# Optical Flow e la Percezione Umana del Movimento

*Eumir Cometti 986666 – Informatica per la Comunicazione Digitale*

*Corso di Principi e Modelli della Percezione – Università degli Studi di Milano*

Sommario

[Optical Flow e la Percezione Umana del Movimento 1](#_Toc158813117)

[Percezione umana del movimento 1](#_Toc158813118)

[Movimento apparente 1](#_Toc158813119)

[Movimento indotto 1](#_Toc158813120)

[Importanza della percezione del movimento 1](#_Toc158813121)

[Optical flow 2](#_Toc158813122)

[Modellizzazione del movimento 2](#_Toc158813123)

[Optical flow e percezione 2](#_Toc158813124)

[Profondità percettiva 2](#_Toc158813125)

[Stabilità percettiva 3](#_Toc158813126)

[Coerenza del movimento 4](#_Toc158813127)

[Principi della Gestalt e Optical Flow 4](#_Toc158813128)

[Principio di continuità 4](#_Toc158813129)

[Principio di somiglianza 4](#_Toc158813130)

[Principio di chiusura 4](#_Toc158813131)

[Algoritmi di calcolo dell’Optical Flow 4](#_Toc158813132)

[Optical Flow denso 5](#_Toc158813133)

[Algoritmo di Farneback 5](#_Toc158813134)

[Optical Flow sparso 5](#_Toc158813135)

[Algoritmo di Lucas-Kanade 5](#_Toc158813136)

# Percezione umana del movimento

Possiamo definire la percezione del movimento come la capacità cognitiva che possiede un organismo che gli permette di cogliere immediatamente il cambiamento di luogo di un oggetto o di un corpo e, allo stesso tempo, di apprendere alcuni attributi legati a questo cambiamento, come la velocità e direzione.

In termini generali, ci sono due principali meccanismi attraverso i quali si percepisce il movimento: Il primo riguarda il rilevamento dei cambiamenti nella posizione relativa delle parti che compongono l'immagine visiva e prende il nome di sistema immagini-retina. Il secondo riguarda l’uso degli occhi per seguire un bersaglio in movimento e prende il nome di sistema occhio-testa

## Movimento apparente

In realtà non è necessario che si verifichi il movimento per avere la sensazione visiva che qualcosa si muova, per questo parliamo di movimento apparente. Per esempio, se si accendono due luci separate da un piccolo spazio e con brevi intervalli di accensione, l'osservatore ha la sensazione di movimento. La cosa più importante nel movimento apparente è l'asincronia tra gli stimoli, cioè il tempo che intercorre tra l'inizio di uno stimolo e il successivo.

## Movimento indotto

Un altro tipo di movimento è quello che definiamo come indotto: l'illusione che appare quando un oggetto che è effettivamente fisso si muove, spostando il quadro di riferimento in cui si trova.

Sembra essere accettato oggi che il movimento sia una percezione primaria del nostro sistema visivo: Il segnale di movimento inizia ad essere elaborato nella retina nella maggior parte degli animali inferiori, dal gatto fino all'uomo, in cui l'elaborazione diventa più complessa e avviene a livello centrale.

# Importanza della percezione del movimento

La percezione del movimento ha una serie di funzioni differenti, che vanno dal fornirci aggiornamenti su cosa sta succedendo, all’aiutarci a percepire le cose, come le forme degli oggetti e gli stati d’animo delle persone. Forse la più importante di tutte, specialmente per gli animali, è il fatto che la percezione del movimento sia strettamente legata alla sopravvivenza: il movimento è la chiave per localizzare la preda e stabilire i movimenti precisi per darle la caccia o, al contrario, per passare inosservato al predatore e fuggire al momento giusto.

Tornando al mondo delle persone, la percezione del movimento è essenziale per la nostra abilità di muoverci attraverso l’ambiente: una fonte di informazioni su dove stiamo andando e quanto velocemente ci stiamo muovendo è il modo in cui gli oggetti nell’ambiente fluiscono davanti a noi mentre ci muoviamo. Quando una persona si muove in avanti, gli oggetti si muovono nella direzione opposta a quella della persona. Questo movimento, chiamato flusso ottico, ci fornisce informazioni sulla direzione e sulla velocità del passante.

Un’ altra abilità del movimento è quella di attirare l’attenzione. Questo effetto non si verifica solamente quando stai consapevolmente cercando qualcosa, ma anche quando stai prestando attenzione a qualcos’altro. Per esempio, mentre stai avendo una conversazione, la tua attenzione può essere improvvisamente catturata da qualcosa in movimento nella tua visione periferica.

# Optical flow

L’optical flow (flusso ottico in italiano) è un concetto che considera il moto di un oggetto all'interno di una rappresentazione visuale digitale.

Tipicamente il moto è rappresentato come un vettore che si origina da (o termina su) un [pixel](https://it.wikipedia.org/wiki/Pixel) ad esempio in una sequenza di frame. Lo scopo dell'optical flow è quello di assegnare ad ogni pixel appartenente al frame corrente un [motion vector](https://it.wikipedia.org/w/index.php?title=Motion_vector&action=edit&redlink=1" \o "Motion vector (la pagina non esiste)) che punta verso la posizione dello stesso pixel in un frame di riferimento successivo.

## Modellizzazione del movimento

La modellizzazione del movimento, in relazione all'optical flow, rappresenta il processo di formulazione matematica e computazionale dei cambiamenti nello spazio e nel tempo dei pixel in un'immagine. Questo processo è finalizzato a catturare e descrivere il movimento apparente degli oggetti nella scena, fornendo un'importante base concettuale per la visione artificiale e la computer vision.

La definizione matematica gioca un ruolo cruciale nella modellizzazione del movimento. Questa coinvolge la creazione di equazioni che descrivono come le intensità luminose dei pixel cambiano nel tempo. Spesso, tali equazioni sono formulate mediante l'uso di metodi differenziali parziali che riescono a catturare la dinamica del flusso ottico.

La modellizzazione del movimento può generare risultati densi o sparsi, a seconda di come viene trattato il flusso ottico dai diversi algoritmi. Un flusso ottico denso fornisce informazioni su ogni pixel nell'immagine, mentre un flusso ottico sparso considera solo un sottoinsieme di punti chiave.

Gli approcci alla modellizzazione del movimento possono variare da classici metodi basati su correlazione, gradienti di intensità o confronto tra frame successivi, a tecniche più moderne basate su reti neurali, come le reti neurali convoluzionali (CNN), che hanno dimostrato efficacia in questo contesto.

La modellizzazione del movimento può essere eseguita a livello locale, dove il flusso ottico viene stimato per piccole regioni dell'immagine, o a livello globale, considerando l'intera immagine. La scelta tra i due approcci dipende dalla complessità del movimento presente nella scena.

La modellizzazione del movimento trova applicazioni in vari contesti pratici, come il tracciamento di oggetti in un video, la stabilizzazione dell'immagine, la percezione del movimento in veicoli autonomi e la creazione di effetti visivi nell'ambito cinematografico.

# Optical flow e percezione

## Profondità percettiva

La percezione della profondità è un aspetto cruciale del sistema visivo umano che ci consente di interpretare la distanza e la disposizione spaziale degli oggetti nell'ambiente circostante. Il nostro cervello fa affidamento su diversi meccanismi e indizi visivi, sia monoculari che binoculari, per generare una percezione tridimensionale del mondo che ci circonda.

La parallasse binoculare è uno di questi meccanismi. Si basa sulla differenza nelle viste oculare destro e sinistro, poiché gli occhi sono posizionati leggermente distanti l'uno dall'altro. Il cervello integra queste informazioni binoculari per stimare la distanza degli oggetti.

Un altro indizio binoculare è la convergenza oculare, che si verifica quando gli occhi si muovono simultaneamente verso il naso quando si guarda un oggetto vicino. La quantità di convergenza necessaria fornisce al cervello un'indicazione sulla distanza dell'oggetto.

L'accomodazione è un processo in cui la lente dell'occhio cambia forma per focalizzarsi su oggetti a diverse distanze. Anche se più efficace per distanze ravvicinate, contribuisce alla percezione della profondità.

L’indizio della prospettiva lineare è basato sul fatto che gli oggetti più lontani sembrano più piccoli rispetto a quelli più vicini nella vista prospettica. Questo elemento è spesso utilizzato inconsciamente per giudicare la distanza.

Le ombre e l'illuminazione giocano un ruolo importante nella percezione della profondità. La direzione e l'intensità delle ombre possono fornire suggerimenti sulla posizione degli oggetti, con ombre lunghe che indicano distanza.

Casi di sovrapposizione o occlusione sono basati sul fatto che un oggetto che copre parzialmente un altro sembra più distante. Questo fattore è spesso sfruttato quando giudichiamo la disposizione relativa degli oggetti.

Un elemento molto importante è il gradiente di tessitura: un gradiente di tessitura si osserva ogni volta che una superficie viene vista in prospettiva invece che direttamente dall’alto. La tessitura diviene più densa e meno dettagliata a mano a mano che la superficie si allontana sullo sfondo, e questa trasformazione aiuta a giudicare la profondità. Per questo gli oggetti con trama più fitta vengono percepiti come più lontani

Altri indizi, come l'atmosfera e la nebbia prospettica, influenzano la percezione della profondità. Gli oggetti lontani possono apparire più sfumati a causa degli effetti atmosferici, come la nebbia o l'umidità nell'aria.

Infine, il movimento relativo fornisce indicazioni sulla distanza. Gli oggetti che si muovono attraverso il campo visivo possono essere interpretati come più vicini rispetto a quelli più stabili.

Il nostro cervello integra simultaneamente queste informazioni provenienti da diverse fonti, come indizi binoculari e monoculari, esperienze passate e contesto visivo, per costruire una percezione coerente e dinamica della profondità nell'ambiente circostante.

## Stabilità percettiva

L’optical flow gioca un ruolo chiave nel mantenere la stabilità percettiva durante i movimenti del corpo e della testa. Come per la profondità percettiva, ci sono più indizi che vengono elaborati in simultanea dal cervello e permettono di mantenere stabile la percezione dell’ambiente circostante.

Durante il movimento della testa i muscoli oculari ci permettono di mantenere il focus sugli oggetti d’interesse. Questa compensazione del movimento oculare è ottenuta eseguendo dei piccoli movimenti dell’occhio, noti come saccadici, che aiutano la retina a ricevere immagini stabili per far si che il sistema nervoso possa continuare ad elaborare informazioni visive coerenti.

Grazie al flusso ottico siamo in grado di anticipare il movimento degli oggetti. Per questo è possibile adattarsi rapidamente ai cambiamenti ambientali e mantenere la stabilità visiva prima ancora che il movimento si verifichi.

Un tipo di adattamento avviene anche a livello neurale: le cellule neurali coinvolte nella percezione visiva presentano un adattamento selettivo al movimento che garantisce una risposta ottimale a un determinato tipo di movimento contribuendo a una percezione più stabile durante moto di corpo e testa.

Tutte queste informazioni vengono integrate con altri flussi informativi provenienti da altri sensori come la visione periferica e la vestibolocettività per correggere eventuali discrepanze.

## Coerenza del movimento

L'essere umano utilizza il flusso ottico per regolare la coerenza del movimento, cioè per percepire e interpretare in modo coerente il movimento degli oggetti nell'ambiente circostante.

La coerenza del movimento si riferisce alla percezione di un movimento uniforme e coerente degli oggetti nella scena. Il cervello umano utilizza il flusso ottico come importante indizio visivo per valutare e regolare questa coerenza del movimento. Quando gli oggetti si muovono in modo coerente, il flusso ottico generato da tali oggetti risulta uniforme e consistente nel tempo, facilitando la percezione di oggetti solidi e distinti che si muovono insieme.

Uno degli aspetti importanti della coerenza del flusso ottico è la sua relazione con la percezione della profondità e della distanza degli oggetti. Quando gli oggetti si muovono in modo coerente, il nostro cervello interpreta questo flusso ottico come indicazione di oggetti situati a una certa distanza e profondità. Questa capacità di percepire la profondità e la distanza degli oggetti in movimento contribuisce alla nostra percezione spaziale dell'ambiente circostante e alla nostra capacità di navigare in esso in modo efficace e sicuro.

Infine, la coerenza del flusso ottico consente al cervello di riconoscere e interpretare modelli di movimento complessi. Ad esempio, possiamo distinguere facilmente il movimento di un gregge di animali dal movimento casuale di oggetti dispersi. Questa capacità di riconoscere e comprendere i modelli di movimento ci aiuta a percepire l'ambiente circostante in modo efficace e coerente, consentendoci di adattarci e reagire alle situazioni in modo appropriato.

# Principi della Gestalt e Optical Flow

I principi della Gestalt e il concetto di optical flow sono correlati in quanto entrambi riguardano la percezione visiva e la nostra comprensione della scena visiva. La Gestalt si concentra sull'organizzazione e sull'interpretazione degli elementi visivi per percepire figure coerenti e significative, mentre l'optical flow si riferisce al movimento apparente dei punti nell'immagine rispetto a un osservatore in movimento.

## Principio di continuità

Il principio della continuità della Gestalt suggerisce che tendiamo a percepire linee continue e forme fluide piuttosto che interruzioni o discontinuità. Questo principio può influenzare la percezione della coerenza del movimento, poiché tendiamo a percepire il movimento come continuo e fluido anche quando gli oggetti sono parzialmente nascosti o oscurati durante il movimento.

## Principio di somiglianza

La legge della somiglianza della Gestalt suggerisce che tendiamo a percepire gli oggetti che si muovono in modo simile come facenti parte dello stesso gruppo o flusso di movimento. Questo principio può influenzare la percezione del movimento come un'unica entità coesa.

## Principio di chiusura

La legge della chiusura della Gestalt suggerisce che tendiamo a percepire gli oggetti come figure chiuse e complete anche quando mancano alcune parti. Questo principio può influenzare la percezione della coerenza del movimento, poiché tendiamo a completare le informazioni mancanti per percepire il movimento come una figura coerente e distinta.

# Algoritmi di calcolo dell’Optical Flow

Esistono due principali classi di algoritmi per il calcolo del flusso ottico. La suddivisone avviene in base a come l’optical flow viene considerato.

## Optical Flow denso

L'optical flow denso coinvolge il calcolo del flusso ottico per ogni singolo pixel nell'immagine. Questo significa che il flusso ottico viene stimato per ogni punto dell'immagine, producendo una mappa di vettori di movimento che copre l'intera immagine. Questo approccio fornisce una quantità significativa di informazioni dettagliate sul movimento in ogni punto dell'immagine, consentendo una comprensione completa del movimento in tutta la scena. Tuttavia, può richiedere molte risorse computazionali, specialmente per immagini ad alta risoluzione.

### Algoritmo di Farneback

L'algoritmo di Farneback è un metodo per il calcolo dell'optical flow che si propone di determinare il movimento dei punti all'interno di un'immagine da un frame temporale all'altro. Questo processo è fondamentale per comprendere il dinamismo presente in sequenze di immagini o video.

Inizialmente, l'algoritmo divide l'immagine in blocchi di dimensioni definite. Per ogni blocco, viene esaminata la variazione dei valori dei pixel tra due frame consecutivi. Attraverso un approccio basato su tecniche di calcolo matematico, l'algoritmo tenta di identificare il modello di trasformazione ottico tra i blocchi corrispondenti nei due frame.

Successivamente, utilizzando il modello di trasformazione trovato per ciascun blocco, l'algoritmo stima il flusso ottico per ogni punto nell'immagine. Questo flusso ottico rappresenta la direzione e la velocità con cui i punti si spostano da un frame all'altro.

Infine, l'algoritmo restituisce una mappa di vettori di movimento, rappresentante il flusso ottico per ogni punto dell'immagine.

## Optical Flow sparso

L'optical flow sparso coinvolge il calcolo del flusso ottico solo per un sottoinsieme limitato di punti chiave o caratteristiche distintive nell'immagine. Questo significa che il flusso ottico viene stimato solo per alcuni punti selezionati, producendo una mappa di vettori di movimento solo per questi punti specifici. Questo approccio riduce notevolmente la quantità di dati di movimento calcolati, consentendo una computazione più efficiente, specialmente su immagini ad alta risoluzione. Tuttavia, l'optical flow sparso può perdere informazioni dettagliate su parti meno significative della scena.

### Algoritmo di Lucas-Kanade

L'algoritmo di Lucas-Kanade è un metodo per il calcolo dell'optical flow che si concentra sull'estrazione e l'analisi del movimento in piccole regioni dell'immagine, comunemente definite come punti chiave o feature points. L’algoritmo opera diverse fasi.

Vengono identificati e selezionati un insieme di punti chiave nell'immagine su cui concentrarsi per calcolare il flusso ottico. Questi punti chiave sono generalmente individuati tramite algoritmi di rilevamento di corner o features, che identificano regioni dell'immagine con forti variazioni di intensità luminosa.

Per ciascun punto chiave selezionato viene calcolato il gradiente spaziale dei pixel circostanti. Questo gradiente fornisce informazioni sulla direzione e l'intensità delle variazioni di colore nell'area intorno al punto chiave, che sono utilizzate per stimare il movimento locale.

L'algoritmo confronta i valori dei pixel nei punti chiave tra due frame temporali consecutivi per calcolare la derivata temporale, cioè la variazione di intensità luminosa nel tempo per ciascun punto. Questa derivata temporale fornisce informazioni sul movimento dei punti chiave nel corso del tempo.

Utilizzando il gradiente spaziale e la derivata temporale calcolati, viene applicata una tecnica di regressione, comunemente nota come metodo dei minimi quadrati, per stimare il flusso ottico per ciascun punto chiave. La tecnica prevede la risoluzione di un sistema di equazioni lineari che approssima il movimento dei punti nell'area circostante.

Tutte queste fasi posso essere eseguite iterativamente per affinare sempre di più la stima del flusso ottico.

Alla fine, l'algoritmo restituisce una mappa di vettori di movimento che rappresenta il flusso ottico stimato per i punti chiave selezionati.