

**ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA**

**CAMPUS DI CESENA**

**SCUOLA DI INGEGNERIA E ARCHITETTURA**

**CORSO DI LAUREA SPECIALISTICA IN INGEGNERIA BIOMEDICA**

**STUDIO DELL'INTEGRAZIONE  
MULTISENSORIALE NELLA CORTECCIA  
ATTRAVERSO RETE NEURALE**

Tesi in

**Sistemi Neurali LM**

**Relatore**

Prof. Mauro Ursino

**Presentata da**

Lucia Damiani

Sessione Terza

Anno Accademico 2015 – 2016





# Indice

---

<b>Introduzione</b>	<b>pag. 9</b>
<b>Capitolo I – L’integrazione multisensoriale</b>	<b>pag. 13</b>
1.1 Cos’è l’integrazione multisensoriale	pag. 13
1.2 I neuroni multisensoriali del collicolo superiore	pag. 18
1.3 I principi dell’integrazione multisensoriale nei neuroni del CS	pag. 19
1.4 Integrazione multisensoriale nella corteccia cerebrale	pag. 23
1.5 Vie dell’elaborazione bimodale	pag. 25
1.5.1 Via dell’elaborazione unimodale visiva	pag. 25
1.5.2 Via dell’elaborazione unimodale acustica	pag. 29
1.6 Effetti on-line della modalità visiva verso stimoli acustici	pag. 30
1.6.1 Ventriloquismo	pag. 31
1.6.2 L’enhancement multisensoriale	pag. 31
1.7 Effetti off-line della modalità visiva verso stimoli acustici	pag. 32
1.8 Approci neuro-computazionali	pag. 33
<b>Capitolo II – Inferenza bayesiana applicata al problema del riconoscimento delle posizioni acustico-visive</b>	<b>pag. 35</b>
2.1 Modelli bayesiani	pag. 35
2.2 Modelli bayesiani di percezione multisensoriale	pag. 37
<b>Capitolo III – Descrizione del modello matematico che è stato applicato</b>	<b>pag. 42</b>
3.1 Concetti generali di stima bayesiana applicati	pag. 42

3.2 La funzione di verosimiglianza	pag. 44
3.3 Un modello di rete neurale per la stima di massima verosimiglianza	pag. 48
3.4 Probabilità a posteriori	pag. 52
3.5 Descrizione della rete neurale	pag. 56
3.5.1 Struttura base della rete	pag. 56
3.5.2 Addestramento della rete	pag. 59
3.6 Modello matematico in Matlab	pag. 62
<b>Capitolo IV – Simulazioni con i valori basali dei parametri</b>	<b>pag. 89</b>
4.1 I campi recettivi	pag. 90
4.2 Le sinapsi cross-modali	pag. 93
4.3 Risposta del modello agli input unimodali	pag. 94
4.4 Modello in risposta agli input cross-modali	pag. 96
4.5 L'effetto "fissione"	pag. 101
<b>Capitolo V – Analisi di sensitività</b>	<b>pag. 102</b>
5.1 Condizioni di base	pag. 102
5.2 Variazioni della precisione della stima	pag. 106
5.3 Variazioni delle sinapsi laterali	pag. 114
5.4 Variazioni della frequenza degli stimoli multisensoriali	pag. 121
<b>Conclusioni</b>	<b>pag. 133</b>
<b>Bibliografia</b>	<b>pag. 138</b>



*Alla mia famiglia*





## Introduzione

Il modo in cui noi riusciamo a percepire ciò che ci circonda è il risultato di un complesso meccanismo di integrazione di informazioni proveniente da differenti canali sensoriali. Il sistema nervoso, grazie alle sue caratteristiche anatomiche e funzionali è una perfetta macchina che riesce nell'elaborazione del messaggio che ci viene offerto dalla realtà che ci circonda. Una tra le più interessanti e studiate forme di integrazione è quella tra informazioni provenienti dal sistema visivo e da quello acustico. La capacità di localizzare uno stimolo acustico nello spazio è meno accurata ed affidabile della localizzazione visiva. Di conseguenza, un segnale visivo è spesso in grado di "catturare" (ventriloquismo) o di incrementare (enhancement multisensoriale) la performance di localizzazione acustica. In questa tesi sono stati trattati entrambi i fenomeni, ma è stata sottoposto ad uno studio mirato soprattutto il primo, quello del ventriloquismo. Abbiamo voluto evidenziare come l'esposizione ripetuta a stimolazione cross-modale visuo-acustica può dare luogo a fenomeni di breve o lungo termine.

Che tutto derivi da una riorganizzazione neurale mediata da una adeguata plasticità? Molto probabilmente sì.

Negli ultimi anni, lo studio dei processi d'integrazione multisensoriale è stato uno degli argomenti più proficuamente studiati nell'ambito delle neuroscienze. Numerose evidenze sperimentali hanno infatti contribuito ad individuare i processi cognitivi e le aree cerebrali responsabili dei fenomeni integrativi; tuttavia molti aspetti sui possibili meccanismi coinvolti restano ancora da chiarire. Da questo punto di vista, un importante contributo può venire dallo studio di modelli neurali in condizioni non basiche. Infatti, tali modelli permettono di simulare l'effetto di alterazioni nei fattori fondamentali del processo, quali la frequenza della stimolazione, le sinapsi laterali e la precisione della stimolazione. Nel presente progetto, i fattori elencati in precedenza sono stati analizzati

attraverso un'analisi di sensitività condotta sui parametri di un modello neurale, con lo scopo di rispondere ad alcune domande:

Come variano le sinapsi cross-modali al variare della frequenza di stimolazione?

Come è influenzato il fenomeno del ventriloquismo dall'azione delle sinapsi laterali inibitorie e eccitatorie?

Come varia la precisione dei risultati al variare della deviazione standard degli stessi stimoli multisensoriali?

Abbiamo cercato di rispondere a questi interrogativi in modo più chiaro possibile basandoci sulle prove computazionali effettuate e sui confronti con grafici relativi a condizioni basali. Il modello sviluppato simula i singoli neuroni (semplici unità di elaborazione) e le loro reciproche connessioni, in modo che le proprietà della rete connessa emergano come risultato della complessa e reciproca interazione tra le singole unità. La metodologia adottata nell'ambito dello studio dei correlati neurali dei processi di integrazione multisensoriale, mira a chiarire quali siano i meccanismi alla base dell'interazione visuo-acustica e dei suoi fenomeni di plasticità, con la possibilità di avere a disposizione un modello matematico sia per riprodurre condizioni che normalmente si verificano in vivo sia per simulare scenari nuovi, con la possibilità di effettuare delle predizioni da testare in successivi esperimenti.

La presente Tesi è organizzata in cinque capitoli:

Il Capitolo I, comprende una panoramica generale sulle caratteristiche dell'integrazione multisensoriale a partire dalle leggi che regolano i processi integrativi, per poi passare alla descrizione delle vie di elaborazione unimodali visiva ed acustica. Pone attenzione sulle caratteristiche anatomiche del nostro sistema neurale e in particolare alla regione del collicolo superiore, essendo tale regione la più studiata dal punto di vista dell'elaborazione multisensoriale.

L'ultima parte del Capitolo esamina l'influenza "on-line" ed "off-line" della modalità visiva sugli stimoli acustici e ne descrive brevemente gli effetti quali ventriloquismo ed enhancement.

Nel Capitolo II sono descritti i modelli bayesiani, prima in generale e poi applicati al problema della localizzazione della posizione audio-visiva trattata in questo progetto di ricerca.

Il Capitolo III illustra il modello matematico al partire dal quale è stata realizzata la rete neurale, argomento di questa tesi. Si è posta molta attenzione a come la struttura di tale modello, in seguito ad addestramento, possa riprodurre le funzioni principali necessarie per eseguire una stima Bayesiana, quali, di densità di probabilità di verosimiglianza e la probabilità a priori.

Il Capitolo IV tratta del comportamento a seguito di addestramento con valori di parametri basali. Sono esposte anche alcune importanti conclusioni che fanno riferimento alla capacità della rete di simulare l'integrazione visuoacustica alla quale siamo soggetti nella vita di tutti i giorni.

Il Capitolo V presenta un'analisi di sensitività condotta su alcuni parametri del modello, mettendo in evidenza il comportamento della rete in risposta a stimoli multimodali e cross-modali, e l'effetto di variazioni parametriche e strutturali.

Al termine dell'ultimo capitolo poi, saranno presentate le conclusioni e indicate le principali linee di sviluppo.



## **Capitolo I – L'integrazione multisensoriale**

La percezione del mondo esterno è basata sull'integrazione di input provenienti da diverse modalità sensoriali. I meccanismi alla base questi primi processi e l'organizzazione delle circuiterie sottostanti sono ancora oggetto di dibattito. Qui di seguito, ci concentreremo in modo particolare le interazioni audiovisive.

### **1.1 Cos'è l'integrazione multisensoriale?**

Gli esseri umani e gli animali si trovano quotidianamente a contatto con un mondo esterno molto complesso. Gli oggetti percepiti durante le esperienze comuni trasmettono un'elevata quantità di informazione, che viene successivamente estrapolata dai vari sensi contemporaneamente. Sicuramente la collaborazione tra i sensi implica notevoli vantaggi, come quello di estrarre informazioni complementari che originano dallo stesso oggetto o evento per migliorarne e facilitarne la percezione, riducendo notevolmente l'incertezza che caratterizzerebbe le misure provenienti dai singoli sensi.

Naturalmente i dati sensoriali trasportati dai singoli sensi devono essere combinati in un'unica via appropriata al fine di migliorare la percezione di oggetti ed eventi. Come conseguenza dell'importanza di questo comportamento, il problema dell'integrazione multisensoriale è stato da tempo riconosciuto come un aspetto cruciale delle neuroscienze (Calvert, Spence, & Stein, 2004; B.E. Stein & Meredith, 1993).

Un tradizionale punto di vista, riportato dai più grandi testi di neuroscienze fino a qualche decennio fa, sostiene che i segnali trasportati dai sensi vengono in primo luogo processati individualmente in aree unisensoriali appropriate, al fine di estrarre le singole informazioni, e solo successivamente ricombinati in aree di associazioni multisensoriale del cervello. Questo punto di vista, chiamato "unisensory before multisensory" è ancora valido in parte. I dati raccolti recentemente hanno però

modificato questa visione, mostrando che anche le aree corticali primarie (come la corteccia visiva primaria, V1, e la corteccia uditiva primaria, A1) ricevono informazioni da altre aree unisensoriali o da aree associative multisensoriali e mostrano dei comportamenti di tipo multisensoriale (Ghazanfar & Schroeder, 2006; Musacchia & Schroeder, 2009; Schroeder & Foxe, 2005).

Detto questo, è necessario introdurre una questione fondamentale: che cos'è la multisensorialità?

Quando possiamo affermare che un neurone è multisensoriale?

A questi quesiti non esiste una singola risposta, ma due differenti quanto corrette definizioni del comportamento multisensoriale. Nel primo caso un neurone multisensoriale è un neurone che risponde a stimoli unisensoriali di almeno due differenti modalità (ad esempio, come nel caso da noi trattato, visiva e uditiva). Ciò significa che questo neurone possiede un campo recettore per ciascuna modalità. Questo tipo di comportamento multisensoriale è tipico dei neuroni appartenenti al collicolo superiore, una profonda struttura subcorticale del mesencefalo coinvolta principalmente nei movimenti di occhi e testa verso gli stimoli esterni, e delle aree corticali associative. Una seconda definizione assume che un neurone mostra un comportamento multisensoriale se la risposta a uno stimolo unisensoriale di una modalità (definita convenzionalmente modalità 1) è affetta (ovvero eccitata o inibita) da uno stimolo proveniente da un'altra modalità (modalità 2). Il punto cruciale di questa definizione è che lo stimolo unisensoriale proveniente dalla modalità 2 non induce una risposta apprezzabile quando agisce da solo; esso può solo condizionare la risposta della modalità opposta durante la stimolazione cross-modale (cioè quando sono presenti gli stimoli di entrambe le modalità).

Questo tipo di multisensorialità appartiene ai neuroni della corteccie primarie, in passato considerati puramente unisensoriali, ma di cui oggi ne abbiamo riconsiderato le caratteristiche.

Malgrado la complessità dell'argomento, possono essere tracciate diverse regole relative all'integrazione multisensoriale e alcuni principi nella guida della modellazione matematica. In particolare i risultati raccolti negli ultimi due decenni hanno ispirato alcune ipotesi teoriche e permesso la formulazione di vari modelli neuro-computazionali che possono essere testati alla luce di questi dati. Come mostrato in vari campi delle neuroscienze, modelli di rete neurale, ispirati dalla neurofisiologia, possono svolgere un ruolo fondamentale nell'esprimere le conoscenze già esistenti in termini quantitativi, per riassumere i dati secondo un'impostazione coerente, per migliorare la conoscenza dei meccanismi complessi coinvolti nella fusione sensoriale, e per formulare e testare il comportamento previsto. Negli ultimi decenni si è sempre più posta attenzione sulle interazioni tra modalità sensoriali, ora largamente indagate in diversi settori di ricerca: numerose evidenze sperimentali provenienti da studi comportamentali, (Bertelson & De Gelder, 2004, Spence & Driver, 2004), neurofisiologici (Stein B.E., 2004), elettrofisiologici (Eimer, 2004) e di neuroimmagine funzionale (Macaluso & Driver, 2005, Calvert, Spence & Stein, 2004) hanno contribuito ad individuare i processi cognitivi e le aree cerebrali responsabili dei fenomeni integrativi. In molti casi, infatti, l'elaborazione di informazioni provenienti da una modalità viene influenzata da informazioni disponibili provenienti da altre modalità, che possono facilitare e/o interferire con i normali compiti cognitivi. Nella maggioranza dei casi, sembrerebbe proprio che avere a disposizione stimoli in più di una modalità faciliterebbe l'elaborazione e il recupero delle informazioni che giungono alla nostra mente. Inoltre, dobbiamo anche considerare il fatto che nella vita di tutti i giorni, raramente ci troviamo di fronte ad oggetti e percezioni in cui le caratteristiche siano riconducibili ad un'unica modalità sensoriale: per la maggior parte del tempo, siamo circondati da informazioni che ci giungono in più modalità percettive, le quali non si sommano banalmente assieme, ma interagiscono in modo fortemente complesso durante l'elaborazione del fenomeno percepito. Dunque, possiamo dire che, essendo la percezione un flusso continuo di informazioni in varie modalità sensoriali, sarebbe riduttivo leggere ogni fenomeno unisensoriale che giunge al nostro

cervello come un evento indipendente. La conclusione a cui ci porta questa riflessione è che i meccanismi cognitivi, nel corso dell'evoluzione umana, debbano essersi adattati sicuramente ad una percezione che è, per sua natura, multisensoriale.

Ma che vantaggi porta alla mente umana queste percezioni di multisensorialità?

Il fatto di avere diversi sensi fornisce evidenti vantaggi: ogni senso ha un'utilità ottimale in una diversa circostanza e, collettivamente, questi incrementano la probabilità di rilevare e di identificare eventi o oggetti di interesse. Tuttavia, tali vantaggi sembrano essere di minore importanza rispetto a quelli offerti dalla capacità di combinare fonti diverse di informazione. In questo caso, il prodotto integrato fornisce più informazioni sulla natura dell'oggetto o dell'evento esterno e lo fa in modo più rapido ed efficace di quanto previsto in base alla somma dei contributi sensoriali individuali. La sinergia, o interazione, tra i sensi, e la fusione del loro contenuto informativo, è definita "integrazione multisensoriale" (Meredith, 2002). Più specificamente, con tale termine si intendono definire i processi neurali implicati nella sintesi dell'informazione proveniente da stimoli cross-modali. L'integrazione multisensoriale è per lo più valutata considerando l'efficacia di una combinazione cross-modale di stimoli, rispetto a quella dei suoi componenti individuali, nell'evocare qualche tipo di risposta dall'organismo. Per esempio, l'ampiezza di una risposta ad un evento bimodale, che ha componenti sia visive che acustiche, è paragonata a quella evocata da stimoli visivi o acustici presentati individualmente (stimoli unimodali). Tale integrazione può comportare sia un incremento che una depressione della risposta neurale (Stein & Meredith, 1993). In entrambi i casi, tuttavia, è la risposta che noi ricerchiamo per avere maggiore chiarezza e conoscenza del mondo esterno. Dal momento che gli stimoli competono per l'attenzione e per l'accesso al sistema motorio, il possibile effetto dell'enhancement multisensoriale (o della depressione multisensoriale) è un aumento (o riduzione, nel caso della depressione) della probabilità di percepire un segnale e di iniziare una risposta motoria verso quest'ultimo. Per quanto concerne l'enhancement multisensoriale, le differenze in ampiezza riflettono le diverse computazioni che stanno alla base dell'elaborazione



dell'informazione sensoriale. Infatti, enhancement di ampiezza più elevata sono la conseguenza di computazioni super-additive di stimoli cross-modali, mentre enhancement di ampiezza ridotta sono dovuti a computazioni sub-additive. Con il termine “super-additivo” si intende una computazione neurale nella quale la risposta multisensoriale è superiore alla somma aritmetica delle risposte agli stimoli che la compongono (ovvero, lo stimolo acustico e visivo insieme offrono maggiori informazioni, rispetto alla percezione individuale di ciascuno). Il termine “sub-additivo”, invece, si riferisce ad una computazione neurale nella quale la risposta multisensoriale è più piccola della somma aritmetica delle risposte agli stimoli che la compongono.

Ma oltre all'ampiezza della risposta dello stimolo, quali sono gli effetti dovuti alla multisensorialità?

L'integrazione multisensoriale, oltre ai cambiamenti nell'ampiezza della risposta, può ridurre l'intervallo temporale tra la codifica sensoriale e la formazione del comando motorio (Bell, Meredith, Van Opstal & Munoz, 2005). In questo modo la risposta multisensoriale risulta avere una latenza significativamente più breve rispetto a quella di entrambe le risposte unisensoriali che la compongono (Rowland, Quessy, Stanford & Stein, 2007). Oltre ad alterare la salienza degli eventi cross-modali, l'integrazione multisensoriale è implicata nella creazione di esperienze percettive unitarie. Questo solleva delle questioni non banali: l'integrazione dell'informazione proveniente da sensi diversi deve render conto non solo della complessità dell'elaborazione dell'informazione in ciascuna modalità, ma anche del fatto che ogni modalità sensoriale ha delle percezioni qualitative specifiche, che non devono esser danneggiate dal processo integrativo.

Benefici dell'integrazione multisensoriale per i comportamenti di orientamento hanno ricevuto ampia attenzione ed hanno fornito molti suggerimenti sui meccanismi neurali sottostanti all'integrazione dell'informazione sensoriale.

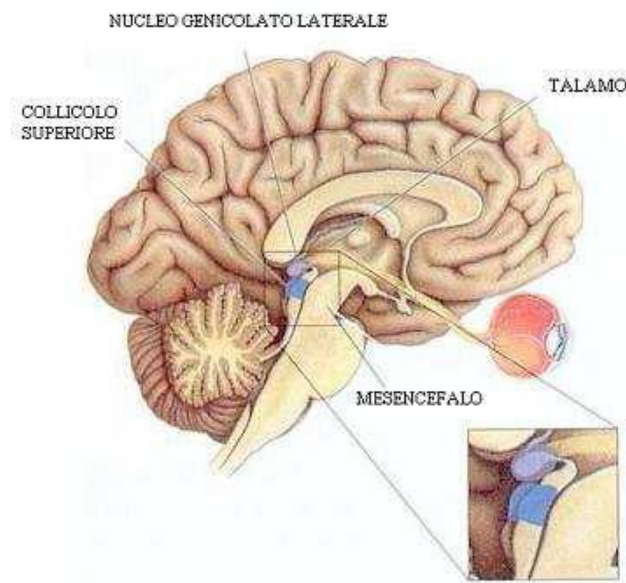
Un neurone multisensoriale è un neurone che risponde a stimoli provenienti da più di una modalità sensoriale. La presenza nel sistema nervoso di tali neuroni offre diversi vantaggi, legati alla

possibilità di integrare informazioni provenienti da fonti diverse. Affinché si realizzi una vera sintesi, la risposta ad uno stimolo multisensoriale deve differire dalla risposta a ciascuna delle singole componenti modalità-specifiche. Presenti in molte aree cerebrali ed in tutti i mammiferi, i neuroni multisensoriali sono particolarmente abbondanti nel Collicolo Superiore (CS) del mesencefalo, regione che per le sue caratteristiche rappresenta il punto di partenza per lo studio della complessità dell'informazione percettiva.

## **1.2 I neuroni multisensoriali nel collicolo superiore**

Nel sistema nervoso la rappresentazione sensoriale è basata su regole organizzative sistematiche, il cui principio sottostante è la creazione di una mappa che rispecchi la diretta ricostruzione spaziale della superficie recettoriale, mantenendo le relazioni di vicinanza esistenti (per cui i recettori di parti adiacenti proiettano a neuroni disposti in parti adiacenti). Ciò è particolarmente evidente per la modalità visiva e per quella tattile, le quali presentano un'organizzazione neurale spaziotopica, in quanto esiste una precisa corrispondenza tra stimoli dell'ambiente e distribuzione dell'attività neurale a livello delle rispettive superfici recettoriali; lo spazio acustico, diversamente, è organizzato in modo tonotopico: la superficie recettoriale del sistema uditivo, infatti, rappresenta in modo ordinato le bande di frequenza a cui l'orecchio è sensibile. In questo lavoro ci focalizzeremo sull'organizzazione spaziale di tipo spaziotopica. Ad ogni livello del sistema nervoso centrale le rappresentazioni dello spazio uditivo, visivo e somatosensoriale occupano regioni spazialmente distinte, definite sia funzionalmente che anatomicamente. Nonostante esistano vie neurali specializzate per il trattamento di specifiche informazioni sensoriali, nel cervello sono presenti aree (corticali e sottocorticali) che ricevono informazioni da diversi canali sensoriali.

Come accennato nel paragrafo precedente, focalizzeremo l'attenzione soprattutto su una regione in particolare del mesencefalo, il collicolo superiore (Figura 1).



**Figura 1** Veduta sagittale del cervello umano rappresentante il collicolo superiore del mesencefalo, il nucleo genicolato

Questa struttura sottocorticale controlla i cambiamenti nell'orientamento (per esempio, gli spostamenti dello sguardo) in risposta a stimoli provenienti dallo spazio visivo controlaterale rispetto al collicolo indagato. I suoi input visivi, acustici e somatosensoriali sono derivati da fibre sensoriali ascendenti e da proiezioni discendenti dalla corteccia, che convergono in vari modi sui neuroni del CS.

### **1.3 I principi dell'integrazione multisensoriale nei neuroni del Collicolo superiore.**

Studi di registrazione da singole unità del gatto (Stein & Meredith, 1993, Burnett, Stein, Perrault & Wallace, 2007) hanno dimostrato come i neuroni multisensoriali del CS obbediscano a tre leggi integrative fondamentali:

- la legge spaziale;
- la legge temporale;
- la legge dell'efficacia inversa.

La legge spaziale è una proprietà di particolare importanza per il ruolo di orientamento del CS. Ogni neurone multisensoriale ha campi recettivi eccitatori multipli, uno per ciascuna modalità a cui risponde. Un campo recettivo è l'area dello spazio sensoriale nella quale la presentazione di uno stimolo induce la risposta di un particolare neurone. La proprietà spaziale dipende dall'organizzazione dei campi recettivi dei neuroni multisensoriali in zone centrali, eccitatorie, e zone periferiche, inibitorie. I campi recettoriali acustici presentano regioni eccitatorie più ampie rispetto a quelli visivi, con confini laterali che possono estendersi ben oltre i corrispondenti campi recettivi visivi.

Stimoli delle due diverse modalità saranno percepiti come originati dalla stessa fonte fintantoché rientreranno nello spazio definito dai loro campi recettivi sovrapposti; non è necessario, quindi, che essi abbiano origine effettivamente dallo stesso punto nello spazio (Kadunce et al. 2001). Se gli stimoli provengono da posizioni spazialmente separate, in modo tale che uno stimolo ricada entro e l'altro al di fuori del campo recettivo del neurone, il secondo stimolo potrà deprimere la risposta neurale del primo o quantomeno non indurre un *enhancement* (Meredith and Stein 1986; Meredith and Stein 1996; Kadunce et al. 1997). La depressione della risposta si ha quando il secondo stimolo si trova entro la regione inibitoria che circonda i campi recettivi eccitatori di alcuni neuroni del CS, e può esser abbastanza potente da sopprimere l'eccitazione evocata dall'altro stimolo. Il principio spaziale dell'integrazione multisensoriale è piuttosto potente ed è evidente in una moltitudine di situazioni percettive nelle quali la posizione di un evento è cruciale. Tuttavia, la necessità di mantenere i campi recettivi allineati spazialmente e l'abilità di muovere indipendentemente ogni organo di senso sembrerebbero essere incompatibili. Una soluzione adottata dal CS è di connettere i

vari campi recettivi modalità-specifici alla posizione degli occhi. Per esempio, muovere gli occhi verso sinistra produce spostamenti compensatori sia dei campi recettivi acustici (Jay and Sparks 1984; Hartline et al. 1995; Peck et al. 1995) che somatosensoriali (Groh and Sparks 1996). Una simile compensazione per i movimenti oculari sembra creare un sistema di coordinate oculocentrico comune, assicurando che i componenti individuali di un evento cross-modale interagiscano per produrre un locus di attività singolo e coerente entro la mappa sensori-motoria del CS.

Un altro principio o legge che deve essere rispettata ai fini dell'integrazione è che i diversi stimoli sensoriali devono anche essere temporalmente contigui (Meredith et al. 1987; Recanzone 2003). In generale, questi stimoli possono raggiungere il sistema nervoso entro una finestra temporale di una certa ampiezza, in qualche caso di diverse centinaia di millisecondi. Questo consente che vi sia integrazione a prescindere dalle diverse latenze di risposta, velocità di conduzione e comparsa degli stimoli visivi, acustici e somatosensoriali. L'ampiezza della risposta integrativa è sensibile alla sovrapposizione temporale delle risposte iniziate da ciascun input sensoriale ed è solitamente massima quando i periodi di picco di attività coincidono. Dunque, in accordo con tale legge temporale, una massima interazione multisensoriale consegue alla presentazione di stimoli temporalmente coincidenti, mentre stimoli temporalmente separati sono processati come eventi indipendenti (Stein & Meredith, 1993). Di conseguenza, livelli massimi di enhancement sono ottenuti qualora due input siano presentati simultaneamente. In realtà è stato riscontrato che la combinazione di stimoli unimodali acustici e visivi a determinati intervalli (50 e 150 ms) produce ugualmente un enhancement della risposta. Sembrerebbe che l'intervallo temporale ottimale per l'integrazione si aggiri attorno ai 100 ms (Meredith, Nemitz & Stein, 1987).

Per quanto riguarda la legge dell'efficacia inversa: "l'enhancement multisensoriale è massimo combinando stimoli unimodali deboli, rispetto alla combinazione di stimoli unimodali potenti" (Stein & Meredith, 1993). Due stimoli forti inducono una risposta maggiore di due stimoli deboli, ma il beneficio della multisensorialità è quasi nullo. L'utilità di questo principio è piuttosto intuitiva:

singoli segnali che siano altamente salienti saranno rilevati e localizzati facilmente. In accordo con questa terza legge si precisa una relazione inversa tra il livello di efficacia degli stimoli e la risposta neurale da loro evocata. La combinazione di stimoli unimodali deboli produce un enhancement maggiore rispetto alla combinazione di stimoli unimodali potenti; ciò significa che la combinazione di due stimoli unimodali, ciascuno dei quali, singolarmente preso, non è in grado di evocare un significativo effetto sull'attività del neurone, può aumentare drammaticamente la risposta nei neuroni multisensoriali (Meredith & Stein, 1986a, Meredith & Stein, 1986b, Stein & Meredith, 1993). L'attività dei neuroni multisensoriali del CS è strettamente dipendente da input eccitatori discendenti da una regione specifica della corteccia associativa. Questi input provengono prevalentemente dal solco ectosilvano anteriore (AES), ma anche dall'adiacente solco soprasilvano rostro-laterale (rLS) (Stein et al. 1983; Wallace and Stein 1994; Wilkinson et al. 1996; Jiang et al. 2001; Jiang et al. 2002; Jiang et al. 2006; Alvarado et al. 2007; Alvarado et al. 2007) (vedi Figura 2).

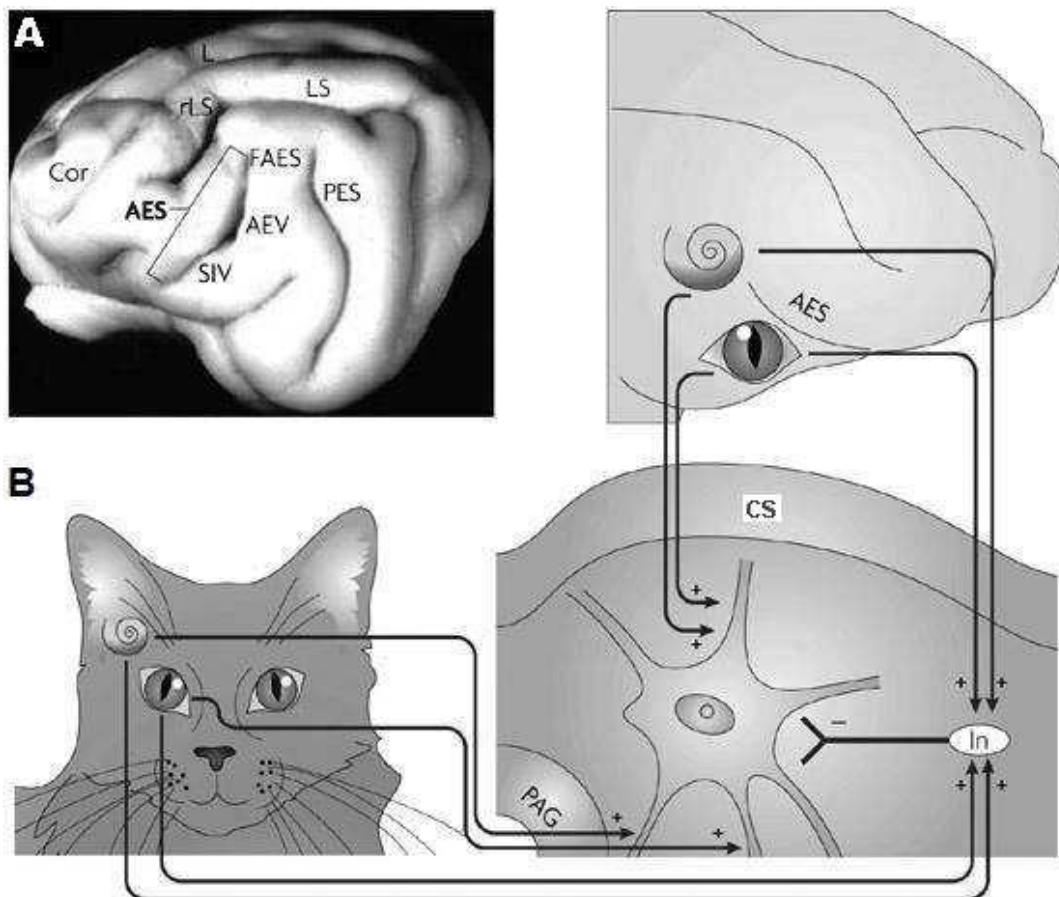
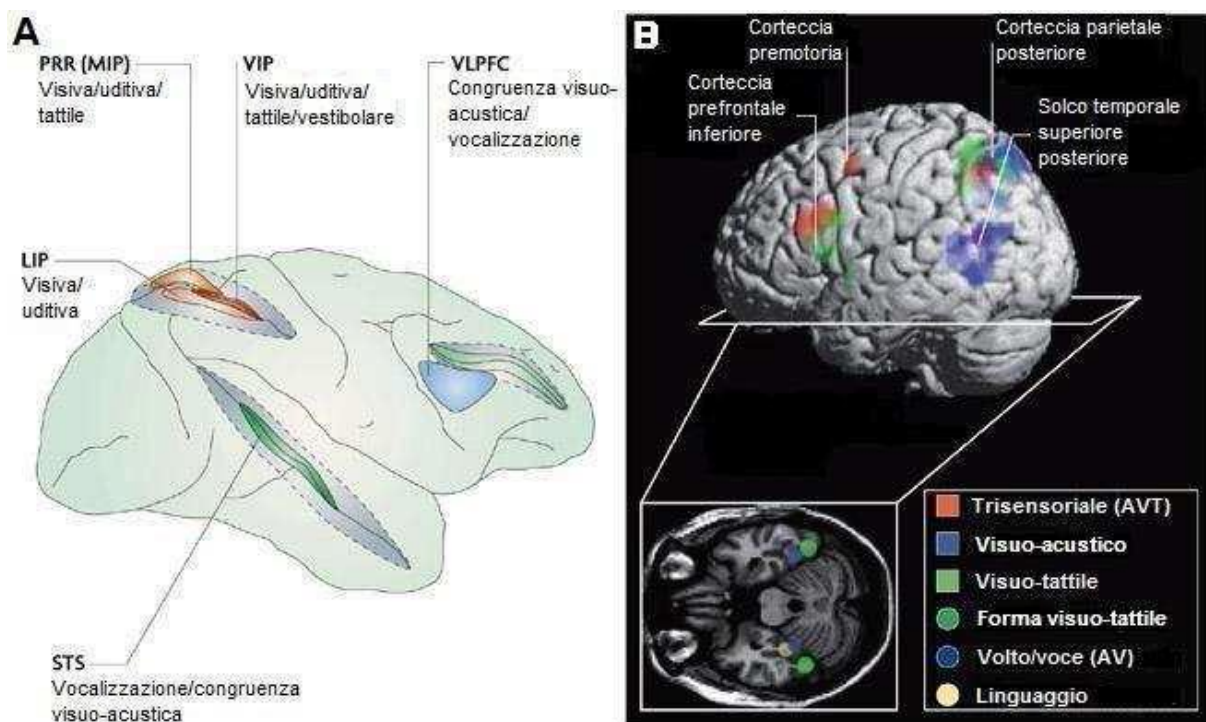


Figura 2 Veduta sagittale del cervello del gatto (A), che evidenzia il solco ectosilvano anteriore (AES). Si notano anche le regioni somatosensoriali (SIV), visive (AEV), visive e uditive (FAES) di AES. (B).

#### 1.4 Integrazione multisensoriale nella corteccia cerebrale.

Nei primati non umani e nell'uomo sono state identificate numerose aree di integrazione multisensoriale, nelle quali la risposta agli eventi cross-modali è governata dalle stesse leggi integrative che regolano la sintesi multisensoriale a livello neurale. In particolare, c'è consenso nell'indicare tra le strutture sottocorticali il CS e l'insula (Bushara, Grafman & Hallet, 2001, Calvert, Hansen, Iversen & Brammer, 2001), mentre a livello corticale sono state individuate l'area STS (Solco Temporale Superiore), principalmente coinvolta nella percezione bimodale del linguaggio (Calvert, Campbell & Brammer, 2000), e l'area IPS (Solco Parietale Inferiore), implicata nei processi di attenzione spaziale (Meienbrock, Naumer, Doehrmann, Singer, & Muckli, 2007).

In Figura 3 sono mostrate le regioni multisensoriali nella corteccia della scimmia. Le aree colorate rappresentano regioni nelle quali sono stati identificati neuroni che rispondono a più di una modalità. Tra queste regioni sono visibili l'area intraparietale laterale (LIP); la regione parietale di reaching (PRR); l'area intraparietale mediale (MIP), l'area intraparietale ventrale (VIP), localizzata nel fondo del solco intraparietale; la corteccia prefrontale ventrolaterale (VLPFC) ed il solco temporale superiore (STS). A destra è visibile una ricostruzione tridimensionale del cervello umano con le presunte aree multisensoriali, definite da criteri di imaging funzionale. Le aree attive a seguito di stimoli visivi, acustici e tattili sono misurate come mostrato. Il colore rosso denota una sovrapposizione trisensoriale (visivo-acustico-tattile); il blu denota regioni di sovrapposizione visiva e acustica; il verde regioni di sovrapposizione visiva e tattile. La sezione orizzontale presente in basso identifica regioni di attivazione per stimoli multisensoriali complessi (oggetti, stimoli linguistici).



**Figura 3**

**Aree multisensoriali nella corteccia della scimmia (A).**

**(B) Cervello umano raffigurante aree multisensoriali putative, come definite dai criteri di neuroimmagine.**



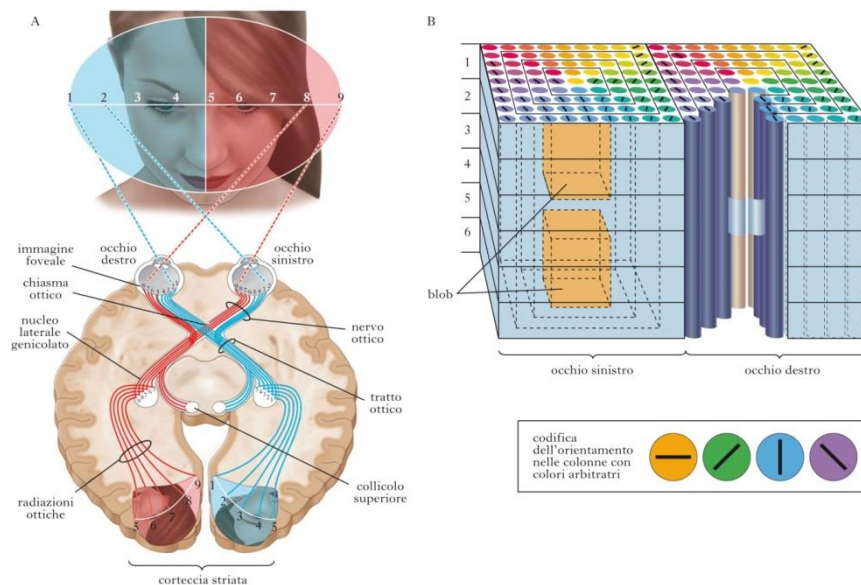
L'interazione audio-visiva non è un'abilità presente già alla nascita, ma viene acquisita progressivamente durante lo sviluppo in ambiente multisensoriale. Lo scopo della Tesi è quello di proporre un modello matematico di rete neurale in grado di simulare l'addestramento nelle cortecce primarie visive e uditive, a partire da condizioni base imposte (in termini fisiologici quelle di un neonato), in seguito a riproduzioni di esperienze audio-visive unisensoriali e multisensoriali (visive e acustiche), e di mostrare gli effetti successivi a tale addestramento (per esempio il ventriloquismo, enhancement multisensoriale ).

## **1.5 Vie dell'elaborazione unimodale**

Nel paragrafo successivo saranno esaminate le modalità di elaborazione unimodale visiva ed acustica con lo scopo di comprendere come si elaborano i segnali unimodali.

### **1.5.1 Via di elaborazione unimodale visiva**

Per comprendere il funzionamento delle vie dell'elaborazione unisensoriali visive, occorre fare una panoramica riguardo alle caratteristiche anatomiche dell'occhio umano. In generale nei mammiferi si possono distinguere tre unità principali di elaborazione: la retina, il nucleo genicolato laterale e la corteccia visiva. Questi tre blocchi costituiscono, in linea generale, la via ottica principale, ovvero quella che garantisce il percorso attraverso cui le informazioni provenienti dal mondo esterno vengono elaborate sino a raggiungere le aree corticali associative, deputate principalmente all'integrazione dell'informazione visiva con le informazioni provenienti dagli altri canali sensoriali. La luce attraverso il cristallino, viene focalizzata sulla cornea, giunge nell'umor vitreo (che rappresenta una cavità dell'occhio) e poi viene assorbita dai fotorecettori presenti nella retina.



**Figura 4 vie dell'elaborazione visiva**

La retina è una membrana fotosensibile costituita da fotorecettori (coni e bastoncelli), cellule che trasformano la luce in stimolo nervoso. I coni sono responsabili della visione diurna mentre i bastoncelli assicurano la visione notturna.

La retina, tuttavia, non si limita a registrare passivamente le immagini formate sulla sua superficie, ma le scompone analizzandone i molteplici parametri fisici.

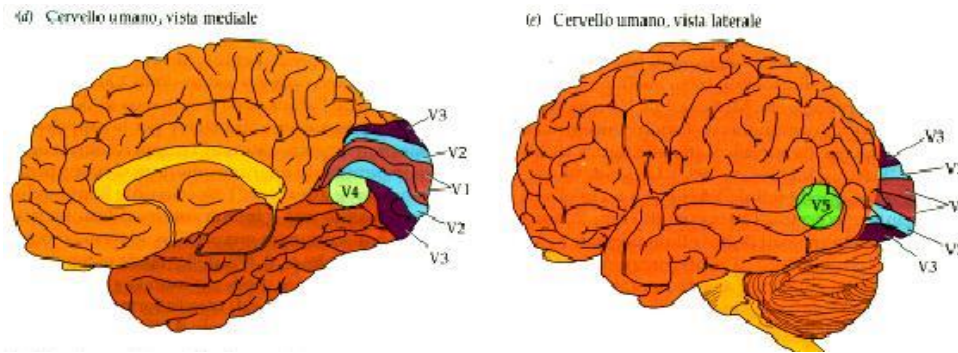
Nella retina, infatti, oltre ai fotorecettori, si possono individuare altre quattro principali classi di neuroni:

- le cellule orizzontali, che trasmettono orizzontalmente i segnali ricevuti dai fotorecettori;
- le cellule bipolari che inviano segnali alle cellule gangliari;
- le cellule amacrine, che hanno funzioni simili alle orizzontali, ma operano in uno strato sottostante;
- le cellule gangliari, che con le loro terminazioni assoniche trasmettono il messaggio nervoso ai centri superiori.

Il segnale parte da qui e arriva fino alle fibre del nervo ottico (quest'ultimo funge da cavo della corrente elettrica che porta l'informazione fino al cervello). Grazie al contributo della retina quindi, l'informazione sensoriale viene pre-elaborata ed è pronta per essere codificata dalle cellule gangliari, quest'ultime infatti la trasmettono lungo i propri assoni e la portano al nucleo genicolato laterale. Il secondo blocco è proprio il nucleo genicolato laterale, un blocco fondamentale per elaborare l'informazione sensoriale. Rappresenta una porzione del talamo attraverso cui quasi tutti gli stimoli sensoriali in ingresso devono passare per poter giungere fino alla corteccia ( fatta eccezione per gli stimoli sensoriali olfattivi). Nel passaggio dalla retina al nucleo genicolato laterale, le caratteristiche topografiche dello stimolo vengono preservate;, questo grazie al fatto che gli assoni che proiettano dalla retina vanno a connettersi con i neuroni del genicolato laterale in maniera precisa ed ordinata, rispettando la configurazione spaziale. Le posizioni relative ai neuroni sulla retina vengono quindi ripetute sul nucleo genicolato laterale mantenendo tutte le informazioni dello stimolo sensoriale iniziale. La funzione di questa stazione intermedia (il nucleo genicolato laterale) non è ancora del tutto nota e, sebbene sembri svolgere una semplice azione di trasferimento, per le sue dimensioni e per la posizione che occupa, potrebbe rivestire un ruolo ben più importante.

Le informazioni provenienti dal nucleo genicolato laterale poi vengono inviate verso la corteccia visiva primaria.

Cosa possiamo dire invece riguardo alla complessa struttura del terzo blocco di elaborazione del segnale visivo?



**Figura 5** aree della corteccia cerebrale

La prima area della corteccia visiva è la cosiddetta area 17 (secondo la classificazione di Brodmann), detta anche “area striata”; qui si innervano le terminazioni provenienti dal nucleo genicolato laterale: in ciascuna di esse viene rappresentato in modo abbastanza dettagliato circa metà campo visivo. Tuttavia esiste anche una seconda via di elaborazione: le informazioni provenienti dai fotorecettori della retina si dirigono verso il collicolo superiore e di qui si dipartono verso la corteccia extra-striata, come mostrato in Figura 1.7. Questa seconda via risulta molto utile nel caso in cui sia presente un danneggiamento della corteccia visiva primaria (V1), essa permette infatti che le informazioni visive possano ancora raggiungere la corteccia exstrastriata ed eventualmente aree associative e multisensoriale nonostante lesioni della corteccia visiva striata.

La corteccia visiva primaria striata (che chiameremo V1, sapendo che in generale la corteccia visiva è composta da altre parti V2,V3,V4,V5 di tipo extra-striato) è una struttura molto complessa in cui le informazioni in uscita dalla retina e dal nucleo genicolato laterale, parzialmente elaborate, vengono separate e categorizzate per un'analisi più elaborata. Essa è suddivisa in diverse aree organizzate gerarchicamente e ciascuna di esse presenta una mappa retinotopica più o meno precisa.

### 1.5.2 Via di elaborazione unimodale acustica

Come appena fatto per la via dell'elaborazione del segnale visivo, occorre fare un accenno anche alle caratteristiche anatomiche dell'orecchio per introdurre la via dell'elaborazione acustica. In generale, quando parliamo di sistema acustico parliamo principalmente della coclea, del nervo cocleare e della via acustica centrale, che dal nucleo cocleare nel tronco encefalico porta alla corteccia del lobo temporale. Già da queste prime veloci descrizioni si può intuire come la via acustica, come anche quella visiva, siano molto elaborate, infatti per entrambe esiste una percezione 2D data rispettivamente dai due occhi e dalle due orecchie. Per localizzare gli stimoli in entrambi i casi è necessaria una rete nervosa molto complessa. La prima stazione cellulare si trova nel nucleo cocleare dove terminano tutte le fibre del nervo cocleare che entrano nel tronco encefalico. Da qui alcune fibre di secondo ordine proiettano al collicolo inferiore dal lato opposto tramite il corpo trapezoidale ed il lemnisco laterale. Il braccio congiuntivo inferiore collega il collicolo inferiore al nucleo genicolato mediale, che proietta alla corteccia acustica primaria posta nel lobo temporale. Un piccolo ma importante contingente di fibre si porta dal nucleo olivare superiore ai centri uditivi superiori.

La Figura 6 mostra l'organizzazione generale della via acustica centrale dall'arrivo dello stimolo fino alla corteccia cerebrale.

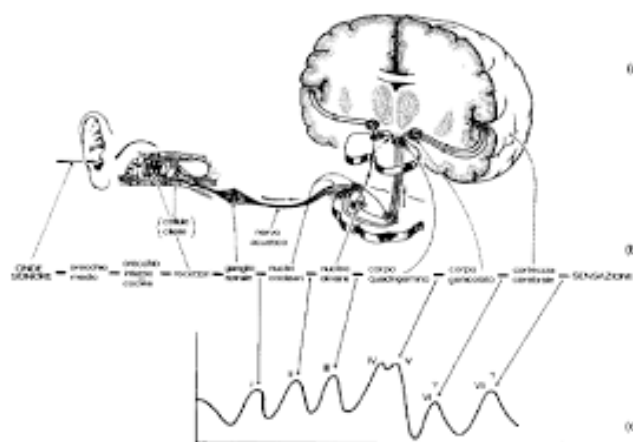


Figura 6 vie dell'elaborazione acustica

Nei paragrafi successivi saranno esaminati gli effetti indotti dalla stimolazione visiva sulla localizzazione di stimoli acustici ed i correlati neurali di queste interazioni cross-modali. Tali effetti sono stati osservati sia nel caso di risposte in condizioni base, sia nel caso di risposte in seguito a variazioni parametriche e strutturali della rete. Inoltre occorre tenere in considerazione sia la condizione on-line che quella off-line della modalità visiva verso gli stimoli acustici.

### **1.6 Effetti on-line della modalità visiva verso stimoli acustici**

Obiettivo di questo studio è contribuire a fare luce, attraverso un modello di rete neurale, sui meccanismi alla base di fenomeni di integrazione multisensoriale. Per tale motivo, dopo avere descritto il funzionamento delle aree unimodali acustica e visiva, si ritiene necessario sottolineare quali siano gli effetti che la modalità visiva produce su quella acustica (e viceversa) in condizioni di stimolazione cross-modale spazialmente coincidente o meno.

### 1.6.1 Il ventriloquismo

L'abilità di localizzare uno stimolo acustico nello spazio è un processo computazionale meno accurato ed affidabile della localizzazione visiva. Di conseguenza, laddove il sistema visivo e quello acustico veicolano informazioni spaziali discordanti, il risultato percettivo che si crea è l'illusione che lo stimolo acustico provenga dalla posizione occupata dallo stimolo visivo, un fenomeno noto come "Effetto Ventriloquismo" (Howard & Templeton, 1966). Tale effetto testimonia come, entro determinate condizioni, l'informazione acustica e quella visiva siano integrate in un percetto unificato: il conflitto spazio-temporale che nasce dalla presenza di stimoli temporalmente coincidenti ma spazialmente disparati, si risolve con l'apparente "cattura" del suono da parte dello stimolo visivo. L'effetto ventriloquismo è stato originariamente studiato in laboratorio chiedendo ai soggetti di effettuare un movimento di pointing verso il target acustico mentre stimoli visivi irrilevanti erano inviati simultaneamente a breve distanza (Bertelson & Radeau, 1981). Ampie disparità spaziali e temporali riducono la grandezza del bias visivo ed annullano la percezione di unicità dei due eventi. Inoltre, la forza dell'effetto si riduce progressivamente per stimoli visivi presentati alla periferia del campo visivo, laddove la loro acuità spaziale decresce. Oltre ad essere indipendente da fattori semantici, quali la familiarità degli stimoli e le informazioni contestuali, il ventriloquismo è la manifestazione di un cambiamento percettivo automatico e non il risultato di aggiustamenti post-percettivi della risposta.

### 1.6.2 L'enhancement multisensoriale

Nel loro insieme, gli studi sul ventriloquismo hanno dimostrato che uno stimolo visivo semplice, sotto determinate circostanze, è in grado di indurre un errore sistematico nel processo di localizzazione di un target acustico, indipendentemente dall'allocazione delle risorse. Tuttavia, le

interazioni cross-modali non si traducono esclusivamente in distorsioni percettive. Al contrario, numerose evidenze suggeriscono che il mantenimento dei sistemi di integrazione cross-modale risponda all'esigenza evolutiva di incrementare le possibilità offerte dai sistemi sensoriali unimodali (Meredith & Stein, 1983, Rowland, Quessy, Stanford & Stein, 2007). A livello fisiologico, stimoli cross-modali presentati in coincidenza spaziale e temporale evocano nei neuroni multisensoriali risposte largamente maggiori di quelle generate dalle singole componenti unisensoriali. A livello comportamentale, è stato documentato un miglioramento nelle risposte di orientamento spaziale in presenza di stimoli cross-modali spazialmente e temporalmente coincidenti, tale fenomeno è noto come enhancement multisensoriale (Laurienti, Burdette, Wallace, Yen, Field, & Stein, 2002, Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004).

### **1.7 Effetti off-line della modalità visiva verso stimoli acustici**

Sotto particolari circostanze uno stimolo visivo può produrre cambiamenti di lunga durata nella percezione dello spazio acustico, un fenomeno noto come "*After-Effect del Ventriloquismo*" (Zwiers, Van Opstal, & Paige, 2003). In questo caso, dopo un periodo di costante esposizione ad una coppia di stimoli visuo-acustici spazialmente separati, ma temporalmente coincidenti, lo stimolo acustico, anche se fornito in assenza di stimolo visivo, viene percepito come originante dalla posizione dello stimolo visivo precedentemente associato ad esso. L'After Effect è stato generalmente considerato un indice di elaborazione percettiva più genuino rispetto a quello osservabile con risposte on-line. Gli After Effect, infatti, sono misurati confrontando le risposte a stimoli unimodali prima e dopo l'esposizione a coppie di stimoli cross-modali. Dal momento che nel test di localizzazione unimodale acustica lo stimolo visivo non è presente, esso non può nemmeno esercitare alcuna influenza sul sistema di risposta. La presenza dell' After Effect, evidente negli animali e nell'uomo, suggerisce come un periodo relativamente breve di esposizione a stimoli



acustici e visivi spazialmente separati, induca in entrambe le specie cambiamenti a lungo termine nella rappresentazione dello spazio acustico. Il fondamento di questa alterazione percettiva è certamente un meccanismo di rapida plasticità neuronale (Bertelson & De Gelder, 2004), conseguente ad un fenomeno di apprendimento percettivo.

I meccanismi neurali di questo effetto non sono ancora del tutto chiari, come evidenziano i risultati dello studio sopracitato (Elisa M, Cristiano C, Mauro U.,2012).

### **1.8 Approcci neuro-computazionali**

La struttura delle connessioni cross-modali nelle aree sensoriali primarie è tuttora controversa. Alcuni possibili meccanismi includono connessioni feedback a partire dalle cortecce associative multisensoriali (Buchel, Price, & Friston, 1998; Macaluso, Frith, & Driver, 2000; McDonald, Teder-Salejarvi, Di Russo, & Hillyard, 2003, 2005) o dalle strutture subcorticali (incluso il collicolo superiore) (Meredith, 2002; Meredith & Stein, 1986b; Mark T. Wallace & Stein, 2007) e connessioni laterali dirette tra aree unisensoriali.

In Figura 7 sono riportate tre possibili schemi strutturali in merito ai meccanismi di connessione sopracitati.

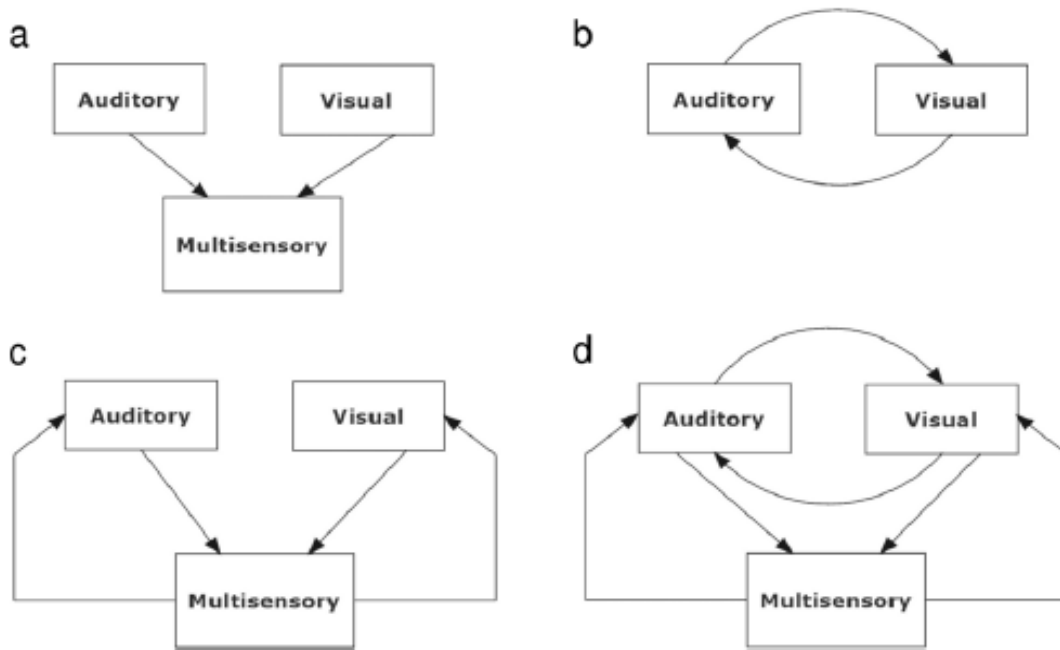


Figura 7 tre possibili schemi strutturali in merito ai meccanismi di connessione sopracitati

Il diagramma nel pannello **a** rappresenta una connessione feedforward puro, che viene spesso utilizzata per simulare l'attività dei neuroni multisensoriali nel collicolo superiore. Il diagramma nel pannello **b** rappresenta connessioni laterali dirette tra due aree unimodali. Il diagramma nel pannello **c** invece, rappresenta sia un collegamento feedforward ad un area sensoriale, sia un collegamento feedback dalla zona multisensoriale a quella unisensoriale. Infine, il diagramma presente nel pannello **d** incorpora tutti i meccanismi precedenti. Questo è stato utilizzato in alcuni modelli corticali (Hoshino, 2011; Magosso et al., 2010). Vale la pena notare che solo nel primo diagramma le due aree ('uditiva' e 'visiva') sono veramente di tipo unisensoriale. Negli altri diagrammi, le due aree corticali primarie sono influenzate da altre modalità mediante connessioni dirette o feedback: ciò ne determina un comportamento multisensoriale.

In questa tesi il diagramma da tenere in considerazione sarà lo schema **b**.

## **Capitolo II – Inferenza bayesiana applicata al problema del riconoscimento delle posizioni acustico-visive.**

Molti dei modelli più recenti per lo studio dell'integrazione multisensoriale si basano su un approccio bayesiano. Il concetto fondamentale di tale approccio è che il nostro cervello “funziona e risponde agli ingressi ambientali in maniera ottimale in condizioni di incertezza”. Questa incertezza deriva da una serie di fattori come il rumore ambientale, l'intrinseca variabilità neurale e rappresentazioni (ad esempio, la densità dei recettori nella retina o nella pelle). Il cervello deve prendere questa “incertezza” in considerazione per fare inferenze accurate sul mondo esterno. Il problema è particolarmente rilevante se si considera l'integrazione di più input sensoriali, poiché l'incertezza (o l'affidabilità) di ciascun ingresso deve essere considerata. I Modelli Bayesiani presuppongono che il cervello esegua una integrazione ottimale, calcolando la probabilità a posteriori dell'evento dato un insieme di informazioni incerte. In primo luogo, noi commenteremo i lavori di letteratura che utilizzano il modello bayesiano per modellare la percezione multisensoriale, fornendo una risposta percettiva di calcolo “Bayes ottimale”. Poi, ci troveremo ad affrontare la principale sfida della teoria bayesiana, vale a dire quali possono essere i circuiti neurali e meccanismi alla base della percezione dell' inferenza bayesiana: a questo scopo, ci si sposterà dal caso più semplice dei singoli neuroni che realizzano inferenza bayesiana al caso più complesso della popolazione di neuroni che realizza un inferenza bayesiana attraverso " Codici di popolazione probabilistici ".

### **2.1 Modelli Bayesiani**

I modelli bayesiani in generale, rappresentano l'incertezza relativa ai parametri tramite funzione di distribuzione di probabilità. Il processo di apprendimento, nel contesto bayesiano, consiste nell'aggiornamento delle opinioni iniziali riguardo al parametro  $\theta$  (rappresentato dalla distribuzione

di probabilità  $p(\vartheta)$ ) alla luce dei dati osservati. Una volta disposta un set di dati  $x$ , si ottiene una nuova distribuzione di probabilità per  $\vartheta$ , detta distribuzione di probabilità a posteriori  $p(\vartheta|x)$ .

Le fondamenta di questo approccio sono rappresentate dal teorema di Bayes, la cui applicazione permette di tenere conto di opinioni e conoscenze eventualmente esistenti a priori sul fenomeno oggetto di studio. Esso connette le distribuzioni a priori e di verosimiglianza con la distribuzione a posteriori in un'unica formula e fornisce un metodo per modificare il livello di fiducia in una data ipotesi alla luce di una nuova informazione. Prima di introdurre il teorema spieghiamo il significato dei seguenti concetti fondamentali per la comprensione di tale modello.

- **Probabilità a priori:** non è altro che il modo di rappresentare l'informazione a priori riguardo ai valori dei parametri prima di osservare i dati. Tale probabilità è indipendente dall'esperienza attuale mentre invece può derivare dalle esperienze passate.
- **verosimiglianza:** rappresenta il collegamento vero e proprio tra i dati e i parametri e introduce i dati alle analisi.
- **probabilità a posteriori:** è il risultato della combinazione dei dati e dell'informazione a priori e può essere letta come la probabilità che il parametro sconosciuto assuma un certo valore alla luce dei dati e dell'informazione che possediamo a priori. In termini probabilistici è la probabilità condizionata che il parametro sconosciuto assuma un certo valore data la misura

Il teorema di Bayes, riportato qui in maniera generale, è il seguente:

$$p(\vartheta|x) = \frac{p(\vartheta)p(x|\vartheta)}{p(x)}$$

Dove  $p(x|\vartheta)$  è la funzione di densità di  $x$  dato  $\vartheta$ .

## 2.2 Modelli Bayesiani di percezione multisensoriale

Recentemente, per interpretare gli esperimenti sull'integrazione multisensoriale, sono stati adottati modelli bayesiani. Questi modelli mirano a formalizzare come diversi segnali sensoriali con diversi livelli di affidabilità riescono a combinarsi a livello percettivo. Nella maggior parte di questi studi i soggetti sono esposti a stimoli cross-modalità presentati a vari gradi di incongruenza nello spazio (ad esempio, in diverse posizioni) o nel tempo (ad esempio, il numero o il tasso di stimoli diversi) e si richiede ai soggetti di esprimere un giudizio sulle caratteristiche degli stimoli esterni a cui sono stati sottoposti (ad esempio, localizzarli spazialmente ecc). In queste condizioni, effetti percettivi cross-modalità sono spesso accompagnati da illusioni come il ventriloquismo, nel dominio spaziale (Alais e Burr, 2004; Battaglia, Jacobs, & Aslin, 2003; Wallace et al., 2004b) o una illusione di fissione / fusione nel dominio temporale (Andersen et al, 2004;.. Shams et al, 2000; Shams, Ma, e Beierholm, 2005). Il modello bayesiano interpreta questi effetti percettivi supponendo che le caratteristiche di una stimolazione esterna  $s$ , chiamiamola “ $\vartheta$ ” (ad esempio, la posizione spaziale degli stimoli o il numero di stimoli) si trasformino in rappresentazioni sensoriali o segnali sensoriali, che chiameremo ad esempio “ $x$ ”. Entrambi le variabili  $s$  e  $x$  inoltre possono essere variabili scalari o vettoriali.

L'osservatore bayesiano ci permette di conoscere la migliore stima di  $\vartheta$  a partire dalle caratteristiche incerte di  $x$ , massimizzando la probabilità a posteriori  $p(\vartheta | x)$  calcolata con la regola di Bayes, cioè,

$$p(\vartheta | x) = p(x | \vartheta) p(\vartheta) / p(x).$$

$p(x | \vartheta)$  è la funzione probabilità di verosimiglianza, che specifica come vengono generati i segnali sensoriali (e tiene conto delle incertezze), e  $p(\vartheta)$  rappresenta la conoscenza a priori circa le variabili  $\vartheta$ .

Fino a poco tempo, i modelli bayesiani di percezione multisensoriale assumevano che i diversi segnali erano causati dalla stessa fonte e poi modellati e combinati per la stima di un singolo attributo fisico sotto la condizione assunta.

Ad esempio, Ernst e Banche studiarono come soggetti umani riuscissero a stimare la larghezza di un oggetto guardandolo e toccandolo.

Ci potremmo in realtà, immaginare diversi modi per risolvere questo problema. Un approccio non probabilistico potrebbe comportare i seguenti passaggi:

- guardare l'immagine ed estrarre una misura della larghezza della barra,
- toccare la sbarra,
- utilizzare la media delle stime visive e tattili.

Il problema fondamentale dell'esempio precedente consiste nell'utilizzo di pesi uguali, il che non è accettabile in molte situazioni reali. Ad esempio, nella più completa oscurità, qualsiasi stima basata sulla visione rifletterà solo rumore e dovrebbe essere ignorata. Così, piuttosto che pesi uguali, ogni stimolo, che sia acustico o visivo, dovrebbe contribuire alla stima finale in modo proporzionale alla sua affidabilità. Questo è precisamente quello che accadrebbe se adottassimo un approccio probabilistico, ma in questo caso, invece di stimare un valore, ci accorgeremo di avere utilizzato la distribuzione di probabilità di informazioni visive e tattili sulla larghezza della barra.

Ad esempio, nel contesto della localizzazione spaziale, Alais e Burr (2004) hanno utilizzato un modello bayesiano per interpretare il giudizio di localizzazione degli stimoli audiovisivi quando all'osservatore è stato chiesto di considerare ogni presentazione bimodale come evento singolo.

La formulazione bayesiana di questo problema è  $p(\vartheta | x_v, x_a) = p(x_v, x_a | \vartheta) p(\vartheta) / p(x_v, x_a)$ , dove  $x_v$  e  $x_a$  sono rispettivamente, gli stimoli di posizione visivo e uditivo, e  $\vartheta$  è la posizione dell'evento bimodale che deve essere stimato.

Supponendo che il rumore associato a ciascun segnale sensoriale sia indipendente e con distribuzione normale e ipotizzando una distribuzione a priori uniforme, la massima stima a posteriori (MAP) coincide con la stima di massima verosimiglianza ed è la somma degli stimoli uditivi e visivi pesati ognuno per la loro affidabilità (cioè, l'inverso della varianza).

Previsioni di questo modello sono in buon accordo con i risultati psicofisici che dimostrano che la localizzazione bimodale è dominata dallo stimolo visivo o uditivo a seconda di quale dei due è più affidabile. Un modello simile a questo, fu applicato anche da Battaglia et al. (2003) per lo stesso problema di localizzazione degli stimoli bimodali (audio-visivi).

Ernst & Banks (2002) hanno scoperto invece, che i soggetti combinano informazioni visive e tattili secondo la Stima di massima verosimiglianza, quando le due informazioni sono assunte provenire dallo stesso oggetto. In questi esperimenti, i due stimoli sono situati vicini lungo la dimensione di interesse (ad esempio spazio, tempo, struttura) fornendo così una forte indicazione riguardo al fatto di appartenere ad un'unica sorgente o causa. Diversi modelli bayesiani però sono stati sviluppati per tenere conto anche di risultati un po' diversi da questi. È stato proposto, per esempio, un modello di questo tipo proprio da Rowland, Stanford, e Stein (2007a) per interpretare il comportamento dei gatti.

Secondo quanto appena detto, dunque, i modelli Bayesiani sono potenti strumenti per prevedere la combinazione degli stimoli a livello comportamentale. Tuttavia, gran parte di questi, considerano il cervello come una scatola nera e non gli forniscono le adeguate basi meccanicistiche. In particolare, sono di centrale importanza due concetti per tutti i modelli bayesiani, la funzione di verosimiglianza e la conoscenza a priori; quindi capire come queste probabilità sono codificate all'interno del sistema nervoso ha da sempre rappresentato una sfida cruciale. Lavori recenti hanno riportato alcune ipotesi riguardo una possibile inferenza bayesiana del modello sia a livello dei singoli neuroni, sia a livello di popolazioni di neuroni.

Un buon esempio di un modello che descrive un singolo neurone, sulla base di un protocollo Bayesiano, è stato presentato da Patton e Anastasio, in una serie di articoli (Anastasio, Patton, e Belkacem Boussaid, 2000; Patton e Anastasio, 2003; Patton, Belkacem-Boussaid, e Anastasio, 2002). Nel loro primo lavoro (Anastasio et al., 2000), gli autori hanno proposto una teoria probabilistica per spiegare l'enhancement e l'efficacia multisensoriale inversa nei neuroni del collicolo superiore, ipotizzando che quelli collocati negli strati più profondi usino la regola di Bayes per calcolare la probabilità che un determinato obiettivo sia presente nel loro campo recettivo. In uno studio successivo, gli stessi autori hanno fornito una semplice implementazione neurale del modello di Bayes. In particolare, hanno mostrato che un singolo neurone che riceve due ingressi condizionatamente indipendenti con rumore di Poisson ed avente una non linearità sigmoideale (con soglia e saturazione) è in grado di calcolare la probabilità a posteriori mostrando enhancement multisensoriale. Al contrario, se gli ingressi hanno una distribuzione gaussiana multivariata, il calcolo della probabilità a posteriori richiede anche la presenza di nodi moltiplicativi (che calcolano i termini quadratici). Un approccio bayesiano a livello di singolo neurone è stato proposto anche da Colonus e Diederich (2004). Questi autori presumevano che in un individuo, il neurone presente nello strato più profondo rappresenta un'unità computazionale che elabora input e calcola il rapporto mediante la regola di Bayes, che garantisce prestazioni ottimali, cioè massimizza la probabilità di rivelazione del bersaglio minimizzandone il falso tasso di allarme. In questi studi, è stata assunta una distribuzione di Poisson bivariata per input visivi-uditivi, sia per le condizioni di bersaglio e non bersaglio (cioè, sia quando un bersaglio visivo è presente sia nel caso complementare quando l'obiettivo visivo non è presente). Gli autori hanno poi confrontato le prestazioni della rete in condizioni multisensoriali e non, scoprendo che i neuroni multisensoriali, forniscono un rilevamento ottimale del bersaglio in caso di stimoli di ingresso crossmodali, mentre i neuroni specifici di modalità unisensoriale risultano comunque precisi, ma tuttavia necessitano dell'azione congiunta



degli altri. Questo risultato è importante dal momento che spiega la necessità di aver presenti entrambe le modalità di stimolazione all'interno della struttura.

## **Capitolo III – Descrizione del modello matematico che è stato applicato**

Un modello matematico è uno schema espresso in linguaggio matematico volto a rappresentare un fenomeno o un insieme di fenomeni. Lo schema può essere costruito mediante uno dei tanti concetti o teorie della matematica (o una loro combinazione): strutture algebriche o geometriche, equazioni algebriche, differenziali (ordinarie o alle derivate parziali), alle differenze finite, stocastiche, teoria delle probabilità, teoria dei giochi, teoria dei sistemi ecc. Nei confronti dei fenomeni cui si riferisce, il modello matematico può avere una funzione meramente descrittiva ovvero ambire a una descrizione più profonda, al fine di consentire una previsione circa il loro andamento futuro. Tale previsione può limitarsi a delineare questo andamento soltanto in termini qualitativi, oppure determinarlo in termini quantitativi esatti (eventualmente attraverso il calcolo numerico assistito dall'elaboratore elettronico). Inoltre, nei confronti di certe classi di fenomeni, il modello matematico può (o deve) assolvere una funzione prescrittiva o di controllo, ovvero indicare in che modo il fenomeno deve svolgersi al fine di rispondere nel modo più efficace a determinati fini.

Il modello matematico che è stato utilizzato in questo lavoro mira a descrivere e studiare il riconoscimento della posizione audio-visiva degli stimoli multisensoriali che vengono proposti al soggetto in varie combinazioni temporali e spaziali.

### **3.1 Concetti generali di stima bayesiana applicati**

Supponiamo che il cervello elabori due ingressi sensoriali diversi (ad esempio un acustico e uno visivo). Nel seguito, rappresenteremo una grandezza appartenente ad una modalità acustica con l'apice A, mentre uno appartenente alla modalità visiva con l'apice V. Inoltre, le lettere maiuscole saranno utilizzate per rappresentare vettori o matrici, mentre lettere minuscole (eventualmente con

un pedice) saranno utilizzati per rappresentare la componente scalare di vettori. Ogni ingresso sensoriale consiste di un vettore con N componenti , che descrive la distribuzione spaziale dello stimolo. Per esempio, la quantità scalare denota la componente j-esima dell'ingresso acustico.

Assumiamo che ciascun componente j (j = 1, 2, ... N) codifichi un particolare posizione spaziale  $g_j$ .

Quindi, un vettore (NX1) di posizioni spaziali sarà definito come:

$$\Theta = [g_1 \ g_2 \ \dots \ g_j \ \dots \ g_N]^T$$

I due vettori  $I^A$  e  $I^V$  sono le rappresentazioni sensoriali dello stimolo esterno che raggiunge il cervello. Entrambi sono influenzati dalla posizione spaziale dello stimolo di ingresso, dalla sfocatura della trasduzione sensoriale e dal rumore. Indichiamo con  $g^A$  e  $g^V$ , rispettivamente, le posizioni relative agli stimoli acustici e visivi applicati che hanno generato le rappresentazioni sensoriali. In termini di stime bayesiani, il problema consiste nel dedurre le posizioni  $g^A$  e  $g^V$  partendo dalla conoscenza delle due rappresentazioni sensoriali iniziali  $I^A$  e  $I^V$  offuscate e affette da rumore.

Il problema è completamente definito, da un punto di vista statistico, se si conoscono le probabilità di verosimiglianza dei due stimoli sensoriali , e la probabilità a priori delle posizioni. Nel seguito, assumeremo che le rappresentazioni sensoriali  $I^A$  siano solo una funzione della posizione dello stimolo acustico  $g^A$ , mentre  $I^V$  sia solo una funzione di  $g^V$ . Inoltre, essi dipendono anche dalla forza dello stimolo e dalla presenza di rumore, ma assumeremo l'indipendenza reciproca.

Con l'ipotesi di indipendenza, possiamo scrivere la seguente espressione generale per la probabilità di verosimiglianza degli ingressi sensoriali:

$$p(I^A, I^V | g^A, g^V) = p(I^A | g^A) p(I^V | g^V) \quad (1)$$

Vale la pena notare che le due rappresentazioni sensoriali non sono indipendenti ,ovvero  $p(I^A, I^V) \neq p(I^A)p(I^V)$  in quanto  $\mathcal{G}^A$  e  $\mathcal{G}^V$  non sono indipendenti e  $p(\mathcal{G}^A, \mathcal{G}^V)$  non è in genere uniforme,  $p(\mathcal{G}^A, \mathcal{G}^V) \neq p(\mathcal{G}^A)p(\mathcal{G}^V)$ .

Secondo la regola bayesiana, e utilizzando l'eq. (1), possiamo scrivere la seguente espressione per la probabilità a posteriori:

$$p(\mathcal{G}^A, \mathcal{G}^V | I^A, I^V) = \frac{p(\mathcal{G}^A, \mathcal{G}^V)p(I^A, I^V | \mathcal{G}^A, \mathcal{G}^V)}{p(I^A, I^V)} = \frac{p(\mathcal{G}^A, \mathcal{G}^V)p(I^A | \mathcal{G}^A)p(I^V | \mathcal{G}^V)}{p(I^A, I^V)} \quad (2)$$

Per avere la stima migliore, dobbiamo massimizzare il numeratore dell'eq. (2), a partire dalla conoscenza delle rappresentazioni sensoriali  $I^A$  e  $I^V$ . In altre parole, le stime (dette rispettivamente  $\hat{\mathcal{G}}_A$  e  $\hat{\mathcal{G}}_V$ ), devono soddisfare la seguente regola:

$$[\hat{\mathcal{G}}_A, \hat{\mathcal{G}}_V] = \arg \max \{p(\mathcal{G}^A, \mathcal{G}^V | I^A, I^V)\} = \arg \max \{p(\mathcal{G}^A, \mathcal{G}^V)p(I^A | \mathcal{G}^A)p(I^V | \mathcal{G}^V)\} \quad (3)$$

Per massimizzare Eq. (2), occorre conoscere la funzione di verosimiglianza e la probabilità a priori.

### 3.2 La funzione di verosimiglianza

Supponiamo ora di conoscere un certo insieme di input sensoriali (di seguito considereremo un generico ingresso sensoriale S, sia con  $S = A$  o  $S = V$ ). Quindi, possiamo scrivere:

$$I^S = [i_1^S \ i_2^S \ \dots \ i_j^S \ \dots \ i_N^S]^T$$

che rappresenta un vettore  $N \times 1$ .

Questo non è altro che l'insieme degli stimoli di ingresso che raggiungono il cervello. Supponiamo poi che la distribuzione a priori di probabilità sia uniforme.

In questo caso, l'Eq. (2) si semplifica:

$$\hat{g}_s = \arg \max \{p(I^S | g^S)\}$$

cioè, bisogna massimizzare la funzione di verosimiglianza.

Nel seguito, assumeremo che l'input sensoriale sia composto da un termine deterministico (chiamato  $M^S(g^S)$  dipendente dalla forza dello stimolo e dalla sua posizione) su cui è sovrapposto il rumore gaussiano bianco a media nulla ( $N^S$ ). Quindi, possiamo scrivere la seguente espressione per la variabile casuale  $I^S$ :

$$\underline{I}^S = M^S + \underline{N}^S$$

o, in forma scalare,:

$$\underline{i}_j^S = m_j^S + \underline{n}_j^S \quad j = 1, 2, \dots, N \quad (5)$$

dove la sottolineatura indica il fatto che tale quantità è casuale.

Se i termini relativi al rumore sono stati generati indipendentemente, anche  $\underline{i}_j^S$  sono variabili indipendenti, dunque per la funzione di verosimiglianza vale la seguente espressione:

$$p(I^s | \mathcal{G}^s) = \prod_{j=1}^N p(i_j^s | \mathcal{G}^s) \quad (6)$$

Il termine deterministico visibile nell'Eq. (5) è un parametro che è funzione della forza dello stimolo (maggiore è la forza, maggiore è la  $i_j^s$ ), e della posizione dello stimolo. In particolare, la componente j-esima di un input sensoriale dovrebbe essere massima quando  $\mathcal{G}_j = \mathcal{G}^s$ , mentre l'ingresso dovrebbe progressivamente diminuire con la distanza. Abbiamo usato una funzione gaussiana per rappresentare le proprietà spaziali deterministiche dell'ingresso.

Possiamo quindi scrivere:

$$m_j^s(\mathcal{G}^s) = i_{Max}^s \exp \left\{ - \frac{d(\mathcal{G}^s, \mathcal{G}_j)^2}{2\sigma^s} \right\} \quad j = 1, 2, \dots, N \quad (7)$$

dove  $d(\mathcal{G}^s, \mathcal{G}_j)$  rappresenta la distanza tra la posizione dello stimolo  $\mathcal{G}^s$  e la posizione effettiva  $\mathcal{G}_j$ .

Invece  $\sigma^s$  è la deviazione standard della funzione gaussiana e qui rappresenta la precisione spaziale dell'ingresso (la  $\sigma^s$  è maggiore quanto più sfocato è lo stimolo) e  $i_{Max}^s$  rappresenta la forza del stimolo esterno. Secondo quanto scritto nell' eq. (7), quando  $d(\mathcal{G}^s, \mathcal{G}_j) = 0$  l'input sensoriale è massimo (non considerando l'effetto del rumore) mentre l'ingresso sensoriale diminuisce progressivamente con la distanza. Nel presente lavoro, la distanza è stata calcolata attraverso un struttura circolare, in modo che ogni input sensoriale riceva un eccitazione simile, indipendentemente che sia esso vicino o lontano dal bordo. Per calcolare la distanza è stata utilizzata la seguente espressione:

$$d(\mathcal{G}^s, \mathcal{G}_j) = \begin{cases} |\mathcal{G}_j - \mathcal{G}^s| & \text{if } |\mathcal{G}_j - \mathcal{G}^s| \leq D/2 \\ D - |\mathcal{G}_j - \mathcal{G}^s| & \text{if } |\mathcal{G}_j - \mathcal{G}^s| > D/2 \end{cases} \quad (8)$$

dove D rappresenta la distanza spaziale complessiva (cioè  $0 < \mathcal{G}_j < D$ ).

A titolo di esempio, assumendo  $D = 180^\circ$ , la posizione  $\vartheta^s = 1^\circ$  è equidistante dalla posizione  $180^\circ$  e dalla posizione  $2^\circ$ ; inoltre, è ugualmente equidistante anche dalla posizione  $179^\circ$  e dalla posizione  $3^\circ$ , etc. Poiché abbiamo ipotizzato che il rumore ha una distribuzione gaussiana con valor medio nullo, la funzione di verosimiglianza  $p(i_j^s | \vartheta^s)$  assumerà la seguente espressione:

$$\begin{aligned}
 p(i_j^s | \vartheta^s) &= \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp \left\{ - \frac{[i_j^s - m_j^s(\vartheta^s)]^2}{2\sigma_N^2} \right\} = \\
 &= \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp \left\{ - \frac{[i_j^s - i_{Max}^s \exp(- (d(\vartheta^s, \vartheta_j))^2 / (2\sigma^s)^2)]^2}{2\sigma_N^2} \right\} \quad j = 1, 2, \dots, N
 \end{aligned} \tag{9}$$

dove  $\sigma_N$  rappresenta la deviazione standard del rumore (quindi, maggiore è  $\sigma_N$  maggiore sarà l'effetto del rumore).

Infine, utilizzando le eq. (6) e (9) insieme, possiamo scrivere l'espressione generale della funzione di verosimiglianza per l'input sensoriale utilizzato in funzione della posizione di stimolo:

$$p(I^s | \vartheta^s) = \prod_{j=1}^N p(i_j^s | \vartheta^s) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp \left\{ - \frac{[i_j^s - i_{Max}^s \exp(- (d(\vartheta^s, \vartheta_j))^2 / (2\sigma^s)^2)]^2}{2\sigma_N^2} \right\} \tag{10}$$

Vale la pena notare che Eq. (7) e (10) non sono solo in funzione della posizione di stimolo  $\vartheta^s$ , ma anche della forza di ingresso (cioè, dai parametri  $i_{Max}^s$ ). Tuttavia, per semplicità quest'ultima dipendenza non è stata espressamente indicata nel membro di sinistra.

La funzione di verosimiglianza è data dall'espressione (10), in cui viene utilizzato uno specifico valore di  $I^S$  (una singola realizzazione del vettore casuale); quindi:

$$l(\mathcal{G}^S) = p(I^S | \mathcal{G}^S) \quad \text{con } I^S \text{ conosciuto}$$

Calcoliamo il logaritmo naturale della funzione di verosimiglianza.

Dall' Eq. (11) abbiamo:

$$\ln(l(\mathcal{G}^S)) = -\sum_{j=1}^N \ln(\sqrt{2\pi\sigma_N^2}) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N \left\{ i_j^S - i_{Max}^S \exp\left[-\frac{(d(\mathcal{G}^S, \mathcal{G}_j))^2}{2\sigma^{S^2}}\right] \right\}^2 \quad (12)$$

La stima di massima verosimiglianza è ottenuta calcolando il valore ( $\hat{\mathcal{G}}^S$ ) che massimizza l' Eq. (10). Equazioni simili valgono per il calcolo della funzione di verosimiglianza nel caso dello stimolo visivo (S = V) e acustico (S = A).

### 3.3 Un modello di rete neurale per la stima di massima verosimiglianza

Il problema è ora quello di trovare un semplice modello di rete neurale, costituito da N neuroni, che possano essere addestrati con una regola che permetta di calcolare automaticamente il massimo dell'eq. (12), dalla conoscenza del vettore di ingresso  $I^S$ . In seguito, ciascun neurone sarà rappresentato attraverso l'indice k. Consideriamo che ogni neurone nella catena abbia una posizione preferita,  $\mathcal{G}_k$  (k = 1, 2, ..., N), cioè, utilizziamo le stesse posizioni precedentemente utilizzati per il vettore sensoriale, questa volta per identificare i nostri neuroni. Ciò può essere ottenuto utilizzando un campo recettivo, per ciascun neurone, centrato nella posizione preferita. Indicheremo ciascun campo recettivo come  $R_k^S$  (k = 1, 2, ..., N); si tratta di un vettore di dimensioni Nx1. L'ingresso al



neurone k-esimo ( $u_k^S$ ) viene calcolato come il prodotto scalare tra l'input sensoriale e il suo campo recettivo.

Possiamo scrivere:

$$u_k^S = \sum_{j=1}^N r_{kj}^S i_j^S \quad (13)$$

Ciascun neurone quindi calcola la sua attività di uscita ( $y_k^S$ ) facendo passare l'ingresso attraverso una funzione monotona crescente non lineare (che imita la presenza di una soglia inferiore e superiore di saturazione dei neuroni). Indicando questa funzione monotona con  $\varphi(u)$  possiamo scrivere:

$$y_k^S = \varphi(u_k^S) = \varphi\left(\sum_{j=1}^N r_{kj}^S i_j^S\right) \quad (14)$$

Nel seguito useremo una funzione sigmoideale, come abitualmente fatto nella modellizzazione delle reti neurali. Tuttavia, per le presenti considerazioni, abbiamo solo bisogno che  $\varphi(u)$  sia monotona crescente. Siamo ora in grado di dimostrare che, al fine di calcolare la stima di massima verosimiglianza, abbiamo bisogno che:

- i. tutti i neuroni abbiano un campo recettivo identico che differisca solo per la posizione preferita;
- ii. dopo l'addestramento il campo recettivo riproduca la riproduzione spaziale dell'input sensoriale.

cioè:

$$r_{kj}^S = r_{\max}^S \exp\left[-\frac{(d(\mathcal{G}_k, \mathcal{G}_j))^2}{2\sigma^S}\right] \quad j = 1, 2, \dots, N \quad (15)$$

Nel prossimo paragrafo analizzeremo come l'eq.(15) possa essere realizzata utilizzando la regola formativa conosciuta sotto il nome di regola di Oja. Se facciamo uso dell'eq. (15) e della (12) otteniamo la funzione di verosimiglianza logaritmica nella posizione  $\mathcal{G}_k$  ( $k = 1, 2, \dots, N$ ):

$$\begin{aligned}
\ln(l(\mathcal{G}_k)) &= -\sum_{j=1}^N \ln\left(\sqrt{2\pi\sigma_N^2}\right) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N \left\{ i_j^S - i_{Max}^S \exp\left[-\frac{(d(\mathcal{G}_k, \mathcal{G}_j))^2}{2\sigma_N^2}\right] \right\}^2 = \\
&= -\sum_{j=1}^N \ln\left(\sqrt{2\pi\sigma_N^2}\right) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N \left[ i_j^S - \frac{i_{Max}^S}{r_{max}^S} r_{kj}^S \right]^2 = \\
&= -\sum_{j=1}^N \ln\left(\sqrt{2\pi\sigma_N^2}\right) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N i_j^{S^2} - \frac{1}{2\sigma_N^2} \left(\frac{i_{Max}^S}{r_{max}^S}\right)^2 \sum_{j=1}^N r_{kj}^{S^2} + \frac{1}{\sigma_N^2} \frac{i_{Max}^S}{r_{max}^S} \sum_{j=1}^N i_j^S r_{kj}^S
\end{aligned} \tag{16}$$

Naturalmente, i primi due termini del membro di destra dell'eq.(16) sono indipendenti da k.

Tuttavia in virtù dell'assunzione i) precedentemente fatta, possiamo affermare che:

$\sum_{j=1}^N r_{kj}^2$  è anche indipendente da k (in particolare, si usa una distanza circolare nell'espressione dei campi recettivi, al fine di evitare qualsiasi effetto di bordo).

Pertanto, possiamo scrivere:

$$\ln(l(\mathcal{G}_k)) = \alpha + \frac{1}{\sigma_N^2} \frac{i_{Max}^S}{r_{max}^S} \sum_{j=1}^N r_{kj}^S i_j^S \tag{17}$$

dove  $\alpha$  rappresenta la somma dei primi tre termini nel membro di destra dell'eq (16), che non dipende da un valore particolare di  $\mathcal{G}_k$ . Di conseguenza, al fine di massimizzare la funzione di verosimiglianza logaritmica, dobbiamo semplicemente massimizzare la quantità:

$$u_k^S = \sum_{j=1}^N r_{kj}^S i_j^S$$

Infine, ricordando che la (14) è una funzione monotona crescente, abbiamo:

$$\hat{g}^s = \arg \max \{ \ln (I(g_k)) \} = \arg \max \{ u_k^s \} = \arg \max \{ y_k^s \} \quad (18)$$

L'eq. (18) ci dice che il neurone con massima attività segnala la posizione dello stimolo, secondo una stima di massima verosimiglianza. Per questo motivo, nel seguito si supporrà che la posizione dello stimolo venga *codificata dal neurone con massima attività*. Affinché l'eq.(8) sia verificata, devono essere verificate le assunzioni i) e ii ), concernenti i campi recettivi,. Per fare in modo che ciò avvenga, occorrono alcune condizioni.

L' assunzione i) implica che la posizione preferita dei neuroni abbia una distribuzione uniforme (cioè le posizioni siano ugualmente rappresentate nella rete) e che, durante la formazione del campo recettivo, tutte le posizioni siano stimolati dal ingresso esterno nello stesso modo (cioè, nessuna posizione riceva un ingresso più forte di un altro o un ingresso più frequente). L' assunzione ii) invece implica che il campo recettivo, dopo l'apprendimento, riproduca la distribuzione spaziale media dell'ingresso nella data posizione preferita ( $r_{kj}^s \propto m_j^s(g_k)$ ). Quest'ultimo requisito può essere raggiunto mediante la regola di Oja.

Questa regola può essere scritta come segue:

$$\Delta R_k^s = \gamma y_k^s (I^s - R_k^s) \quad (19)$$

dove  $\Delta R_k^s$  rappresenta la variazione di campo recettivo dopo che un neurone ha ricevuto uno stimolo. Secondo quanto scritto nell' eq.(19), un neurone ad elevata attività di uscita può modificare il suo campo recettivo spostandolo verso l'ingresso effettivo, perdendo però parte dei valori precedenti. Al contrario, i neuroni silenziosi con scarsa attività di uscita non modificano in modo

apprezzabile il loro campo recettivo. Dopo un lungo addestramento, il campo vettoriale recettivo  $R_k^S$  sarà posizionato vicino al *baricentro* degli ingressi che attivano in modo significativo il neurone. Nel nostro caso, questo baricentro è proporzionale a  $M^S(\mathcal{g}_k)$ , come previsto dall'eq. (7) in cui  $\mathcal{g}^S = \mathcal{g}_k$  (lo stimolo che meglio attiva il neurone k-esimo) e utilizzando il valore medio della resistenza di ingresso  $i_{Max}^S$ .

Quindi, abbiamo:

$$R_k^S \propto M^S(\mathcal{g}_k) \quad (20)$$

che non è altro che la forma vettoriale dell'eq.(15).

Affiché si attui l'eq.(20), si deve però verificare nella rete neurale una dinamica del tipo “il vincitore prende tutto” (“winner takes all”), in modo che solo pochi neuroni siano in grado di vincere la concorrenza, mentre la maggior parte dei neuroni rimangano in silenzio. In questo modo, un neurone modifica le sinapsi solo quando l'ingresso è realmente vicino alla sua posizione spaziale preferita.

Per questo motivo, come solito si usa fare quando si lavora con questo tipo di reti, abbiamo introdotto sinapsi laterali in modo da attuare una competizione tra i neuroni.

### 3.4 La probabilità a posteriori

Consideriamo ora il caso in cui i due ingressi sensoriali ( $I^A$  e  $I^V$ ) non siano indipendenti, come nel caso precedente, ma collegati dalla probabilità a priori  $p(\mathcal{g}^A, \mathcal{g}^V)$ . In questo caso, si deve massimizzare il numeratore dell'eq. (2), che può essere riscritta in forma logaritmica.

Abbiamo allora:

$$\begin{aligned} [\hat{\mathcal{g}}_A, \hat{\mathcal{g}}_V] &= \arg \max \{ p(\mathcal{g}^A, \mathcal{g}^V) p(I^A | \mathcal{g}^A) p(I^V | \mathcal{g}^V) \} = \\ &= \arg \max \{ \ln(p(\mathcal{g}^A, \mathcal{g}^V)) + \ln(p(I^A | \mathcal{g}^A)) + \ln(p(I^V | \mathcal{g}^V)) \} \end{aligned} \quad (21)$$

Naturalmente, nel caso in cui la densità di probabilità a priori fosse uniforme, l'eq.(21) sarebbe la stessa della stima della verosimiglianza eseguita separatamente su  $\mathcal{G}^A$  e  $\mathcal{G}^V$ . Nel seguito, indicheremo la funzione da massimizzare  $\psi(\mathcal{G}^A, \mathcal{G}^V)$  come:

$$\psi(\mathcal{G}^A, \mathcal{G}^V) = \ln(p(\mathcal{G}^A, \mathcal{G}^V)) + \ln(p(I^A | \mathcal{G}^A)) + \ln(p(I^V | \mathcal{G}^V)) \quad (22)$$

Consideriamo ora che le funzioni di verosimiglianza seguino l'espressione (7) usata sopra, ma con parametri diversi per il campo recettivo a seconda delle due diverse modalità di stimolazione usata (in particolare, si assume che i due stimoli  $\sigma^A$  e  $\sigma^V$  abbiano accuratezza spaziale diversa e  $i_{Max}^A \neq i_{Max}^V$  (cioè, i due stimoli possano avere una forza diversa). Utilizzando l'eq.(16) entro l'eq.(22), possiamo calcolare la seguente espressione per la funzione  $\psi(\mathcal{G}^A, \mathcal{G}^V)$ , valutata in due posizioni diverse (una posizione  $\mathcal{G}_k^A$  per il k-esimo neurone uditivo e una posizione  $\mathcal{G}_h^V$  per l'h-esimo neurone visivo):

$$\begin{aligned} \psi(\mathcal{G}_k^A, \mathcal{G}_h^V) = & - \sum_{j=1}^N \ln(\sqrt{2\pi\sigma_N^2}) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N i_j^{A^2} - \frac{1}{2\sigma_N^2} \left( \frac{i_{Max}^A}{r_{max}^A} \right)^2 \sum_{j=1}^N r_{kj}^{A^2} + \frac{1}{\sigma_N^2} \frac{i_{Max}^A}{r_{max}^A} \sum_{j=1}^N i_j^A r_{kj}^A \\ & - \sum_{j=1}^N \ln(\sqrt{2\pi\sigma_N^2}) - \frac{1}{2\sigma_N^2} \sum_{j=1}^N i_j^{V^2} - \frac{1}{2\sigma_N^2} \left( \frac{i_{Max}^V}{r_{max}^V} \right)^2 \sum_{j=1}^N r_{hj}^{V^2} + \frac{1}{\sigma_N^2} \frac{i_{Max}^V}{r_{max}^V} \sum_{j=1}^N i_j^V r_{hj}^V + \ln(p(\mathcal{G}_k^A, \mathcal{G}_h^V)) \end{aligned} \quad (23)$$

Consideriamo ora che tutti i neuroni in una determinata modalità abbiano campi recettivi identici. In questa condizione, il terzo e il settimo termine dell'eq.(23) risultano indipendenti dai particolari valori di k e h. Quindi, la massimizzazione dell'eq.(23) corrisponde alla massimizzazione della seguente equazione (trascurando anche tutti gli altri termini che non dipendono da k ed h):

$$\frac{1}{\sigma_N^2} \frac{i_{Max}^A}{r_{max}^A} \sum_{j=1}^N i_j^A r_{kj}^A + \frac{1}{\sigma_N^2} \frac{i_{Max}^V}{r_{max}^V} \sum_{j=1}^N i_j^V r_{hj}^V + \ln(p(g_k^A, g_h^V)) \quad (24)$$

Come detto sopra, nel presente lavoro si assume che l'uscita massima dalla rete uditiva segnali la posizione dello stimolo uditivo, mentre l'uscita massima dalla rete visiva segnali la posizione di quello visivo. Tuttavia, se si assume che l'uscita del neurone dipenda solo dal seguente termine :

$$u_k^S = \sum_{j=1}^N r_{kj}^S i_j^S \quad \text{with } S = A \text{ or } V$$

non si riesce a tenere conto della probabilità a priori (cioè del terzo termine nell'eq.(24)). Come risulta dall'eq.(24), la stima dei valori ottimali richiede che gli ingressi ai neuroni uditivi e visivi vengano modificati, per tenere conto appunto proprio di questa probabilità a priori. In altre parole, sia i neuroni uditivi che quelli visivi devono ricevere un termine di "cross-talk" dall' altra modalità, che rifletta la conoscenza a priori. Una condizione tipica si verifica quando i due stimoli (quello uditivo e visivo) provengono da uno stesso evento spaziale, quindi le due posizioni in questo caso sono correlate. Supponendo che siano possibili piccole distanze, possiamo utilizzare la seguente espressione per la probabilità a priori:

$$p(g^A, g^V) = \beta_1 p'(g^A, g^V) + \beta_2 p''(g^A, g^V) \quad (26)$$

Secondo quanto riportato nell'eq.(25), la probabilità a priori è data dalla somma ponderata di una distribuzione uniforme ( $p'(g^A, g^V)$ ), che riflette la possibilità che uno stimolo visivo ed uno uditivo

siano prodotti da eventi diversi, e un secondo termine ( $p''(g^A, g^V)$ ) che riflette la probabilità (più forte) che gli eventi uditivi e visivi siano stati originati dalla stessa fonte.

Possiamo scrivere:

$$p'(g^A, g^V) = \frac{1}{D^2} \quad (\text{distribuzione uniforme}) \quad (26)$$

$$p''(g^A, g^V) = p(g^A)p''(g^V|g^A) = \frac{1}{D} \frac{1}{\sqrt{2\pi\sigma^{AV^2}}} \exp\left(-\frac{d(g^A, g^V)^2}{2\sigma^{AV^2}}\right) \quad (27)$$

Occorre precisare che l'eq.(27) è stata scritta assumendo che la singola posizione abbia una distribuzione uniforme (ovvero  $p(g^A) = 1/D$ ); la probabilità della seconda posizione, nel caso ci sia una singola fonte per entrambi gli stimoli, diminuisce drasticamente con la distanza. Il parametro  $\sigma^{AV}$  riflette l'accuratezza spaziale della sovrapposizione dei due stimoli, quando provengono dalla stessa fonte. Naturalmente, l'eq.(25) integrata sull'intero spazio delle possibili posizioni (ad esempio, tra 0 e D), deve soddisfare l'assioma fondamentale della probabilità.

A tal fine, deve essere utilizzato il seguente vincolo:  $\beta_1 + \beta_2 = 1$ , quindi  $\beta_2 = 1 - \beta_1$ .

Si ottiene così:

$$p(g^A, g^V) = \beta_1 \frac{1}{D^2} + (1 - \beta_1) \frac{1}{D} \frac{1}{\sqrt{2\pi\sigma^{AV^2}}} \exp\left(-\frac{d(g^A, g^V)^2}{2\sigma^{AV^2}}\right) \quad (28)$$

Il parametro  $\beta_1$  rappresenta la frazione di stimoli cross-modali provenienti da fonti indipendenti. Al contrario,  $1 - \beta_1$  rappresenta invece, la frazione di stimoli cross-modali provenienti da una singola fonte.

### 3.5 Descrizione della rete neurale

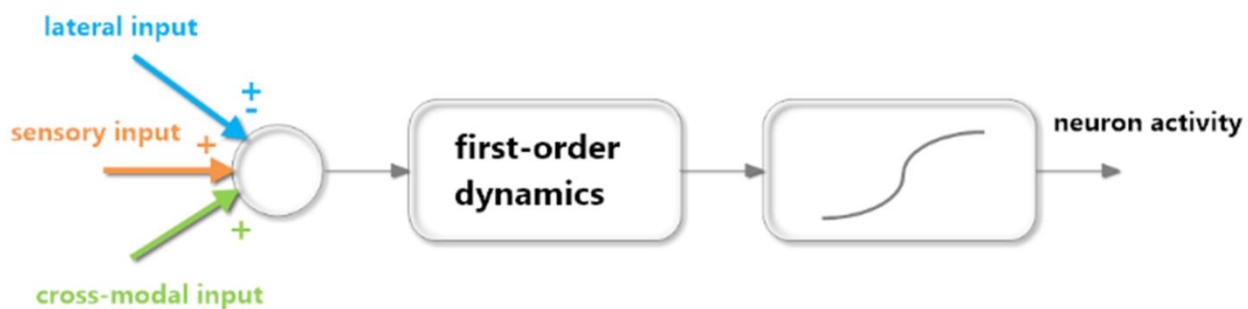


Figura 8 Ogni neurone è descritto con un dinamica lineare del primo ordine e una sigmoide. Ogni neurone riceve tre tipi di ingressi: l'input sensoriale (attraverso il suo campo recettivo), un ingresso laterale da altri neuroni nella stessa area, un ingresso cross-modale dai neuroni in altra area.

#### 3.5.1 Struttura base della rete

Il modello di rete neurale che è stato utilizzato in questa tesi consiste in due catene di  $N$  neuroni unisensoriali. Ogni neurone codifica per una particolare posizione spaziale nella sua modalità prestabilita. Inoltre, ogni catena è topologicamente organizzata, ovvero, i neuroni prossimali vengono considerati vicini e così via. Nel seguito, indicheremo con un apice una particolare area (audio o visiva) e con un indice la posizione del neurone all'interno di quell'area. Ogni neurone riceve tre diversi tipi di ingressi: un input sensoriale dall'ambiente (che chiameremo  $u$ ), un ingresso laterale dai neuroni della stessa modalità (chiamato  $l$ ) e un ingresso cross-modale dai neuroni dell'altra modalità (che per esempio indicheremo con  $c$ ). L'ingresso globale (pari alla somma dei precedenti tre contributi) viene fatto passare attraverso una relazione di tipo sigmodale per riprodurre



la presenza di una soglia inferiore e di una possibile saturazione superiore dell'attività neuronale, e attraverso un filtro passa-basso del primo ordine con costante di tempo  $\tau$ , che rappresenta la capacità integrativa del neurone. Quindi, per un generico neurone  $k$ -esimo nella modalità  $S$  ( $S = A$  o  $V$  per le modalità uditive e visive, rispettivamente) possiamo scrivere:

$$\tau \frac{dy_k^S}{dt} = -y_k^S + \phi(u_k^S + c_k^S + e_k^S) \quad (29)$$

Dove  $y_k^S$  rappresenta l'uscita del neurone, e la relazione sigmoideale è descritta dalla seguente equazione:

$$\phi(x) = \frac{1}{1 + \exp(k(x - x_0))} \quad (30)$$

$k$  e  $x_0$  sono parametri, che stabiliscono la pendenza e la posizione della relazione sigmoideale. Secondo l'eq.(30), l'attività di uscita del neurone è normalizzata tra 0 e 1 (dove il numero zero significa un neurone silenzioso, mentre il numero 1 indica un neurone massimamente attivato).

Vale la pena notare che, per semplicità, abbiamo utilizzato gli stessi parametri  $x$ ,  $k$  and  $x_0$ ) per tutti i neuroni indipendentemente dalla loro modalità visiva o uditiva che sia. Questa scelta è stata fatta per cercare di mantenere il numero di assunzioni del modello ad un numero minimo possibile.

L'espressione degli input sensoriali è stata calcolata come il prodotto scalare tra la rappresentazione sensoriale dello stimolo ( $I^S = [i_1^S \ i_2^S \ \dots \ i_k^S \ \dots \ i_N^S]^T$ ) ed il campo recettivo del neurone (

$$R_k^S = [r_{k1}^S \ r_{k2}^S \ \dots \ r_{kj}^S \ \dots \ r_{kN}^S]^T):$$

$$u_k^S = \sum_{j=1}^N r_{kj}^S i_j^S \quad (31)$$

Abbiamo ipotizzato inoltre, che il campo recettivo del neurone abbia inizialmente una grande estensione, descritta con una funzione gaussiana, e poi progressivamente si restringa durante

l'addestramento, in base alla larghezza della ingresso esterno (vedere il paragrafo successivo "Addestramento del modello").

L'ingresso laterale è calcolata come segue:

$$I_k^S = \sum_{j=1}^N v_{kj} y_j^S \quad (32)$$

Dove  $v_{kj}$  rappresenta una sinapsi laterale intra-area che collega il neurone presinaptico  $j$  al neurone postsinaptico  $k$  nella stessa zona. Qui abbiamo utilizzato la disposizione “a cappello messicano” classica: un neurone è eccitato dai neuroni prossimali nella stessa zona, e inibito da quelli più distali.

$$v_{kj} = v_{ex} \exp\left(-\frac{\delta(\mathcal{g}_j, \mathcal{g}_k)^2}{2\sigma_{ex}^2}\right) - v_{in} \exp\left(-\frac{\delta(\mathcal{g}_j, \mathcal{g}_k)^2}{2\sigma_{in}^2}\right) \quad (33)$$

dove  $v_{ex}, v_{in}, \sigma_{ex}, \sigma_{in}$  sono parametri che fissano la forza e la larghezza della porzione eccitatoria e inibitoria del cappello messicano. In particolare abbiamo  $v_{ex} > v_{in}$  and  $\sigma_{ex} < \sigma_{in}$ .

Inoltre  $d(\mathcal{g}_j, \mathcal{g}_k)$  rappresenta la distanza, calcolata come segue:

$$d(\mathcal{g}_j, \mathcal{g}_k) = \begin{cases} |\mathcal{g}_j - \mathcal{g}_k| & \text{if } |\mathcal{g}_j - \mathcal{g}_k| \leq D/2 \\ D - |\mathcal{g}_j - \mathcal{g}_k| & \text{if } |\mathcal{g}_j - \mathcal{g}_k| > D/2 \end{cases} \quad (34)$$

Vale la pena notare che abbiamo usato la stessa espressione di sinapsi laterali (vedi eq.31) sia nelle aree uditive che in quelle visive, per cercare di limitare il numero di assunzioni del modello.

Infine, il termine cross-modale nell'eq.(29) viene calcolato come la convoluzione del vettore delle sinapsi modali trasversali e l'attività nell'altra zona unisensoriale, cioè:

$$c_k^S = \sum_{j=1}^N w_{kj}^{SQ} y_j^Q \quad \text{with } S = A \text{ or } V \quad Q = A \text{ or } V \quad \text{with } S \neq Q \quad (34)$$

dove  $w_{kj}^{SQ}$  rappresenta le sinapsi cross-modali che partono dal neurone pre-sinaptico  $j$  nell'area  $Q$  al neurone post-sinaptico  $k$  nella zona  $S$ . Ricordiamo che si presumono le sinapsi cross-modali inizialmente nulle e poi apprese progressivamente durante la fase di addestramento.

### 3.5.2 Addestramento della rete

Partendo dal valore base iniziale delle sinapsi, la rete è stata addestrata attraverso un periodo di formazione in cui le rappresentazioni degli input sensoriali (ad esempio,  $I^A$  e  $I^V$ ) sono state date con una distribuzione casuale. In particolare, abbiamo ipotizzato che gli ingressi sensoriali siano composti da un termine deterministico, che rappresenta la distribuzione spaziale dell'ingresso centrata nella posizione spaziale dello stimolo, e da un termine relativo al rumore bianco gaussiano (con valor medio nullo e deviazione standard assegnata). Quindi:

$$i_k^S = i_{Max}^S \exp \left( - \frac{(d(\mathcal{G}^S, \mathcal{G}_k))^2}{2\sigma^{S^2}} \right) + n_k^S \quad \text{with } S = A, V \quad (35)$$

dove  $\mathcal{G}^S$  rappresenta la posizione spaziale dello stimolo,  $i_{Max}^S$  è la forza dello stimolo (pari al valore dello stimolo nella sua posizione centrale in assenza di rumore) e  $\sigma^S$  è la deviazione standard della rappresentazione spaziale. Secondo la fisiologia, abbiamo ipotizzato che gli input visivi siano spazialmente più accurati di quelli uditivi, quindi abbiamo impostato  $\sigma^V < \sigma^A$ . Inoltre, abbiamo ipotizzato che la deviazione standard del rumore ( $\sigma_N$ ) sia una data frazione della forza di ingresso.

Le posizioni dei due stimoli ( $\mathcal{G}^A$  e  $\mathcal{G}^V$  vedi eq.35) sono state generate in maniera casuale dalla

distribuzione della probabilità a priori indicate dall'eq.28 attraverso la scelta di diversi valori per il parametro  $\beta_l$  .

Le sinapsi che descrivono il campo recettivo  $r_{kj}^S$  , e quelle che descrivono il collegamento cross-modale tra le due aree  $w_{kj}^{SQ}$  sono state addestrate utilizzando la regola Oja.

Possiamo scrivere, in forma scalare:

$$\Delta r_{kj}^S = \gamma y_k^S (i_j^S - r_{kj}^S) \quad \text{with } S = A, V \quad (36)$$

$$\Delta w_{kj}^{SQ} = \gamma y_k^S (y_j^Q - w_{kj}^{SQ}) \quad \text{with } S = A, V \quad (37)$$

Le eq.(36) e (37) sono state applicate, ad ogni passo, utilizzando i valori finali di equilibrio del neurone di uscita (cioè, quando i fenomeni transitori fossero esauriti).

All'inizio dell'addestramento tutte le sinapsi cross-modali sono assunte pari a zero. Viceversa, le sinapsi che descrivono i campo-ricettivi hanno un'ampia estensione spaziale e ampiezza moderata, identica per le due modalità, ovvero:

$$r_{kj}^S = r_0 \exp \left( - \frac{(d(\mathcal{G}_j, \mathcal{G}_k))^2}{2\sigma_R^2} \right) \quad \text{with } S = A, V \quad (38)$$

dove  $r_0$  rappresenta la forza iniziale del campo recettivo e  $\sigma_R$  la sua estensione spaziale (noi assumeremo  $\sigma_R > \sigma_A$  e  $\sigma_R > \sigma_V$  ,ovvero, supponiamo un alto valore di campi recettivi iniziali) .

Naturalmente l'eq.(38) varrà solamente al primo passo di addestramento.

Nelle figure qui sotto sono rappresentate le due tipologie di stimolazioni utilizzate, in ordine quella unimodale e quello multimodale:

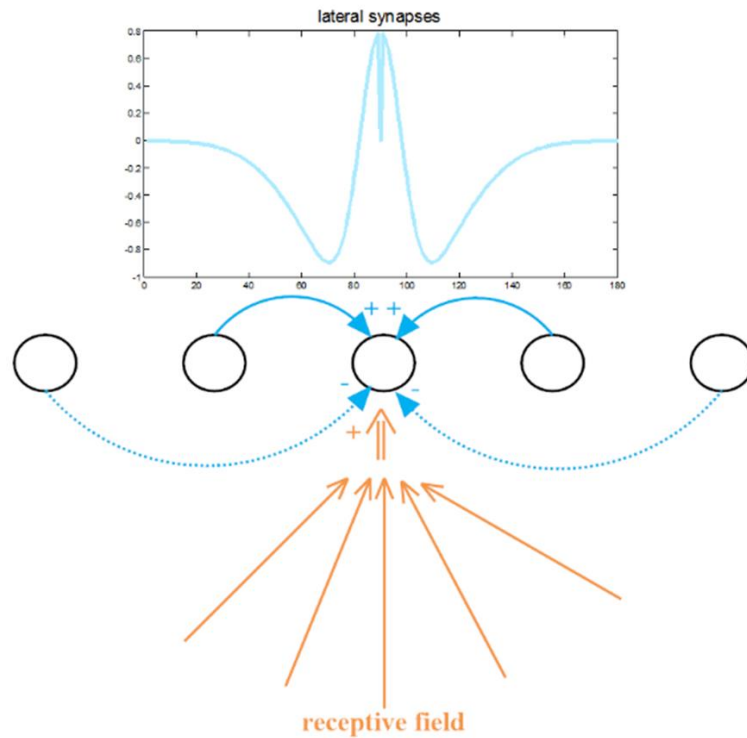


Figura 9 modello di stimolazione unimodale. Ogni neurone codifica per una posizione spaziale, le sinapsi laterali sono addestrate tramite la regola “a cappello messicano” e le sinapsi in ingresso che attuano il campo recettivo, sono formate sulla base della regola Oja.

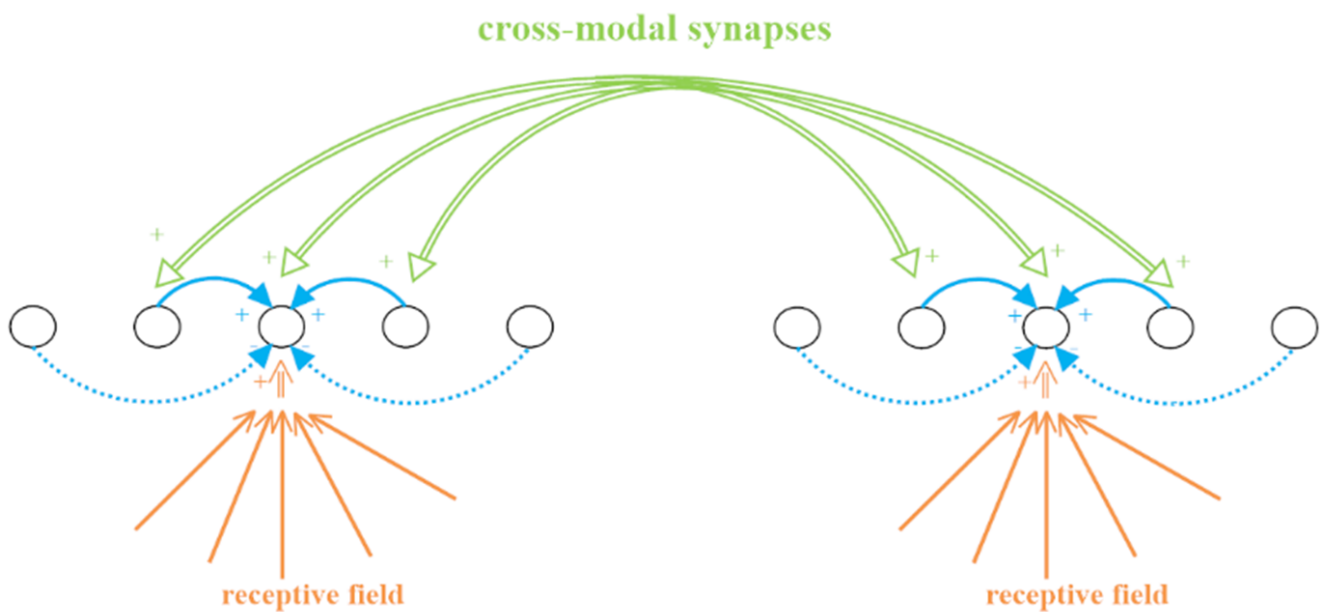


Figura 10 modello di stimolazione multimodale con la presenza di sinapsi cross-modali generate attraverso la regola di Oja.

### 3.6 Modello matematico in Matlab

Il programma che è stato utilizzato per addestrare la rete al comportamento voluto, prende il nome di “addestra\_uno\_ogni\_tre\_random”. Questa versione di programma, genera in particolare 2/5 di input solo visivi, 2/5 di input solo acustici e 1/5 di stimoli crossmodali nella stessa posizione. Dunque solamente uno stimolo ogni tre risulta essere di tipo cross-modale. Sono stati utilizzati stimoli random.

```
% a acustici; v visivo
```

```
global sigma_a sigma_v I0_a I0_v sigma_na sigma_nv N Index
```

```
% numero i neuroni
```

```
N = 180;      % numero di neuroni
```

```
Index = [1:N]; % indice che individua ogni neurone
```

```
%caratteristica dei neuroni
```

```
phi=16;
```

```
pend=0.7;
```

```
tau=5;
```

```
Ingressi
```

```
%genero la matrice delle sinapsi laterali
```

```
%(sono disposte in riga; ogni riga è un neurone post sinaptico)
```

```
for i = 1:N,
```

```

DD = abs(Index - i);    % distanza non circolare
D = DD.*(DD <= N/2) + (N - DD).*(DD > N/2); %distanza con calcolo circolare
% sinapsi laterali uditive
L_a(i,:)=Lex0_a*exp(-D.*D/2/sigex/sigex)-Lin0_a*exp(-D.*D/2/sigin/sigin);
L_a(i,i)=0;    %non riceve sinapsi da se stesso
% sinapsi laterali visive
L_v(i,:)=Lex0_v*exp(-D.*D/2/sigex/sigex)-Lin0_v*exp(-D.*D/2/sigin/sigin);
L_v(i,i)=0;    %non riceve sinapsi da se stesso

end

% genero i campi recettori visivi e uditivi (ho messo gli stessi valori di partenza per entrambi)
%(sono disposte in riga; ogni riga è un neurone post sinaptico)
Rec = zeros(N,N);
Rec0 = 1.5;
sig_r = 30;
for l = 1:N,
    DD = abs(Index - l);    % distanza non circolare
    D = DD.*(DD <= N/2) + (N - DD).*(DD > N/2); %distanza con calcolo circolare
    Rec(l,:)=Rec0*exp(-D.*D/2/sig_r/sig_r);
end

Rec_a=Rec;

```

```

Rec_v=Rec;

%Grafico un input visivo e uno uditivo nella stessa posizione prima dell'addestramento

pos_a = 90;

pos_v = 90;

[I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % gli input sono generati come vettori riga

Input_a = Rec_a*I_a'; % calcolo l'input auditivo passato per il campo recettivo

Input_v = Rec_v*I_v'; % calcolo l'input visivo passato per il campo recettivo

Y_a =1./(1+exp(-(Input_a-phi)*pend));

Y_v =1./(1+exp(-(Input_v-phi)*pend));

figure(1)

subplot(2,1,1)

plot(Index,Y_a,'r')

subplot(2,1,2)

plot(Index,Rec_a(90,:),'b',Index,Rec_a(50,:),'r',Index,Rec_a(170,:),'g') %plotto le sinapsi
auditive prima dell'addestramento

figure(2)

subplot(2,1,1)

plot(Index,Y_v,'r')

subplot(2,1,2)

plot(Index,Rec_v(90,:),'b',Index,Rec_v(50,:),'r',Index,Rec_v(170,:),'g') %plotto le sinapsi
visive prima dell'addestramento

pause

```



```
D_Rec_a = zeros(N,N);
D_Rec_v = zeros(N,N);
W_av = zeros(N,N);    % sinapsi da v ad a
W_va = zeros(N,N);    % sinapsi da a a v
DW_av = zeros(N,N);
DW_va = zeros(N,N);
```

```
num_epoche = 40;
```

```
Matrice_rec50_a = zeros(num_epoche,N);
Matrice_rec90_a = zeros(num_epoche,N);
Matrice_rec170_a = zeros(num_epoche,N);
Matrice_rec50_v = zeros(num_epoche,N);
Matrice_rec90_v = zeros(num_epoche,N);
Matrice_rec170_v = zeros(num_epoche,N);
Matrice_Wav50 = zeros(num_epoche,N);
Matrice_Wav90 = zeros(num_epoche,N);
Matrice_Wav170 = zeros(num_epoche,N);
Matrice_Wva50 = zeros(num_epoche,N);
Matrice_Wva90 = zeros(num_epoche,N);
Matrice_Wva170 = zeros(num_epoche,N);
```

```
gate_a(1:5:900)=1;
gate_a(2:5:900)=0;
gate_a(3:5:900)=1;
```

```

gate_a(4:5:900)=0;
gate_a(5:5:900)=1;
gate_v(1:5:900)=0;
gate_v(2:5:900)=1;
gate_v(3:5:900)=0;
gate_v(4:5:900)=1;
gate_v(5:5:900)=1;

dt=0.2;           % passo di campionamento
T_stimolo = 17/dt;   % durata dello stimolo
t=[0:T_stimolo]*dt; % asse dei tempi

LL=length(t);

for epoca = 1:num_epoche,
    epoca

P1 =randi(180,1,180);
P2 =randi(180,1,180);
P3 =randi(180,1,180);
P4 =randi(180,1,180);
P5 =randi(180,1,180);

PP(1:5:900)=P1;
PP(2:5:900)=P2;
PP(3:5:900)=P3;
PP(4:5:900)=P4;
PP(5:5:900)=P5;

```

```

for kk = 1: length(PP),
    pos_a = PP(kk);    %posizione input acustico

    pos_v = PP(kk)+ randn(1,1)*1;    %posizione input visivo, ammetto che il visivo possa
differire di 1-2 grdi dall'acustico

    %genero l'input

    [I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % generato come riga

    Input_a = Rec_a*I_a'*gate_a(kk);          % calcolo l'input auditivo passato per il campo
recettivo (adesso è colonna)

    Input_v = Rec_v*I_v'*gate_v(kk);          % calcolo l'input visivo passato per il campo
recettivo (adesso è colonna)

    Y_a = zeros(N,LL);
    Y_v = zeros(N,LL);

    for kt =1:LL-1,

        Y_a(:,kt+1) = Y_a(:,kt) + (dt/tau)*(-Y_a(:,kt)+1./(1+exp(-
(Input_a+L_a*Y_a(:,kt)+W_av*Y_v(:,kt)-phi)*pend)));

        Y_v(:,kt+1) = Y_v(:,kt) + (dt/tau)*(-Y_v(:,kt)+1./(1+exp(-
(Input_v+L_v*Y_v(:,kt)+W_va*Y_a(:,kt)-phi)*pend)));

    end

    Yfinale_a = Y_a(:,LL);
    Yfinale_v = Y_v(:,LL);

    if (pos_a == 90) && (gate_a(kk)==1)&&(gate_v(kk)==1)

```

```

figure(1)

subplot(2,1,1)

plot(Index,Yfinale_a','b');

title('attività neurone auditivo')

end

if (pos_a == 90)&& (gate_a(kk)==1)&&(gate_v(kk)==1)

figure(2)

subplot(2,1,1)

plot(Index,Yfinale_v','r');

title('attività neurone visivo')

end

% aggiorno i campi recettori visivi e uditivi (regola di Oya)

gamma = 0.004;

gamma1 = 0.004;

for i = 1:N,

    D_Rec_a(i,:) = gamma*Yfinale_a(i)*(I_a-Rec_a(i,:));

    D_Rec_v(i,:) = gamma*Yfinale_v(i)*(I_v-Rec_v(i,:));

% faccio a meno della saturazione Wmax

%    DW_av(i,:) = gamma1*(Wmax-sum(W_av(i,:)))/Wmax*Yfinale_a(i)*(Yfinale_v'-

W_av(i,:));

```

```

% DW_va(i,:) = gamma1*(Wmax-sum(W_va(i,:)))/Wmax*Yfinale_v(i)*(Yfinale_a'-
W_va(i,:));

DW_av(i,:) = gamma1*Yfinale_a(i)*(Yfinale_v'-W_av(i,:));

DW_va(i,:) = gamma1*Yfinale_v(i)*(Yfinale_a'-W_va(i,:));

end

```

```

Rec_a=Rec_a+D_Rec_a;

```

```

Rec_v=Rec_v+D_Rec_v;

```

```

W_av=W_av+DW_av;

```

```

W_va=W_va+DW_va;

```

```

Matrice_rec50_a(epoca,:) = Rec_a(50,:);

```

```

Matrice_rec90_a(epoca,:) = Rec_a(90,:);

```

```

Matrice_rec170_a(epoca,:) = Rec_a(170,:);

```

```

Matrice_rec50_v(epoca,:) = Rec_v(50,:);

```

```

Matrice_rec90_v(epoca,:) = Rec_v(90,:);

```

```

Matrice_rec170_v(epoca,:) = Rec_v(170,:);

```

```

Matrice_Wav50(epoca,:) = W_av(50,:);

```

```

Matrice_Wav90(epoca,:) = W_av(90,:);

```

```

Matrice_Wav170(epoca,:) = W_av(170,:);

```

```

Matrice_Wva50(epoca,:) = W_va(50,:);

```

```

Matrice_Wva90(epoca,:) = W_va(90,:);

```

```

Matrice_Wva170(epoca,:) = W_va(170,:);

```

**end**

**figure(1)**

**subplot(2,1,2)**

**plot(Index,Rec\_a(90,:), 'b', Index, Rec\_a(50,:), 'r', Index, Rec\_a(170,:), 'g')**

**title('campo recettivo acustico')**

**figure(2)**

**subplot(2,1,2)**

**plot(Index,Rec\_v(90,:), 'b', Index, Rec\_v(50,:), 'r', Index, Rec\_v(170,:), 'g')**

**title('campo recettivo visivo')**

**grid**

**figure(3)**

**subplot(2,1,2)**

**plot(Index,W\_av(90,:), 'b', Index, W\_av(50,:), 'r', Index, W\_av(170,:), 'g')**

**title('sinapsi cross-modali entranti acustico')**

**subplot(2,1,1)**

**plot(Index,W\_va(90,:), 'b', Index, W\_va(50,:), 'r', Index, W\_va(170,:), 'g')**

**%pause(0.2)**

**end**

**%confronto i campi recettivi con un input a rumore nullo**

**pos\_a = 90;**

```

pos_v = 90;

sigma_na = 0;

sigma_nv = 0;

[I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % gli input sono generati come vettori riga

figure

plot(Index,Rec_a(90,:),Index,I_a)

figure

plot(Index,Rec_v(90,:),Index,I_v)

save prova_ultima W_av W_va Rec_a Rec_v

save prova_ultima N Matrice_rec50_a Matrice_rec90_a Matrice_rec170_a Matrice_rec50_v
Matrice_rec90_v Matrice_rec170_v -append

save prova_ultima Matrice_Wav50 Matrice_Wav90 Matrice_Wav170 Matrice_Wva50
Matrice_Wva90 Matrice_Wva170 -append

```

In seguito all'ADDESTRAMENTO, i dati salvati nel file "ultima\_prova" sono stati richiamati in memoria e quindi dati in ingresso ad un altro programma, denominato "usa\_doppiacatena", per poter effettuare la prova vera e propria con gli stimoli in ingresso e graficare come le due curve di attivazione relative ai due diversi stimoli sensoriali (la curva rossa per lo stimolo acustico, quella blu per quello visivo) si influenzino a vicenda al cambiare delle dinamiche del sistema.

File "usa\_doppiacatena" :

```

clear

clc

close all

% x acustici; y visivo

```

**load prova\_ultima**

**% a acustici; v visivo**

**global sigma\_a sigma\_v I0\_a I0\_v sigma\_na sigma\_nv N Index**

**% numero i neuroni**

**N = 180;        % numero di neuroni**

**Index = [1:N];    % indice che individua ogni neurone**

**%caratteristica dei neuroni**

**phi=16;**

**pend=0.7;**

**tau=5;**

**%Ingressi**

**%genero la matrice delle sinapsi laterali**

**%(sono disposte in riga; ogni riga è un neurone post sinaptico)**

**for i = 1:N,**

**DD = abs(Index - i);        % distanza non circolare**

**D = DD.\*(DD <= N/2) + (N - DD).\*(DD > N/2);    %distanza con calcolo circolare**

**% sinapsi laterali uditive**

**L\_a(i,:)=Lex0\_a\*exp(-D.\*D/2/sigex/sigex)-Lin0\_a\*exp(-D.\*D/2/sigin/sigin);**



```

L_a(i,i)=0; %non riceve sinapsi da se stesso

% sinapsi laterali visive

L_v(i,:)=Lex0_v*exp(-D.*D/2/sigex/sigex)-Lin0_v*exp(-D.*D/2/sigin/sigin);

L_v(i,i)=0; %non riceve sinapsi da se stesso

end

pos_a = input('posizione input acustico: ');
pos_v = input('posizione input visivo: ');
dt=0.2; % passo di campionamento
T_stimolo = 120/dt; % durata dello stimolo
t=[0:T_stimolo]*dt; % asse dei tempi
LL=length(t);

%genero l'input
[I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % generato come riga
Input_a = Rec_a*I_a'; % calcolo l'input auditivo passato per il campo recettivo
(adesso è colonna)
Input_v = Rec_v*I_v'; % calcolo l'input visivo passato per il campo recettivo
(adesso è colonna)
Y_a = zeros(N,LL);
Y_v = zeros(N,LL);
% simulo generando un grafico dinamico
figure(1)

```

```

for kt =1:LL-1,
    Y_a(:,kt+1) = Y_a(:,kt) + (dt/tau)*(-Y_a(:,kt)+1./(1+exp(-
(Input_a+L_a*Y_a(:,kt)+W_av*Y_v(:,kt)-phi)*pend)));
    Y_v(:,kt+1) = Y_v(:,kt) + (dt/tau)*(-Y_v(:,kt)+1./(1+exp(-
(Input_v+L_v*Y_v(:,kt)+W_va*Y_a(:,kt)-phi)*pend)));
    plot(Index,Y_a(:,kt+1),'r',Index,Y_v(:,kt+1),'b')
axis([60 120 0 1.1])
pause(0.03)
end

Yfinale_a = Y_a(:,LL);
Yfinale_v = Y_v(:,LL);

grid

%-----
%calcolo posizione stimolo auditivo
if pos_a < 90
    ascissa= [ [1:1:pos_a+89] [pos_a-90:1:0] ];
end
if pos_a > 90
    ascissa= [ [181:1:pos_a+90] [pos_a-89:1:180] ];
end
if pos_a==90
    ascissa = 1:1:180;

```

**end**

**baricentro\_a=sum(Yfinale\_a'.\*ascissa)/sum(Yfinale\_a)    %metodo del baricentro**

**[M\_a Index\_a] = max(Yfinale\_a);**

**pos\_max\_a = ascissa(Index\_a)**

**%-----**

**% calcolo posizione stimolo visivo**

**if pos\_v < 90**

**ascissa= [ [1:1:pos\_v+89] [pos\_v-90:1:0] ];**

**end**

**if pos\_v > 90**

**ascissa= [ [181:1:pos\_v+90] [pos\_v-89:1:180] ];**

**end**

**if pos\_v==90**

**ascissa = 1:1:180;**

**end**

**baricentro\_v=sum(Yfinale\_v'.\*ascissa)/sum(Yfinale\_v)    %metodo del baricentro**

**[M\_v Index\_v] = max(Yfinale\_v);**

**pos\_max\_v = ascissa(Index\_v)**

**[baricentro\_a-pos\_a baricentro\_v-pos\_v]**

In seguito, lo studio è proseguito verso vari aspetti relativi al problema, graficando il ventriloquismo relativo alla situazione corrente, le sinapsi cross-modali e i campi recettivi.

Calcola\_Ventriloquismo:

**load prova\_ultima**

**% a acustici; v visivo**

**global sigma\_a sigma\_v I0\_a I0\_v sigma\_na sigma\_nv N Index**

**% numero i neuroni**

**N = 180;        % numero di neuroni**

**Index = [1:N];    % indice che individua ogni neurone**

**% caratteristica dei neuroni**

**phi=16;**

**pend=0.7;**

**tau=5;**

**Ingressi**

**% Inserisco le deviazioni standard del rumore**

**sigma\_na = 0;    %I0\_a/5;    % Deviazione standard del rumore auditivo**

```
sigma_nv = 0; %I0_v/5; % Deviazione standard del rumore visivo
```

```
dt = 0.2;
```

```
L_a = zeros(N,N);
```

```
L_v = zeros(N,N);
```

```
L_a = zeros(N,N);
```

```
L_v = zeros(N,N);
```

```
%genero la matrice delle sinapsi laterali
```

```
%(sono disposte in riga; ogni riga è un neurone post sinaptico)
```

```
for i = 1:N,
```

```
    DD = abs(Index - i); % distanza non circolare
```

```
    D = DD.*(DD <= N/2) + (N - DD).*(DD > N/2); %distanza con calcolo circolare
```

```
    % sinapsi laterali uditive
```

```
    L_a(i,:)=Lex0_a*exp(-D.*D/2/sigex/sigex)-Lin0_a*exp(-D.*D/2/sigin/sigin);
```

```
    L_a(i,i)=0; %non riceve sinapsi da se stesso
```

```
    % sinapsi laterali visive
```

```
    L_v(i,:)=Lex0_v*exp(-D.*D/2/sigex/sigex)-Lin0_v*exp(-D.*D/2/sigin/sigin);
```

```
    L_v(i,i)=0; %non riceve sinapsi da se stesso
```

```
end
```

```
shift = [-30 -25 -20 -15 -10 -5 -2 +2 +5 +10 +15 +20 +25 +30]; % shift studiati dal programma
```

```

Ventriloquismo_baricentro_a = zeros(180,length(shift));
Ventriloquismo_baricentro_v = zeros(180,length(shift));
Ventriloquismo_massimo_a = zeros(180,length(shift));
Ventriloquismo_massimo_v = zeros(180,length(shift));

for colonna = 1:length(shift), % metto in colonna gli shift;
    colonna
    for pos_a = 1: 180, % metto nelle righe le posizioni dello stimolo acustico
        pos_v = pos_a + shift(colonna);

%   if pos_v > 180
%       pos_v = pos_v - 180;
%   end           % non c'è bisogno di questi if, ci pensa
%               % già la funzione Funzione_Genera_Input
%   if pos_v < 0
%       pos_v = 180 + pos_v;
%   end

T_stimolo = 120/dt;      % durata dello stimolo (mi limito a 120 ms)
t=[0:T_stimolo]*dt;    % asse dei tempi
LL=length(t);

%genero l'input

```

```

[I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % generato come riga

Input_a = Rec_a*I_a';          % calcolo l'input auditivo passato per il campo recettivo
(adesso è colonna)

Input_v = Rec_v*I_v';          % calcolo l'input visivo passato per il campo recettivo
(adesso è colonna)

Y_a = zeros(N,LL);
Y_v = zeros(N,LL);

% simulo

for kt =1:LL-1,

    Y_a(:,kt+1) = Y_a(:,kt) + (dt/tau)*(-Y_a(:,kt)+1./(1+exp(-
(Input_a+L_a*Y_a(:,kt)+W_av*Y_v(:,kt)-phi)*pend)));

    Y_v(:,kt+1) = Y_v(:,kt) + (dt/tau)*(-Y_v(:,kt)+1./(1+exp(-
(Input_v+L_v*Y_v(:,kt)+W_va*Y_a(:,kt)-phi)*pend)));

end

Yfinale_a = Y_a(:,LL);
Yfinale_v = Y_v(:,LL);

grid

%-----
%calcolo posizione stimolo auditivo

if pos_a < 90

    ascissa= [ [1:1:pos_a+89] [pos_a-90:1:0] ];

end

```

```

if pos_a > 90
    ascissa= [ [181:1:pos_a+90] [pos_a-89:1:180] ];
end

if pos_a==90
    ascissa = 1:1:180;
end

baricentro_a=sum(Yfinale_a'.*ascissa)/sum(Yfinale_a);    %metodo del baricentro
[M_a Index_a] = max(Yfinale_a);
pos_max_a = ascissa(Index_a);

%-----
% calcolo posizione stimolo visivo

if pos_v < 90
    ascissa= [ [1:1:pos_v+89] [pos_v-90:1:0] ];
end

if pos_v > 90
    ascissa= [ [181:1:pos_v+90] [pos_v-89:1:180] ];
end

if pos_v==90
    ascissa = 1:1:180;
end

baricentro_v=sum(Yfinale_v'.*ascissa)/sum(Yfinale_v);    %metodo del baricentro

```



```
[M_v Index_v] = max(Yfinale_v);
```

```
pos_max_v = ascissa(Index_v);
```

```
Ventriloquismo_baricentro_a(pos_a,colonna) = baricentro_a - pos_a;
```

```
Ventriloquismo_baricentro_v(pos_a,colonna) = baricentro_v - pos_v;
```

```
Ventriloquismo_massimo_a(pos_a,colonna) = pos_max_a - pos_a;
```

```
Ventriloquismo_massimo_v(pos_a,colonna) = pos_max_v - pos_v;
```

```
end
```

```
end
```

```
Media_Ventriloquismo_baricentro_a = mean(Ventriloquismo_baricentro_a);
```

```
Media_Ventriloquismo_baricentro_v = mean(Ventriloquismo_baricentro_v);
```

```
Media_Ventriloquismo_massimo_a = mean(Ventriloquismo_massimo_a);
```

```
Media_Ventriloquismo_massimo_v = mean(Ventriloquismo_massimo_v);
```

```
% disegno la figura finale
```

```
Width = 1.5;
```

```
Marker = 12;
```

```
Font = 12;
```

```
figure(1)
```

```
plot(shift,Media_Ventriloquismo_baricentro_a,'r*-','linewidth',Width,'markersize',Marker)
```

```
hold on
```

```
plot(shift,Media_Ventriloquismo_baricentro_v,'bo--','linewidth',Width,'markersize',Marker)
```

```
xlabel('shift (degree)','fontsize',Font)
```

```
ylabel('error in perceived position (degree)','fontsize',Font)
```

```
title('barycenter method','fontsize',Font)
```

```

set(gca,'fontsize',Font)
axis([-30 30 -10.1 10.1])
%
figure(2)
plot(shift,Media_Ventriloquismo_massimo_a,'r*-', 'linewidth',Width,'markersize',Marker)
hold on
plot(shift,Media_Ventriloquismo_massimo_v,'bo--', 'linewidth',Width,'markersize',Marker)
xlabel('shift (degree)', 'fontsize',Font)
ylabel('error in perceived position (degree)', 'fontsize',Font)
title('maximum method', 'fontsize',Font)
set(gca,'fontsize',Font)
axis([-30 30 -10.1 10.1])

%save risultati_ventriloquismo shift Media_Ventriloquismo_baricentro_a
Media_Ventriloquismo_baricentro_v Media_Ventriloquismo_massimo_a
Media_Ventriloquismo_massimo_v

Grafica_sinapsi_crossmodali:

clear all
close all
clc

load prova_ultima

```

```

Index = [1:1:180];

figure

Width = 1.5;

Marker = 8;

Font = 14;

subplot(221)

for kk = 1:2:20,

plot(Index,Matrice_Wav90(kk,:),'g--','linewidth',Width)

hold on

end

plot(Index,Matrice_Wav90(end,:),'r','linewidth',1.5*Width)

xlabel('position (deg)','fontsize',Font)

title('cross-modal to auditory','fontsize',Font)

axis([0 180 -0.1 0.36])

set(gca,'fontsize',Font)

grid

subplot(222)

for kk = 1:2:20,

plot(Index,Matrice_Wva90(kk,:),'g--','linewidth',Width)

hold on

end

plot(Index,Matrice_Wva90(end,:),'r','linewidth',1.5*Width)

xlabel('position (deg)','fontsize',Font)

title('cross-modal to visual','fontsize',Font)

```

```
axis([0 180 -0.1 0.36])
```

```
set(gca,'fontsize',Font)
```

```
grid
```

Grafica\_campi\_recettivi:

```
clear all
```

```
close all
```

```
clc
```

```
global sigma_a sigma_v I0_a I0_v sigma_na sigma_nv N Index
```

```
% calcolo gli ingressi
```

```
pos_a = 90;
```

```
pos_v = 90;
```

```
N = 180;
```

```
Index = [1:1:180];
```

```
Ingressi
```

```
sigma_na = 0;
```

```
sigma_nv = 0;
```

```
[I_a I_v]= Funzione_Genera_input(pos_a,pos_v); % gli input sono generati come vettori riga
```

```
load prova_ultima
```

```
Rec_a = Matrice_rec90_a(end,:);
```

```

Rec_v = Matrice_rec90_v(end,:);

figure

Width = 1.5;

Marker = 8;

Font = 14;

subplot(221)

for kk = 1:2:20,

plot(Index,Matrice_rec90_a(kk,:),'--g','linewidth',Width)

hold on

end

plot(Index,Matrice_rec90_a(end,:),'r','linewidth',1.5*Width)

xlabel('position (deg)','fontsize',Font)

title('auditory RF','fontsize',Font)

axis([0 180 -0.1 1.5])

set(gca,'fontsize',Font)

grid

subplot(222)

for kk = 1:2:20,

plot(Index,Matrice_rec90_v(kk,:),'--g','linewidth',Width)

hold on

end

plot(Index,Matrice_rec90_v(end,:),'r','linewidth',1.5*Width)

xlabel('position (deg)','fontsize',Font)

title('visual RF','fontsize',Font)

```

```

axis([0 180 -0.1 1.5])
set(gca,'fontsize',Font)

grid

figure

Width = 1.5;
Marker = 8;
Font = 12;

subplot(221)

plot(Index,Rec_a,'r',Index,I_a,'b','linewidth',Width)

xlabel('position (deg)','fontsize',Font)

title('auditory RF (red) - auditory input (blue)','fontsize',Font)

axis([0 180 -0.1 1.0])

set(gca,'fontsize',Font)

grid

subplot(222)

plot(Index,Rec_v,'r',Index,I_v,'b','linewidth',Width)

xlabel('position (deg)','fontsize',Font)

title('visual RF (red) - visual input (blue)','fontsize',Font)

axis([0 180 -0.1 2.0])

set(gca,'fontsize',Font)

grid

```

Ricordiamo ancora una volta inoltre, che gli input utilizzati in tale modello sono stati generati in maniera random da una “funzione\_genera\_input” che è stata richiamata nel programma di addestramento precedentemente trascritto.

Funzione\_genera\_input:

Questa funzione è stata utilizzata per generare lo stimolo in ingresso alle catene di neuroni auditivi e visivi, immaginando estensione spaziale gaussiana e rumore gaussiano.

```
function [I_a I_v]= Funzione_Genera_input(pos_a,pos_v)  
global sigma_a sigma_v I0_a I0_v sigma_na sigma_nv N Index  
  
% calcolo le distanze circolari  
DD_a=abs(Index - pos_a); % distanza auditiva non circolare  
DD_v=abs(Index - pos_v); % distanza visiva non circolare  
D_a = DD_a.*(DD_a <= N/2) + (N - DD_a).*(DD_a > N/2); %distanza auditiva con calcolo  
circolare  
D_v = DD_v.*(DD_v <= N/2) + (N - DD_v).*(DD_v > N/2); %distanza auditiva con calcolo  
circolare  
  
% genero i vettori medi  
M_a = I0_a*exp(-D_a.*D_a/2/sigma_a/sigma_a); % vettore medio dello stimolo auditivo  
(rappresenta la distribuzione media nello spazio)  
M_v = I0_v*exp(-D_v.*D_v/2/sigma_v/sigma_v); % vettore medio dello stimolo auditivo  
(rappresenta la distribuzione media nello spazio)
```

**% genero gli input**

**I\_a = sigma\_na\*randn(1,180) + M\_a;**

**% ingresso ai 180 neuroni auditivi**

**I\_v = sigma\_nv\*randn(1,180) + M\_v;**

**% ingresso ai 180 neuroni visivi**

Questo programma è stato utilizzato per studiare il comportamento della rete in condizioni imposte di base e in condizioni modificate, in modo da poter ottenere una simulazione del l'effetto dei vari parametri sulla localizzazione degli stimoli acustici e visivi.



## Capitolo IV – Simulazioni con i valori basali dei parametri

Al fine di testare i concetti espressi nel primo capitolo, nel presente lavoro la rete è stata sottoposta a 40 epoche di addestramento. Ogni epoca è composta a sua volta, da 2 ingressi uditivi unisensoriali per ogni posizione (ad esempio, sono state eseguite 360 diverse prove), 2 stimoli visivi unisensoriali per ogni posizione (360 prove) e 1 stimolo cross-modale con input visivi e uditivi coincidenti sempre per ciascuna posizione (180 prove).

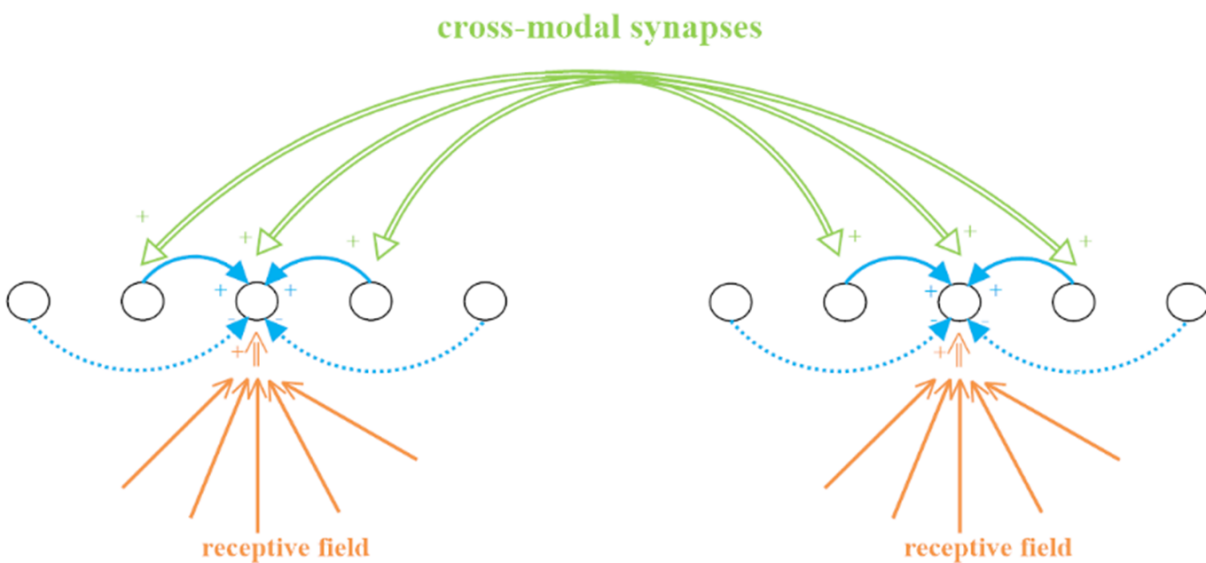


Figura 11 Modello di rete neurale utilizzata

In breve, ogni epoca ha contato un numero di 900 prove con distribuzione spaziale uniforme, con 2/5 relativi alla stimolazione unisensoriale visiva, 2/5 relativi a quella unisensoriale uditiva e 1/5 relativa agli ingressi cross-modali coincidenti. Pertanto, il numero complessivo di prove dopo 40 epoche è stato di ben 36000. Al termine di ogni singola epoca, quando i neuroni risultavano in condizioni stazionarie, le sinapsi sono state addestrate usando le regole Oja sopra qui sotto nuovamente riportata:

$$\Delta w_{ij} = \gamma y_i (y_j - w_{ij})$$

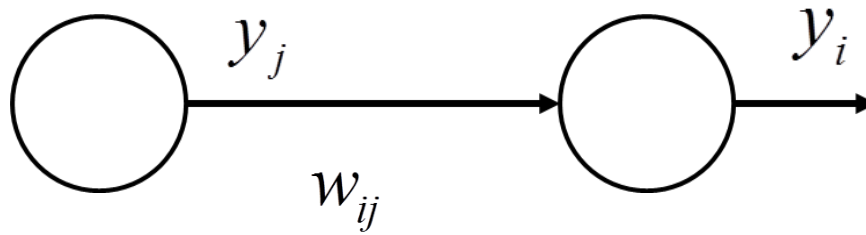


Figura 12 grafico relativo alla regola di Oja

#### 4.1 I Campi recettivi

La figura sottostante mostra come, usando i valori di base per i parametri del modello, gli RFS siano stati progressivamente influenzati dall'addestramento. All'inizio, ai campi recettivi di tutti i neuroni sono stati dati gli stessi valori (sia nella rete uditiva e visiva), con una grande deviazione standard spaziale. Questo, per simulare la presenza di RFS immaturi. Per quanto riguarda i due neuroni uditivi e visivi rappresentativi, si nota che gli RFS progressivamente tendono a ridursi durante l'addestramento, fino a raggiungere una configurazione stabile finale. Gli RFS dei neuroni uditivi rimangono piuttosto grandi, mentre gli RFS di quelli visivi diventano molto più stretti: questa differenza riflette la deviazione standard degli stimoli esterni scelta durante l'addestramento. Un'importante conseguenza della regola di Oja è che, dopo l'addestramento, la RF di ogni neurone riflette l'ingresso medio a quel neurone. Nel nostro modello, questo è fornito da una funzione gaussiana con deviazione standard assegnata. Per verificare questa previsione, nella Figura 13 confronto gli RFS finali dei due neuroni presi da esempio con le funzioni gaussiane che rappresenta gli ingressi medi utilizzati durante l'addestramento. L'accordo è molto buono per quanto riguarda la RF uditiva; viceversa, la RF visiva mostra un buon accordo dal punto di vista spaziale, ma una certa differenza di ampiezza.

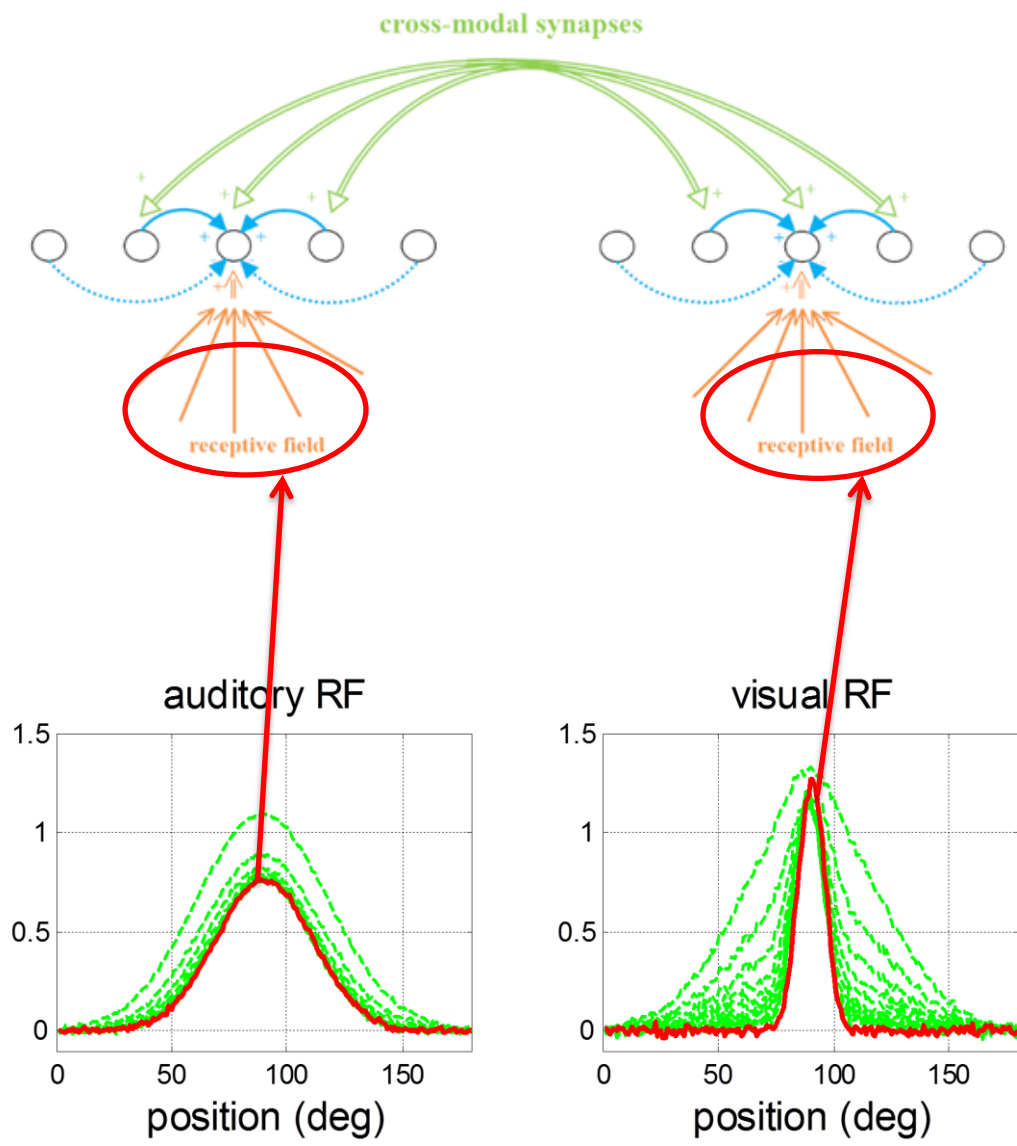


Figura 13 Durante l'addestramento i campi recettivi tendono a ridursi

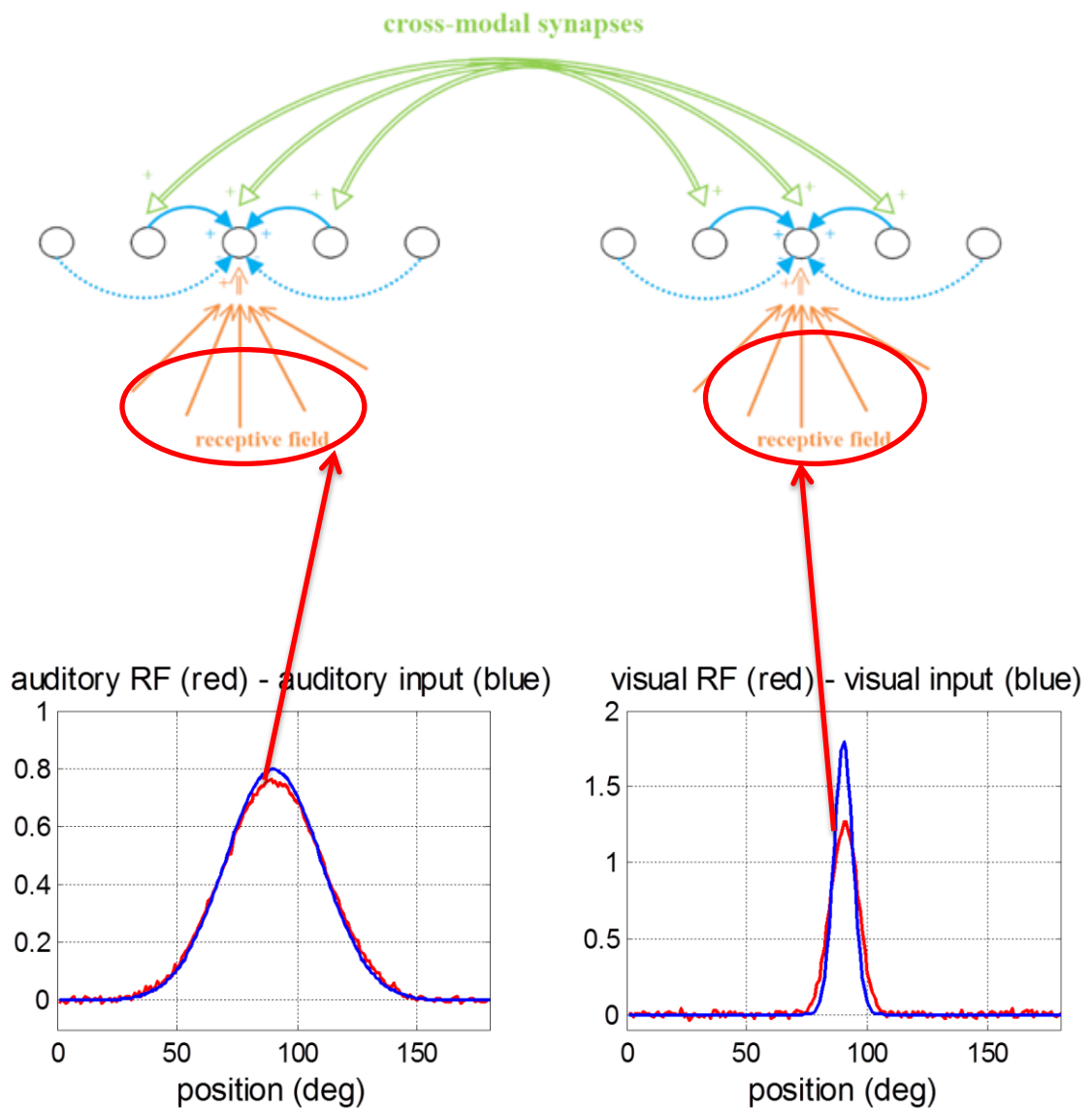


Figura 14 Dopo l'addestramento i campi recettivi diventano simili allo stimolo medio nella posizione indicata

In conclusione, possiamo dire che, dopo l'addestramento con la regola Oja, la SD di RFS riflette attentamente la SD media degli stimoli in ingresso, con i neuroni visivi più spazialmente concentrati rispetto a quelli uditivi.

## 4.2 Le sinapsi Cross-modali

Se tutti gli stimoli di ingresso che abbiamo fornito alla rete fossero stati di tipo unimodale, le sinapsi Cross-modali non si sarebbero generate e le due reti si sarebbero comportate in modo puramente unisensoriale. In questa condizione, l'unico effetto relativo all'addestramento della rete sarebbe stato il cambiamento nelle RFS. Invece, grazie alla presenza di una certa quantità di stimoli multisensoriali, vengono create le sinapsi Cross-modali tra le due reti. Queste ultime sono illustrate nella Figura sottostante, che mostra come l'accoppiamento delle due aree unisensoriali produca sinapsi Cross-modali. Le sinapsi cross-modali iniziano da zero e progressivamente aumentano assumendo una forma campanulare. Alla fine dell'addestramento, il neurone in una modalità unisensoriale (acustica o visiva) riceve sinapsi dai neuroni relativi all'altra modalità unisensoriale che approssimativamente hanno la stessa posizione spaziale.

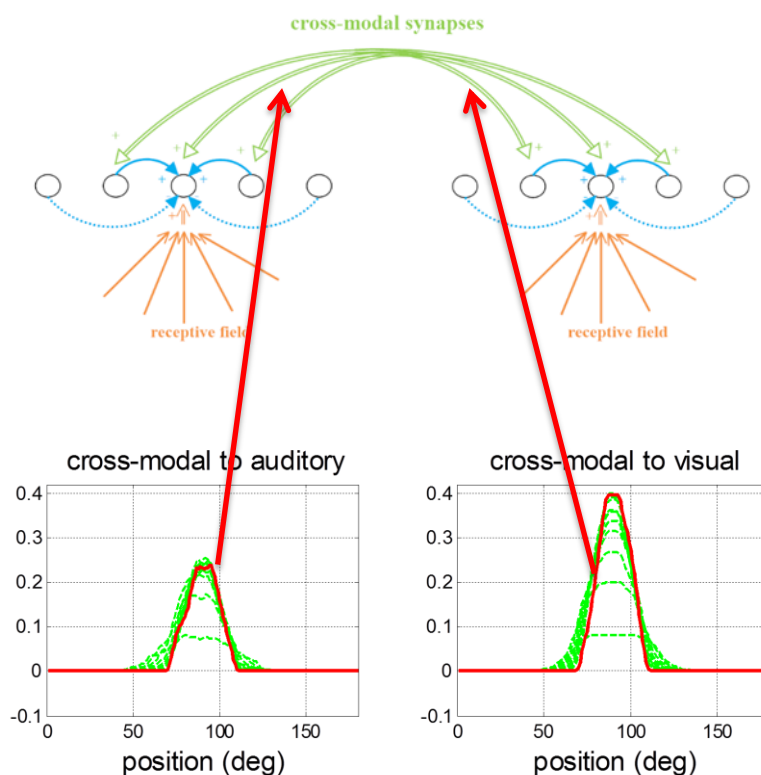


Figura 15 Sinapsi Cross-modali che sono prodotte dall'accoppiamento delle due aree unimodali

### 4.3 Risposta del modello agli input unimodali

Dopo l'addestramento, abbiamo testato il comportamento della rete con una varietà di stimoli in ingresso. In primo luogo, abbiamo simulato la risposta agli input unimodali. In questa condizione, a causa dei bassi valori di sinapsi Cross-modali, solo la rete in una modalità risultava eccitata, mentre l'altra rete taceva. Come descritto nella sezione precedente in relazione agli aspetti teorici, in condizioni di stimolo unimodale l'ingresso esterno ai neuroni (ovvero, il prodotto tra lo stimolo e il campo recettivo) è una approssimazione della funzione di verosimiglianza. A titolo di esempio, la Figura sottostante mostra l'ingresso esterno di tutti i neuroni della rete, in risposta ad uno stimolo uditivo (pannello di sinistra) o uno stimolo visivo (pannello di destra) collocato alla posizione  $90^\circ$ . I risultati mostrano, come previsto teoricamente, che le due grandezze sono strettamente simili. Pertanto, in assenza di connessioni intra-strato laterali, la rete unisensoriale sarebbe strettamente predisposta a riprodurre uno stimatore di massima verosimiglianza, semplicemente scegliendo il neurone con massima attività.

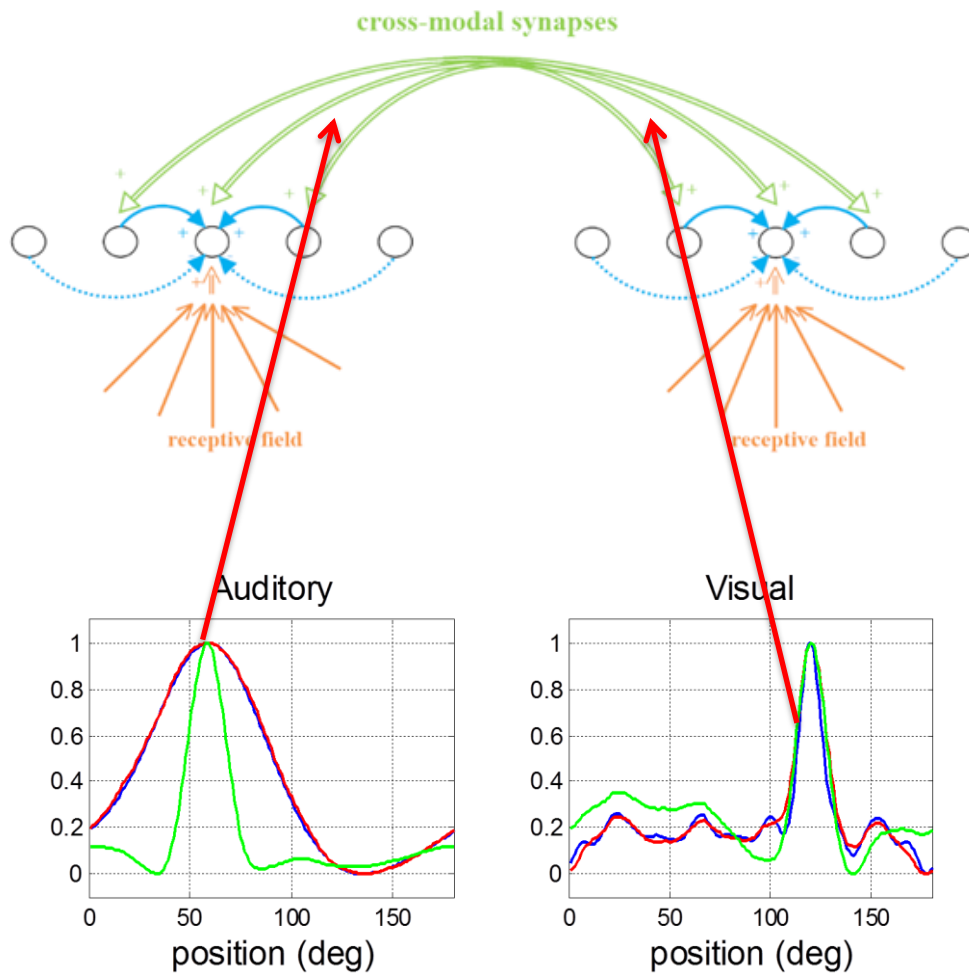


Figura 16 Modello di risposta agli input unimodali

**Blue:** logaritmo della funzione di verosimiglianza  
**Red:** input esterno alla rete (dal campo recettivo)  
**Green:** input totale (incluse le sinapsi intra-area laterali)

In realtà però, la somma degli ingressi totali ai neuroni è differente a causa della presenza di connessioni laterali. Come noto dalla letteratura, la presenza di connessioni inibitorie tende a compattare la risposta della rete riducendo la semi-larghezza a metà altezza della curva di sintonia;

questo è mostrato in Figura, dove l'input totale di neuroni è raffigurato con una linea verde. Tuttavia, questo processo non modifica sostanzialmente la posizione del massimo. Per simulare accuratamente la SMV, abbiamo effettuato 180 prove diverse per ciascuna modalità, dando un impulso di ingresso rumoroso in ciascuna posizione. Confrontando i risultati ottenuti sperimentalmente con le previsioni ottenute dalla rete utilizzando la posizione del neurone massimamente attivo, siamo giunti ad alcuni risultati soddisfacenti. I risultati, sottolineano le forti analogie tra la MLE e la risposta della rete. In conclusione, si può affermare che, in condizioni unimodale, ciascuna rete addestrata realizza una stima di massima verosimiglianza della posizione dello stimolo.

#### **4.4 Modello in risposta agli input cross-modali**

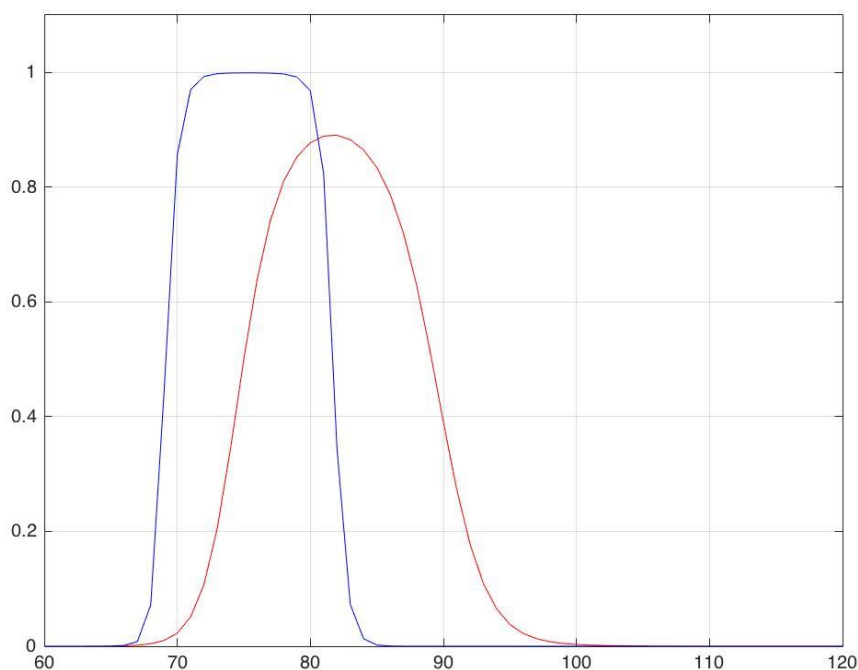
Successivamente, abbiamo simulato il comportamento della rete in risposta a stimoli cross-modali. A tal fine, abbiamo stimolato la rete con stimoli provenienti da differenti posizioni spaziali. Più precisamente, si è passato uno stimolo uditivo dalla posizione 1 alla 180 e ad ognuno abbiamo aggiunto un secondo stimolo, questa volta di tipo visivo, ad una distanza che andava da  $-30^\circ$  a  $+30^\circ$  rispetto quello uditivo. Un aspetto fondamentale, da tenere in considerazione, è il fatto che in queste simulazioni entrambe le reti sono state simultaneamente eccitate, infatti a causa della presenza di sinapsi cross-modali, l'attività di ogni rete è influenzata dall'attività nell'altra. La principale conseguenza è che la posizione percepita degli stimoli è spostata rispetto a quella reale. Una sintesi dei risultati è presentata in Figura, dove si mostra lo spostamento tra la posizione reale dello stimolo e quella percepita (tracciata contro la distanza tra uditivo e stimoli visivi). Ogni punto in figura rappresenta la media di 180 prove.

Come è evidente da questi grafici, la posizione percepita dello stimolo uditivo è significativamente spostata in direzione di quello visivo (distanze positivi significano che lo stimolo visivo è alla destra di quello uditivo e viceversa). Questo errore di percezione è massimo (circa  $8^\circ$ ) quando la distanza



audiovisiva è a circa  $20^\circ$ . La percezione della posizione dello stimolo visivo viene anche spostata in direzione di quello uditivo, ma l'errore in questo caso risulta essere molto più piccolo (errore massimo di circa  $1,5^\circ$ ).

In conclusione, possiamo osservare una sorta di "effetto ventriloquismo". I risultati sono in buon accordo con quelli comportamentali che ci saremmo potuti aspettare. L'unica differenza significativa è che, nel modello, l'errore di percezione diminuisce ad una distanza maggiore di  $25^\circ$ , dove invece i dati comportamentali presentano ancora un errore significativo.



**Figura 17** Stimolazione cross-modale: il modello simula l'effetto ventriloquismo, lo **stimolo visivo** posizionato a  $90^\circ$  **stimolo acustico** a  $75^\circ$ .

Questa figura mostra come l'attività nella rete uditiva, inizialmente posizionata intorno alla posizione del vero stimolo uditivo, si sposta progressivamente in direzione di quello visivo in conseguenza all'azione degli ingressi cross-modali.

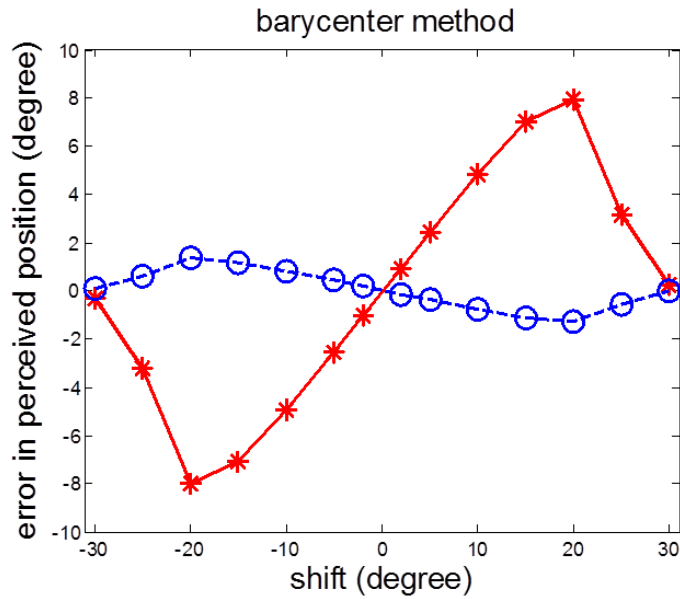


Figura 18 Stimolazione cross-modale: il modello simula l'effetto del ventriloquismo (la posizione è stata calcolata sia con il baricentro sia attraverso l'attività massima) dove con il **BLU** si esprime l'errore relativo alla posizione uditiva e con il **ROSSO** quello relativo alla posizione visiva.

Confrontiamo i risultati ottenuti dal nostro progetto relativi al fenomeno del ventriloquismo con quelli relativi alla letteratura:

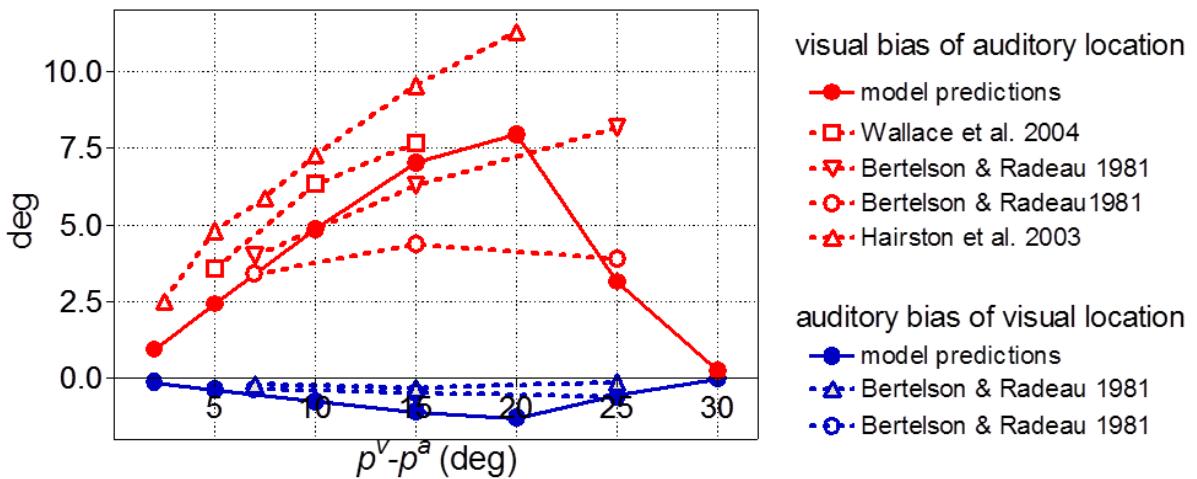


Figura 19 confronto dei risultati ottenuti sul ventriloquismo con i risultati della letteratura.

**Blu:** errore posizione visiva  
**Rosso:** errore posizione uditiva

Infine, abbiamo confrontato le previsioni del modello sulla posizione degli stimoli uditivi e visivi , con quelle ottenuti con uno stimatore Bayesiano (con stima di massima probabilità a posteriori). A tal fine, tuttavia, è necessario fornire una espressione per la probabilità a priori. Durante l'addestramento abbiamo sempre usato due stimoli cross-modalità coincidenti. Tuttavia, nel calcolare la probabilità a priori ora, abbiamo ipotizzato che gli ingressi cross-modalità visivi e uditivi abbiano possano avere una distanza molto piccola (cioè, una differenza spaziale minima) e che vi sia una residua probabilità, anche se molto piccola, di stimoli indipendenti fra loro. Quindi, la probabilità a priori è stata rappresentata con una funzione gaussiana della distanza tra i due stimoli ed una piccolissima  $M$  costante. Nella presente simulazione, abbiamo utilizzato una deviazione standard di  $0,8^\circ$ , il che sta a significare che lo stimolo visivo e uditivo debbano essere quasi coincidenti (solo raramente possono avere una distanza di un grado, quasi mai due gradi di distanza).

I risultati, illustrati in figura, mostrano che l'accordo tra gli errori di posizione, ottenuti con il modello, e la BE è abbastanza buona. Entrambi prevedono che la stima dell' errore aumenti con la distanza tra gli stimoli audiovisivi fino ad una distanza di circa 15 gradi. Nella gamma tra 0-15 le previsioni dello stimatore Bayesiano e del modello sono abbastanza vicine. Una differenza significativa, tuttavia, è evidente sopra i 15 gradi.

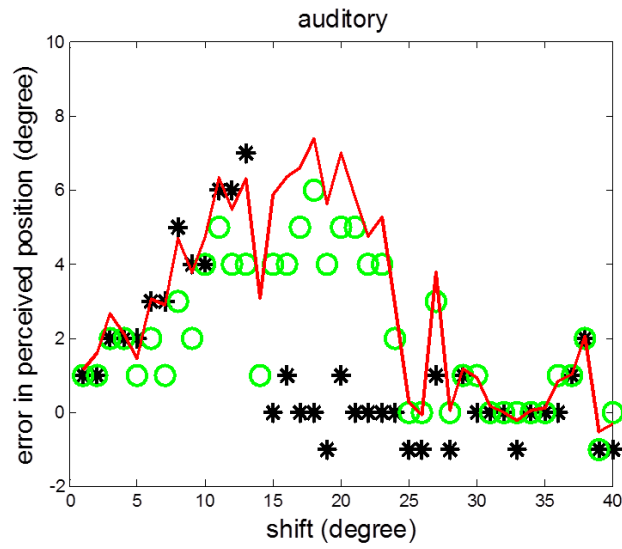


Figura 20 accordi tra errori di predizione e stima bayesiana per lo stimolo acustico

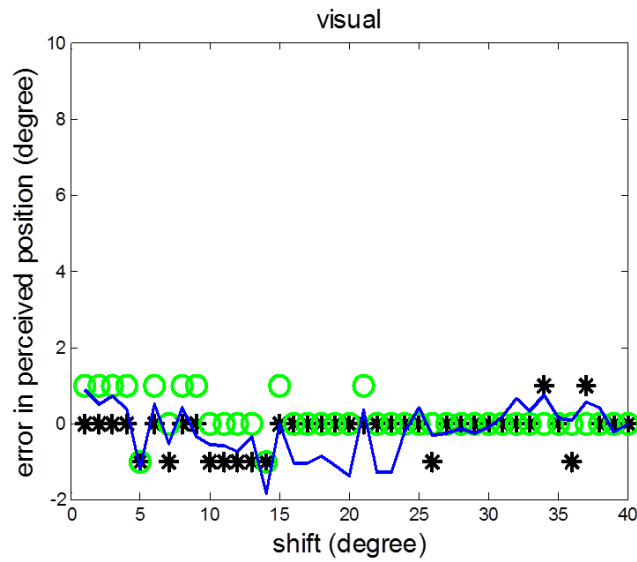


Figura 21 accordi tra errori di previsione e stima bayesiana per lo stimolo visivo

○: model (maximum activity) ,  
linea continua: model barycenter  
 \*: Bayesian estimate;

#### 4.5 L'effetto “fissione”

Una delle conseguenze della presenza di sinapsi cross-modali tra neurone uditivo e la rete visiva è il verificarsi di illusioni audio-visive. In particolare, non è presente solo l'effetto del ventriloquismo analizzato sopra (dove lo stimolo visivo cattura spazialmente quello uditivo), ma anche l'Effetto fissione (dove lo stimolo uditivo domina nel dominio temporale). In un precedente lavoro computazionale (Cristiano C., Mauro U., Nadia B., Giuseppe V., Elisa M., (2014)), che citeremo solo in questa tesi, era stato dimostrato che entrambi le illusioni possono essere simulate con lo stesso modello, fornendo alla risposta uditiva una costante di tempo più veloce rispetto a quella visiva.

La disposizione temporale dei segnali acustici e del flash erano le stesse e le ampiezze di stimolo erano state scelte in modo che ogni singolo ingresso produca un'attività di picco transitoria vicino alla saturazione. I risultati mostrano che la rete produce la comparsa di un secondo flash illusorio, cioè, il noto effetto “dell'illusione del flash”.

In questo lavoro, oltre che studiare il comportamento della rete in condizioni base sottoposta a stimoli multimodali, si sono volute intraprendere delle modifiche parametriche e procedurali al fine di comprendere, in linea generale, quale sia l'influenza reale che l'addestramento può avere sulla localizzazione spaziale dei due stimoli unimodali.

## Capitolo V – Analisi di sensitività

Nel corso di questa tesi sono stati trattati vari aspetti riguardanti il riconoscimento della posizione audio-visiva in seguito a vari tipi di simulazioni di tipo unisensoriale e multisensoriale.

### 5.1 Condizioni di base

Partendo dalle condizioni imposte di base, ovvero :

**sigma\_a =20;**     % deviazione standard ampiezza spaziale auditiva

**sigma\_v = 4;**     % deviazione standard ampiezza spaziale visiva

#### Forza dello stimolo

**I0\_a = 36/(sqrt(2\*pi\*sigma\_a\*sigma\_a));**     % forza input uditivo

**I0\_v = 20/(sqrt(2\*pi\*sigma\_v\*sigma\_v));**     % forza input visivo

#### Deviazioni standard del rumore

**sigma\_na = I0\_a/3;**     % deviazione standard del rumore auditivo

**sigma\_nv = I0\_v/3;**     % deviazione standard del rumore visivo

#### Parametri delle sinapsi laterali

**sigmaex=12;**     %deviazione standard delle sinapsi laterali eccitatorie

**sigmain=24;**     %deviazione standard delle sinapsi laterali inibitorie

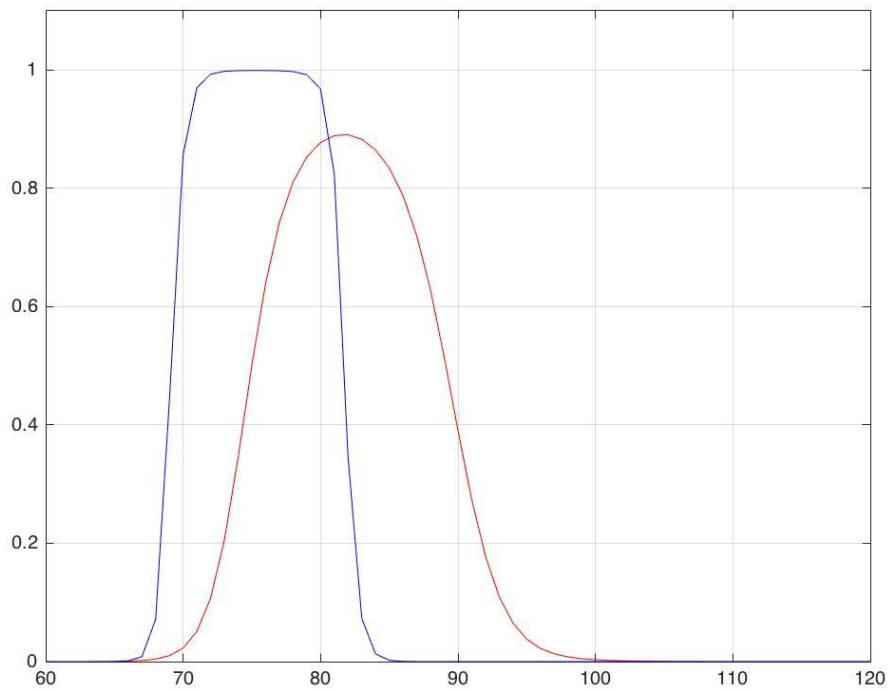
**Lex0\_a=1.9;** %forza delle sinapsi laterali eccitatorie uditive

**Lin0\_a=1.85;** %forza delle sinapsi laterali inibitorie uditive

**Lex0\_v=1.9;** %forza delle sinapsi laterali eccitatorie visive

**Lin0\_v=1.85;** %forza delle sinapsi laterali inibitorie visive

Grafico della prova:



**Figura 22** curve di risposta agli stimoli visivo (blu) e acustico (rosso)

Dove la curva rossa rappresenta la stimolo acustico, mentre quella blu lo stimolo visivo.

Le posizioni spaziali date inizialmente ai due stimoli era  $90^\circ$  per quello acustico e  $75^\circ$  per quello visivo. Si può notare come l'attività visiva tende a spostarsi verso la posizione di quella acustico a causa della presenza delle sinapsi cross-modali, che hanno effetto sul comportamento della rete.

Campi recettivi:

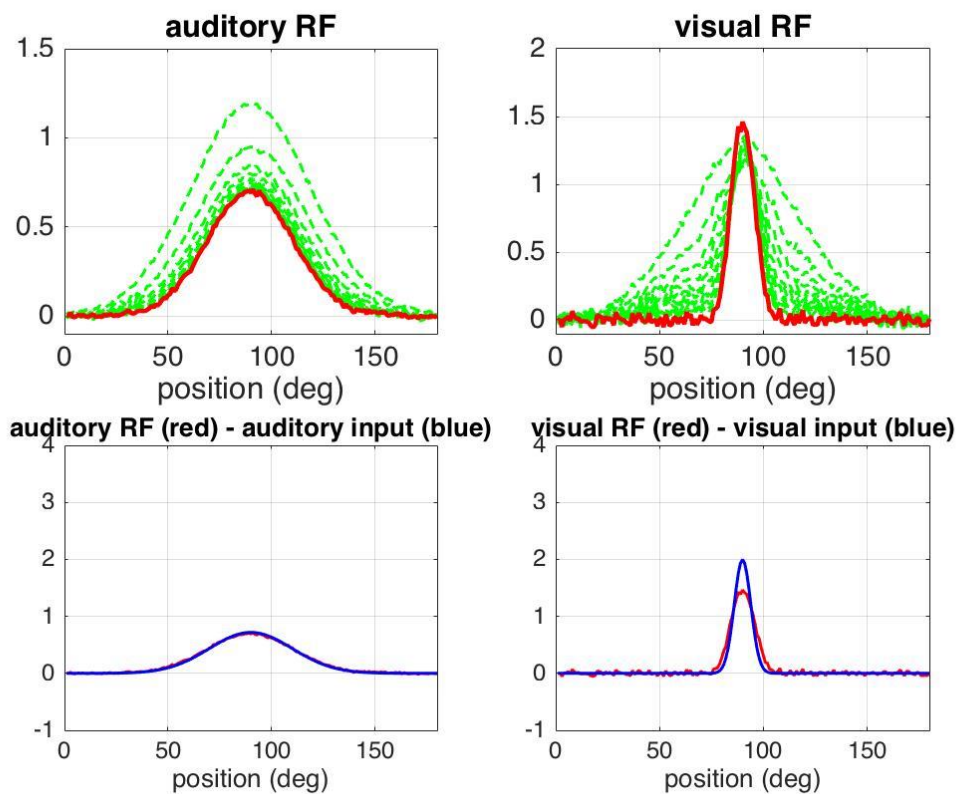


Figura 23 Campi recettivi

I campi recettivi sono ben ricostruiti.

Sinapsi Cross-modali:

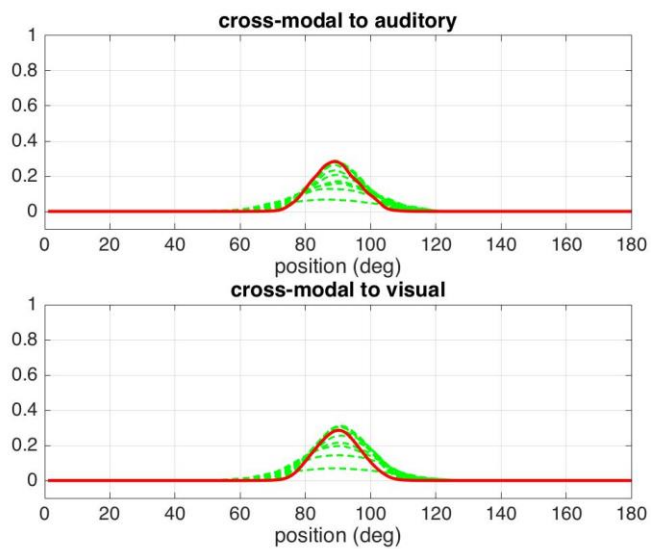
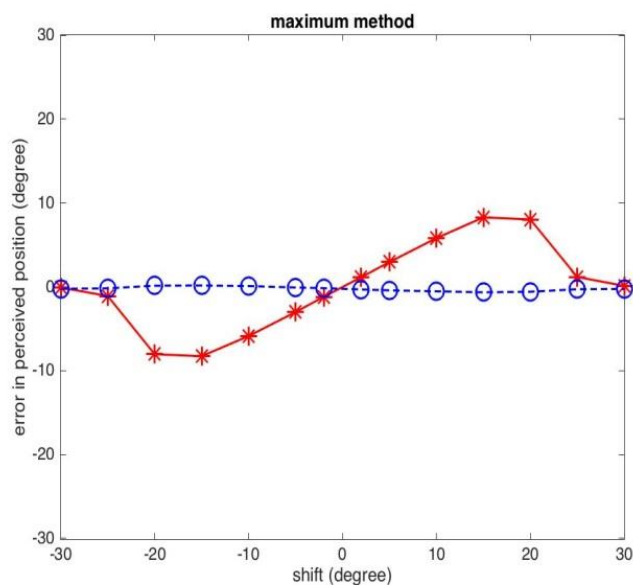


Figura 24 sinapsi cross-modali

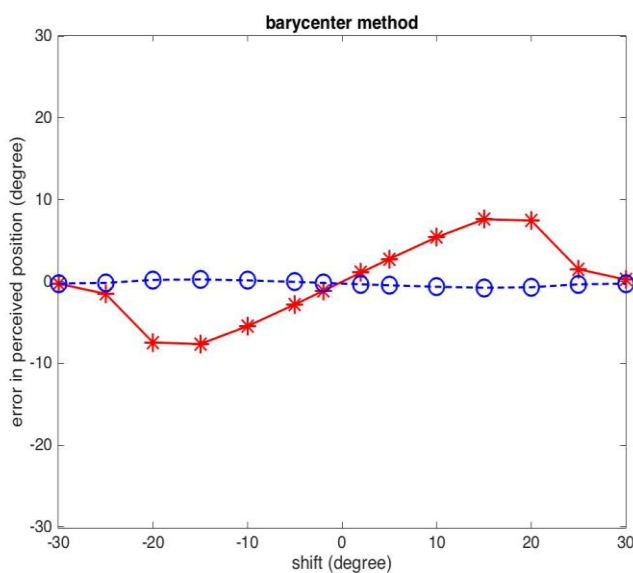


Le Sinapsi Cross-modali sono abbastanza alte.

Il Ventriloquismo è stato calcolato rispettivamente con due metodi: il metodo che sfrutta la massima verosimiglianza (Figura 23) ed il metodo del baricentro (figura 24).



**Figura 25** ventriloquismo calcolato con il metodo della massima verosimiglianza



**Figura 26** ventriloquismo calcolato con il metodo del baricentro

Nonostante siano stati utilizzati due metodi differenti è ben visibile che il risultato è equivalente in ambedue i casi: il fenomeno del ventriloquismo è molto accentuato, risultando in uno shift dello stimolo acustico di oltre 8°.

## 5.2 Variazioni della precisione della stima

Sono state attuate delle modifiche, per prima cosa, riguardanti la precisione degli stimoli, che naturalmente si riflette sulla precisione dei parametri stimati. Le  $\sigma^A$  e  $\sigma^V$  sono state scelte in base a specifici criteri che ci hanno permesso di evidenziare l'influenza dell'accuratezza degli input sulle stime.

➤  $\sigma^A = 20$   $\sigma^V = 8$  :

La deviazione standard delle sinapsi uditive è stata lasciata inalterata, mentre è stata aumentata, anche se di poco, quella delle sinapsi visive.

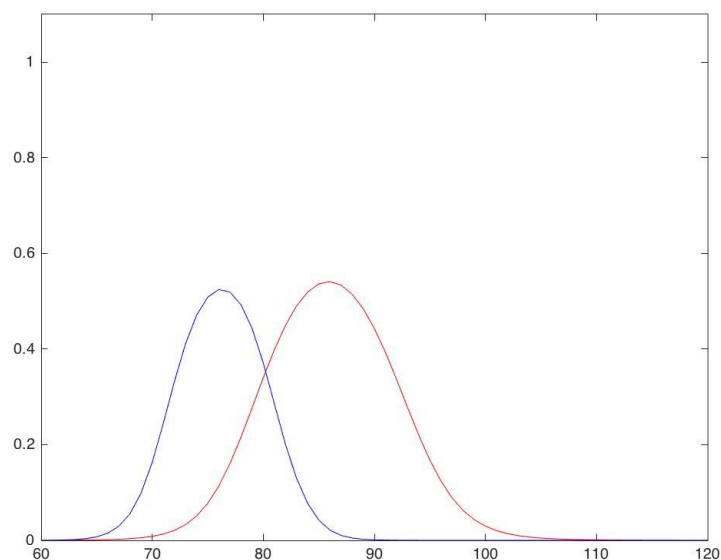


Figura 27 curve delle attività acustica (rosso) e visiva (blu)

dove la curva rossa rappresenta la stimolo acustico, mentre quella blu lo stimolo visivo.

Le posizioni spaziali date inizialmente ai due stimoli erano  $90^\circ$  per quello acustico e  $75^\circ$  per quello visivo. Le curve sono molto più basse ma lo spostamento dei due stimoli l'uno rispetto all'altro rimane all'incirca costante.

I campi recettivi risultano anche in questo caso ben ricostruiti:

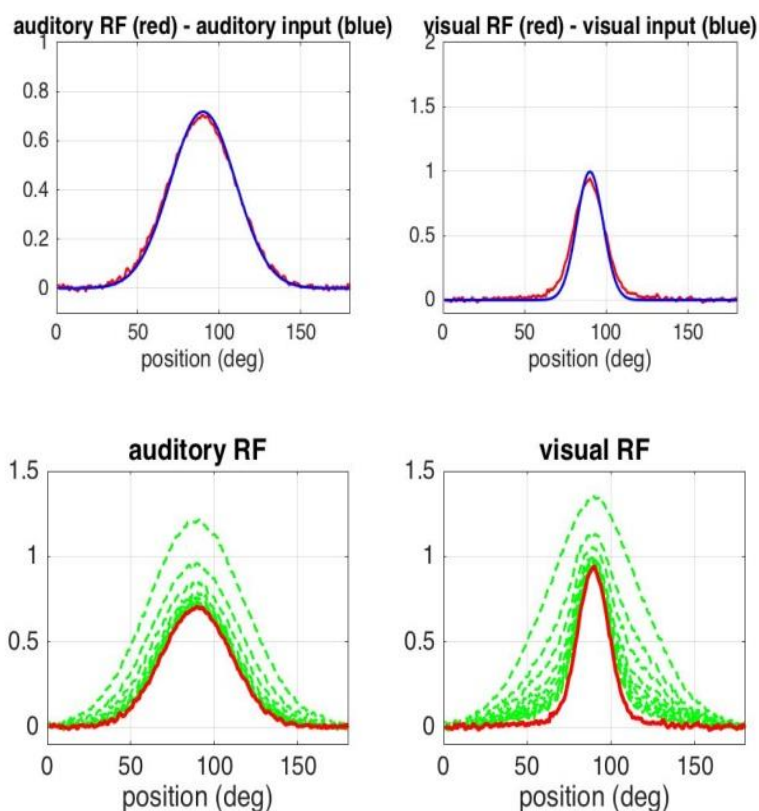


Figura 28 Campi recettivi

Notiamo anche che il ventriloquismo, sempre calcolato con entrambi i metodi, si attenua in modo significativo (da circa  $8^\circ$  a  $2^\circ$ ).

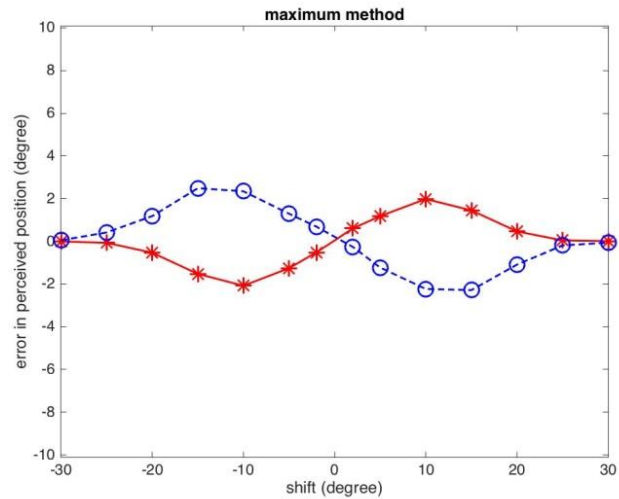


Figura 29 ventriloquismo calcolato con il metodo della massima verosimiglianza

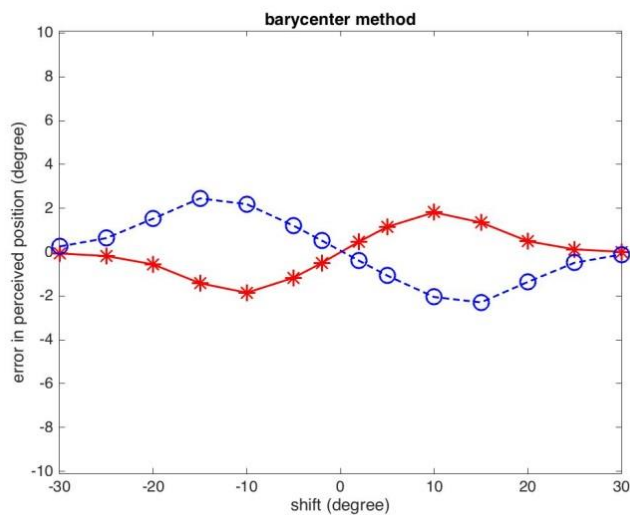


Figura 30 ventriloquismo calcolato con il metodo del baricentro

➤  $\sigma^A = 30$   $\sigma^V = 4$  :

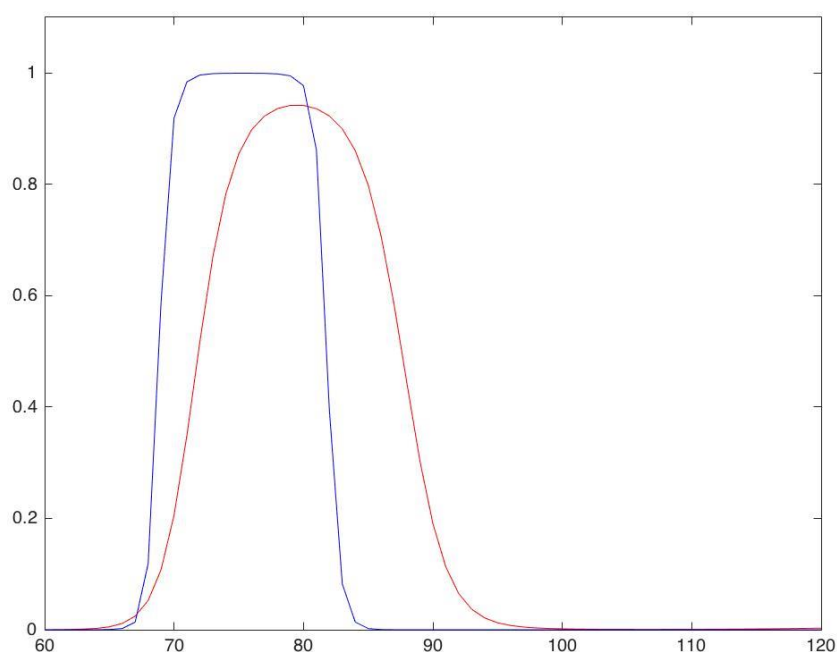
La deviazione standard delle sinapsi uditive è stata aumentata, mentre è stata lasciata inalterata quella delle sinapsi visive. Per avere risultati soddisfacenti, in questo caso, abbiamo bisogno di

aumentare la forza dello stimolo acustico in modo tale che abbia l'ampiezza adeguata per stimolare sufficientemente i neuroni dell'area acustica.

Forza degli stimoli acustici e visivi scelti in ingresso nel caso corrente:

$$I0\_a = 45/(\text{sqrt}(2*\text{pi}*\text{sigma}_a*\text{sigma}_a));$$

$$I0\_v = 20/(\text{sqrt}(2*\text{pi}*\text{sigma}_v*\text{sigma}_v));$$



**Figura 31** Curva di risposta allo stimolo acustico (rosso) e visivo (blu)

dove la curva rossa rappresenta la stimolo acustico, mentre quella blu lo stimolo visivo.

Le posizioni spaziali date inizialmente ai due stimoli era  $90^\circ$  per quello acustico e  $75^\circ$  per quello visivo. Le due curve, di ampiezza maggiore rispetto alla precedente simulazione, rispondono bene alle nostre aspettative. Lo shift dello stimolo acustico verso quello visivo tende ad accentuarsi, riflettendo la minore precisione acustica conseguente alle variazioni effettuate nei parametri.

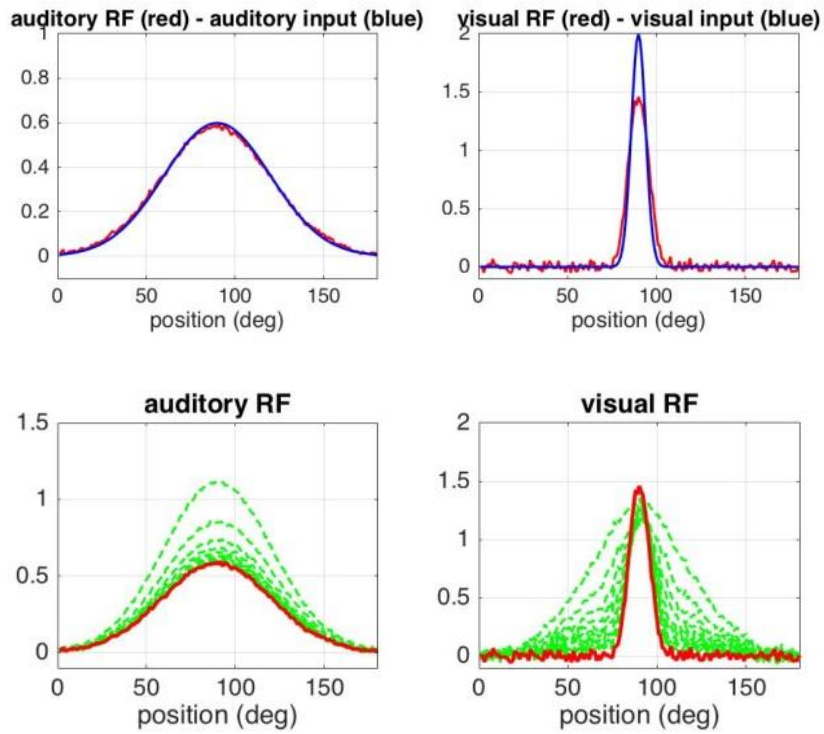


Figura 32 campi recettivi

I campi recettivi, nonostante i cambiamenti effettuati, riescono bene a ricostruire la forza e la forma dello stimolo acustico e visivo.

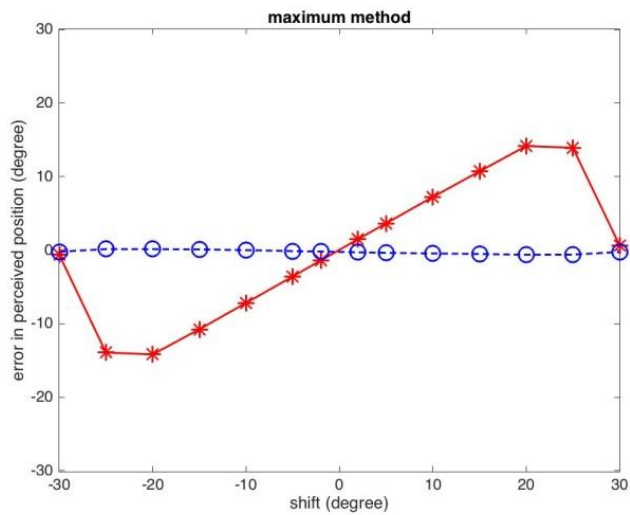
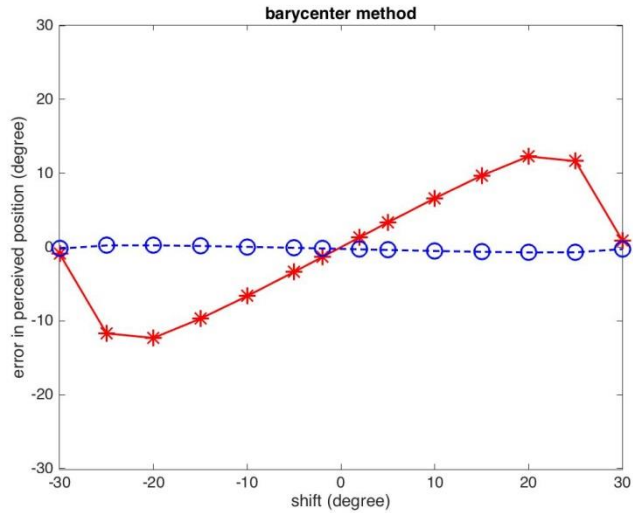


Figura 33 ventriloquismo calcolato con il metodo della massima verosimiglianza



**Figura 34 ventriloquismo calcolato con il metodo del baricentro**

Il ventriloquismo, calcolato con entrambi i metodi già nominati, cresce e si allarga.

Il suo effetto aumenta di molto.

➤  $\sigma^A = 40$   $\sigma^V = 4$ ;

La deviazione standard delle sinapsi uditive è stata aumentata, mentre quella delle sinapsi visive è stata lasciata inalterata. Per avere risultati soddisfacenti, in questo caso abbiamo bisogno, non solo di aumentare l'ampiezza della forza dello stimolo acustico ma anche di ridurre il rumore relativo agli ingressi iniziali. Infatti, essendo lo stimolo acustico spazialmente molto piatto, un eccessivo rumore può portare alla presenza di più vincitori nell'area acustica, cioè alla formazione di più zone attivate lontane fra loro, falsando i risultati.

Le variazioni adottate, sono state dunque, oltre che nelle deviazioni standard, nella forza degli stimoli:

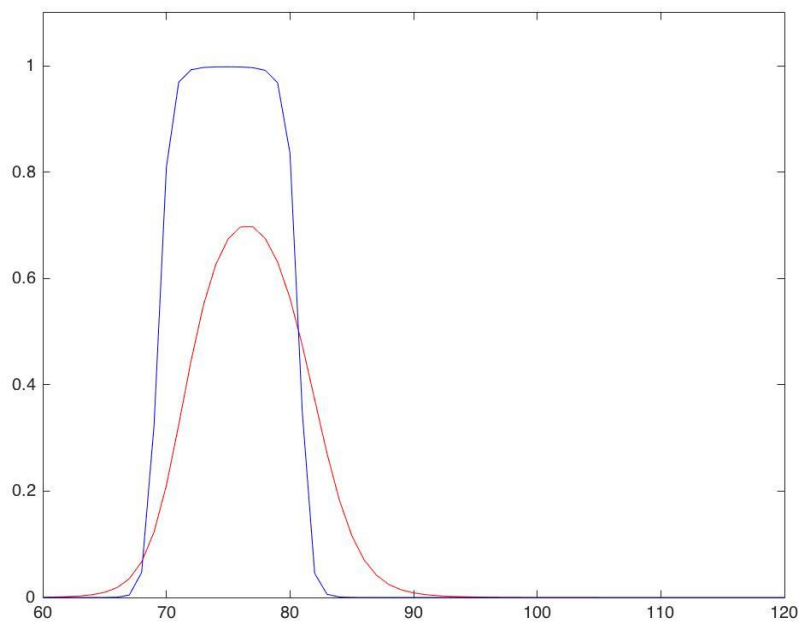
$I0_a = 45 / (\text{sqrt}(2 * \text{pi} * \text{sigma}_a * \text{sigma}_a));$

$$I0_v = 20/(\text{sqrt}(2*\text{pi}*\text{sigma}_v*\text{sigma}_v));$$

e nelle deviazioni standard del rumore:

$$\text{sigma}_{na} = I0_a/4; \quad \% \text{ Deviazione standard del rumore auditivo}$$

$$\text{sigma}_{nv} = I0_v/4; \quad \% \text{ Deviazione standard del rumore visivo}$$



**Figura 35** curva degli stimoli acustico (rosso) e visivo (blu)

dove la curva rossa rappresenta la stimolo acustico, mentre quella blu lo stimo visivo.

Le posizioni spaziali date inizialmente ai due stimoli era  $90^\circ$  per quello acustico e  $75^\circ$  per quello visivo. Nonostante il peggioramento drastico nella precisione dello stimolo acustico, aumentando la sua forza e diminuendo il rumore che gli apparteneva in ingresso, siamo riusciti ad ottenere un grafico tuttavia soddisfacente. I campi recettivi, ricostruiscono abbastanza bene i due stimoli.



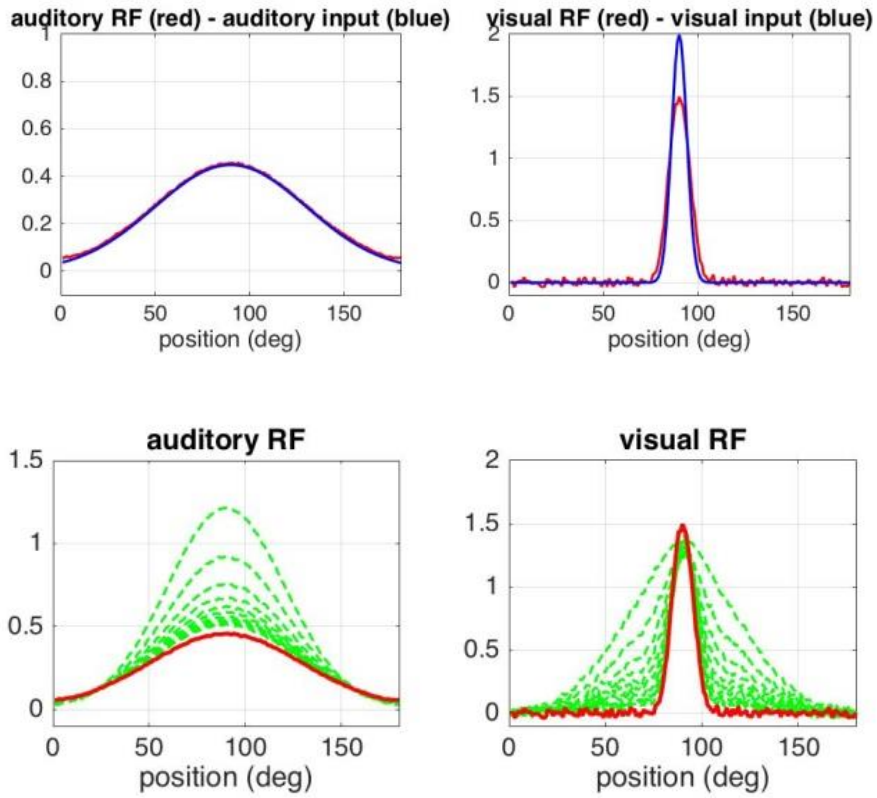


Figura 36 Campi recettivi

In queste condizioni il ventriloquismo diviene “perfetto”, cioè la localizzazione acustica va a coincidere con quella visiva.

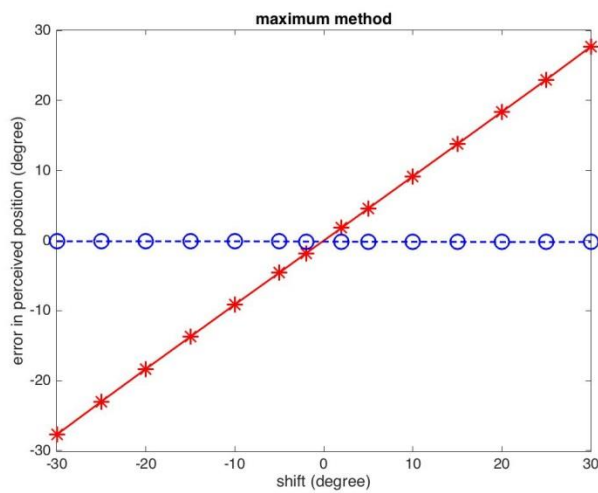
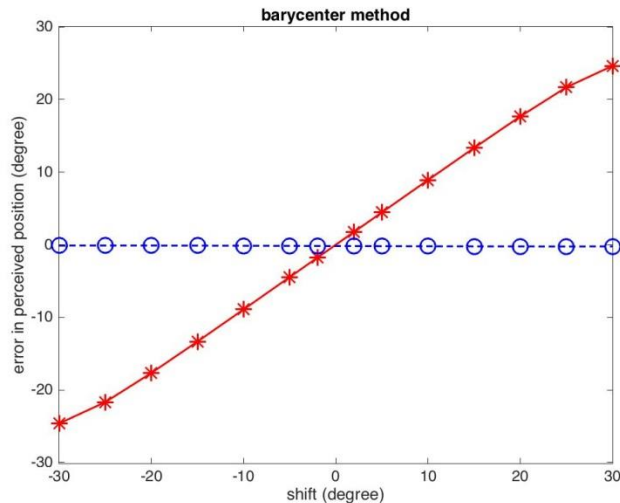


Figura 37 ventriloquismo calcolato con il metodo della massima verosimiglianza



**Figura 38 ventriloquismo calcolato con il metodo del baricentro**

Abbiamo effettuato prove fino ad un massimo di  $\sigma^A = 50$  e  $\sigma^V = 4$  ed abbiamo ottenuto la riprova che la curva relativa al ventriloquismo all'aumentare della larghezza dello stimolo acustico, scenda prima. In queste condizioni il ventriloquismo diventa fortissimo.

### 5.3 Variazioni delle sinapsi laterali

La seconda analisi di sensitività che è stata effettuata in questa tesi ha preso in considerazione la deviazione standard delle sinapsi laterali eccitatorie ed inibitorie. I risultati ottenuti hanno evidenziato un cambiamento riguardante soprattutto l'effetto del ventriloquismo al variare dell'inibizione e dell'eccitazione.

In generale infatti, questo risulta essere più forte al ridursi dell'inibizione.

➤  $\sigma_{ex} = 10 \quad \sigma_{in} = 30 :$

In questo caso, (ricordando quelli che erano I valori imposti di base  $\sigma_{ex}=12$  e  $\sigma_{in}=24$ ) entrambe le sinapsi laterali inibitorie e eccitatorie sono state leggermente alzate. Si nota sul fenomeno del ventriloquismo, una grossa influenza delle variazioni effettuate. La curva risulta infatti più stretta e tende a scendere sempre più velocemente, effetto dovuto probabilmente alle sinapsi laterali inibitorie che sono cresciute. Il picco del ventriloquismo rimane pressoché inalterato.

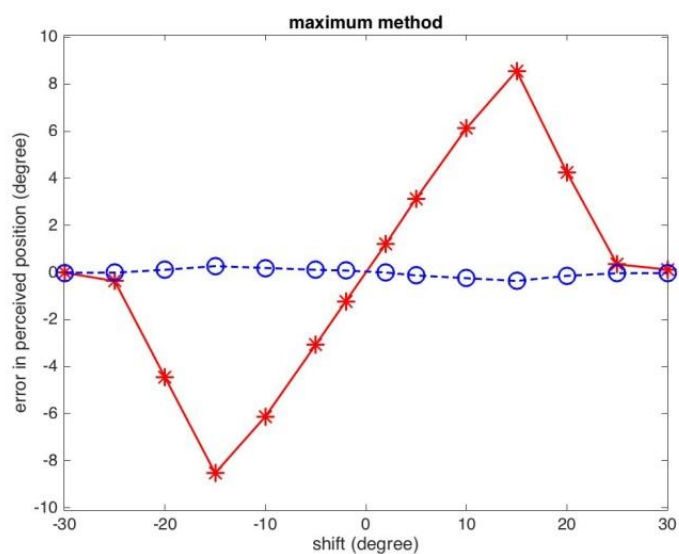


Figura 39 ventriloquismo calcolato con il metodo della massima verosimiglianza

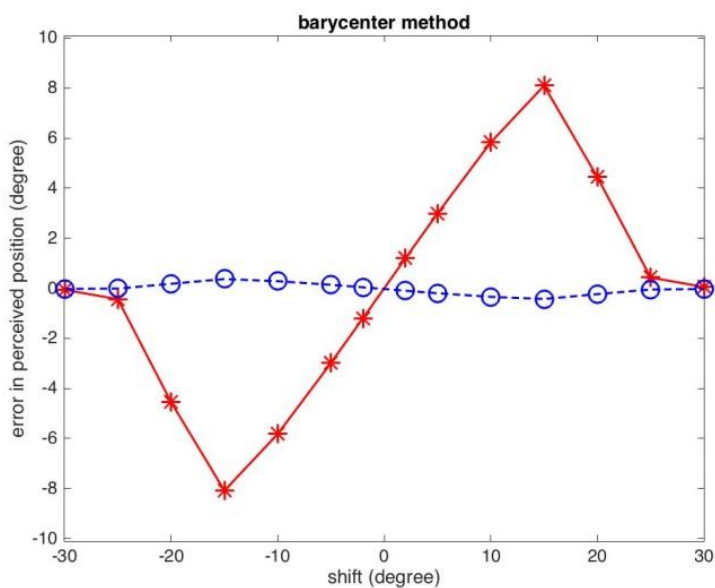


Figura 40 ventriloquismo calcolato con il metodo del baricentro

➤  $\sigma_{ex}=12$   $\sigma_{in}=50$  :

Provando, rispetto al caso precedentemente simulato, ad aumentare ulteriormente le sinapsi laterali inibitorie lasciando inalterate quelle eccitatorie, verificiamo quello che avevamo già intuito. Ovvero la curva relativa al fenomeno del ventriloquismo scende più velocemente all'aumentare dell'inibizione. Il grafico risulta come “schiacciato” su valori già precedentemente visti. Possiamo allora affermare che il ventriloquismo si verifica per distanze spaziali minori di quelle ottenute con i valori basali dei parametri.

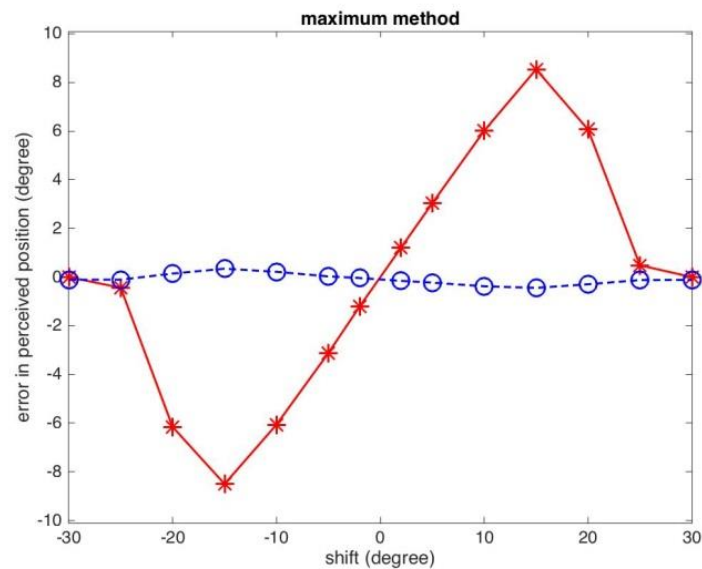


Figura 41 ventriloquismo calcolato con il metodo della verosimiglianza

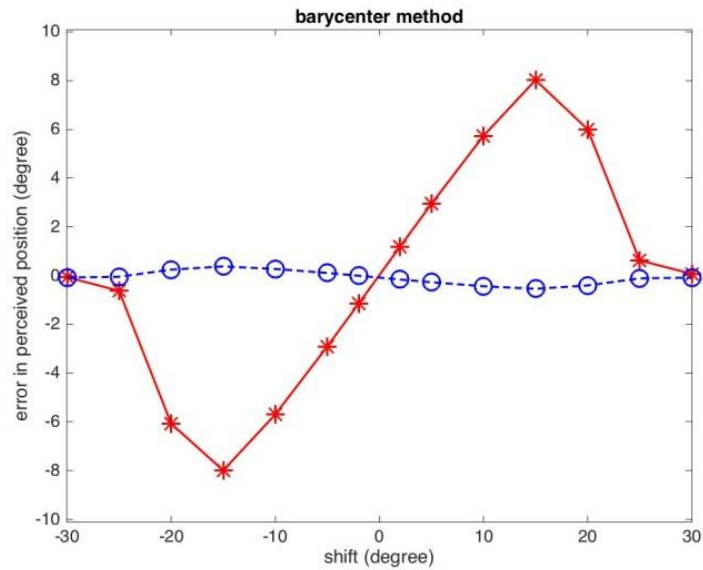
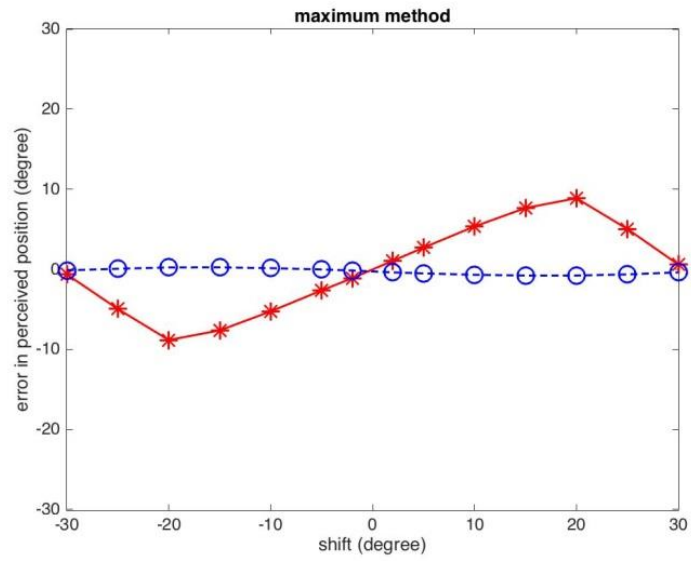


Figura 42 ventriloquismo calcolato con metodo del baricentro

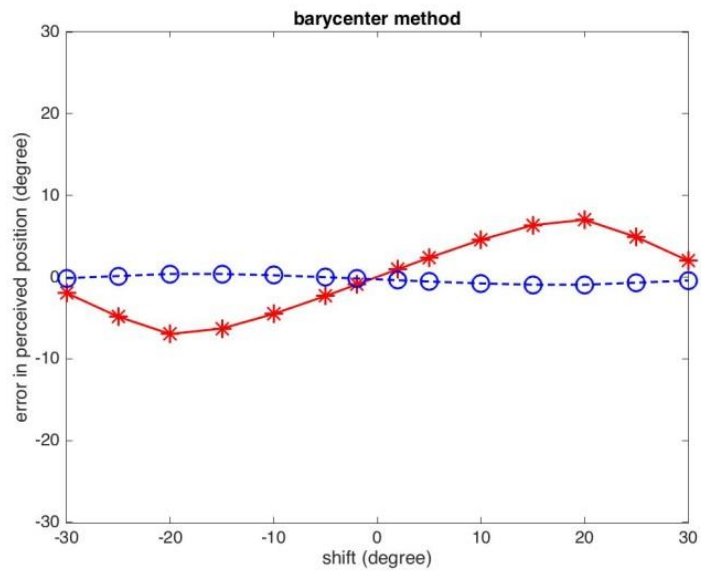
➤  $\sigma_{ex} = 12$   $\sigma_{in} = 16$  :

In questa prova, invece, si sono mantenute costanti le sinapsi laterali eccitatorie, mentre si sono abbassate quelle inibitorie. L'effetto di questa scelta mira ad evidenziare come con tali valori il ventriloquismo si allarghi rispetto ai casi precedentemente trattati. La curva risulta molto più allungata e tende a scendere più tardi spazialmente parlando.

Il ventriloquismo, diminuendo l'effetto inibitorio delle sinapsi laterali, aumenta.



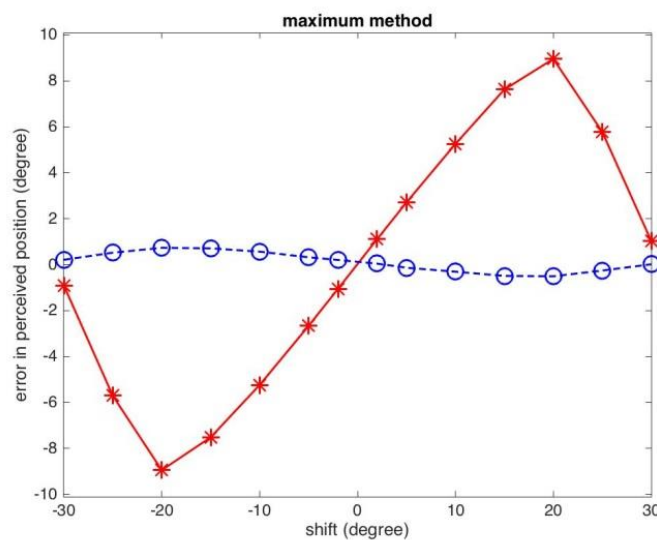
**Figura 43** ventriloquismo calcolato con il metodo della massima verosimiglianza



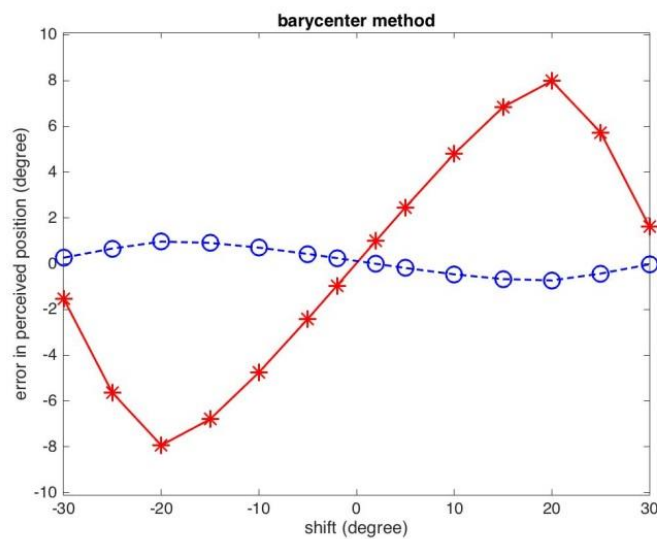
**Figura 44** ventriloquismo calcolato con il metodo del baricentro

➤  $\sigma_{ex} = 16 \sigma_{in} = 24$  :

Vediamo ora invece cosa succede se lasciando inalterate le sinapsi laterali inibitorie, aumentiamo quelle eccitatorie. Il ventriloquismo risulta praticamente inalterato; i risultati ottenuti rivelano che l'aumento delle sinapsi laterali eccitatorie non influisca in modo apprezzabile sull'entità del fenomeno del ventriloquismo.



**Figura 45** ventriloquismo calcolato con il metodo della massima verosimiglianza



**Figura 46** ventriloquismo calcolato con il metodo del baricentro

Come conclusione di questa seconda analisi di sensitività possiamo affermare che l'allargamento spaziale del fenomeno ventriloquismo si è ottenuto con i parametri delle sinapsi laterali eccitatorie e inibitorie in  $\sigma_{ex}=12$   $\sigma_{in}=16$ , ovvero lasciando inalterata l'eccitazione e diminuendo l'inibizione laterale.

Osserviamo in unico grafico le variazioni conseguenti all'utilizzo di dei diversi parametri:

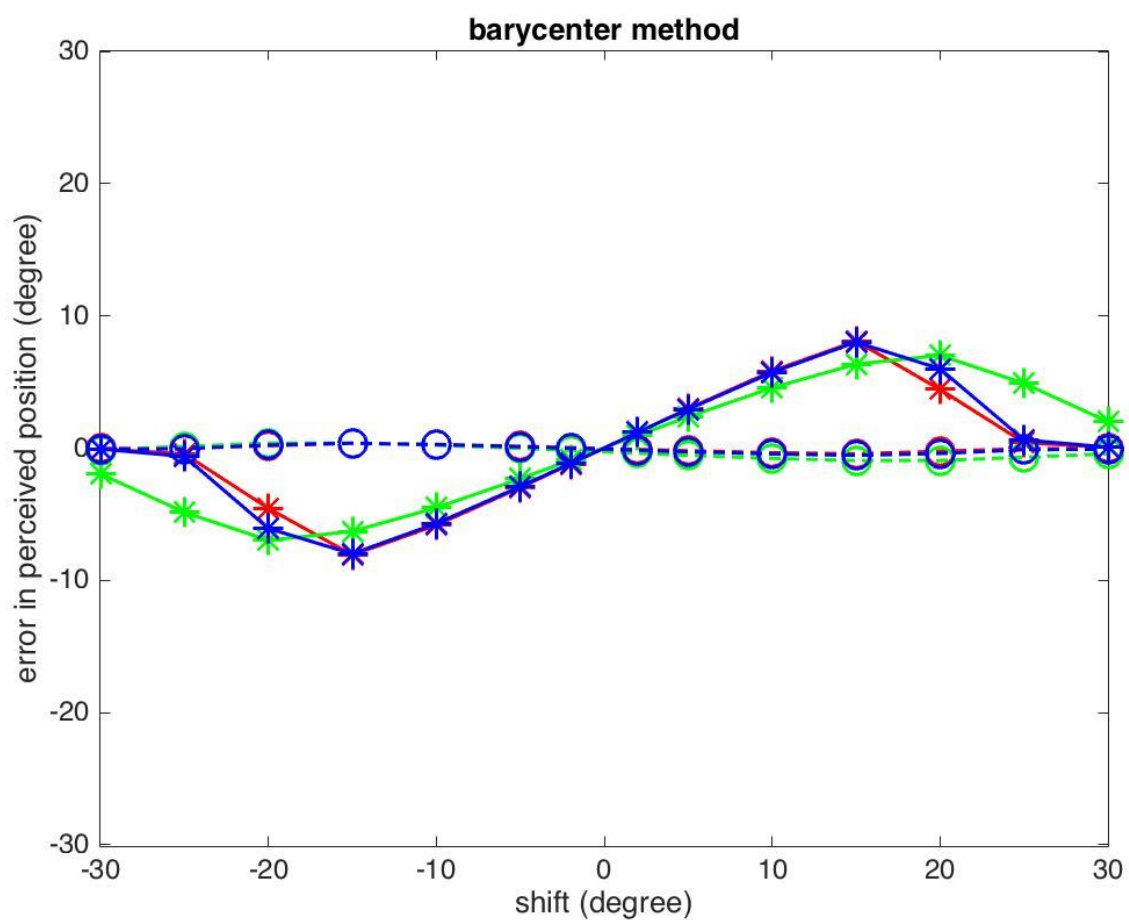
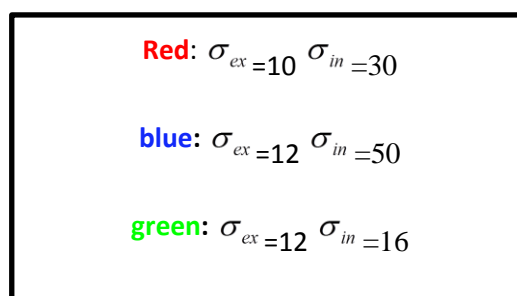


Figura 47 ventriloquismo a confronto: calcolato con diverse variazioni dei parametri. Si è rappresentato con il colore rosso





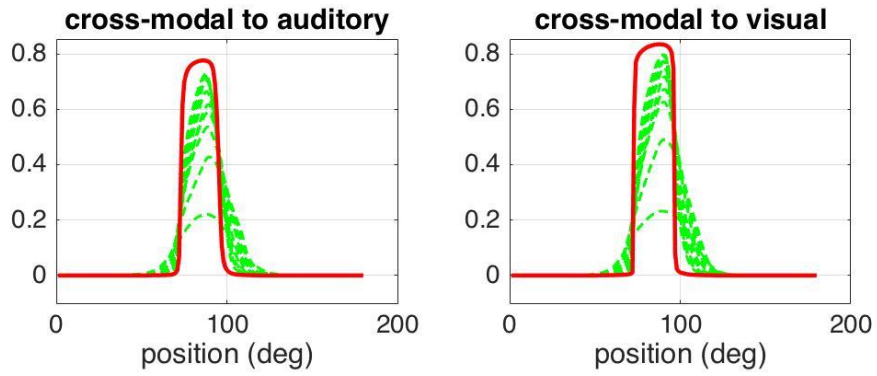
#### 5.4 variazioni della frequenza degli stimoli multisensoriali

L'ultima analisi di sensitività effettuata nel corso di questo progetto è stata eseguita variando la frequenza degli stimoli cross-modalità rispetto a quelli unimodali e lasciando inalterati il resto dei parametri che erano stati invece variati nelle precedenti simulazioni. La frequenza di stimolazione multisensoriale è stata modificata scegliendo per ogni prova un opportuno numero di "gate di stimolazione". Nella prova di base erano 5 ( stimolazione acustica (A), stimolazione visiva (V) ,stimolazione acustica (A), stimolazione visiva (V), stimolazione multisensoriale(AV)), corrispondenti quindi a una frequenza di stimoli cross-modalità pari a 1/3. Ovvero, ogni neurone (acustico o visivo) riceveva uno stimolo cross-modale per ogni due stimoli unimodali. In queste nuove simulazioni sono stati presi come oggetto di analisi i casi in cui i gate di stimolazione erano:

- 5 (A)(V)(AV)(AV)(AV)
- 6 (A)(V)(A)(V)(AV)(AV)
- 7 (A)(V)(A)(V)(A)(V)(AV)

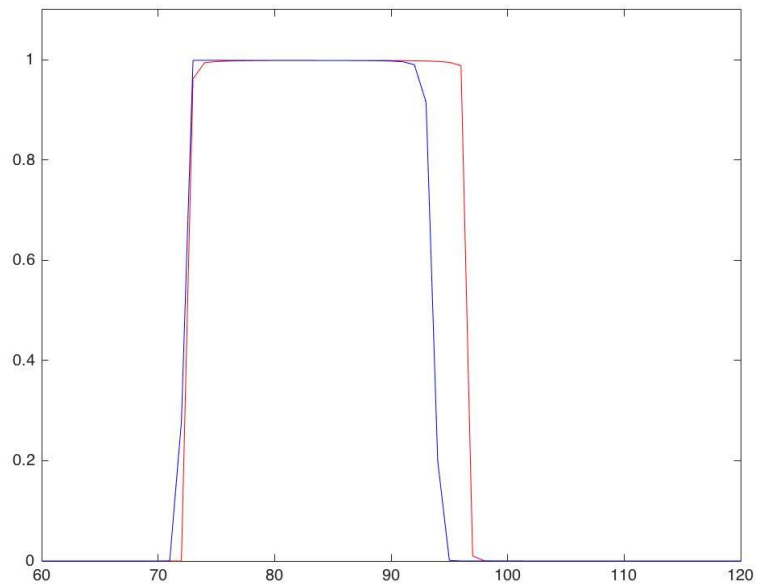
La scelta di avere ben 3 stimoli cross-modalità ogni 5 ( (A)(V)(AV)(AV)(AV) ) (frequenza pari a 3/4) ha mostrato che non sono presenti forti sinestesie nonostante la presenza di stimoli forti.

Le sinapsi cross-modalità come prevedibile, sono molto alte, a causa della presenza di tre stimolazioni multisensoriali rispetto ad una sola relativa invece alle condizioni di base.



**Figura 48 Sinapsi Cross-modali**

Nel grafico della prova ottenuto attraverso l'utilizzo del programma "usa\_doppiacatena" ci aspetteremo dunque che la stimolazione acustica sia sopra quella visiva.



**Figura 49 curva delle due stimolazioni acustica (rossa) e visiva (blu)**

La maggioranza di stimolazioni multisensoriali ha un forte effetto sui risultati di questa simulazione. Il ventriloquismo risulta essere molto forte e tende a portare lo stimolo acustico su quello visivo tanto da riuscirli a vedere quasi sovrapposti. Ovvero, Il ventriloquismo è quasi perfetto.

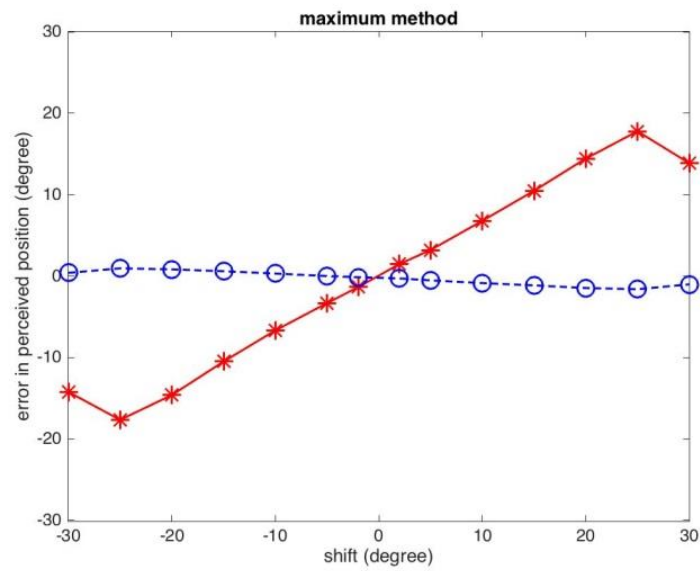


Figura 50 ventriloquismo calcolato con il metodo della massima verosimiglianza

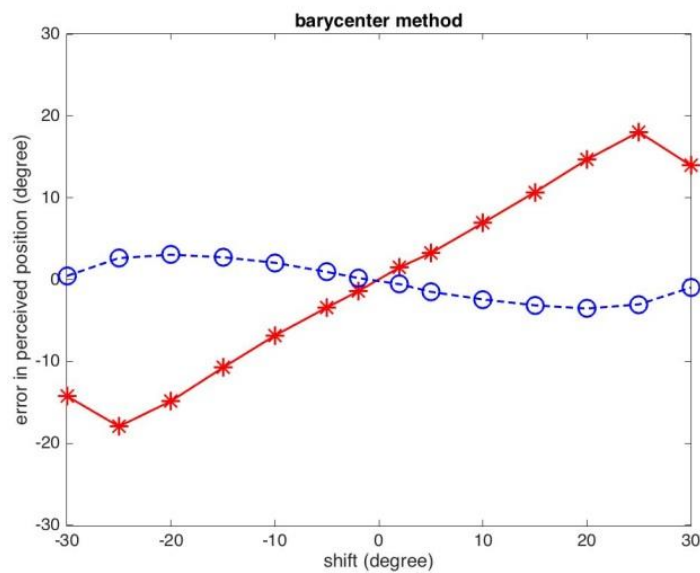
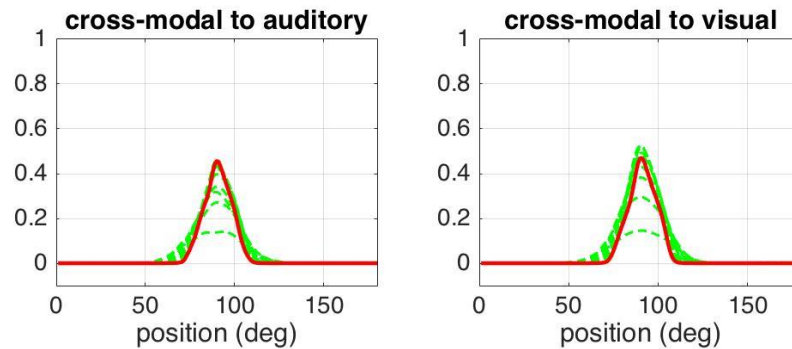


Figura 51 ventriloquismo calcolato con il metodo del baricentro

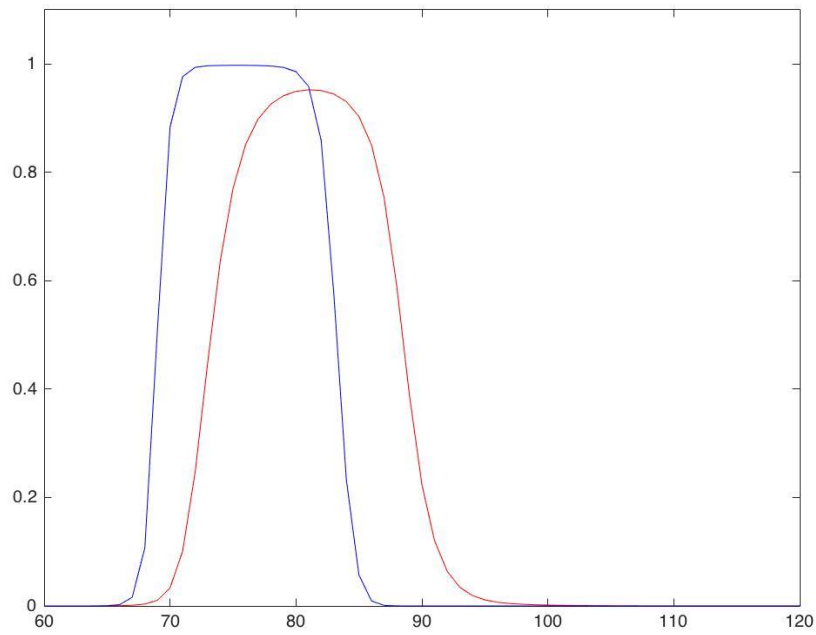
In seguito lo stesso programma è stato fatto funzionare con ben **6 gate** ( (A)(V)(A)(V)(AV)(AV) ), ovvero con una stimolazione multisensoriale in più rispetto le condizioni di base (frequenza pari a 1/2).

Le sinapsi crossmodali tendono circa al valore di 0.5.



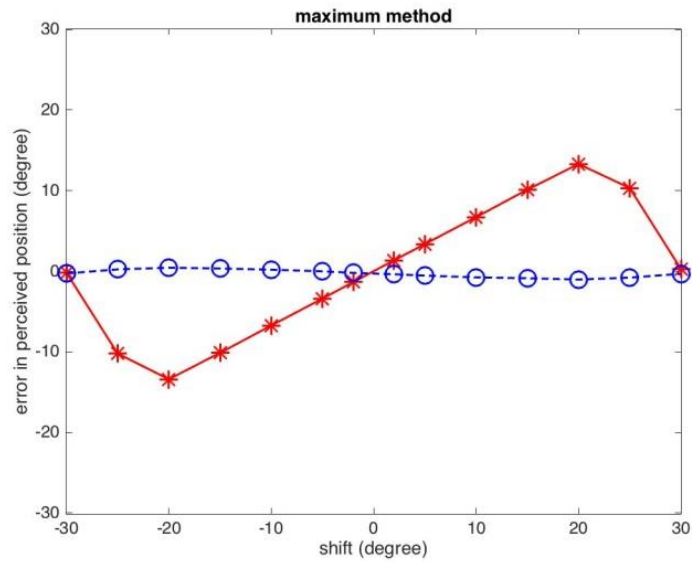
**Figura 52 Sinapsi cross-modali**

Lo stimolo acustico e visivo si avvicinano molto ma non sono così sovrapposti come invece succederebbe se utilizzassi più stimolazioni multisensoriali (vedi stimolazione precedente con 5 gate ((A)(V)(AV)(AV)(AV))).

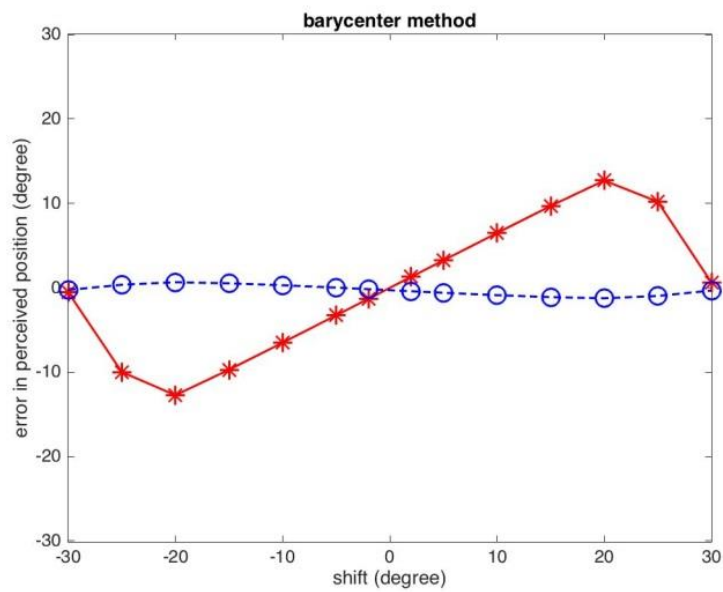


**Figura 53** Curva della stimolazione acustica (rossa) e visiva (blu)

Il ventriloquismo anche in questo caso cresce rispetto a quello relativo alle condizioni di base, anche se tuttavia risulta essere non così perfetto come invece l'avevamo visto nella simulazione precedente con frequenza  $3/4$ , dove la presenza di più stimoli multisensoriali rispetto a quelli unisensoriali influenzava molto tale fenomeno.



**Figura 54 ventriloquismo calcolato con il metodo di massima verosimiglianza**

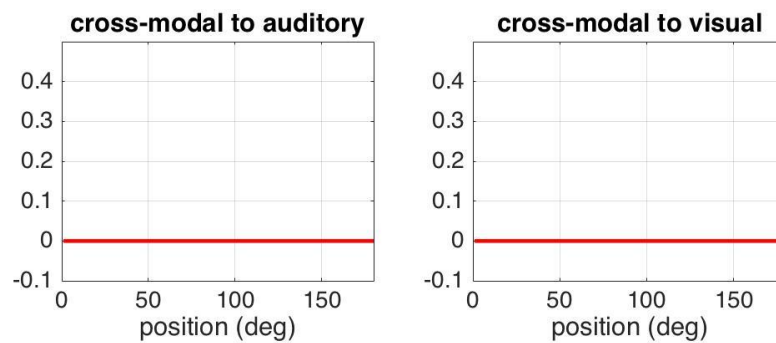


**Figura 55 ventriloquismo calcolato con il metodo del baricentro**

L'ultima simulazione relativa alla variazione della frequenza di stimolazione cross-modale è stata effettuata con 7 gate ((A)(V)(A)(V)(A)(V)(AV)).

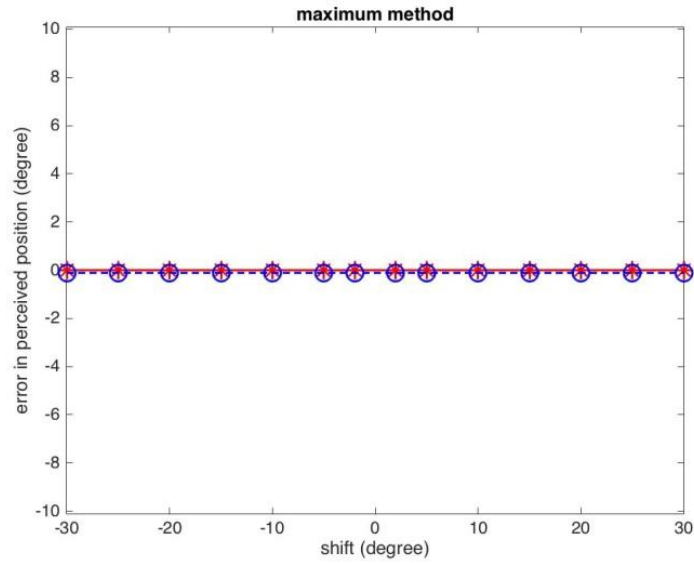
Rispetto la condizione di base ((A)(V)(A)(V)(AV)) sono state aggiunte due stimolazioni unisensoriali, una acustica e una visiva, riducendo così la frequenza di input cross-modali per ogni neurone a 1/4.

Questa variazione ha messo in evidenza il fatto che la legge di addestramento sotto un certo limite non riesce più a creare sinapsi cross-modali. Infatti, se gli input cross-modali risultano troppo poco frequenti rispetto quelli unisensoriali, le sinapsi cross-modali non vengono create in numero adeguato. Le sinapsi cross-modali non si creano.

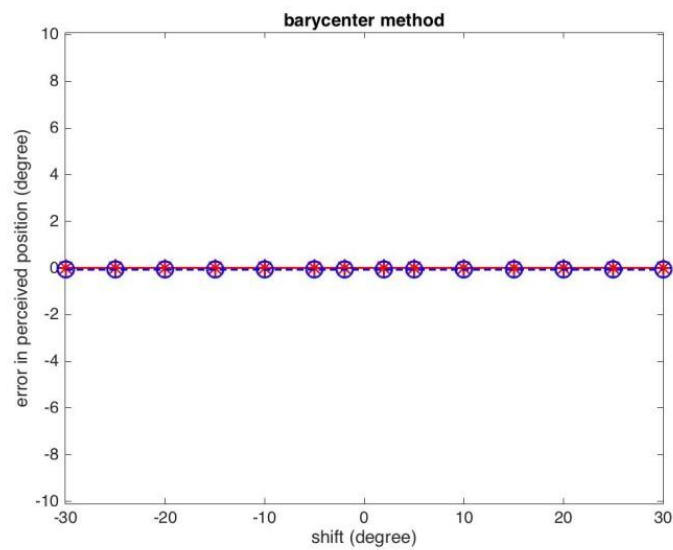


**Figura 56 sinapsi Cross-modali**

Il ventriloquismo non si presenta: il numero elevato di stimolazioni unisensoriali rispetto a quello delle stimolazioni multisensoriali rende questo fenomeno nullo.



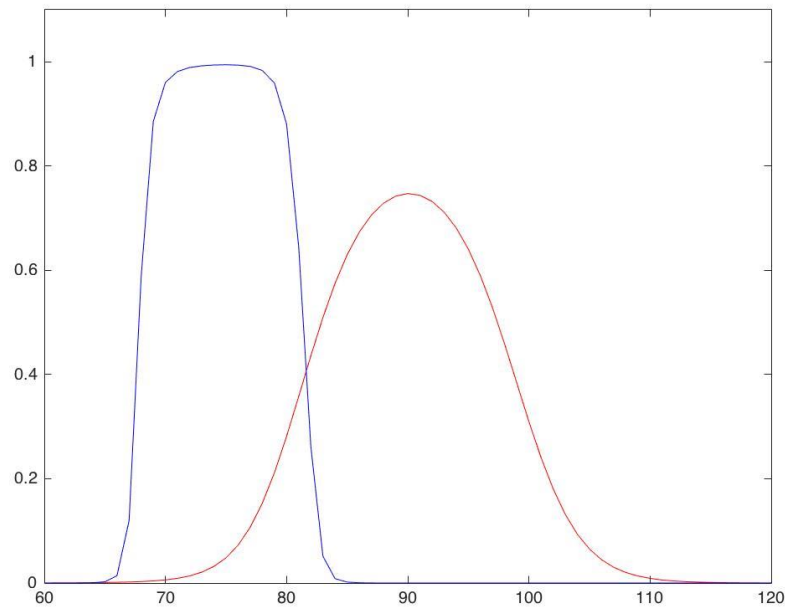
**Figura 57 ventriloquismo calcolato con il metodo della massima verosimiglianza**



**Figura 58 ventriloquismo calcolato con il metodo del baricentro**

I due stimoli, quello acustico e quello visivo, si muovono relativamente poco l'uno rispetto all'altro proprio a causa della mancanza delle sinapsi cross-modali.





**Figura 59** curva della stimolazione acustica (rossa) e visiva (blu)

Per cercare di risolvere questa problematica, si è cercato di modificare a sua volta questa simulazione modificando dei parametri quali le  $\sigma^A$  e  $\sigma^V$  prima mantenuti fisse a valori base. Osserviamo i risultati ottenuti quindi, sempre utilizzando 7 gate ((A)(V)(A)(V)(A)(V)(AV)) ma questa volta utilizzando  $\sigma^A=30$   $\sigma^V=4$  (al posto di  $\sigma^A=20$   $\sigma^V=4$  imposte nelle condizioni base). Questa volta il ventriloquismo torna ad essere presente, le sinapsi cross-modali infatti riescono a crearsi grazie all'aumento del parametro  $\sigma^A$ . Gli stimoli unisensoriali acustico e visivo riescono visibilmente meglio a spostarsi l'uno sull'altro, nonostante lo stimolo acustico risulti però molto debole a causa della variazione parametriche aggiuntiva effettuata.

I grafici rendono ben visibile il miglioramento apportato dalle nostre variazioni parametriche.

Le sinapsi cross-modali:

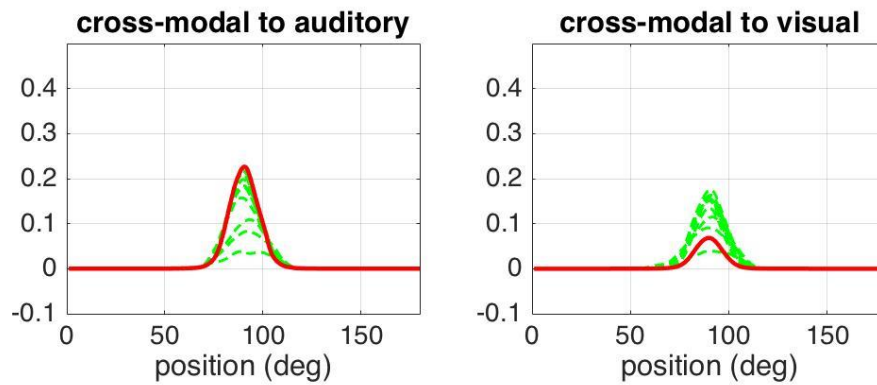


Figura 60 sinapsi Cross-modali

Occorre notare che in questa condizione le sinapsi diventano fortemente asimmetriche: sono molto più forti quelle dirette verso l'uditivo, probabilmente a causa del basso valore di attività auditiva rispetto a quella visiva.

La prova:

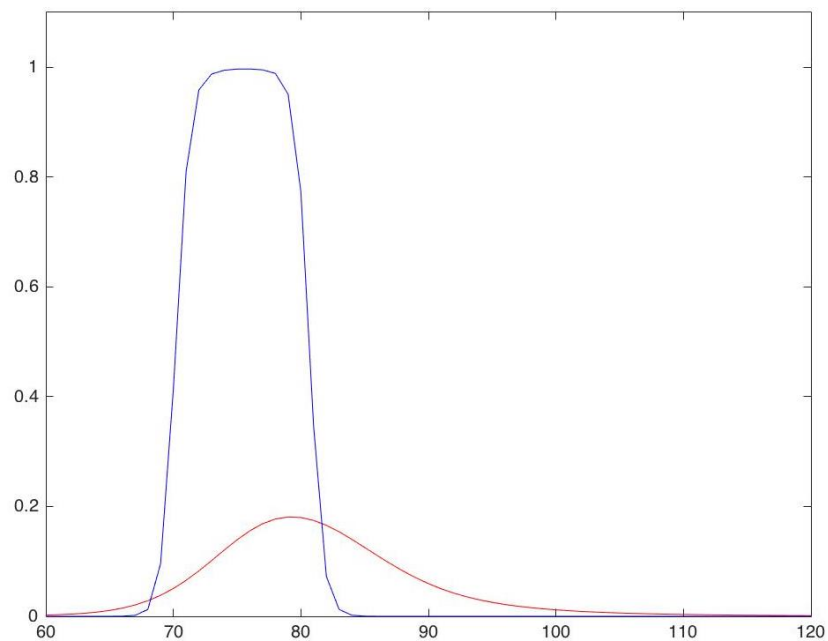


Figura 61 curva della risposta alla stimolazione acustica (rossa) e visiva (blu)

Il ventriloquismo:

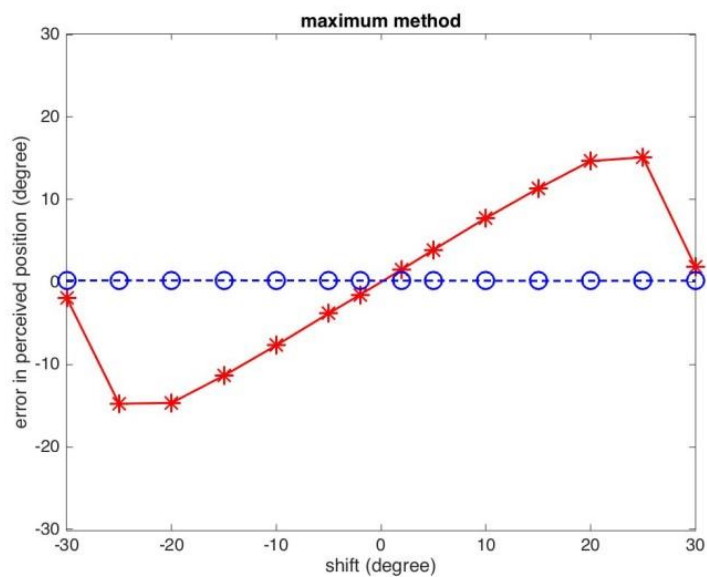


Figura 62 ventriloquismo calcolato con il metodo della massima verosimiglianza

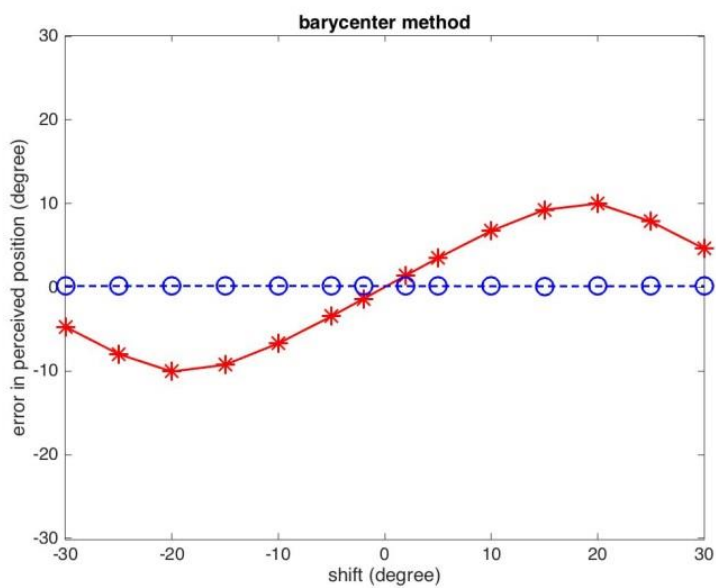


Figura 63 ventriloquismo calcolato con il metodo del baricentro



## Conclusioni

Neuroni in grado di rispondere a stimoli di diversa modalità sensoriale sono stati individuati in diverse regioni cerebrali (Stein & Meredith, 1993, Calvert, Hansen, Iversen, & Brammer, 2001, Calvert, Spence & Stein, 2004, Macaluso & Driver, 2005). Fino a poco tempo fa gli studi sull'integrazione multisensoriale erano concentrati su un numero ridotto di regioni cerebrali, principalmente sul collicolo superiore e su particolari aree associative. Studi più recenti hanno mostrato che l'integrazione multisensoriale avviene anche a livello della corteccia primaria, e che quindi, in questo caso specifico, le cortecce primarie della modalità visiva e uditiva presentano delle funzioni multisensoriali. L'integrazione delle informazioni che questi neuroni operano sembra svolgere un ruolo fondamentale nell'orientare il comportamento sulla base delle informazioni provenienti dall'esterno. L'interazione audio-visiva non è un'abilità presente già alla nascita, ma viene acquisita progressivamente durante lo sviluppo in ambiente multisensoriale. Infatti i neuroni visivi e acustici di un neonato sono costituiti da campi recettori molto larghi e risposte cross-modali molto deboli (Stein et al. 1973a; Stein et al. 1973b). Durante la loro maturazione tali neuroni sviluppano le loro capacità di integrazione multisensoriale e la loro risposta diventa più robusta (Wallace et al., 2004; Wallace and Stein, 1997).

Il modello matematico proposto nella presente Tesi ha cercato di ricreare una rete neurale che simulasse uno stimatore bayesiano in grado di comprendere e predire come una rete neurale si comportasse in un ambiente cross-modale. Una volta addestrata, la rete è stata utilizzata per simulare gli effetti dell'integrazione audio-visiva nella corteccia primaria, quali per esempio il ventriloquismo spaziale. Il modello è stato addestrato inizialmente con sequenze di input visivi, acustici e audio-visivi in condizioni basali cercando di simulare il comportamento di uno stimatore bayesiano per permettere di comprendere il funzionamento della rete e predirne il comportamento a fronte di successive variazioni.

Questa simulazione in particolare, rispecchia in pieno gli obiettivi posti in questa Tesi:

- I campi recettori della rete mostrano un ottimo addestramento se sottoposti a sequenze di stimoli unisensoriali (visivi e acustici) e cross-modali, permettendo in particolare di discriminare la percezione spaziale degli input con maggior precisione.
- Le sinapsi cross-modali apprendono che sussiste una correlazione tra la posizione dello stimolo acustico e di quello visivo.
- La rete addestrata si comporta secondo le previsioni, riproducendo alcuni effetti dell'integrazione audio-visiva nella corteccia primaria, tra cui il ventriloquismo spaziale da noi studiato.

La rete in seguito è stata addestrata in tre casi diversi (“Analisi di sensitività”, Capitolo V), con sinapsi laterali visive e acustiche variabili, con frequenza di stimolazione diversa e modificando la precisione della stimolazione.

In tutte e tre le simulazioni si è potuto notare come:

- nel primo caso, variando l'ampiezza della sinapsi laterali eccitatorie e inibitorie la rete risponde in modo più forte al fenomeno del ventriloquismo con il diminuire dell'inibizione laterale;
- nel secondo caso, attraverso la variazione della frequenza di stimolazione, si è evidenziato il ruolo delle sinapsi Cross-modali. Un ventriloquismo più forte si è ottenuto infatti, all'aumentare di quest'ultime. Simulazioni effettuate con stimolazioni multisensoriali poco frequenti hanno evidenziato un ventriloquismo nullo, mentre al contrario, una forte componente Cross-modale, ottenuta da prove con elevata frequenza di stimoli cross-modali, ci ha permesso una rappresentazione del fenomeno del ventriloquismo più accentuata;
- nel terzo caso, attraverso la variazione della precisione degli stimoli, ovvero variando le deviazioni standard rispettivamente dello stimolo acustico e visivo, si è giunti alla

conclusione che la precisione della stimolazione sensoriale è inversamente proporzionale al fenomeno del ventriloquismo da noi studiato. Uno stimolo acustico più preciso infatti, ha portato ad un indebolimento evidente del ventriloquismo mentre uno con deviazione standard più elevata (e quindi meno preciso) ha mostrato un ventriloquismo più forte.

Secondo quanto detto fino ad ora, il modello simulato nella prima prova (Capitolo IV) con valori basali, e la successiva analisi di sensitività (Cap. V) possono contribuire matematicamente a comprendere i meccanismi neurali fisiologici alla base dell'apprendimento audio-visivo in seguito a esperienze cross-modali e unisensoriali.

In tale modo è stato reso chiaro che:

- Con una regola di apprendimento realistica, la rete neurale può imparare la funzione di verosimiglianza (campi recettivi);
- le Sinapsi Cross-modali includono informazioni a priori sulla co-occorrenza degli stimoli;
- i segnali sono ponderati in base alla loro affidabilità (effetto ventriloquismo, effetto fissione) e le interazioni cross-modali aiutano a pesarli automaticamente favorendo quelli più affidabili.

Molto importante risulta soprattutto questo ultimo punto, dove è evidente il ruolo determinante che ha assunto lo studio del fenomeno del ventriloquismo in questo lavoro. Lo stimolo acustico viene infatti percepito non nella sua reale posizione bensì spostato verso quello visivo. Non è invece presente alcuna influenza dello stimolo acustico nei confronti di quello visivo, come dimostra la quasi totale assenza di spostamento della localizzazione visiva. Il ventriloquismo emerge dall'interazione tra l'area A e l'area V e dalla presenza di sinapsi laterali all'interno di ciascuna area,

le uniche ipotesi introdotte sono che tali aree siano spazialmente organizzate. I fenomeni di integrazione multisensoriale non sono introdotti a priori nella rete ma emergono dalla interazione e dalle connessioni tra le diverse aree, a partire da poche ipotesi inserite.

I punti di forza del presente modello sono tanti: per prima cosa è interamente basato su meccanismi plausibili dal punto di vista neurobiologico; inoltre utilizzando un unico set di parametri (condizione di base) è in grado di simulare diversi fenomeni di integrazione visuoacustica in accordo con i risultati di letteratura e di riprodurre un'ampia casistica di studi comportamentali neuropsicologici; infine cambiando i parametri del modello, ma mantenendo intatta l'organizzazione, è possibile riprodurre la variabilità del comportamento in vivo ed è in grado di effettuare delle predizioni; il tutto avendo tuttavia un grado di complessità moderato. Il modello fornisce quindi un valido strumento per interpretare la realtà biologica.

La rete, tramite opportune estensioni, potrebbe essere utilizzata più approfonditamente in futuro per riprodurre uno stimatore bayesiano in condizioni in cui i due stimoli cross-modali possano provenire anche da sorgenti diverse (e quindi essere collocati a maggiore distanza spaziale) permettendoci di avere una previsione ancora più completa del modello. Potrebbero inoltre essere effettuati test soggettivi per verificare realmente l'efficacia di tale stimatore bayesiano e generalizzarlo. Non è da escludere anche una possibile analisi del problema da un punto di vista diverso, quello del casual interference. In tale ambito, il modello dovrebbe prima inferire la presenza di una o due sorgenti per gli stimoli cross-modali e solo successivamente, sulla base di tale stima, stimare la posizione dell'input visivo e auditivo. Fino ad ora infatti, il modello è stato simulato con una causa alla volta; sarebbe interessante in futuro utilizzare ad esempio due cause e creare una rete neurale che riesca a simulare uno stimatore bayesiano che consideri la possibilità di una o due cause e ci fornisca le giuste predizioni.





## Bibliografia:

Alais, D., & Burr, D. (2004). *The ventriloquism effect results from near-optimal integration*. *Current Biology* (14), 257-262.

Alexandre Pouget, Jeffrey M Beck, Wei Ji Ma & Peter E Latham (2013). *Probabilistic brains: knowns and unknowns*. Review by nature neuroscience.

Alexandre Pouget, Peter Dayan and Richard S. Zemel (2003). *INFERENCE AND COMPUTATION WITH POPULATION CODES*. Department of Brain and Cognitive Sciences, Meliora Hall, University of Rochester, Rochester, New York, 14627; Gatsby Computational Neuroscience Unit, Alexandra House, 17 Queen Square, London WC1N 3AR, United Kingdom; Department of Computer Science, University of Toronto, Toronto, Ontario M5S 1A4.

Alvarado, JC, Stanford, TR, Vaughan, JW and Stein, BE (2007). *Cortex mediates multisensory but not unisensory integration in superior colliculus*. *J Neurosci* 27(47): 12775-12786.

Alvarado, JC, Vaughan, JW, Stanford, TR and Stein, BE (2007). *Multisensory versus unisensory integration: contrasting modes in the superior colliculus*. *J Neurophysiol* 97(5): 3193-3205.

Bell, A., Meredith, M., Van Opstal, A., & Munoz, D. (2005). *Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements*. *Journal of Neurophysiology* (93), 3659-3673.

Bertelson, P., & De Gelder, B. (2004). The psychology of multisensory perception. In S. C. & D. J., *Crossmodal space and crossmodal attention* (p. 141–177). Oxford: Oxford University Press.

Bertelson, P., & Radeau, M. (1981). *Cross-modal bias and perceptual fusion with auditory-visual spatial*. *Perception & Psychophysics* (29), 578-584.

Bolognini, N., Leo, F., Passamonti, C., Stein, B., & Làdavas, E. (2007). *Multisensory-mediated auditory localization*. *Perception* (36), 1477-1485.

Bolognini, N., Miniussi, C., Savazzi, S., Bricolo, E., & Maravita, A. (2009). *TMS modulation of visual and auditory processing in the posterior parietal cortex*. *Exp Brain Res* , 195, 509–517.

Burnett, L., Stein, B., Perrault, J., & Wallace, M. (2007). *Excitotoxic lesions of the superior colliculus preferentially impact multisensory neurons and multisensory integration*. *Experimental Brain Research* (179), 325-338.

Bushara, K., Grafman, J., & Hallet, M. (2001). *Neural correlates of auditory-visual stimulus onset asynchrony detection*. *Journal of Cognitive Neuroscience* (21), 300-304.

Calvert, G., Campbell, R., & Brammer, M. (2000). *Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex*. *Current Biology* (10), 649–657.

Calvert, G., Hansen, P., Iversen, S., & Brammer, M. (2001). *Detection of multisensory integration sites by application of electrophysiological criteria to BOLD response*. *Neuroimage* (14), 427-438.

Calvert, G.A., Spence, C., & Stein, B.E. (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

Colonus, H and Diederich, A (2006). *The race model inequality: interpreting a geometric measure of the amount of violation*. Psychological Review 113: 148-154.

Cuppini C., Magosso E., Bolognini N., Vallar G., Ursino.M.(2014). *NeuroImage. A neurocomputational analysis of the sound-induced flash illusion*. Department of Electrical, Electronic and Information Engineering, University of Bologna, Bologna, Italy; Department of Psychology, University of Milano-Bicocca, Milan, Italy;IRCCS Istituto Auxologico Italiano, Milan, Italy.

David Alais and David Burr (2004). *The Ventriloquist Effect Results from Near-Optimal Bimodal Integration*. Istituto di Neuroscienze del CNR 56127 Pisa,Italy.

David Alais, Fiona N. Newell and Pascal Mamassian (2010). *Multisensory Processing in Review: from Physiology to Behaviour*. School of Psychology, University of Sydney, Australia; School of Psychology and Institute of Neuroscience, Trinity College Dublin, Ireland; Laboratoire Psychologie de la Perception, Université Paris Descartes, France.

Eimer, M. (2004). *Multisensory integration: how visual experience shapes spatial perception*. Current Biology (3), 115-117.

Ernst, MO and Banks, MS (2002). *Humans integrate visual and haptic information in a statistically optimal fashion*. Nature 415: 429-433.

Fodor, J. (1983). *The modularity of the mind*. MIT Press.

Ghazanfar, A., & Schroeder, C. (2006). *Is neocortex essentially multisensory?* Trends in Cognitive Sciences (10), 278–285.

Groh, JM and Sparks, DL (1996). *Saccades to somatosensory targets. III. eye-position-dependent somatosensory activity in primate superior colliculus.* J Neurophysiol 75(1): 439-453.

Hartline, PH, Vimal, RL, King, AJ, Kurylo, DD and Northmore, DP (1995). *Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats.* Experimental brain research. Experimentelle Hirnforschung 104(3): 402-408.

Howard, I., & Templeton, W. (1966). *Human spatial orientation.* London.

Jay, MF and Sparks, DL (1984). *Auditory receptive fields in primate superior colliculus shift with changes in eye position.* Nature 309(5966): 345-347.

Jiang, W, Jiang, H and Stein, BE (2002). *Two corticotectal areas facilitate multisensory orientation behavior.* Journal of Cognitive Neuroscience 14: 1240-1255.

Jiang, W, Jiang, H and Stein, BE (2006). *Neonatal cortical ablation disrupts multisensory development in superior colliculus.* J Neurophysiol 95(3): 1380-1396.

Jiang, W, Wallace, MT, Jiang, H, Vaughan, JW and Stein, BE (2001). *Two cortical areas mediate multisensory integration in superior colliculus neurons.* Journal of Neurophysiology 85: 506-522.

Kadunce, D., Vaughan, J., Wallace, M., & Stein, B. (2001). *The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus.* Experimental Brain Research (139), 303-310.

- Kadunce, DC, Vaughan, JW, Wallace, MT, Benedek, G and Stein, BE (1997). *Mechanisms of within- and cross-modality suppression in the superior colliculus*. *J Neurophysiol* 78(6): 2834-2847.
- Laurienti, P., Burdette, J., Wallace, M., Yen, Y., Field, A., & Stein, B. (2002). *Deactivation of sensory-specific cortex by cross-modal stimuli*. *Journal of Cognitive Neuroscience* (14), 420–429.
- Laurienti, P., Kraft, R., Maldjian, J., Burdette, J., & Wallace, M. T. (2004). *Semantic congruence is a critical factor in multisensory behavioral performance*. *Experimental Brain Research* (158), 405-414.
- Macaluso, E., & Driver, J. (2005). *Multisensory spatial interactions: a window onto functional integration in the human brain*. *Trends in Neuroscience* (28), 264–271.
- Magosso, E., Zavaglia, M., Serino, A., di Pellegrino, G., & Ursino, M. (2010). *Visuotactile representation of peripersonal space: a neural network study*. *Neural Computation*, 22, 190-243.
- McDonald, J. J., Teder-Salejarvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). *Neural substrates of perceptual enhancement by cross-modal spatial attention*. *Journal of Cognitive Neuroscience*, 15, 10-19.
- McDonald, J. J., Teder-Salejarvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). *Neural basis of auditory-induced shifts in visual time-order perception*. *Nature Neuroscience*, 8, 1197-1202.
- Meienbrock, A., Naumer, M., Doehrmann, O., Singer, W., & Muckli, L. (2007). *Retinotopic effects during spatial audio-visual integration*. *Neuropsychologia* (45), 531-539.

Meredith, M. (2002). *On the neural basis for multisensory convergence: a brief overview*. Cognitive Brain Research (14), 31-40.

Meredith, M., & Stein, B. (1986a). *Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration*. Journal of Neurophysiology (156), 640-662.

Meredith, M., & Stein, B. (1986b). *Spatial factors determine the activity of multisensory neurons in cat superior colliculus*. Brain Research (19), 350-304.

Meredith, M., Nemitz, J., & Stein, B. (1987). *Determinants of multisensory integration in superior colliculus: temporal factors*. Journal of Neuroscience (10), 3215-3229

Meredith, MA and Stein, BE (1996). *Spatial determinants of multisensory integration in cat superior colliculus neurons*. J Neurophysiol 75(5): 1843-1857.

Peck, CK, Baro, JA and Warder, SM (1995). *Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus*. Experimental brain research. Experimentelle Hirnforschung 103(2): 227-242.

Recanzone, G. H. (1998). *Rapidly induced auditory plasticity: the ventriloquism aftereffect*. Proceedings of the National Academy of Sciences of the United States of America, 95, 869-875.

Recanzone, GH (2003). *Auditory influences on visual temporal rate perception*. J Neurophysiol 89(2): 1078-1093.

- Rowland, B., Quessy, S., Stanford, T., & Stein, B. (2007). *Multisensory integration shortens physiological response latencies*. *Journal of Neuroscience* (22), 5879-5884.
- Rowland, B., Quessy, S., Stanford, T., & Stein, B. (2007). *Multisensory integration shortens physiological response latencies*. *Journal of Neuroscience* (22), 5879-5884.
- Schroeder, C. E. & Foxe, J. (2005). *Multisensory contributions to low-level, 'unisensory' processing*. *Current Opinion in Neurobiology*, 15, 454-458.
- Shams, L, Ma, WJ and Beierholm, U (2005). *Sound-induced flash illusion as an optimal percept*. *Neuroreport* 16(17): 1923-1927.
- Shams, L., & Kim, R. (2010). *Crossmodal influences on visual perception*. *Physics of Life Reviews*, 7, 269-284.
- Stein, B. E., Huneycutt, W., & Meredith, M. (1988). *Neurons and behavior: the same rules of multisensory integration apply*. *Brain Research* (448), 355–358.
- Stein, B., & Meredith, M. A. (1993). *Merging of senses*. Cambridge: MIT Press.
- Stein BE, Labos E, Kruger L. (1973a) *Determinants of response latency in neurons of superior colliculus in kittens*. *Journal of Neurophysiology* 36:680–689. [PubMed: 4713314]
- Stein BE, Labos E, Kruger L. (1973b) *Sequence of changes in properties of neurons of superior colliculus of the kitten during maturation*. *Journal of Neurophysiology* 36:667–679. [PubMed: 4713313]
- Stein, B. E. & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press



- Ursino M., Magosso E., Cuppini C. (2015). *A neurocomputational model of audio-visual multisensory integration in early cortical areas: Hebbian training and relationship with Bayesian estimators*. Department of Electrical, Electronic and Information Engineering, University of Bologna, Italy.
- Vroomen, J., & de Gelder, B. (2004). *Temporal Ventriloquism: Sound Modulates the Flash-Lag Effect*. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 513-518.
- Wallace, M. T., and Stein, B. E. (1997). *Development of multisensory neurons and multisensory integration in cat superior colliculus*. *Neuroscience*. 17, 2429–2444.
- Wallace, M. T., Perrault, T. J. Jr., Hairston, W. D., and Stein, B. E. (2004). *Visual experience is necessary for the development of multisensory integration*. *J. Neurosci.* 24, 9580–9584.
- Wallace, M., Wilkinson, L., & Stein, B. (1996). *Representation and integration of multiple sensory inputs in primate superior colliculus*. *Journal of Neurophysiology* (2), 1246-1266.
- Wallace, MT and Stein, BE (1994). *Cross-modal synthesis in the midbrain depends on input from cortex*, *Journal of Neurophysiology* 71: 429-432.
- Wei Ji Ma, Alexandre Pouget. (2008). *Linking neurons to behavior in multisensory perception: A computational review*. Department of Brain and Cognitive Sciences, University of Rochester, Rochester NY 14627, USA.
- Wilkinson, LK, Meredith, MA and Stein, BE (1996). *The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior*. *Experimental Brain Research* 112: 1-10.

Zwiers, M., Van Opstal, A., & Paige, G. (2003). *Plasticity in human sound localization induced by compressed spatial vision*. *Nature Neuroscience* (6), 175–181.



## **Ringraziamenti:**

**Il più grosso ringraziamento che possa fare va alla mia famiglia, ai miei genitori e a mio fratello che hanno sempre creduto in me fin dal primo momento in cui ho iniziato questo percorso e nonostante le difficoltà mi hanno sempre spronato a non mollare e a perseguire i miei obiettivi.**

**Un altro grosso ringraziamento va alla mia seconda famiglia, i miei amici, che hanno saputo sostenermi in ogni momento ricordandomi sempre che erano lì accanto a me con affetto e stima qualsiasi decisione prendessi.**

**Ai miei colleghi, compagni di università, a loro devo questi due anni meravigliosi che solo al ricordo mi si riempono gli occhi di lacrime, sono stati due anni intensi e bellissimi che ricorderò sempre con il sorriso.**

**A tutte le persone che hanno sempre creduto in me anche quando credevo di non farcela, a tutti quelli che quando abbassavo la testa sconsolata mi hanno fatto forza ricordandomi le mie capacità, a tutti coloro che mi vogliono bene veramente e a cui oggi spero di aver regalato almeno un po' di quell'orgoglio che loro mi fanno provare costantemente nell'averli accanto.**

**Grazie,**

**Luci.**