

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA  
CAMPUS DI CESENA  
SCUOLA DI INGEGNERIA E ARCHITETTURA

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA  
BIOMEDICA

**MODELLO MATEMATICO  
DELL'INTEGRAZIONE AUDIO-VISIVA E DEL  
SUO SVILUPPO IN AMBIENTE  
MULTISENSORIALE**

Tesi in  
**Sistemi Neurali LM**

**Relatore**  
Prof. Mauro Ursino

**Presentata da**  
Luca Baietta

Sessione Terza  
Anno Accademico 2012 – 2013



# Indice

<b>INTRODUZIONE.....</b>	<b>5</b>
<b>CAPITOLO 1</b>	
<b>INTEGRAZIONE AUDIO-VISIVA NELLA CORTECCIA PRIMARIA.....</b>	<b>9</b>
1.1 Multisensorialità: aspetti generali.....	9
1.2 Effetti dell'integrazione audio-visiva nella corteccia primaria.....	11
1.2.1 Il ventriloquismo.....	12
1.2.2 L'enhancement audio-visivo.....	14
1.3 Approcci neuro-computazionali.....	14
1.4 Modelli di riferimento.....	16
<b>CAPITOLO 2</b>	
<b>MODELLO MATEMATICO DELL'INTEGRAZIONE AUDIO-VISIVA.....</b>	<b>21</b>
2.1 Descrizione qualitativa del modello.....	21
2.2 Descrizione matematica del modello.....	23
<b>CAPITOLO 3</b>	
<b>ADDESTRAMENTO DELLA RETE.....</b>	<b>33</b>
3.1 Caso 1: sinapsi laterali uguali.....	34
3.1.1 Prima simulazione.....	34
3.1.2 Seconda simulazione.....	38
3.2 Caso 2: sinapsi laterali diverse.....	42
3.3 Caso 3: auto-addestramento delle sinapsi laterali.....	45
3.4 Studi sul ventriloquismo.....	50
3.5 Osservazioni.....	58

## **CAPITOLO 4**

<b>ESTENSIONE DEL MODELLO E COMPORTAMENTO DELLA RETE.....</b>	<b>59</b>
4.1 Riduzione progressiva delle costanti di apprendimento.....	59
4.2 Simulazioni con auto-addestramento delle sinapsi laterali.....	61
4.2.1 Prima simulazione.....	62
4.2.2 Seconda simulazione.....	67
4.3 Simulazione con sinapsi laterali diverse.....	71
4.4 Simulazione con sinapsi laterali uguali.....	75
4.5 Enhancement audio-visivo.....	79

## **CAPITOLO 5**

<b>CONCLUSIONI.....</b>	<b>83</b>
-------------------------	-----------

<b>APPENDICI.....</b>	<b>87</b>
-----------------------	-----------

Appendice A.....	87
------------------	----

Appendice B.....	92
------------------	----

Appendice C.....	97
------------------	----

Appendice D.....	101
------------------	-----

<b>BIBLIOGRAFIA.....</b>	<b>105</b>
--------------------------	------------

# Introduzione

La percezione del mondo esterno è il risultato di un processo di integrazione di informazioni provenienti da diverse modalità sensoriali. La capacità del sistema nervoso centrale di integrare queste informazioni provenienti da differenti sorgenti sensoriali è definita Integrazione Multisensoriale (Stein & Meredith, 1993). L'integrazione multisensoriale agisce a diversi livelli del sistema nervoso: nelle strutture subcorticali (es. il Collicolo Superiore), a elevati livelli delle corteccie associative (es. le regioni parietali posteriori) e anche nelle regioni corticali primarie (come la corteccia visiva primaria) in passato ritenute puramente unisensoriali.

Lo studio dell'integrazione multisensoriale si basa su differenti ma complementari approcci metodologici. La ricerca neurofisiologica sugli animali analizza le proprietà dei neuroni multimodali in specifiche aree corticali e subcorticali per far luce sui principi base che governano l'integrazione multisensoriale a livello del singolo neurone (Graziano et al., 1997; Kanduce et al., 1997; Perrault et al., 2005). La psicologia e la psicofisica sperimentale caratterizzano i processi multisensoriali a livello comportamentale, confrontando le risposte ottenute in presenza di stimoli unimodali e cross-modali (Driver & Spence, 1998; Farnè & Làdavas, 2002; Frassinetti et al., 2002; Haggard et al., 2007). Misure elettroencefalografiche, che registrano eventi legati a variazioni del potenziale d'azione, e tecniche di imaging (es. PET) permettono di ricostruire le strutture cerebrali e i meccanismi neurali coinvolti nei processi multisensoriali (Macaluso et al., 2000; Calvert, 2001; Eimer & Van Velzen, 2002). Tutte queste tecniche sono state applicate in soggetti neurologicamente sani e su pazienti con disturbi sensoriali o spaziali (Farnè & Làdavas, 2002; Frassinetti et al., 2005; Sarri et al., 2006), che potrebbero quindi incidere sulle abilità multisensoriali, per ottenere ulteriori dati sulle correlazioni neurali dell'integrazione multisensoriale.

Gli approcci visti in precedenza hanno contribuito a fornire un'enorme mole di dati sulla caratterizzazione delle proprietà dell'integrazione multisensoriale e a

identificare le aree cerebrali maggiormente coinvolte in questo fenomeno. Tuttavia la comprensione dei meccanismi neurali che coinvolge il sistema nervoso è ancora insufficiente. Questa limitazione potrebbe essere in parte imputata alla complessità dei meccanismi coinvolti; infatti l'integrazione multisensoriale è dovuta a popolazioni di neuroni interconnessi, nelle quali diversi fattori come le caratteristiche dei singoli neuroni, le disposizioni delle connessioni e la topologia della rete contribuiscono a determinare i fenomeni osservati. Chiariti questi aspetti, è piuttosto limitativo basarsi esclusivamente sui dati sperimentali. In più la mancanza di una conoscenza adeguata della topologia delle connessioni neurali sottolinea il fatto che la comprensione del meccanismo dell'addestramento neurale, attraverso il quale vengono acquisite capacità di integrazione multisensoriale, è molto limitata. Infatti diversi dati riportano che l'abilità di integrare informazioni sensoriale non è innata, ma si sviluppa gradualmente e si modifica plasticamente con l'esperienza; ovvero il contatto del soggetto col mondo esterno, ricco di stimoli cross-modali, modificherebbe la rete neurale in maniera funzionalmente rilevante. Le regole di apprendimento e le condizioni che portano allo sviluppo dell'integrazione multisensoriale nel sistema nervoso sono ancora poco conosciute.

Allo scopo di migliorare la comprensione dei principi computazionali e dei meccanismi neurali dell'integrazione multisensoriale (detta anche interazione cross-modale), negli ultimi anni sono state sviluppate delle tecniche di simulazione digitale su dei modelli neuro-computazionali. Un approccio molto utilizzato in questa analisi è quello di creare delle situazioni in cui l'informazione viene fornita da due differenti modalità sensoriali. Una tra le più interessanti e studiate forme di integrazione è quella di segnali spazialmente localizzati provenienti dal sistema visivo e da quello acustico. A tal proposito un fenomeno molto conosciuto che si registra quando a un osservatore viene presentato contemporaneamente un input visivo e uno acustico spazialmente separati è il ventriloquismo: esso si manifesta in una "cattura" dello stimolo acustico da parte dello stimolo visivo, ovvero lo stimolo acustico è percepito spostato verso la posizione di applicazione dello stimolo visivo (Bertelson & Radeau, 1981; Howard & Templeton, 1966).

I neuroni visivi e acustici di un neonato sono costituiti da campi recettori molto larghi e risposte unimodali e 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). Questo processo richiede diversi mesi di esperienze cross-modali prima di raggiungere la condizione di un adulto. La presente Tesi si pone come obiettivo quello di sviluppare un modello matematico di rete neurale che possa contribuire a chiarificare i meccanismi neurali alla base dell'integrazione audio-visiva nelle corteccie primarie a partire da condizioni base imposte (in termini fisiologici quelle di un neonato), e che possa descrivere come la rete riesca a svilupparsi in un ambiente multisensoriale. In particolare lo studio propone di analizzare l'addestramento dei campi recettori dei neuroni visivi e acustici, dell'attività cross-modale e, talvolta, delle sinapsi laterali in seguito a simulazioni di esperienze cross-modali (audio-visive) e unisensoriali (visive e acustiche), il tutto attraverso delle regole matematiche di apprendimento abbiano. Una volta addestrata la rete, il modello sarà successivamente testato per simulare gli effetti dell'integrazione audio-visiva nella corteccia primaria, tra i quali il già citato fenomeno del ventriloquismo e l'enhancement.

Il modello di rete neurale proposto in questo lavoro è costituito da due catene di neuroni unisensoriali, rispettivamente visivi e acustici. Ogni singolo neurone riceve degli input dall'esterno, comunica con i neuroni della stessa area sensoriale attraverso delle sinapsi laterali e con quelli dell'area sensoriale opposta mediante delle connessioni inter-catena.

Il Capitolo 1 affronta gli aspetti generali dell'integrazione multisensoriale, per poi concentrarsi particolarmente sull'integrazione audio-visiva che sussiste nella corteccie primarie. Quindi vengono esaminati gli effetti dell'integrazione audio-visiva nella corteccia primaria, quali il ventriloquismo e l'enhancement, e vengono introdotti dei modelli matematici utilizzati nello studio di questi fenomeni, dai quali il modello proposto in questa Tesi trae spunto.

Il Capitolo 2 fornisce una descrizione qualitativa e matematica del modello di rete neurale sviluppato nella presente Tesi.

## Introduzione

Il Capitolo 3 tratta il comportamento del modello in seguito all'addestramento in un ambiente multisensoriale in tre differenti condizioni fisiologiche. Vengono quindi mostrati i primi risultati relativi al ventriloquismo.

Nel Capitolo 4 il modello viene esteso con l'introduzione di alcune varianti. Di seguito viene ripetuto l'addestramento della rete con le nuove varianti e rappresentati i risultati sul ventriloquismo e sull'enhancement audio-visivo.

Nel Capitolo 5 vengono esposte le conclusioni che sintetizzano i risultati ottenuti durante le varie simulazioni, elencando i pregi e i difetti di ognuna. Al termine del Capitolo sarà mostrata la simulazione che ha soddisfatto maggiormente gli obiettivi posti nella Tesi e i possibili sviluppi futuri.

# Capitolo 1

## Integrazione audio-visiva nella corteccia primaria

### 1.1 Multisensorialità: aspetti generali

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 verrà 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, riducendone 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 della neuroscienza (Calvert, Spence, & Stein, 2004; B.E. Stein & Meredith, 1993).

Un tradizionale punto di vista, riportato dai più grandi testi di neuroscienza 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 sono ricombinate 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 questo 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 visiva e uditiva, oppure visiva e tattile). 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 considerate puramente unisensoriali.

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 considerazioni 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, i modelli di rete neurale, ispirati sulla 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.

In questo capitolo verranno enunciate conoscenze e teorie computazionali relative integrazione multisensoriale audio-visiva nella corteccia primaria. Inizialmente sarà presentato il tema dell'integrazione audio-visiva nelle aree corticali primarie, in passato ritenute puramente unisensoriali, superando la visione classica secondo la quale l'informazione proveniente dai vari sensi sia unificata solo successivamente in particolari aree associative, e verrà mostrato come la percezione di un oggetto (o un evento) in una modalità possa essere affetta da un'altra modalità (es. il ventriloquismo). Successivamente saranno introdotti degli approcci computazionali e modelli studiati recentemente in grado di simulare gli effetti dell'integrazione audio-visiva.

L'interazione audio-visiva non è un'abilità presente già alla nascita, ma verrà 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 corteccie 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 e unisensoriali (visive e acustiche), e di mostrare gli effetti successivi a tale addestramento (ventriloquismo, enhancement).

## **1.2 Effetti dell'integrazione audio-visiva nella corteccia primaria**

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 situate agli alti livelli della corteccia nel lobo frontale, parietale e temporale. Infatti la concezione tradizionale riteneva che gli stimoli unisensoriali erano processati separatamente nelle corteccie primarie, senza una significativa interazione cross-modale, e che i processi multisensoriali

avvenivano solo successivamente nelle corteccie associative più alte in cui convergevano segnali provenienti da differenti aree unimodali. Gli studi effettuati recentemente hanno permesso di superare questa visione, suggerendo che l'integrazione multisensoriale avviene anche a livello della corteccia primaria, e che le corteccie primarie della modalità visiva e uditiva hanno delle funzioni multisensoriali (Ghazanfar & Schroeder, 2006).

Molti studi sono stati effettuati sulle proprietà multisensoriali della corteccia uditiva, mostrando che gli stimoli visivi hanno una notevole influenza sui campi recettori dei neuroni uditivi. Questa influenza è dovuta a una convergenza anatomica di input visivi con i neuroni acustici ed è correlata alle dinamiche temporali delle risposte sensoriali.

Per quello che riguarda la corteccia visiva, le prime analisi di Morrell (Morrell, 1972) riportano che circa il 41% dei neuroni visivi nel gatto possono essere influenzati dagli stimoli acustici. Gli studi anatomici sulle scimmie hanno evidenziato la convergenza audio-visiva nella corteccia visiva primaria; al di sopra dei bassi livelli delle aree visive, le tecniche di mappatura su larga scala rivelarono che gli stimoli acustici attivavano parte della corteccia infero temporale (Ghazanfar & Schroeder, 2006; Poremba, et al., 2003).

## 1.2.1 Il ventriloquismo

Gli effetti che la modalità visiva produce su quella acustica (e viceversa) in condizioni di stimolazione cross-modale sono stati studiati mediante l'analisi delle illusioni percettive, che si registrano laddove il sistema visivo e quello acustico veicolano informazioni discordanti. In particolare:

- gli input visivi, a causa della loro elevata precisione spaziale, migliorano la localizzazione spaziale di quelli acustici (Alais & Burr, 2004; Alais et al., 2010);
- gli input acustici, a causa della loro elevata precisione temporale, migliorano la localizzazione temporale dei quelli visivi (Alais et al., 2010; Shams, Kamitani, & Shimojo, 2002).

L'abilità di localizzare uno stimolo acustico nello spazio è un processo computazionale altamente complesso, meno accurato ed affidabile della localizzazione visiva. Quando a un osservatore viene presentato un input visivo e uno acustico, temporaneamente coincidenti ma spazialmente disparati, la localizzazione dello stimolo acustico viene percepita shiftata in direzione della posizione di quello visivo (Bertelson & Radeau, 1981; Radeau & Bertelson, 1987): questo fenomeno è noto come "Effetto Ventriloquismo". Alcuni studi hanno dimostrato che la percezione shiftata dello stimolo acustico non è dovuta solo da fattori cognitivi, ma anche da un fenomeno di attrazione automatica del suono da parte di un simultaneo ma spazialmente disparato input visivo. L'esposizione a questo tipo di conflitto spaziale audio-visivo, oltre a produrre un immediato spostamento "online" dello stimolo acustico verso quello visivo, produce un effetto "offline" chiamato "After-Effect del Ventriloquismo" (Frissen, Vroomen, de Gelder & Bertelson, 2005; Lewald, 2002; G.H. Recanzone, 1998): in seguito a questa esposizione, gli input acustici unimodali vengono percepiti shiftati nella stessa direzione dello stimolo visivo precedentemente presentato, riportando così un adattamento plastico.

Nel dominio temporale, al contrario, lo stimolo acustico domina su quello visivo. Un fenomeno relativo ad esso è rappresentato dal "ventriloquismo temporale", un effetto cross-modale in cui un input acustico può alterare il tempo di percezione di uno stimolo visivo (Freeman & Driver, 2008; Getzmann, 2007; Parise & Spence, 2008; Vroomen & de Gelder, 2004). Un altro esempio in cui la modalità uditiva domina su quella visiva è un'illusione audio-visiva studiata da Shams e collaboratori (Shams, Kamitani, & Shimojo, 2000), chiamata "sound-induced flash illusion": quando un flash singolo si presenta assieme a due o più beep, gli osservatori riportano la visione di due o più flash ("fission"); quando invece un singolo beep si presenta assieme a due flash, allora avviene la fusione di questi ultimi due in un unico stimolo ("fusion") (Andersen, et al., 2004; Bolognini, Senna, Maravita, Pascual-Leone, & Merabet, 2010; Innes-Brown & Crewther, 2009; Shams & Kim, 2010).

## **1.2.2 L'Enhancement audio-visivo**

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. 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).

Inoltre alcuni autori hanno ipotizzato che uno stimolo di una modalità predisponga un'altra modalità, migliorandone così la risposta a un input successivo (es. la visione di un bicchiere che cade dal tavolo può predisporre la corteccia acustica a un susseguente suono di rottura). A tal proposito, un possibile meccanismo riguarda l'eccitazione sottosoglia di una modalità più rapida da parte di una modalità più lenta. Dal momento che la risposta visiva è più lenta di quella acustica, questa può eccitare la regione uditiva sottosoglia migliorando così la percezione dello stimolo acustico che si presenta in un'immediata finestra temporale.

## **1.3 Approcci neuro-computazionali**

La raccolta di un'elevata quantità di dati introdusse una serie di problemi sulla modellazione dell'integrazione multisensoriale nella corteccia primaria, in particolare sulla struttura delle connessioni cross-modali nelle aree sensoriali

primarie, che tuttora è ancora 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. La Figura 1.1 riporta tre possibili schemi strutturali in merito ai meccanismi di connessione sopracitati.

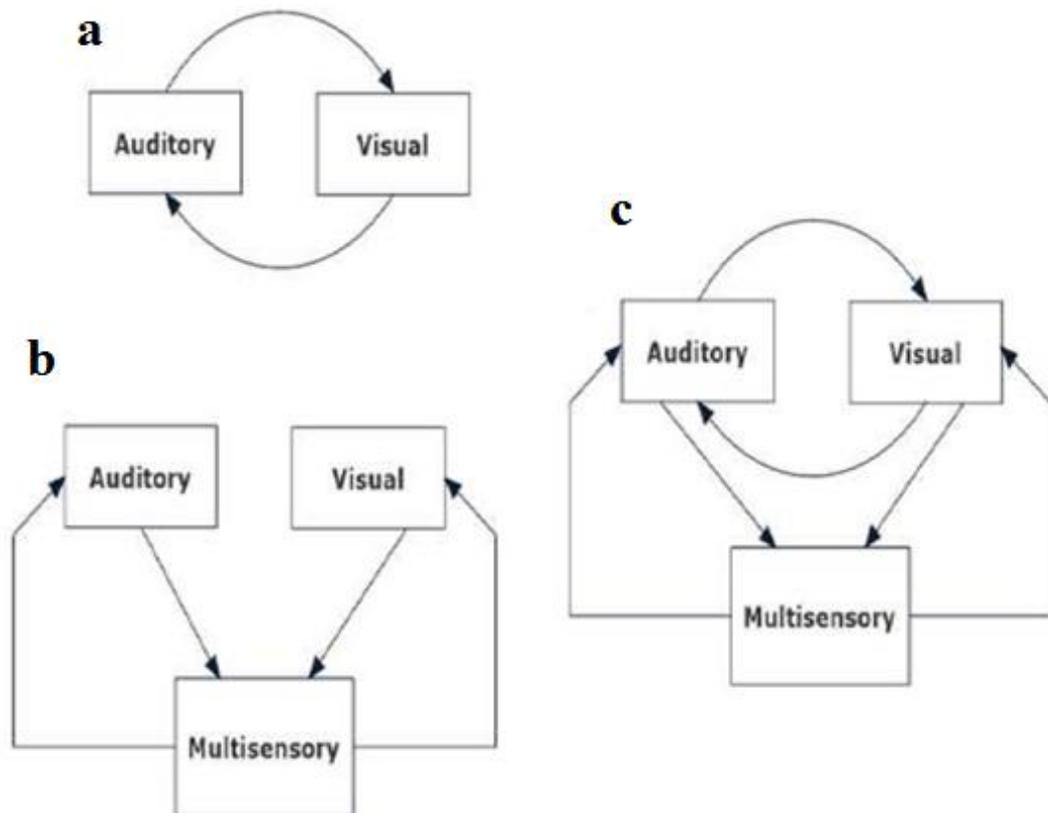


Figura 1.1: Tre schemi di strutture neurali della corteccia primaria

Il diagramma del pannello a assume connessioni laterali dirette tra le due aree unimodali; questo schema è stato utilizzato per simulare illusioni percettive audio-

visive tra le aree unisensoriali primarie (Magosso et al., 2012). Il modello studiato nella presente Tesi si basa proprio su questo tipo di schema, come si dirà in seguito.

Gli schemi rappresentati nei pannelli b e c, non utilizzati nello studio di questo modello di integrazione audio-visiva, includono sia connessioni feedforward dalle aree unisensoriali verso quelle multisensoriali (es. il collicolo superiore) e sia connessioni feedback dalle aree multisensoriali a quelle unisensoriali. In particolare lo schema b è stato utilizzato per simulare la rappresentazione multisensoriale dello spazio peripersonale, ovvero lo spazio immediatamente circostante al corpo (Magosso, Ursino, et al., 2010; Magosso, Zavaglia, et al., 2010). Il modello appena citato è formato da due sottoreti, ognuna delle quali rappresenta due aree unimodali di neuroni, rispettivamente visivi e tattili, sinapticamente connessi a un'area multisensoriale visuo-tattile.

Infine il pannello c rappresenta un diagramma simile a quello dello schema b, con in più l'introduzione di connessioni laterali dirette tra le due aree unimodali già viste nel pannello a. Questo è stato utilizzato in diversi modelli corticali, ad esempio nell'analisi dell'interazione visuo-tattile (Magosso, 2010) e dell'enhancement cross-modale sottosoglia (Hoshino, 2011).

Come si può ben notare in tutti e tre gli schemi riportati, le due aree corticali primarie (visive e acustiche) sono affette da altre modalità mediante connessioni dirette o feedback: ciò ne determina un comportamento multisensoriale.

## 1.4 Modelli di riferimento

Come già detto in precedenza, il modello studiato nella presente Tesi si basa sullo schema rappresentato nel pannello a della Figura 1.1. Diversi modelli precedenti ad esso hanno utilizzato un approccio simile, da cui se ne trarrà spunto.

Un modello molto semplice, particolarmente indirizzato nell'analisi dell'integrazione audio-visiva nelle aree unisensoriali, è stato proposto da Magosso, Cuppini e Ursino (Magosso, Cuppini, & Ursino, 2012). Tale modello consiste in due catene di  $N$  neuroni acustici e  $N$  neuroni visivi (con  $N=180$ ), rispettivamente (Figura 1.2) ed è stato realizzato con l'intento di spiegare i

meccanismi neurali alla base dell'effetto ventriloquismo (fenomeno online) e dell'after effect del ventriloquismo (fenomeno offline).

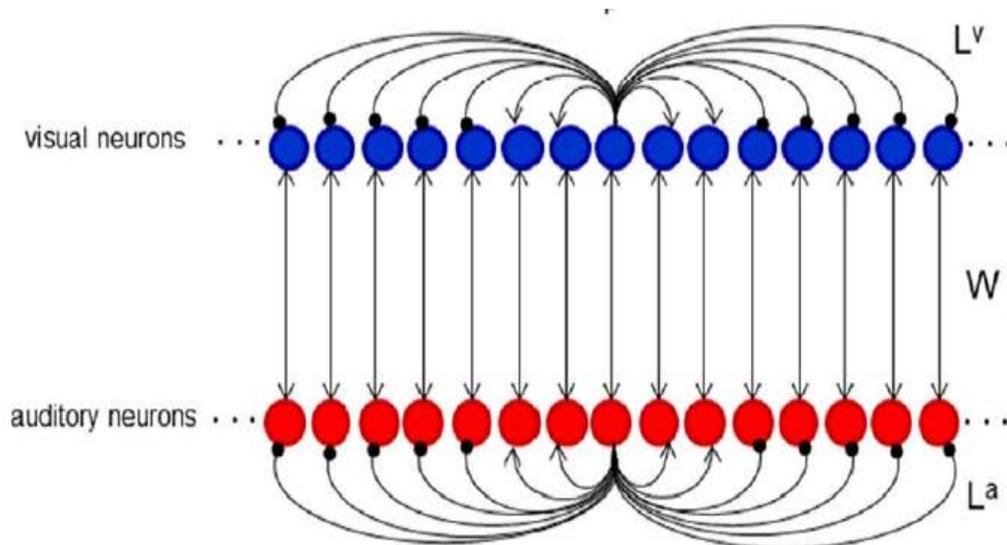


Figura 1.2: Schematizzazione del modello di Magosso, Cuppini, Ursino

Gli elementi neurali all'interno di ciascun strato sono connessi tra loro attraverso delle sinapsi laterali (sinapsi intra-area). Inoltre questi ricevono un input anche dai neuroni dell'area opposta tramite delle sinapsi inter-area con arrangiamento uno ad uno. In tal modo, l'ingresso ricevuto da un neurone risulta essere la somma di tre contributi: un input esterno, un input ricevuto dalle sinapsi laterali provenienti dai neuroni all'interno della stessa area unisensoriale ed un input cross-modale proveniente dai neuroni della modalità sensoriale opposta. Il comportamento del modello con i parametri in condizioni basali è stato testato per mostrare quali siano le condizioni che generano il ventriloquismo (Figure 1.3-1.4). Inoltre in questo modello, assumendo che le sinapsi laterali siano plastiche e che possano essere addestrate durante la fase di training mediante un meccanismo di apprendimento hebbiano, è stato analizzato anche il fenomeno dell'after effect del ventriloquismo (non trattato nel modello presentato in questa Tesi).

## Capitolo 1 – Integrazione audio-visiva nella corteccia primaria

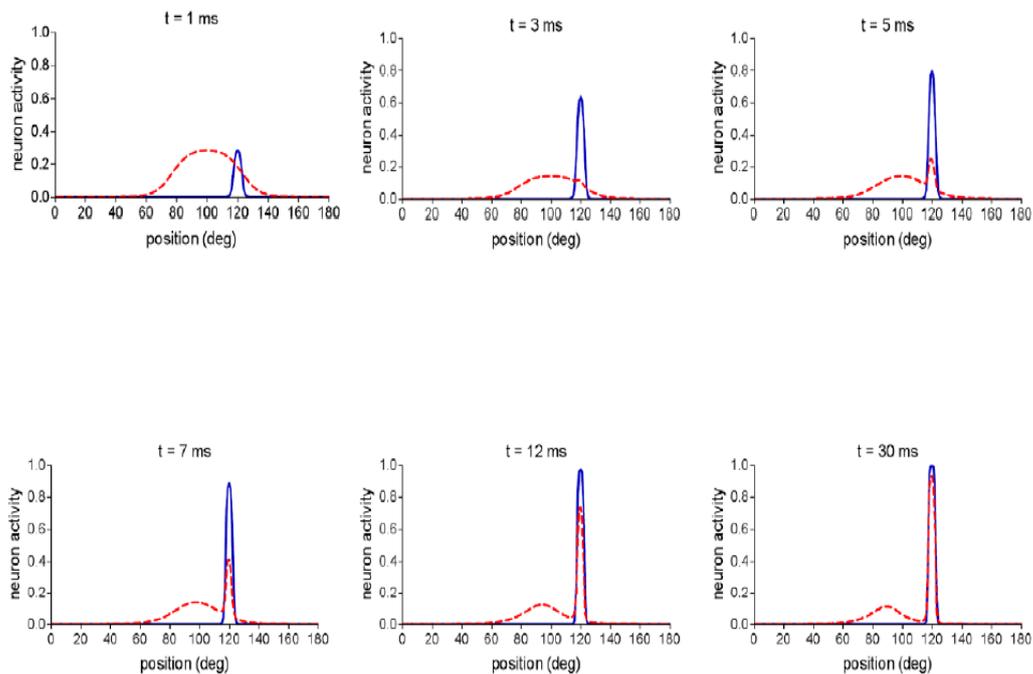


Figura 1.3: Fenomeno del ventriloquismo a diversi istanti di tempo (modello di Magosso, Cuppini Ursino), ottenuto applicando uno stimolo visivo in posizione 120 e uno stimolo acustico in posizione 100. La linea rossa tratteggiata indica l'attività neurale acustica, quella blu continua l'attività neurale visiva

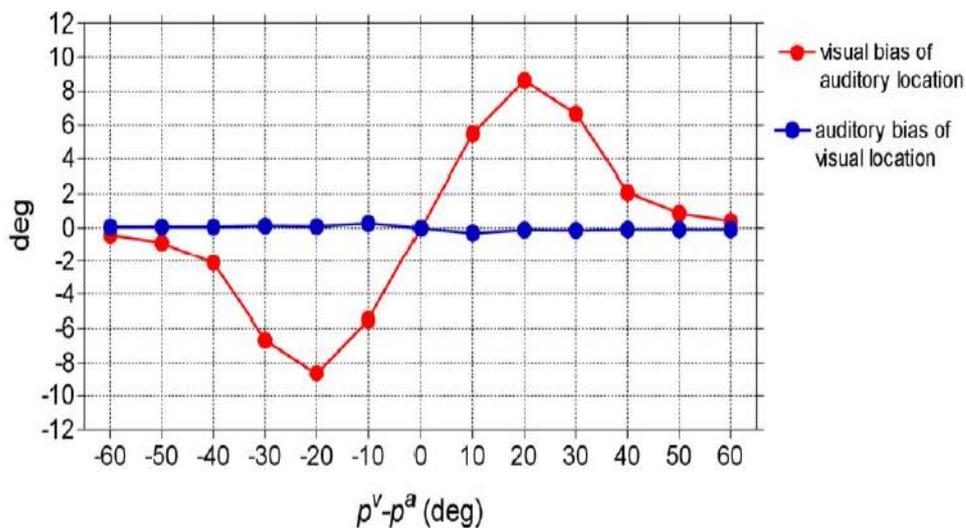


Figura 1.4: Shift di localizzazione dello stimolo acustico (rosso) e dello stimolo visivo (blu), in funzione della loro separazione angolare (modello di Magosso, Cuppini Ursino)

Una versione simile al modello visto in precedenza (Cuppini, Magosso, Bolognini, Vallar & Ursino, under review) introduce delle dinamiche temporali diverse per i neuroni visivi e acustici, così da simulare dei campi recettori spazio-temporali più attendibili per entrambe le modalità, in modo tale che la modalità visiva abbia una miglior risoluzione spaziale, mentre quella uditiva una miglior risoluzione temporale. Con questa nuova assunzione il modello può spiegare non solo l'effetto ventriloquismo, dove uno stimolo visivo domina su uno stimolo acustico spazialmente disparato (Figura 1.5), ma anche i fenomeni di fission e fusion (non trattati nella Tesi), dove gli stimoli acustici prevalgono su quelli visivi nel dominio temporale.

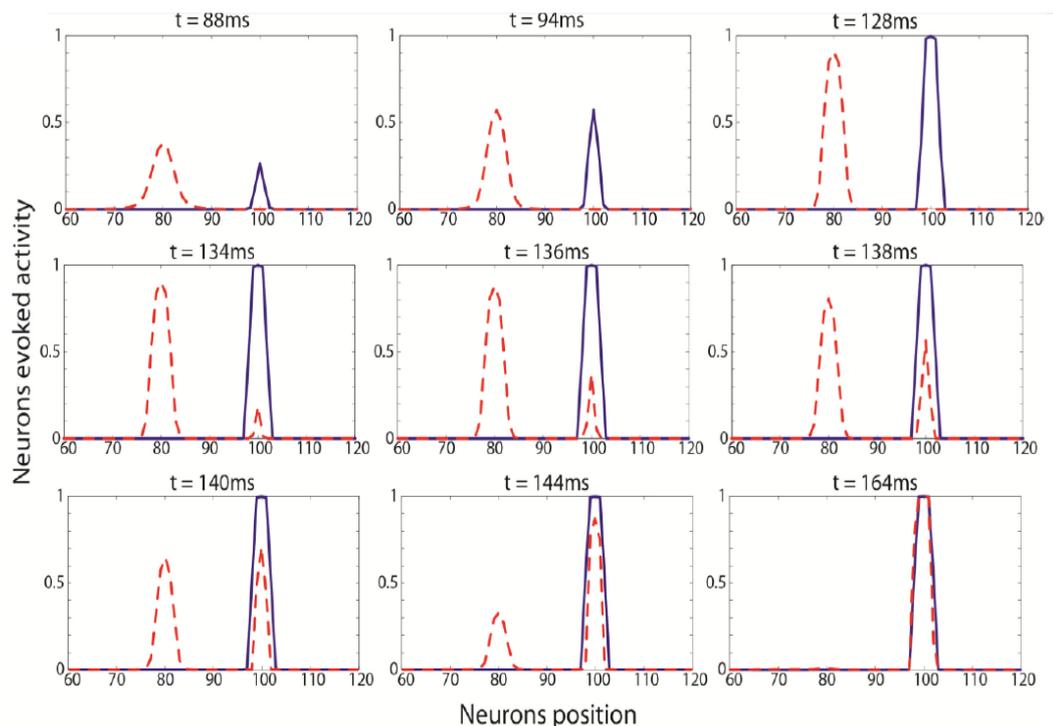


Figura 1.5: Fenomeno del ventriloquismo a diversi istanti di tempo (modello di Cuppini, Magosso, Bolognini, Vallar & Ursino), ottenuto applicando uno stimolo visivo in posizione 100 e uno stimolo acustico in posizione 80. La linea rossa tratteggiata indica l'attività neurale acustica, quella blu continua l'attività neurale visiva

Il modello in questo lavoro di Tesi è stato implementato prendendo spunto da questi due modelli di riferimento appena descritti, con lo scopo di:

- analizzare come la rete neurale si addestri (mediante un meccanismo di apprendimento hebbiano) in un ambiente multisensoriale mediante delle simulazioni di esperienze cross-modali e unisensoriali (visive o acustiche), mostrando in particolare lo sviluppo dei campi recettori, dell'attività cross-modale e, in alcune simulazioni, delle sinapsi laterali;
- mostrare gli effetti dell'integrazione audio-visiva (ventriloquismo e enhancement) in seguito all'addestramento della rete per un determinato numero di epoche;
- ricercare varianti matematiche e parametri appropriati in modo tale che la rete neurale raggiunga la stabilità ad un elevato numero di epoche.

## **Capitolo 2**

# **Modello matematico dell'integrazione audio-visiva**

### **2.1 Descrizione qualitativa del modello**

Il modello di rete neurale sviluppato nella presente Tesi consiste in due catene di  $N$  neuroni visivi e  $N$  neuroni acustici ( $N=180$ ), rispettivamente (Figura 2.1). Ogni neurone codifica le informazioni provenienti da una specifica posizione nello spazio. Tutti i neuroni sono topologicamente allineati, in modo tale che quelli che occupano posizioni vicine nell'array codificano posizioni prossimali nello spazio. Si è assunta una distanza di  $1^\circ$  tra neuroni adiacenti, in modo tale che ogni strato di neuroni copra un'area di  $180^\circ$  dello spazio.

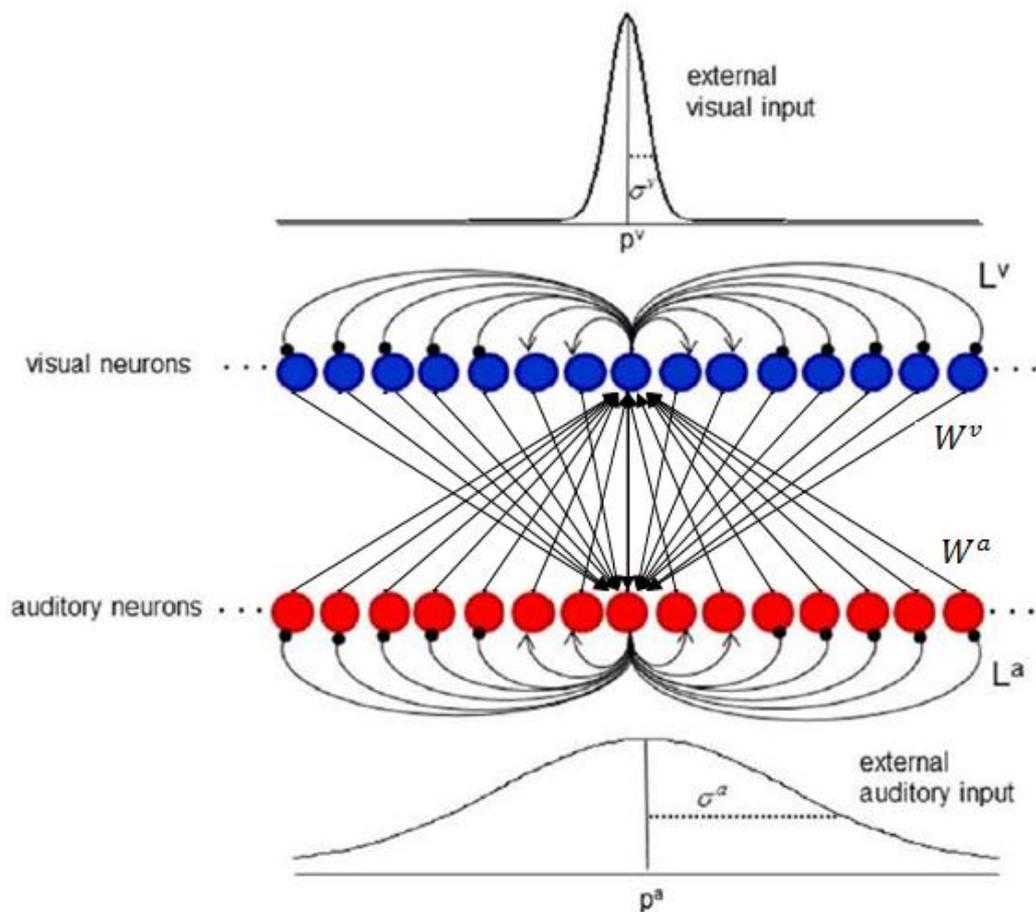


Figura 2.1: Schema rappresentativo del modello di rete neurale in esame

La risposta neurale è descritta mediante un'equazione differenziale con una dinamica del primo ordine, che simula le proprietà integrative della membrana cellulare, e da una caratteristica statica sigmoideale, che stabilisce la presenza di una soglia inferiore e di una saturazione superiore per l'attivazione neurale. Il valore di saturazione è stato imposto a 1, quindi tutte le uscite sono normalizzate al valore massimo.

Gli elementi neurali all'interno di ciascuno strato sono connessi tra loro attraverso sinapsi laterali (sinapsi intra-area). In più ogni neurone riceve degli input anche dai neuroni dell'area sensoriale opposta (sinapsi inter-area). Così l'ingresso totale ricevuto da ogni neurone risulta la somma di tre contributi: un input esterno, un input proveniente dai neuroni della stessa area unisensoriale attraverso le sinapsi

lateralmente, un input cross-modale proveniente dai neuroni dell'altra modalità sensoriale.

L'input esterno rappresenta uno stimolo visivo o acustico che originano da una posizione limitata dello spazio (ad esempio un "flash" visivo o un "beep" acustico); questo stimolo locale viene simulato attraverso una funzione Gaussiana. Il punto centrale di tale funzione ( $p^a$  e  $p^v$ , rispettivamente per lo stimolo acustico e lo stimolo visivo) corrisponde alla posizione di applicazione dello stimolo dal mondo esterno, mentre la deviazione standard ( $\sigma^a$  e  $\sigma^v$ ) è in relazione con la larghezza dello stimolo locale. Nel modello si considera come ipotesi fondamentale che lo stimolo visivo sia spazialmente più preciso rispetto a quello acustico, quindi  $\sigma^v < \sigma^a$ . Una volta in ingresso alla rete, gli input esterni vengono filtrati spazialmente dai campi recettori dei neuroni.

Gli input laterali originano dalle connessioni intra-area della stessa area unimodale. Queste sinapsi laterali includono un contributo eccitatorio e uno inibitorio, e hanno una disposizione classica a "cappello messicano", realizzato da una differenza di due funzioni gaussiane. Così ogni neurone eccita (ed è eccitato da) quelli vicini, e inibisce (ed è inibito da) quelli più distanti. In questo modo stimoli distali della stessa modalità tendono a reprimersi reciprocamente mediante un meccanismo competitivo.

L'input cross-modale è unicamente eccitatorio: ogni neurone riceve uno stimolo proveniente dai neuroni dell'area sensoriale opposta. L'eccitazione è massima in corrispondenza di quei neuroni della modalità opposta situati in prossimità spaziale, mentre per quelli più distali si ha un contributo inferiore.

## 2.2 Descrizione matematica del modello

In seguito, ciascun neurone sarà descritto attraverso un apice,  $m$ , che ne indica la specifica area corticale ( $m = a$  o  $v$ , dove  $a$  è riferita all'area auditiva e  $v$  a quella visiva), e da un pedice,  $j$ , che si riferisce alla posizione spaziale relativa a quell'area ( $j = 1, 2, \dots, 180$ ).  $u(t)$  e  $y(t)$  sono utilizzati per rappresentare, rispettivamente, l'input e l'output totale di un particolare neurone all'istante di

tempo  $t$ .  $y_j^m$ , che rappresenta quindi la risposta del neurone  $j$ -esimo appartenente alla modalità  $m$ , può essere descritta dalla seguente equazione differenziale del primo ordine:

$$\tau \frac{dy_j^m(t)}{dt} = -y_j^m(t) + F(u_j^m(t))$$

dove  $\tau$  è la costante di tempo del neurone e  $F(u_j^m(t))$  rappresenta una relazione sigmoideale:

$$F(u_j^m) = \frac{1}{1 + e^{-s(u_j^m - \theta)}}$$

dove  $s$  e  $\theta$  sono parametri che stabiliscono, rispettivamente, la pendenza e la posizione centrale della funzione sigmoideale. La caratteristica sigmoideale è visibile in Figura 2.2. Il valore di saturazione è stato imposto a 1, così le attività neurali sono normalizzate al valore massimo.

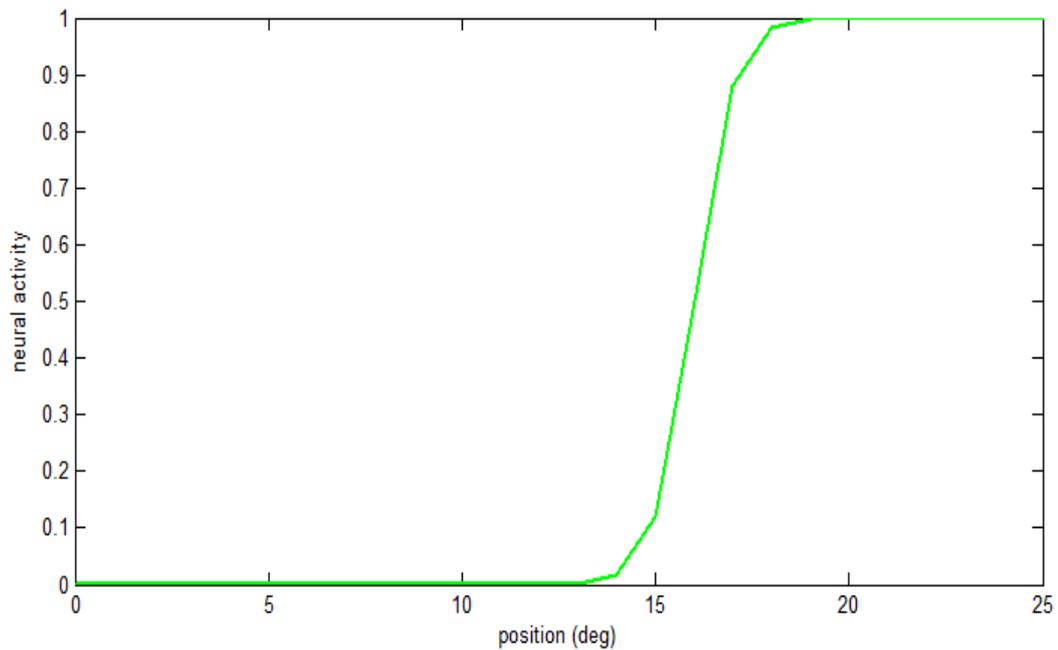


Figura 2.2: Caratteristica sigmoideale dei neuroni visivi e acustici

In questo lavoro tutti i neuroni appartenenti a entrambe le aree sono stati definiti utilizzando gli stessi parametri e la stessa costante di tempo.

L'input totale,  $u_j^m(t)$ , che ogni neurone riceve in ingresso è dato dalla somma di tre componenti: un contributo esterno,  $e_j^m(t)$ , un input proveniente dalle sinapsi laterali della stessa area unisensoriale,  $l_j^m$ , e un input cross-modale proveniente dai neuroni dell'altra modalità,  $c_j^m$ :

$$u_j^m(t) = e_j^m(t) + l_j^m(t) + c_j^m$$

La componente esterna è stata simulata utilizzando una funzione gaussiana, che definisce uno stimolo esterno locale visivo (Figura 2.3) o acustico (Figura 2.4) che origina da una posizione limitata dello spazio, filtrata spazialmente dal campo recettore del neurone:

$$e_j^m(t) = \sum_{k=1}^N RF_{jk}^m E_k^m$$

dove  $E_k^m$  rappresenta l'intensità dello stimolo esterno visivo ( $m = v$ ) o acustico ( $m = a$ ) centrato in posizione  $j$  percepita dal neurone  $k$ -esimo:

$$E_k^m = E_0^m e^{-\frac{(d_{jk}^m)^2}{2(\sigma^m)^2}}$$

Dove  $E_0^m$  indica l'intensità dello stimolo esterno in posizione centrale  $j$ ,  $d_{jk}^m$  è la distanza tra il neurone  $j$ -esimo e  $k$ -esimo e  $\sigma^m$  definisce la larghezza dello stimolo esterno. Come già detto in precedenza, dal momento che il sistema visivo ha una risoluzione spaziale migliore rispetto a quello acustico, risulta  $\sigma^v < \sigma^a$ . Con questo accorgimento lo stimolo visivo esterno risulta più localizzato e, al contrario, quello acustico risulta più diffuso.

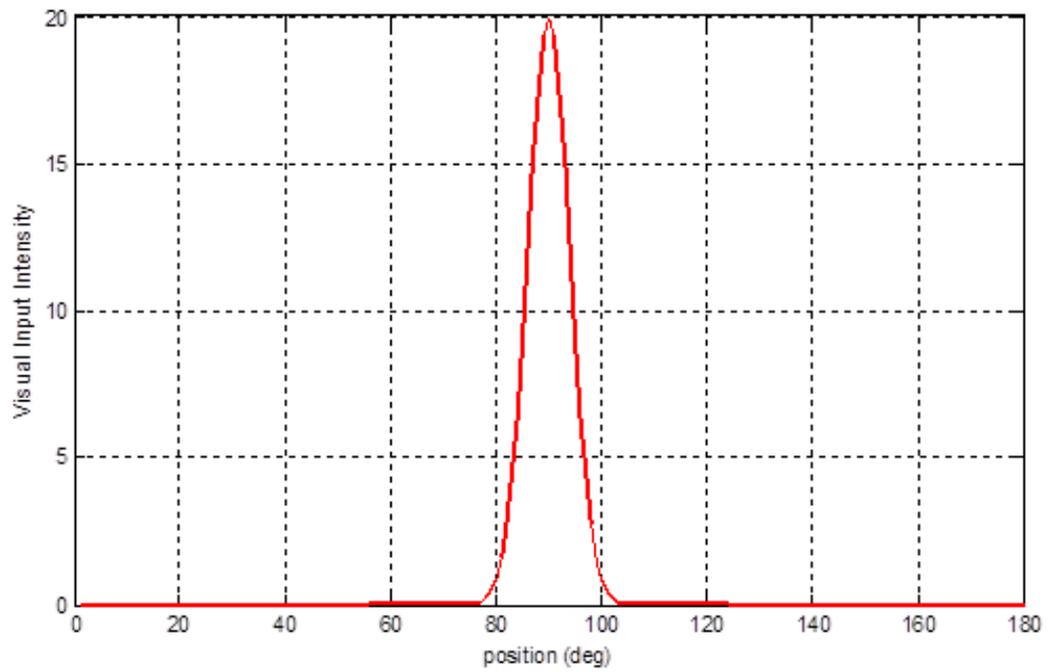


Figura 2.3: Stimolo esterno visivo centrato in posizione 90

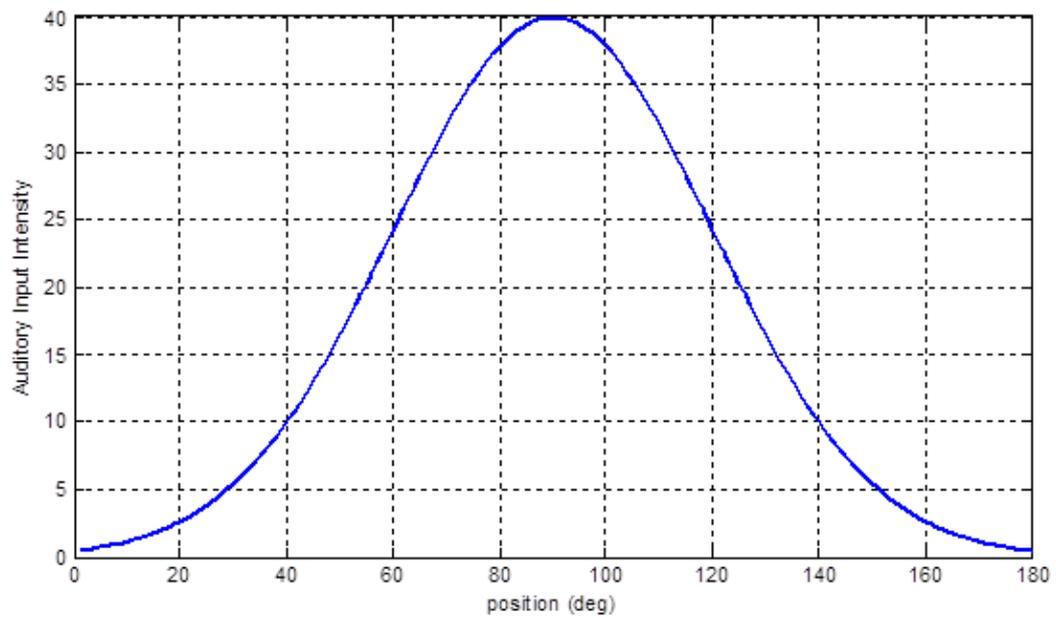


Figura 2.4: Stimolo esterno acustico centrato in posizione 90

Gli elementi situati all'estremità della catena potrebbero non ricevere gli stessi input dagli altri, producendo così degli effetti di bordo indesiderati. Per evitare questa complicazione si assume che l'array di neuroni abbia una struttura circolare. Di conseguenza, per la distanza  $d_{jk}^m$  si assume la seguente definizione:

$$d_{jk}^m = \begin{cases} |j - k| & \text{se } |j - k| \leq N/2 \\ N - |j - k| & \text{se } |j - k| > N/2 \end{cases}$$

$RF_{jk}^m$  indica l'intensità del campo recettore relativo al neurone  $j$ -esimo percepita dal neurone in posizione  $k$ . Tale campo tende a modificarsi progressivamente durante la simulazione attraverso una particolare regola di addestramento di tipo hebbiano, che sarà illustrata in seguito. Il campo recettore viene definito a livello basale (ovvero prima dell'addestramento) da una funzione Gaussiana (Figura 2.5), allo stesso modo per tutti i neuroni di entrambe le modalità:

$$RF_{jk}^m = RF_0 e^{-\frac{(d_{jk})^2}{2(\sigma_v)^2}}$$

Dove  $RF_0$  rappresenta l'ampiezza del campo recettore in posizione centrale  $j$ ,  $\sigma_v$  è la deviazione standard correlata con la larghezza del campo recettore basale, mentre  $d_{jk}$  è la distanza del neurone  $j$ -esimo dal neurone in posizione  $k$ , definita utilizzando la forma circolare già vista in precedenza, ovvero:

$$d_{jk}^m = \begin{cases} |j - k| & \text{se } |j - k| \leq N/2 \\ N - |j - k| & \text{se } |j - k| > N/2 \end{cases}$$

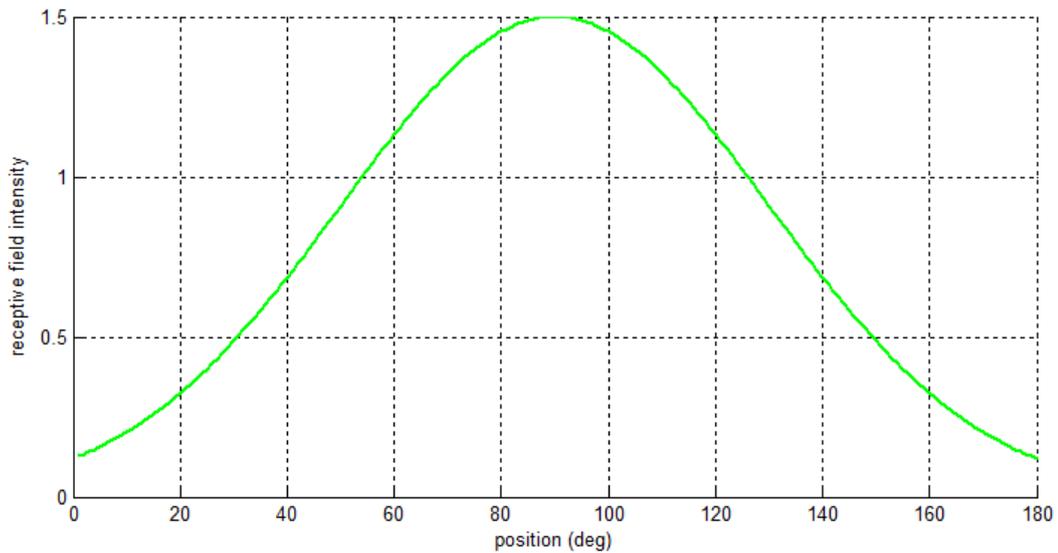


Figura 2.5: Campo recettore basale del neurone (visivo e acustico) in posizione 90

L'input che ogni neurone unisensoriale riceve dagli altri elementi della stessa area origina dalle connessioni laterali presenti all'interno della stessa area unimodale. Per implementare questo meccanismo è stata utilizzata la seguente espressione:

$$l_j^m(t) = \sum_{k=1}^N L_{jk}^m y_k^m(t)$$

dove  $y_k^m$  è l'attività del neurone pre-sinaptico di modalità  $m$  in posizione  $k$  e  $L_{jk}^m$  indica la forza delle sinapsi laterali che vanno dal neurone pre-sinaptico, posizionato in  $k$ , al neurone post-sinaptico in posizione  $j$ . Le sinapsi laterali sono date dalla differenza tra un contributo eccitatorio ed uno inibitorio, come rappresentato in Figura 2.6 e descritto dalla seguente equazione:

$$L_{jk}^m = L_{ex,jk}^m - L_{in,jk}^m$$

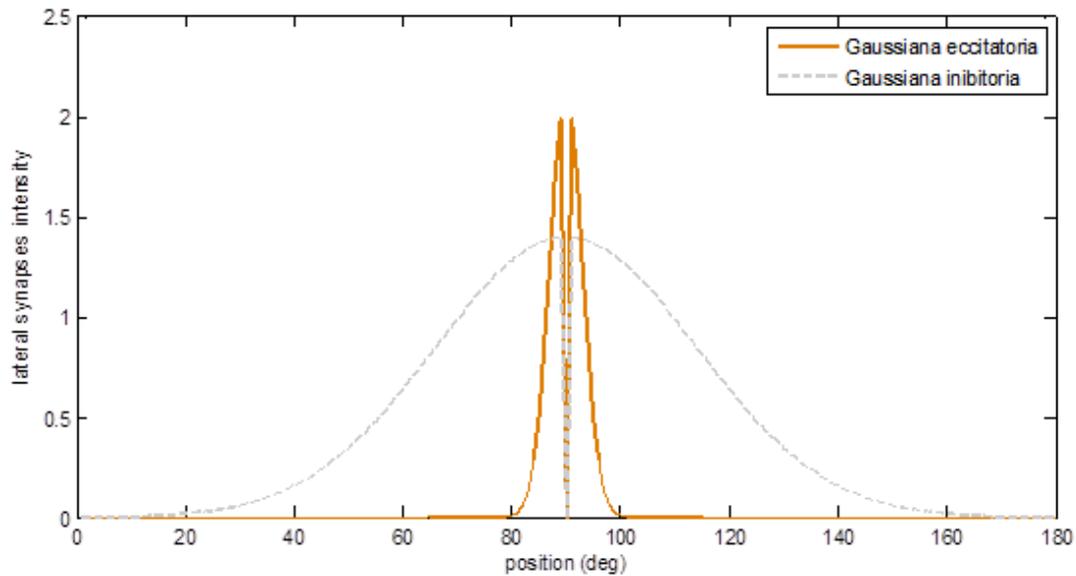


Figura 2.6: Esempio di Gaussiana eccitatoria e inibitoria relative neurone in posizione 90 utilizzate per la costruzione delle sinapsi laterali

In condizioni basali queste sinapsi laterali hanno una disposizione a “cappello messicano” (Figura 2.7), realizzato come differenza di due gaussiane:

$$L_{ex,jk}^m = \begin{cases} L_{ex0} e^{-\frac{(d_{jk})^2}{2\sigma_{ex}^2}}, & \text{se } d_{jk} \neq 0 \\ 0, & \text{se } d_{jk} = 0 \end{cases}$$

$$L_{in,jk}^m = \begin{cases} L_{in0} e^{-\frac{(d_{jk})^2}{2\sigma_{in}^2}}, & \text{se } d_{jk} \neq 0 \\ 0, & \text{se } d_{jk} = 0 \end{cases}$$

dove  $L_{ex0}$  e  $L_{in0}$  sono parametri che stabiliscono la forza della sinapsi eccitatoria ed inibitoria,  $\sigma_{ex}$  e  $\sigma_{in}$  sono le deviazioni standard che indicano la velocità con cui la sinapsi decresce con la distanza e  $d_{jk}$  è la distanza tra il neurone in posizione  $j$  e  $k$ , anch'essa definita attraverso la forma circolare. In queste equazioni è stata esclusa la presenza di autoeccitazione e autoinibizione.

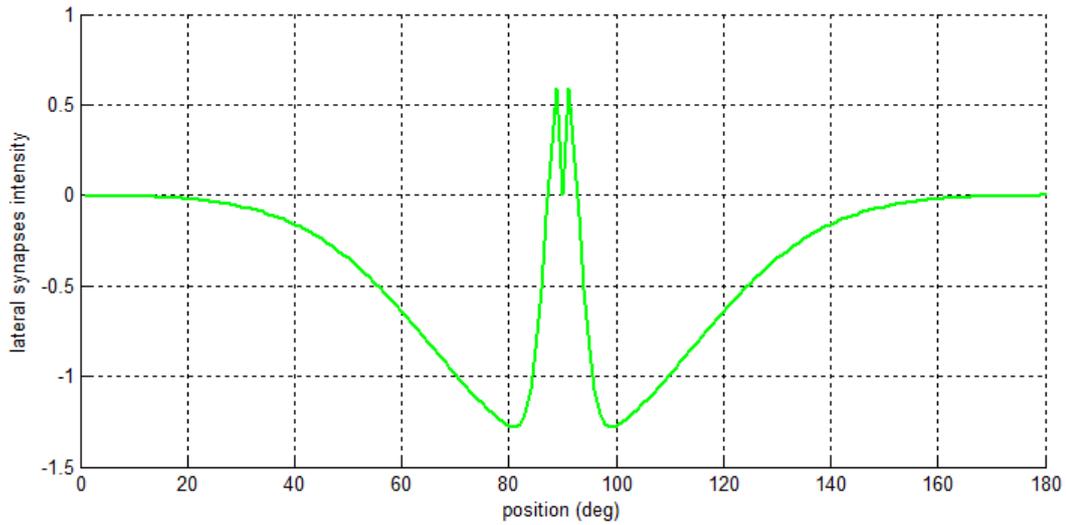


Figura 2.7: esempio di sinapsi laterali entranti nel neurone in posizione 90

Per ottenere una funzione a “cappello messicano” si è reso necessario imporre:  $L_{ex0} > L_{in0}$  e  $\sigma_{ex} < \sigma_{in}$ . Ogni neurone eccita quelli vicini e inibisce, ed è inibito, da quelli più distanti. In questo modo stimoli della stessa modalità che sono distanti tendono a reprimersi reciprocamente mediante un meccanismo competitivo.

Per ciò che riguarda l'input cross-modale  $c_j^m$ , ogni neurone riceve stimoli provenienti dai neuroni dell'altra modalità. Esso è unicamente eccitatorio, in cui si avrà un contributo massimo da parte degli elementi prossimali alla posizione spaziale del neurone *j-esimo*. In base a questo si ha:

$$c_j^a(t) = \sum_{k=1}^N W_{jk}^v y_k^v(t)$$

$$c_j^v(t) = \sum_{k=1}^N W_{jk}^a y_k^a(t)$$

dove  $W_{jk}^v$  e  $W_{jk}^a$  indicano, rispettivamente, le sinapsi cross-modali visive e acustiche relative al neurone *j-esimo* percepite dal neurone *k-esimo* (in condizioni basali tale valore sarà posto a 0), mentre  $y_k^v$  e  $y_k^a$  rappresentano l'attività pre-

sinaptica del neurone *k-esimo* di entrambe le modalità. Come per i campi recettori, anche le sinapsi cross-modali si sviluppano durante la simulazione attraverso un tipo di addestramento hebbiano, come sarà illustrato in seguito.

In Tabella 2.1 sono riportati i parametri del modello utilizzati in condizioni basali.

<b>Neurons</b>			
<b>N = 180</b>	<b><math>\tau = 5</math></b>	<b>s = 2</b>	<b><math>\theta = 16</math></b>
<b>Basal Receptive field (RF)</b>			
<b><math>RF_0 = 1.5</math></b>		<b><math>\sigma_v = 40</math></b>	
<b>Basal cross-modal synapses</b>			
<b><math>W_{jk}^v = 0</math></b>		<b><math>W_{jk}^a = 0</math></b>	

Tabella 2.1: Parametri del modello in condizioni basali

# Capitolo 3

## Addestramento della rete

Le equazioni differenziali sono state implementate e risolte numericamente attraverso il software Matlab, imponendo condizioni iniziali nulle e utilizzando il metodo di integrazione di Eulero. E' stato utilizzato un passo di integrazione di 0.2 millisecondi, sufficientemente piccolo da garantire una buona accuratezza.

Per testare il comportamento del modello durante l'addestramento è stato fornito in ingresso alla rete, in sequenza, uno stimolo visivo, uno stimolo acustico (non necessariamente provenienti dalla stessa posizione spaziale) e uno stimolo cross-modale di tipo audio-visivo (AV), ovvero uno stimolo acustico e uno visivo applicati simultaneamente nella stessa posizione spaziale. Questo è possibile attraverso l'introduzione nel modello di tre diverse permutazioni, ognuna formata da 180 elementi, che generano rispettivamente stimoli acustici, visivi e AV centrati in posizioni spaziali differenti, tutti con la stessa durata temporale di 17 millisecondi.

Una volta definiti gli ingressi si è proceduto all'addestramento della rete per diverse epoche, ognuna delle quali composta da 540 input, di cui 180 visivi 180 uditivi e 180 cross-modali. In particolare si è assunto che i campi recettori di ogni neurone visivo e acustico e le loro sinapsi cross-modali potessero essere addestrate, sfruttando la correlazione tra l'attività neurale pre-sinaptica e post-sinaptica.

Più precisamente è stato adottato un addestramento che si basa sulla regola di Oja, ovvero un addestramento di tipo hebbiano a cui viene sottratto un fattore di dimenticanza. A questo proposito si ottiene:

$$\begin{aligned}\partial RF_{jk}^m &= \gamma_{RF} y_j^m (E_k^m - RF_{jk}^m) \\ \partial W_{jk}^m &= \gamma_{W0} y_j^m (y_k^n - W_{jk}^m), \quad \text{con } m \neq n\end{aligned}$$

Dove  $\gamma_{RV}$  e  $\gamma_{W0}$  sono le costanti di apprendimento,  $y_j^m$  è l'attività neurale post-sinaptica che viene moltiplicata per l'attività neurale pre-sinaptica (regola di Hebb) e per un altro termine che contribuisce a formare il fattore di dimenticanza. Con questa procedura di training sono state effettuate delle simulazioni in tre casi differenti, che in seguito verranno esposti in maniera dettagliata.

## 3.1 Caso 1: sinapsi laterali uguali

In questo caso la rete neurale è stata addestrata con sinapsi laterali visive e acustiche identiche tra loro, mantenute costanti durante la simulazione. Quindi si avrà una distribuzione a “cappello messicano” uguale in entrambe le modalità.

A questo proposito sono state effettuate due diverse simulazioni: nella prima sono state utilizzate sinapsi laterali molto strette, con un'elevata eccitazione dei neuroni prossimali e una scarsa inibizione di quelli distali (Figura 3.1); nella seconda, al contrario, sono state introdotte sinapsi laterali più larghe, molto inibitorie e con una scarsa eccitazione dei neuroni prossimali (Figura 3.6).

### 3.1.1 Prima simulazione

Come già detto in precedenza, in questa prima simulazione sono state introdotte sinapsi laterali visive e uditive costanti e identiche tra loro, in cui la componente eccitatoria è molto forte e localizzata rispetto a quella inibitoria (Figura 3.1).

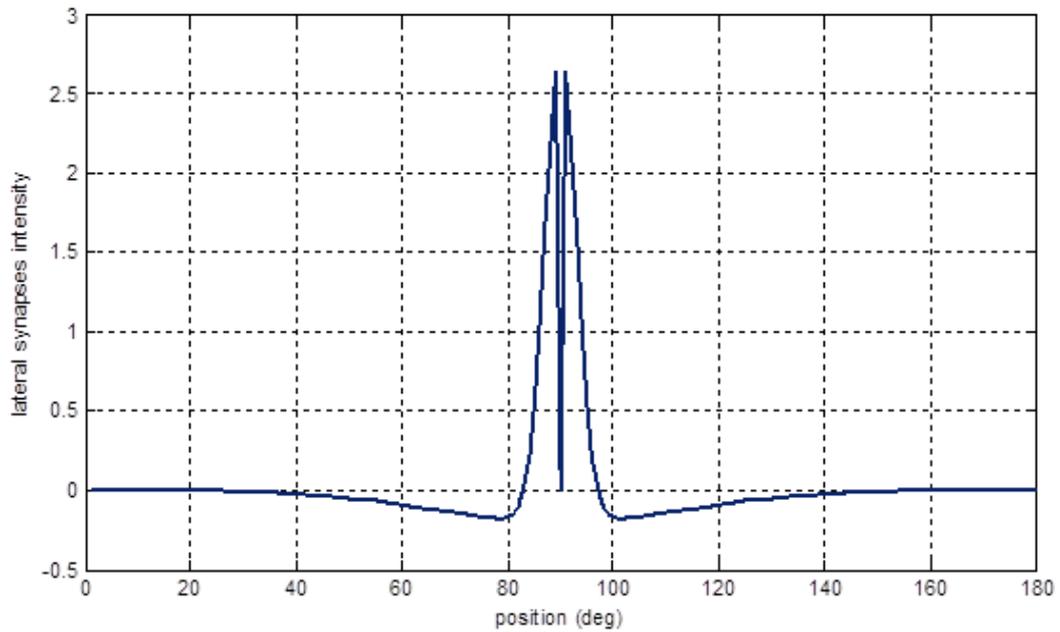


Figura 3.1: Sinapsi laterali visive e uditive coincidenti utilizzate nella prima simulazione del caso 1, relative al neurone in posizione 90

Una volta assegnati i parametri (Tabella 3.1), la rete è stata addestrata per un numero complessivo di 50 epoche, in cui è stato analizzato lo sviluppo dei campi recettori visivi e acustici (Figure 3.2-3.3) e quello delle sinapsi cross-modalità di entrambe le modalità (Figura 3.4-3.5).

External input			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 40$	$\sigma^a = 30$
Synaptic connections			
$L_{ex,0} = 3$	$\sigma_{ex} = 3 \text{ deg}$	$L_{in,0} = 0.2$	$\sigma_{in} = 24 \text{ deg}$
Hebbian rules			
$\gamma_{RV} = 0.002$		$\gamma_{W0} = 0.0002$	

Tabella 3.1: Parametri relativi alla prima simulazione del caso 1

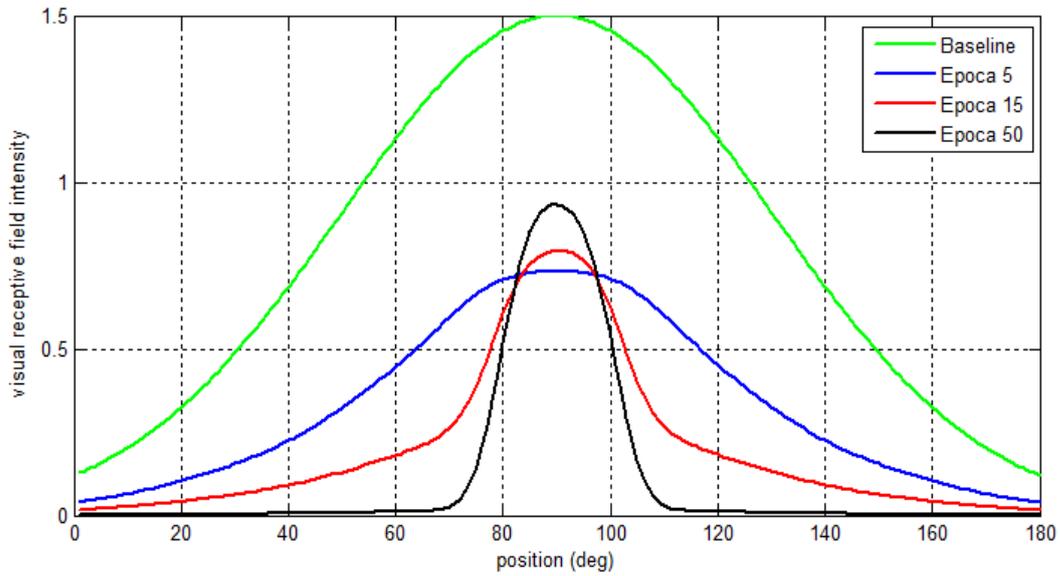


Figura 3.2: Addestramento del campo recettore visivo relativo al neurone in posizione 90

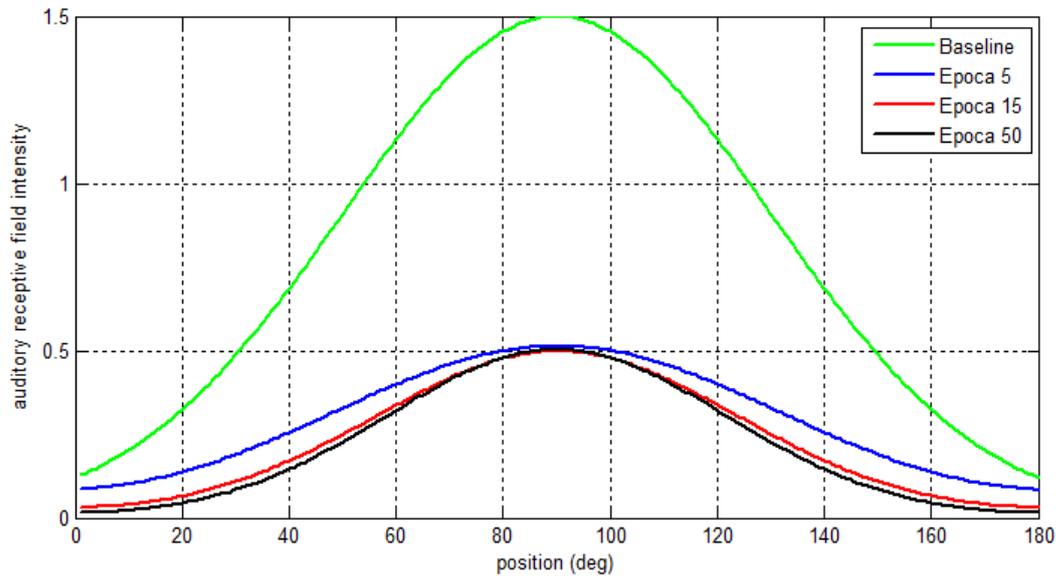


Figura 3.3: Addestramento del campo recettore acustico relativo al neurone in posizione 90

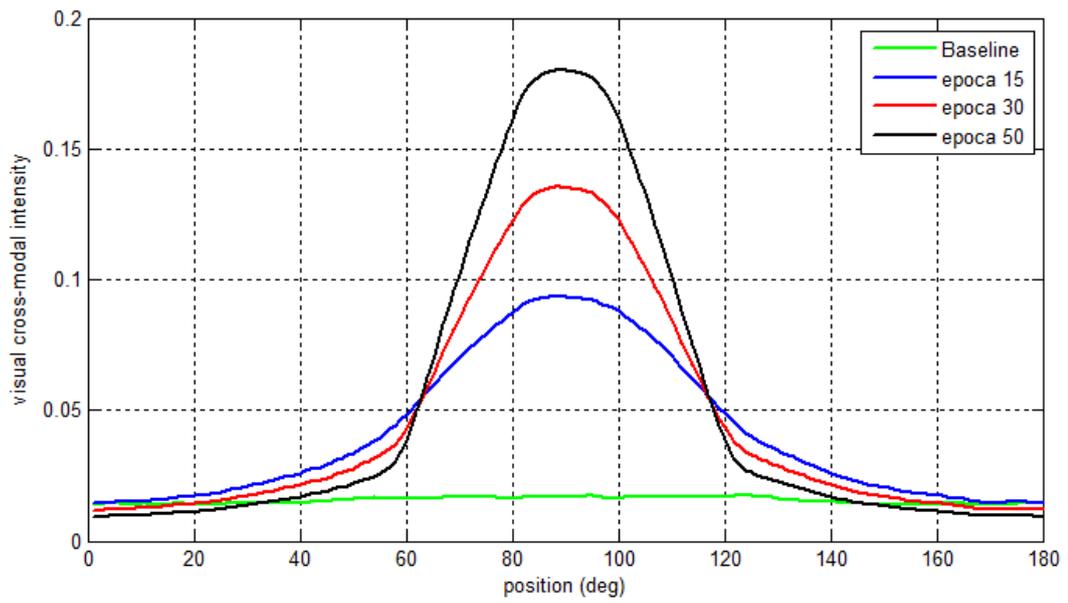


Figura 3.4: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

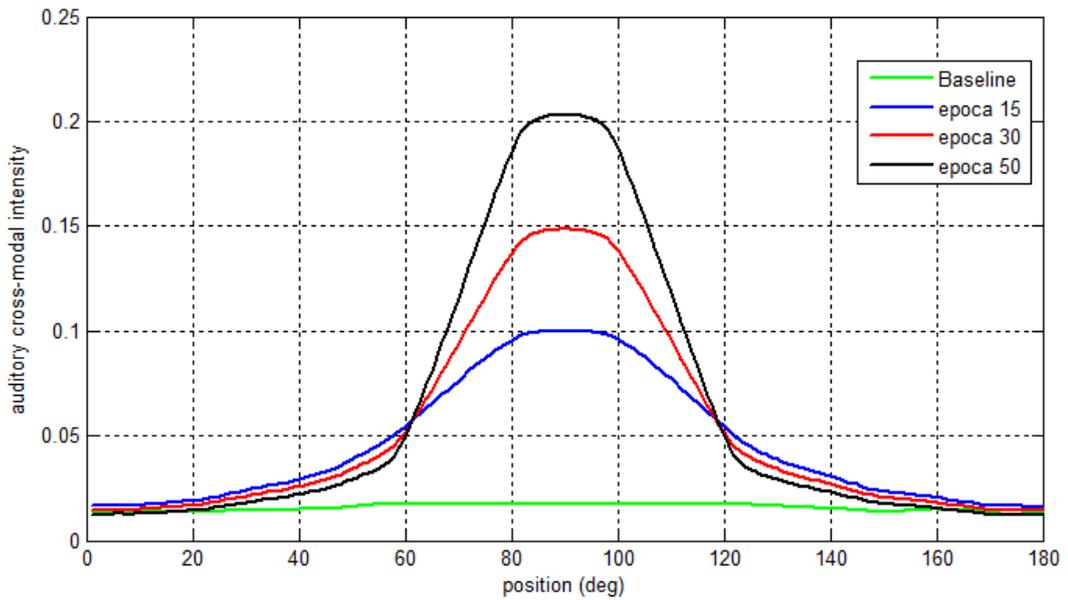


Figura 3.5: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

### 3.1.2 Seconda simulazione

Questo addestramento è stato simulato sulla falsa riga di quello precedente, per un numero complessivo di 100 epoche, con la variazione di alcuni parametri (Tabella 3.2).

External input			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 60$	$\sigma^a = 30$
Synaptic connections			
$L_{ex,0} = 2$	$\sigma_{ex} = 20 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 50 \text{ deg}$
Hebbian rules			
$\gamma_{RV} = 0.002$		$\gamma_{W0} = 0.0002$	

Tabella 3.2: Parametri relativi alla seconda simulazione del caso 1

A differenza della simulazione precedente, in questo caso specifico sono state utilizzate sinapsi laterali più larghe, con un'eccitazione molto debole da parte dei neuroni prossimali e un'importante inibizione di quelli distali (Figura 3.6)

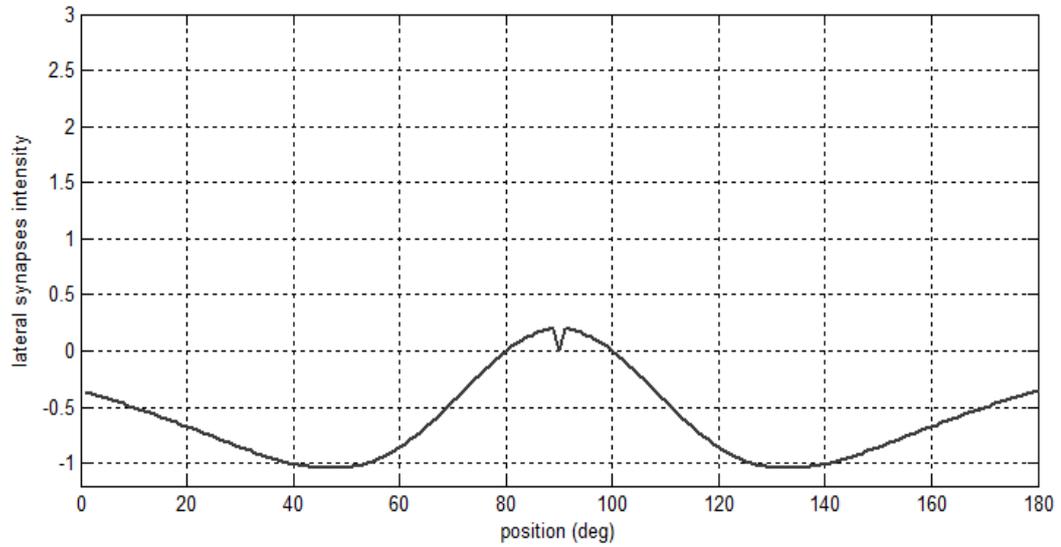


Figura 3.6: Sinapsi laterali visive e uditive coincidenti utilizzate nella seconda simulazione del caso 1, relative al neurone in posizione 90

Anche in questa simulazione è stato analizzato l'addestramento dei campi recettori visivi e uditivi (Figure 3.7-3.8) e delle sinapsi cross-modali delle due modalità (Figure 3.9-3.10).

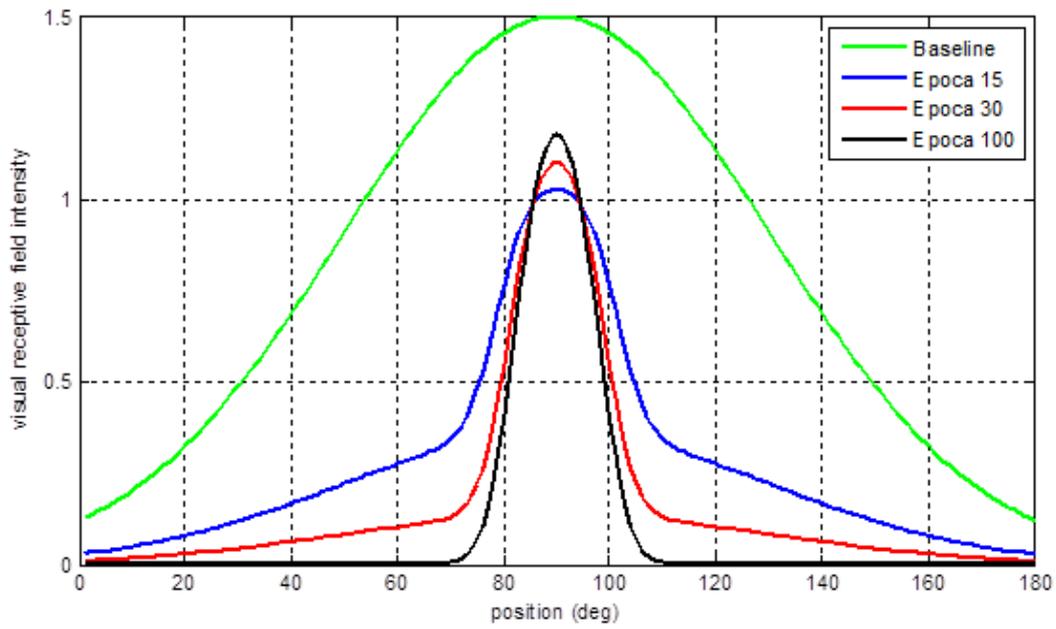


Figura 3.7 Addestramento del campo recettore visivo relativo al neurone in posizione 90

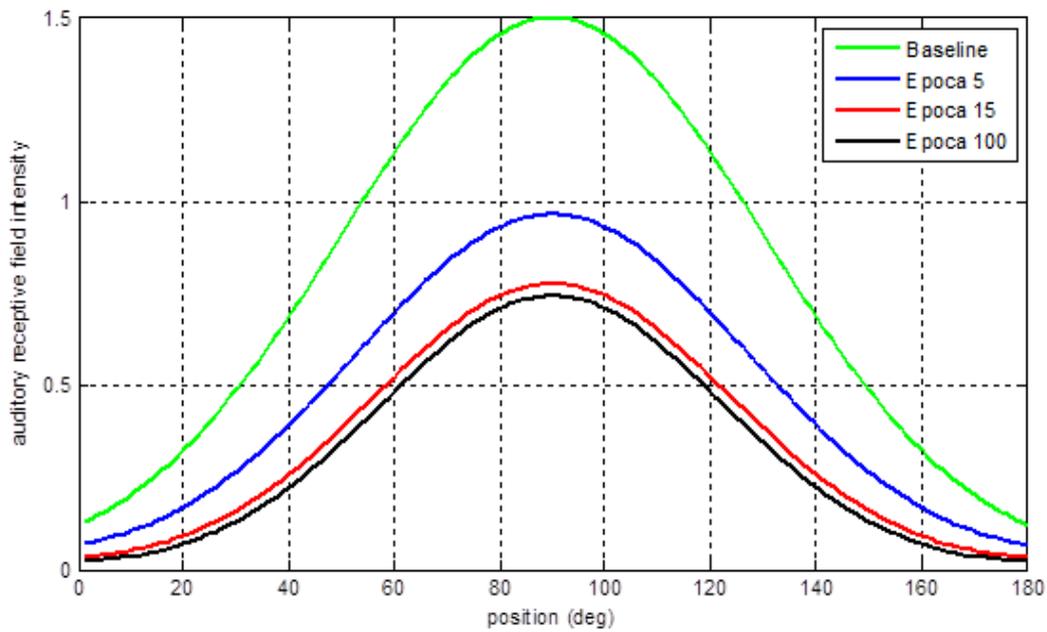


Figura 3.8: Addestramento del campo recettore acustico relativo al neurone in posizione 90

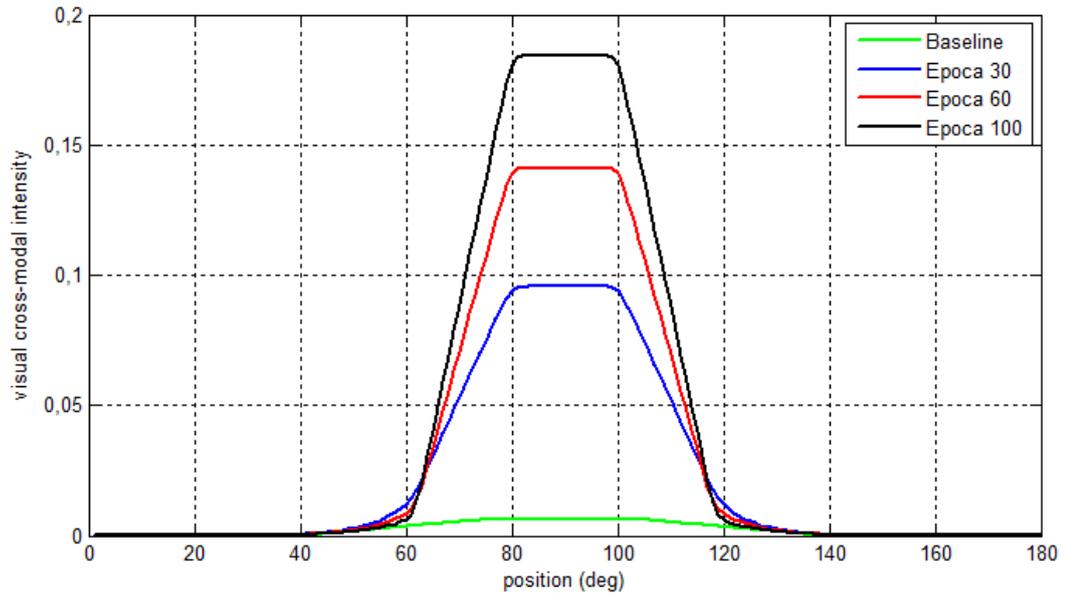


Figura 3.9 Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

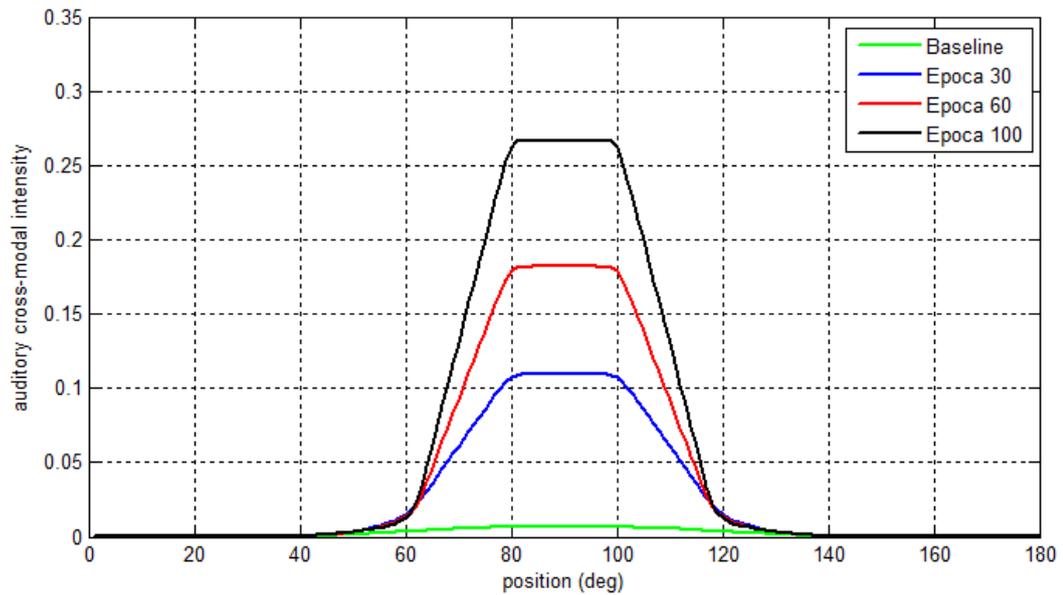


Figura 3.10: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

## 3.2 Caso 2: sinapsi laterali diverse

In questa simulazione la rete è stata addestrata con le sinapsi laterali visive e uditive diverse tra loro, mantenute costanti durante la simulazione (Figura 3.11).

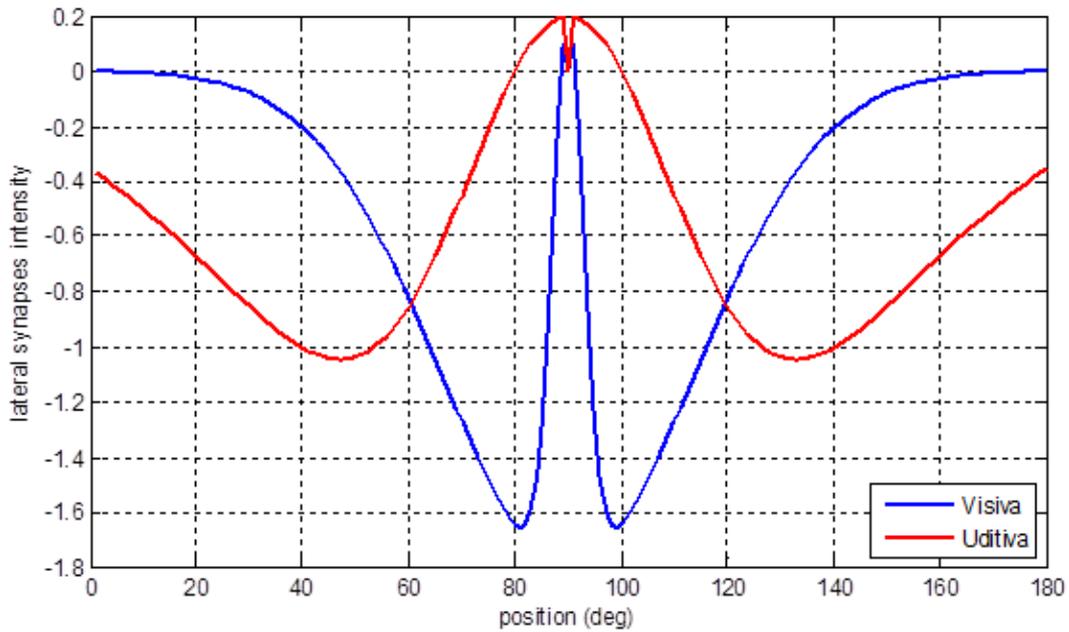


Figura 3.11: Sinapsi laterali visive (in blu) e uditive (in rosso) entranti nel neurone in posizione 90

L'addestramento è stato simulato per un numero complessivo di 100 epoche, in cui sono stati utilizzati i parametri riportati in Tabella 3.3.

External input			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 60$	$\sigma^a = 30$
Visual lateral synapses			
$L_{ex,0} = 2$	$\sigma_{ex} = 3 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 24 \text{ deg}$
Auditory lateral synapses			
$L_{ex,0} = 2$	$\sigma_{ex} = 20 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 50 \text{ deg}$
Hebbian rules			
$\gamma_{RV} = 0.002$		$\gamma_{W0} = 0.0006$	

Tabella 3.3: Parametri relativi alla simulazione del caso 2

Di seguito sono mostrati i grafici relativi allo sviluppo dei campi recettori (Figure 3.12-3.13) e delle sinapsi cross-modali (Figure 3.14-3.15) di entrambe le modalità.

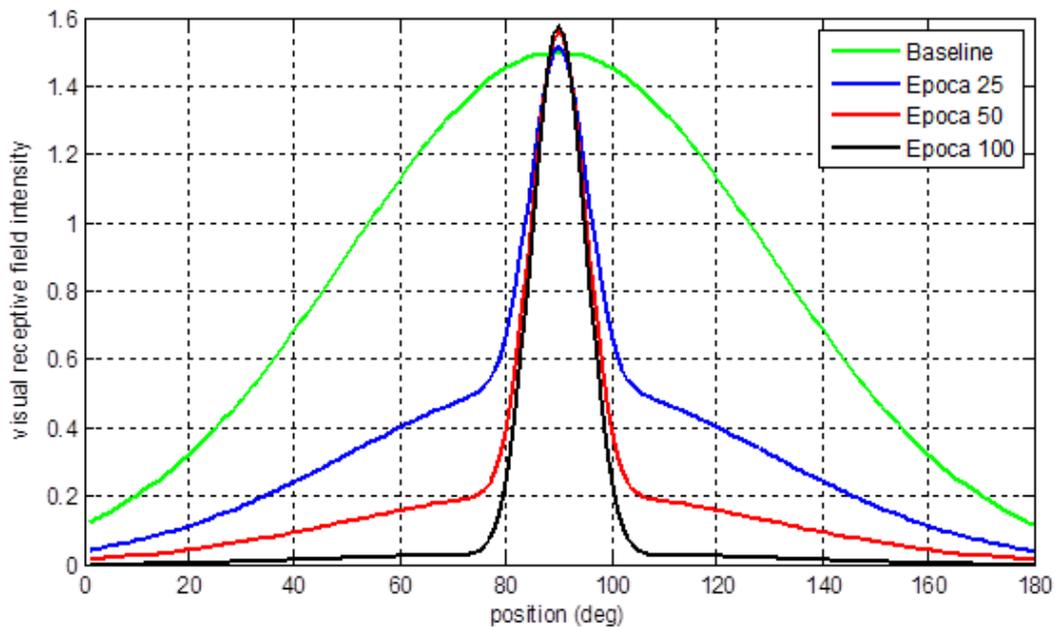


Figura 3.12: Addestramento del campo recettore visivo relativo al neurone in posizione 90

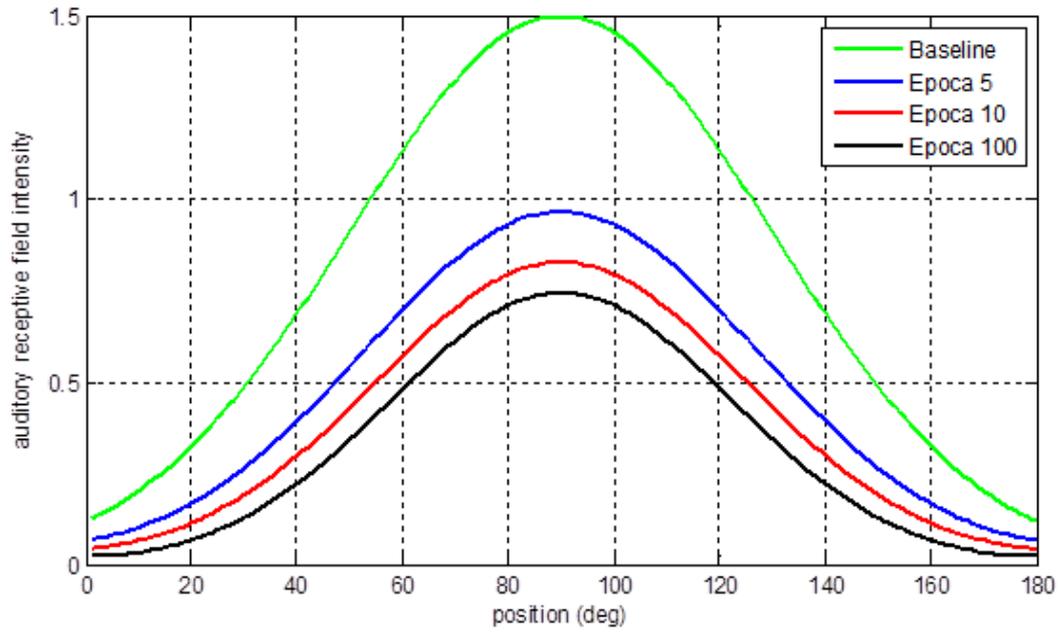


Figura 3.13: Addestramento del campo recettore acustico relativo al neurone in posizione 90

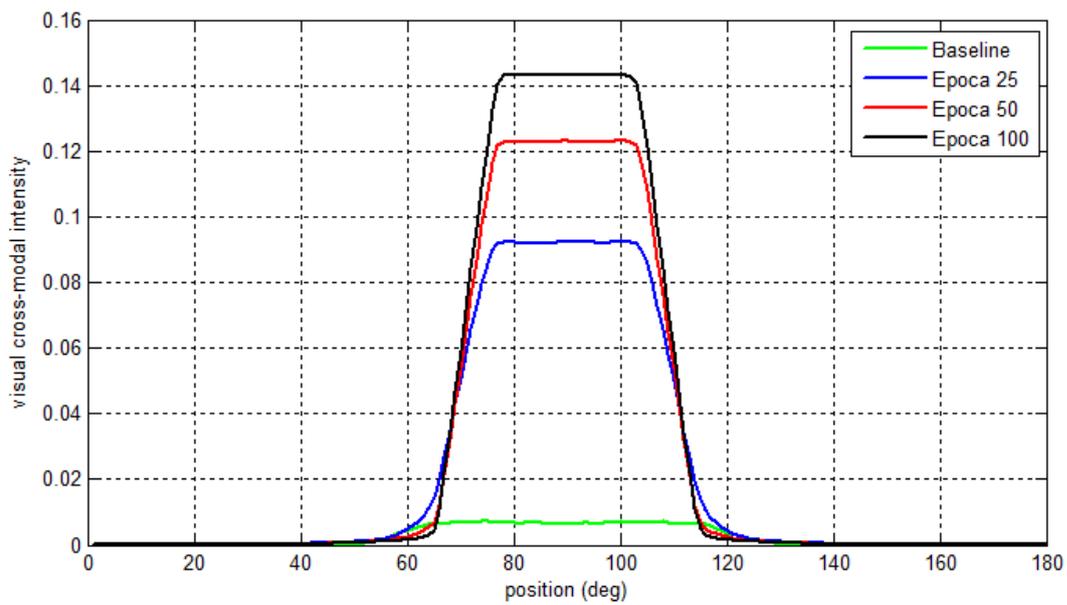


Figura 3.14: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

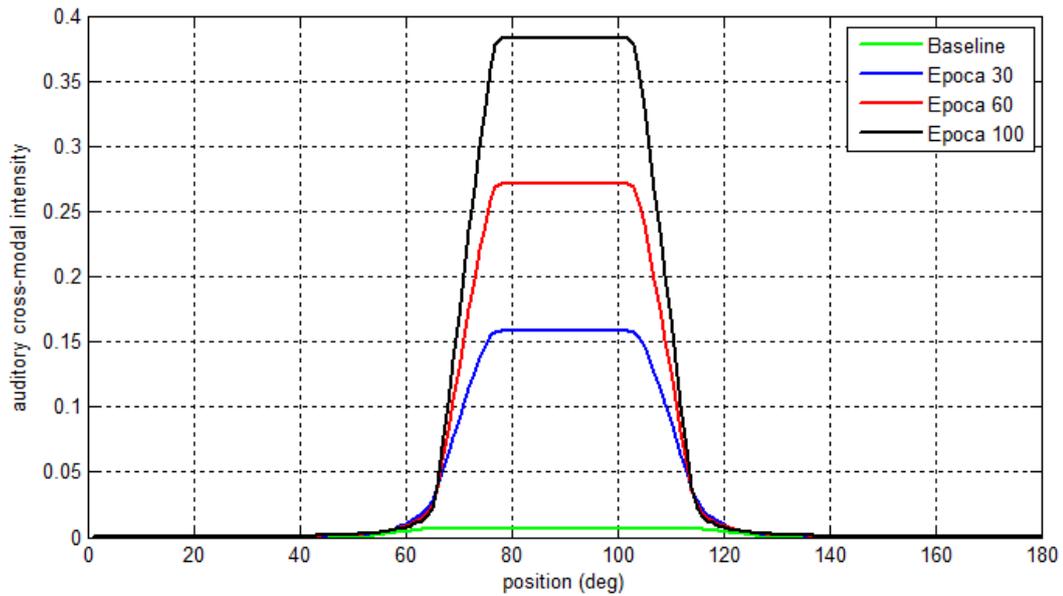


Figura 3.15: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

### 3.3 Caso 3: auto-addestramento delle sinapsi laterali

Per questa specifica simulazione si è assunto che le sinapsi laterali di entrambe le modalità siano plastiche e che possano essere addestrate durante la fase di training sulla base della correlazione tra l'attività neurale post-sinaptica e l'attività neurale pre-sinaptica. Come già visto in precedenza per i campi recettori e le sinapsi cross-modali, anche per le sinapsi laterali è stato utilizzato un addestramento di tipo hebbiano che si basa sulla regola di Oja; in questo frangente le sinapsi si addestrano solo nel caso in cui l'attività post-sinaptica superi una determinata soglia imposta ( $\theta_L$ ). Si sono usati, inoltre, valori diversi per l'attività pre-sinaptica, in modo da indurre un'eccitazione più forte ma più stretta, e un'inibizione a più larga ma a più ampio raggio:

$$\partial L_{ex,jk}^m = \gamma_{L0} [y_j^m - \theta_L]^+ ([y_k^m - \theta_L]^+ - L_{ex,jk}^m)$$

$$\partial L_{in,jk}^m = \gamma_{L0} [y_j^m - \theta_L]^+ (0,8 y_k^m - L_{in,jk}^m)$$

dove  $\gamma_{L0}$  rappresenta la costante di apprendimento delle sinapsi eccitatorie e inibitorie mentre  $[\ ]^+$  è una funzione parte positiva che caratterizza l'addestramento delle sinapsi laterali, ovvero:

$$[y_j^m - \theta_L]^+ = \begin{cases} y_j^m - \theta_L, & \text{se } y_j^m \geq \theta_L \\ 0, & \text{se } y_j^m < \theta_L \end{cases}$$

Per ottenere questa funzione in Matlab è stato applicato il seguente accorgimento matematico:

$$\frac{y_j^m - \theta_L + |y_j^m - \theta_L|}{2} = \begin{cases} y_j^m - \theta_L, & \text{se } y_j^m \geq \theta_L \\ 0, & \text{se } y_j^m < \theta_L \end{cases}$$

La simulazione è stata eseguita per un numero complessivo di 76 epoche, in cui, oltre ai campi recettori (Figure 3.16-3.17) e alle sinapsi cross-modali (Figure 3.18-3.19), è stato analizzato anche l'addestramento delle sinapsi laterali di entrambe le modalità (Figure 3.20-3.21), utilizzando i parametri riportati in Tabella 3.4.

<b>External input</b>			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 43$	$\sigma^a = 30$
<b>Basal lateral synapses</b>			
$L_{jk}^v = 0$		$L_{jk}^a = 0$	
<b>Hebbian rules</b>			
$\gamma_{RV} = 0.002$	$\gamma_{W0} = 0.0002$	$\gamma_{L0} = 0.0005$	
<b>Threshold</b>			
$\theta_L = 0.8$			

Tabella 3.4: Parametri relativi alla simulazione del caso 3

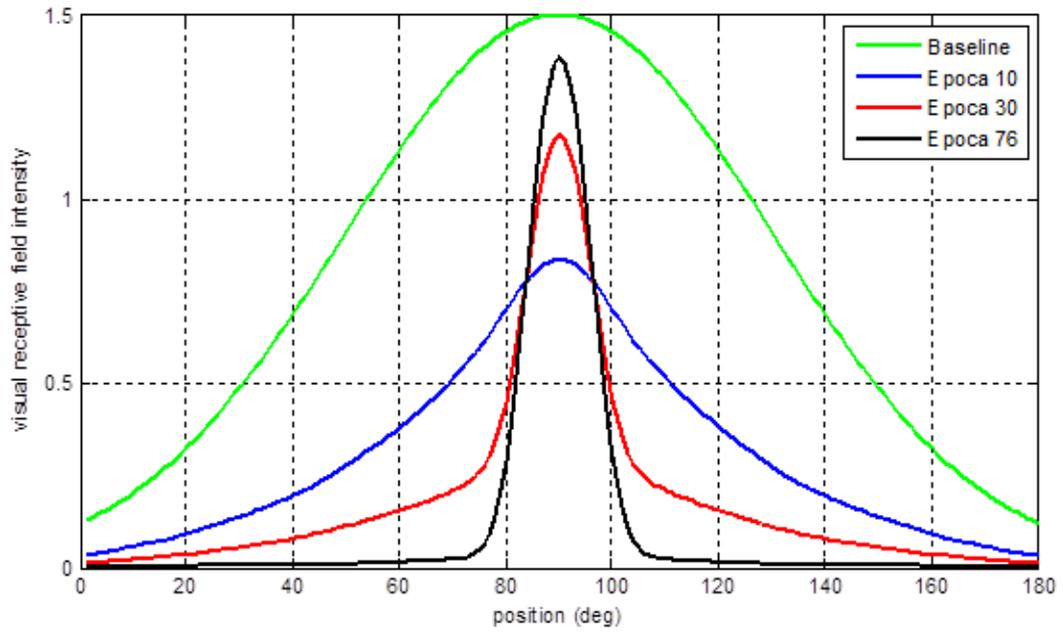


Figura 3.16: Addestramento del campo recettore visivo relativo al neurone in posizione 90

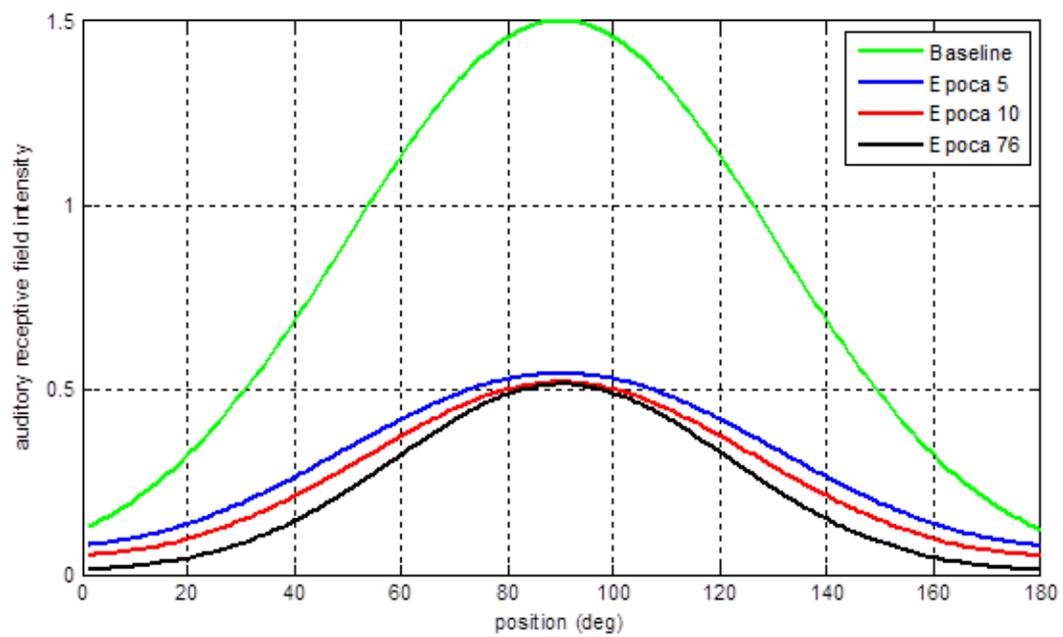


Figura 3.17: Addestramento del campo recettore acustico relativo al neurone in posizione 90

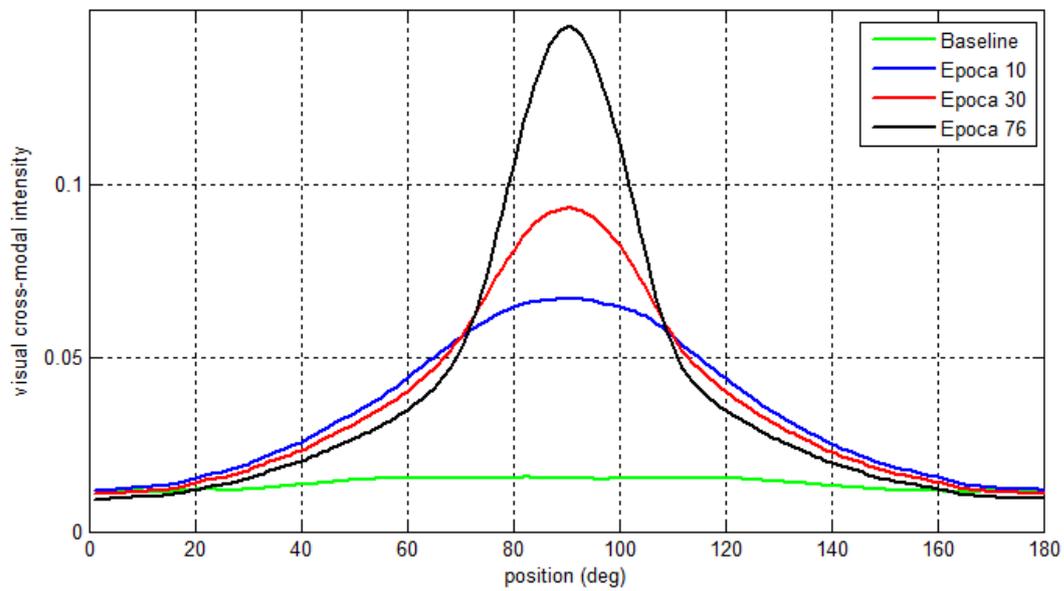


Figura 3.18: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

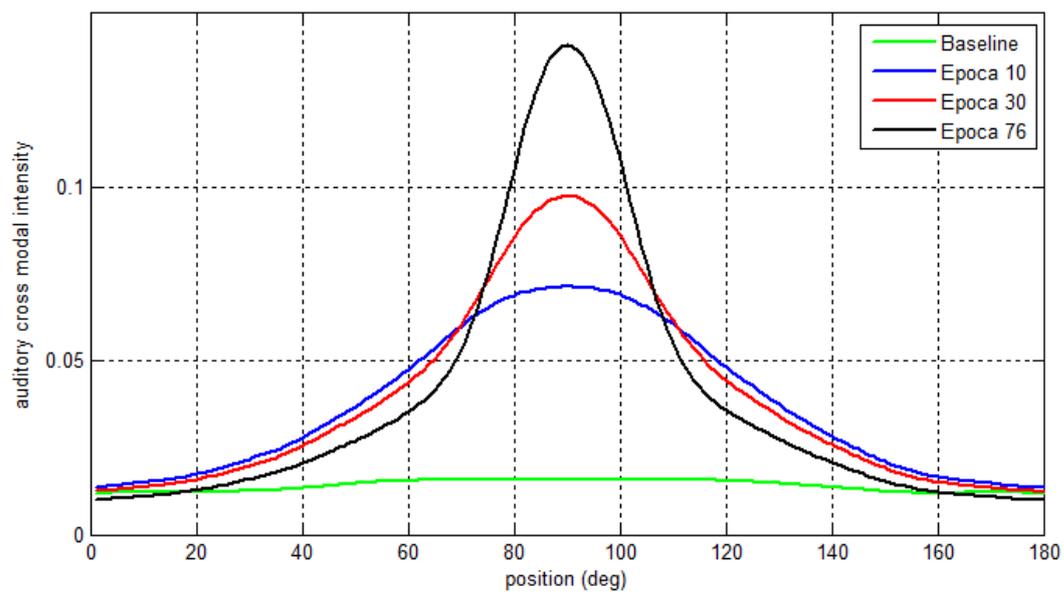


Figura 3.19: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

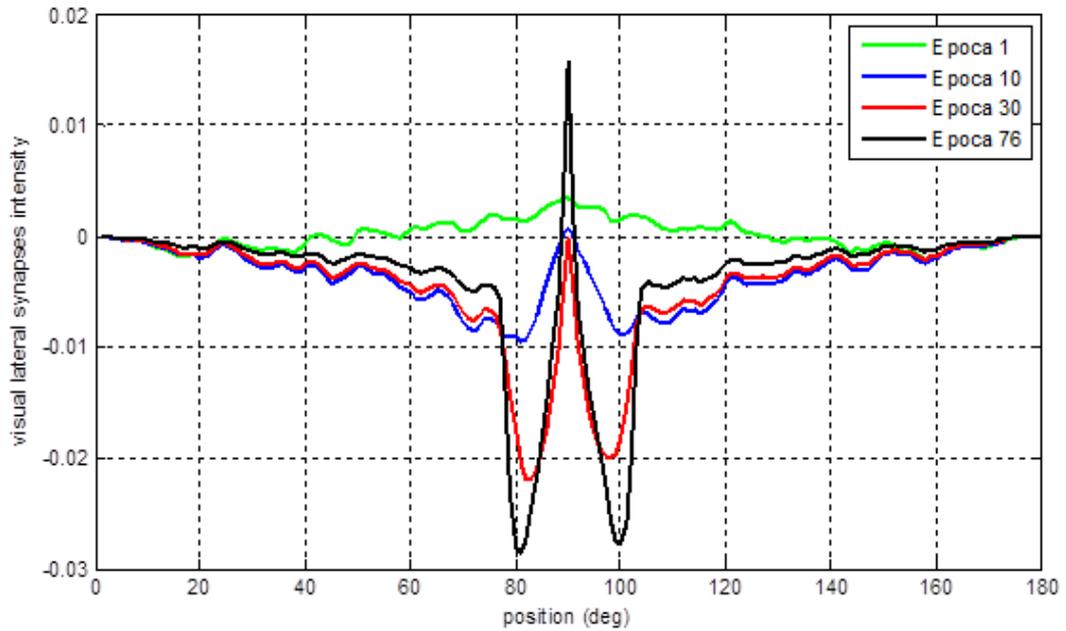


Figura 3.20: Addestramento delle sinapsi laterali visive relative al neurone in posizione 90

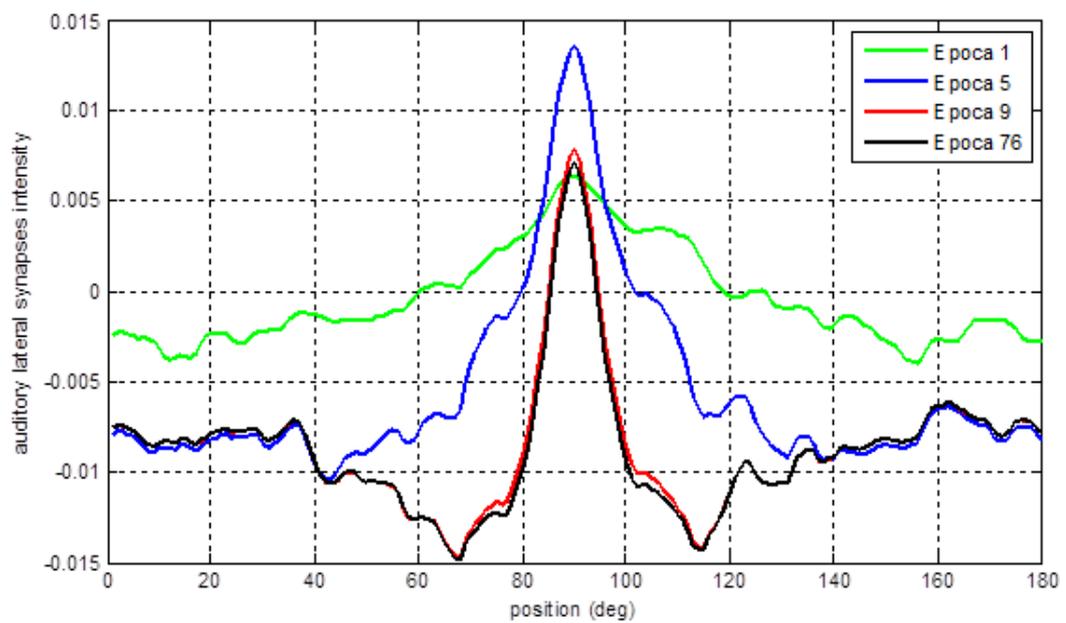


Figura 3.21: Addestramento delle sinapsi laterali acustiche relative al neurone in posizione 90

### 3.4 Studi sul ventriloquismo

Una volta addestrata la rete in tutti e tre i casi appena esposti, è stato verificato se essa era in grado di riprodurre gli effetti dell'integrazione audio-visiva nella corteccia primaria, focalizzando principalmente l'attenzione sul ventriloquismo spaziale. Per simulare tale fenomeno sono stati imposti alla rete uno stimolo esterno visivo e, contemporaneamente, uno stimolo esterno acustico in disparate posizioni spaziali ( $p^v \neq p^a$ ), entrambi della durata di 50 millisecondi. In questo modo lo stimolo acustico dovrebbe progressivamente shiftare in direzione dello stimolo visivo.

In generale, secondo i dati sperimentali raccolti in seguito a studi inerenti, se i due eventi accadono a grande distanza tra loro (superiore a  $40^\circ$ ) le attività neurali si comportano come aree individuali senza un'interazione apprezzabile, mentre se avvengono a una distanza modesta lo stimolo acustico tende a spostarsi verso quello visivo (effetto del ventriloquismo), mostrando uno shift molto alto per un angolo di separazione audio-visivo tra  $20^\circ$ - $30^\circ$ . Al contrario, la percezione dello stimolo visivo mostra generalmente uno shift molto basso in direzione dello stimolo acustico.

Per comprendere meglio la dinamica del ventriloquismo, le Figure 3.22-3.25 riportano i risultati di alcune istantanee dell'attività neurale visiva e acustica durante la presentazione dei due stimoli, relativamente alla rete addestrata nei tre casi visti precedentemente.

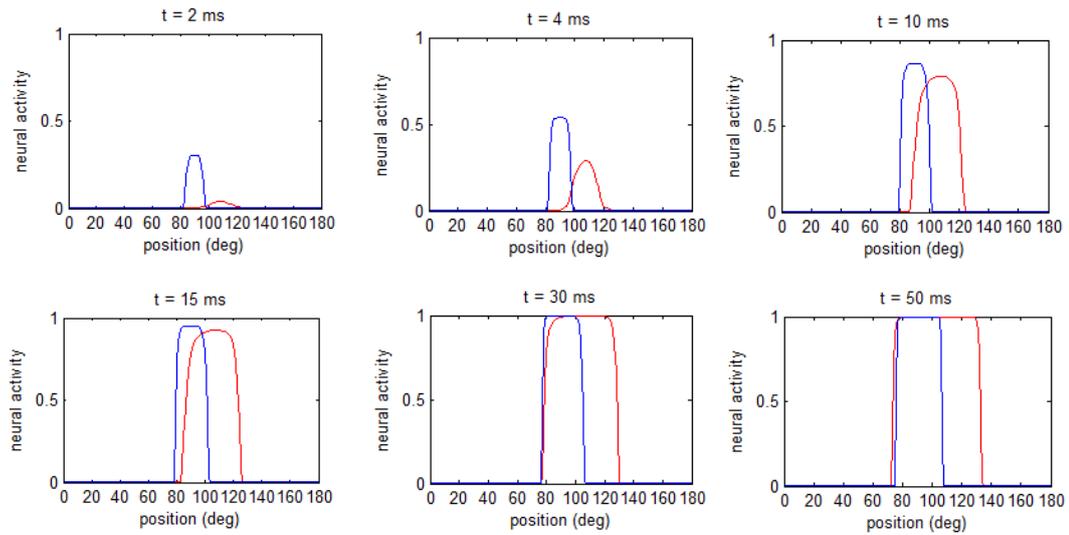


Figura 3.22: Diverse istantanee della risposta neurale visiva (in blu) e acustica (in rosso), relativamente alla rete addestrata nella prima simulazione del caso 1

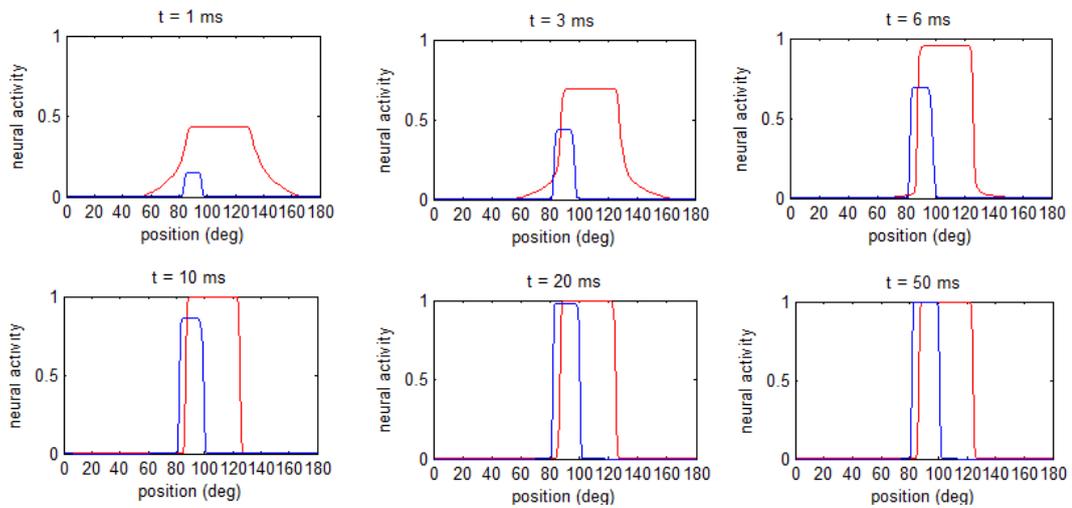


Figura 3.23: Diverse istantanee della risposta neurale visiva (in blu) e acustica (in rosso), relativamente alla rete addestrata nella seconda simulazione del caso 1

### Capitolo 3 – Addestramento della rete

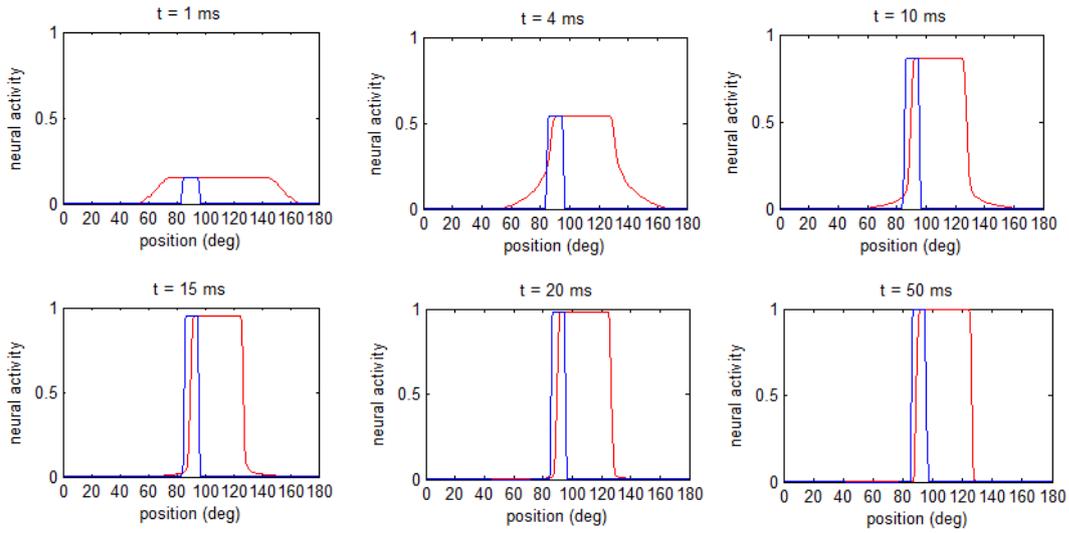


Figura 3.24: Diverse istantanee della risposta neurale visiva (in blu) e acustica (in rosso), relativamente alla rete addestrata nel caso 2

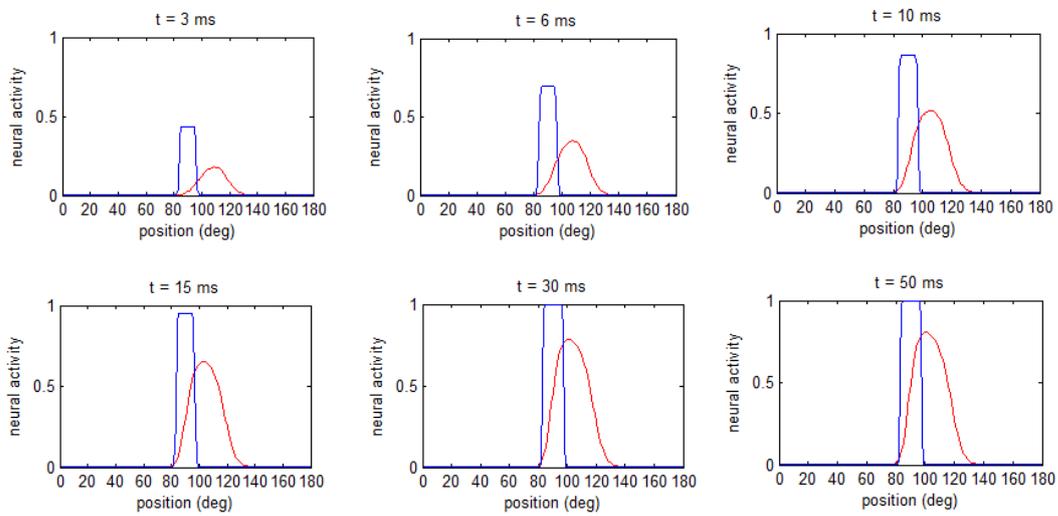


Figura 3.25: Diverse istantanee della risposta neurale visiva (in blu) e acustica (in rosso), relativamente alla rete addestrata nel caso 3

I grafici delle Figure 3.22-3.25 sono stati ottenuti applicando alla rete uno stimolo visivo e uno stimolo acustico temporaneamente coincidenti e spazialmente disparati ( $p^v = 90^\circ$  e  $p^a = 110^\circ$ ). La curva blu rappresenta l'attività neurale visiva, mentre la curva rossa l'attività neurale uditiva. Tutte le simulazioni sono state effettuate con gli stessi ingressi utilizzati nell'addestramento della rete. In tutte e quattro le Figure si osserva che l'attività neurale visiva è minimamente affetta dallo stimolo acustico, mentre l'attività neurale uditiva è maggiormente influenzata dallo stimolo visivo.

Come indice quantitativo della posizione percepita dello stimolo è stata adottata la metrica del baricentro. In accordo con tale metrica, lo stimolo percepito è calcolato come il valor medio (baricentro) dell'attività dell'intera popolazione di neuroni:

$$z_{bar}^m = \frac{\sum_{k=1}^N y_k^m k}{\sum_{k=1}^N y_k^m}$$

dove  $z_{bar}^m$  indica la coordinata del baricentro calcolata nella regione  $m$ ,  $y_k^m$  è l'attività del neurone  $k$ -esimo in tale regione.

In realtà la formula scritta in questa maniera non tiene conto degli effetti di bordo indesiderati, quindi è stato necessario applicare una modifica. Inizialmente si definisce una nuova ascissa ( $z$ ) su cui effettuare il calcolo del baricentro, in base alla posizione centrale dello stimolo:

$$z = \begin{cases} 1,2, \dots, p^m + 89, p^m - 90, p^m - 89, \dots, 0, & \text{se } p^m < 90 \\ 181, 182, \dots, p^m + 90, p^m - 89, p^m - 88, \dots, 180, & \text{se } p^m > 90 \\ 1,2, \dots, 180, & \text{se } p^m = 90 \end{cases}$$

quindi la formula può essere riscritta in questo modo:

$$z_{bar}^m = \frac{\sum_k y_k^m z_k}{\sum_k y_k^m}$$

Lo shift di localizzazione è stato calcolato come la differenza tra la posizione percepita dello stimolo e la posizione di applicazione dello stimolo.

Le Tabelle 3.5 A-D rappresentano i risultati relativi all'analisi del ventriloquismo, in cui vengono riportati i dati relativi all'influenza visiva sulla localizzazione acustica e, viceversa, all'influenza acustica sulla localizzazione visiva in funzione dell'angolo di separazione tra i due stimoli applicati ( $p^a - p^v$ ).

$p^a - p^v$	<i>Shift acustico</i>	<i>Shift visivo</i>
5°	-1.99°	0.3°
10°	-3.78°	0.58°
20°	-6.89°	1.3°
25°	-7.95°	1.59°
30°	-8.33°	1.7°
40°	-5.92°	1.28°

Tabella 3.5 A: Shift di localizzazione visivo e acustico in funzione della separazione angolare tra i due stimoli applicati, relativamente alla prima simulazione del caso 1

$p^a - p^v$	<i>Shift acustico</i>	<i>Shift visivo</i>
5°	-1.57°	0.44°
10°	-2.96°	0.81°
20°	-4.8°	1.37°
25°	-4.87°	1.43°
30°	-4.47°	1.26°
40°	-2.22°	0.63°

Tabella 3.5 B: Shift di localizzazione visivo e acustico in funzione della separazione angolare tra i due stimoli applicati, relativamente alla seconda simulazione del caso 1

$p^a - p^v$	<i>Shift acustico</i>	<i>Shift visivo</i>
5°	-1.28°	0.48°
10°	-2.15°	0.8°
20°	-2.3°	0.74°
25°	-2.13°	0.82°
30°	-2.01°	0.94°
40°	-0.69°	0.33°

Tabella 3.5 C: Shift di localizzazione visivo e acustico in funzione della separazione angolare tra i due stimoli applicati, relativamente al caso 2

$p^a - p^v$	<i>Shift acustico</i>	<i>Shift visivo</i>
5°	-1.8°	0.15°
10°	-3.73°	0.22°
20°	-6.3°	0.25°
25°	-5.8°	0.2°
30°	-3.9°	0.09°
40°	-1.5°	0°

Tabella 3.5 D: Shift di localizzazione visivo e acustico in funzione della separazione angolare tra i due stimoli applicati, relativamente al caso 3

Le successive Figure 3.26-3.27 rappresentano dei grafici riassuntivi che mostrano come gli stimoli visivi e acustici, applicati simultaneamente, si influenzano reciprocamente variando il proprio angolo di separazione. In particolare viene evidenziato lo shift di percezione dello stimolo visivo e di quello acustico, ovvero la differenza tra la posizione percepita e la posizione originale di applicazione dello stimolo, in funzione della distanza  $p^v - p^a$ .

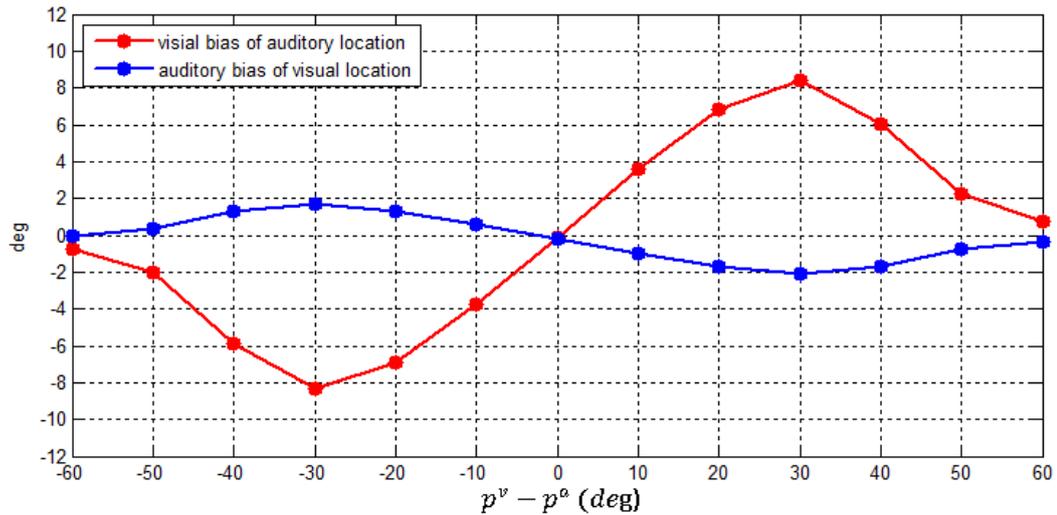


Figura 3.26: Shift di localizzazione dello stimolo acustico (in rosso) e dello stimolo visivo (in blu) in funzione del loro angolo del loro angolo di separazione, relativamente alla prima simulazione del caso 1

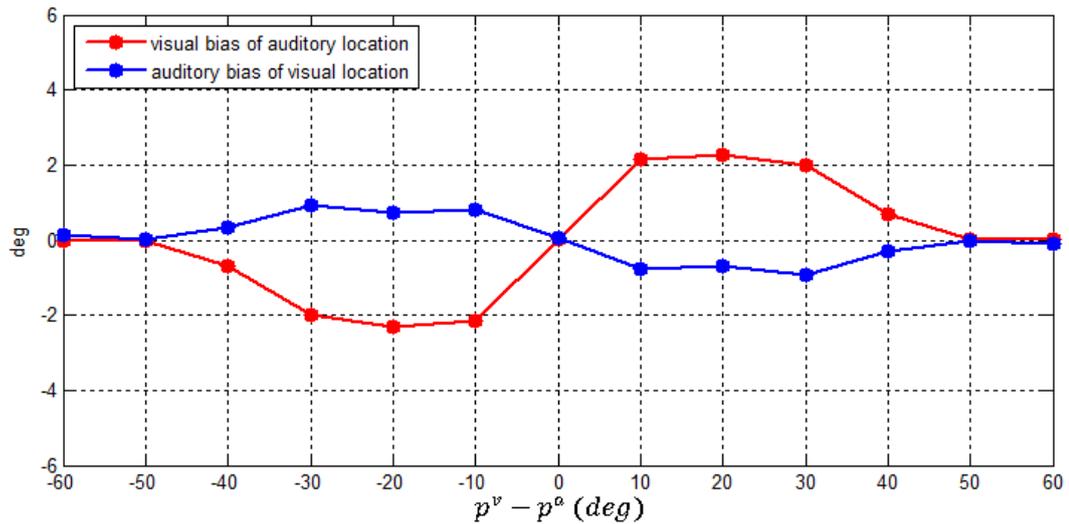


Figura 3.27: Shift di localizzazione dello stimolo acustico (in rosso) e dello stimolo visivo (in blu) in funzione del loro angolo del loro angolo di separazione, relativamente al caso 2

Le Figure 3.26-3.27 sono state ottenute mantenendo lo stimolo visivo in una posizione fissa ( $p^v = 90^\circ$ ) e variando la posizione dello stimolo acustico ( $p^a$ ) da  $30^\circ$  a  $150^\circ$ .

Dalle Tabelle 3.5 A-D e dalle Figure 3.26-3.27 si osserva che nel caso 1 si ottiene generalmente un'ottima simulazione del ventriloquismo spaziale, notando un'elevata influenza visiva sulla modalità acustica e una scarsa influenza acustica sulla modalità visiva, in accordo con i dati sperimentali. Anche nel caso 3 viene evidenziata una buona riproduzione del ventriloquismo, ma a questo proposito è necessario approfondire maggiormente lo studio sull'addestramento delle sinapsi laterali. Invece nel caso 2 si evince una scarsa influenza tra le due modalità: in questo frangente la riproduzione del ventriloquismo è in disaccordo con i dati sperimentali.

### 3.5 Osservazioni

In tutti i casi visti fino ad ora sono stati ottenuti degli ottimi risultati relativi all'addestramento dei campi recettori visivi e acustici, notando come essi si restringano in modo tale che la percezione spaziale degli input diventi più precisa. In alcune prove si ottiene anche una buona simulazione dell'effetto ventriloquismo. Tuttavia ci sono alcuni punti ancora irrisolti. In primo luogo, come già detto, nell'auto-addestramento delle sinapsi laterali si nota che esse tendono a rimanere molto deboli e quindi poco attendibili dal punto di vista fisiologico. Altre prove svolte (non riportate), in cui sono state utilizzate delle costanti di apprendimento più alte, hanno dimostrato come tali sinapsi continuerebbero ad addestrarsi indefinitamente fino, talvolta, a mostrare instabilità. Inoltre dai grafici delle sinapsi cross-modali si osserva che, volendo addestrare la rete per un numero di epoche molto alto (situazione simile a quella fisiologica), esse tenderebbero a crescere ulteriormente all'infinito, e quindi anche i risultati relativi al ventriloquismo sarebbero diversi. Per superare questi limiti occorre quindi estendere il modello con l'introduzione di nuove varianti.

# Capitolo 4

## Estensione del modello e comportamento della rete

### 4.1 Riduzione delle costanti di apprendimento

Il modello implementato e simulato nei tre casi esaminati in precedenza fornisce già degli ottimi risultati, in modo particolare sullo sviluppo dei campi recettori e, in alcune prove, sulla riproduzione del ventriloquismo.

Dai grafici delle sinapsi cross-modali di tutte le prove eseguite in precedenza si nota come esse, nel caso si volesse proseguire la simulazione per un numero molto elevato di epoche (riproducendo così una reale condizione fisiologica), continuerebbero a crescere progressivamente all'infinito. Ciò non è fisiologicamente corretto, in quanto è vero che le sinapsi cross-modali si addestrano progressivamente nel tempo, ma per un elevato numero di epoche queste dovrebbero saturare a un valore massimo.

Secondo quanto detto fino ad ora l'idea è quella di introdurre nel modello una variante, in modo tale che le sinapsi cross-modali non possano superare un determinato valore massimo imposto di saturazione fisiologica ( $W_{max}$ ), permettendo così alla costante di apprendimento hebbiana di decrescere progressivamente ad ogni epoca.

La regola di addestramento delle sinapsi cross-modali è stata riformulata con l'introduzione di un nuovo fattore di apprendimento ( $\gamma_W$ ):

$$\partial W_{jk}^m = \gamma_W y_j^m (y_k^n - W_{jk}^m), \quad \text{con } m \neq n$$

$$\text{dove } \gamma_W = \gamma_{W0} (W_{max} - \sum_k W_{jk}^m)$$

Da quest'ultima equazione si deduce che  $\gamma_W$  decresce progressivamente all'aumentare della relativa sinapsi cross-modale.

Per testare la nuova variante del modello è stata effettuata una simulazione relativa al caso 2 (Tabella 3.3), ovvero con sinapsi laterali visive e acustiche fisse e diverse tra loro. In questo esempio la prova è stata eseguita per un numero complessivo di 70 epoche, in cui è stato analizzato l'addestramento delle sinapsi cross-modali visive (Figura 4.1), imponendo come valore massimo di saturazione  $W_{max} = 1$ .

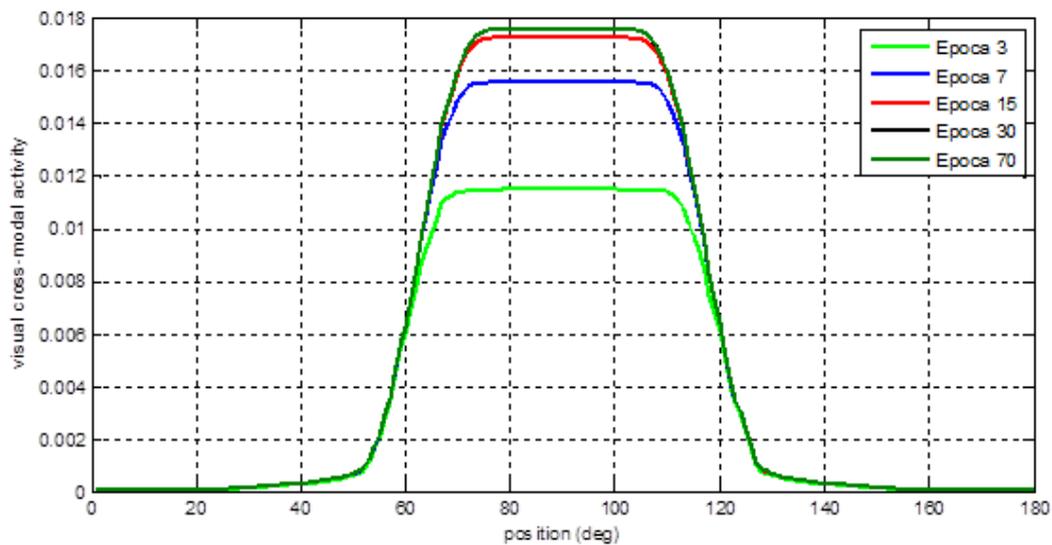


Figura 4.1: Esempio di sinapsi cross-modali visive relative al neurone in posizione 90, con riduzione progressiva del fattore di apprendimento

Come previsto, dalla Figura 4.1 si può notare come le sinapsi cross-modali inizialmente si addestrino in maniera forte, per poi stabilizzarsi all'aumentare del numero di epoche a causa della riduzione progressiva del fattore di apprendimento ( $\gamma_W$ ), fino a tendere a 0 (cioè quando  $W_{max} = \sum_k W_{jk}^m$ ). In questo modo, imponendo un valore massimo appropriato di saturazione, è possibile ottenere un addestramento adeguato delle sinapsi cross-modali anche per un elevato numero di epoche.

Un approccio simile è stato utilizzato anche nell'auto-addestramento delle sinapsi laterali. Infatti nel caso 3, studiato in precedenza, è stata fornita in ingresso una costante di apprendimento molto bassa, ottenendo così delle sinapsi laterali molto deboli e poco facilmente addestrabili, come si può notare dalle Figure 3.20-3.21. Altre prove svolte (non riportate), in cui sono state utilizzate delle costanti di apprendimento più alte, hanno dimostrato come tali sinapsi continuerebbero ad addestrarsi per un numero infinito di epoche, fino a mostrare instabilità. Per far fronte a questo inconveniente sono stati introdotti due nuovi fattori di apprendimento per le sinapsi laterali eccitatorie e inibitorie,  $\gamma_{L,ex}$  e  $\gamma_{L,in}$  rispettivamente, che possano decrescere all'aumentare della forza sinaptica:

$$\gamma_{L,ex} = \gamma_{L0} (L_{ex,max} - \sum_k L_{ex,jk}^m)$$

$$\gamma_{L,in} = \gamma_{L0} (L_{in,max} - \sum_k L_{in,jk}^m)$$

di conseguenza la regola di addestramento delle sinapsi laterali con l'introduzione dei nuovi fattori di apprendimento diventa:

$$\partial L_{ex,jk}^m = \gamma_{L,ex} [y_j^m - \theta_L]^+ ([y_k^m - \theta_L]^+ - L_{ex,jk}^m)$$

$$\partial L_{in,jk}^m = \gamma_{L,in} [y_j^m - \theta_L]^+ (0,8 y_k^m - L_{in,jk}^m)$$

in termini fisiologici, nel modello sono state introdotte delle saturazioni in modo tale che le sinapsi laterali eccitatorie e inibitorie non possano superare un valore massimo imposto, rispettivamente  $L_{ex,max}$  e  $L_{in,max}$ .

## 4.2 Simulazioni con auto-addestramento delle sinapsi laterali

Una volta introdotte nel modello le due varianti appena descritte, sono state effettuate due simulazioni in cui si è assunta, per entrambe, la plasticità e l'auto-

addestramento delle sinapsi laterali durante la fase di training, analizzando in particolare il comportamento della rete in seguito alla variazione dei due parametri  $\theta_L$  e  $\gamma_{L0}$ . In queste due prove sono stati modificati diversi parametri del modello in condizioni basali, come viene riportato nella Tabella sottostante.

<b>Neurons</b>			
<b>N = 180</b>	<b><math>\tau = 5</math></b>	<b>s = 0.5</b>	<b><math>\theta = 16</math></b>
<b>Basal Receptive field (RF)</b>			
<b><math>RF_0 = 0.7</math></b>		<b><math>\sigma_v = 40</math></b>	
<b>Basal cross-modal synapses</b>			
<b><math>W_{jk}^v = 0</math></b>		<b><math>W_{jk}^a = 0</math></b>	
<b>Basal lateral synapses</b>			
<b><math>L_{ex,jk}^v = 0</math></b>		<b><math>L_{in,jk}^a = 0</math></b>	

Tabella 4.1: Parametri del modello in condizioni basali

## 4.2.1 Prima simulazione

In questa prima simulazione la rete è stata addestrata per un numero complessivo di 70 epoche, in cui è stato analizzato lo sviluppo dei campi recettori (Figure 4.2-4.3), delle sinapsi cross-modali (Figure 4.4-4.5) e delle sinapsi laterali (Figura 4.6). I parametri utilizzati nella prova sono riportati in Tabella 4.2.

External stimuli			
$E_0^v = 30$	$\sigma^v = 4$	$E_0^a = 60$	$\sigma^a = 30$
Hebbian rules			
$\gamma_{RV} = 0.002$	$\gamma_{W0} = 0.0004$	$\gamma_{L0} = 0.001$	
Threshold			
$\theta_L = 0.85$	$L_{ex,max} = 100$	$L_{in,max} = 175$	$W_{max} = 30$

Tabella 4.2: Parametri relativi alla prima simulazione con auto-addestramento delle sinapsi laterali

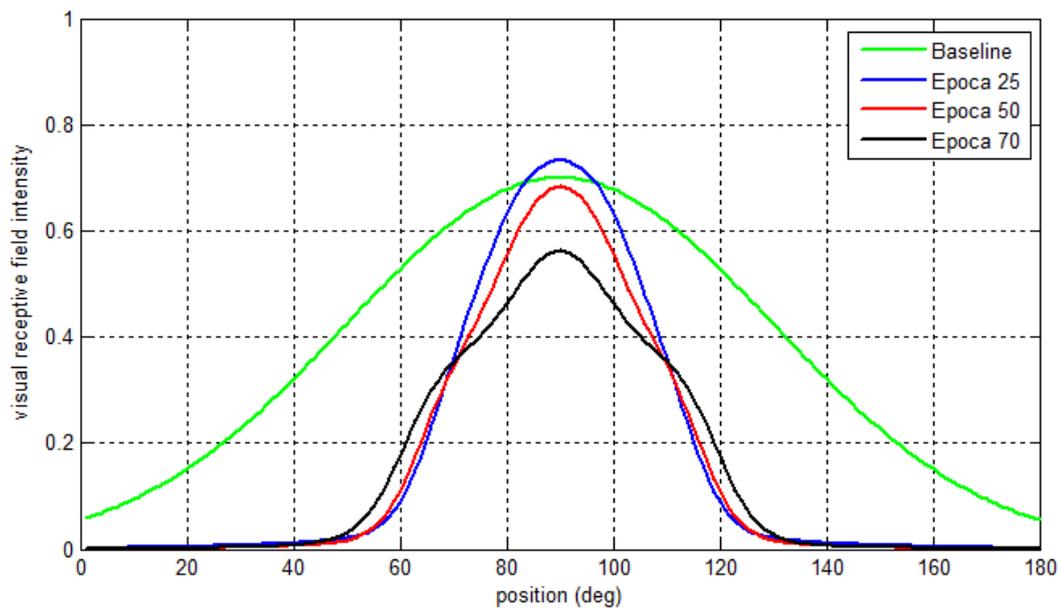


Figura 4.2: Addestramento del campo recettore visivo relativo al neurone in posizione 90

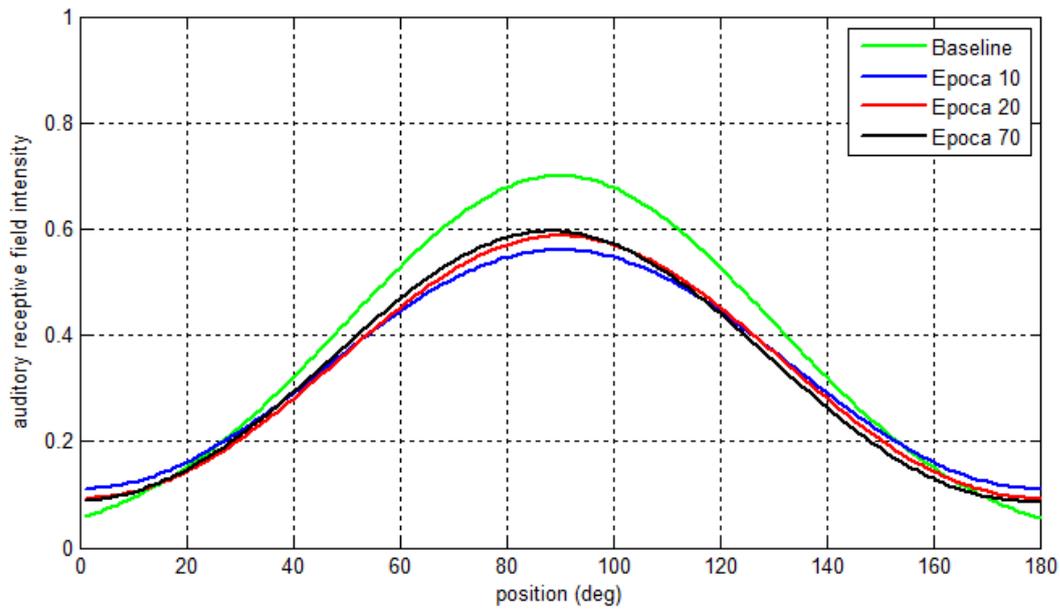


Figura 4.3: Addestramento del campo recettore acustico relativo al neurone in posizione 90

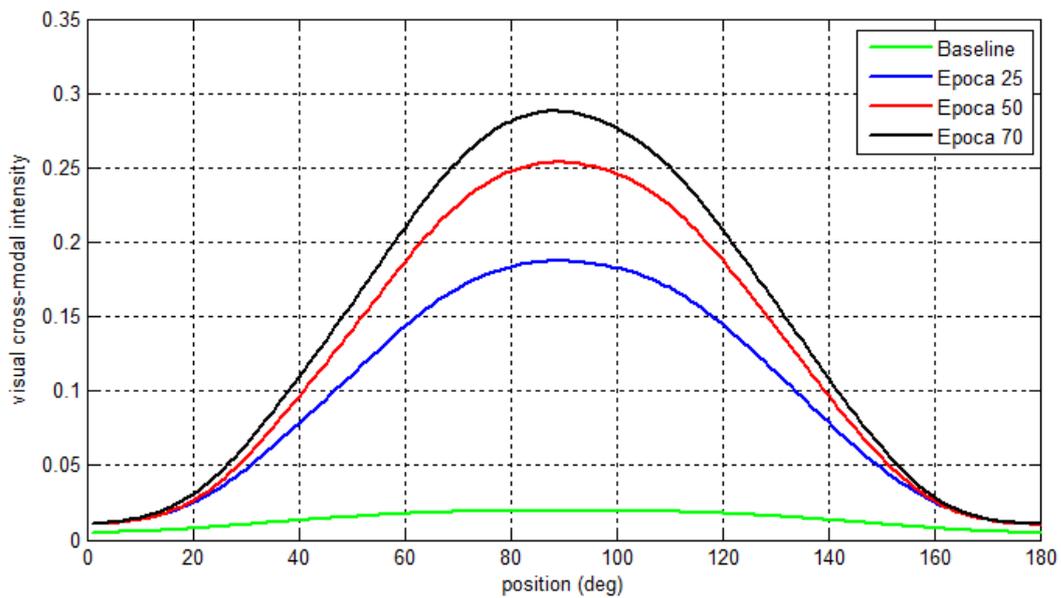


Figura 4.4: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

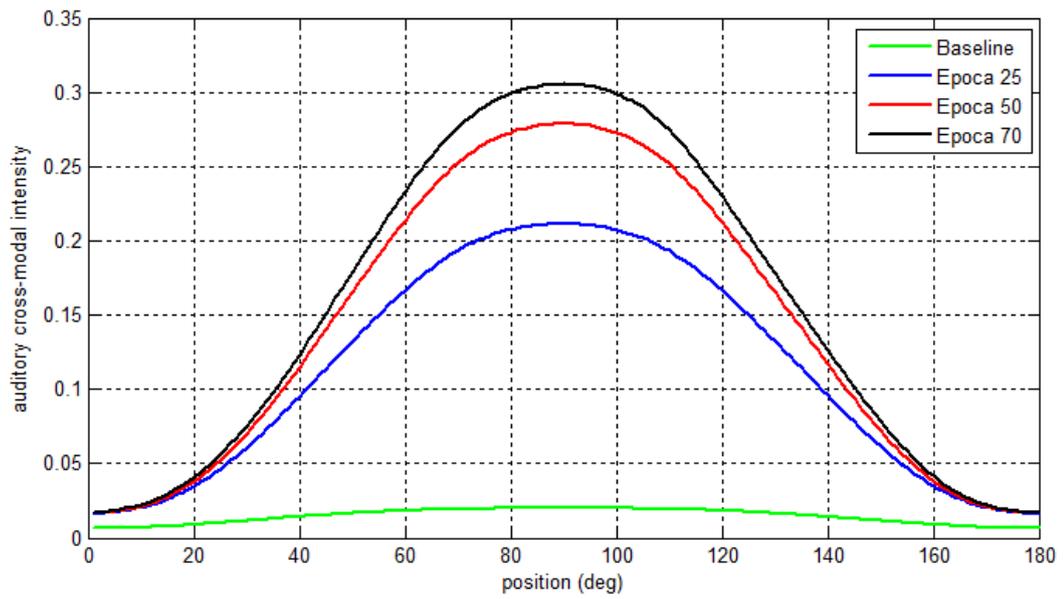


Figura 4.5: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

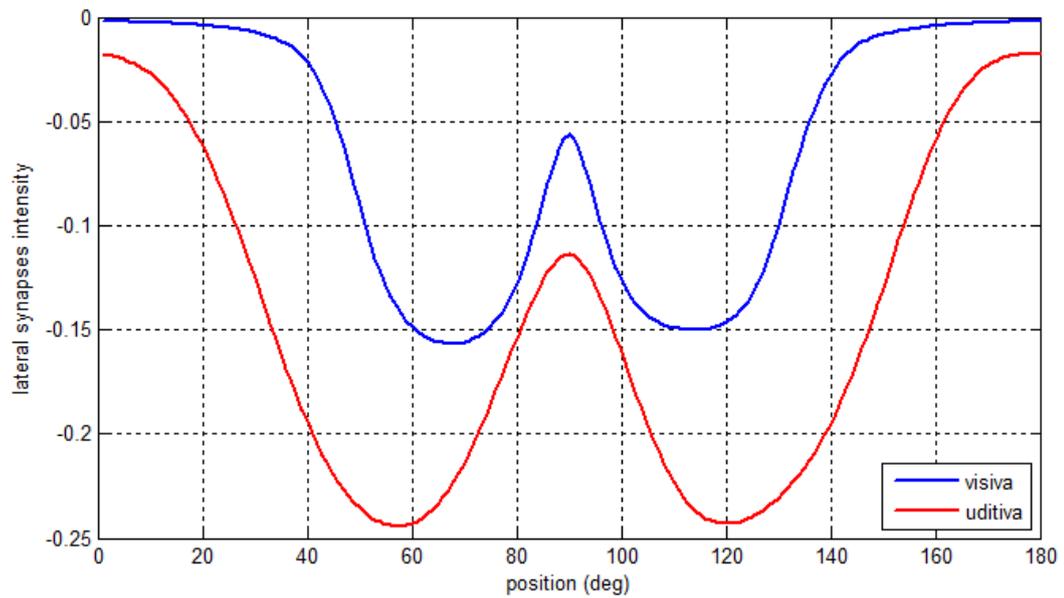


Figura 4.6: Risultato finale dell'addestramento delle sinapsi laterali visive (in blu) e acustiche (in rosso), relative al neurone in posizione 90

Dai grafici di questa prova si osserva un ottimo addestramento delle sinapsi laterali, di gran lunga migliore rispetto a quello simulato nel caso 3 della precedente versione del modello (Paragrafo 3.3, Figure 3.20-3.21 ), ma si presentano fortemente inibitorie e prive di capacità eccitatorie. In più le sinapsi cross-modali rimangono molto larghe rispetto alle simulazioni precedenti, ma rispecchiano in pieno la variante introdotta nel modello, notando come esse tendano a stabilizzarsi al valore massimo di saturazione per un elevato numero di epoche. Il risultato più strano è rappresentato dall'addestramento del campo recettore visivo: dopo un certo numero di epoche esso si allarga e si abbassa in maniera anomala, diversamente da quanto visto nelle precedenti simulazioni. Alla luce dei risultati ottenuti si può già prevedere una riproduzione negativa del ventriloquismo, come mostrato in figura 4.7.

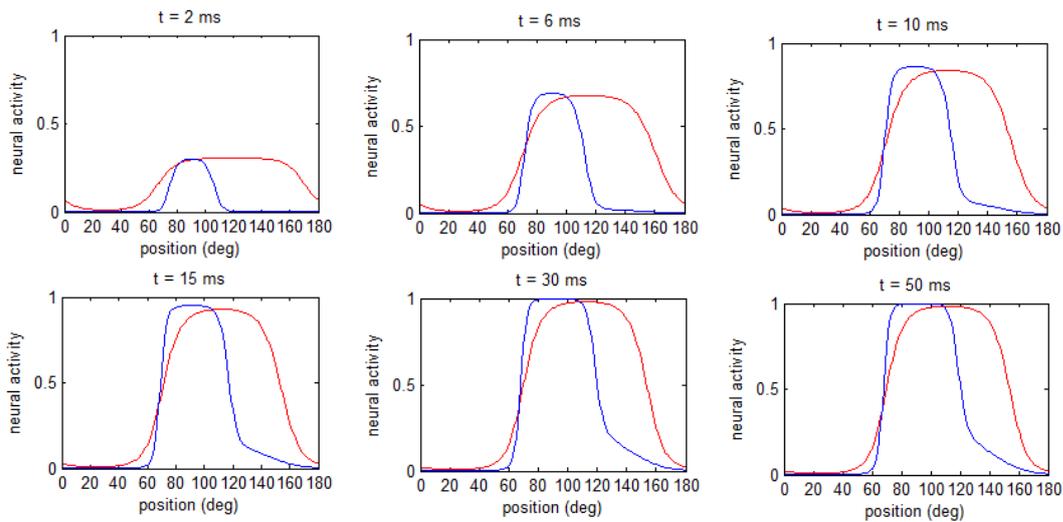


Figura 4.7: Diverse istantanee della risposta neurale visiva (in blu) e acustica (in rosso) relative alla prima simulazione con auto-addestramento delle sinapsi laterali

La Figura 4.7 è stata ottenuta applicando alla rete uno stimolo visivo e uno stimolo acustico temporaneamente coincidenti ma spazialmente disparati ( $p^v = 90^\circ$  e  $p^a = 120^\circ$ ). La curva blu rappresenta l'attività neurale visiva, mentre la curva rossa l'attività neurale uditiva.

Dalla simulazione riportata in Figura 4.7 è stato calcolato uno shift di percezione dello stimolo acustico di circa  $-8^\circ$ , ma allo stesso tempo uno shift di percezione visivo di  $7.6^\circ$ , molto elevato rispetto alla reale influenza acustica sulla modalità visiva. Quindi i risultati scaturiti dall'analisi del ventriloquismo sono molto difforni da quelli precedenti e dissimili rispetto ai dati sperimentali, da cui ci si aspetta una larga influenza visiva sulla modalità acustica e, al contrario, una scarsa influenza acustica sulla modalità visiva.

## 4.2.2 Seconda simulazione

Questo addestramento è stato simulato nelle stesse condizioni della prima prova, mantenendo quasi tutti i parametri utilizzati precedentemente ad eccezione di  $\theta_L$  e  $\gamma_{L0}$ , come riportato in Tabella 4.3.

External stimuli			
$E_0^v = 30$	$\sigma^v = 4$	$E_0^a = 60$	$\sigma^a = 30$
Hebbian rules			
$\gamma_{RV} = 0.002$	$\gamma_{W0} = 0.0004$	$\gamma_{L0} = 0.0008$	
Threshold			
$\theta_L = 0.8$	$L_{ex,max} = 100$	$L_{in,max} = 175$	$W_{max} = 30$

Tabella 4.3: Parametri relativi alla seconda simulazione con auto-addestramento delle sinapsi laterali

Anche in questo caso la prova è stata simulata per un numero complessivo di 70 epoche ed è stato analizzato l'addestramento dei campi recettori (Figure 4.8-4.9), delle sinapsi cross-modali (Figure 4.10-4.11) e delle sinapsi laterali (Figura 4.12).

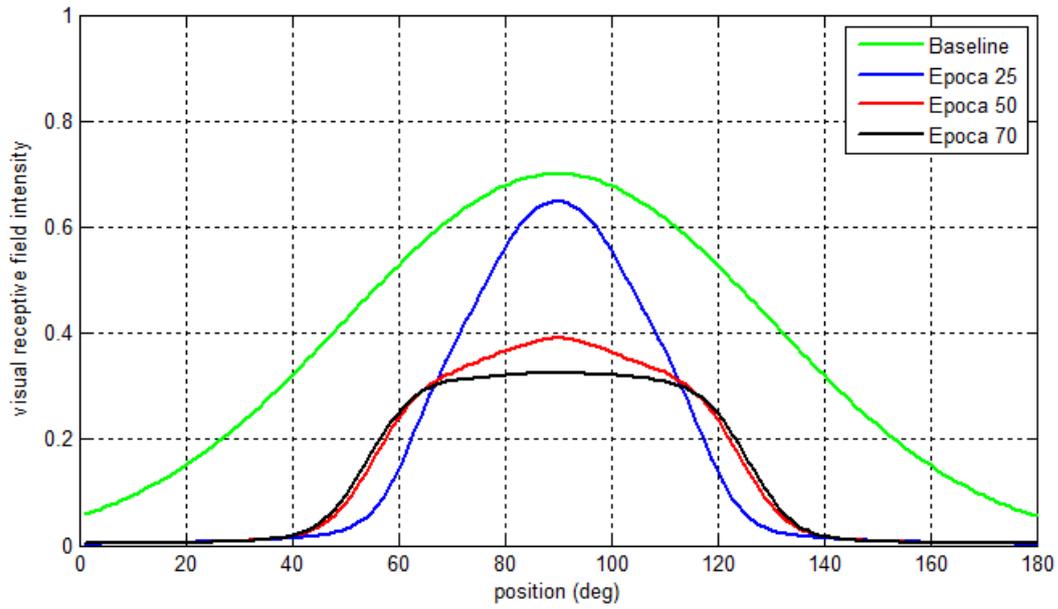


Figura 4.8: Addestramento del campo recettore visivo relativo al neurone in posizione 90

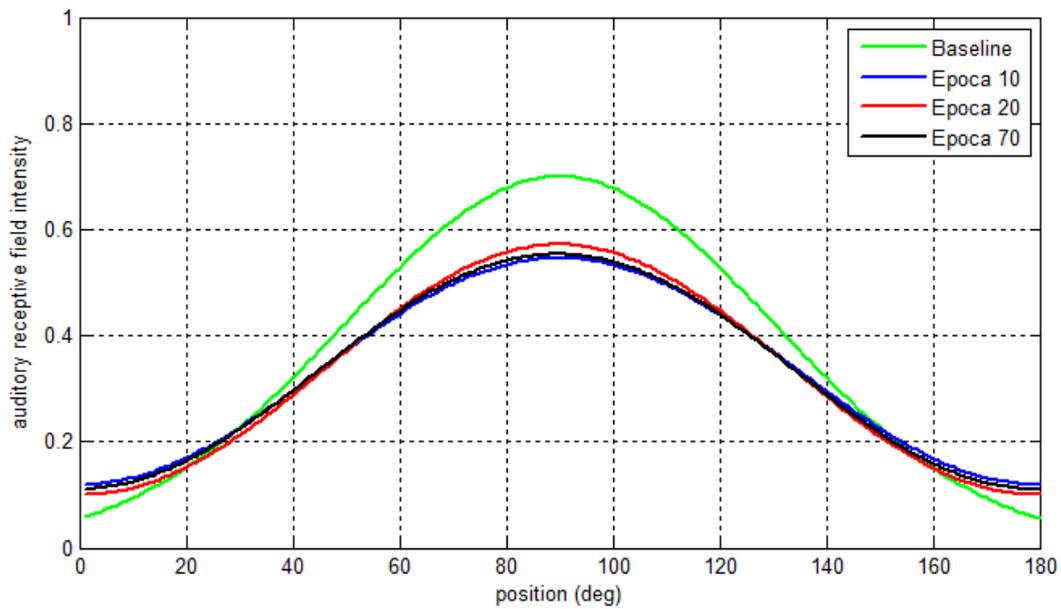


Figura 4.9: Addestramento del campo recettore acustico relativo al neurone in posizione 90

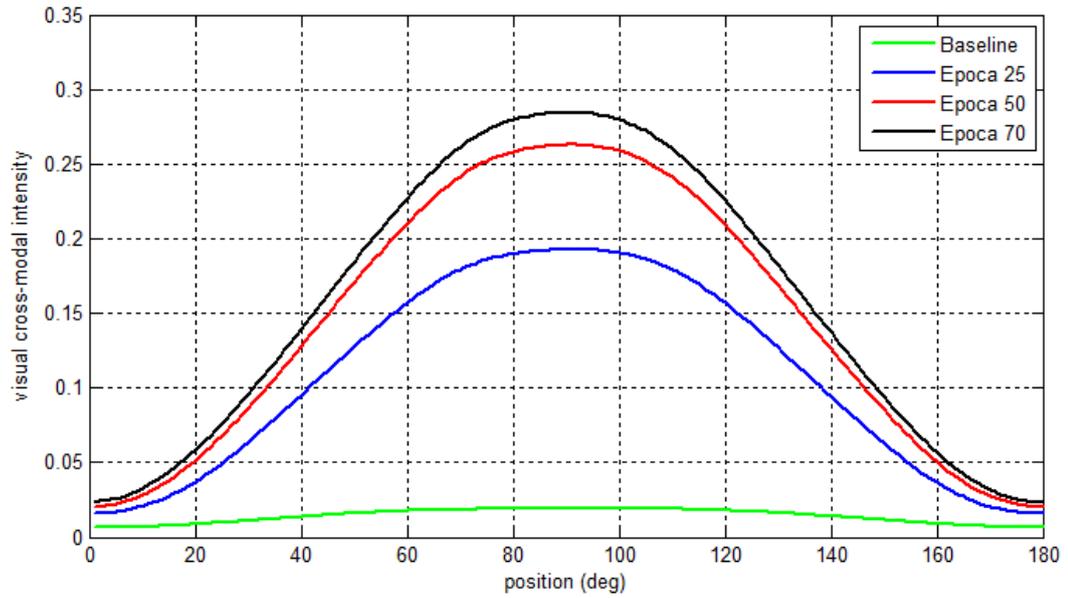


Figura 4.10: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

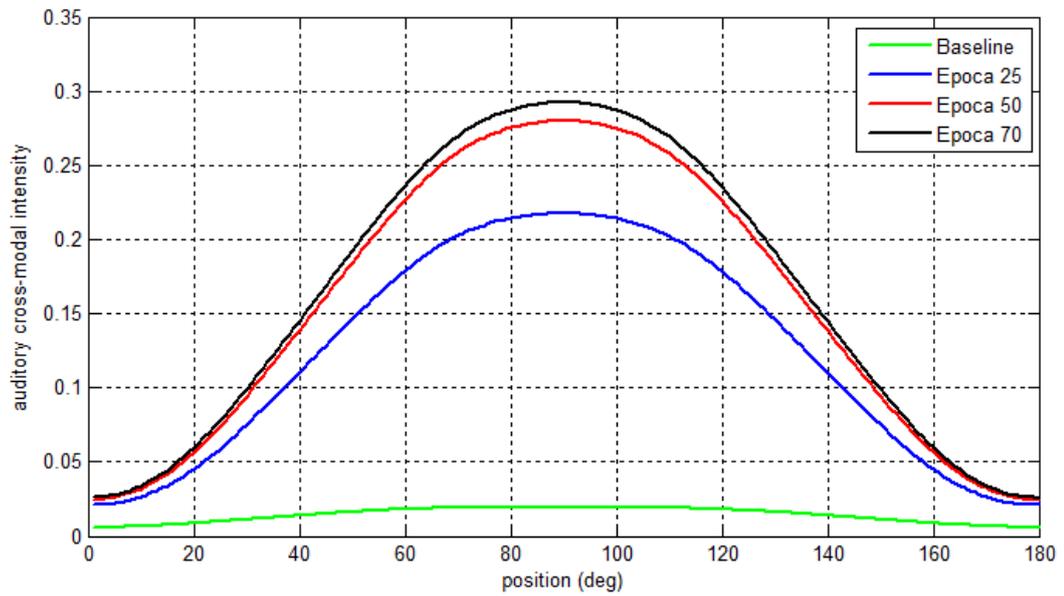


Figura 4.11: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

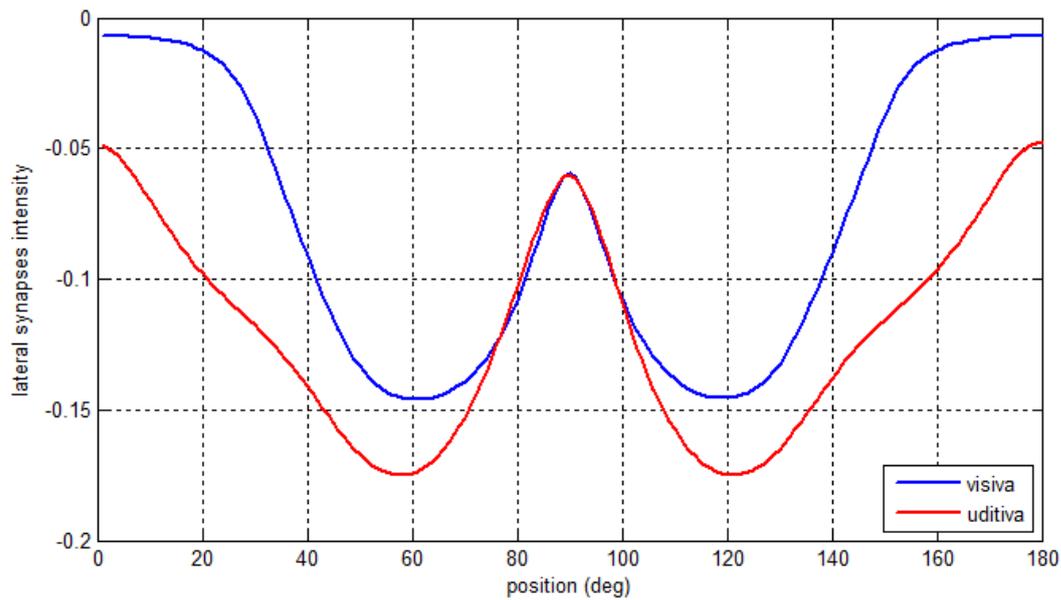


Figura 4.12: Risultato finale dell'addestramento delle sinapsi laterali visive (blu) e acustiche (rosso), relative al neurone in posizione 90

Dai grafici appena presentati si notano dei risultati simili alla precedente simulazione: le sinapsi laterali rimangono fortemente inibitorie, le sinapsi cross-modali sono molto larghe e il campo recettore visivo si allarga e si abbassa in maniera anomala già a partire dalla 30esima epoca.

In questo addestramento non è stata effettuata l'analisi del ventriloquismo, ma alla luce della simulazione eseguita nella prova precedente e dai grafici poco confortanti ottenuti anche in questa, si può prevedere che i risultati sul ventriloquismo saranno pessimi.

Dagli scarsi risultati emersi in queste due prove si può così concludere che il comportamento della rete con auto-addestramento delle sinapsi laterali risulta inadeguato.

### 4.3 Simulazione con sinapsi laterali diverse

Questo addestramento è stato eseguito nelle stesse circostanze già esaminate in precedenza nel caso 2, ovvero con sinapsi laterali visive e acustiche diverse tra loro e mantenute costanti durante la simulazione. Non avendo considerato la plasticità delle sinapsi laterali, l'unica variante introdotta in questo caso nel modello è la progressiva riduzione del fattore di apprendimento delle sinapsi cross-modali  $\gamma_W$ .

La rete è stata addestrata per un numero complessivo di 100 epoche, analizzando lo sviluppo dei campi recettori (Figure 4.13-4.14) e delle sinapsi cross-modali (Figure 4.15-4.16). La Tabella 4.4 riporta i parametri utilizzati in questa prova, mentre i parametri in condizioni basali sono quelli assegnati in Tabella 2.1.

<b>External stimuli</b>			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 60$	$\sigma^a = 30$
<b>Visual lateral synapses</b>			
$L_{ex,0} = 2$	$\sigma_{ex} = 3 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 24 \text{ deg}$
<b>Auditory lateral synapses</b>			
$L_{ex,0} = 2$	$\sigma_{ex} = 20 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 50 \text{ deg}$
<b>Hebbian rules</b>			
$\gamma_{RV} = 0.002$		$\gamma_{W0} = 0.0006$	
<b>Threshold</b>			
$W_{max} = 20$			

Tabella 4.4: Parametri relativi alla simulazione con sinapsi laterali diverse

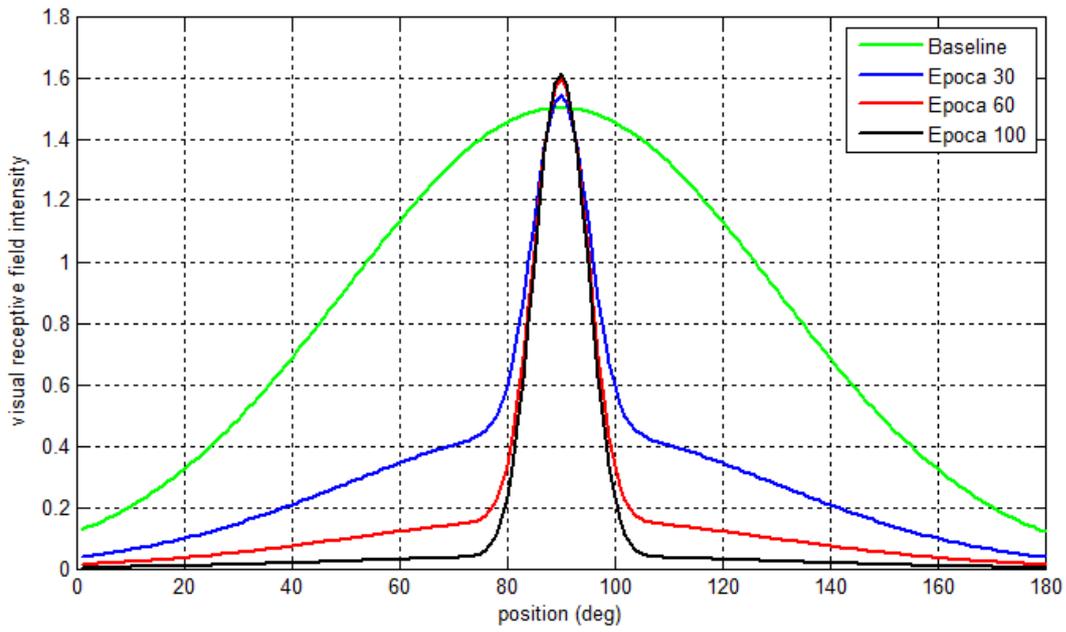


Figura 4.13: Addestramento del campo recettore visivo relativo al neurone in posizione 90

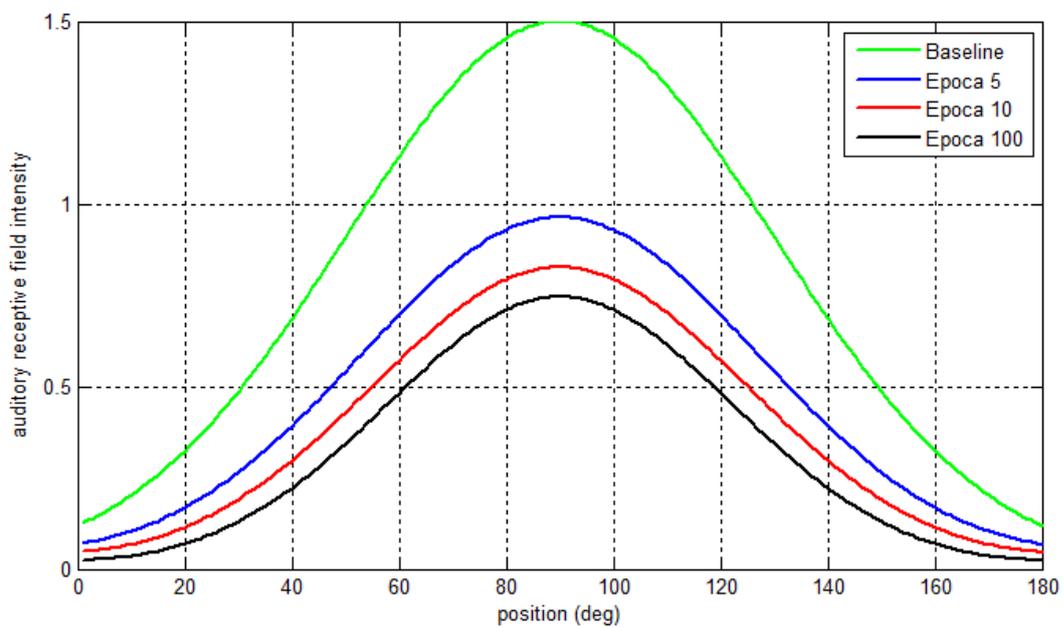


Figura 4.14: Addestramento del campo recettore acustico relativo al neurone in posizione 90

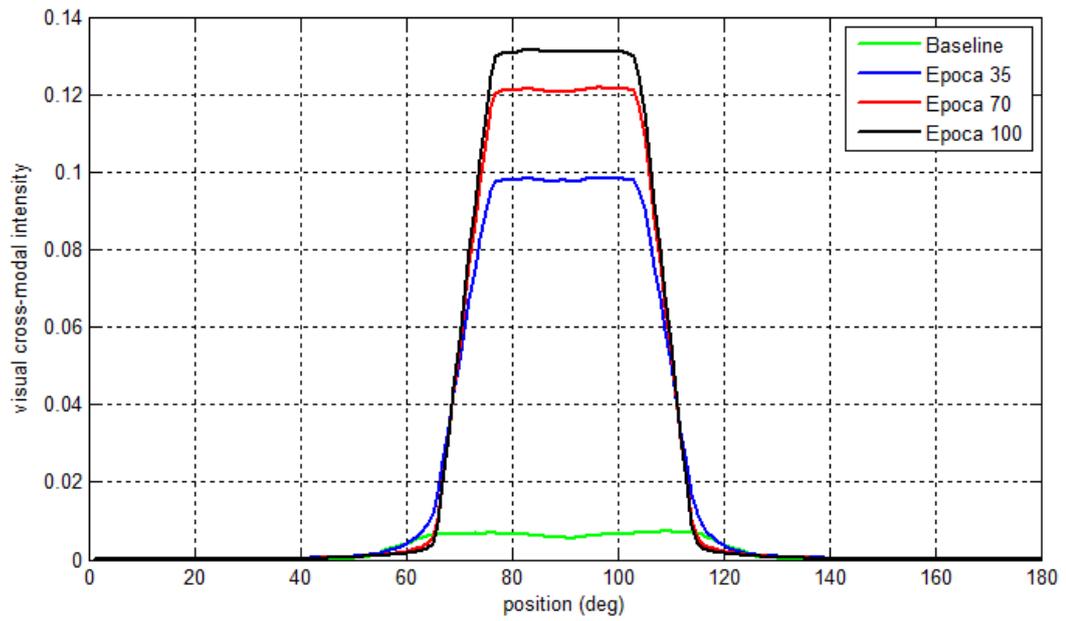


Figura 4.15: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

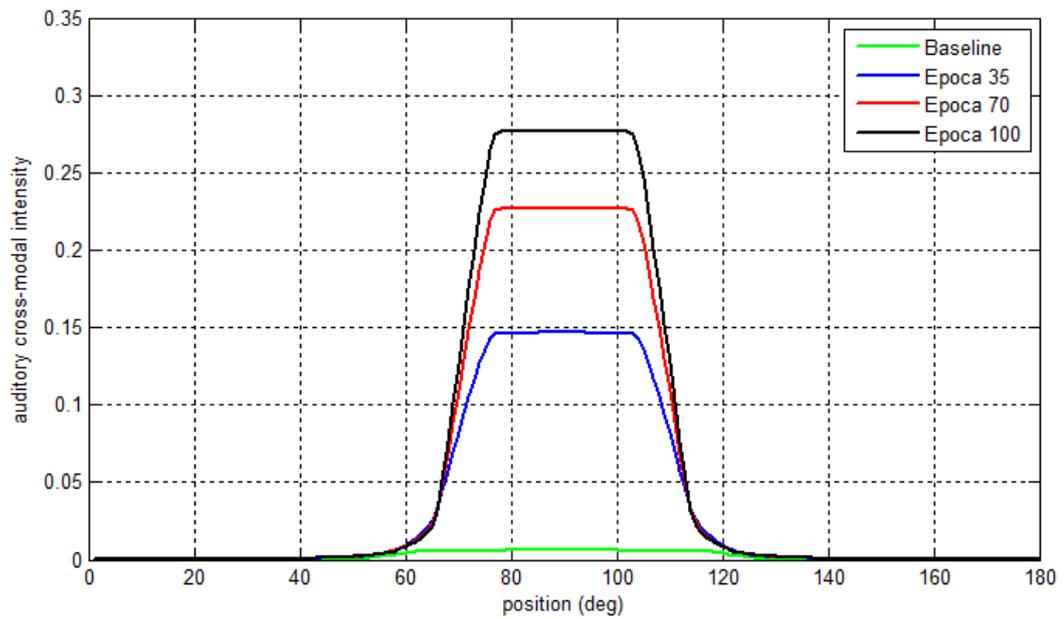


Figura 4.16: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

Da questa simulazione emerge un ottimo sviluppo dei campi recettori visivi e acustici e delle sinapsi cross-modali, notando come quest'ultime si mantengono molto strette e come il fattore di apprendimento decresca nel tempo.

Il miglioramento fisiologico appena osservato non è però seguito dai risultati scaturiti dalla simulazione del ventriloquismo, come si può notare in Figura 4.17.

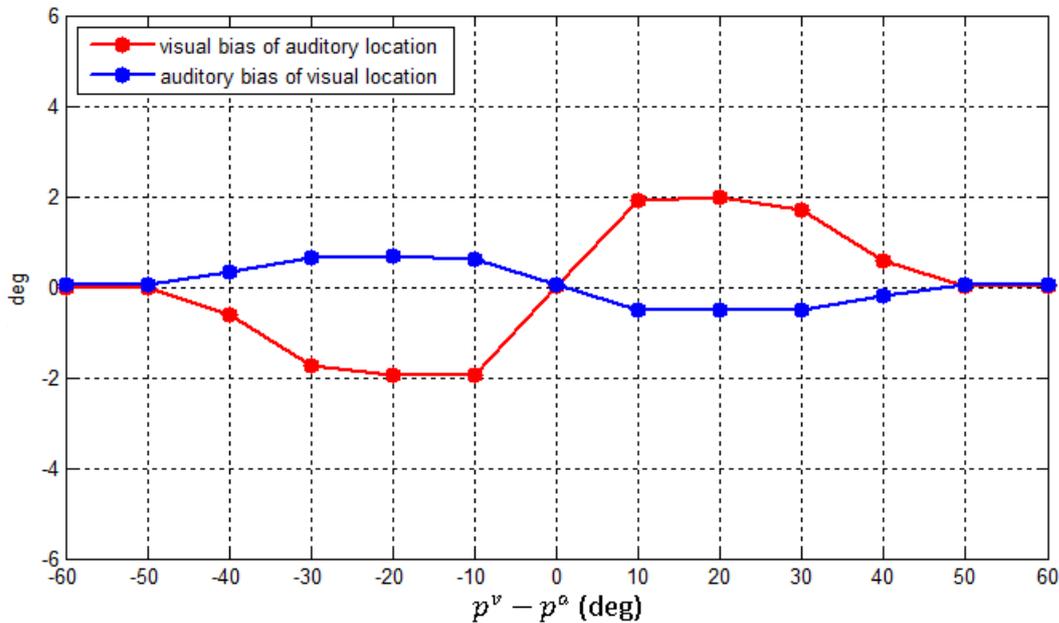


Figura 4.17: Shift di localizzazione dello stimolo acustico (in rosso) e dello stimolo visivo (in blu) in funzione del loro angolo di separazione.

La Figura 4.17 è stata ottenuta mantenendo lo stimolo visivo in una posizione fissa ( $p^v = 90^\circ$ ) e variando la posizione dello stimolo acustico ( $p^a$ ) da  $30^\circ$  a  $150^\circ$ . Da questa Figura si nota una scarsa influenza acustica sulla modalità visiva, correttamente da quanto riportato dai dati sperimentali, ma anche una scarsa influenza visiva sulla modalità acustica, contrariamente alle aspettative.

Pertanto anche il comportamento della rete addestrata con sinapsi laterali visive e uditive fisse e diverse tra loro risulta inadeguato.

## 4.4 Simulazione con sinapsi laterali uguali

In questa simulazione la rete è stata addestrata in condizioni simili al caso 1, ovvero con sinapsi laterali visive e acustiche fisse e uguali tra loro. In particolare sono state introdotte le stesse sinapsi laterali utilizzate nella seconda simulazione (Paragrafo 3.1.2, Figura 3.6), molto larghe con una debole eccitazione dei neuroni prossimali e una buona inibizione di quelli distali. Anche in questa prova l'unica variante introdotta è la progressiva riduzione del fattore di apprendimento delle sinapsi cross-modali  $\gamma_W$ .

L'addestramento è stato eseguito per un numero complessivo di 100 epoche, visualizzando il relativo sviluppo dei campi recettori (Figure 4.18-4.19) e delle sinapsi cross-modali (Figure 4.20-4.21). La Tabella sottostante riporta i parametri utilizzati durante la simulazione, mentre i parametri in condizioni basali fanno riferimento alla Tabella 2.1.

<b>External stimuli</b>			
$E_0^v = 20$	$\sigma^v = 4$	$E_0^a = 45$	$\sigma^a = 30$
<b>Lateral synapses</b>			
$L_{ex,0} = 2$	$\sigma_{ex} = 20 \text{ deg}$	$L_{in,0} = 1.8$	$\sigma_{in} = 50 \text{ deg}$
<b>Hebbian rules</b>			
$\gamma_{RV} = 0.002$		$\gamma_{W0} = 0.0005$	
<b>Threshold</b>			
$W_{max} = 8$			

Tabella 4.5: Parametri relativi alla simulazione con sinapsi laterali uguali

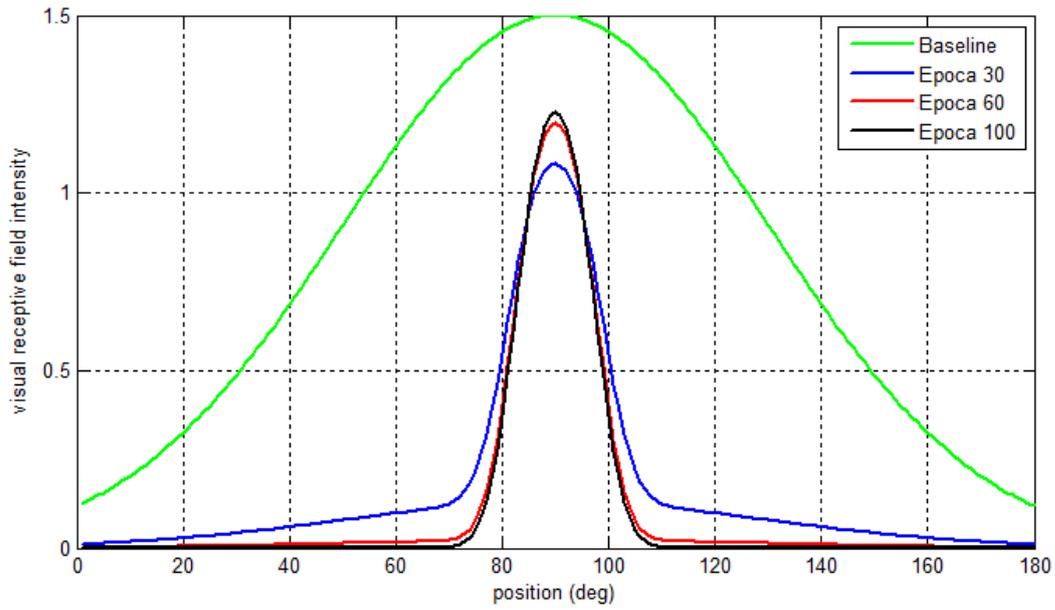


Figura 4.18: Addestramento del campo recettore visivo relativo al neurone in posizione 90

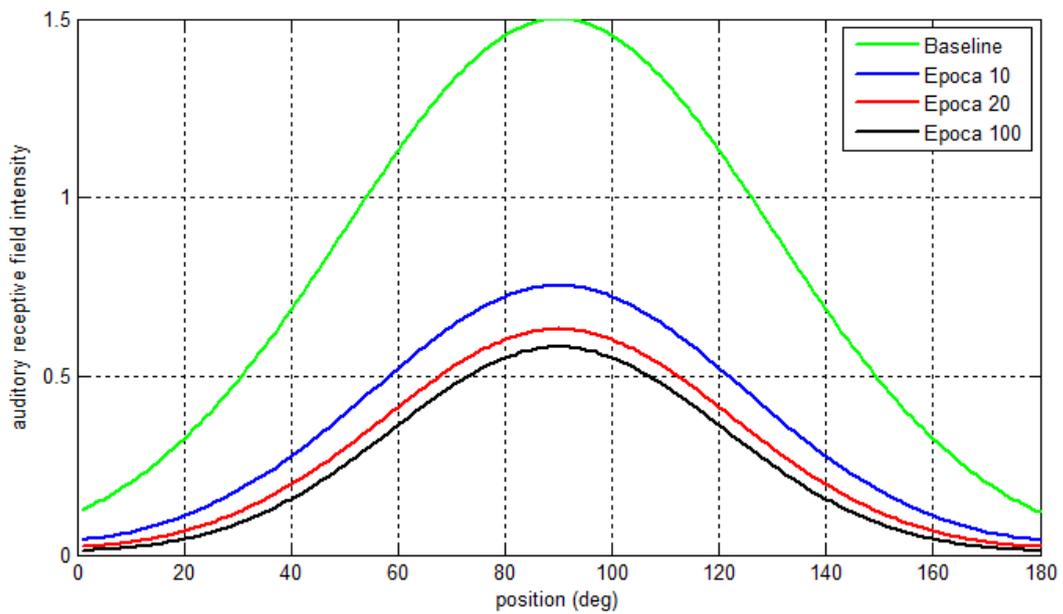


Figura 4.19: Addestramento del campo recettore visivo relativo al neurone in posizione 90

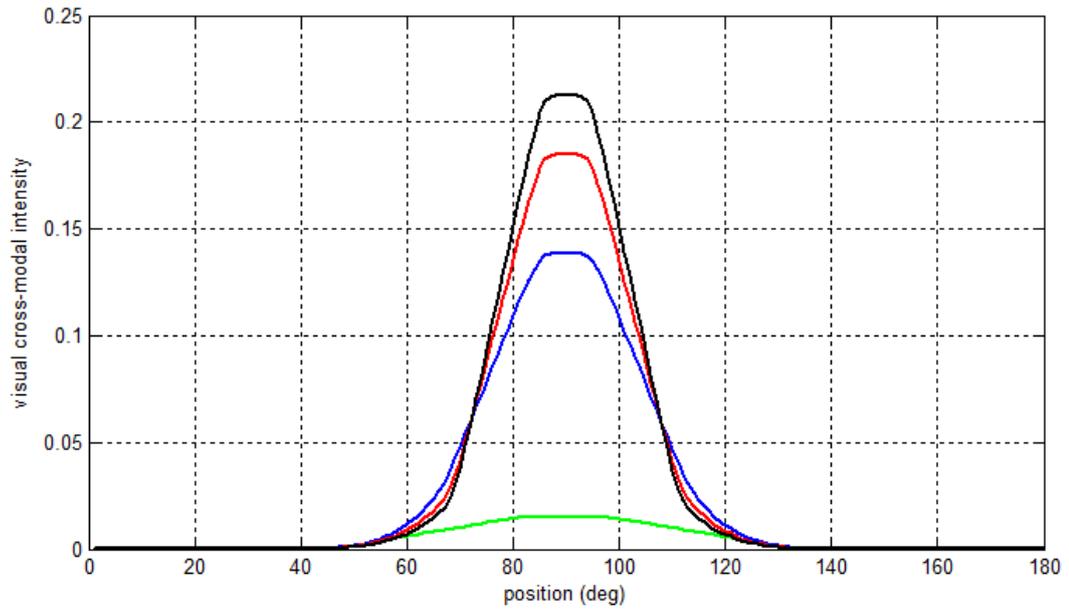


Figura 4.20: Addestramento delle sinapsi cross-modali visive relative al neurone in posizione 90

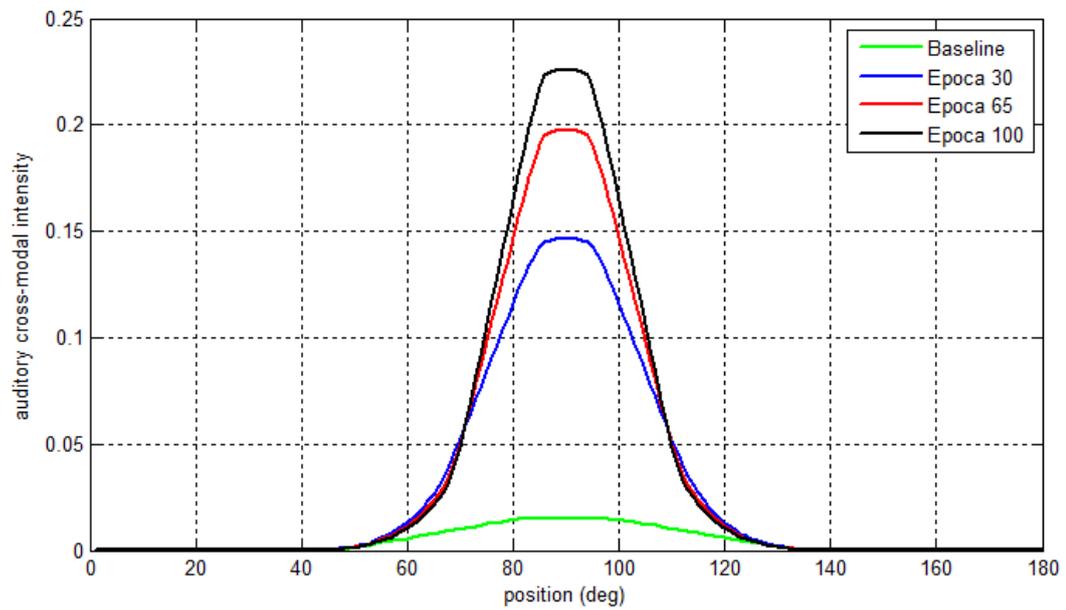


Figura 4.21: Addestramento delle sinapsi cross-modali acustiche relative al neurone in posizione 90

Anche in questa prova si registra un ottimo addestramento dei campi recettori visivi e acustici. Le sinapsi laterali si mantengono molto strette e si addestrano coerentemente alla nuova variante matematica introdotta, notando come tendano a saturare nel tempo. A differenza della simulazione con sinapsi laterali diverse, in questa si osserva anche un'ottima riproduzione del ventriloquismo spaziale, come mostrato in Figura 4.22.

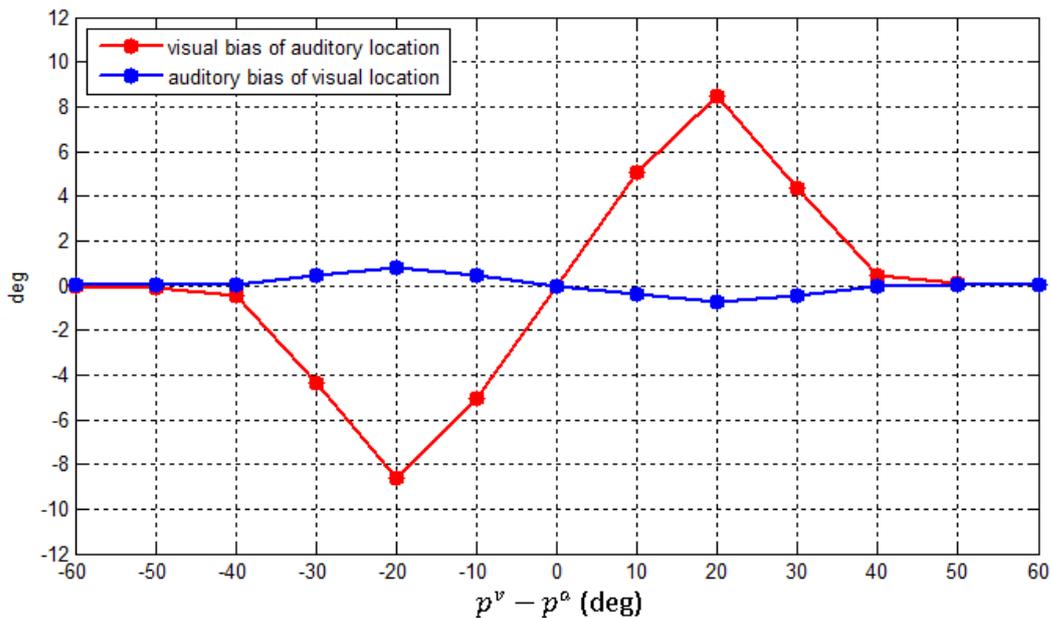


Figura 4.22: Shift di localizzazione dello stimolo acustico (rosso) e dello stimolo visivo (blu) in funzione del loro angolo di separazione.

La Figura 4.22 è stata ottenuta mantenendo lo stimolo visivo in una posizione fissata ( $p^v = 90^\circ$ ) e variando la posizione dello stimolo acustico ( $p^a$ ) da  $30^\circ$  a  $150^\circ$ . Dal grafico si nota, oltre ad una scarsa influenza acustica sulla localizzazione dello stimolo visivo, un'elevata influenza visiva sulla localizzazione dello stimolo acustico, soprattutto quando la distanza tra gli stimoli applicati è compresa tra  $10^\circ$  e  $30^\circ$ . In particolare, se i due eventi accadono a grande distanza tra loro (superiore a  $40^\circ$ ) le attività neurali si comportano come aree individuali senza un'interazione apprezzabile, mentre se avvengono a una

distanza modesta lo stimolo acustico si sposta in direzione di quello visivo (ventriloquismo), mostrando uno shift molto alto. Al contrario, la percezione dello stimolo visivo mostra generalmente uno shift molto basso in direzione dello stimolo acustico. Si può notare come il grafico in Figura 4.22 sia praticamente simile a quello ottenuto dal modello di Magosso, Ursino e Cuppini (Paragrafo 1.4, Figura 1.4).

## **4.5 Enhancement audio-visivo**

Come già enunciato nel Paragrafo 1.2.2, l'enhancement è un altro effetto dell'integrazione audio-visiva nella corteccia primaria in cui stimoli cross-modali, temporaneamente e spazialmente coincidenti, evocano risposte neurali maggiori rispetto a quelle generate dalle singole componenti unisensoriali.

Questo rafforzamento è stato testato in un'istanza successiva all'addestramento della rete con sinapsi laterali uguali, mostrato nel Paragrafo 4.4. A questo proposito sono state effettuate due prove, in cui vengono poste a confronto le rispettive risposte neurali: nella prima è stato applicato nel modello un solo input unimodale acustico della durata di 50 millisecondi (Figura 4.23), nella seconda sono stati invece applicati uno stimolo visivo e uno acustico spazialmente e temporaneamente coincidenti, anche essi della durata di 50 millisecondi (Figura 4.24).

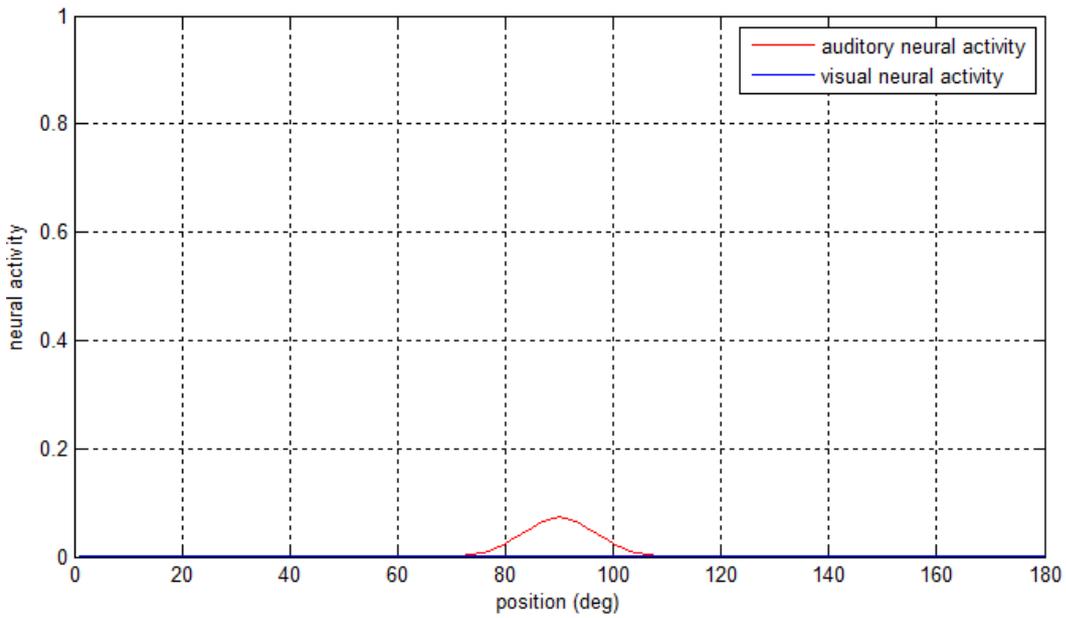


Figura 4.23: Risposta neurale in seguito all'applicazione di uno stimolo unimodale acustico in posizione 90

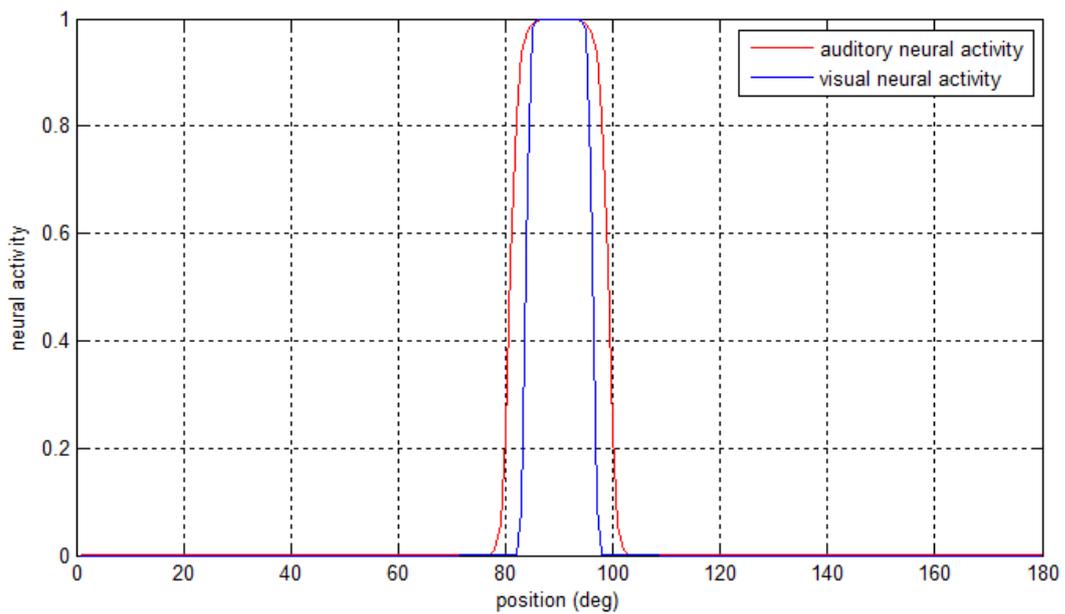


Figura 4.24: Risposta neurale in seguito all'applicazione di uno stimolo visivo e uno acustico in posizione 90

Nella figura 4.23 è stata analizzata l'attività neurale in seguito all'applicazione, alla rete già addestrata, di uno stimolo acustico di bassa intensità ( $E_0^a = 35$ ) in posizione 90; in questa Figura si apprezza una lieve attività neurale uditiva (in rosso), mentre non si segnala alcuna attività neurale visiva. La Figura 4.24 invece monitora la risposta neurale visiva (in blu) e uditiva (in rosso) in seguito all'applicazione di un input acustico della stessa intensità di quella utilizzata nella precedente prova e di un input visivo di intensità  $E_0^v = 15$ , temporaneamente coincidenti e centrati in posizione 90.

Dalla Figura 4.24 si registra un rinforzamento dell'attività neurale acustica in presenza di un input visivo contemporaneo e spazialmente coincidente, mostrando una risposta più forte rispetto a quella ottenuta con l'applicazione di uno stimolo acustico singolo della stessa intensità (Figura 4.23).



# Capitolo 5

## Conclusioni

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'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 unimodali e 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). Questo processo richiede diversi mesi di esperienze cross-modali prima di raggiungere la condizione di un adulto.

Il modello matematico proposto nella presente Tesi ha contribuito a chiarificare i meccanismi neurali alla base dell'integrazione audio-visiva nelle cortecce primarie, in particolare come la rete neurale si sviluppa in un ambiente cross-modale. Quindi, una volta addestrata la rete, il modello è stato utilizzato in un'istanza successiva per simulare gli effetti dell'integrazione audio-visiva nella corteccia primaria, quali il ventriloquismo spaziale e l'enhancement audio-visivo.

Il modello è stato addestrato con sequenze di input visivi, acustici e audio-visivi in diverse condizioni fisiologiche. La rete infatti è stata simulata in tre casi diversi, ovvero con sinapsi laterali visive e acustiche uguali, con sinapsi laterali diverse e con auto addestramento delle sinapsi laterali. In tutti e tre i casi si è ottenuto un ottimo addestramento dei campi recettori e delle attività cross-modali. Quindi è stata effettuata la simulazione del ventriloquismo in cui si è notata un'ottima

riproduzione soprattutto nella rete con sinapsi laterali uguali, mentre nella rete con sinapsi laterali diverse il ventriloquismo è molto debole. Anche nel caso con auto-addestramento delle sinapsi si è ottenuta un'ottima riproduzione del ventriloquismo, anche se le sinapsi laterali tendono a rimanere molto deboli. In tutti e tre i casi simulati si nota però che le sinapsi cross-modali continuerebbero a crescere progressivamente all'infinito, e ciò non è fisiologicamente corretto. Inoltre, per quanto riguarda l'auto-addestramento delle sinapsi laterali, come già detto, tali sinapsi tendono a rimanere molto deboli e poco facilmente addestrabili nel caso in cui si usino basse costanti di apprendimento, mentre l'uso di costanti di apprendimento più alte induce instabilità e, molto spesso, un'inibizione eccessivamente forte.

Nelle successive simulazioni del modello sono state introdotte due varianti. In primo luogo si è imposto un valore massimo di saturazione fisiologica in modo tale che la costante di apprendimento delle sinapsi cross-modali potesse progressivamente decrescere ad ogni epoca. È stato utilizzato un approccio simile nel caso dell'auto-addestramento delle sinapsi laterali al fine di ridurre progressivamente i fattori di apprendimento eccitatori e inibitori. In quest'ultimo caso però i risultati ottenuti sono stati deludenti; infatti si è riscontrata una forte instabilità dei campi recettori visivi, le sinapsi laterali tendono a rimanere molto inibitorie e nella simulazione del ventriloquismo si nota una forte influenza acustica sulla modalità visiva. Nel caso di sinapsi laterali fisse e diverse tra loro l'addestramento dei campi recettori e delle sinapsi cross-modali è ottimo, ma nel ventriloquismo si nota una scarsa influenza visiva sulla modalità acustica. Questo limite è stato superato nella simulazione del modello con sinapsi laterali uguali tra loro, in cui, oltre a un ottimo addestramento dei campi recettori e delle sinapsi cross-modali, è stata ottenuta un'ottima riproduzione del ventriloquismo, con risultati simili al modello studiato da Magosso, Ursino e Cuppini.

Quest'ultima simulazione 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. In più, tali sinapsi tendono a stabilizzarsi su un valore massimo di saturazione fisiologica per un numero di epoche molto alto.
- La rete addestrata si comporta secondo le previsioni, riproducendo alcuni effetti dell'integrazione audio-visiva nella corteccia primaria, tra cui il ventriloquismo spaziale e l'enhancement

Secondo quanto detto fino ad ora, il modello simulato nell'ultima prova (Paragrafo 4.4) con sinapsi laterali uguali, larghe, con un'importante inibizione dei neuroni distali e una bassa eccitazione di quelli prossimali (Figura 3.6) può contribuire matematicamente a comprendere i meccanismi neurali fisiologici alla base dell'apprendimento audio-visivo in seguito a riproduzioni di esperienze cross-modali e unisensoriali. La rete inoltre potrebbe essere utilizzata più approfonditamente in futuro per simulare non solo il ventriloquismo a livello spaziale e l'enhancement, ma anche tutti gli altri effetti dell'integrazione audio-visiva nella corteccia primaria citati nel Capitolo 1, quali l'After-Effect del ventriloquismo (offline), il ventriloquismo temporale e le illusioni audio-visive ("fission" e "fusion").



# Appendici

## Appendice A

Codice matlab utilizzato nell'addestramento della rete nel caso 1 (sinapsi laterali fisse e uguali tra loro)

```
clear
clc
close all
% x acustico; y visivo

N=180; %numero di neuroni

% sinapsi laterali
sigex=20; %std delle sinapsi laterali eccitatorie
sigin=50; %std delle sinapsi laterali inibitorie
Lex0_y=2.0; %forza delle sinapsi laterali eccitatorie
visive
Lin0_y=1.8; %forza delle sinapsi laterali inibitorie
visive
Lex0_x=2.0; %forza delle sinapsi laterali eccitatorie
uditive
Lin0_x=1.8; %forza delle sinapsi laterali inibitorie
uditive

%caratteristica dei neuroni
phi=16; %posizione centrale della funzione sigmoidale
pend=2; %pendenza della funzione sigmoidale
tau=5; %costante di tempo

% input
E0y=20; % forza input visivo
sy=4; % std input visivo
E0x=45; % forza input uditivo
sx=30; % std input uditivo

k=[1:1:180];

%genero le sinapsi laterali
for i = 1:N,
    distanza=abs(k-i);
    controllo_dist=double((abs(k-i)>(N/2)));

    % matrice delle distanze al quadrato FORMA CIRCOLARE
    D2=((distanza-sign(controllo_dist)*N).^2);
    % sinapsi laterali visive
```

## Appendici

```
    Ly(i,:)=Lex0_y*exp(-D2/2/sigex/sigex)-Lin0_y*exp(-
D2/2/sigin/sigin);
    Ly(i,i)=0;      %non riceve sinapsi da se stesso
    % sinapsi laterali uditive
    Lx(i,:)=Lex0_x*exp(-D2/2/sigex/sigex)-Lin0_x*exp(-
D2/2/sigin/sigin);
    Lx(i,i)=0;      %non riceve sinapsi da se stesso

end

% genero i campi recettori visivi e uditivi
RV = zeros(N,N);
RV0 = 1.5;          % forza dei campi recettori in condizioni
basali
sig_v = 40;        % std dei campi recettori in condizioni
basali
j=[1:1:N];
for l = 1:N,
    distanza=abs(j-1);
    controllo_dist=double((abs(j-1)>(N/2)));

    % matrice delle distanze al quadrato FORMA CIRCOLARE
    D2=((distanza-sign(controllo_dist)*N).^2);
    RV(j,l,1)=RV0*exp(-D2/2/sig_v/sig_v);

end

RVy=RV;
RVx=RV;

%Grafico un input visivo e uno uditivo nella stessa posizione
prima dell'addestramento
Position = 90;      %posizione input visivo e uditivo
%genero l'input visivo e uditivo in posizione 90
distanza=abs(k-Position);
controllo_dist=double((abs(k-Position)>(N/2)));
% vettore delle distanze al quadrato FORMA CIRCOLARE
D2=((distanza-sign(controllo_dist)*N).^2);
Vy=E0y*exp(-D2/2/sy/sy)/(sqrt(2*pi*sy*sy)); %input visivo
normalizzato
Vx=E0x*exp(-D2/2/sx/sx)/(sqrt(2*pi*sx*sx)); %input acustico
normalizzato
Iy0 = RVy(:, :, 1)*Vy';
Ix0 = RVx(:, :, 1)*Vx';
Y =1./(1+exp(-(Iy0-phi)*pend)); %risposta neurale
visiva
X =1./(1+exp(-(Ix0-phi)*pend)); %risposta neurale
acustica
figure(1)
subplot(2,1,1)
plot(j,Y', 'r')
subplot(2,1,2)
plot(j,RVy(80, :, 1), 'b', j, RVy(50, :, 1), 'r', j, RVy(170, :, 1), 'g')
%plotto i campi recettori visivi prima dell'addestramento
figure(2)
subplot(2,1,1)
```

```

plot(j,X','r')
subplot(2,1,2)
plot(j,RVx(80,:,1),'b',j,RVx(50,:,1),'r',j,RVx(170,:,1),'g') %
plotto i campi recettori uditivi prima dell'addestramento
pause

D_RVy = zeros(N,N);
D_RVx = D_RVy;
D_Wy = zeros(N,N);
D_Wx = D_Wy;
Wy = zeros(N,N);
Wx = Wy;
gatex(1:3:540)=1;
gatex(2:3:540)=0;
gatex(3:3:540)=1;
gatey(1:3:540)=0;
gatey(2:3:540)=1;
gatey(3:3:540)=1;
dt=0.2; % passo di campionamento
L_stimolo = 17/dt; % lunghezza dello stimolo
t=[0:L_stimolo]*dt; % asse dei tempi
LL=length(t);
for epoca = 1:100,
    epoca

P1 =randperm(180); % permutazione degli input acustici
P2 =randperm(180); % permutazione degli input visivi
P3 =randperm(180); % permutazione degli input AV
PP(1:3:540)=P1;
PP(2:3:540)=P2;
PP(3:3:540)=P3;
for kk = 1: length(PP),
    Position_y = PP(kk); %posizione input visivo
    Position_x = PP(kk); %posizione input uditivo
    %genero l'input visivo
    distanza_y=abs(k-Position_y);
    controllo_dist_y=double((abs(k-Position_y)>(N/2)));
    D2_y=((distanza_y-sign(controllo_dist_y)*N).^2);
    Vy=E0y*gatey(kk)*exp(-D2_y/2/sy/sy)/(sqrt(2*pi*sy*sy)); %input
    visivo normalizzato
    %genero l'input uditivo
    distanza_x=abs(k-Position_x);
    controllo_dist_x=double((abs(k-Position_x)>(N/2)));
    D2_x=((distanza_x-sign(controllo_dist_x)*N).^2);
    Vx=E0x*gatex(kk)*exp(-D2_x/2/sx/sx)/(sqrt(2*pi*sx*sx)); %input
    uditivo normalizzato

% attività neurale visiva e uditiva
y=zeros(N,LL);
x=y;
for kt =1:LL-1,
    y(:,kt+1) = y(:,kt) + (dt/tau)*(-y(:,kt)+1./(1+exp(-
(RVy(:, :, epoca)*Vy'+Ly*y(:,kt)+Wy(:, :, epoca)*x(:,kt)-phi)*pend)));
    x(:,kt+1) = x(:,kt) + (dt/tau)*(-x(:,kt)+1./(1+exp(-
(RVx(:, :, epoca)*Vx'+Lx*x(:,kt)+Wx(:, :, epoca)*y(:,kt)-phi)*pend)));
end
Y = y(:,LL);

```

## Appendici

```
X = x(:,LL);

if Position_y == 90
    figure(1)
    subplot(2,1,1)
    plot(j,Y', 'r');
end

if Position_x == 90
    figure(2)
    subplot(2,1,1)
    plot(j,X', 'r');
end

% aggiorno i campi recettori visivi e uditivi e le sinapsi cross-
modali (addestramento hebbiano)

gamma = 0.002;           %fattore di apprendimento dei campi
recettori
gamma1 = 0.0005;        %fattore di apprendimento dell'attività
cross-modale
Wmax = 8;               %valore massimo delle sinapsi cross-modali

    for i = 1:N,
        D_RVy(i,:) = gamma*Y(i) * (Vy-RVy(i,:,epoca));
        D_RVx(i,:) = gamma*X(i) * (Vx-RVx(i,:,epoca));
        D_Wy(i,:) = gamma1*(Wmax-
sum(Wy(i,:,epoca))/Wmax*Y(i) * (X'-Wy(i,:,epoca));
        D_Wx(i,:) = gamma1*(Wmax-
sum(Wx(i,:,epoca))/Wmax*X(i) * (Y'-Wx(i,:,epoca));
    end

RVy(:, :, epoca)=RVy(:, :, epoca)+D_RVy;
RVx(:, :, epoca)=RVx(:, :, epoca)+D_RVx;
Wy(:, :, epoca)=Wy(:, :, epoca)+D_Wy;
Wx(:, :, epoca)=Wx(:, :, epoca)+D_Wx;

end
RVy(:, :, epoca+1)=RVy(:, :, epoca);
RVx(:, :, epoca+1)=RVx(:, :, epoca);
Wy(:, :, epoca+1)=Wy(:, :, epoca);
Wx(:, :, epoca+1)=Wx(:, :, epoca);
figure(1)
subplot(2,1,2)
plot(j,RVy(80,:,epoca+1), 'b',j,RVy(50,:,epoca+1), 'r',j,RVy(170,:,e
poca+1), 'g')
figure(2)
subplot(2,1,2)
plot(j,RVx(80,:,epoca+1), 'b',j,RVx(50,:,epoca+1), 'r',j,RVx(170,:,e
poca+1), 'g')
figure(3)
subplot(2,1,1)
plot(j,Wy(90,:,epoca+1), 'r')
subplot(2,1,2)
```

```

plot(j,Wx(90, :,epoca+1), 'b')
sum(Wy(i, :,epoca+1))
sum(Wx(i, :,epoca+1))
end
grid

```

## Appendice B

Codice matlab utilizzato nell'addestramento della rete nel caso 2 (sinapsi laterali fisse e diverse tra loro).

```

clear
clc
close all
% x acustici; y visivo

N=180; %numero di neuroni

% sinapsi laterali
sigex_y=3; %std delle sinapsi laterali eccitatorie visive
sigin_y=24; %std delle sinapsi laterali inibitorie visive
sigex_x=20; %std delle sinapsi laterali eccitatorie acustiche
sigin_x=50; %std delle sinapsi laterali inibitorie acustiche
Lex0_y=2.0; %forza delle sinapsi laterali eccitatorie visive
Lin0_y=1.8; %forza delle sinapsi laterali inibitorie visive
Lex0_x=2.0; %forza delle sinapsi laterali eccitatorie uditive
Lin0_x=1.8; %forza delle sinapsi laterali inibitorie uditive

%caratteristica dei neuroni
phi=16; %posizione centrale della funzione sigmoidale
pend=2; %pendenza della funzione sigmoidale
tau=5; %costante di tempo

% input
E0y=20; % forza input visivo
sy=4; % std input visivo
E0x=60; % forza input uditivo
sx=30; % std input uditivo

k=[1:1:180];
%genero le sinapsi laterali
for i = 1:N,
    distanza=abs(k-i);
    controllo_dist=double((abs(k-i)>(N/2)));

    % matrice delle distanze al quadrato FORMA CIRCOLARE
    D2=((distanza-sign(controllo_dist)*N).^2);
    % sinapsi laterali visive

```

```

    Ly(i,:)=Lex0_y*exp(-D2/2/sigex_y/sigex_y)-Lin0_y*exp(-
D2/2/sigin_y/sigin_y);
    Ly(i,i)=0;      %non riceve sinapsi da se stesso
    % sinapsi laterali uditive
    Lx(i,:)=Lex0_x*exp(-D2/2/sigex_x/sigex_x)-Lin0_x*exp(-
D2/2/sigin_x/sigin_x);
    Lx(i,i)=0;      %non riceve sinapsi da se stesso

end

% genero campi recettori visivi e uditivi
RV = zeros(N,N);
RV0 = 1.5;
sig_v = 40;
j=[1:1:N];
for l = 1:N,
    distanza=abs(j-1);
    controllo_dist=double((abs(j-1)>(N/2)));

    % matrice delle distanze al quadrato FORMA CIRCOLARE
    D2=((distanza-sign(controllo_dist)*N).^2);
    RV(j,l,1)=RV0*exp(-D2/2/sig_v/sig_v);

end

RVy=RV;
RVx=RV;

%Grafico un input visivo e uno uditivo nella stessa posizione
prima dell'addestramento
Position = 90;      %posizione input visivo e uditivo
%genero l'input visivo e uditivo in posizione 90
distanza=abs(k-Position);
controllo_dist=double((abs(k-Position)>(N/2)));
% vettore delle distanze al quadrato FORMA CIRCOLARE
D2=((distanza-sign(controllo_dist)*N).^2);
Vy=E0y*exp(-D2/2/sy/sy)/(sqrt(2*pi*sy*sy));    %input visivo
normalizzato
Vx=E0x*exp(-D2/2/sx/sx)/(sqrt(2*pi*sx*sx));    %input uditivo
normalizzato
Iy0 = RVy*Vy';
Ix0 = RVx*Vx';
Y =1./(1+exp(-(Iy0-phi)*pend));    %risposta neurale visiva
X =1./(1+exp(-(Ix0-phi)*pend));    %risposta neurale uditiva
figure(1)
subplot(2,1,1)
plot(j,Y', 'r')
subplot(2,1,2)
plot(j,RVy(80,:,1), 'b', j,RVy(50,:,1), 'r', j,RVy(170,:,1), 'g')
%plotto le sinapsi visive prima dell'addestramento
figure(2)
subplot(2,1,1)
plot(j,X', 'r')
subplot(2,1,2)
plot(j,RVx(80,:,1), 'b', j,RVx(50,:,1), 'r', j,RVx(170,:,1), 'g')    %
plotto le sinapsi uditive prima dell'addestramento

```

## Appendici

```
pause

D_RVy = zeros(N,N);
D_RVx = D_RVy;
D_Wy = zeros(N,N);
D_Wx = D_Wy;
Wy = zeros(N,N);
Wx = Wy;
gatex(1:3:540)=1;
gatex(2:3:540)=0;
gatex(3:3:540)=1;
gatey(1:3:540)=0;
gatey(2:3:540)=1;
gatey(3:3:540)=1;
dt=0.2; % passo di campionamento
L_stimolo = 17/dt; % lunghezza dello stimolo
t=[0:L_stimolo]*dt; % asse dei tempi
LL=length(t);
for epoca = 1:100,
    epoca
        sum(Wy(90,:,epoca))
        sum(Wx(90,:,epoca))

P1 =randperm(180); %permutazione degli input uditivi
P2 =randperm(180); %permutazione degli input visivi
P3 =randperm(180); %permutazione degli input AV
PP(1:3:540)=P1;
PP(2:3:540)=P2;
PP(3:3:540)=P3;
for kk = 1: length(PP),
    Position_y = PP(kk); %posizione input visivo
    Position_x = PP(kk); %posizione input uditivo
    %genero l'input visivo
    distanza_y=abs(k-Position_y);
    controllo_dist_y=double((abs(k-Position_y)>(N/2)));
    D2_y=((distanza_y-sign(controllo_dist_y)*N).^2);
    Vy=E0y*gatey(kk)*exp(-D2_y/2/sy/sy)/(sqrt(2*pi*sy*sy)); %in
    questo modo l'area vale esattamente
    %genero l'input uditivo
    distanza_x=abs(k-Position_x);
    controllo_dist_x=double((abs(k-Position_x)>(N/2)));
    D2_x=((distanza_x-sign(controllo_dist_x)*N).^2);
    Vx=E0x*gatex(kk)*exp(-D2_x/2/sx/sx)/(sqrt(2*pi*sx*sx)); %in
    questo modo l'area vale esattamente E0x

%attività neurale visiva e acustica
y=zeros(N,LL);
x=y;
for kt =1:LL-1,
    y(:,kt+1) = y(:,kt) + (dt/tau)*(-y(:,kt)+1./(1+exp(-
(RVy(:, :, epoca)*Vy'+Ly*y(:,kt)+Wy(:, :, epoca)*x(:,kt)-phi)*pend)));
    x(:,kt+1) = x(:,kt) + (dt/tau)*(-x(:,kt)+1./(1+exp(-
(RVx(:, :, epoca)*Vx'+Lx*x(:,kt)+Wx(:, :, epoca)*y(:,kt)-phi)*pend)));
end
Y = y(:,LL);
X = x(:,LL);
```

```

if Position_y == 90
    figure(1)
    subplot(2,1,1)
    plot(j,Y', 'r');
end

if Position_x == 90
    figure(2)
    subplot(2,1,1)
    plot(j,X', 'r');
end

% aggiorno i campi recettori visivi e uditivi e le sinapsi cross-
modali (addestramento hebbiano)

gamma = 0.002;           %fattore di apprendimento dei campi
recettori
gamma1 = 0.0006;        %fattore di apprendimento delle sinapsi
cross-modali
Wmax = 20;              %valore massimo delle sinapsi cross-modali

    for i = 1:N,
        D_RVy(i,:) = gamma*Y(i)*(Vy-RVy(i,:,epoca));
        D_RVx(i,:) = gamma*X(i)*(Vx-RVx(i,:,epoca));
        D_Wy(i,:) = gamma1*(Wmax-
sum(Wy(i,:,epoca))/Wmax*Y(i)*(X'-Wy(i,:,epoca)));
        D_Wx(i,:) = gamma1*(Wmax-
sum(Wx(i,:,epoca))/Wmax*X(i)*(Y'-Wx(i,:,epoca)));
    end

RVy(:, :, epoca)=RVy(:, :, epoca)+D_RVy;
RVx(:, :, epoca)=RVx(:, :, epoca)+D_RVx;
Wy(:, :, epoca)=Wy(:, :, epoca)+D_Wy;
Wx(:, :, epoca)=Wx(:, :, epoca)+D_Wx;

end
RVy(:, :, epoca+1)=RVy(:, :, epoca);
RVx(:, :, epoca+1)=RVx(:, :, epoca);
Wy(:, :, epoca+1)=Wy(:, :, epoca);
Wx(:, :, epoca+1)=Wx(:, :, epoca);
figure(1)
subplot(2,1,2)
plot(j,RVy(80,:,epoca+1), 'b', j,RVy(50,:,epoca+1), 'r', j,RVy(170,:,e
poca+1), 'g')
figure(2)
subplot(2,1,2)
plot(j,RVx(80,:,epoca+1), 'b', j,RVx(50,:,epoca+1), 'r', j,RVx(170,:,e
poca+1), 'g')
figure(3)
subplot(2,1,1)
plot(j,Wy(90,:,epoca+1), 'r')
hold on
subplot(2,1,2)
plot(j,Wx(90,:,epoca+1), 'b')

```

## Appendici

```
hold on
sum(Wy(i, :, epoca+1))
sum(Wx(i, :, epoca+1))
end
grid
```

# Appendice C

Codice matlab utilizzato nell'addestramento della rete nel caso 3 (auto-addestramento delle sinapsi laterali).

```

clear
clc
close all
% x acustico; y visivo

N=180;                %numero di neuroni

%caratteristica dei neuroni
phi=16;              %posizione centrale della funzione sigmoidale
pend=0.5;           %pendenza della funzione sigmoidale
tau=5;              %costante di tempo

% input
E0y=30;              % forza input visivo
sy=4;                % std input visivo
E0x=60;              % forza input uditivo
sx=30;                % std input uditivo

k=[1:1:180];
%genero le sinapsi laterali

Lx=zeros(180,180);
Lexx=zeros(180,180);
Linx=zeros(180,180);
Ly=zeros(180,180);
Lexy=zeros(180,180);
Liny=zeros(180,180);
DLexx=zeros(180,180);
DLinx=zeros(180,180);
DLexy=zeros(180,180);
DLiny=zeros(180,180);

% genero campi recettori visivi e uditivi
RV = zeros(N,N);
RV0 = 0.7;          %forza del campi recettori in condizioni basali
sig_v = 40;        %std dei campi recettori in condizioni basali
j=[1:1:N];
for l = 1:N,
    distanza=abs(j-1);
    controllo_dist=double((abs(j-1)>(N/2)));

    % matrice delle distanze al quadrato FORMA CIRCOLARE
    D2=((distanza-sign(controllo_dist)*N).^2);
    RV(j,l,1)=RV0*exp(-D2/2/sig_v/sig_v);
end

```

## Appendici

```
RVy=RV;
RVx=RV;

%Grafico un input visivo e uno uditivo nella stessa posizione
prima dell'addestramento
Position = 90;           %posizione input visivo e uditivo
%genero l'input visivo e uditivo in posizione 90
distanza=abs(k-Position);
controllo_dist=double((abs(k-Position)>(N/2)));
% vettore delle distanze al quadrato FORMA CIRCOLARE
D2=((distanza-sign(controllo_dist)*N).^2);
Vy=E0y*exp(-D2/2/sy/sy)/(sqrt(2*pi*sy*sy));   %input visivo
normalizzato
Vx=E0x*exp(-D2/2/sx/sx)/(sqrt(2*pi*sx*sx));   %input acustico
normalizzato
Iy0 = RVy*Vy';
Ix0 = RVx*Vx';
Y =1./(1+exp(-(Iy0-phi)*pend));           %attività neurale visiva
X =1./(1+exp(-(Ix0-phi)*pend));           %attività neurale acustica

D_RVy = zeros(N,N);
D_RVx = D_RVy;
D_Wy = zeros(N,N);
D_Wx = D_Wy;
Wy = zeros(N,N);
Wx = Wy;
gatex(1:3:540)=1;
gatex(2:3:540)=0;
gatex(3:3:540)=1;
gatey(1:3:540)=0;
gatey(2:3:540)=1;
gatey(3:3:540)=1;
dt=0.2;           % passo di campionamento
L_stimolo = 17/dt; % lunghezza dello stimolo
t=[0:L_stimolo]*dt; % asse dei tempi
LL=length(t);
for epoca = 1:70,
    epoca

P1 =randperm(180); %permutazione degli input acustici
P2 =randperm(180); %permutazione degli input visivi
P3 =randperm(180); %permutazione degli input AV
PP(1:3:540)=P1;
PP(2:3:540)=P2;
PP(3:3:540)=P3;
for kk = 1: length(PP),
    Position_y = PP(kk); %posizione input visivo
    Position_x = PP(kk); %posizione input uditivo
%genero l'input visivo
distanza_y=abs(k-Position_y);
controllo_dist_y=double((abs(k-Position_y)>(N/2)));
D2_y=((distanza_y-sign(controllo_dist_y)*N).^2);
Vy=E0y*gatey(kk)*exp(-D2_y/2/sy/sy)/(sqrt(2*pi*sy*sy)); %input
visivo normalizzato
%genero l'input uditivo
distanza_x=abs(k-Position_x);
```

```

controllo_dist_x=double((abs(k-Position_x)>(N/2)));
D2_x=((distanza_x-sign(controllo_dist_x)*N).^2);
Vx=E0x*gatex(kk)*exp(-D2_x/2/sx/sx)/(sqrt(2*pi*sx*sx)); %input
acustico normalizzato

%attività neurale visiva e acustica
y=zeros(N,LL);
x=y;
for kt =1:LL-1,
    y(:,kt+1) = y(:,kt) + (dt/tau)*(-y(:,kt)+1./(1+exp(-
(RVy(:, :,epoca)*Vy'+Ly(:, :,epoca)*y(:,kt)+Wy(:, :,epoca)*x(:,kt)-
phi)*pend)));
    x(:,kt+1) = x(:,kt) + (dt/tau)*(-x(:,kt)+1./(1+exp(-
(RVx(:, :,epoca)*Vx'+Lx(:, :,epoca)*x(:,kt)+Wx(:, :,epoca)*y(:,kt)-
phi)*pend)));
end
Y = y(:,LL);
X = x(:,LL);

if (Position_y == 90) && (gatex(kk)==1)&&(gatey(kk)==1)
    figure(1)
    subplot(2,1,1)
    plot(j,Y', 'b');
    subplot(2,1,2)
    plot(j,X', 'r');
end

% aggiorno i campi recettori visivi e uditivi, le sinapsi cross-
modali e le sinapsi laterali (addestramento hebbiano)

gamma = 0.002; %fattore di apprendimento dei campi
recettori
gamma1 = 0.0004; %fattore di apprendimento delle sinapsi
cross-modali
gammaL=0.001; %fattore di apprendimento delle sinapsi
laterali
Lex_max = 100; %valore massimo delle sinapsi laterali
eccitatorie
Lin_max = 225; %valore massimo delle sinapsi laterali
inibitorie
Wmax=30.0; %valore massimo delle sinapsi cross-modali
tetaL=0.85; %soglia dell'attività post sinaptica

Ypos=(Y-tetaL+abs(Y-tetaL))/2/(1-tetaL); %normalizzato fra 0 e
1
Xpos = (X-tetaL+abs(X-tetaL))/2/(1-tetaL); %normalizzato fra 0 e
1

for i = 1:N,
    D_RVy(i, :) = gamma*Y(i) * (Vy-RVy(i, :,epoca));
    D_RVx(i, :) = gamma*X(i) * (Vx-RVx(i, :,epoca));

```

## Appendici

```
D_Wy(i,:) = gamma1*(Wmax-
sum(Wy(i,:,epoca))/Wmax*Y(i)*(X'-Wy(i,:,epoca)));
D_Wx(i,:) = gamma1*(Wmax-
sum(Wx(i,:,epoca))/Wmax*X(i)*(Y'-Wx(i,:,epoca)));
DLexx(i,:) = gammaL*(Lex_max-
sum(Lexx(i,:,epoca))/Lex_max*Xpos(i)*(Xpos -Lexx(i,:,epoca)));
DLinx(i,:) = gammaL*(Lin_max-
sum(Linx(i,:,epoca))/Lin_max*Xpos(i)*(X'*0.8-Linx(i,:,epoca)));
DLexy(i,:) = gammaL*(Lex_max-
sum(Lexy(i,:,epoca))/Lex_max*Ypos(i)*(Ypos -Lexy(i,:,epoca)));
DLiny(i,:) = gammaL*(Lin_max-
sum(Liny(i,:,epoca))/Lin_max*Ypos(i)*(Y'*0.8-Liny(i,:,epoca)));
%equivale a un fattore di apprendimento più basso, ma stesso
fattore di dimenticanza
end

RVy(:, :, epoca)=RVy(:, :, epoca)+D_RVy;
RVx(:, :, epoca)=RVx(:, :, epoca)+D_RVx;
Wy(:, :, epoca)=Wy(:, :, epoca)+D_Wy;
Wx(:, :, epoca)=Wx(:, :, epoca)+D_Wx;
Lexx(:, :, epoca) = Lexx(:, :, epoca) + DLexx;
Linx(:, :, epoca) = Linx(:, :, epoca) + DLinx;
Lx(:, :, epoca) = Lexx(:, :, epoca) - Linx(:, :, epoca);
Lexy(:, :, epoca) = Lexy(:, :, epoca) + DLexy;
Liny(:, :, epoca) = Liny(:, :, epoca) + DLiny;
Ly(:, :, epoca) = Lexy(:, :, epoca) - Liny(:, :, epoca);

end
RVy(:, :, epoca+1)=RVy(:, :, epoca);
RVx(:, :, epoca+1)=RVx(:, :, epoca);
Wy(:, :, epoca+1)=Wy(:, :, epoca);
Wx(:, :, epoca+1)=Wx(:, :, epoca);
Lexx(:, :, epoca+1) = Lexx(:, :, epoca);
Linx(:, :, epoca+1) = Linx(:, :, epoca);
Lx(:, :, epoca+1) = Lx(:, :, epoca);
Lexy(:, :, epoca+1) = Lexy(:, :, epoca);
Liny(:, :, epoca+1) = Liny(:, :, epoca);
Ly(:, :, epoca+1) = Ly(:, :, epoca);
figure(2)
subplot(2,1,1)
plot(j,RVy(80,:,epoca+1),'b',j,RVy(50,:,epoca+1),'r',j,RVy(170,:,e
poca+1),'g')
subplot(2,1,2)
plot(j,RVx(80,:,epoca+1),'b',j,RVx(50,:,epoca+1),'r',j,RVx(170,:,e
poca+1),'g')
figure(3)
plot([1:180],Lx(90,:,epoca+1),'r',[1:180],Ly(90,:,epoca+1),'b')
sum(Lexx(90,:,epoca+1))
sum(Linx(90,:,epoca+1))
sum(Lexy(90,:,epoca+1))
sum(Liny(90,:,epoca+1))
end
```

# Appendice D

Codice matlab utilizzato nella simulazione del ventriloquismo e dell'enhancement

```

clear
clc
close all
% x acustico; y visivo

N=180; %numero di neuroni

%caratteristica dei neuroni
phi=16;
pend=2;
tau=5;

% input
E0y=20; % forza input visivo
sy=4; % std input visivo
E0x=43; % forza input uditivo
sx=30; % std input uditivo

k=[1:1:180];

%leggo le sinapsi
load caso4 % file.mat in cui sono salvati i dati relativi
all'addestramento

%Grafico un input visivo e uno uditivo nella stessa posizione
prima dell'addestramento
Position_y = 90; %posizione input visivo
Position_x = 110; %posizione input uditivo
%genero l'input visivo
distanza_y=abs(k-Position_y);
controllo_dist_y=double((abs(k-Position_y)>(N/2)));
D2_y=((distanza_y-sign(controllo_dist_y)*N).^2);
Vy=E0y*exp(-D2_y/2/sy/sy)/(sqrt(2*pi*sy*sy));

%genero l'input uditivo
distanza_x=abs(k-Position_x);
controllo_dist_x=double((abs(k-Position_x)>(N/2)));
D2_x=((distanza_x-sign(controllo_dist_x)*N).^2);
Vx=E0x*exp(-D2_x/2/sx/sx)/(sqrt(2*pi*sx*sx));

dt=0.2; % passo di campionamento
L_stimolo = 50/dt; % lunghezza dello stimolo
t=[0:L_stimolo]*dt; % asse dei tempi
LL=length(t);
y=zeros(N,LL);
x=y;

```

## Appendici

```
% risposta neurale visiva e acustica in 50 ms
figure(1)
for kt =1:LL-1,
    y(:,kt+1) = y(:,kt) + (dt/tau)*(-y(:,kt)+1./(1+exp(-
(RVy(:, :, 76)*Vy'+Ly(:, :, 76)*y(:,kt)+Wy(:, :, 76)*x(:,kt)-
phi)*pend)));
    x(:,kt+1) = x(:,kt) + (dt/tau)*(-x(:,kt)+1./(1+exp(-
(RVx(:, :, 76)*Vx'+Lx(:, :, 76)*x(:,kt)+Wx(:, :, 76)*y(:,kt)-
phi)*pend)));
    plot(k,x(:,kt+1), 'r',k,y(:,kt+1), 'b')
    axis([0 180 0 1.1])
    pause(0.02)
end

grid

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

xtot = sum(x(:,LL:end),2); %sommo i valori temporali da 100ms
alla fine
posizione_audio=sum(xtot'.*ascissa)/sum(xtot) %metodo del
baricentro

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

ytot = sum(y(:,LL:end),2); %sommo i valori temporali da 100ms
alla fine
posizione_visivo=sum(ytot'.*ascissa)/sum(ytot) %metodo del
baricentro

xytot=xtot+ytot;
ascissa=1:1:180;
```

```
posizione_totaleo=sum(xytot'.*ascissa)/sum(xytot) %metodo  
del baricentro
```



# Bibliografia

Alais, D. & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current biology : CB*, 14, 257-262.

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory Processing in Review: from Physiology to Behaviour. *Seeing and Perceiving*, 23, 3-38.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, 21, 301-308.

Bertelson P, Radeau M (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept Psychophys* 29: 578–584.

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

Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., & Merabet, L. B. (2010). Auditory enhancement of visual phosphene perception: The effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, 477, 109-114.

Buchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, 394, 274-277.

Calvert, G. A. (2001). Cross-modal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110–1123.

## Bibliografia

- Calvert, G.A., Spence, C., & Stein, B.E. (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Cuppini, C., Magosso, E., Bolognini, M., Vallar, G., & Ursino, M. (under review). A Neurocomputational Analysis of the Sound-Induced Flash Illusion. *Neuroimage*.
- Driver, J., and Spence, C. (1998). Crossmodal links in spatial attention. *Philos. Trans. R. Soc. Lond. B* 353, 1319–1331.
- Eimer, M., and Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: evidence from event-related potentials. *Psychophysiology* 39, 437–449.
- Farné, A., and Làdavas, E. (2002). Auditory peripersonal space in humans. *J. Cogn. Neurosci.* 14, 1030–1043.
- Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147, 332–343.
- Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Làdavas, E. (2005). Audiovisual integration in patients with visual deficit. *J. Cogn. Neurosci.* 17, 1442–1452.
- Freeman, E., & Driver, J. (2008). Direction of Visual Apparent Motion Driven Solely by Timing of a Static Sound. *Current Biology*, 18, 1262–1266.
- Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2005). The aftereffects of ventriloquism: generalization across sound-frequencies. *Acta Psychologica*, 118, 93–100.

- Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception*, 36, 1089- 1103.
- Ghazanfar, A. A. & Schroeder, C. E. (6-1-2006). Is neocortex essentially multisensory? *Trends in cognitive sciences* 10[6], 278-285.
- Graziano, M. S., Hu, X. T., and Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *J. Neurophysiol.* 77, 2268–2292.
- Haggard, P., Christakou, A., and Serino, A. (2007). Viewing the body modulates tactile receptive fields. *Exp. Brain Res.* 180, 187–193.
- Hoshino, O. (2011). Neuronal responses below firing threshold for subthreshold cross-modal enhancement. *Neural Computation*, 23, 958-983.
- Howard, I. P. & Templeton, W. B. (1966). *Human spatial orientation*. Oxford, England: John Wiley & Sons.
- Innes-Brown, H., & Crewther, D. (2009). The Impact of Spatial Incongruence on an Auditory Visual Illusion. *Plos One*, 4, e6450.
- Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G., and Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol.* 78, 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.

## Bibliografia

- Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learning & Memory*, 9, 268-278.
- Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289, 1206-1208.
- Macaluso, E., Frith, C. D., and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208.
- Magosso, E. (2010). Integrating information from vision and touch: a neural network modeling study. *IEEE Transactions on Information Technology in Biomedicine*, 14, 598-612.
- Magosso, E., Cuppini, C., & Ursino, M. (2012). A Neural Network Model of Ventriloquism Effect and Aftereffect. *Plos One*, 7, e42503.
- Magosso, E., Ursino, M., di Pellegrino, G., Ladavas, E., & Serino, A. (2010). Neural bases of peri hand space plasticity through tool-use: insights from a combined computational-experimental approach. *Neuropsychologia*, 48, 812-830.
- 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.

- Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: a brief overview. *Brain Research. Cognitive Brain Research*, 14, 31-40.
- Meredith, M. A., & Stein, B. E. (1986b). Visual, Auditory, and Somatosensory Convergence on Cells in Superior Colliculus Results in Multisensory Integration. *J Neurophysiol*, 56, 640-662.
- Meredith, M., & Stein, B. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science* (221), 389-391.
- Morrell, F. (1972). Visual system's view of acoustic space. *Nature*, 238, 44-46.
- Musacchia, G. & Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hearing Research*, 258, 72-79.
- Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, 442, 257-261.
- Perrault, T. J. Jr., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2005). Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J. Neurophysiol.* 93, 2575–2586.
- Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L., & Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science*, 299, 568-572.
- Radeau, M., & Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs – Thomas (1941) Revisited. *Psychological Research*, 49, 17-22.

## Bibliografia

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.

Rowland, B., Quessy, S., Stanford, T., & Stein, B. (2007). Multisensory integration shortens physiological response latencies. *Journal of Neuroscience* (22), 5879-5884.

Sarri, M., Blankenburg, F., and Driver, J. (2006). Neural correlates of crossmodal visual-tactile extinction and of tactile awareness revealed by fMRI in a right-hemisphere stroke patient. *Neuropsychologia* *44*, 2398–2410.

Schroeder, C. E. & Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Current Opinion in Neurobiology*, *15*, 454-458.

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews*, *7*, 269-284.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*, 788-788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*, 147-152.

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

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. *J. Neurosci.* 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.