(stCount) = (nnz(RPASImg == 1)*100/N);
end
```