

MACCHINE PERCETTIVE 1

Prof. Giacomo Mario Bisio
Tel. 010-353.2756

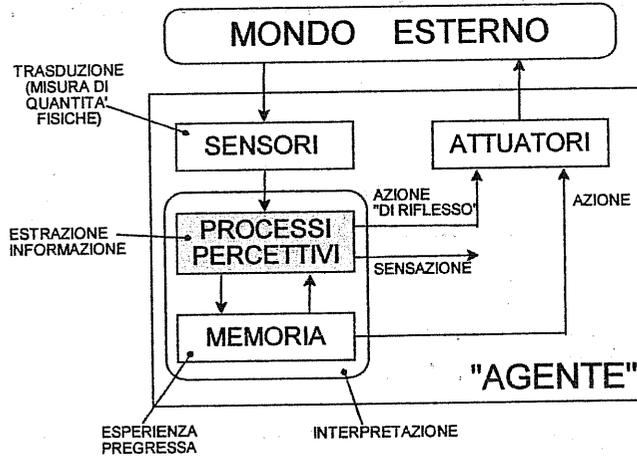
Prof. Silvio Paolo Sabatini
Tel. 010-353.2092

Appunti raccolti nell'a.a. 2004-05 e redatti da Matteo Bruno

INTRODUZIONE

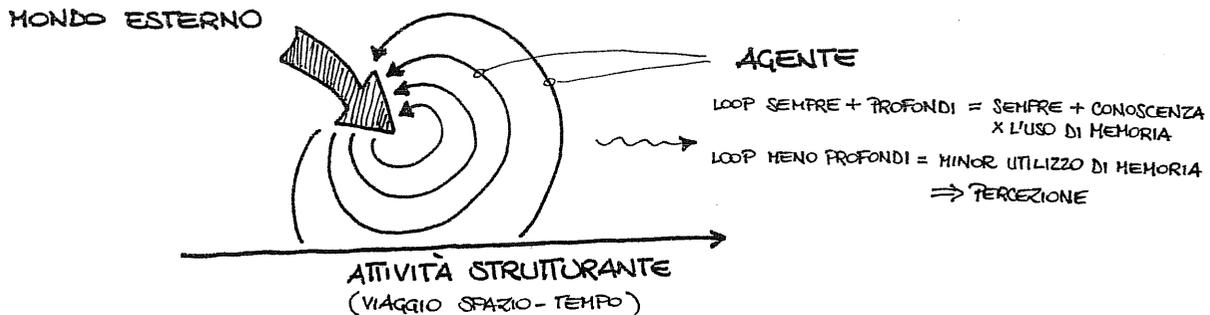
SISTEMI PERCETTIVI: parte costitutiva di un oggetto intelligente capace di prendere dati dall'esterno, elaborarli ed agire in base ad essi.

Vediamo lo SCHEMA PERCEZIONE - AZIONE:



Possiamo dividere le attività in:

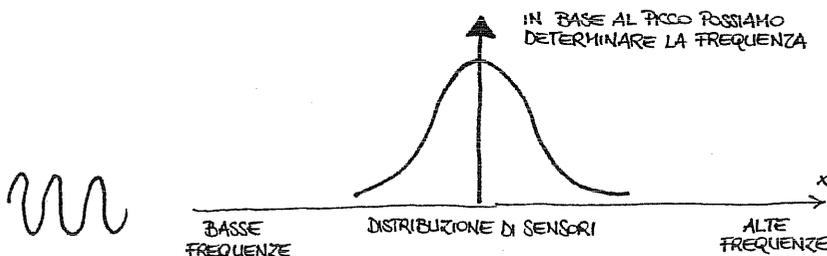
- ATTIVITÀ
- SENSORIALE: trasposizione - traduzione di una grandezza fisica in un'altra grandezza
 - PERCETTIVA
 - COGNITIVA: legata all'esperienza pregressa, è l'interpretazione dei nostri sensi.



Supponiamo di avere un certo % di sensori collocati in \neq punti dello spazio con caratteristiche \neq : si può pensare che essi possono misurare direttamente alcune caratteristiche, poiché tali sensori registrano cose diverse.

Questa strategia è adottata nei sistemi naturali (es.: coclea x il suono): in tal caso il sensore è fatto in modo tale da risuonare in \neq parti a seconda della frequenza del segnale.

ELABORAZIONE PER STRUTTURA



1. VISIONE COME PROCESSO DI ELABORAZIONE

L'elaborazione avviene tramite l'interazione tra i sensori e la memoria x mezzo della percezione.

Esistono 2 diverse strategie x mediare l'interazione tra i dati sensoriali e la memoria:

① DATA DRIVEN [BOTTOM-UP] (James Gibson - visione ecologica)
Conosciamo meglio e abbiamo una partecipazione attiva nella scena

② GOAL DIRECTED [BOTTOM-DOWN] (Richard Gregory)

- Aspettativa
- Verifica di previsioni
- Strategie di pianificazione

IN GENERALE SU MODELLI DI CONOSCENZA A PRIORI

Siamo maggiormente interessati alla prima strategia, e quindi a trarre conclusioni in base ai dati.

1.1. LIVELLI DI RAPPRESENTAZIONE

Esistono 3 livelli di rappresentazione:

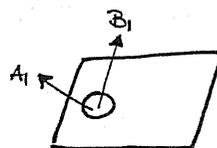
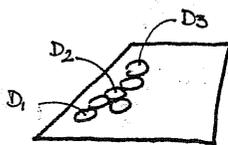
① EARLY VISION

Trasformazione di immagini in altre immagini.

Abbiamo immagini parametriche, rappresentazione iconica (ancora analogica) e l'estrazione di alcune proprietà fisiche della scena, quali ORIENTAMENTO, COLORE, MOVIMENTO (velocità) e PROFONDITÀ.

② VISIONE DI MEDIO LIVELLO - INTEGRAZIONE

Da un'immagine estraiamo ≠ info in varie posizioni: dobbiamo integrare x vedere se alcune regioni hanno caratteristiche simili



DIVERSI P^{II} DI VISTA DI UNA STESSA PORZIONE DI SPAZIO, POI INTEGRATI.

Questo metodo ci fa affrontare la segmentazione di un'immagine e il compartamento dei contorni, e l'integrazione di ≠ analisi x la profondità.

La finalità è quella di ricostruire la scena con l'orientamento e altre caratteristiche.

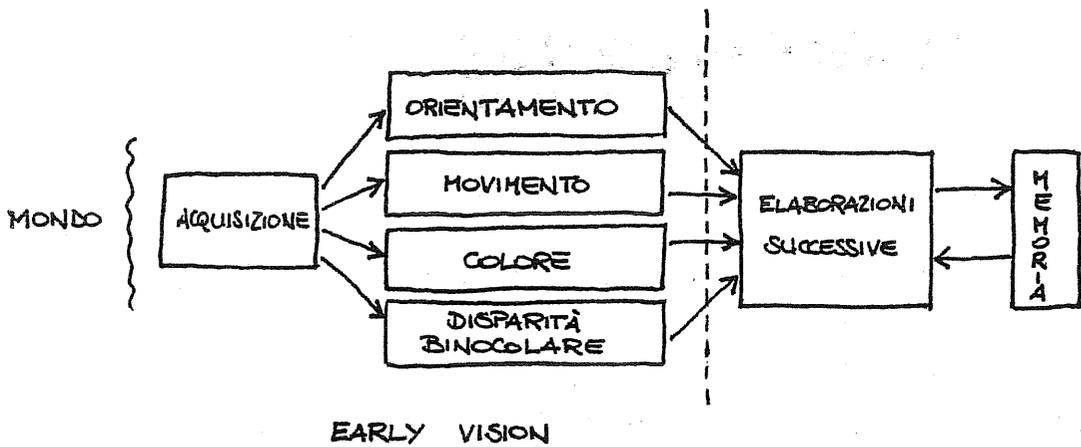
③ VISIONE AD ALTO LIVELLO

Effettua l'analisi della scena: in pratica è un pb di riconoscimento degli oggetti nella scena o un semplice controllo del comportamento.

1.2. EARLY VISION

Dalla distribuzione di intensità luminosa formata sulla retina vogliamo estrarre le info + elementari possibili, che stimino un segnale a seconda di come sono combinate.

EARLY VISION è, in pratica, la prima analisi che scompone le immagini che si creano
 da retina in caratteristiche elementari x poi poterle combinare + complessamente:



EARLY VISION

ARATTERIZZAZIONE DEL SEGNALE VISIVO

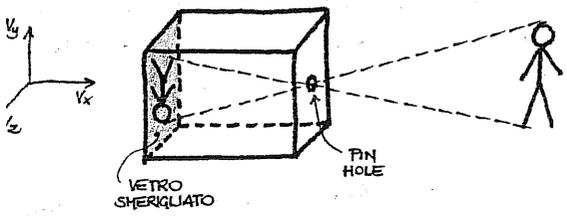
si definisce

$$P(x, y, t, \lambda, V_x, V_y, V_z)$$

FUNZIONE OTTICA
COMPLETA (= FOC)

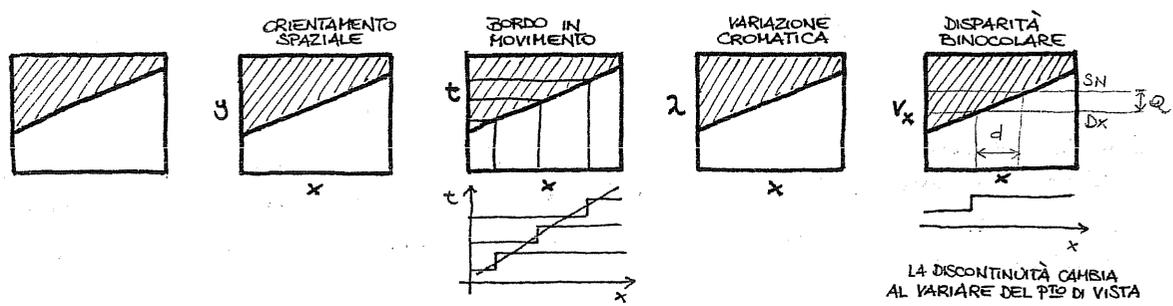
- dove x, y = distribuzione spaziale
- t = istante di tempo dell'intensità luminosa
- λ = lunghezza d'onda
- V_x, V_y, V_z = punto di vista (= posizione della camera nel mondo)

In pratica la FOC rappresenta l'insieme delle immagini che possiamo raccogliere
 tramite un sensore ottico (o una camera oscura) che si sposta nello spazio:



Al momento dell'acquisizione non si
 effettua l'interpretazione del segnale
 entrante: è solo al momento dell'EARLY
 VISION che il segnale luminoso viene
 scomposto in simboli.

facendo la proiezione di P su \neq piani coordinati, troviamo l'info che ci serve
 osservando le variazioni tra le varie proiezioni di P.



il bordo ha significati \neq a seconda del piano su cui si trova: in (t, x) , ad es,
 significa che al variare del tempo è una discontinuità lungo l'asse x in movi-
 mento (EDGE); in (x, y) è un bordo di un oggetto fisico orientato nello spazio
 (BORDO EDGE 2D).

In P sono importanti le ANISOTROPIE, ovvero le grandezze fisiche che variano a
 seconda della direzione considerata. L'avvertire la presenza di anisotropie è
 il primo passo da compiere x l'interpretazione del segnale visivo.

La dimensionalità dell'immagine dipende dal * di pixel e dalle tonalità di grigio che caratterizzano ognuno di questi.

Per caratterizzare la FOC dobbiamo considerare la derivata direzionale $\frac{\partial P}{\partial \mu}$.

Facendo l'espansione di Taylor su P troviamo info locali del segnale; x conoscere il segnale possiamo farne la derivata puntuale, ma è poco affidabile x il segnale è corrotto dal rumore.

Poichè normalmente osserviamo corpi estesi, è + opportuno fare una misura (derivata) locale facendo la derivata direzionale della media locale di P, ovvero

$$\frac{\partial}{\partial \mu} \{ \text{media locale di } P \}$$

Il pb del rumore è che, derivando, possiamo mettere in evidenza dei picchi che appaiono significativi, ma che nel segnale non esistono.

Come precauzione x ridurre il rumore passiamo, prima di derivare, un filtro passa basso G:

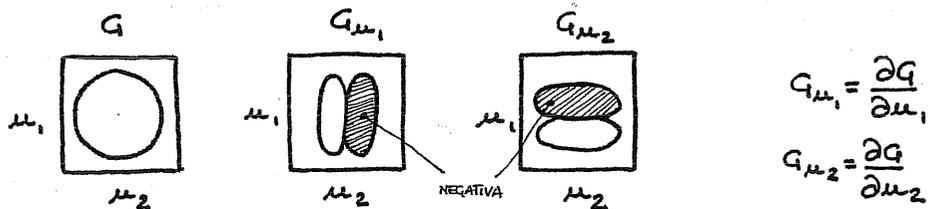
$$\frac{\partial}{\partial \mu} (P * G)$$

dove $G = G(\mu, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{\mu_1^2 + \mu_2^2}{2\sigma^2}\right] = \text{GAUSSIANA CIRCOLARE [LPF]}$

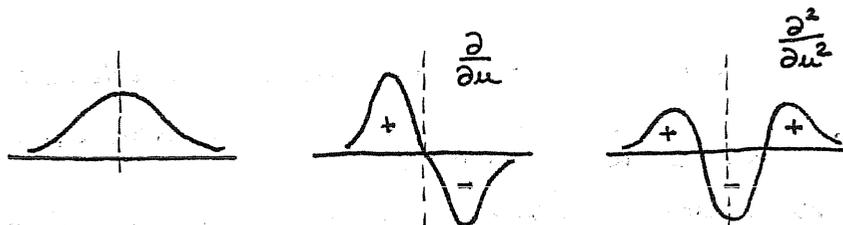
Per la linearità degli operatori possiamo trasformare la derivata in

$$\frac{\partial}{\partial \mu} (P * G) = P * \left(\frac{\partial}{\partial \mu} G \right)$$

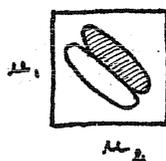
e dunque avremo



Nel caso (1-D) sarà



Combinando le derivate rispetto ad μ_1 e μ_2 si ottiene, invece, un filtro orientato



$$G_{\mu_1 + \mu_2} = \frac{\partial G}{\partial \mu_1} + \frac{\partial G}{\partial \mu_2}$$

Applicando delle DOG e componendole otteniamo dei filtri selettivi ai contorni in direzioni diverse. Tali operatori vengono chiamati FEATURE DETECTOR (= estrattori di caratteristiche): essi in realtà non estraggono, bensì misurano le proprietà della funzione P.

CONSIDERAZIONI

SCALA (= dimensione dell'apertura/kernel)

Poiché è un'apertura fisica la sua dimensione non potrà mai $\rightarrow 0 \Rightarrow$ no dei vate puntuali, \hat{x} come kernel usiamo la gaussiana.

ASSENZA DI INFO A PRIORI IMPLICA:

- ① Assenza di modelli, assenza di memoria \Rightarrow LINEARITÀ [\hat{x} ELABORAZIONE SEMPLICE]
- ② Assenza di localizzazione preferita (= non c'è un pto da analizzare meglio) \Rightarrow SPAZIO-INVARIANZA
- ③ Assenza di orientamento preferito \Rightarrow ISOTROPIA
- ④ Assenza di dimensione preferita \Rightarrow INVARIANZA ALLA SCALA

obbiamo legare la dimensione di G alla risoluzione che vogliamo ottenere: eriamo l'invarianza alla scala con la gaussiana

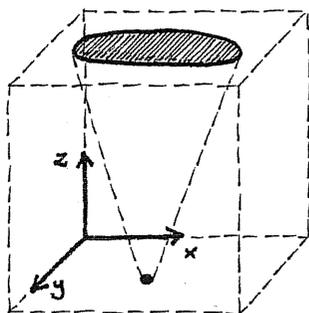
$$G(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

unque il parametro libero che regola l'apertura senza cambiare il kernel è σ .

ta: il fatto di scegliere una gaussiana come forma d'apertura ci toglie il pb di avere falsi lati, che invece una forma squadrata ci procurerebbe.

PROCESSO MULTISCALA LINEARE

spazio generato nell'applicazione dell'operatore G alle diverse scale viene etto LINEAR-SCALE-SPACE



Ogni pixel genera una diffusione.

L'interpretazione del LINEAR SCALE-SPACE mediante un'eq^{ne} di diffusione è

$$\frac{\partial L}{\partial s} = \nabla \cdot \nabla L = \overset{\text{LAPLACIANO}}{\Delta} L = L_{xx} + L_{yy}$$

dove L = luminosanza immagine

generale la caratterizzazione di P richiede derivate direzionali di ordine n -mo

N -JET LOCAL = insieme di ordine n delle $1-N$ derivate applicate per ogni pto della funzione P .

nsiderando l'insieme delle derivate della gaussiana alle diverse scale avremo -JET MULTISCALA.

ICORA SU FILTRI GAUSSIANI

bbiamo visto che, derivando G , otteniamo operatori orizzontali e verticali, mentre n l'approccio dei FILTRI DIREZIONABILI possiamo ottenere tutte le orientazioni ssibili tramite la combinazione di 2 filtri base.

ediamo ora analiticamente:

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$\frac{\partial G}{\partial x} = G_x(x, y; \sigma) = -\frac{1}{2\pi\sigma^2} \cdot \frac{x}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Esprimiamo G in coordinate polari:

$$G(r, \theta; \sigma) \xrightarrow{\frac{\partial}{\partial x}} G_x(r, \theta) = -\pi e^{-\frac{\pi^2}{2\sigma^2}} \cos \theta$$

SEPARAZIONE DELLE VARIABILI
COMPONENTE RADIALE COMPONENTE ANGOLARE

Ruotando la G_x di $\frac{\pi}{2}$ (\square), quindi $(\theta - \frac{\pi}{2})$ nella formula, otteniamo

$$G_x(r, \theta - \frac{\pi}{2}) = -\pi e^{-\frac{\pi^2}{2\sigma^2}} \cos(\theta - \frac{\pi}{2})$$

$$G_y(r, \theta) = -\pi e^{-\frac{\pi^2}{2\sigma^2}} \sin \theta$$

e dunque, generalizzando,

$$\boxed{G_\alpha(r, \theta) = -\pi e^{-\frac{\pi^2}{2\sigma^2}} \cos(\theta - \alpha) = -\pi e^{-\frac{\pi^2}{2\sigma^2}} (\cos \theta \cos \alpha + \sin \theta \sin \alpha) = (\cos \alpha \quad \sin \alpha) \begin{pmatrix} G_x \\ G_y \end{pmatrix}}$$

La derivata prima direzionale richiede dunque una base di dimensione 2.
Data

$$\begin{matrix} \nearrow \\ \text{BASE A} \\ (n+1 \text{ DIMENSIONI}) \end{matrix} \frac{\partial^m G}{\partial x^m} \xrightarrow{\mathcal{F}} (-j\omega)^m \mathcal{F}\{G\}$$

1.3. TEORIA DELLA REGOLARIZZAZIONE APPLICATA AI PB DI EARLY VISION

Il pb di passare da una visione (3-D) ad una (2-D) è un pb di ottica classica; partire da immagini (2-D) e ricostruire le (3-D) come pb di OTTICA INVERSA: è un pb difficile, e non è detto che la soluzione sia unica.

Un pb inverso viene anche detto PB MAL POSTO: a trovarne la soluzione dobbiamo porre dei vincoli sulla soluzione stessa, in modo da restringere il campo delle possibilità.

La risoluzione dei pb mal posti può avvenire tramite

- INVERSE GENERALIZZATE
- FORMULAZIONI VARIAZIONALI

Un pb è ben posto se valgono le seguenti condizioni:

- la soluzione esiste
- la soluzione è unica
- la soluzione dipende con continuità dai dati

Inoltre la soluzione deve essere stabile nonostante il rumore sui dati.

Non è detto che un pb continuo sia anche stabile: se consideriamo un operatore lineare A e la seguente relazione

$$\begin{matrix} \nearrow \\ \text{SOLUZIONE} \end{matrix} y = Az \begin{matrix} \nwarrow \\ \text{DATI} \end{matrix}$$

questo è un pb diretto ben posto. Se invece consideriamo

$$\begin{matrix} \nwarrow \\ \text{DATI} \end{matrix} y = Az \begin{matrix} \nearrow \\ \text{SOLUZIONE} \end{matrix}$$

questo è un pb mal posto, e quando A non è invertibile siamo di fronte ad un pb inverso.

TECNICHE DI REGOLAZIONE servono, appunto, per trasformare un PB MAL POSTO in un BEN POSTO.

fornito il pb

$$y = Az$$

dichiamo la pseudo inversa con

$$z^+ = A^+ y$$

ove z^+ = pseudosoluzione che, a sua natura, non sarà regolarizzata

A^+ = generalizzazione della "matrice inversa di Moore-Penrose"

$$A^+ = (A^T A)^{-1} A^T \quad \text{PSEUDOINVERSA}$$

vuole trovare

$$\|Az - y\|^2 = \text{minimo} \quad \left[\begin{array}{l} \text{PB AI MINIMI} \\ \text{QUADRATI} \end{array} \right]$$

generale, se A è un operatore lineare (e non necessariamente una matrice), può definire come

$$A^+ = (A^* A)^{-1} A^*$$

ove A^* = operatore aggiunto

e siamo in uno spazio vettoriale l'operatore aggiunto viene definito da

$$\left. \begin{array}{l} \underline{u} : X \rightarrow Y \\ \underline{v} : X \rightarrow Y \\ \langle \underline{u}, \underline{v} \rangle = \text{prodotto scalare} \end{array} \right\} \Rightarrow \langle A \underline{u}, \underline{v} \rangle = \langle \underline{u}, A^* \underline{v} \rangle$$

La soluzione z^+ (non regolarizzata) non garantisce la regolarità della soluzione (= SMOOTHNESS), quindi può non valere la condizione di continuità.

3.1. METODO DI REGOLARIZZAZIONE DI TIKHONOV

costituisce A^+ come limite di una successione di operatori R_λ (con $\lambda > 0$) t.c.

$$\lim_{\lambda \rightarrow 0} R_\lambda = A^+ \Rightarrow \lim_{\lambda \rightarrow 0} R_\lambda y = A^+ y = z^+$$

invece i dati sono corrotti dal rumore:

$$y_E = y + m_E \quad \uparrow \text{RUMORE}$$

pb allora diventa

$$R_\lambda y_E = \underbrace{R_\lambda y}_{z^+ \text{ (PER } \lambda \rightarrow 0)} + \underbrace{R_\lambda m_E}_{\substack{\uparrow \text{ SOLUZIONE} \\ \text{INSTABILE} \Rightarrow \text{ DIVERGE}}}$$

La ricerca di una soluzione tramite l'operatore è difficile da regolarizzare il rumore è amplificato. Bisogna trovare un compromesso tra la soluzione esatta

il disturbo dovuto al rumore: alla soluzione ai minimi quadrati aggiungiamo un termine di vincolo fisico

$$\underbrace{\|Az - y\|^2}_{\text{SOLUZIONE PB}} + \underbrace{\lambda \|Pz\|^2}_{\substack{\text{PARAMETRO DI REGOLAZIONE} \\ \text{VINCOLO STABILIZZATORE}}} = \text{minimo} \quad (*)$$

in q. modo abbiamo costruito una funzione di costo che tiene conto di certe info. a priori e avere una certa regolarità e restringere il campo delle soluzioni.

L'espressione + generale che ha proposto Tikhonov x il termine di regolazione è

$$\|Pz\|^2 = \sum_{r=0}^p \int c_r(\xi) \left| \frac{d^r z}{d\xi^r} \right|^2 d\xi$$

TERMINE DI
FEDELTA'
AI DATI

dove $c_r(\xi)$ è una funzione strettamente positiva.

La precisione dipende dall'ordine di derivate a cui si arriva. In realtà avremo solo somme di derivate: + le derivate hanno valori elevati e meno la funzione è smooth.

Si può dimostrare che la soluzione R_λ è esprimibile come

$$R_\lambda = (A^*A + \lambda P^*P)^{-1} A^*$$

Nota: se $\lambda=0$, allora avremo $R_\lambda = (A^*A)^{-1} A^*$ (minimi quadrati)

Formalmente la soluzione regolarizzata si trova cercando la soluzione z tra tutte quelle

$$\|Pz\| < E$$

(E = costante fissata)

che minimizzano il pb di partenza

$$\|Az - y_E\| \Rightarrow \|Az - y_E\|^2 + \lambda \|Pz\|^2 = \text{minimo}$$

↑
DATI RUMOROSI
↑
MOLTIPLICATORE
DI LAGRANGE

Questa funzione può essere riassunta come

$$F(x, y(x), z(x), \dot{z}, \ddot{z}, \dots) = \|Az - y_E\|^2 + \lambda \|Pz\|^2$$

Per risolverla servono le eqⁿⁱ di Eulero-Lagrange, che nei casi + semplici sono del tipo

$$F_u = -\frac{d}{dx} F_{\dot{u}} = 0$$

$$F_u - \frac{d}{dx} F_{\dot{u}} + \frac{d^2}{dx^2} F_{\ddot{u}} = 0$$

1.3.2. REGOLARIZZAZIONE E FILTRAGGIO

$$z_\lambda = R_\lambda y$$

$$\text{con } R_\lambda = (A^*A + \lambda P^*P)^{-1} A^*$$

Supponiamo che

$$P \equiv I$$

A = operatore di convoluzione

e quindi

$$Az = y \Rightarrow (Az)(x) = \int_{-\infty}^{+\infty} k(x-\xi) z(\xi) d\xi = y(x)$$

Tentiamo di trovare la soluzione attraverso R_λ applicando le trasformate e passando nel dominio delle frequenze:

$$\left. \begin{aligned} K(\omega) &= \mathcal{F}\{k(x)\} \\ Y(\omega) &= \mathcal{F}\{y(x)\} \\ Z(\omega) &= \mathcal{F}\{z(x)\} \end{aligned} \right\} \Rightarrow KZ = Y \Rightarrow Z^*(\omega) = \frac{Y(\omega)}{K(\omega)}$$

dove K riassume le caratteristiche computazionali nel dominio delle trasformate, e in esse equivale all'operatore A.

unque avremo

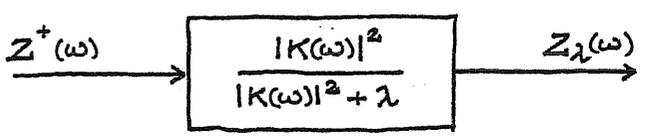
$$R_2(\omega) = \frac{1}{|K(\omega)|^2 + \lambda} \quad K^*(\omega) = \frac{1}{|K(\omega)|^2 + \lambda} \frac{K^*(\omega) K(\omega)}{K(\omega)}$$

perciò

$$z_2(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|K(\omega)|^2}{|K(\omega)|^2 + \lambda} \cdot \frac{Y(\omega)}{K(\omega)} e^{j\omega x} d\omega = \mathcal{F}^{-1}\{Z_2(\omega)\}$$

$\Phi(\omega, \lambda) = \begin{cases} \text{FILTRO APPLICATO ALLA SOLUZIONE NON REGOLARIZZATA} \\ \text{FILTRO APPLICATO ALLA SOLUZIONE NON REGOLARIZZATA} \end{cases}$
 $Z^+(\omega)$ (NON REGOLARIZZATA)

pratica, applicando il filtro alla soluzione non regolarizzata avremo



possiamo vedere il tutto come una funzione di trasferimento

$$\frac{|K(\omega)|^2}{|K(\omega)|^2 + \lambda} \cdot \frac{1}{|K(\omega)|}$$

applicata ai dati, ovvero a Y(omega).

e definiamo

$$\psi(x, \lambda) = \mathcal{F}^{-1}\left\{\frac{\Phi(\omega, \lambda)}{K(\omega)}\right\}$$

ora otterremo

$$z_2(x) = \int_{-\infty}^{+\infty} \underbrace{\psi(x-\xi, \lambda)}_{\substack{\text{FILTRO EQUIVALENTE COMPLETO CHE,} \\ \text{APPLICATO AI DATI, REGOLARIZZA LA SOLUZIONE}}} y(\xi) d\xi$$

SEMPIO: ESTRAZIONE DEI BORDI (EDGE)

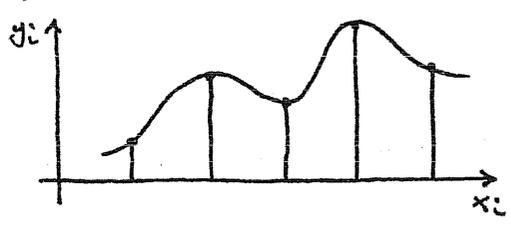
la derivata è soggetta a pb di mal condizionamento

$$y_\epsilon(x) = y(x) + \epsilon \cos(\omega_0 x) \quad \epsilon \ll 1$$

$$\frac{dy_\epsilon}{dx} = \frac{dy}{dx} - \underbrace{\epsilon \omega_0 \sin(\omega_0 x)}_{\text{TERMINE AMPLIFICATO SE } \omega_0 \gg \epsilon}$$

unque il rapporto segnale/rumore, dopo la derivazione, è diminuito.

supponiamo di avere i campioni y_i (= pixel dell'immagine): vogliamo trovare una funzione interpolante per poi fare la derivata di tale funzione continua.



Data $A =$ operatore di campionamento si desidera che

$$Az = z(x_i) = z_i$$

quindi

$$\sum_i |y_i - z_i|^2 + \lambda \|Pz_i\|^2 = \text{minimo}$$

TERMINE DI UGUAGLIANZA DATI TERMINE DI REGOLARIZZAZIONE

dati che abbiamo non sono regolarizzati: l'interpolazione non è esatta nei p_i e la funzione è + dolce + regolarizzata.

$$z_2(x) = \int_{-\infty}^{+\infty} \psi(x-\xi, \lambda) y(\xi) d\xi$$

$$\frac{d}{dx} z_2(x) = \int_{-\infty}^{+\infty} \underbrace{\frac{d}{dx} \psi(x-\xi, \lambda)}_{\text{FILTRO}} y(\xi) d\xi$$

Come termine regolarizzante dobbiamo fare in modo che le derivate di z siano piccole

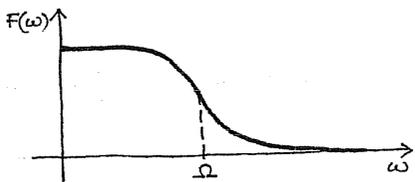
$$\|Pz\| = \sum_i \left(\frac{z_{i+1} - z_i}{h} \right)^2$$

Otteniamo un filtro con frequenza di taglio λ : in realtà λ può essere scelta in base a considerazioni matematiche o fisiche in modo tale che la soluzione sia fisicamente possibile.

In q. ottica si cercano funzioni che abbiano derivata 2^a piccola.

CASO EDGE DETECTION

Se non ci fosse rumore, l'immagine sarebbe un segnale a BANDA LIMITATA, quindi ci sarebbe un legame con la derivata prima



Se $f(x)$ è BAND-LIMITED con $F(\omega) = \mathcal{F}\{f(x)\}$

dove Ω = frequenza di CUT-OFF di $F(\omega)$

$$M = \sup \{F(\omega)\}$$

allora avremo

$$f' < \Omega M$$

In pratica se $f(x)$ è band-limited, allora $f'(x)$ è limitata e perciò $f''(x)$ è piccola. Ne segue che, nell'EDGE DETECTION,

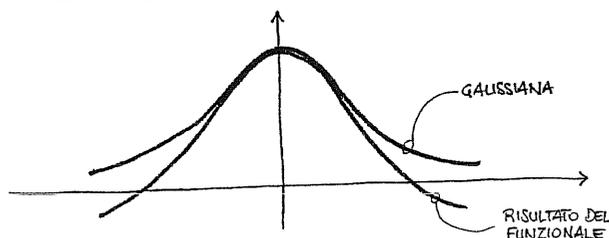
$$\sum_i |y_i - z_i|^2 + \lambda \sum_i \underbrace{\left(\frac{z_{i+1} + z_{i-1} - 2z_i}{2h} \right)^2}_{\text{PICCOLO}}$$

e, in forma continua,

$$\int (y-z)^2 dx + \lambda \int \left(\frac{d^2 z}{dx^2} \right)^2 dx = \text{minimo}$$

[NOTA: TUTTI I SISTEMI FISICI POSSONO TROVARE UNA FORMULAZIONE ENERGETICA]

Troviamo dunque una soluzione



che è molto simile alla gaussiana, quindi è naturale mettere in relazione il risultato ottenuto con le derivate di gaussiane.

OSSERVAZIONE: DERIVATA DI GAUSSIANE E REGOLARIZZAZIONE

$$y_E(x) = y(x) + E \cos(\omega_0 x)$$

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

GAUSSIANA

$$\underbrace{G_x(x, \sigma)}_{\text{NOSTRO FILTRO}} = \frac{1}{\sqrt{2\pi\sigma^2}} \left(-\frac{2x}{2\sigma^2} \right) e^{-\frac{x^2}{2\sigma^2}}$$

DERIVATA DI GAUSSIANA

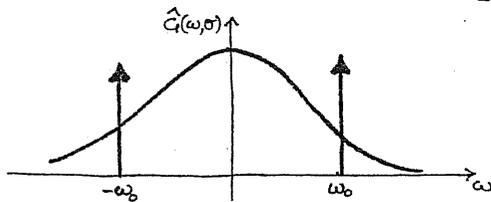
Sappiamo che q. filtraggio significa fare una convoluzione nello spazio:

$$y_E(x) * G_x(x, \sigma) = \underbrace{y(x) * G_x(x, \sigma)}_{\frac{d}{dx}(y * G)} + \underbrace{\varepsilon \cos(\omega_0 x) * G_x(x, \sigma)}_{\frac{d}{dx}(\varepsilon \cos(\omega_0 x) * G(x, \sigma)) = (*)}$$

per analizzarlo + semplicemente passiamo a Fourier:

$$G(x, \sigma) \xrightarrow{\mathcal{F}} \hat{G}(\omega, \sigma) \triangleq e^{-\frac{\omega^2 \sigma^2}{2}}$$

$$\begin{aligned} \mathcal{F} \left\{ \frac{d}{dx} [\varepsilon \cos(\omega_0 x) * G(x, \sigma)] \right\} &= j\omega \left\{ \varepsilon \cdot \frac{1}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \hat{G}(\omega, \sigma) \right\} \\ &= j\omega_0 \varepsilon \hat{G}(\omega_0) [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \end{aligned}$$



ornando nello spazio avremo

$$(*) = j\omega_0 \varepsilon e^{-\frac{\omega_0^2 \sigma^2}{2}} \cos(\omega_0 x)$$

Dobbiamo mettere a posto j : sappiamo che $j = e^{j\pi/2}$, dunque fa cambiare fase

$$\begin{aligned} (*) &= \varepsilon \omega_0 e^{-\frac{\omega_0^2 \sigma^2}{2}} e^{j\pi/2} \cos(\omega_0 x) \\ &= \varepsilon \omega_0 e^{-\frac{\omega_0^2 \sigma^2}{2}} \cos(\omega_0 x + \frac{\pi}{2}) \\ &= -\varepsilon \omega_0 e^{-\frac{\omega_0^2 \sigma^2}{2}} \sin(\omega_0 x) \end{aligned}$$

Se $\sigma = 0 \Rightarrow$ derivata puntuale

$\sigma \neq 0 \Rightarrow$ smorza l'effetto negativo di $\varepsilon \omega_0$

L'uso delle derivate di gaussiane ha introdotto il termine $e^{-\frac{\omega_0^2 \sigma^2}{2}}$ che consente di smorzare l'effetto negativo di prima, e quindi di effettuare una regolarizzazione.

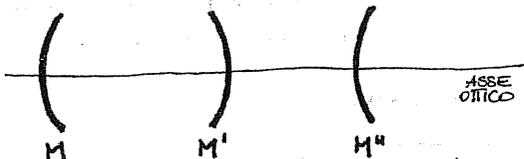
② ANALISI DEI SISTEMI OTTICI

In q. parte del corso tratteremo 2 aspetti:

- ① FORMAZIONE DELLE IMMAGINI
- ② ASPETTI ENERGETICI

2.1. FORMAZIONE DELLE IMMAGINI

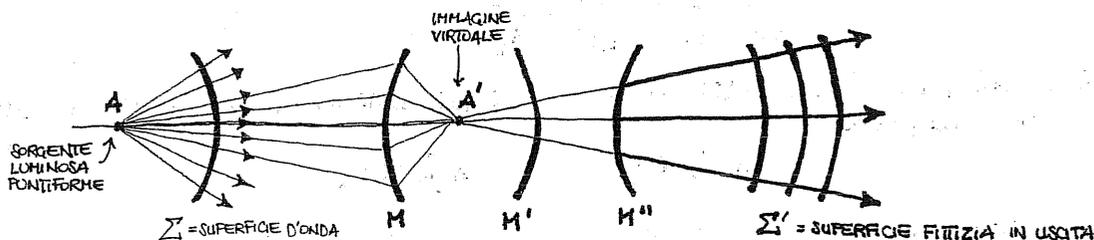
Definiamo il SISTEMA OTTICO come successione di superfici RIFRANGENTI o RIFLET-
TENTI:



In genere tali superfici sono sferiche, oppure piane e condizioni di raggio infinito.

Parliamo di SISTEMA OTTICO CENTRATO se i centri delle superfici sferiche stanno su una stessa retta, detta ASSE OTTICO. In q. caso il sistema risulterà simmetrico rispetto alle notazioni dell'asse ottico stesso.

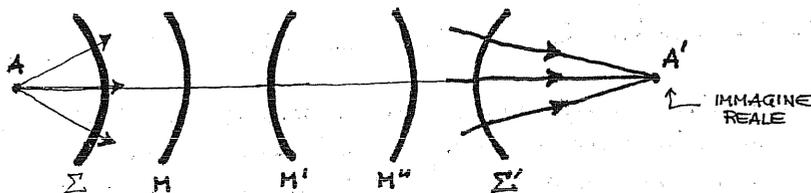
Analizziamo ora il concetto di STIGMATISMO/ASTIGMATISMO: esso riguarda le trasformazioni che subiscono i raggi che escono da un cosiddetto p^{to} luminoso e che formano la SUPERFICIE D'ONDA, cioè il luogo dei p^{ti} dello spazio in cui la fase è costante



I raggi incidono sul sistema ottico dando origine ad una superficie Σ'.
I raggi associati a tale superficie possono

- convergere in un p^{to} A' (= IMMAGINE D'A) ⇒ SISTEMA STIGMATICO
La superficie Σ' è sferica ed i raggi sono ad essa ⊥ ;
l'immagine inoltre è puntiforme come la sorgente
- non convergere in un p^{to} ⇒ SISTEMA ASTIGMATICO

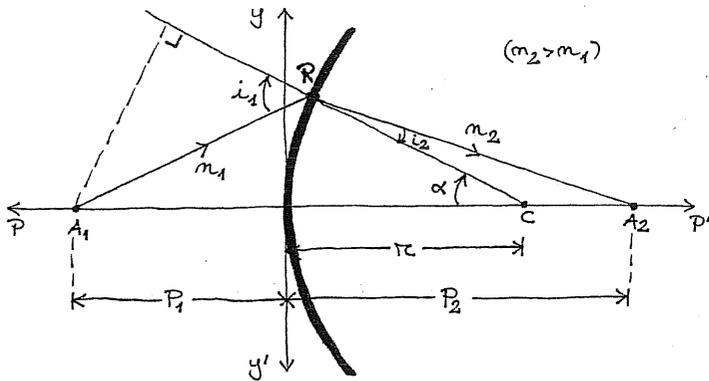
Nel caso in cui i raggi convergono in A' si parla di IMMAGINE VIRTUALE, mentre ci sono altre situazioni in cui le superfici Σ' hanno cavità opposte e in q. caso si parlerà di A' come di IMMAGINE REALE



Nota: nel caso di immagine virtuale, anche se possiamo vederla, non la possiamo raccogliere; nel caso di immagine reale la possiamo raccogliere, mettendo A' su un foglio, una pellicola sensibile.

2.1.1. DIOTIRO SFERICO

DIOTIRO SFERICO = superficie sferica che delimita 2 mezzi nello spazio, aventi 2 indici di rifrazione diversi.



Definiamo un sistema di assi (P, y) x descrivere la posizione delle sorgenti luminose e (P', y') x le loro immagini.

Preso A_1 ne costruiamo l'immagine.

Il raggio che si muove lungo l'asse (P, P') proseguirà il suo percorso senza deviazioni.

Per un raggio che diverge rispetto all'asse ottico vale

$$m_1 \sin i_1 = m_2 \sin i_2$$

LEGGE DI SNELL
(o DI RIFRAZIONE)

Se consideriamo il triangolo A_1RC , dal th dei seni otterremo

$$\overline{A_1R} \sin i_1 = \overline{A_1C} \sin \alpha$$

$$\overline{A_2R} \sin i_2 = \overline{A_2C} \sin \alpha$$

Da tali relazioni ricaviamo che

$$\frac{\sin i_1}{\sin i_2} = \frac{\overline{A_1C}}{\overline{A_1R}} \cdot \frac{\overline{A_2R}}{\overline{A_2C}}$$

Quando i raggi presi in considerazione si discostano poco dall'asse ottico si può effettuare l'approx

$$\overline{A_1R} \cong P_1 \quad \overline{A_2R} \cong P_2$$

RAGGI PARASSIALI

e dunque sostituendo nella relazione precedente otteniamo

$$\frac{\sin i_1}{\sin i_2} \cong \frac{P_1 + r}{P_1} \cdot \frac{P_2}{P_2 - r} \stackrel{\text{LEGGE DI SNELL}}{\cong} \frac{m_2}{m_1}$$

Manipolando quest'ultima relazione

$$\frac{m_2}{m_1} = \frac{1 + r/P_1}{1 - r/P_2} \Rightarrow m_2 - \frac{m_2 r}{P_2} = m_1 + \frac{m_1 r}{P_1} \Rightarrow \frac{m_2 - m_1}{r} = \frac{m_1}{P_1} + \frac{m_2}{P_2}$$

Questa relazione esprime un legame fisico tra la posizione della sorgente e quella dell'immagine in base alle caratteristiche del DIOTIRO SFERICO.

Mentre tale relazione dipende, ad es, dal raggio, non dipende dall'angolo α , dunque è la stessa x tutti i raggi parassiali

Conclusione: il diotiro sferico è un sistema stigmatico x i raggi parassiali

Nota: la legge di rifrazione dipende dalle costanti dielettriche, che quindi cambiano al cambiare della frequenza, con conseguente spostamento del pte: ne segue che, x avere un sistema stigmatico, la sorgente luminosa deve essere monocromatica.

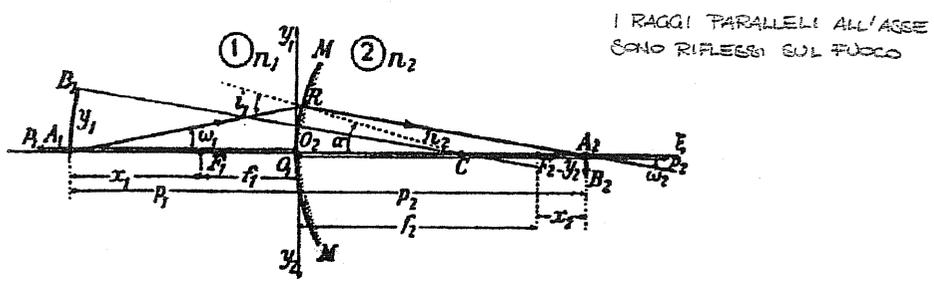


FIG. 2.1

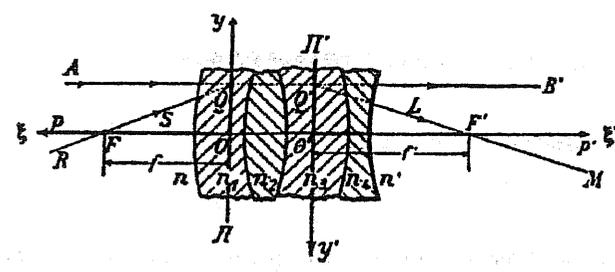


FIG. 2.2

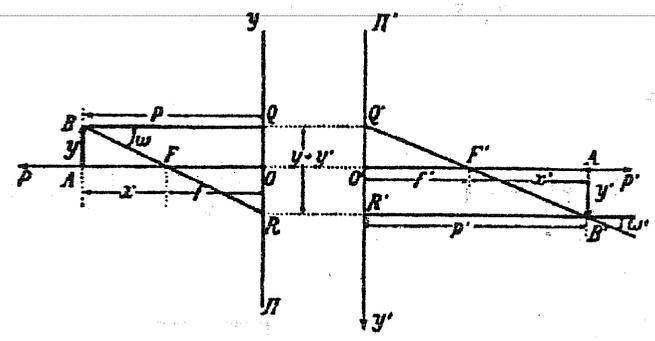


FIG. 2.3

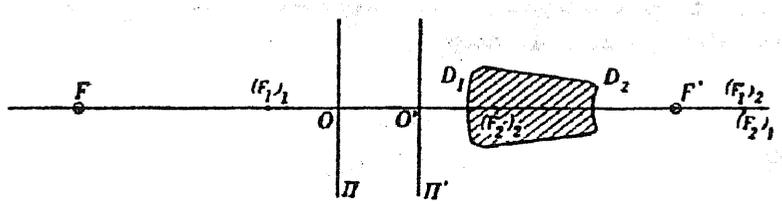
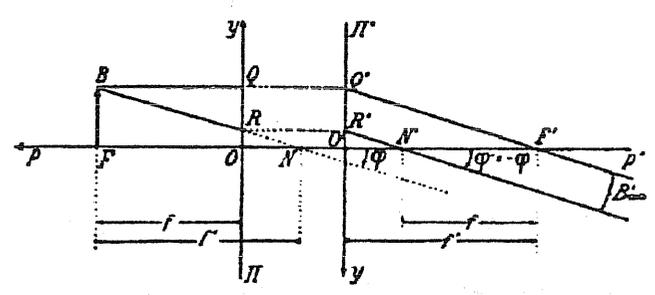


FIG. 2.4



SE LA SorgENTE È NEL FUOCO, L'IMMAGINE È A +∞

FIG. 2.5

INGRANDIMENTO LINEARE TRASVERSALE ED ININGRANDIMENTO ANGOLARE

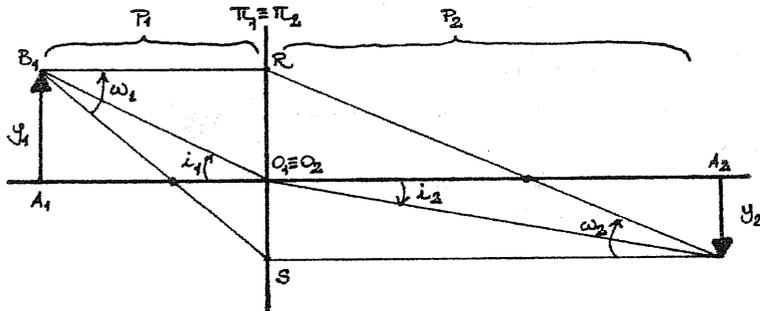


FIG. 1

$$G_y \triangleq \frac{y_2}{y_1} = \text{INGRANDIMENTO LINEARE TRASVERSALE}$$

$$G_\omega \triangleq \frac{\omega_2}{\omega_1} = \text{INGRANDIMENTO ANGOLARE}$$

(1) Tenendo conto che del raggio B_1O_1 , si può calcolare direttamente la rifrazione si ha (sempre nell'ambito dell'approx parassiale)

$$G_y = \frac{y_2}{y_1} \approx \frac{P_2 \sin i_2}{P_1 \sin i_1} = \frac{P_2}{P_1} \cdot \frac{n_1}{n_2}$$

(2) Dalla FIG. 1 si ha

$$\frac{y_1 + y_2}{P_1} = \tan \omega_1 \approx \omega_1$$

$$\frac{y_1 + y_2}{P_2} = \tan \omega_2 \approx \omega_2$$

e di qui

$$G_\omega = \frac{\omega_2}{\omega_1} = \frac{P_2}{P_1} \stackrel{(1)}{=} \frac{n_1}{n_2} \cdot \frac{1}{G_y} \Rightarrow \boxed{G_\omega G_y = \frac{n_1}{n_2}}$$

Ne segue che ciò che si guadagna in ingrandimento (lineare trasversale) si perde in apertura del fascio di raggi.

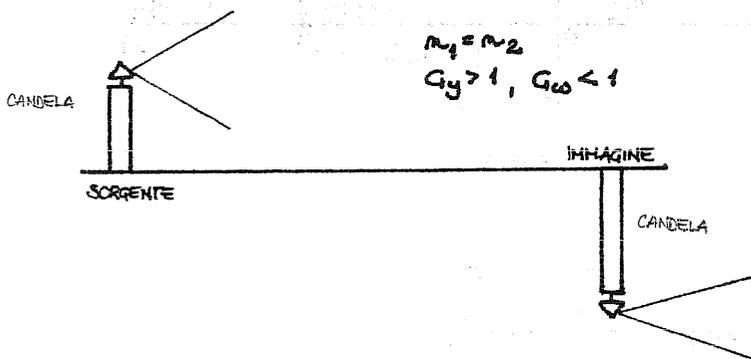


FIG. 2

ALL'INGRANDIMENTO DELLA IMMAGINE CORRISPONDE UNA DIMINUZIONE ANGOLARE

NB: (a) Nel caso di FIG. 1, tracciata con $n_2 > n_1$, risulta $G_y < 1$ e $G_\omega < 1$

(b) Se in FIG. 1 consideriamo il punto O, ovvero il piano π , risulta

$$G_y \Big|_{\pi} = 1 \quad \text{e} \quad G_\omega = \frac{n_1}{n_2} = \frac{\sin i_2}{\sin i_1} \approx \frac{i_2}{i_1}$$

RIFRAZIONE RAGGI PARASSIALI

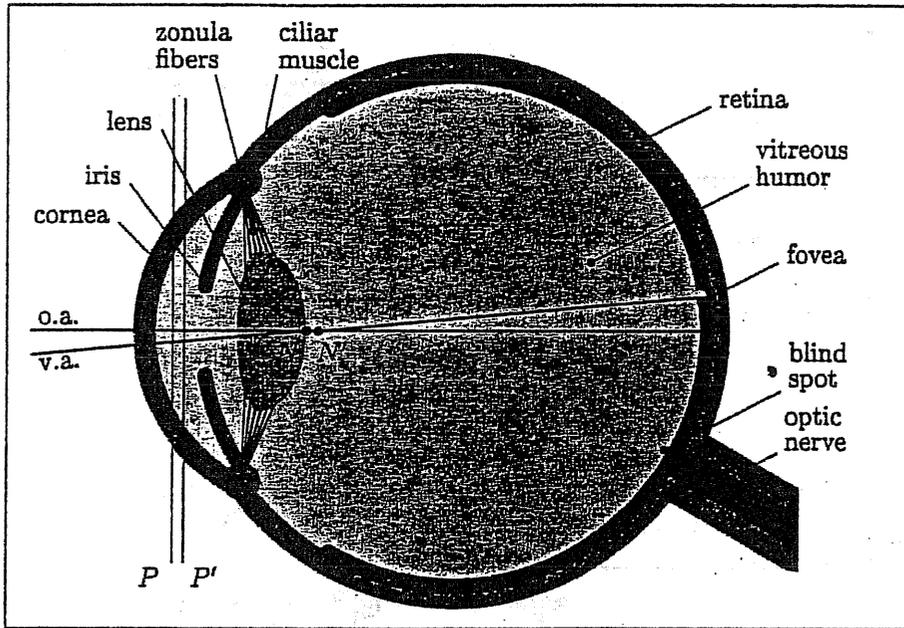
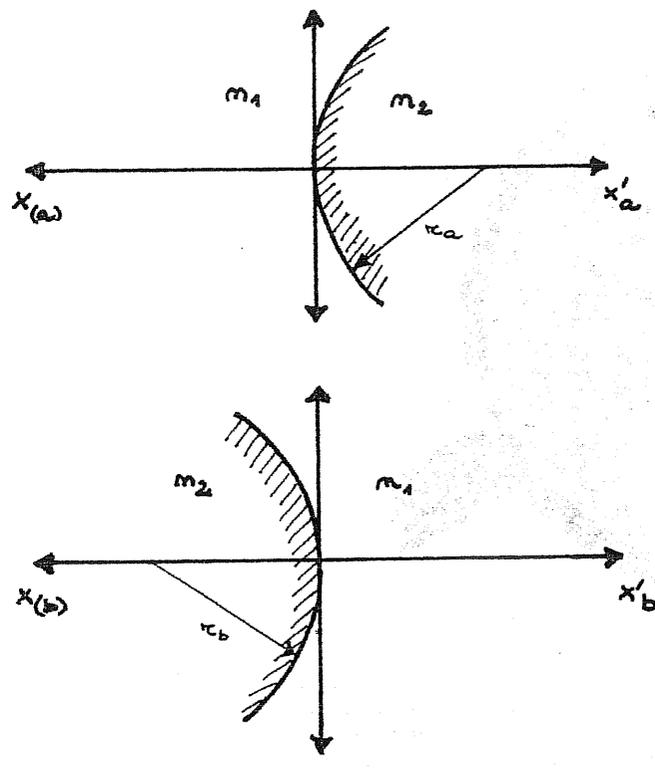


FIG. 2.6. Simplified horizontal section through a right eye accommodated for distant vision. P, P' : principal planes; N, N' : nodal points. o.a.: optical axis. v.a.: visual axis (direction in which the fovea looks). Not labeled: anterior chamber (between the cornea and the iris) and posterior chamber (between the iris and the lens). These two chambers are filled with the aqueous humor.

	accommodation		
	far	any	near
<i>index of refraction</i>			
cornea		1.376	
aqueous humor		1.336	
lens	varying, 1.376 - 1.406		
vitreous humor		1.336	
<i>radius of curvature in mm</i>			
cornea: front side		7.7	
cornea: back side		6.8	
lens: front side	10.0		5.33
lens: back side	-6.0		-5.33
<i>position in mm</i>			
cornea: front side		0.0	
cornea: back side		0.5	
lens: front side	3.6		3.2
lens: back side		7.2	
object focal point F	-15.707		-12.397
image focal point F'	-24.387		21.016
object nodal point N	7.078		6.533
image nodal point N'	7.332		6.847
object principal point P	1.348		1.772
image principal point P'	1.602		2.086
fovea		24.0	
<i>refractive power in diopters</i>			
cornea		43.053	
lens	19.11		33.06
eye	58.636		70.57

TABLE 2.1. Some optical data on the human eye, after Gullstrand (1908), from Le Grand & El Hage (1980). The refractive index of the lens can not be described by a single quantity, since it varies within the lens.

DIOTTRO SFERICO



Vogliamo trovare le coordinate del p^{to} immagine e quelle della sorgente

$$x' = \frac{f_2 x}{x - f_1} = \frac{Ax + B}{Cx + D}$$

TRASFORMAZIONE BILINEARE

SPECIFICA DEL DIOTTRO

$$f_{1a} = \frac{m_1}{m_2 - m_1} r_a$$

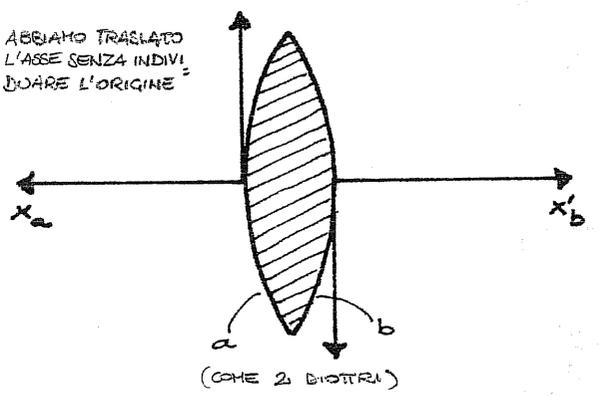
$$f_{2a} = \frac{m_2}{m_2 - m_1} r_a$$

$$f_{1b} = \frac{m_2}{m_1 - m_2} r_b$$

$$f_{2b} = \frac{m_1}{m_1 - m_2} r_b$$

LENTE GOTTILE (SPESORE TRASCURVABILE)

ABBIAHO TRASLATO L'ASSE SENZA INDIVIDUARE L'ORIGINE



dove $x_b = -x'_a = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} * x'_a$

$$x'_b = \begin{pmatrix} f_{2b} & 0 \\ 1 & -f_{1b} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f_{2a} & 0 \\ 1 & -f_{1a} \end{pmatrix} * x_a$$

$$= \begin{pmatrix} f_{2b} & 0 \\ 1 & -f_{1b} \end{pmatrix} \begin{pmatrix} -f_{2a} & 0 \\ 1 & -f_{1a} \end{pmatrix} * x_a$$

$$x'_b = \begin{pmatrix} -f_{2b} f_{2a} & 0 \\ -f_{2a} - f_{1b} & f_{1a} f_{1b} \end{pmatrix} * x_a = \begin{pmatrix} \frac{m_1 m_2}{(m_2 - m_1)^2} r_a r_b & 0 \\ \frac{m_2}{m_2 - m_1} (r_b - r_a) & -\frac{m_1 m_2}{(m_2 - m_1)^2} r_a r_b \end{pmatrix} * x_a$$

$$= \frac{\left[\frac{m_1 m_2}{(m_2 - m_1)^2} r_a r_b \right] x_a}{\left[\frac{m_2}{m_2 - m_1} (r_b - r_a) \right] x_a - \frac{m_2 m_1}{(m_2 - m_1)^2} r_a r_b} = \frac{f x_a}{x_a - f}$$

avendo definito

$$f \triangleq \frac{m_1}{m_2 - m_1} \cdot \frac{r_a r_b}{r_b - r_a}$$

$$\frac{1}{f} = \left(\frac{m_2}{m_1} - 1 \right) \left(\frac{1}{r_a} - \frac{1}{r_b} \right)$$

Il distro sferico ci consente di trattare anche pti che non si trovano sull'asse ottico ma, ad es, su una superficie sferica di centro C e raggio $P_1 + r$ [v. FIG. 2.1.]

Consideriamo il pto B_1 , fuori dall'asse: l'immagine di q. pto può essere calcolata come quella x A_1 , x anche $\overline{CB_1}$, è un asse ottico del sistema (detto ASSE OTTICO SECONDARIO poiché non passa x il centro).

L'immagine si forma in B_2 che, x costruzione, è sulla superficie sferica di centro C e raggio $\overline{CA_2}$.

Se i pti di B si discostano poco dall'asse, possiamo confonderli con superfici tangenti all'asse stesso, e quindi $\overline{CB_2} \cong \overline{CA_2}$.

Il distro sferico è dunque un SISTEMA OTTICO ORTOSCOPICO x, limitatamente ai pti parassiali, i piani sorgente sono in corrispondenza con i piani im- magine.

Definiamo ora

$$G_y \triangleq \frac{y_2}{y_1} = \frac{P_2 - r}{P_1 + r} = \text{INGRANDIMENTO LINEARE TRASVERSALE}$$

Possiamo trovare la coordinata longitudinale dell'immagine rispetto alla sorgente.

• $P_1 = \infty$	\Rightarrow	$P_2 = \frac{m_2 r}{m_2 - m_1} \triangleq f_2$	<u>DISTANZA FOCALE NELLO SPAZIO IMMAGINI</u>	} POSSONO ESSERE NEGATIVE (0 x LO È r, 0 x I VALORI DI m_1, m_2)
• $P_2 = \infty$	\Rightarrow	$P_1 = \frac{m_1 r}{m_2 - m_1} \triangleq f_1$	<u>DISTANZA FOCALE NELLO SPAZIO SORGENTI</u>	

Ricaviamo che

$$\frac{f_1}{f_2} = \frac{m_1}{m_2} > 0 \Rightarrow f_2 - f_1 = r \quad \left[\text{IL SEGNO DELLE DISTANZE FOCALI È LO STESSO} \right]$$

Con q. definizioni possiamo scrivere

$$\boxed{\frac{f_1}{P_1} + \frac{f_2}{P_2} = 1} \quad \text{FORMULA DI HUYGENS}$$

Tornando all'ingrandimento lineare trasversale, avremo che

$$G_y = \frac{y_2}{y_1} = \frac{P_2 - r}{P_1 + r} = \frac{P_2 - (f_2 - f_1)}{P_1 + (f_2 - f_1)} = \frac{P_2}{P_1} \cdot \frac{1 - \frac{f_2}{P_2} + \frac{f_1}{P_2}}{1 + \frac{f_2}{P_1} - \frac{f_1}{P_1}}$$

$$\stackrel{\text{HUYGENS}}{=} \frac{P_2}{P_1} \cdot \frac{\frac{f_1}{P_1} + \frac{f_1}{P_2}}{\frac{f_2}{P_1} + \frac{f_2}{P_2}} = \frac{P_2}{P_1} \cdot \frac{f_1}{f_2}$$

Possiamo anche scrivere diversamente:

$$G_y = \frac{P_2}{P_1} \cdot \frac{f_1}{f_2} = \frac{P_2}{f_2} \cdot \frac{f_1}{P_1} \stackrel{\text{HUYGENS}}{\uparrow} \frac{P_2}{f_2} (1 - \frac{f_2}{P_2}) = \frac{P_2 - f_2}{f_2} = \frac{f_1}{P_1} \cdot \frac{1}{1 - \frac{f_1}{P_1}} = \frac{f_1}{P_1 - f_1}$$

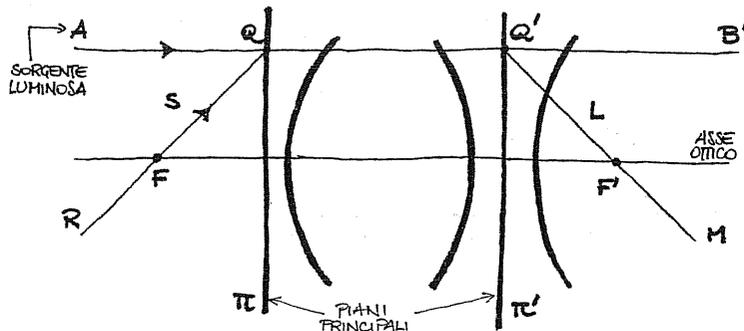
e dunque, riassumendo,

$$\boxed{G_y = \frac{y_2}{y_1} = \frac{P_2}{P_1} \cdot \frac{f_1}{f_2} = \frac{P_2 - f_2}{f_2} = \frac{f_1}{P_1 - f_1}} \quad \text{FORMULE DI INGRANDIMENTO}$$

NB: si ricordi che, se la sorgente è nel fuoco, l'immagine è a $+\infty$.

Se il sistema è costituito da tante superfici riflettenti, allora rimane ortoscopico & ogni superficie è considerata come un diotro sferico con una sorgente ed una immagine che sarà a sua volta sorgente x il diotro successivo.

Dato l'asse ottico e le superfici rifrangenti, prendiamo una retta parallela all'asse. Il raggio incidente intersecherà l'asse ottico nel fuoco secondario F' :



Infatti il raggio incidente A sarà come un raggio complessoivamente rifratto LM che incontrerà l'asse in un pt: dunque A ed LM sono coniugati.

La stessa cosa accade a B' , il cui raggio viene riflesso sul fuoco F' .

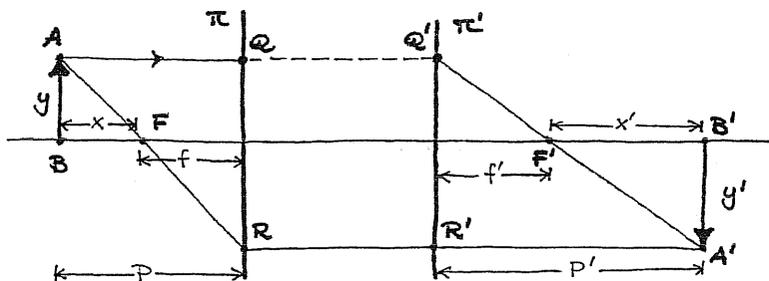
Se Q (= sorgente virtuale) e Q' (= immagine virtuale) sono intersezione di raggi coniugati, allora sono a loro volta punti coniugati.

Possiamo ora definire un sistema di riferimento composto dai piani π, π' . Il pt F' è l'intersezione di 2 raggi che partono paralleli (l'asse ottico e il raggio AB' , poi rifratto), quindi è l'immagine di una sorgente posta all' $\infty \Rightarrow$ FOCO SECONDARIO.

Per lo stesso ragionamento avremo che F = FOCO PRIMARIO.

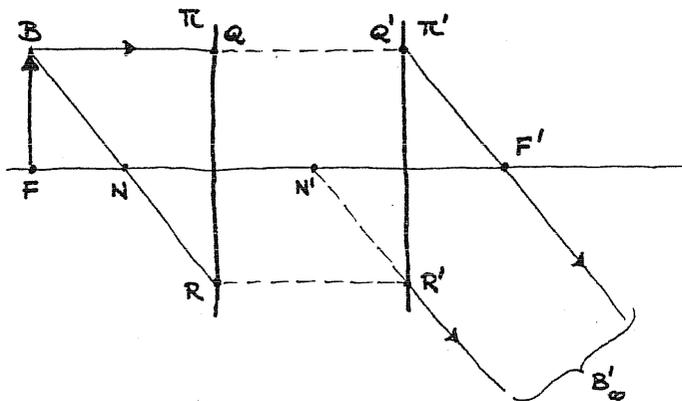
Per tutti i pt dei piani π e π' valgono le stesse proprietà che vigono tra Q e Q' .

Vediamo ora come avviene la formazione dell'immagine: segue la stessa metodologia del diotro sferico, xò la costruzione delle immagini avviene su π' , mentre quella delle sorgenti su π .



Per quanto complesso sia il sistema di lenti, il raggio segue sempre le stesse leggi, dipendenti dalla distanza focale.

Un altro parametro che può risultare necessario x caratterizzare il sistema ottico è quello dei PUNTI NODALI



Poichè il fascio è proprio dove sta l'immagine otteniamo 2 raggi paralleli x B' , immagine di B , è all' ∞ .

N, N' = PUNTI NODALI: sono coniugati x hanno l'asse ottico in comune e inoltre sono in corrispondenza di raggi coniugati.

Un raggio che entra in N viene prolungato dal piano π' , e riemerge indovato andando a colpire l'asse ottico in N' (mantiene quindi la stessa angolazione).

Se le distanze focali sono le stesse, N sta sul piano π ed N' su π' .

DIAFRAMMI

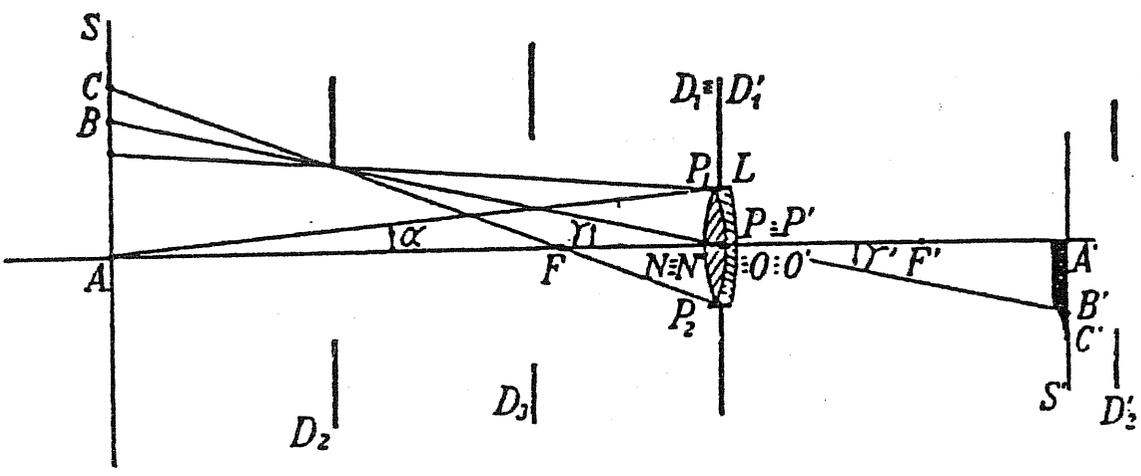


FIG. 2.7.

LA DIFFERENZA NEGLI ANGOLI α E α' RIFLETTE LA DIFFERENZA IN INGRANDIMENTO (LINEARE E TRASVERSALE)

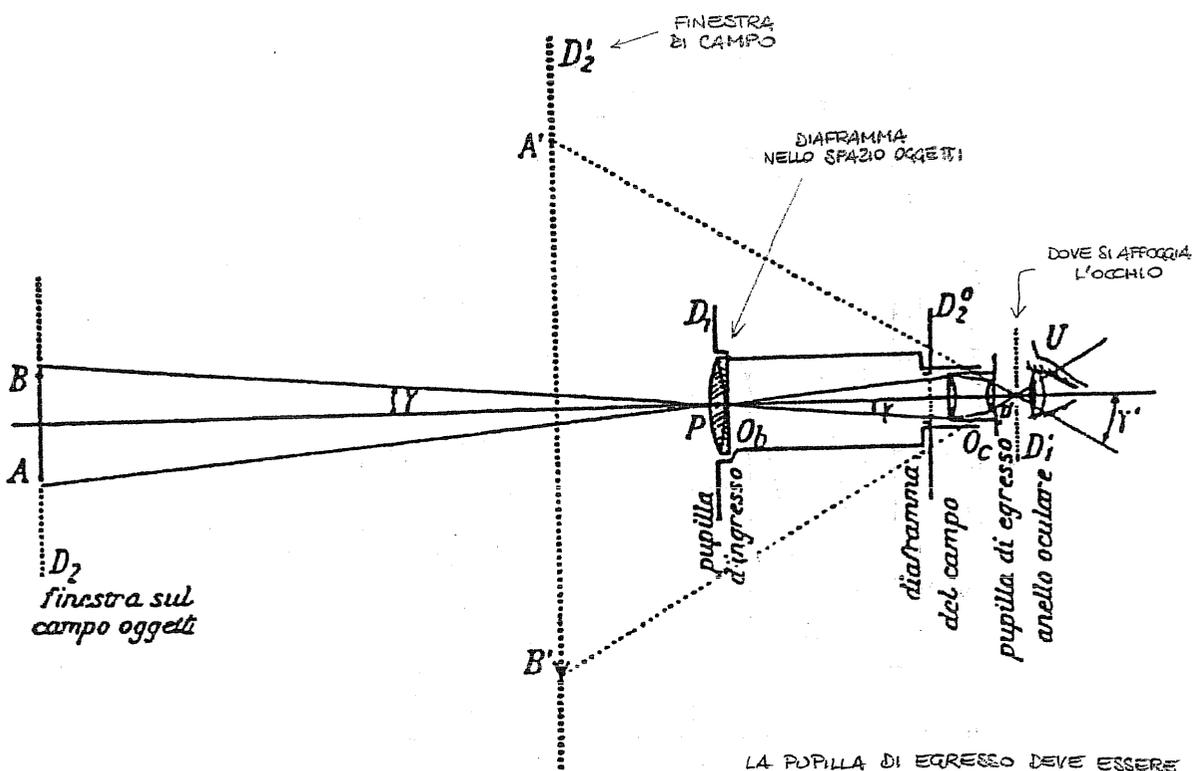


FIG. 2.8.

LA PUPILLA DI EGRESO DEVE ESSERE UGUALE IN DIMENSIONE A QUELLA DELL'OCCHIO

Da un p^{te} di vista quantitativo possiamo trattare sistemi costituiti da diverse lenti tramite il seguente procedimento:

$$\left. \begin{aligned} x' &= \frac{Ax+B}{Cx+D} \\ x'' &= \frac{Ex'+F}{Gx'+H} \end{aligned} \right\} \Rightarrow \boxed{x'' = \frac{\hat{A}x + \hat{B}}{\hat{C}x + \hat{D}}}$$

[V. ALLEGATO 42#]

TRASFORMAZIONE BILINEARE TOTALE

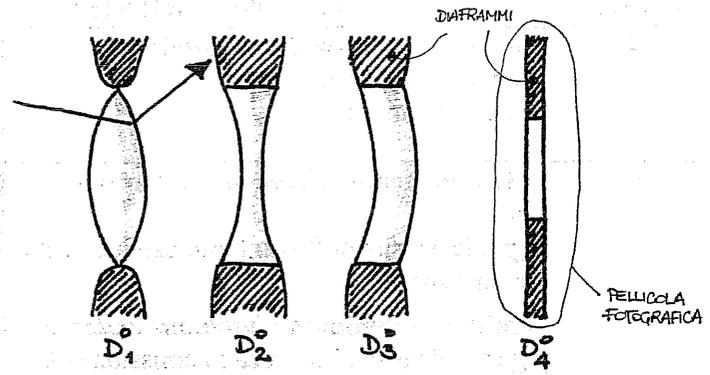
dove $\begin{pmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{pmatrix} = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}$

NB: quando un p^{te} immagine è considerato come sorgente dobbiamo cambiare l'orientamento degli assi e traslare l'origine. Questo metodo può essere usato x qualsiasi superficie di rifrazione, permettendo di combinare i risultati.

2.1.2. DIAFRAMMI

DIAFRAMMA = tutto ciò che limita il passaggio di raggi luminosi in un sistema ottico (= schermo opaco).

Il * minimo di diaframmi coincide con il * di lenti o elementi rifrangenti.



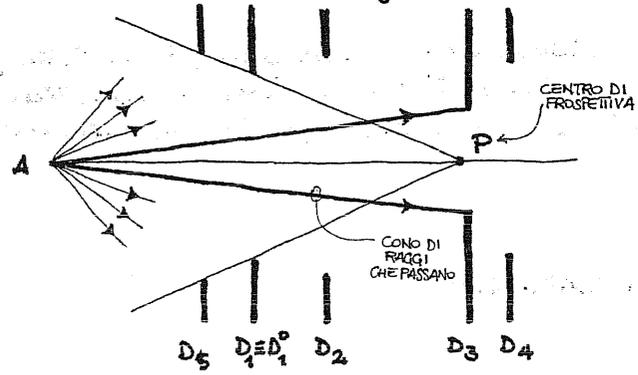
Abbiamo rappresentato uno schermo opaco con un foro di diametro regolato, ma possono essere sup. porti x le lenti, fori regolati o l'esterno di una pellicola fotografica.

Quando forniamo un'immagine non abbiamo bisogno di tutti i raggi, ma ne scegliamo solo un fascio.

Ora ci poniamo 2 domande:

- Quali raggi passano tra quelli del sistema
- Quali sorgenti vede il sistema ottico

Prendiamo un p^{te} A sorgente e riportiamo tutti i diaframmi nello spazio delle sorgenti. Con le leggi (matriciali) dell'ottica possiamo riportarli nelle nuove coordinate.

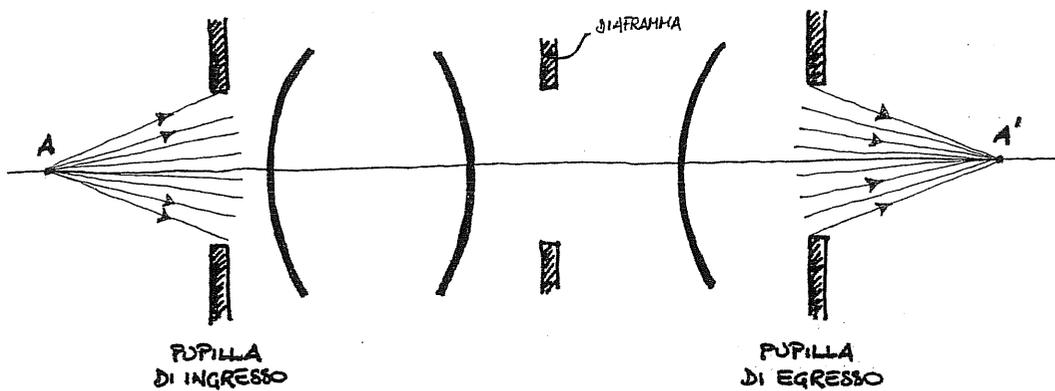


Solo alcuni raggi passeranno, cioè quelli all'interno del cono: essi passano attraverso il diaframma + ommolante (= IRIDE), il cui foro è detto PUPILLA

Il centro di prospettiva P ci dice quali sono i p^{te} i cui raggi possono entrare nel sistema ottico. In q. caso la zona dello spazio che vediamo è determinata da D₁, che diventa la nostra FINESTRA DI CAMPO.

Se stringiamo D₃ abbiamo sempre lo stesso campo visivo, ma minor quantità di luce

iamo regolare la quantità di raggi in ingresso o in uscita attraverso, rispettivamente, la PUPILLA DI INGRESSO o la PUPILLA DI UGESSO



1. CARATTERISTICHE ENERGETICHE E FOTOMETRICHE DELLE IMMAGINI

siamo fare riferimento ad alcune grandezze energetiche x la valutazione delle immagini:

ENERGETICHE

PERCETTIVE

UNITÀ RADIOMETRICHE	SIMBOLI	UNITÀ FOTOMETRICHE
<p>potenza radiante, flusso raggianti / radiant flux (Watt)</p> <p> ENERGIA ASSORBITA/EMESSA PER UNITÀ DI TEMPO [J/S]</p>	Φ	<p>potenza luminosa, flusso luminoso / luminous flux (lumen)</p> <p>flusso luminoso</p>
<p>irradianza, irradiazione / irradiance (Watt/m²)</p> <p>{per la densità di potenza incidente su una superficie (PER UNITÀ DI AREA [W/m²])}</p> <p>emittanza radiante / radiant exitance</p> <p>{per la densità di potenza in emissione (da sorgenti estese)} (PER UNITÀ DI AREA [W/m²])</p>	E	<p>illuminamento / illuminance (lumen/m² = lux)</p> <p>{per la densità di flusso luminoso incidente su una superficie}</p> <p>emittanza luminosa / luminous exitance</p> <p>{per la densità di potenza in emissione (da sorgenti estese)}</p>
<p>radianza / radiance (Watt / sr m²)</p> <p>ENERGIA EMESSA NELL'UNITÀ DI TEMPO DALL'UNITÀ DI SUPERFICIE NELLO ARCO SOLIDO</p> <p> STERADIANTE</p>	<p>L</p> <p>(B)</p> <p>(Σ)</p>	<p>luminanza / luminance (lumen / sr m²)</p> <p>[brillanza / brightness {non più usati}]</p>
<p>intensità radiante, intensità energetica / radiant intensity (Watt / sr)</p> <p></p>	I	<p>intensità luminosa / luminous intensity</p> <p>(lumen / sr = candela)</p>

le grandezze possono essere prese in considerazione anche dal pto di vista foto metrico, con nomi ed unità di misura differenti: le corrispondenze tra energia e fotometria sono state ricavate con formule apposite in base alla sensibilità dell'occhio.

ata

$$P_e = \int_0^\lambda p(\lambda) d\lambda$$

↑ ENERGETICA ↑ DENSITÀ DI POTENZA POTENZA RADIANTE

si definisce

$$P_f \stackrel{\Delta}{=} K_{max} \int_0^\lambda p(\lambda) V(\lambda) d\lambda$$

↑ FOTOMETRICA POTENZA LUMINOSA

dove K_{max} = costante dimensionale = $[\frac{cd}{W/sr}]$
 $V(\lambda)$ = efficacia luminosa spettrale

CANDELA = intensità luminosa (in una data direzione) di una sorgente che emette una radiazione monocromatica di frequenza pari a $540 \cdot 10^{12}$ Hz con una intensità radiante di $\frac{1}{683} \frac{W}{sr}$.

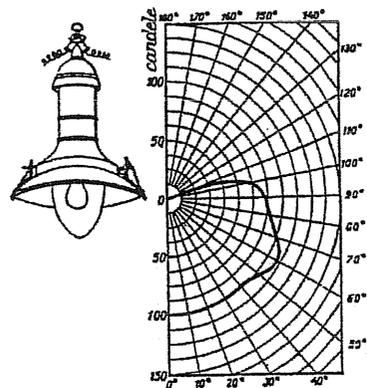
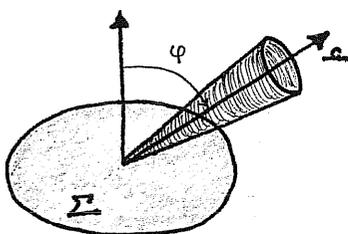


FIG. 1 lampada da tavolo con il suo diagramma di radiazione nelle \neq direzioni (energia nell'unità di tempo ρ_0 nell'unità di angolo solido)

LUMINANZA = flusso luminoso (in uscita) che passa attraverso una superficie unitaria perpendicolarmente alla direzione assegnata, ed è compreso in un angolo solido unitario.



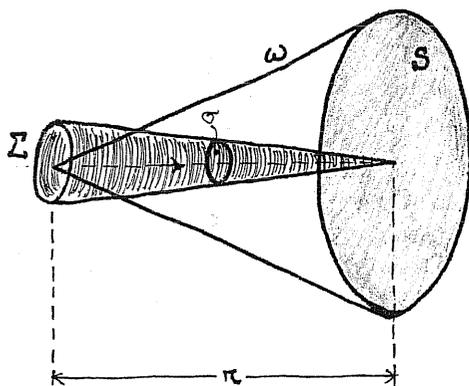
Data una superficie Σ , scegliamo una direzione di emissione che forma un angolo φ con la normale a Σ .

Considerando l'angolo solido Ω ed il flusso ϕ avremo

$$L = \frac{\phi}{\Sigma \cos \varphi \Omega} \quad \text{LUMINANZA} \quad \left[\frac{\text{lumen}}{\text{sr} \cdot \text{cm}^2} \right]$$

Le superfici che sono perfetti diffusori hanno una luminanza indipendente da φ . Ciò vale anche x la radianza.

ILLUMINAMENTO = data una superficie S, l'illuminamento di S è dato dal prodotto della luminanza media per l'angolo solido ω sotto il quale è vista la sorgente Σ .



$$E = \frac{\phi}{S} = \frac{L \sigma \frac{S}{r^2}}{S} = L \frac{\sigma}{r^2} = L \omega \quad \text{ILLUMINAMENTO}$$

In un SISTEMA OTTICO PERFETTAMENTE TRASPARENTE avremo

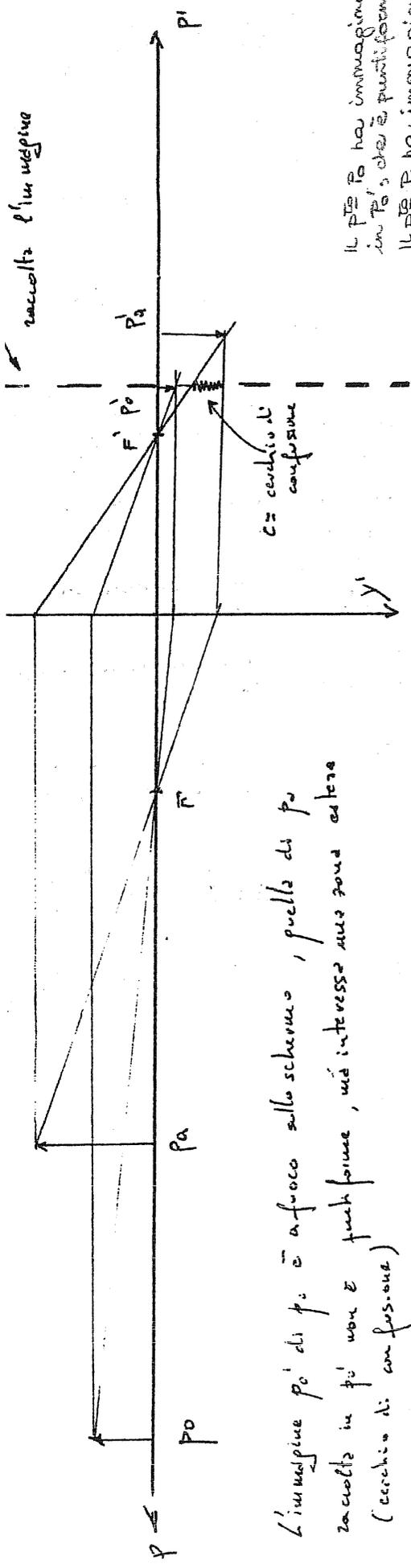
- ① $L_{\text{immagine}} = L_{\text{sorgente}}$
- ② $\frac{\text{ILLUMINAMENTO IMMAGINE}}{\text{LUMINANZA Sorgente}} = \frac{\pi}{4} \cdot \frac{h}{f}$

dove f = distanza focale
 h = diametro pupilla di ingresso

In relazione a ciò si parla di

$$N = f\text{-numero} \quad \text{o} \quad f\text{-stop} = f_{\#} = \frac{f}{h}$$

Nota: obiettivi luminosi hanno N piccolo (2,4,6).
 Il diaframma caratterizza la qualità delle immagini.

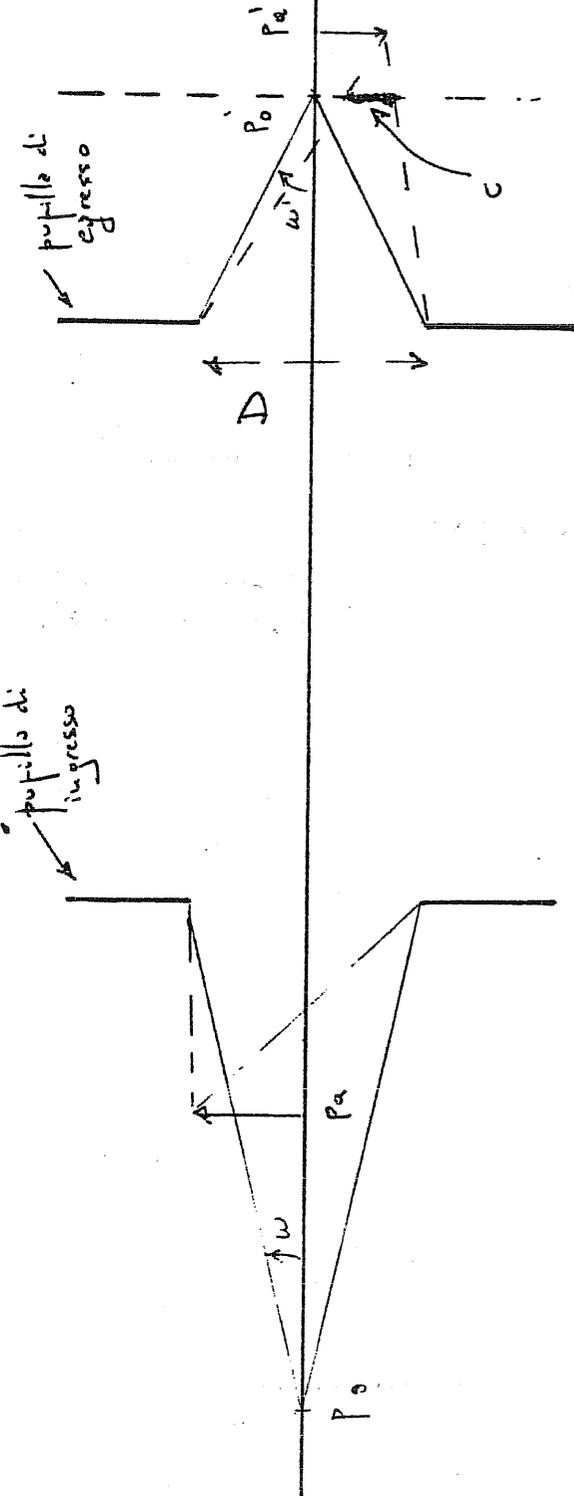


raccolta l'immagine

L'immagine p_0' di p_0 è afuoco sullo schermo, quello di p_0 raccolto in p_0' non è punto focale, né interesse suo zona estera (cerchio di confusione)

Il P_0' ha immagine in P_0 , che è punto focale.
 Il P_0 ha immagine in P_0' , ma non la raccogliamo in un P_0' , bensì forma un cerchio (CERCHIO DI CONFUSIONE), poiché lo schermo di raccolta della immagine è prima.
 Se allarghiamo il diaframma (pupilla di ingresso), otteniamo un cerchio di confusione maggiore.

pupilla di ingresso



$$c = D \frac{p_A - p_0'}{p_A}$$

Caratterizzazione degli obiettivi e loro "prestazioni"

Si consideri un obiettivo con *iris* e *fuoco* manuali.

Siamo interessati a caratterizzare le "prestazioni" di un obiettivo in relazione ai seguenti tre parametri:

- campo di messa a fuoco (profondità di campo)
- fattore di ingrandimento
- angolo di vista

0.1 Campo di messa a fuoco

Parametri caratteristici dell'obiettivo:

f distanza focale [mm]

N apertura

Altre definizioni:

S_o posizione del punto a fuoco nello spazio oggetti

S_i posizione del piano immagine rispetto alla lente

Vale la seguente relazione:

$$\frac{1}{S_o} + \frac{1}{S_i} = \frac{1}{f}$$

Esempio: Considerato un obiettivo con $f = 25\text{mm}$, e mettendo a fuoco a ∞ la posizione della lente dal piano immagine è pari a 25mm , infatti:

$$\frac{1}{\infty} + \frac{1}{S_i^{\infty}} = \frac{1}{25 \cdot 10^{-3}\text{m}} \quad \rightarrow \quad S_i^{\infty} = 25 \cdot 10^{-3}\text{m}.$$

Mettendo a fuoco un oggetto a 0.5m la distanza della lente dal piano immagine si allunga:

$$\frac{1}{0.5\text{m}} + \frac{1}{S_i^{0.5\text{m}}} = \frac{1}{25 \cdot 10^{-3}\text{m}} \quad \rightarrow \quad S_i^{0.5\text{m}} \simeq 26.31 \cdot 10^{-3}\text{m}.$$

Considerato una messa a fuoco all' ∞ , si può definire la *distanza iperfocale* h :

$$h = \frac{f^2}{N \cdot c}$$

dove c indica il diametro del *cerchio di confusione* e dà una misura del grado di accettabilità della "messa a fuoco". Tipicamente $c = 0.03\text{mm}$. QUESTA DIMENSIONE È LA SOGLIA DELLA PERCEZIONE \Rightarrow PROFONDITÀ E CAMPO

Esempio: Posto $N = 1.5$ e $f = 25\text{mm}$

$$h = \frac{25^2 \cdot 10^{-6}\text{m}^2}{1.5 \cdot 0.03 \cdot 10^{-3}\text{m}} = \frac{25^2 \cdot 10^{-3}\text{m}}{1.5 \cdot 30} \simeq 13.88\text{m}$$

All'aumentare dell'apertura N (iris), aumenta la profondità di campo e la distanza iperfocale risulta più corta.

Se consideriamo una "messa a fuoco" ad una distanza S_o , torna utile definire i parametri S_{close} e S_{far} che indicano rispettivamente la distanza minima e massima dei punti a fuoco nello spazio oggetti:

$$S_{close} = \frac{hS_o}{h + S_o - f} \quad S_{far} = \frac{hS_o}{h - S_o + f}$$

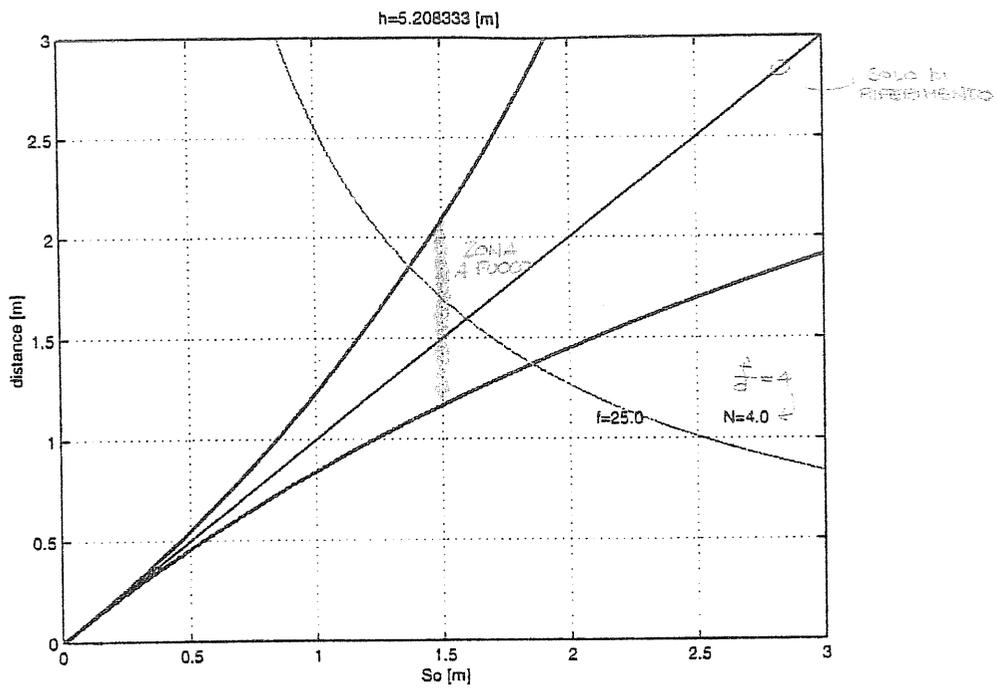
0.2 Fattore di ingrandimento

$$M = \frac{f}{S_o - f}$$

0.3 Angolo di vista

3)

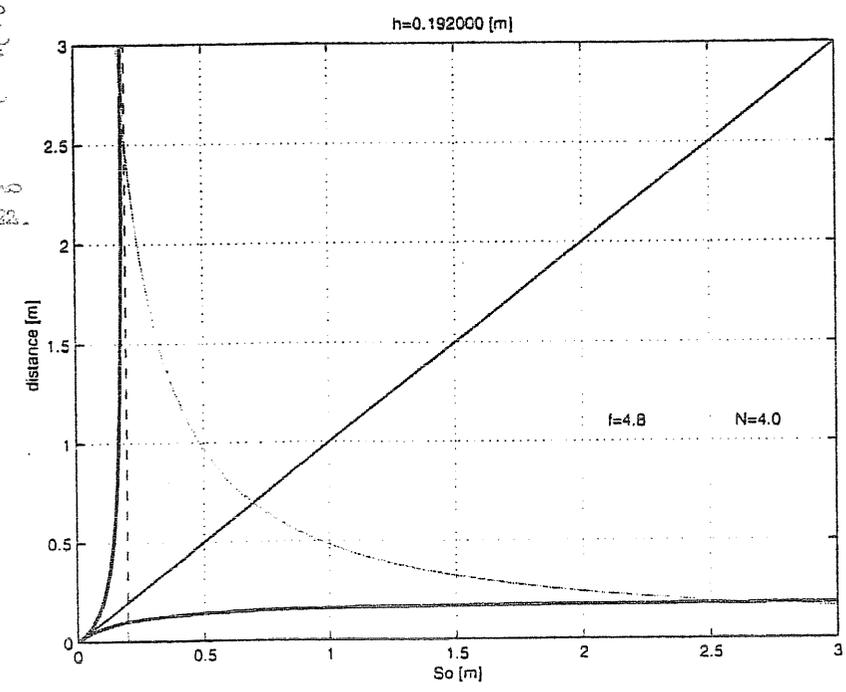
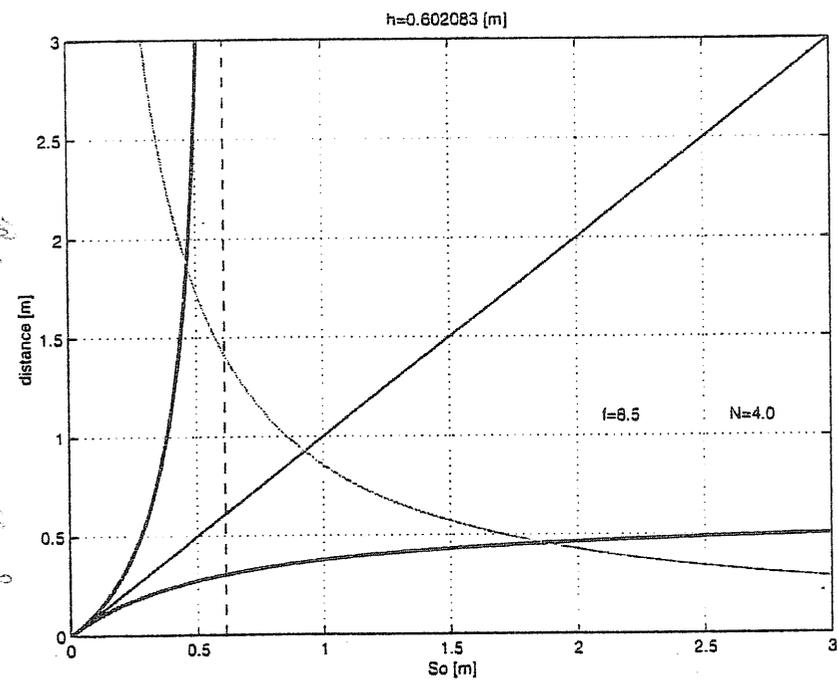
ESEMPI DI PROFONDITÀ DI CAMPO : la portata, e la posizione dell'oggetto. Le curve (o Trombe) segnano i limiti della messa a fuoco



Se la portata è + infinita di 0.6 m sono a fuoco quasi tutti i pt. compresi tra 0.0 m e +∞. L'info sulla profondità e portata di campo ci dice la posizione che può essere osservata, e quali oggetti possono essere visti bene.

Per migliorare la profondità di campo si stringe il diaframma (c'è minor luce ⇒ maggior tempo di esposizione).

Una pupilla molto piccola, inoltre, intro-duce effetti di diffrazione, poiché la sua dimensione diventa confrontabile con la lunghezza d'onda x obiettivi che abbiano un f-numero pari a 22.



3. ELABORAZIONE DEL MOVIMENTO

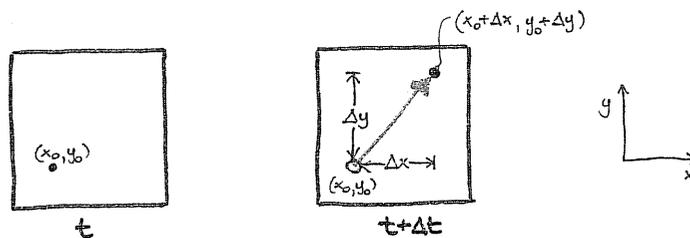
Interpretare ed elaborare una sequenza di immagini avviene attraverso

- (1.) STIMA DEL CAMPO DI MOTO: scalare che definisce, punto x punto, cosa abbiamo sulla scena [RAFFRETTAZIONE]
- (2.) Analisi dei MOVIMENTI RELATIVI tra oggetti e scena [INTERPRETAZIONE]

3.1. PERCEZIONE DEL MOTO

Il movimento è un'informazione $I(x, y, t)$ legata alla variazione spazio-temporale della luminanza.

Supponiamo di avere 2 frames



Se Δx e Δy hanno piccoli range temporali, li riusciamo a percepire come movimento, e cioè come MOTO APPARENTE.

Avremo che

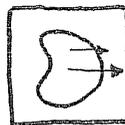
$$\begin{cases} \Delta x, \Delta y < 15' \text{ (di grado)} \\ \Delta t < 80 \text{ msec} \end{cases}$$

CONDIZIONE DELLA PERCEZIONE DEL MOTO

Nota: q. tipo di moto può essere stimato attraverso operatori locali.

IMPLICAZIONI FUNZIONALI:

- ① ROBOTICA: oggetto in movimento, si vuole seguire una traiettoria
 - TRACKING
 - TEMPO ALL' IMPATTO (= TIME TO CONTACT)
 - TEMPO x OCCUPARE UNA POSIZIONE (= TIME TO HEADING)
- ② RICOSTRUZIONE 3-D: Tentativo di ricostruire una scena tridimensionale a partire da un'immagine: $2D + t \rightarrow 3D$
- ③ SEGMENTAZIONE: tecnica x individuare, specie nell'ambito medico, particolari porzioni di immagini (2D o 3D) in genere statiche, ma non solo (x colore, tessitura, movimento)



Si possono rilevare parti di scena con moto coerente o incoerente rispetto allo sfondo: le parti con moto coerente vengono rilevate.

Ci sono inoltre pt² che si muovono di moto coerente x direzione ma non in modulo, e viceversa: i primi sono i + importanti.

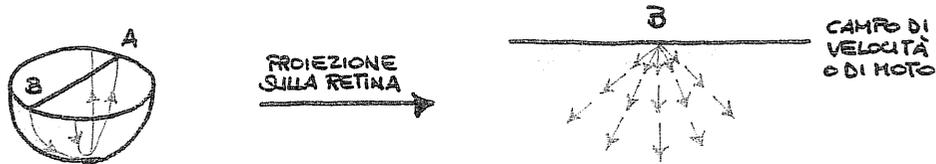
Oltre a q. 3 fondamentali applicazioni ci sono altre cose di cui tener conto, come l'equilibrio, il ritorno dello sguardo in un pt² di default, etc.

2. STIMA DEL CAMPO DI MOTO

Attraverso una diretta osservazione delle immagini raccogliamo info sul campo di moto.

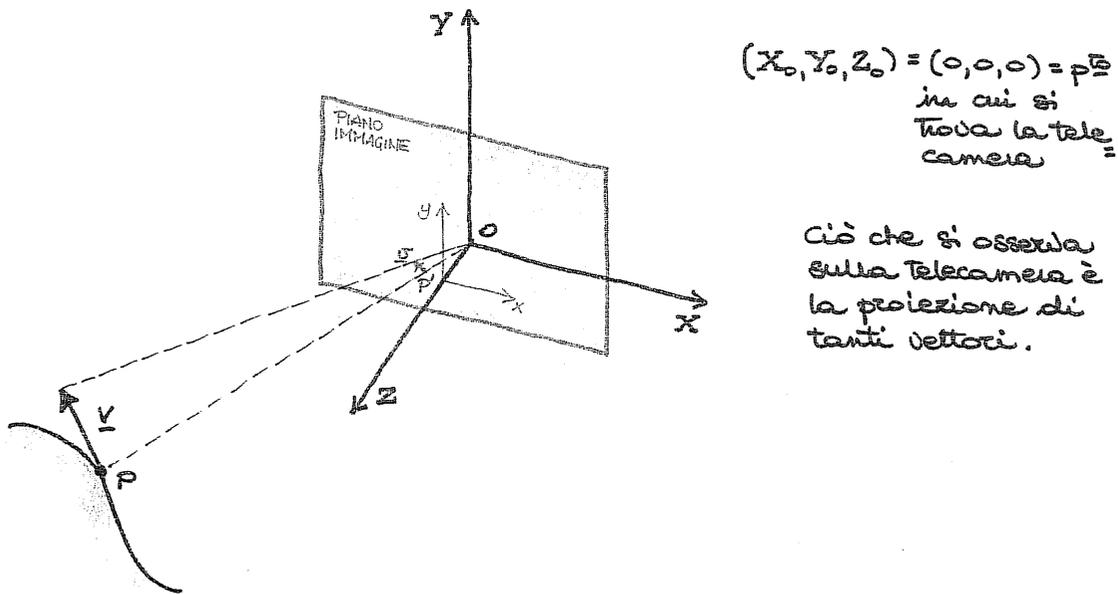
HELMOLTZ (prima) e GIBSON (dopo) avevano cercato di dimostrare che la percezione della tridimensionalità è proprio legata al movimento: prima l'approccio alle immagini era statico.

Essi posero l'attenzione sull'OPTICAL ARRAY, una visione a tutto tondo dell'ambiente che si contrappone all'immagine retinica (RETINAL IMAGE): in q. modo l'immagine non è piatta, bensì su una sfera che circonda l'osservatore.



CAMPO DI MOTO = proiezione sulla superficie della retina del movimento degli oggetti (e delle loro velocità) nell'optical array.

Stia una terna di coordinate (X, Y, Z) dello spazio oggetti ed un oggetto P in movimento, tale $P(t)$ si muove con velocità V :

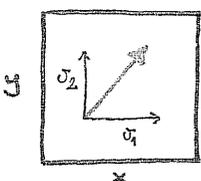


3. FLUSSO OTTICO

Senza potendo disporre della conoscenza della scena, e la stima del campo di moto dovremo ricorrere al calcolo del FLUSSO OTTICO.

FLUSSO OTTICO = campo osservabile \underline{v} , generalmente vicino al campo di moto (ma cmq diverso).

Es quanto ci riguarda considereremo il campo di moto \equiv flusso ottico, ma in realtà ciò non vale sempre. Se, ad es, nella scena gli oggetti sono fissi ma le luci in movimento, allora si percepisce il movimento della scena: in q. caso campo di moto \neq flusso ottico.



Dato $\underline{v}(v_1, v_2) \rightarrow \underline{v}(x, y)$, cioè dobbiamo trovare per ogni $P(t)$ un vettore velocità

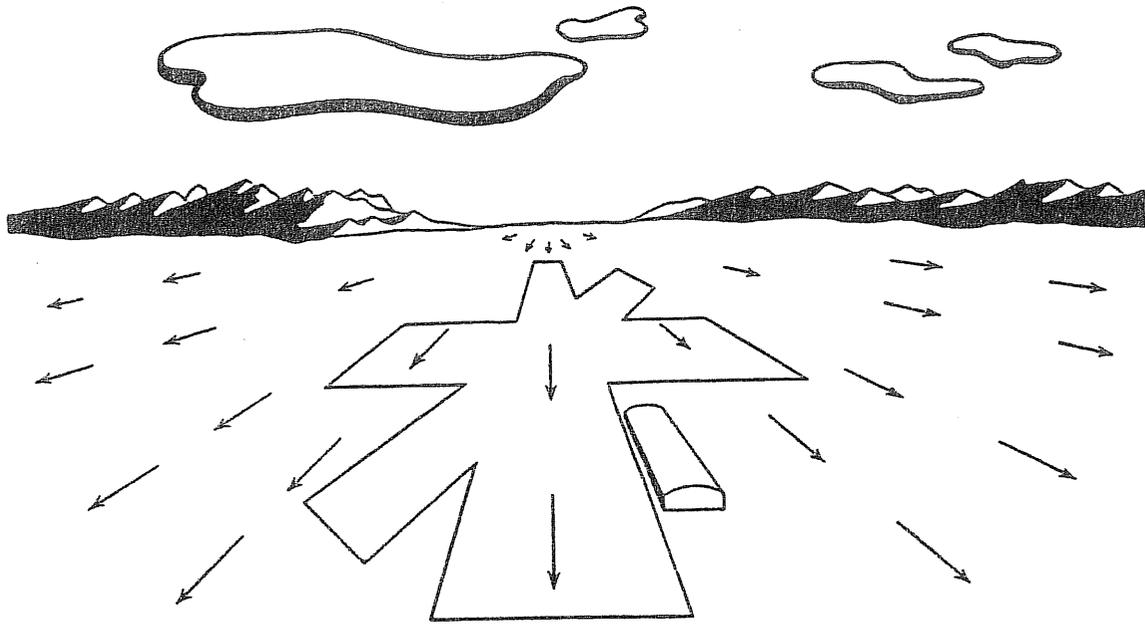


FIG. 3.1. Motion Perspective in the Visual Field Ahead
 PROIEZIONE RETINICA DELL'OPTIC ARRAY (VISIONE PIANTA)

[JAMES J. GIBSON, 1950]

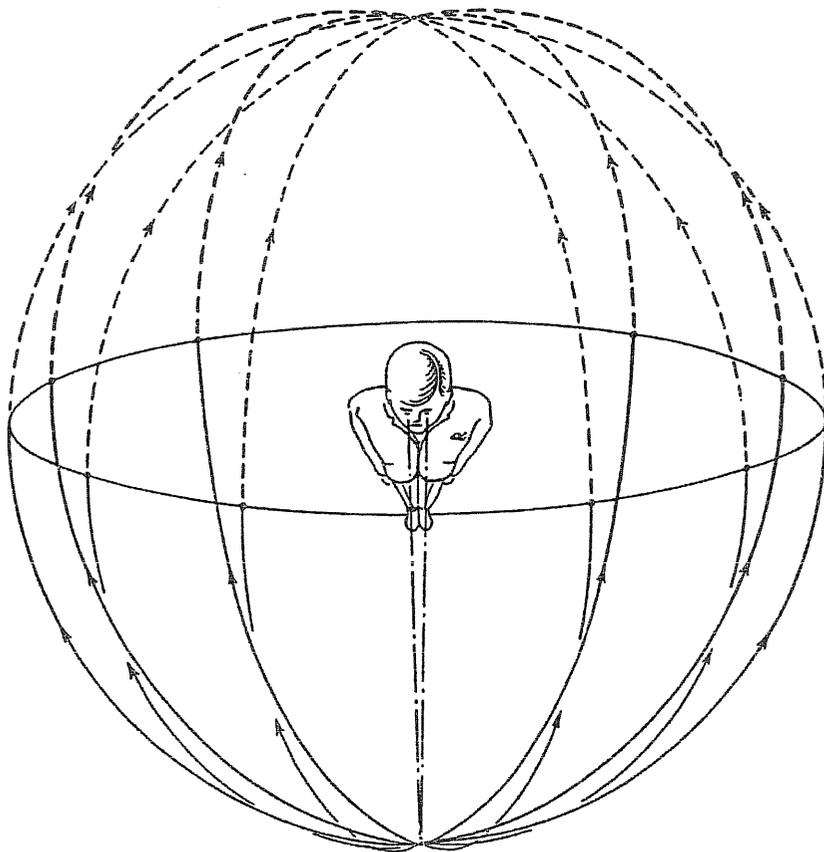
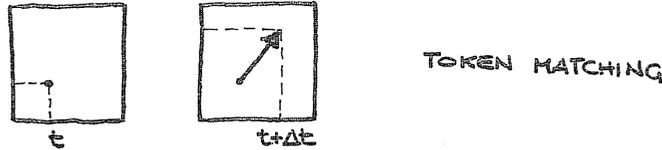


FIG. 3.2. The Directions of Deformations in the Visual Field during Forward Locomotion, as Projected on a Spherical Surface around the Head

Per il calcolo del flusso ottico ci affidiamo al FB DELLA RICERCA DI CORRISPONDENZE (= TOKEN MATCHING) degli spigoli (= edge).



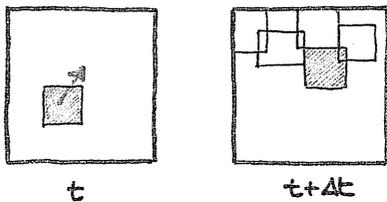
Questo tipo di procedimento genera dei campi sparsi, ed inoltre può capitare di prendere uno spigolo x un altro.

Per avere campi + densi si possono usare altri metodi come

- ① METODI DI CORRELAZIONE
- ② METODI DIFFERENZIALI
- ③ METODI BASATI SUL FILTRAGGIO SPAZIO-TEMPORALE

3.3.1. METODI DI CORRELAZIONE

Simili al discorso del TOKEN MATCHING, dove si cerca la corrispondenza di un particolare, fanno una correlazione di aree, e non di un bordo o di un incrocio.



Una porzione dell'immagine (all'istante t) è messa a confronto con molte regioni (allo istante t+Δt) parzialmente sovrapposte: la parte di area con maggiore correlazione è quella che si è spostata. Anche in q. caso c'è il pb dei falsi positivi, ma è meno rischioso.

3.3.2. METODI DIFFERENZIALI

Si basano sull'hyp della CONSERVAZIONE DELLA LUMINANZA, ovvero che le sorgenti luminose sono fisse, e gli unici movimenti sono dovuti allo spostamento.

Ciò, dal pt di vista matematico, significa

$$\frac{dI(x,y,t)}{dt} = 0 \Rightarrow \underbrace{\frac{\partial I}{\partial x} \cdot \frac{dx}{dt}}_{\sigma_1} + \underbrace{\frac{\partial I}{\partial y} \cdot \frac{dy}{dt}}_{\sigma_2} + \frac{\partial I}{\partial t} = 0$$

Definiamo ora

$$\left. \begin{aligned} \frac{\partial I}{\partial x} &\triangleq I_x \\ \frac{\partial I}{\partial y} &\triangleq I_y \\ \frac{\partial I}{\partial t} &\triangleq I_t \end{aligned} \right\} \Rightarrow (I_x, I_y) = \nabla I$$

$$\left. \begin{aligned} \frac{dx}{dt} &= \sigma_1 \\ \frac{dy}{dt} &= \sigma_2 \end{aligned} \right\} \Rightarrow \underline{\sigma} = (\sigma_1, \sigma_2)$$

Allora possiamo riscrivere

$$I_x \sigma_1 + I_y \sigma_2 + I_t = 0$$

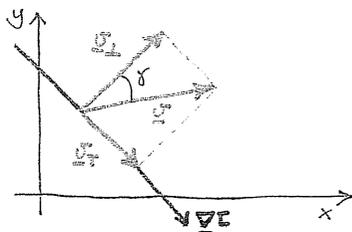
$$\boxed{\nabla I \cdot \underline{\sigma} + I_t = 0}$$

EQNE DEL GRADIENTE SU BASE PUNTUALE

Troviamo I_t conoscendo 2 frame successivi e costruendo il rapporto incrementale; la stessa cosa si fa per $\nabla I = (I_x, I_y)$, mentre il discorso non vale per $\underline{\sigma}$:

$$|\nabla I| \cdot \underbrace{|\underline{v}| \cos \delta}_{v_{\perp}} = -I_t \quad (*)$$

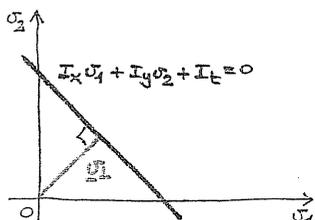
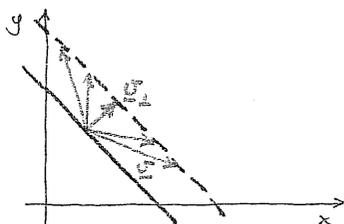
quindi otteniamo solo la proiezione di \underline{v} sulla direzione del gradiente



\underline{v}_{\perp} = componente ortogonale alla direzione del gradiente, è determinata da (*)

$\underline{v}_{\parallel}$ = componente tangenziale alla direzione del gradiente, è indeterminata

se \underline{v} conosciamo solo \underline{v}_{\perp} : per \underline{v}_{\perp} abbiamo ∞ $\underline{v}_{\parallel}$, quindi ∞ \underline{v} :



Il vettore \underline{v}_{\perp} , ortogonale alla retta, nell'origine è dato da

$$\underline{v}_{\perp} = -\frac{I_t}{\sqrt{I_x^2 + I_y^2}}$$

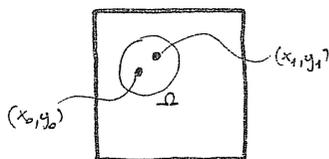
per rimuovere l'indeterminazione si può procedere in 2 modi:

- ① INTRODURRE ALTRI VINCOLI [LUCAS - KANADE, 1981]
- ② ASSUMERE LA REGOLARITÀ (SMOOTHNESS) DELLA SOLUZIONE \underline{v} [HORN - SCHUNK, 1981]

realizziamoli nel dettaglio.

① INTRODUZIONE DI ALTRI VINCOLI

Combiniamo le eqⁿⁱ del gradiente in ptⁱ vicini



Se i 2 ptⁱ sono vicini e non appartenenti ai bordi, è probabile che appartengano allo stesso oggetto: in q. caso essi avranno la stessa velocità.

Scriviamo il gradiente x 2 ptⁱ vicini:

$$\begin{cases} I_x(x_0, y_0, t) v_1 + I_y(x_0, y_0, t) v_2 + I_t(x_0, y_0, t) = 0 \\ I_x(x_1, y_1, t) v_1 + I_y(x_1, y_1, t) v_2 + I_t(x_1, y_1, t) = 0 \end{cases}$$

Q. sistema è risolvibile, xò prendere solo 2 ptⁱ potrebbe non bastare, x sta del numero di cui l'immagine è affetta.

Per ovviare al pb, al posto di 2 ptⁱ prendiamo un piccolo intorno Ω e costruiamo una funzione che analizziamo ai minimi quadrati:

$$E(v_1, v_2) = \sum_{(x,y) \in \Omega} [I_x v_1 + I_y v_2 + I_t]^2 \text{ minimo} \quad (*)$$

Una possibile variante è:

$$E(v_1, v_2) = \sum_{(x,y) \in \Omega} \underbrace{w^2(x,y)}_{\text{FUNZIONE DI PESO GAUSSIANA}} [v \cdot \nabla I + I_t]^2 \text{ minimo} \quad (**)$$

Consideriamo l'eq^{me} (*) come un pb variazionale. Poiché non ci sono derivate prime avremo

$$E_v + \frac{d}{dx} E_{v'} = 0$$

perciò

$$\begin{cases} \frac{\partial E(\sigma_1, \sigma_2)}{\partial \sigma_1} = 0 \\ \frac{\partial E(\sigma_1, \sigma_2)}{\partial \sigma_2} = 0 \end{cases} \Rightarrow \begin{cases} \sum z [I_x \sigma_1 + I_y \sigma_2 + I_t] I_x = 0 \\ \sum z [I_x \sigma_1 + I_y \sigma_2 + I_t] I_y = 0 \end{cases}$$

e dunque

$$\begin{cases} \sigma_1 \sum I_x^2 + \sigma_2 \sum I_x I_y + \sum I_x I_t = 0 \\ \sigma_1 \sum I_x I_y + \sigma_2 \sum I_y^2 + \sum I_y I_t = 0 \end{cases} \Leftrightarrow \boxed{M \underline{\sigma} = -\underline{b}} \quad \begin{array}{l} \text{EQNE DEL} \\ \text{GRADIENTE} \\ \text{SU BASE LOCALE} \end{array}$$

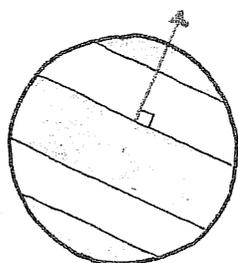
dove

$$M \hat{=} \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

Poichè M è spesso non-invertibile, andiamo lo stesso incontro ad indeterminazione: ciò è dovuto a come sono fatte le immagini x l'apertura.

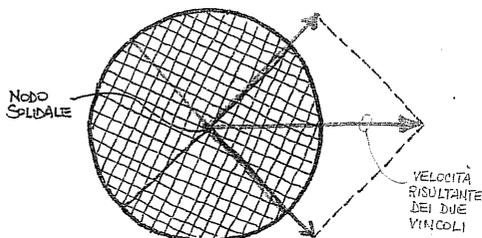
PS DELL'APERTURA

[= SE PRENDIAMO UNA FINESTRA + PICCOLA DEL INTERNO]
[= IMPOSSIBILE DETERMINARE IL VERO MOTTO DI U]



In q. caso si percepisce sempre il movimento ortogonale, anche se in realtà non è quello il movimento.

M è non-invertibile e singolare. Vediamo solo $\underline{\sigma}_T$.



In q. caso abbiamo una tessitura + fine e la presenza di incroci: avremo perciò 2 vincoli.

Con i 2 vincoli sulla velocità riusciamo ad avere $\underline{\sigma}_T$.

② ASSUNZIONE DI REGOLARITÀ (SMOOTHNESS)

L'assunzione di regolarità è un vincolo di continuità del flusso ottico. Abbiamo visto che

$$\sigma_1 = \sigma_1(x, y) \quad \sigma_2 = \sigma_2(x, y) \quad \underline{\sigma} = (\sigma_1, \sigma_2)$$

Die che vogliamo una soluzione normalizzata può essere esplicitato prendendo piccole le variazioni del campo vettoriale:

$$\begin{cases} \frac{\partial \sigma_1}{\partial x} = \frac{\partial \sigma_1}{\partial y} = 0 \\ \frac{\partial \sigma_2}{\partial x} = \frac{\partial \sigma_2}{\partial y} = 0 \end{cases} \Leftrightarrow \underline{\nabla} \cdot \underline{\sigma} + I_t = 0$$

Dobbiamo costruire un funzionale di costo quadratico che, oltre a q. 2 eq^{me}, tenga anche conto dell' eq^{me} del gradiente:

$$\int [(\underline{\nabla} \cdot \underline{\sigma} + I_t)^2 + \lambda P^2 \underline{\sigma}] dx dy = \text{minimo}$$

METODO DI HORN-SCHUNK

dove $P^2 \underline{\sigma} = \left(\frac{\partial \sigma_1}{\partial x}\right)^2 + \left(\frac{\partial \sigma_1}{\partial y}\right)^2 + \left(\frac{\partial \sigma_2}{\partial x}\right)^2 + \left(\frac{\partial \sigma_2}{\partial y}\right)^2$

La soluzione di q. funzionale è di tipo approx ed iterativo, ricavata con Eulero - Lagrange:

$$\begin{cases} \sigma_1^{(m+1)} = \sigma_1^{(m)} - \frac{I_x \sigma_1^{(m)} + I_y \sigma_2^{(m)} + I_t}{1 + \lambda^2 (I_x^2 + I_y^2)} \cdot I_x \\ \sigma_2^{(m+1)} = \sigma_2^{(m)} - \frac{I_x \sigma_1^{(m)} + I_y \sigma_2^{(m)} + I_t}{1 + \lambda^2 (I_x^2 + I_y^2)} \cdot I_y \end{cases}$$

Il pt^o critico è la scelta della condizione iniziale $(\sigma_1^{(0)}, \sigma_2^{(0)})$: viene suggerito di utilizzare quella trovata mediante l'applicazione del vincolo del gradiente. Lo svantaggio è che bisogna determinare anche il TEMPO DI CONVERGENZA.

3. METODI BASATI SUL FILTRAGGIO SPAZIO-TEMPORALE

andiamo

$$I(x, y, t) \xleftrightarrow{\mathcal{F}} \hat{I}(f_x, f_y, f_t) \quad \text{DISTRIBUZIONE LUMINOSA}$$

CASO 1D

Consideriamo il caso di una sola dimensione ed il tempo:

$$I(x, t)$$

e che l'immagine si muove con una velocità costante σ_1 :

$$I(x, t) \longrightarrow I(x - \sigma_1 t, t)$$

Il nostro obiettivo è ora quello di caratterizzare $\hat{I}(f_x, f_t)$. Appliciamo una trasformazione di coordinate

$$\begin{aligned} \underline{a} = \begin{pmatrix} x \\ t \end{pmatrix} &\longrightarrow \underline{a}' = \begin{pmatrix} x - \sigma_1 t \\ t \end{pmatrix} \\ \underline{b} = \begin{pmatrix} f_x \\ f_t \end{pmatrix} &\longrightarrow \underline{b}' = ? \end{aligned}$$

Avremo quindi che

$$\underline{a}' = A \underline{a} \quad \text{dove } A = \begin{bmatrix} 1 & -\sigma_1 \\ 0 & 1 \end{bmatrix} \quad [\det(A) = 1]$$

Sappiamo che

$$f(\alpha t) \xrightarrow{\mathcal{F}} \frac{1}{|\alpha|} \hat{f}\left(\frac{1}{\alpha} f_t\right) \quad \text{[PROPRIETÀ DI CAMBIO SCALA NELLE TRASFORMATE DI FOURIER]}$$

Dunque, partendo da

$$I(A \underline{a}) \xrightarrow{\mathcal{F}} \hat{I}((A^{-1})^T \underline{b}) \quad \text{dove } (A^{-1})^T = \begin{bmatrix} 1 & 0 \\ \sigma_1 & 1 \end{bmatrix}$$

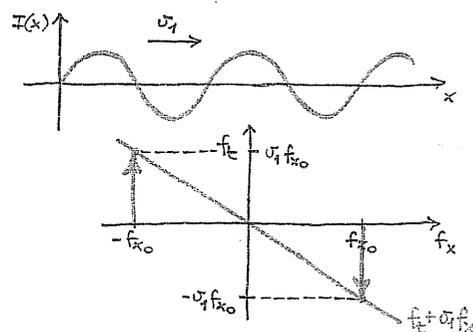
allora otteniamo

$$\hat{I}(f_x, f_t + \sigma_1 f_x) \implies \underline{b}' = \begin{pmatrix} f_x \\ f_t + \sigma_1 f_x \end{pmatrix}$$

In conclusione la frequenza spaziale rimane inalterata, mentre quella temporale risulta traslata.

Es.: $I(x) = \cos(f_{x_0} x)$
 $I(x, t) = \cos(f_{x_0} (x - \sigma_1 t))$

L'immagine subisce una deformazione che tende a portare le frequenze su una retta.



2. CASO 2D

In q. caso avremo

$$I(x, y, t) \rightarrow I(x - v_1 t, y - v_2 t, t)$$

Anche stavolta effettuiamo una trasformazione di coordinate:

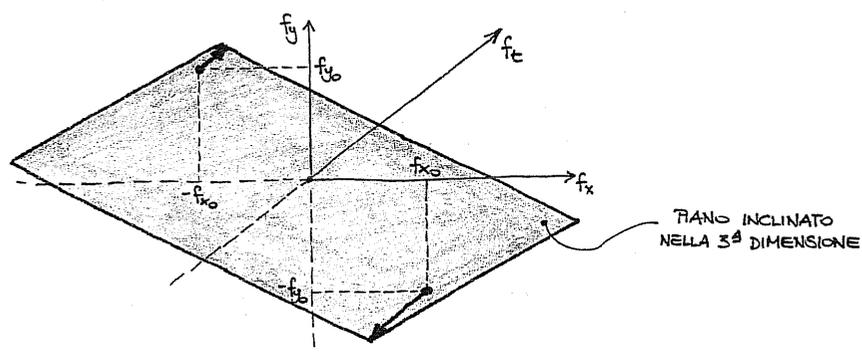
$$\underline{a} = \begin{pmatrix} x \\ y \\ t \end{pmatrix} \rightarrow \underline{a}' = \begin{pmatrix} x - v_1 t \\ y - v_2 t \\ t \end{pmatrix} = A \underline{a} \quad \text{con } A = \begin{bmatrix} 1 & 0 & -v_1 \\ 0 & 1 & -v_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Ne segue perciò che

$$\underline{b}' = \begin{pmatrix} f_x \\ f_y \\ f_t + v_1 f_x + v_2 f_y \end{pmatrix}$$

dunque, nel caso 2D, non abbiamo + l'eq^{me} di una retta, bensì di un piano.

ES.: $I(x, y, t) = \cos(f_{x_0} x + f_{y_0} y)$



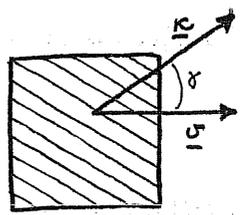
Se poniamo $\underline{k} = \begin{pmatrix} f_x \\ f_y \end{pmatrix}$ e $\underline{v} = (v_1, v_2)$ allora avremo

$$f_t = -v_1 f_x - v_2 f_y \iff f_t = -\underline{v} \cdot \underline{k} \quad \text{EQ^{NE} PIANO}$$

Anche nel caso del filtraggio non riusciamo a trovare una soluzione unica. Operativamente filtriamo la $I(x, y, t)$ con un BPF con frequenza di picco k_0 , riportandoci così ad una situazione ideale con un unico coseno.

Possiamo ancora ricavare

$$f_t = -\underline{v} \cdot \underline{k} \implies |\underline{v}| \cos \delta = -\frac{f_t}{|k_0|} \quad \text{EQ^{NE} IN FREQUENZA PB APERTURA}$$

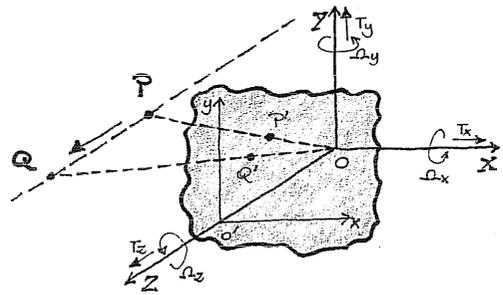


In conclusione la relazione $f_t = -\underline{v} \cdot \underline{k}$ ci dà la proiezione.

3.4. INTERPRETAZIONE DEL FLUSSO OTTICO

Ci si chiede come ricavare dal flusso ottico le info sul moto relativo tra osservatore e scena.

Prendiamo un asse di riferimento fisso, uno x occhio della telecamera. Abbiamo poi un piano immagine (x, y) ed un pt P nello spazio oggetti, con proiezione P' nel piano immagine.



Supponiamo che P si muova verso Q , ed inoltre che il mondo sia fisso e che siamo noi a muoverci.

La telecamera, poi, ruota intorno alla scena a velocità $\Omega_x, \Omega_y, \Omega_z$.

finiamo

$\underline{T} = (T_x, T_y, T_z) =$ VETTORE VELOCITÀ TRASLAZIONALI

$\underline{\Omega} = (\Omega_x, \Omega_y, \Omega_z) =$ VETTORE VELOCITÀ ROTAZIONALI

$\underline{r} = (X, Y, Z) =$ VETTORE POSIZIONE

moto nello spazio oggetti obbedisce alla relazione

$$\dot{\underline{r}} = -\underline{T} - \underline{\Omega} \wedge \underline{r}$$

LEGGI DI MOVIMENTO

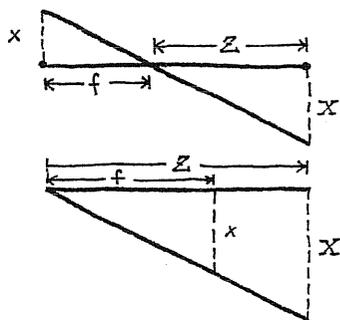
[= VARIAZIONE NEL TEMPO DEL VETTORE POSIZIONE]

esplicitandola in forma scalare

$$[\underline{\Omega} \wedge \underline{r} = \begin{pmatrix} \Omega_x Y - \Omega_y X \\ \Omega_y Z - \Omega_z Y \\ \Omega_z X - \Omega_x Z \end{pmatrix}]$$

$$(*) \begin{cases} \dot{X} = -T_x - \Omega_y Z + \Omega_z Y \\ \dot{Y} = -T_y + \Omega_x Z - \Omega_z X \\ \dot{Z} = -T_z - \Omega_x Y + \Omega_y X \end{cases}$$

usando la proiezione prospettica e supponendo $\overline{OO'} = f =$ distanza focale del sistema, la proiezione ci aiuta nella risoluzione:



Riusciamo in q. modo ad ottenere una relazione che lega X ad x:

$$-\frac{x}{f} = \frac{X}{Z} \Rightarrow \frac{x}{f} = \frac{X}{Z}$$

Ritornando al pb, avremo che

$$(v_1, v_2) \equiv (x', y') \equiv \left(\frac{dx}{dt}, \frac{dy}{dt} \right)$$

e, ponendo la distanza focale $f=1$, otteniamo

$$\begin{cases} x = \frac{X}{Z} \\ y = \frac{Y}{Z} \end{cases} \Rightarrow \begin{cases} \frac{dx}{dt} = \frac{d}{dt} \left(\frac{X}{Z} \right) = \dot{X} \cdot \frac{1}{Z} - \frac{X}{Z^2} \dot{Z} \\ \frac{dy}{dt} = \frac{d}{dt} \left(\frac{Y}{Z} \right) = \dot{Y} \cdot \frac{1}{Z} - \frac{Y}{Z^2} \dot{Z} \end{cases}$$

casiamo dunque un legame del tipo

$$(v_1, v_2) \longleftrightarrow (\underline{T}, \underline{\Omega}, \underline{Z})$$

ostituendo la (*) otteniamo

$$\dot{x} = -\frac{T_x}{Z} - \Omega_y + \Omega_z \left(\frac{Y}{Z} \right) - x \left(-\frac{T_z}{Z} - \Omega_x \left(\frac{Y}{Z} \right) + \Omega_y \left(\frac{X}{Z} \right) \right)$$

accogliamo i termini traslatori e rotazionali:

$$\dot{x} = \underbrace{-\frac{T_x}{Z} + x \frac{T_z}{Z}}_{\text{TRASLATORI}} - \underbrace{\Omega_y(x^2+1) + xy\Omega_x + \Omega_z y}_{\text{ROTAZIONALI}}$$

ocedendo analogamente

$$\dot{y} = \underbrace{-\frac{T_y}{Z} + y \frac{T_z}{Z}}_{\text{TRASLATORI}} + \underbrace{\Omega_x(1+y^2) - xy\Omega_y - \Omega_z x}_{\text{ROTAZIONALI}}$$

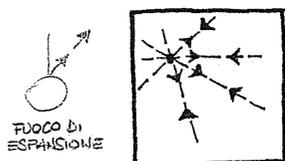
Siccome

$$\begin{cases} \dot{x} = \sigma_1 = \sigma_1^T + \sigma_1^R \\ \dot{y} = \sigma_2 = \sigma_2^T + \sigma_2^R \end{cases} \Rightarrow \begin{cases} \sigma_1^T = -\frac{T_x}{Z} + x \frac{T_z}{Z} \\ \sigma_1^R = -\Omega_y(1+x^2) + xy\Omega_x + y\Omega_z \\ \sigma_2^T = -\frac{T_y}{Z} + y \frac{T_z}{Z} \\ \sigma_2^R = \Omega_x(1+y^2) - xy\Omega_y - x\Omega_z \end{cases}$$

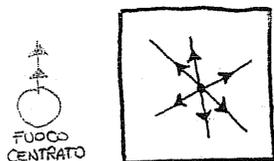
Le componenti notazionali non contengono Z: se noi avessimo telecamere che si muovono solo ruotando, non avremmo info sulla profondità della scena.

Consideriamo ora alcuni casi limite:

- $\underline{\Omega} = 0$: le componenti σ^T del campo di moto sono caratterizzate dallo essere dei vettori sulla linea di fuga.

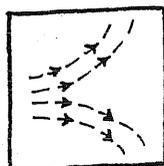


Posiamo andare $\left\{ \begin{array}{l} \text{avanti } (T_z > 0) \Rightarrow \text{ESPANSIONE DEL MOTO} \\ \text{indietro } (T_z < 0) \Rightarrow \text{COMPRESSIONE DEL MOTO} \end{array} \right.$



È il caso di un uomo che si muove nella direzione di spostamento: in q. caso abbiamo un **FUOCO CENTRATO**

- $\underline{T} = 0$: non avremo espansione; le linee di fuga scompaiono ed al loro posto avremo dei rami di iperbole



COMPONENTI ROTAZIONALI

In alcuni casi le linee di rotazione sono piatte, quindi la loro combinazione con un moto traslazionale con fuoco centrale ci porta ad una traslazione del fuoco di espansione.

Se basiamo il nostro allineamento occhio - fuoco attraverso un procedimento che cerca la posizione del fuoco otteniamo uno sfasamento

$$\underline{\sigma}^T(x, y) = \begin{pmatrix} \sigma_1^T(x, y) \\ \sigma_2^T(x, y) \end{pmatrix} = \begin{pmatrix} \frac{-T_x + T_z x}{Z} \\ \frac{-T_y + T_z y}{Z} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vogliamo ricavare le coordinate del fuoco di espansione:

$$\begin{cases} -\frac{T_x}{Z} + \frac{T_z}{Z} x = 0 \\ -\frac{T_y}{Z} + \frac{T_z}{Z} y = 0 \end{cases} \Rightarrow \text{FE} \equiv \begin{cases} x_0 = \frac{T_x}{T_z} \\ y_0 = \frac{T_y}{T_z} \end{cases} \quad \text{FUOCO DI ESPANSIONE (FOE)}$$

NB: il FUOCO DI ESPANSIONE indica il p^{to} nello spazio oggetti verso cui si sta muovendo l'osservatore.

$\underline{F} = (x_0, y_0)$ è un p^{to} a velocità nulla, ma estremamente numeroso.

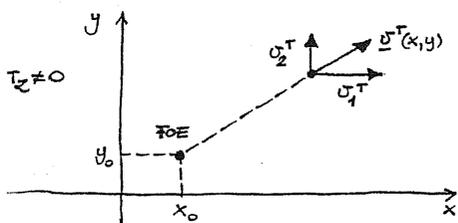
Se siamo di fronte al caso $\Omega = 0$, allora tutti i vettori del campo di moto giacciono su rette passanti x il FOE.

Dimostriamolo:

[HYP: $T_z \neq 0$]

$$\begin{cases} \sigma_1^T = -\frac{T_x}{T_z} + \frac{T_x}{T_z} x = \frac{T_x}{T_z} \left(x - \left(\frac{T_x}{T_z} \right) \right) = \frac{T_x}{T_z} (x - x_0) \\ \sigma_2^T = -\frac{T_y}{T_z} + \frac{T_y}{T_z} y = \frac{T_y}{T_z} \left(y - \left(\frac{T_y}{T_z} \right) \right) = \frac{T_y}{T_z} (y - y_0) \end{cases}$$

graficamente sarà:



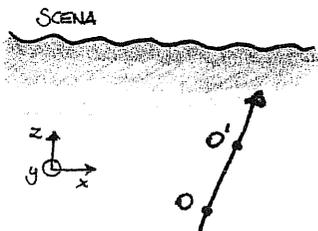
Congiungendo il FOE con il p^{to}, il vettore $\sigma^T(x, y)$ è formato da σ_1^T e σ_2^T tali che

$$\frac{\sigma_2^T}{\sigma_1^T} = \frac{y - y_0}{x - x_0} \quad \text{RETTA CONGIUNGENTE FOE E PUNTO}$$

con coefficiente angolare dipendente da σ^T .

5. CALCOLO DEL FUOCO DI ESPANSIONE DAL MOTO DI PARALLASSE

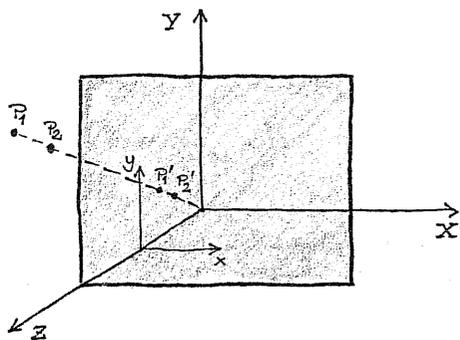
Per ricostruire una scena potremmo muoverci al suo interno: così facendo come se siamo la scena da 2 p^{ti} di vista diversi, proprio x il nostro spostamento.



IL PARALLASSE indica la posizione spostata degli oggetti dovuta al moto dell'osservatore.

Ciò permette di stimare la posizione del FOE anche nel caso di $\Omega \neq 0$.

Supponiamo di prendere 2 p^{ti} che sono "quasi" sullo stesso asse ottico, ma a profondità diverse



Il campo di moto sarà

$$\begin{cases} \sigma_{P1}^R = \sigma_{P1}^T + \sigma_{P1}^R \\ \sigma_{P2}^R = \sigma_{P2}^T + \sigma_{P2}^R \end{cases} \quad \text{con} \begin{cases} X_1 \approx X_2 & x_1 \approx x_2 \\ Y_1 \approx Y_2 & y_1 \approx y_2 \\ Z_1 \neq Z_2 & z_1 \neq z_2 \end{cases}$$

Di fatto avremo

$$\sigma_{P1}^R = \sigma_{P2}^R$$

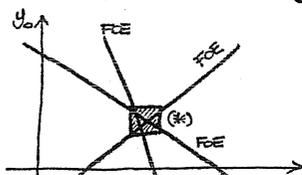
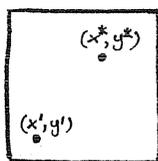
Se costruiamo il campo di moto attraverso la differenza delle 2 velocità σ_{P1}^R e σ_{P2}^R otteniamo un campo di moto in cui non abbiamo + riferimenti del moto rotazionale:

$$\begin{cases} \sigma_1^R - \sigma_1^R = (-T_x + T_z x) \left(\frac{1}{Z_1} - \frac{1}{Z_2} \right) \\ \sigma_2^R - \sigma_2^R = (-T_y + T_z y) \left(\frac{1}{Z_1} - \frac{1}{Z_2} \right) \end{cases} \quad \text{PROFONDITÀ}$$

unque

$$\frac{\sigma_2^R - \sigma_2^R}{\sigma_1^R - \sigma_1^R} = \frac{-T_y + T_z y}{-T_x + T_z x} = \frac{T_z (y - y_0)}{T_z (x - x_0)} = \frac{y - y_0}{x - x_0} \quad \text{EQNE RETTA}$$

dove le incognite sono, ovviamente, (x_0, y_0) .



$$\frac{\sigma_2(x^*, y^*) - \sigma_2(x^* + \epsilon, y^* + \epsilon)}{\sigma_1(x^*, y^*) - \sigma_1(x^* + \epsilon, y^* + \epsilon)}$$

Per stimare la posizione del fuoco di espansione, in assenza di rumore ci bastano 3 pt; altrimenti ne serve un * maggiore, che ci permette di trovare un fascio di rette che individua una regione ristretta (*) in cui il FOE si trova.

3.6. CALCOLO DEL TEMPO DI COLLISIONE

Detti

Z = distanza dalla telecamera

Tz = velocità di moto dell'osservatore

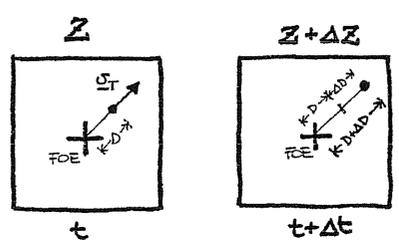
e ricordando che

$$\begin{cases} v_1 = (x-x_0) \frac{T_z}{Z} \\ v_2 = (y-y_0) \frac{T_z}{Z} \end{cases}$$

si definisce

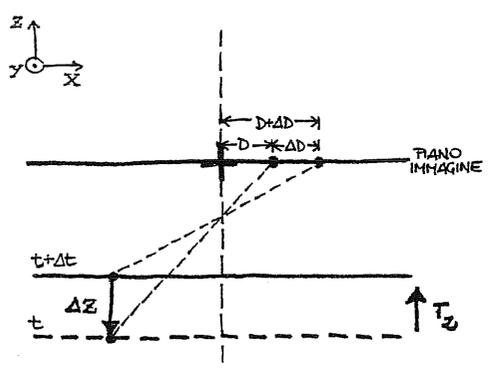
$$\tau_c = \frac{Z}{T_z} \quad \text{TEMPO DI COLLISIONE}$$

È importante la variazione della distanza tra fuoco d'espansione e punto immagine: supponiamo di guardare nel fuoco di espansione



All'istante (t+Δt) il punto è spostato, secondo il moto di espansione, di una distanza D proporzionale a Z.

Ciò può essere meglio visualizzato così:



Dalle proporzioni tra triangoli simili

$$\frac{Z}{\Delta Z} = \frac{D}{\Delta D} \Rightarrow \frac{Z}{\Delta Z} = \frac{D}{\Delta D} \Rightarrow \frac{Z}{T_z} = \frac{D}{\Delta D}$$

e quindi

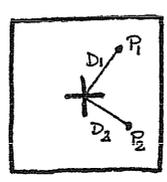
$$\tau_c = \frac{Z}{T_z} = \frac{D}{|v_T|}$$

NOTO X COSTRUZIONE (SAPRAI MOLA POSIZIONE DEL FOE)
LO STIMIAMO

Se conosciamo Tz, possiamo ricavare la profondità dell'oggetto con la relazione

$$Z(x,y) = T_z \frac{D}{|v_T(x,y)|} \quad \text{PROFONDITÀ}$$

Dati 2 pt nella scena



$$\left. \begin{aligned} T_z &= \frac{Z_1}{D_1} |v_1^T| \\ T_z &= \frac{Z_2}{D_2} |v_2^T| \end{aligned} \right\} \Rightarrow \frac{Z_1}{Z_2} = \frac{D_2}{D_1} \frac{|v_1^T|}{|v_2^T|}$$

Analogamente, se conosciamo Z2 = Z(x2, y2), possiamo ricavare

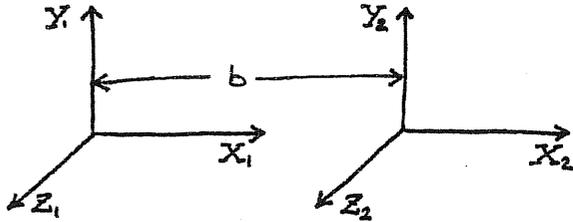
$$Z(x_1, y_1) = Z(x_2, y_2) \frac{D_2}{D_1} \frac{|v_1^T|}{|v_2^T|}$$

NB: vale solo se...

• Sono legati tra le informazioni stereo e quelle di moto: l'uso combinato di 2 entità ci permette di togliere le ambiguità legate all'interpretazione di una scena.

ciò viene detto DATA FUSION (STEREO + MOTO).

Ad es, le info sul moto di profondità possono essere legate alla misura di $\Delta v = v_L - v_R =$ differenza tra le velocità della telecamera su e dx



④ VISIONE STEREOSCOFICA

La VISIONE STEREOSCOFICA serve alla percezione della PROFONDITÀ (= distanza dall'osservatore) degli oggetti nella scena.

Nella VISIONE MONOCULARE [v. par. 2.1.1.] abbiamo visto che, data la posizione della sorgente nella scena e le corrispondenti coordinate immagine (P, y) e (P', y') , avremo

$$\frac{1}{P} + \frac{1}{P'} = \frac{1}{f} \quad (1)$$

$$\frac{y'}{y} = \frac{f}{P-f} = \frac{P'-f}{f} \quad (2)$$

Dunque se conosciamo P' e y' possiamo determinare P e y .

Tuttavia, se $P \gg f$, allora la determinazione di P a partire da P' risulta essere inaccurata, poiché $P' \approx f$.

Infatti, dalla (1):

$$\frac{\partial P}{\partial P'} = - \frac{1}{\underbrace{\left(\frac{P'}{f} - 1\right)^2}_{\approx 0}}$$

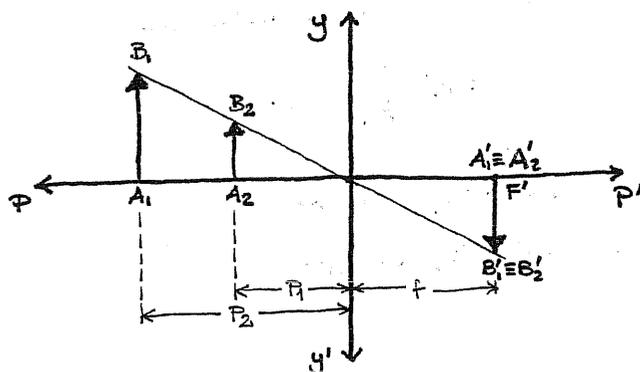
Insomma, se $P \gg f$, si pone

$$P' = f \quad (1')$$

$$y' = y \frac{f}{P} \quad (2')$$

Nota: in genere, quando è soddisfatta $P \gg f$, si parla di FOCALI CORTE.

Dalle eqⁿⁱ (1') e (2') si evidenzia che la visione monoculare non consente di individuare univocamente la distanza degli oggetti:



NON SI VERIFICA
DISCRIMINAZIONE
TRA DISTANZE \neq

Nella VISIONE STEREOSCOFICA invece, si utilizza il fatto che l'immagine della sorgente è raccolta da 2 posizioni dell'osservatore distinte spazialmente: dunque nei 2 piani immagine (dx e sn) si raccolgono immagini \neq della stessa sorgente.

Sia

$$I_L(x) = I_R(x - d(x))$$

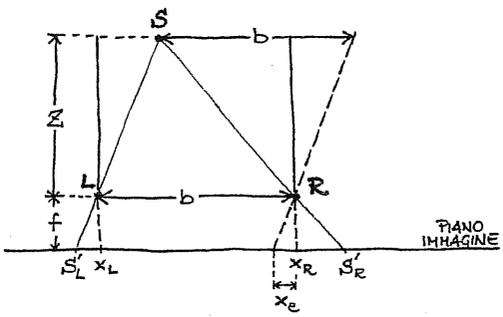
dove $I(x)$ = distribuzione dell'intensità di illuminamento nell'immagine
 $d(x)$ = DISPARITÀ BINOCOLARE: spostamento lineare nella distribuzione dell'intensità di illuminamento.

∴ non consideriamo la disparità fotometrica e geometrica, né pb di occlusione.

1. GEOMETRIA DELLA VISIONE STEREOSCOPICA

1) SISTEMI OTTICI CON ASSI PARALLELI

Preso un pto sorgente S, consideriamo la sua immagine con la geometria di tipo prospettico



Definiamo

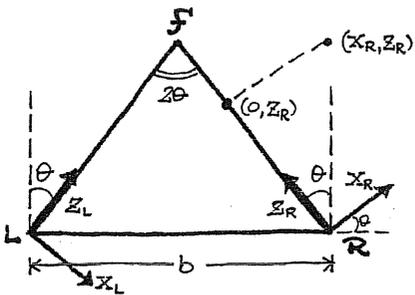
$$d \triangleq x_L - x_R = -b \frac{f}{Z} \quad \text{DISPARITÀ ORIZZONTALE}$$

Se S fosse fuori dal piano individuato dagli assi avremmo

$$y_{L(R)}' = y_{L(R)} \frac{f}{Z_{L(R)'}}$$

Nel caso di assi paralleli $Z_L \equiv Z_R$, quindi $Y_L \equiv Y_R$: ne segue che non c'è DISPARITÀ VERTICALE.

2) SISTEMI OTTICI CON ASSI CONVERGENTI IN F', PTO DI FISSAZIONE



Dato un pto immagine su R di coordinate x_R , il suo coniugato nello spazio $so_{x,z}$ è

$$X_R = \frac{x_R z_R}{f} \quad (1)$$

← DISTANZA FOCALE

Da x_R otteniamo una moltitudine di X_R , a seconda della z_R .

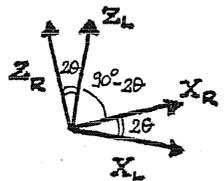
Il pto (X_R, z_R) avrà un'immagine su L data da

$$x_L = X_L \frac{f}{z_L} \quad (2)$$

dove

$$\begin{cases} X_L = X_L' \cos(2\theta) - z_L' \sin(2\theta) \\ z_L = X_L' \sin(2\theta) + z_L' \cos(2\theta) \end{cases} \quad (3)$$

Considerando



dobbiamo fare una rotazione ed una traslazione con

$$\begin{cases} X_L' = X_R + b \cos \theta \\ z_L' = z_R - b \sin \theta \end{cases} \quad (4)$$

Dalla (2) si ricava che

$$x_L \triangleq \frac{m(x_R)}{m(x)} = f \frac{\left(x_R \frac{z_R}{f} + b \cos \theta\right) \cos(2\theta) - (z_R - b \sin \theta) \sin(2\theta)}{\left(x_R \frac{z_R}{f} + b \cos \theta\right) \sin 2\theta + (z_R - b \sin \theta) \cos 2\theta}$$

Come possiamo notare, nel caso di assi paralleli

$$\theta = 0 \Rightarrow x_L = f \cdot \frac{x_R \frac{z_R}{f} + b}{z_R} = x_R + f \frac{b}{z_R} \Rightarrow x_L - x_R = f \frac{b}{z_R}$$

Può essere interessante considerare $m(x)|_{x_R=0}$, ossia allineati ad uno degli assi. Prendiamo il p^{to} $(0, Z_R)$, a cui corrisponde il p^{to} (x_L, Z_L) , e sostituiamo nelle formule:

$$m(x)|_{x_R=0} = f \frac{b \cos \theta \cos 2\theta - (Z_L - b \sin \theta) \sin 2\theta}{b \cos \theta \sin 2\theta - (Z_L - b \sin \theta) \cos 2\theta} \quad \left[\text{P}^{\text{to}} \text{ GENERICO SULL'ASSE } x_R \right]$$

$d \triangleq x_L - x_R = m(x)$

Per il p^{to} di fissazione f ($\Rightarrow Z_R = \frac{b}{2 \sin \theta}$) avremo

$$m(x)|_{\substack{x_R=0 \\ Z_R = \frac{b}{2 \sin \theta}}} = f \frac{\cos \theta \cos 2\theta - \sin 2\theta \sin \theta}{\cos 4\theta} = 0$$

Studiamo $m(x)$ nell'intorno del p^{to} di fissazione, ponendo $Z_R = \frac{b}{2 \sin \theta} + \Delta Z$. Sostituendo q. relazione, senza fare approx, otteniamo

$$m(x)|_{\substack{x_R=0 \\ Z_R = \frac{b}{2 \sin \theta} + \Delta Z}} = f \frac{-\frac{\Delta Z}{b} \overset{\approx \frac{b}{Z_R}}{\sin 2\theta} \cdot \overset{\approx \frac{b}{Z_R}}{2 \sin \theta}}{1 - 2 \sin^2 2\theta - \frac{\Delta Z}{b} 2 \sin \theta \cos 2\theta} \approx 1$$

$$\theta \text{ piccoli} \Rightarrow m(x)|_{\substack{x_R=0 \\ Z_R = \frac{b}{2 \sin \theta} + \Delta Z}} \approx f \left(-b \frac{\Delta Z}{Z_R^2} \right)$$

DA QUI SI DEDUCE COME VARIA LA DISPARITÀ NELLO INTORNO DEL PUNTO DI FISSAZIONE f .

Allora

$$d(x) \triangleq x_L - x_R = m(x) - 0 \approx -fb \frac{\Delta Z}{Z_R^2}$$

Nota: può essere interessante anche descrivere il comportamento della disparità in termini angolari.

4.2. ALGORITMI PER LA VISIONE STEREOSCOFICA

Con $x_0 = p^{\text{to}}$ generico dell'immagine, dati $I_R(x)$ e $I_L(x)$ si vuole determinare la disparità $d(x_0)$.

Per quanto riguarda gli algoritmi, possiamo suddividerli in 2 categorie:

- ① ALGORITMI BASATI SULLA CORRISPONDENZA ESPLICITA
- ② ALGORITMI SENZA CORRISPONDENZA ESPLICITA

Analizziamoli + in dettaglio

① ALGORITMI BASATI SULLA CORRISPONDENZA ESPLICITA

Si determinano le caratteristiche dell'immagine s_n e d_n , e utilizzando bordi e linee si ricava la disparità [v. FIG. 4.8.].

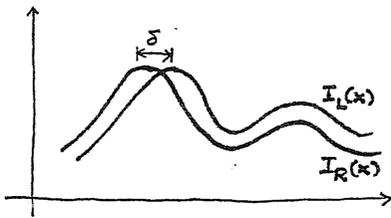
Questo approccio genera valori di disparità in p^{ti} sparsi dell'immagine (se se ne individuano troppe, si ha difficoltà nel determinare le appartenenze).

② ALGORITMI SENZA CORRISPONDENZA ESPLICITA

Con q. metodo si analizzano le funzioni $I_R(x)$ e $I_L(x)$ x individuare le porzioni di segnale corrispondenti alla stessa sorgente.

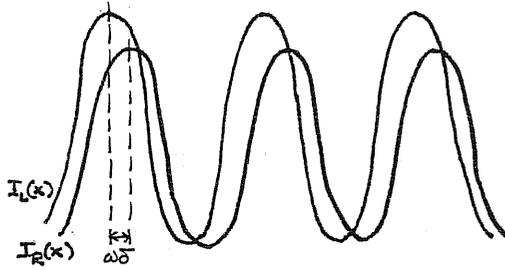
la corrispondenza può essere:

- BASATA SULL'AMPIEZZA
- BASATA SULLA FASE



BASATO SULL'AMPIEZZA

δ = VALORE DI SCORRIMENTO PER IL QUALE LA DIFFERENZA $I_R - I_L$ È MINIMA



BASATO SULLA FASE

$\omega\delta$ = SFASAMENTO SPAZIALE
ANALIZZA COME SONO SPAZIALI LE COMPONENTI ARMONICHE.

[METODO + ROBUSTO]

Nota: il profilo dei 2 segnali non è mai lo stesso. Esso dipende, ad es, dalla posizione della sorgente.

Bisogna sempre fare un'analisi locale, non globale, delle corrispondenze tra le distribuzioni di intensità. In genere si considerano delle finestre spaziali.

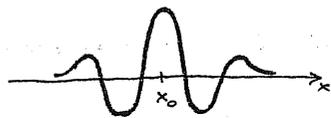
3.1. METODI BASATI SULLA FASE

Da la distribuzione $I(x)$, avremo che

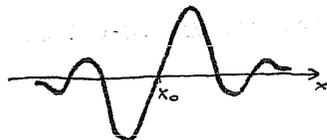
$$\begin{aligned} I(x) &\xrightarrow{\mathcal{F}} i(\omega) \\ I(x+\Delta x) &\xrightarrow{\mathcal{F}} i(\omega) e^{-j\omega\Delta x} \end{aligned}$$

$d(x_0)$
DIFFERENZA DI FASE

Per fare un'analisi spettrale su base locale (nelle coordinate spaziali e in quelle frequenziali), introduciamo allora le FUNZIONI DI GABOR



cos



sin

Utilizziamo dunque la nostra immagine attraverso una funzione pari (cos) e una dispari (sin):

$$C_R(x_0) = \int_{\mathbb{R}} I_R(x) G(x_0 - x) dx$$

\mathbb{C} \mathbb{R} \mathbb{C}

$$C_L(x_0) = \int_{\mathbb{R}} I_L(x) G(x_0 - x) dx$$

\mathbb{C} \mathbb{R} \mathbb{C}

[$G(x_0 - x)$ È UNA FUNZIONE COMPLESSA
PICHÈ FILTRIAMO SIN E COS]

Le quantità C_L e C_R sono complesse, e ne possiamo considerare la differenza di fase

$$\angle C_L(x_0) - \angle C_R(x_0) \triangleq \Delta\phi(x_0) \cong \omega_0 d(x_0)$$

Possiamo motivare meglio q. risultato: definiamo

$$\begin{aligned} \pi(\omega) &= \mathcal{F}\{I_R(x)\} = \mathcal{F}\{R(x)\} \\ L(\omega) &= \mathcal{F}\{I_L(x)\} = \mathcal{F}\{L(x)\} \end{aligned}$$

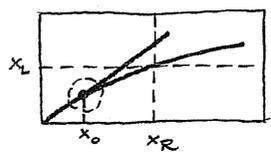
Calcoliamo

$$C_L = \int G(x) L(x_0 - x) dx$$

Di qui in poi il calcolo è esplicitato nell'allegato A.

$$I_L(x) = R(m(x)) \quad \text{con } x_L = m(x) = m(x_R)$$

Non è necessario conoscere esplicitamente la funzione $m(x)$, poiché è localmente approssimabile con Taylor:



$$m(x) = Ax + B$$

NELL'INTORNO DEL GENERICO PUNTO x_0 APPROX LA CURVA CON UNA RETTA (1° ORDINE)

In tal modo l'eq^{me} finale diventa

$$\begin{aligned} C_L &= \frac{1}{2\pi} \int g(A\omega - \omega_0) \pi(\omega) e^{j\omega Ax_0} e^{j\omega B} d\omega \\ C_R &= \frac{1}{2\pi} \int g(\omega - \omega_0) \pi(\omega) e^{j\omega x_0} d\omega \end{aligned}$$

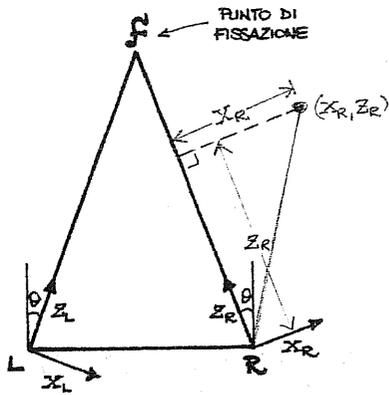
↑
FUNZIONE DI GABOR

La differenza tra le 2 eq^{mi} è data da A (=ingrandimento) nell'argomento della funzione g e dalla fase dell'esponenziale.

Considerando che $A \approx 1$, portando fuori il fattore $e^{j\omega B}$, avremo una differenza di fase

$$e^{j\omega B} \approx e^{j\omega_0 B} \Rightarrow \Delta\phi \cong \omega_0 B \quad \text{SFASAMENTO TRA R E L}$$

Questo metodo serve a estrarre l'info di fase, partendo dalla dislocazione delle 2 immagini.



Dato il punto immagine su R di coordinate x_R , il suo coniugato nello spazio delle sorgenti sarà

$$X_R = x_R \frac{Z_R}{f}$$

Le coordinate immagine su L di questa stessa sorgente sono:

$$x_L = f \frac{X_L}{Z_L}$$

dove

$$\begin{cases} X_L = X'_L \cos 2\theta - Z'_L \sin 2\theta \\ Z_L = X'_L \sin 2\theta - Z'_L \cos 2\theta \end{cases} \quad \text{con} \quad \begin{cases} X'_L = X_R + b \cos \theta \\ Z'_L = Z_R - b \sin \theta \end{cases}$$

Sostituendo si ricava

$$x_L = f \frac{\left(x_R \frac{Z_R}{f} + b \cos \theta\right) \cos 2\theta - (Z_R - b \sin \theta) \sin 2\theta}{\left(x_R \frac{Z_R}{f} + b \cos \theta\right) \sin 2\theta + (Z_R - b \sin \theta) \cos 2\theta}$$

Normalizzando le coordinate immagine rispetto a f e le coordinate sorgente rispetto a b, si ha

$$\frac{x_L}{f} = \frac{(x Z + \cos \theta) \cos 2\theta - (Z - \sin \theta) \sin 2\theta}{(x Z + \cos \theta) \sin 2\theta + (Z - \sin \theta) \cos 2\theta} \quad \leftarrow \text{dove} \quad \begin{cases} x \hat{=} x_R/f \\ Z \hat{=} Z_R/b \end{cases}$$

Da questa si ha

LA DIPENDENZA DI x_L NORMALIZZATO È SOLO DA θ

$$d(x) = \text{DISPARITÀ NORMALIZZATA} \hat{=} \frac{x_L}{f} - \frac{x_R}{f} = \frac{\cos \theta - (x^2 + 1) Z \sin 2\theta - x \sin \theta}{Z \cos 2\theta + x Z \sin 2\theta + \sin \theta} \quad (*)$$

Nota: il risultato (*) può essere verificato considerando come un punto sorgente generico di coordinate (X_R, Z_R) si proietta sulle coordinate L.

Per un punto immagine x , nota $d(x)$, si può risalire a Z ($\hat{=} \frac{Z_R}{b}$):

$$Z = \frac{\cos \theta - x \sin \theta}{d(x) (\cos 2\theta + \sin \theta) + \sin 2\theta (x^2 + 1)} \quad \left(Z = \frac{\cos \theta - x \sin \theta + d \sin \theta}{d (\cos 2\theta + \sin 2\theta) + (x^2 + 1) \sin 2\theta} \right)$$

Considerando le variazioni di Z rispetto a $Z_0 \hat{=} \frac{1}{2 \sin \theta}$, coordinata $\frac{Z_R}{b}$ del punto di fissazione si ha

$$\Delta Z \hat{=} Z \Big|_{x=0} - Z_0 = \frac{d(x) (\cos 2\theta - \sin \theta)}{[d(x) (\cos 2\theta + \sin \theta) + \sin 2\theta] 2 \sin \theta}$$

↑
PUNTO DI FISSAZIONE $Z_0(f) = \frac{b}{2 \sin \theta}$

$$\frac{d}{f} \triangleq \frac{x_L}{f} - \frac{x_R}{f} = \frac{x_L}{f} - x = \frac{N}{D} = \frac{[(xZ + \cos \theta) \cos 2\theta] - (Z - \sin \theta) \sin 2\theta}{(xZ + \cos \theta) \sin 2\theta + (Z - \sin \theta) \cos 2\theta} =$$

$$= \frac{[(xZ + \cos \theta) \cos 2\theta - (Z - \sin \theta) \sin 2\theta] - [(x^2 Z + x \cos \theta) \sin 2\theta + (xZ - x \sin \theta) \cos 2\theta]}{(xZ + \cos \theta) \sin 2\theta + (Z - \sin \theta) \cos 2\theta}$$

Alora

$$\left[\begin{array}{l} \sin 2\theta = 2 \sin \theta \cos \theta \\ \cos 2\theta = \cos^2 \theta - \sin^2 \theta \end{array} \right]$$

$$\begin{aligned} D &= xZ \sin 2\theta + \cos \theta \sin 2\theta + Z \cos 2\theta - \sin \theta \cos 2\theta = \\ &= xZ \sin 2\theta + Z \cos 2\theta + 2 \cos^2 \theta \sin \theta - \sin \theta \cos^2 \theta + \sin^3 \theta = \\ &= xZ \sin 2\theta + Z \cos 2\theta + \cos^2 \theta (2 \sin \theta - \sin \theta) + \sin \theta (1 - \cos^2 \theta) = \\ &= xZ \sin 2\theta + Z \cos 2\theta + \sin \theta \end{aligned}$$

$$\begin{aligned} N &= [(xZ + \cos \theta) \cos 2\theta - (Z - \sin \theta) \sin 2\theta] - Dx = \\ &= [(xZ + \cos \theta) \cos 2\theta - (Z - \sin \theta) \sin 2\theta] - x^2 Z \sin 2\theta - xZ \cos 2\theta - x \sin \theta = \\ &= xZ \cos 2\theta + \cos \theta \cos 2\theta - Z \sin 2\theta + \sin \theta \sin 2\theta - x^2 Z \sin 2\theta - xZ \cos 2\theta - x \sin \theta = \\ &= \cos^3 \theta - \cos \theta \sin^2 \theta + 2 \sin^2 \theta \cos \theta - (x^2 + 1) Z \sin 2\theta - x \sin \theta = (*) \\ &= \cos \theta (\cos^2 \theta + \sin^2 \theta) + \sin^2 \theta (-\cos \theta + \cos \theta) + (x^2 + 1) Z \sin 2\theta - x \sin \theta = \\ &= \cos \theta - (1 + x^2) Z \sin 2\theta - x \sin \theta \end{aligned}$$

Segue dunque che

$$d_m = \frac{N}{D} = \frac{\cos \theta - (1 + x^2) Z \sin 2\theta - x \sin \theta}{xZ \sin 2\theta + Z \cos 2\theta + \sin \theta} \quad \begin{array}{l} \text{DISPARITÀ} \\ \text{NORMALIZZATA} \end{array}$$

$$\begin{aligned} (*) : \quad \cos^3 \theta - \cos \theta \sin^2 \theta + 2 \sin^2 \theta \cos \theta &= \cos \theta (\cos^2 \theta - \sin^2 \theta + 2 \sin^2 \theta) = \\ &= \cos \theta (\underbrace{\cos^2 \theta + \sin^2 \theta}_1) = \cos \theta \end{aligned}$$

```
%Disparita`
```

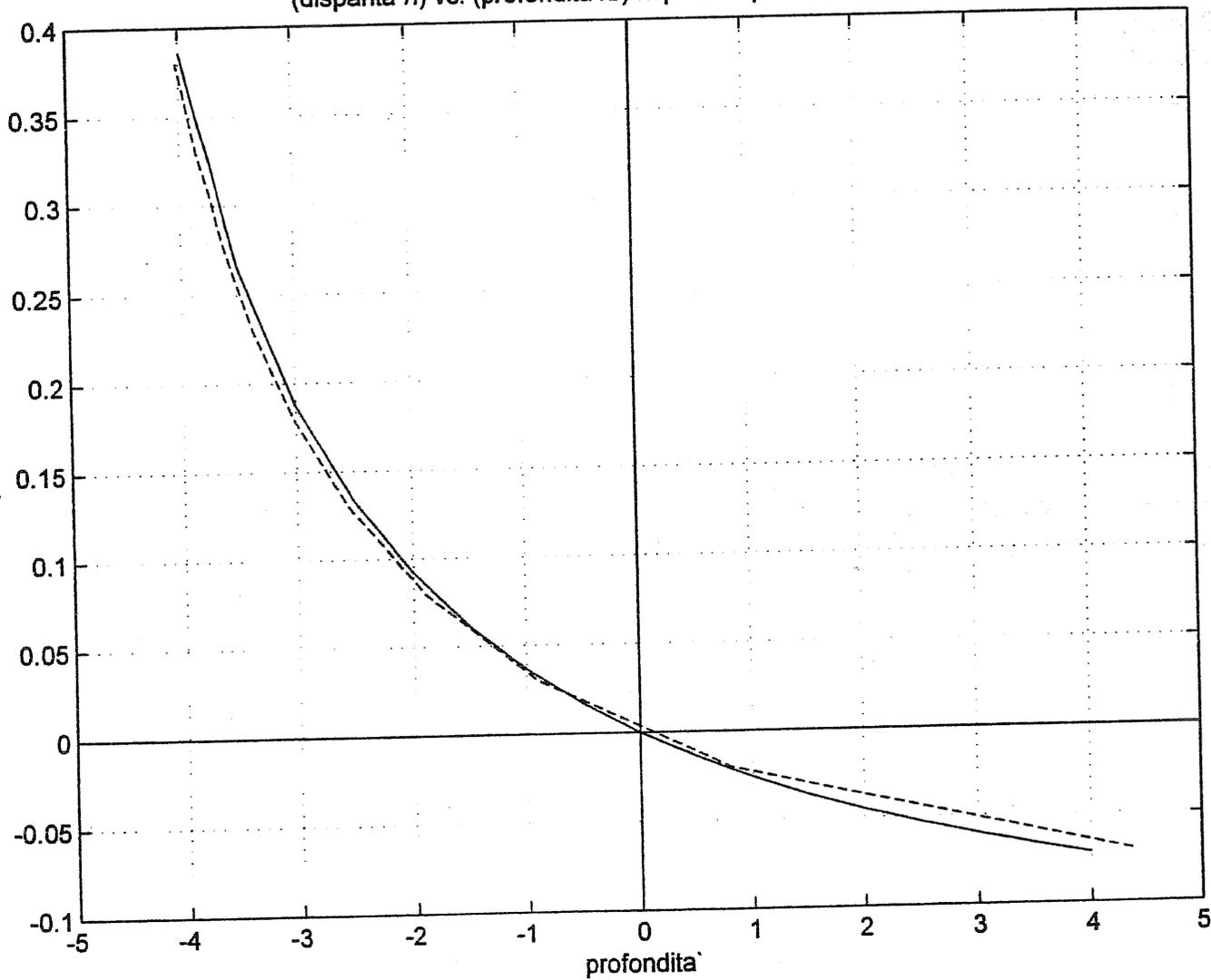
```
vergenza=10;  
teta=(pi/180)*(vergenza/2);  
x=0;  
DZ=-4:0.5:+4;  
Z=1./(2*sin(teta))+DZ;  
XL=cos(teta)-Z*sin(2*teta)+Z*x*cos(2*teta);  
ZL=sin(teta)+Z*cos(2*teta)+x*Z*sin(2*teta);  
xL=XL./ZL;  
d=xL-x;  
  
d1=-0.07:0.05:0.4;  
DZ1=(-(cos(2*teta)+sin(teta)).*d1)./(((cos(2*teta)+sin(teta)).*d1  
1...  
+sin(2*teta))*2*sin(teta));
```

```
figure
```

```
plot(DZ,d, DZ1,d1,'--') ,grid on, axis auto  
xlabel('profondita`')  
ylabel('disparita`')  
title('(disparita`/f) vs. (profondita`/b) rispetto al punto di f  
issazione')  
set(gca,'XTick',[-5:1:5])  
line ([0,0],[-0.1,0.4]), line ([-5,5],[0,0])
```

-- = RELAZIONE INVERSA: DATA LA DISPARITA'
SI RISALE ALLA PROFONDITA'

(disparita' / f) vs. (profondita' / b) rispetto al punto di fissazione



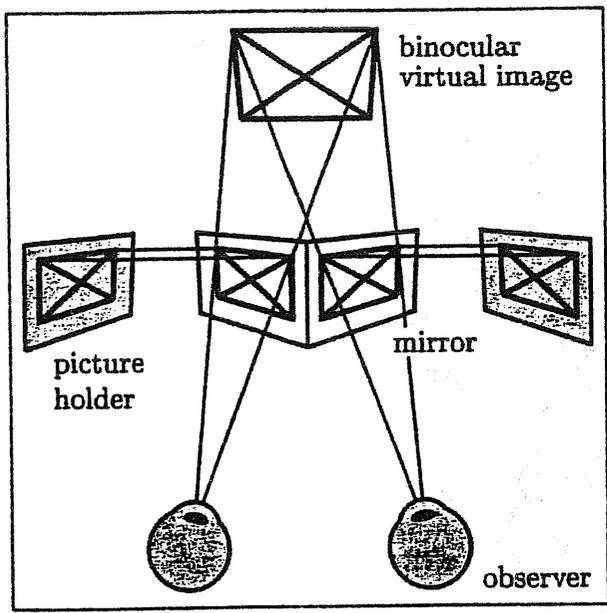


FIG. 4.1. The mirror stereoscope as devised by Wheatstone. The two half-images are inserted into the image holders and observed in the two mirrors. A fused image results, with a strong stereoscopic impression of depth.

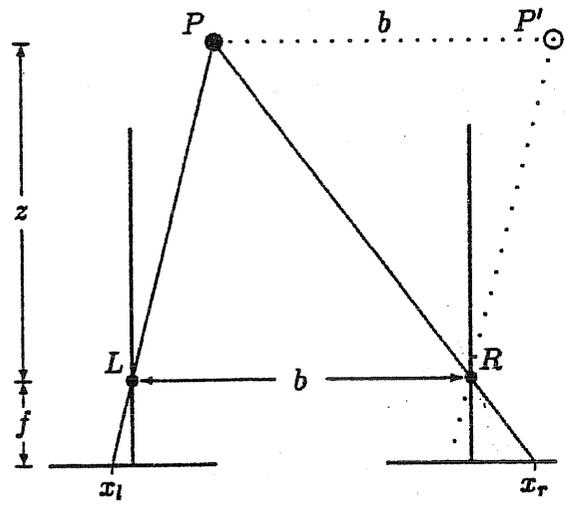


FIG. 4.2. Horizontal disparities with parallel camera axes. The point P is imaged in the left and right images with the x coordinates x_l and x_r . The disparity is $d = x_l - x_r$. Parallel transposition of the left ray from P to P' and application of parallel proportionality theorem results in: $d = -f \cdot b / z$.

$\delta = \frac{\alpha}{2} + \varphi_c = \frac{1}{2}(\varphi_l + \varphi_r) = \text{ANGOLO DI VERSIONE}$
 $\alpha, \delta = \text{COORDINATE ANGOLARI per descrivere la posizione dei punti sorgente}$

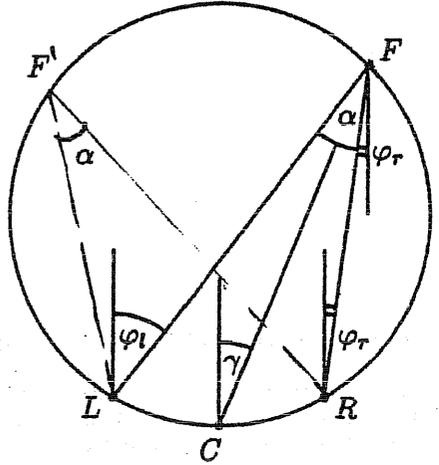


FIG. 4.3. Binocular geometry with convergent camera alignment. The angle between the lines from the fixation point F to the two nodal points (L, R) is the vergence angle α . All points on the circle through L, R and F (Vieth-Müller circle) have the same vergence angle. The angles between the current direction of view and the rest position of the eyes ("straight ahead" relative to the orientation of the head) are φ_l and φ_r . C is the "cyclopean point," which bisects the arc between the two nodal points. The direction of view measured with reference to this point is called the angle of version and is denoted by γ .

Individuazione dei pt nello spazio in funzione delle coordinate angolari.
 L e R sono 2 telecamere con assi ottici non paralleli, ma orientati con angoli φ_l e φ_r .
 Gli assi si incontrano in F = PUNTO DI FISSAZIONE.
 Per proprietà geometriche
 $\varphi_c = \alpha + \varphi_r$
 $\Rightarrow \alpha = \varphi_l - \varphi_r$

Tutti i pt F' individuati dalla circonferenza come intersezione degli assi ottici formano angoli α , poiché insistono sullo stesso arco.

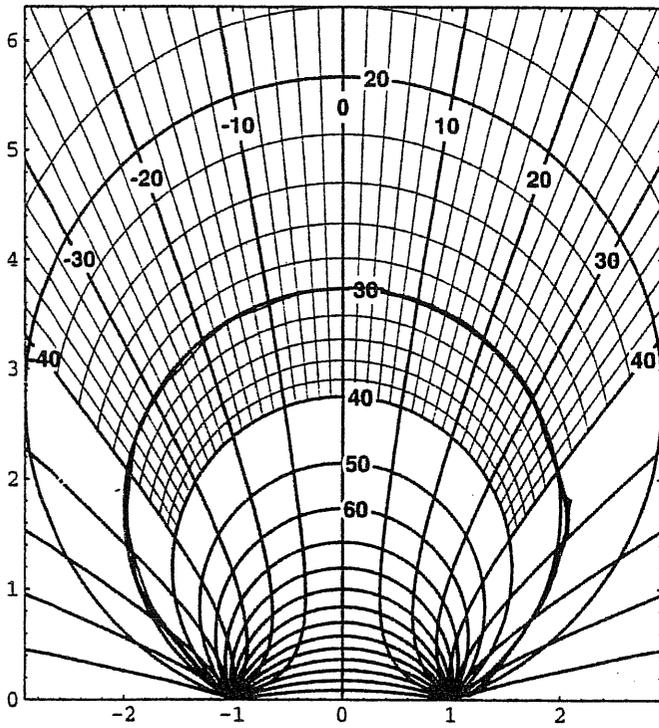


FIG. 4.4. Hering coordinates in the horizontal plane. The circles are curves of constant vergence (Vieth-Müller circles), and the hyperbolas are curves of constant version (Hillebrand hyperbolas). The nodal points of the eyes are at $R, L = (\pm 1, 0)$; the proportions become those of the human visual system if all distances are multiplied by 3.25 cm. The thick lines form a 10° grid; the thin lines, a 2° grid. The coarser spacing with increasing distance from the nodal points corresponds to the decreasing accuracy of stereoscopic vision with increasing distance.

La struttura additiva della relazione in FIG. 4.6. consente di utilizzare la FIG. 4.4. per il calcolo della disparità.

Preso il pto P_1 :

$$\alpha = \alpha_{P_1} \Rightarrow \delta(P_1) = \alpha_{P_1} - \alpha_P$$

$\alpha = \cos^2 \Rightarrow$ CERCHI

$\gamma = \cos^2 \Rightarrow$ IPERBOLI

Se L, R sono PII NODALI (il raggio me esce indistinto):

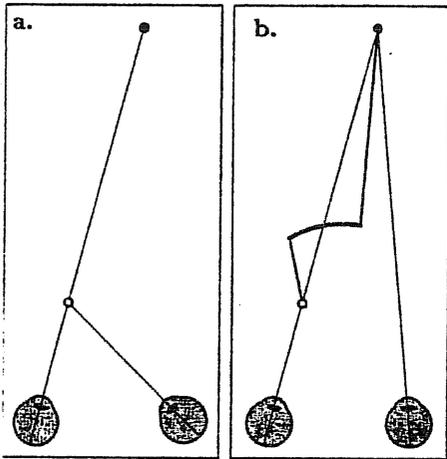
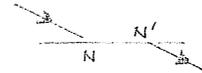
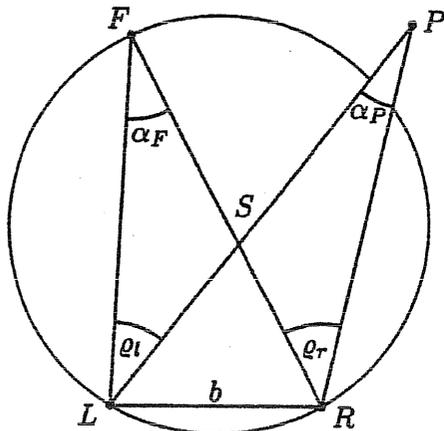


FIG. 4.5. Vergence and version as internal variables of binocular eye movements (Hering's law). **a.** Both eyes fixate on the point indicated by \bullet . Looking from the left eye, the next fixation point (\bullet) lies directly behind the first one, so that no movement of the left eye is actually necessary. **b.** When moving fixation to the more distant point, the left eye first turns to the left, then to the right and again to the left, until it finally reaches the same position it had initially. The intersection of the axes of vision moves alternately along curves of constant version and of constant vergence.



Dato S:

$$\alpha_F + \rho_L = \alpha_P + \rho_R$$

$$\rho_L - \rho_R = \alpha_P - \alpha_F$$

$\delta = \text{DISPARITÀ ANGOLARE}$

FIG. 4.6. Calculation of horizontal disparity δ as an angle. $\alpha = \alpha_P$: vergence. P : any point which is present in the peripheral visual field while fixation is on F . α_P : Target vergence of P . ρ_L, ρ_R : Azimuth of the point P in camera-centered polar coordinates. S : Intersection of the straight lines \overline{LP} and \overline{RF} . Simple trigonometry yields: $\delta = \rho_R - \rho_L = \alpha_P - \alpha_F$

$\rho = \text{ECCENTRICITÀ} = \text{posizione angolare di P rispetto all'asse ottico (R/L)}$

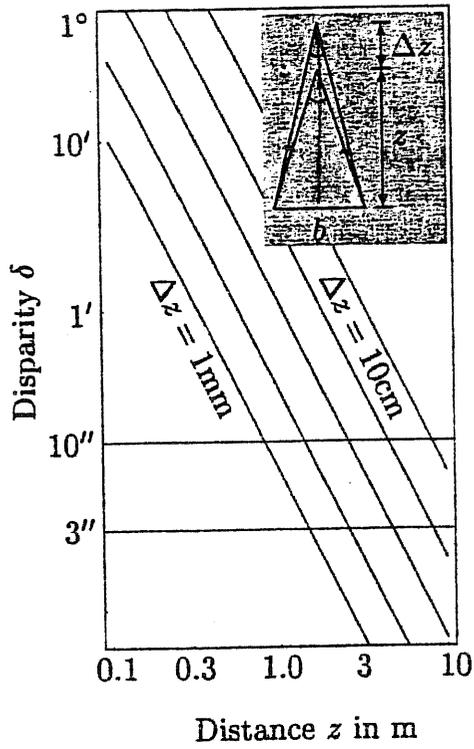


FIG. 4.7. Relationship between depth and horizontal disparity in the median plane. Log-log scale. The lines show the disparity for a fixed depth offset Δz viewed at the distance z ; this disparity decreases approximately as the square of distance. The base line is assumed to be $b = 6.5$ cm. The five lines correspond to the depth offsets $\Delta z = 1$ mm, 3 mm, 1 cm, 3 cm and 10 cm. The horizontal lines indicate the threshold of perception. At a distance of 3 meters, depth variations of less than 1 cm can no longer be resolved stereoscopically. The inset drawing shows the geometric construction which the main graph represents. The disparity corresponds to the difference between the two apex angles.

Calcolo esplicito della disparità con un calcolo geometrico analitico

$$\begin{aligned}
 \delta(z, \Delta z) &= 2 \left[\arctg\left(\frac{b/2}{z}\right) - \arctg\left(\frac{b/2}{z+\Delta z}\right) \right] \approx b \frac{\Delta z}{z^2} \\
 \uparrow \quad \uparrow \\
 F \quad P=z+\Delta z
 \end{aligned}$$

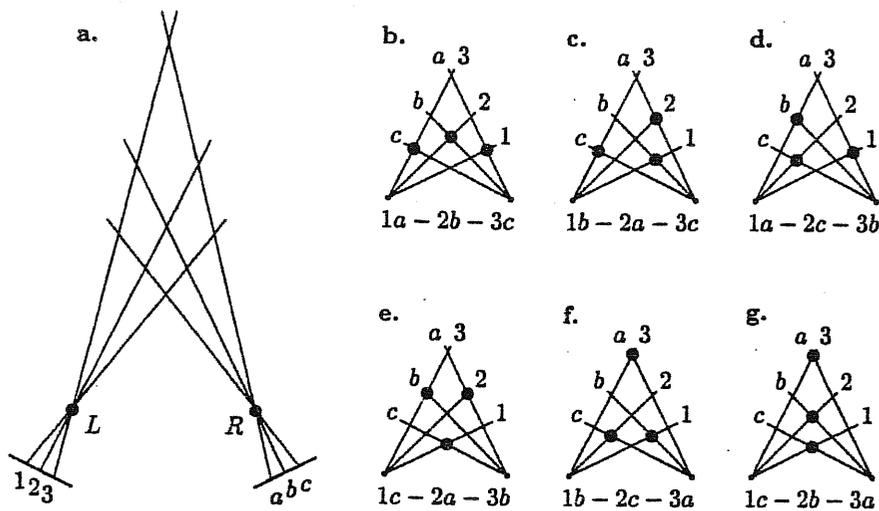


FIG. 4.8. The correspondence problem of stereoscopic vision. **a.** The left image has features at the locations 1, 2, 3, the right image at *a, b, c*. There are $3! = 6$ possible ways to match these three pairs of features (**b. - g.**), which meet the requirements of the uniqueness and completeness. Each match leads to a different spatial interpretation. *L, R*: Nodal points of the left and right eyes.

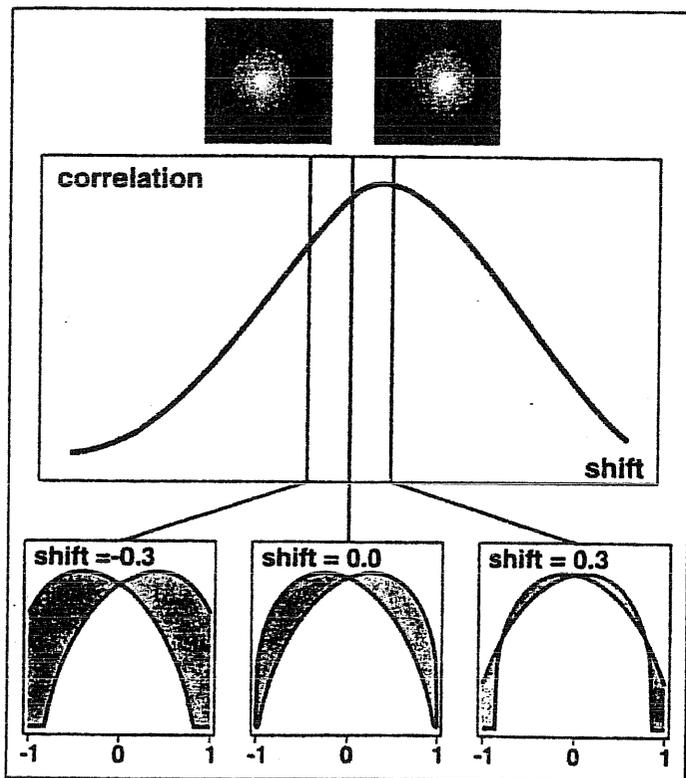


FIG. 4.9. Diagram illustrating the correlation algorithm for determining disparities based on intensities. The lower inset graphs show intensity profiles along horizontal lines across the half-images (top) for different displacements or shifts s . The gray-shaded areas represent the difference between the two gray value functions. The correlation function shown in the larger graph is maximized when the displacement $s = 0.3$. Except for an additive constant, this function corresponds to the function $-\Phi_c(x, s)$ in Equation 6.8.

```

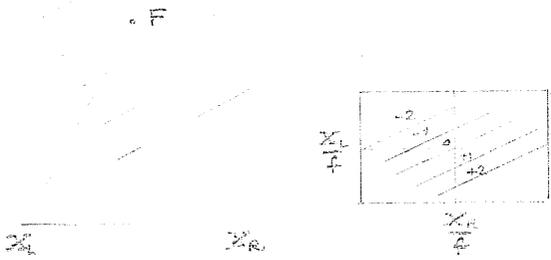
%Disparita`
clear D
clear Z
clear IL
vergenza=15;
teta=(pi/180)*(vergenza/2);
x=-1:0.01:+1;
for i=1:5;
    DZ(i)=(i-3);
    Z(i)=1./(2*sin(teta))+DZ(i);
    XL=cos(teta)-Z(i)*sin(2*teta)+x*Z(i)*cos(2*teta);
    ZL=sin(teta)+Z(i)*cos(2*teta)+x*Z(i)*sin(2*teta);
    xL=XL./ZL;
    D(i,:)=xL;

    IL(i,:)=cos(xL*pi*10);
end
IR=cos(x*pi*10);

figure
subplot(2,1,1), plot(x,D(1,:),x,D(2,:),x,D(3,:),x,D(4,:),x,D(5
,:),:)) , grid on
axis auto
title('xL vs. xR at various depths')
xlabel('xR/f')
ylabel('xL/f')
line ([0,0],[-1.5,1]), line ([-1,1],[0,0])

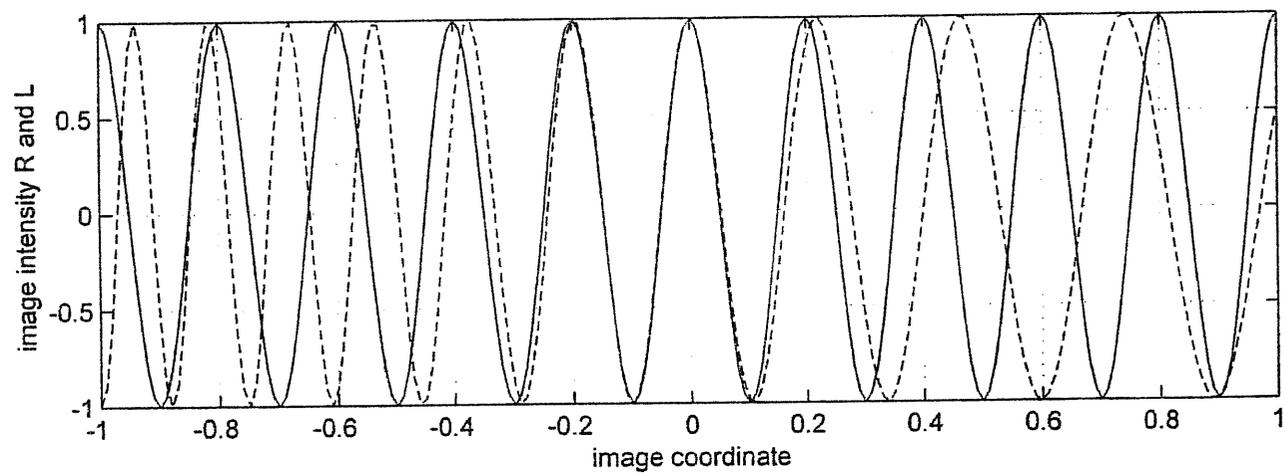
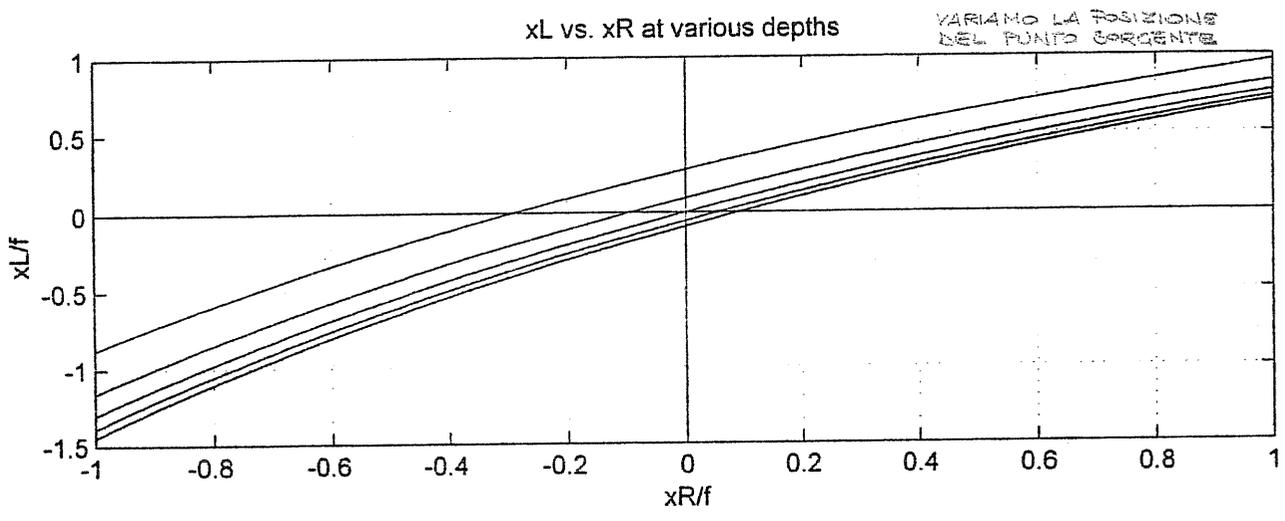
subplot(2,1,2), plot(x,IR,x,IL(3,:),'--'), grid on, axis auto,
hold on,
set(gca,'XTick',[-1:0.20:1])
xlabel('image coordinate')
ylabel('image intensity R and L')

```



$DE(x) = -2 \div 2$
 Variazione a passi di 1

I p.e. hanno Z_R specificata, ma X_R variabile



Se $X_R = X_L$ i 2 grafici sono identici, e non sfasati.
 Inoltre lo sfasamento non è uniforme, è aumentato la differenza tra le 2 coordinate.

Picchi e valli ci danno info sulle difformità

$$G * L(x_0) = \int G(x) L(x_0 - x) dx = \int G(x) \left[\frac{1}{2\pi} \int \pi(\omega) e^{jm(x_0 - x)} d\omega \right] dx =$$

INVERTENDO
L'ORDINE DI
INTEGRAZIONE

$$\stackrel{\downarrow}{=} \frac{1}{2\pi} \int \pi(\omega) \left[\int G(x) e^{j\omega m(x_0 - x)} dx \right] d\omega \stackrel{\downarrow}{=}$$

CAMBIAMENTO DI VARIABILI:

$$z \hat{=} m(x_0 - x)$$

$$x = x_0 - m^{-1}(z)$$

$$dx = -\frac{d}{dz}(m^{-1}(z)) dz$$

$$= \frac{1}{2\pi} \int \pi(\omega) \left[\int G(x_0 - m^{-1}(z)) e^{j\omega z} \left(-\frac{d}{dz}(m^{-1}(z))\right) dz \right] d\omega =$$

IN PARTICOLARE SE:

$$m(x) = Ax + B$$

$$z \hat{=} m(x_0 - x) =$$

$$= A(x_0 - x) + B$$

$$x = x_0 - \frac{z - B}{A}$$

$$dx = -\frac{dz}{A}$$

$$\stackrel{\downarrow}{=} \frac{1}{2\pi} \int \pi(\omega) \left[\int G\left(x_0 - \frac{z - B}{A}\right) e^{j\omega z} \left(-\frac{dz}{A}\right) \right] d\omega \stackrel{\downarrow}{=}$$

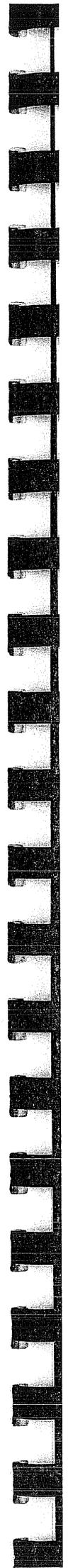
PONENDO $y \hat{=} -\frac{z}{A}$

$$= \frac{1}{2\pi} \int \pi(\omega) \left[\int G\left(x_0 + y + \frac{B}{A}\right) e^{-j\omega A y} dy \right] d\omega \stackrel{\downarrow}{=}$$

PONENDO

$$x \hat{=} x_0 + y + \frac{B}{A} \Rightarrow Ay = Ax - B - Ax_0$$

$$= \frac{1}{2\pi} \int \pi(\omega) \left[\underbrace{\int G(x) e^{-j\omega Ax} dx}_{g(A\omega - \omega_0)} \right] e^{j\omega Ax_0} e^{j\omega B} d\omega$$



Definite

$$f(\omega) = \int_{-\infty}^{+\infty} F(x) e^{-j\omega x} dx \quad (\mathcal{F})$$

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{j\omega x} d\omega \quad (\mathcal{F}^{-1})$$

sia

$$\underbrace{e^{-\frac{x^2}{2\sigma^2}}}_{\text{GAUSSIANA}} \xrightleftharpoons[\mathcal{F}^{-1}]{\mathcal{F}} e^{-\frac{\omega^2}{2\tau^2}} \quad (\tau \triangleq \frac{1}{\sigma})$$

Pertanto sia

$$g(x; \omega_0) \triangleq e^{-\frac{x^2}{2\sigma^2}} e^{j\omega_0 x} \xrightleftharpoons[\text{TEOREMA MODULAZIONE}]{\downarrow} g(\omega - \omega_0) \triangleq e^{-\frac{(\omega - \omega_0)^2}{2\tau^2}}$$

$$g(x - x_0; \omega_0) \triangleq e^{-\frac{(x - x_0)^2}{2\sigma^2}} e^{j\omega_0(x - x_0)} \xrightleftharpoons[\text{TEOREMA TRASLAZIONE}]{\downarrow} g(\omega - \omega_0; x_0) \triangleq e^{-\frac{(\omega - \omega_0)^2}{2\tau^2}} e^{-j(\omega - \omega_0)x}$$

**VISIONE
ATTIVA**

VISIONE ATTIVA

VISIONE ATTIVA = PROCESSO DI ACQUISIZIONE (INTER-)ATTIVA DELL'INFO VISIVA NON SOLO ELABORAZIONE, MA ANCHE ESPLORAZIONE

PURPOSIVE VISION: non si risolve un pb di direzione se non si conosce l'obiettivo dell'assegnazione.

- INTERAZIONE TRA IL SOGGETTO E L'AMBIENTE
- ADATTABILITÀ DELLO SCHEMA DI ACQUISIZIONE

INDICE

- MOTIVAZIONI
- CARATTERISTICHE COMPORTAMENTALI DELLA VISIONE ATTIVA
- STRATEGIE MORFOLOGICHE ED ARCHITETTURALI PER LA VISIONE ATTIVA
- COMPORTAMENTI OPTOCINETICI

MOTIVAZIONI

- IL TERMINE "**SENSORE ATTIVO**" GENERALMENTE SI ASSOCIA AD UN SENSORE CHE TRASMETTE UN ^(RADIAZIONE E.M.) SEGNALE NELL'AMBIENTE E MISURA IL SEGNALE RIFLESSO (TIPICAMENTE UNA RADIAZIONE ELETTROMAGNETICA):

- RADAR
- SONAR
- ULTRASUONI
- MICROONDE
- LUCE COLLIMATA

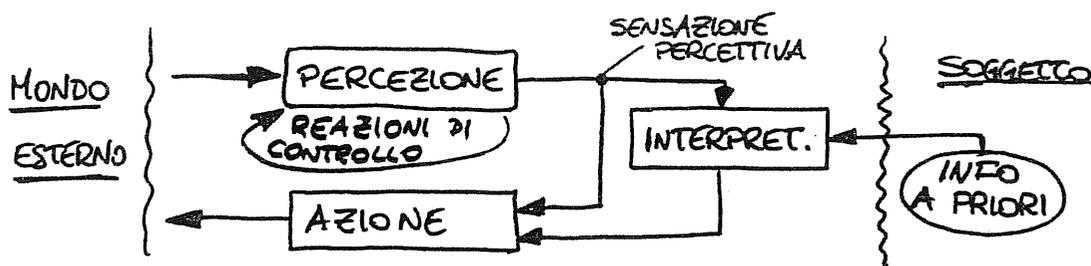
È invasiva: ci sono segnali che inquinano l'ambiente

- VISIONE ATTIVA ≠ SENSORI ATTIVI

IN GENERALE UN PROCESSO DI VISIONE ATTIVA COINVOLGE UN'INTERAZIONE CON L'AMBIENTE.

TALE INTERAZIONE AVVIENE PER MEZZO DI SENSORI PASSIVI UTILIZZATI IN MODO ATTIVO, MODIFICANDO LO STATO DEI SENSORI IN FUNZIONE DEL COMPITO DA ASSolverE.

• PERCEPTION-ACTION CYCLE



- SI POSSONO EVIDENZIARE DELLE REAZIONI (O RIFLESSI) QUALI SCORCIATOIE TRA PERCEZIONE E AZIONE
- LA MAGGIOR PARTE DI TALI REAZIONI È RAPPRESENTATA DA REAZIONI DI CONTROLLO CHE FORMANO CICLI CHIUSI COI QUALI IL SOGGETTO SI ADATTA AL COMPITO PERCETTIVO. OVVERO, MODIFICA LA SUA ATTITUDINE O PREDISPOSIZIONE ALLA PERCEZIONE IN RELAZIONE AI CAMBIAMENTI
 - DELL'AMBIENTE ESTERNO
 - DELLE RELAZIONI TRA IL SOGGETTO E L'AMBIENTE ESTERNO

Il compito che abbiamo cambia il nostro modo di percepire.

OSS.: LE RELAZIONI DI CONTROLLO POSSONO ESSERE:

*2 TRA GLI
COMPORTAMENTI
ATTIVI*

- INTENZIONALI
(AD EX.: RICERCA DI PARTICOLARI NELLA SCENA)
- COMPENSATIVE
(AD EX.: STABILIZZAZIONE DELLE IMMAGINI, COMPENSAZIONE DEL MOTION BLURRING)

LA COMPrensIONE DI TALI REAZIONI DI CONTROLLO È ESTREMAMENTE IMPORTANTE DAL PUNTO DI VISTA DELLA COSIDDETTA VISIONE COMPORTAMENTALE (= VISUALLY-GUIDED BEHAVIOUR)

- NEI SISTEMI NATURALI → COMPrensIONE DELLE STRATEGIE ADOTTATE
- NEI SISTEMI ARTIFICIALI → DOTARE UN SISTEMA ROBOTICO DI FUNZIONALITÀ PERCETTIVE AUTONOME

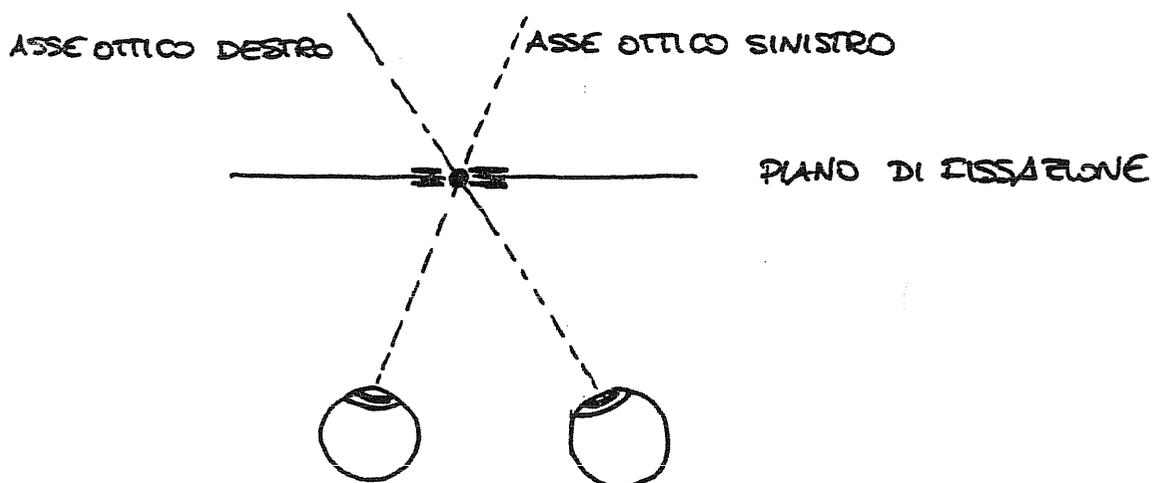
OSSERVAZIONI CONCLUSIVE

1. IL PROCESSO PERCETTIVO APPARE PIÙ COMPLESSO DI QUANTO CI ASPETTAVAMO : COINVOLGE “MISURE DINAMICHE” DEL SEGNALE VISIVO (PROCESSI DI ADATTAMENTO, ETC.)
2. SI METTE IN EVIDENZA CHE PERCEZIONE NON VUOL DIRE SOLO “PREPARAZIONE AL RICONOSCIMENTO”
3. IMPORTANZA DEI PERCORSI SUBCORTICALI DEL SISTEMA VISIVO

CARATTERISTICHE COMPORTAMENTALI DELLA VISIONE ATTIVA

GAZE CONTROL = CAPACITÀ DI MANTENERE O INDIRIZZARE LO SGUARDO SU UNA PARTICOLARE REGIONE DEL CAMPO VISIVO IN MODO DA MANTENERE LA SUA IMMAGINE SULLA RETINA CENTRATA RISPETTO ALLA FOVEA (=ZONA CENTRALE DELLA RETINA).

IN GENERALE, LA STABILIZZAZIONE DELL'IMMAGINE RISPETTO ALLA FOVEA SI OTTIENE PER I PUNTI DI UNA REGIONE ATTORNO AL PUNTO DI FISSAZIONE DEFINITO COME IL PUNTO DI INTERSEZIONE DEI DUE ASSI OTTICI:



WHAT/WHERE BEHAVIOURS: NEI PROCESSI PERCETTIVI SI POSSONO DISTINGUERE 2 DIFFERENTI MODALITÀ:

- IDENTIFICAZIONE →

WHAT

ANALIZZARE IN DETTAGLIO IL SEGNALE VISIVO RACCOLTO DALLA FOVEA DURANTE LA FISSAZIONE

- LOCALIZZAZIONE →

WHERE

DIRIGERE LO SGUARDO VERSO NUOVI "BERSAGLI"

DOVUTI A 2
CANALI ≠
HA // DI
ELABORAZIONE
DEL NOSTRO
CERVELLO

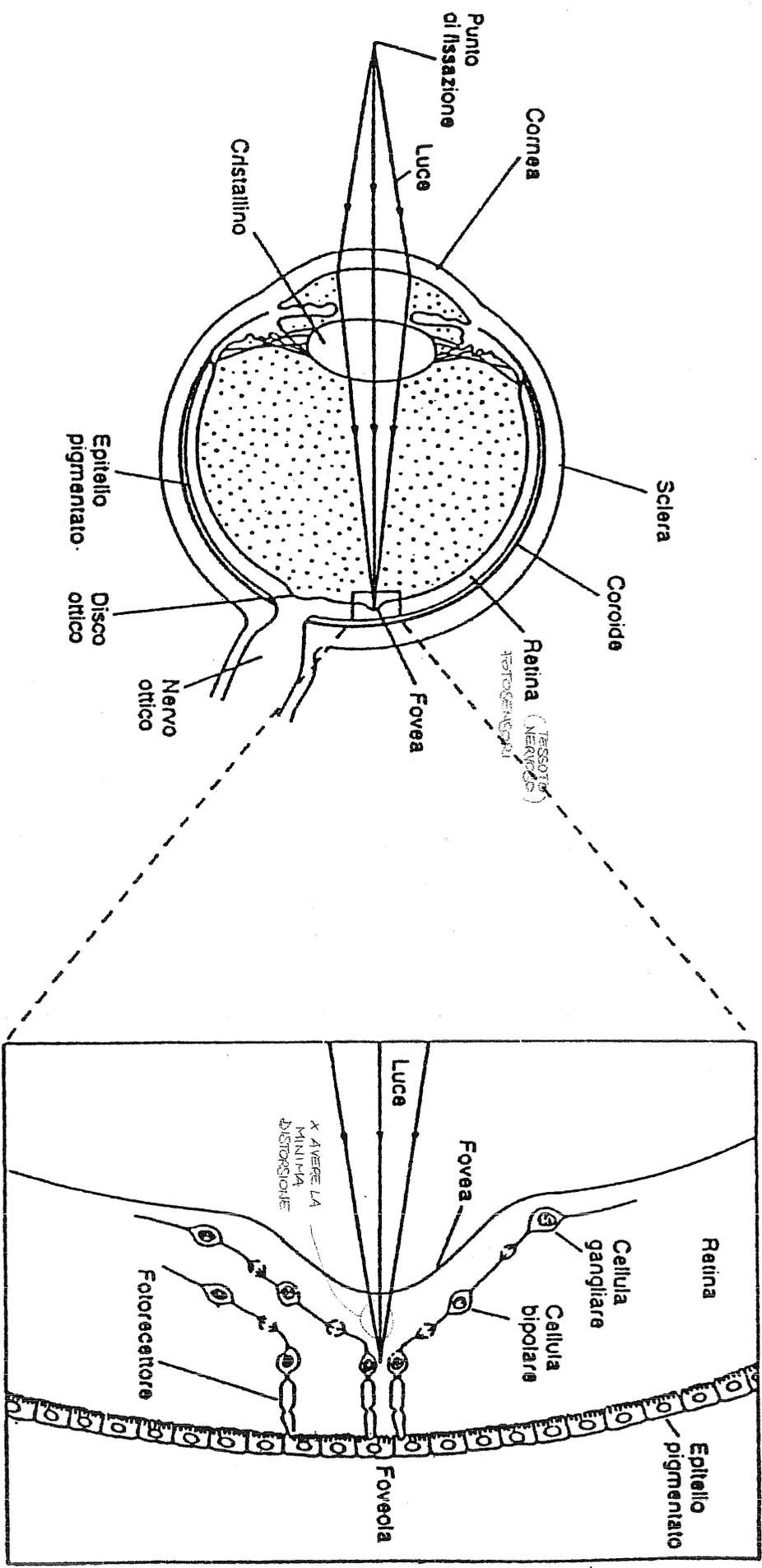


FIG. 1. I fotorecettori sono localizzati nella retina. La posizione della retina nell'occhio è illustrata nella figura a sinistra. I particolari della retina a livello della fovea sono illustrati nella figura a destra (lo schema è stato semplificato togliendo le connessioni laterali stabilite dalle cellule orizzontali e amacrine. In quasi tutta la retina la luce deve passare attraverso gli strati delle cellule nervose coi loro processi, prima di raggiungere i fotorecettori. Nel centro della fovea, la così detta foveola, gli elementi nervosi prossimali sono invece spostati lateralmente in modo che la luce possa raggiungere i fotorecettori direttamente. Di conseguenza, le immagini visive che si formano nella foveola hanno il minor grado di distorsione.

FOVEA = DEPRESSIONE DELLA RETINA CON UNA MAGGIOR CONCENTRAZIONE DI CONI

CONI E BASTONCELLI LAVORANO IN ZONE DIVERSE:

• CONI: CON + LUMINOSITÀ (GIORNO)

• BASTONCELLI: CON POCA LUCE

In zona si ha max concentrazione di cono e min di bastoncelli ⇒ VISIONE FOVEALE (=riconoscimento forme); la VISIONE PERIFERICA dà invece info di allerta.

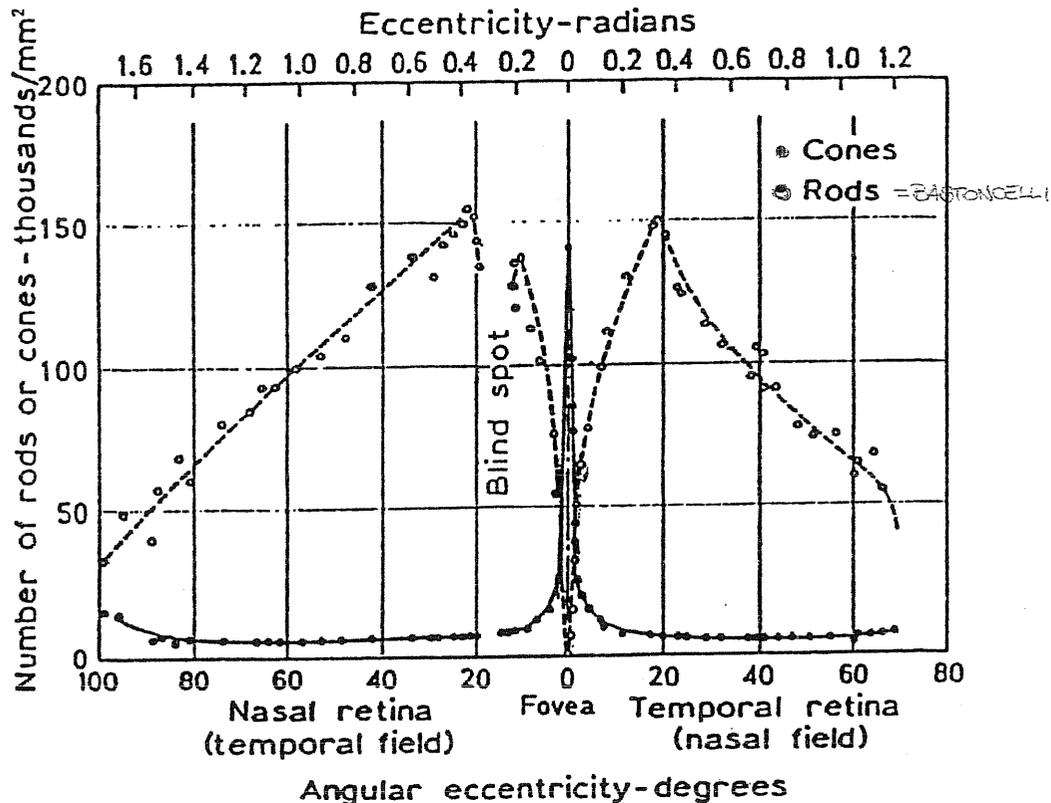


Fig. 2.6. Distribution of the rods and cones across the human retina in the horizontal meridian. (Reproduced from Østerberg, courtesy of Munksgaard, Copenhagen).

- **VISIONE FOVEALE** → ELEVATA RISOLUZIONE SPAZIALE
LEGATA ALL'ANALISI DEI DETTAGLI
 - RICONOSCIMENTO DI FORME
 - TRACKING

- **VISIONE PERIFERICA** → BASSA RISOLUZIONE SPAZIALE
FUNZIONALITÀ DI ALLERTA E DI GUIDA

APPARATO OCULOMOTORE

MUSCOLATURA DEI BULBI OCULARI

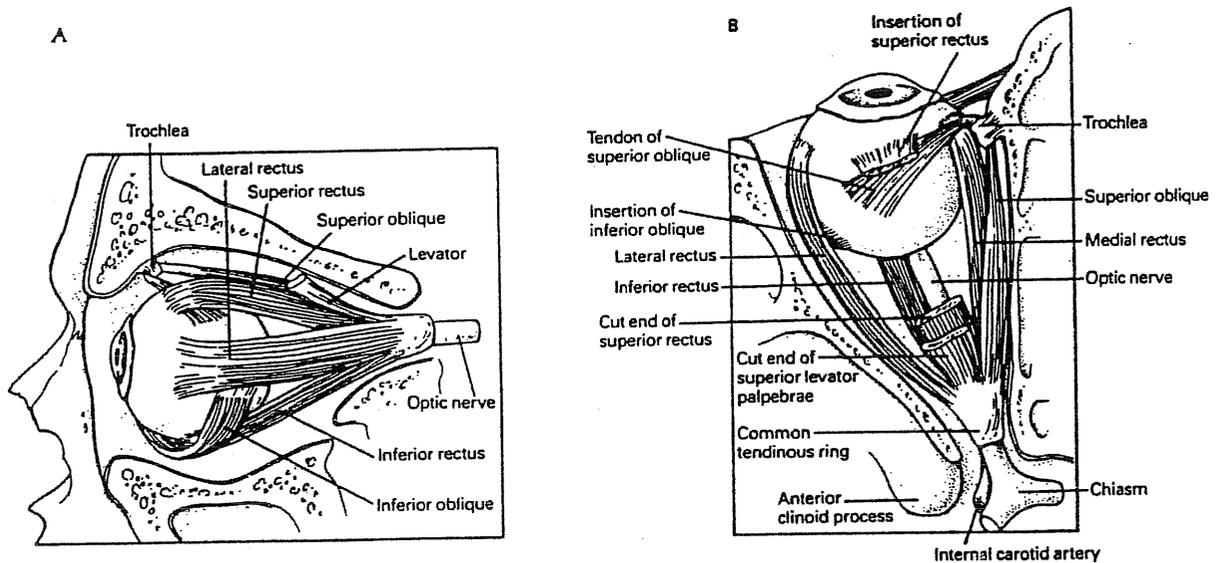


FIG. 2. (A) Cutaway drawing of the eye in its orbit and the three pairs of eye muscles that control its movement. Lateral view of the lateral rectus, superior and inferior rectus, and superior and inferior oblique muscles (medial rectus, note shown, is on the other side of the eyeball). (B) Cutaway view showing the inferior rectus on the medial surface of the orbit, and the origins and insertions of the muscles. The optic nerve follows the eyeball as it gyrates in its gimbal. The optic nerves from both eyes join at the chiasm (from Gouras, 1985)

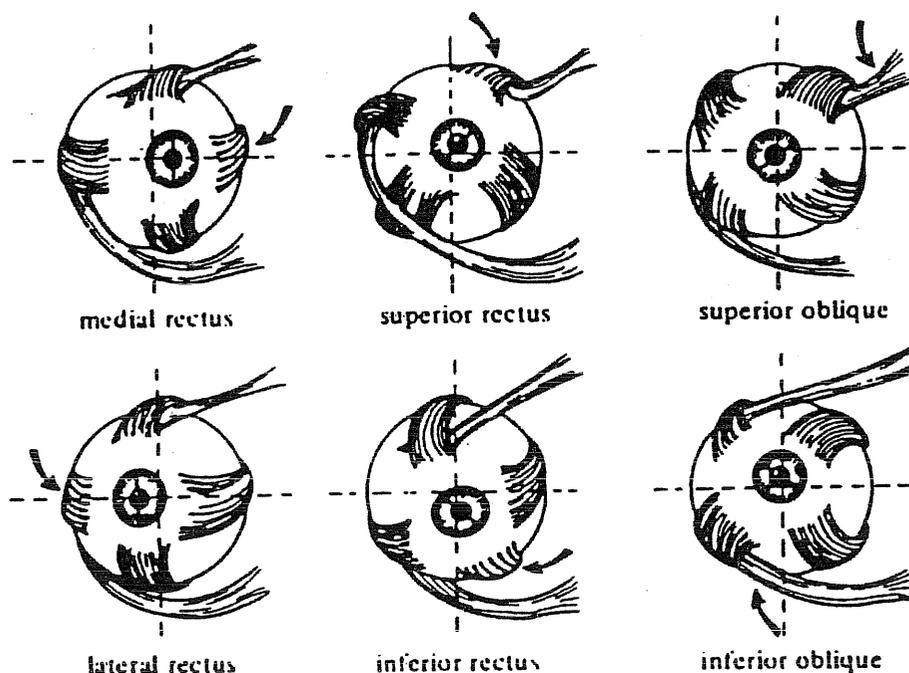


FIG. 3. I muscoli estrinseci degli occhi

NUCLEI OCULOMOTORI

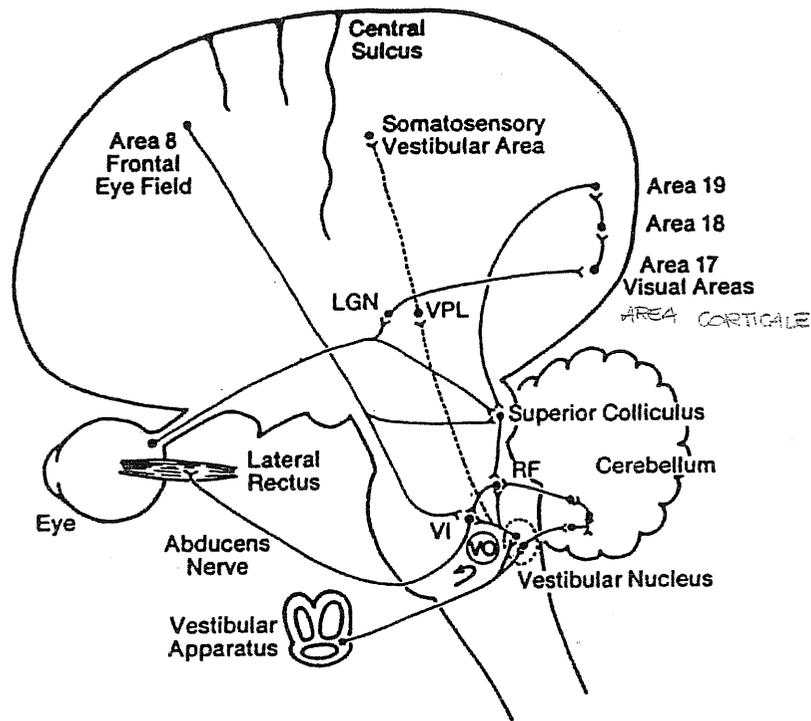
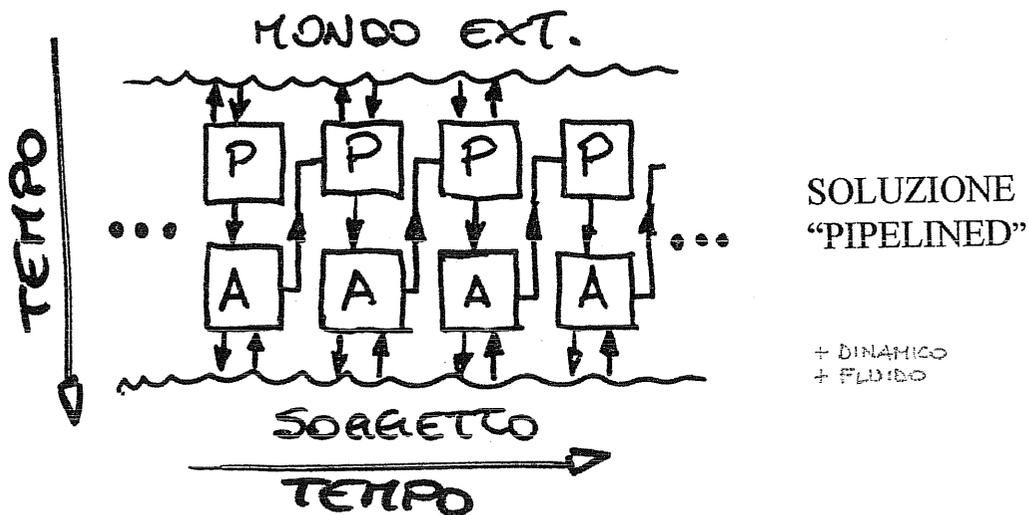
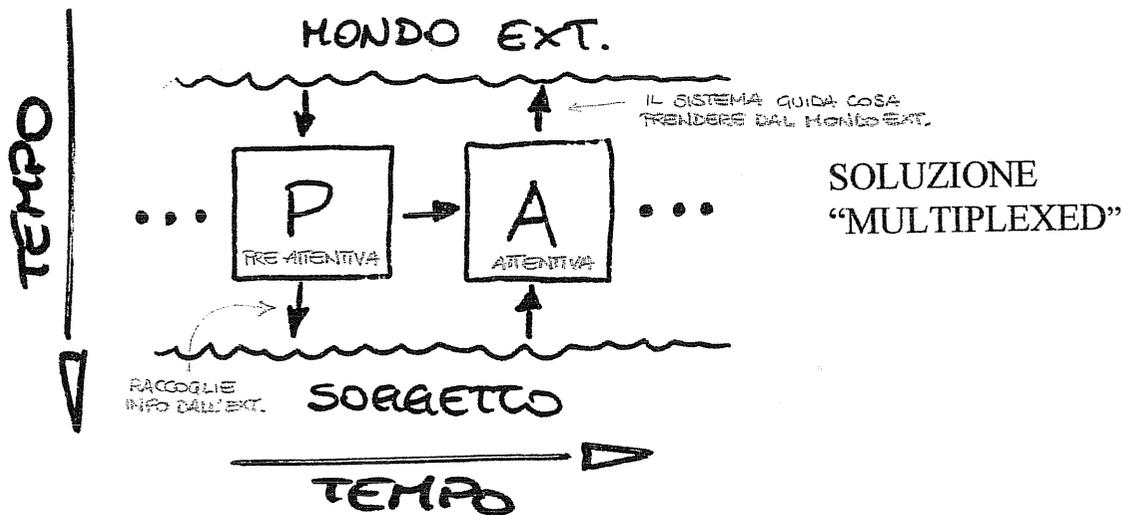


FIG. 4. A few of the circuits involved in the control of eye movements. For simplicity, the control of only one muscle, the lateral rectus, is shown. This muscle is innervated by motoneurons in the abducens nucleus, via the 6th cranial nerve (VI), the abducens nerve. The shortest vestibulo-ocular loop (VO) from the vestibular apparatus to the lateral rectus has two synapses, one in the vestibular nucleus and one in the abducens nucleus. Nuclei in the reticular formation (RF) receive converging inputs from the superior colliculus, the vestibular nucleus, and the cerebellum. The superior colliculus carries signals that are important for saccadic eye movements and receives direct inputs from retinal ganglion cells. The frontal eye fields also have a role in generating saccadic eye movements and project to brain-stem nuclei. The retina also projects to the lateral geniculate nucleus (LGN) in the thalamus, and from there to visual cortex. Projections from extrastriate cortical areas project to the oculomotor centers in the brain stem to provide information about visual targets. Motion signals in visual cortex are used to track moving targets by information that is sent to the cerebellum via pontine nuclei (not shown) in the brain stem (from Shepherd, 1987)

STRATEGIE FUNZIONALI: VISIONE ATTENTIVA & PREATTENTIVA

VISIONE ATTENTIVA → DETTAGLI → WHAT
 VISIONE PREATTENTIVA → ANALISI "GREZZA" → WHERE

CONCATENAZIONE TEMPORALE TRA PROCESSI PREATTENTIVI E ATTENTIVI



ASSIOMATICAMENTE I DUE PROCESSI SEGUONO UNA CONCATENAZIONE SEQUENZIALE, MA SECONDO UNA SOLUZIONE "PIPELINED"

SPOSTAMENTI DEL PUNTO DI FISSAZIONE

SONO RESPONSABILI DEL CONTROLLO DEL FOCUS OF ATTENTION (= LOCALIZZAZIONE SPAZIALE DELL'ATTENZIONE)

PARTE DELLA SCENA CHE
CADE NELLA FOVEA

MECCANISMI DI CONTROLLO

1. PREATTENTIVO – GUIDATO DAI DATI (= DATA-DRIVEN)

STRUTTURE SPAZIALI O EVENTI TEMPORALI NON ATTESI ATTRAGGONO IL FOCUS OF ATTENTION PER ESTRARRE MAGGIORI INFORMAZIONI

2. ATTENTIVO – GUIDATO DAL SISTEMA (= SYSTEM-DRIVEN)

IL FOCUS OF ATTENTION È INDIRIZZATO VERSO REGIONI DEGNE DI INTERESSE IN BASE A PREDIZIONI OTTENUTE DA INFORMAZIONI GIÀ ACQUISITE O IN BASE A CONOSCENZE A PRIORI (MODELLI)

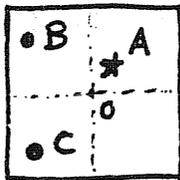
3. ABITUDINE – GUIDATO DAI DATI (=DATA-DRIVEN)

DURANTE UNA PROLUNGATA ANALISI VISIVA DI UNA SCENA STATICA SI RIDUCE L'INFLUENZA DELLE STRUTTURE DELL'IMMAGINE NEL GENERARE SPOSTAMENTI DEL FOCUS OF ATTENTION

MECCANISMI PREATTENTIVI (PREATTENTIONAL SHIFTS)

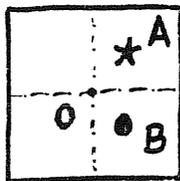
IN GENERALE SONO SOLLECITATI DA DISCONTINUITÀ (GRADIENTI) SPAZIO-TEMPORALI DEL SEGNALE VISIVO (ZONE AD ELEVATO CONTRASTO O MOVIMENTO)

CRITERI DI SCELTA DEL PUNTO DI FISSAZIONE (ALCUNI ESEMPI)



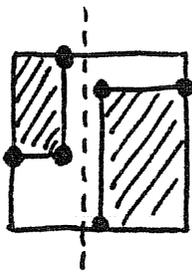
DISTANZA DAL CENTRO DEL CAMPO VISIVO

$$\overline{OA} < \overline{OB} < \overline{OC}$$



$$\overline{OA} = \overline{OB}$$

A PARITÀ DI DISTANZA SONO PREFERITI SPOSTAMENTI VERSO L'ALTO A SPOSTAMENTI VERSO IL BASSO



CARATTERISTICHE DELL'IMMAGINE

- ANGOLI, PUNTI VICINI AD INCROCI
- SE SONO PRESENTI DELLE SIMMETRIE, IL PUNTO DI FISSAZIONE TENDE AD ESSERRE SCELTO LUNGO L'ASSE DI SIMMETRIA

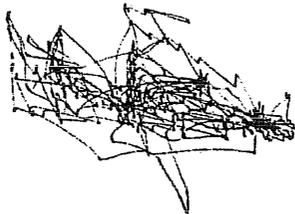
CARATTERISTICHE DELL'IMMAGINE

UNO STIMOLO IMPROVVISO NEL CAMPO PERIFERICO SPESSO PROVOCA UNO SHIFT DEL PUNTO DI FISSAZIONE VERSO TALE STIMOLO

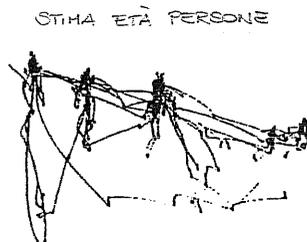
OSS.: IL CRITERIO DI SCELTA DEL PUNTO DI FISSAZIONE PUÒ ESSERE MODIFICATO DAL CONTESTO E DAL PARTICOLARE COMPITO DA ASSOLVERE



COSA FACEVA LA FAMIGLIA PRIMA DELL'ARRIVO VISITATORE.



(b)



(a)

POSIZIONE DI PERSONE ED OGGETTI NELLA STANZA.



(c)

FIG. 5. Reproduction from I.E. Repin's picture "An Unexpected Visitor" and three records of eye movements. The subject examined the reproduction with both eyes for three minutes each time. Before the recording sessions, the subject was asked to: (a) give the ages of the people; (b) surmise what the family had been doing before the arrival of the unexpected visitor; and (c) remember the position of the people and objects in the room

OSS.: È PIUTTOSTO LABILE IL CONFINE TRA CONTROLLO PREATTENTIVO E CONTROLLO ATTENTIVO (RIFLESSI PURI vs. MOVIMENTI INTENZIONALI)

COMPORAMENTI OPTOCINETICI

PERMETTONO:

1. STABILIZZAZIONE DELLE IMMAGINI CHE SI FORMANO SULLA RETINA
2. ATTIVAZIONE DELLA VISIONE COMPORTAMENTALE
(COMPORTAMENTO GUIDATO DALLA VISIONE)

RIGUARDANO:

1. ACCOMODAMENTO
LEGATO ALLA MODIFICA DELLA CURVATURA DEL CRISTALLINO.
REGOLA LA MESSA A FUOCO
2. MOBILITÀ OCULARE:

- MOVIMENTI SACCADICI
(*SACCADES*)
- MOVIMENTI LENTI DI INSEGUIMENTO
(*SMOOTH PURSUIT MOVEMENTS*)
- NISTAGMO OPTOCINETICO
(*OPTOKINETIC RESPONSE*)
- RIFLESSI VESTIBOLO-OCULARI
(*VOR*)

CONIUGATI

- MOVIMENTI DI VERGENZA
(*VERGENCE*)

DISGIUNTI

SACCADI = MOVIMENTI DI TIPO BALISTICO (SACCADI DI CORREZIONE)
SONO VELOCI E NON MODIFICABILI FINO A CHE NON SONO COMPLETATI

LATENZA ~ 200 msec
v ~ 700 °/sec
PERIODO DI FISSAZIONE ~ (150 ÷ 2000) msec

} MOVIMENTI ESTREMAMENTE RAPIDI

OSS.: IL PERIODO DI FISSAZIONE È CARATTERIZZATO DA UN TREMORE OCULARE DETTO MICRONISTAGMO FISIOLÓGICO (10 ÷ 200 Hz)

Fa in modo che l'immagine non sia mai ferma, stimolando i neuroni (che si attivano a variazioni)

MOVIMENTI LENTI D'INSEGUIMENTO

- SERVONO A MANTENERE LA FOVEA SULL'OGGETTO FISSATO, CHE PUÒ ESSERE STAZIONARIO O IN MOVIMENTO
- RISULTANO ACCURATI PER
$$V_{\text{STIMOLO}} < (60 \div 80) \text{ } ^\circ/\text{sec}$$
PER VELOCITÀ SUPERIORI SONO NECESSARIE SACCADI DI COMPENSAZIONE

NISTAGMO OPTO-CINETICO

- SONO COMPOSTI DA UNA SERIE ALTERNATA DI MOVIMENTI LENTI DI INSEGUIMENTO E DI SACCADI IN SENSO CONTRARIO
- PROLUNGANDO IL TEMPO IN CUI L'IMMAGINE STABILIZZATA (NON DEGRADATA DAL MOTION BLURRING) PUÒ ESSERE ANALIZZATA

MOVIMENTI DI VERGENZA

- CONTROLLANO IN MODO DISGIUNTO LA DIREZIONE DEI DUE ASSI OTTICI
- DURANTE I MOVIMENTI DI VERGENZA GLI OCCHI COMPIONO ROTAZIONI IN SENSO OPPOSTO. LA FUNZIONE DI QUESTI MOVIMENTI È QUELLA DI MANTENERE L'IMMAGINE DI UNA MIRA VISIVA IN PUNTI CONIUGATI DELLA RETINA QUANDO LA MIRA SI AVVICINA O SI ALLONTANA

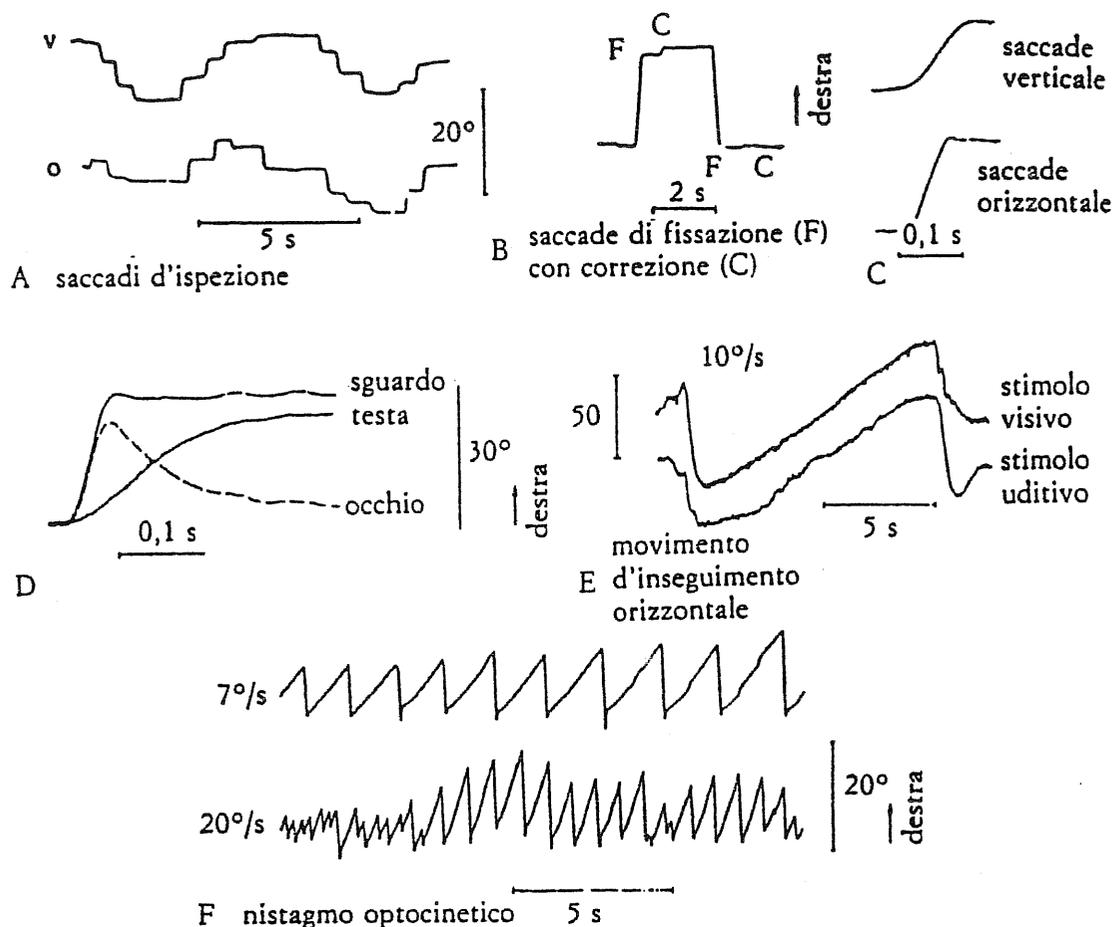


FIG. 6. Esempi di registrazioni elettrooculografiche di movimenti oculari nell'uomo. A: movimenti saccadici verticali ed orizzontali mentre il soggetto esplora l'ambiente; B: movimento saccadico eseguito per fissare un punto e movimento di correzione; C: saccade verticale ed orizzontale. Da notare che la saccade verticale è piú lenta; D: movimento coordinato degli occhi e della testa in seguito alla comparsa di uno stimolo visivo periferico (registrazione nella scimmia); E: movimento lento d'inseguimento di uno stimolo visivo e di uno stimolo uditivo che si spostano con uguale velocità; F: nistagmo optocinetico provocato da strisce bianconere che si muovono.

Fonte: adattamento da Schmidt-Thews [1985]

TABLE 1. Summary of primate gaze control systems

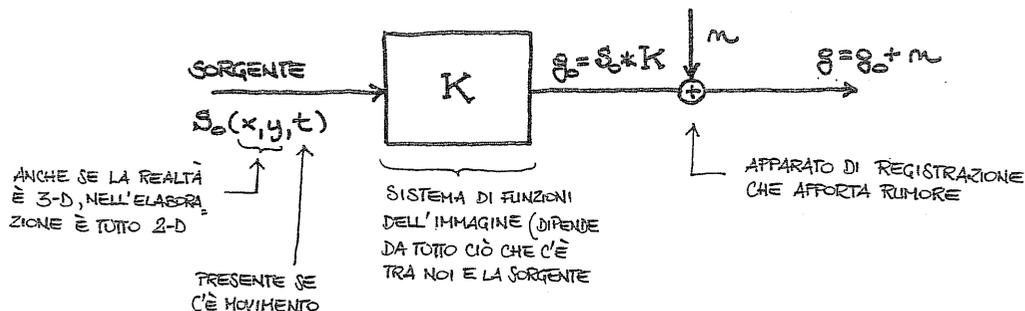
Hold gaze	Fixed target	Vestibular-ocular reflex (VOR). A system that uses knowledge of accommodation and vergence state together with head accelerations to stabilize the gaze vector.
	Moving target	Vergence. A binocular system for locking both foveas over the same three-dimensional target.
		Pursuit. A system for tracking moving objects by generating smooth-velocity control signals.
Change gaze	Saccades. High-speed, precomputed movements that rapidly change gaze over small to very large visual angles.	

From Ballard (1991). *Animate vision. Artificial Intelligence*

ESERCITAZIONI

1. ELABORAZIONE D'IMMAGINI

1.1. FORMAZIONE DI IMMAGINI/SEQUENZE VIDEO



Se consideriamo il tempo abbiamo a che fare con il moto, dunque con sequenze video.

Il fatto di proiettare la scena su 2 dimensioni comporta la perdita della profondità visiva: a tale motivo si utilizza la visione STEREOSCOFICA (con 2 occhi e relative visioni \neq) e con precisi algoritmi si deriva la profondità.

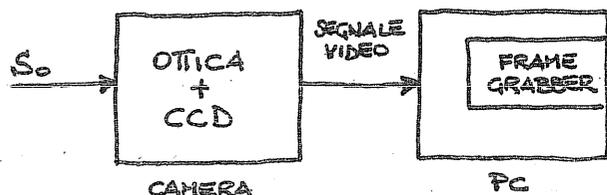
Il sistema di funzioni K è molto importante, poiché fornisce informazioni sull'ambiente e l'atmosfera in cui sono effettuate le misurazioni.

Giacché $g(x, y, t)$ è continua, dobbiamo digitalizzarla attraverso 3 operazioni:

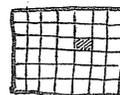
- ① CAMPIONAMENTO SPAZIALE: (da x, y) fornisce i pixel ed influisce sulla RISOLUZIONE
- ② CAMPIONAMENTO TEMPORALE: (da t) genera i frames, la cui successione darà il movimento.
 Pur partendo da immagini statiche si vede il movimento e la retina ha bisogno di un TEMPO DI INTEGRAZIONE (25 frame/sec)
- ③ QUANTIZZAZIONE DELL'AMPIEZZA DI g : quantizziamo ogni pixel con un livello di colore su 256 ($= 2^8$) possibili (usiamo 1 byte).

Ora avremo una matrice bidimensionale con indici discreti; + matrici ci daranno la sequenza che varia nel tempo.

Dal pt^a di vista hardware, q. blocchi non sono così staccati, ma saranno fatti in q. modo:



CCD = dispositivo elettronico i cui sensori rilevano la luce che lo colpisce



FRAME GRABBER = scheda del PC che digitalizza i frames che arrivano

CCD abbiamo in uscita delle matrici, che vengono nuovamente trasformate in analogico x poter passare nel cavo video. Dalle matrici, che si sono riga x riga a formare un segnale monodimensionale, si passa analogico.

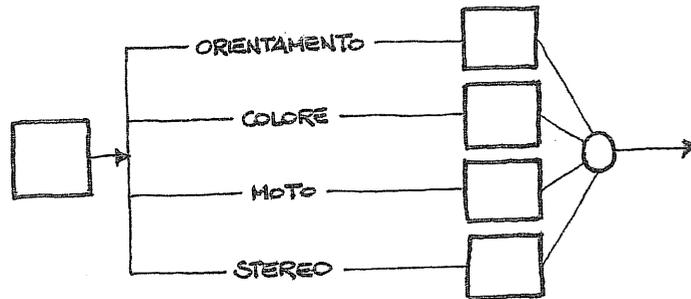
OPERAZIONI SULLE IMMAGINI

e immagini si possono fare \neq operazioni:

① RICOSTRUZIONE DELLA S_0 : molte volte da g si cerca di ricostruire S_0 , poiché dalla nitidezza vengono maggiori info. Si può decondurre ed usare dei regolatori x ridurre il rumore, ma bisogna conoscere perfettamente K .

② RICONOSCIMENTO: utilizza tecniche basate sulla probabilità

③ ESTRAZIONE DI CARATTERISTICHE: abbiamo un segnale in ingresso e cerchiamo di scomporlo in altre branche:
(BASSO LIVELLO)



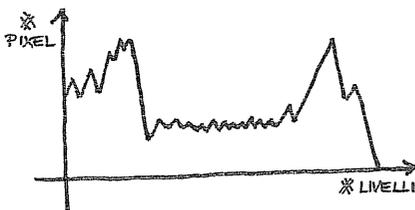
Poi si rimettono insieme x agire nella scena stessa. In pratica si fa una semplificazione in parallelo ed una ricostruzione ad alto livello.

Ciò viene fatto mediante l'uso di operatori che lavorano con matrici: ci sono 3 tipi diversi, ovvero OPERATORI

- PUNTUALI: agiscono sul singolo pixel
- LOCALI: agiscono su un insieme di pixel adiacenti
- GLOBALI: agiscono su tutti i pixel insieme

2.1. OPERATORI GLOBALI

- primo e + semplice operatore globale è l'ISTOGRAMMA (v. FIG. 1)



Sulle ascisse si mette il $\#$ dei livelli (\equiv intensità) e sulle ordinate il $\#$ di pixel che hanno tale intensità.

Si possono anche cambiare le variabili x avere altre informazioni.

Il secondo operatore è la TRASFORMATA DI FOURIER (operativamente la FFT), che fornisce le frequenze dell'immagine.

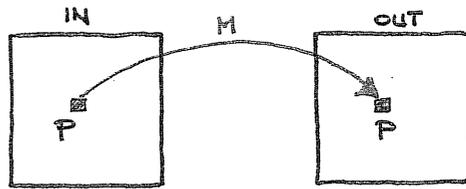
È bidimensionale, quindi fornisce frequenze spaziali in x e y ; guardando il modulo vediamo com'è distribuita l'energia (v. FIG. 2)

In (0,0) c'è la componente continua, mentre la frequenza + alta è 0.5
 (1 pixel bianco ed 1 nero $\Rightarrow T = 2 \text{ pixel} \Rightarrow f = \frac{1}{T} = \frac{1}{2} = 0.5$).

Anche la fase è molto importante, perchè fornisce la posizione delle frequenze (fase in frequenza = ritardo nello spostamento).

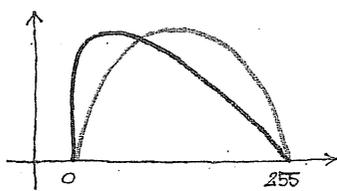
1.2.2. OPERATORI PUNTUALI

Trasformano il valore di un pixel in un altro pixel

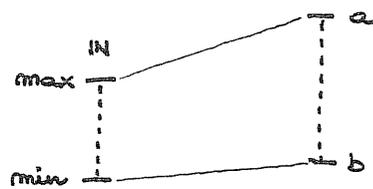


$$P_{ij}^{OUT} = M(P_{ij}^{IN})$$

① NORMALIZZAZIONE: serve a migliorare il contrasto o a riportare i valori dei pixel in $[0, 255]$ (v. Fig. 3)



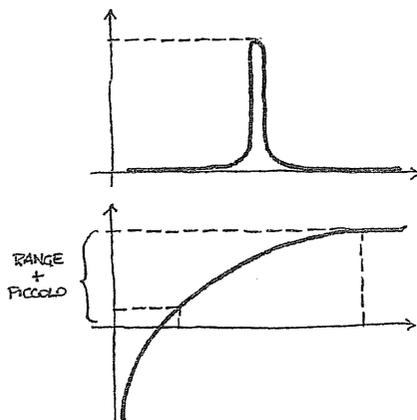
— = UTILIZZA VALORI DI INTENSITÀ TROPPO BASSI
 — = MIGLIOR CONTRASTO



[UNSIGNED CHAR. \rightarrow FLOAT]

$$P_{ij}^{OUT} = \frac{|P_{ij}^{IN} - \min|}{\max - \min} (b - a) + a$$

② COMPRESSIONE LOGARITMICA: si utilizza quando si ha una grande variazione di range nelle ampiezze (v. Fig. 4)



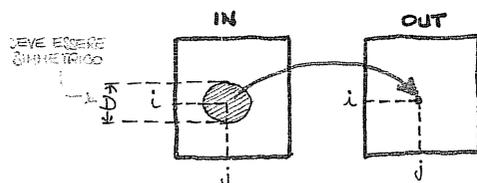
Si come il logaritmo sale molto lentamente, se prendiamo l'ampiezza e la facciamo entrare nel log otteniamo una compressione.

$$P_{ij}^{OUT} = C \log(1 + |P_{ij}^{IN}|)$$

Ciò consente di vedere cose + fini che prima non vedevamo.

1.2.3. OPERATORI LOCALI

Trasformano il valore di un insieme di pixel adiacenti



$$P_{ij}^{OUT} = M(P_{ij}^{IN} \in [])$$

Ogni pixel in uscita non dipenderà da un solo pixel in ingresso, ma da un insieme: ciò è quello che fa la convoluzione.

Detta K la superficie che va a pesare il pixel in ingresso, avremo

$$O(x, y) = \iint K(x', y') i(x-x', y-y') dx' dy'$$

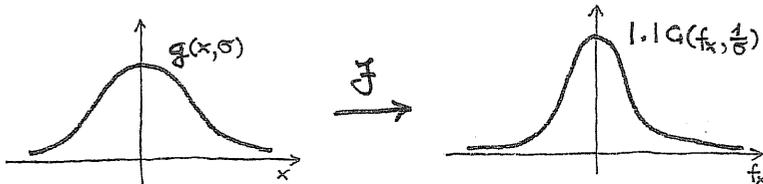
hanno discretizzare passando dagli integrali alle sommatorie:

$$O[i][j] = \sum_{-D/2}^{D/2} \sum_{-D/2}^{D/2} K[m][n] i[i+m][j+n]$$

una convoluzione significa fare una media dei pixel e, in particolare, ci interessa la convoluzione con una funzione GAUSSIANA (v. FIG.5)

$$\frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}} \quad \text{[NOTA: } \sigma_x = \sigma_y \Rightarrow \text{FORMA TONDA]}$$

gaussiana trasformata con Fourier è ancora una gaussiana, però se è grande nello spazio è + piccola in frequenza



gaussiana in pratica è un filtro PASSA-BASSO che viene utilizzato x togliere rumore dall'immagine.

secondo del filtro che usiamo cambiano le caratteristiche che riusciamo a estrarre.

Elaborazioni di immagini

Operatori globali

Istogramma dell'intensità

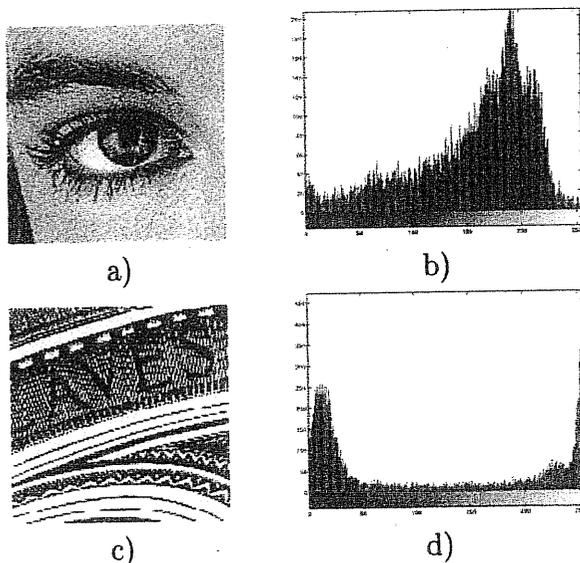


Figure 1: Le immagini a) e c) sono 128×128 pixel a 256 livelli di grigio; b) e d) sono i relativi istogrammi.

Trasformata di Fourier

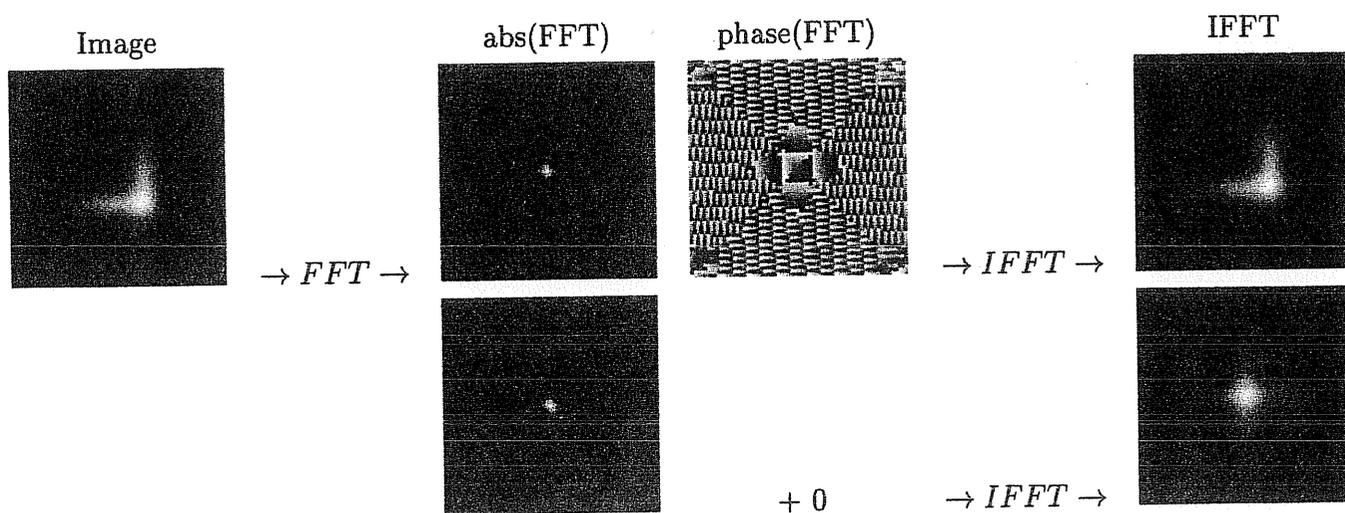


Figure 2: Le immagini sono 65×65 pixel a 256 livelli di grigio.

IN D, FIDOLA ABBIAMO
PERDUTO L'INFO DI SINTESI
(NON ABBIAMO ANTI,
TRASFORMATO LA FASE),
MA L'ENERGIA È LA
STESSA.

Operatori puntuali

Contrast stretching

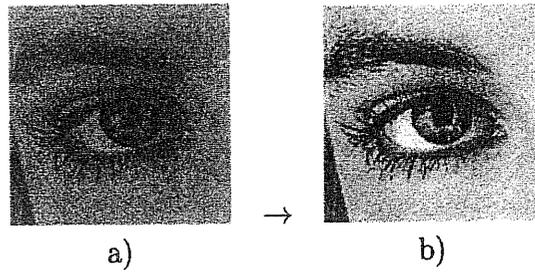


Figure 3: L'immagine b) è il risultato ottenuto dalla normalizzazione del contrasto dell'immagine a).

Compressione logaritmica

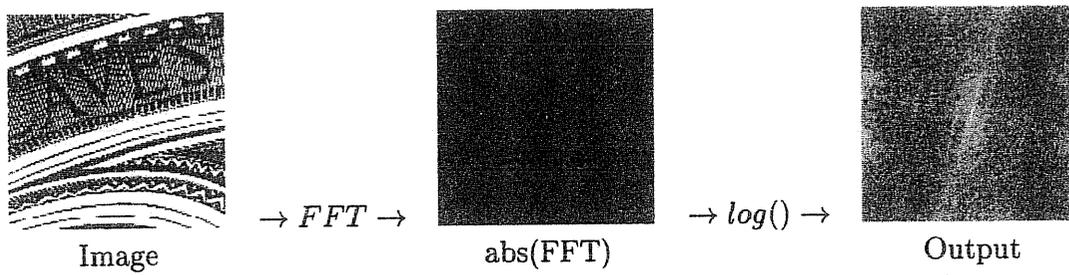


Figure 4: In Output si possono cogliere particolari non visibili in $\text{abs}(FFT)$, che utilizza una scala lineare per l'ampiezza.

Operatori locali

Convoluzione con funzione Gaussiana

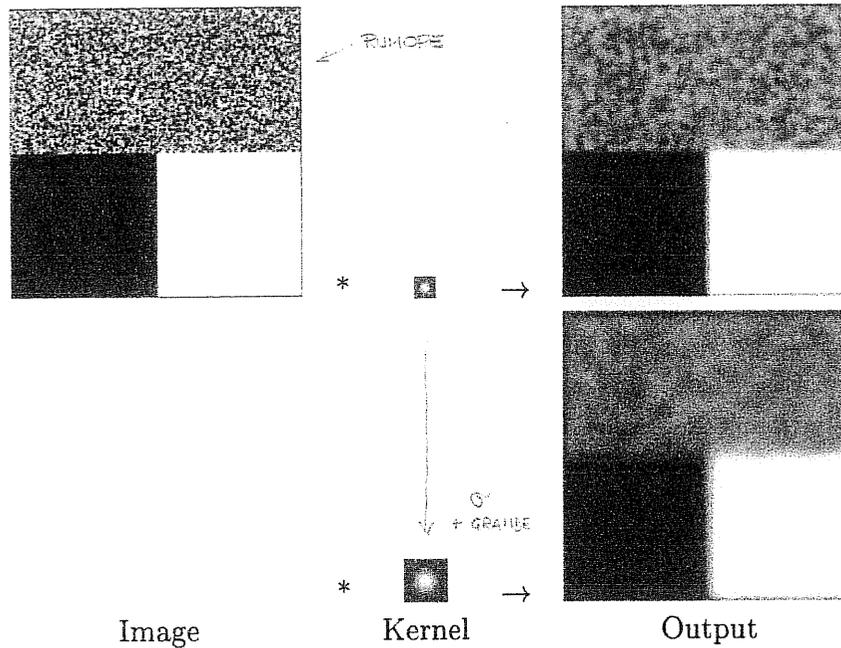


Figure 5: Image e Output sono 128×128 pixel, Kernel è 7×7 e 15×15 pixel e sono utilizzati 256 livelli di grigio.

Convoluzione con derivata di funzione Gaussiana

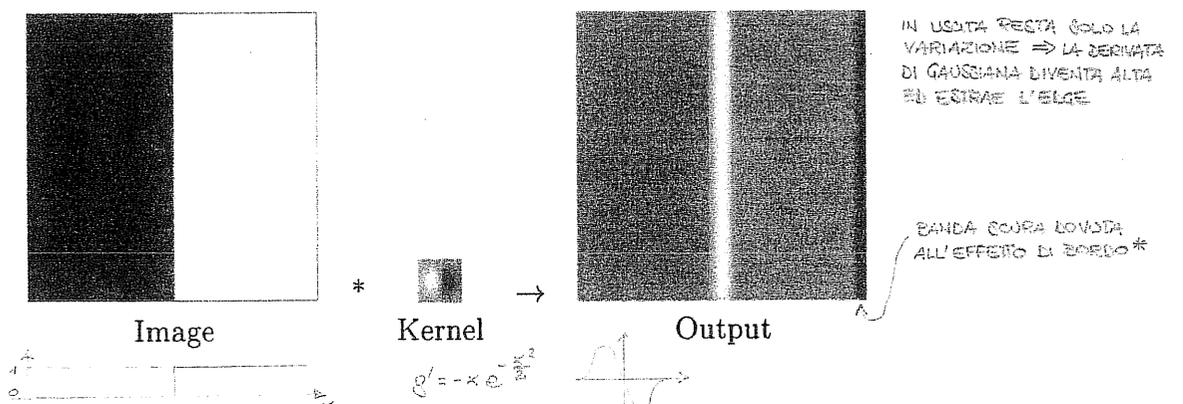


Figure 6: Image e Output sono 128×128 pixel, Kernel è 15×15 pixel e sono utilizzati 256 livelli di grigio.

*: Quando facciamo la convoluzione abbiamo degli effetti di bordo, poiché a ottenere tutta l'immagine dobbiamo amplificarla in ogni bordo di $1/2$, altrimenti otteniamo solo una figura interna.
 Il bordo aggiunto viene riempito di $0 \Rightarrow$ NERO



1.3. INTRODUZIONE ALL'ESERCITAZIONE 1

• RICHIAMI DI C

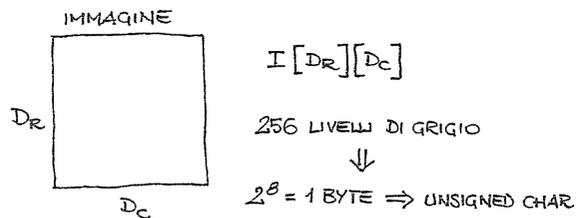
```
* define D 128
```

```
main ()
{
    float M[D][D];
    float **M;           [PUNTIPORE DEDICATA ALLA MEMORIA]
    M = malloc (sizeof (*float) * D);
    for (i=0; i<D; i++)
        M[i] = malloc ((float) * D);
    :
}
```

LETTURA DA FILE

```
* define D 128
```

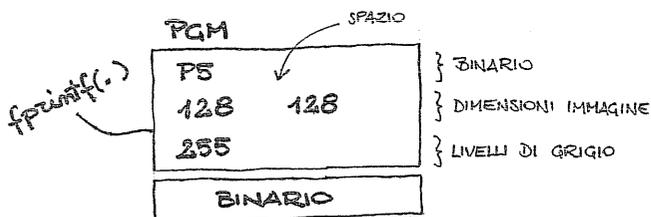
```
main ()
{
    unsigned char U[D][D];
    FILE *fp;
    fp = fopen ("eye.128", "rb");
    if (fp = NULL)           [LETTURA BINARIA]
    {
        printf (...);
        return 0;
    }
    fread (U, sizeof (unsigned char), D*D, fp);
    fclose (fp)
    :
}
```



SCRITTURA SU FILE

```
:
fwrite (V, sizeof (unsigned char), D*D, fp)
:
```

Il formato + semplice è il PGM. Prima di scrivere l'immagine bisogna creare questo HEADER:

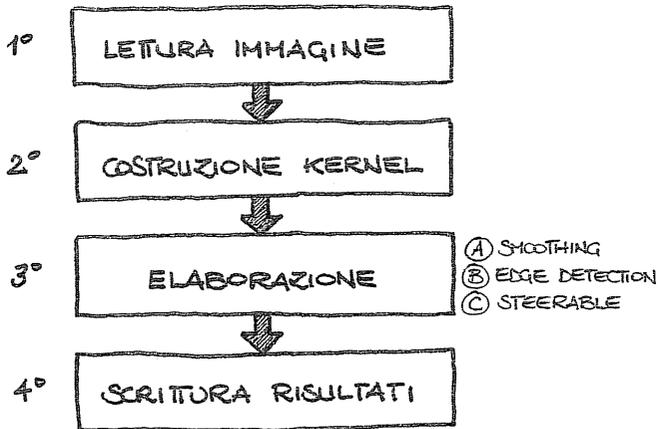


Questo file va letto con `fscanf(...)` x conoscere i parametri dello header, e successivamente si fa la `malloc`

```
fscanf (%s, %d, ...);
```

5TO ESERCITAZIONE

programmata in C potrebbe essere composta da questi blocchi:



Il primo blocco è sicuramente quello che effettua LETTURA e SCRITTURA.

Il 1° e il 4° sono gli unici blocchi in cui lavoriamo con UNSIGNED CHAR, poiché l'elaborazione avviene in FLOAT.

fread (...)

↓
(cast) unsigned char → FLOAT

$$\begin{bmatrix} 0 \div 255 & I_{IN} \\ MIN \div MAX & I_{OUT} (FLOAT) \end{bmatrix}$$

↓
normalizzazione

↓
fwrite (...)

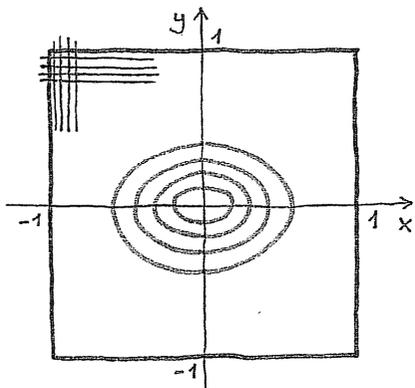
SMOOTHING (A)

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}}$$

GAUSSIANA BIDIMENSIONALE (FILTRO UTILIZZATO)

Usiamo il campionamento \times generale il nostro filtro digitale.

Supponiamo, nel piano cartesiano normale, di avere un intervallo $[-1,+1]$ e disegniamo la gaussiana: da q. funzione continua ne vogliamo ottenere una discreta.



$$step = \frac{2}{N-1}$$

FASCE DISPARI CON $\frac{N-1}{2} < 0$, $\frac{N-1}{2} > 0$ E LO 0 CENTRALE

$$x = -1$$

$$y = +1$$

2 CICLI FOR ANNIDATI

```

for (x=-1; x <= 1; x=x+step)
  for (y=1; y >= -1; y=y-step)
  
```

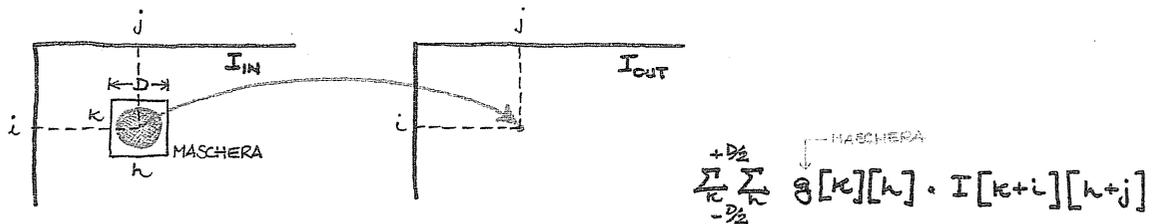
Verrà generata una matrice in cui ogni elemento ha il valore della gaussiana campionata. Se prendiamo $\sigma_x = \sigma_y$ avremo una gaussiana tonda, altrimenti con $\sigma_x \neq \sigma_y$ sarà allungata.

Possiamo anche determinare un angolo θ con cui ruotare la gaussiana attraverso le formule

$$\begin{cases} x_\theta = x \cos \theta - y \sin \theta \\ y_\theta = x \sin \theta + y \cos \theta \end{cases}$$

per poi ricampionarla.

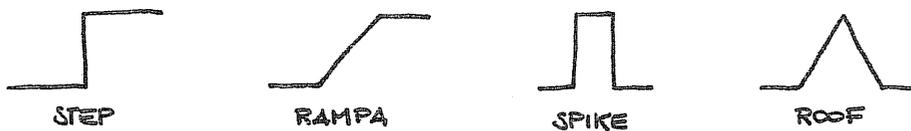
③ CONVOLUZIONE



Salvo errori nel risultato, facendo la convoluzione tra l'impulso e l'immagine d'uscita otteniamo il kernel.

EDGE DETECTION (B)

Esistono 4 tipi ideali di EDGE:



In realtà fisica q. tipi di edge saranno corrotti dal rumore.

Per rilevare gli edge, normalmente, si procede in 3 passi:

- (1.) FILTRAGGIO: toglie il rumore [⇒ CONVOLUZIONE CON GAUSSIANA]
- (2.) DIFFERENZIAZIONE: evidenzia le variazioni di intensità [⇒ DERIVATA]
- (3.) RILEVAMENTO: trova gli edge [⇒ PUNTI A MAX DERIVATA]

Vediamoli in dettaglio:

(1.)
$$g(x,y) = e^{-\frac{x^2}{2} - \frac{y^2}{2}}$$
 [PASSA BASSO: TOGLIE LE ALTE FREQUENZE]

(2.)
$$I_c = I * g \Rightarrow \frac{\partial}{\partial x} I_c = \frac{\partial}{\partial x} (I * g) = I * g_x$$

Quindi

$$g_x = -x e^{-\frac{x^2+y^2}{2}} \quad g_y = -y e^{-\frac{x^2+y^2}{2}}$$

Essendo in 2-D, dobbiamo derivare rispetto a x e y: l'operatore che lo fa è il GRADIENTE

gradiente
$$I(x,y) = (I_x(x,y), I_y(x,y))$$

$$M(x,y) = \sqrt{I_x^2 + I_y^2}$$

↑
COMPRESA

- (3.) Vogliamo trovare dove è max il modulo del gradiente: possiamo normalizzarlo da 0 a 100:

$$N(x,y) = \frac{M(x,y)}{\max_{x,y} M(x,y)} \cdot 100$$

e definire

$$E(x,y) = \begin{cases} 1 & |N(x,y)| > T \\ 0 & \text{altrimenti} \end{cases}$$

EDGE MAP

La EDGE MAP sarà sempre un'immagine, che ci dà, appunto, la mappa di dove sono gli edge.

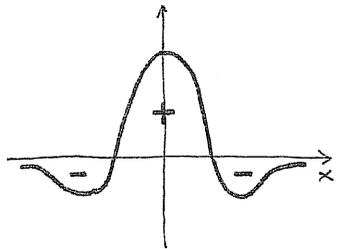
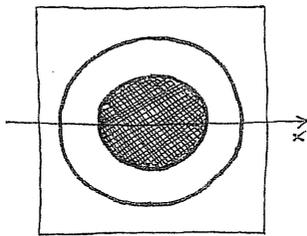
Ciò che abbiamo appena fatto è con le DERIVATE DI GAUSSIANA: un altro metodo usa invece il LAPLACIANO DI GAUSSIANA. Esso, invece di n letture il max della derivata 1°, rileva i passaggi x lo 0 della derivata 2°: ciò dovrebbe funzionare meglio su variazioni + piatte.

Si ha il vantaggio che non bisogna scegliere una soglia e priori, ed è ciò che avviene nei primi stadi del sistema visivo (cellule corticali).

Utilizziamo sempre la gaussiana x togliere il rumore:

$$L_G = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = k \left(2 - \frac{x^2 + y^2}{\sigma^2} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

Visivamente:

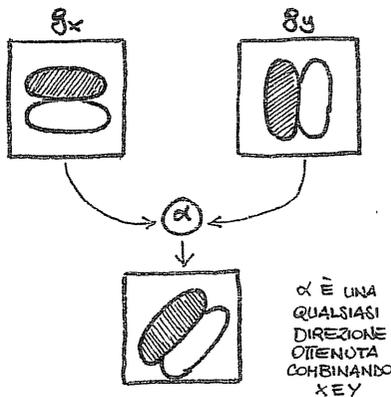


- Costruire L_G
- Convoluzione $I * L_G$
- Rilevare i passaggi x lo zero delle EDGE MAP (x riga e colonna)

4 tipi di passaggio

- $\{-, 0, +\}$
- $\{+, 0, -\}$
- $\{+, -\}$
- $\{-, +\}$

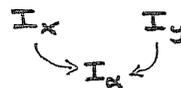
FILTRI STEERABLE (C)



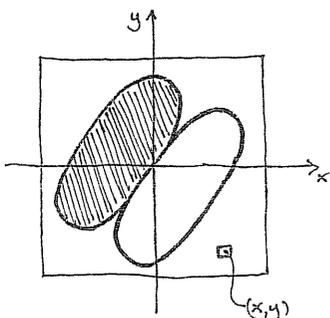
Possiamo derivare rispetto ad α :

$$g_\alpha = \cos \alpha g_x + \sin \alpha g_y$$

ed ottenere la convoluzione di I_α da I_x e I_y senza pb:



CHIAMI DI MATLAB



Il campionamento è fatto in automatico:

$$x = -1 : \text{step} : 1 ;$$

mentre la funzione MESHGRID riporta la funzione su 2 piani:

$$[X, Y] = \text{meshgrid}(x, y) \quad \left[\begin{array}{l} \text{GENERA SPAZIO} \\ \text{CAMPIONATO IN} \\ \text{FORMA MATRICIALE} \end{array} \right]$$

Poi possiamo richiamare direttamente X e Y nella generazione della gaussiana:

$$Z = \exp(-(X.^2 - Y.^2)/2\sigma)$$

Se la vogliamo visualizzare come immagine:

figure, imagesc(Z), colormap gray

mentre se vogliamo visualizzarla in 3-D sarà:

figure, mesh(Z)

Per le funzioni:

function $[o1, o2] = nome(i1, i2, i3)$ [DEFINIZIONE]

VALORI USCITA
PARAMETRI INGRESSO

$[a1, a2] = nome(b1, b2, b3)$ [CHIAMATA]

Per leggere un file binario:

IDENTIFICATORE DEL FILE → $fid = fopen(nome_file, "r");$

MATRICE → $Y = fread(fid, [dr, dc], 'uchar');$

$fclose(fid)$

$Y = Y';$ [BISOGNA FARE LA TRASPOSTA, POICHÈ LEGGE X COLONNE E NON X RIGHE]

Per scrivere su un file:

$a = a';$

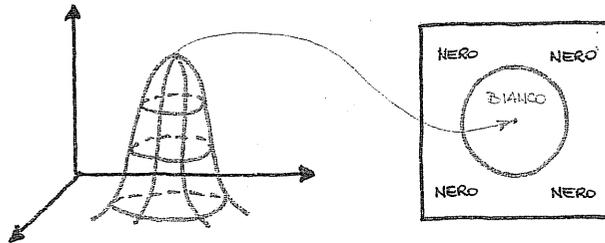
$fwrite(fid, a, ' ');$

TIPO
 ↑ { UNSIGNED CHAR
 CHAR
 FLOAT
 ...

2. GENERAZIONE DI STIMOLI SINTETICI

Gli STIMOLI SINTETICI sono immagini/sequenze create a partire da una descrizione analitica. Queste immagini, già note, servono a verificare la correttezza degli algoritmi.

Abbiamo visto che la struttura dati di un'immagine è una matrice indicizzata x raggiungere ogni pixel dell'immagine. Può anche essere vista come una funzione di 2 indici:



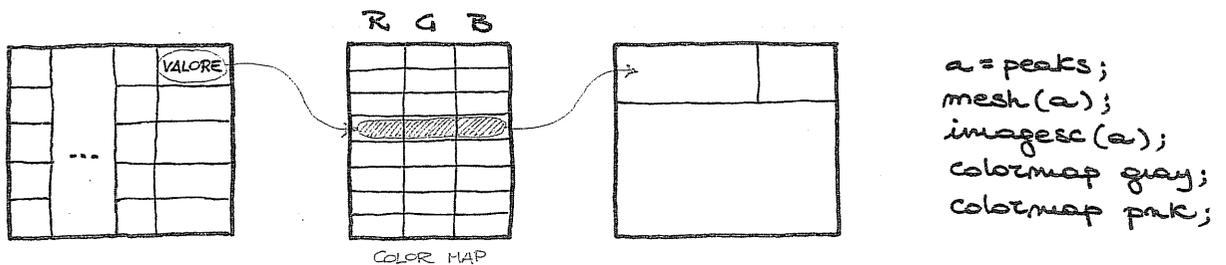
Esistono 2 tipi di immagini:

- ① INDICIZZATE
- ② RGB

Analizziamo + in dettaglio

① INDICIZZATE

Un valore dell'immagine (utilizzando la COLOR MAP e la LOOK-UP TABLE) va ad indicizzare un x di colonne e di righe e poi genera il colore



② RGB

Non c'è + la COLOR MAP: forniamo proprio il colore desiderato attraverso una matrice tridimensionale (R,G,B)

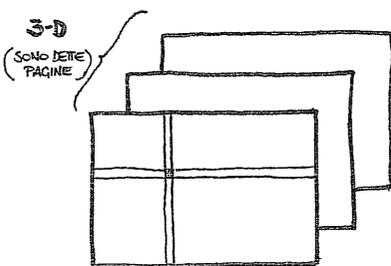


image → solo una RGB con $R \in [0,1]$

imagesc → scala l'immagine

Possiamo avere tipi diversi di stimoli:

① STIMOLI STATICI 2-D (IMMAGINI)

• EDGE



• GRATICCI



↑
USATI NELLE
STIMOLAZIONI
VISIVE

$$\sin(2\pi f_x x)$$

$$a = -1 : 0,01 : 1;$$

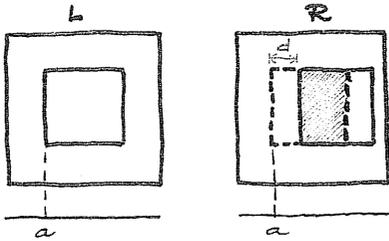
$$[X, Y] = \text{meshgrid}(a);$$

PARLANDO DI SIN SI PARLA DI FREQUENZA = $1/T$, DOVE IL PERIODO È IN PIXEL ⇒ MAX FREQ DIGITALE = $1/2$ CIÒ È 1 PIXEL BIANCO E 1 PIXEL NERO

- NOISE : si usa la funzione $RAND()$, poiché lo spettro di rumore è piatto e si stimola con tutte le frequenze.

STIMOLI STATICI (2 IMMAGINI-STEREO)

Abbiamo proiettato la terza dimensione su un piano : x ricavarla dobbiamo spostarci leggermente. Dalla differenza tra la 1^a proiezione e la 2^a ricaviamo le info circa la profondità.



$$I_L(x) = I_R(x+d) \quad \text{dove } d = \text{DISPARITÀ}$$

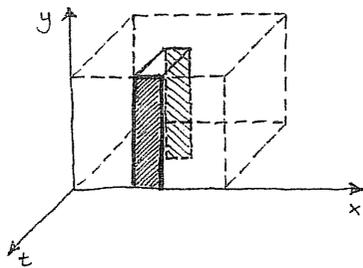
Quindi generiamo un'immagine con rumore e ci mettiamo l'immagine spostata.

In MATLAB:

```
L = rand(101);
R = L;
R(20:70; 20:70) = L(20:70; 25:75);
anaglifo
A(i, i, 1) = L
A(i, i, 3) = R
```

[ANAGLIFO: IMMAGINE RGB CHE CONTIENE PARTE SIN E DX, SERVE PER VEDERE LA PROFONDITÀ]

STIMOLI DINAMICI 3-D (SEQUENZE VIDEO)



$L(x, y, t)$ = variazione di luminosità che varia nel tempo



v = pixel/frame