

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova
Dipartimento di Psicologia Generale

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE PSICOLOGICHE
INDIRIZZO PERCEZIONE E PSICOFISICA
CICLO XXI

Un modello computazionale di integrazione dei segnali di movimento

Direttore della Scuola: Ch.ma Prof.ssa Clara Casco

Coordinatore d'indirizzo: Ch.mo Prof. Umberto Castiello

Supervisore: Ch.ma Prof.ssa Clara Casco

Dottorando: Michele Guerreschi

Indice

0. Introduzione
1. Fondamenti e questioni epistemologiche e metodologiche
 - 1.0 Introduzione
 - 1.1 L'oggetto e i metodi di studio: la percezione del movimento e l'articolazione degli approcci scientifici
 - 1.2 Un approccio metateorico unificante di matrice psicofisica
 - 1.3 Modelli formali e computazionali
 - 1.4 Linee argomentative, euristiche e posizioni epistemologiche
2. La prospettiva psicofisica
 - 2.0 Introduzione
 - 2.1 I segnali statici bidimensionali
 - 2.2 Il concetto di segnale di movimento
 - 2.3 Il flusso ottico
 - 2.4 L'asse concettuale Globale-Locale
 - 2.5 Il campo associativo spaziale
 - 2.6 Il campo associativo spazio-temporale
 - 2.7 Il problema dell'integrazione dei segnali di movimento
3. La prospettiva neurofisiologica
 - 3.0 Introduzione
 - 3.1 Unità di analisi strutturale e procedurale: neurone e potenziale d'azione
 - 3.2 Le rappresentazioni neurali, i processi di coding e decoding, i campi recettivi
 - 3.3 Risposte cellulari specifiche per direzione, frequenza, velocità. Cellule semplici e cellule complesse
 - 3.4 Il campo recettivo come entità intrinsecamente spazio-temporale
 - 3.5 Locale e globale: basi neurofisiologiche del *campo associativo* e struttura colonnare della corteccia striata
 - 3.6 Il campo associativo spazio-temporale
 - 3.7 Le vie neurali della percezione del movimento
4. La prospettiva matematica
 - 4.0 Introduzione
 - 4.1 Gli approcci potenzialmente slegati da vincoli neurofisiologici o psicologici
 - 4.2 I modelli neurofisiologici, psicologici e psicofisiologici
5. Un modello di integrazione dei segnali di movimento
 - 5.0 Introduzione
 - 5.1 Il caso bidimensionale statico
 - 5.1.1. Il flusso *feedforward* dall'input alle cellule complesse
 - 5.1.2. Il flusso tra le cellule complesse: connessioni laterali e corto e lungo raggio
 - 5.2 L'estensione del modello allo spazio-tempo tridimensionale
 - 5.3 Alcune simulazioni
 - 5.3.1. Simulazione 1. Onde
 - 5.3.2. Simulazione 2. Stimoli *à la Michotte*
6. Conclusioni

§0. Introduzione

Il fine di questo scritto è la costruzione di un modello matematico computazionale che si occupi dell'integrazione dei segnali visivi di movimento e sia plausibile dal punto di vista psicofisico e neurofisiologico.

Il primo capitolo verte su alcune questioni epistemologiche. Lo scopo principale è di introdurre un approccio formale che permetta una trattazione unitaria, nella spiegazione della percezione del movimento, dei diversi livelli empirici coinvolti: fisico, psicofisico e neurofisiologico.

Nel secondo capitolo sono introdotti alcuni concetti fondamentali di natura psicofisica, tra cui spicca quello di campo associativo.

Il terzo capitolo è dedicato alla trattazione del livello neurofisiologico. Particolare enfasi è posta sulla possibilità di considerare il campo recettivo delle cellule semplici e complesse come “un'entità intrinsecamente spazio-temporale”.

Il quarto capitolo concerne l'aspetto formale. Principalmente sono trattati i modelli classici di percezione del movimento, di cui il modello oggetto di questa tesi aspira ad essere naturale estensione.

Nel quinto capitolo, infine, è presentato il modello. Partendo dal caso tipico di elaborazione di stimoli bidimensionali, si costruiscono le prime forme di generalizzazione del classico *modello di energia*. Successivamente, grazie ai costrutti introdotti nella esposizione del caso bidimensionale, si passa agevolmente alla trattazione del caso spazio-temporale tridimensionale, in cui il modello accetta filmati come input. Infine si presentano gli esiti di alcune simulazioni.

Nel sesto si sintetizzano le conclusioni.

§1. Fondamenti e questioni epistemologiche e metodologiche

§1.0. Introduzione

Scopo principale del capitolo è proporre le basi da cui partire per costruire e valutare il modello oggetto della tesi. Queste basi riguardano sia l'approccio scelto, che si propone di integrare tutti gli aspetti salienti nello studio della percezione – ovverosia quello psicologico, quello biologico e quello fisico – sia il modo in cui si intende articolarne i contributi: si propone l'uso di uno schema formale e generale di operazionalizzazione. Tale schema è solo abbozzato ma già permetterebbe di enfatizzare le similarità tra i diversi approcci classici e, in un certo senso, si propone come argomento di studio a sé stante: per dirla con una metafora informatica, mira ad essere *piattaforma indipendente*. Come i linguaggi Java e C# sono usati per creare programmi che "girano" su tutte le macchine, così lo schema formale proposto vorrebbe disinteressarsi dei criteri specifici di implementazione (di come saranno misurate le variabili, per esempio; o se queste siano fisiche, biologiche o psicologiche). Chiaramente questo distacco richiede che nello strumento formale siano tradotte le specificità empiriche dei livelli di realtà a cui intende applicarsi o, almeno, che tale traduzione sia in linea di principio possibile.

§1.1. L'oggetto e i metodi di studio: la percezione del movimento e l'articolazione degli approcci scientifici

Se indichiamo con t la misura del tempo in cui osserviamo un oggetto – ottenibile, per esempio, con un cronometro – e con x, y, z , le misure della posizione spaziale, dette *coordinate spaziali*, in cui tale oggetto si trova, eseguibili con un metro, è possibile prendere nota di dove e quando si trovi tale oggetto assegnandogli diverse quadruple del tipo (x, y, z, t) . Una quadrupla siffatta rappresenta lo *stato spazio-temporale* dell'oggetto; si dice *evento* un insieme di stati. L'ovvio principio fisico di non-ubiquità, secondo cui un medesimo oggetto non può stare in diversi luoghi nello stesso tempo ⁽¹⁾, permette di rappresentare un evento come una terna di funzioni, ciascuna delle quali indica una coordinata spaziale in funzione del tempo. Il *movimento* di un corpo, dal punto di vista fisico, è il fatto che esiste almeno una coordinata spaziale che assume valori differenti in istanti differenti. A partire dalla rappresentazione spazio-temporale di un oggetto, è possibile ricavare la descrizione del suo movimento, nei termini di osservabili fisiche come la *velocità* e l'*accelerazione*. La velocità è definita come il rapporto tra la distanza percorsa da un corpo ed il tempo impiegato in tale spostamento. L'*accelerazione*, in maniera simile, è definita come il rapporto tra due velocità possedute dal corpo ed il tempo che intercorre tra le due.

La *meccanica* è la branca della fisica che si occupa di studiare il movimento dei corpi [1]. Essa si suddivide a sua volta in *cinematica* e *dinamica*. Nella prima si enunciano le regole con cui si legano tra di loro le variabili dinamiche (posizione, tempo, velocità e accelerazione) in una prospettiva geometrico-analitica, senza trattare delle cause del moto. Nella seconda si introducono i concetti di massa e di forza con cui si spiega come i corpi interagiscano tra di loro e si determinino le equazioni del moto. Per i nostri scopi sarà opportuno e sufficiente trattare di alcuni concetti fondamentali della cinematica. Ciò che vale la pena enfatizzare è che il movimento è una qualità

¹ Valga una volta per tutte quanto segue. Le nozioni di fisica che saranno introdotte in questo scritto non hanno la pretesa né la possibilità di essere esaustive. Il loro fine è limitato alla descrizione di un modo di rappresentare gli stimoli, quello fisico, appunto. Si tratta di cenni e semplificazioni e, in quanto, tali passibili di critica. L'affermazione secondo cui in fisica non varrebbe il principio di ubiquità, per esempio, potrebbe essere criticata da quanti conoscessero la meccanica quantistica. Così come altre semplificazioni e notazioni più avanti utilizzate.

fisica *osservabile*, ossia qualcosa a cui possiamo associare *misure* tramite opportuni strumenti o ragionamenti che costituiscono il processo di *misurazione*.

Da un punto di vista ecologico ed evolutivo, il valore adattivo della possibilità che un organismo abbia di rilevare il movimento è evidente e più volte sottolineato dai ricercatori [2-4]: basti pensare al ruolo che ha la percezione del movimento per un animale predatore quando caccia o per le api quando, danzando, comunicano la posizione di una fonte di cibo. Per quanto concerne i meccanismi biologici, varie branche delle neuroscienze (neuropsicologia e neurofisiologia cliniche e sperimentali, in particolare) hanno mostrato l'esistenza di una specializzazione anatomico-funzionale nell'elaborazione dei segnali di movimento [5-8].

L'esistenza di meccanismi neuronali specifici per l'elaborazione dei segnali di movimento ed il loro valore adattivo hanno un ovvio corrispondente fenomenico: il movimento è una delle qualità primarie del percepire, come lo sono la forma e il colore [9-11].

Nei tre capoversi precedenti abbiamo sommariamente introdotto le tre componenti salienti dello studio scientifico della percezione del movimento: la prospettiva fisica, quella biologica e quella psicologica. A quelle tre si aggiungono, formando complessi intrecci, gli apporti della matematica e delle scienze dell'informazione e, più in generale, gli approcci formali e modellistici. Alcuni si situano chiaramente nell'area multidisciplinare in cui si trovano anche le neuroscienze: esempi a riguardo sono gli studi della geometria dello spazio visivo [12, 13]; la ricerca di formalizzazione delle capacità geometriche del sistema visivo [14, 15]; la costruzione di modelli connessionisti che intendono simulare parte dei processi percettivi [16-20]. Altri sono di chiara matrice ingegneristico-applicativa e principalmente finalizzati a costruire software che permettano ad automi di interagire in maniera appropriata con l'ambiente *come se* percepissero, piuttosto che a spiegare come il cervello elabori i segnali visivi [21, 22]. È chiaro che vi sono parecchie sfumature ed i confini tra le discipline non sono sempre così netti.

Esiste un approccio teorico e metodologico che permette di organizzare e facilitare l'integrazione delle diverse prospettive succitate pertinenti allo studio della percezione del movimento. Mostriamo che l'approccio in questione deriva da una versione della psicofisica, in cui si rendano espliciti assunti e tendenze della disciplina, e lo utilizzeremo come falsariga meta-teorica.

§1.2.

Un approccio metateorico unificante di matrice psicofisica

Nelle sue linee generali, la psicofisica si occupa di mettere in relazione le misure fisiche e quelle fenomeniche [23, 24]. Il processo di misurazione può essere suddiviso in due componenti; una, quella prettamente fisica, consiste nella interazione tra due o più enti fisici; l'altra, logica, riguarda la concettualizzazione che permette di confrontare diversi eventi siffatti di interazione e di utilizzarli come dati [25-27].

Per fare un esempio di misurazione, prendiamo la temperatura. Per misurarla vanno posti a contatto due corpi, uno dei quali, per esempio una bacinella d'acqua, è il sistema termodinamico di cui si vuole misurare la temperatura e l'altro, per esempio un termometro, è lo strumento di misura. Tale interazione determina un cambiamento nello strumento di misura – un numero che compare su un display digitale o l'espansione o contrazione della colonnina di mercurio – la quale variazione rappresenta l'esito del processo di misurazione, ovverosia nel caso in oggetto la misura della temperatura dell'acqua nella bacinella. Un altro esempio è costituito dalla valutazione neuropsicologica di un soggetto; in questo caso l'interazione coinvolge il paziente e la coppia diagnosta-test. Essa consta di un articolato scambio di informazioni tra il paziente ed il diagnosta, organizzato e vincolato dal test, che permette di ottenere una stringa di dati, per esempio un numero compreso tra uno e dieci che rappresenta la gravità di un certo disturbo; oppure uno di tre o quattro

enunciati che denotano un giudizio clinico (deficit assente, deficit moderato, deficit grave) [28-30]. Un altro esempio, infine, concerne ancora la temperatura, solo che i due enti che interagiscono sono una bacinella d'acqua ed un individuo. In questo caso l'individuo può fornire informazioni sulla temperatura dell'acqua semplicemente dicendo che è calda o fredda; oppure mettendo in ordine di temperatura dalla più fredda alla più calda diverse bacinelle d'acqua appositamente preparate dallo sperimentatore; o, ancora, può eseguire un compito di discriminazione di bacinelle d'acqua secondo la loro temperatura. Ciò che accomuna questi ed altri modi è che il soggetto sperimentale fornisce informazioni – più o meno legate alla sua consapevolezza e volontà – su quale sia il suo stato interno a seguito dell'interazione con la bacinella d'acqua. Attraverso alcuni metodi, detti metodi psicofisici, lo sperimentatore determina il valore di una *variabile fenomenica*, per esempio quale sia la *temperatura percepita* dell'acqua dal soggetto o quale sia la minima differenza di temperatura percepibile dal soggetto. Questo esempio, nella sua semplicità, permette di fare due importanti osservazioni. La prima è che, nel processo di misurazione fenomenica, l'individuo risulta l'analogo del termometro utilizzato nella misurazione fisica. La seconda è che lo sperimentatore che esegue ed utilizza la misura può estendere l'analogia assumendo che il soggetto sperimentale risponda alla temperatura *come se* avesse a sua disposizione una sorta di *scala graduata interna unidimensionale*, simile alla colonnina di mercurio. Tale *continuum fenomenico* è un costrutto teorico che non richiede di supporre che il soggetto ne abbia una rappresentazione cosciente né che a tale “colonnina fenomenica” corrisponda una struttura mentale o neurofisiologica. *Mutatis mutandis*, quanto detto vale anche quando il soggetto sperimentale debba misurare lunghezze o complesse configurazioni spaziali: una eventuale rappresentazione geometrica utilizzata dallo sperimentatore per spiegare le risposte del soggetto non implica che tale rappresentazione *abbia un corrispettivo isomorfo nel cervello del soggetto sperimentale*. Se succedesse che la neurofisiologia dimostrasse l'esistenza di tali rappresentazioni isomorfe (cosa che in effetti è successa: si pensi alla retinotopia della vie visive [5-11]), lo sperimentatore potrebbe allora cercare di mettere in relazione la rappresentazione ottenuta dalla misurazione fenomenica con quella ottenuta dalla misurazione neurofisiologica. Va comunque sottolineato che questo successivo passaggio, benché auspicabile, non è epistemologicamente necessario.

Gli esempi riportati nel capoverso precedente sono centrati sulla interazione tra due entità: l'oggetto da misurare e lo strumento di misurazione; necessariamente è stato introdotto anche un terzo attore, cioè l'osservatore che compie la misurazione. Esso non è, in generale, componente necessaria nel processo di misurazione ma fondamentale per quanto concerne l'acquisizione e l'utilizzo dei dati che scaturiscono dalla misurazione, cioè le *misure*. Secondo la natura della qualità fisica o fenomenica da misurare, l'osservatore utilizza determinate *scale di misura* [31-33], che constano, formalmente, di insiemi strutturati di simboli con cui si rappresentano i diversi possibili esiti di processi di misurazione. In fisica si usano precipuamente i numeri *reali* (il peso di un corpo, la distanza tra due punti, la velocità di un corpo ecc.). In psicologia vi è ampio uso anche di scale non continue (in sostanza di numeri interi) e, addirittura, di scale i cui elementi non sono tra loro comparabili per ordine (dette *scale nominali*).

Riassumendo, ciò che conta degli esempi sopra riportati è l'astrazione che permettono: un *esperimento* consiste nell'interazione, in condizioni controllate, tra un ente o evento (un corpo fisico, un animale, una persona) ed uno strumento di misura che fornisce, appunto, *misure*, cioè *dati*. Tali dati sono concettualizzabili come elementi di un insieme matematico strutturato. Teorie e modelli coinvolgono relazioni tra *dati* e, quindi, tra strutture matematiche. Una misurazione psicofisica consiste nell'ottenere dati che riguardano l'interazione di un soggetto sperimentale con un oggetto o un evento. Una relazione psicofisica – di cui le ben note *leggi psicofisiche* come la *legge di Weber* sono casi particolari – si instaura tra dati fisici e dati psicofisici.

Già Fechner [24-26] distingueva tra psicofisica interna e psicofisica esterna. La seconda è ciò che consideriamo la psicofisica *tout court*; la prima si occuperebbe di argomenti di cui oggi si occupano la psicofisiologia, la neuropsicologia e le neuroscienze in generale: essa

riguarda le relazioni tra dati fisici, dati neurofisiologici e dati fenomenici.

Fatte queste premesse possiamo introdurre lo schema prototipico in cui ha luogo la costruzione e validazione di un modello matematico nelle neuroscienze.

Lo stimolo fisico distale. Riprendiamo la descrizione cinematica con cui abbiamo iniziato il capitolo. Se studiamo il movimento [1] bidimensionale di un oggetto, possiamo denotare con S_ϕ l'insieme dei possibili esiti di misurazione dell'evento-stimolo. Ogni elemento s_ϕ appartenente a S_ϕ rappresenta lo *stato* dell'evento od oggetto fisico come una collezione strutturata di simboli (in genere numeri) ciascuno dei quali è misura di una determinata caratteristica fisica; nella fattispecie la terna $s_\phi = (x_\phi, y_\phi, t_\phi)$ indica dove e quando si trovi l'oggetto in movimento. Come già accennato possiamo affermare che le posizioni spaziali x_ϕ e y_ϕ sono in funzione del tempo e denotiamo questo fatto scrivendo $x_\phi = x_\phi(t_\phi)$ e $y_\phi = y_\phi(t_\phi)$; in cinematica tali funzioni sono dette costituire la *legge oraria* del moto. Per praticità abbandoniamo il pedice ϕ , dando per inteso che stiamo parlando di osservabili fisiche. Prendiamo due istanti distinti $t_1 \neq t_2$ e denotiamo con x_1, x_2 ed y_1, y_2 le rispettive posizioni lungo le due dimensioni spaziali di riferimento; si definisce *velocità media* lungo l'asse x il rapporto $(x_2 - x_1)/(t_2 - t_1)$ e, similmente, *velocità media* lungo l'asse y il rapporto $(y_2 - y_1)/(t_2 - t_1)$. Gli strumenti del calcolo infinitesimale permettono di ottenere una misura *istantanea* della velocità a partire dalla legge oraria. Da un punto di vista operativo, si cerca di misurare le posizioni spaziali x ed y con un'alta frequenza temporale, in modo che si abbiano misure per istanti differenti *molto vicini* tra di loro. Da un punto di vista analitico-geometrico, a partire dalla legge oraria del moto $x = x(t)$ e $y = y(t)$, si definisce la *velocità istantanea* lungo l'asse x , denotata con v_x , come il limite del rapporto $(x_2 - x_1)/(t_2 - t_1)$ per t_2 che tende a t_1 . Tale limite è detto *derivata rispetto a t di $x(t)$* calcolata in t_1 ed è solitamente denotata con $\dot{x}(t_1)$ o $\frac{dx}{dt}(t_1)$. Similmente si ha che la velocità istantanea lungo l'asse y è rappresentata da $v_y = \dot{y}(t_1)$ o $\frac{dy}{dt}(t_1)$. Si chiama *velocità vettoriale* o, più semplicemente, *velocità* – e si indica con \mathbf{v} – la coppia di funzioni (v_x, v_y) . La *velocità scalare* v è il concetto che più si avvicina a quello naïve di velocità ed è la lunghezza del vettore velocità \mathbf{v} , ossia $v = |\mathbf{v}| = \sqrt{v_x^2 + v_y^2}$. Esso rappresenta l'intensità del moto; la direzione del movimento, invece, è data dalla *direzione* del vettore \mathbf{v} : tale direzione è un angolo, che qui denotiamo con θ , e corrisponde all'arco-tangente del rapporto tra v_y e v_x , cioè si ha che $\theta = \arctan\left(\frac{v_y}{v_x}\right)$.

Riassumendo, la precedente caratterizzazione fisica del moto bidimensionale concerne sette osservabili fisiche, ovvero quantità misurabili con strumenti e metodi fisici: il tempo (t), le due componenti della posizione spaziale (x, y), le due componenti della velocità (v_x, v_y) e le due componenti della accelerazione (a_x, a_y). Ciascuna sequenza di queste sette misurazioni è lo stato cinematico istantaneo di un corpo. Un insieme di sequenze siffatte ($t, x, y, v_x, v_y, a_x, a_y$) è un evento. Per determinare un evento non sono necessarie tutte e sette le osservabili: quattro su sette – le due componenti della velocità e le due componenti della accelerazione – possono essere ricavate dalla legge oraria attraverso metodi geometrico-analitici (differenziazione).

Lo stimolo fisico prossimale. L'interazione tra il mondo fisico e quello fenomenico è permessa dagli organi di senso.

Quando si preme il tasto di un pianoforte, si fa vibrare una corda metallica; tali vibrazioni si propagano nell'aria come un'onda in uno stagno quando gli si getti un sasso; tali onde constano di variazioni di pressione dell'aria nel tempo e nello spazio; quando arrivano all'orecchio, determinano la vibrazione di una membrana (il timpano) che, a sua volta, tramite tre ossicini (incudine, martello e staffa), provoca la vibrazione della membrana avvolta all'interno della coclea e la conseguente attivazione delle cellule cigliate ad essa legate. La vibrazione di una corda (lo stimolo fisico distale)

si accompagna, quindi, ad una vibrazione di una articolata struttura biologica (lo stimolo fisico prossimale) che *trasduce* – ovverosia trasforma, codifica – lo stimolo fisico in segnale nervoso. Nel caso della percezione visiva, l'organo di senso è l'occhio e la trasduzione avviene a livello retinico. Lo *stimolo prossimale visivo* è il pattern di radiazioni elettromagnetiche (luce in senso fisico) che giunge sulla retina; dal punto di vista ottico-geometrico tale pattern determina una proiezione bidimensionale del mondo tridimensionale. Nel caso di uno stimolo visivo bidimensionale, se si fa uso di alcune semplificazioni [27], tra cui il considerare la retina piatta e continua, lo stimolo prossimale è una semplice versione in scala dello stimolo distale. Per questo motivo risulta in questa sede marginale la caratterizzazione dello stimolo prossimale.

Lo stimolo fenomenico. Un soggetto che guardi un corpo muoversi è in grado di valutarne la direzione del moto, la velocità, la posizione in diversi istanti ed eventualmente [34] anche l'accelerazione. Attraverso i metodi psicofisici [23-25] è possibile misurare il corrispettivo fenomenico delle osservabili fisiche sopra introdotte. In altri termini è possibile costruire una rappresentazione S_ψ , costituita da terne $s_\psi = (x_\psi, y_\psi, t_\psi)$, della percezione. Quando la qualità percettiva da misurare è il movimento, nella pratica, non si misura tanto la *posizione percepita* nel *tempo fenomenico* dello stimolo quanto, piuttosto, la sua velocità percepita, dalla quale è in linea di principio possibile ottenere la rappresentazione spazio-temporale.

Una funzione psicofisica mette in relazione misure fisiche e misure fenomeniche associate al medesimo referente empirico (²).

Lo stimolo neurale. Come già accennato, attraverso il processo di trasduzione gli organi di senso permettono che il mondo fisico divenga fenomenico: i segnali fisici sono codificati in attività nervosa, la quale è il corrispettivo biologico dell'esperienza (³). L'attività e la struttura nervose sono passibili di essere misurate. Vi sono vari metodi che possiamo a grandi linee distinguere secondo il grado di invasività e la risoluzione spaziale o temporale che implicano [29]: alcuni esempi sono la risonanza magnetica funzionale (fMRI), la magnetoencefalografia, i potenziali evocati, la misurazione in vivo dell'attività di singoli neuroni. Chiaramente i metodi invasivi sono utilizzati solo su animali non umani (tranne che in alcuni casi di misurazione in vivo su pazienti di neurochirurgia, finalizzate a scoprire le funzioni di alcune zone del cervello per limitare i danni dovuto all'asportazione di sue parti [5-8, 29]).

Quale che sia il metodo utilizzato, si fa in genere riferimento a due rappresentazioni. Una riguarda l'attività neurale *di base*, cioè il pattern di attivazione elettrochimica di neuroni singoli o di popolazioni di neuroni. Essa è direttamente legata alla presentazione di uno stimolo e strettamente associata al metodo e allo strumento di misurazione utilizzati. L'analisi di tale configurazione di attività neuronale è finalizzata alla comprensione del processo di *coding* [35], cioè di come le caratteristiche del mondo esterno divengano segnale neurale. L'altra rappresentazione è ricavata dalla precedente attraverso un processo detto di *decoding* [*ibidem*] e si prefigge di individuare i correlati neurali di alcune variabili fisiche che caratterizzano lo stimolo. Per omogeneità di notazione indichiamo quest'ultima con S_v , i cui elementi $s_v = (x_v, y_v, t_v)$ denotano il corrispettivo neurale degli stati fisici $s_\phi = (x_\phi, y_\phi, t_\phi)$ e fenomenici $s_\psi = (x_\psi, y_\psi, t_\psi)$, i quali costituiscono, rispettivamente, l'evento fisico S_ϕ e quello fenomenico S_ψ . Nel paragrafo §3.2 vedremo come sia possibile mettere in relazione le due forme (ottenute attraverso i processi di *coding* e *decoding* succitati) di rappresentazione dell'attività nervosa.

2 La questione della identità del referente nei due processi di misurazione è alquanto complessa. Per una trattazione articolata si vedano [25-27].

3 La relazione tra il mondo fenomenico e l'accadere neurale è assai dibattuta, soprattutto dal punto di vista epistemologico. Oramai si conviene che esista una stretta relazione tra il funzionamento e le strutture del sistema nervoso centrale e l'insieme dei processi e delle esperienze comunemente etichettati con il termine "mente". Sussistono, invece, questioni sulla natura di tale relazione, per esempio se di tipo causale [36, 37].

§1.3.

Modelli formali e computazionali

Nel paragrafo precedente abbiamo enfatizzato la possibilità di rappresentare eventi fisici, esperienza e accadimenti neurali in modi *formalmente simili*. Ciò ha due implicazioni fondamentali. La prima è che ci permette di sottolineare che l'approccio psicofisico classico e gli altri studi di neuroscienze sono parti naturalmente interagenti di una medesima prospettiva nelle scienze della percezione. La seconda è che l'identificazione di un *linguaggio formale* comune di rappresentazione dei dati acquisibili ai diversi livelli – dati fisici, dati fenomenici e dati neurali – ha forti implicazioni teoriche. In un certo qual modo il linguaggio in questione diviene uno strumento di ricerca a sé stante, con cui è possibile creare teorie che hanno implicazioni in tutti e tre i livelli empirici e li vincolano reciprocamente.

Il sintagma “linguaggio formale”, usato al posto di “matematica”, è volontariamente vago, principalmente per i seguenti motivi. Innanzitutto, per quanto mi è dato di sapere, non esiste ancora una prospettiva unitaria e ampiamente accettata ed utilizzata dalla comunità scientifica di riferimento, che si concreti in una sorta di *neuroscienza matematica*, disciplina che si proporrebbe come evidente analogo della sorella anziana *fisica matematica*. In secondo luogo, non solo non vi sarebbe una prospettiva unitaria ma, pure, non esisterebbe una teoria condivisa, articolata e sufficientemente generale da cui partire.

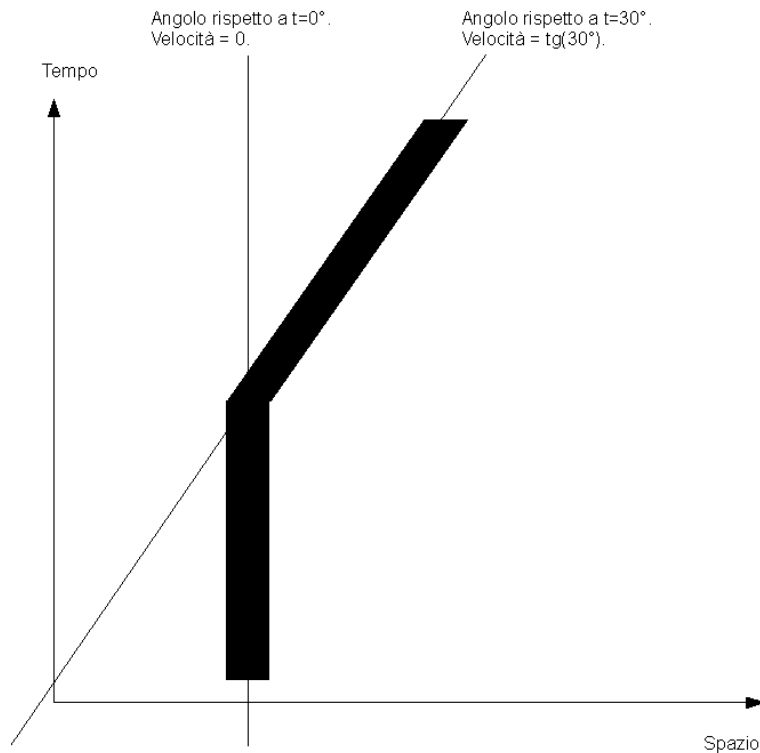
Proprio la mancanza di un ampio terreno comune e di una prassi epistemologica, implica l'esigenza di alcune precisazioni e la costruzione di una falsariga metateorica, seppur provvisoria, con cui dare senso al nostro discorso ⁽⁴⁾.

Abbiamo individuato tre livelli empirici: quello fisico, quello fenomenico e quello neurale. In ciascuno di questi livelli abbiamo fissato un insieme di variabili ed un sistema di rappresentazione, che si articola in regole con cui scegliere e manipolare simboli. La scelta delle variabili è stata fatta in modo che fossero sensatamente considerabili corrispondere nei diversi livelli empirici: si suppone cioè che sia possibile *mettere in relazione*, per esempio, lo spazio-tempo fisico W_ϕ , quello fenomenico W_ψ e quello neurale W_ν , in modo che tale relazione sia applicabile anche ad altre osservabili. Con “mettere in relazione” si intende la costruzione o identificazione di regole (ovvero funzioni) che permettano di legare formalmente le rappresentazioni dei tre livelli: sia quelle all'interno di ciascun livello sia quelle di livelli differenti. Vi devono quindi essere almeno *due* tipi di regole. Per capire meglio perché, dovremmo articolare ulteriormente i criteri di rappresentazione utilizzati per ogni dominio, ma ci limitiamo ad un esempio. Riprendiamo a tal fine il livello fisico dei dati e lo studio cinematico. Identifichiamo anzitutto un *dominio* $W_\phi = X_\phi \times Y_\phi \times T_\phi$, lo spazio-tempo fisico. Si usa porre ciascun asse come l'insieme dei numeri reali \mathbb{R} , cioè si ha che $X_\phi = Y_\phi = T_\phi = \mathbb{R}$ e, quindi, $W_\phi = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$. Ciò fatto, definiamo posizione spazio-temporale o *stato cinematico* s_ϕ un elemento del dominio, cioè una terna di valori (x_ϕ, y_ϕ, t_ϕ) presa da W_ϕ . Attraverso un opportuno sistema di misura possiamo associare ad un oggetto fisico bidimensionale lo stato cinematico s_ϕ se nell'istante t_ϕ l'oggetto (o il suo centro di massa) si trova nel punto di coordinate spaziali (x_ϕ, y_ϕ) . Lo stato di moto bidimensionale di un oggetto – ciò che abbiamo chiamato *evento* – è, quindi, un insieme di stati cinematici, cioè un insieme $S_\phi = \{\dots, s_\phi, \dots\}$ che è sottoinsieme del dominio W_ϕ . L'assioma di non-ubiquità pone un primo vincolo su quali siano i sottoinsiemi di W_ϕ che possono rappresentare eventi e quelli che non possono. Infatti, affinché un insieme S_ϕ possa rappresentare un evento è necessario che al suo interno non vi siano due stati cinematici distinti $s_\phi' = (x_\phi', y_\phi', t_\phi')$ e $s_\phi'' = (x_\phi'', y_\phi'', t_\phi'')$ per i quali si abbia che $t_\phi' = t_\phi''$ ma $x_\phi' \neq x_\phi''$ o $y_\phi' \neq y_\phi''$, cioè, prosaicamente, non deve succedere che un oggetto sia nel medesimo istante ($t_\phi' = t_\phi''$) in luoghi differenti ($x_\phi' \neq x_\phi''$ o $y_\phi' \neq y_\phi''$). Questo vincolo ha una implicazione: permette di affermare che l'evento è rappresentabile da una coppia di funzioni del tempo che, in ogni istante, determinano

4 Quanto scriveremo si basa essenzialmente sui lavori di Burigana [25-27] e né è, secondo i casi, una semplice riproposizione o una naturale estensione.

la posizione spaziale assunta dall'oggetto; formalmente si usa scrivere $x_\varphi = x_\varphi(t_\varphi)$ e $y_\varphi = y_\varphi(t_\varphi)$, cioè che, noto un istante temporale t_φ' , è possibile determinare univocamente dove si trovi l'oggetto, ossia nel punto di coordinate (x_φ', y_φ') , con $x_\varphi' = x_\varphi(t_\varphi')$ e $y_\varphi' = y_\varphi(t_\varphi')$. Questa rappresentazione permette di formalizzare anche un altro assioma implicito nella trattazione cinematica, quello di continuità. Esso richiede che il movimento di un oggetto non proceda “a salti” nello spazio (per intenderci, si esclude una sorta di *teletrasporto*), cioè che presi due istanti temporali *molto vicini*, le rispettive posizioni siano *sufficientemente vicine* anch'esse. In termini grafici significa che la linea che rappresenta il movimento nello spazio-tempo non abbia tagli. In figura 1 è rappresentato il grafico spazio-tempo del movimento di un oggetto (la linea nera). Esso prima è fermo e poi, improvvisamente, si muove.

Figura 1. Rappresentazione spazio-temporale del moto di un oggetto (evento)



Le poche considerazioni precedenti forniscono un esempio del primo tipo di regole: all'interno dell'insieme delle rappresentazioni *formalmente possibili* per un evento – nel nostro esempio tutti i sottoinsiemi di \mathcal{W}_φ – solo alcune lo saranno anche empiricamente, cioè quelle che soddisfano gli assiomi di non-ubiquità e continuità.

Un altro tipo di regole sono quelle che più spesso ricaviamo per induzione dai dati o per deduzione dalle teorie e concernono la relazione tra diversi livelli empirici. Formalmente è come definire una funzione per estensione (elencandone gli elementi) o per intensione (definendone le proprietà). Una funzione tra un insieme A ed un insieme B è, infatti, un sottoinsieme del loro prodotto cartesiano. Un insieme, sottoinsieme di uno più ampio, può essere caratterizzato in due modi. Tramite elencazione degli elementi (estensione) o definendone la funzione caratteristica (proprietà). Se immaginiamo che un insieme rappresenti le varie declinazioni di un fenomeno fisico – per esempio la relazione tra pressione e volume di un gas a temperatura costante – possiamo raccogliere quante più coppie possibili di valori (volume, temperatura) e considerarle come rappresentanti della legge che le lega; oppure possiamo proporre una qualche relazione tra i simboli ed affermare, per qualche motivo più o meno sensato, che le coppie possibili (volume, temperatura) sono tutte e sole quelle per le quali, fissato il particolare gas utilizzabile per le misurazioni, il

prodotto tra temperatura e volume è una costante. È chiaro che i due modi sono strettamente legati: raccogliere esempi di coppie di valori (volume, temperatura) può aiutare a comprendere la regola generale che li lega (induzione); la sensatezza di una regola che lega i valori di volume a quelli di temperatura, quale che sia l'origine della formulazione di questa regola (un messaggio extraterrestre, un'analogia azzardata, un sogno), si basa sulla possibilità di raccogliere coppie concrete di valori (volume, temperatura) e vedere se la regola li prevede oppure no (ragionamento ipotetico-deduttivo).

Un esempio a riguardo è la costruzione di relazioni psicofisiche. Nel caso della percezione del movimento, ci si riferisce a quell'ipotetica funzione, che denotiamo con $f_{\phi,\psi}$ (dove i pedici intendono indicare che agisce su rappresentazioni fisiche, del livello ϕ , per ottenere rappresentazioni fenomeniche, del livello ψ), tale che, nota una rappresentazione S_ϕ di un evento, permettesse di determinare la rappresentazione fenomenica S_ψ dello stesso o una distribuzione di probabilità sulle rappresentazioni fenomeniche possibili. Limitiamoci, per semplicità ⁽⁵⁾, al caso in cui la relazione sia una vera e propria funzione; formalmente avremmo quindi che $S_\psi = f_{\phi,\psi}(S_\phi)$: in parole povere, essa indica che, se presento lo stimolo S_ϕ , l'organismo esperisce [ovvero si comporta come se vedesse] S_ψ . Una funzione siffatta può essere ottenuta per induzione raccogliendo, per esempio, una opportuna collezione di n eventi fisici $\{S_\phi^1, S_\phi^2, \dots, S_\phi^j, \dots, S_\phi^n\}$ ed una collezione di n eventi fenomenici $\{S_\psi^1, S_\psi^2, \dots, S_\psi^j, \dots, S_\psi^n\}$ associati (ossia ottenuti misurando con metodi psicofisici la risposta di un soggetto che veda gli eventi fisici). A questo punto, con metodi statistici ed utilizzando alcuni vincoli di sensatezza fisica e percettiva, è in linea di principio possibile stimare la relazione funzionale psicofisica $f_{\phi,\psi}$. Tale stima, se la caratteristica fisica e quella fenomenica sono entrambe unidimensionali, come nell'esempio della temperatura percepita di un oggetto, prende la forma di una interpolazione grafica. Siano T_ϕ^j e T_ψ^j , rispettivamente, la temperatura fisica e quella percepita nella j -esima prova ⁽⁶⁾. Su un grafico cartesiano si segnano gli n punti (T_ϕ^j, T_ψ^j) ($j=1\dots n$) e si cerca la curva migliore, nel senso geometrico, che li unisca. Tale curva è una *regola* che mette in relazione due livelli empirici distinti – quello fisico e quello fenomenico: una relazione psicofisica che lega la temperatura fisica T_ϕ (variabile indipendente) alla temperatura percepita T_ψ (variabile dipendente).

Vi è pure un altro modo di costruire regole tra diversi domini empirici. Essa utilizza lo strumento formale in un modo diverso da quanto sia, per esempio, l'interpolazione geometrica o l'induzione statistica. Una esemplificazione paradigmatica è il modo in cui si può ricavare la *legge di Fechner* dalla *legge di Weber*. Siano I_ψ e I_ϕ , rispettivamente, l'intensità percepita e quella fisica di uno stimolo (una macchia luminosa o un suono, ad esempio); la legge di Fechner stabilisce che, a meno di costanti che dipendono dalle unità di misura, $I_\psi = \log(I_\phi)$ [24]. Tale relazione, confermata sperimentalmente, può essere ricavata formalmente [*ibidem*] partendo dal risultato sperimentale detto “legge di Weber”, che stabilisce quanto segue: l'incremento minimo ΔI_ϕ di intensità dello stimolo, necessario affinché lo stesso stimolo di intensità $I_\phi + \Delta I_\phi$ sia discriminabile da quello di intensità pari a I_ϕ , è proporzionale al valore I_ϕ medesimo e la costante di proporzionalità è la stessa per tutti i valori I_ϕ . In altri termini si ha che $\Delta I_\phi / I_\phi = \text{costante}$. Sia f la funzione psicofisica tale che, eliminando i pedici, $I_\psi = f(I_\phi)$, ossia la funzione che associa al valore fisico dello stimolo la misura della sua percezione; abbiamo, per definizione di minima differenza discriminabile, che il più piccolo I_ψ' tale che $I_\psi' > I_\psi$ è pari a $I_\psi + \Delta I_\psi = f(I_\phi + \Delta I_\phi)$, dove $I_\psi = \Delta I_\phi / I_\phi$. Se assumiamo che il dominio fenomenico sia un *continuum* come quello fisico, possiamo azzardarci a ritenere

5 Il lettore interessato agli aspetti matematici avrà notato che si può considerare $f_{\phi,\psi}$ come una relazione tra gli spazi vettoriali S_ϕ ed S_ψ . Semplici considerazioni empiriche portano a ritenere che difficilmente essa possa essere una funzione *iniettiva*, visto che sono caratteristiche note del sistema visivo quelle di non avere una sensibilità infinita e di andare oltre l'informazione data [43]: entrambe fanno sì che ad un percolato corrispondano più stimoli differenti possibili in grado di determinarlo. Inoltre, tolto il caso di visione nei sogni, immaginazione, allucinazioni e simili – a rigore esperienze visive senza stimolo attuale – la relazione in questione deve essere considerata suriettiva: cioè ad ogni stimolo fisico corrisponde almeno un percolato. Quanto detto in questa nota vuole essere un'ulteriore dimostrazione di come il processo di formalizzazione sia sufficientemente potente da cogliere al suo interno le diverse sfumature empiriche nei termini di vincoli formali da applicare sulle strutture matematiche in gioco.

6 Chiaramente questa struttura sperimentale è, presa alla lettera, insensata. La misurazione di un valore fenomenico T_ψ richiede diverse prove e, quindi, dovremmo avere m misurazioni per ognuno degli n livelli dello stimolo fisico, così che il numero finale di prove sarebbe mn .

infinitesimi i due incrementi ΔI_ψ e ΔI_ϕ e denotarli con dI_ψ e dI_ϕ , ottenendo così che $dI_\psi = dI_\phi/I_\phi$ da cui, integrando, si ottiene la relazione cercata $I_\psi = \log(I_\phi)$.

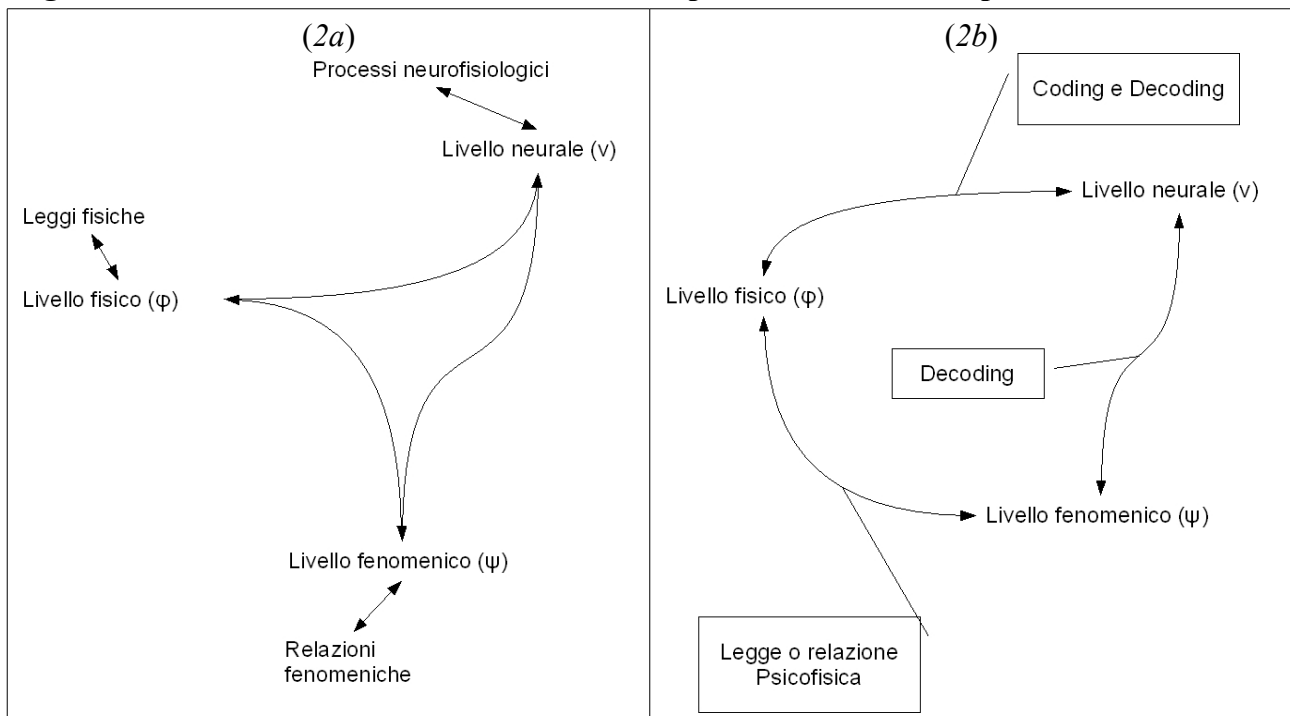
L'esempio riportato nel capoverso precedente è degno di essere qui menzionato per due ragioni. La prima è che mostra ancora una volta che l'approccio da noi adottato è una semplice generalizzazione dell'apparato concettuale classico. La seconda è che abbiamo prova di come la matematica possa essere usata nelle scienze della percezione alla stregua di come è usata in fisica o in biologia, cioè non solo come strumento per analizzare i dati ma, pure, come linguaggio per esprimere, o creare, modelli o teorie. Considerazioni analoghe valgono anche per le relazioni che sussistono tra il livello fisico e quello neurale e tra il livello neurale e quello fenomenico.

Una volta che si siano posti alcuni vincoli sulle *forme matematiche possibili*, il teorico ha a disposizione un apparato concettuale enorme, articolato e potente – la matematica, appunto – con cui costruire teorie o modelli percettologici.

I vincoli sono da considerare come assiomi, cioè come condizioni di senso da cui partire (⁷); quando esplicitati, sono di natura formale e riprendono alcuni assunti di sensatezza empirica relativa a ciascun livello o alle loro interazioni; tuttavia sono spesso impliciti, nel senso che non si dà loro alcuna forma – né nel linguaggio formale né linguaggio naturale –, principalmente perché o troppo complessi o troppo numerosi.

Ricapitolando siamo interessati a due tipi di regole (relazioni, funzioni): un tipo è quello che lega le variabili all'interno di un dominio empirico (Fig. 2a); l'altro è quello che lega le variabili di differenti domini empirici (Fig. 2b); inoltre alcune sono date a priori mentre altre sono l'oggetto della ricerca percettologica.

Figura 2. Schema della relazione entro i livelli empirici e tra i livelli empirici



Finora abbiamo trattato tanto della possibilità quanto della utilità di usare un linguaggio comune per i diversi livelli empirici che le scienze della visione mettono in relazione. Possiamo quindi abbozzare la definizione di modello matematico di un processo visivo. Nel nostro caso vi

⁷ La storia della scienza ci insegna che le rivoluzioni scientifiche riguardano principalmente questi assunti [36]. Uno dei concetti che si vuole qui esprimere è che l'uso più incisivo della matematica, nelle scienze percettive in generale e nelle scienze della visione in particolare, non è legato ad una questione teorica ma, piuttosto, metateorica. Non si dovrebbe parlare di teoria matematica, come se potesse esservi un'altra teoria, quanto piuttosto di teoria espressa con la matematica. Questo modo di fare permetterebbe un dialogo stretto e proficuo tra sperimentalisti e teorici.

sono tre livelli empirici; associamo a ciascuno livello un insieme di variabili, ciascuna delle quali è diretta controparte formale di un'osservabile empirica; organizziamo tali variabili in strutture matematiche, possibilmente in modo che vi sia corrispondenza tra le strutture utilizzate nei diversi livelli (così che, per esempio, se introduciamo lo spazio vettoriale formato dalle terne spazio-temporali fisiche $(x, y, t)_\phi$, vi sia anche lo spazio vettoriale delle terne spazio-temporali fenomeniche $(x, y, t)_\psi$); identifichiamo un insieme di assiomi ed introduciamoli nel modello nei termini di vincoli tra le variabili. In queste condizioni un *modello matematico della percezione visiva* è un insieme di relazioni che legano tra di loro strutture o variabili di livelli empirici differenti o dei medesimi livelli empirici. Un modello assume al ruolo di *teoria* quando il numero e l'importanza delle variabili in gioco è tale da potersi dire che tratta dei fatti e dei problemi fondamentali della percezione visiva.

In senso stretto, la differenza tra modello matematico, o formale, *tout-court* e modello computazionale è legata al linguaggio in cui le parti salienti del modello sono espresse ed al modo in cui si svolgono i calcoli: un modello potrebbe coinvolgere solo matematica simbolica ed i vari suoi passaggi essere ottenuti attraverso passaggi formali. Oppure richiedere di essere in parte o in tutto tradotto in simulazioni numeriche. In senso lato il riferimento stereotipico del sintagma "modello computazionale" applicato alla visione sono i sistemi connessionisti, ovvero le reti neurali. Nella sua connotazione principale, un modello computazionale deve da qualche parte mostrare funzioni tra variabili ⁽⁸⁾ che sembrano flussi di attivazione tra strati di neuroni artificiali.

Qui per "modello computazionale" si intende semplicemente un modello matematico passibile di essere tradotto in una simulazione.

§1.4.

Linee argomentative, euristiche e posizioni epistemologiche

Abbiamo enfatizzato la curiosità per lo strumento formale e la convinzione che si possa farne un uso più *invasivo*, in analogia a quanto avviene da secoli in fisica e da un po' di tempo in biologia.

Assumiamo che il sistema nervoso abbia un numero limitato di meccanismi di calcolo. Ciò ha significato evolutivo ed un'enorme utilità teoretica. Infatti se trovassimo uno strumento formale che ci permettesse di spiegare l'elaborazione di segnali bidimensionali, ipotizzeremmo che il medesimo strumento sia utilizzato dal sistema nervoso anche per elaborazione tridimensionale. Questa euristica di generalizzazione è già stata ampiamente utilizzata in neurofisiologia funzionale: si pensi alla tonotopia o alla codifica spaziale della conoscenza. Il corrispettivo formale di questa euristica è tipico in geometria: quando è possibile si identificano regole indipendenti dalla dimensione n dello spazio di riferimento (almeno fintantoché n è un intero finito). Utilizzeremo principalmente questa euristica per estendere i risultati noti nell'elaborazione neurale di segnali statici bidimensionali al caso di segnali spazio-temporali tridimensionali.

Un breve cenno merita la diatriba tra Dennett e, tra gli altri, Pessoa [38, 39], centrata sull'interpretazione del *filling-in*, ma di portata generale. La questione è la seguente: i processi di completamento – per esempio figurale o di bordi – agiscono a livello neurale in maniera isomorfa a ciò che accade a livello fenomenico? In altri termini è possibile spiegare i fattori di integrazione del segnale visivo globali solo con la attivazione simultanea di aree cerebrali corrispondenti alle zone integrate? Il segnale visivo si diffonde nella corteccia visiva come alcuni modelli farebbero pensare? Noi adottiamo una prospettiva operazionista che ci permette, in un certo qual modo, di glissare: non ci chiediamo, in altri termini, quale sia la *vera rappresentazione fenomenica* dei processi di completamento (di bordi o figure, poco importa qui). Piuttosto vincoliamo tra di loro le rappresentazioni scelte nei diversi livelli empirici e ci riteniamo tanto più soddisfatti quanto più l'articolazione di questa struttura di relazioni è stabile, coerente e, magari, elegante.

8 Omomorfismi tra spazi vettoriali

§2. La prospettiva psicofisica

§2.0. Introduzione

La psicofisica classica si occupa di mettere in relazione misure fisiche e fenomeniche. In questo capitolo sono brevemente trattati alcuni costrutti psicofisici fondamentali per la descrizione del movimento percepito. Si introduce il concetto di "segnale di movimento" come elaborazione di quello di "segnale spazio-temporale", a sua volta generalizzazione del ben noto segnale visivo bidimensionale.

Il *flusso ottico* è la forma più nota di segnale di movimento ed utilizzato spesso anche in discipline non neuroscientifiche. La sua determinazione è spesso di carattere locale. Si introducono, quindi, le ragioni fenomenologiche che portano alla ricerca di fattori di elaborazione *globale* del segnale visivo. In tale prospettiva, punto di partenza fondamentale è il costrutto di *campo associativo*, elaborato nell'ambito dello studio dei processi di organizzazione figurale bidimensionale: se ne propone una estensione allo spazio-tempo che permetta di tenere conto dell'integrazione anche di segnali di movimento.

§2.1. I segnali statici bidimensionali

Un segnale visivo è, dal punto di vista fisico, la *variazione* della quantità (luminanza) o della qualità (cromaticità) della luce che incide su almeno un fotorecettore retinico.

Supponiamo di aver fissato un sistema di riferimento spaziale ed una opportuna scala di misura per la luminanza. Se ci si occupa di soli stimoli acromatici bidimensionali statici, in visione monoculare [27], la descrizione *distale* di uno stimolo può essere denotata tramite una funzione I_ϕ che riassume il fatto che nel punto di coordinate spaziali (x, y) , vi è una quantità di luce (radiante o riflessa che sia) pari a $I_\phi(x, y)$. Determinata la posizione dell'occhio e la direzione della retina, attraverso le leggi della geometria proiettiva [*ibidem*] è possibile determinare la descrizione dello *stimolo prossimale*, la quale consta di una funzione $I_p(x, y)$ che rappresenta la quantità di luce incidente sul punto retinico di coordinate (x, y) .

Per quanto concerne il versante fenomenico, si conviene di rappresentare il percepito relativo allo stimolo in uno spazio *molto simile* a quello usato per rappresentare lo stimolo fisico (sia distale sia prossimale), tramite una funzione $I_v(x, y)$ che indica il valore (in un'opportuna scala psicofisica) di chiarezza (brightness) presente nello spazio fenomenico di coordinate (x, y) .

Fatta questa precisazione, risulta agevole definire il concetto di "segnale di movimento".

§2.2. Il concetto di segnale di movimento

Se all'esempio del paragrafo precedente aggiungiamo la dimensione temporale – denotata dalla variabile t –, per estensione diciamo che un *segnale spazio-temporale* è rappresentato da una funzione $I_\phi(x, y, t)$, se se ne enfatizza il versante distale, o dalla funzione $I_p(x, y, t)$ nella sua controparte *prossimale*; dal punto di vista fenomenico, $I_v(x, y, t)$ indica il valore di brightness percepito nel punto (x, y, t) .

La locuzione "segnale di movimento" sottende un concetto più complesso di quello di *presenza di una quantità di luce (o contrasto)*, precisamente quello di *velocità*, il cui miglior

analogo bidimensionale statico è quello di *orientamento di un bordo* o *dell'asse principale di una figura*. Quindi dobbiamo introdurre una nuova funzione che rappresenta la misura distale di un corpo in movimento, la quale mette in relazione un punto (x, y, t) non più ad una sola misura (la quantità di luce) ma a due misure: la direzione del moto e la sua velocità. Denotiamo tale funzione con $v_{\phi}(x, y, t) = (v_{\phi, x}(x, y, t), v_{\phi, y}(x, y, t))$, essendo la direzione del movimento pari a $\arctan\left(\frac{v_{\phi, y}(x, y, t)}{v_{\phi, x}(x, y, t)}\right)$ e la velocità (velocità scalare o *speed*) di valore $\sqrt{v_{\phi, x}(x, y, t)^2 + v_{\phi, y}(x, y, t)^2}$. Con una notazione simile a quella precedente, $v_{\rho}(x, y, t)$ e $v_{\psi}(x, y, t)$ denotano, rispettivamente, la velocità retinica (livello prossimale) e quella percepita (livello fenomenico).

Un *segnale di movimento* è proprio la funzione $v(x, y, t)$, talvolta intesa nel suo referente fisico (v_{ϕ} o v_{ρ}), talaltra, seppur impropriamente, in quello fenomenico (v_{ψ}).

§2.3.

Il flusso ottico

Il concetto di flusso ottico si rivela centrale non solo nei lavori di Gibson [10, 11] ma, pure, nelle trattazioni ingegneristiche ed informatiche ricollegabili all'ambito della *computer vision* [40-42]. Con la notazione sopra introdotta il flusso ottico è una caratterizzazione $v_{\rho}(x, y, t)$ dello stimolo prossimale, di cui il pattern di velocità $v_{\psi}(x, y, t)$ è la controparte fenomenica.

In altri termini il flusso ottico è un campo vettoriale che ad ogni punto dello spazio (retinico e bidimensionale) e ad ogni istante associa un vettore velocità.

Il calcolo standard del flusso ottico si basa sul gradiente spazio-temporale del pattern di luminanza che stimola l'occhio (o, nel caso ingegneristico, il sistema ottico artificiale, come una telecamera). Siano $I_{\phi}(x, y, t)$ la funzione che rappresenta un filmato in bianco e nero e $I_{\rho}(x, y, t)$ la sua proiezione retinica. Si assume che gli oggetti – ossia ciò che si muove – abbiano luminanza costante e che siano continui; ciò significa che, nello spazio tempo, le componenti dell'immagine in movimento sono quelle per cui $I_{\rho}(x, y, z, t) = I_{\rho}(x+dx, y+dy, t+dt)$, ossia che si *mantengono* in un piccolo intorno spazio-temporale (dx, dy, dt) di (x, y, t) . La formula di derivazione per campi scalari ci dice che $I_{\rho}(x+dx, y+dy, t+dt) = I_{\rho}(x, y, t) + (\nabla I_{\rho}) \cdot (dx, dy, dt)$ e, quindi, che $(\nabla I_{\rho}) \cdot (dx, dy, dt) = 0$. Se dividiamo per dt ed esplicitiamo la forma del gradiente otteniamo l'equazione

$\frac{\partial I_{\rho}}{\partial x} \frac{dx}{dt} + \frac{\partial I_{\rho}}{\partial y} \frac{dy}{dt} + \frac{\partial I_{\rho}}{\partial t} \frac{dt}{dt} = 0$ dalla quale, ponendo $v_{\rho, x} := dx/dt$ e $v_{\rho, y} := dy/dt$, si ottiene l'equazione che vincola il flusso ottico ($v_{\rho, x}, v_{\rho, y}$)

$$v_{\rho, x} \frac{\partial I_{\rho}}{\partial x} + v_{\rho, y} \frac{\partial I_{\rho}}{\partial y} = - \frac{\partial I_{\rho}}{\partial t} \quad (1)$$

Ad ogni punto (x, y, t) possiamo quindi associare l'equazione (1) che presenta due incognite (le componenti del flusso ottico, appunto). Ciò significa che non è possibile determinare con il solo calcolo del gradiente il flusso ottico; il fatto che vi sia questo *grado di libertà* residuo nella determinazione analitica del flusso ottico è l'equivalente formale del *problema dell'apertura* noto in psicofisica. In *computer vision* la sua soluzione comporta l'introduzione di ulteriori vincoli tra le variabili (così da ottenere un numero di equazioni pari al numero delle incognite). Al percettologo interessa, invece, sapere quali siano i vincoli che usa il sistema visivo per risolvere questo problema; nei termini del linguaggio del percettologo, egli è interessato alla trasformazione che lega il flusso ottico (variabile distale o prossimale) alla velocità percepita (variabile fenomenica).

§2.4.

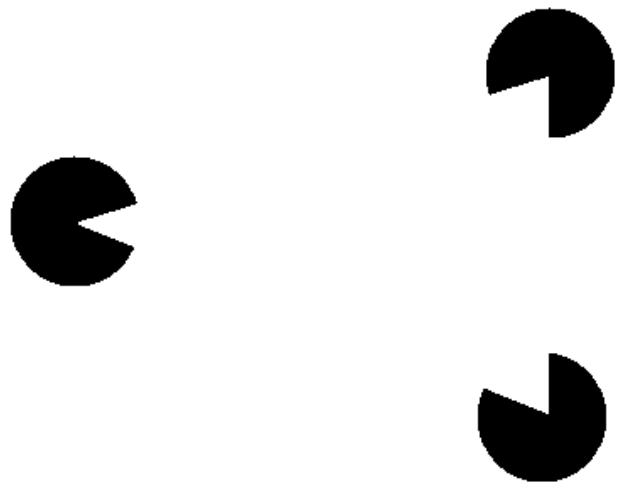
L'asse concettuale Globale-Locale

Abbiamo introdotto le osservabili salienti della psicologia della percezione del movimento: il segnale spazio-temporale, il segnale di movimento, il flusso ottico e il movimento percepito. Prima di trattare dei modi con cui si ipotizza i tre suddetti eventi siano collegati, è opportuno far menzione di un tema implicitamente introdotto dal metodo con cui si vincola la forma del flusso ottico. La determinazione del campo vettoriale (v_p, x, v_p, y) parte dai dati *locali* del pattern di luminanza incidente sulla superficie fotosensibile. In termini molto semplici ciò è reso possibile dall'assunzione delle continuità degli oggetti e dalla sufficienza delle informazioni presenti nell'*immediato intorno* di un punto a determinarne le qualità fisiche e percepibili.

Da un punto di vista percettologico, le condizioni locali di stimolazione possono non essere – ed in genere non sono – sufficienti a determinare il percepito. Kanizsa [43] ne dà svariati esempi bidimensionali, uno dei quali è il noto *triangolo di Kanizsa* (Figura 3): dal punto di vista fisico sono presenti solo tre figure nere su sfondo bianco: si tratta di tre cerchi a ciascuno dei quali "è stata tolta una fetta". Dal punto di vista fenomenico si vede un triangolo bianco i cui vertici sovrastano le tre figure nere. Non solo: ciascun "pacman" si completa *amodalmente* in un cerchio e la chiarezza del triangolo è maggiore di quella dello sfondo bianco, benché la sua luminanza, invece, non ne differisca.

La lezione è che i processi percettivi richiedono si prendano in considerazione sia le condizioni *locali* sia quelle *globali* dello stimolo come variabili *indipendenti* nella determinazione del percepito: le famose *leggi della Gestalt*, o *fattori di unificazione figurale* [9-11, 43], concretano questa istanza e propongono regole che hanno come oggetto lo stimolo distale, o quello prossimale, con il fine di spiegare quali parti dello stimolo siano associate a formare qualità percettive salienti (bordi, figure, sfondi, qualità cromatiche e acromatiche). Le principali sono la vicinanza, la somiglianza, il destino comune, la buona continuazione, la chiusura, l'esperienza passata [*ibidem*]. Per i nostri scopi assume notevole rilievo la buona continuazione, perché Field [44] né fornisce una fondamentale caratterizzazione.

Figura 3. Triangolo di Kanizsa



§2.5.

Il campo associativo spaziale

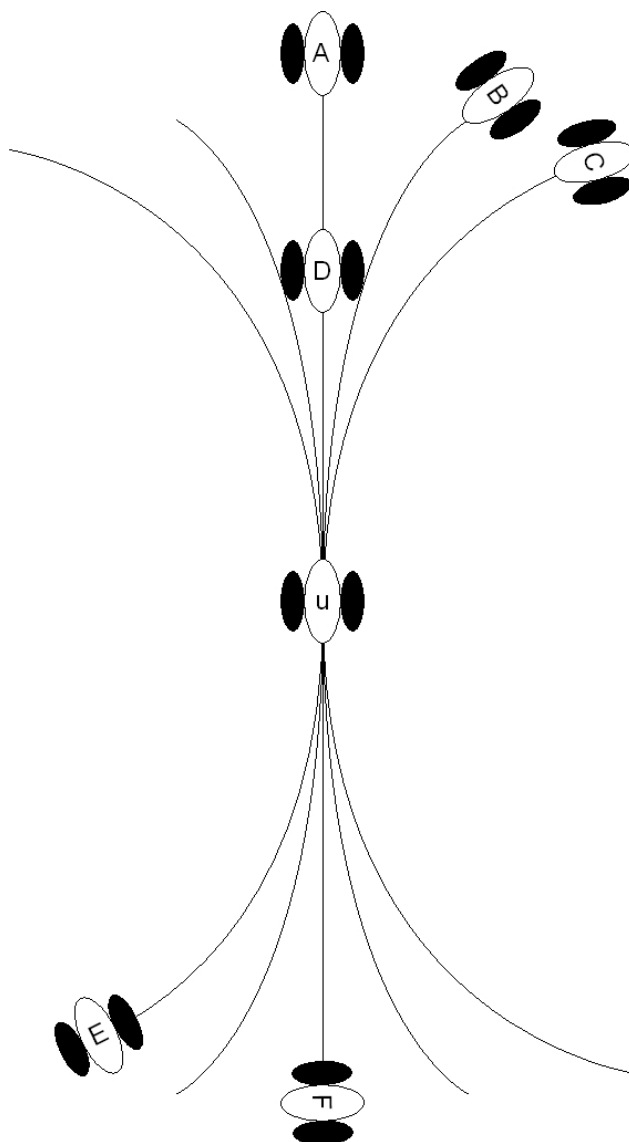
Il classico lavoro di Field *et al.* del 1993 [44] permette di gettare un ponte tra le leggi di carattere prettamente psicofisico e quelle neurofisiologiche. Con cinque esperimenti mostra che segnali locali bidimensionali stazionari sono integrati dal sistema percettivo lungo linee preferenziali compatibili con la succitata regola gestaltista di *buona continuazione*. Tali esperimenti sono tutti riconducibili alla medesima forma prototipica di seguito riassunta. Per una durata dell'ordine delle centinaia di millisecondi, è presentata ai soggetti una immagine costituita da un numero N di Gabor bidimensionali variamente disposte (Gabor patch) ed orientate; di queste N_S sono disposte in modo che i rispettivi assi principali siano collineari ad una linea curva continua *immaginaria* – nel senso che non è fisicamente presente ma può essere inferita geometricamente o percepita per completamento –, mentre i rimanenti $N_R = N - N_S$ presentano una orientazione casuale. Gli stimoli sono presentati a coppie, con opportuno schema di casualizzazione, di modo che un elemento della coppia abbia solo distrattori ($N_S = 0$) mentre l'altro abbia anche il target ($N_S \neq 0$). Compito del soggetto è di identificare lo stimolo in cui è presente la serie di Gabor patch disposte in modo da formare una curva 'campionata'. Diverse variabili sono manipolate, tra le quali: la distanza tra i patch, la loro dimensione, la relazione angolare tra gli assi di simmetria dei patch e la curva immaginaria, la fase dei patch. I risultati principali sono i seguenti: (i) il completamento della linea target avviene per elementi che hanno orientazioni che differiscono fino a 60° ; (ii) la distanza tra gli elementi che sono uniti nel processo di segregazione può essere anche fino a 7 volte la loro dimensione; (iii) la fase dei gabor-patch è irrilevante rispetto alla loro distanza ed alla loro orientazione.

Complessivamente tali risultati portano a postulare l'esistenza di un *campo associativo* (figure 4, 5), presente in ogni punto dello spazio visivo bidimensionale, che permetta di legare elementi bidimensionali tra loro distanti secondo tre variabili principali: la curvatura della linea di associazione, la sua lunghezza e la relazione tra l'asse principale degli elementi figurati da associare e la direzione della curva nel punto in cui incontra gli elementi figurati.

Discutiamo nel dettaglio la figura 4. Parafrasando quanto appena detto, sperimentalmente Field dimostra che segnali direzionali locali statici, in punti distinti dello spazio, interagiscono positivamente (correlano) quando sono tangenti (V1) ad una *buona curva* che passa per i punti in questione; tale associazione è tanto meno forte (V2) quanto maggiore è la curvatura della curva e (V3) quanto maggiore è la distanza tra i due punti. L'elemento figurale indicato con la lettera "u" si associa, quindi, positivamente con gli elementi figurati "A", "B", "C" e "D", mentre non è associabile con gli elementi "E" e "F":

- L'elemento "F" dista da "u" quanto l'elemento "A" (stesso peso del fattore V3), tuttavia l'asse principale di "F" è ortogonale alla curva che lo unisce ad "u", mentre l'asse principale di "A" è parallelo alla curva che lo unisce a "u", quindi, per il fattore V1, "u" si associa ad "A" ma non a "F". Ragionamento simile vale per il confronto tra "C" ed "E".
- Il fattore V2, la curvatura, implica che l'associazione tra "u" ed "A" sia più forte di quella tra "u" e "B" e quest'ultima sia a sua volta più forte di quella tra "u" e "C": ciascuno dei tre elementi "A", "B" e "C" ha (circa) la stessa distanza da "u", ma si trovano su linee di congiunzione a curvatura crescente.
- Per il fattore V3 – la distanza – l'elemento figurale "D" si associa maggiormente ad "u" rispetto ad "A", perché, benché siano sulla stessa linea e non differiscano, quindi, per il fattore di curvatura, "D" è più vicino ad "u" di "A".

Figura 4. Rappresentazione del campo associativo postulato da Field *et al.* (1993).



Il campo associativo ipotizzato da Field si presta ad essere formalizzato, come vedremo e come è in effetti avvenuto, in modelli di *computer vision*, in modelli psicofisici ed in modelli neurofisiologici di organizzazione figurale ed integrazione di immagini statiche, ossia in modelli sullo spazio bidimensionale. Nel paragrafo seguente ne ipotizziamo una estensione allo spazio-tempo tridimensionale.

§2.6.

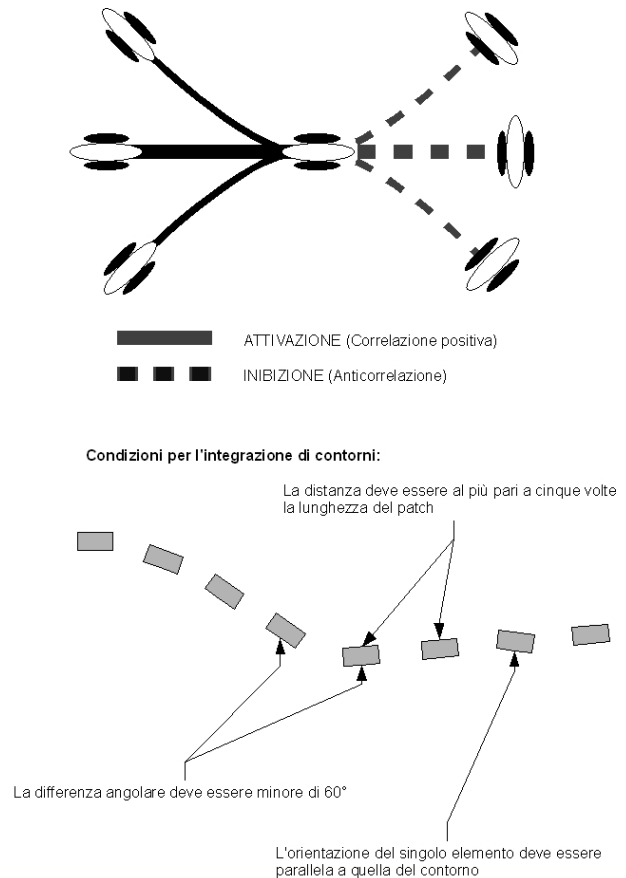
Il campo associativo spazio-temporale

Ci occupiamo del movimento bidimensionale. Da un punto di vista formale – per quanto concerne le rappresentazioni, quindi – ciò equivale ad aggiungere la dimensione temporale alle due dimensioni spaziali. Pare naturale, perciò, chiedersi se si possa estendere allo spazio-tempo il costrutto di campo associativo, in modo che quello postulato da Field ne sia la sezione spaziale.

L'euristica di generalizzazione si dimostra sensata in questa sede, anche alla luce della teoria gestaltista. Infatti la legge del destino comune, come già aveva osservato Wertheimer e sottolinea Vicario [10], può essere interpretata come l'estensione allo spazio-tempo della legge di buona

continuazione.

Figura 5. Campo associativo e suoi vincoli ricavati sperimentalmente (con modifiche da [69])



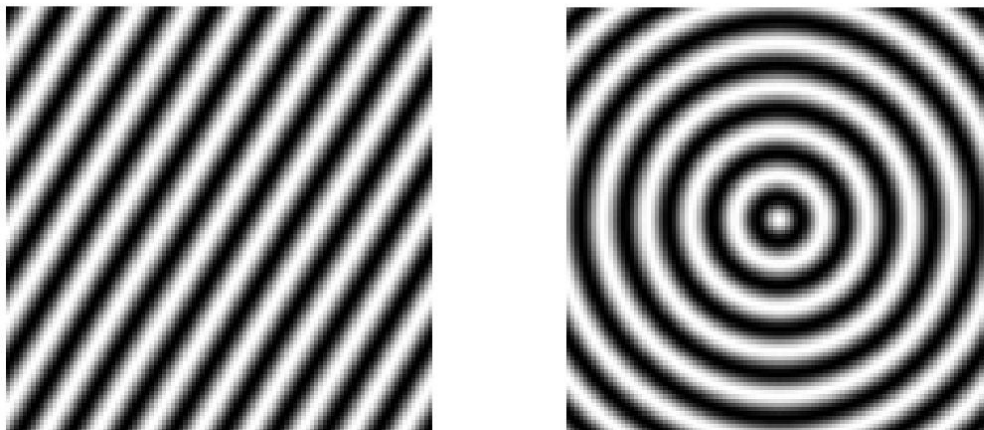
Quali sono le configurazioni stimolo, per così dire, *preferite* dal campo associativo spaziale? Se noi postuliamo che le linee di associazione uscenti da un punto siano a curvatura costante – per intenderci linee o archi di cerchio –, le configurazioni "preferite" – le buone forme – sono rette e circonferenze. Non solo; anche la loro giustapposizione a diversi valori di luminanza sono buone forme: griglie sinusoidali e onde circolari [Fig. 6]. Da un punto di vista formale, le leggi della Gestalt sarebbero implementate in un una sorta di scatola nera, un operatore, che agisce come un *campo di forze (percettive)* sulla rappresentazione dello stimolo distale per fornire la rappresentazione fenomenica. Quando già lo stimolo è una buona forma, è lecito supporre che l'azione del campo sia ininfluenza, dato che trova già le condizioni che vorrebbe creare. Questa informazione è molto utile, come vedremo, per formalizzare l'azione del campo, sia dal punto di vista psicofisico, sia da quello psicobiologico. Se il campo associativo spaziale porta a ritenere che le buone forme sia come quelle in figura 6, pare naturale che le *buone forme* spazio-temporali siano le medesime quando rappresentate nella sezione bidimensionale spazio-temporale. L'equivalente percettivo sarebbero griglie sinusoidali in movimento ed onde sferiche spazio-temporali, cioè di onda circolari piane i cui fronti d'onda prima si allargano e poi si restringono.

§2.7. Il problema dell'integrazione dei segnali di movimento

Nel paragrafo precedente abbiamo postulato l'esistenza di un campo associativo spazio-temporale per ragioni teoriche: come una naturale estensione del caso bidimensionale spaziale al caso tridimensionale spazio-temporale. Importanti questioni empiriche ci portano nella stessa direzione.

Un esempio è fornito da alcuni risultati dell'applicazione del paradigma dei *Random Dot Kinematograms* (RDK) [9, 45-50]. Negli esperimenti con RDK si fa uso di filmati-stimolo i cui singoli frame sono composti da configurazioni casuali di punti che, tuttavia, correlano tra i diversi frame, di modo che i singoli punti sono percepiti muoversi. Il fatto interessante è che in talune condizioni il sistema visivo vede non tanto il movimento dei singoli punti, quanto un movimento *globale*; ad una velocità media. Con uno sforzo di immaginazione pensiamo alla rappresentazione spazio-temporale dello stimolo: adottiamo la prospettiva dell'*homunculus*. Immaginiamo la condizione in cui i frame sono tutti uguali al primo. In questo caso non si percepisce movimento e la rappresentazione spazio-temporale – in cui gli assi x e y sono spaziali e l'asse verticale z è il tempo – assomiglia ad un canneto. Se nei diversi frame i punti si spostano, la rappresentazione assomiglia ad un canneto spazzato dal vento: alcune canne saranno più piegate di altre – ossia alcuni punti andranno più veloci – ma, se la densità delle canne è sufficiente, si percepisce un effetto *globale* del vento e le canne sembrano piegate tutte nello stesso modo: insomma si vede la foresta e non gli alberi.

Figura 6. Possibili forme pregnanti



In questa sede non ci interessano tanto i fattori che modulano l'integrazione spazio-temporale, quanto il fatto che essa esista e si possa associarla a meccanismi simili a quelli che agiscono nelle sole due dimensioni spaziali.

§3. La prospettiva neurofisiologica

§3.0. Introduzione

Si dà una rapida descrizione degli *atomi* del sistema nervoso: i neuroni. La loro attività è determinata dalle caratteristiche del mondo esterno all'organismo, per le quali esso possiede meccanismi di trasduzione: funzionalmente, quindi, i neuroni forniscono una *codifica* dello stato fisico del mondo. Quando un neurofisiologo si trovi di fronte ad una misurazione dell'attività cerebrale di un organismo, il suo interesse è comprendere come tradurla per inferire quali siano le condizioni del mondo esterno all'organismo: cerca un sistema di *decodifica*. *Codifica* e *decodifica* neurali sono processi di grande interesse anche per il percettologo.

Vi sono dati robusti sulla capacità di codifica delle unità visive. I principali sono che l'unità visiva prototipica risponde in maniera specifica ad alcune variabili fisiche, nel senso che, data una variabile fisica V , l'attività di una cellula visiva ha un picco quando lo stimolo possiede un fissato valore V_{pref} della variabile fisica; la risposta cellulare è, invece, tanto meno intensa quanto maggiore è la differenza tra il valore attuale della variabile fisica e quello preferito V_{pref} . In questa sede ci importa che le unità visive danno risposte specifiche, tra l'altro, per la frequenza spaziale, l'orientazione, la grandezza, la direzione di movimento e la velocità. Tali differenti specificità possono essere spiegate associando ad una unità un campo recettivo che possiede determinate caratteristiche.

I neuroni visivi sono organizzati in popolazioni secondo alcune regole. La principale è che unità che hanno preferenze simili per una variabile fisica sono più vicine tra loro nella corteccia di unità che hanno, invece, diversi valori preferiti della variabile fisica in oggetto.

Prendendo spunto dalle caratteristiche di base del sistema nervoso, si ricercano i corrispettivi neurali dei concetti trattati a livello psicofisico, principalmente dei noti meccanismi di integrazione dei segnali statici e di quelli supposti agire per integrare i segnali di movimento. Si pone particolare enfasi sulla naturalezza della estensione formale dalle due dimensioni spaziali alle tre dimensioni spazio-temporali, ogni qual volta i dati neurofisiologici si rivelino ad essa compatibili.

Nell'intero capitolo si fa costante ed implicito riferimento a [5-11, 29, 35, 51].

§3.1. Unità di analisi strutturale e procedurale: neurone e potenziale d'azione

Il neurone è l'unità funzionale del sistema nervoso. Esso, nella sua versione prototipica (⁹), è costituito da tre componenti: il soma, l'albero dendritico e l'assone. Il soma, o corpo cellulare, contiene il nucleo e la maggior parte della massa citoplasmatica; i dendriti e l'assone sono estensioni filiformi del soma. I dendriti sono i canali di input del neurone, mentre l'assone il canale di output. L'osservabile biofisica attraverso la quale si misura l'attività del neurone è il *potenziale di membrana*, definito come la differenza di potenziale elettrico tra l'esterno (liquido extracellulare) e l'interno del neurone. Quando il neurone è a riposo, cioè non è sottoposto a sollecitazione da altri neuroni che proiettano le loro terminazioni assoniche sul suo albero dendritico, il potenziale di membrana è negativo (circa -70 mV). Ciò significa che, in tale condizione, all'interno del neurone vi sono più cariche negative, ovvero meno cariche positive, che all'esterno. Tale differenza di potenziale è mantenuta dalla cellula tramite meccanismi attivi – che usano energia (ATP) – di

⁹ Vi sono vari tipi di neuroni e vari modi di classificarli. Qui si intende fornire una rappresentazione sufficientemente generale che enfatizzi la natura computazionale del parenchima cerebrale.

scambio ionico (la pompa sodio-potassio); ed è modulata da meccanismi elettrochimici, che coinvolgono canali proteici che attraversano la membrana cellulare e regolano lo scambio ionico tra il liquido cellulare ed extracellulare. A seguito della stimolazione chimica che agisce sull'albero dendritico ed è determinata dall'attività di altri neuroni, i canali proteici cambiano la loro conformazione; ciò determina uno scambio ionico – una corrente elettrica – tra l'interno e l'esterno della cellula, ed una variazione repentina (della durata di circa 1-2 ms) del potenziale di membrana, che raggiunge il valore massimo di circa 30 mV, per poi riposizionarsi al valore iniziale. Tale evento è detto *spike* o *potenziale d'azione*. Esso è preceduto dalla stimolazione chimica delle appendici dendritiche ed è seguito, a livello della terminazione assonica, dal rilascio di composti chimici (neurotrasmettitori) che a loro volta influenzano l'attività di altri neuroni.

I neuroni, quindi, comunicano tra di loro attraverso messaggeri chimici, mentre la loro attività di base consta di potenziali d'azione, ossia di repentine variazioni del potenziale di membrana rispetto ad un valore di base. Secondo il livello di analisi prescelto, la neurofisiologia può occuparsi del potenziale di membrana come variabile dinamica, ossia della sua variazione nel tempo [51]; oppure spostare il proprio interesse sulla presenza di potenziali d'azione in funzione del tempo o, ancora, sulla frequenza temporale degli stessi. Quest'ultima osservabile, detta *frequenza di scarica*, è centrale negli studi in cui si misura *in vivo* l'attività cerebrale di animali non umani in risposta a stimoli visivi (o di altre modalità sensoriali). La frequenza di scarica è importante anche perché è il modo in cui il neurone modula la sua risposta agli stimoli. Il potenziale d'azione, infatti, è un fenomeno stereotipato del tipo *tutto-o-nulla*: consiste sempre nel passaggio ad un picco di differenza di potenziale di 30 ms e dura sempre circa 2 ms. La maggiore o minore stimolazione a cui è sottoposto un neurone è codificata con una variazione del numero di spike per unità di tempo, cioè una variazione di frequenza di scarica.

Da un punto di vista computazionale, un neurone è una funzione che ha tanti canali di input quanti sono i suoi dendriti ed un singolo canale di output [19, 20, 35]. Attraverso i canali di input (dendriti) riceve informazioni dall'attività di altri neuroni (nei termini della loro frequenza di scarica) e, in funzione di esse, modula la propria frequenza di scarica che diviene input per altri neuroni a cui proietta tramite l'appendice assonica.

§3.2.

Le rappresentazioni neurali, i processi di coding e decoding, i campi recettivi

Come accennato in [§1.3] bisogna far riferimento a due tipi di rappresentazione dell'attività neurale. Uno è strettamente legato alle osservabili neurofisiologiche mentre l'altro ne costituisce una successiva elaborazione.

Il cervello umano contiene circa $10^{10} \div 10^{11}$ neuroni. Siccome l'osservabile biofisica che definisce l'attività neuronale è il potenziale di membrana, possiamo dire che lo stato istantaneo di attività neurale è rappresentabile da una sequenza (vettore) di N numeri reali $a_1, \dots, a_j, \dots, a_N$, dove N è il numero di neuroni del cervello in questione e il numero reale a_j rappresenta l'*attività*, cioè il potenziale di membrana, del j -esimo neurone. Il fatto che sia impossibile praticamente misurare istantaneamente e contemporaneamente l'attività di tutti gli N neuroni di un cervello non preclude la possibilità di ritenere questa rappresentazione come una sequenza di osservabili e, quindi, essa stessa una osservabile complessa ⁽¹⁰⁾. L'attività del cervello, cioè la sua variazione di stato nel tempo, è rappresentabile come un insieme di N funzioni del tempo $a_1(t), \dots, a_j(t), \dots, a_N(t)$, con $a_j(t)$ che denota il potenziale di membrana del j -esimo neurone nel tempo t ; oppure, in maniera analoga, come un insieme di sequenze composte da $N+1$ valori $a_1, \dots, a_j, \dots, a_N, t$, ciascuna delle quali denota lo stato neurale $(a_1, \dots, a_j, \dots, a_N)$, nell'istante t . Possiamo ritenere questa la rappresentazione di base dell'attività neurale. Le due trasformazioni iniziali a cui è possibile sottoporre tale rappresentazione

¹⁰ La meccanica statistica (sia quella classica sia quella quantistica) è un esempio dell'utilizzo di vettori di stato siffatti.

sono, come già accennato, quella con cui sostituiamo al potenziale di membrana a_j , come osservabile di interesse, la presenza o assenza del potenziale d'azione (una variabile binaria); e quella con cui sostituiamo la frequenza di scarica. A quest'ultima possiamo sostituirla con un'altra, con cui l'attività del j -esimo neurone è rappresentata da un numero $a_j(t)$ che indica la *probabilità* ⁽¹¹⁾ che il neurone sia attivo nell'istante t ⁽¹²⁾. Ciò che conta è che l'attività neuronale sia rappresentata da numeri in relazione monotonicamente crescente con la frequenza di spike e il potenziale di membrana ⁽¹³⁾. Ciò che si evince, e ci interessa, è che godiamo di una certa libertà nello scegliere i numeri con cui rappresentare l'attività di un neurone: ciò che conta è che ad un neurone più attivo di un altro si associ un numero più grande.

In termini geometrici, la dinamica dell'attività neuronale è rappresentata da una *linea* in uno spazio euclideo $N+1$ dimensionale. Chiameremo tale linea *evento neurale*.

Il *processo di coding* [35] è quello che permette di associare ad una rappresentazione fisica dello stimolo il relativo insieme di stati neurali [assunti dal cervello di un organismo che interagisca con lo stimolo tramite gli organi di senso]. Quindi è una funzione che, ad una rappresentazione dello stimolo nei termini di $M+1$ variabili fisiche $x_{\phi,1}, x_{\phi,2}, \dots, x_{\phi,j}, \dots, x_{\phi,M}, t_{\phi}$, associa una linea nello spazio $N+1$ dimensionale degli stati neurali o una distribuzione di probabilità sulle linee possibili. Idealmente la funzione di coding dovrebbe permettere di conoscere l'evento neurale senza misurazioni, semplicemente conoscendo le caratteristiche dello stimolo. Simmetricamente, possiamo chiederci come siano rappresentate le caratteristiche fisiche di uno stimolo, data la conoscenza di un evento neurale: in tal caso ci occupiamo del cosiddetto *processo di decoding*. Esso è una funzione che ad un evento neurale di base associa un evento nei termini di $M+1$ variabili $x_{v,1}, x_{v,2}, \dots, x_{v,j}, \dots, x_{v,M}, t_v$ che rappresentano la versione neurale dello stato od evento fisico.

Un classico esempio di funzione di decoding è il metodo utilizzato da Georgopoulos [5-8, 52] per interpretare il pattern di attivazione di neuroni motori come predittore del movimento eseguito dall'animale sperimentale.

La determinazione del *campo recettivo* di un neurone visivo è, invece, un esempio di studio del *processo di coding*. Come detto, ad un neurone è associabile un valore di potenziale di riposo, che ne rappresenta lo stato di attività in assenza di stimolazione. Anche l'osservabile *frequenza di scarica* possiede un valore di base: esso costituisce una sorta di *rumore di fondo* dell'attività neurale ed è utilizzato nel classico paradigma di misurazione *in vivo* come *baseline*, ovvero come valore di riferimento rispetto al quale si determina se un neurone aumenta o diminuisce la propria attività.

Si definisce *campo recettivo* di un neurone visivo la porzione di campo visivo, ovvero di superficie retinica, la stimolazione della quale è associata ad una variazione nella frequenza di scarica del neurone rispetto alla *baseline*. Hubel e Wiesel (1959, 1962, 1963, 1965, 1968, 1972, 1974, 1977, cit. in [5-8]) sono stati i primi a determinare i campi recettivi di neuroni a diversi livelli del sistema nervoso centrale e a fonderne la nomenclatura. La rappresentazione tipica del campo recettivo di un neurone è una funzione, che qui denotiamo con f , che, ad ogni punto (x, y) del piano, associa un numero $f(x, y)$ secondo il seguente schema ⁽¹⁴⁾:

- si presenta uno stimolo luminoso puntiforme nel punto (x, y) di intensità L ;
- se si ha una *variazione* nella frequenza di scarica del neurone, allora
 - se si ha un *aumento* nella frequenza di scarica del neurone, si pone $f(x, y)$ uguale ad un

11 Si tratta di una *densità di probabilità* se il tempo t non è discretizzato ma considerato variabile continua.

12 La frequenza di scarica non può essere ritenuta, almeno in senso stretto, una variabile temporale, visto che in un determinato istante un potenziale d'azione è presente oppure assente e la determinazione della frequenza richiede di calcolare la presenza di potenziali d'azione in un certo *intervallo di tempo*. In altri termini le due osservabili, dal punto di vista dinamico, sono incompatibili: la massima precisione nel determinare se vi sia o non vi sia un potenziale d'azione è associata alla minima informazione sulla frequenza e viceversa. In termini pratici, tuttavia, il tempo è discretizzato, così che ha senso una analisi tempo-frequenza del segnale neurale.

13 Più alto è il potenziale di membrana più è probabile raggiunga la soglia affinché si generi uno spike; quindi più probabile è che vi sia uno spike; fissato un determinato intervallo di tempo, maggiore è la frequenza di generazione di spike in esso, maggiore è la probabilità che in un istante preso a caso nell'intervallo in questione vi sia uno spike. Se l'intervallo viene preso molto piccolo, abbiamo una densità di probabilità.

14 Lo schema sperimentale classico è assai più complesso e, comunque, sostituito principalmente dall'uso di stimoli non più puntiformi ma casuali e dalla successiva analisi attraverso il metodo della *reverse correlation* [53, 54]. Tuttavia lo schema è comunque istruttivo e mantiene la logica generale, secondo la quale l'unità neuronale è un filtro lineare dell'immagine [35].

- valore positivo proporzionale all'aumento di attività misurato;
- se si ha una *diminuzione* nella frequenza di scarica, si assegna a $f(x, y)$ un valore *negativo* proporzionale all'aumento di attività misurato;
- altrimenti si assegna ad $f(x, y)$ il valore 0.

Nelle figure 7, 8 e 9 sono riportate le rappresentazioni di alcuni campi recettivi di cellule della corteccia visiva primaria.

I campi recettivi sono un esempio importante di funzione di coding anche perché mostrano come, in alcuni casi, la rappresentazione dell'attività neurale di base mantenga una *stretta analogia* con la rappresentazione dello stimolo fisico: è ciò a cui ci si riferisce quando, in termini neurofisiologici, si parla di *organizzazione retinotopica* di alcune aree cerebrali. Riprendiamo le notazioni utilizzate in precedenza. Immaginiamo di stendere la corteccia cerebrale dell'area visiva come fosse un foglio di gomma e che essa abbia sostanzialmente solo due dimensioni; associamole un sistema di riferimento in modo che ad ogni punto (x_v, y_v) corrisponda una e una sola unità cellulare. L'organizzazione dell'area corticale è detta *retinotopica* se, prese due unità distinte che si trovano in due punti *molto vicini* di coordinate corticali $(x_v, y_v)_1$ e $(x_v, y_v)_2$, anche i centri dei rispettivi campi recettivi, di coordinate spaziali (o retiniche) $(x_\phi, y_\phi)_1$ e $(x_\phi, y_\phi)_2$, sono *molto vicini* ⁽¹⁵⁾.

Nella continuazione eviteremo di essere pedanti riguardo alla distinzione tra le due forme di rappresentazione dell'attività neurale – quella dell'attività di base e quella legata alle osservabili fisiche –; tuttavia è importante comprenderla e ricordarla. Essa ci permetterà, infatti, di distinguere le seguenti variabili, che si differenziano per metodi e logica di misurazione ma sono legate per il referente empirico:

- (i) la velocità fisica del moto di un oggetto rappresentato in un filmato (v_ϕ);
- (ii) la corrispondente velocità percepita (v_ψ);
- (iii) il pattern di attivazione neurale in risposta al filmato (a_v);
- (iv) la velocità dello stimolo codificata dal pattern di risposta neurale (v_v);

§3.3.

Risposte cellulari specifiche per direzione, frequenza, velocità. Cellule semplici e cellule complesse

Il concetto di campo recettivo può essere esteso, nel senso che non si applica solo allo spazio visivo, cioè all'insieme delle possibili posizioni bidimensionali ma, anche, ad altri *spazi*, ovvero insiemi strutturati di *variabili*. Uno di questi è lo spazio delle *frequenze*: sperimentalmente [5-11] si osserva che le cellule di V1 rispondono in maniera preferenziale ad alcune frequenze spaziali piuttosto che ad altre.

Esistono almeno due modi di formalizzare in maniera unitaria la specificità di ogni cellula per l'orientazione, la frequenza e la scala. Uno utilizza le funzioni Gabor e l'altro le derivate della gaussiana. Il primo è quello più noto [11, 55] e trattato in maniera approfondita da [56, 57]; il secondo [58-60] permette notevolmente di snellire i calcoli nelle simulazioni: useremo quest'ultimo, ma nelle figure 7 e 8 sono rappresentati alcuni campi recettivi con il primo metodo. Ogni figura ha una struttura matriciale; in ogni cella c'è una coppia di campi recettivi Gabor, dei quali quello a sinistra è dispari (gaussiana moltiplicata per una funzione seno) rispetto all'asse di simmetria – cioè è di segno differente da una parte e dall'altra dell'asse di simmetria –, mentre quello a destra è pari (gaussiana moltiplicata per una funzione coseno), ovvero sia ha lo stesso segno da una parte e

¹⁵ Si tratta di una funzione tra spazi topologici e si richiede che sia continua. La trasformazione reale è log-polare, cioè al punto di coordinate (ρ, θ) retiniche corrisponde il punto $(\log(\rho), \theta)$ di coordinate corticali.

dall'altra dell'asse di simmetria. Nelle righe vi sono campi recettivi con lo stesso valore di scala – in parole povere con la stessa estensione; nelle colonne, campi recettivi con la stessa orientazione (l'angolo indica l'orientazione dell'asse principale di simmetria rispetto all'asse x secondo la convenzione antioraria). In figura 8 i campi recettivi hanno frequenza doppia rispetto a quelli di figura 7.

Figura 7. Campi recettivi Gabor-like a bassa frequenza spaziale.

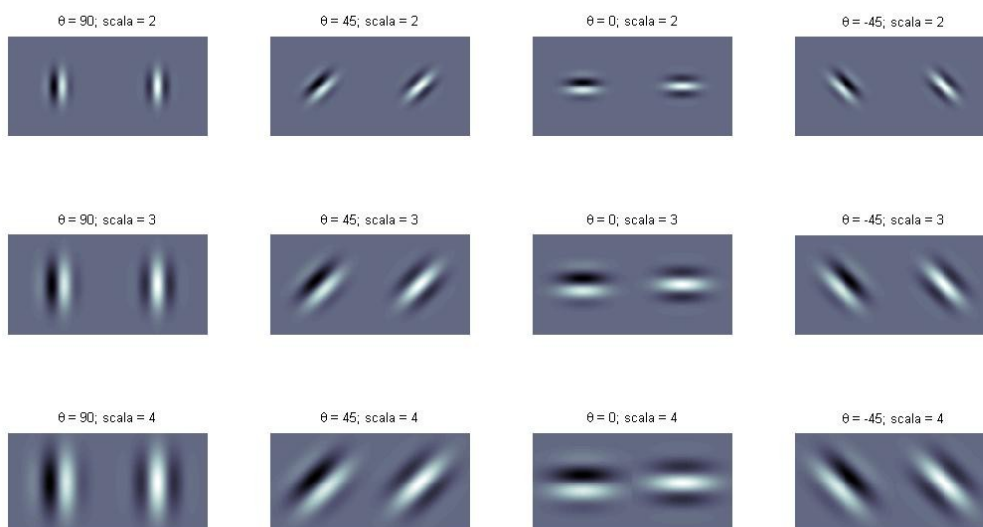
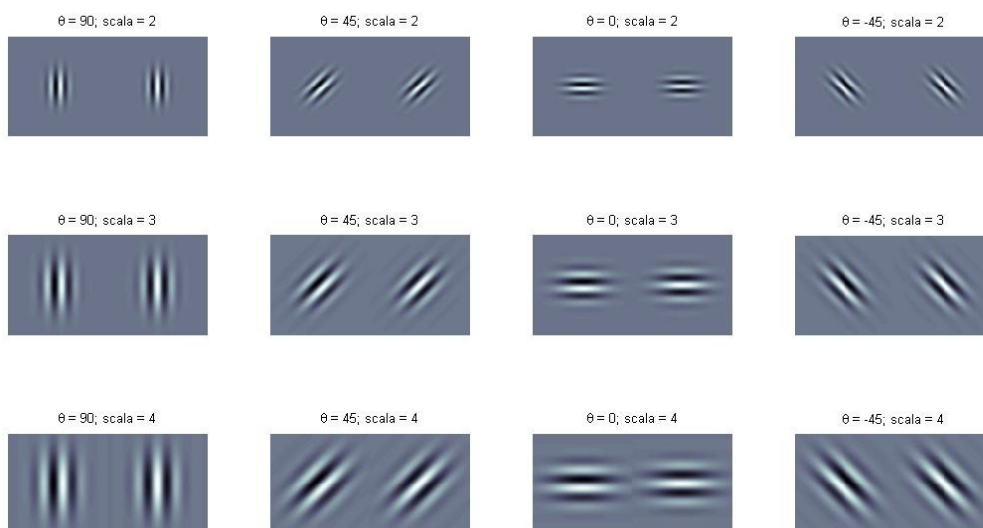


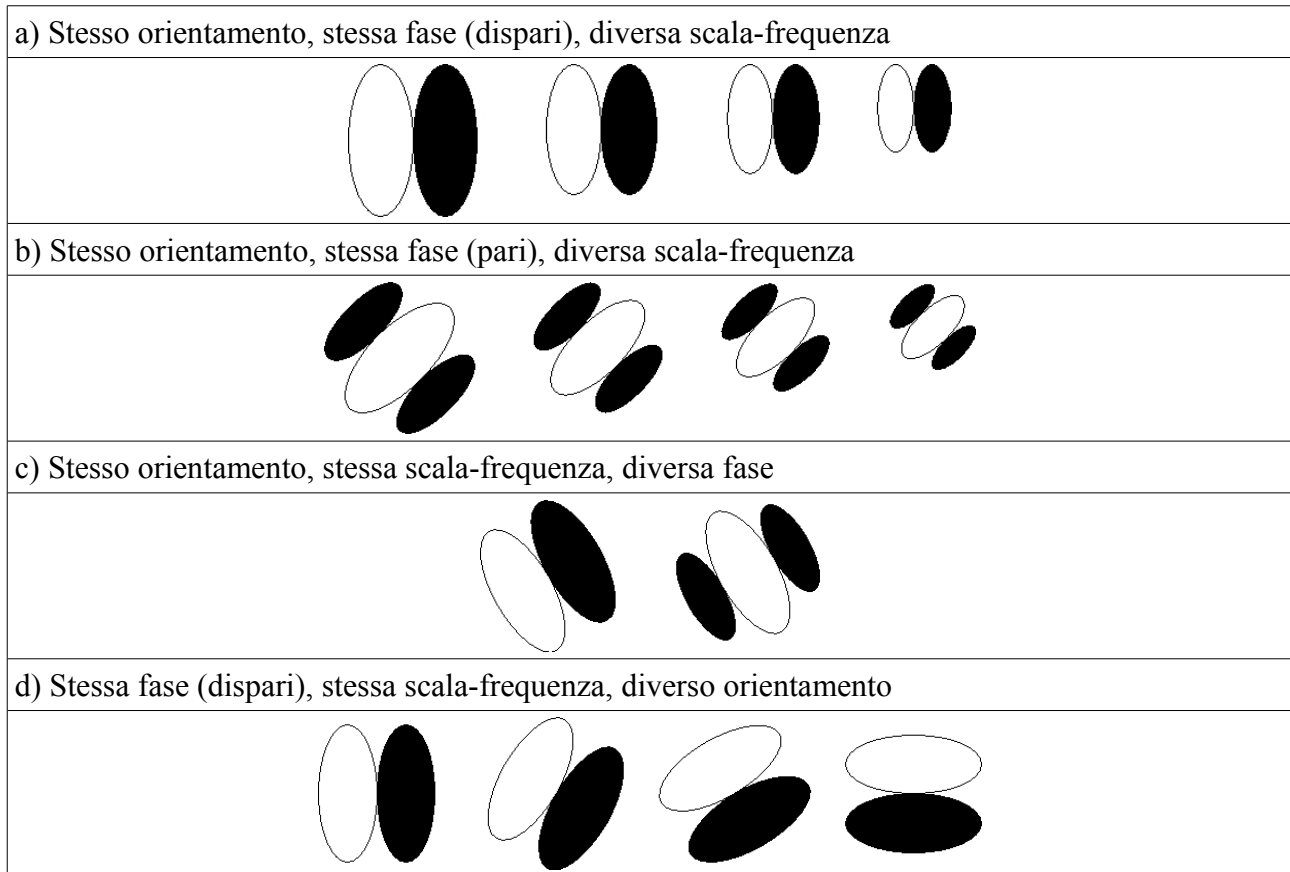
Figura 8. Campi recettivi Gabor-like ad alta frequenza spaziale.



Si è dimostrato che esiste un vincolo tra la frequenza preferita da una unità e l'ampiezza del suo campo recettivo [56, 57]: approssimativamente ampiezza e frequenza sono legate in maniera

inversamente proporzionale. Ciò significa che cellule sensibili ad alte frequenze spaziali hanno campi recettivi piuttosto piccoli mentre cellule con campi recettivi più ampi sono sensibili a frequenze spaziali minori. In figura 9 sono riportati schematicamente i confronti salienti per distinguere le specificità di risposta delle unità secondo i relativi campi recettivi.

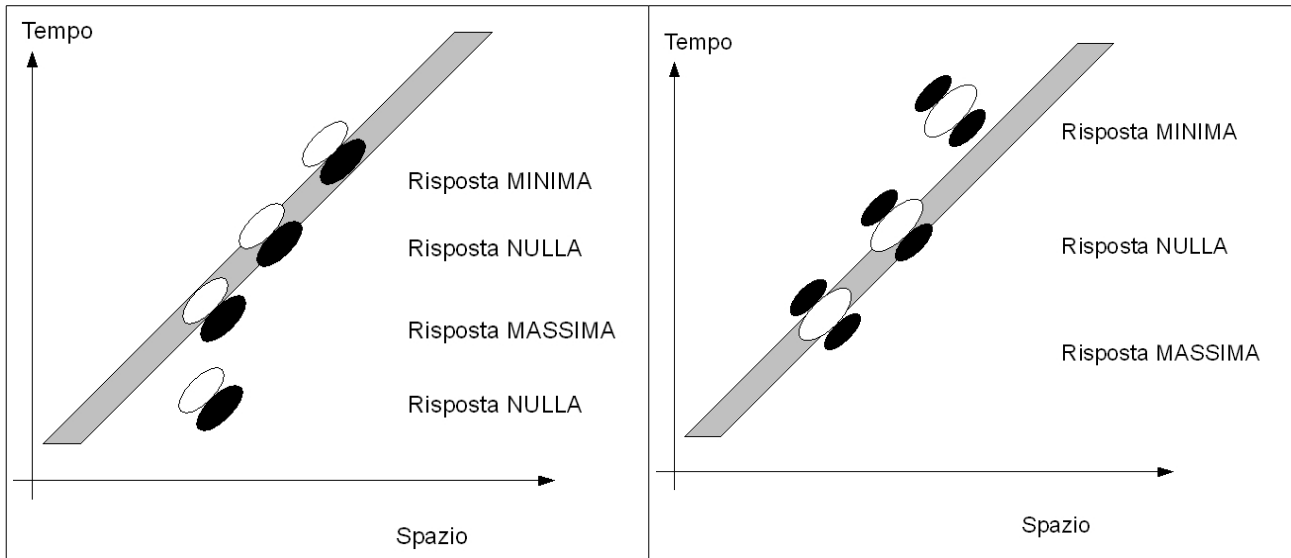
Figura 9. Confronti tra campi recettivi



Le cellule con campi recettivi pari e dispari, come riportati in figg. 7-9, sono dette *cellule semplici* [5-11]. Esse sono sensibili alla frequenza dello stimolo (che si rivela nella grandezza dei lobi), alla fase (il numero dei lobi) e all'orientazione. Vi è un'altra categoria di cellule di cui si deve far menzione: le *cellule complesse*. Come le cellule semplici, sono sensibili alla frequenza spaziale dello stimolo e alla sua direzione ma non alla fase. Si ritiene che la loro risposta sia l'esito della integrazione non lineare della risposta di due cellule semplici che hanno medesime frequenza spaziale e direzione preferite di quelle della cellula complessa ma differiscono nella fase (ossia una è pari e l'altra è dispari) [72, 73].

Come agiscono i campi recettivi? Ossia quale informazione sullo stimolo ci fornisce l'attivazione di una unità con un determinato campo recettivo avente particolari frequenza, ampiezza ed orientazione preferiti? In figura 10 è schematizzata la posizione relativa tra un campo recettivo dispari ed un segnale visivo statico. Si noti che la risposta dell'unità legata al campo recettivo è massima in valore assoluto (quindi massima o minima) quando l'asse di simmetria del campo recettivo, e, quindi, il punto spaziale che rappresenta il centro del campo recettivo, sono su un bordo, cioè su una linea che rappresenta un rapido salto di luminanza; si noti inoltre che la risposta è nulla quando il campo recettivo copre regioni di luminanza uniforme (lo sfondo o la figura). I campi recettivi dispari, quindi, sono *rilevatori di bordi*. I campi recettivi pari, invece, rispondono in maniera preferenziale a barrette orientate.

Figura 10. Come un campo recettivo entra nella codifica neurale di un segnale visivo.

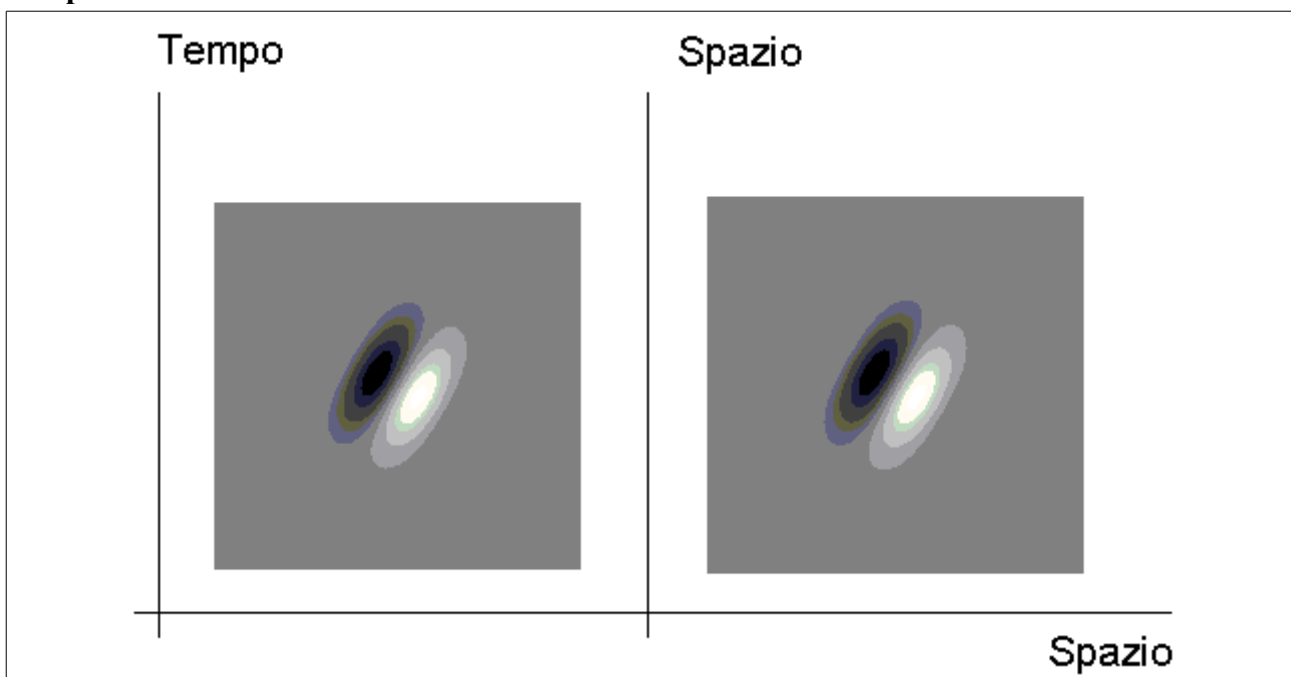


§3.4.

Il campo recettivo come entità intrinsecamente spazio-temporale

Il campo recettivo è stato presentato come una entità statica. Con tecniche di misurazione simili a quelle usate per delineare campi recettivi bidimensionali puramente spaziali [53-54] è possibile ricavare campi recettivi spazio-temporali [61-65].

Figura 11. Confronto tra campo recettivo puramente spaziale e campo recettivo spazio-temporale

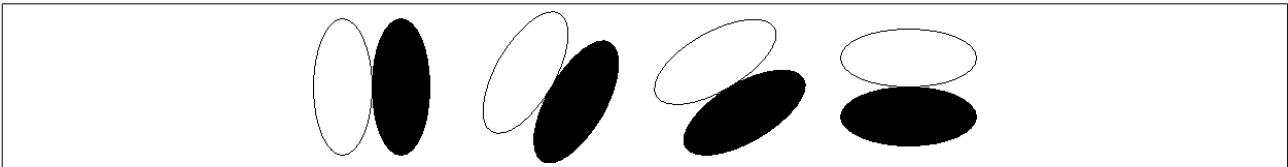


Per semplicità di rappresentazione utilizziamo una sola dimensione spaziale. Ciò che sperimentalmente si ottiene è che la forma spazio-temporale dei campi recettivi è molto simile a

quella spaziale bidimensionale. In figura 11 si enfatizza tale somiglianza (¹⁶).

Anche la funzione neurocomputazionale è simile tra i filtri spaziali e quelli spazio-temporali. La differenza è che un bordo nello spazio-tempo è un segnale di velocità. Quindi l'orientazione spazio-temporale di un campo recettivo rende l'unità cui appartiene sensibile al movimento, cioè alla direzione e all'intensità del moto. La velocità a cui risponde preferenzialmente la cellula è proporzionale all'inclinazione del campo recettivo rispetto all'asse temporale. In figura 12 si riprende la figura 9d; se rappresentiamo il tempo sull'asse verticale e lo spazio su quello orizzontale, i campi recettivi, da sinistra a destra, rispondono in maniera preferenziale a segnali spazio-temporali viepiù veloci.

Figura 12. Cellule semplici sintonizzate su velocità crescenti (da sinistra a destra)



§3.5.

Locale e globale: basi neurofisiologiche del *campo associativo* e struttura colonnare della corteccia striata

Field stesso [44] ha ipotizzato una possibile base neurofisiologica del costrutto psicofisico "campo associativo" e molti lavori si sono occupati di approfondire l'argomento, tra i quali [66-71]. Complessivamente sembra di poter affermare che i correlati neurofisiologici del campo associativo siano le connessioni orizzontali a lungo raggio che attraversano V1, benché si possa immaginare anche un ruolo importante giocato dal *feedback* su V1 proveniente da altre aree.

Il campo associativo, quindi, avrebbe una implementazione neurofisiologica che si presta ad essere interpretata in termini fortemente analogici: le linee di associazione del campo sarebbero l'equivalente funzionale di connessioni assoniche a medio e lungo raggio.

Per comprendere come possa essere strutturato il pattern di connessioni tra le unità visive in modo che siano soddisfatti i vincoli psicofisici del campo associativo, bisogna comprendere come siano organizzati i neuroni visivi in relazione alle loro specificità di risposta. La corteccia visiva primaria è dal punto di vista funzionale suddivisibile in blocchi [5-11], ciascuno dei quali ha un'altezza pari allo spessore della corteccia ed una base, tangenziale alla corteccia, pari ad un quadrato di circa 2 mm × 2 mm. Ognuno di questi blocchi è costituito da cellule che hanno il medesimo centro spaziale del campo recettivo ma differiscono per l'orientazione spaziale preferita (¹⁷), coprendo le direzioni che vanno da 0° a 180°; essi sono detti *moduli corticali* o *ipercolonne*. Le cellule appartenenti al medesimo modulo corticale – ve ne sono sia di semplici sia di complesse – si distinguono per l'orientazione [11] e la frequenza [74]. È interessante notare che all'interno di una iper-colonna, le unità sono funzionalmente organizzate in modo che cellule con campi recettivi di simile orientazione siano vicine tra di loro. Dal punto di vista funzionale, come si ipotizza in alcuni modelli [11, 14, 74, 75, 76], le cellule all'interno di una iper-colonna sarebbero tra di loro organizzate come fossero su un cerchio. Infatti immaginiamo di dover ordinare le unità su una linea mettendo vicine quelle che hanno campi recettivi con orientazione simile (supponiamo che lo spazio delle direzioni sia campionato a blocchi di 15°); avremmo 0° 15° 30° 45° 60° 75° 90° 105° 120° 135° 150° 165°. L'ultima sarebbe 180° ma siccome il periodo campionato è proprio 180°, l'unità il cui campo recettivo ha direzione di 180° è la stessa che ha direzione 0°; quindi l'unico

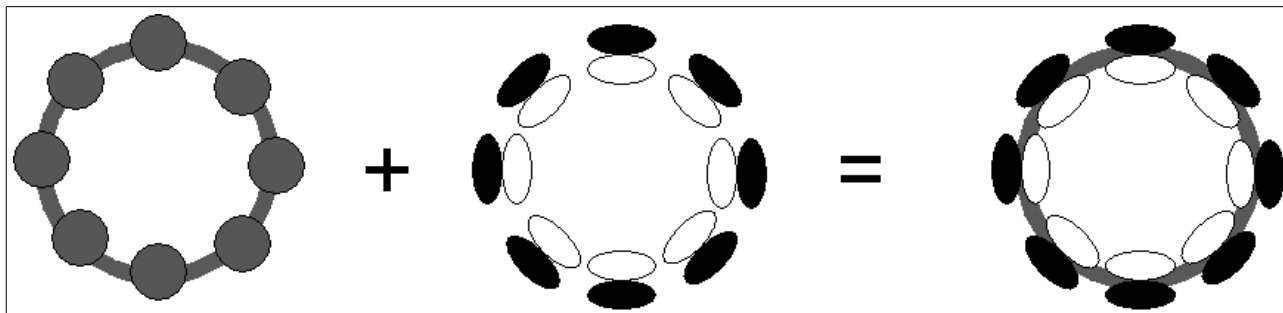
¹⁶ In §5, invece, si discuterà delle *differenze* tra la situazione spaziale e quella temporale. Per mantenere la linearità argomentativa evitiamo di introdurre sin d'ora le specificità della dimensione temporale rispetto a quella spaziale.

¹⁷ Differiscono anche per la dominanza oculare ma noi limitiamo il nostro interesse alla visione ciclopica.

modo di mantenere la vicinanza tra cellule che hanno orientamenti simili e di connettere l'unità specifica per 165° a quella specifica per 0° , organizzandole su un cerchio.

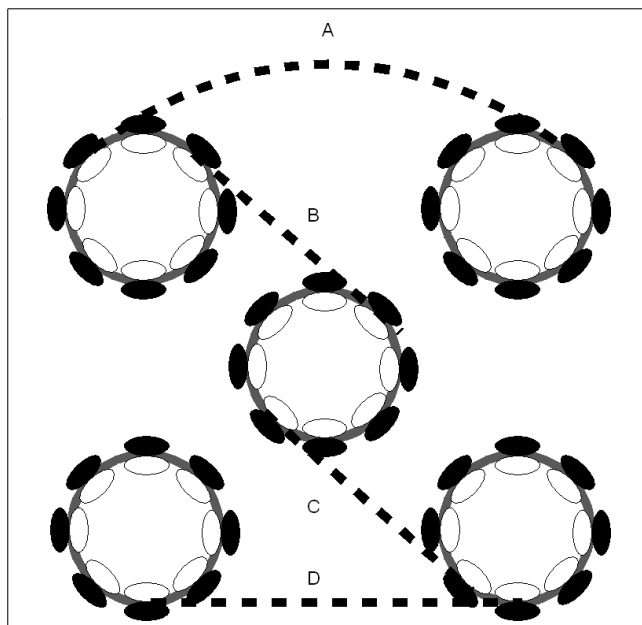
Ora possiamo spiegare come si ipotizza siano collegate cellule di *diverse* iper-colonne affinché tale pattern di connessione fornisca una base neurofisiologica al campo associativo. Anzitutto supponiamo che in ogni colonna si coprano gli interi 360° e non solo 180° . Figura 13 rappresenta in maniera schematica tale situazione. Nel primo "addendo" ogni punto grigio rappresenta una unità neuronale e l'intero cerchio il range di unità (differenti per direzione preferita) di una medesima iper-colonna. Il secondo "addendo" rappresenta i rispettivi campi recettivi. La loro sovrapposizione dà una versione stenografica dell'organizzazione di connessioni *interne ad una iper-colonna* tra unità che differiscono per orientazione preferita.

Figura 13. Rappresentazione funzionale di una iper-colonna.



Un modello neurofisiologico del campo associativo è ragionevole se spiega come interagiscono tra di loro le unità all'interno di una iper-colonna e tra iper-colonne differenti, in modo che queste interazioni siano compatibili con i risultati psicofisici. In figura 14 è schematizzata la versione tipica nell'ambito dei modelli matematici [14, 76, 77].

Figura 14. Rappresentazione della interazione tra diverse iper-colonne, compatibile con il costrutto psicofisico di campo associativo.

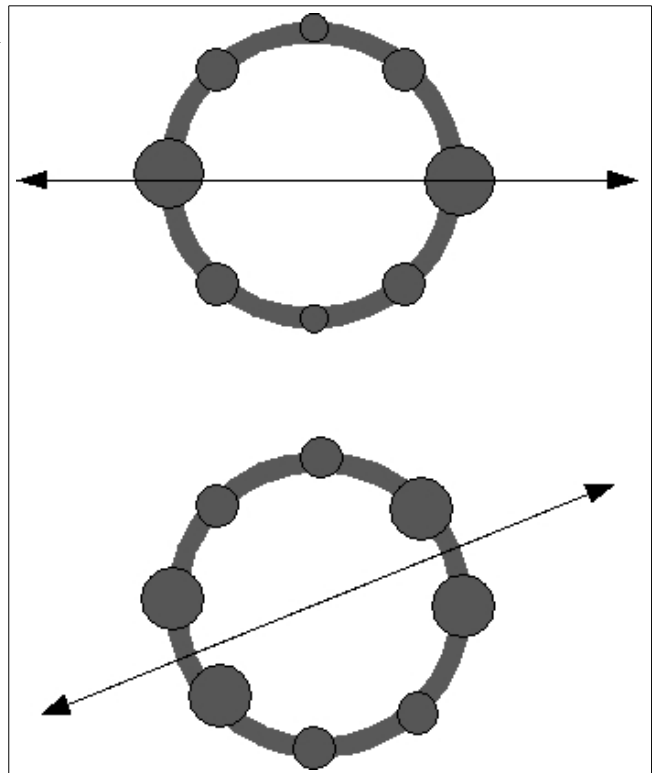


Ciascun cerchio rappresenta un differente modulo corticale e, quindi, un distinto punto sul piano corticale, in organizzazione retinotopica. Le connessioni indicate con A, B, C e D associano tra loro unità neuronali appartenenti a diverse iper-colonne. Le connessioni B, C e D implementano la buona continuazione di contorni o bordi lineari; la connessione A, invece, rappresenta l'associazione tra unità appartenenti ad un contorno o bordo *curvo*. La presenza di una linea di associazione tra due unità denota, in termini neurofisiologici, la presenza di due connessioni sinaptiche tra le stesse (una

in una direzione e l'altra nella direzione opposta) che fanno sì che le attività delle due unità correlino.

Per comprendere, invece, come dovrebbero interagire le unità *entro* una medesima iper-colonna, bisogna considerare due fatti. Anzitutto – osserviamo la figura 4 – come abbiamo detto il patch u è associato con il patch A ma non con il patch F; quindi, passando dai patch alle unità di una iper-colonna, è sensato supporre che unità che hanno orientazioni preferite *perpendicolari* siano non associate o, addirittura, si inibiscano [14, 74, 76, 77, 78]. Inoltre prendiamo in considerazione la *decodifica* del pattern di attivazione di unità all'interno di un modulo corticale. In figura 15 sono rappresentati due pattern di attività di una iper-colonna. Come si nota, per il fatto che in una iper-colonna sono campionati solo 180° e non 360° , sul cerchio il pattern di attività ha periodo 180° (cioè si ripete ogni 180°). La risposta del modulo corticale nel suo complesso è una *direzione complessiva*, che può essere pensata come quella di un asse principale di inerzia di una massa filiforme a forma di circonferenza (un anello sottile), la cui densità lineare di massa sia proporzionale all'attività neuronale. Tale "direzione complessiva" rappresenta la direzione locale del bordo dell'immagine (la sua tangente o la sua normale: è indifferente in due dimensioni) e, quindi, il sistema di *dialogo interno* tra le unità di una iper-colonna, deve essere tale che le unità tra di loro perpendicolari siano non-associate o, addirittura, in reciproca inibizione.

Figura 15. Decoding della risposta di un modulo corticale.



§3.6.

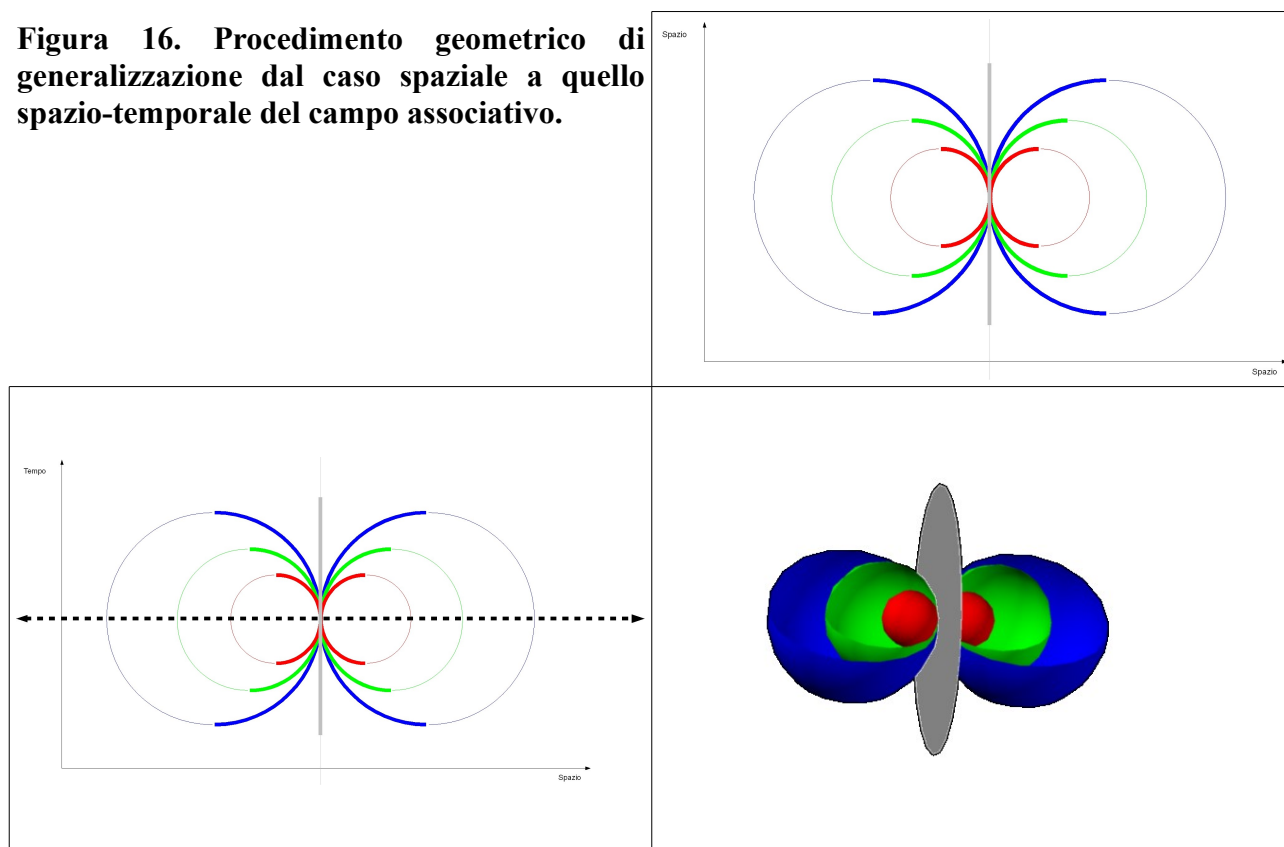
Il campo associativo spazio-temporale

Nel caso psicofisico abbiamo postulato l'esistenza di un campo associativo spazio-temporale. Supponiamo che ne esista il corrispettivo neurale e che questo possa essere ottenuto per generalizzazione di quello bidimensionale trattato nel paragrafo precedente. In figura 16 è presentata una versione psicofisica di tale generalizzazione. Sono rappresentate due versioni schematiche del campo associativo: una nello spazio bidimensionale, l'altra nella sezione spazio-temporale dello spazio-tempo tridimensionale: esse sono formalmente identiche. La forma del

campo agente nello spazio-tempo tridimensionale si ottiene come solido di rotazione attorno all'asse tratteggiato in figura. Come già accennato nell'ambito della trattazione psicofisica, alle linee di associazione si sostituiscono piani e, alle curve (archi di circonferenza), sfere (calotte sferiche). Le linee di associazione curve sono state completate (con linee di minor spessore) nelle circonferenze di cui sono parte per enfatizzare, dopo la rotazione, la tridimensionalità del campo associativo.

Da un punto di vista neurofisiologico come possiamo ipotizzare sia implementato questo campo associativo tridimensionale? Come varia la relazione tra neuroni appartenenti al medesimo modulo corticale? Il modello che proponiamo dà una risposta in tal senso [§5.2].

Figura 16. Procedimento geometrico di generalizzazione dal caso spaziale a quello spazio-temporale del campo associativo.



§3.7.

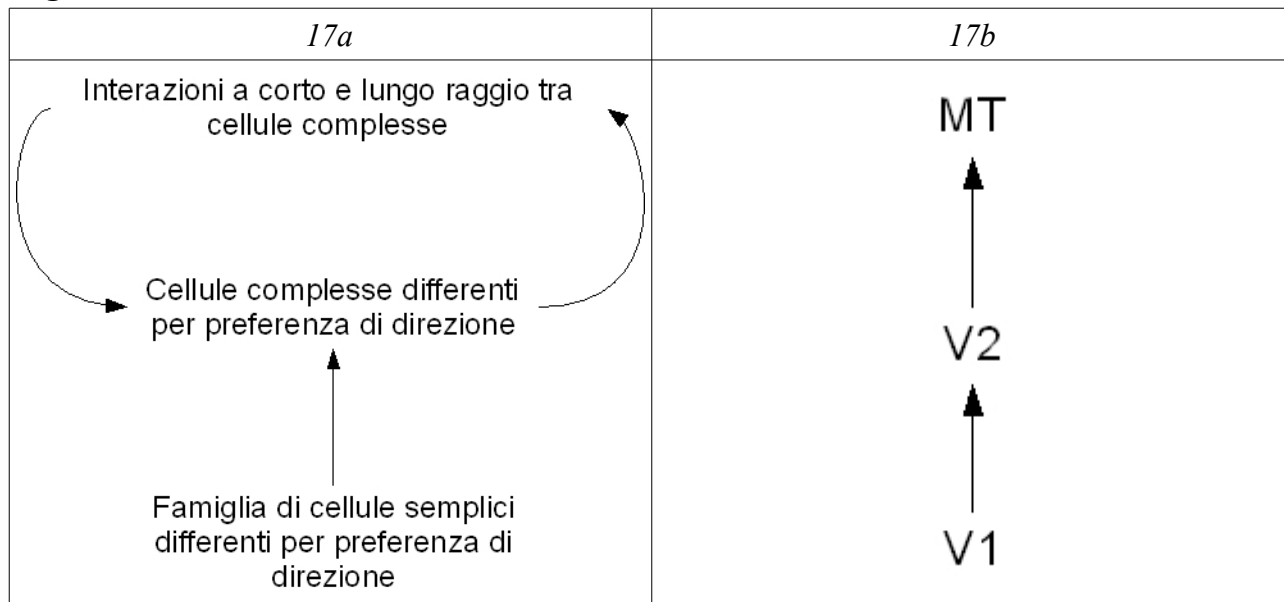
Le vie neurali della percezione del movimento

Il percorso neurale di elaborazione dei segnali di movimento [11, 38, 78] può essere semplificato come nello schema di figura 17. Bisogna notare che lo schema di sinistra non è parallelo a quello di destra, nel senso che le cellule complesse sarebbero in MT e quelle semplici in V2 e V1: infatti vi sono cellule complesse in V1 e vi sono cellule sensibili al movimento sia in V1 sia in MT. Noi ci limitiamo a descrivere due livelli possibili di schematizzazione del flusso di informazione: uno tra neuroni e l'altro tra aree.

Per i nostri scopi è più interessante la schematizzazione di [Fig. 17a]: il modello proposto in questa tesi, infatti, è suddiviso nei tre livelli delineati. A livello delle cellule semplici avviene una prima elaborazione dello stimolo: le unità rispondono in maniera preferenziale a bordi e barre a diverse orientazioni e frequenze spaziali. Coppie di cellule semplici che hanno medesime frequenza ed orientazione preferite, ma diversa fase, sono associate ad un'unica cellula complessa: il loro output viene da questa sommato dopo essere sottoposto ad una forma di non linearità. Le cellule complesse, quindi, rappresentano le stesse informazioni delle cellule semplici (stimano in maniera

distribuita la direzione di contorni) ma perdono metà periodo [11, 73]. Parte dell'input complessivo alle cellule complesse è di tipo feedforward e proviene dalle cellule semplici; l'altra parte è dovuto alle connessioni *lateral*i che legano le cellule complesse tra di loro; tali connessioni sono di due tipi: quelle che legano cellule complesse appartenenti a moduli corticali distinti implementano il costruito psicofisico di "campo associativo"; le connessioni, che legano cellule complesse appartenenti alla stessa iper-colonna, servono ad implementare la procedura di integrazione di segnali locali e di filtraggio dal rumore.

Figura 17. Le vie visive del movimento.



§4. La prospettiva matematica

§4.0. Introduzione

Al livello psicofisico si tratta dei dati e delle variabili fisiche e fenomeniche. Al livello neurofisiologico, di dati e variabili sia fisiche sia biofisiche e neurofisiologiche. Il primo capitolo è centrato sull'argomentare ed enfatizzare che modelli e teorie agiscono con le rappresentazioni e sulle rappresentazioni; ciò ha due implicazioni. La prima è che la matematica è estremamente utile al percettologo; la seconda è che tale utilità può palesarsi anche con una certa indipendenza dal versante empirico.

L'indipendenza dell'approccio matematico nell'*Image Processing* è più evidente nelle discipline, come la *Computer Vision*, che non hanno come interesse fondamentale la plausibilità neurofisiologica o quella psicologica. Ciò nonostante processi postulati per risolvere problemi ingegneristici di elaborazione dei segnali sembra siano stati usati anche dalla natura.

Vi sono diverse generazioni di modelli matematici che si propongono di integrare psicofisica e neurofisiologia. Il capostipite è considerato il modello di Reichardt. Il più recente Modello di Energia ne è una generalizzazione ed è più specifico su come sarebbero implementati i processi di elaborazione del segnale di movimento. Il nostro scopo è di aumentare tale specificità.

§4.1. Gli approcci potenzialmente legati da vincoli neurofisiologici o psicologici

Gli ambiti della *Computer Vision* e dell'*Image Processing* sono prettamente ingegneristici e finalizzati all'elaborazione di immagini per identificare in maniera automatica alcuni *token*, ossia caratteri primari dell'immagine, come bordi o componenti figurali distinte dallo sfondo [92]. Questi sono anche i compiti del sistema visivo umano – si pensi al *primal sketch* di Marr –, viene perciò naturale chiedersi se vi sia differenza alcuna con l'approccio neuroscientifico o percettologico. La questione è complessa, tuttavia una semplificazione dignitosa è la seguente: nella *Computer Vision* e nell'*Image Processing*, la qualità degli algoritmi si basa precipuamente solo sulla loro efficienza. Importa poco che le parti in cui la procedura è suddivisa sia passibile di interpretazione nei termini di processi psicofisici o neurofisiologici: ciò che conta è che con il minor carico computazionale possibile, nei termini di spazio e tempo-macchina necessari per i calcoli, la procedura algoritmica in questione permetta di ottenere i *token* per cui è stata costruita. Per esempio un ingegnere del suono potrebbe impegnarsi a progettare sistemi di filtraggio per eliminare il rumore di fondo da registrazioni, oppure per classificare i pattern vocali presenti in modo da poterli confrontare con altri e scoprire se la voce nella registrazione in questione è la stessa, registrata in condizioni controllate, di un certo individuo. Nel fare ciò potrebbe essere totalmente disinteressato al fatto che dalle cellule cigliate alla corteccia uditiva succeda qualcosa di simile. Considerazioni simili valgono per eventuali algoritmi che permettano di produrre perizie calligrafiche o riconoscimento di volti.

Il fatto interessante è che non di rado questi algoritmi si rivelano neurofisiologicamente, cognitivamente o fenomenologicamente plausibili. Un esempio è l'uso di filtri ottenuti prendendo il laplaciano della gaussiana (LoG): essi hanno una forma che è sovrapponibile a quella dei campi recettivi delle cellule del nucleo genicolato laterale ed i punti di *zero-crossing* dell'esito del filtraggio, ossia i punti in cui si ha un passaggio da un valore positivo ad uno negativo, appartengono ai bordi delle figure dell'immagine.

Un altro esempio è fornito dal modello di *Tensor Voting* di Medioni [79], che comunque, va

detto, esplicitamente fa riferimento a vincoli fenomenologici in parte del suo sviluppo. In esso l'informazione *locale* è codificata da una struttura matematica (e geometrica), chiamata *tenore simmetrico*, che raccoglie l'informazione che, dal punto di vista neurofisiologico, è elaborata in maniera distribuita da una iper-colonna: la stima della direzione di un bordo e la *credibilità* di tale stima. Quindi la rappresentazione usata da Medioni è un *campo tensoriale*, ossia una funzione che ad ogni punto dello spazio associa un tensore: un ente matematico che codifica quale tipo di token possa esserci nel punto in questione (un bordo o un solo punto, nel caso di elaborazione di immagini) e la relativa *verisimiglianza*. Egli costruisce un campo sullo spazio che permette ai tensori stimati di interagire tra di loro, similmente a quanto fanno i percettologi quando postulano un campo associativo (psicofisico o neurofisiologico) che permette a diversi moduli corticali (iper-colonne) di influenzarsi vicendevolmente. Va enfatizzato che la sua ipotesi di campo si basa, come anticipato, su considerazioni percettive; precisamente sulle Leggi della Gestalt, in particolar modo su quella di buona continuazione.

§4.2.

I modelli neurofisiologici, psicologici e psicofisiologici

L'ambito delle reti neurali è sterminato [16-20]. Vale la pena citare due lavori [80, 81] in cui si usa un meccanismo di analisi statistica (ICA, Independent Component Analysis), implementato in una rete neurale, per stimare i filtri ideali di elaborazione di un certo insieme di immagini o filmati (nella fattispecie immagini o filmati naturalistici, ossia che rappresentano paesaggi). Ci sono due risultati interessanti: i filtri spaziali che si ottengono sono simili (gabor o derivate di gaussiane) a quelli con cui si modellano le cellule semplici, così come i filtri spazio-temporali.

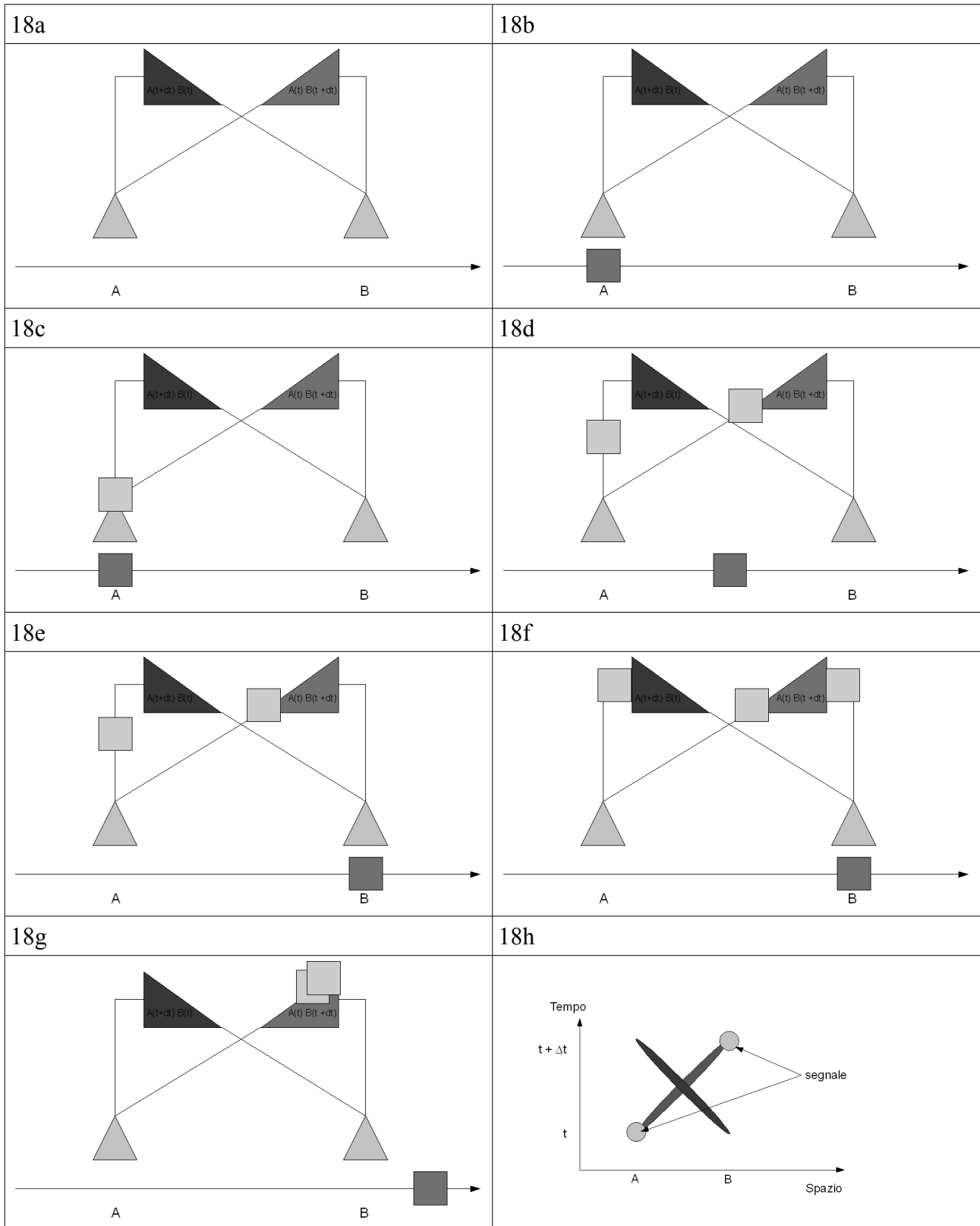
Il primo, almeno per autorevolezza e popolarità, modello neurofisiologico che si è proposto di fornire le basi biologiche della percezione del movimento è quello di Reichardt [9, 11, 82, 83]. Quello più recente e che gode di maggior prestigio è il *modello di energia* di Adelson e Bergen [9, 11, 84, 85, 86]: esso contempla il modello di Reichardt come caso particolare ed è quello su cui si basa il nostro.

In figura 18 è riportata una versione schematica del modello di Reichardt. Nei grafici 18a-g sono presentati in sequenza diversi istanti temporali e il relativo flusso del segnale verso il detettore di Reichardt. Uno stimolo parte dal punto A ed arriva a B. I due comparatori rispondono alla presenza di un segnale in due parti dello spazio, in modo da dare risposta massima quando tra i due c'è un certo ritardo temporale. Il comparatore di sinistra risponde massimamente quando in un istante t è presente un segnale nel punto B e nell'istante $t + \Delta t$ è presente un segnale nel punto A. In altri termini il comparatore di sinistra risponde in maniera specifica ad un segnale di movimento costituito da un corpo che si muove verso *sinistra* ad una velocità scalare pari a $(B - A)/\Delta t$. Simmetricamente, il comparatore a destra è specifico per un movimento alla medesima velocità scalare, ma verso *destra*.

Se utilizziamo una rappresentazione spazio-temporale, si nota che i due moduli che fungono da comparatori possono essere rappresentati come *filtri* che agiscono sullo spazio-tempo ed associano, appunto, punti spazio-temporali. La figura 18h mostra questa interpretazione.

Il modello di energia è applicato con efficacia alla modellizzazione della percezione di immagini statiche [11] e la sua estensione ai meccanismi di percezione del movimento è una semplice generalizzazione che segue quella sopra citata [§3.4] sulla forma dei campi recettivi delle cellule semplici, già definiti come “entità intrinsecamente spazio-temporali”, e le considerazioni testé riportate sulla interpretazione del modello di Reichardt in termini di correlatori spazio-temporali.

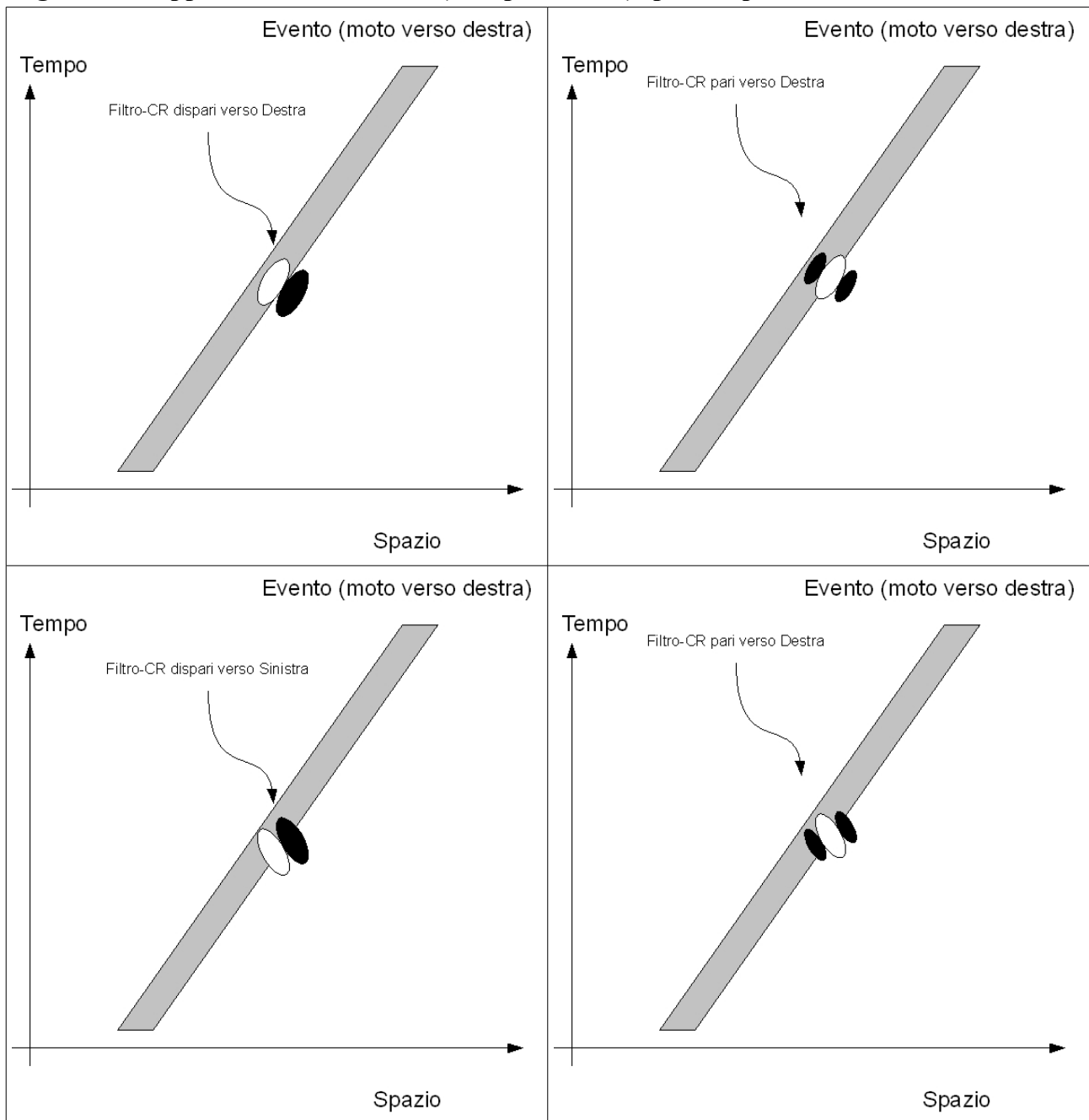
Figura 18. Schematica versione del modello di Reichardt.



In figura 19 è riportato, in analogia a [Fig. 10], un segnale di movimento in rappresentazione spazio-temporale. Ad esso sono sovrapposte due coppie di campi recettivi (CR) ovvero, dal punto di vista neurocomputazionale, *filtri*: due sono specifici per il movimento verso destra e due per il movimento verso sinistra. Ogni coppia è formata da due filtri di uguali scala e direzione, ma fase

differente: un filtro è pari e l'altro dispari.

Figura 19. Rappresentazione di filtri (Campi recettivi) specifici per il movimento.



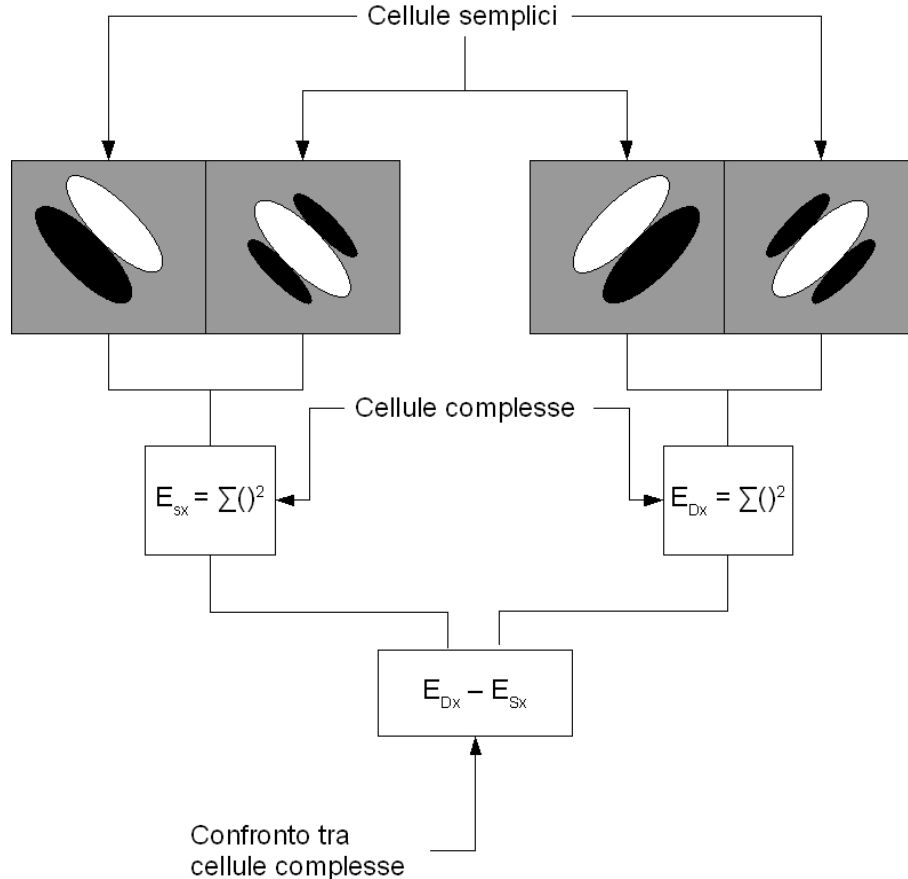
Il modello di energia, nella sua versione classica, è applicato a rappresentazioni spazio-temporali *bidimensionali* del movimento (cioè con una sola dimensione spaziale) ed opera nel seguente modo:

- (i) ad ogni punto dello spazio-tempo sono associate *due* coppie di filtri; una coppia di filtri è sensibile al movimento verso destra e l'altra verso sinistra; all'interno di ciascuna coppia i filtri differiscono solo per la fase: uno è pari (tre lobi) e l'altro è dispari (due lobi).
- (ii) Per ogni coppia, la risposta di ciascun filtro della coppia viene elevata al quadrato e sommata all'altro. Per ogni punto dello spazio-tempo si ottengono così due numeri, uno per ogni coppia, ciascuno dei quali rappresenta quanto lo stimolo abbia, in quel determinato

punto spazio-temporale, la pendenza spazio-temporale codificata dalla coppia cui il valore si riferisce.

- (iii) La risposta complessiva del sistema è ottenuta sottraendo, in ogni punto dello spazio-tempo, la stima del movimento verso sinistra a quella del movimento verso destra. Un numero positivo è interpretato come la presenza di un segnale di movimento verso destra; simmetricamente un numero negativo denota la presenza di un segnale di movimento verso sinistra.

Figura 20. Il modello di Energia classico.



In figura 20 è proposta una versione schematica del modello di energia classico. La risposta del primo stadio del modello è una funzione $E(x, t, \theta)$ – dove (x, t) è il punto spazio temporale in cui si trova il centro del campo recettivo con orientazione θ –, il cui valore si ottiene sommando il quadrato dell'attività della cellula *pari* al quadrato dell'attività della cellula *dispari*. Nella sua versione più semplice – quella che possiamo ritenere standard [86] – il modello di energia usa, per ogni punto dello spazio, solo due direzioni ($\theta = \pm 45^\circ$) per ogni tipo di campo recettivo (pari o dispari); si ottengono così due funzioni $E(x, t, -45^\circ)$ ed $E(x, t, +45^\circ)$ che possiamo, per comodità, denotare rispettivamente con $E_{DX}(x, t)$ e $E_{SX}(x, t)$. La predizione del modello si basa sul confronto, punto per punto, tra l'energia verso destra E_{DX} e l'energia verso sinistra E_{SX} . Precisamente si costruisce una terza funzione, che possiamo chiamare $m(x, t)$, ponendo $m(x, t) = E_{DX}(x, t) - E_{SX}(x, t)$. Se $m(x, t)$ è maggiore di zero significa che il punto (x, t) è parte di un segnale-bordo che pende più verso destra che verso sinistra – ossia in (x, t) c'è un segnale di velocità verso destra; viceversa se $m(x, t)$ è minore di zero.

Le limitazioni principali del modello classico sono le seguenti:

- (i) si usano solo due dimensioni: una spaziale e una temporale;
- (ii) per ogni punto si usano solo due direzioni spazio-temporali;
- (iii) si fa riferimento a un solo valore di scala (ampiezza del campo recettivo-filtro);
- (iv) si analizza solo l'integrazione locale dei segnali di movimento, senza occuparsi della integrazione globale o a lungo raggio.

Diversi lavori, tra cui [17, 18, 87, 88, 89, 90, 91], si sono proposti di rispondere ai quesiti impliciti nei punti (i-iv) sopra riportati. Il nostro modello va nella medesima direzione e si occupa di (i), (ii) e (iv).

§5. Un modello di integrazione dei segnali di movimento

5.0. Introduzione

Il primo paragrafo ha lo scopo di introdurre le nozioni basilari del modello di integrazione dei segnali di movimento nel caso bidimensionale. Già in due dimensioni, infatti, è possibile rendere conto dei tre costrutti principali che generalizzano il Modello Classico di Energia, ovverosia l'utilizzo di connessioni a lungo (tra moduli corticali distinti) e corto (all'interno del medesimo modulo corticale) raggio tra cellule complesse ed il metodo di *decoding* del pattern di risposta della popolazione di unità complesse appartenenti al medesimo modulo corticale.

La relativa semplicità, nel caso bidimensionale, di rappresentazione grafica dei costrutti del modello permette di introdurre con più efficacia la parte formale.

Nel secondo paragrafo è presentato il modello oggetto della tesi, che si propone di elaborare filmati, cioè segnali di input spazio-temporali a tre dimensioni. Esso è proposto come naturale estensione del caso bidimensionale, resa possibile dall'aver trattato il tempo in maniera analoga a come è trattato lo spazio.

Il terzo paragrafo, infine, raccoglie i risultati di alcune simulazioni svolte per valutare il modello.

In tutto il capitolo è costante il riferimento a [59, 60, 94-103].

§5.1. Il caso bidimensionale statico

Riassumiamo anzitutto quali sono i vincoli da cui dobbiamo partire. Vogliamo generalizzare il modello di energia classico lungo due direzioni: la dimensionalità (dallo spazio-tempo bidimensionale allo spazio-tempo tridimensionale) e l'ampiezza del campo di influenza sulle cellule complesse (dal semplice flusso *feedforward* proveniente dalle cellule semplici alle connessioni laterali *tra* cellule complesse).

Partiamo, perché più semplice, dal caso di input bidimensionale *statico*.

§5.1.1. Il flusso feedforward dall'input alle cellule complesse

Nel modello di energia classico vi sono tre stadi di computazione. L'input è rappresentato da una funzione $I(x, y)$ che denota la luminanza o il contrasto presente nel punto (x, y) . Il secondo stadio è la codifica operata dalle cellule semplici, dispari e pari, rappresentata dalla coppia di funzioni $A^S(x, y, \zeta, \sigma)$ e $A^C(x, y, \zeta, \sigma)$; $A^S(x, y, \zeta, \sigma)$ rappresenta l'attivazione della cellula semplice dispari, il cui campo recettivo è centrato nel punto (x, y) , ha orientazione ζ e ampiezza σ ; $A^C(x, y, \zeta, \sigma)$ rappresenta l'attivazione della cellula pari corrispondente. Coppie siffatte di cellule semplici proiettano i loro output su una cellula complessa $E(x, y, \zeta, \sigma)$.

Conveniamo di denotare con $n(x, y, \zeta, \sigma)$ l'unità neuronale che abbia campo recettivo centrato nel punto spaziale (x, y) , direzione preferita ζ e valore σ di scala-frequenza del campo recettivo. Eventualmente aggiungeremo apici o pedici per distinguere le cellule semplici da quelle complesse – usando, perciò, una notazione del tipo $n^F(x, y, \zeta, \sigma)$ per indicare una cellula complessa oppure $n^{AC}(x, y, \zeta, \sigma)$ per una cellula semplice con campo recettivo pari. L'enfasi sulla differenza tra

l'unità e la sua attivazione non sarà sempre necessaria e, quindi, utilizzeremo spesso la notazione nei termini dell'attivazione per riferirsi alla unità: scriveremo così “la cellula complessa $E(x, y, \xi, \sigma)$ ” al posto del più corretto “la cellula complessa $n(x, y, \xi, \sigma)$ la cui attivazione è $E(x, y, \xi, \sigma)$ ”.

La relazione tra le funzioni I, A^S, A^C ed E implementa il flusso di elaborazione *feedforward* che dalla retina va alla corteccia visiva primaria. L'attivazione delle cellule semplici è mediata dai loro campi recettivi.

I campi recettivi delle cellule semplici sono posti nella forma di derivate direzionali della funzione gaussiana $g_\sigma(x, y)$; precisamente, posto

$$g_\sigma := e^{-\frac{x^2+y^2}{\sigma^2}} \quad (2)$$

ed identificato con \mathbf{u} il versore con direzione ξ ,

$$\mathbf{u} := (\cos(\xi), \sin(\xi)) \quad , \quad (3)$$

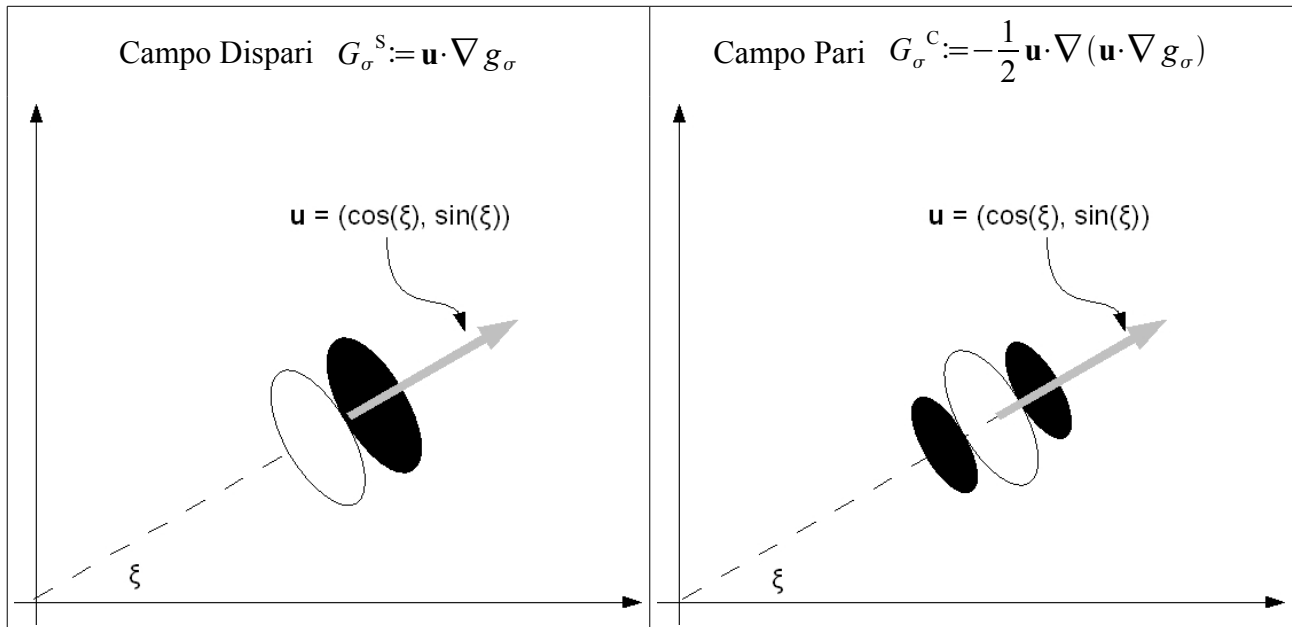
si rappresentano i campi recettivi G_σ^S e G_σ^C [Fig. 21], rispettivamente dispari e pari, delle cellule semplici tramite

$$G_\sigma^S := \mathbf{u} \cdot \nabla g_\sigma \quad (4)$$

e

$$G_\sigma^C := -\frac{1}{2} \mathbf{u} \cdot \nabla (\mathbf{u} \cdot \nabla g_\sigma) \quad . \quad (5)$$

Figura 21. Campi recettivi pari e dispari



Per comodità identifichiamo la coppia di cellule semplici A^S, A^C tramite un solo ente matematico, la funzione complessa $A = iA^S + A^C$. Similmente poniamo $G_\sigma = iG_\sigma^S + G_\sigma^C$.

L'attivazione delle cellule semplici è implementata tramite una operazione detta

convoluzione tra i loro campi recettivi ed il segnale di input I . Formalmente si ha che

$$A = G_\sigma * I = (G_\sigma^C + i G_\sigma^S) * I = G_\sigma^C * I + i G_\sigma^S * I = A^C + i A^S \quad (6)$$

L'attivazione delle cellule complesse è ottenuta sommando i quadrati dell'attivazione delle cellule semplici corrispondenti, cioè:

$$E = A \bar{A} = (A^C)^2 + (A^S)^2 \quad (7)$$

Il valore $E(x, y, \zeta, \sigma)$ rappresenta la stima della direzione della normale al bordo dell'immagine I nel punto (x, y) al valore di scala σ . In figura 22 è presentato lo schema generale di attivazione feedforward delle cellule complesse. Il fatto che l'attivazione delle cellule complesse avvenga tramite i quadrati dell'attivazione di corrispondenti cellule semplici determina la perdita dell'informazione sul segno del bordo, come rappresentato nelle figure 23 e 24. Ciò fa sì che, all'interno di un modulo corticale, le cellule complesse forniscano un campionamento delle direzioni preferite con periodo dimezzato rispetto a quello delle cellule semplici: di π rad (180°), invece che di 2π rad (360°): a riguardo si confrontino, sia in figura 23 sia in figura 24, le risposte per filtri con preferenza a -90° e a $+90^\circ$.

Figura 22. Flusso feedforward: schema generale

$I(x, y)$	*	$G^S(x, y, \zeta, \sigma)$	=	$A^S(x, y, \zeta, \sigma)$	$(\rightarrow)^2$	$A^S(x, y, \zeta, \sigma)^2$
						+
$I(x, y)$	*	$G^C(x, y, \zeta, \sigma)$	=	$A^C(x, y, \zeta, \sigma)$	$(\rightarrow)^2$	$A^C(x, y, \zeta, \sigma)^2$
						=
						$E(x, y, \zeta, \sigma)$

§5.1.2.

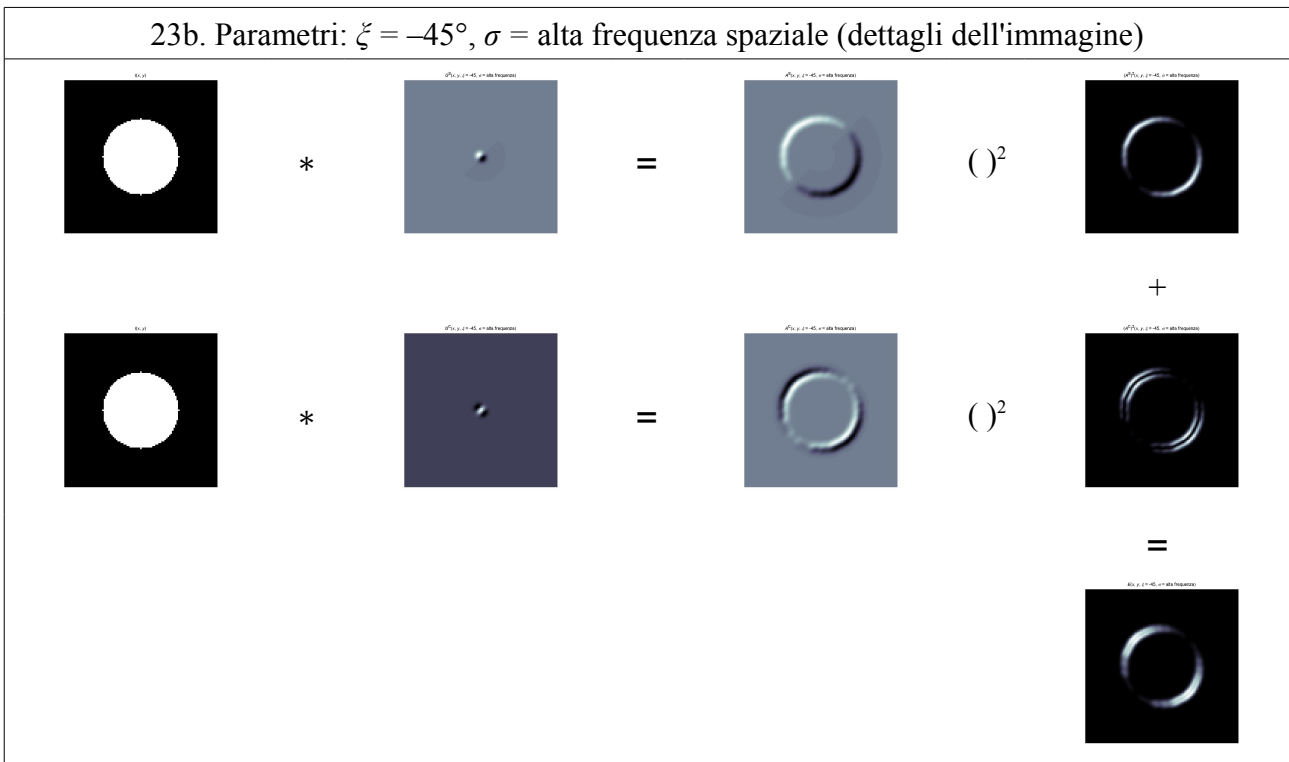
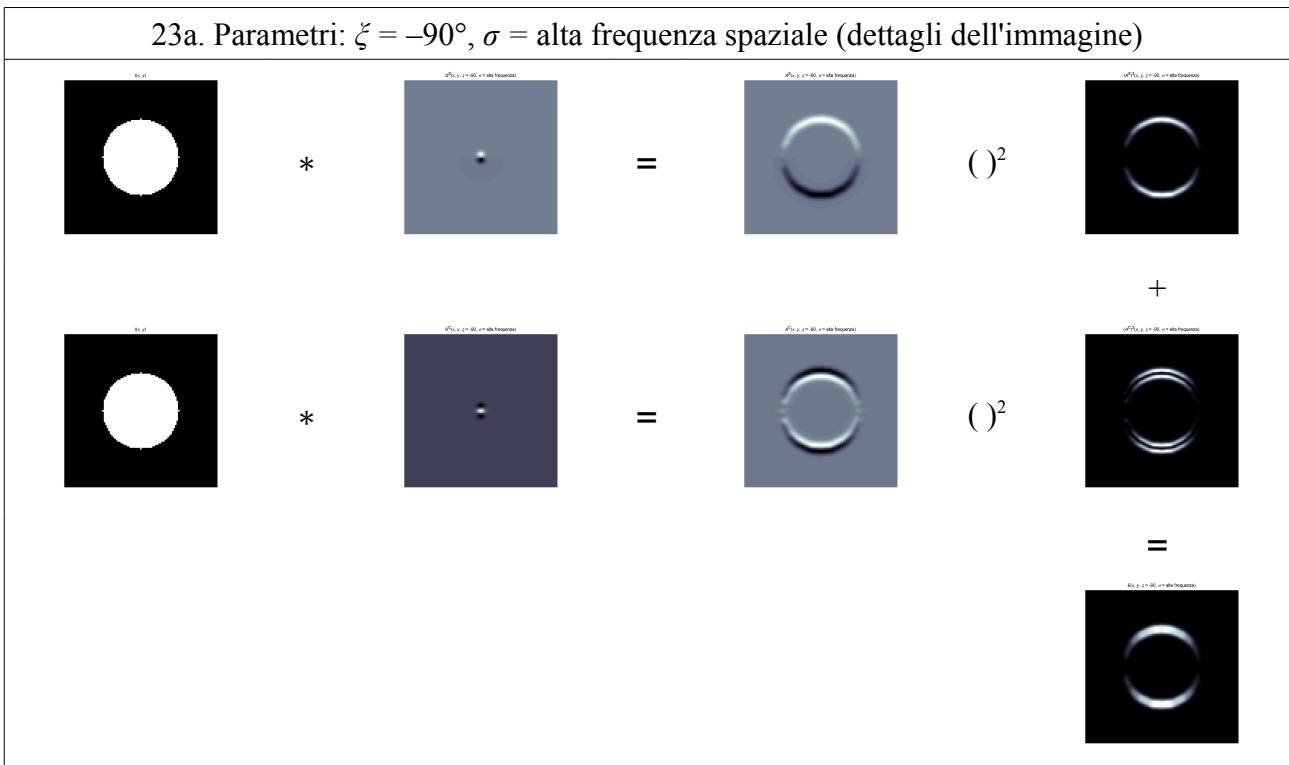
Il flusso tra le cellule complesse: connessioni laterali e corto e lungo raggio

Ora introduciamo i costrutti con cui generalizziamo il modello di energia classico. Essi si riferiscono alle relazioni tra cellule complesse distinte, ossia alle connessioni ricorrenti attraverso cui le cellule complesse si influenzano vicendevolmente. Da un punto di vista formale rappresentiamo tali connessioni tramite una funzione $K_\sigma(x_0, y_0, \zeta_0, x_1, y_1, \zeta_1)$ che denota l'associazione (il peso sinaptico) tra la cellula complessa $n(x_1, y_1, \zeta_1, \sigma)$ e la cellula complessa $n(x_0, y_0, \zeta_0, \sigma)$, cioè quanto l'attività $E(x_1, y_1, \zeta_1, \sigma)$ influenza l'attività $E(x_0, y_0, \zeta_0, \sigma)$. Il modello prevede la più semplice delle interazioni – quella lineare –, quindi il pattern K_σ di pesi sinaptici è utilizzato per far evolvere lo stato delle cellule complesse nel seguente modo:

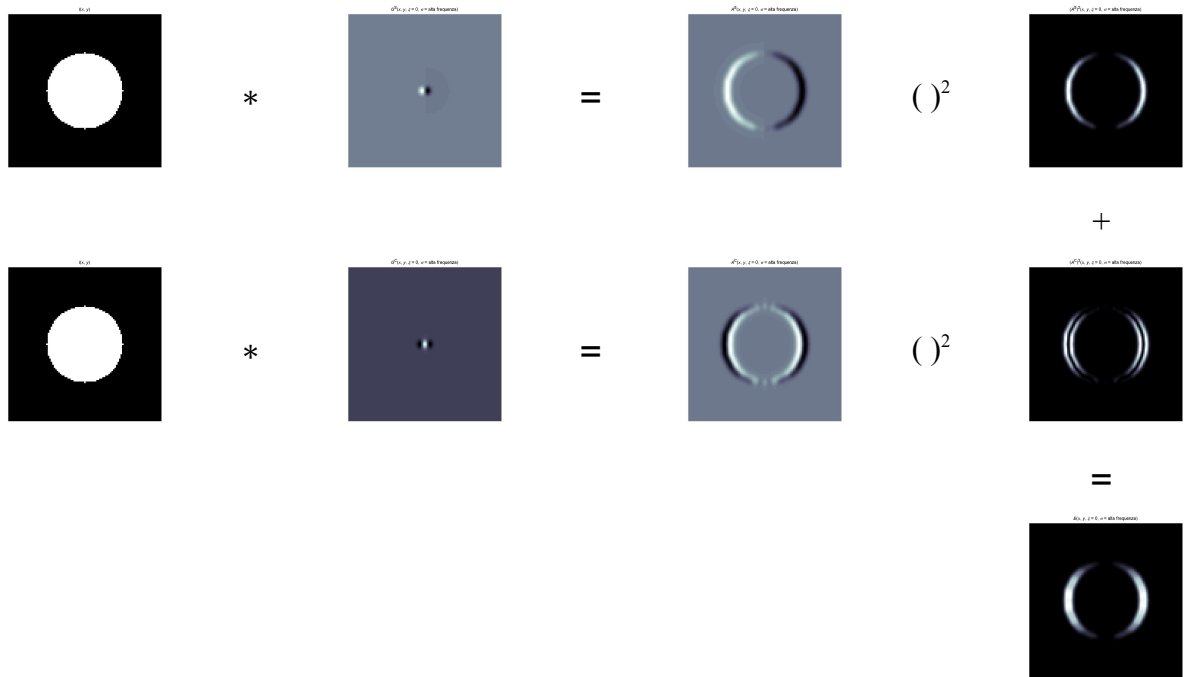
- (i) si pone $E_{j=0}$ = stato delle cellule complesse a seguito del flusso feedforward dalle cellule semplici;
- (ii) $E_{j+1}(x_0, y_0, \zeta_0, \sigma) = \iiint dx_1 dy_1 d\zeta_1 \{K_\sigma(x_0, y_0, \zeta_0, x_1, y_1, \zeta_1) E(x_1, y_1, \zeta_1, \sigma)\}$

Il campo associativo K_σ è costruito distinguendo due tipi di connessioni: quelle interne ad uno stesso modulo corticale e quelle tra moduli corticali distinti.

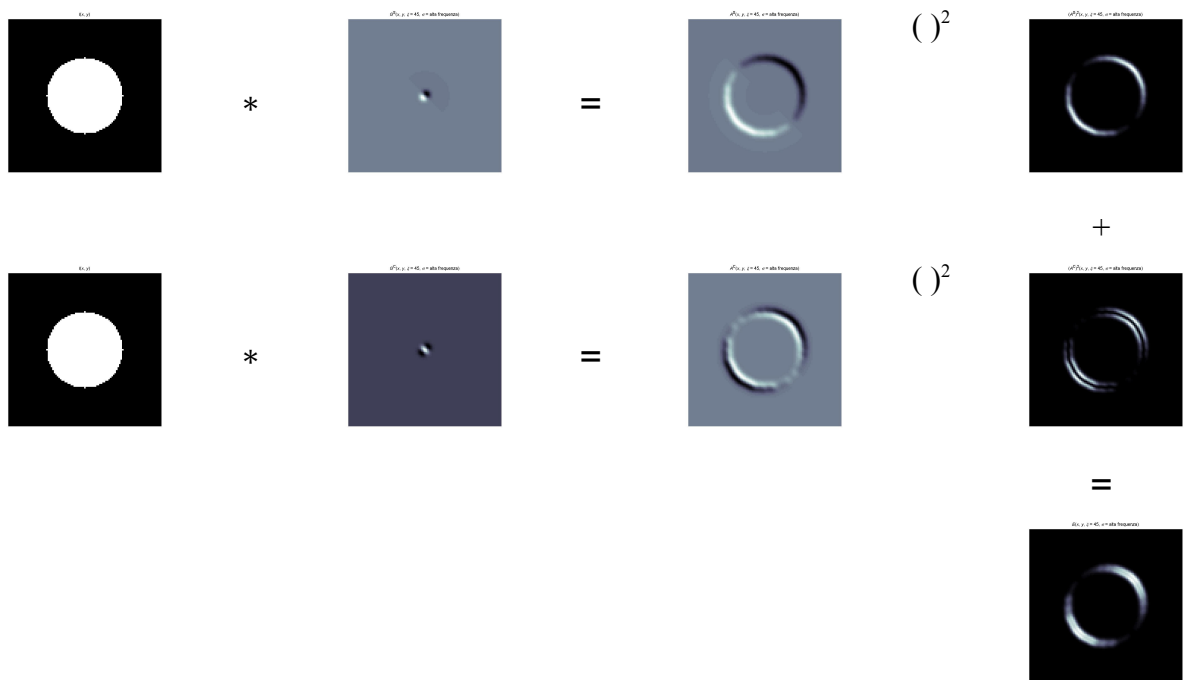
Figura 23. Flusso feedforward ad alta frequenza spaziale (piccolo valore di scala, dettagli fini)



23c. Parametri: $\zeta = 0^\circ$, $\sigma =$ alta frequenza spaziale (dettagli dell'immagine)



23d. Parametri: $\zeta = +45^\circ$, $\sigma =$ alta frequenza spaziale (dettagli dell'immagine)



23e. Parametri: $\zeta = +90^\circ$, $\sigma =$ alta frequenza spaziale (dettagli dell'immagine)

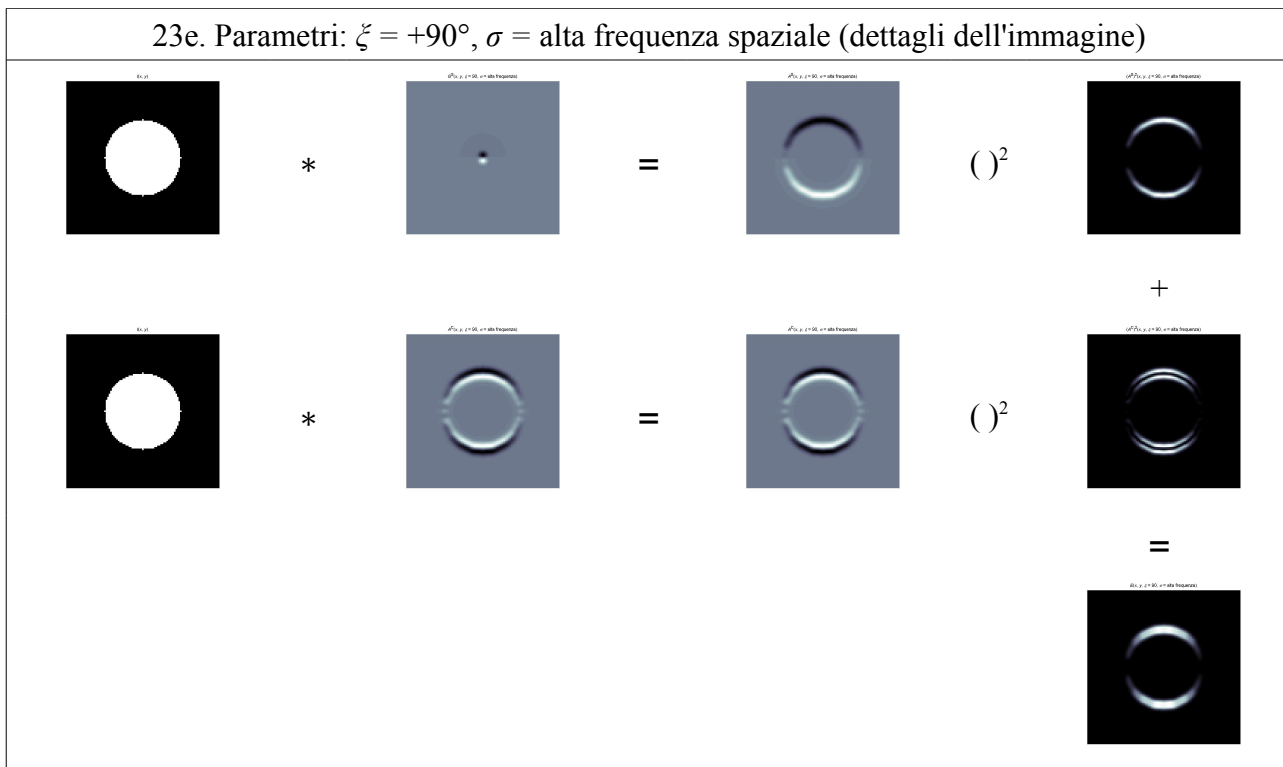
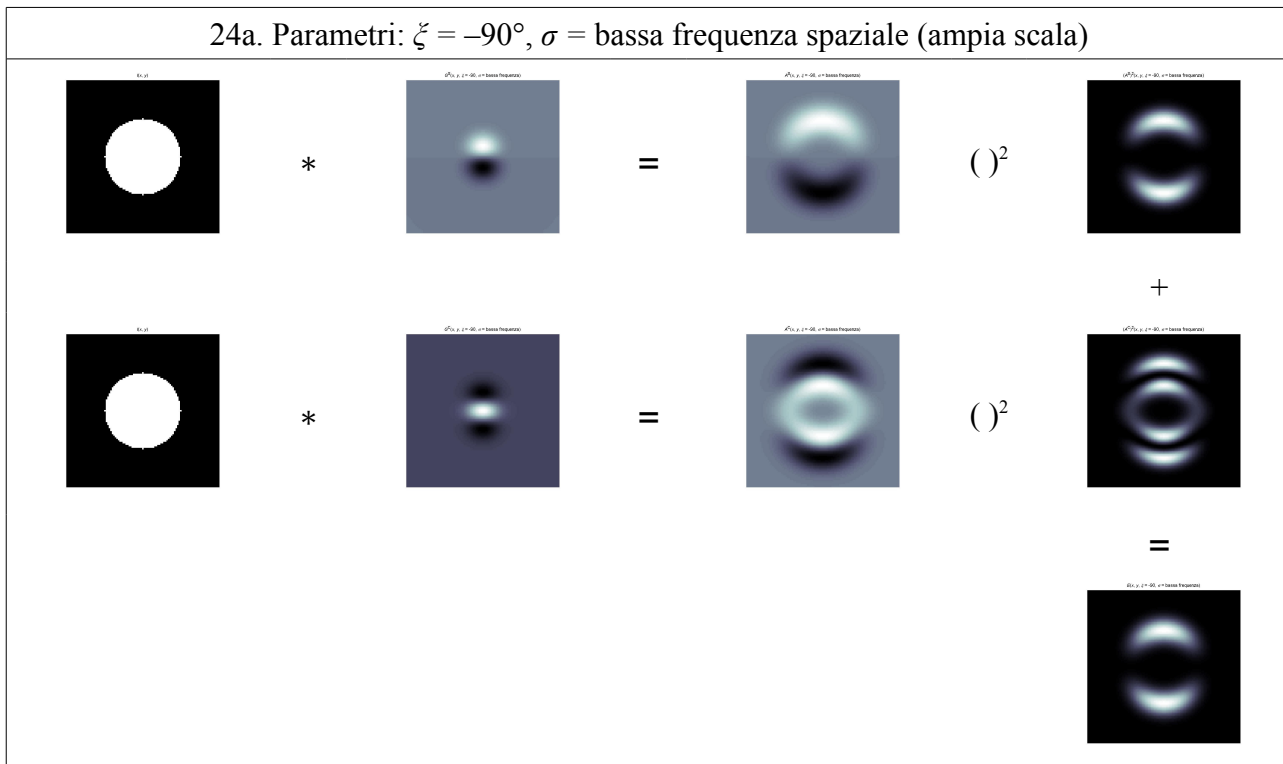
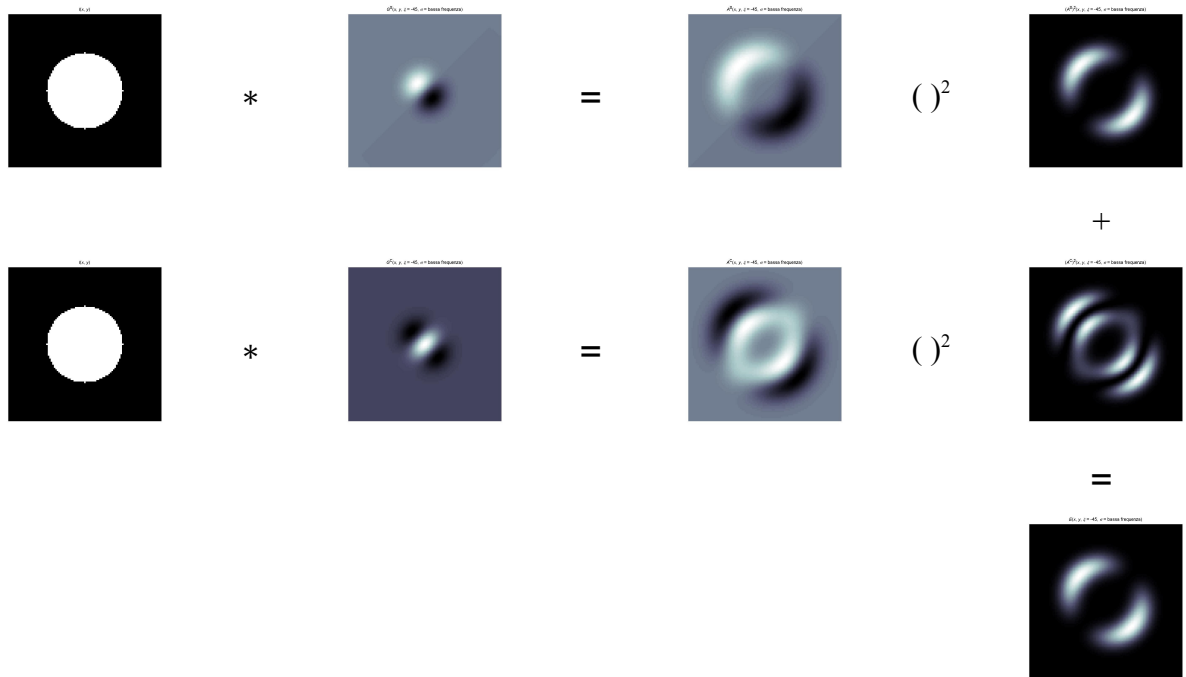


Figura 24. Flusso feedforward a bassa frequenza spaziale (grande valore di scala)

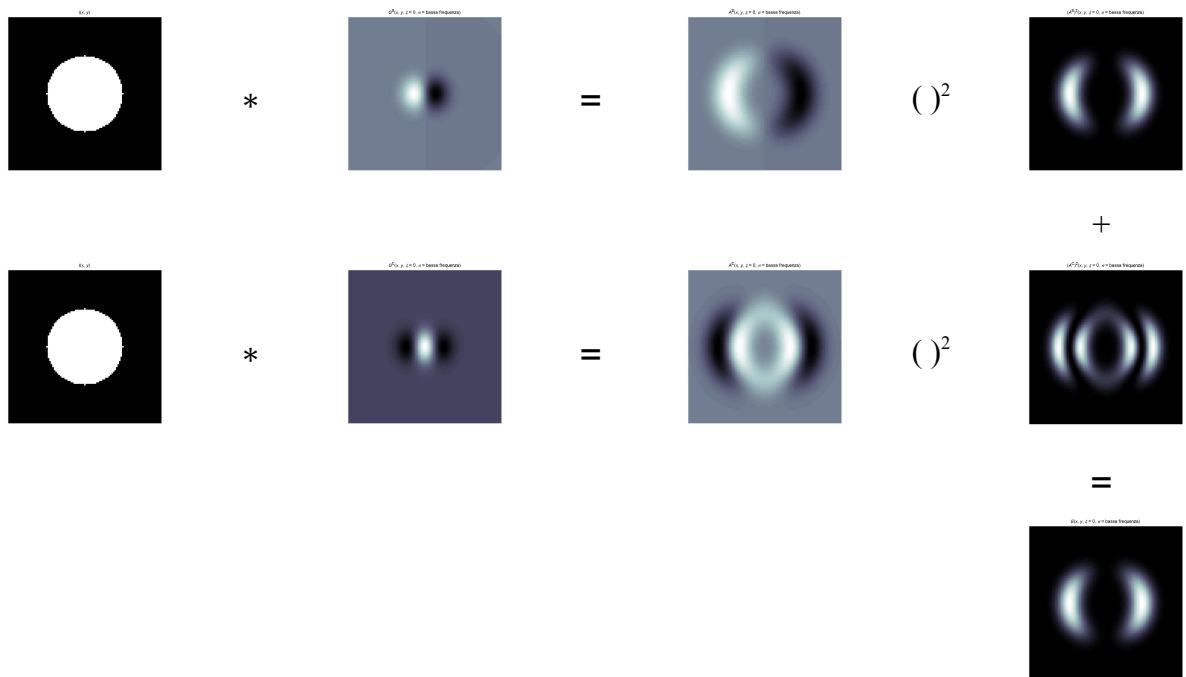
24a. Parametri: $\zeta = -90^\circ$, $\sigma =$ bassa frequenza spaziale (ampia scala)



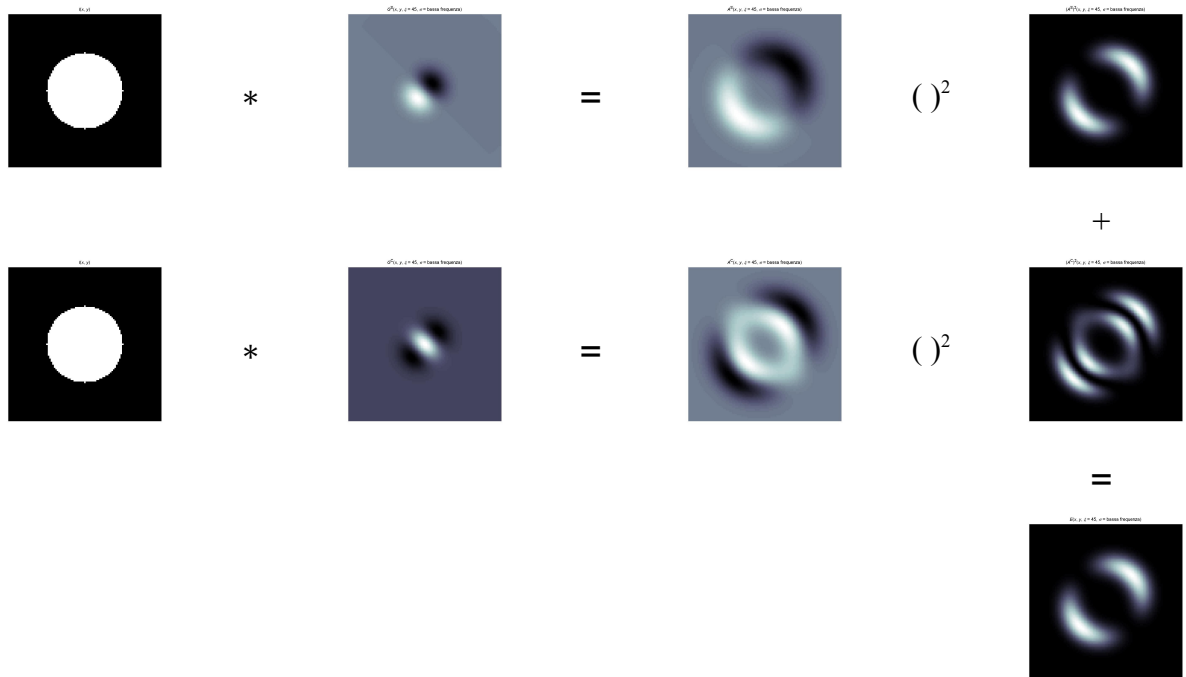
24b. Parametri: $\xi = -45^\circ$, $\sigma =$ bassa frequenza spaziale (ampia scala)



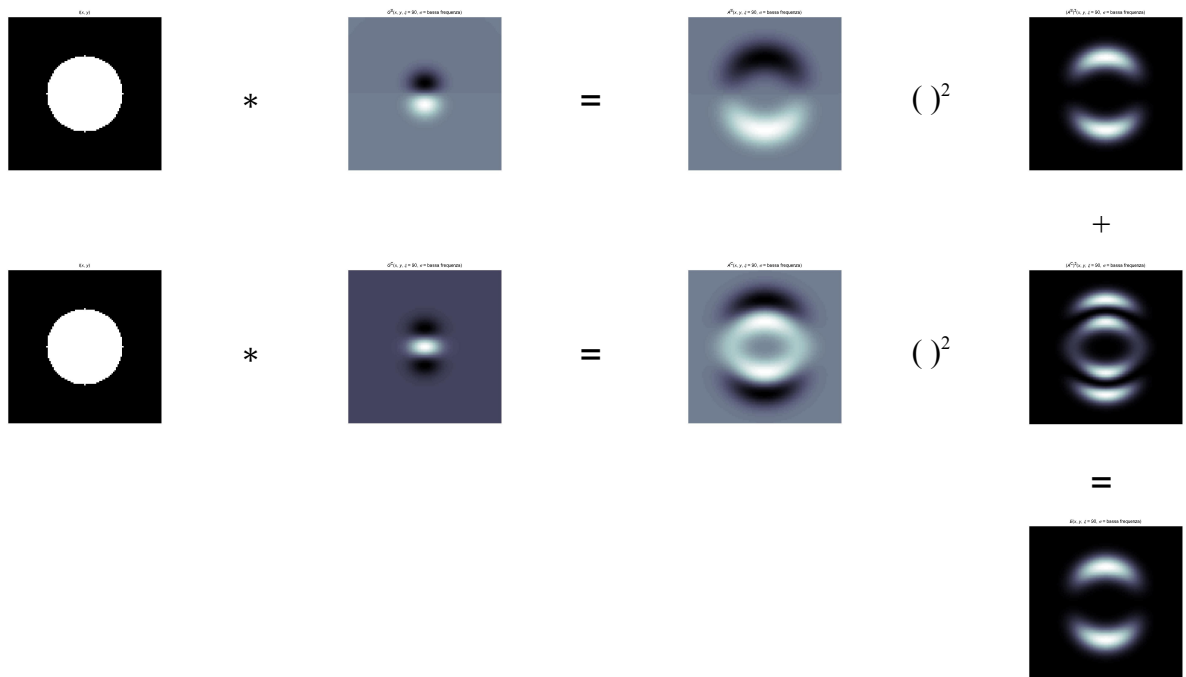
24c. Parametri: $\xi = 0^\circ$, $\sigma =$ bassa frequenza spaziale (ampia scala)



24d. Parametri: $\zeta = 45^\circ$, $\sigma =$ bassa frequenza spaziale (ampia scala)



24e. Parametri: $\zeta = +90^\circ$, $\sigma =$ bassa frequenza spaziale (ampia scala)



Le connessioni tra unità complesse del medesimo modulo corticale devono soddisfare i seguenti requisiti:

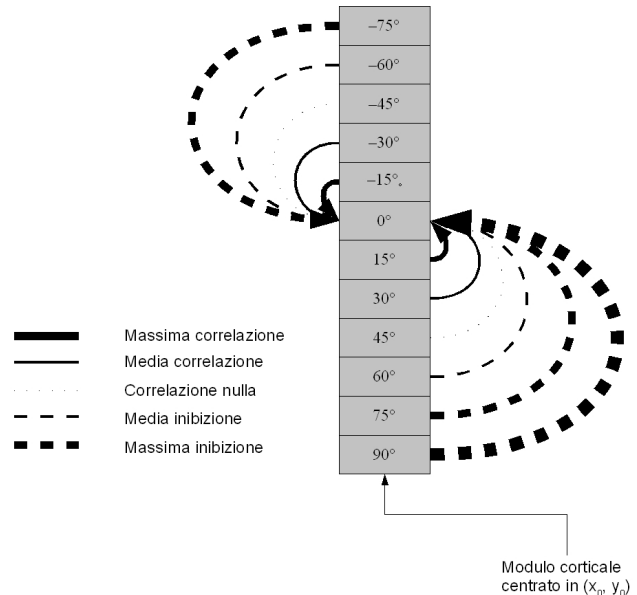
- (i) devono trattare le unità come fossero organizzate su una circonferenza;
- (ii) devono essere π -periodiche;
- (iii) devono dipendere dalla *distanza angolare* in modo che sia *minima* l'associazione tra unità che hanno direzioni preferite che differiscono per $\pi/2$ radianti (90°).

Le funzioni $\cos(\Delta\xi)^2$ e $\cos(2\Delta\xi)$ soddisfano tutti e tre i criteri. Preferiamo la seconda perché implementa l'equivalente, sulla circonferenza, della ubiquitaria *inibizione laterale* presente tra le cellule visive. Abbiamo, quindi, che le connessioni tra cellule complesse, appartenenti al medesimo modulo corticale centrato nel punto (x, y) , sono rappresentate dalla funzione

$$D(\xi_0, \xi_1) := \cos(2\Delta\xi) = \cos(2(\xi_1 - \xi_0)) \quad (8)$$

In figura 25 è disegnato un campo di interazione tra unità dello stesso modulo corticale (Precisamente sono rappresentate le proiezioni – di altre cellule complesse dello stesso modulo corticale – sulla unità complessa che risponde preferenzialmente a 0° gradi).

Figura 25. Campo di interazione tra cellule complesse appartenenti al medesimo modulo corticale



Le connessioni tra moduli corticali distinti mettono in relazione unità sintonizzate su direzioni *compatibili* ovvero associabili secondo le regole del campo associativo. In figura 26 è fornita una versione schematica del pattern di connessioni.

Prese comunque due unità complesse appartenenti a moduli corticali distinti, non è detto esista una connessione *diretta* che le lega, mentre è vero che per ogni unità di un modulo c'è esattamente una unità dell'altro ad essa direttamente collegata (ammesso che i moduli siano *sufficientemente vicini*). Prendere due moduli corticali distinti equivale, formalmente, a fissare due punti (x_0, y_0) e (x_1, y_1) . Prendiamo l'unità complessa nel modulo centrato in (x_0, y_0) che abbia direzione preferita ζ_0 : ci chiediamo quale sia l'unità del modulo centrato in (x_1, y_1) che, compatibilmente al campo associativo, proietta direttamente ad essa (Fig. 27). Per facilitare la spiegazione introduciamo una variazione nella notazione in analogia a (3): al posto di identificare una unità complessa all'interno di un modulo corticale tramite la sua orientazione preferita ζ , ci riferiremo al corrispondente versore preferito \mathbf{u} , dove $\mathbf{u} = (\cos(\zeta), \sin(\zeta))$.

Abbiamo postulato che le connessioni del campo associativo facilitino il completamento di contorni (linee, bordi) a curvatura costante, cioè segmenti (rette) ed archi di circonferenza (circonferenze). Geometricamente, quindi, dobbiamo connettere i punti (x_0, y_0) e (x_1, y_1) con una linea tale che: (i) sia a curvatura costante e (ii) sia perpendicolare, nel punto (x_0, y_0) , al vettore $\mathbf{u}_0 = (\cos(\xi_0), \sin(\xi_0))$. L'unità cercata nel modulo centrato in (x_1, y_1) sarà quella che, in tale punto, ha direzione preferita anch'essa perpendicolare alla linea di campo che congiunge (x_0, y_0) e (x_1, y_1) : il suo vettore preferito sarà denotato con \mathbf{u}_0' . Conveniamo di porre

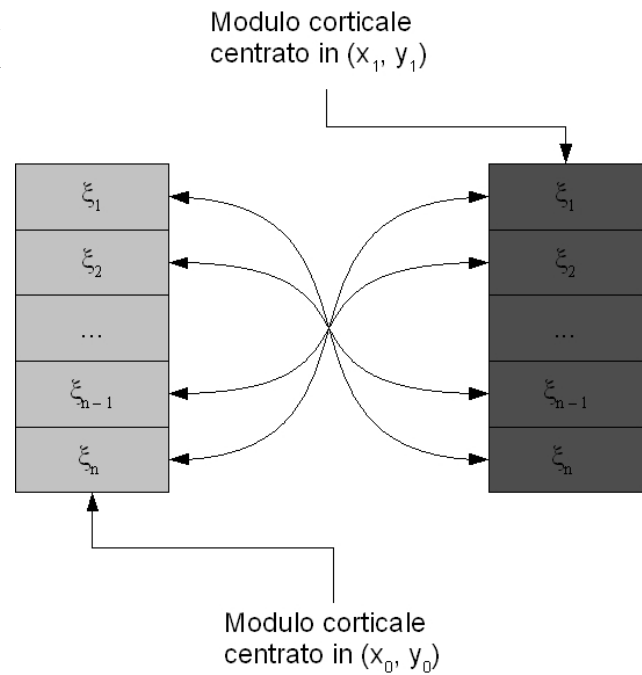
$$\mathbf{s} := (x_1, y_1) - (x_0, y_0) = (\Delta x, \Delta y) \quad (9)$$

Formalmente \mathbf{u}_0' è determinato in funzione di \mathbf{s} e \mathbf{u}_0 . Ragioni di simmetria [Fig. 27] ci facilitano i calcoli ⁽¹⁸⁾ che ci portano ad affermare che

$$\mathbf{u}_0' = \mathbf{M} \mathbf{u}_0 := \mathbf{u}_0 - 2 \frac{\mathbf{u}_0 \cdot \mathbf{s}}{\mathbf{s} \cdot \mathbf{s}} \mathbf{s} \quad (10)$$

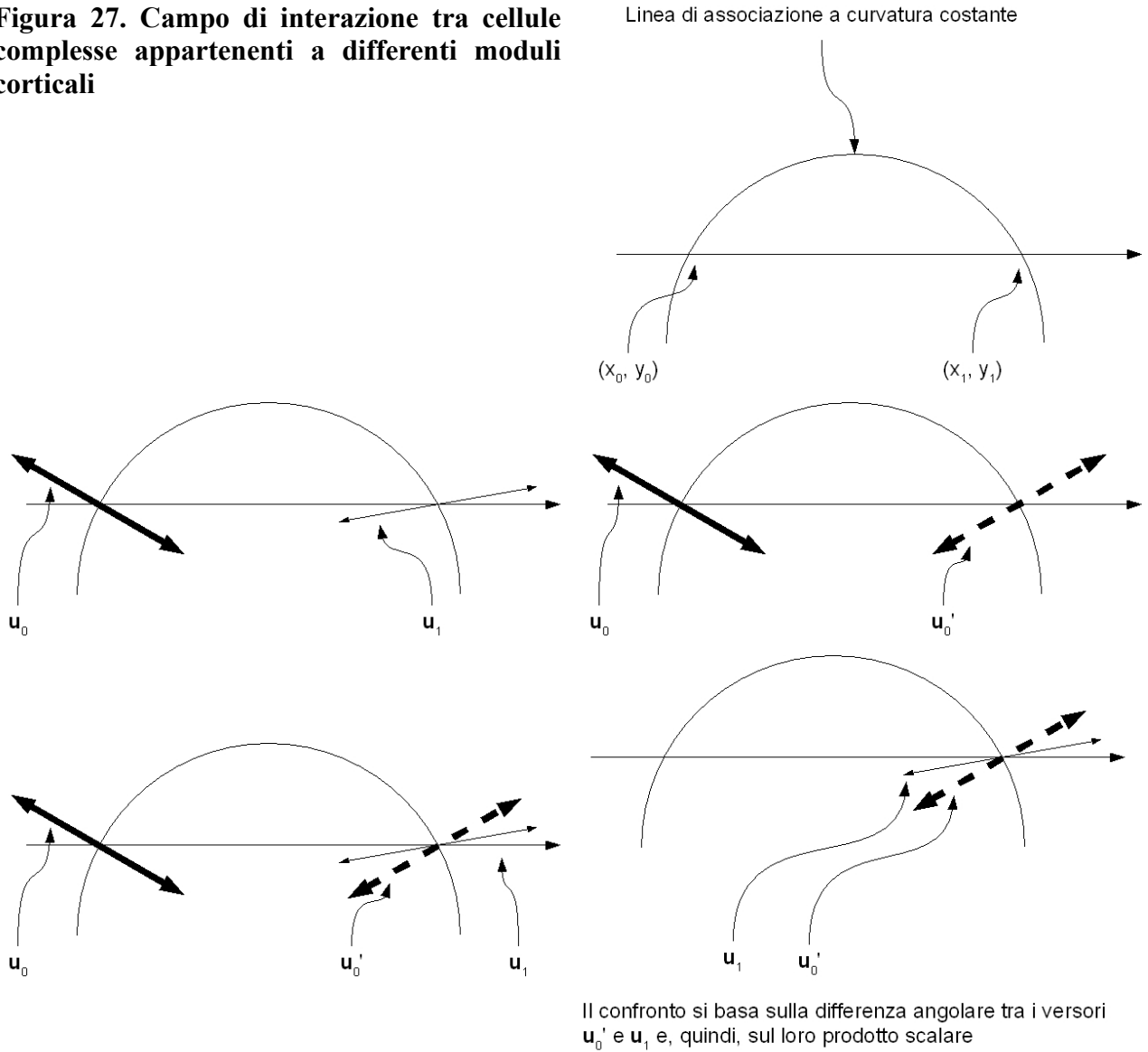
La matrice \mathbf{M} in (10) denota la regola con cui da \mathbf{u}_0 si passa a \mathbf{u}_0' , cioè il modo in cui determinare il candidato ideale del modulo corticale centrato in (x_1, y_1) ad essere associabile alla cellula complessa appartenente al modulo centrato in (x_0, y_0) e con vettore preferito \mathbf{u}_0 .

Figura 26. Rappresentazione schematica dell'implementazione formale del campo di interazione tra cellule complesse



¹⁸ Cambiamo sistema di riferimento ed utilizziamo la retta che congiunge i due (x_0, y_0) e (x_1, y_1) come asse delle ascisse e il vettore ad esso normale come asse delle ordinate; poniamo, inoltre, l'origine in (x_0, y_0) . Sia $\mathbf{n} = (-\Delta y, \Delta x) \perp \mathbf{s}$; i corrispondenti versori $\mathbf{u}_N = \frac{\mathbf{n}}{|\mathbf{n}|}$ e $\mathbf{u}_T = \frac{\mathbf{s}}{|\mathbf{s}|}$ definiscono la nuova base cercata. In tale base si ha che $\mathbf{u}_0 = (\mathbf{u}_0 \cdot \mathbf{u}_N) \mathbf{u}_N + (\mathbf{u}_0 \cdot \mathbf{u}_T) \mathbf{u}_T$ e si deve avere che $\mathbf{u}_0' = (\mathbf{u}_0 \cdot \mathbf{u}_N) \mathbf{u}_N - (\mathbf{u}_0 \cdot \mathbf{u}_T) \mathbf{u}_T$. Dalla prima equazione si ottiene $(\mathbf{u}_0 \cdot \mathbf{u}_N) \mathbf{u}_N = \mathbf{u}_0 - (\mathbf{u}_0 \cdot \mathbf{u}_T) \mathbf{u}_T$, e quindi, sostituendo nella seconda che $\mathbf{u}_0' = \mathbf{u}_0 - 2 (\mathbf{u}_0 \cdot \mathbf{u}_T) \mathbf{u}_T$.

Figura 27. Campo di interazione tra cellule complesse appartenenti a differenti moduli corticali



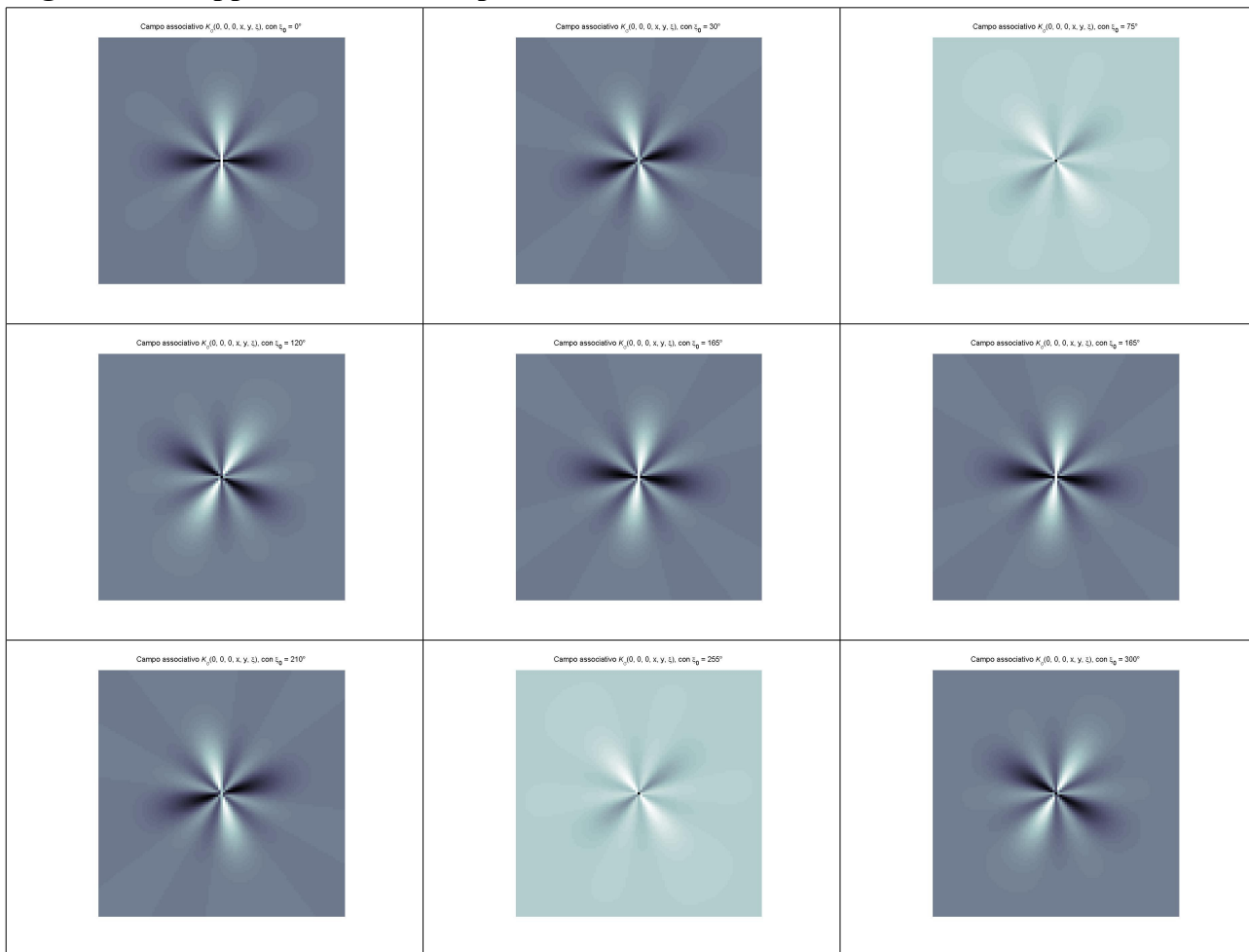
Denotiamo con $C(x_0, y_0, \xi_0, x_1, y_1, \xi_0', \sigma)$ il peso dell'associazione tra l'unità $n(x_0, y_0, \xi_0, \sigma)$ – appartenente al modulo centrato in (x_0, y_0) , con direzione preferita ξ_0 al valore di scala σ – e l'unità appartenente al modulo centrato in (x_1, y_1) ad essa massimamente compatibile, secondo le regole del campo associativo: l'unità $n(x_1, y_1, \xi_0', \sigma)$, la cui scala-frequenza d'azione è sempre σ ed ha direzione preferita ξ_0' . La funzione C tiene conto della *buona continuazione* perché congiunge unità con direzioni preferite massimamente compatibili (ξ_0 e ξ_0') secondo le linee del campo. Nella determinazione della forza di associazione secondo il campo associativo di Field hanno un peso anche la *distanza* della curva che separa due unità (massimo circa 5 volte l'ampiezza dei gabor patch) e la sua *curvatura* (massimo circa 60°); un metodo standard che tiene conto di entrambi i fattori – e semplifica notevolmente i calcoli – è quello di usare una funzione *a doppia clava* [Fig. 28], la cui formula è la seguente:

$$C(x_0, y_0, \xi_0, x_1, y_1, \xi_0', \sigma) = e^{-\frac{s \cdot s}{4\sigma^2}} \left(1 - 2 \frac{(\mathbf{u}_0 \cdot \mathbf{s})^2}{\mathbf{s} \cdot \mathbf{s}} \right) \quad (11)$$

Notiamo che il denominatore presente nell'argomento della componente esponenziale è

posto pari a $(2\sigma)^2$, dove σ è il fattore di scala a livello delle cellule semplici. Esso, come già accennato, dà una misura della ampiezza dei filtri implementati dalle cellule semplici. Intuitivamente σ può essere concepito come l'*unità di misura* del sistema, cioè il minimo discriminabile. È come se il sistema, per determinare le distanze, avesse a disposizione un regolo rigido sul quale le tacche distano σ le une dalle altre. Se σ è un millimetro, per esempio, non si possono misurare i decimi di millimetro; mentre si possono misurare lunghezze pari a un certo numero di millimetri. Se il campo associativo connette le unità perché si correlino informazioni provenienti da punti *distanti* del campo percettivo, essi devono essere *abbastanza distanti*, almeno *n volte* la distanza minima discriminabile: qui abbiamo scelto l'*n* minimo: due.

Figura 28. "Doppia clava" del campo associativo



Con (10), dati due moduli corticali [centrati rispettivamente in] (x_0, y_0) e (x_1, y_1) ed una unità, nel primo, che ha orientazione preferita ξ_0 , cui corrisponde il versore $\mathbf{u}_0 = (\cos(\xi_0), \sin(\xi_0))$, abbiamo individuato l'unità del secondo modulo che risponde in maniera preferenziale ad un'orientazione massimamente compatibile con il principio di buona continuazione, essendo tale orientazione ottimale pari a ξ_0' , cui corrisponde il versore $\mathbf{u}_0' = (\cos(\xi_0'), \sin(\xi_0'))$. L'apporto dell'attività della cellula complessa $E(x_1, y_1, \xi_0', \sigma)$ nell'influenzare l'attività $E(x_0, y_0, \xi_0, \sigma)$ è *pesato* dalla funzione $C(x_0, y_0, \xi_0, x_1, y_1, \xi_0', \sigma)$. All'interno di un medesimo modulo corticale, invece, vale la (8) che, sostituendo la notazione nei termini di \mathbf{u}_k a quella con ξ_k , possiamo riscrivere come

$$D(\mathbf{u}_0, \mathbf{u}_1) := D(\xi_0, \xi_1) = \cos(2 \Delta \xi) = \cos(2(\xi_1 - \xi_0)) = 2(\mathbf{u}_0 \cdot \mathbf{u}_1)^2 - 1 \quad , \quad (12)$$

nella quale sono stati usati la formula trigonometrica $\cos(2\alpha) = 2\cos^2(\alpha) - 1$ ed il fatto che il prodotto scalare di due vettori è pari al coseno dell'angolo che formano.

Ora possiamo mettere assieme i pezzi e stimare come agisca la funzione che rappresenta il peso dell'associazione $K_\sigma(x_0, y_0, \xi_0, x_1, y_1, \xi_1)$ tra due cellule complesse qualsiasi $E(x_1, y_1, \xi_1, \sigma)$ e $E(x_0, y_0, \xi_0, \sigma)$. L'attività di $n(x_1, y_1, \xi_1, \sigma)$ influenza l'attività di $n(x_0, y_0, \xi_0, \sigma)$ tramite due altre unità. Una è l'unità appartenente al medesimo modulo corticale di $n(x_1, y_1, \xi_1, \sigma)$ e massimamente associata a $n(x_0, y_0, \xi_0, \sigma)$ tramite le linee del campo, cioè $n(x_1, y_1, \xi_0', \sigma)$. L'altra è l'unità appartenente al modulo centrato in (x_0, y_0) e massimamente associata a $n(x_1, y_1, \xi_1, \sigma)$, cioè $n(x_0, y_0, \xi_1', \sigma)$. L'influenza di $n(x_1, y_1, \xi_1, \sigma)$ su $n(x_0, y_0, \xi_0, \sigma)$, mediata da $n(x_1, y_1, \xi_0', \sigma)$, è pari a

$$C(x_0, y_0, \xi_0, x_1, y_1, \xi_0', \sigma) D(\xi_1, \xi_0') \quad (13)$$

mentre quella mediata da $n(x_0, y_0, \xi_1', \sigma)$ è

$$C(x_1, y_1, \xi_1, x_0, y_0, \xi_1', \sigma) D(\xi_0, \xi_1') \quad (14)$$

Definiamo il campo associativo K_σ come la media tra (13) e (14), cioè, usando anche (10), (11) e (12),

$$K_\sigma(x_0, y_0, \xi_0, x_1, y_1, \xi_1) = \frac{1}{2} e^{\frac{s \cdot s}{4\sigma^2}} \left(2(\mathbf{u}_0 \mathbf{M} \mathbf{u}'_1)^2 - 1 \right) (\mathbf{u}_0 \mathbf{M} \mathbf{u}'_0 + \mathbf{u}_1 \mathbf{M} \mathbf{u}'_1) \quad (15)$$

L'azione del campo K_σ modifica lo stato delle cellule complesse facendole interagire. Conveniamo di indicare con $E_j(x, y, \xi, \sigma)$ l'attività dell'unità $n(x, y, \xi, \sigma)$ al j -esimo passo di computazione, cioè dopo j volte che le cellule complesse hanno interagito tramite le connessioni laterali a lungo e corto raggio implementate da K_σ . Lo stato iniziale è l'attivazione determinata dal solo flusso feedforward, si ha cioè che $E_0 = |A|^2$. In tali condizioni l'evoluzione dello stato delle cellule complesse è

$$E_{j+1}(x_0, y_0, \xi_0, \sigma) = \iiint dx_1 dy_1 d\xi_1 E_j(x_1, y_1, \xi_1, \sigma) K_\sigma(x_0, y_0, \xi_0, x_1, y_1, \xi_1) \quad (16)$$

Si noti che nel modello qui presentato non si prende posizione riguardo alla relazione tra il ciclo di computazione, che separa lo stato j -esimo dal $j+1$ -esimo, ed il tempo. In altri termini non abbiamo posto vincoli – che sicuramente esistono – sulla durata di un ciclo di interazioni implementato da K_σ . Ciò equivale a supporre la durata praticamente infinitesima; oltretutto si semplifica assumendo che l'azione di K_σ sia istantanea.

Finora abbiamo trattato del livello di *coding*. Ora ci troviamo nella situazione di un neurofisiologo che voglia dare senso ad un pattern di attivazione neurale. In altri termini, dato un particolare stato E di attivazione delle cellule complesse, quale è l'informazione sull'immagine di input I che possiamo ottenere? Per determinare la *direzione* rappresentata dal pattern di attivazione delle cellule complesse appartenenti ad un'iper-colonna, come già detto in precedenza, le consideriamo organizzate su una circonferenza dotata di massa. Il valore di attivazione di ciascuna, dopo essere stato opportunamente scalato, rappresenta la densità lineare di massa della circonferenza. La direzione *complessiva* è, quindi, identificata dall'asse principale di inerzia della configurazione (si veda la figura 15). In altri metodi si trovano ragionamenti simili [52, 62, 93].

Si rivela opportuna un'osservazione riferita alla necessità di *scalare opportunamente* lo stato di attivazione delle cellule complesse. Il campo associativo formale, definito in eq. (15) e rappresentato in figura 28, varia da -1 a 1 ; siccome agisce come i pesi di una somma (è un

operatore lineare), la sua applicazione fa sì che le unità complesse possano assumere anche valori negativi. In genere, nelle reti neurali, si preferisce evitare di usare numeri negativi per rappresentare lo stato di una unità cellulare. Nel nostro caso il problema non si pone per quanto detto in [§3.2, secondo capoverso]. Avremmo potuto introdurre qualche forma di non-linearità con cui trasformare il range delle risposte possibili delle cellule semplici dall'intero asse reale a, per esempio, l'intervallo $[0, 1]$ ⁽¹⁹⁾; questo, tuttavia, avrebbe solo complicato i calcoli. Per interpretare il pattern di attivazione di un'iper-colonna in termini di "densità lineare di massa", è sufficiente portare il suo range (l'intervallo dei valori acquisiti) sull'intervallo $[0, 1]$ ed è possibile farlo *dopo* la computazione della risposta del sistema: ciò che conta per il calcolo dell'asse di inerzia rimane invariato ⁽²⁰⁾.

Il fatto che il pattern di attivazione delle cellule complesse di un modulo corticale sia interpretabile nei termini di una *direzione stimata* del gradiente, non implica che in quel punto *vi sia un gradiente*, cioè che quel punto sia in prossimità di un bordo. Infatti il gradiente è un *vettore* e, come tale, caratterizzato da due informazioni: direzione e modulo. È il modulo che ci dice se nel punto vi sia oppure no un gradiente, cioè se il punto sia oppure non sia su un bordo. Se ricordiamo quale è lo stato iniziale delle cellule complesse, cioè prima (eq. 7) che sia simulata l'azione delle connessioni laterali a corto e lungo raggio, si può ragionevolmente stimare il modulo del gradiente tramite la radice quadrata della somma dei quadrati dei valori di attivazione. Come già accennato, Medioni *et al.* [79] risolvono questo problema usando una rappresentazione locale di bordi in termini di *tensori* al posto che di *vettore*. Concettualmente la soluzione è simile: servono, nello spazio, due numeri – direzione e modulo – per rappresentare un contorno: il primo denota, appunto, la direzione stimata dell'*eventuale* contorno, mentre la seconda la verisimiglianza che vi sia un contorno. Medioni parla di *salienza* e direzione, ma il concetto è lo stesso. È interessante notare che un tensore simmetrico – l'ente matematico usato da Medioni – rappresenta, nel piano, una conica, precisamente un'ellisse. Possiamo rappresentare un tensore siffatto con due segmenti perpendicolari, interpretabili come gli assi principali di un'ellisse, appunto. Quando i due segmenti sono di uguale lunghezza, l'ellisse degenera in un cerchio. In tale condizione il punto su cui il tensore è applicato *non* fa sicuramente parte di un bordo. Quando, invece, la differenza tra le lunghezze dei due segmenti aumenta, il punto a cui è associato il tensore fa parte di un bordo, la cui tangente è proprio nella direzione del segmento più lungo. La differenza tra le due lunghezze è proprio la salienza del contorno, ovvero la probabilità che il punto spaziale in questione stia realmente su un contorno. Nel caso limite – quello di massima sicurezza che il punto stia su un bordo – il tensore diviene degenero, nel senso che uno dei due segmenti si contrae sino ad annullarsi e l'altro – la direzione della tangente – acquisisce lunghezza massima. La rappresentazione nei termini di "massa filiforme chiusa a circonferenza" – di anello, per intenderci – che diamo dell'attività delle cellule complesse di una iper-colonna possiede evidenti analogie con il metodo di codifica delle informazioni usato da Medioni ⁽²¹⁾. Infatti i due assi principali di inerzia associabili a questo anello sono interpretabili come i due segmenti che costituiscono un tensore simmetrico. L'equivalente della salienza sarebbe la stima del modulo del gradiente decodificato con il metodo sopra riportato.

§5.2.

L'estensione del modello allo spazio-tempo tridimensionale

Lo scopo è di proporre un modello di elaborazione dei segnali di movimento, quindi, relativo allo spazio-tempo (fisico, neurale e fenomenico) tridimensionale. Il paragrafo precedente ha ragione d'essere solo perché permette di associare alla trattazione formale un dominio più semplice da rappresentare con immagini. Il nostro percorso argomentativo qui si conclude con il seguente

19 Di solito si usa a tal fine la funzione sigmoide e si pone $\text{attivazione} = 1/(1+\exp(-\text{input}))$ [14, 19, 20].

20 Una discussione più articolata sull'uso di numeri negativi per rappresentare lo stato di un neurone artificiale richiederebbe nozioni ed assunti di topologia che esulano dallo scopo principale di questo scritto.

21 Ringrazio la Professoressa Citti per avermi suggerito questa similarità.

schema:

- 1 Abbiamo proposto un modello bidimensionale di elaborazione dei segnali visivi che
 - 1.1 generalizza il modello di energia di Adelson e Bergen [84],
 - 1.2 possiede buona plausibilità neurofisiologica [11, 58-61, 66, 74-78, 81] ed
 - 1.3 è computazionalmente maneggevole.
- 2 Abbiamo dichiarato
 - 2.1 di voler far uso dell'euristica di generalizzazione a dimensioni superiori (§1) e
 - 2.2 di valutarne gli esiti sia secondo la plausibilità neurofisiologica
 - 2.3 sia secondo l'esito di simulazioni.
- 3 Quindi assumiamo che quanto detto riguardo all'elaborazione dei segnali statici bidimensionali valga anche ne [o, meglio, sia caso particolare de] l'elaborazione di segnali spazio-temporali tridimensionali. In questa prospettiva:
 - 3.1 aggiungiamo le opportune modifiche ed estensioni affinché la *forma* abbia senso anche nel nuovo spazio di *contenuti*,
 - 3.2 forniamo una interpretazione neurofisiologicamente plausibile dei nuovi costrutti, o evidenziamo eventuali problemi, e, infine,
 - 3.3 produciamo alcune simulazioni

Se fissiamo il valore di scala, nel caso di elaborazione degli stimoli bidimensionali abbiamo a che fare con tre variabili: le dimensioni x e y dello stimolo e la dimensione ζ che indica l'orientazione preferita delle cellule. Per trattare della elaborazione del movimento dobbiamo usare cinque variabili, cioè aggiungerne due alle precedenti. Una è il tempo t e l'altra è l'orientazione preferita rispetto all'asse temporale, che denotiamo con η .

Il significato neurofisiologico e la necessità fenomenica di η divengono più chiari se osserviamo le figure 29 e 30, nelle quali sono rappresentati i campi recettivi dispari spazio-temporali di cellule semplici. La terna (x, y, t) rappresenta il centro del campo recettivo e la coppia (ξ, η) gli angoli che definiscono in coordinate polari il versore del gradiente. I campi recettivi spazio-temporali delle cellule semplici sono ottenuti esattamente, come in §5.1, tramite le equazioni (4) e (5) – per campi recettivi, rispettivamente, dispari e pari – solo che si sostituisca la gaussiana bidimensionale (2) con la gaussiana tridimensionale

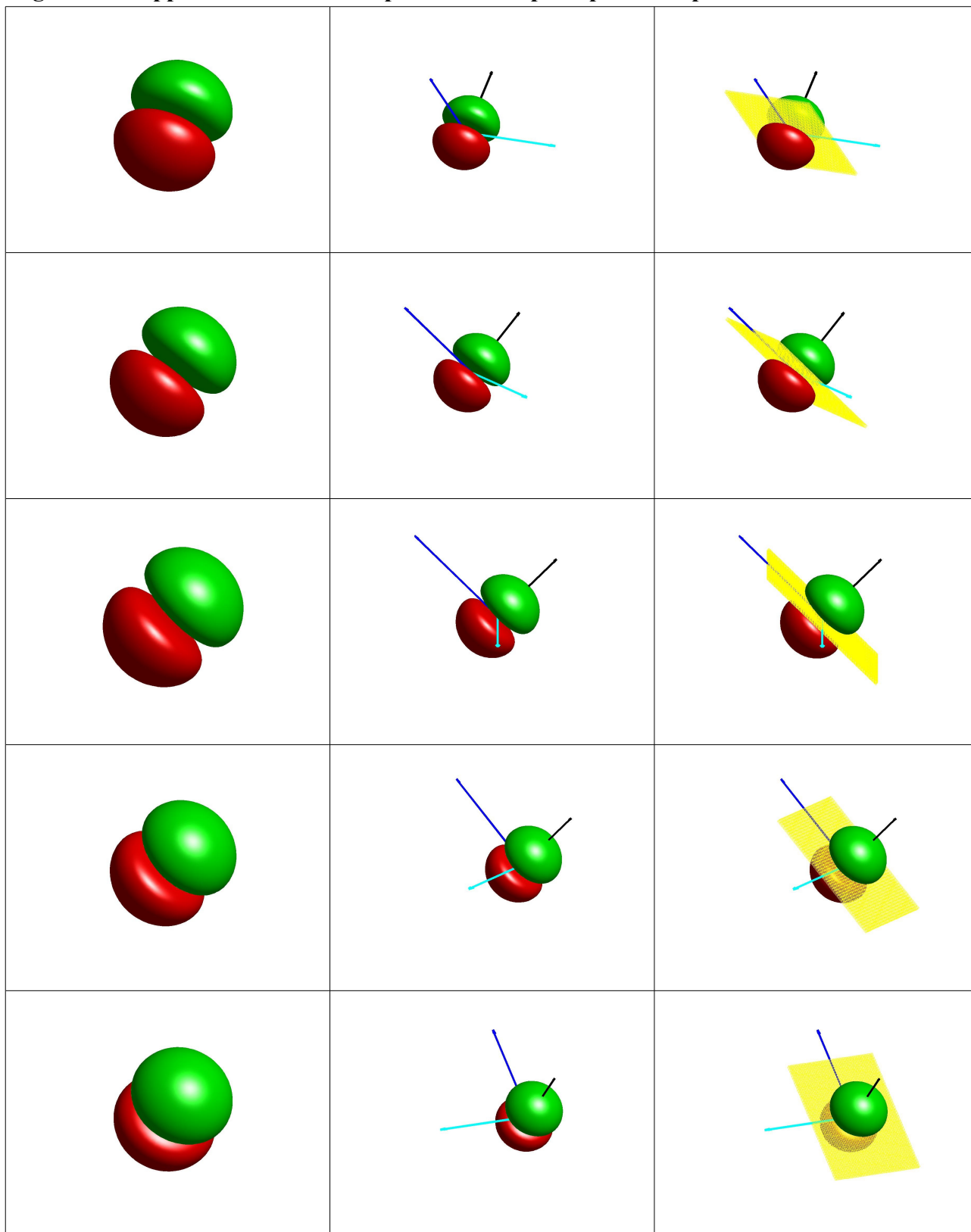
$$g_{\sigma} := e^{-\frac{x^2 + y^2 + t^2}{\sigma^2}} \quad (17)$$

ed il versore \mathbf{u} di (3) con

$$\mathbf{u} := (\cos(\xi)\sin(\eta), \sin(\xi)\sin(\eta), \cos(\eta)) \quad , \quad (18)$$

In figura 29 è raffigurato il campo recettivo spazio-temporale *dispari* di una cellula semplice. Nelle diverse *righe* si ha un cambiamento di punto di vista per enfatizzare la tridimensionalità. Nella terza riga abbiamo la rappresentazione in cui l'asse verticale è il tempo t e quello orizzontale e la dimensione x dello spazio. Nelle *colonne* 2 e 3 sono aggiunte informazioni alle rispettive immagini in colonna 1. In colonna 2 si aggiungono i versori di riferimento per descrivere il filtro. Come evidenziato in figura 29, questi tre versori indicano le direzioni in cui avviene la comunicazione tra unità distinte. E possono essere interpretati anche per ciò che codificano *localmente*. Ciascuno filtro agisce sull'immagine tridimensionale in completa analogia a quanto mostrato nelle figure 19 e 22 per il caso bidimensionale. L'attivazione della cellula complessa $E(x, y, t, \xi, \eta, \sigma)$ corrispondente è interpretabile come la verisimiglianza che nel punto spazio-temporale (x, y, t) , al livello di scala σ , il gradiente dell'immagine di input I abbia direzione parallela (o antiparallela) al versore \mathbf{u} definito in (18).

Figura 29. Rappresentazione di campi recettivi dispari spatio-temporali



Esiste un altro modo di caratterizzare l'attività $E(x, y, t, \zeta, \eta, \sigma)$. Associamo al vettore \mathbf{u} , in maniera rigida, altri due, ad esso perpendicolari. Uno, che possiamo indicare con \mathbf{v} , denota la velocità spatio-temporale stimata dell'immagine nel punto (x, y, t) ; l'altro, che indichiamo con \mathbf{w} , è perpendicolare sia ad \mathbf{u} sia a \mathbf{v} ed è parallelo alla direzione stimata del bordo spaziale dell'immagine

nel punto (x, y, t) . Si ha che

$$\mathbf{v} := (-\cos(\xi)\cos(\eta), -\sin(\xi)\cos(\eta), \sin(\eta)) \quad (19)$$

e

$$\mathbf{w} := (-\sin(\xi), \cos(\xi)) \quad (20)$$

Il concetto di *vettore velocità spazio-temporale* si differenzia da quello noto di *vettore-velocità spaziale* perché si ha bisogno di una dimensione in più. Infatti nello spazio-tempo ogni piccolo elemento (dx, dy, dt) è uguale a $(dx/dt, dy/dt, 1)dt$ che, a sua volta, per definizione di velocità, può essere riscritto come $(v_x, v_y, 1)dt$. Quindi nello spazio tempo il vettore velocità (v_x, v_y) può essere rappresentato con $(v_x, v_y, 1)$. Ciò che importa comprendere è l'informazione che tale vettore veicola. Prendendo il versore \mathbf{v} definito in (19), possiamo farlo diventare della forma "(qualcosa_x, qualcosa_y, 1)" semplicemente dividendo la prima e la seconda componente per la terza; possiamo perciò affermare che un filtro definito dal versore \mathbf{u} , nello stesso modo in cui è definito in (4) e (5), ma aggiungendo una dimensione spazio-temporale ed una angolare, è sensibile – risponde in maniera specifica – ad un segnale spazio-temporale il cui gradiente sia parallelo o antiparallelo ad \mathbf{u} o, in maniera equivalente, ad un segnale che, all'istante t nel punto spaziale (x, y) , abbia velocità

$$\mathbf{v}^{classica} := (-\cos(\xi)\cot(\eta), -\sin(\xi)\cot(\eta)) \quad (21)$$

In poche parole, l'unità complessa $E(x, y, t, \zeta, \eta, \sigma)$, associata al campo recettivo in questione (definito tramite (4) o (5) e con \mathbf{u} da (18)), è sintonizzata su segnali di movimento che vanno in *direzione* $(-\cos(\zeta), -\sin(\zeta))$ a *velocità* di modulo $\cot(\eta)$. Tipicamente tali segnali sono linee la cui direzione è perpendicolare a ζ .

La descrizione precedente delle osservabili fisiche codificate da unità con campi recettivi spazio-temporali richiede di occuparsi del significato neurofisiologico di unità per cui $\eta=0$ e, quindi, $\cot(\eta) = \infty$. Ha senso postulare l'esistenza di unità siffatte? Sperimentalmente si è dimostrato [60] che esistono. Dal punto di vista teorico, inoltre, esse si rivelano utili, per codificare informazioni su flash luminosi, ed eleganti, perché rendono naturale il processo di generalizzazione.

È opportuno osservare che anche il modo in cui è trattato il tempo risulta a prima vista problematico. Infatti i filtri che usiamo sono *non causali*, cioè si assume implicitamente che il tempo sia *isotropo* e trattato dal punto di vista neurale come una dimensione in tutto e per tutto uguale a quelle spaziali. L'isotropia del tempo trattato in questo modo diviene epistemologicamente sensata se trattiamo il complesso delle unità come costituissero un *buffer* di memoria – una sorta di memoria iconica, per intenderci. In tal caso, infatti, l'area di integrazione spazio-temporale che si trova nel *futuro* rispetto alla coordinata t del centro del campo recettivo delle cellule semplici denoterebbe il tempo di attesa prima che il resto del sistema utilizzi l'informazione codificata.

Ora possiamo mettere definitivamente in pratica il processo di generalizzazione del modello. Abbiamo, finora, sostituito solo due equazioni a quelle introdotte nel paragrafo precedente: in (17) abbiamo ridefinito la gaussiana (2) aggiungendo la dimensione temporale ed in (18) abbiamo ridefinito il versore (3), aggiungendo una componente angolare. Le (4-7) rimangono invariate, la (8) e la (9) divengono, rispettivamente,

$$D(\mathbf{u}_0, \mathbf{u}_1) = D(\xi_0, \eta_0, \xi_1, \eta_1) := 2(\mathbf{u}_0 \cdot \mathbf{u}_1)^2 - 1 \quad , \quad (22)$$

e

$$\mathbf{s} := (x_1, y_1, t_1) - (x_0, y_0, t_0) = (\Delta x, \Delta y, \Delta t) \quad (23)$$

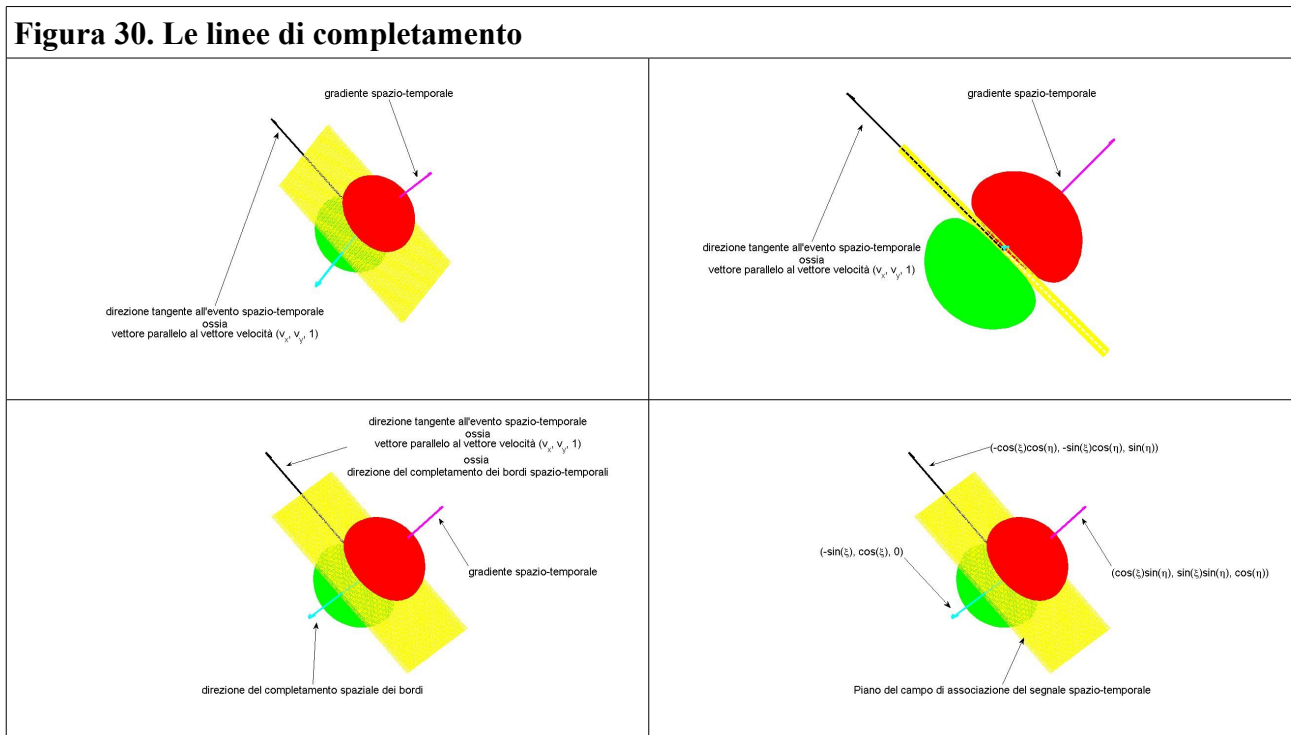
Con questi cambiamenti, la formula del campo associativo nel caso spazio-temporale tridimensionale rimane invariata, precisamente

$$K_\sigma(x_0, y_0, \xi_0, \eta_0, x_1, y_1, \xi_1, \eta_1) = \frac{1}{2} e^{-\frac{\mathbf{s} \cdot \mathbf{s}}{4\sigma^2}} \left(2(\mathbf{u}_0 \mathbf{M} \mathbf{u}_1^t)^2 - 1 \right) (\mathbf{u}_0 \mathbf{M} \mathbf{u}_0^t + \mathbf{u}_1 \mathbf{M} \mathbf{u}_1^t) \quad (24)$$

L'equazione (16), che rappresenta nel caso bidimensionale l'evolversi dello stato delle cellule complesse, a seguito dell'interazione permessa dal campo K_σ , diviene

$$E_{j+1}(x_0, y_0, t_0, \xi_0, \eta_0, \sigma) = \iiint dx_1 dy_1 dt_1 d\xi_1 \sin \eta_1 d\eta_1 E_j(x_1, y_1, t_1, \xi_1, \eta_1, \sigma) K_\sigma(x_0, y_0, \xi_0, x_1, y_1, \xi_1) \quad (25)$$

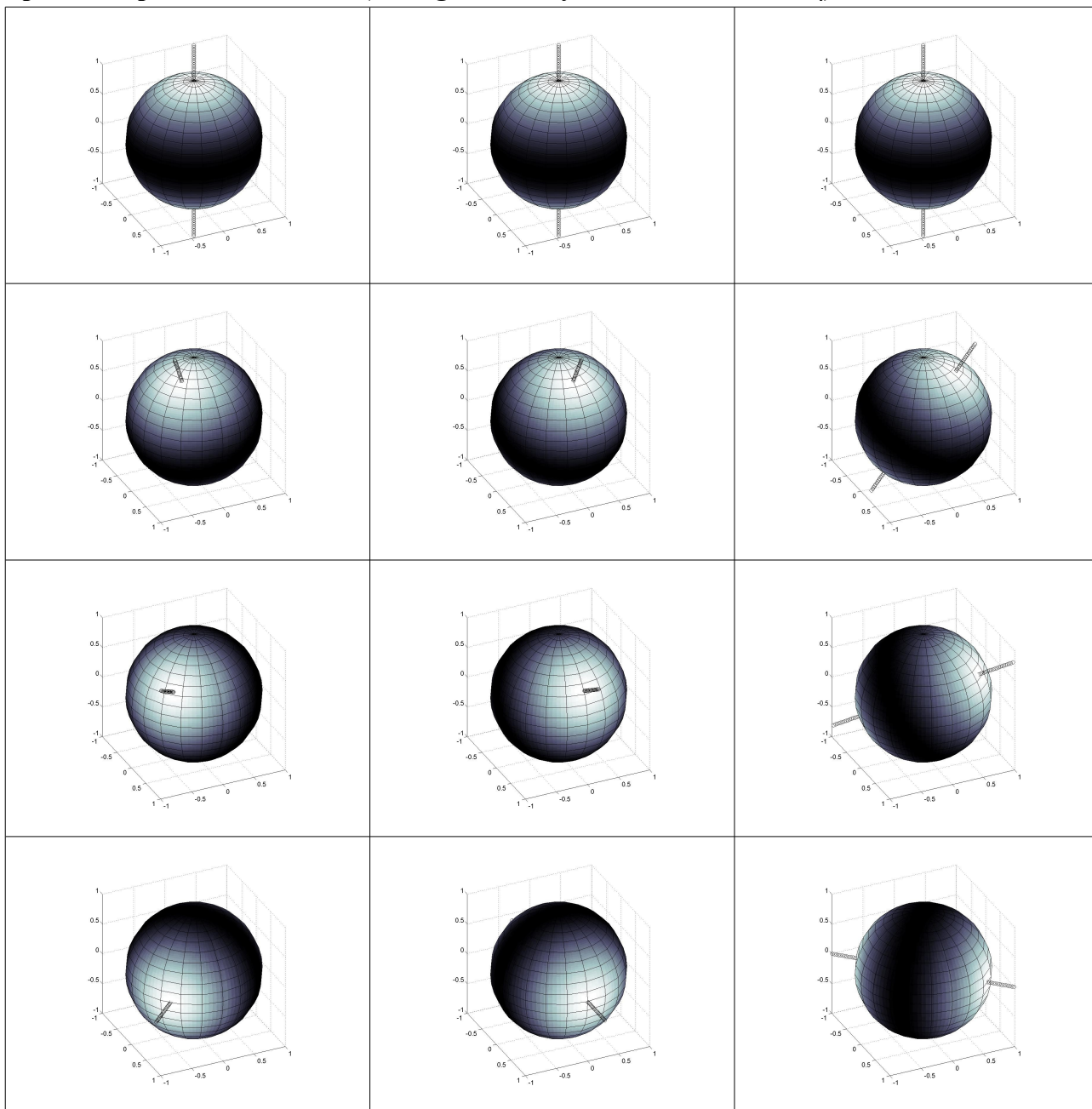
Dove, in analogia a (16), indichiamo con $E_j(x, y, t, \xi, \eta, \sigma)$ l'attività dell'unità complessa $n(x, y, t, \xi, \eta, \sigma)$ al j -esimo passo di computazione. Anche in questo caso $E_0 = |A|^2$, ossia lo stato iniziale consta dello stato delle cellule complesse a seguito della sola attivazione *feedforward*.



È interessante confrontare il campo all'interno delle ipercolonne nel caso spazio-temporale tridimensionale con quello bidimensionale: là è una *circonferenza*, qui una *sfera* [Infatti il fattore $\sin(\eta_1)$ è stato aggiunto in (25), rispetto a (16), per tenere conto della topologia sferica del pattern di associazione tra elementi del medesimo modulo corticale]. Per il resto la logica è la stessa. In figura 31 riportiamo alcuni esempi delle relazioni di muta attivazione o inibizione tra le cellule complesse. La linea uscente denota, in forma vettoriale, la cellula complessa *target*, quella, cioè, su cui verte il

campo intra-modulare (ossia i contributi provenienti dalle altre cellule del modulo corticale). Le unità sono disposte sulla superficie sferica e il colore associato ad ogni punto rappresenta il *peso sinaptico* della unità corrispondente nei confronti della cellula complessa target. Il bianco è massimamente eccitatorio ed il nero massimamente inibitorio. Si osservi che quando la cellula target codifica per un movimento a velocità vicina a 1 (terza riga, $\eta \approx 45^\circ$), l'unità che codifica per il movimento alla stessa velocità ma in direzione opposta inibisce massimamente l'unità target (22).

Figura 31. Campo di interazione tra unità del medesimo modulo corticale nel caso dello spazio-tempo tridimensionale (la longitudine è ξ , la latitudine è $\pi/2 - \eta$).



22 Le unità di misura non sono introdotte ma il senso è che, per velocità né troppo alte né troppo basse, le unità all'interno di una iper-colonna si *inibiscono reciprocamente* come richiesto dal motion after-effect: la sovrastimolazione di unità sintonizzate sul movimento in una direzione *libera* dalla inibizione quelle sintonizzate sul movimento in direzione opposta e, quindi, un successivo stimolo statico è visto muoversi in direzione opposta a quella in cui si muoveva lo stimolo di adattamento.

§5.3. Alcune simulazioni

Riportiamo brevemente i risultati di due simulazioni che abbiamo svolto. Esse si differenziano principalmente per il fatto che una testa la sensibilità del modello ad alcune *singolarità* della visione [26]. Dal punto di vista dell'implementazione esse sono state svolte con gli stessi parametri:

- (i) Il cubo spazio-temporale ha lato pari a 31, ossia i due assi spaziali e l'asse temporale sono stati suddivisi in 31 intervalli.
- (ii) I valori di ζ sono 12 in totale e variano da 0° e 345° a passo di 30° . Anche i valori di η sono 12 in totale e variano da 0° e 165° a passo di 15° .
- (iii) σ è fissato a 2.
- (iv) Per ogni stimolo, abbiamo fermato l'elaborazione al quinto ciclo, cioè abbiamo memorizzato $E_j(x, y, t, \zeta, \eta, \sigma)$ per $j = 0, 1, \dots, 6$ ⁽²³⁾.
- (v) Per rendere più veloci i cicli di calcolo, abbiamo utilizzato una versione modificata del campo associativo, in analogia a quanto fatto in [14]. Tale semplificazione consiste nello svolgere separatamente il calcolo dell'interazione all'interno di un modulo corticale ed il calcolo tra moduli corticali distinti.
- (vi) La risposta di ciascuna modulo-corticale è stata calcolata come l'asse principale di inerzia del pattern di attivazione delle cellule complesse, considerate come disposte su una sfera e dotate di massa proporzionale al loro stato di attivazione. Il relativo vettore velocità è stato calcolato come in (19): precisamente, detto $\mathbf{u} = (u_x, u_y, u_t)$ l'asse ottenuto, il vettore velocità che si ritiene decodificato è pari a $\mathbf{v} = (-u_x u_t, -u_y u_t, 1) / \sqrt{(u_x^2 + u_y^2)}$. La *saliienza* del vettore velocità così ottenuto è stimata prendendo la norma di \mathbf{u} , cioè $\sqrt{(u_x^2 + u_y^2 + u_t^2)}$.

§5.3.1. Simulazione 1. Onde

Sono stati usati due tipi di stimoli. Lo stimolo $I(x, y, t)$ può essere una semplice sinusoidale spazio-temporale – una sinusoidale che cambia di fase nel tempo – o un'onda sferica. Si rimanda il lettore alla figura 6 per la loro rappresentazione ⁽²⁴⁾. Nel caso della sinusoidale si vede un movimento coerente verso una direzione. Per quanto concerne l'onda sferica, si vedono una sequenza di fronti d'onda muoversi prima dal centro verso la periferia e poi in direzione contraria.

Entrambe le simulazione danno i risultati sperati. In figura 32 si mostra l'esito della simulazione con l'onda sferica in input (La simulazione con lo stimolo sinusoidale dà risultati simili e non è riportata). Da sinistra a destra e dall'alto in basso sono rappresentati i campi vettoriali di velocità in diversi istanti (frame). Come si nota i primi sono costituiti da frecce centrifughe, a denotare un movimento dal centro alla periferia. Nelle ultime, come atteso, la direzione del movimento cambia e le frecce divengono centripete.

§5.3.2. Simulazione 2. Stimoli à la Michotte

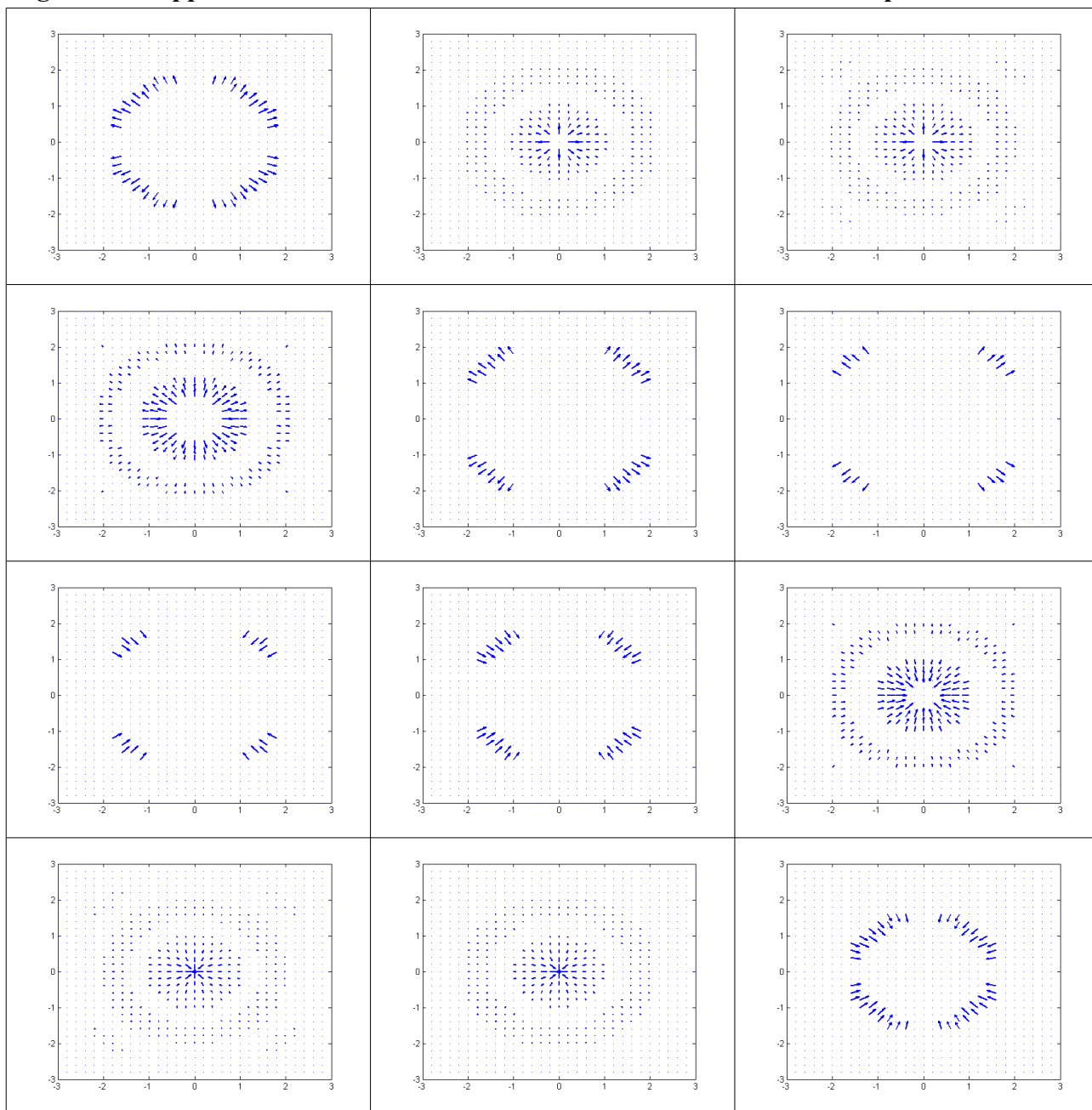
Sono utilizzati 4 tipi di stimolo che denoteremo, rispettivamente, con I^C , I^A , I^B e I^{AB} (Figure 33, 34). Lo stimolo I^{AB} è una tipica configurazione di Michotte. I rimanenti tre sono utilizzati come stimoli di confronto. In I^C è rappresentato un quadrato C che si muove a velocità costante verso destra. In I^A

²³ Ricordiamo che $E_{j=0}$ rappresenta lo stato di attivazione delle cellule semplici determinato dal solo flusso *feedforward*.

²⁴ Le rappresentazioni di figura 6 sono puramente spaziali. Se sull'asse verticale, al posto di una dimensione spaziale, mettiamo il tempo otteniamo le sezioni spazio-temporali degli stimoli in oggetto.

il quadrato si muove alla stessa velocità di C e si ferma dopo un certo intervallo di tempo, rimanendo fermo sino alla fine del filmato. Simmetricamente a ciò che accade nello stimolo I^A , il quadrato B in I^B prima è fermo e poi, improvvisamente, inizia a muoversi a velocità pari a quella di C . Lo stimolo I^{AB} , infine, è costruito accostando A e B . Gli stimoli I^A e I^B sono stati costruiti in modo che, nella configurazione I^{AB} , A si fermi esattamente quando tocca B e quest'ultimo si muova poco dopo l'arrivo di A , come in uno stimolo tipico di Michotte [104].

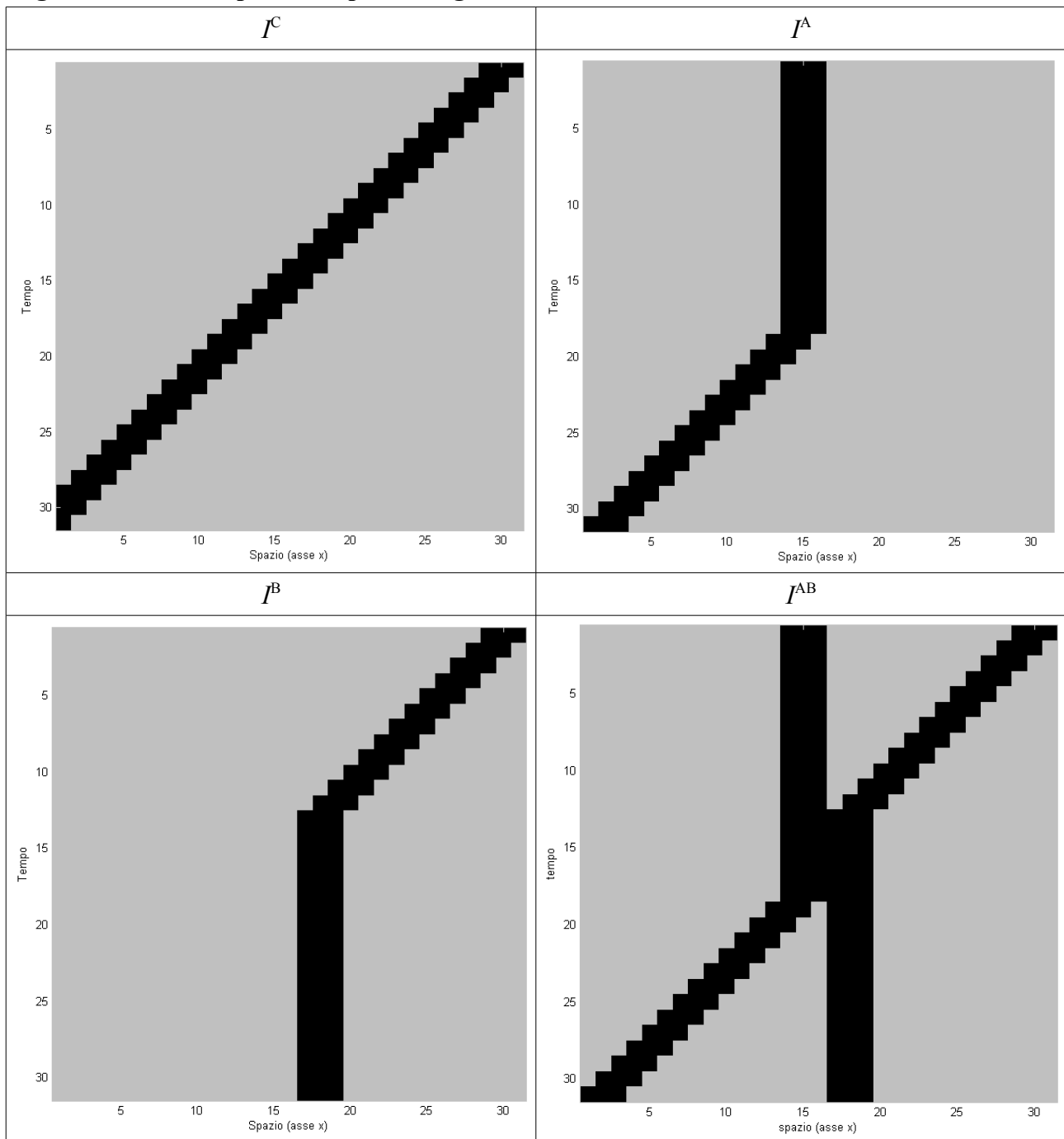
Figura 32. Rappresentazione dei vettori velocità stimati in funzione del tempo



Denotiamo con v_{ψ}^C , v_{ψ}^{0B} e v_{ψ}^{AB} , le velocità percepite, rispettivamente, di C , di B quando è presente da solo e di B quando il suo movimento è anticipato da quello di A nello stimolo I^{AB} . Si osserva che la velocità di B quando parte è percepita maggiore di quella che raggiunge dopo un certo intervallo di tempo dall'inizio del moto. Inoltre si ha una sovrastima della velocità di B quando il suo movimento è anticipato da A [104]. Formalmente si ha, quindi, che $v_{\psi}^C \leq v_{\psi}^{0B} \leq v_{\psi}^{AB}$. Ci aspettiamo, per quanto osservato in [§3.2], che una relazione identica valga tra le *velocità neurali*,

ottenute dopo la procedura di *decoding*, ossia che $v_v^C \leq v_v^{0B} \leq v_v^{AB}$.

Figura 33. Sezione spazio-temporale degli stimoli AB, A, B e C.

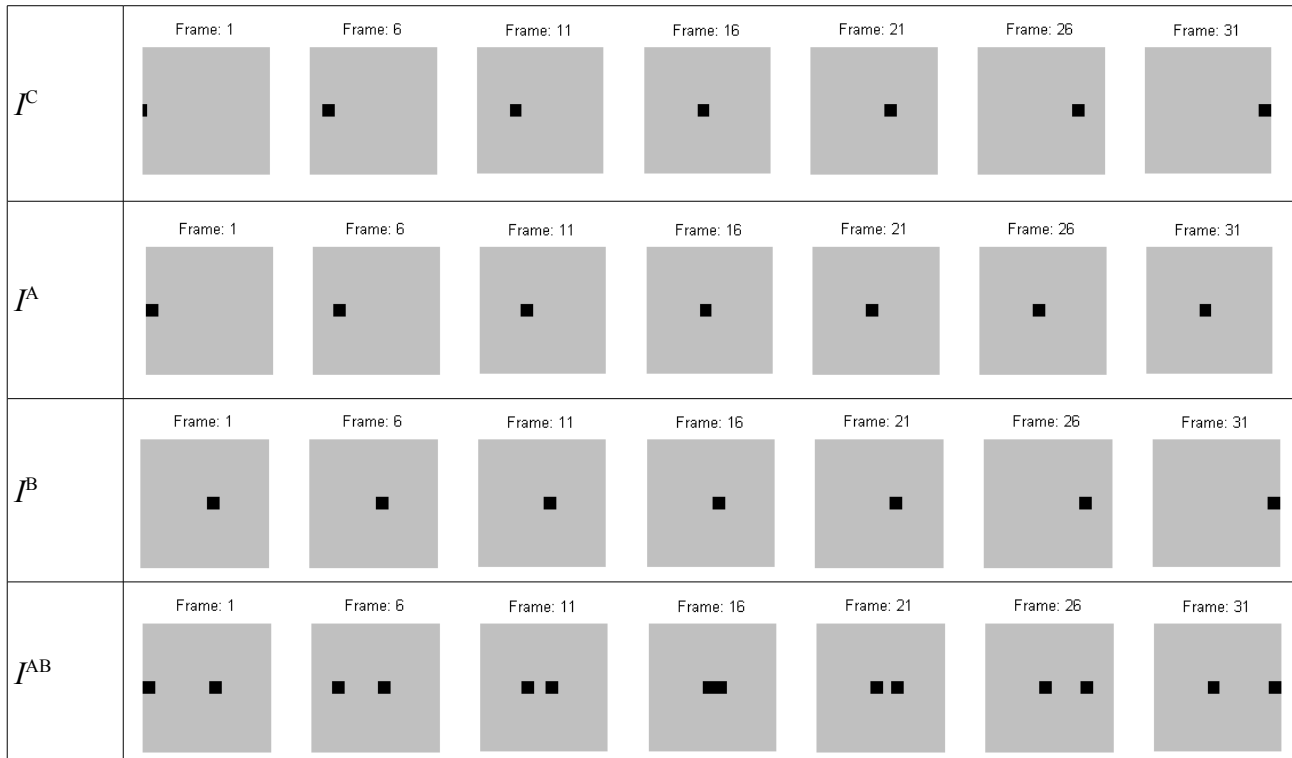


Come accennato, l'esito della simulazione è un campo vettoriale che ad ogni punto associa un vettore velocità ed un numero che rappresenta la *salienza* dello stesso, cioè una stima del fatto che in quel punto vi sia oppure non vi sia un vettore velocità (un segnale di movimento).

Per $Q = AB, A, B, C$, chiamiamo $V_v^Q(x, y, t)$ la velocità scalare stimata dal modello nel punto (x, y, t) e $S_v^Q(x, y, t)$ la salienza di tale stima, ossia una valutazione di quanto sia verosimile vi sia un segnale di velocità in (x, y, t) . Da ogni V_v^Q costruiamo una rappresentazione R_v^Q secondo il seguente criterio: $R_v^Q(x, y, t) = V_v^Q(x, y, t)$ se $S_v^Q(x, y, t) > \text{soglia}(S_v^Q)$, altrimenti $R_v^Q(x, y, t) = 0$. Abbiamo posto $\text{soglia}(S_v^Q) = \text{media di } S_v^Q + \text{una deviazione standard di } S_v^Q$, dove media e deviazione standard sono

state calcolate su tutto lo spazio-tempo tridimensionale. In altri termini R_v^Q mantiene solo i valori di V_v^Q che sono molto salienti. Infine sommiamo su y – cioè calcoliamo $\sum_y R_v^Q(x, y, t)$ – ed otteniamo la rappresentazione finale $v_v^Q(x, t)$ ponendo $v_v^Q(x, t) = \sum_y R_v^Q(x, y, t)$ se e solo se $(x, 0, t)$ appartiene alla componente di B che ne rappresenta il movimento; in altri termini teniamo solo le stime di velocità scalare associate ad un punto dell'immagine di input che rappresenta il movimento di B .

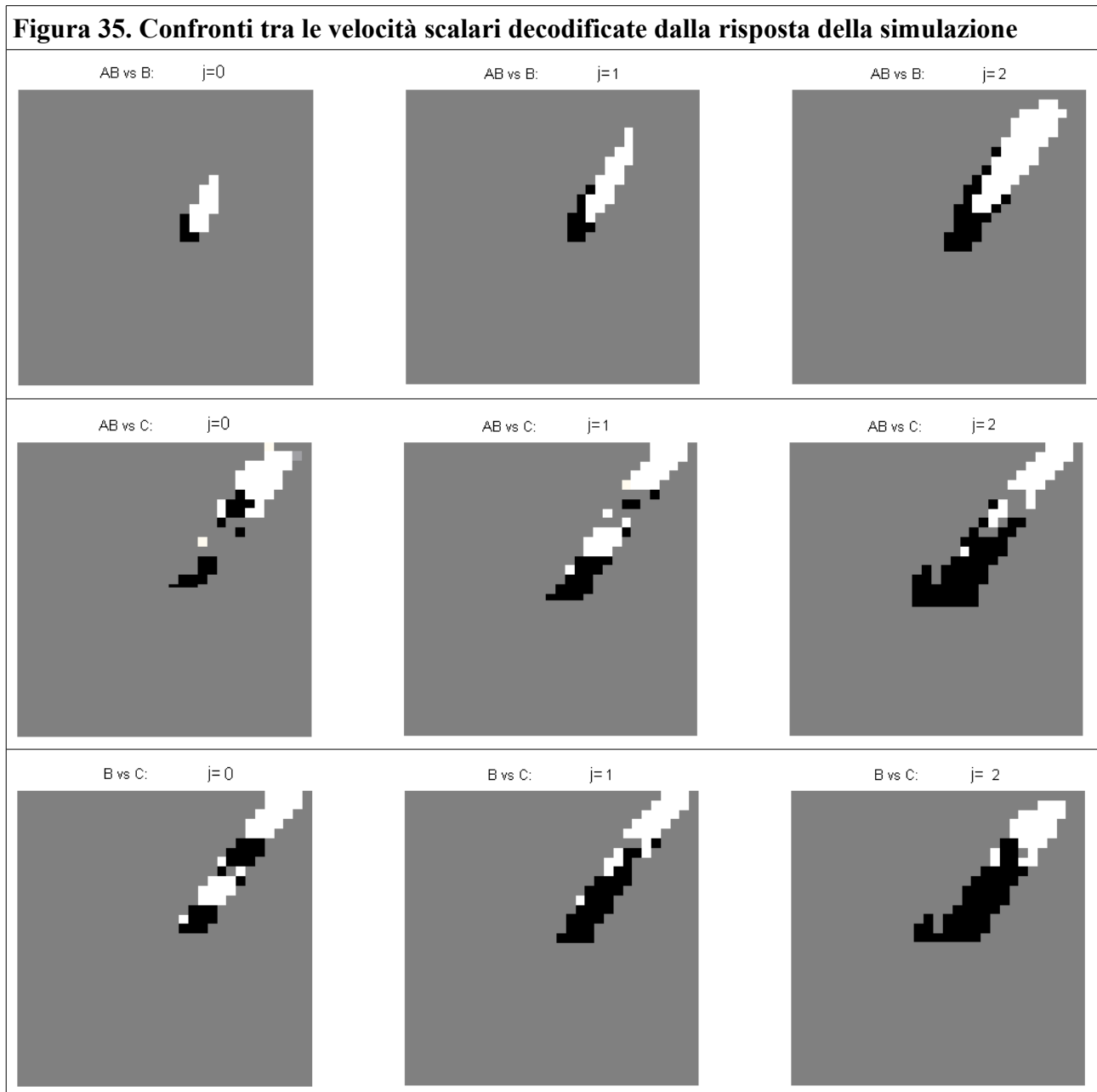
Figura 34. Frame degli stimoli AB, A, B, e C.



Nella prima riga di figura 35 sono rappresentate, per i passi computazionali sino al terzo, le differenze $v_v^{AB}(x, t) - v_v^A(x, t) - v_v^B(x, t)$. Si è posto di mettere il colore *bianco* nei voxel in cui c'è un valore maggiore di zero, il colore grigio per i valori uguali a zero ed il colore *nero* per i valori minori di zero: le aree bianche degli stimoli sono, in altri termini, quelle in cui vi è una sovrastima della velocità nella condizione AB , rispetto alla somma delle condizioni A e B . Le immagini nella seconda e terza riga sono state costruite con gli stessi criteri usati per le immagini della prima riga. Nella seconda riga sono rappresentate le differenze $v_v^{AB}(x, t) - v_v^C(x, t)$, cioè la differenza di velocità stimata tra la condizione AB e la condizione C . Nella terza riga, infine, sono rappresentate le differenze tra la condizione B e la condizione C , ossia $v_v^B(x, t) - v_v^C(x, t)$.

Guardando la prima riga di figura 35, sembrerebbe che il modello sia in grado di spiegare anche la sovrastima di velocità del corpo B dovuta al corpo A . Anche nelle figure relative al confronto tra lo stimolo AB e lo stimolo C (seconda riga) e in quelle dell'ultima riga, che presentano il confronto tra lo stimolo B e lo stimolo C , vi sono aree bianche e, quindi, zone di sopravvalutazione congruenti con l'ipotesi che $v_v^C \leq v_v^{OB} \leq v_v^{AB}$. Un esame più attento, tuttavia, ci pone di fronte alla complessità della questione. Infatti se calcoliamo la velocità *media* delle aree spazio-temporali riportate nei grafici, ovvero sia associamo, prendendo ad esempio il caso AB , a ciascun punto bianco o nero il reale valore $v_v^{AB}(x, t)$ e ne facciamo la media, otteniamo che le velocità medie nelle tre condizioni sono molto simili, per tutti e tre i passi di computazione raggiunti, come rappresentato in figura 36. La situazione è problematica poiché, secondo le differenze medie di velocità, la condizione C determinerebbe un segnale neurale compatibile con l'esperienza di vedere C più veloce (i valori sono negativi) di quanto sembri B in AB ; parimenti la

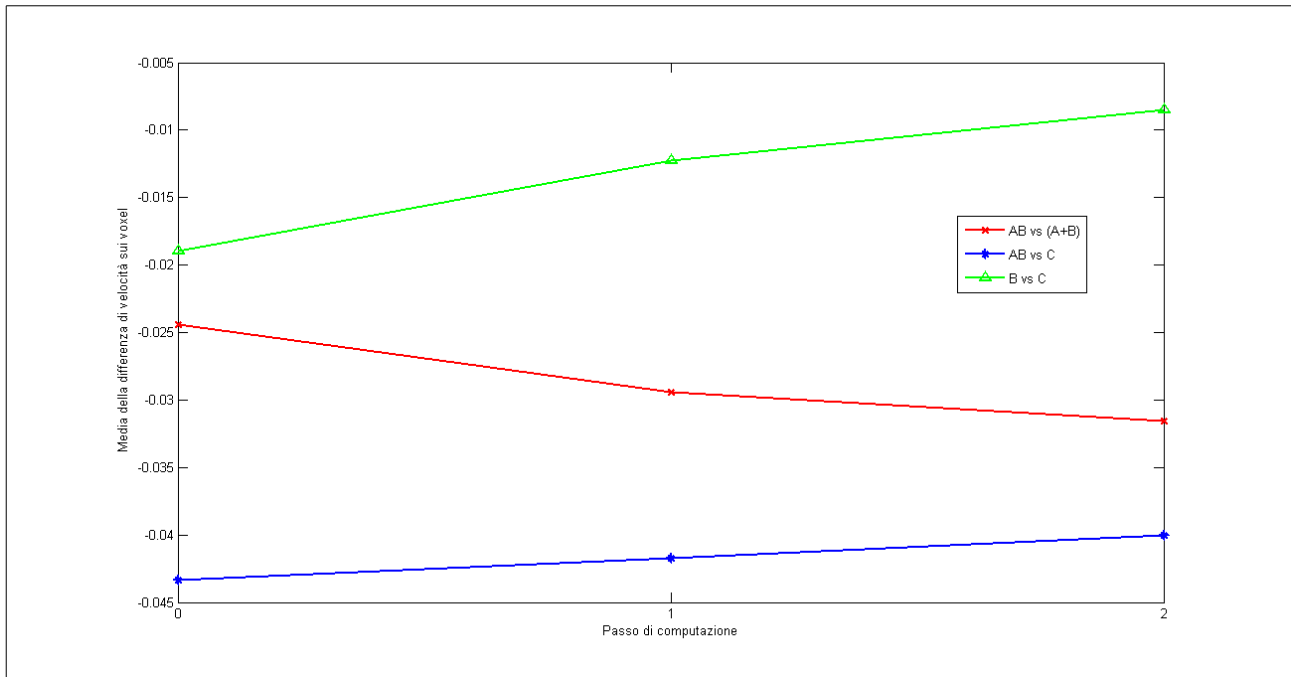
seconda linea del grafico implicherebbe che il movimento di B è visto più lento quando è nel contesto assieme ad A (stimolo di Michotte) piuttosto che quando è da solo.



È possibile spiegare questa ambivalenza nella risposta del modello, darle un senso? Io credo vi possano essere quattro tipi di spiegazione. Una concerne il fattore di scala, che non è stato manipolato ma lasciato fisso: può darsi che il prendere in considerazioni le interazioni a livelli di scala maggiori possa mostrare l'efficacia del modello anche nell'integrazione dei segnali nella configurazione di Michotte. Un altro fattore è la scelta di implementare il modello in maniera analoga a [14]: si ha un forte risparmio di risorse computazionali ma, forse, si perde la possibilità di implementare la risposta di sovrastima allo stimolo di Michotte. Invece la spiegazione secondo la quale l'esito di decoding cambierebbe se si prendessero più cicli di computazione non regge. In un altro esperimento, qui non riportato, il sistema ha elaborato iterativamente lo stato delle cellule complesse sino a che ha raggiunto un punto di convergenza (in circa 30 cicli) ma il pattern di risposta – per quanto concerneva le velocità stimate – era simile a quello che abbiamo riportato

dopo 2 cicli di computazione. Infine, può darsi che il particolare esito fenomenico della configurazione di Michotte richieda l'introduzione di ulteriori vincoli o stadi di computazione, ovvero che, dal punto di vista neurale, coinvolga livelli di elaborazione successivi.

Figura 36. Confronto delle medie delle differenze di velocità rappresentate nelle tre condizioni AB vs (A+B), AB vs C e B vs C.



§6. Conclusioni

Abbiamo mostrato la possibilità di utilizzare il linguaggio matematico per organizzare una prospettiva operazionista unitaria in cui psicofisica, neurofisiologia computazionale e modellistica interagiscano proficuamente.

Abbiamo mostrato come sia possibile estendere il modello classico di energia in due direzioni: utilizzando lo spazio-tempo tridimensionale come dominio di input ed aggiungendo le connessioni laterali tra unità complesse per implementare i processi di organizzazione globale dell'elaborazione dello stimolo.

È stato posto in evidenza che questa generalizzazione possiede una formulazione abbastanza semplice, quando si ritenga la dimensione temporale equivalente ad una dimensione spaziale. Tale equivalenza consiste nell'assegnare al tempo una isotropia che non ha senso fisico, a meno che non si ritenga che il sistema di unità organizzate nello spazio-tempo formi una sorta di *buffer visivo* – come una memoria iconica.

Le simulazioni forniscono risultati incoraggianti per quanto concerne stimoli che non implicano una *singolarità percettiva*, che, cioè, non danno un esito fenomenico inaspettato o curioso che spesso viene chiamato *illusione*. I risultati con la configurazione stimolo di Michotte sono complessi e, a prima vista, incongruenti; ciò può essere dovuto al fatto che non si è approfondito il ruolo del fattore di scala o alla scelta, per ragioni di economia di calcolo, di implementare una versione semplificata del modello, ammesso che il livello di analisi scelto sia sufficiente per spiegare la sovrastima di velocità nelle configurazioni-stimolo di Michotte.

Vi sono diverse vie da seguire per articolare ed approfondire ulteriormente il modello proposto, tra le quali le più salienti sono:

- Studiare il ruolo del fattore di scala, in modo da poter costruire un campo che spieghi i legami tra unità che differiscono per esso. Una proposta in tal senso, ma nel caso statico bidimensionale, è proposta in [74].
- Analizzare le caratteristiche formali dell'operatore lineare implementato dal campo associativo. Esso è simmetrico e, quindi, possiede interessanti proprietà formali che consistono nel fatto che vi sono alcuni stati di attivazione delle cellule complesse che rimangono invariati anche a seguito dell'interazione tramite il campo. Tali stati costituirebbero una sorta di *base* con cui costruire tutti gli stati possibili, fornendo possibili semplificazioni delle procedure di calcolo. Nel caso bidimensionale è già stato fatto in [14].
- Approfondire la questione dell'isotropia della dimensione temporale e valutare se utilizzare altri tipi di invarianze per caratterizzarla. Infatti – anche se non se ne è fatta menzione per ragioni di economia dell'esposizione – l'operatore che implementa il campo associativo è invariante per roto-traslazioni ed è proprio il fatto di poter ruotare il campo lungo la dimensione temporale come fosse una dimensione spaziale a costituire l'analogo formale della postulata isotropia del tempo. Alla luce di ciò, si potrebbe valutare se introdurre un altro gruppo di trasformazioni che caratterizzino diversamente la dimensione temporale, per esempio il gruppo di Lorentz.

Bibliografia

- [1] Rosati, S. (1986). *Fisica generale*. Ambrosiana, Milano.
- [2] Manning, A. (1979). *Il comportamento animale*. Bollati Boringhieri, Torino
- [3] Lorenz, K. (1980). *L'etologia*. Il Mulino, Bologna.
- [4] Vallortigara, G. (2000). *Altre Menti. Lo studio comparato della cognizione animale*. Il Mulino, Bologna.
- [5] Kandel, E. R., Schwartz, J. H., Jessell, T. M. (1994). *Principi di Neuroscienze*. Ambrosiana, Milano.
- [6] Bear, M. F., Connors B. W., Paradiso M. A. (1999). *Neuroscienze*, Masson.
- [7] Umiltà, C. (a cura di) (1999). *Manuale di neuroscienze*, Il Mulino.
- [8] Rosenzweig, M. R., Leiman A. L., Breedlove S. M. (2002). *Psicologia Biologica*, Casa Editrice Ambrosiana.
- [9] Casco, C. (1992). *Vedere*. Bollati Boringhieri, Torino.
- [10] Purghé, F., Stucchi, N., Olivero, A. (1999). *La percezione visiva*, UTET.
- [11] Bruce, V., Green, P.R., Georgeson, M.A., (2003). *Visual Perception*, Psychology Press.
- [12] Yamazaki, T. (1987). "Non-Riemannian approach to geometry of visual space: An application of affinely connected geometry to visual alleys and horopter", *Journal of Mathematical Psychology*, 31: 270-298.
- [13] Gogel, W.C. (1990). "A theory of phenomenal geometry and its applications", *Perception and Psychophysics*, 48: 105-123.
- [14] Bressloff, P.C., Cowan, J.D., Golubitsky, M., Thomas, P.J., Wiener, M.C. (2001). "Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex", *Philosophical Transactions of the Royal Society London B*, 356:299-330.
- [15] Petitot, J. (2003). "An introduction to the Mumford-Shah segmentation model", *Journal of Physiology*, 97: 335-342.
- [16] Baloch, A.A., Grossberg, S. (1997). "A neural model of high-level motion processing: line motion and formotion dynamics", *Vision Research*, 37: 3037-3059.
- [17] Chey, J., Grossberg, S., Mingolla, E. (1998). "Neural dynamics of motion processing and speed discrimination", *Vision Research*, 38: 2769-2786.
- [18] Mingolla, E. (2003). "Neural models of motion integration and segmentation", *Neural networks*, 16: 939-945.
- [19] Rumelhart, D.E., McClelland, J.L. (1986). *Parallel Distributed Processing. Explorations in the microstructure of cognition, (Vol I. Foundations. Vol II. Psychological and Biological Models)*. The MIT Press, Cambridge.
- [20] Floreano, D. (1996). *Manuale sulle reti neurali*. Il Mulino, Bologna.
- [21] Florack, L.M.J., ter Haar Romeny B.M., Koenderink J.J., Viergever, M.A. (1992). "Scale and differential structure of images", *Image and Vision Computing*, 10(6): 376-388.
- [22] Xu, C., Prince, J.L. (1998). "Snakes, shapes and gradient vector flow", *IEEE transactions on Image Processing*, 7(3): 359-369.
- [23] Vidotto, G. in Vicario, G.B. (a cura di) (1988). *Psicologia sperimentale*, CLEUP Editore, Padova.
- [24] Purghé, F. (1997). *Metodi di psicofisica e di scaling unidimensionale*. Bollati Boringhieri, Torino.
- [25] Burigana, L. (1986). *Lineamenti del progetto psicofisico*, CLEUP Editore, Padova.
- [26] Burigana, L. (1996). *Singularità della visione*, UPSEL Domeneghini Editore.
- [27] Burigana, L. (1999). *From optical to visual relational constraints*, Guerini Studio.
- [28] Folstein, M.F., Folstein S.E., McHugh, P.R. (1975). "Mini-mental state': a practical method for grading the cognitive state of patients for the clinician", *Journal of Psychiatric Research*, 12: 189-198.
- [29] Denes, G., Pizzamiglio, L. (1996). *Manuale di neuropsicologia*. Zanichelli.
- [30] Othmer, E., Othmer S.C. (2004). *L'intervista clinica con il DSM-IV-TR*. Raffaello Cortina Editore.
- [31] Cristante, F., Lis, A., Sambin, M. (1992). *Fondamenti teorici dei metodi statistici in psicologia*. UPSEL Domeneghini Editore, Padova.
- [32] Pedon, A. (1995). *Metodologia per le scienze del comportamento*. Il Mulino, Bologna.
- [33] Luccio, R. (1996). *Tecniche di ricerca e analisi dei dati in psicologia*. Il Mulino, Bologna.
- [34] Gottsdanker, R.M. (1956). "The ability of human operators to detect acceleration of target motion", *Psychological Bulletin*, 53: 477-487.
- [35] Dayan, P., Abbott L.F. (2001). *Theoretical Neuroscience*, MIT Press.
- [36] Gava, G. (1996). *Il riduzionismo della scienza*. Guerini Scientifica.
- [37] Edelman, G.M. (1993). *Sulla materia della mente*. Adelphi.
- [38] Pessoa, L., De Weerd, P. (2003). *Filling in. From perceptual completion to cortical reorganization*. Oxford University Press.
- [39] Pessoa, L., Thompson, E., Noë, A. (1998). "Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception", *Behavioral and Brain Sciences*, 21: 723-802.
- [40] Horn, B.K.P., Schunck, B.G. (1981). "Determining optical flow", *Artificial Intelligence*, 17: 185-203.
- [41] Bergen, J.R., Anandan, P., Hanna, K.J., Hingorani, R. (1992). "Hierarchical model-based motion estimation", *Lecture Notes in Computer Science*, 588: 237-252.
- [42] Simoncelli, E.P., Adelson, E.H., Heeger, D.J., (1991). "Probability distributions of optical flow", *IEEE Conference*

on computer vision and pattern recognition, Maui, Hawaii.

- [43] Kanizsa, G. (1980). *Grammatica del vedere*, Il Mulino.
- [44] Field, D.J., Hayes, A., Hess, R.F. (1993). "Contour integration by the human visual system: evidence for a local 'association field'", *Vision Research*, 33(2): 173-193.
- [45] Snowden, R.J., Braddick, O. (1991). "The temporal integration and resolution of velocity signals", *Vision Research*, 31: 907-914.
- [46] Watamaniuk, S.N.J., Duchon, A. (1991). "The human visual system averages speed information", *Vision Research*, 32: 931-941.
- [47] Verghese, P., McKee, S.P. (2000). "Stimulus configuration determines the detectability of motion signals in noise", *Journal of the Optical Society of America A*, 17(9): 1525-1534.
- [48] Vreven, D., Verghese, P. (2002). "Integration of Speed Signals in the Direction of Motion", *Perception & Psychophysics*, 64(6): 996-1007.
- [49] Verghese, P., McKee, S.P. (2002). "Predicting future motion", *Journal of Vision*, 2: 413-423.
- [50] Verghese, P., McKee, S.P. (2006). "Motion grouping impairs speed discrimination", *Vision Research*, 46: 1540-1546.
- [51] Izhikevich, E.M. (2007). *Dynamical Systems in Neuroscience*, The MIT Press.
- [52] Georgopoulos, A., Kalaska, J. Caminiti, R., Massey, J. (1982). "On the relations between the direction of two dimensional arm movements and cell discharge in primate motor cortex", *Journal fo Neuroscience*, 2: 1527-1537.
- [53] Ringach, D.L. (2002). "Spatial Structure and Symmetry of Simple-Cell Receptive Fields in Macaque Primary Visual Cortex", *Journal of Neurophysiology*, 88: 455-463.
- [54] Ringach, D.L. (2004). "Mapping receptive fields in primary visual cortex", *Journal of Physiology*, 558(3):717-728.
- [55] Marčelja, S. (1980). "Mathematical description of the responses of simple cortical cells", *Journal of the Optical Society of America*, 70(11):1297-1300.
- [56] Lee, T. S., (1996). "Image representation using 2D Gabor wavelets", *IEEE Trans. Inform. Theory*, 18(10):1-13.
- [57] Lee T. S. (2003). "Computations in the early visual cortex", *Journal of Physiology*, 97:121-139.
- [58] Soodak, R.E. (1986). "Two-dimensional modeling of visual receptive fields using Gaussian subunits", *Proceedings of the National Academy of Sciences of the United States of America*, 83: 9259-9263.
- [59] Young, R.A., Lesperance, R.M., Weston Meyer, W. (2001). "The gaussian derivative model for spatial-temporal vision: I. Cortical model", *Spatial Vision*, 14(3,4): 261-319.
- [60] Young, R.A., Lesperance, R.M. (2001). "The gaussian derivative model for spatial-temporal vision: I. Cortical data", *Spatial Vision*, 14(3,4): 321-389.
- [61] DeAngelis, G.C., Ohzawa, I., Freeman R.D. (1995). "Receptive-field dynamics in the central visual pathways", *Trends in Nerurosciences*, 18: 451-458.
- [62] Priebe, N. J., Lisberg, S. G. (2004). "Estimating target speed from population response in visual area MT", *The Journal of Neuroscience*, 24(8): 1907-1916.
- [63] Priebe N.J., Cassanello C.R., Lisberger S.G. (2003). "The Neural Representation of Speed in Macaque Area MT/V5", *The Journal of Neuroscience*, 23(13):5650 -5661.
- [64] Priebe, N.J., Lisberger, S.G., Movshon, J.A. (2006). "Tuning for spatio-temporal frequency and speed in directionally selective neurons of macaque striate cortex", *The Journal of Neuroscience*, 26(11): 2941-2950.
- [65] Pack C. C., Conway B.R., Born R.T., Livingstone M.S. (2006). "Spatiotemporal Structure of Nonlinear Subunits in Macaque Visual Cortex", *The Journal of Neuroscience*, 26(3): 893-907.
- [66] Li Z. (1998). "A neural model of contour integration in the primary visual cortex", *Neural Computation*, 10: 903-940.
- [67] Chavane, F., Monier, C., Binguier, V., Baudot, P., Borg-Graham, L., Lorenceau, J., Frégnac, Y. (2000). "The visual cortical association field: A Gestalt concept or a psychophysiological entity?", *Journal of Physiology – Paris*, 94: 333-342.
- [68] Angelucci, A., Levitt, J.B., Walton, E.J.S., Hupe, J., Bullier, J., Lund, J.S. (2002). "Circuits for Local and Global Signal Integration in Primary Visual Cortex", *The Journal of Neuroscience*, 22(19): 8633-8646.
- [69] Hess, R.F., Hayes, A., Field, D.J. (2003). "Contour integration and cortical processing", *Journal of Physiology – Paris*, 97: 105-119.
- [70] Choe, Y., Miikkulainen, R. (2004). "Contour integration and segmentation with self-organized lateral connections", *Biological Cybernetics*, 90: 75-88.
- [71] Schwabe, L., Obermayer, K., Angelucci, A., Bressloff, P.C. (2006). "The Role of Feedback in Shaping the Extra-Classical Receptive Field of Cortical Neurons: A Recurrent Network Model", *The Journal of Neuroscience*, 26(36): 9117-9129.
- [72] Martinez, L.M., Alonso, J.M.(2001). "Construction of Complex Receptive Fields in Cat Primary Visual Cortex found in the direction from simple cells to complex cells but not in the opposite direction", *Neuron*, 32: 515-525.
- [73] Shams, L., von der Malsburg, C. (2002). "The role of complex cells in object recognition", *Vision Research*, 42: 2547-2554.
- [74] Bressloff, P.C., Cowan, J.D. (2002). "A spherical model for orientation and spatial-frequency tuning in a cortical

- hypercolumn”, *Philosophical Transactions of the Royal Society B*, 357:1643-1667.
- [75] Citti, G, Sarti, A. (2006). “A Cortical Based Model of Perceptual Completion in the Roto-Translation Space”, *Journal of Mathematical Imaging and Vision*, 24: 307-326.
- [76] Lehar, S. (2003). “Directional Harmonic Theory: A Computational Gestalt Model to Account for Illusory Contour and Vertex Formation”, *Perception*, 32(4): 423-448.
- [77] Ben-Shahar, O., Zucker, S. (2004). “Geometrical Computations Explain Projection Patterns of Long-Range Horizontal Connections in Visual Cortex”, *Neural Computation*, 16: 445-476
- [78] Koch, C., Wang, H.T., Mathur, B.(1989). “Computing motion in the primate's visual system”, *Journal of Experimental Biology*, 146: 115-139.
- [79] Medioni, G., Lee, M.S., Tang C.K. (2000). *A Computational Framework for Segmentation and Grouping*, Elsevier Science B.V.
- [80] van Hateren, J.H., van der Schaaf, A. (1998). “Independent component filters of natural images compared with simple cells in primary visual cortex”, *Proc. R. Soc. Lond. B.*, 265: 359-366.
- [81] van Hateren, J.H., Ruderman, D.L. (1998). “Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex”, *Proc. R. Soc. Lond. B*, 265: 2315-2320.
- [82] Reichardt, W. (1969). Movement perception in insects. In W. Reichardt, *Processing of optical data by organisms and by machines*. New York: Academic Press.
- [83] Beghi L., Saviolo N., Xausa E., Zanforlin M., (1983). “Perception of Movement, the Correlation Model and the Landing Reaction of the Flesh Fly (Callphoridae)”, *Biological Cybernetics*, 46:81-91.
- [84] Adelson, E.H., Bergen, J.R. (1985). “Spatiotemporal energy models for the perception of motion”, *Journal of the Optical Society of America A*, 2(2):284-299.
- [85] Adelson E. H., Bergen, J.R. (1991). “The plenoptic function and the elements of early vision”, in Landy, M. & Movshon A. (eds), (1991). *Computational models of visual processing*, Cambridge, MA: MIT Press
- [86] Emerson, R.C., Bergen, J.R., Adelson, E.H. (1992). “Directionally Selective Complex Cells and the Computation of Motion Energy in Cat Visual Cortex”, *Vision Research*, 32: 203-218.
- [87] Watson, A.B., Ahumada, A.J. (1985). “Model of human visual-motion sensing”, *Journal of the Optical Society of America A*, 2(2):1986-2007.
- [88] Chubb, C., Sperling G. (1988). “Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception”, *Journal of the Optical Society of America A*, 5(11):1986-2007.
- [89] Bowns, L. (2002). “Can spatio-temporal energy models of motion predict feature motion?”, *Vision Research*, 42: 1671-1681.
- [90] Chey, J., Grossberg, S., Mingolla, E. (1997). “Neural dynamics of motion grouping: from aperture ambiguity to object speed and direction”, *Journal of the Optical Society of America A*, 14(10): 2570-2594.
- [91] Grossberg, S., Mingolla, E., Viswanathan, L. (2001). “Neural dynamics of motion integration and segmentation within and across apertures”, *Vision Research*, 41: 2521-2553
- [92] Scateni, P., Cignoni, P., Montani, C., Scopigno, R. (2005). *Fondamenti di grafica tridimensionale interattiva*. McGraw-Hill.
- [93] Shadlen M.N., Britten K.H., Newsome W.T., Movshon J.A. (1996). “A Computational Analysis of the Relationship between Neuronal and Behavioral Responses to Visual Motion”, *The Journal of Neuroscience*, 76(4):1486-1510.
- [94] Apostol, T.M. (1978). *Calcolo*. Bollati Boringhieri: Torino.
- [95] Pauli, W. (1981). *Theory of relativity*. Dover.
- [96] Tricomi, F.G. (1985). *Integral equations*. Dover.
- [97] Dettman, J.W. (1988). *Mathematical methods in physics and engineering*. Dover.
- [98] Schouten, J.A. (1989). *Tensor analysis for physicists*. Dover.
- [99] Biondi, E., Morasso, P., Tagliasco, V. (A cura di) (1991). *Neuroscienze e scienze dell'artificiale: dal neurone all'intelligenza*. Pàtron Editore: Bologna.
- [100] Gilardi, G. (1994). *Analisi tre*. MacGraw-Hill.
- [101] Henle, M. (1994). *A combinatorial introduction to topology*. Dover.
- [102] Hausner, M. (1998). *A vector space approach to geometry*. Dover.
- [103] Davies, E.B. (2007). *Linear operators and their spectra*. Cambridge University Press.
- [104] Parovel, G., Casco, C. (2006), “The psychophysical law of speed estimation in Michotte's causal events”, *Vision Research*, 46:4134-4142.