

ALMA MATER STUDIORUM UNIVERSITÀ DI
BOLOGNA CAMPUS DI CESENA

SCUOLA DI INGEGNERIA E ARCHITETTURA
Corso di Laurea Magistrale in Ingegneria Biomedica

“INDIVIDUAZIONE DEI CAMPI RECETTORI
DI CELLULE IN CORTECCIA VISIVA
PRIMARIA (V1), ATTRAVERSO LO STUDIO
DI DATI AD ALTA DIMENSIONALITÀ”

Tesi in:
Bioimmagini e Visione LM

Relatore:
Prof. Ing. MAURO URSINO

Presentata da:
LUCA BENINI

Correlatori:
Prof. Ing. ALESSANDRO
SARTI
Ing. GIACOMO COCCI

SESSIONE III
ANNO ACCADEMICO 2012–2013

PAROLE CHIAVE:

campi recettori

cellule complesse

correlazione inversa

Indice

Introduzione	vii
Capitolo I	1
1 Il sistema visivo	1
1.1 Caratteristiche e funzionamento	2
1.2 Il percorso visivo primario	9
1.3 Il campo recettivo	21
1.3.1 Organizzazione	22
1.3.2 Separabilità	28
Capitolo II	28
2 Ricostruzione di campi recettivi	31
2.1 Campo recettivo e risposta neurale	32
2.2 Spike-triggered average (STA)	35
2.3 Spike-triggered covariance (STC)	39
2.3.1 Utilizzo della STC su stimoli naturali	42
2.4 Ricostruzione di campi recettivi	44
2.5 Data set di Ringach (V1 del macaco)	53
2.5.1 Il Clustering	57
2.5.2 Ricostruzione dei campi recettivi	65
Capitolo III	66
3 Individuazione del campo recettivo	67
3.1 I filtri di Gabor	69
3.2 L'energia di orientazione	80

3.3	Artefatti ai bordi	84
3.4	La scale space	89
3.5	Lo Z-score	98
3.6	Risultati ottenuti	102
3.7	Conclusioni	112
4	Ringraziamenti	115
	Bibliografia	117

Introduzione

Grazie ai suoi sensi l'uomo riesce a percepire il mondo e lo spazio che lo circonda, a muoversi ed interagire con esso.

L'argomento che andrò a trattare in questo lavoro di tesi si concentrerà sulla *vista*, una caratteristica indispensabile nella vita quotidiana di ogni essere vivente, responsabile dell'elaborazione delle immagini che percepiamo.

Negli anni '50 e '60 vennero effettuati i primi studi elettro-fisiologici relativi ai metodi attraverso i quali il nostro organismo elabora le immagini e determina l'esistenza di stimoli; la luce che arriva alla retina, viene trasformata in segnale elettrico, il quale a sua volta viene inviato al talamo e poi alla corteccia visiva primaria (V1), dove l'informazione viene elaborata e ricostruita secondo le caratteristiche geometriche spazio-temporali dello stimolo^[9,12].

E' stato proposto, e nel tempo accettato dalla maggior parte del mondo accademico, un modello riduzionalistico per descrivere il comportamento dei neuroni della corteccia visiva primaria (V1). In essa si trovano cellule *semplici* selettive in orientazione, frequenza spaziale e velocità di moto, le quali calcolano linearmente una somma pesata di parte del segnale visivo, ben approssimabile con un filtro di Gabor bidimensionale (minimizzando così l'incertezza tra localizzazione nello spazio reale e quello armonico). Le altre cellule chiamate *complesse* sommano input derivanti da varie cellule semplici aventi proprietà selettive simili, ma con differenze in fase o posizione spaziale. E' stato dimostrato inoltre che la corteccia visiva primaria è organizzata in modo da minimizzare sia l'indeterminazione sulla misura dell'orientazione dello stimolo, sia la misura della velocità locale degli stimoli visivi, e che è possibile ben modellare i profili recettivi spazio-temporali con funzioni di Gabor tridimensionali^[2,7,35].

Il progetto di ricerca che ho svolto in questi 10 mesi si è focalizzato sull'integrazione dei risultati raggiunti grazie all'elaborazione di nuovi dati sperimentali. Questi sono stati prelevati dalla corteccia visiva di macachi, attraverso l'utilizzo di tecniche di registrazione elettrofisiologiche mediante array di micro-elettrodi^[25], durante la presentazione di alcuni filmati (sequenze di immagini o frames).

Attraverso la tecnica del *clustering*, dalle registrazioni degli esperimenti sono stati raggruppati gli spike appartenenti ad uno stesso neurone, sfruttando alcune caratteristiche come la forma del potenziale d'azione^[32]. Da questa elaborazione è stato possibile risalire a quali stimoli hanno prodotto una risposta neurale.

I dati messi a disposizione da Ringach^[25] non potevano essere trattati direttamente con le tecniche della *spike-triggered average* e della *spike-triggered covariance* a causa di alcune loro caratteristiche. Utilizzando *filtri di Gabor* bidimensionali e l'*energia di orientazione* è stato però possibile modellare la risposta di cellule complesse in corteccia visiva primaria. Applicare questi modelli su dati ad alta dimensionalità (immagini molto grandi), sfruttando la tecnica di *standardizzazione* (Z-score), ha permesso di individuare la regione, la scala e l'orientazione all'interno del piano immagine dei profili recettivi delle cellule di cui era stata registrata l'attività neurale. Ritagliare tale regione e applicare la *spike-triggered covariance* su dati della giusta dimensionalità, permetterebbe di risalire ai profili recettivi delle cellule eccitate in un preciso momento, da una specifica immagine e ad una precisa scala e orientazione.

Se queste ipotesi venissero confermate si potrebbe marcare e rafforzare la bontà del modello utilizzato per le cellule complesse in V1 e comprendere al meglio come avviene l'elaborazione delle immagini.

Capitolo 1

Il sistema visivo

In questo primo capitolo ho analizzato la fisiologia del sistema visivo, andando a suddividere e a studiare in modo approfondito ogni sua singola parte.

La luce che dal mondo esterno entra nell'occhio e va a colpire i fotorecettori viene trasformata in informazione, per poi compiere un lungo percorso e numerose elaborazioni prima di arrivare al cervello. Qui termina il suo straordinario percorso e milioni di cellule lavorando in sinergia ci permettono di distinguere i colori, gli spazi, la profondità... . In modo specifico mi sono occupato delle cellule presenti in corteccia visiva primaria (semplici e complesse), analizzando il concetto di *campo recettivo* e soffermandomi sulle sue caratteristiche ed organizzazione.

1.1 Caratteristiche e funzionamento

La vista é uno dei cinque sensi del nostro corpo. Da sempre l'uomo la utilizza, spesso anche inconsciamente, per i compiti che svolge durante la giornata. Il suo funzionamento è molto articolato e complesso e per questo non ancora completamente noto; riuscire a comprendere a pieno il sistema visivo in ogni sua singola parte, permetterebbe di riparare o sostituire quelle parti che funzionano male o cessano di funzionare.

Gli occhi umani sono posizionati nelle due cavità orbitare del cranio e rappresentano l'organo principale del **sistema visivo**. La superficie esterna di ciascun occhio è una sfera incompleta e opaca. Essa ha il compito di ricavare informazioni sull'ambiente circostante grazie alla luce che proviene dall'esterno, regolarne l'intensità attraverso un diaframma, focalizzarla attraverso un sistema regolabile di lenti per formare un'immagine e trasformarla in segnali elettrici che attraverso il nervo ottico vengono inviati al cervello per l'elaborazione e l'interpretazione.

Partendo dall'esterno si trova la **congiuntiva**, una mucosa trasparente che ricopre la superficie interna delle palpebre e la superficie anteriore del globo oculare. Ha funzioni protettive, difensive e riparatrici per l'occhio^[3].

La **cornea** è una membrana trasparente formata da proteine, più convessa rispetto al resto dell'occhio. Viene suddivisa in 5 strati attraverso i quali la luce penetra all'interno dell'occhio per finire sulla retina. Dietro la cornea c'è una cavità (camera anteriore) che contiene un liquido trasparente e acquoso, l'*umor acqueo*, utile per mantenere la pressione e la forma del globo oculare; è racchiuso da uno strato membranoso sottile, la membrana ialoidea e composto essenzialmente da acqua, ma anche da vitamina C, glucosio, acido lattico e proteine. Sono poi presenti tre strati di tessuto che, partendo dall'esterno, prendono il nome di:

- **sclera**: un rivestimento protettivo che ricopre circa l'80% della superficie oculare. E' unita alla cornea nella parte anteriore.

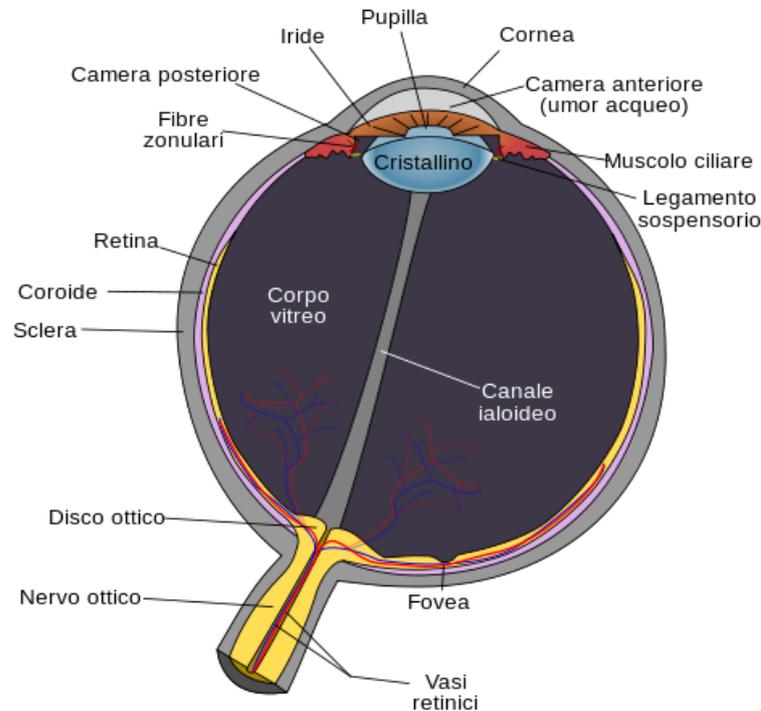


Figura 1.1: Anatomia dell'occhio.

- **coroide**: molto ricca di vasi sanguigni; riveste il 60% della parte posteriore del bulbo oculare. E' unita al corpo ciliare.
- **iride**: una membrana muscolare circolare posta dietro la cornea e davanti al cristallino che determina il colore dei nostri occhi in base alla pigmentazione di cui è formata. Presenta un'apertura circolare al centro, la *pupilla*, le cui dimensioni sono controllate da un muscolo posto sul suo margine. Contraendosi e rilassandosi, questo muscolo fa allargare o rimpicciolire la pupilla stessa, controllando la quantità di luce che penetra nell'occhio.

Il **cristallino** è la lente interna dell'occhio, posta dietro l'iride,

di forma circolare, è formato da un gran numero di fibre trasparenti disposte in strati. La sua funzione fondamentale è di far convergere i raggi luminosi sul piano retinico e quindi consentire una visione nitida sia da lontano che da vicino. Per fare questo sfrutta la sua capacità di accomodazione, ossia di modificare la sua forma per effetto della contrazione del *muscolo ciliare* a cui è collegato da alcuni legamenti che ne modificano la lunghezza focale, cioè la distanza alla quale esso mette a fuoco le immagini^[3].

Dietro il cristallino, il bulbo oculare contiene una sostanza gelatinosa trasparente, elastica e filamentosa l'*umor vitreo*, che riempie l'occhio per il 90% del suo volume. E' costituito in massima parte da acqua (99%) oltre che da fibrille collagenose, acido ialuronico, proteine solubili ed elettroliti. La presenza dell'umor vitreo contribuisce al mantenimento di una giusta pressione endoculare e quindi a tenere il bulbo oculare disteso, in più, è capace di ammortizzare gli urti e mantenere la retina nella sua posizione^[3].

Il cuore dell'occhio e' costituito dalla *retina*, una membrana fotosensibile di 0,2 mm, formata in gran parte da due tipi di cellule nervose stratificate (in totale oltre 120 milioni di elementi) che poggiano, dal lato esterno della retina, su uno strato pigmentato:

- *coni*: responsabili della visione diurna, permettono un'alta acutezza visiva. Hanno la massima concentrazione nella *fovea*, fino a 160.000 per mm^2 , e presiedono alla percezione del colore e alla nitidezza dei contrasti.
- *bastoncelli*: sono più sensibili dei precedenti alla luce e permettono la visione crepuscolare. Hanno massima densità nella parte periferica.

Per distinguere i colori, i coni non sono tutti uguali, infatti, ne sono presenti di tre tipi sensibili ai tre colori fondamentali: il rosso, il verde e il blu.

In ogni occhio vi sono circa 6 milioni di coni e 120 milioni di bastoncelli. I coni sono presenti in numero nettamente inferiore rispetto ai

bastoncelli, ciò significa che il nostro occhio ha maggiore potere risolutivo per punti luminosi piuttosto che per punti colorati. A meno di variazioni individuali, infatti, è possibile distinguere due punti luminosi su sfondo scuro se separati da 1 mm e osservati dalla distanza di 3,44 m. Se i due punti fossero di diverso colore, invece, non saremmo in grado di apprezzare anche la differenza cromatica se non almeno raddoppiando la distanza fra di essi.

Sulla retina, in perfetta opposizione alla pupilla, si trova una piccola zona ellissoidale del diametro di circa 2,5 mm, di colore giallo, chiamata *macula*, essa corrisponde alla zona di massima acutezza visiva dell'occhio, ossia con maggior concentrazione di coni al centro della quale è presente la *fovea* in cui le cellule fotosensibili sono rappresentate solo da coni. Allontanandosi dalla fovea, compaiono sempre più bastoncelli che, procedendo verso la periferia, sostituiscono i coni completamente, all'estremità esterna.

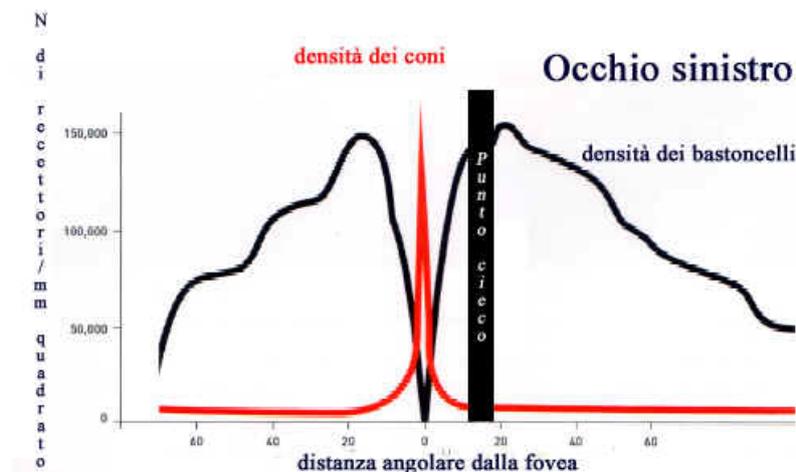


Figura 1.2: Distribuzione di coni e bastoncelli nella retina.

La visione totale dell'occhio fermo abbraccia un campo di 140° in senso orizzontale e di circa 120° in senso verticale, mentre quello della fovea poco più di 1° ^[5].

× × × × × × × × ×



Figura 1.3: Test per la verifica della presenza del punto cieco dell'occhio: coprire l'occhio sinistro con la mano, partendo dalla 'X' a sinistra fissare tutte le 'X'. Su una di queste si può notare che il grosso punto nero sulla destra scompare, per poi riapparire subito non appena si sposta nuovamente lo sguardo.

Il “cavo” che collega l'occhio al cervello è il *nervo ottico*; è lungo circa 5 cm ed ha un diametro che varia dai 3 ai 7 mm; è costituito da circa un milione di fibre provenienti dalle cellule gangliari retiniche. Nel punto in cui il nervo ottico penetra nel bulbo oculare si trova una piccola zona rotonda di retina priva di cellule fotosensibili, la papilla ottica, che rappresenta il punto cieco dell'occhio. Questo avviene perchè, l'asse visivo non coincide con la direzione nella quale il nervo ottico s'innesta sulla retina^[1]. E' possibile fare un piccolo test per evidenziare questo difetto rappresentato in figura 1.3.

Per un funzionamento ottimale l'occhio deve essere in grado di spostare rapidamente lo sguardo fra gli oggetti posti a distanze differenti. Questi movimenti vengono prodotti da sei muscoli che spostano il bulbo oculare in alto, in basso, a sinistra, a destra e in senso obliquo garantendo la stabilità visiva. Sono innervati da tre paia di nervi cranici, dalla pupilla che si dilata o restringe, tramite il muscolo dell'iride, e dal cristallino che cambia il suo raggio di curvatura tramite il muscolo ciliare (questi ultimi due muscoli sono involontari). I movimenti dei muscoli oculari sono molto precisi: è stato stimato che gli occhi possono essere mossi per mettere a fuoco non meno di 100.000 punti distinti del campo visivo. I muscoli oculari, lavorando insieme, hanno anche l'importante funzione di far convergere entrambi gli occhi sullo stesso punto, in modo che le immagini percepite dai due occhi coincidano. Questo lavoro, unito all'incessante elaborazione delle immagini ricevute da parte del cervello, garantisce la fusione delle immagini,

la valutazione visiva delle dimensioni, della tridimensionalità e della distanza dell'oggetto osservato, l'effetto di profondità e spessore. Nell'occhio la messa a fuoco viene ottenuta con l'appiattimento o l'arrotondamento del cristallino: tale processo viene chiamato *accomodazione*. Quando l'occhio è rilassato, l'accomodazione non è necessaria per vedere oggetti lontani. Per gli oggetti più vicini, il cristallino viene progressivamente arrotondato dalla contrazione del corpo ciliare, che fa rilassare il legamento. I due occhi, lavorando come un unico organo, localizzano un oggetto stimolando i punti corrispondenti (elementi visivi dei due occhi che condividono la direzione visiva) delle due retine così che esso possa apparire unico. L'input visivo viene considerato dal sistema posturale se la mira visiva è distante massimo 5 metri (campo visivo vicino). In caso contrario le informazioni derivanti dai recettori visivi verranno trascurate^[4].

I neuroni binoculari permettono di stimare la distanza degli oggetti valutando il diverso angolo formato dallo stesso stimolo visivo sui due occhi. La differenza di angolo si chiama *disparità*. La disparità è importante per valutare la distanza a meno di 30 metri, oltre utilizziamo altri indizi e quindi abbiamo la stessa capacità valutativa in visione monoculare che in visione binoculare.

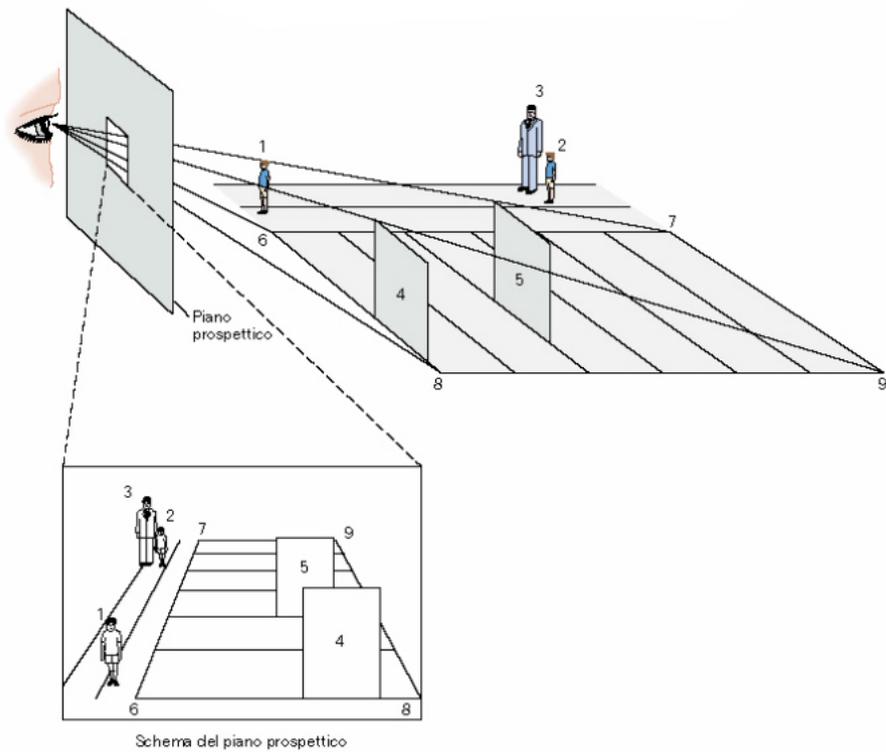


Figura 1.4: La visione monoculare

1.2 Il percorso visivo primario

L'analisi dell'informazione visiva è allo stesso tempo parallela e gerarchica. L'uomo è in grado di percepire solo le radiazioni elettromagnetiche comprese nella gamma di lunghezze d'onda tra 380 e 780 nm. La superficie sensibile dell'occhio è costituita dai fotorecettori (i bastoncelli ed i coni), il cui compito è quello di trasformare in impulsi elettrici le informazioni ricevute dalle reazioni fotochimiche che vengono attivate dalla radiazione luminosa e di inviare questi segnali ai neuroni retinici - le cellule orizzontali, bipolari, amacrine e ganglionari - che sono variamente connessi fra di loro ed effettuano una prima elaborazione del segnale visivo. Dalla figura 1.5 è possibile vedere come la luce arriva allo strato dei fotorecettori solo dopo aver attraversato l'intero spessore della retina.

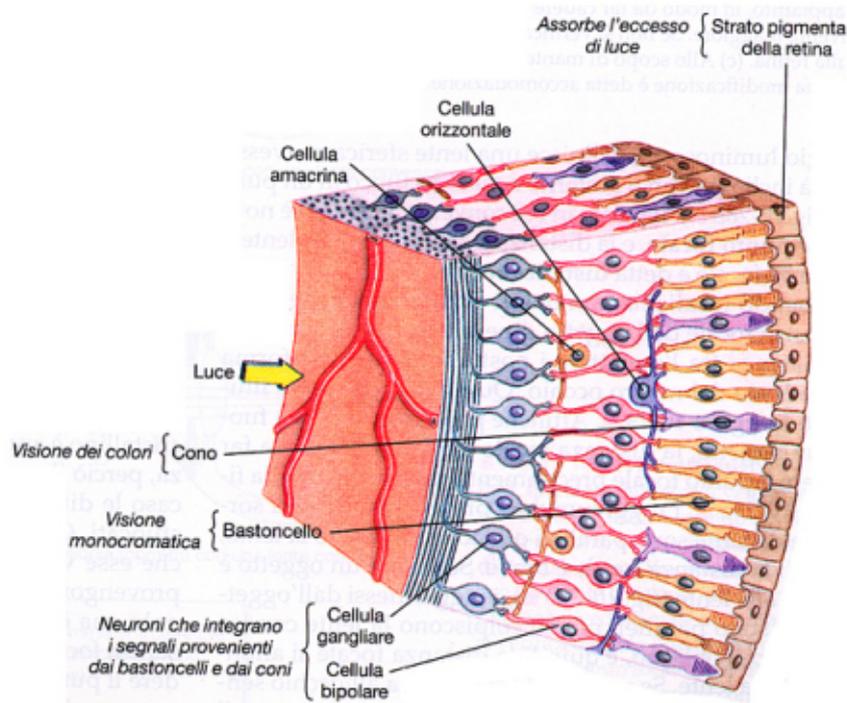


Figura 1.5: Struttura della retina.

Il meccanismo della visione comporta la sensibilizzazione delle cellule della retina da parte di un pigmento fotosensibile che nei bastoncelli prende il nome di rodopsina e nei coni di iodopsina; questo è il fenomeno che dà li via ad una serie di reazioni chimiche e stimolazioni nervose, il cui esito finale è la percezione di luci e colori.

I fotopigmenti, a contatto con un fotone, cambiano la propria struttura molecolare scatenando un'iperpolarizzazione che rende fortemente negativo il potenziale di membrana: questa iperpolarizzazione determina una diminuzione nella produzione di glutammato da parte del fotorecettore, alterazione che determina la produzione di stimoli nervosi visivi. Per la produzione della proteina implicata, la rodopsina, è necessaria la vitamina A. La rodopsina viene inattivata per azione della luce e deve essere riformata dai bastoncelli in condizioni di oscurità ma questo non avviene istantaneamente. Si può osservare questo effetto in particolar modo, passando dalla luce del sole a una stanza buia, percependo per qualche istante una cecità quasi completa.

I ***bastoncelli*** possiedono un'alta amplificazione del segnale e sono in grado di rispondere, al massimo della loro sensibilità (visione scotopica), addirittura al singolo fotone di luce, in virtù del tempo di integrazione della risposta così lunga. Il rovescio della medaglia si riscontra nella saturazione della risposta a stimoli troppo intensi. Quindi funzionano al loro meglio quando il livello di luminosità generale è molto basso, visto che possono amplificare enormemente la poca luce presente.

Infine, possedendo un solo fotopigmento, i bastoncelli possono fornire una visione unicamente monocromatica: la rodopsina assorbe a 498 nm, quindi la luce più efficiente a stimolare i bastoncelli è sostanzialmente compresa fra il blu e il verde.

I ***coni*** non sono sensibili come i bastoncelli: al massimo della loro sensibilità hanno bisogno di almeno un centinaio di fotoni per iniziare a rispondere alla luce; in compenso hanno una risposta più veloce di quella dei bastoncelli e saturano per quantità di luce molto più alte di quelle per cui saturano i bastoncelli.

I coni possiedono 3 tipi di pigmenti visivi che mediano le informazioni

sui colori che noi percepiamo (ogni cono della retina contiene solo uno dei tre fotopigmenti) i coni cosiddetti "blu" (short) assorbono prevalentemente la luce a 437 nm, che è nel blu-violetto appunto, come evidenziato nella figura 1.6; i coni "verdi" (middle) hanno un picco di assorbimento della luce a 533 nm, corrispondente a un azzurro-verde; i coni chiamati "rossi" (long) hanno in realtà un picco d'assorbimento a 564 nm che, di fatto, è un colore giallo-verde. I diversi tipi di coni non sono distribuiti uniformemente: il 60% circa dei coni contiene il pigmento "rosso", mentre il 30% contiene il "verde" e solo il 10% contiene il pigmento "blu".

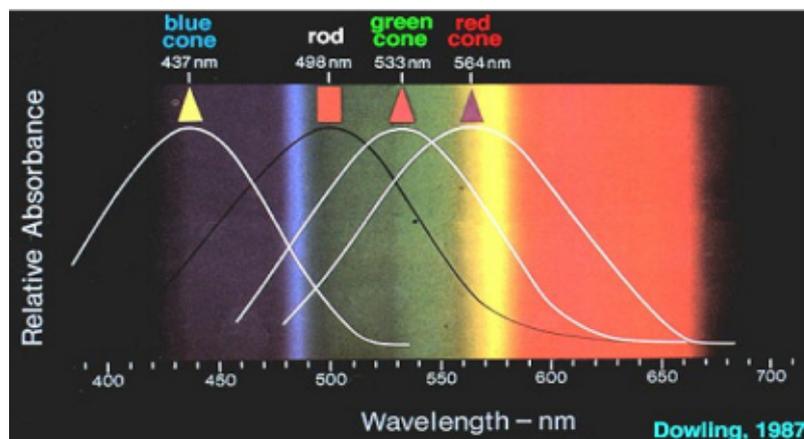


Figura 1.6: Sensibilità dei tre coni e dei bastoncelli.

La sensazione di colore è legata ai rapporti delle attività neurali dovute ai tre fotopigmenti: ad esempio un raggio luminoso monocromatico della lunghezza d'onda di 450 nm produce una risposta forte dal fotopigmento "blu", una risposta più debole dal "verde" ed una più debole ancora dal "rosso"^[2].

Lo strato esterno di pigmento bruno su cui poggia la superficie esterna della retina protegge i coni da un'eccessiva esposizione alla luce.

I fotorecettori sono connessi alle *cellule gangliari* alle quali mandano le informazioni. Le cellule gangliari non sono completamente silenti, ma hanno un'attività spontanea che viene modulata dalla luce. Ogni

cellula gangliare risponde ad una specifica area della retina, chiamata *campo recettivo*, il quale è di forma circolare. Il campo recettivo di una cellula gangliare è quell'area della retina dove la stimolazione di un recettore causa una modifica nella sua frequenza di scarica. Le cellule gangliari possono essere distinte, come mostrato in figura 1.7, in cellule M (per *magnae* o grandi) e cellule P (per *parvae* o piccole).

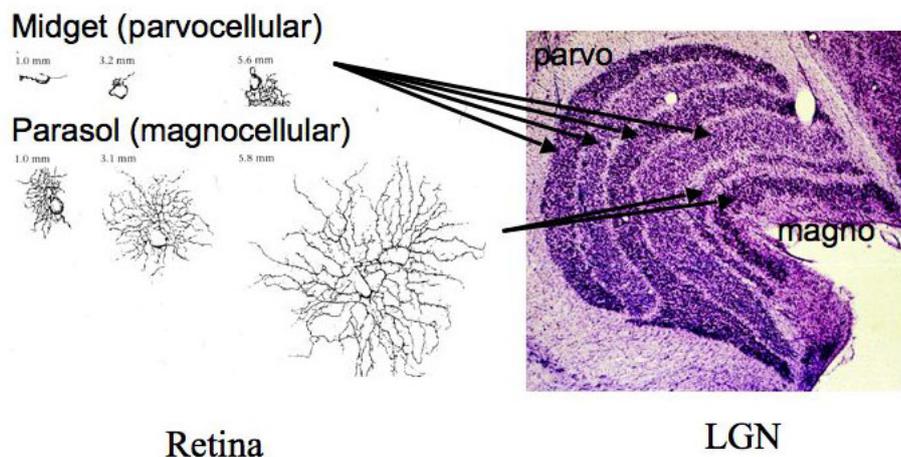


Figura 1.7: Strati del nucleo genicolato laterale (LGN); cellule Magno e Parvo.

Le cellule P sono selettive per la lunghezza d'onda e per le alte frequenze spaziali e hanno una risposta lenta e sostenuta. Sono connesse con un minor numero di recettori e sono adatte a fornire informazioni sulla forma e sul colore di un oggetto. Le cellule M non sono selettive per le lunghezze d'onda, ma lo sono per le basse frequenze ed hanno una risposta fasica e una maggior velocità di conduzione. Sono connesse con un gran numero di coni e bastoncelli e per questo sono in grado di fornire informazioni sul movimento di un oggetto o su rapide variazioni del contesto luminoso. Le cellule gangliari possono essere

distinte in due classi sulla base della loro risposta a piccoli spot di luce applicati al centro del loro campo recettivo in:

- “*On center*”; generano pochi potenziali d’azione al buio; la luce diretta al centro del loro campo recettivo aumenta la loro frequenza di scarica (sono eccitate quando la luce è accesa). La luce applicata alla periferia inibisce l’effetto prodotto dall’illuminazione del centro. Lo stimolo inibitorio più efficace è costituito da un anello di luce che colpisce l’intera periferia.
- “*Off center*”; sono inibite dalla luce applicata al centro del loro campo recettivo. La loro frequenza di scarica è massima per un breve periodo dopo che la luce è stata spenta (sono eccitate quando la luce è spenta).

Le proprietà dei campi recettivi delle cellule gangliari si mantengono costanti per gran parte delle intensità luminose. Tuttavia, dopo l’adattamento al buio estremo o a luce molto debole (ad es. la luce delle stelle) per più di un’ora, le proprietà cambiano, e l’illuminazione della periferia cessa di inibire la risposta all’illuminazione del centro. Le cellule gangliari On e Off sono presenti in un numero approssimativamente uguale, e forniscono due vie parallele per l’analisi dell’informazione sensoriale. Questo è dovuto al fatto che ogni fotorecettore manda output a entrambi i tipi di cellule gangliari. Avendo canali separati (On e Off center), che lavorano parallelamente, risulta un aumento dell’efficienza del sistema visivo soprattutto nell’identificare i cambiamenti repentini del livello di illuminazione. Una funzione di questa elaborazione è probabilmente quella di migliorare la capacità dei centri superiori di rilevare oggetti che contrastano debolmente con lo sfondo e rapidi cambi dell’immagine visiva^[7].

I diversi parametri dello stimolo sono analizzati in due vie parallele e separate: la via del “cosa” (what) e la via del “dove” (where). La prima codifica le informazioni circa le caratteristiche dello stimolo (come forma e colore) e l’identità degli oggetti, e può essere suddivisa in due

ulteriori percorsi, colore e forma. La via *where*, invece, analizza le informazioni spaziali relative all'oggetto ed è generalmente suddivisa in analisi del movimento e analisi della forma derivante dal movimento. Questa suddivisione inizia ad essere evidente a livello retinico^[6].

Il collegamento diretto tra i recettori e le cellule gangliari è rappresentato dalle **cellule bipolari**. L'informazione visiva viene trasferita dai coni alle cellule gangliari in due modi diversi: diretto e indiretto. Nel primo caso, i coni, che sono al centro del campo recettivo gangliare, stabiliscono un contatto sinaptico diretto con cellule bipolari che a loro volta contattano direttamente le cellule gangliari. Il segnale che viene dai coni posti alla periferia del campo recettivo viene invece trasferito alle cellule gangliari in modo indiretto, ovvero per mezzo di **cellule orizzontali** e **cellule amacrine**.

Le cellule orizzontali e bipolari non posseggono canali sodio voltaggio-dipendenti capaci di generare potenziali d'azione, e trasmettono i segnali solo passivamente, per via elettrotonica^[7].

Gli assoni di tutte le cellule gangliari si riuniscono a formare il **nervo ottico** che fuoriesce dall'occhio e che conduce l'informazione visiva fuori dalla retina fino ai centri superiori, dapprima al corpo genicolato laterale e da qui alle aree corticali. I nervi dei due occhi si congiungono prima di raggiungere il LGN e formano il **chiasma ottico**.

L'informazione sensoriale così elaborata viene codificata dalle cellule gangliari e trasmessa lungo gli assoni di queste principalmente al nucleo genicolato laterale. Il tratto ottico di destra è composto dalle fibre nervose provenienti dall'emiretina temporale dell'occhio destro e dall'emiretina nasale dell'occhio sinistro; viceversa, il tratto ottico di sinistra è composto dalle fibre provenienti dall'emiretina temporale dell'occhio sinistro e dall'emiretina nasale dell'occhio destro. In pratica, ciascun tratto ottico contiene la rappresentazione completa dell'emicampo visivo controlaterale. Si avrà dunque, che la corteccia visiva destra "vedrà" il campo visivo di sinistra, mentre la corteccia visiva sinistra "vedrà" il campo visivo di destra. L'informazione visiva può essere acquisita dal sistema nervoso centrale solo durante le fissazioni, dato che durante i movimenti saccadici sono attivi meccanismi

inibitori^[5].

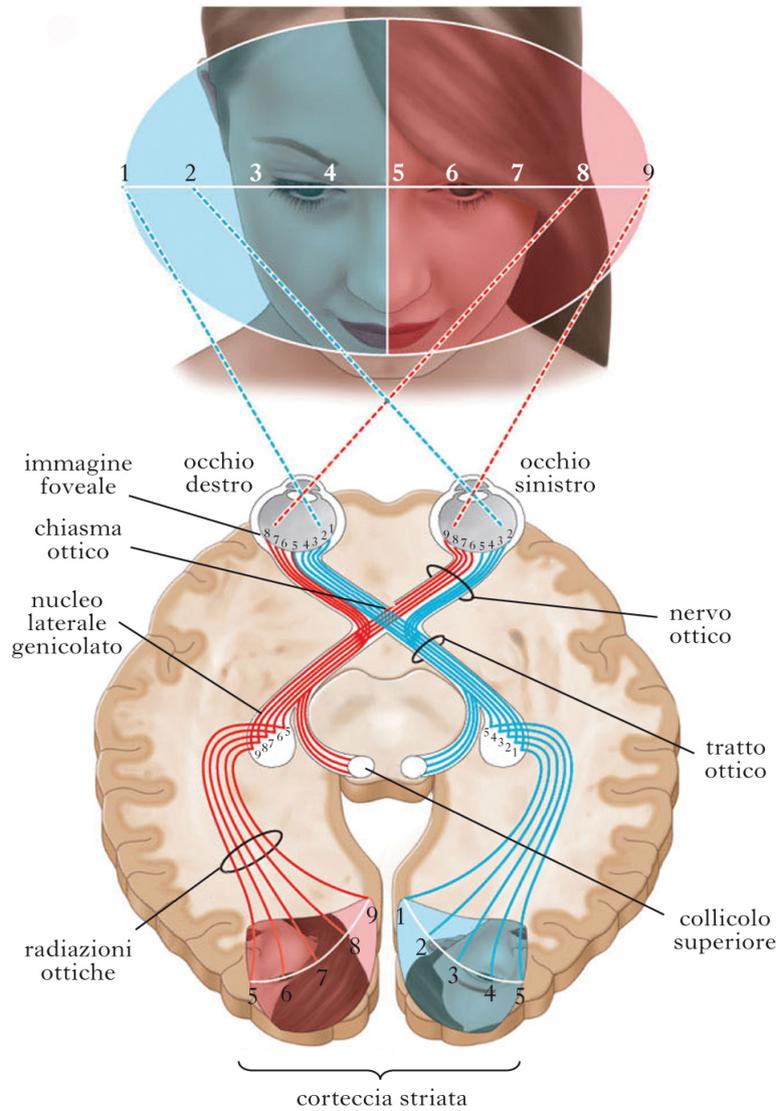


Figura 1.8: Separazione delle due parti del campo visivo e percorso dell'informazione visiva.

Il *nucleo genicolato laterale (LGN)* è una porzione del talamo preposta al trattamento dell'informazione visiva proveniente dalla

retina ed attraverso il quale passano quasi tutti gli input sensoriali verso la corteccia ad eccezione di quelli olfattivi.

Le fibre nervose provenienti da punti diversi della retina si dirigono verso punti diversi del nucleo genicolato (LGN) e della corteccia, ricreando così una mappa cerebrale della retina nel cervello. Nella proiezione dalla retina al LGN vengono preservate le caratteristiche topografica, infatti gli assoni provenienti dalla retina vanno a connettersi con i neuroni del LGN secondo un ordine preciso che dipende dalla posizione dei relativi neuroni sulla retina. La funzione di questa stazione intermedia non è ancora del tutto nota e, sebbene sembri svolgere una semplice azione di trasferimento, per le sue dimensioni e per la posizione che occupa, potrebbe rivestire un ruolo ben più importante. In particolare, l'esistenza di un certo numero di sinapsi inibitorie e la consistente presenza di terminazioni di retroazione che dalla corteccia si innervano nel talamo, fa supporre che il LGN abbia una funzione importante nell'integrazione spazio-temporale dell'informazione^[6].

Dai neuroni del LGN partono le afferenze nervose dirette alla corteccia cerebrale, dove l'elaborazione dei segnali visivi viene perfezionata rendendo possibile la percezione cosciente. A livello del corpo genicolato laterale le fibre delle cellule gangliari terminano in modo estremamente ordinato in sei strati; ciascuno di questi riceve afferenze da un solo occhio; le fibre provenienti dalla retina nasale controlaterale prendono contatto con gli strati 6, 4, 1, mentre quelle che si originano dalla retina temporale terminano negli strati 5, 3, 2. Nel corpo genicolato sinistro si ha la proiezione dell'occhio destro e in quello di destra del sinistro. Queste strisce visive, larghe circa mezzo millimetro, possono essere messe in evidenza usando coloranti specifici, che permettono di conoscere la loro estensione e la loro forma. Si è visto così che la zona della fovea occupa un'area molto grande, come del resto ci si poteva aspettare. Tutto ciò è molto importante, perchè sta ad indicare che il cervello è capace di analizzare ogni stimolo separatamente e di lavorare in parallelo. Il LGN raccoglie l'informazione direttamente dalla retina, ed invia delle proiezioni nella corteccia visiva primaria dove, fra l'altro, avviene l'integrazione. Una parte dell'informazione però

torna indietro, infatti il nucleo genicolato laterale riceve più connessioni dalla V1 che dalla stessa retina: le fibre che dalla V1 giungono al LGN sono quattro volte maggiori di quelle che da quest'ultimo si dipartono verso la corteccia.

La **corteccia visiva primaria (V1)**, è l'area 17 di Brodmann nel lobo occipitale. Furono Hubel e Wiesel agli inizi degli anni '60 i primi a riscontrare un'organizzazione ordinata nella corteccia visiva^[12]; questa è la prima stazione del sistema visivo in cui compaiono cellule che ricevono informazione da entrambi gli occhi (cellule binoculari). Queste cellule sono eccitabili dalla stimolazione di una sola piccola area di campo visivo (campo recettivo) che però è vista da entrambi gli occhi. Le cellule binoculari sono in grado di valutare la profondità e permettono di stimare la distanza a cui siamo dagli oggetti.

La V1 si articola in sei strati principali disposti in bande parallele e riceve le informazioni dalla parte controlaterale del campo visivo. Le innervazioni sono perfettamente retinotopiche, ovvero spazialmente organizzate come i recettori sulla retina.

Le afferenze del nucleo genicolato laterale arrivano principalmente al 4° strato (A, B, C α , C β). Le fibre del cammino magno (M) arrivano allo strato 4C α , da dove sono proiettate al 4B. Al 4C β e al 4A giungono le afferenze del cammino parvo (P). L'organizzazione in strati è fondamentale, non solo dal punto di vista anatomico ma anche funzionale. Infatti penetrando la corteccia, immersi in questi livelli, vi sono due tipi principali di cellule: le **cellule semplici** (strati 4 e 6) e le **cellule complesse** (strati 1, 2, 3 e 5)^[6].

In quest'area della corteccia si procede all'estrazione di bordi a partire dall'immagine proveniente dal nucleo genicolato laterale. Ogni cellula semplice riceve in ingresso le informazioni di luminosità e le organizza in una zona centrale (eccitatoria) e una periferica (inibitoria) a formare un nuovo *campo percettivo*. Questo è organizzato in modo da permettere il riconoscimento di una linea opportunamente inclinata, posizionata nella zona eccitatoria. Le cellule complesse a loro volta hanno come ingresso l'informazione prodotta dalle cellule semplici. Ciascun campo percettivo viene affiancato ed in parte sovrapposto, a

formare delle aree prive di una specifica zona di attivazione/inibizione, ma specifiche per il riconoscimento di bordi, secondo una specifica inclinazione, a prescindere dalla posizione nel campo percettivo.

Le cellule che formano i *blobs* (fig. 1.9) sono specializzate nell'elaborazione dell'informazione cromatica, ma totalmente insensibili alle direzioni. Le aree che circondano i blobs sono dette regioni "interblob". Qui la gran parte delle cellule risponde a stimoli con un particolare orientamento, come linee o barre, e ha campi recettivi piccoli. Le cellule risultano quindi organizzate in colonne. Ogni colonna è formata dagli strati della corteccia, in ciascuno dei quali le cellule si attivano per una stessa direzione specifica con lo stesso campo percettivo come riferimento. Inoltre nelle zone adiacenti studi elettroencefalografici hanno dimostrato l'esistenza di colonne che codificano per le altre direzioni, ma per lo stesso campo percettivo. La sovrastruttura che comprende tutte le colonne di un determinato campo percettivo prende il nome di *ipercolonna*; ogni ipercolonna è affiancata a quella la cui afferenze provengono dall'area omologa dell'occhio controlaterale.

Le cellule con campi recettivi nella stessa area di campo visivo sono raggruppate e formano un'ipercolonna che contiene le cellule che rispondono ad entrambi gli occhi, a tutti gli orientamenti e un blob per l'analisi del colore. Come è possibile vedere dalla figura 1.10, i moduli con proprietà simili sono connessi tra loro da connessioni orizzontali.

Gli esperimenti sembrano confermare che esistano almeno 4 vie che lavorano in parallelo per l'estrazione e l'elaborazione delle informazioni riguardanti l'immagine visiva; le afferenze della V1 sono dirette sia ad altre zone corticali (dagli strati 2, 3 e 4B) sia ai centri profondi per le funzioni di feedback al nucleo genicolato laterale e al collicolo superiore (dagli strati 5 e 6). La destinazione principale delle fibre uscenti da V1 è la corteccia visiva secondaria (V2). Sia la parte riguardante il colore che quella riguardante le forme proiettano a V4, e sembra che la separazione tra queste due suddivisioni continui in V4. La maggior parte delle cellule in V3 è selettiva per l'orientamento e si pensa sia coinvolta nell'analisi delle forme dinamiche. V5 è essenziale per elaborare informazioni circa il movimento e la profondità^[6].

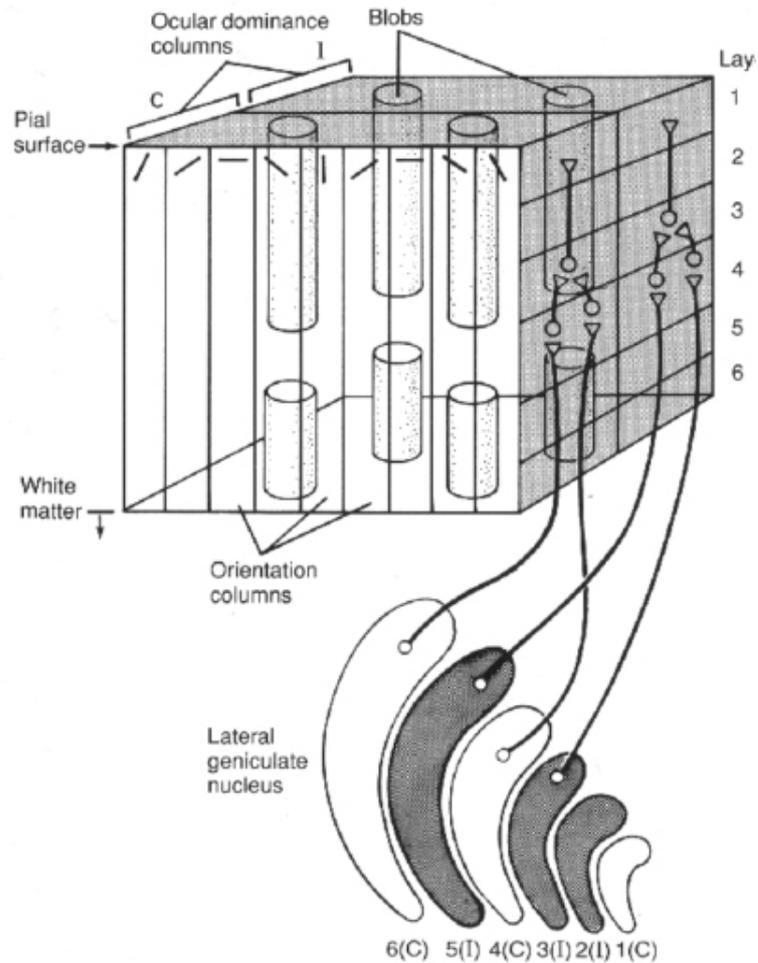


Figura 1.9: L'informazione dal LGN arriva alle cellule in V1 divise in 6 strati. In figura è possibile riconoscere l'organizzazione in blobs specializzati nel riconoscimento dei colori, gli interblob (che circondano i blob) e in colonne che rispondono a particolari orientazioni.

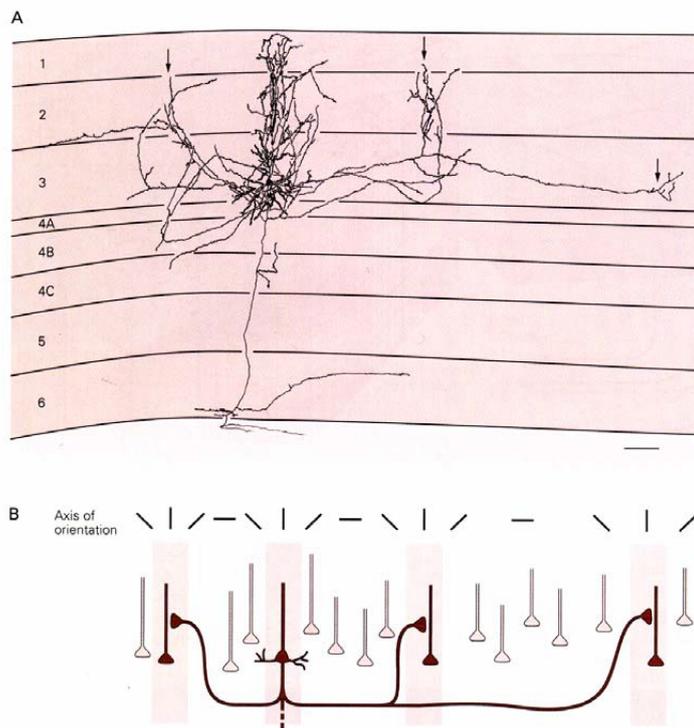


Figura 1.10: Rappresentazione dei collegamenti tra cellule: le parti con proprietà simili sono connesse tra loro tramite collegamenti orizzontali.

1.3 Il campo recettivo

Il *campo recettivo* (*CR*) è una caratteristica di tutte le cellule nervose ed è definito come la parte dell'area visiva all'interno della quale deve cadere lo stimolo per attivare la risposta del neurone interessato^[9]. Gli esperimenti tipici che consentono di determinare i profili dei campi recettivi consistono nel misurare l'attività elettrica del neurone, mediante un micro-elettrodo, mentre vengono presentati vari tipi di stimoli visivi in sedi differenti del campo visivo. Ad esempio, gli stimoli che ho utilizzato in questo lavoro sono sequenze di frames (filmati) ad alta dimensionalità mostrati a macachi; questi stimoli naturali hanno prodotto un'attività a livello della corteccia visiva primaria che è stata registrata attraverso l'utilizzo di micro elettrodi. Il concetto di campo recettivo è al centro dello studio delle risposte dei neuroni visivi, perchè caratterizza la trasformazione delle immagini in attività neurale. Ad eccezione dei fotorecettori e degli elementi delle prime stazioni retiniche, che rispondono semplicemente alla presenza o assenza di luce, la risposta dei neuroni delle vie più profonde (LGN e V1) sono più complesse. E' possibile infatti definire il CR sia in funzione delle sole coordinate spaziali sia come funzione di spazio e tempo.

La via ottica principale è costituita da unità organizzate in modo gerarchico e queste sono a loro volta costituite da strati di cellule intimamente connesse. Le cellule appartenenti al primo strato del sistema visivo sono eccitate in modo diretto dagli stimoli luminosi e i loro campi recettivi sono determinati esclusivamente dal modo in cui tali cellule interagiscono tra loro sullo stesso strato. Le cellule degli strati successivi, invece, sono eccitate dagli stimoli luminosi solo per via indiretta: lo stimolo viene alterato da elaborazioni compiute dagli strati precedenti. Dunque i campi recettivi di cellule appartenenti a strati successivi al primo, sono determinati sia dalle elaborazioni degli strati precedenti, sia dal modo in cui esse interagiscono tra loro. Salendo a livelli superiori dell'apparato visivo, i campi recettivi si sovrappongono, le loro dimensioni aumentano e lo stimolo efficace diviene progressivamente più complesso (da stimoli puntiformi a barre

luminose orientate sino a tessiture vere e proprie).

Una numerosa classe di neuroni corticali risponde solamente ad una barra luminosa o scura posta su uno sfondo contrastante; per un dato neurone la barra deve essere posta in una regione particolare del campo visivo e possedere una particolare orientazione rispetto all'asse orizzontale. Altri neuroni corticali, per rispondere, richiedono non solo che lo stimolo abbia una particolare orientazione, ma anche una particolare frequenza spaziale oppure che si muova in una determinata direzione o che possieda dimensioni peculiari^[8].

1.3.1 Organizzazione

Generalmente il campo recettivo non è omogeneo, ma è caratterizzato da sottoregioni "ON" e "OFF" la cui stimolazione luminosa separata ha effetti opposti sulla risposta cellulare. Un primo esempio viene dalle cellule gangliari della retina e dai neuroni del LGN; i loro campi recettivi hanno una forma arrotondata e sono organizzati in *centro-periferia*. E' possibile osservare due tipi di configurazioni:

- la prima, in cui il centro del campo recettivo risponde a stimoli luminosi è chiamata *On-Center*, mentre la periferia risponde a stimoli bui. La stimolazione della sottoregione ON con un fascio luminoso provoca un aumento nell'attività elettrica della cellula (depolarizzazione), mentre uno stimolo scuro provoca un decremento della stessa (iperpolarizzazione).

- l'altra, chiamata *Off-Center*, in cui le polarità sono complementari rispetto a quelle viste prima e in cui il centro del campo recettivo risponde a stimoli scuri, mentre la periferia a quelli luminosi (vedi figura 1.11). Il comportamento del neurone a seguito della stimolazione scura di una sottoregione OFF accresce la sua attività spontanea, mentre la comparsa della luce inibisce la risposta^[8].

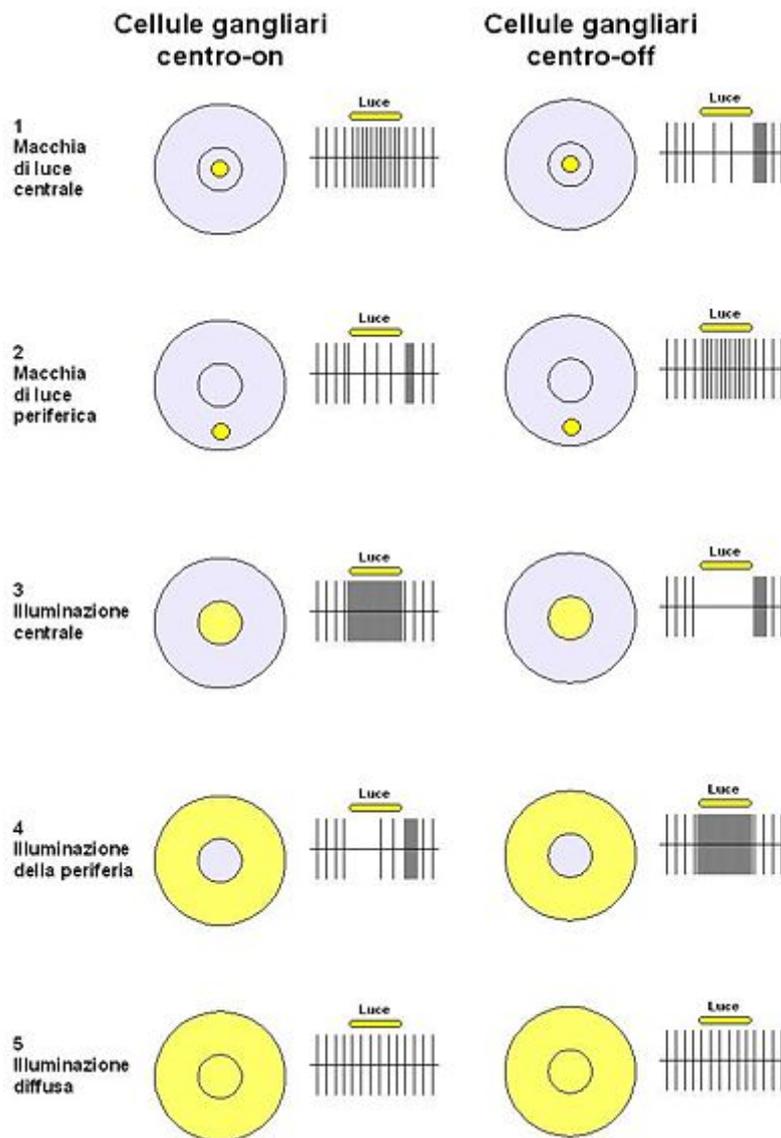


Figura 1.11: Numero di potenziali di azione al secondo (posti a destra dei profili recettori) prodotti per uno stesso tipo di stimolo luminoso che investe il profilo recettore di una cellula gangliare ON center e OFF center.

Ben più interessanti sono i campi recettivi delle cellule corticali, i quali non sono più circolari, ma tendono a diventare ellittici; di conseguenza lo stimolo migliore per attivare la cellula non è più un puntino luminoso ma una linea, una barra o un margine chiaro-scuro (edge). Inoltre emerge nella corteccia anche la necessità che lo stimolo luminoso abbia una determinata orientazione spaziale. Ciò vuol dire, ad esempio, che una linea per essere efficace deve avere un determinato orientamento rispetto alle coordinate principali. Così, se una linea posta ad esempio a $\pi/4$ rispetto alla verticale rappresenta uno stimolo efficace quando cade nel profilo recettore di una determinata cellula corticale, nel momento in cui la sua inclinazione viene cambiata a $\pi/6$ la stessa cellula corticale non sarà influenzata con la stessa forza. La preferenza per stimoli allungati e con un certo orientamento è una caratteristica comune delle cellule di V1. Ogni cellula ha un suo orientamento preferito e la sua capacità di rispondere diminuisce a mano a mano che l'orientamento dello stimolo si discosta da quello ottimale (fig. 1.13). La preferenza per stimoli allungati e con un certo orientamento è una caratteristica comune delle cellule dell'area 17. Hubel e Wiesel furono i primi a studiare questi tipi di cellule già nel 1965; il loro lavoro li portò a vincere il premio Nobel nel 1981. Essi distinsero le cellule della corteccia visiva in tre categorie, sulla base delle proprietà di risposta:

1. **semplici** la maggior parte delle quali ricevono l'informazione dai neuroni del LGN e hanno un'orientazione spaziale. Come mostrato in figura 1.12 presentano sub-regioni di forma allungata alternate che rispondono a stimoli luminosi o bui. E' possibile distinguere due o tre zone parallele On e Off ben definite, anche se non concentriche come nel LGN. Recentemente è stato dimostrato che i profili dei campi recettivi di molte cellule semplici presentano delle frange laterali in aggiunta alle due o tre zone centrali. Studiando più attentamente questa differenza di fase, si può osservare che nella maggior parte dei casi cellule adiacenti sono sfasate di 90 gradi.

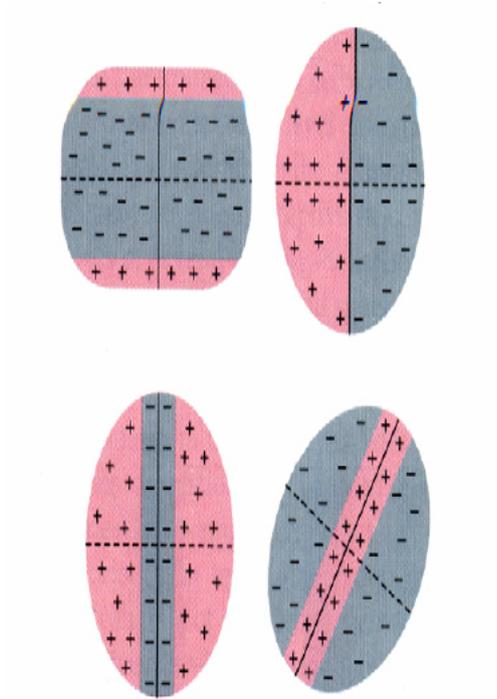


Figura 1.12: Esempio di campo recettivo di cellule semplici

Le cellule semplici possiedono zone antagoniste eccitatorie ed inibitorie simili a quelle presenti nel corpo genicolato o nelle cellule gangliari, anche se più grandi; esse presentano zone eccitatorie ed inibitorie di forma allungata e con un asse di allungamento caratteristico; sono in grado di rilevare l'inclinazione di un'immagine, posta in una certa zona del campo visivo^[8].

E' possibile predire la direzione preferenziale di movimento di una cellula semplice dal suo profilo spazio-temporale $x-t$. Inoltre stime della velocità di movimento possono essere misurate dall'inclinazione delle sub-regioni del profilo $x-t$ ^[9].

2. **complesse** sono l'altro tipo di cellule presenti in gran numero in corteccia. Rispondono a stimoli di luce in ogni punto all'interno del loro campo recettivo. Questo è formato da combinazioni non lineari di sub-unità simili a quelle delle cellule semplici^[9]. Le cellule complesse hanno CR in genere più grandi delle semplici, ma le zone On e Off non sono chiaramente separate. Inoltre la maggior parte delle complesse risponde in modo più marcato a stimoli in movimento, il quale deve avvenire in una certa direzione (selettività alla direzione di spostamento), come mostrato nella figura sottostante. Le cellule complesse rivelano la posizione di un oggetto indipendentemente dalla sua posizione nel campo visivo; rispondono infatti sia ad una linea orientata secondo una precisa direzione, sia ad una linea con lo stesso orientamento ma in movimento. Un certo numero di cellule complesse è poi in connessione con altre cellule, dette *ipercomplesse*^[8].

Questo tipo di cellule hanno un campo recettivo approssimabile con funzioni del secondo ordine, sono rappresentate attraverso la struttura di sub-unità combinate tra loro. Recenti studi hanno mostrato che questi profili danno accurata predizione su orientazione e frequenza spaziale di cellule complesse nelle scimmie^[9].

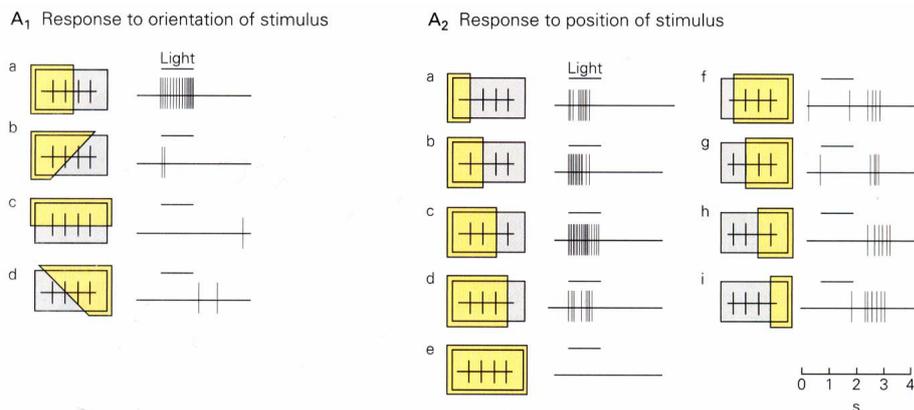


Figura 1.13: Esempio di campo recettivo di cellule complesse

3. *ipercomplesse* mentre le prime due si trovano in prevalenza nell'area 17 di Brodmann, queste sono in maggior numero nelle aree 18 e 19. Le cellule ipercomplesse si differenziano dalle precedenti perchè richiedono stimoli di una certa lunghezza e larghezza per una risposta ottimale. La selettività per la lunghezza o la larghezza di una linea o di una barra è dovuta alla presenza di porzioni inibitorie poste al di fuori del campo recettivo eccitatorio. Quando lo stimolo luminoso invade tali zone si ha una riduzione o abolizione della scarica. Recentemente, questo comportamento è stato osservato anche in certi tipi di cellule semplici e complesse e quindi il termine ipercomplesse tende a non essere più usato e si parla di cellule semplici e complesse con inibizione terminale (simple end-stopped e complex end-stopped)^[8].

Le cellule semplici ricevono informazioni da cellule retiniche gangliari che hanno i loro campi recettivi allineati, contraggono connessioni con le cellule del nucleo genicolato e, quando le cellule gangliari sono tutte attivate, provocano l'attivazione della cellula della corteccia primaria, che quindi risponde alla direzione dello stimolo (barretta orizzontale o verticale, per esempio). Si può allora pensare che, analogamente, le cellule complesse ricevano informazioni da più cellule semplici, che rispondono ad immagini che presentano la stessa orientazione ma presenti in posizioni diverse del campo visivo, in modo che la cellula complessa possa acquisire l'informazione circa la direzione del movimento della barretta così orientata.

1.3.2 Separabilità

L'organizzazione del campo recettivo non è statica; se esaminato nel dominio spazio-temporale, i campi recettivi di molte cellule presenti nel nucleo genicolato laterale e in corteccia mostrano un'impressionante dinamismo. Ma come spazio e tempo interagiscono tra di loro per determinare le proprietà della risposta cellulare? Esistono due possibilità: il CR può essere spazio-tempo *separabile* o *inseparabile*. Nel primo caso le tre dimensioni del campo recettivo ($R(x,y,t)$) possono venire descritte come il prodotto di due funzioni indipendenti: un profilo spaziale ($G(x,y)$) ed un profilo temporale ($H(t)$) (es: $R(x,y,t)=G(x,y)H(t)$). Se invece una cellula ha un campo recettivo spazio-tempo inseparabile, non è possibile dividerlo nelle due componenti. Per le cellule che hanno questo comportamento non è facile ottenere un accurato ritratto del campo recettivo. Per la cellula di figura 1.14 A, il campo recettivo spazio-tempo è separabile, ciò significa che la disposizione spaziale delle sub-regioni è fissata ma la loro intensità e la loro polarità sono modulate sul tempo. Per la cellula di figura 1.14 B, il campo recettivo spazio-tempo è inseparabile, in quanto l'organizzazione spaziale cambia col tempo. Questa disposizione persiste poi per il resto della durata della risposta della cellula. In questo modo è possibile predire la selettività di velocità di cellule semplici attraverso il loro profilo x-t, cosa impossibile partendo dalla loro mappatura On/Off^[9].

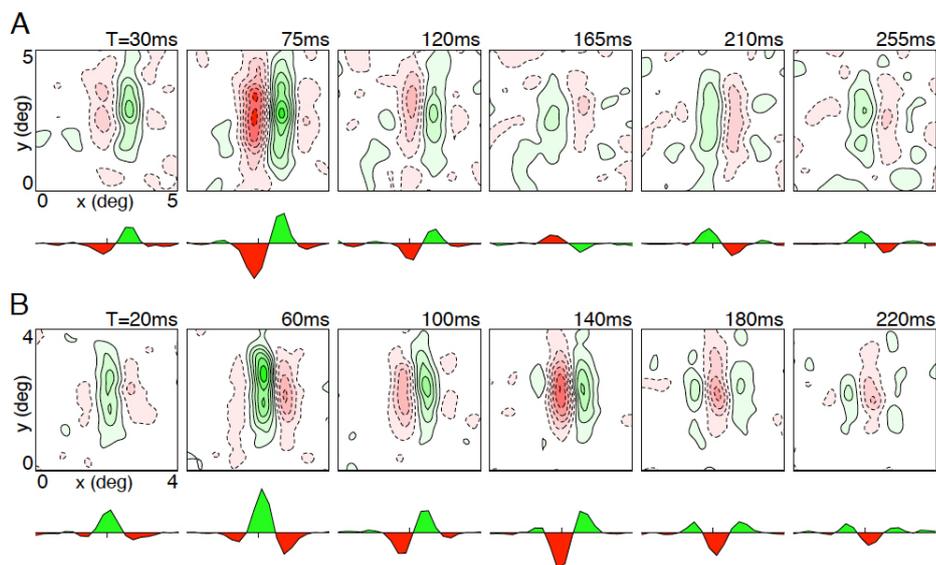


Figura 1.14: Dinamica della struttura del campo recettivo per due cellule prese dalla corteccia striata di gatto. Questi dati sono stati ottenuti utilizzando la tecnica della reverse correlation. Per ogni cellula, è mostrato il profilo spaziale 2-D (x - y), mappato mantenendo costante l'ampiezza per sei valori di tempo "t". Sotto ogni immagine è presente un profilo di campo recettivo 1-D ottenuto integrando il profilo 2-D lungo l'asse y , che è parallelo all'orientazione preferenziale della cellula. Spostamenti positivi nel profilo 1-D indicano sub-regioni eccitate da stimoli luminosi; spostamenti negativi, invece, corrispondono a sub-regioni eccitate da stimoli scuri. (A) Il campo recettivo di questa cellula semplice è approssimabile come spazio-tempo separabile. Da $T=30$ a 120 ms, il profilo recettore ha due sub-regioni dominanti che sono disposte con la regione eccitatoria a stimoli scuri sulla sinistra. Questa regione è molto marcata a $T=75$ ms. Tra $T=120$ e 165 ms, il CR cambia la sua polarità, così che ora sulla sinistra c'è la sub-regione sensibile a stimoli luminosi. Notare che, per tutti i valori di t , il campo recettivo 1-D è circa a simmetria dispari. (B) Un differente tipo di comportamento spazio-temporale è mostrato in figura. Per queste cellule il CR spazio-temporale è inseparabile, e questo significa che l'organizzazione spaziale del CR cambia nel tempo. A $T=20$ ms il profilo 1-D è circa a simmetria pari, mentre a $T=100$ ms, il profilo del CR è a simmetria dispari. Più tardi, a $T=180$ ms, è ancora a simmetria pari, ma il profilo è invertito rispetto a $T=20$ ms.

Capitolo 2

Ricostruzione di campi recettivi

In questo capitolo sarà prima introdotto un modello di *reverse correlation*. Attraverso questa tecnica è possibile estrapolare dati neuro-fisiologici (i profili recettori) dalle cellule della corteccia visiva primaria visti nel capitolo precedente. Saranno prima introdotti i modelli utilizzati per descrivere il comportamento dei neuroni della V1 e poi analizzate la *spike-triggered average* e la *spike-triggered covariance*, varianti della tecnica della reverse correlation.

Per rendere più chiari questi concetti, saranno ricostruiti e mostrati campi recettivi grazie all'elaborazione degli stimoli naturali messi a disposizione da De Angelis e il suo team^[28]. Infine sarà introdotto il data set delle registrazioni elettrofisiologiche effettuate da Ringach^[31], su cui si baserà questo lavoro di tesi e sui quali sarà effettuato il *clustering*, tecnica attraverso la quale verranno suddivisi e raggruppati gli spike appartenenti ad uno stesso neurone. Attraverso questa suddivisione è possibile poi risalire a quali stimoli hanno prodotto una risposta da parte di questa cellula.

2.1 Campo recettivo e risposta neurale

Come si ottengono i campi recettivi visti nel capitolo 1? Qual'è il metodo che permette di ricavare i profili recettori delle cellule in corteccia visiva primaria?

Per ottenere questi dati neuro-fisiologici viene utilizzata una tecnica chiamata “reverse correlation”. Prima di descriverla però, bisogna comprendere come modellare la risposta neurale delle cellule appartenenti al sistema visivo e quale teoria sta alla base di questa tecnica.

I neuroni appartenenti a percorsi del sistema sensoriale, compresi quelli nella corteccia visiva primaria, possono essere visti come una sorta di trasduttori stimolo-risposta che operano su segnali provenienti dal mondo esterno. In assenza di stimolazione un neurone è generalmente silente, mentre quando uno stimolo con le giuste caratteristiche è proiettato all'interno del campo recettivo, l'attività neurale viene incrementata o soppressa. Quando la stimolazione cessa, la cellula torna al suo stato naturale. E' ovvio che questa breve descrizione non può essere usata per esprimere la relazione stimolo-risposta di tutte le cellule corticali.

I trasduttori più semplici con i quali è possibile lavorare sono quelli lineari, con un unico ingresso ed una sola uscita. Trasduttori di questo tipo sono generalmente chiamati LSISOS (Linear Single-Input, Single-Output Systems). Essi rispondono nel tempo ad un breve impulso di area unitaria con una risposta $h(t)$, chiamata *risposta all'impulso*. Un sistema è caratterizzato dalla sua risposta all'impulso e, se il sistema è lineare, conoscere la risposta all'impulso significa conoscere tutto sul suo funzionamento. Quindi conoscendo la risposta all'impulso del sistema è possibile prevedere come il sistema risponderà alla presentazione di un qualsiasi stimolo arbitrario.

Ogni stimolo può essere decomposto in una serie di brevi impulsi di differente ampiezza e a differenti istanti di tempo; è sicuramente corretto dire che la risposta di un sistema LSISO può essere predetto semplicemente sommando molte risposte all'impulso propriamente scalate in ampiezza e traslate nel tempo (Fig. 2.1). Questa sommatoria, chia-

mata in modo più comune “convoluzione”, è lo strumento principale per costruire la risposta neurale a fronte di qualunque stimolo visivo. In questo caso, è chiaro che conoscendo la risposta all’impulso $h(t)$ di un sistema lineare, significa avere una descrizione completa di come il sistema lavora^[35].

Ma com’è possibile misurare la risposta all’impulso di un sistema LSI-SO con ingresso $S(t)$ e output $r(t)$?

Esistono diversi metodi ma per ragioni pratiche e teoriche è conveniente calcolare la risposta all’impulso del sistema attraverso la funzione di cross-correlazione tra ingresso e uscita. Questa funzione, per sistemi stazionari, è definita come:

$$C_{sr}(t') = \langle S(t)r(t+t') \rangle$$

dove $\langle \cdot \rangle$ rappresenta la media nel tempo. In ogni istante di tempo è calcolato il prodotto tra lo stimolo $S(t)$ e la risposta che questo evocerà $r(t+t')$. Facendo la media di tutti i risultati si ottiene facilmente la correlazione tra stimoli e risposta per un determinato offset di tempo t' . Un metodo equivalente per calcolare la funzione di cross-correlazione è definire una variabile temporale $u=t+t'$ e, per ogni istante u , calcolare il prodotto tra la risposta $r(u)$ e lo stimolo che lo precede di t' cioè $S(u-t')$. I due algoritmi guardano il problema da due posizioni differenti ma sono assolutamente equivalenti e portano agli stessi risultati. In un caso vengono considerati prima gli stimoli, nell’altro viene privilegiata la risposta.

$$C_{sr}(t') = \langle S(t)r(t+t') \rangle = \langle S(u-t)r(u) \rangle$$

Per ragioni puramente computazionali il secondo metodo viene generalmente preferito. Questa maniera di calcolare la funzione di cross-correlazione di stimolo-risposta viene chiamata *reverse correlation*.

I neuroni della corteccia visiva possono essere visti come un trasduttore lineare a multipli ingressi, uno per ogni porzione di spazio all’interno del campo recettivo, ed un’unica uscita. Sistemi di questo tipo vengono definiti LMISOS (Linear Multi Input, Single Output Sy-

stem) e sono completamente caratterizzati dalla risposta all'impulso in tre dimensioni $h(x, y, t)$. Così, conoscendo la risposta all'impulso spazio-temporale, è possibile predire l'attività neurale in risposta ad un qualsiasi stimolo spazio-temporale.

Sfortunatamente però un neurone che viene eccitato non può essere considerato un puro sistema LMISOS. La funzione che identifica il "firing rate" di un neurone è infatti fortemente non lineare. Uno dei meccanismi più utilizzati per modellare la convoluzione del campo recettivo con uno stimolo passa attraverso un generatore casuale di Poisson che modella la generazione di spike.

Per ricostruire il campo recettivo di un neurone visivo esistono diversi metodi raggruppati sotto la denominazione di "reverse correlation" come la *spike-triggered average* (STA) e la *spike-triggered covariance* (STC) che saranno studiate nelle prossime sezioni. E' dimostrato^[40] che se l'input ha distribuzione statistica simmetrica, questi metodi riescono a ricostruire il campo recettivo a prescindere dalla non linearità rappresentata dai meccanismi di generazione di spike. Questi metodi si sono dimostrati estremamente utili per caratterizzare le funzioni di base dei sistemi sensoriali; essi richiedono che lo stimolo venga tratto da una distribuzione a simmetria sferica per poter produrre una stima imparziale del campo recettivo. Anche se questo vincolo può essere soddisfatto da stimoli artificiali (come il rumore bianco gaussiano), è violato dalle correlazioni e asimmetrie tipicamente presenti negli stimoli naturali e così la stima dei campi recettivi attraverso la tecnica della reverse correlation, sotto determinate condizioni, permette di calcolare le risposte a stimoli naturali preventivamente.

Recenti studi^[35] suggeriscono che esistono fondamentali differenze tra risposte neurali a stimoli artificiali rispetto a stimoli naturali. Gli stimoli naturali, per esempio, sono codificati in modo più efficiente rispetto a quelli artificiali sia nel sistema visivo che in quello uditivo. In più, è evidente che modelli di elaborazione sensoriale derivanti da risposte a stimoli artificiali non sono sufficienti per prevedere la risposta neurale a stimoli naturali.

Inoltre stimoli naturali visivi e uditivi hanno proprietà statistiche mol-

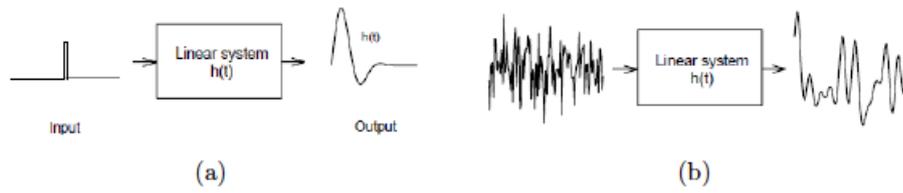


Figura 2.1: (a) La risposta all'impulso può essere misurata usando un breve impulso di area unitaria come input. (b) E' possibile ricostruire la risposta all'impulso $h(t)$ cross-correlata ad un ingresso di rumore bianco con il corrispondente output.

to complesse. Per esempio, gli stimoli naturali prestano tipicamente una forte correlazione tra spazio e tempo. Sfortunatamente queste proprietà che distinguono le due tipologie di stimoli, complicano l'uso di risposte neurali a stimoli naturali nella ricerca del campo recettivo spazio-temporale del neurone.

2.2 Spike-triggered average (STA)

Nella sezione precedente è stato mostrato com'è possibile modellare la risposta neurale agli stimoli visivi. Questi inducono risposte di neuroni in tutto il percorso sensoriale, per questo il loro utilizzo offre opportunità per comprendere la codifica sensoriale. Per neuroni aventi una relazione risposta/stimolo di tipo lineare, caratteristiche visive rilevanti possono essere identificate calcolando i loro campi recettivi attraverso il metodo della STA.

La *spike-triggered average* (STA) è uno strumento per caratterizzare le proprietà di risposta di un neurone usando gli spike emessi in risposta ad uno stimolo variabile nel tempo. La STA fornisce una stima del campo recettivo lineare di un neurone.

Matematicamente, la STA è la media degli stimoli che precedono uno spike. Per calcolarla, vengono estratti gli stimoli dalla finestra tem-

porale che precede ogni spike e ne viene calcolata la media (figura 2.2). La STA fornisce una stima non distorta del campo recettivo di un neurone solo se la distribuzione dello stimolo è a simmetria sferica (per esempio, rumore bianco gaussiano).

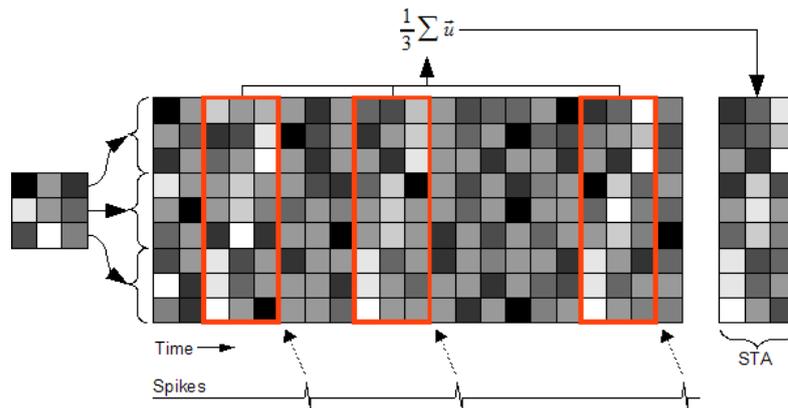


Figura 2.2: La figura mostra come viene calcolata la STA. Per ogni stimolo presentato (formato da 9 valori di pixel per ogni fotogramma), vengono registrate le risposte del neurone. Gli stimoli in un certo intervallo di tempo che precede ogni spike (qui formati da 3 pixel) sono selezionati (riquadri rossi) e poi mediati per ottenere la STA. La STA (mostrata a destra) indica che il neurone è selettivo per un gruppo di 3 pixel bianchi che cambiano posizione spaziale su ciascuno dei tre fotogrammi consecutivi.

Lo scopo di questa tecnica è di caratterizzare la relazione tra stimoli e risposta neurale. La teoria su cui si basa la spike-triggered average assume che la probabilità che un neurone si ecciti è governata solo dagli stimoli più recenti. Più precisamente il modello di risposta è assunto tale che la *firing rate* è una funzione degli stimoli presentati durante una finestra temporale di durata fissata. L'analisi di dati sperimentali è ridotta nell'esaminare le proprietà degli stimoli all'interno di una finestra temporale precedente ad ogni registrazione di spike è conosciuta come "spike-triggered stimulus ensemble"^[23]. E' possibile rappresentare l'intero insieme di tutti gli stimoli e quello degli stimoli che hanno eccitato la cellula come due nuvole di pun-

ti nello spazio. Il compito di stimare la funzione di risposta neurale equivale a descrivere il modo in cui queste due nuvole differiscono (fig. 2.3).

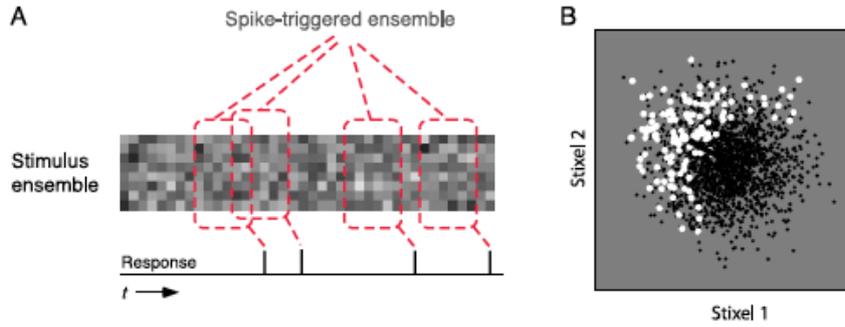


Figura 2.3: (A) Sequenza di stimoli discretizzati e osservazione della risposta neurale (treno di spike). Ad ogni istante di tempo lo stimolo è rappresentato da un array di valori random. Questo può essere rappresentato, per esempio, variando l'intensità di ogni pixel sullo schermo. La risposta neurale in ogni istante di tempo risulta essere completamente determinata da stimoli (segmenti) accorsi durante uno specifico tempo nel passato. In figura, il segmento copre sei step. La spike-triggered ensemble consiste nel set di segmenti associati allo spike. (B) I segmenti di stimoli spike-triggered (punti bianchi) costituiscono un sottogruppo dei segmenti di tutti gli stimoli presentati (punti neri).

La semplice deviazione tra la spike-triggered e la distribuzione degli stimoli naturali è un cambio nella media. Assumendo che gli stimoli naturali abbiano media zero, questa può essere stimata calcolando la media degli spike-triggered ensemble (STA):

$$STA = \frac{1}{N} \sum_{n=1}^N s(t_n) \quad (2.1)$$

dove t_n è il tempo dell' n -esimo spike, $s(t_n)$ è un vettore che rappresenta gli stimoli presentati durante la finestra temporale precedente a questo istante di tempo e N è il numero totale di spike. Praticamente

il tempo t_n fa da contenitore; se due o più finestre temporali (box rossi) sono sovrapposti, allora il vettore degli stimoli per quell'intervallo sarà moltiplicato per il numero di spike che si sono presentati. In figura 2.4 è illustrato il metodo della STA.

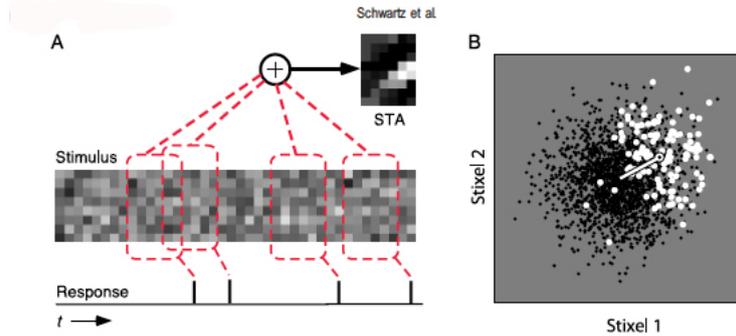


Figura 2.4: Due illustrazioni alternative di STA. (A) La STA è costruita dalla media dei segmenti degli stimoli della spike-triggered (box rossi in figura), sottraendogli la media di tutti i set di segmenti degli stimoli. (B) Rappresentazione geometrica (vettore spazio) della spike-triggered average in due dimensioni. I punti neri indicano gli stimoli naturali, quelli bianchi indicano stimoli che hanno eccitato la cellula. La STA, indicata con una linea in figura, corrisponde alla differenza tra la media (centro di massa) della spike-triggered ensemble e la media dell'insieme di tutti gli stimoli naturali.

2.3 Spike-triggered covariance (STC)

Nella corteccia visiva però la maggior parte dei neuroni sono delle cellule complesse, con una relazione risposta/stimolo di tipo non lineare. Per questo motivo con la STA non sarebbe possibile caratterizzare questo tipo di comportamento. Bisogna allora utilizzare un metodo alternativo, come la *spike-triggered covariance* (STC).

Questa tecnica confronta la varianza lungo diverse dimensioni nello spazio d'ingresso di tutti gli stimoli e di tutti gli stimoli che hanno provocato una risposta. L'analisi con spike-triggered covariance è uno strumento per caratterizzare le proprietà di risposta di un neurone utilizzando la covarianza di stimoli che hanno eccitato il neurone. La STC è legata alla STA ma, diversamente, può essere utilizzata per identificare una caratteristica spaziale multidimensionale di un neurone. La spike-triggered covariance identifica le caratteristiche dello stimolo che interessano la risposta di un neurone attraverso una decomposizione degli autovettori della matrice di covarianza spike-triggered. Autovettori significativamente grandi o piccoli di autovalori della covarianza di stimoli naturali corrispondono alla risposta neurale migliore o soppressa.

Assumendo che gli stimoli naturali abbiano covarianza sferica è possibile calcolare la matrice della STC in questo modo:

$$STC = \frac{1}{N-1} \sum_{n=1}^N [s(t_n) - STA][s(t_n) - STA]^T \quad (2.2)$$

dove T indica il trasposto del vettore. Ancora una volta t_n rappresenta un contenitore e questo significa che ogni termine deve essere moltiplicato per il numero di spikes che si sono presentati nella finestra temporale associata.

La matrice STC rappresenta la struttura multidimensionale della varianza relativa alla spike-triggered ensemble. In modo specifico, la varianza di tutti gli stimoli (ensemble) in ogni direzione specificata da un vettore unitario \hat{u} è semplicemente $\hat{u}^T C \hat{u}$. La superficie spazzata

da tutti i vettori unitari, scalata per la radice quadrata della loro varianza è un ellissoide multidimensionale. Gli assi principali di questo ellissoide, assieme alle rispettive varianze, possono essere trovati come gli autovettori e gli autovalori associati alla matrice *STC*. E' possibile vedere questo in figura 2.5. Solitamente la *STA* viene sottratta agli stimoli prima di calcolare la *STC*.

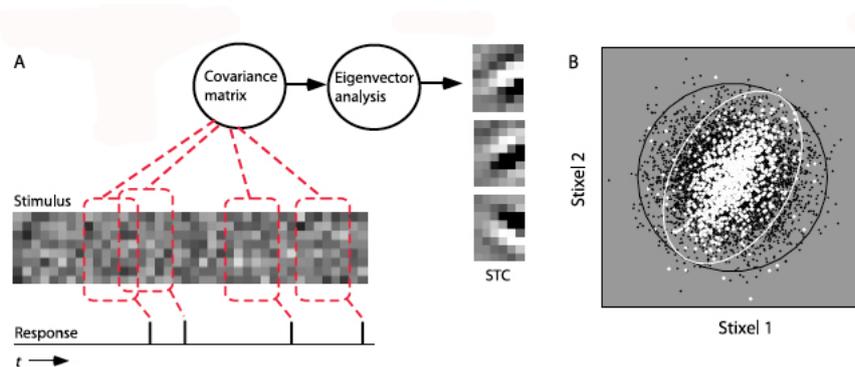


Figura 2.5: Due illustrazioni di *STC*. (A) L'*STC* è calcolata trovando la covarianza della spike-triggered degli stimoli naturali e poi facendo l'analisi degli autovettori della matrice di covarianza. (B) Rappresentazione geometrica della *STC*. I punti neri rappresentano gli stimoli naturali, quelli bianchi stimoli che hanno provocato uno spike. L'ellisse rappresenta la covarianza di tutto l'insieme. La distanza dall'origine dell'ellisse lungo ogni direzione è la deviazione standard di tutti gli stimoli in quella direzione. Gli stimoli naturali sono distribuiti in modo circolare (Gaussiana). Gli stimoli spike-triggered sono distribuiti su un'ellisse con una minor varianza (relativa a stimoli naturali) lungo l'asse minore. Questa direzione è meno significativa, in quanto stimoli che hanno grandi componenti lungo questa direzione hanno minor probabilità di generare uno spike.

Una variante della STC è rappresentato dalla *reverse correlation*. Anche questa tecnica serve per studiare come i neuroni sensoriali sommano i segnali provenienti da diverse posizioni dei loro campi recettivi, e anche il modo in cui valutare stimoli presentati in tempi diversi e che hanno generare una risposta. Questa tecnica si basa proprio sui metodi della STA e della STC visti precedentemente. Per utilizzare la reverse correlation come primo passo si deve calcolare la matrice di covarianza “C” relativa agli spike che hanno provocato una risposta (C^{spike}) e la matrice di covarianza di tutti gli stimoli presentati (C^{stim}):

$$C_{ij}^{spike} = \frac{1}{N_{spike} - 1} \sum_{t=1}^{N_{spike}} (s(t)_i - STA_i^{spike})(s(t)_j - STA_j^{spike}) = STC_{ij}^{spike} \quad (2.3)$$

$$C_{ij}^{stim} = \frac{1}{N - 1} \sum_{t=1}^N (s(t)_i - STA_i^{stim})(s(t)_j - STA_j^{stim}) = STC_{ij}^{stim} \quad (2.4)$$

dove N_{spike} è il numero di spike registrati, N è il numero di stimoli (frames) mostrati, $s(t)_i$ è il valore degli stimoli lungo la i -esima direzione al tempo t .

Il secondo step è quello di calcolare la differenza tra le due matrici di covarianza e cercare gli autovalori che sono significativamente differenti da zero:

$$\Delta C = C^{spike} - C^{stim} \quad (2.5)$$

Per determinare significatività statistiche degli autovalori bisogna confrontarli con la distribuzione nulla, che è la distribuzione degli autovalori della matrice:

$$\Delta C^{null} = C^{null} - C^{stim} \quad (2.6)$$

Le matrici ΔC^{null} e C^{null} sono formate assumendo che non vi siano associazioni tra gli stimoli e la risposta neurale. Per fare questo, per esempio, è possibile utilizzare spikes presi in modo random nel tempo con la stessa frequenza trovata per i neuroni reali (dall'esperimento di Ringach^[31]). Gli autovalori significativi di ΔC possono essere sia positivi che negativi.

L'ultima parte prevede il calcolo della STC.

2.3.1 Utilizzo della STC su stimoli naturali

I dati naturali possono presentare un certo grado di correlazione tra loro. Questa caratteristica può generare alcuni problemi nel calcolo della reverse correlation; è quindi consigliato, prima di applicare tale tecnica, di effettuare una trasformazione di *whitening* cioè una trasformazione di decorrelazione che trasforma un insieme di variabili casuali, aventi una matrice di covarianza M nota, in una serie di nuove variabili aleatorie la cui covarianza è la matrice identità (questo significa che le variabili sono incorrelate e hanno tutte varianza 1).

La trasformazione si chiama "sbiancamento" perchè trasforma il vettore d'ingresso in un vettore di rumore bianco. Si differenzia da una trasformazione di decorrelazione generale in quanto, quest'ultima, ha covarianza uguale a zero, in modo che la matrice di correlazione possa essere qualsiasi matrice diagonale.

Quindi prima di effettuare la reverse correlation, vogliamo che i dati siano decorrelati. Un metodo per rimuovere la correlazione tra stimoli dagli autovettori di ΔC è quello del *pre-whitening*. Con questa tecnica è possibile rimuovere la correlazione da ognuno dei vettori degli stimoli prima di diagonalizzare ΔC .

La semplice matrice di covarianza degli stimoli può essere scritta in termini di autovalori λ_n e autovettori $f^{(n)}$ come:

$$C_{ij}^{stim} = \sum_{n=1}^k \frac{1}{\lambda_n} f_i^{(n)} f_j^{(n)} \quad (2.7)$$

E' possibile allora definire la matrice $C_{ij}^w = \sum_{n=1}^p \lambda_n^{-1/2} f_i^{(n)} f_j^{(n)}$. In questo modo l'analogo di ΔC diventa:

$$C^\Delta = C^w C^{spike} C^w \quad (2.8)$$

Questa procedura equivale a “sbiancare” ognuno dei frame usati come stimolo in modo indipendente (attraverso la moltiplicazione con C^w) e poi calcolando la spike-triggered covarianza.

Nel limite di dati infiniti, l'ipotesi nulla corrisponde a $C^{spike} = C^{stim}$. In questo caso $C^\Delta = []$. Per un data set di dimensioni finite, la distribuzione nulla viene calcolata da molte realizzazioni della matrice

$$C_{null}^\Delta = C_w C^{null} C_w \quad (2.9)$$

in cui C^{null} è definita dall'equazione 2.6. La maggior parte degli autovalori di C^Δ hanno valori vicini a 1^[27].

Un metodo alternativo per rimuovere la correlazione tra gli stimoli prevede di moltiplicarli per la pseudo-inversa di C^{stim} . Questo approccio, che cerca le dimensioni rilevanti degli stimoli, è equivalente a cercare gli autovettori della seguente matrice:

$$\begin{aligned} \Delta \tilde{C}_{ij} = & \frac{1}{N_{spike} - 1} \sum_{t=1}^{N_{spike}} (s(t)_i - STA_i^{stim})(s(t)_j - STA_j^{stim}) - \\ & \frac{1}{N - 1} \sum_{t=1}^N (s(t)_i - STA_i^{stim})(s(t)_j - STA_j^{stim}) \end{aligned} \quad (2.10)$$

che è riconducibile a:

$$\Delta \tilde{C}_{ij} = STC_{ij}^{spike} - STC_{ij}^{stim} \quad (2.11)$$

La matrice $\Delta \tilde{C}_{ij}$ descrive un cambio tra la distribuzione di stimoli che provocano uno spike e tutti gli stimoli presentati e poi le viene sottratta la media degli stimoli STA^{stim} . Nonostante $\Delta C \neq \Delta \tilde{C}$, i

loro autovettori coincidono.

$$C_{ij}^{-1}(k) = \sum_{n=1}^k \frac{1}{\lambda_n} f_i^{(n)} f_j^{(n)} \quad (2.12)$$

2.4 Ricostruzione di campi recettivi

Per cercare di comprendere meglio i concetti visti finora, in questa sezione sarà descritto il lavoro svolto a partire dai dati messi a disposizione dal team Izumi Ohzawa, Gregory C. De Angelis e Ralph D. Freeman^[29]. Lo scopo del loro lavoro era lo studio di *campi recettivi* spazio-temporali valutati per cellule semplici della corteccia striata di gatto partendo da stimoli artificiali che verranno descritti in seguito. Per ottenere questa mappatura, la tecnica di *reverse correlation* che è stata utilizzata è la STA.

Gli stimoli utilizzati erano sequenze random di barre luminose o più scure, presentate una alla volta in diverse posizioni del campo recettivo. Il posizionamento della barra è stato scelto in modo random su una griglia bi-dimensionale 20x20. Gli assi della griglia sono stati orientati a posteriori seguendo l'orientazione del campo recettivo, in modo che visivamente l'orientazione fosse sempre verticale ($\theta = \pi/2$). Per ogni presentazione, il valore della barra (luminoso o scuro) è stato scelto in modo random. Per ogni sequenza degli stimoli, barre chiare e scure sono state presentate una sola volta in ogni punto della griglia e per ogni presentazione, il valore delle barre è stato anch'esso scelto in modo random. Gli esperimenti sono stati ripetuti per diverse volte. Nella figura 2.6 è mostrata una sequenza di stimoli e l'andamento spazio-temporale del campo recettivo ricostruito. La parte alta della figura rappresenta un segmento di stimoli, contenente circa 40 stimoli random consecutivi. Questa sequenza può essere pensata come una serie nel tempo in due dimensioni (x, y) e mostrata in figura da una successione di "lastre" ordinate da destra a sinistra. Ogni placca corrisponde ad un singolo stimolo mostrato, contenente uno stimolo (barra

luminosa o scura) che si trova all'interno della griglia 20x20. Alcuni di questi stimoli cadono nella regione eccitatoria del campo recettivo ed eccitano la cellula. Un treno di spike è disegnato in figura, poco prima della sequenza di stimoli^[29].

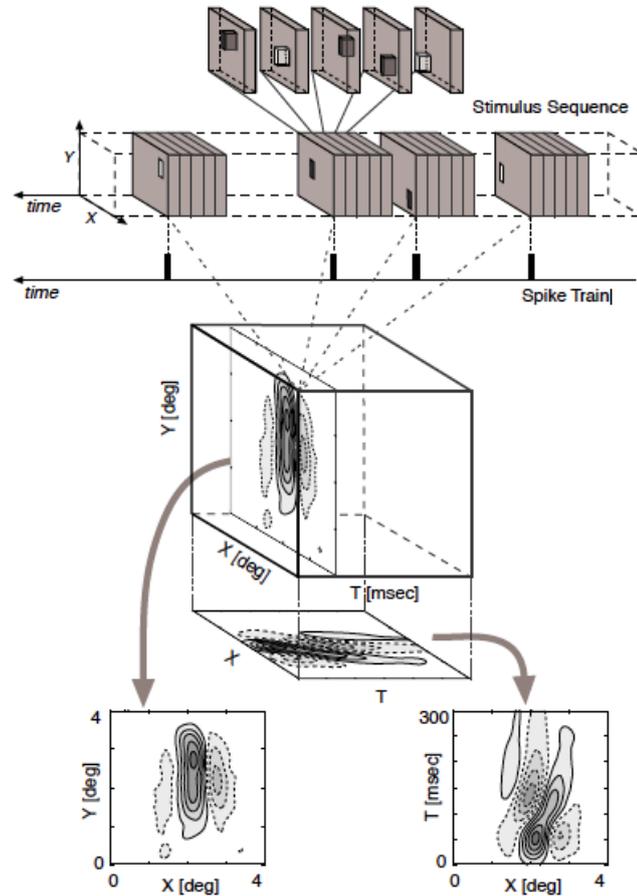


Figura 2.6: Tecnica della reverse correlation usata per misurare i campi recettivi spazio-temporali di cellule semplici. Alcune sequenze di stimoli sono illustrate nella parte alta della figura. La parte centrale è mostrata come “esplosione” per evidenziare gli stimoli separatamente: la loro durata è di circa 30-50 ms. Ogni stimolo è una barra scura o luminosa presentata su una griglia 20x20 di color grigio. La posizione e il valore della barra sono presi in modo random. La tecnica della reverse correlation può essere utilizzata per ottenere una media dei profili degli stimoli che hanno causato lo sparo della cellula. Questa è ottenuto sommando sequenze di stimoli che precedono ogni spike, per tutti gli spike generati. Il risultato è una mappa tridimensionale di x , y e t (due coordinate spaziali ed una temporale) che indica l’efficacia di uno stimolo in posizione x , y a causare lo sparo della cellula “ t ” ms dopo la presentazione dello stimolo. Questa è proprio la definizione di campo recettivo in spazio e tempo. Sul fondo dell’immagine sono mostrate gli stessi profili solo nello spazio (cross-sezioni).

Per ogni spike generato, si deve quindi guardare indietro per trovare quali stimoli lo hanno generato. Ma quanto tempo indietro bisogna guardare? Tipicamente per cellule nell'area 17 di gatti adulti (come per questo esperimento), solo gli stimoli presentati qualche centinaia di millisecondi prima contribuiscono alla risposta^[29]. Notare che gli stimoli sono presentati continuamente nella parte alta della figura 2.6, ma non sono mostrati per chiarezza.

Gli stimoli nella parte centrale della figura 2.6 sono sommati tra loro; sono presenti solo quegli stimoli che hanno generato una risposta della cellula, valutando +1 in caso di barra (stimolo) luminoso, -1 in caso di stimolo scuro e per lo sfondo. Ogni presentazione di spike ha una durata limitata. Questo produce una media del profilo degli stimoli (dopo aver diviso per il numero di cubi sommati) che hanno eccitato il neurone. Il profilo di questa media è disegnato all'interno del cubo (x , y , t) al centro della figura 2.6. L'elemento al suo interno rappresenta l'insieme di stimoli che hanno contribuito alla generazione di uno spike; per esempio uno stimolo presentato "t" millisecondo prima nella posizione spaziale (x , y) contribuisce alla generazione di spike dalla quantità indicata dal valore dell'elemento nel dominio x , y , t .

Siccome è difficile mostrare dati tridimensionali, è mostrata la cross-sezione o la proiezione dell'immagine in due dimensioni. Per esempio la cross-sezione presa al tempo t è una mappa spaziale del CR a quell'istante, ed è mostrata sulla parte bassa della figura 2.6. La cross-sezione lungo la dimensione y da il campo recettivo spazio temporale (x - t), come mostrata nella parte in basso a destra della figura 2.6^[29].

Il lavoro che ho svolto sui dati raccolti dal team di De Angelis si è sviluppato a partire da due tipologie di file: il primo di tipo informativo, in cui erano memorizzate tutte le condizioni sperimentali e i valori usati per ogni stimolazione; l'altro invece, conteneva la registrazione dell'evento. Per distinguere i vari esperimenti tra loro, i segnali registrati sono stati divisi seguendo tre criteri: il numero seriale dell'animale, il codice di sperimentazione e il numero di cellule.

Il pacchetto dati messo a disposizione da DeAngelis e il suo team^[28] comprendeva un software (RCAnalyzer) a cui devono essere presentati

in ingresso i dati relativi ad un preciso esperimento e, automaticamente, viene calcolata la *reverse correlation* restituendo in uscita, per ogni istante di campionamento (15 ms), due matrici 20x20, una relativa al campo recettivo eccitatorio ed una relativa a quello inibitorio. Partendo da queste due matrici ne ho calcolata una terza come differenza tra quella eccitatoria e quella inibitoria, che sono visualizzate nelle prossime figure agli istanti di campionamento da $t=15$ ms fino a 90 ms.

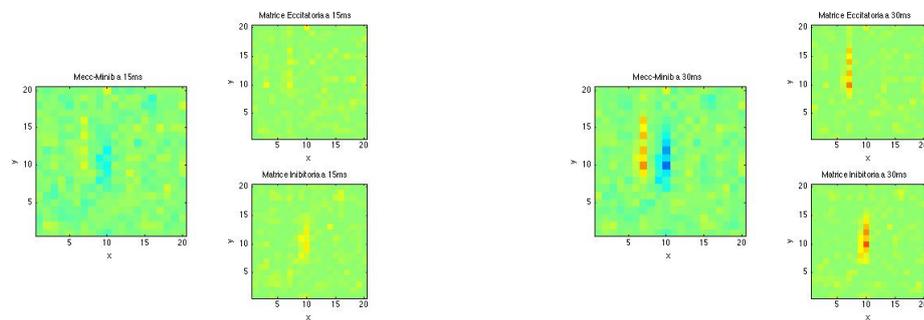


Figura 2.7: Le due immagini fanno riferimento a due istanti di campionamento differenti: $t=15$ ms (a sinistra) e $t=30$ ms (a destra). All'interno di queste si possono vedere: a sinistra il campo recettivo risultante dalla differenza tra quello eccitatorio e quello inibitorio che sono mostrati in due immagini separate nella parte destra sopra e sotto. Dati relativi all'esperimento "kd121r17.02e.0".

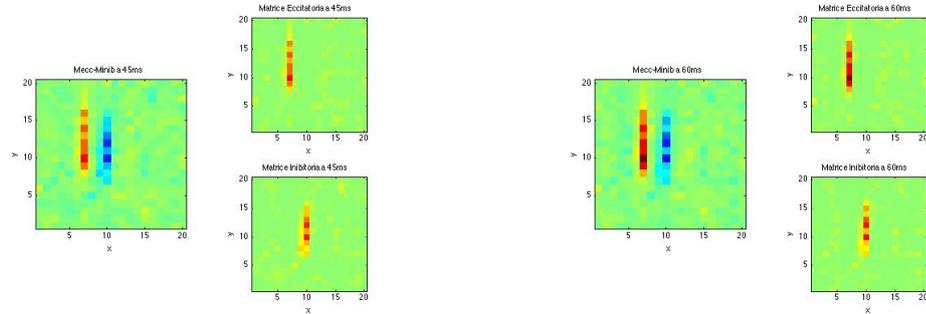


Figura 2.8: Le due immagini fanno riferimento a due istanti di campionamento differenti: $t=45$ ms (a sinistra) e $t=60$ ms (a destra). All'interno di queste si possono vedere: a sinistra il campo recettivo risultante dalla differenza tra quello eccitatorio e quello inibitorio; questi sono mostrati in due immagini separate nella parte destra sopra e sotto. Dati relativi all'esperimento "kd121r17.02e.0".

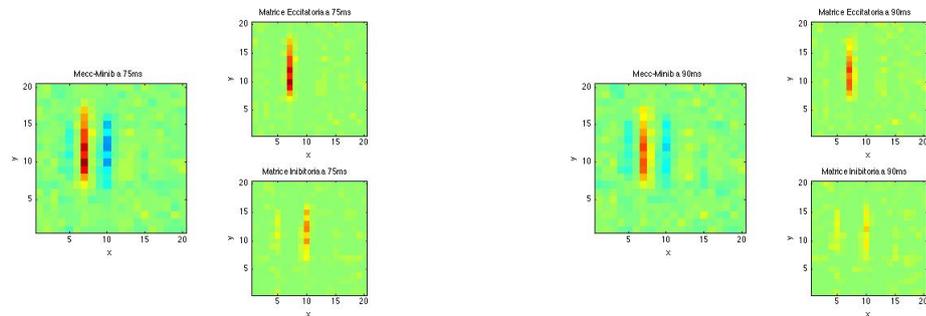


Figura 2.9: Le due immagini fanno riferimento a due istanti di campionamento differenti: $t=75$ ms (a sinistra) e $t=90$ ms (a destra). All'interno di queste si possono vedere a sinistra il campo recettivo risultante dalla differenza tra quello eccitatorio e quello inibitorio che sono mostrati in due immagini separate nella parte destra sopra e sotto. Dati relativi all'esperimento "kd121r17.02e.0".

Nella fase successiva ho studiato i valori di intensità dei campi recettivi. E' possibile distinguere le due regioni On e Off descritte precedentemente. In figura 2.10 si distinguono chiaramente le due regioni, a sinistra mostrate per valore di intensità (x, y, i) , mentre a destra sono visualizzate nello spazio (x, y) . In figura 2.11 invece, sono rappresentate le stesse regioni però come isosuperfici. Anche in questo caso è facile riconoscerle. Le isosuperfici sono superfici tridimensionali che uniscono i punti aventi lo stesso valore all'interno di un volume di spazio. Di seguito si possono osservare le due figure appena descritte per $t=60$ ms.

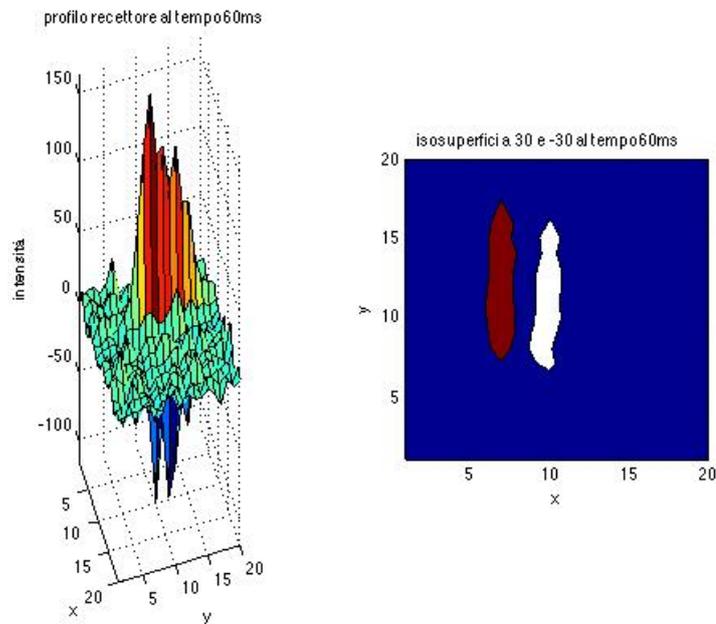


Figura 2.10: Visualizzazione dei campi recettivi (eccitatorio e inibitorio) all'istante di campionamento $t=60$ ms. A sinistra sono visualizzati per valori di intensità. Valori positivi corrispondono alla parte eccitatoria, valori negativi a quella inibitoria. A destra lo stesso campo recettivo visualizzato nello spazio (x, y) .

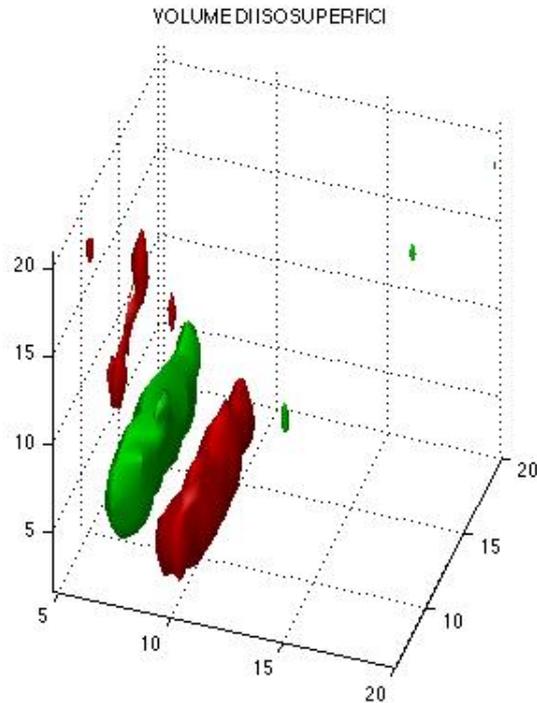


Figura 2.11: Isosuperfici che rappresentano l'intero profilo spazio-temporale.

Infine ho analizzato i campi recettivi nel dominio spazio-tempo $x-t$ (gradi visivi-secondi). Per ottenere questo e un campo recettivo mono dimensionale ho integrato lungo "y" la funzione spazio-temporale (x, y, t); il profilo ottenuto rappresenta lo spostamento delle sotto-regioni (eccitatoria ed inibitoria) in pixel nel tempo. E' possibile notare che le due regioni sono inclinate verso destra (la freccia rossa nella figura 2.12 indica le direzione di spostamento). Dal grafico è possibile risalire alla variazione infinitesima di velocità con cui si muove il campo recettivo attraverso la divisione tra le variazione infinitesima di spazio (pixel) e tempo: $dv=dx/dt$. Ottenere una linea dritta e pulita è impossibile senza un grosso residuo di rumore. Per questo motivo, per questi scopi,

è generalmente più utilizzata l'analisi nel dominio delle frequenze.

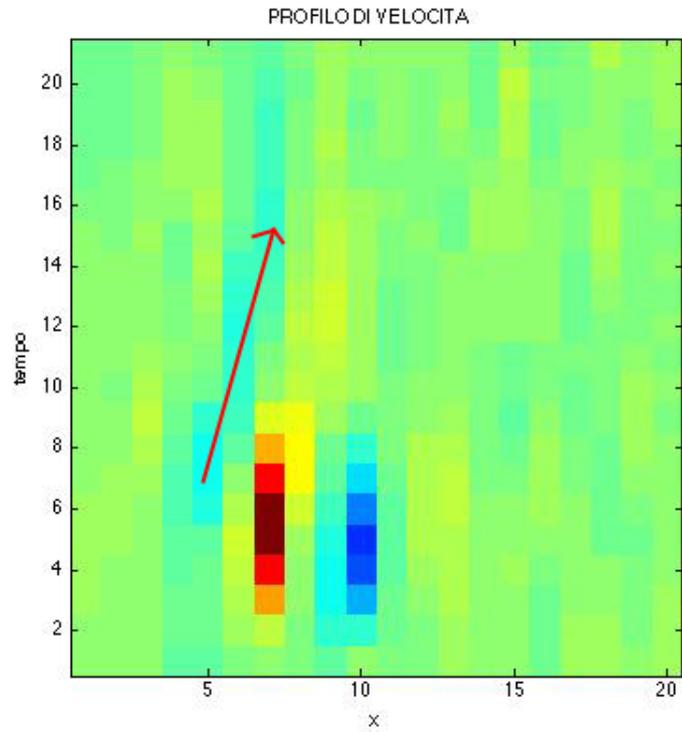


Figura 2.12: Campi recettivi valutati nel dominio spazio-tempo ($x-t$). La freccia rossa indica la direzione di spostamento.

2.5 Data set di Ringach (V1 del macaco)

In questo lavoro di tesi sono stati utilizzati i dati messi a disposizione da Ringach e il suo team^[31]; questi sono formati dalle risposte a sequenze d'immagini naturali provenienti da neuroni appartenenti alla corteccia visiva primaria (V1). Tutte le registrazioni sono state eseguite su macachi attraverso un impianto di micro-elettrodi nella corteccia visiva primaria^[39].

Agli animali sono state mostrate diverse sequenze d'immagini riprodotte da stimoli naturali (in particolare, scene di film), proiettando dei video assicurandosi che fossero allineati con il loro punto di fissazione. I frames sono stati campionati ad una risoluzione di 320x240 pixels e ad una frequenza temporale di 30 Hz. Trenta differenti segmenti della durata di 30 s circa sono stati campionati da quattro differenti film (Sleeper, Benji, Goldfinger and Sheakespeare in Love) per un totale di 20 minuti di video. Questi sono stati compressi e salvati su un disco per essere in seguito riprodotti su un monitor con una frequenza di refresh di 90 Hz. Ogni immagine è stata riprodotta per tre frames consecutivi.

Ogni singolo frame mostrato è stato salvato come immagine "jpg"; tutti questi dati sono stati divisi per filmato e a sua volta per segmenti. Ci sono circa 900 frame per ogni segmento che rappresentano 30 s di stimolazione.

Il k-esimo frame dell'immagine è stato presentato durante l'intervallo $(kT, (k+1)T)$, dove T rappresenta l'intervallo di refresh dei frames, (non del monitor) e vale: $T = \frac{3}{90Hz} = 33.33$ ms.

Tutte le caratteristiche dei dati e dell'esperimento sono state salvate come file Matlab. Qui sono presenti gli spezzoni relativi alle forme d'onda registrate. Per questo esperimento è stato settato un valore di soglia a basso livello; in questo modo è possibile che anche una parte di rumore sia stato registrato. Nella figura sotto sono mostrate tutte queste forme d'onda sovrapposte. Attraverso una semplice analisi SVD è possibile notare che ci sono due nuvole di punti, una vicino all'origine (indicato dal punto rosso): questa nube di puntini

rappresenta semplicemente il rumore. Si tratta di segnali che hanno attraversato in modo casuale la soglia e sono stati registrati. Tuttavia, si può vedere una nube di punti chiaramente separati lontana dalla nuvola del rumore. Spesso attraverso un semplice algoritmo di clustering è possibile classificare automaticamente le forme d'onda.

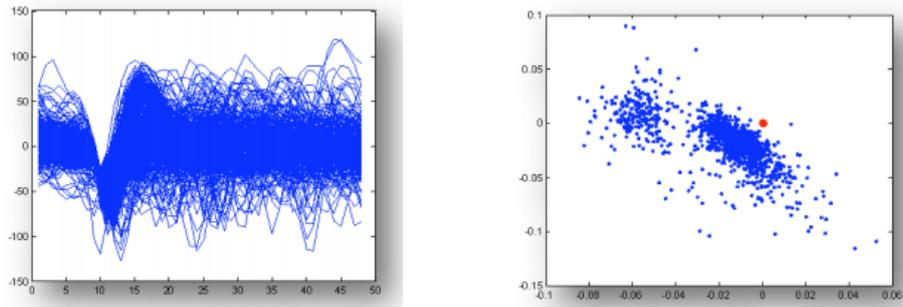


Figura 2.13: A sinistra è possibile notare la sovrapposizione di tutte le forme d'onda registrate. A destra sono rappresentate le stesse forme d'onda ma come nuvole di punti grazie all'analisi SVD. Il puntino rosso indica l'origine.

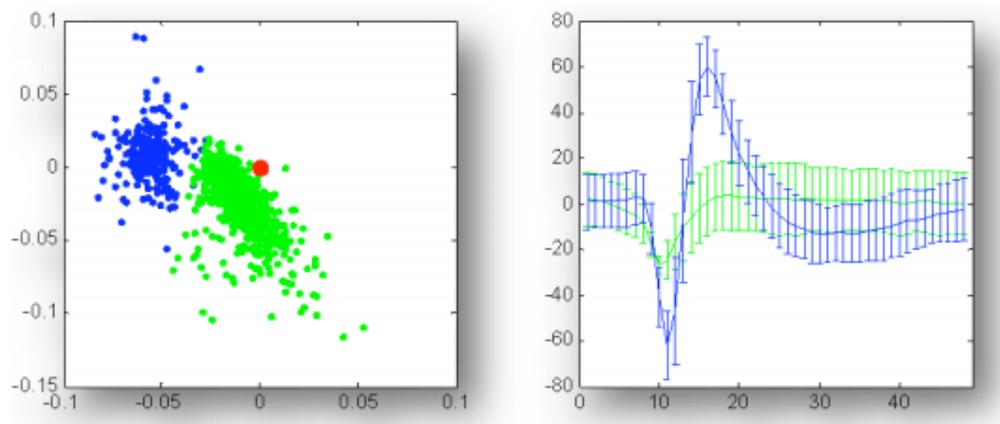


Figura 2.14: Sfruttando un algoritmo di clustering le forme d'onda vengono associate a due distinte nuvole di punti (figura a destra). Nella parte sinistra sono mostrate le forme d'onda relative alle due nuvole con la rispettiva deviazione standard. I punti e la forma d'onda di colore blu rappresentano il segnale utile, quelle verdi, invece, il rumore.

Le registrazioni elettro-fisiologiche fatte dal team di Ringach, riguardavano la risposta neurale delle cellule in corteccia visiva primaria ed erano relative alla stimolazione attraverso i dati descritti ad inizio sezione. L'array 10x10 di micro-elettrodi utilizzati in questo esperimento sono stati inseriti ad una profondità di circa 1 mm nella superficie corticale degli animali sfruttando un'installazione ad alta velocità per ridurre i danni dovuti all'inserimento dell'impianto. La procedura d'impianto è illustrata nella figura sotto^[39].

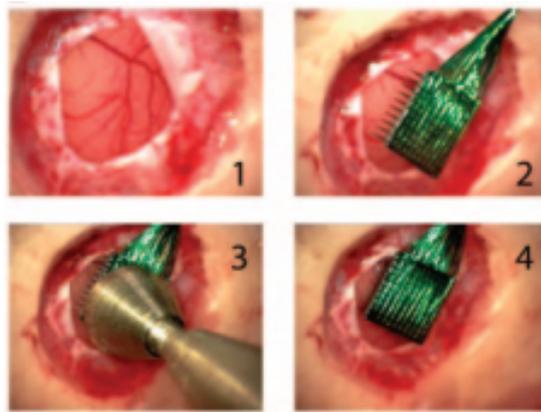


Figura 2.15: Sequenza d'inserzione dell'array di elettrodi: (1) locazione, (2) posizionamento dell'array, (3) inserzione pneumatica, e (4) inserzione dell'array.

Tutte le risposte di tutti i neuroni che sono stati eccitati dalla presentazione degli stimoli sono state registrate. Quindi, per ogni stimolo, più segnali (potenziali d'azione) derivanti da diversi neuroni sono stati registrati. Per riuscire a distinguere i neuroni che hanno risposto, si utilizza una tecnica chiamata clustering che, sfruttando diverse caratteristiche delle forme d'onda registrate, associa ogni spike al relativo neurone. Una volta fatto questo essendo noti l'istante di tempo in cui è stato presentato ogni stimolo e l'istante in cui si è registrata la risposta, è possibile risalire a quali stimoli hanno eccitato il neurone.

2.5.1 Il Clustering

Molte questioni in neuro-scienza dipendono dall'analisi dell'attività neurale registrata sotto diverse condizioni e comportamenti. Per questo motivo, acquisire simultaneamente più dati provenienti da più neuroni è indispensabile per chiarire i principi del processamento dell'informazione neurale. Attraverso il *clustering* è possibile classificare e organizzare i dati visti nella sezione precedente in modo che sia possibile eseguire le varie elaborazioni.

I nuovi sistemi di acquisizione consentono di registrare centinaia di canali contemporaneamente e l'affidabilità di questi dati dipende dall'accuratezza con cui vengono identificati i neuroni singolarmente^[32]. Questa tecnica è diventata molto comune e permette di avere un'alta risoluzione temporale. Ogni canale, e quindi ogni elettrodo, riceve il segnale da più neuroni che possono poi venire raggruppati attraverso alcune tecniche matematiche. E' possibile che differenti canali sentano spikes provenienti dallo stesso neurone ma con vari gradi d'attenuazione, dipendenti dalla distanza tra il canale ed il neurone; fare questo raggruppamento è comunque abbastanza semplice^[33].

Gli step alla base dell'algoritmo per classificare gli spike sono i seguenti:

- 1 - rilevamento spike.
- 2 - estrazione di differenti caratteristiche dalla forma degli spike.
- 3 - raggruppamento degli spikes utilizzando le caratteristiche trovate.

I metodi utilizzati nello *spike sorting* si basano normalmente sul raggruppamento sfruttando alcune caratteristiche della forma degli spike, come l'ampiezza picco-picco, la larghezza del segnale, o le componenti principali. Non è comunque possibile conoscere a priori quale caratteristica sia migliore per discriminare gli spike in un determinato data set. In più anche utilizzando un'ottima procedura per

classificare i neuroni, servono regolazioni aggiuntive da fare manualmente, le quali introducono errori.

Un approccio alternativo è quello di definire le classi di spike attraverso un set di soglie pre-definite manualmente. Questo metodo, purtroppo, è efficiente ed affidabile solamente quando il rapporto segnale/rumore è molto elevato.

I problemi che si riscontrano nella tecnica dello spike sorting sono solitamente legati alle forme d'onda extra-cellulari, come quella mostrata in figura 2.16. Com'è possibile notare ci sono differenti tipi di potenziale d'azione. Questi sono legati a differenti neuroni? Se è così, bisogna stabilirne la corrispondenza? Le forme d'onda di figura mostrano anche un notevole quantità di rumore di fondo che potrebbe corrispondere a rumore dovuto all'amplificazione del segnale oppure a picchi minori dovuti a neuroni più lontani. Come si possono allora classificare i neuroni in presenza di questo rumore? Inoltre i picchi di alcune cellule potrebbero sovrapporsi. Come si possono dividere questi picchi^[34]?

Per poter rispondere a queste domande occorre introdurre le varie tecniche di clustering.

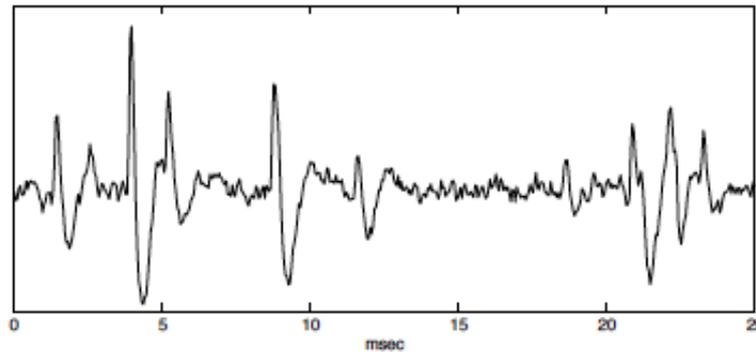


Figura 2.16: La forma d'onda extra-cellulare mostra diversi potenziali d'azione generati da un numero imprecisato di neuroni.

Superparamagnetic clustering

Sui dati a disposizione era già stata applicata la tecnica dello spike-sorting: attraverso un software, gli spike erano già stati raggruppati. Esistono molte applicazioni che eseguono spike-sorting, ognuna delle quali segue metodi e algoritmi differenti. I dati utilizzati sono stati elaborati con la tecnica del “*superparamagnetic clustering*” (SPC)^[35], una nuova procedura sviluppata in un contesto meccaico-statistico. L’algoritmo completo comprende tutti e tre gli step visti nella sezione precedente: nel primo, gli spikes sono rilevati utilizzando una soglia di ampiezza automatica su dati precedentemente filtrati con un filtro passa alto, nel secondo step, un piccolo set di coefficienti d’onda viene scelto come input per l’algoritmo di clustering; infine il clustering superparamagnetic classifica gli spike utilizzando il set dei coefficienti selezionati.

La trasformata d’onda (TO) utilizzata in questo metodo è una rappresentazione tempo-frequenza del segnale che porta due grossi vantaggi: 1) ha un’ottima risoluzione sia nel dominio del tempo che in quello delle frequenze e 2) elimina il bisogno di avere un segnale stazionario. Esso è definito dalla convoluzione tra il segnale $x(t)$ e la funzione d’onda $\psi_{a,b}(t)$:

$$W_{\psi}X(a, b) = \langle x(t) | \psi_{a,b}(t) \rangle \quad (2.13)$$

dove $\psi_{a,b}(t)$ è la versione dilatata (contratta) e shiftata dell’unica funzione d’onda $\psi(t)$:

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi \left(\frac{t-b}{a} \right) \quad (2.14)$$

in cui a e b sono rispettivamente i parametri di scala e traslazione^[32]. Quello che si vuole fare è mettere in relazione gli spikes della forma d’onda registrata con l’attività di una popolazione di neuroni. Non è sempre vero ma, solitamente, i neuroni generano potenziali d’azione con una ben precisa forma. Per molti neuroni, la caratteristica principale della forma dello spike è la sua ampiezza o il valore di picco.

Uno dei metodi più semplici per misurare l'attività di un neurone è utilizzare un valore di soglia legato alla tensione. L'attività neurale viene quindi misurata attraverso la generazione di un impulso ogni volta che la tensione misurata supera la soglia. Il riconoscimento di spike è eseguito attraverso sogliatura dopo aver filtrato il segnale con un filtro passa-banda. La soglia (Thr) viene automaticamente settata come:

$$Thr = 4\sigma_n; \quad \text{con} \quad \sigma_n = \text{median} \left\{ \frac{|x|}{0,6745} \right\} \quad (2.15)$$

dove x è il segnale filtrato e σ_n è la stima della deviazione standard del rumore di fondo.

Il vantaggio che offre questa tecnica è la richiesta di hardware e software minimi; come svantaggio non sempre è possibile ottenere un isolamento accettabile^[34].

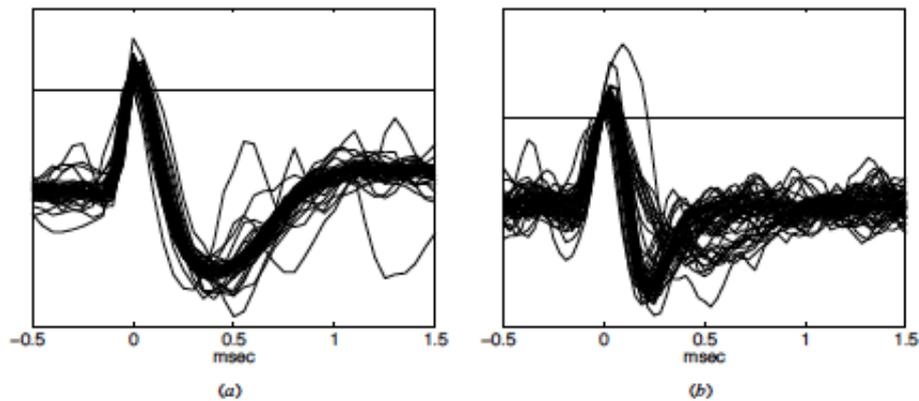


Figura 2.17: (a) Una tipica traccia di un neurone ben isolato. Una traccia viene visualizzata ogni volta che il valore di tensione supera la soglia indicata dalla linea orizzontale. Sono inoltre visualizzate le forme d'onda precedenti. I dati sono della stessa registrazione della traccia mostrata in figura 2.16. (b) Una traccia di un neurone mal isolato.

La figura 2.17 mostra l'esempio di un neurone ben isolato (a) e di uno male isolato (b). Nel primo caso è possibile notare la presenza di altri spikes di fondo, ma questi hanno solo un piccolo effetto sulla qualità dell'isolamento. Nel secondo caso, si possono notare due distinti spikes ed è impossibile impostare un valore di soglia che permetta di isolarne uno.

Molto spesso non è possibile separare alcuni spike dal rumore di fondo con accuratezza. Il livello di soglia determina il trade-off tra gli spike persi (falsi negativi) ed il numero di eventi di fondo che oltrepassano la soglia (falsi positivi) come illustrato in figura 2.18. Se la soglia è regolata al livello "A", tutti gli spikes dell'unità 1 sono riconosciuti, ma ci sarà un grosso numero di falsi positivi (gli spike dell'unità 2). Se la soglia viene portata al livello "B", solo gli spikes dell'unità 1 saranno riconosciuti, ma un vasto numero sarà sotto soglia (falsi negativi). Idealmente la soglia deve essere settata per ottimizzare il rapporto desiderato tra falsi positivi e falsi negativi.

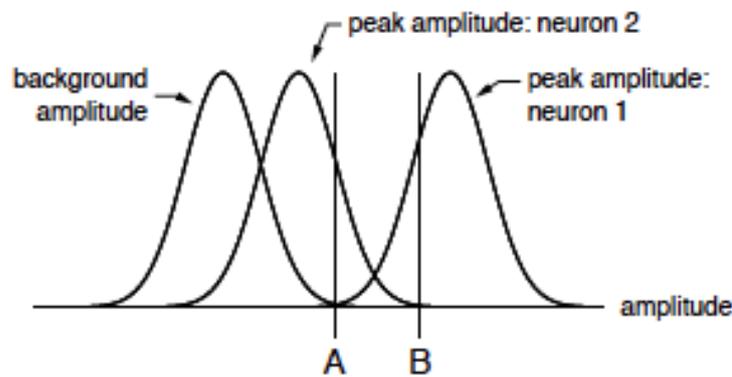


Figura 2.18: La figura mostra la distribuzione delle ampiezze per l'attività di fondo e le ampiezze di picco degli spikes da due unità. L'ampiezza è lungo l'asse orizzontale. Impostare il livello di soglia per la posizione A introduce un gran numero di spikes dall'unità 2. Aumentando la soglia fino a B riduce il numero di spikes che sono erroneamente classificati, a discapito di molti spikes che vengono persi.

Se il livello di rumore di fondo è piccolo rispetto all'ampiezza degli spikes e le distribuzioni d'ampiezza sono ben separate, entrambi questi errori tenderanno a zero e la posizione della soglia perderà d'importanza^[34].

Il clustering superparamagnetic è un nuovo metodo per fare clustering implementato come modello di Potts iterato con il metodo Monte Carlo^[32] ma che non tratterò nello specifico in quanto non implementato direttamente in questo lavoro di tesi. La figura che segue (2.19) riassume brevemente i tre passi principali di questo algoritmo:

- 1) riconoscimento automatico degli spikes attraverso soglie in ampiezza;
- 2) le forme d'onda sono calcolate per ogni spike e il miglior coefficiente per separarli in classi viene selezionato automaticamente;
- 3) i coefficienti delle forme d'onda selezionati sono utilizzati come input dell'algoritmo della superparamagnetic clustering.

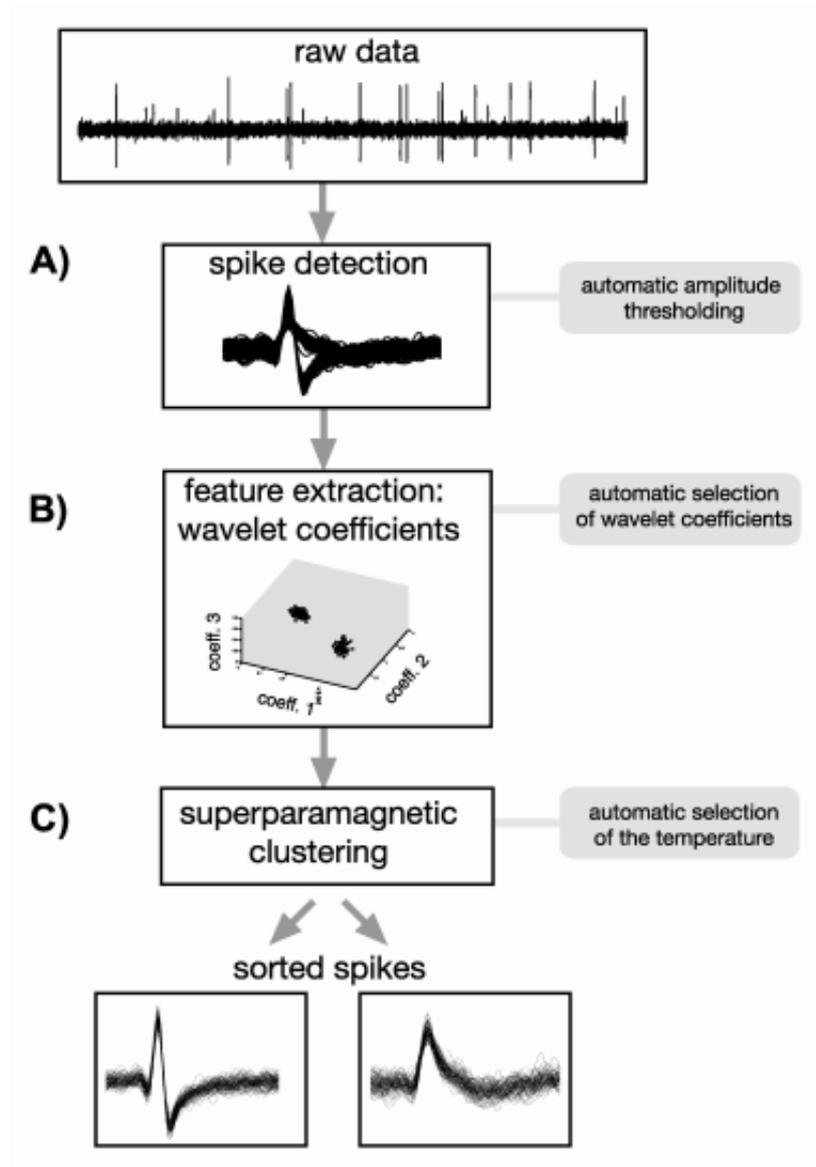


Figura 2.19: Rappresentazione schematica del metodo spike-sorting. Gli spike sono rilevati dalla soglia in ampiezza di dati naturali. Nell'estrazione delle caratteristiche, i picchi rilevati (* contrassegnati con un asterisco) sono stati trasformati in onde e dai coefficienti della distribuzione di TO con picchi multipli sono stati selezionati per il clustering degli spike utilizzando una variazione del metodo di Bayes. Infine, i dati nello spazio di funzioni multidimensionali sono stati raggruppati per spikes appartenenti a diversi neuroni.

La figura 2.20 mostra una simulazione di un data set in presenza di rumore. La parte “A” della figura rivela le tre forme di spike che si aggiungono al rumore di fondo. E’ possibile notare meglio questa caratteristica nella figura “B”. In basso (C) è mostrata una parte del segnale di figura 2.20(B) con una miglior risoluzione temporale. Le differenze in forma e ampiezza tra gli spikes appartenenti ad una stessa classe (identificati con un marker dello stesso livello di grigio), sono dovuti al rumore di fondo.

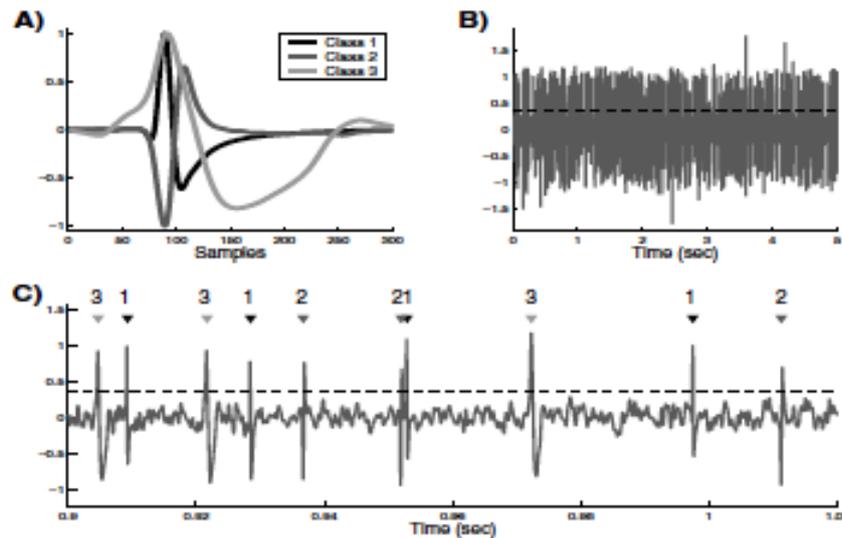


Figura 2.20: simulazione di data set usato per lo spike sorting. (A) Le tre forme di spikes. (B) Il segnale completo con spikes e rumore di fondo. (C) Gli stessi dati scalati nel tempo.

2.5.2 Ricostruzione dei campi recettivi

Una volta che i dati sono stati organizzati e raggruppati attraverso la tecnica del clustering, è stato possibile applicare a questi la tecnica della reverse correlation ma è sorto un problema.

Non essendo possibile conoscere a priori a quali cellule (semplici o complesse) appartengono le registrazioni elettro-fisiologiche effettuate da Ringach, non è stato possibile utilizzare la spike-triggered average. Con questa tecnica, infatti, è possibile intercettare solo il comportamento di neuroni aventi una relazione risposta/stimolo di tipo lineare come le cellule semplici. Questo metodo è stato largamente usato per misurare i campi recettivi spazio-temporali di neuroni all'inizio del percorso visivo. Nella corteccia visiva però la maggior parte dei neuroni sono delle cellule complesse, con una relazione risposta/stimolo di tipo non lineare ed il metodo della spike-triggered covariance sarebbe ideale da questo punto di vista, in quanto permetterebbe di ricostruire i campi recettivi sia di cellule complesse che di cellule semplici. Gli stimoli utilizzati però, sono immagini ad alta dimensionalità e questo richiede una potenza di calcolo non banale. Come spiegato inizialmente, i frame utilizzati erano matrici 320x240 pixel e, per ogni risposta (spikes), si voleva calcolare il profilo spazio-temporale. Per fare questo sono necessari diversi frame; considerando di utilizzare quelli mostrati in un range di 400 ms prima della risposta (non è possibile infatti sapere con certezza quale stimolo ha provocato la risposta e nemmeno il tempo di risposta di una cellula che può variare da neurone a neurone), e che la frequenza di visualizzazione dei frame è di 33.33 ms, si sarebbero presi circa 12 frames: $320 \times 240 \times 12 = 921.600$. Ottenendo una matrice di covarianza "C" 921.600×921.600 . Anche salvando ogni elemento come "single", cioè utilizzando per ogni elemento una memoria di 4 byte, servirebbero comunque più di 3000 Gb di memoria.

Per immagini di questo tipo non è quindi possibile applicare direttamente la reverse correlation. Bisogna considerare un approccio differente che sarà analizzato nel prossimo capitolo.

Capitolo 3

Individuazione del campo recettivo

Non essendo possibile elaborare direttamente gli stimoli utilizzati da Ringach con la tecnica della reverse correlation a causa dell'alta dimensionalità che li caratterizza, in questo capitolo sarà valutato un approccio alternativo con lo scopo di individuare la zona dello spazio in cui il profilo recettore è collocato.

Per fare questo entrano in gioco le *funzioni di Gabor* e l'*energia di orientazione* che modellano il comportamento di cellule semplici e complesse in corteccia visiva primaria. Per intercettare le caratteristiche funzionali di ogni cellula, questo approccio prevede l'utilizzo dei modelli appena citati valutati su diverse orientazioni e scale. Per individuare la regione in cui si trova il campo recettivo è stata utilizzata la Z-score (standardizzazione), una tecnica di analisi statistica. Infine saranno mostrati ed analizzati i risultati ottenuti da questa elaborazione, valutando approcci differenti e sviluppi futuri.

Come visto alla fine del capitolo precedente non è stato possibile applicare sui dati^[31] le tecniche della STA e della STC in modo diretto. Per questo motivo si è studiato un metodo alternativo.

L'idea seguita per risolvere questo problema si basa sulla riduzione della dimensionalità dei dati. Invece di applicare la reverse correlation su tutta l'immagine, si è pensato di individuarne quella porzione che ha eccitato la cellula. Si vuole cioè trovare quella regione dell'immagine che ha provocato lo spike del neurone. Come si fa a fare questo? Per riprodurre il comportamento di cellule semplici e complesse, si possono utilizzare le funzioni di Gabor e l'energia di orientazione. Questi modelli vanno però applicati su più scale e per diverse orientazioni, in quanto non è noto a priori nè dov'è posizionato nè com'è orientato il profilo recettore. Una volta individuata tale zona, è possibile ritagliarla e utilizzare questa come una nuova immagine su cui applicare le tecniche viste nel capitolo precedente per ricostruire il profilo recettivo.

L'obiettivo di questa tesi è quindi quello di trovare e studiare le tecniche che permettono di individuare la regione in cui è posizionato il campo recettivo. Una volta localizzato con precisione, basterà poi applicare le tecniche viste precedentemente.

3.1 I filtri di Gabor

Come già illustrato, differenti neuroni si trovano, in grandissimo numero, all'interno della corteccia visiva. A seguito di numerosi studi, è emerso che la selettività di un neurone visivo all'orientazione è una delle caratteristiche determinanti del loro comportamento. Questa selettività fa sì che il "firing rate" di un neurone sia altamente dipendente dall'orientazione che uno stimolo assume^[12].

Come visto nel primo capitolo, i neuroni della corteccia visiva sono organizzati in strutture colonnari. Ogni colonna presenta cellule con la medesima orientazione, colonne adiacenti differiscono tra loro in maniera quasi continua, fino a ricoprire tutte le orientazioni possibili. Questa architettura funzionale è alla base della struttura della corteccia visiva, e sembra giocare un ruolo cruciale nel determinare come il nostro cervello interpreti e decodifichi i segnali provenienti dal mondo esterno e, in particolare, l'elaborazione e il riconoscimento di alcune caratteristiche specifiche dell'immagine, come ad esempio i contorni. Le funzioni di Gabor furono inizialmente proposte da Dennis Gabor nel 1946 come strumento di analisi tempo-frequenza per la decomposizione di un segnale mono-dimensionale. Egli rilevò che la funzione di Gabor può ottenere la migliore risoluzione sia nel dominio temporale (o spaziale) che in quello delle frequenze, mostrando che queste funzioni elementari minimizzano proprio l'incertezza tempo-frequenza. Attorno agli anni '90 Daugman estese queste funzioni su un dominio bi-dimensionale e mostrò la stretta similarità tra la loro forma e i campi recettivi della corteccia visiva umana dovuta alla loro selettività all'orientazione^[7]. Da quel momento le funzioni di Gabor sono state largamente usate nell'analisi delle immagini digitali.

Per determinare i contorni in una scena, il sistema visivo umano usa un set di filtri passa banda a diverse orientazioni. I *filtri di Gabor* sono eccellenti filtri lineari per segnali mono-dimensionali. Con i filtri bi e tri-dimensionali è inoltre possibile modellare i campi recettivi delle cellule della corteccia visiva primaria.

Non tutte le funzioni di Gabor sono utili, ma solo una sotto famiglia

sensibile alle orientazioni. La formulazione mono-dimensionale di un filtro di Gabor può essere scritta come:

$$g_{\alpha, u_0, x_0}(x) = e^{-\pi\alpha^2(x-x_0)^2} e^{-i[2\pi u_0(x-x_0)+\phi]} \quad (3.1)$$

Questa formula può essere vista come il prodotto tra una Gaussiana avente scala α e centrata in x_0 , moltiplicata per un'onda sinusoidale complessa di frequenza u_0 (anch'essa centrata in x_0) e avente fase ϕ .

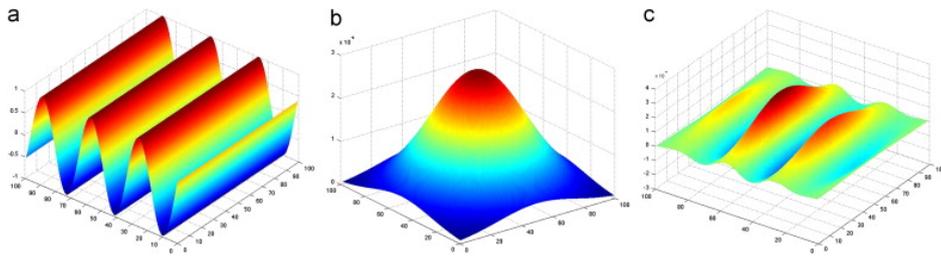


Figura 3.1: Un filtro di Gabor (c) è il prodotto tra un'onda sinusoidale (a) ed una Gaussiana (b)

Essendo questa formula complessa, è possibile dividerla in due parti:

- *Reale:*

$$Re\{g_{\alpha, u_0, x_0}(x)\} = e^{-\pi\alpha^2(x-x_0)^2} \cos(2\pi u_0(x-x_0) + \phi) \quad (3.2)$$

- *Immaginaria:*

$$Im\{g_{\alpha, u_0, x_0}(x)\} = e^{-\pi\alpha^2(x-x_0)^2} \sin(2\pi u_0(x-x_0) + \phi) \quad (3.3)$$

Le due parti (Re ed Im) sono in quadratura di fase tra loro, cioè la differenza di fase è pari a $\pi/2$. Questa caratteristica rende le funzioni di Gabor particolarmente adatte per modellare i profili recettori di cellule della corteccia visiva primaria ed in particolare di cellule complesse, come sarà mostrato in seguito.

Una funzione di Gabor bi-dimensionale nel dominio spaziale (x,y) è così scritta:

$$g_{\alpha,\beta,P_0,p_0}(x,y) = e^{-\pi[\alpha^2(x-x_0)^2+\beta^2(y-y_0)^2]} e^{-i[2\pi(u_0(x-x_0)+v_0(y-y_0))+\phi]} \quad (3.4)$$

in questo caso la funzione di Gabor è centrata in $p_0=(x_0,y_0)$ e formata da una Gaussiana bi-dimensionale moltiplicata per un'onda piana. Una Gaussiana bi-dimensionale è ottenibile, per la proprietà di separabilità di cui godono le funzioni Gaussiane, attraverso la moltiplicazione tra due Gaussiane monodimensionali, una lungo l'asse x , di larghezza α e centrata in x_0 e l'altra lungo l'asse delle y , di larghezza β e centrata in y_0 ^[11]. Come nel caso mono-dimensionale, anche in quello bi-dimensionale possiamo dividerla in due parti:

- *Reale:*

$$Re\{g(x,y)\} = e^{-\pi[\alpha^2(x-x_0)^2+\beta^2(y-y_0)^2]} \cos(2\pi(u_0(x-x_0)+v_0(y-y_0))+\phi) \quad (3.5)$$

- *Immaginaria:*

$$Im\{g(x,y)\} = e^{-\pi[\alpha^2(x-x_0)^2+\beta^2(y-y_0)^2]} \sin(2\pi(u_0(x-x_0)+v_0(y-y_0))+\phi) \quad (3.6)$$

Gabor dimostrò che con queste funzioni era possibile ottenere la miglior risoluzione sia nel dominio del tempo che nel dominio delle frequenze; sono quindi funzioni che risolvono in maniera ottima il vincolo posto dal principio di indeterminazione sulla localizzazione di segnali nei domini duali spazio-frequenza. Questa funzione possiede infatti il più piccolo prodotto tra estensione nello spazio (Δx) ed estensione nel dominio delle frequenze (Δu). In questo modo è possibile realizzare filtri con ottime proprietà di localizzazione sia nel dominio spaziale che in quello delle frequenze. Inoltre, questo particolare sottoinsieme di filtri risultano selettivi per le orientazioni, andando a rispondere con maggiore forza ad un segnale in ingresso avente orientazione solidale a quella del filtro stesso^[12].

Attraverso tecniche sperimentali^[30] è stato possibile studiare la rispo-

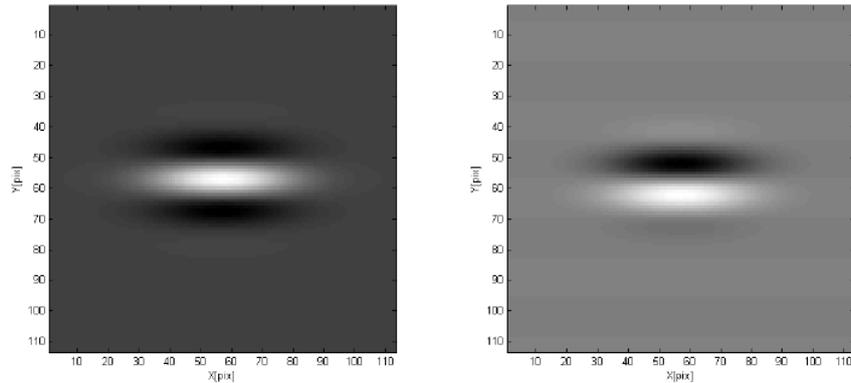


Figura 3.2: Parte reale ed immaginaria di una funzione di Gabor nel dominio spaziale (x, y) . In questo caso i parametri assumono i seguenti valori: $x_0 = y_0 = 0$, $\alpha^2 = \frac{1}{128\pi}$ e $\epsilon = \frac{\beta}{\alpha} = 2$

sta di cellule semplici, sottoposte a stimoli luminosi puntuali generati in punti diversi, all'interno di una griglia di dimensioni tali da coprire l'intero profilo recettore di ogni cellula. Da questi esperimenti Palmer e Jones^[30] mostrarono come la struttura dei profili recettori determinata sperimentalmente, abbia forme molto simili a quelle di parti reali o immaginarie di filtri realizzati mediante l'utilizzo di funzioni di Gabor. Risulta quindi possibile affermare che questa tipologia di filtri approssima con buona precisione la struttura geometrica di profili recettori di cellule semplici. Inoltre, parte reale e parte immaginaria di una funzione di Gabor sono, rispettivamente, funzioni pari e dispari. Proprio questa tipologia di funzioni è quella che meglio si adatta a descrivere la struttura di profili recettori appartenenti a coppie di cellule semplici in quadratura di fase tra loro ma con medesima orientazione preferenziale del profilo recettore.

E' quindi lecito affermare che la parte reale e quella immaginaria di funzioni di Gabor bi-dimensionali sono in grado di modellare i profili recettori di questa tipologie di cellule con buona approssimazione. In particolare, la parte reale può essere utilizzata per modellare profili recettori appartenenti ad una cellula semplice che presenta una simme-

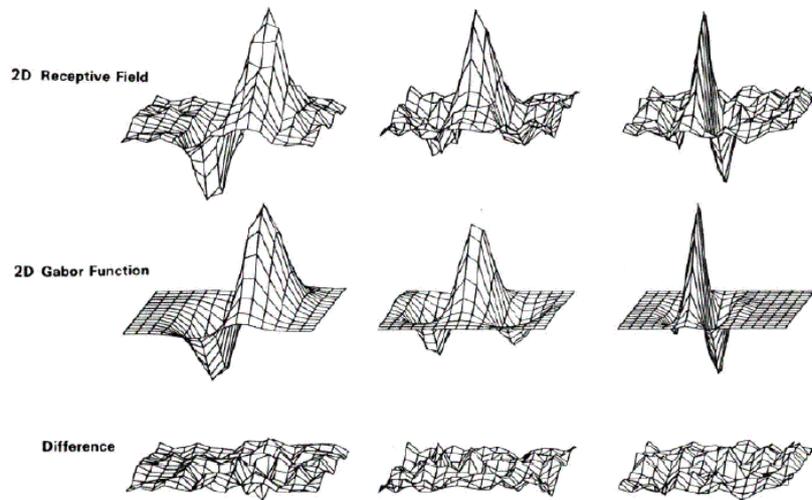


Figura 3.3: Immagine tratta dallo studio di Palmer e Jones^[30]. La riga in alto della figura mostra tre profili recettori 2D di cellule semplici determinati sperimentalmente con la tecnica della reverse correlation; la seconda mostra le funzioni di Gabor 2D reali o immaginarie che meglio approssimano i profili recettori determinati sperimentalmente; nella riga in basso è mostrato l'errore residuo del fitting tra le prime due.

tria pari, mentre la parte immaginaria per modellare profili recettori di una cellula con simmetria dispari^[16]. Bisogna ricordare però che in corteccia le cellule semplici non sono divise in maniera così netta in cellule pari e dispari quindi, nel modellarle, il parametro ϕ della fase gioca un ruolo fondamentale.

Il lavoro svolto sui filtri di Gabor mi ha portato prima a studiare le loro caratteristiche e a comprendere il loro funzionamento e poi ad implementarli attraverso l'utilizzo di Matlab. Questa particolare famiglia di filtri, come spiegato sopra, può essere utilizzata nell'elaborazione di immagini, in modo particolare nel riconoscimento dei contorni, in quanto sono specifici per una determinata orientazione. Uno dei parametri più studiati in letteratura è l' "indice di sotto-regioni" (ISR)^[13]; questo parametro indica il numero di sotto regioni On e Off che si tro-

vano all'interno dell'area del campo recettivo e che provocano l'effetto eccitatorio o inibitorio di un neurone corticale. Attraverso l'equazione 3.7 è possibile calcolare l'ISR come il numero di semi-cicli (metà lunghezza d'onda), all'interno di due deviazioni standard della Gaussiana costituente il filtro di Gabor:

$$ISR = 8\sigma u_0 \quad (3.7)$$

in cui σ è la lunghezza spaziale o scala, u_0 è la frequenza spaziale e ISR è il numero di sottoregioni all'interno di una campana di Gauss. Per il principio di indeterminazione di Heisenberg applicato alla teoria dell'informazione, meno risoluzione si vuole nello spazio e più risoluzione si ottiene nelle frequenze. Per semi-ciclo si fa riferimento alla metà di un periodo di un'onda sinusoidale. Lavorando con immagini si deve fare riferimento al valore del pixel, in questo caso della luminosità. In un periodo l'intensità cambia valore due volte, è quindi logico che in un solo pixel non è possibile rappresentare due valori. La massima frequenza spaziale che posso usare è quindi 0.5 cicli/pixel. In generale poi, i profili recettori in corteccia, hanno un rapporto σ/u_0 che rimane circa costante ad un valore "k" e questo fa sì che sotto la Gaussiana si abbiano "k" sotto regioni. σ è la deviazione standard della campana di Gauss e da esso dipende la larghezza della campana. Il valore attorno cui oscilla l'ISR è 2.72, per questo motivo, data in ingresso la scala, è possibile calcolare facilmente la frequenza spaziale^[13].

La fase successiva del mio lavoro ha visto l'implementazione della parte reale e di quella immaginaria dei filtri di Gabor, che sono poi state utilizzate per elaborare agli stimoli utilizzati da Ringach. Si è valutata la parte reale per quattro orientazioni ($\theta = 0, \pi/4, \pi/2$ e $3/4\pi$) essendo la funzione coseno simmetrica rispetto l'origine; questa caratteristica non appartiene alla funzione seno che, essendo una funzione dispari, non presenta simmetria rispetto l'origine e quindi è stata valutata per un range compreso tra 0 e 2π per otto orientazioni di θ . Dalle immagini che seguono è possibile vedere come la parte immaginaria dei filtri di Gabor abbia polarità inversa per orientazioni opposte: os-

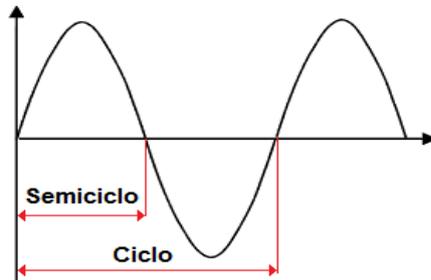


Figura 3.4: La figura mostra un'intero periodo di un'onda sinusoidale. Partendo da zero, si ha un aumento fino ad un picco positivo e poi un ritorno a zero; si raggiunge un massimo negativo, e poi un'altra volta a zero. Ogni ripetizione, dall'inizio (valore 0) fino a tornare al punto di partenza è chiamato ciclo. La metà di un ciclo si chiama semi-ciclo.

servando lo stesso punto dell'immagine per $\theta = 0$ e $-\pi$, $+\pi/2$ e $-\pi/2$... si ottengono infatti gli stessi valori ma di segno opposto.

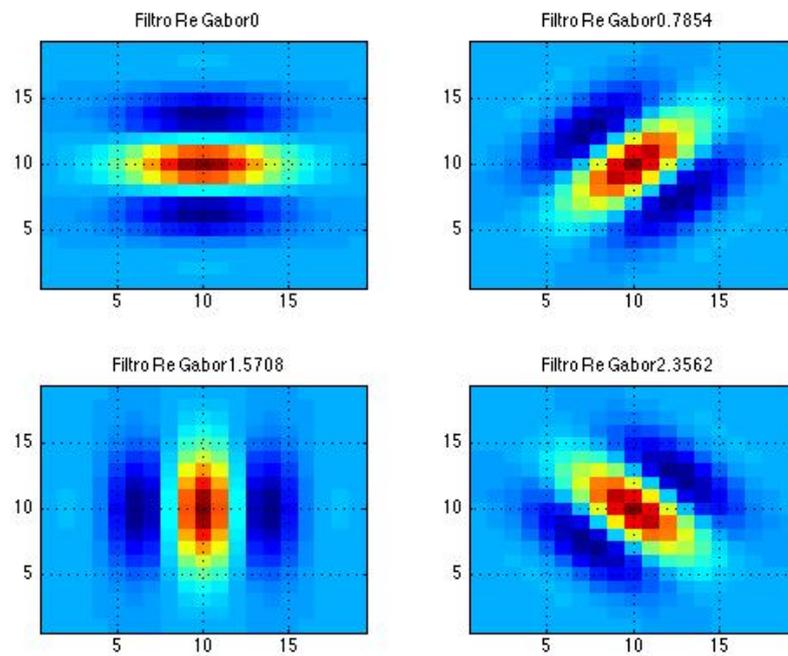


Figura 3.5: Implementazione della parte reale del filtro di Gabor. E' possibile osservare le quattro orientazioni di θ rispetto l'asse x: 0 , $\pi/4$, $\pi/2$ e $3/4\pi$.

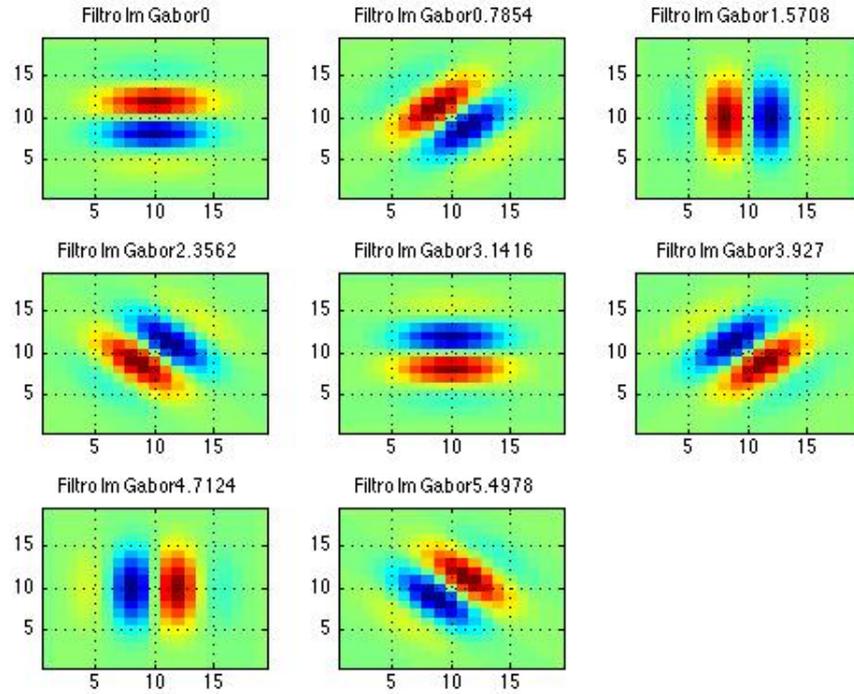


Figura 3.6: Sviluppo della parte immaginaria del filtro di Gabor. Non presentando simmetria rispetto l'origine, questa funzione è stata valutata per otto orientazioni di θ rispetto l'asse x: $0, \pi/4, \pi/2, 3/4\pi, \pi/2, -3/4\pi, -\pi/2$ e $-\pi/4$. A causa di questa caratteristica, orientazioni opposte mostrano risposte di ugual valore ma segno opposto.

Per ottenere un risultato più chiaro, prima di convolvere le immagini con i filtri creati, è possibile trasformare le immagini da RGB in scala di grigio. Nelle figure seguenti sono mostrati i risultati ottenuti da queste convoluzioni partendo dall'immagine originale.



Figura 3.7: Immagine originale.

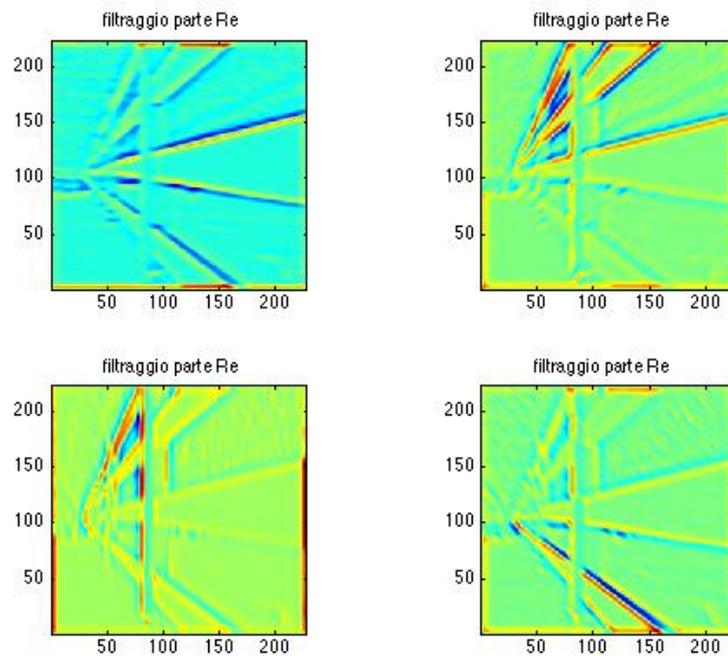


Figura 3.8: Risultati della convoluzione tra l'immagine 2.12 e la parte reale dei filtri di Gabor visti nell'immagine 2.10.

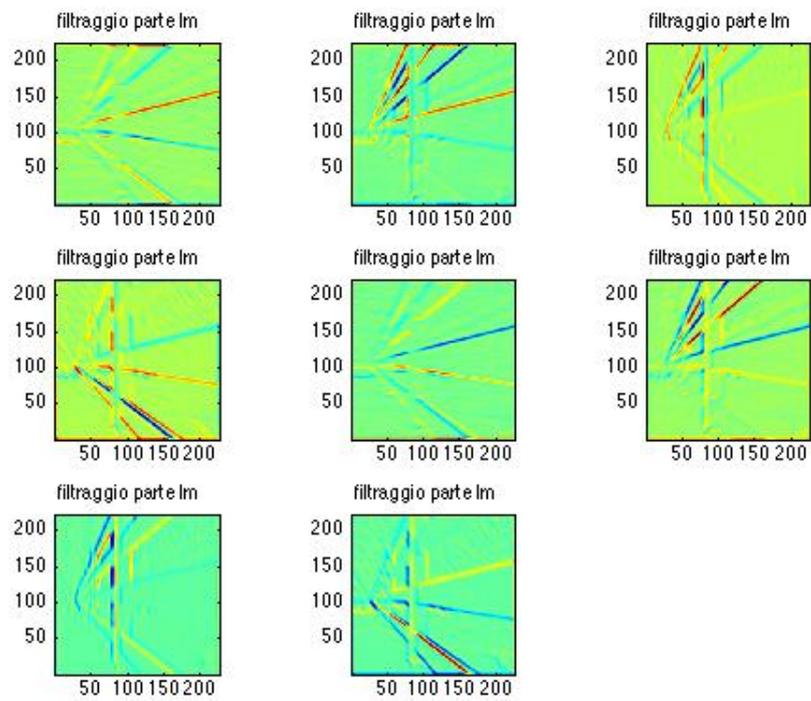


Figura 3.9: Risultati della convoluzione tra l'immagine originale e la parte immaginaria dei filtri di Gabor creati precedentemente.

3.2 L'energia di orientazione

Un problema in cui s'incorre utilizzando cellule semplici come filtri, e quindi parte reale e parte immaginaria di funzioni di Gabor, è dato dalla sensibilità alla variazione di fase: a causa della struttura del profilo recettore di queste cellule, diviso in regioni contigue di opposta polarità, la loro risposta ad uno stimolo dipende da come lo stimolo tende ad allinearsi col profilo recettore della cellula in ogni istante^[14]. In generale si può affermare che cellule semplici presentano una fase che varia da una all'altra in maniera arbitraria, tranne per coppie di cellule semplici vicine e aventi la medesima orientazione preferenziale del profilo recettore, legate, da una relazione di quadratura di fase. Risulta quindi necessario un livello di elaborazione successivo in grado di estrarre informazioni che siano indipendenti dalla polarità e dalla fase momentanea dello stimolo. Per fare questo occorre considerare le cellule complesse che prendono in ingresso input provenienti da una coppia di cellule semplici con la medesima orientazione preferenziale del profilo recettore, ma in quadratura di fase tra loro. Questo meccanismo può essere descritto utilizzando come modello per i profili recettori di due cellule semplici, la parte immaginaria (cellula con simmetria dispari) e quella reale (cellula con simmetria pari) di una funzione di Gabor; la funzione di Gabor complessa rappresenta invece una cellula complessa:

- *cellula semplice pari:*

$$\operatorname{Re}\{g(x, y)\} = e^{-\pi[\alpha^2(x-x_0)^2 + \beta^2(y-y_0)^2]} \cos(2\pi(u_0(x-x_0) + v_0(y-y_0)) + \phi) \quad (3.8)$$

- *cellula semplice dispari:*

$$\operatorname{Im}\{g(x, y)\} = e^{-\pi[\alpha^2(x-x_0)^2 + \beta^2(y-y_0)^2]} \sin(2\pi(u_0(x-x_0) + v_0(y-y_0)) + \phi) \quad (3.9)$$

- cellula complessa:

$$g(x, y) = e^{-\pi[\alpha^2(x-x_0)^2 + \beta^2(y-y_0)^2]} \cdot e^{-i[2\pi(u_0(x-x_0) + v_0(y-y_0)) + \phi]} \quad (3.10)$$

Se si vuole ottenere una misura dell'orientazione di uno stimolo insensibile alla sua fase è possibile utilizzare l'**energia di orientazione**, ricavabile dalla somma dei quadrati dei due output provenienti da due cellule semplici in quadratura di fase. Questa energia sarà sempre positiva e tanto più elevata quanto più l'orientazione dello stimolo sarà solidale con quella delle coppie di cellule semplici. In questo modo avremo una risposta sensibile all'orientazione dello stimolo ma insensibile alla sua fase (figura 3.10). Partendo da due unità che presentano un comportamento lineare rappresentanti una coppia di cellule semplici, una con simmetria pari e l'altra con simmetria dispari, è possibile affermare che l'output di entrambe queste unità sarà affetto da una sensibilità alla variazione di fase. Tuttavia, sommando il quadrato di ciascuno dei due output, provenienti dalle due cellule semplici in quadratura di fase, risulta possibile estrarre l'energia di orientazione, come dimostrato da Adelson e Bergen nel loro studio riguardante la percezione di contorni in movimento^[14].

Questa procedura trae vantaggio dal fatto che, dato che le due funzioni che rappresentano la coppia di cellule semplici sono costituite da una funzione $\sin(x)$ e da una funzione $\cos(x)$, entrambe sono moltiplicate per la stessa Gaussiana. L'operazione di somma di quadrati appena descritta permette di estrarre l'energia di orientazione. Risulta così possibile modellare il comportamento di cellule corticali in maniera relativamente precisa e con una notevole rilevanza dal punto di vista fisiologico. In pratica, questo metodo di elaborazione, facilita l'estrazione di informazioni spettrali locali bi-dimensionali (texture, scala, frequenza spaziale) senza sacrificare l'estrazione di informazioni concomitanti riguardanti il campo visivo^[15].

Partendo da questa teoria ho estrapolato per la figura 3.7 l'energia di orientazione. Le parti di colore blu sono quelle a energia ridotta,

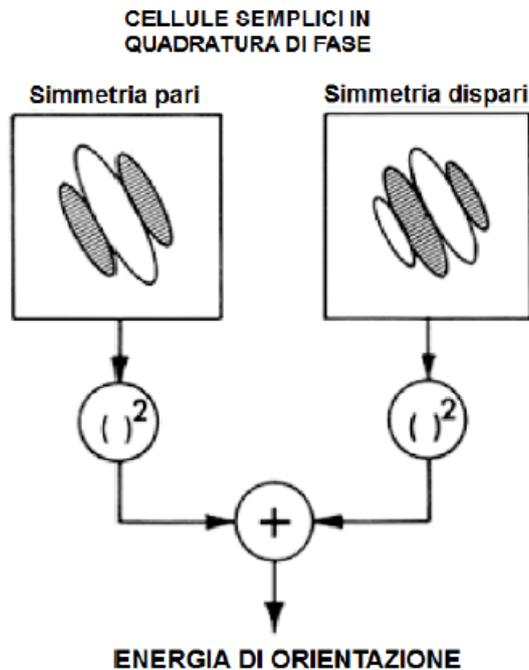


Figura 3.10: Due filtri lineari rappresentanti una coppia di cellule semplici con la stessa orientazione preferenziale, ma con una differenza di fase di 90° . Se le loro risposte vengono elevate al quadrato e sommate tra loro, il segnale risultante rappresenterà una misura, indipendente dalla fase l'energia di orientazione.

mentre quelli di colore rosso sono ad energia elevata. Colori intermedi corrispondono a livelli intermedi di energia.

L'energia di orientazione così calcolata sarà tanto più elevata quanto più l'orientazione dello stimolo sarà solidale con quella della coppia di cellule semplici (modellate dalla parte reale ed immaginaria di un filtro di Gabor con una determinata orientazione).

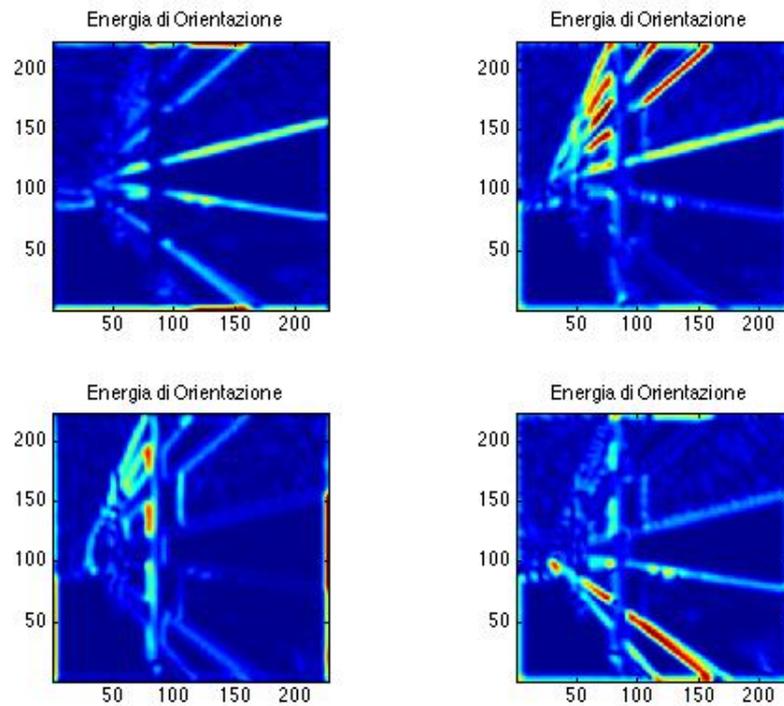


Figura 3.11: Rappresentazione dell'energia di orientazione. Colori tendenti al rosso rappresentano alti valori di energia, mentre colori tendenti al blu indicano valori bassi. I colori intermedi tra i due indicano valori di energia intermedia.

3.3 Artefatti ai bordi

Quando si eseguono elaborazioni d'immagini è possibile che si presentino errori ed artefatti indesiderati. Durante le elaborazioni viste nella sezione precedente mi sono imbattuto su “artefatti ai bordi” nati dalla convoluzione tra i filtri di Gabor e l'immagine stessa e, successivamente, anche nel calcolo dell'energia di orientazione. Come è possibile vedere dalla figura 3.13 ed in modo particolare per $\theta = 0$ e $\pi/2$, il filtro ha riconosciuto come bordo orientato anche il contorno della figura. Quando deve calcolare i pixel vicini ai quattro bordi (due laterali, uno superiore ed uno inferiore) dell'immagine, la maschera del filtro si trova in parte “fuori” dal dominio dell'immagine. Matlab utilizza la tecnica dello “zero padding” che pone a zero il valore dei pixel in quelle posizioni. Questo valore potrebbe discostarsi in modo importante dal valore che hanno pixel sul bordo dell'immagine e quindi essere rilevato un contorno anche se in realtà non lo è. Per ovviare a questo problema esistono varie soluzioni, le più utilizzate sono:

I : “*symmetric*”: consiste nello “specchiare” l'immagine originale lungo tutti i suoi bordi in modo che i pixel al di fuori dei limiti dell'immagine abbiano valori molto simili a quelli nelle immediate vicinanze.

II : “*replicate*”: i valori mancanti sono sostituiti da quelli dell'immagine stessa per prolungamento periodico. In pratica si “piastrava” il piano con l'immagine e quando il filtro passa sui bordi si considerano i pixel dell'immagine adiacente. Nell'immagine 3.12 sotto è possibile vedere l'applicazione di questa tecnica.

Non esiste una tecnica migliore dell'altra, caso per caso si deve valutare quale utilizzare.

Per cercare di mettere in pratica questi concetti, ho preso l'immagine di figura 3.7 e ho calcolato l'energia di orientazione. Come detto nella sezione precedente, l'energia di orientazione da una risposta maggiore

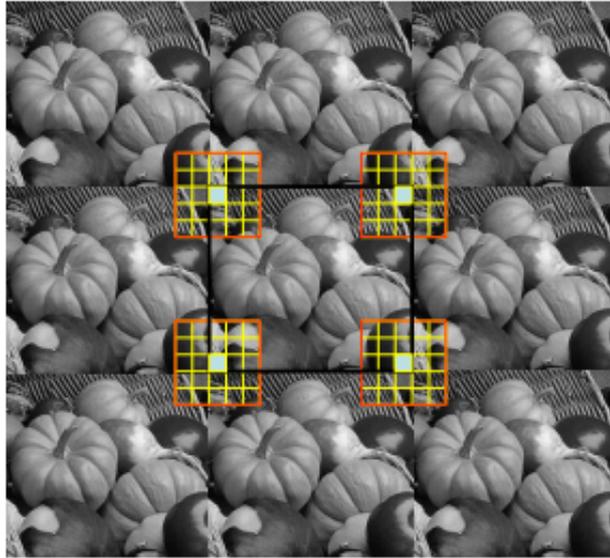


Figura 3.12: Applicazione della tecnica “*replicate*”.

nelle zone in cui è possibile riscontrare linee o bordi diretti in certe orientazioni. Se i quattro bordi dell’immagine non vengono trattati in modo appropriato, è possibile che durante l’elaborazione vengano scambiati per contorni: linee orizzontali ($\theta = 0$ e π) e verticali ($\theta = \pm\pi/2$), ed ottenere in queste zone una risposta (rettangoli rossi in figura 3.13).

Applicare le tecniche “*symmetric*” e “*replicate*” appena viste per calcolare l’energia di orientazione, permette di attenuare o ridurre notevolmente il problema degli artefatti ai bordi.

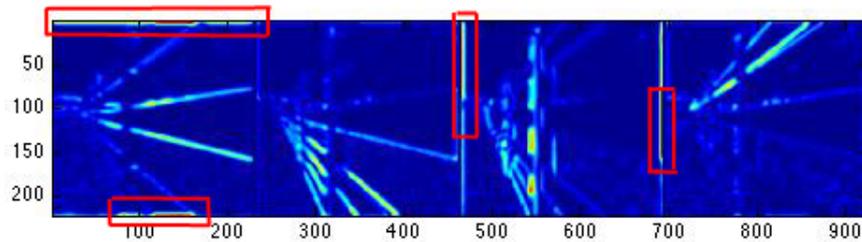


Figura 3.13: Calcolo dell'energia di orientazione sull'immagine 2.12 utilizzando la pura convoluzione. L'elaborazione riconosce un bordo anche sul contorno della figura (rettangoli rossi).

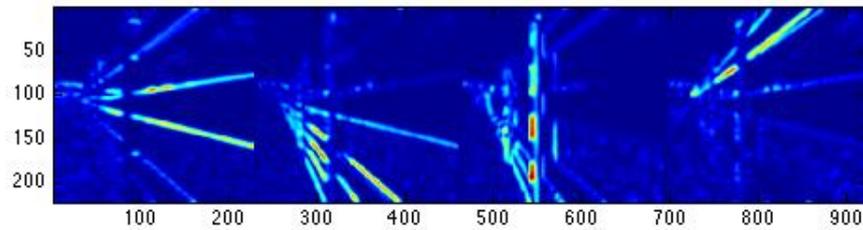


Figura 3.14: Calcolo dell'energia di orientazione utilizzando la tecnica "*symmetric*". Come è possibile vedere, gli artefatti sui bordi che erano presenti in nella figura 2.18 sono spariti.

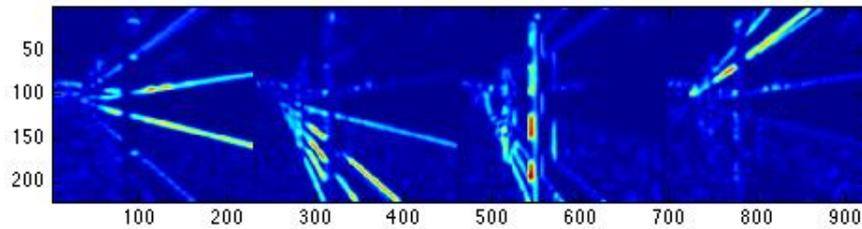


Figura 3.15: Energia di orientazione calcolata utilizzando la tecnica “*replicate*”. Anche in questo caso, gli artefatti ai bordi vengono enormemente attenuati.

3.4 La scale space

I campi recettivi non sono tutti uguali e, soprattutto, non hanno tutte le stesse dimensioni. Quelli più piccoli sono utilizzati per una ricostruzione ad alta risoluzione, mentre i più grandi scattano una “foto” sfocata del mondo esterno. Chiameremo la dimensione del campo recettivo la sua *scala*. L’immagine che entra dall’occhio e arriva alla retina è contemporaneamente presentata su più livelli. Ogni scala è campionata come una nuova misura indipendente dalle altre, ma devono comunque essere disponibili nello stesso momento. La retina misura un certo numero di immagini: una “*scale space*” (x, y, θ) .

Abbiamo già visto nel capitolo 1 che la densità dei campi recettivi non è omogenea. Il principio di invarianza della scala dice che “tutte le scale devono essere prese in modo identico, non c’è alcuna preferenza di scala”. Questo significa che possiamo aspettarci la stessa capacità di elaborazione per ogni scala. Questo si riduce ad una pari quantità di campi recettivi per ognuna di esse. La sovrapposizione più densa in 2D di aree circolari uguali è una matrice esagonale. Poichè gli occhi possono muoversi, è più logico inserire le matrici esagonali di campi recettivi tutti nel centro della retina, sovrapposti l’uno sull’altro in modo da ottenere tante sovrapposizioni. La piastrella esagonale dei campi recettivi con più piccola scala costituisce la fovea. I campi recettivi con scala leggermente più grande hanno altrettanti campi recettivi ma, naturalmente, fanno riferimento ad un’area leggermente più grande. Così via, fino a quando non si arriva alla scala maggiore, i cui campi recettivi coprono l’intera area disponibile della retina^[20].

Il modello afferma che viene fatto un campionamento simultaneo dell’immagine a tutte le scale utilizzando l’“invarianza di scala”, cioè il campionamento viene fatto senza alcuna preferenza per una particolare scala. In questo modo quello che viene misurato e inviato al cervello non è una singola immagine, ma una sequenza (pila) di immagini. La prima fase del front-end visivo è quindi un dispositivo di campionamento multi-scala. Questo modello è in buon accordo con le recenti misure delle dimensioni dei campi recettivi.

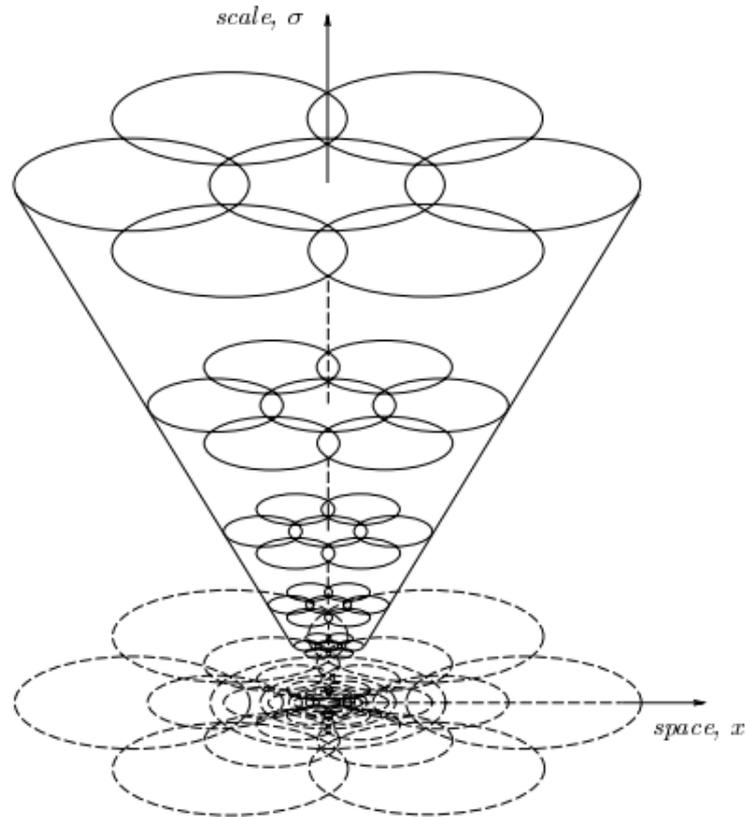


Figura 3.16: Illustrazione schematica della distribuzione dei campi recettivi su tutte le scale per il sistema della fovea. E' possibile vedere la relativa sovrapposizione tra i campi recettivi vicini fissati a differenti scale.

Per riprodurre il modo in cui il cervello lavora contemporaneamente su più scale è stato implementato il concetto di “scale-space” (spazio della scala). Questa tecnica prevede che la struttura dell’immagine venga studiata su diverse scale, ottenendo una famiglia di immagini “smoothed”, parametrizzate dalla dimensione del kernel usato. Gli oggetti nel mondo appaiono in modi diversi a seconda della scala di osservazione; proprio il concetto di scala e la nozione di rappresen-

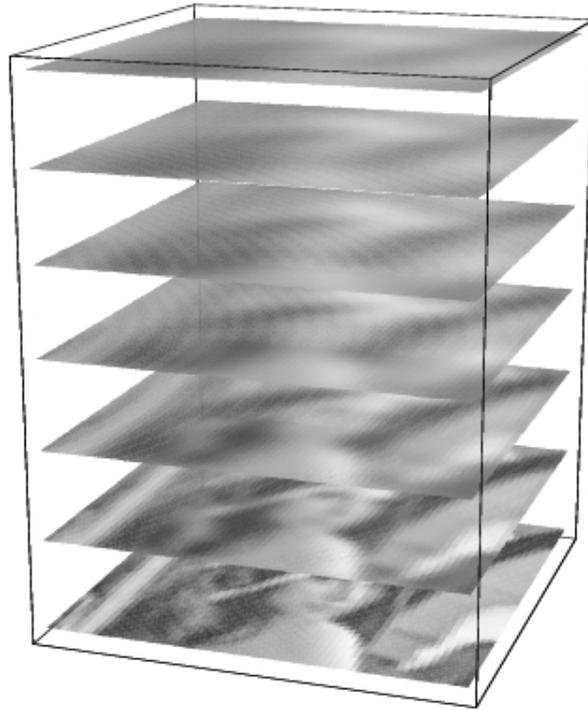


Figura 3.17: La retina non misura una singola immagine, ma una serie d'immagini a differenti scale. La figura mostra in pila l'immagine "Lena" in x, y, θ . La scala (θ) è la dimensione verticale.

zione multi-scala hanno un'importanza cruciale. Quest'ultimo è ben noto nella cartografia; le mappe vengono prodotte a differenti gradi di astrazione. La mappa del mondo contiene grandi continenti, isole e, possibilmente, alcune delle più grandi città, mentre i paesi e le isole più piccole sono mostrate nelle mappe di uno stato. Passando alla guida di una città, il livello d'astrazione cambia considerevolmente e include strade, palazzi... . E' chiaro che per estrapolare ogni tipo di informazione è necessario aver a che fare con certi tipi di operatori. Il tipo d'informazione che può essere ottenuta è determinata dalla relazione tra la dimensione della struttura e la dimensione (risoluzione) degli operatori^[19].

Le applicazioni in cui può essere utile questo tipo di rappresentazione sono tantissime, da macchine e robots con abilità visive, riconoscimento di oggetti, manipolazione di oggetti, navigazione guidata... . Per riprodurre computazionalmente il primo stage del processo visivo è consigliato utilizzare un kernel Gaussiano e le sue derivate. Per semplificare tale processo si rimuovono i dettagli non necessari e di disturbo.

L'idea di rappresentare un segnale multi-scala non è nuovo. Alcuni lavori di Rosenfeld e Thurston (1971) mostravano vantaggi utilizzando operatori con differenti dimensioni per l'individuazione dei bordi. Negli anni successivi sono stati implementati nuovi approcci per rappresentare le immagini, come l'utilizzo di differenti livelli di risoluzione spaziale e del "down-sampling". Queste idee, portate avanti soprattutto da Burt e Crowley, sono alla base di quella che oggi è la tecnica più utilizzata, la "*piramide*".

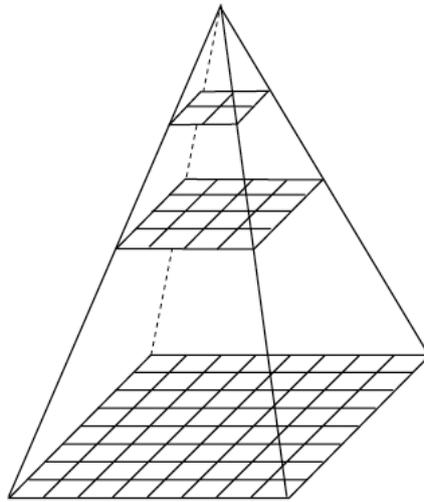


Figura 3.18: Rappresentazione della piramide ottenuta attraverso successive riduzioni delle dimensioni dell'immagine attraverso la combinazione delle tecniche di smoothing e down-sampling.

Il “*down-sampling*” è l’operazione di riduzione della frequenza di campionamento dei segnali. Tale tecnica viene impiegata solitamente per comprimere la velocità di trasmissione o la dimensione dei dati in formato digitale. Se i dati sono di tipo grafico (ad esempio, immagini digitali), l’operazione prende il nome di *down-sampling*, e in quel caso consiste nella riduzione dei pixel che compongono l’immagine.

La maggior parte degli spazi della scala sono di tipo lineare (Gaussiano) con un’ampia gamma di proprietà. Il termine “scale-space” è applicato a segnali con un numero arbitrario di variabili. Il caso più comune è quello di applicarlo ad immagini bi-dimensionali. Presa un’immagine $f(x,y)$, la sua rappresentazione scale-space lineare (Gaussiana) è una famiglia di segnali derivati $L(x,y;t)$ definiti dalla convoluzione di $f(x,y)$ con una kernel Gaussiano bi-dimensionale:

$$g(x, y; t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t} \quad (3.11)$$

così che:

$$L(\cdot, \cdot; t) = g(\cdot, \cdot; t) * f(\cdot, \cdot) \quad (3.12)$$

in cui le prime colonne di “ L ” indicano che la convoluzione avviene solo tra le variabili x e y , mentre il parametro di scala “ t ” indica quale livello di scala è definito. Quest’ultimo rappresenta la varianza del filtro gaussiano ed è simile a un limite: per $t = 0$ il filtro “ g ” diventa una funzione impulsiva di tipo $L(x,y;0) = f(x,y)$, che non è altro che la rappresentazione della scale-space al livello di scala $t = 0$, cioè l’immagine originale f . All’aumentare di “ t ”, “ L ” è il risultato dello sfocamento di “ f ”, rimuovendo un numero sempre maggiore di dettagli dall’immagine. I dettagli che hanno un valore significativamente più piccolo rispetto alla deviazione standard del filtro (\sqrt{t}) vengono rimossi dall’immagine con ha parametro di scala “ t ”. Per chiarire questi concetti sono riportati alcuni esempi.



Figura 3.19: A sinistra è visualizzata l'immagine originale ($t = 0$), a destra l'immagine $L(x,y;t)$ con $t = 1$.



Figura 3.20: Nella prima immagine $L(x,y;t)$ il valore di t è stato impostato a 4, mentre nella seconda immagine $t = 16$. Per questo motivo l'immagine di sinistra risulta più sfocata, ed ha perso un numero maggiore di dettagli rispetto all'immagine originale



Figura 3.21: A sinistra è visualizzata l'immagine a cui è stato applicato $t = 64$; a destra l'immagine $L(x,y;t)$ con $t = 256$ in cui, senza conoscere l'immagine originale è quasi impossibile distinguere anche gli oggetti più grandi.

Le immagini utilizzate negli esperimenti da Ringach erano costituite da frame consecutivi di un video. Per poterle elaborare attraverso la scale-space, queste immagini sono prima state trasformate da RGB a scala di grigi, per poi essere implementate attraverso l'analisi scale-space. Per prima cosa, ogni frame è stato convoluto con un kernel Gaussiano normalizzato che agisce come un filtro passa basso. I dettagli presenti nelle immagini si trovano alle alte frequenze; se queste vengono filtrate è possibile utilizzare regole meno stringenti per campionare il segnale (Th. Shannon): saranno quindi sufficienti meno campioni per rappresentare l'informazione. Se il filtro passa basso è una Gaussiana ad una determinata scala questo permette di campionare, per esempio, un elemento ogni due e di ottenere un'immagine di dimensioni dimezzate ma con la stessa quantità d'informazione. In prima approssimazione, è possibile vedere una funzione di Gabor come la derivata direzionale di una Gaussiana, per questo motivo filtrare un'immagine con un filtro di Gabor è come calcolare la derivata direzionale dell'immagine stessa. Se invece di derivare e filtrare un'immagine con una Gaussiana la si convolve con una funzione di Gabor si ottiene quindi lo stesso risultato. Calcolare diverse scale significa convolvere l'immagine con Gaussiane sempre più grandi (oppure convolverla con una Gabor sempre più grande) per poi fare la derivata. Per motivi puramente computazionali però, conviene invece di ingrandire il filtro, diminuire le dimensioni dell'immagine. Dal risultato di questa convoluzione sono state prese solamente le righe e le colonne dispari.

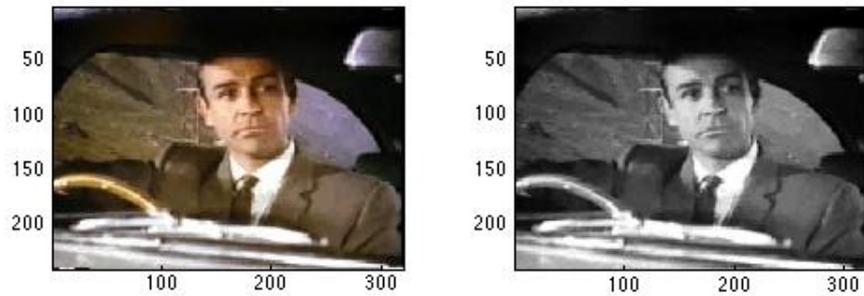


Figura 3.22: Sulla sinistra immagine originale di un frame, sulla destra la stessa immagine proposta in scala di grigi.

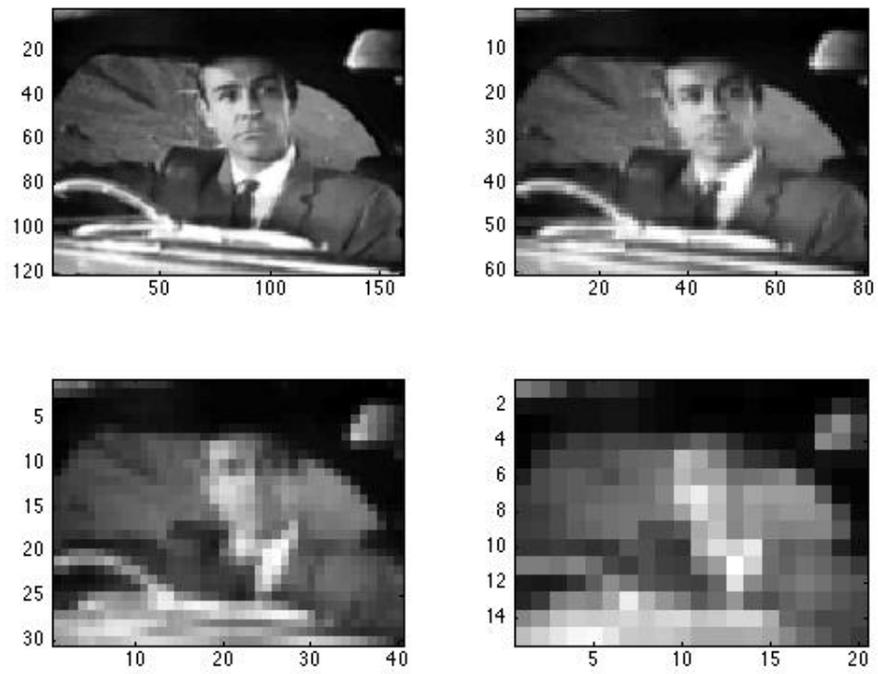


Figura 3.23: Immagine a cui è stato applicato il filtro Gaussiano e poi ridotta attraverso il metodo dello “*scale-space*”.

Attraverso l'utilizzo delle tecniche viste fin qui, si può modellare la risposta di cellule della corteccia visiva primaria. Come detto precedentemente, le cellule complesse prendono in ingresso input provenienti da una coppia di cellule semplici con la medesima orientazione preferenziale del profilo recettore, ma in quadratura di fase tra loro. La risposta di queste cellule è modellabile attraverso i filtri di Gabor, mentre attraverso l'energia di orientazione (che non è altro che la somma dei quadrati degli output provenienti da cellule semplici) è possibile modellare il comportamento di cellule complesse in corteccia. Non sapendo però le dimensioni del campo recettivo di queste cellule è necessario utilizzare più scale.

Il mio lavoro si è quindi concentrato nel convolvere tutti gli stimoli utilizzati da Ringach con la parte reale ed immaginaria dei filtri dei Gabor. Da questa elaborazione ho potuto calcolare l'energia di orientazione e valutarla per diverse scale. Le immagini ottenute sono state salvate su disco come "Gabor Pyramid" e rappresentavano la risposta di cellule complesse ai vari stimoli, ognuna valutata per diverse scale e orientazioni.

Attraverso l'utilizzo di metodi di analisi statistica, che saranno analizzati nella prossima sezione e applicati alle "Gabor Pyramid", si vuole ora andare ad individuare quale zona delle immagini ha eccitato la cellula.

3.5 Lo Z-score

Come detto alla fine del capitolo precedente, essendo i dati ad alta dimensionalità, non possiamo valutare in modo diretto su di essi la reverse correlation. Serve quindi un approccio che riduca questa dimensionalità senza perdere informazione ed ora che le immagini sono state valutate su più scale, si vuole andare ad individuare quella zona dell'immagine che ha prodotto una risposta cellulare. Per fare questo è stato utilizzato una semplice tecnica statistica, chiamata standardizzazione o *Z-score*.

La “standard score” è il numero di deviazioni standard di un’osservazione o di un dato che si trova sopra la media. Così una standard score positiva rappresenta un gruppo di dati al di sopra della media, mentre una standard score negativa rappresenta dei dati al di sotto della media. E’ una quantità adimensionale ottenuta sottraendo la media della popolazione dal valore reale e dividendo il risultato per la deviazione standard della popolazione. Questo procedimento viene chiamato “Z-values”, “Z-scores”, “normal scores”, o “standardized variables”; l’uso della “Z” è dovuto alla distribuzione normale che è anche conosciuta come “Z distribution”. Viene usata frequentemente per ricondurre una variabile aleatoria distribuita secondo una media “ μ ” e varianza σ^2 , ad una variabile aleatoria con distribuzione “standard”, ossia di media zero e varianza pari a 1.

Il procedimento prevede di sottrarre alla variabile aleatoria la sua media e dividere il tutto per la deviazione standard (per σ e non per σ^2), ovvero utilizzando la formula utile a trovare i punti zeta (Z-score):

$$Z = \frac{X - \mu}{\sigma} \quad (3.13)$$

dove μ è la media e σ la deviazione standard della popolazione.

Il valore assoluto di “Z” rappresenta la distanza tra il valore naturale e la media della popolazione calcolata in deviazione standard. Z è negativo quando il valore naturale è al di sotto della media, altrimenti è positivo.

Per individuare la regione che ha eccitato la cellula si sono prese le immagini valutate con scala e orientazione (Gabor Pyramid) ed applicata questa tecnica di standardizzazione. E’ stato utilizzato cioè un modello di risposta neurale insieme ad una tecnica di analisi statistica. Attraverso questa tecnica, è stata valutata la STA delle “Gabor Pyramid”, prendendo solo le immagini che avevano eccitato la cellula. A questa è stata sottratta una STA di “Gabor Pyramid” però valutata con stimoli random. Infine il risultato ottenuto è stato diviso per la deviazione standard calcolata sulle “Gabor Pyramid” random. In pratica, è stata fatta una STA decorrelata dall’energia di orientazio-

ne. Questo ha permesso di trovare la regione dell'immagine a cui il neurone complesso risponde (e la scala, e l'orientazione) a prescindere da tutte le fasi dello stimolo che lo possono eccitare.

Riassumendo, per raggiungere lo scopo di questa tesi, sono state utilizzate:

- la STA per trovare la regione d'interesse;
- il modelli di energia di orientazione per varie scale e orientazioni per modellare il comportamento di cellule complesse;
- la standardizzazione statistica per avere una misura il più possibile "neutrale" e vedere quanto la popolazione "scelta" si discostava dalla popolazione "media".

Un esempio di elaborazione attraverso la tecnica della Z-score è riportato in figura 3.24 dov'è possibile individuare la zona di massima risposta.

Nella prossima sezione saranno invece valutati e riportati i risultati ottenuti da questa elaborazione.

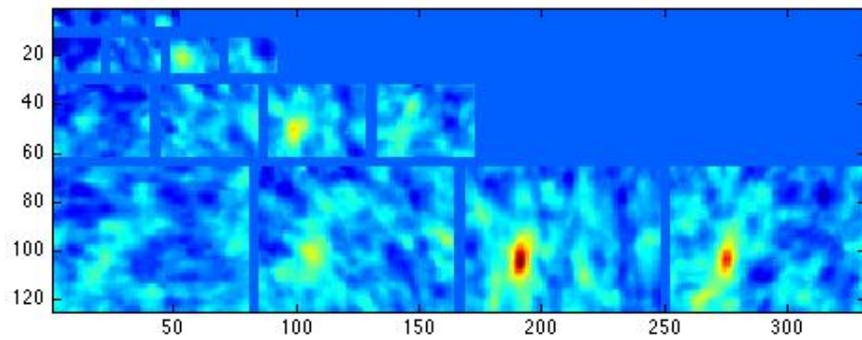


Figura 3.24: Risultato dell'elaborazione delle immagini con Z-score per l'esperimento 1, elettrodo 46, neurone "a".

3.6 Risultati ottenuti

In questo capitolo verranno mostrati i risultati ottenuti dall'elaborazione dei dati fatte seguendo le tecniche viste in questo capitolo. Sfruttando i filtri di Gabor, l'energia di orientazione, la scale space e la Z-score è stato possibile modellare la risposta agli stimoli presentati a cellule della corteccia visiva primaria valutando più scale e diverse orientazioni. Gli stimoli utilizzati sono quelli presentati 100 ms prima della registrazione dello spike.

L'energia di orientazione ha permesso di modellare la risposta delle cellule presenti in corteccia visiva, valutata per differenti scale e orientazioni. Sfruttando poi una tecnica di standardizzazione (Z-score) è stato possibile decorrelare la STA dell'energia di orientazione. Questo ha permesso di individuare la zona dell'immagine in cui il profilo recettore è collocato (obiettivo della tesi). Come si può notare dalle figure che saranno mostrate in seguito, è possibile distinguere facilmente delle zone ben definite ad elevata energia (di colore giallo-rosso), che rappresentano proprio la regione dell'immagine che ha prodotto un'eccitazione da parte della cellula. Tutta la restante parte dell'immagine, di colore azzurro-blu è a bassa energia, e rappresenta quelle zone che non hanno (o hanno ma con poca frequenza) eccitato i neuroni.

All'interno di ogni figura è possibile distinguere sedici immagini che, partendo dall'apice in basso a sinistra rappresentano: lungo l'asse x quattro differenti orientazioni valutate per $\theta = 0, \pi/4, \pi/2$ e $3/4\pi$; lungo l'asse y quattro differenti scale analizzate con parametro di scala "t" = 4, 8, 16, 32.

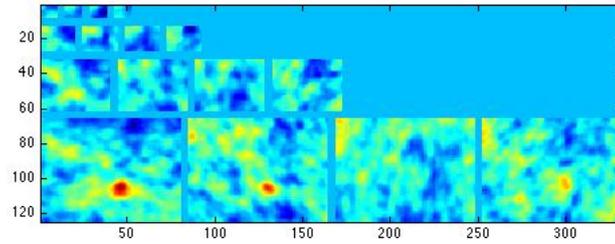


Figura 3.25: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 5, neurone "a".

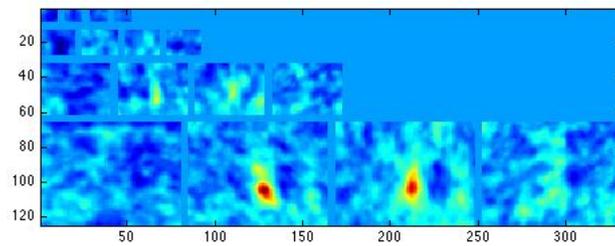


Figura 3.26: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 15, neurone "a".

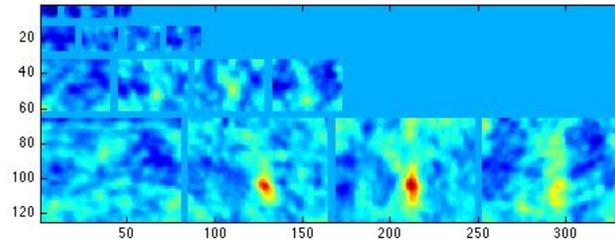


Figura 3.27: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 15, neurone "c".

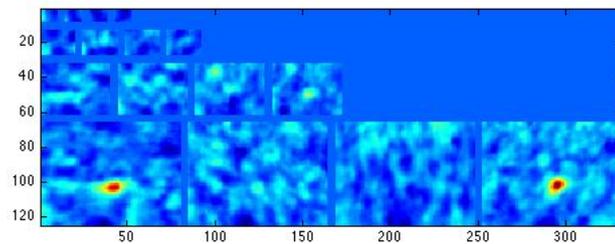


Figura 3.28: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 22, neurone "a".

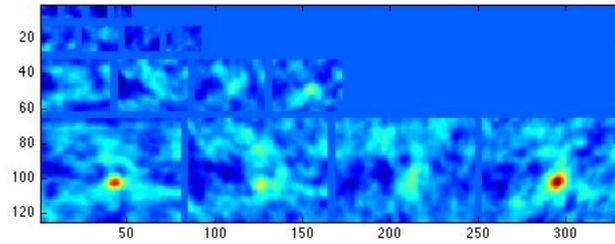


Figura 3.29: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 22, neurone "b".

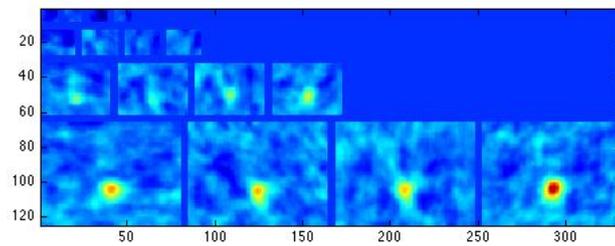


Figura 3.30: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 26, neurone "b".

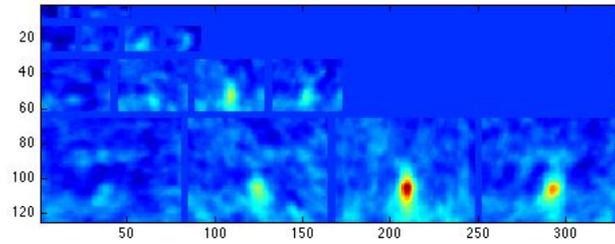


Figura 3.31: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 47, neurone "b".

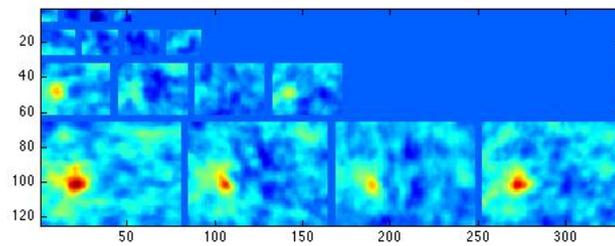


Figura 3.32: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 13, neurone "b".

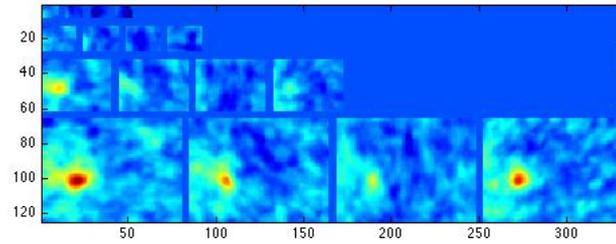


Figura 3.33: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 13, neurone "c".

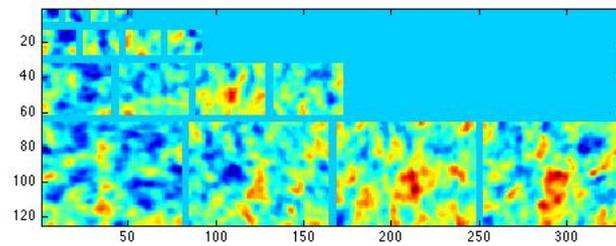


Figura 3.34: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 10, neurone "b".

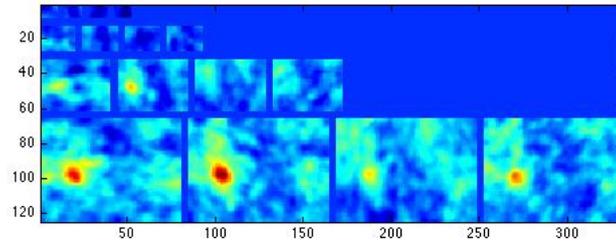


Figura 3.35: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 20, neurone "b".

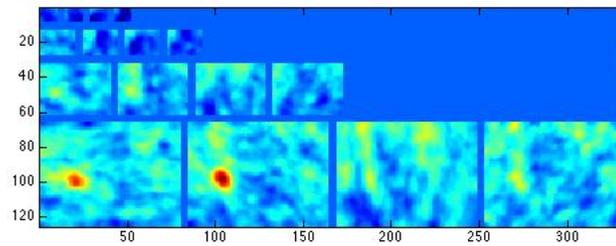


Figura 3.36: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 20, neurone "c".

Non tutte le elaborazioni hanno però portato a buoni risultati evidenziando una regione con contorni così ben definiti. Bisogna tenere in considerazione che con il clustering sono stati raggruppati gli spike di uno stesso neurone ma, come descritto nella sezione relativa a questa tecnica, è spesso presente anche una componente di rumore. Tra i vari gruppi formati per ogni esperimento, saranno presenti errori di raggruppamento, rumore, errori dovuti all'elaborazione (per esempio dovuti ad una scelta non ottimale della soglia), che non restituiscono un risultato significativo.

Nelle figure sottostanti saranno mostrati questi risultati. Si possono notare zone a più alta energia di colore giallo rosso mischiate a zone ad energia più limitata di colore azzurro-blu mischiate tra loro in modo casuale e senza creare regioni ben definite come quelle presenti nelle immagini viste fin'ora.

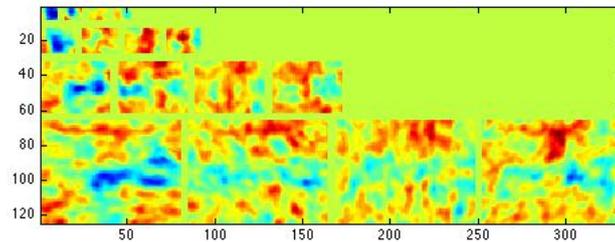


Figura 3.37: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 26, neurone "a"

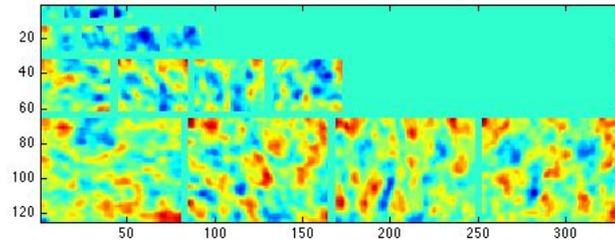


Figura 3.38: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 46, neurone "b"

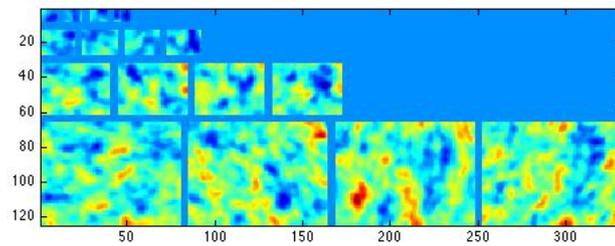


Figura 3.39: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 1, elettrodo 54, neurone "b"

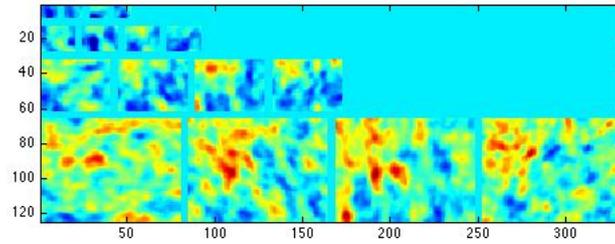


Figura 3.40: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 15, neurone "a"

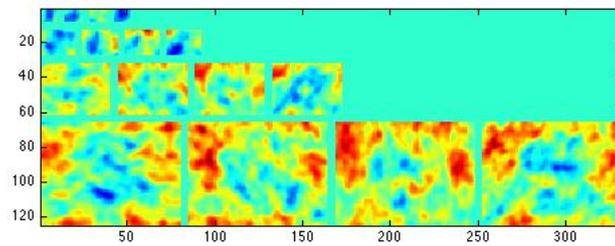


Figura 3.41: Individuazione della regione in cui è posizionato il campo recettivo per l'esperimento 3, elettrodo 29, neurone "a"

3.7 Conclusioni

Vorrei iniziare questa sezione sottolineando che l'**obiettivo** preposto in questa tesi, cioè l'individuazione nello spazio degli stimoli naturali messi a disposizione da Ringach della regione in cui è situato il campo recettivo delle cellule neurali, è stato raggiunto con **ottimi risultati**. Com'è possibile vedere dalle immagini nella sezione precedente sono infatti ben definite le aree in cui si trova il campo recettivo.

Per arrivare a questi risultati si è inizialmente fatta una descrizione anatomica e fisiologica dell'occhio, descrivendo il funzionamento di ogni sua parte. E' stata analizzata l'organizzazione delle cellule corticali e la loro struttura iper-colonnare dipendenti dall'orientazione. E' stato poi introdotto il concetto di campo recettivo analizzando le sue principali caratteristiche di organizzazione e separabilità focalizzando questo concetto sulle cellule della corteccia visiva primaria (V1).

Nel capitolo seguente è stato introdotto il metodo con cui si ricavano questi dati neuro-fisiologici (profil recettori). Sono quindi stati introdotti i modelli per descrivere gli stimoli naturali e in seguito illustrata la tecnica della reverse correlation e due sue varianti: la spike triggered average e la spike-triggered covariance. Inoltre, attraverso la trasformazione di pre-whitening utilizzata prima di applicare la reverse correlation, è possibile decorrelare i dati.

Gli stimoli e le registrazioni elettro-fisiologiche utilizzate in questa tesi sono stati messi a disposizione da Ringach e il suo staff. Il clustering (non sviluppato in questa tesi) ha permesso di associare tra tutti gli spikes registrati quelli dovuti all'eccitamento di uno stesso neurone. Una volta organizzati i dati, si volevano trovare i campi recettivi dei neuroni della V1. Come ci si aspettava non è stato possibile applicare la reverse correlation sui dati in quanto la STA modella il comportamento di cellule semplici il cui comportamento può essere descritto con una funzione lineare. Non essendo noto a quali tipi di neuroni facevano riferimento le registrazioni utilizzate, questo approccio non poteva essere utilizzato. Per quanto riguarda la STC, invece, non è

stato possibile utilizzarla a causa dell'alta dimensionalità degli stimoli; la richiesta di memoria era troppo elevata.

Per ovviare a questo problema si è scelto di provare a ridurre la dimensionalità degli stimoli, andando ad individuare proprio su queste immagini (320x240 pixels) la regione che ha eccitato la cellula. Per prima cosa, sono stati studiati e analizzati i modelli che riproducono il comportamento di cellule semplici, i filtri di Gabor, e di cellule complesse, l'energia di orientazione. Non conoscendo la dimensione e l'orientazione del campo recettivo, l'analisi dei dati è stata fatta su diverse scale e per differenti orientazioni.

Sfruttando infine l'analisi statistica della Z-score, sono stati ottenuti ottimi risultati che mostrano in modo chiaro la zona in cui il profilo recettore è collocato. E' possibile notare questo nelle figure della sezione precedente.

A questo punto i possibili studi futuri possono seguire due strade: una che prevedere di continuare l'elaborazione dei dati, l'altra invece, tenta di migliorare i risultati ottenuti e prova ad utilizzare metodi differenti. Nel primo caso, una volta individuate le zone d'interesse, queste devono essere ritagliate per ottenere nuove immagini a ridotta dimensionalità. Bisogna considerare che le immagini su cui sono stati ottenuti i risultati sono state precedentemente "scalate", per questo, l'area del ritaglio va aggiustata utilizzando un "fattore moltiplicativo" per ottenere le dimensioni corrette sull'immagine originale. Si otterrà così una nuova immagine, a dimensionalità notevolmente ridotta, su cui è possibile applicare la STC. In questo modo sarà richiesta una potenza computazionale di molto inferiore. Se per esempio considerassimo la nuova immagine di dimensioni 50x50 pixels, si otterrebbe una matrice di covarianza "C" 2.500x2.500 che, se salvata come "single" (cioè con 4 byte), richiederebbe una memoria di appena 25 Mb. La seconda strada è quella di un'analisi più minuziosa e accorta su alcune tematiche trattate, in cui, a causa della mancanza di tempo, inesperienza o perchè comunque una scelta andava fatta, è possibile fare analisi più approfondite, utilizzando metodi differenti (non per forza migliori) o strade alternative. Per esempio, la tecnica di cluste-

ring sulla quale si è basata questa tesi (superparamagnetic clustering) non è l'unica esistente, e non è stata implementata in questo lavoro di tesi. E' possibile cercare soglie migliori che riescano a distinguere più chiaramente il segnale utile dal rumore. Inoltre la tecnica utilizzata non prende in considerazione l'influenza che un neurone può avere su quelli vicini e le interazioni che lo sparo di uno può provocare su quelli a cui è connesso. In questo modo si considera il neurone come unità singola e non come parte di un gruppo di unità che interagiscono e lavorano in sinergia (come si presume) tra loro. Un altro esempio riguarda lo stimolo che ha eccitato la cellula; in questa elaborazione, è stato considerato lo stimolo presentato 100 ms prima della presentazione dello spike^[29]. Questo valore, preso dalla letteratura, non è valido per tutti i neuroni, ma è un valore relativo alle cellule della corteccia visiva primaria. In alternativa si potrebbero utilizzare un treno di stimoli appartenenti ad un intervallo di tempo più lungo prima della presentazione dello spike, mediati tra loro.

Un'altra idea per sviluppi futuri, potrebbe essere di aggiungere altre variabili come tempo e velocità a quelle utilizzate (scala e orientazione) per individuare la posizione del profilo recettore con maggiore precisione.

Anche se sicuramente sono necessari studi e ricerche ulteriori per comprendere maggiormente nel dettaglio la struttura e il funzionamento di profili recettivi neuronali, i risultati ottenuti da questo lavoro di tesi fanno ben sperare per una non lontana e migliore comprensione di questo astratto ma affascinante mondo che riguarda la visione.

Capitolo 4

Ringraziamenti

Dopo anni di duro lavoro e di sacrifici sono finalmente giunto alla “*FINE*” di questo percorso. I miei primi ringraziamenti vanno alla mia famiglia che oltre al sostegno economico non mi ha MAI fatto mancare il suo sostegno appoggiando SEMPRE le mie scelte.

Un ringraziamento al Professor Alessandro Sarti che prima mi ha entusiasmato a lezione di “bioimmagini” e poi mi ha permesso di addentrarmi in questa affascinante materia e di realizzare questo lavoro di tesi. Un grossissimo ringraziamento anche all’Ing. Giacomo Cocci che con costanza e pazienza mi ha guidato per tutti questi mesi e traghettato fino qui.

Per rimanere nell’ambiente voglio ringraziare un altro ingegnere, MARTA, compagna di vita, di studi e di esami =) Grazie per avermi aiutato in questo ultimo periodo di stress, e soprattutto di aver sopportato i miei continui cambi d’umore, le mie pazzie e i miei momenti isterici. Grazie per avermi portato al fiume a Santa per farmi rilassare, in fiaschetta a distrarmi con un buon bicchiere di vino e a conoscere la tua magnifica famiglia e tanti nuovi e cari amici (Cate, Ceci, Matti, Ari, Luca, Fabar...). Spero che da oggi inizino per noi tante nuove avventure...qui, a Bertinoro, ou dans la belle ville de Paris.

Un ringraziamento a tutti gli amici che mi hanno fatto sorridere e divertire, facendomi sentire meno pesanti anche i periodi più duri.

Grazie ai compagni di studio (e di alfa) che si sono dimostrate perso-

ne stupende trasformandosi in amici. Non posso dimenticare gli amici della triennale che continuo tutt'oggi a vedere e sentire a cui si sono aggiunte nuove amicizie da tutta Italia nate durante la magistrale Andrea, Daniele, Gaetano, Lisa, Roberta, Matteo.... e tutti i ragazzi che ho avuto la fortuna di conoscere in questi anni.

Un grazie anche a Ingegneria che, nonostante mi abbia fatto passare nottate in bianco, mi ha insegnato a non mollare davanti alle difficoltà e che con l'impegno è possibile superare ogni ostacolo (esame).

Grazie a tutti quelli che anche solo per qualche secondo mi sono stati vicino in questo percorso e che, per emozione, distrazione e per limiti di spazio ho dimenticato in questi ringraziamenti.

GRAZIE DI CUORE A TUTTI.

Voglio terminare ricordando qualche evento che, tra 30 anni, quando rileggerò queste righe mi faranno sorridere.

Oggi 27 marzo sono ZIO da poco più di 4 mesi di Pietro <3 nato da Luca e Chiara. Un mese prima circa è nata Carlotta...BENVENUTI. Benvenuto anche a Edoardo di Orlando e Silvia.

Tra poco tempo, invece, Giovanni e Valeria, Simone e Monia e Alessandro e Cecilia diventeranno genitori di Amalia, Diego e presumibilmente.....di Giulio. Anche Riccardo e Valentina diventeranno genitori, abbiamo appena scoperto il sesso (M) ma per il nome è ancora troppo presto.

Matteo e Daniela si sposano.

Io e Marta andiamo a vivere a PARIGI.....

.....**E QUESTO E' SOLO L'INIZIO!!!!**

Bibliografia

- [1] “A computational theory of visual receptive fields”, Tony Lindeberg. Department of Computational Biology, School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden. (2013)
- [2] “The two-dimensional spatial structure of simple receptive fields in cat striate cortex”, Jones J P, Palmer L A. . J Neurophysiol 58: 1187-1210. (1987).
- [3] “Grammatica del vedere”, Kanisza G. . Il mulino, Bologna (1980).
- [4] “Discharge patterns and functional organization of mammalian retina”, Kuer S. W. J. . Neurophysiol 16(1): 37-68. (1953).
- [5] “The primate retina contains two types of ganglion cells, with high and low contrast sensitivity”, Kaplan E., Shapley R. M. .PNAS 83 : 2755- 2757 (1986).
- [6] “Fisiologia, un approccio integrato”, Silverthorn D. U. . (2007).
- [7] “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical

- filters”, Daugman J. G. J. . Opt Soc Am: 1160-1169. (1985).
- [8] “Spatiotemporal energy models for the perception of motion”, Adelson E. H., Bergen J. R. . Journal of the Optical Society of America 2: 284-299. (1985).
- [9] “Receptive-field dynamics in the central visual pathways”, Gregory C. DeAngelis, Izumi Ohzawa and Ralph D. Freeman. Group in Vision Science, School of Optometry, University of California, Berkeley, California 94720-2020.
- [10] “Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences”, Dario L. Ringach, Michael J. Hawken and Robert Shapley. Journal of Vision 2, 12-24. (2002).
- [11] “Theory of communication”, Gabor D. . IEE 93: 429-459. (1946).
- [12] “Eye, brain, and vision”, Hubel D. H. . (1995).
- [13] “Spatiotemporal receptive fields of cells in V1 are optimally shaped for stimulus velocity estimation”, Giacomo Cocci, Davide Barbieri e Alessandro Sarti. J. Opt. Soc. Am. A / Vol. 29, No. 1. (2012).
- [14] “Spatiotemporal energy models for the perception of motion”, Adelson E. H., Bergen J. R. . Journal of the Optical Society of America 2: 284-299. (1985).

-
- [15] “Complete disabbellendoete 2-D Gabor transforms by neural networks for image analysis and compression”, Daugman J. G., *Acoustics, Speech and Signal Processing* 36: 1169-1179. (1988).
- [16] “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters”, Daugman J. G. . *Journal Optical Society of America* 36: 1160-1169. (1985).
- [17] “A review of methods for spike sorting: the detection and classification of neural action potentials”, Michael S Lewicki. Howard Hughes Medical Institute, Computational Neurobiology Laboratory, The Salk Institute (1998).
- [18] “Designing complex Gabor filters”, Bill Christmas. (2007).
- [19] “Scale-space theory: A basic tool for analysing structures at different scales”, Tony Lindeberg. *Journal of Applied Statistics*, vol. 21, no. 2, pp. 225 270. (1994).
- [20] “Front-End Vision and Multi-Scale Image Analysis”, Bart M. ter Haar Romeny. Kluwer Academic Publishers. (2002).
- [21] “Computational Identification of Receptive Fields”, Tatyana O. Sharpee. Center for Theoretical Biological Physics, University of California at San Diego, La Jolla, California 92093.
- [22] “On the relation between receptive field structure and stimulus selectivity in the tree shrew primary visual cortex”, Julia Veit, Anwesha Bhattacharyya, Robert Kretz and Gregor Rainer.

Visual cognition laboratory, University of Fribourg, Switzerland.

- [23] “Spike-triggered neural characterization”, Odelia Schwartz, Jonathan W. Pillow, Nicole C. Rust, Eero P. Simoncelli. *Journal of Vision* 6, 484507. (2006).

- [24] “Modification of Visual Cortical Receptive Field Induced by Natural Stimuli”, Yingjie Zhu and Haishan Yao. Institute of Neuroscience, State Key Laboratory of Neuroscience, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China and Graduate School of Chinese Academy of Sciences, Shanghai, China. (2012).

- [25] “Reverse correlation in neurophysiology”, Dario Ringach and Robert Shapley. *Cognitive Science* 28 147166. (2004)

- [26] “Morphological scale-space in image processing”, Joseph H. Bosworth and Scott T. Acton. Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22904, USA.

- [27] “Spike Triggered Covariance in Strongly Correlated Gaussian Stimuli”, Johnatan Aljadeff, Ronen Segev, Michael J. Berry, Tatyana O. Sharpee. *PLOS Computational Biology*. Volume 9. (2013)

- [28] URL: <http://visiome.neuroinf.jp>

- [29] “Encoding of Binocular Disparity by Simple Cells in the Cat’s Visual Cortex”, Izumi Ohzawa, Gregory C. DeAngelis and Ralph D. Freeman. Group in Vision Science, School of Optometry,

University of California, Berkeley, California 94720-2020.

- [30] “An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex”, Jones J. P., Palmer L. A. . *Journal of Neurophysiology* 58: 1233-1258.
- [31] URL: <http://crcns.org/data-sets/vc/pvc-1>
- [32] “Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering”, R. Q. Quiroga, Z. Nadasdy and Y. Ben-Shaul. *Neural Computational* 16, 1661-1687. Massachusetts Institute of technology. (2004).
- [33] “Accurate spike sorting for multi-unit recordings”, Takashi Takekawa, Yoshikazu Isomura and Tomoki Fukai. *European Journal of Neuroscience*, Vol. 31, pp. 263272. (2010).
- [34] “A review of methods for spike sorting: the detection and classification of neural action potentials”, Michael S. Lewicki. *Network: Comput. Neural Syst.* 9 R53R78. (1998).
- [35] “Spatio-temporal models of cell receptive fields in the primary visual cortex”, G. Cocci.
- [36] *www.xcorr.net*, URL: <http://xcorr.net/2009/10/22/crcns-data-set-pvc-1-by-ringach-lab-getting-something-to-work>
- [37] URL: <http://hubel.med.harvard.edu/book/bcontext.htm>

- [38] URL: http://matlabserver.cs.rug.nl/edgedetectionweb/web/edgedetection_params.html
- [39] “Precise Alignment of Micromachined Electrode Arrays With V1 Functional Maps”, Ian Nauhaus , Dario L. Ringach. *Journal of Neurophysiology* Published. Vol. 97no. 3781-3789. (2007).
- [40] “A simple white noise analysis of neuronal light responses”, Chichilnisky E J. . *Comput Neural Syst* 12: 199-213. (2001).