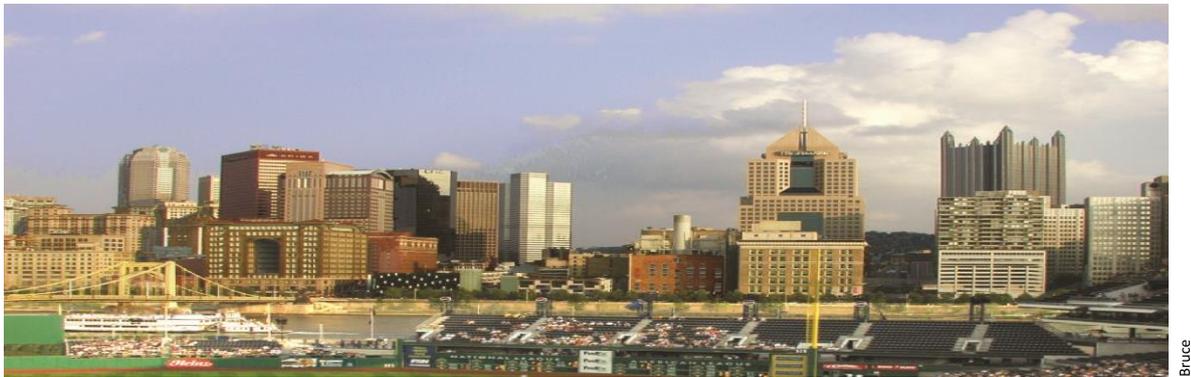
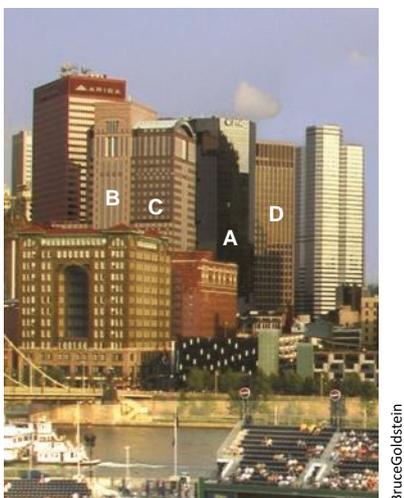


## Capitolo 5. Perceiving Objects and Scenes

Seduto nella parte superiore dello stadio "PNC Park", casa dei Pittsburgh Pirates, Roger guarda verso la città (Figura 5.1). Sulla sinistra vede un gruppo di circa 10 edifici e può facilmente distinguere un edificio dall'altro. Guardando di fronte a sé, vede un edificio piccolo davanti ad uno più grande, e non ha difficoltà ad affermare che sono due edifici distinti. Guardando in basso verso il fiume, sulla destra, nota una banda gialla orizzontale sopra le gradinate del campo. Gli risulta ovvio che questa non faccia parte del campo da baseball ma che sia posizionata dall'altra parte del fiume. Tutte le percezioni arrivano a Roger in modo naturale e richiedono uno sforzo minimo. Ma quando guardiamo la scena da vicino, diventa evidente che essa ponga molti "enigmi". Le seguenti dimostrazioni ne mettono in evidenza alcuni.



**Figura 5.1** È facile affermare che ci sono più edifici diversi tra loro sulla destra e che, di fronte, c'è un edificio basso e rettangolare davanti ad un edificio più alto. È possibile riconoscere che la banda gialla orizzontale sopra le gradinate si trova oltre il fiume. Queste sono percezioni semplici per gli esseri umani ma possono essere abbastanza complesse per un sistema di visione artificiale.



**Figura 5.2** "puzzle" di una città. Vedi la Dimostrazione per le

istruzioni.

### DIMOSTRAZIONE

Enigmi percettivi in una scena

Le domande sottostanti fanno riferimento alle aree etichettate nella Figura 5.2. Il tuo compito è quello di rispondere a ciascuna domanda e indicare il motivo della tua risposta:

- Che cos'è l'area scura di A?
- Le superfici di B e C sono rivolte nella stessa o diversa direzione?
- Le aree B e C di trovano sullo stesso edificio o su edifici differenti?
- L'edificio D si estende dietro all'edificio A?

Anche se potrebbe essere stato semplice rispondere alle domande, è probabile che sia stato in qualche modo più impegnativo indicare quale fosse la motivazione. Per esempio, come sai che l'area scura in A è un'ombra? Potrebbe essere un edificio di colore scuro che si trova di fronte ad un edificio di colore chiaro. Oppure sulla base di cosa hai deciso che l'edificio D si estende dietro all'edificio A? Dopo tutto, potrebbe semplicemente finire esattamente dove inizia A. Potremmo porre domande simili a proposito di qualsiasi cosa in questa scena perché, come vedremo, un particolare schema di forme può essere creato da un gran numero di oggetti.

Uno dei messaggi di questo capitolo è che è necessario andare oltre al modello di chiaro-scuro che la scena crea sulla retina per determinare che cosa sia "là fuori". Un modo per apprezzare l'importanza di questo processo dell'"andare oltre" è considerare quanto sia stato difficile programmare anche i computer più potenti per eseguire compiti percettivi che gli umani realizzano con facilità.

Consideriamo, per esempio, i veicoli robotici che sono stati progettati per competere nella corsa "Urban Challenge" che si è tenuta il 3 Novembre del 2007 a Victorville, in California. Questa corsa, che è stata sponsorizzata dalla *Defense Advanced Research Projects Agency* (DARPA), richiedeva che i veicoli guidassero per 55 miglia attraverso un percorso che assomigliasse alle strade di una città, insieme ad altri veicoli in movimento, segnali stradali e indicazioni. I veicoli dovevano compiere questa impresa da soli, con un coinvolgimento

umano limitato all'inserimento delle coordinate della posizione globale del tracciato del percorso nel sistema di guida del veicolo. I mezzi dovevano rimanere nel circuito ed evitare il traffico imprevedibile senza alcun intervento umano, basandosi solo sul funzionamento di sistemi informatici a bordo.

Il vincitore della corsa, un veicolo della Carnegie Mellon University, è riuscito a rimanere sul tracciato e ad evitare le altre macchine mentre viaggiava alla velocità di 14 miglia all'ora. Anche i mezzi di Stanford, Virginia Tech, MIT, Cornell e dell'Università della Pennsylvania sono riusciti a portare a termine con successo la gara, per un totale di 11 squadre che si sono qualificate per la gara finale.

L'impresa di navigare attraverso l'ambiente, specialmente quando contiene ostacoli che si muovono, è estremamente complessa. Comunque, nonostante questi veicoli robotici fossero in grado di evitare gli ostacoli lungo un percorso definito, non erano in grado di identificare la maggior parte degli oggetti che stavano evitando. Per esempio, anche se la macchina del Carnegie-Mellon era stata capace di evitare un ostacolo in mezzo alla strada, non avrebbe saputo dire se l'ostacolo fosse un cumulo di pietre, un cespuglio o un cane.

specificatamente per determinare la posizione delle pareti di una stanza e per individuare i mobili all'interno della stessa, sono in grado di svolgere questi compiti approssimativamente per alcune fotografie, come nella **Figura 5.4a**, ma spesso commettono degli errori, come nella **Figura 5.4b** (Pero et al., 2011).

Anche se la posizione e l'estensione del letto nella Figura 5.4.b possono essere percepiti facilmente da una persona, non è così scontato che lo siano per un computer, nonostante il programma sia stato specificatamente ideato per rilevare oggetti definiti da linee rette, come il letto. Anche se riuscisse a trovare i bordi del letto, determinare l'identità di altri oggetti nella stanza va al di là delle capacità di questo programma allo stato d'arte.



**Figura 5.3** Un computer o una persona possono determinare se le due prospettive frontali in (a) e (b) rappresentano la stessa persona, ma gli esseri umani hanno prestazioni migliori rispetto ai computer quando vengono paragonate una prospettiva frontale e una ad angolo, come in (c).

Un oggetto che ha ricevuto moltissima attenzione dai ricercatori della visione computerizzata è il viso, nel tentativo di sviluppare un sistema computerizzato di sorveglianza che possa riconoscere i volti. Con una grande quantità di ricerca investita nei sistemi di riconoscimento facciale computerizzati, sono stati sviluppati nuovi programmi che possono determinare, tanto quanto gli esseri umani, se due volti visti frontalmente, come nella **Figura 5.3a e b**, appartengono alla stessa persona o a persone differenti (O'Toole, 2007; O'Toole et al., 2007). Comunque, quando uno dei volti è visto di scorcio, gli esseri umani hanno prestazioni migliori rispetto ai computer. Infine, i sistemi di visione artificiale programmati

distinti. Guardando in basso verso il fiume, sulla destra, nota una banda gialla orizzontale sopra le gradinate del campo. Gli risulta ovvio che questa non faccia parte del campo da baseball ma che sia posizionata dall'altra parte del fiume. Tutte le percezioni arrivano a Roger in modo naturale e richiedono uno sforzo minimo. Ma quando guardiamo la scena da vicino, diventa evidente che essa ponga molti "enigmi". Le seguenti dimostrazioni ne mettono in evidenza alcuni.

## Perché è così difficile progettare una macchina in grado di percepire?

Ora descriveremo alcune delle difficoltà legate alla progettazione di una “macchina che percepisce” (Perceiving Machine). Bisogna ricordare che il punto è che, sebbene questi problemi rappresentino delle difficoltà per i computer, gli esseri umani li risolvono facilmente.

### Lo stimolo sui recettori è ambiguo

Quando osserviamo la pagina di questo libro, l'immagine proiettata dai bordi della pagina sulla nostra retina, è ambigua. Può sembrare strano dirlo, perché una volta che conosciamo la forma dell'oggetto e la sua distanza dall'occhio, determinare la proiezione dell'oggetto sulla retina è un semplice problema geometrico, che può essere risolto estendendo dei “raggi” dagli angoli dell'oggetto fino all'interno dell'occhio, come mostrato nella **Figura 5.5**. Ma il sistema percettivo non si occupa di determinare l'immagine dell'oggetto sulla retina. La percezione *inizia* con l'immagine sulla retina e il suo lavoro è quello di determinare quale sia l'oggetto “là fuori” che ha creato l'immagine. Il compito di determinare quale sia l'oggetto responsabile di una particolare immagine sulla retina è chiamato **inverse projection problem**, perché implica che si inizi con l'immagine retinica e si estendano i raggi *fuori* dall'occhio. Quando lo facciamo, come mostrato nella Figura 5.6, vediamo che la pagina rettangolare (in rosso) potrebbe aver creato l'immagine retinica, ma anche un trapezio, un rettangolo più largo ed un infinito numero di altri oggetti, potrebbero aver creato la stessa immagine. Quando consideriamo che un'immagine particolare sulla retina può essere creata da molti oggetti presenti nell'ambiente, è facile capire perché affermiamo che l'immagine sulla retina sia ambigua.



(a)



(b)

sistema percettivo risolve il problema della proiezione inversa e determina quale oggetto, fra molti, sia responsabile di una particolare immagine sulla retina.

L'ambiguità dell'immagine sulla retina viene illustrata anche nella **Figura 5.7a**, che, se osservata da un luogo specifico, crea un'immagine circolare sulla retina e sembra essere un cerchio di rocce. Comunque, cambiare punto di vista rivela che le rocce, dopotutto, non sono disposte in cerchio (**Figura 5.7b**). Quindi, proprio come un'immagine che risulta rettangolare sulla retina può essere creata da trapezi e altri oggetti non rettangolari, così un'immagine circolare sulla retina può essere creata da oggetti non rotondi.

**Figura 5.4** (a) Le linee rappresentano il tentativo di un programma di visione artificiale di individuare gli angoli della stanza e i punti in cui le pareti, il soffitto e il pavimento si

incontrano. In questo esempio, il computer svolge un lavoro abbastanza buono. (b) Un altro esempio in cui è stato usato il medesimo programma di visione artificiale, in cui le indicazioni del programma riguardo alla posizione dei contorni rettilinei della stanza sono inaccurate.

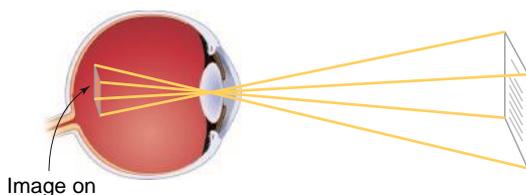
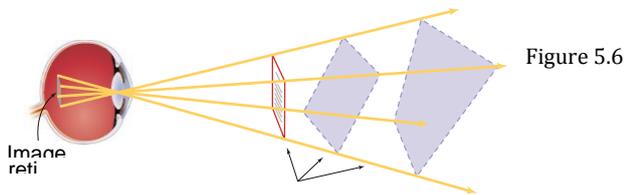


Image on

**Figura 5.5** Determinare la proiezione di un oggetto sulla retina è solo una questione di estensione di raggi dall'oggetto fino all'occhio.

La “environmental rock sculpture”, nella Figura 5.7, è progettata per ingannare lo spettatore creando una condizione particolare (viene osservata da un posto specifico) che si traduce in una percezione errata. Ma la maggior parte delle volte non si verificano percezioni errate come questa, il



**Figura 5.6** Il principio che sta alla base del problema della proiezione inversa, in cui il compito consiste nel determinare quale oggetto crea l'immagine sulla retina. È difficile risolvere tale problema perché la stessa immagine sulla retina può essere creata da un'infinità di oggetti, tra cui i due rettangoli e il trapezio, come raffigurato qui. È il motivo per cui afferiamo che l'immagine sulla retina è ambigua



(a)



(b)

**Figura 5.7** Scultura ambientale di Thomas Macaulay . (a) Quando viene osservata esattamente dal punto di vista corretto (Il balcone del secondo piano della Blackhawk Mountain School of Art di Black Hawk, n Colorado), e pietre sembrano essere disposte in cerchio. (b) Osservare le pietre dal piano terra fornisce delle informazioni riguardo alla loro reale configurazione.

Courtesy of Thomas Macaulay, Blackhawk Mountain School of Art, Blackhawk, CO

Tuttavia, per quanto sia facile per il sistema percettivo umano, risolvere il problema della proiezione inversa pone difficili sfide al sistema visivo del computer.

## Gli oggetti possono essere nascosti o offuscati

Alle volte gli oggetti sono nascosti o indistinti. Potete trovare la matita e gli occhiali da vista nella **Figura 5.8**? (Fermatevi e provatelo prima di leggere oltre). Sebbene questo possa richiedere una piccola ricerca le persone possono trovare la matita in primo piano e la montatura degli occhiali dietro al computer, questo anche se è visibile solo una piccola porzione di questi oggetti.

Questo problema degli oggetti nascosti si verifica ogni qualvolta un oggetto ne nasconde parte di un altro. Ciò succede frequentemente nell'ambiente circostante, ma le persone sono facilmente in grado di capire che la parte dell'oggetto che viene coperta da altro continua ad esistere, sono quindi in grado di usare le loro conoscenze riguardo all'ambiente per determinare cos'è probabile succeda nel presente.

Le persone sono anche in grado di riconoscere oggetti che non son messi perfettamente a fuoco, come le facce in **Figura 5.9**. Guardate quante persone potete identificare, consultate poi le risposte a pagina 123.

Nonostante la natura degradata di queste immagini spesso le persone possono identificarne la maggior parte, mentre i computer performano male a questo compito (Sinha, 2002).

## Da diversi punti di vista gli oggetti appaiono differenti

Un altro problema che deve affrontare una qualsiasi macchina che percepisce è che spesso gli oggetti sono visti da diverse angolazioni. Questo significa che le immagini degli oggetti cambiano in continuazione a seconda dell'angolazione dalla quale sono visti.

Anche se gli esseri umani continuano a percepire l'oggetto in **Figura 5.10** come la stessa sedia vista da diverse angolazioni, ciò non è così ovvio per un computer. L'abilità di riconoscere un oggetto visto da diversi punti di vista è chiamata **invarianza del punto di vista**. Abbiamo già visto che l'invarianza del punto di vista consente alle persone di poter dire se facce viste da angolazioni differenti sono la stessa persona, questo compito è difficoltoso per i computer (fare riferimento alla Figura 5.3).

Le difficoltà che deve affrontare una macchina che percepisce illustrano che il processo di percezione è più



BruceGoldstein

**Figura 5.8** Una porzione del disordine sulla scrivania dell'autore. Potete localizzare la matita nascosta (facile) e gli occhiali dell'autore (difficile)?



**Figura 5.9** Chi sono queste persone? Le soluzioni a pagina 123.

Da Sinha, P. (2002). Riconoscere pattern complessi. *Nature Neuroscience*, 5, 1093-1097. Riprodotto con permesso.

complesso di quel che sembra (qualcosa già lo sapete dal processo percettivo in Figura 1.1 e dal materiale psicologico nei Capitoli 2-4).

Ma gli esseri umani come superano queste complessità? Cominceremo a rispondere a questa domanda considerando l'*organizzazione percettiva*.



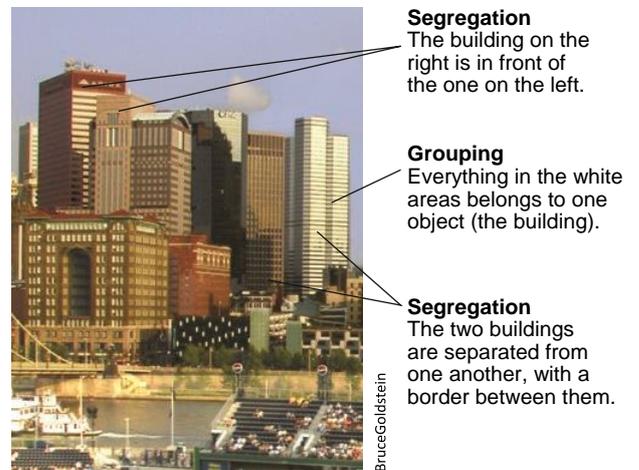
Bruce

**Figura 5.10** La vostra abilità di riconoscere ciascuna di queste rappresentazioni come appartenenti alla stessa sedia è un esempio di invarianza del punto di vista.

### Organizzazione percettiva

L'**organizzazione percettiva** è il processo per il quale gli elementi dell'ambiente circostante vengono percettivamente raggruppati per creare le nostre percezioni degli oggetti. Durante questo processo le stimolazioni in entrata sono organizzate in unità coerenti come oggetti. Il processo dell'organizzazione percettiva include due componenti, *classificazione* e *segregazione* (**Figura 5.11**; Peterson & Kimchi, 2012). La *classificazione* è il processo per il quale eventi visivi sono "messi assieme" in unità o in oggetti. Così, quando Roger vede i palazzi di Pittsburgh come un'unità individuale ha raggruppato gli elementi visivi della scena per creare ogni singolo edificio. Se siete in grado di percepire il Dalmata in **Figura .12**, avete raggruppato percettivamente alcune delle zone scure per formare un Dalmata, con le altre aree scure viste come ombre nello sfondo.

Il processo del raggruppamento lavora assieme a quello di segregazione, il quale consente di separare un'area o un oggetto da altro. Così, il vedere due edifici di **Figura 5.11** come separati tra loro, con i margini che indicano dove l'uno finisce a l'altro inizia, implica la segregazione.



BruceGoldstein

**Figura 5.11** Esempi di raggruppamenti e segregazioni in una città



Photograph by R. C. James

**Figura 5.12** Alcune forme bianche e nere che vengono organizzate percettivamente in un Dalmata

## L'approccio Gestaltista al raggruppamento percettivo

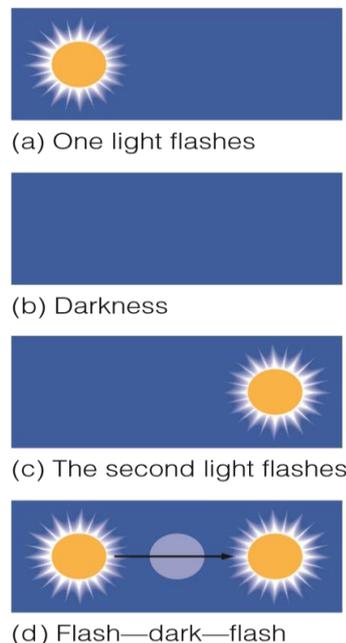
Cos'è che fa sì che alcuni elementi vengano raggruppati in modo tale per cui facciano parte di uno stesso oggetto? Alcune risposte a questa domanda furono date all'inizio del 1900 da parte della **psicologia della Gestalt**- dove *Gestalt*, approssimativamente tradotto, vuol dire *configurazione*. "Come fanno le configurazioni a essere formate da elementi più piccoli?", si chiesero gli psicologi della Gestalt.

**Strutturalismo** Possiamo comprendere l'approccio Gestaltista considerando un approccio antecedente alla psicologia della Gestalt, chiamato *strutturalismo*, il quale fu proposto da Wilhelm Wundt, il quale fondò il primo laboratorio di psicologia scientifica nel 1879 all'Università di Lipsia. Lo **strutturalismo** distingue tra **sensazioni**- processi elementari che si verificano grazie alla stimolazione dei sensi- e **percezioni**, esperienze consapevoli più complesse così come la nostra consapevolezza degli oggetti. Le sensazioni potrebbero essere collegate a esperienze molto semplici, come il vedere un singolo flash di una luce, ma la percezione conta la maggior parte delle nostre esperienze sensoriali. Ad esempio, quando guardate la **Figura 5.13** percepite un volto, ma il punto di partenza in accordo con lo strutturalismo, sarebbero molte sensazioni, le quali sono rappresentate dai piccoli puntini.



**Figura 5.13** Secondo lo strutturalismo, un numero di sensazioni (rappresentato dai puntini) che ammonta crea la nostra percezione di un volto

Gli strutturalisti vedevano le sensazioni come l'analogo degli atomi in chimica. Così come gli atomi si combinano per creare strutture molecolari complesse, le sensazioni si combinano per creare percezioni complesse. Un altro principio dello strutturalismo è che la combinazione delle sensazioni che avviene per formare le percezioni è supportata dall'osservazione delle esperienze passate. Gli psicologi della Gestalt hanno respinto l'idea che le percezioni fossero formate dall'ammontare di sensazioni e hanno respinto anche il concetto che le esperienze passate giochino un ruolo fondamentale nella percezione. Per capire il motivo per il quale gli psicologi della Gestalt ritenevano che la percezione non poteva essere spiegata dalla somma di sensazioni, si consideri l'esperienza dello psicologo Max Wertheimer, il quale nel 1911 era in vacanza e prese un treno che attraversava la Germania (Boring, 1942). Quando a Francoforte scese dal treno per stirarsi i muscoli delle gambe, comprò un gioco stroboscopico da un venditore che stava vendendo giocattoli lungo il binario. Lo stroboscopio, un congegno meccanico che crea un'illusione di movimento dovuto al rapido alterarsi di due immagini leggermente differenti tra loro, spinse Wertheimer a interrogarsi su come l'idea strutturalista secondo la quale l'esperienza è creata da sensazioni, potesse spiegare l'illusione del movimento che osservò.



**Figura 5.14** Le condizioni necessarie per creare un movimento apparente. (a) una luce lampeggiante, seguita da (b) un breve periodo di buio, seguito da (c) un'altra luce lampeggiante situata in un altro punto rispetto alla prima. La percezione risultante, raffigurata in (d), è una luce che si muove da sinistra a destra. Il movimento è visto tra le due luci anche se nello spazio tra queste c'è solo il buio.



CorbisBridge/Alamy

**Figura 5.15** Il maxischermo azionario a Time Square, New York. Per muoversi, lettere e i numeri che vi appaiono sopra, sono creati da centinaia di piccole luci lampeggianti.

**Movimento apparente** La **Figura 5.14** illustra il principio che sta dietro all'illusione del movimento creato dallo stroboscopio, questo principio è chiamato **movimento apparente** perché anche se viene percepito un movimento non c'è niente che si muove. Ci sono tre componenti per gli stimoli che creano un movimento apparente:

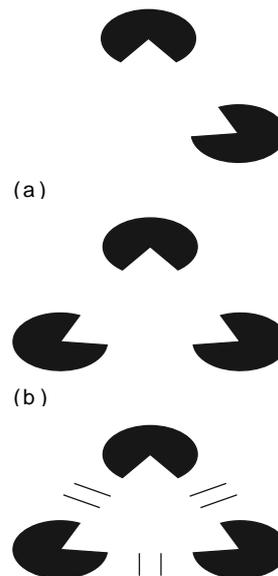
- (1) Un'immagine il cui flash si alterna (on-off) (**Figura 5.14a**);
- (2) Un momento di oscurità della durata di una frazione di secondo (**Figura 5.14b**);
- (3) Si alterna il flash della seconda immagine (**Figura 5.14c**).

Fisicamente, quindi, ci sono due immagini il cui flash si alterna tra on e off, le quali sono separate da un periodo di oscurità. Tuttavia, noi non lo vediamo questo momento perché il nostro sistema percettivo aggiunge qualcosa durante il periodo di buio- la percezione di un'immagine che si muove attraverso lo spazio tra le immagini dei punti (1) e (3), (**Figura 5.14d**).

Un esempio moderno del movimento apparente è rappresentato dalle insegne elettroniche come quella in **Figura 5.15**, le quali mostrano pubblicità o i titoli delle notizie in movimento. La percezione del movimento in questi display è così efficace che è difficile immaginare che queste scritte siano fatte da luci stazionarie che si accendono e si spengono.

Wertheimer trasse delle conclusioni a partire dal fenomeno del movimento apparente. Primo, il movimento apparente non può essere spiegato con sensazioni perché non c'è niente nello spazio scuro tra le due immagini che lampeggiano. Secondo, *il tutto è diverso dalla somma delle parti*, perché il sistema percettivo crea la percezione del movimento quando questo non c'è niente. Quest'idea che il tutto sia diverso dalla somma delle parti è diventata il grido di battaglia degli psicologi della Gestalt. Il "tutto era dentro, le "sensazioni" erano fuori!

**Contorni illusori** Un'altra dimostrazione che si argomenta contro le sensazioni e per l'idea che il tutto sia diverso dalla somma delle parti è mostrato in **Figura 5.16**. Questa dimostrazione include cerchi con una "bocca" tagliata, i quali ricordano le figure di Pac Man dal classico videogioco introdotto negli anni del 1980. Cominciamo dal Pac Man in **Figura 5.16a**. Potreste vedere un margine congiunge due "bocche" consecutive dei Pac Men, ma se coprite una delle due questo margine scompare. Questo singolo bordo diventa pare di un triangolo quando aggiungiamo un terzo Pac Man, **Figura 5.16b**. I tre Pac Men hanno creato la percezione di un triangolo, il quale diventa ancor più chiaro aggiungendo delle linee come in **Figura 5.16c**. I margini che creando il triangolo sono chiamati **contorni illusori** perché in realtà non ci sono dei bordi fisici presenti. Le sensazioni non possono spiegare i contorni illusori, perché non ce ne sono di presenti lungo i bordi. Questa dimostrazione fornisce ulteriori prove che il tutto è diverso dalla somma delle parti.

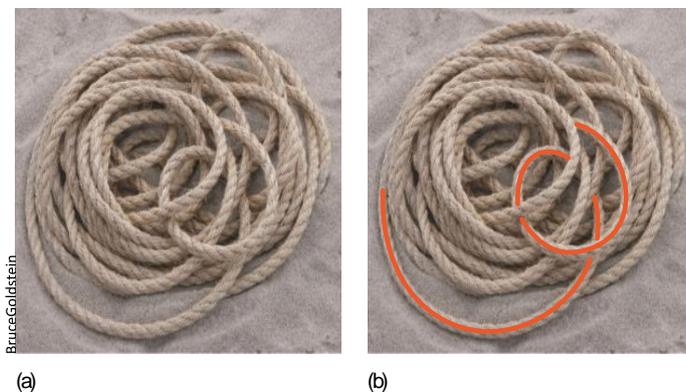


**Figura 5.16** I contorni illusori chiaramente visibili in (b) e in (c) non possono essere causati dalle sensazioni perché vi è solo del bianco.

## I principi organizzativi della Gestalt

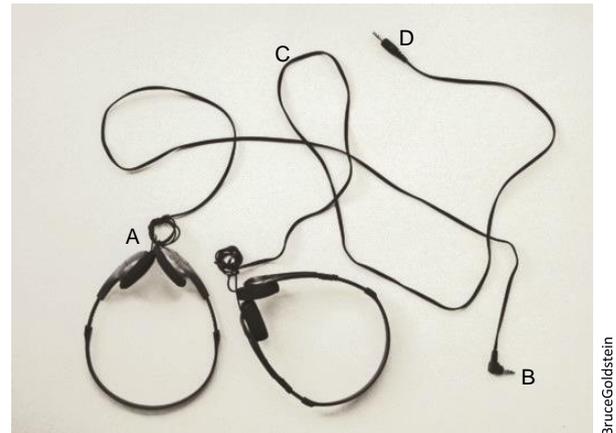
Avendo messo in discussione l'idea secondo la quale le percezioni sono create dal sommarsi di sensazioni, gli psicologi gestaltisti proposero che la percezione dipenda da un numero di **principi organizzativi** i quali determinano come gli elementi di una scena vengano raggruppati tra loro. I punti di partenza di questi principi vanno ricercati nell'ambiente circostante. Considerate, ad esempio, il modo in cui percepite la fune di **Figura 5.17a**. Nonostante ci siano molti posti in cui un filo si sovrappone ad un altro probabilmente percepite la fune non come un numero di pezzetti separati tra loro ma come una corda continua, come viene illustrato dal pezzo di corda evidenziato in **Figura 5.17b**.

Gli psicologi della Gestalt essendo degli osservatori acuti delle percezioni, usavano questa tipologia di osservazioni per formulare il *principio di buona continuazione*.

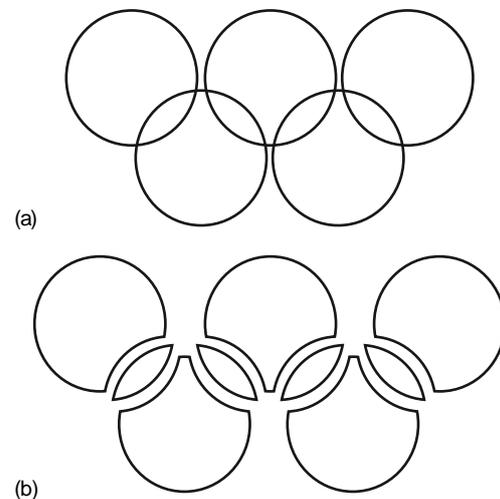


**Figura 5.17** (a) Una corda in spiaggia. (b) La buona continuazione ci aiuta a percepire la fune come un pezzo unico.

**La buona continuazione** Secondo il **principio di buona continuazione**, i punti, che quando sono connessi tra loro appaiono come una retta o come una linea leggermente curva, vengono visti come un tutt'uno e queste linee tendono a essere viste in modo tale da seguire il percorso più regolare. Questo principio funziona anche con le superfici: gli oggetti che sono parzialmente coperti da altri oggetti sono visti come continuativi dietro a quest'ultimi. La fune della Figura 5.17 mostra come degli oggetti nascosti siano visti come continuativi dietro all'oggetto che li copre. Il filo (**Figura 5.18**) che partendo dal punto A scorre fino al punto B è un esempio delle linee che percorrono il percorso più scorrevole. Il percorso che parte da A non giunge a C o a D perché questi percorsi violerebbero la buona continuazione avendo delle curve brusche.



**Figura 5.18** La buona continuazione ci aiuta a percepire due fili distinti, anche se questi sono sovrapposti.

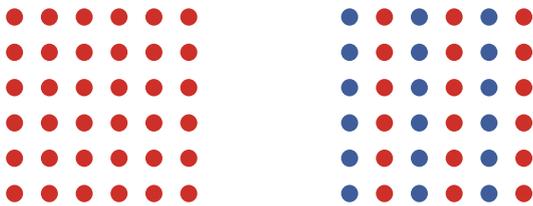


**Figura 5.19** (a) questa figura di solito è percepita come un'immagine contenente 5 circonferenze e non un'immagine con 9 figure (b)

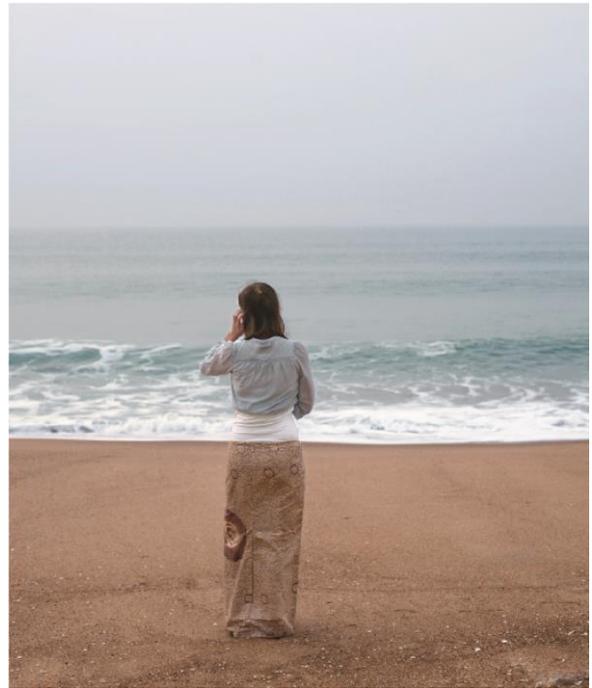
Qui ha una funzione utile anche la buona continuazione che crea la percezione di morbide linee curve che risultano circonferenze

**SOMIGLIANZA** la maggior parte delle persone percepisce la figura 5.20a o come righe orizzontali di cerchietti, colonne verticali di cerchietti o un quadrato riempito di puntini equidistanti fra loro ma, quando cambiamo il colore di alcune colonne come nella figura 5.20b, la maggior parte delle persone percepisce delle colonne verticali di cerchietti. Questo fenomeno di percezione illustra il principio di somiglianza per cui “cose simili vengono raggruppate insieme”. Questo principio spiega perché i cerchi dello stesso colore vengono raggruppati insieme. Un suggestivo esempio di raggruppamento per somiglianza è dimostrato nella figura 5.21; il raggruppamento può venire anche a causa della somiglianza di forma, dimensione o direzione. Il raggruppamento può verificarsi anche per stimoli uditivi per esempio, note che hanno tono simile e si susseguono in modo coeso nel tempo possono essere percepite in gruppo a formare una melodia. Considereremo questo fenomeno e altri effetti di raggruppamento uditivo quando descriveremo il processo organizzativo dell’udito nel capitolo 12.

**PROSSIMITÀ (VICINANZA)** la nostra percezione della figura 5.22 come 3 gruppi distinti di candele illustra il principio di prossimità o vicinanza: oggetti vicini tra loro vengono raggruppati tra loro.



**Figura 5.20** percepita come righe orizzontali, colonne verticali o entrambi. (b) percepita come colonne verticali



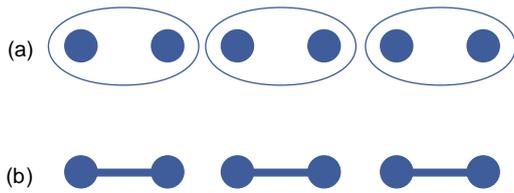
Courtesy of Wilma Hurskainen

**Figura 5.21** Questa fotografia, “Onde”, di Wilma Hurskainen, è stata scattata nel momento esatto in cui la costa del mare bianco è allineata all’area bianca del vestito della donna. La similarità dei colori causa il raggruppamento; le differenti aree colorate del vestito sono raggruppate percettivamente con gli stessi colori della scena. È da notare anche come il bordo frontale dell’acqua crei un raggruppamento grazie la buona continuazione attraverso il vestito della donna.



Bruce Goldstein

**Figura 5.22** le candele sono raggruppate per prossimità creando tre gruppi separati. Riesci a identificare altri principi della Gestalt nel modello della menorah?



**Figura 5.23** (a) raggruppamento per regione comune; (b) connessione

**DESTINO COMUNE** secondo il principio di destino comune gli oggetti che si muovono nella stessa direzione sembrano essere raggruppati insieme. Così come quando si vede uno stormo di centinaia di uccelli volare insieme, si tende a percepire lo stormo come un'unità e se, alcuni di questi uccelli cominciassero a volare in un'altra direzione, creerebbero una nuova unità. È da evidenziare che questo principio può presentarsi anche nel caso gli oggetti non si assomiglino, la chiave del destino comune è il fatto che un gruppo di oggetti si muova nella stessa direzione.

I principi appena descritti sono stati proposti dagli psicologi della Gestalt all' inizio del XX secolo, i principi aggiunti successivamente sono stati proposti dagli psicologi moderni della percezione

**ZONA COMUNE** figura 5.23a illustra il principio della zona comune per il quale elementi che sono vicini nella stessa zona sembrano come raggruppati tra loro. Nonostante i cerchi dentro gli ovali siano tra loro più lontani rispetto al cerchio interno all'ovale vicino, percepiamo i cerchi dentro all'ovale come appartenenti alla stessa zona; questo si verifica perché ogni ovale è visto come una regione di spazio separata (Palmer, 1992; Palmer e Rock, 1994). Bisogna prestare attenzione al fatto che in questo esempio la prossimità è sopraffatta dal fenomeno della zona comune perché per la prossimità cerchi vicini verrebbero percepiti insieme ma anche se i cerchi che sono in zone differenti sono più vicini tra loro, non si raggruppano fra loro come nella figura 5.22

**CONNESSIONE UNIFORME** secondo il principio della connessione comune una regione connessa di una stessa proprietà visiva come possono essere la luce, il colore, texture o moto è percepita come una singola unità (Palmer e Rock, 1994). Per esempio, nella figura 5.23b, i cerchi connessi sono percepiti come un gruppo unico come lo erano quando si trovavano nella stessa regione nella figura 5.23° e di nuovo la prossimità viene sopraffatta; questa volta dalla connessione.

I principi della Gestalt descritti in precedenza predicono cosa noi percepiremo in base a cosa succede solitamente nell'ambiente. Molti dei miei studenti reagiscono a quest'idea rispondendo che i principi della Gestalt non sono perciò nulla di speciale perché quello che questi principi fanno è descrivere le cose ovvie che vediamo ogni giorno. Quando lo dicono io gli ricordo che la ragione per cui noi percepiamo scene come gli edifici della città nella figura 5.2 o la scena nella figura 5.24 così facilmente è perché utilizziamo le osservazioni sulle proprietà dell'ambiente che si verificano normalmente per organizzare le scene.



BruceGoldstein

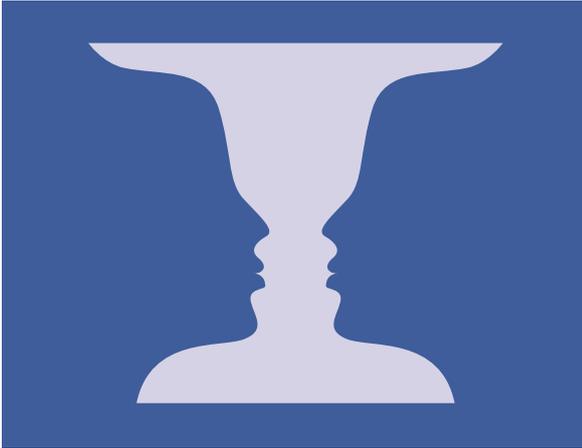
**Figura 5.24** un evento usuale nell'ambiente: gli oggetti (le gambe degli uomini) sono parzialmente nascosti da un altro oggetto (le assi grigie). In quest'esempio, le gambe continuano in linea dritta e sono dello stesso colore sopra e sotto le assi così è molto probabile che continuino dietro le assi.

Così assumiamo per vero, senza nemmeno pensarci su, che le gambe degli uomini nella figura 5.24 si estendano dietro le assi grigie, perché sappiamo che generalmente nell'ambiente quando due parti visibili di un oggetto (come le gambe degli uomini) hanno lo stesso colore e sono "allineate" esse appartengono allo stesso oggetto e si estendono dietro qualsiasi cosa che blocchi la visione della loro interezza. Le persone di solito non pensano a come noi percepiamo situazioni tipo queste come basate su assunzioni, ma è così che infatti succede. La ragione per cui l'assunzione sembra così ovvia è perché nell'ambiente abbiamo avuto molte esperienze con cose come questa. Il fatto che l'assunzione sia attualmente pressoché una cosa sicura potrebbe farci prendere per garantiti i principi della Gestalt etichettandoli come "ovvi" ma la realtà è che i principi della Gestalt non sono altro che la caratteristica operativa di base del nostro sistema visivo che determina come il nostro sistema percettivo organizza gli elementi dell'ambiente in unità più grandi.

## Segregazione Figura-Sfondo

Gli psicologi della Gestalt erano interessati anche a determinare le caratteristiche dell'ambiente responsabili della segregazione percettiva: la separazione percettiva di un oggetto da un altro, come si verifica quando si guardano gli edifici nella figura 5.1, e si percepiscono come separati gli uni dagli altri.

La questione di cosa causa la segregazione percettiva è spesso riferita al problema della segregazione della figura-sfondo. Quando noi vediamo un oggetto separato, di solito è visto come una figura che sta fuori dal retroscena (chiamato sfondo). Per esempio, probabilmente avrai visto un libro o dei fogli sulla scrivania come figure e la superficie della

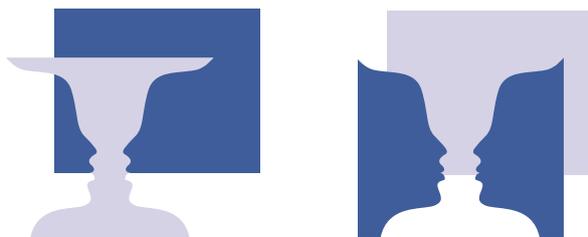


**Figura 5.25** una versione della figura reversibile sfondo-figura di Rubin

scrivania come lo sfondo. Gli psicologi della Gestalt erano interessati a determinare le proprietà della figura e dello sfondo e a capire cosa ci causa il percepire un'area come figura e l'altra come sfondo.

**PROPRIETÀ DI FIGURA E SFONDO** un modo in cui gli psicologi della Gestalt hanno studiato le proprietà della figura e dello sfondo è stato considerando modelli come quello nella figura 5.25 introdotto dallo psicologo danese Edgar Rubin nel 1915. Questo modello è un esempio di un'immagine figura-sfondo reversibile perché può essere percepita alternativamente sia come due facce blu scuro che si guardano tra loro davanti ad uno sfondo grigio, sia come un vaso grigio con uno sfondo blu scuro. Alcune delle proprietà della figura e dello sfondo sono:

- La figura è più riconoscibile dello sfondo. Così, quando si vede il vaso come figura, apparirà come un oggetto che sarà ricordato più tardi. Però, quando vedi la stessa area di luce come lo sfondo, non apparirà come oggetto ma solo come retroscena e perciò non particolarmente memorizzabile.
- La figura è vista come posta davanti allo sfondo. Così quando il vaso è visto come figura appare davanti ad uno sfondo scuro (**Figure 5.26**) e quando sono le facce viste come figure, esse sono sopra la luce dello sfondo.



(a)

(b)

(a) Quando il vaso è percepito come figura, è visto davanti ad un omogeneo sfondo scuro. (b) Quando le

facce sono viste come figure, sono viste davanti ad un omogeneo sfondo chiaro.

- Vicino ai confini che condivide con la figura, lo sfondo è visto come un materiale indefinito, senza una forma specifica e sembra si estenda dietro la figura. Questo non è per dire che gli sfondi mancano interamente di forma, essi sono spesso formati da bordi distanti da quelli che condividono con la figura. Per esempio, gli sfondi nella figura 5.26 sono quadrati.
- Il bordo che separa la figura dallo sfondo sembra appartenere alla figura. Si consideri per esempio la figura di Rubin nell'immagine 5.25: quando le due facce sono viste come figura, il bordo che separa le facce blu dallo sfondo grigio appartiene alle facce. Questa peculiarità del bordo di appartenere ad un'area è chiamata **proprietà del confine**. Quando la percezione si sposta a vedere il vaso come figura, così si sposta anche la proprietà di confine, e così ora il confine appartiene alla faccia.

**FATTORI BASATI SULL'IMMAGINE CHE DETERMINANO QUALE AREA È LA FIGURA.**

Gli psicologi della Gestalt specificarono un numero di fattori entro l'immagine che determinano che aree vengono percepite come figure. Quest'idea che l'informazione nell'immagine determini la percezione, è simile all'approccio che gli psicologi della Gestalt usarono per i raggruppamenti, nei quali i loro principi sono tutti riferiti a come le proprietà dell'immagine determinano quali elementi vengono visti come raggruppati insieme. Uno dei fattori basati sull'immagine proposto dai Gestaltisti è che è più probabile che aree più in basso nel campo visivo siano percepite come figure (Ehrenstein, 1930; Kofka, 1935).



(a)



(b)

Quest'idea è stata confermata sperimentalmente anni dopo da Shaun Vecera e collaboratori (2002) che fece lampeggiare

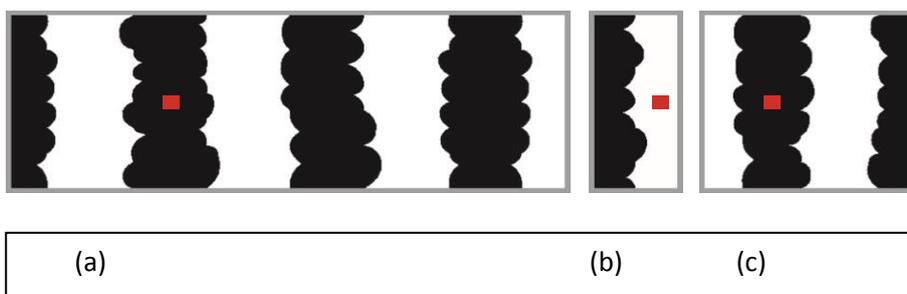


**Figura 5.28** la scena nella metà bassa del campo visivo è vista come figura. Il cielo nella parte alta del campo visivo è visto come sfondo

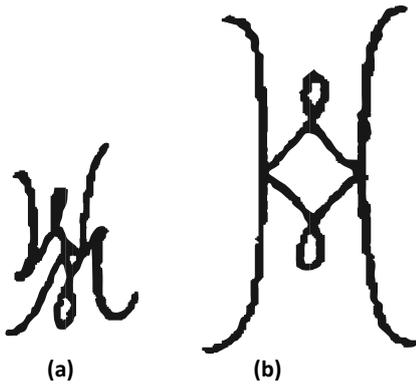
stimoli come quello della figura 5.27a per 150 ms e determinò che area venne vista come figura, se quella rossa o quella verde. I risultati che si vedono nella figura 5.27b indicano che nei display alto- basso, gli osservatori erano più propensi a percepire l'area in basso come figura, ma nel display destra-sinistra hanno mostrato una piccola preferenza per la regione sinistra. Per questi risultati, Vecera concluse che non c'è una preferenza nella destra o nella sinistra per determinare la figura ma c'è una spiccata preferenza nel vedere oggetti nella parte bassa dello schermo come figura. La conclusione di questo esperimento, cioè che la parte più bassa di un'area tende ad essere vista come figura, ha senso quando consideriamo una scena come nella figura 5.28 nella quale la parte bassa dell'immagine è la figura e il cielo è lo sfondo. Quello che è significativo di questa scena è che è una scena tipica che percepiamo quotidianamente. Nella nostra esperienza normale, la figura è molto più probabilmente appartenente all'orizzonte.

Un'altra proposta gestaltista riguardava il fatto che le figure erano più probabilmente percepite nel lato convesso dei confini (bordi che fuoriescono) (Kanizsa e Gerbino, 1976). Mary Peterson e Elizabeth Salvagio (2008) dimostrarono ciò presentando schermi come quello nella figura 5.29a e chiedendo agli osservatori di indicare se il quadratino rosso era percepito o meno come figura. Così, se loro avessero percepito l'area scura come figura avrebbero detto "sì".

Se invece avessero percepito l'area scura come sfondo avrebbero detto "no". Il risultato, d'accordo con la proposta gestaltista, fu che regioni convesse (come l'area scura nella figura 5.29a) venivano percepite come figure l'89% delle volte; ma Peterson e Salvagio andarono oltre la semplice conferma della proposta gestaltista presentando schermi come quello nella figura 5.29b e c con meno componenti. Compiendo questo grande decremento, la probabilità che schermi convessi venissero visti come figure, con la regione convessa contenente il quadratino rosso nello schermo a due componenti venisse vista come figura era solo del 58%. Secondo Peterson e Salvagio questo risultato significa che per capire come la segregazione si verifica dobbiamo andare oltre la semplice identificazione dei fattori come la convessità. Apparentemente, la segregazione è determinata non solo da cosa succede ad un singolo bordo ma da cosa succede nell'intera scena. Questo ha senso quando consideriamo che la percezione generalmente si verifica in scene che si estendono in una vasta area. Ritourneremo su quest'idea più tardi nel capitolo in cui consideriamo come percepiamo le scene.



**Fig. 5.29** stimoli dall'esperimento di Peterson e Salvagio (2008) (a) schermo a 8 componenti, (b) schermo a 2 componenti, (c) schermo a 4 componenti. Il quadratino rosso appare su diverse aree e diversi trial. La prova del partecipante era giudicare se il quadratino rosso era su una figura o sullo sfondo



**Fig. 5.30** (a) una "W" sopra a una "M". (b) Quando vengono ricombinate, emerge una nuova struttura inibendo le lettere significative

**FATTORI SOGGETTIVI CHE DETERMINANO QUALE AREA SIA LA FIGURA.** Ricordiamo che gli psicologi della Gestalt erano in disaccordo con l'idea strutturalista per cui le percezioni erano create sommando le sensazioni.

Erano inoltre in disaccordo con l'idea che l'esperienza passata di una persona abbia un ruolo importante nel determinare la percezione, quindi elementi visivi che sono stati raggruppati in precedenza sarebbero più probabilmente raggruppati insieme quando visti di nuovo. La seguente dimostrazione del gestaltista Max Wertheimer (1912) illustra come la psicologia gestaltista minimizzi l'esperienza. Wertheimer osserva che tendiamo a percepire il display nella figura 5.30 a come una "W" seduta sulla punta della "M", in gran parte per la nostra esperienza passata con queste due lettere. Però quando la W e la M sono posti come nella figura 5.30b vediamo due linee verticali con un motivo in mezzo. Nonostante noi possiamo dire dove si trovano la W e la M se guardiamo attentamente, la fantasia con le due linee verticali è la percezione predominante. Ritornando all'organizzazione dei principi della Gestalt, Wertheimer disse che le due verticali sono create dal principio della buona continuazione e che questo principio sovrasta qualsiasi effetto dell'esperienza passata che decade dall'aver visto la W o la M prima.

L'idea gestaltista che l'esperienza passata e il significato degli stimoli (come la M e la W) giochino un ruolo minore nell'organizzazione percettiva è illustrata anche dalla proposta della Gestalt che una delle prime cose che si verificano nel processo percettivo è la separazione della figura dallo sfondo. Sostengono che la figura debba emergere dallo sfondo per essere riconosciuta. In altre parole, la figura dev'essere separata dallo sfondo prima di riuscire ad assegnare un significato alla figura; ma Bradley Gibson e Mary Peterson (1994) fecero un esperimento che contestava quest'idea mostrando che la formazione figura-sfondo può essere influenzata dal significato di uno stimolo. L'hanno dimostrato presentando uno schermo come quello nella figura 5.31a che può essere percepito in due modi: (1) una donna in piedi (la parte nera dello schermo) o (2) una forma senza significato (la parte bianca dello schermo).



**Fig. 5.32** "la foresta ha gli occhi" di Bey Dolittle (1984). Riesci a trovare tutte e 13 i volti?



(a)

(b)

**Fig. 5.31** stimolo di Gibson e Peterson (1994). (a) la zona nera è più probabilmente percepita come figura perché ha

### DIMOSTRAZIONE

#### Trovare Facce in un Paesaggio

Consideriamo l'immagine nella **Figura 5.32**. A primo impatto, questa scena sembra contenere principalmente alberi, rocce, e acqua. Ma ad un'ispezione più accurata, potete vedere alcune facce tra gli alberi nello sfondo, e se guardate più da vicino, potete vedere che un certo numero di facce sono formate da diversi gruppi di rocce. Vedete se siete capaci tutte le 13 facce nascoste in quest'immagine.

Alcune persone trovano difficile percepire le facce in un primo momento, ma poi all'improvviso ci riescono. Il cambiamento in percezione da "rocce in un ruscello" o "alberi in una foresta" a "facce" è un cambiamento nell'organizzazione percettiva delle rocce e degli alberi. Le due forme che all'inizio percepite come due rocce separate nel ruscello diventano percettivamente raggruppate assieme quando esse diventano gli occhi destro e sinistro di una faccia. Infatti, una volta che percepite un particolare gruppo di rocce come una faccia, è spesso difficile *non* percepirle in questo modo -esse sono diventate permanentemente organizzate in una faccia. Questo è simile al processo che abbiamo osservato per il dalmata. Una volta che vediamo il dalmata, è difficile non percepirlo.

I principi e ricerca che abbiamo descritto si sono concentrate largamente su come la nostra percezione di singoli oggetti dipenda dall'organizzazione di principi e dai principi che determinano quali parti di un'immagine saranno viste come figure e quali saranno viste come sfondo. Se guardate indietro alle illustrazioni in questa sezione, noterete che molte di loro sono semplici immagini ideate per uno specifico principio di organizzazione percettiva.

Per capire però come veramente la percezione si manifesta nell'ambiente, dobbiamo espandere la nostra vista per considerare non solo oggetti individuali, ma anche scene più complesse.

### METTITI ALLA PROVA 5.1

1. Quali sono alcuni dei problemi che rendono la percezione degli oggetti difficile per i computer ma non per gli umani?
2. Cos'è lo strutturalismo, e perché la psicologia della Gestalt propose un'alternativa a questa definizione di percezione?
3. Come gli psicologi della Gestalt spiegano l'organizzazione percettiva?
4. Come hanno definito la segregazione figura-sfondo gli psicologi della Gestalt? Quali sono alcune proprietà di base di figura e di sfondo?
5. Quali proprietà basate su immagini di uno stimolo tendono a favorire la percezione di un'area come "figura"? Sii sicuro di comprendere l'esperimento di Vacera il quale ha mostrato che la regione più bassa di un'immagine tende ad essere percepita come figura, e perché Peterson e Salvagio hanno affermato che per comprendere come avviene la segregazione dobbiamo considerare cosa sta succedendo nell'intera scena.
6. Descrivi le idee della Gestalt riguardo al ruolo del significato e dell'esperienza passata nel determinare la segregazione figura-sfondo.
7. Descrivi l'esperimento di Gibson e Peterson che ha mostrato che il significato può giocare un ruolo nella segregazione figura-sfondo.
8. Cosa dimostra la scena Bev Doolittle nella Figura 5.32?

## Percepire Scene e Oggetti nelle Scene

La nostra discussione di organizzazione e figura-sfondo ha descritto come la nostra percezione è influenzata da caratteristiche come vicinanza, buona continuazione, e somiglianza; più in alto o più in basso nel campo visivo; convessità o concavità dei margini. Ma alla fine di questa discussione, abbiamo anche notato che il *significato* di uno stimolo può influenzare entrambe le formazioni figura-sfondo (esperimento di Gibson e Peterson) e la nostra percezione degli oggetti in una scena (la dimostrazione "Trova le Facce"). Il significato ora prende il centro del palcoscenico nella nostra discussione, in quanto descriviamo la ricerca moderna su come gli osservatori percepiscono oggetti e scene. Una *scena* è una visione di un ambiente del mondo reale che contiene (1) elementi di fondo e (2) molteplici oggetti che sono organizzati in modo casuale rispetto agli altri e lo sfondo (Epstein, 2005; Henderson & Hollingworth, 1999). Un modo di distinzione tra oggetti e scene è che gli oggetti sono definiti e sono "agiti su", mentre le scene sono estese nello spazio e

sono "agite dentro". Per esempio, se noi stiamo camminando per la strada e spediamo una lettera, noi staremmo *agendo sulla* cassetta delle lettere (un oggetto), e *agendo all'interno* della strada (la scena).

## Percepire l'Essenza di una Scena

La percezione delle scene presenta un paradosso. In una mano, le scene sono spesso larghe e complesse. Ad ogni modo, nonostante questa taglia e complessità, potete identificare molte scene dopo averle visualizzate per solo una frazione di secondo. Questa descrizione generale del tipo di scena è chiamata *l'essenza di una scena*. Un esempio delle vostre abilità per percepire rapidamente l'essenza di una scena è il caso in cui potete rapidamente passare da un canale ad un altro, anche senza comprendere il significato di ogni immagine poiché balena via -inseguimento d'auto, contendenti in un quiz, o una scena d'esterno con le montagne- nonostante potreste vedere ogni immagine per un secondo o meno e quindi non sareste in grado di identificare oggetti specifici. Quando fate ciò, state percependo l'essenza di ogni scena (Olivia & Torralba, 2006).

Ricerche hanno dimostrato come sia possibile percepire l'essenza di una scena in una frazione di secondo. Mary Potter (1976) mostrò un'immagine target agli osservatori e poi chiese loro di indicare se avessero visto quell'immagine dato che loro avevano visto una sequenza di 16 immagini presentate rapidamente. I suoi osservatori riuscirono a farlo con un'accuratezza prossima al 100% anche quando le immagini vennero presentate per soli 250 ms (millisecondi; ¼ di secondo). Anche quando l'immagine target era specificata solamente da una descrizione scritta, come ad esempio "ragazza che applaude", gli osservatori registrarono un'accuratezza circa del 90% (Figura 5.33).

Un altro approccio per determinare quanto rapidamente le persone possono percepire le scene fu usato da Li Fei-Fei e colleghi (2007), i quali presentarono immagini di scene con un range di esposizione da 27 ms a 500 ms e chiesero agli osservatori di scrivere una descrizione di cosa avevano visto. Questo metodo di determinare il responso degli osservatori è un buon esempio del metodo

fenomenologico, descritto a pagina 17. Fei-Fei usò una procedura chiamata "*masking*" (o *mascheramento* in italiano) per essere sicura che gli osservatori vedessero l'immagine esattamente per la durata desiderata.

### METODO

#### Usare una Maschera per Ottenere Brevi Presentazioni dello Stimolo

Come bisogna fare se vogliamo presentare uno stimolo che sia visibile per soli 100 ms? Nonostante si possa pensare che il modo per farlo potrebbe essere proiettare uno stimolo per 100 ms, ciò non funzionerà per via di un fenomeno chiamato *persistenza della visione*- la percezione di uno stimolo visivo continua per circa 250 ms (1/4 di secondo) dopo che lo stimolo si è estinto. Quindi, un'immagine che viene presentata per 100 ms sarà *percepita* all'incirca per 350 ms. Ma la persistenza di visione può essere eliminata presentando uno *stimolo di mascheramento visivo*, solitamente un campione casuale che compre lo stimolo originale, così se un'immagine è proiettata per 100 ms seguita immediatamente da uno stimolo di mascheramento, l'immagine è visibile per soli 100 ms. Uno stimolo di mascheramento è dunque spesso presentato immediatamente dopo uno stimolo del test per bloccare la persistenza della visione dall'incremento della durata dello stimolo del test.

Tipici risultati dell'esperimento di Fei-Fei sono mostrati nella **Figura 5.34**. A brevi durate, gli osservatori videro solo aree chiare e scure dell'immagine. Da 67 ms poterono identificare alcuni oggetti larghi (una persona, un tavolo), e quando la durata fu alzata a 500 ms (mezzo secondo) furono capaci di identificare oggetti più piccoli e dettagli (il ragazzo, il portatile). Per l'immagine di un ornato salotto dell'Ottocento, gli osservatori furono capaci a 67 ms di identificare l'immagine come una stanza di una casa e a 500 ms di identificare dettagli, come sedie e ritratti. Quindi, l'essenza complessiva della scena è percepita prima, seguita dalla percezione di dettagli e oggetti più piccoli nella scena

### Ragazza che applaude



Descrizione

250 ms



250 ms



250 ms

**Figura 5.33** Procedura per l'esperimento di Potter (1976). Lei prima presentò sia una foto target o, come mostrato qui, una descrizione, e poi presentate rapidamente 16 immagini per 250 ms ciascuna. Il compito degli osservatori era di indicare se l'immagine target fosse presente. In questo esempio, vengono mostrate solo 3 delle 16 immagini, con la seconda immagine presentata come immagine target. In altre prove, l'immagine target non è compresa nella serie di 16 immagini.



- 27 ms    Sembra che qualcosa di nero nel centro con quattro linee dritte esca da esso in contrasto con uno sfondo bianco.  
(Soggetto: AM)
- 40 ms    La prima cosa che potei riconoscere era una macchia scura al centro. Poteva essere di forma rettangolare, con il lato superiore curvo, ma è solo un'ipotesi.  
(Soggetto: KM)
- 67 ms    Una persona, credo, seduta o rannicchiata. Affacciati alla parte sinistra dell'immagine. Vediamo principalmente il loro profilo. Erano ad un tavolo o dove alcuni oggetti erano di fronte a loro (alla loro parte sinistra nell'immagine).  
(Soggetto: EC)
- 500 ms    Questo sembra un padre o qualcun altro che aiuta un ragazzino. L'uomo aveva qualcosa nelle sue mani, come uno schermo LCD o un portatile. Sembravano stare in un cubicolo.  
(Soggetto: WC)

**Figura 5.34** La descrizione degli osservatori di un fotografo presentato nell'esperimento di Fei Fei (2007). I tempi di visualizzazione sono indicati sulla sinistra.

Cosa permette agli osservatori di percepire così rapidamente l'essenza di una scena? Aude Oliva e Antonio Torralba (2001, 2006) propongono che gli osservatori usino un'informazione chiamata **rappresentazioni globali dell'immagine**, le quali possono essere percepite rapidamente e sono associate con specifici tipi di scene. Alcune delle rappresentazioni globali dell'immagine proposte da Oliva e Torralba sono:

- *Grado di naturalezza.* Scenari naturali, come l'oceano e la foresta in **Figura 5.35**, hanno zone strutturate e contorni ondulati. Scenari create dall'uomo, come la strada, sono dominati da linee dritte, orizzontali e verticali.
- *Grado di apertura.* Scenari all'aperto, come l'oceano, spesso hanno una linea d'orizzonte visibile e contengono pochi oggetti. Anche lo scenario della strada è all'aperto, anche se non quanto lo scenario dell'oceano. La foresta è un

esempio di uno scenario con un basso grado d'apertura.

- *Grado di ruvidezza.* Scenari lisci (a bassa ruvidezza) come l'oceano contengono pochi piccoli elementi. Scenari con alta ruvidezza come la foresta contengono molti piccoli elementi e sono più complessi.
- *Grado di espansione.* La convergenza di linee parallele, come quelle che vedi quando guardi le rotaie che appaiono per svanire nella distanza, o nello scenario stradale in Figura 5.35, indica un alto grado di espansione. Questa caratteristica è dipendente specialmente dal punto di vista dell'osservatore. Per esempio, nello scenario stradale, guardando direttamente ad un lato di un edificio potrebbe risultare in bassa espansione.
- *Colore.* Alcuni scenari hanno colori caratteristici, come lo scenario dell'oceano (blu) e la foresta (verde e marrone) (Goffaux et al., 2005).

Le rappresentazioni globali dell'immagine sono *olistiche e rapidamente percepite*. Sono proprietà delle scene come un tutt'uno e non dipendono dal processo di consumo di tempo come percepire piccoli dettagli, riconoscere oggetti individuali, o separare un oggetto dall'altro. Un'altra proprietà delle rappresentazioni



globali dell'immagine è che esse contengono informazioni riguardo l'organizzazione di una scena e la disposizione spaziale. Per esempio, il grado di apertura e il grado di espansione riferiscono direttamente a caratteristiche dell'organizzazione di una scena, e anche la naturalezza riferisce informazione sull'organizzazione che derivano dalla consapevolezza se una scena è "naturale" o contiene "strutture create dall'uomo".

Le proprietà globali dell'immagine non solo aiutano a spiegare come noi possiamo percepire l'essenza della scena basata su caratteristiche che possono essere viste in una breve esposizione, esse illustrano anche la seguente proprietà generale della percezione: le nostre esperienze passate nella percezione di proprietà dell'ambiente giocano un ruolo nel determinare le nostre percezioni. Impariamo, per esempio, che il blu è associato con il cielo aperto, che i paesaggi sono spesso verdi e morbidi, e che verticali e orizzontali sono associate con gli edifici. Caratteristiche dell'ambiente come queste, che ricorrono frequentemente, sono chiamate **regolarità nell'ambiente**. Ora descriveremo queste regolarità in dettaglio.

### Regolarità nell'ambiente: Informazioni per Percepire

Sebbene le persone facciano uso di regolarità nell'ambiente per aiutarsi a percepire, spesso sono inconsapevoli della specifica informazione che stanno usando. Questo aspetto di percezione è simile a ciò che succede quando utilizziamo il linguaggio. Anche se le persone legano facilmente le parole assieme per formare una frase nelle conversazioni, esse potrebbero non sapere le regole di grammatica che specificano come queste parole vengano combinate. Similmente, noi usiamo facilmente la nostra conoscenza di regolarità nell'ambiente per aiutarci a percepire, nonostante possiamo non essere in grado di identificare l'informazione specifica che stiamo usando.

Possiamo distinguere due tipi di regolarità: **regolarità fisiche** e **regolarità semantiche**.

**Figura 5.35** Tre scene che hanno diverse proprietà globali dell'immagine. Vedi il testo per la descrizione.



**Figura 5.36** Gli oggetti nell'ambiente hanno colori omogenei. Vedi il testo per i dettagli

**Regolarità Fisiche** Le **regolarità fisiche** sono verificate regolarmente come proprietà fisiche dell'ambiente. Per esempio, ci sono più orientamenti verticali e orizzontali nell'ambiente rispetto a orientamenti obliqui. Ciò si verifica negli ambienti creati dall'uomo (per esempio, gli edifici contengono molte orizzontali e verticali) e anche negli ambienti naturali (alberi e piante sono probabilmente più verticali e orizzontali che inclinate) (Coppola et al., 1998). Non ci sono, perciò, coincidenze se le persone riescono a percepire orizzontali e verticali più facilmente degli altri orientamenti- l'effetto obliquo che abbiamo introdotto nel Capitolo 1 (v. pagina 11) (Appelle, 1972; Campbell et al., 1966; Orban et al., 1984).

Un'altra regolarità fisica è che gli oggetti nell'ambiente hanno spesso colori omogenei e gli oggetti circostanti hanno colori differenti. Così, se prendiamo un punto nella scena della **Figura 5.36** (come A) e ci allontaniamo leggermente da questo punto (su B), è probabile che i due punti siano sullo

stesso oggetto se il colore è lo stesso. Se, però, ci spostiamo su C, il colore cambia, il che significa che è probabile che stiamo guardando un oggetto diverso. Mentre guardate questa scena, vedete se siete capaci di trovare esempi di buona continuazione e buona figura.

La seguente dimostrazione illustra ancora un'altra irregolarità fisica.

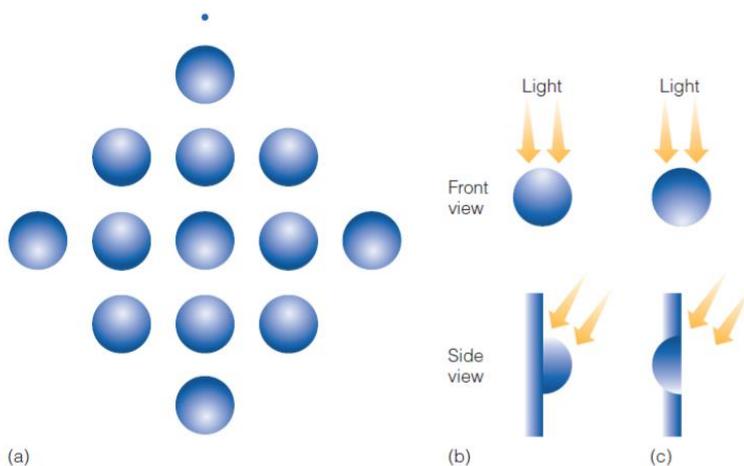
---

### DIMOSTRAZIONE

#### Forma Dall'Ombreggiatura

Cosa percepite nella **Figura 5.37a**? Alcuni dei dischi sembrano sporgere, come parti di sfere tridimensionali, e altre appaiono come incavate? Se vedete i dischi in questo modo, notate che quelli che appaiono sporgenti sono disposti in un quadrato. Dopo aver osservato ciò, capovolgete la pagina in modo che il puntino sia in basso. Questo cambia la vostra percezione?

Le Figure 5.37 b e c mostrano che se noi presupponiamo che la luce venga dall'alto (il che è il caso solito nell'ambiente), poi modelli come i cerchi che sono chiari sulla parte superiore verrebbero creati da un oggetto che sporge (**Figura 5.37b**), ma un modello come i cerchi illuminati sul fondo verrebbero creati da un'incavatura nella superficie (**Figura 5.37c**). L'assunzione che la luce venga dall'alto venne chiamata **l'assunzione della luce-dall'alto** (Kleffner & Ramachandran, 1992). Apparentemente, le persone ricorrono all'assunzione della luce-dall'alto perché il più della luce nel nostro ambiente viene dall'alto. Questo include il sole, come anche la maggior parte delle luci artificiali.



**Figura 5.37** (a) Alcuni di questi dischi sono percepiti come sporgenti e alcuni sono percepiti incavati. (b) Spiegazione del principio di spiegazione di (a). La luce proveniente dall'alto illuminerà la cima di una forma che sporge e (c) il fondo di un'incavatura.



(a)



(b)

**Figura 5.38** Perché in (a) sembrano esserci delle rientranze nella sabbia e (b), invece, assomiglia a cumuli di sabbia? Guarda il testo per la spiegazione.

Un altro esempio di ipotesi di luce dall'alto viene fornita da due immagini nella Figura 5.38.

La Figura 5.38a mostra delle rientranze create da persone camminando sulla sabbia. Ma quando noi giriamo l'immagine al contrario, come nella Figura 5.38b, poi le rientranze nella sabbia diventano cumuli di sabbia.

È chiaro dagli esempi di regolarità fisiche nell'ambiente che una delle ragioni per cui gli umani sono in grado di percepire e riconoscere oggetti e scene molto meglio dei robot guidati da computer, è che il nostro sistema è personalizzato per rispondere alle caratteristiche fisiche del nostro ambiente. Ma questa personalizzazione va oltre le caratteristiche fisiche.

Si verifica anche perché noi abbiamo imparato che tipo di oggetti tipicamente servono in specifici tipi di scene.

### Regolarità semantiche

Nel linguaggio, semantico si riferisce al significato delle parole o frasi. Applicato alle percezioni delle scene, semantico di riferisce al significato di una scena,

Questo significato è spesso relazionato alla funzione di una scena, cosa succede senza questa.

Per esempio, la preparazione del cibo, cucinare, e forse mangiare avviene in cucina; attendere, comprare biglietti, controllare la valigia, e attraversare i controlli di sicurezza accade in aeroporto.

Le regolarità semantiche sono delle caratteristiche associate con le funzioni svolte in differenti tipi di scene.

Un modo per dimostrare che le persone sono consapevoli delle regolarità semantiche è semplicemente quello di chiedere loro di immaginare un particolare tipo di scena o oggetto, come nella seguente dimostrazione.

La maggior parte delle persone che sono cresciute in una società moderna hanno piccoli problemi nella visualizzazione di un ufficio o di una sezione di un grande magazzino.

Quello che è importante riguardo le abilità, per i nostri scopi, è che la parte di questa visualizzazione implica dettagli all'interno di queste scene.

La maggior parte delle persone vedono un ufficio come l'avere una scrivania con il computer sopra, librerie, e una sedia.

La scena del grande magazzino dovrebbe contenere scaffali di vestiti, uno spogliatoio e forse un registratore di cassa. Cosa hai visto quando hai visualizzato il microscopio o il leone? Molte persone affermano che non vedono solo un singolo oggetto, ma un oggetto all'interno di uno scenario. Forse tu percepisci il microscopio posizionato su un banco di laboratorio o in un laboratorio e il leone nella foresta o in una savana o in uno zoo.

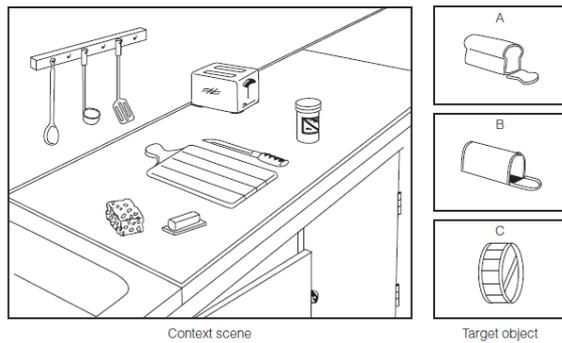
Un esempio di conoscenza che noi abbiamo delle cose che tipicamente appartengono ad alcune scene è fornito da un classico esperimento di Stephen Palmer (1975), usando stimoli come l'immagine nella Figura 5.39. Prima di tutto Palmer presentava un contesto di scena come quella a sinistra e poi brevemente ha fatto lampeggiare una delle immagini bersaglio a destra.

Quando Palmer chiedeva agli osservatori di identificare l'oggetto nell'immagine bersaglio, loro correttamente identificavano un oggetto come un pezzo di pane (il quale è appropriato alla scena della cucina) l'80% delle volte, ma correttamente identificavano la cassetta delle lettere o il tamburo (due oggetti che non rientrano nella scena) solo il 40% delle volte.

Apparentemente, gli osservatori di Palmer stavano usando la loro conoscenza delle cucine per aiutare loro a percepire in un breve flash il pezzo di pane.

L'effetto delle regolarità semantiche è anche illustrato nella Figura 5.40, la quale viene chiamata "le personalità multiple di una macchia" (Oliva & Torralba, 2007). La macchia è percepita come differenti oggetti a seconda del suo orientamento e del contesto all'interno del quale è vista. Sembra essere un oggetto su un tavolo in (b), una scarpa su una persona che si china in (c), e una macchina e una persona attraversare la strada in (d), anche se ha la stessa forma in tutte le immagini.

## Il ruolo dell'inferenza nella percezione



Le persone usano la loro conoscenza della regolarità fisica e semantica, come ad esempio il primo che noi abbiamo descritto, per inferire cosa è presente in un luogo.

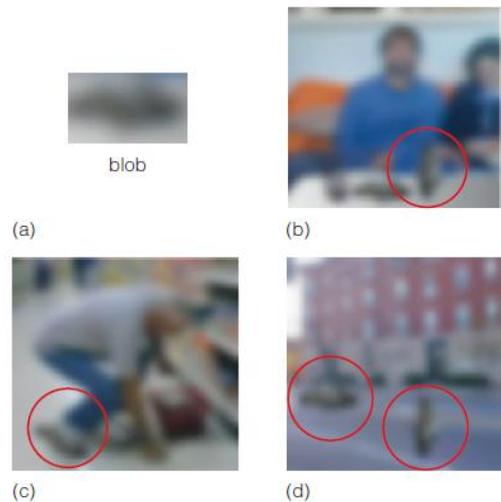
L'idea che la percezione comporti inferenze non è nuova; è stata proposta nel diciottesimo secolo da Hermann von Helmholtz (1866/1911), il quale era uno dei preminenti psicologi e fisici dei suoi giorni.

Helmholtz ha fatto molte scoperte in psicologia e fisica, ha sviluppato l'oftalmoscopio (il dispositivo che un optometrista o oftalmologo usa per guardare dentro gli occhi), e ha proposto teorie sulla percezione di oggetti, visione a colori, e udito.

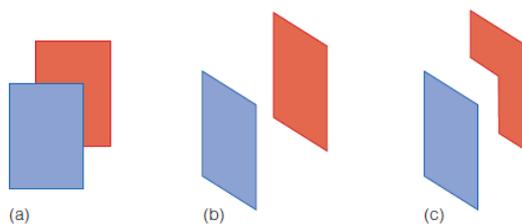
Una delle sue proposte riguardo la percezione è un principio chiamato la "teoria dell'inferenza inconscia", la quale afferma che alcune delle nostre percezioni sono il risultato di supposizioni inconscie che noi facciamo riguardo l'ambiente.

**Figura 5.39** Stimoli usati nell'esperimento di Palmer (1975).

La scena a sinistra è presentata prima, poi si chiede all'osservatore di identificare uno degli oggetti a destra.



**Figura 5.40** Cosa ci aspettiamo di vedere in differenti contesti che influiscono sulla nostra interpretazione dell'identità della macchia dentro ai cerchi.



**Figura 5.41** La visualizzazione in (a) è di solito interpretata come (b) un rettangolo blu davanti un rettangolo rosso. Tuttavia, potrebbe essere (c) un rettangolo blu e una disposizione adeguata di una figura rossa a sei lati.

La teoria dell'inferenza inconscia era stata proposta per spiegare la nostra capacità di creare percezioni da informazioni di stimoli che possono essere viste in più di un modo.

Per esempio, cosa vedete nel display della **Figura 5.41a**?

Più persone percepiscono un rettangolo blu davanti ad un rettangolo rosso, come mostrato nella **Figura 5.41b**.

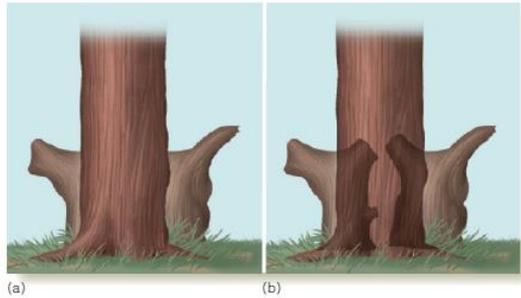
Ma come indicato nella **Figura 5.41c**, questa visualizzazione potrebbe essere stata causata da una forma a sei lati rossa posizionata o di fronte o dietro il rettangolo blu.

Secondo la teoria dell'inferenza inconscia, noi deduciamo che la **Figura 5.41a** sia un rettangolo che copre un altro rettangolo a causa delle esperienze che noi abbiamo fatto con situazioni simili in passato

Un corollario della teoria dell'inferenza inconscia è il "principio della probabilità". Afferma che noi percepiamo l'oggetto che ha causato molto probabilmente lo schema di stimoli che noi abbiamo ricevuto.

Quindi, noi percepiamo la **Figura 5.41a** come un rettangolo blu davanti un rettangolo rosso perché molto probabilmente, sulla base della nostra esperienza passata, hanno causato quello schema.

Un motivo per cui Helmholtz ha proposto il principio della probabilità è per far fronte all'ambiguità degli stimoli percettivi che noi abbiamo descritto all'inizio del capitolo. Helmholtz ha considerato il processo di percezione come simile al processo coinvolto nella risoluzione di un problema



**Figura 5.42** (a) Cosa si nasconde dietro l'albero? Sono due forme strane di tronchi di albero.

Per la percezione, il compito è determinare quale oggetto ha causato un particolare schema di stimolazione e questo problema viene risolto da un processo, dove l'osservatore porta a carico la sua conoscenza di ambiente al fine di dedurre che cosa può essere l'oggetto.

Consideriamo, per esempio, la seguente situazione: se voi state facendo un'escursione nel bosco, vi fermate di punto in bianco sulle vostre orme perché non troppo lontano davanti voi vedete quello che sembra essere un animale nascosto dietro un albero (Figure 5.42a). Questa conclusione è basata su un numero di fattori. Guardando le forme a sinistra e destra dell'albero come un singolo oggetto si desume dall'organizzazione Gestalt, i principi di similarità (entrambe le forme sono dello stesso colore quindi sembra che loro siano parte dello stesso oggetto) e una buona continuazione (la linea attraverso la cima dell'oggetto si estende senza difficoltà da un lato dell'albero all'altro).

Inoltre, l'immagine ricorda gli animali che avete visto prima. Dato che temete che l'animale possa essere pericoloso, voi prendete un altro percorso.

Come la vostra deviazione vi porta intorno all'albero, avete notato che le forme scure non sono un animale dopo tutto, ma sono due tronchi d'albero con una forma strana.

Quindi in questo caso, il risultato del vostro processo inferenziale vi ha lasciato fuori rotta.

Questo non dovrebbe sorprendere, perché non tutte le inferenze sono vere. Ma perché le inferenze percettive si basano sul nostro vasto magazzino di informazioni riguardo cosa solitamente avviene nell'ambiente, loro hanno ragione il più delle volte.

(E cosa succederebbe se l'oggetto dietro l'albero fosse una creatura non amichevole? Forse giocare sul sicuro è la miglior politica in questo caso.)

L'idea che l'inferenza sia coinvolta nella percezione è simile all'idea che la percezione implichi fare delle supposizioni, che noi abbiamo discusso in merito ai principi della Gestalt.

L'idea che le supposizioni e le inferenze siano importanti per la percezione ha avuto una ricaduta nel corso della ricerca della storia sotto varie forme, da Helmholtz ai principi della Gestalt, alle regolarità dell'ambiente.

Molto recentemente, la psicologia moderna ha quantificato l'idea della percezione inferenziale usando una tecnica statistica chiamata "inferenza Bayesiana" che tiene in considerazione delle probabilità (Kersten et al., 2004; Yuille & Kersten, 2006).

Per esempio, supponiamo che noi vogliamo determinare quanta probabilità c'è che poverà domani. Se noi sappiamo che oggi ha piovuto, questo aumenta le possibilità che poverà domani, perché se piove un giorno è più probabile che piova anche il giorno successivo.

Sulla base di questo ragionamento sulla percezione, noi possiamo chiederci, per esempio, se un dato oggetto in una cucina sia un pezzo di pane o una cassetta della posta. Dal momento che è più probabile che un pezzo di pane sarà in cucina, il sistema percettivo conclude che il pane è presente. Le statistiche Bayesiane prevedono questo tipo di ragionamento, espresso in formule matematiche che noi non vogliamo descrivere qui (vedi Geisler, 2008, 2011; Kersten et al., 2004).

## Collegamento di attività neurale e percezione di oggetti

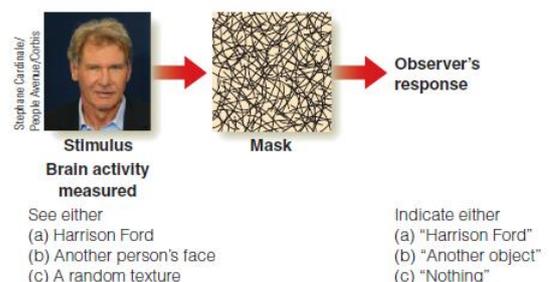
Finora nella nostra discussione di oggetti e luoghi, noi ci siamo focalizzati su come la percezione è determinata da aspetti di stimoli.

In effetti, le parole "neurone" e "cervello" non sono apparse nemmeno una volta! Ora è il momento di considerare la relazione tra i processi psicologici e la percezione di oggetti. Questa relazione è stata studiata in un numero di modi diversi, sia in animali (soprattutto scimmie) sia in esseri umani.

## Attività cerebrale e identificazione di un'immagine

Nel primo esperimento che noi descriveremo, Kalanit Grill-Spector e collaboratori (2004) erano interessati nel determinare la relazione tra attivazione del cervello che si verifica quando si guarda un oggetto e un'abilità di una persona di identificare un oggetto. L'"oggetto" che loro usavano erano immagini della faccia di Harrison Ford. Loro presentavano immagini, come mostrato nella Figura 5.43, all'osservatore in uno scanner cerebrale misurando le risposte dell'area facciale fusiforme (FFA) in un lobo temporale per ciascuna immagine.

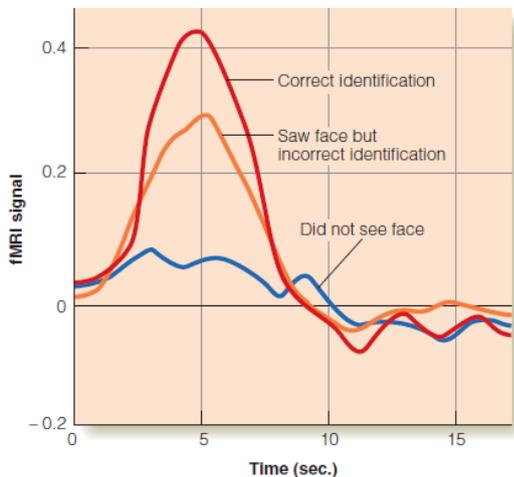
Per ogni prova, gli osservatori guardavano sia (a) una foto di Harrison Ford, (b) una foto di un'altra faccia di una persona, o (c) un texture casuale



**Figura 5.43** Procedura per l'esperimento Grill-Spector et al. (2004).

è importante perché mostra come il cervello reagisce ad uno stimolo come viene presentato determinando le nostre abilità di identificare lo stimolo.

Ognuno di questi stimoli era presentato brevemente (circa 50 ms), seguito immediatamente da un modello casuale di maschera, la cui visibilità è limitata ad ogni stimolo per 50 ms (vedi Metodo: Usare una maschera per ottenere brevi presentazioni di stimoli, pagina 109). Limitare la visibilità a 50ms fa gli stimoli difficili da identificare, quindi il compito dell'osservatore, è di indicare se per ogni foto era "Harrison Ford", "un altro oggetto", o "niente", non era facile. La Figura 5.44 mostra il corso dell'attivazione del cervello per le prove in cui la faccia di Harrison Ford era presente. La curva superiore (rossa) mostra che l'attivazione era maggiore quando l'osservatore identificava correttamente la foto come la faccia di Harrison Ford. La curva successiva mostra che l'attivazione era minore quando loro rispondevano "un altro oggetto" con la faccia di Harrison Ford. In questo caso, loro hanno rilevato la foto come una faccia ma non erano in grado di identificarla come la faccia di Harrison Ford.



**Figure 5.44** I risultati dell'esperimento di Grill-Spector et al. (2004) per le prove nelle quali la faccia di Harrison Ford era presentata.

Ricordiamo che tutte le curve nella Figura 5.44 rappresentano l'attività del cervello che avveniva durante la presentazione della faccia di Harrison Ford. Questi risultati quindi mostrano che l'attività neurale che si verifica, quando una persona sta guardando uno stimolo, è collegata alla capacità della persona di identificare lo stimolo. Un'elevata risposta neurale è associata con il trattamento che risulta nell'abilità di identificare lo stimolo; una minor risposta, con la rilevazione degli stimoli; e un'assenza di risposta con una mancanza complessiva dello stimolo. Questo

## Attività cerebrale e vista

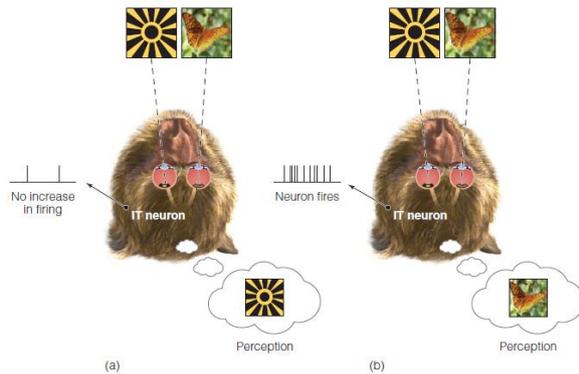
L'esperimento di Harrison Ford ha presentato stimoli che venivano mostrati velocemente, quindi erano difficili da vedere. Un altro metodo per studiare la relazione tra l'attività cerebrale e la vista è da cercare per collegamenti tra l'attività del cervello e stimoli che sono facili da vedere. Noi descriveremo due esperimenti, uno riguardo una scimmia e uno riguardo un uomo, entrambi fanno uso di una tecnica nella quale differenti foto sono presentate agli occhi, sinistro e destro.

Nella percezione quotidiana, i nostri due occhi ricevono leggermente differenti immagini perché gli occhi si trovano in due punti un po' diversi.

Queste due immagini sono, tuttavia, abbastanza simili così possono essere combinate in una percezione singola dal cervello. (Noi descriveremo questo processo per unire queste foto, che si chiama "fusione binoculare", nel capitolo 10, quando affrontiamo nel dettaglio la percezione.)

Ma se ogni occhio riceve immagini totalmente diverse, il cervello non può fondere le due immagini e si verifica una condizione chiamata "rivalità binoculare", dove l'osservatore percepisce o l'immagine dell'occhio destro o quella dell'occhio sinistro, ma non entrambi nello stesso momento. D.L. Sheinberg e Nikos Logothetis (1997) hanno usato questo principio, presentano il modello del raggio di sole all'occhio sinistro della scimmia e nello stesso momento presentano una foto di una farfalla all'occhio destro della scimmia (Figura 5.45). Presentare differenti foto a ciascun occhio causa la visione di un raggio di sole per parte del tempo alla scimmia e per parte del tempo alla farfalla, ma non insieme. Per determinare cosa stesse ricevendo la scimmia, loro la addestrarono a premere una leva quando percepiva il modello del raggio di sole e un'altra leva quando percepiva la farfalla. Come la scimmia faceva rapporto di cosa stava ricevendo, gli sperimentatori al contempo registravano le attività di un neurone nella corteccia della circonvoluzione temporale inferiore (IT) che in precedenza era stata mostrata per rispondere alla farfalla ma non al raggio di sole.

Il risultato dell'esperimento è indicato dallo schema di attivazione nella Figura 5.45: ogni qualvolta la scimmia percepiva il raggio di sole, la velocità di attivazione del neurone era bassa, ma quando la percezione della scimmia si trasferiva alla farfalla, l'attivazione aumentava.



**Figura 5.45** Nell'esperimento della fusione binoculare di Sheinberg e Logothetis lo stimolo "raggio di sole" era presentato all'occhio sinistro della scimmia, e la farfalla era presentata all'occhio destro. Entrambi gli occhi venivano stimolati nello stesso momento, ma a causa della fusione binoculare la scimmia percepiva o il raggio di sole o la farfalla in un dato momento. (a) Quando la scimmia percepiva il raggio di sole, la velocità di attivazione del neurone nella corteccia IT rimaneva la stessa. (b) Quando la percezione si trasferiva alla farfalla, l'attivazione aumentava. In entrambe le situazioni, le immagini nella retina rimanevano le stesse. Quello che cambiava era la percezione della scimmia e l'attivazione neurale.

Considera cosa è successo in questo esperimento. Le immagini sulle retine della scimmia sono rimaste le stesse durante l'esperimento - la solarizzazione è stata sempre rappresentata sulla retina sinistra e la farfalla è stata sempre rappresentata sulla retina destra. Il cambiamento della percezione dallo "sprazzo di sole" alla "farfalla" doveva quindi aver avuto luogo nel cervello della scimmia, e questi cambiamenti nella percezione erano collegati ai cambiamenti nell'accensione di un neurone nel cervello. Questa procedura di rivalità binoculare è stata anche utilizzata per collegare la percezione e la risposta neurale negli esseri umani usando l'fMRI. Frank Tong e collaboratori (1998) hanno presentato l'immagine del volto di una persona ad un occhio e l'immagine di una casa all'altro occhio, osservando le immagini attraverso gli occhiali colorati, come mostrato nella figura 5.46. Gli occhiali colorati facevano sì che il viso fosse presentato all'occhio sinistro e la casa all'occhio destro. Poiché ogni occhio ha ricevuto un'immagine diversa, si è verificata la rivalità binoculare, quindi mentre le immagini sono rimaste le stesse sulla retina, osserva solo il volto o solo la casa, e tali percezioni si alternano avanti e indietro ogni pochi secondi. I soggetti hanno premuto un pulsante quando hanno percepito la casa ed un altro pulsante quando hanno percepito il viso, mentre Tong ha utilizzato la risonanza magnetica per misurare l'attività nell'area para-ippocampale del soggetto (PPA) e dell'area facciale fusiforme (FFA). Quando gli

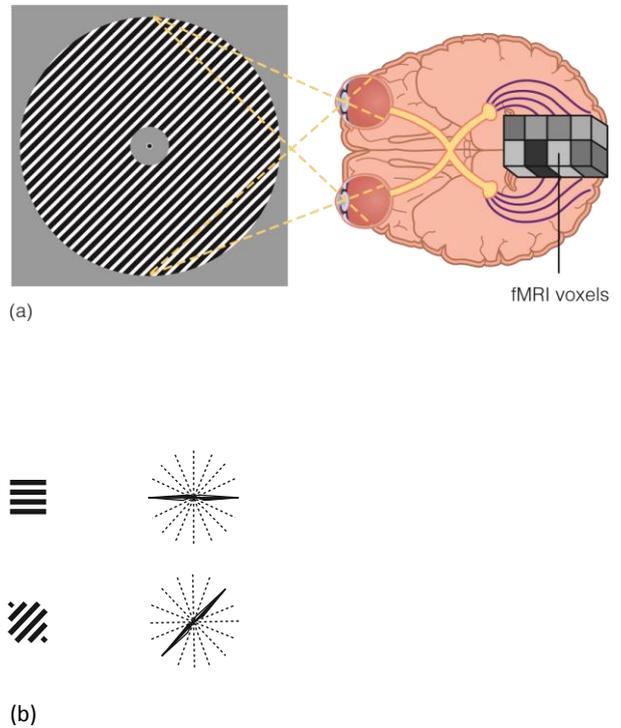
osservatori stavano percependo la casa, l'attività aumentava nel PPA (e diminuiva nel FFA); quando stavano percependo il volto, l'attività aumentava nel FFA (e diminuiva nel PPA). Questo risultato è quindi simile a quello che Sheinberg e Logothetis hanno trovato nella scimmia singolo neurone farfalla esperimento di solarizzazione. Nonostante le immagini sulla retina rimasero le stesse durante l'esperimento, l'attività nel cervello cambiò a seconda di ciò che la persona stava vivendo. Questi esperimenti hanno generato una grande quantità di eccitazione tra i ricercatori del cervello perché misuravano simultaneamente l'attivazione e la percezione del cervello e hanno dimostrato una relazione dinamica tra la percezione e l'attività cerebrale in cui i cambiamenti nella percezione e nell'attività cerebrale si rispecchiavano l'un l'altro.

## Leggere il cervello

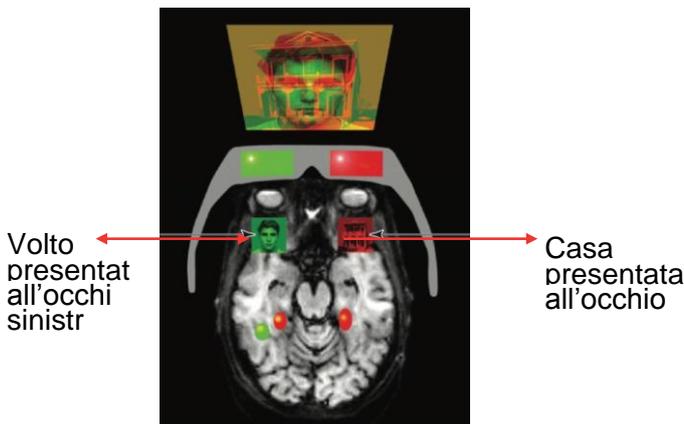
Dopo il successo degli esperimenti sulla rivalità binoculare, i ricercatori hanno fatto il passo successivo chiedendo se è possibile determinare che cosa sta vedendo una persona analizzando il modello di attività nel cervello. Raggiungere ciò implica misurare l'attività cerebrale di una persona mentre sta vedendo qualcosa e poi in qualche modo "decodificare" quell'attività per determinare la percezione associata con essa. Un primo esperimento di "lettura del cervello" ha semplificato il problema limitando ciò che il soggetto vede per stimoli reticolanti orientati come quello della figura 5.47. Kamitani e Tong (2005) hanno registrato le risposte fMRI del loro soggetto a un numero di reticoli con diversi orientamenti (quello in figura 5.47a in 45 gradi a destra) e determinato la risposta alle griglie in un numero di voxel fMRI. Un voxel è una piccola area del cervello a forma di cubo di circa 2 o 3 mm su ciascun lato. (La dimensione di un voxel dipende dalla risoluzione dello scanner fMRI. Sono in fase di sviluppo scanner in grado di risolvere aree inferiori a 2 o 3 mm su un lato). Kamitani e Tong hanno determinato lo schema dell'attività del voxel generato da ciascun orientamento e hanno usato la relazione tra l'attività del voxel e l'orientamento per creare un "decodificatore di orientamento". Questo decodificatore è stato progettato per determinare, dalla risposta del cervello di una persona, quale orientamento stesse vedendo la persona.

Per testare il decodificatore, hanno presentato dei reticoli orientati a un soggetto, hanno alimentato la risposta fMRI risultante nel decodificatore e il decodificatore ha previsto l'orientamento del reticolo. I risultati, mostrati nella Figura 5.47b, indicano che il decodificatore predisse accuratamente gli orientamenti presentati. Lo sviluppo di un decoder in grado di determinare l'orientamento che una persona stava percependo era un risultato impressionante. Ma che dire degli stimoli complessi come le scene nell'ambiente? Espandere il nostro set di stimoli da otto orientamenti di reticolo a ogni scena possibile nell'ambiente è un bel salto! Ma il lavoro recente verso la creazione di un tale "decodificatore di scene" ha avuto un certo successo. Thomas Naselaris e coworkers (2009)

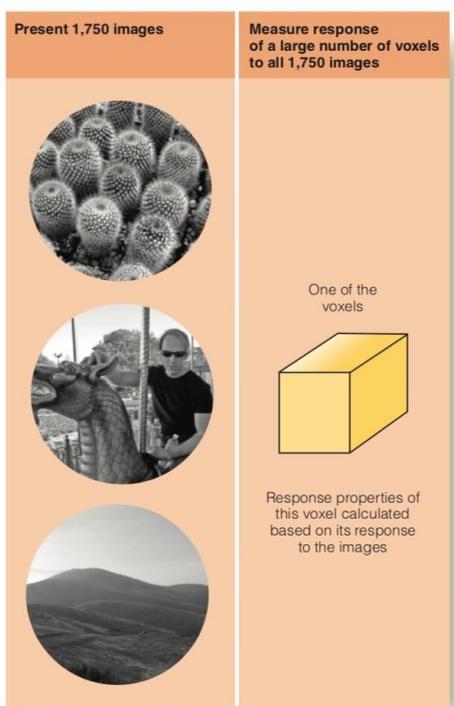
hanno creato un dispositivo per la lettura del cervello sviluppando due metodi per analizzare i pattern di attivazione del voxel registrati dalle aree visive del cervello di un osservatore. Il primo metodo, chiamato **codifica strutturale**, si basa sulla relazione tra l'attivazione del voxel e le caratteristiche strutturali di una scena, come linee, contrasti, forme, e trame. Proprio come il decodificatore di orientamento di Kamitani e Tong è stato calibrato determinando i pattern di attivazione del voxel generati da otto diversi orientamenti, il decodificatore strutturale di Naselaris è stato calibrato presentando un grande numero di immagini, come quelle nella **Figura 5.48**, a un osservatore e determinando come un numero elevato dei voxel ha risposto alle caratteristiche specifiche di ciascuna scena, come l'orientamento della linea, i dettagli e la posizione dell'immagine. Una volta che l'encoder strutturale è stato calibrato, è stato "invertito" per fare previsioni nella direzione opposta, utilizzando gli schemi delle risposte voxel per prevedere le caratteristiche dell'immagine che il soggetto stava visualizzando. Il secondo metodo, chiamato **codifica semantica**, si basa sulla relazione tra l'attivazione del voxel e il significato o la categoria di una scena. L'encoder semantico viene calibrato misurando il modello di attivazione del voxel su un gran numero di immagini che sono state precedentemente classificate in categorie come "folla", "ritratto", "veicolo" e "esterno". Da questa calibrazione, viene determinata la relazione tra il pattern di attivazione del voxel e la categoria dell'immagine. Ci rende possibile utilizzare le risposte voxel che si verificano quando un soggetto sta visualizzando una scena per fare previsioni sul tipo di scena che il soggetto sta visualizzando.



**Figura 5.47** (a) Osserva nell'esperimento di Kamitani e Tong (2005) le griglie orientate come quella di sinistra. I cubi nel cervello rappresentano la risposta di 6 voxel. L'attività di 400 voxel è stata monitorata nell'esperimento. (b) Risultati per due orientamenti. Le grate sono lo stimolo presentato all'osservatore. La linea è l'orientamento previsto dal decodificatore di orientamento. Il decodificatore è stato in grado di prevedere con precisione la presentazione di tutti e otto gli orientamenti testati



**Figura 5.46** Osservate nell'esperimento di Tong et al.(1998) la sovrapposizione della casa rossa e della faccia verde attraverso gli occhiali rosso-verdi, quindi l'immagine della casa è stata presentata all'occhio destro e l'immagine del volto all'occhio sinistro. A causa della rivalità binoculare, la percezione degli osservatori si alterna avanti e indietro tra il viso e la casa. Quando gli osservatori hanno percepito la casa, l'attività si è verificata nell'area del luogo paraippocampale (PPA) negli emisferi sinistro e destro (ellissi rosse). Quando osserva il volto, l'attività si è verificata nell'area della faccia fusiforme (FFA) nell'emisfero sinistro (ellissi verdi).



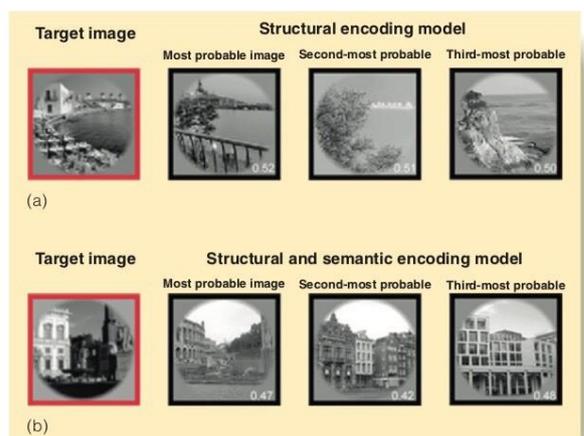
**Figura 5.48** Calibrazione del decodificatore strutturale di Naselaris. Le immagini naturali vengono presentate a un soggetto e viene determinato il modo in cui un gran numero di voxel rispondono alle caratteristiche contenute in ciascuna immagine. Solo tre delle immagini e un voxel sono mostrati qui.

Le informazioni fornite dal decodificatore strutturale e dal decodificatore semantico forniscono un indizio su ciò che il soggetto sta vedendo. Ad esempio, l'encoder strutturale potrebbe indicare che ci sono linee rette di vari orientamenti sulla sinistra della scena, che ci sono contorni curvi in un'altra area. Il decodificatore semantico, che fornisce un diverso tipo di informazioni, potrebbe indicare che il soggetto sta guardando una scena esterna. Ma conoscere le caratteristiche di una scena e il tipo di scena non ci dice come appare realmente la scena. Questo passaggio si ottiene quando il decodificatore consulta un database di 6 milioni di immagini naturali e preleva le immagini che corrispondono maggiormente alle informazioni determinate dall'analisi dell'attività cerebrale della persona. La Figura 5.49a mostra i risultati quando è stato utilizzato solo l'encoder strutturale. L'encoder ha selezionato le tre immagini sulla destra come la migliore corrispondenza per l'immagine di destinazione nella casella rossa, che è l'immagine che la persona stava osservando. La struttura di tutte le immagini corrispondenti è simile, con oggetti che appaiono sulla sinistra dell'immagine e spazi aperti nella parte centrale e destra. Tuttavia, mentre l'immagine di destinazione contiene edifici, gli edifici sono assenti o difficili da vedere nelle immagini corrispondenti. Pertanto, l'encoder strutturale fa solo un buon lavoro di abbinamento alla *struttura* dell'immagine di destinazione, ma un lavoro scarso di corrispondenza con il *significato* dell'immagine di destinazione. L'aggiunta dell'encoder semantico migliora le prestazioni, come mostrato nella

**Figura 5.49b.** È facile vedere l'effetto del codificatore semantico, perché ora

il significato delle immagini della partita è molto più vicino all'immagine di prova, con tutte le corrispondenze che mostrano i lati degli edifici. La capacità di scattare foto che si avvicinano così tanto alla corrispondenza di cui una persona sta guardando, basandosi solo sull'analisi del modello di attivazione del cervello della persona, è piuttosto un risultato. Un motivo per cui le immagini selezionate come corrispondenze non sono esattamente le stesse dell'obiettivo è che le immagini di destinazione non sono contenute nel database di immagini da 6 milioni di immagini da cui è stato selezionato l'encoder. Alla fine, secondo Naselaris, un database di immagini molto più grande produrrà partite molto più vicine all'obiettivo. L'accuratezza aumenterà anche man mano che apprendiamo di più su come l'attività neurale di varie aree del cervello rappresenti le caratteristiche delle scene ambientali.

Naturalmente, il decodificatore finale non avrà bisogno di confrontare il suo output con enormi database di immagini. Analizzerà solo i pattern di attivazione del voxel e ricrea l'immagine delle scene. Attualmente, c'è solo un "decodificatore" che ha raggiunto questo, e questo è il tuo cervello! (Anche se vale la pena notare che il tuo cervello fa uso di un "database" di informazioni sull'ambiente, come sappiamo dal ruolo delle regolarità dell'ambiente nel percepire le scene). Raggiungere questo ultimo decodificatore in laboratorio rientra nella categoria del "progetto scientifico" descritto all'inizio di questo libro (vedi pagina 3) ed è ancora lontano dall'essere raggiunto. Tuttavia, i decoder che attualmente esistono sono risultati sorprendenti, che solo di recente potrebbero essere stati classificati come "fantascienza".



**Figura 5.49** (a) Il soggetto ha visualizzato l'immagine nel quadrato rosso. Il decodificatore strutturale ha scelto le altre tre immagini come la migliore corrispondenza dal database di 6 milioni di immagini. (b) Un'altra immagine vista dal soggetto e le tre immagini scelte dai decodificatori strutturali e semantici come le migliori corrispondenze dal database

