

INDICE

SOMMARIO	3
1. INTRODUZIONE	4
1.1 La retina dell'uomo e della mosca.....	7
2. IL SENSORE RETINICO	18
2.1 La storia del sensore	18
2.2 Caratteristiche del sensore	19
2.3 Le immagini log-polari	22
2.4 Il peso dell'immagine log-polare.....	28
2.5 Calcolo del peso dell'immagine log-polare.....	32
3. GLI INVARIANTI GEOMETRICI NELLE IMMAGINI RETINICHE.....	37
3.1 Introduzione.....	37
3.2 Gli invarianti nelle trasformazioni geometriche	41
3.3 Invarianze geometriche in immagini prospettiche.....	42
3.4 Invarianze geometriche in immagini retiniche	44
3.5 Invarianze geometriche in immagini omnidirezionali retiniche.....	53
3.6 Prospetto riassuntivo delle trasformazioni log polari	56
4. EDGE DETECTION SU IMMAGINI LOG-POLARI	69
4.1 Introduzione	69
4.2 Space variant edge detection	71
5.SPECCHIO DEDICATO A TELECAMERA RETINICA	77
5.1 Introduzione	77
5.2 Algoritmo per la creazione del profilo dello specchio.....	77
5.3 Caratteristiche progettuali.....	78
5.4 Specchio con suddivisione degli insiemi nel piano lineare	83
5.5 Specchio con suddivisione degli insiemi nel piano esponenziale.....	86
5.6 Specchio per sensore retinico a errore relativo costante.....	89

CONCLUSIONI	95
APPENDICE	97
SOFTWARE DI ACQUISIZIONE	97
INDICE DELLE FIGURE.....	99
INDICE DELLE TABELLE	103
BIBLIOGRAFIA	104
RINGRAZIAMENTI.....	107

SOMMARIO

Nella presente tesi viene studiato e analizzato inizialmente il funzionamento e la struttura del sensore retinico, che è parte integrante della telecamera costituente nel complesso il sensore di visione omnidirezionale retinico. Vengono analizzate le immagini log polari, acquisite tramite la telecamera retinica e vengono ricavati alcuni invarianti geometrici presenti in queste particolari immagini. Sarà inoltre implementato un nuovo tipo di edge detection dedicato appositamente alle immagini retiniche.

Infine viene progettato uno specchio omnidirezionale con l'obbiettivo di permettere al sensore di visione omnidirezionale retinico di localizzare gli oggetti visualizzati nell'ambiente esterno, commettendo nella misura delle distanze un errore relativo (nel caso peggiore) costante.

1. INTRODUZIONE

Gli apparati di visione sono senza dubbio i sensori più utilizzati da robot mobili che svolgono compiti di navigazione in ambienti strutturati e non strutturati, in quanto permettono l'acquisizione di una grande quantità di informazione essenziale nello sviluppo di molteplici applicazioni. Può essere comunque utile affiancare agli apparati di visione altri tipi di sensori (come gli ultrasuoni, gli infrarossi) che potenzino le risorse disponibili dal robot in modo da ottimizzare i compiti svolti.

Il sistema di visione risulta molto più efficace se riesce a dare una descrizione più vasta della scena in esame. Ecco perché, in questi ultimi anni, si sono sviluppati i sistemi di visione omnidirezionale. Questi sono basati su un sensore catadiottrico, composto da una telecamera rivolta verso l'alto che acquisisce l'immagine riflessa da uno specchio convesso disposto sopra di essa. Il termine catadiottrico deriva da "diottrica", la disciplina che studia gli elementi rifrangenti (come le lenti) e la "catottrica", la disciplina delle superfici riflettenti. Normalmente il sensore interno alla telecamera che viene usata, è un sensore tradizionale, ovvero un sensore i cui fotorecettori sono organizzati in matrice e possiedono tutti la stessa dimensione.

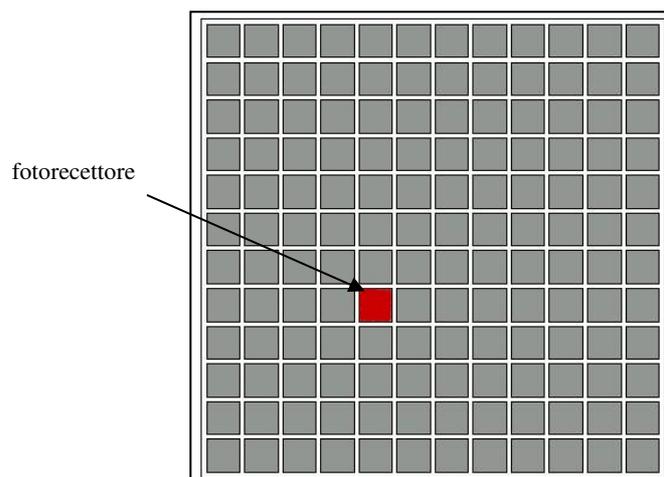


Figura 1. Struttura del sensore tradizionale. I fotorecettori hanno tutti la stessa dimensione.

Nel sensore di visione omnidirezionale trattato in questa tesi, la telecamera utilizzata sfrutta un nuovo tipo di sensore, il sensore retinico, strutturalmente diverso

dai sensori tradizionali. Viene costituito quindi nel complesso, un sensore di visione omnidirezionale retinico, che costituisce l'obiettivo di questa tesi.

Il sensore retinico è un sensore di nuova generazione la cui struttura si ispira al funzionamento della retina umana. Infatti nell'occhio umano, al centro della retina si ha la maggiore acuità visiva, in una zona detta fovea, mentre allontanandosi dalla fovea l'acutezza visiva decresce. Imitando questo aspetto biologico il sensore presenta una densità molto alta di fotorecettori al centro permettendo un'ottima risoluzione nel centro dell'immagine, e una densità di fotorecettori decrescente dal centro alla periferia del sensore, dove la risoluzione dell'immagine diviene minima.

Il beneficio più importante introdotto dall'uso della telecamera omnidirezionale è, come detto, l'aumento del campo visivo, che associato alla notevole riduzione del peso dell'immagine dovuto al sensore retinico, può comportare un'efficienza computazionale notevole nell'ambito di numerose applicazioni. Le caratteristiche intrinseche e gli invarianti nelle immagini log polari possono inoltre essere sfruttate per lo sviluppo di algoritmi di vario genere nell'ambito della robotica ma anche in altre svariate applicazioni come la sorveglianza, la teleconferenza, lo scambio di dati nella rete.

Il sensore di visione omnidirezionale retinico

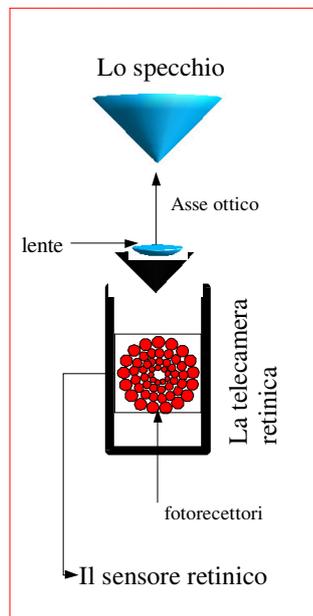


Figura 2. Struttura della telecamera omnidirezionale retinica

La prima fase del lavoro che ha portato alla stesura della tesi, è stata l'analisi della struttura e del funzionamento in dettaglio del sensore retinico. Successivamente, una

volta apprese queste nozioni, è stato implementato un software per gestire le immagini acquisite con la telecamera retinica, le cosiddette immagini log polari.

Sono stati ricavati analiticamente degli invarianti geometrici nelle immagini log polari come ad esempio gli invarianti in seguito a rotazioni e a scalature dell'immagine. Queste proprietà poi, sono state verificate su immagini acquisite (sia omnidirezionali retiniche che retiniche) sfruttando dei pattern prestabiliti. Le immagini log polari possono essere rimappate nelle corrispondenti immagini cartesiane dove si evidenzia la risoluzione spazio variante indotta dal sensore retinico. Per quest'ultimo tipo di immagini è stato studiato e implementato un particolare tipo di edge detection, basato sull'algoritmo di Canny [15]. Nell'edge detection implementato viene considerata la variazione di risoluzione nell'immagine, che può comportare un fallimento nella ricerca degli edge, ecco quindi che l'edge detection viene adattato alla zona di immagine trattata, a minore o maggiore risoluzione.

Come ultimo passo di questo lavoro di tesi, è stato progettato uno specchio omnidirezionale da abbinare alla telecamera retinica. L'algoritmo per la creazione dello specchio è lo stesso usato in [3] e in [4]. L'obiettivo principale è stato quello di progettare lo specchio in modo da localizzare gli oggetti visualizzati nell'ambiente circostante tramite l'accesso ad una tabella di corrispondenze: pixel nell'immagine – distanze sul piano, e soprattutto di garantire un errore relativo costante (nel caso peggiore) nella misura delle distanze degli oggetti visualizzati.

La struttura della tesi è articolata quindi in questo modo: il primo Capitolo analizza dettagliatamente la struttura del sensore e le immagini che vengono acquisite dalla telecamera retinica. Il secondo Capitolo evidenzia gli invarianti geometrici e le proprietà intrinseche delle immagini log polari, sia retiniche che omnidirezionali retiniche. Nel terzo Capitolo si tratta del particolare tipo di edge detection adattato alle immagini log polari. Nel quarto Capitolo viene descritta la progettazione dello specchio omnidirezionale dedicato alla telecamera retinica.

1.1 LA RETINA DELL'UOMO E DELLA MOSCA

INTRODUZIONE

NOTE SULL'OCCHIO¹

La retina umana è ritenuta la parte più importante del complesso organo visivo umano. Questo consta essenzialmente di 3 formazioni: l'occhio o globo oculare, il nervo ottico e gli annessi oculari. Questi ultimi sono: pareti della cavità orbitaria, muscoli estrinseci dell'occhio, fascia del bulbo, corpo adiposo dell'orbita, sopracciglia e palpebre, congiuntiva, apparato lacrimale.

La funzione dell'occhio è quella di captare le radiazioni luminose provenienti dal mondo esterno e di trasformarle in impulsi nervosi che saranno utilizzati a scopi percettivi (visione) e riflessi.

Il bulbo oculare è costituito da tre membrane oculari sovrapposte e da un nucleo contenuto in esse. La membrana più superficiale è la tonaca fibrosa, che in un piccolo segmento anteriore, a contorno circolare, è trasparente e prende il nome di cornea; nella rimanente parte è opaca e si chiama sclera o sclerotica; le due zone sono separate da un solco anulare superficiale: il solco della sclera, sito a livello del sesto anteriore dell'occhio. All'interno della tonaca fibrosa è applicata la tonaca vascolare, ricca di vasi sanguigni e di pigmento scuro, e contenente elementi muscolari lisci. Questa, nello strato più esterno, il più esteso, ha una disposizione semplice e regolare e prende il nome di coroide; lo strato più interno invece ha una struttura più complicata e forma, unendosi alla retina, il corpo ciliare e, più avanti, l'iride.

L'iride è un diaframma con un foro centrale circolare detto pupilla; questa è posta posteriormente ed a una certa distanza dalla cornea e la sua funzione è quella di regolare la quantità di luce che entra nell'occhio. La più profonda delle tre membrane, la retina, è di natura nervosa e da essa origina il nervo ottico. La sua faccia posteriore, la più estesa, è disposta regolarmente sulla faccia profonda della coroide e prende il nome di parte ottica della retina, essendo responsabile della funzione visiva. La parte anteriore, invece, non contiene elementi di natura nervosa, non partecipa alla formazione del nervo ottico e prende il nome di parte cieca della

¹ Tratto da [1], e da [2].

retina. Aderisce alla faccia profonda del corpo ciliare e dell'iride e concorre alla formazione di questi organi, distinguendosi in parte ciliare e parte iridea. Le due parti sono separate da un avvallamento, noto come ora serrata. La Figura 3 illustra in dettaglio la struttura delle membrane oculari e del nucleo del bulbo oculare.

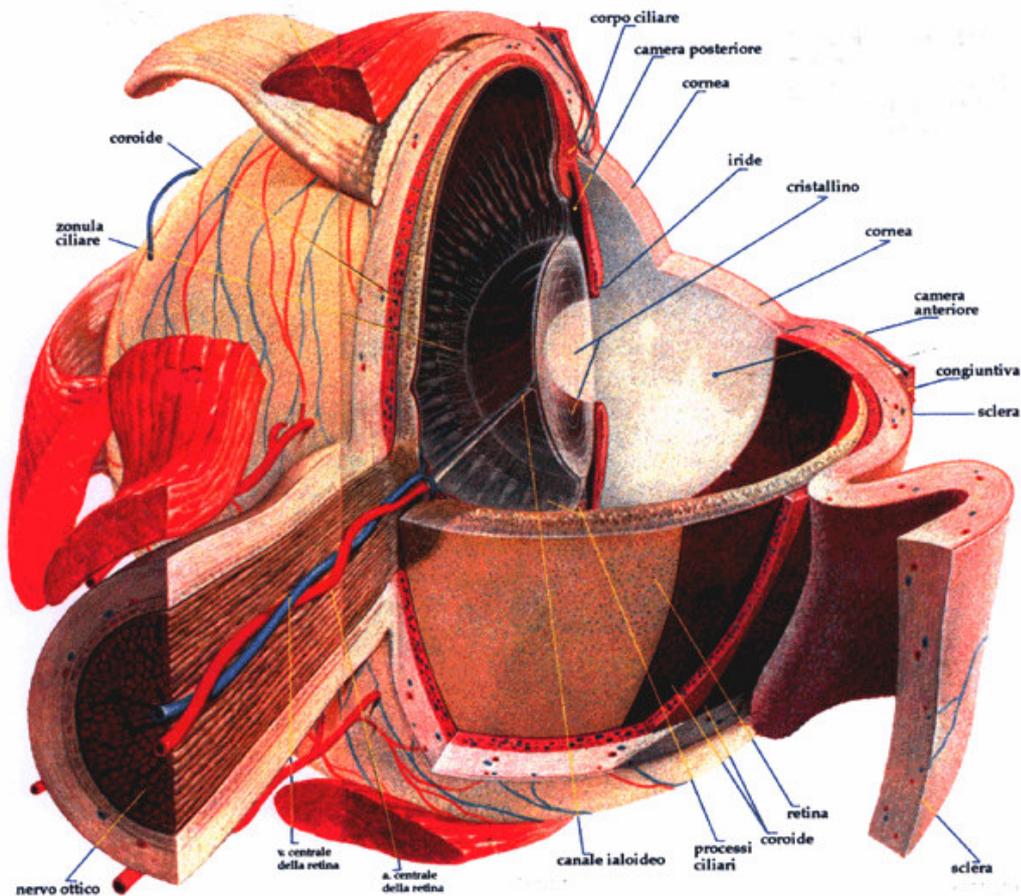


Figura 3. Spaccato dell'occhio. Immagine tratta da [1]

All'interno del globo oculare, fra la cornea e l'iride, si trova uno spazio riempito di un liquido incolore, trasparente e fluido come l'acqua, che prende il nome di umor acqueo, esso ha lo scopo, mediante una lieve pressione sulla parete interna della cornea, di mantenerne la forma del cristallino. Il cristallino si trova posteriormente alla camera anteriore, dietro l'iride e la pupilla, è un corpo trasparente e biconvesso connesso, per mezzo di un particolare apparato di sospensione, (la zonula ciliare) con la regione ciliare. Il cristallino funziona come una lente biconvessa con curvature

differenti ed è costituito da diversi strati sovrapposti a guisa di cipolla. E' connesso all'interno dell'occhio da fibre muscolari che gli permettono di variarne la curvatura di modo da far cadere costantemente sulla retina il piano focale dell'immagine (accomodamento del cristallino). Tra l'iride, il corpo ciliare ed il cristallino è compreso uno spazio anulare, la camera posteriore, comunicante con la camera anteriore lungo il margine della pupilla, ripiena anch'essa dello stesso umor acqueo. Il rimanente, ampio spazio del bulbo oculare, limitato dal cristallino col suo apparato sospensore e dalla retina, è occupato dal corpo vitreo, formato da sostanza di consistenza gelatinosa, incolore e trasparente.

Cornea, umor acqueo, cristallino e corpo vitreo costituiscono i mezzi diottrici (o sistema o apparato diottrico) dell'occhio, che può essere considerato come una lente convergente dotata di un notevole potere rifrattivo: circa 60 diottrie. Il centro ottico viene a trovarsi a circa 7 mm dalla cornea, mentre il punto dove convergono i raggi luminosi paralleli (fovea), si trova a 24 mm dalla cornea.

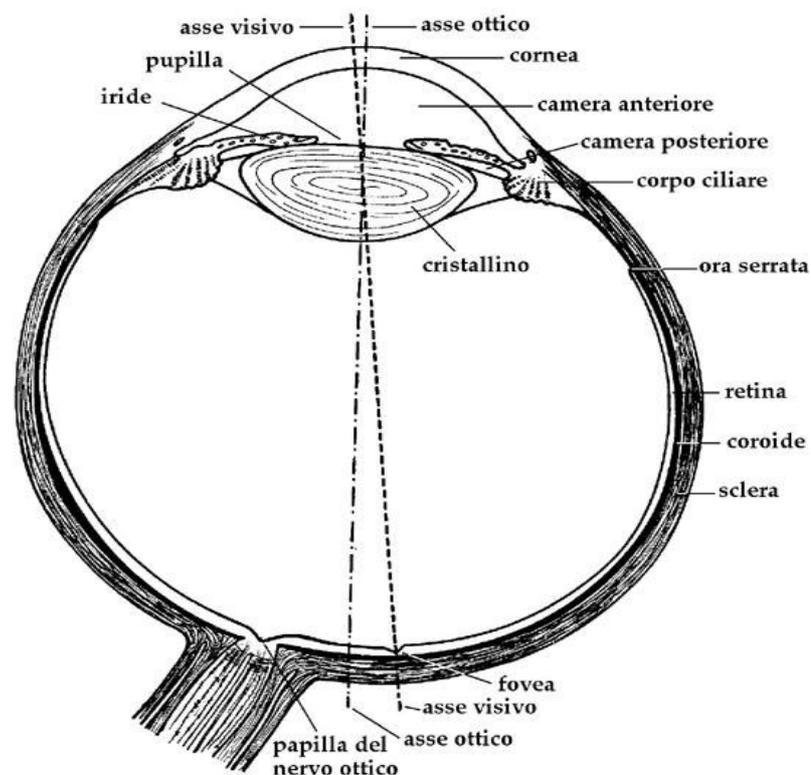


Figura 4. Sezione dell'occhio. Immagine tratta da [2]

LA RETINA

STRUTTURA

La retina è la più interna delle tre membrane che formano le pareti del globo oculare e si estende dal punto di entrata del nervo ottico fino al margine pupillare dell'iride. Grazie alla sua struttura principale e per le sue connessioni con il nervo ottico, deve essere considerata come di natura nervosa. Risulta composta, in tutta la sua estensione, di due foglietti sovrapposti, un foglietto esterno ed un foglietto interno. Possiamo suddividere la retina in due parti.

La zona posteriore detta parte ottica, si estende dall'entrata del nervo ottico ed arriva fino all'ora serrata; il suo foglietto interno ha conseguito, nello sviluppo, una struttura complicatissima e serve alle funzioni visive. Il foglietto esterno, detto strato pigmentato, è formato da un unico strato di cellule che presentano caratteristiche epiteliali, applicate sulla membrana basale della coroide e ricche di pigmento scuro (fucsina). Queste cellule presentano una sezione esagonale su tagli perpendicolari alla superficie della retina, e sono provviste, dalla parte rivolta verso la retina, di numerosi e sottili prolungamenti, o frange, che si insinuano fra i coni ed i bastoncelli.

La zona anteriore, detta anche parte cieca, risulta formata di una parte ciliare, che aderisce al corpo ciliare, e di una parte iridea, che aderisce all'iride. In entrambe le parti, la retina è sottile e di struttura molto semplice; risulta formata in ambedue i foglietti da cellule con carattere epiteliale ed è insensibile alla luce.

La retina propriamente detta, costituita dal foglietto interno della parte ottica della retina, è una membrana regolarmente applicata, mediante la sua superficie convessa, alla faccia profonda dello strato pigmentato, con l'intermezzo del quale è in rapporto con la coroide; è inoltre a contatto, mediante la sua superficie profonda, col corpo vitreo. Lo spessore della retina va lentamente diminuendo da dietro in avanti: sul contorno della papilla del nervo ottico misura 0,4 mm, diventando, all'ora serrata, di 0,1 mm. Vi si descrivono tre regioni particolari: la papilla del nervo ottico, la macula lutea e l'ora serrata.

La papilla del nervo ottico corrisponde al punto dove si raccolgono, convergendo da ogni parte, le fibre nervose che hanno preso origine nella retina e che vanno a costituire il nervo ottico. La papilla appare come un piccolo disco di color biancastro, circolare o leggermente ellittico con l'asse maggiore trasversale, del diametro di 1,5-

1,7 mm. E' situata a 3-4 mm dal polo posteriore del bulbo, 1 mm più in basso. Essa presenta al centro una depressione, più o meno estesa e profonda, nota come escavazione fisiologica, dalla quale emergono i vasi retinici che è insensibile alla luce.

La macula lutea o fovea centralis è una piccola regione della retina che si trova presso il polo posteriore dell'occhio, lateralmente ad esso. Il suo centro si trova a circa 3,91 mm lateralmente al centro della papilla e 0,78 mm più in basso; al suo centro corrisponde l'estremità posteriore della linea della visione. Ha forma leggermente ellittica; il diametro medio è di circa 2 mm. La sua parte centrale presenta una fossetta, la fovea centrale, lunga 0,2-0,4 mm, dove la retina è molto sottile (0,1-0,08 mm). Nella fovea, lo strato delle cellule visive, che è il più profondo, a causa della riduzione degli strati antistanti è quasi allo scoperto, ed è perciò più accessibile all'azione diretta dei raggi luminosi; la fovea è la regione della visione distinta. L'ora serrata segna il margine anteriore della parte ottica della retina che, assottigliandosi bruscamente e modificandosi nella struttura, si continua nella parte ciliare. Appare come una linea circolare, formata da una serie di festoni le cui sporgenze corrispondono agli intervalli fra i processi ciliari.

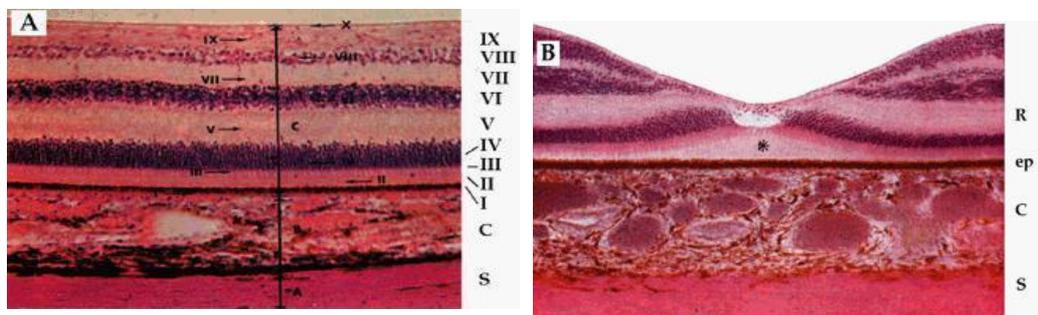


Figura 5. A. Sezioni istologiche della retina umana. Da 1 a 9 gli strati della retina con sclera e coroide. B.Fovea

Possiamo distinguere nella retina vari strati sovrapposti, dalla superficie esterna fino a quella interna. Essi sono:

- Coni e bastoncelli (strato I Figura 5 A)
- Membrana limitante esterna (strato II Figura 5 A)
- Granuli esterni (strato III Figura 5 A)
- Plessi forme esterno (strato IV Figura 5 A)
- Granuli interni (strato V Figura 5 A)
- Plessi forme interno (strato VI Figura 5 A)
- Cellule multipolari (strato VII Figura 5 A)
- Fibre nervose (strato VIII Figura 5 A)
- Membrana limitante interna. (strato IX Figura 5 A)

FUNZIONAMENTO

Nella retina vengono focalizzate le immagini visive, la luce è focalizzata dalla cornea e dal cristallino e attraversa l'umor vitreo prima di essere assorbita dai fotorecettori. I fotorecettori sono appunto i coni e i bastoncelli. Data la posizione di questi, la luce deve attraversare tutti gli altri strati della retina prima di colpirli; i neuroni parte degli strati retinici superficiali sono abbastanza trasparenti, cosicché la luce arriva ai fotorecettori senza venire assorbita o distorta in maniera apprezzabile. Nella fovea comunque, gli strati neuronici sovrastanti i ricettori sono spostati lateralmente cosicché i fotorecettori foveali ricevono le immagini visive nella forma meno distorta.

I coni e i bastoncelli sono cellule altamente specializzate il cui nome deriva proprio dalla loro forma, essi come detto sono gli elementi recettivi del sistema visivo atti a tradurre l'energia luminosa in potenziali elettrici.

Nella maggior parte della retina si presentano gruppi di bastoncelli, rigorosamente paralleli ed accostati gli uni agli altri, e ad ogni intervallo fra questi gruppi si trova un cono. Complessivamente, i bastoncelli sono 75 milioni, i coni sono 3 milioni. Nella fovea centrale, fino a 0,25 mm dal suo centro, esistono solo coni; al di là cominciano ad intercalarsi i bastoncelli che aumentano progressivamente, fino a che, a 3-4 mm dal centro della fovea, dove sono divenuti venti volte più numerosi dei coni.

Questa proporzione rimane sensibilmente costante fino alle immediate vicinanze dell'ora serrata, dove i bastoncelli diminuiscono di numero, senza che aumenti quello dei coni; in questo punto i coni e i bastoncelli non sono più a stretto contatto fra loro, ma spazieggiati.

Possiamo notare dalla tabella sottostante la differenza più significativa tra i coni e i bastoncelli e i loro sistemi neurali.

	CONI	BASTONCELLI
Sensibilità	bassa:poco pigmento, poca amplificazione, specializzati per la visione diurna (fotopica)	elevata:molto pigmento, molta amplificazione, specializzati per la visione notturna (scotopica)
Risoluzione temporale	elevata (fino a 55 Hz)	bassa (fino a 12Hz)
Acuità visiva	elevata: particolarmente concentrati nella fovea, vie retiniche poco convergenti	bassa: assenti nella fovea, vie retiniche molto convergenti
Colori	Cromatici: tre tipi di coni, ciascuno con un diverso pigmento	Acromatici: tutti con un solo tipo di pigmento

Tabella 1. Caratteristiche peculiari dei coni e dei bastoncelli

I coni e i bastoncelli come si evince dalla tabella 1 sono dotati di pigmenti visivi nel loro strato esterno, la luce va ad interagire su queste zone. Ogni pigmento è una molecola di piccole dimensioni, capace di assorbire la luce.

Nella figura sottostante lo schema specifico funzionale dei coni e dei bastoncelli.

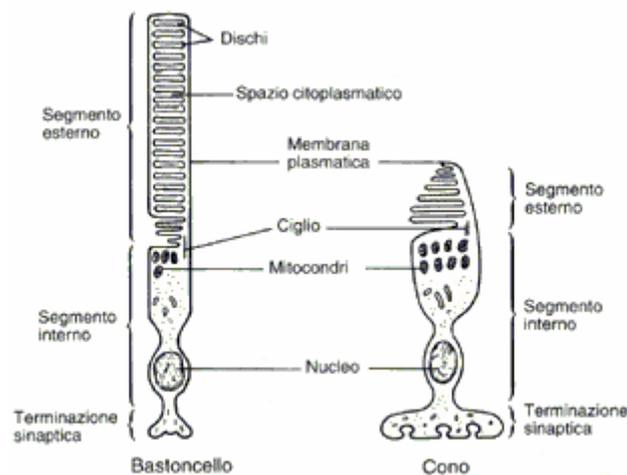


Figura 6. Schema specifico funzionale dei coni e dei bastoncelli.

Immagine tratta da [2]

Il segnale pervenuto ai fotorecettori viene modificato ed elaborato prima di essere inviato al sistema nervoso centrale, l'uscita della retina è acquisita dalle cellule gangliari che fanno parte del nervo ottico. In particolare il segnale viene trasferito alle cellule gangliari attraverso tre tipi di interneuroni: cellule bipolari, le cellule orizzontali e le cellule amacrine. In questo passaggio l'informazione viene separata in due vie parallele dette via centro on e via centro off. Le cellule gangliari centro-on vengono eccitate quando la luce stimola il centro dei loro campi recettivi e inibite quando ne viene stimolata la periferia; le cellule gangliari centro-off presentano risposte opposte. Queste trasformazioni dell'informazione visiva fanno sì che i centri superiori possano mettere in evidenza piccole differenze e rapide variazioni di luminosità. Altri tipi di cellule gangliari hanno altre funzioni quali ad esempio la rilevazione di dettagli di colore dell'immagine visiva, o anche la caratterizzazione del movimento dell'immagine.

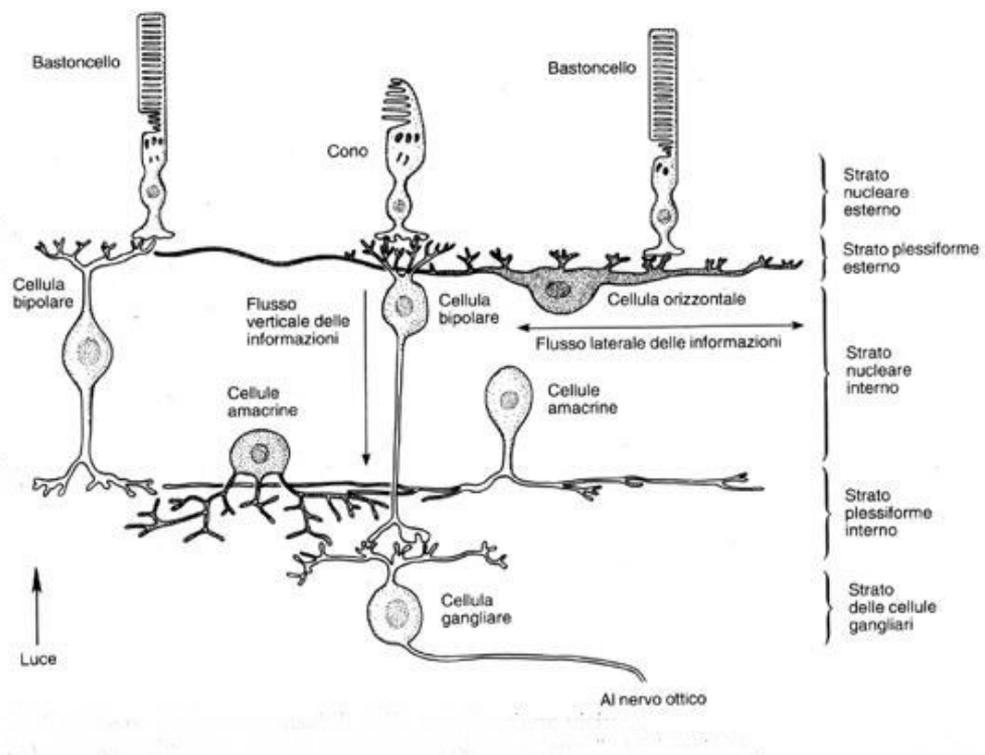


Figura 7. Flusso delle informazioni attraverso le cellule nervose sino al nervo ottico.

Immagine tratta da [2]

LA RETINA NELLA MOSCA

Il sistema visivo di questo insetto è molto interessante ed è stato oggetto di studi (e lo è anche al giorno d'oggi) al fine di trarre spunto da esso al fine di creare algoritmi per l'elaborazione delle immagini o anche hardware dedicato.

Come ampiamente trattato in [3] possiamo ricavare le caratteristiche principali del sistema visivo di questo insetto.

La mosca possiede due occhi composti situati sopra la testa, sporgenti da essa e di forma emisferica oltre che di dimensioni molto grandi se paragonati al resto del corpo. La superficie dell'occhio è rivestita da protuberanze di forma più o meno esagonale da cui parte e si sviluppa l'unità visiva fondamentale chiamata Ommatide.

I singoli si estendono interamente nella superficie dell'occhio e sono circondati da un rivestimento chiamato capsula oculare che li separa dalla testa.

Il numero di ommatidi che ricoprono gli occhi della mosca è notevole: dai 3000 ai 5000. Hanno singolarmente un diametro di 20 o 30 μm . Oltretutto si è notato che i maschi hanno una quantità di ommatidi maggiore delle femmine e questo potenzia ulteriormente il loro sistema visivo. Il campo visivo del singolo ommatide varia da 1° a 3° e tenendo conto dell'insieme totale degli ommatidi e della disposizione dei due occhi si trae la conclusione che il campo visivo totale è di circa 270°

Ogni Ommatide ha una struttura assai complessa. La sommità di esso consiste in uno strato corneale che funge da lente convergente seguito dal cosiddetto cono cristallino, un corpo piriforme trasparente che può essere sia solido che liquido. Alla base di quest'ultimo troviamo le "semper cell" da qui partono le celle retiniche, di fatto la retina della mosca. Le celle retiniche sono 8 delle quali però solo 6 sono collegate alla percezione del moto dedotto dalla visione che ha l'insetto. Esse sono raggruppate in un fascio cilindrico nel cui centro è disposta una sostanza particolare di struttura complessa chiamata Rhabdomere. Questo ultimo corpo contiene prevalentemente il pigmento fotosensibile che assorbe la luce che entra nell'ommatide. La retina quindi è collegata tramite fibre nervose al cervello dell'insetto che elabora le informazioni acquisite. Una funzione peculiare del sistema visivo della mosca è quello di analizzare, al fine di seguire un oggetto in movimento,

la variazione di luminosità che si genera tra punti vicini in seguito ad un movimento dell'oggetto².

Il primo ganglio dove viene eseguita una prima analisi ed elaborazione dell'informazioni giunte dalla retina è chiamato Lamina. Nella Lamina giungono segnali da 6 ommatidi vicini: è in questo momento che inizia la prima analisi precedentemente detta. Dopo la Lamina segue un altro centro neurale molto importante quale il complesso ganglio della Medulla. Ed infine successivamente l'informazione converge nel centro neurale della Lobula diviso in due sedi separate. La Figura 9 evidenzia quanto specificato relativamente alla struttura del singolo ommatide.

Altre funzioni derivate dal sistema visivo di questo insetto sono operazioni di prefiltraggio dell'immagine visualizzata, funzione operata a livello di Lamina.

E' bene ricordare però che il funzionamento del sistema visivo della mosca, data la complessità, non è stato ancora totalmente assimilato ciò nonostante i vari studi che si sono susseguiti lungo gli anni, in più nel simulare reti neurali che si ispirano a tale sistema si attua un enorme semplificazione dato il numero e la complessità di connessioni che questo presenta in loco.

La Figura 8 mostra l'occhio della mosca e nel dettaglio a destra il singolo ommatide. La Figura 9 mostra una sezione trasversale di ommatidi con i collegamenti ai centri nervosi, con evidenziate i gangli Medulla, Lamina e Lobula.

² Tratto da [4]

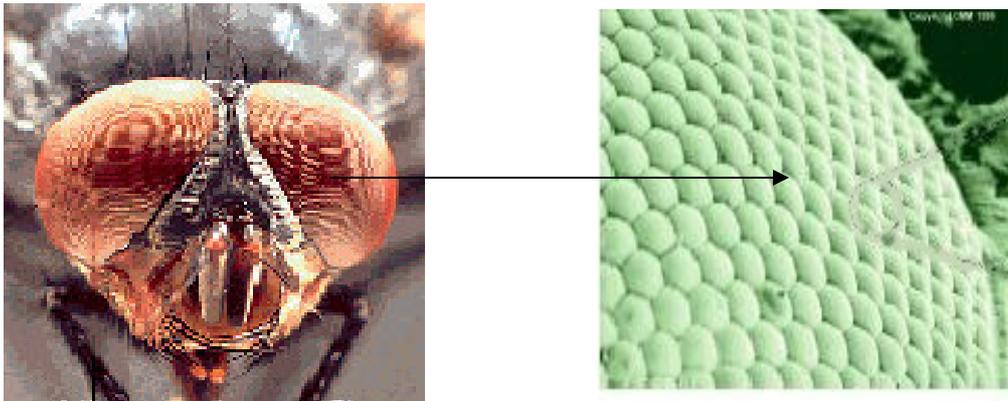


Figura 8. Dagli occhi della mosca al dettaglio del singolo Ommatide.

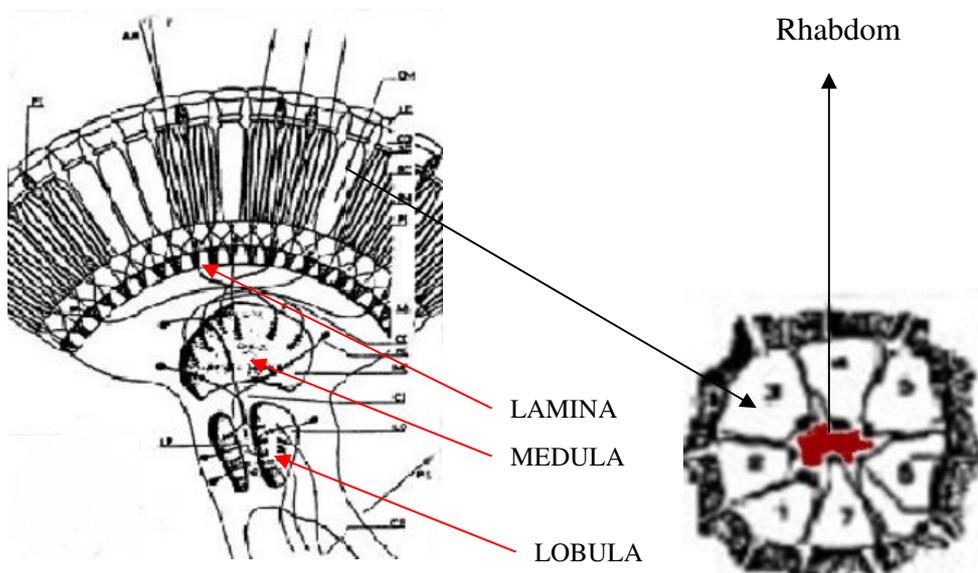


Figura 9. A sinistra, sezione trasversale di ommatidi, a destra sezione longitudinale della formazione retinulare.

2. IL SENSORE RETINICO

2.1 LA STORIA DEL SENSORE

Il sensore retinico fu sviluppato per la prima volta agli inizi degli anni 90 in Belgio in collaborazione con l'università di Genova. Già negli anni 80 però, era nata l'idea di produrre immagini log polari e queste ovviamente erano generate da un software che simulava il sensore retinico. Per la costruzione di questo primo sensore venne usata la tecnologia CCD(charge couplet device). Con la tecnologia dell'epoca, il diametro del sensore retinico era di 94 mm mentre la dimensione del pixel più piccolo e del pixel più grande erano rispettivamente di 30 μm e 412 μm . Il numero totale di fotorecettori era 2022 di cui 1920 disposti sugli anelli esterni alla fovea mentre quest'ultima possedeva 102 fotorecettori.

Nelle successive implementazioni del sensore venne invece usata la tecnologia CMOS.(Complementary Metal Oxide Semiconductor) La tecnologia CMOS consente una qualità di immagini superiore e una trasmissione dei dati più veloce; inoltre questa nuova tecnologia usata, ha minori dimensioni e consente la costruzione di apparecchi più piccoli.

Nel 1996 venne creato il primo sensore retinico con questa tecnologia, il progetto venne finalizzato con la collaborazione del consorzio IMEC, dell'università di Genova e Aitek. In questo tipo di sensore il diametro era di 8 mm, il pixel più piccolo aveva una dimensione di 14 μm , il numero di fotorecettori totali era di 8013 di cui 7168 nella parte spazio variante del sensore esterna alla fovea, mentre 845 fotorecettori all'interno della fovea

L'ultima realizzazione del sensore risale al 1997. Sviluppato nella città di Tower, in Israele, dall'Unione Europea, all'interno di un progetto di ricerca chiamato Svisca. Questo tipo di sensore (realizzato sempre con tecnologia CMOS) aveva un diametro di 7.1 mm. Il pixel più piccolo aveva una dimensione di 6.5 μm , il pixel più grande invece di 7.5 μm .

L'evoluzione nella tecnologia ha portato lungo gli anni a realizzare dei sensori retinici con una risoluzione sempre maggiore, apportando quindi un vantaggio nella qualità delle immagini acquisite tramite la telecamera retinica sfruttante il sensore

retinico. Oltre al vantaggio poi, dovuto alle dimensioni sempre più piccole dei sensori che consente, come detto, la realizzazione di apparecchi sempre più piccoli.

2.2 CARATTERISTICHE DEL SENSORE

Il sensore retinico è caratterizzato da una risoluzione spazio variante che come detto imita la distribuzione dei fotorecettori presenti nella retina umana³. La densità dei fotorecettori è minore nella periferia del sensore e aumenta avvicinandosi al centro del sensore, dove la densità dei fotorecettori è massima. Questi ultimi sono disposti lungo anelli concentrici e la spaziatura fra questi non è costante ma diminuisce anch'essa dalla periferia verso il centro del sensore. Schematizzeremo sempre nelle figure successive i fotorecettori come dei cerchi mentre utilizzeremo il termine anelli per identificare dove sono disposti i fotorecettori.

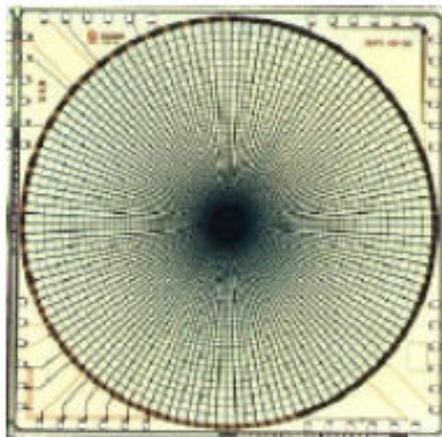


Figura 10⁴. Layout del sensore. La densità dei fotorecettori è notevolissima al centro.

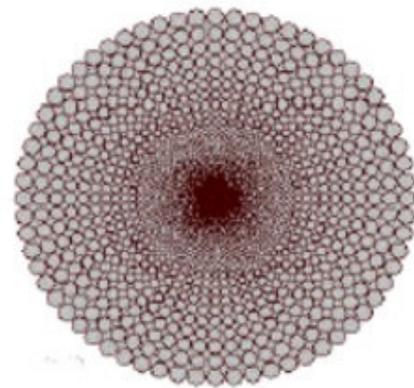


Figura 11. Esempio Distribuzione dei fotorecettori. 64 anelli per 48 cerchi (fotorecettori).

Conseguentemente alla diminuzione della distanza fra due anelli consecutivi, si verifica anche un altro fenomeno: i fotorecettori hanno una dimensione che è massima nell'anello più esterno, poi passando all'anello successivo più interno essa diminuisce ma il numero dei fotorecettori per anello rimane lo stesso. Questo

³ Giulio Sandini e Giorgio Metta, "Retina like sensors: motivations, technology and application.[5]

⁴ Immagini tratte da: "Giotto, retina like camera", Consorzio Unitek [6]

fenomeno permane sino all'anello che delimita la fovea, dove i fotorecettori hanno raggiunto la dimensione minima. Nella fovea quindi i fotorecettori hanno una disposizione diversa rispetto ai fotorecettori disposti sugli anelli esterni alla fovea. Nella prima realizzazione del sensore retinico con tecnologia Cmos il numero di fotorecettori sugli anelli della fovea veniva dimezzato da un anello al successivo quando necessario ed inoltre passando da un anello al successivo la dimensione del singolo fotorecettore aumentava. Possiamo vedere l'esempio di layout di questo sensore nella Figura 12. Nel caso specifico la fovea contiene 10 anelli con 64 fotorecettori, 5 anelli con 32 fotorecettori, 2 anelli con 16 fotorecettori ed 1 anello con 8, 4, 1 fotorecettore rispettivamente.

Nel sensore retinico di ultima costruzione invece, dall'anello della fovea sino al centro del sensore fotorecettori presentano una dimensione costante ma di conseguenza diminuiscono nel numero lungo gli anelli. La Figura 13 esemplifica il layout di questo sensore di ultima generazione, mentre la Figura 14 ne mostra il layout della fovea.

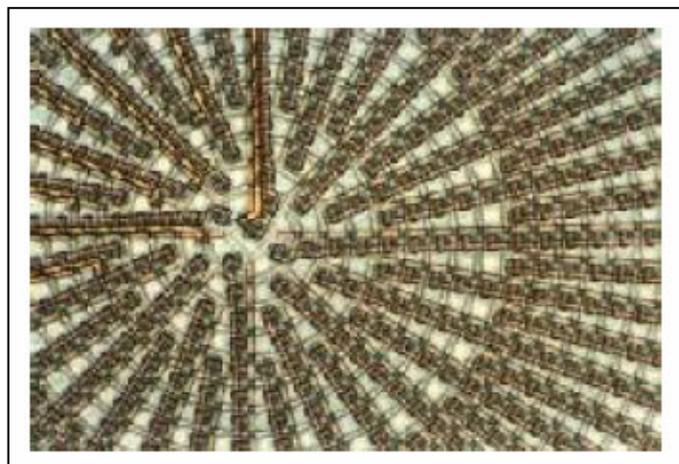


Figura 12. Layout della fovea nel primo sensore retinico realizzato con tecnologia Cmos.

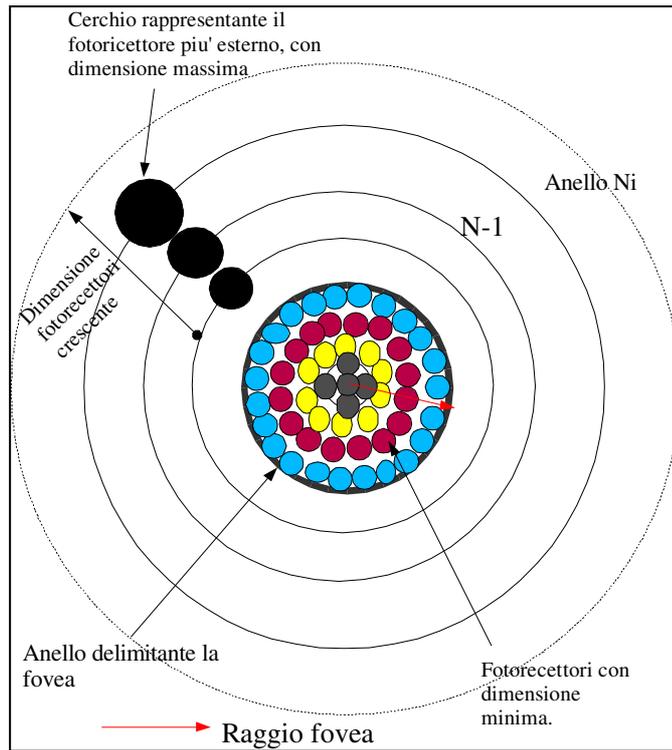


Figura 13. Schematizzazione degli anelli nel sensore e dei fotorecettori disposti sugli anelli. Notare la diminuzione della dimensione dei fotorecettori dall'anello piu' esterno verso il centro, e la costanza della dimensione all'interno della fovea.

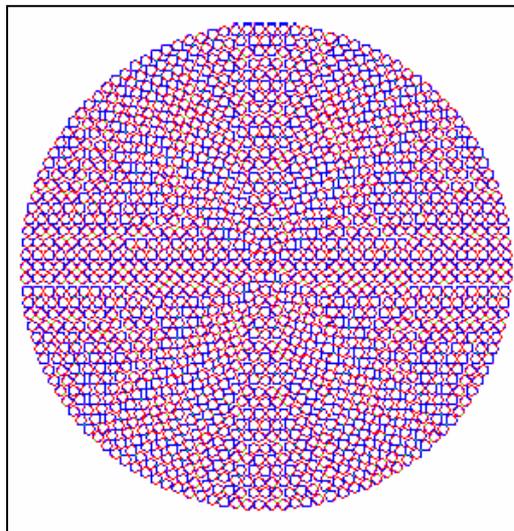


Figura 14. Layout della fovea dell'ultimo sensore Cmos realizzato. I pixel hanno dimensione costante

La telecamera utilizzata per l'acquisizione delle immagini presenta i dettagli tecnici visualizzati nella tabella 2, i dati sono stati raccolti da [6]; la tabella contiene anche i dettagli tecnici della geometria del sensore

GEOMETRIA DEL SENSORE	
- diametro del sensore	7.1 mm
- numero totale di anelli	152
- numero di anelli in periferia	110
- numero anelli fovea	42
- numero di pixel in periferia	27720
- numero di pixel nella fovea	5473
- raggio della fovea	0.273 mm
- dimensione del pixel più piccolo	0.0065 mm
TELECAMERA GIOTTO	
- Software	Windows Nt
- Interfaccia	Porta Parallela EPP
- Frame Rate	50 Frame /s
- Bit per pixel	8 bit (scala di grigi)
- Tipo lenti	C mount
- dimensioni	diametro 5.5 cm, lunghezza 7 cm

Tabella 2. Dettagli tecnici del sensore retinico e della telecamera Giotto

2.3 LE IMMAGINI LOG POLARI

Le immagini prodotte dal sensore retinico, o immagini log polari, ad una prima osservazione possono apparire molto strane in quanto presentano una disposizione dei pixel che comporta una interpretazione molto complicata della scena inquadrata dalla telecamera retinica. Guardiamo ad esempio l'immagine sottostante: per chiunque sarebbe difficile identificare la scena visualizzata dall'immagine log-polare. Scena che invece si chiarifica nell'immagine successiva.

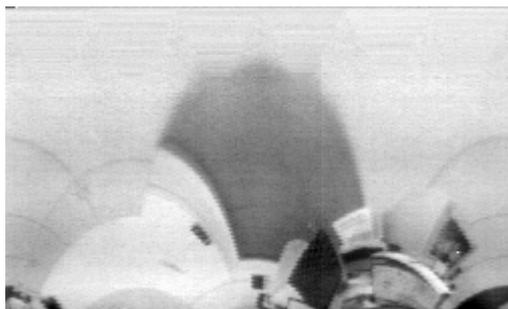


Figura 15. Immagine log-polare acquisita con la telecamera retinica.



Figura 16. L'immagine log-polare trasformata in cartesiana appare ora facilmente interpretabile.

A parte qualche forma o figura che rimane invariante, ovvero che non viene modificata rispetto alla rappresentazione tradizionale dell'immagine (acquisita tramite un sensore tradizionale), il resto delle forme e figure appare distorto. Questo effetto è dovuto al fatto che i pixel dell'immagine vengono disposti in un sistema di riferimento diverso: il cosiddetto piano corticale con coordinate del tipo $(\log \rho, \theta)$. Consideriamo il piano cartesiano con coordinate (x, y) ; possiamo ricavare la posizione di un punto utilizzando le coordinate polari (ρ, θ) , possiamo quindi identificare il pixel corrispondente ad un determinato fotorecettore con la sua posizione in coordinate polari. Il pixel quindi verrà mappato nel piano $((\log \rho, \theta))$ generando l'immagine log-polare.

Questo tipo di mappa è stato suggerito come citato in [7] dallo studio della corteccia visiva primaria dei primati dove l'insieme dell'informazione visiva recuperata dai fotorecettori presenti nella retina, viene mappata nella corteccia visiva. La rappresentazione della fovea è amplificata, ovvero una maggiore quantità di neuroni è dedicata ad essa.

La trasformazione dal piano cartesiano a quello log-polare segue specificatamente le seguenti equazioni:

$$Ni = \log_a(\rho / \rho_0)$$

$$Vi = k\theta$$

dove con Ni , Vi si indicano le coordinate nel piano log-polare, ρ è il raggio corrispondente all'anello su cui è disposto il pixel, θ è l'angolo. Chiariamo ora il significato dei parametri fissi: a , ρ_0 , k . Il primo parametro è la base del logaritmo che in pratica determina la spaziatura tra i vari anelli concentrici del sensore. Infatti dall'equazione precedente possiamo verificare che:

$$\rho_0 \times a^{Ni} = \rho_i$$

$$\Rightarrow \rho_i - \rho_{i-1} =$$

$$\rho_0(a^{Ni} - a^{Ni-1})$$

ovvero la distanza tra due anelli consecutivi dipende dal valore di a .

Il secondo parametro fisso ρ_0 rappresenta come già evidenziato il valore del raggio dell'anello più interno della rappresentazione log-polare. Possiamo notare che l'equazione di trasformazione dal piano cartesiano al piano log polare presenta una discontinuità nell'origine, cioè quando ρ_0 assume il valore zero. Questo significa che il fotorecettore più interno del sensore dovrebbe avere una dimensione zero ma questo non è fisicamente possibile, quindi il raggio ρ_0 è limitato dalla dimensione del pixel più piccolo. Poiché però abbiamo visto nel Capitolo 1 che il sensore segue la rappresentazione log-polare per i pixel esterni alla fovea, ma non all'interno di questa (dove i pixel hanno dimensione costante) ,il raggio ρ_0 è il raggio della fovea. Infine

il valore K determina la risoluzione angolare in quanto $1/k$ è il valore minimo dell'angolo rappresentabile.

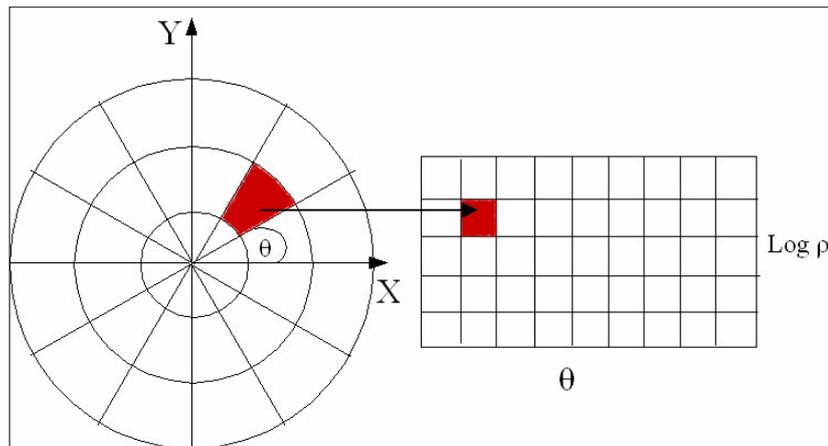


Figura 17. Viene rappresentato il mapping di un pixel dal piano cartesiano a quello log-polare

L'immagine log-polare prodotta dalla telecamera retinica può essere rimappata nella corrispondente immagine cartesiana evidenziando la risoluzione spazio variante; in questo modo è possibile avere una rappresentazione interpretabile facilmente della scena, anche se con risoluzione variabile. E' bene notare un particolare molto importante. Nel rappresentare l'immagine cartesiana in pratica viene simulata visivamente la struttura del sensore log-polare; prendiamo ad esempio un fotorecettore nella periferia (con dimensione quindi molto grande), esso campiona una zona della scena più vasta rispetto ad un fotorecettore più vicino al centro del sensore, ma nel rappresentare il pixel del sensore corrispondente al fotorecettore più grande dobbiamo utilizzare più pixel e replicarli tanto più, quanto più grande è il pixel (infatti i pixel di un immagine sono tutti uguali, non è possibile avere dei pixel più grandi e dei pixel più piccoli). La Figura 18 chiarisce ulteriormente questo ultimo passaggio molto importante.

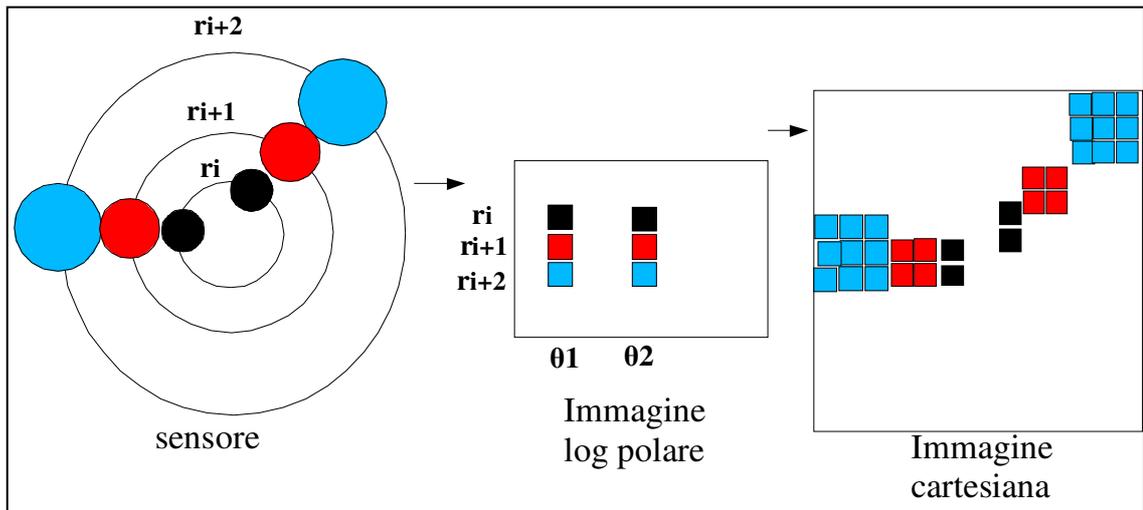


Figura 18. Per rappresentare l'informazione prodotta dal fotorecettore del sensore, viene replicato n volte un pixel nell'immagine cartesiana. Possiamo notare la transizione sensore-immagine log-polare- immagine cartesiana

Nella Figura 19 sottostante possiamo notare l'effetto di simmetria radiale indotta nell'immagine log-polare: i petali del fiore disposti lungo raggi nell'immagine cartesiana a sinistra vengono mappati lungo le righe dell'immagine log-polare. In questi caso, data la simmetria radiale i petali vengono mappati lungo le righe dell'immagine log-polare.



Figura 19. A sinistra l' immagine log-polare, a destra l'immagine cartesiana rimappata.

Si può notare un ulteriore caratteristica molto interessante e importante dell'immagine log-polare. I pixel disposti sugli anelli vengono mappati come si è visto lungo le righe del piano log-polare in quanto il raggio per questi pixel è costante; poiché però i pixel nella fovea rappresentano una area molto piccola, nel piano log-polare il valore dei pixel viene di fatto "dilatato" ovvero come prima replicato tante volte quanto basta per completare l'intera riga dell'immagine log-

polare. Ovvero anche per l'immagine log-polare esiste una variazione di risoluzione dalle prime righe di essa (dove è minore) fino alle ultime (dove è maggiore) Questo fenomeno provoca un ulteriore fattore di distorsione nelle prime righe dell'immagine log-polare in cui viene di fatto mappata la fovea. Prendiamo come esempio l'immagine sottostante, i quadrati centrali vanno ad interessare la fovea del sensore e quando sono mappati nell'immagine log-polare compaiono più grandi e deformati nelle prime righe dove vi è minore risoluzione.

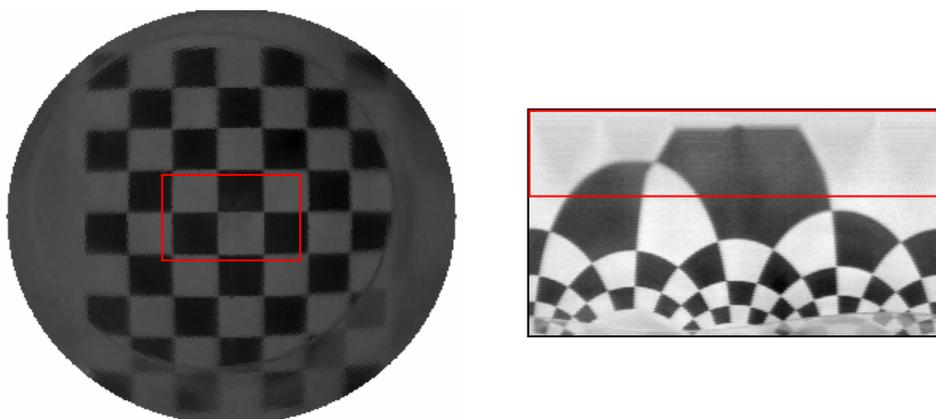


Figura 20. In questa immagine si nota che i quadrati che vanno ad interessare la fovea del sensore in sede di mappaggio log-polare compaiono più grandi e deformati perché si collocano nelle prime righe dove la risoluzione è minore.

Possiamo chiarificare ulteriormente questo passaggio tramite la Figura 21 in cui viene schematizzato il mapping tra i pixel identificati dai fotorecettori presenti nella fovea e l'immagine log-polare.

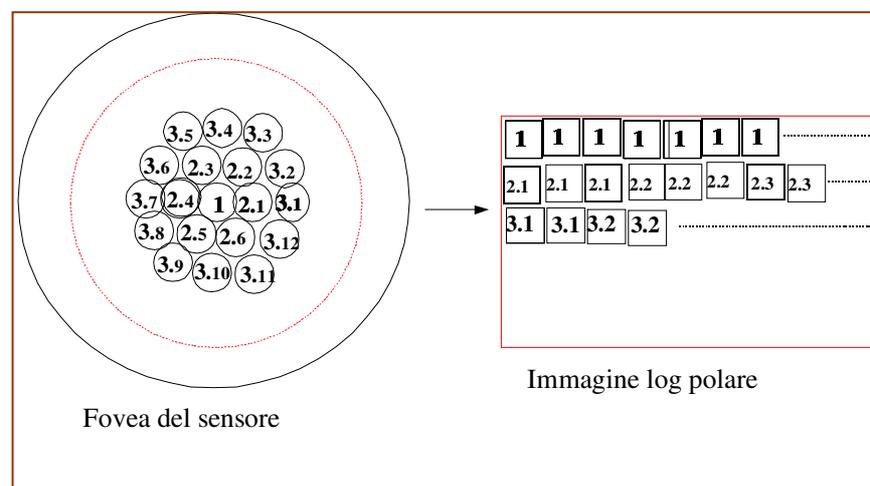


Figura 21. Mapping dei pixel nella fovea nell'immagine log-polare.

2.4 IL PESO DELL'IMMAGINE LOG-POLARE

INTRODUZIONE

Si è posta precedentemente notevole attenzione sui vantaggi che derivano dall'utilizzo di un sensore retinico piuttosto che di uno standard. Focalizziamo ora l'attenzione sul beneficio più importante offerto da questa classe di sensori ovvero la riduzione del peso dell'immagine. Questo vantaggio può essere utilizzato in notevoli applicazioni in cui è cruciale il peso dei dati trattati, come ad esempio nella telesorveglianza dove è importante la velocità di trasferimento dati tra le varie telecamere e l'elaborazioni degli stessi; in applicazioni robotiche dove la riduzione del peso dei byte trattati diminuisce la complessità degli algoritmi utilizzati, in applicazioni spaziali dove si ha l'intento di ridurre il consumo di potenza dei dispositivi e quindi anche in questo caso può essere utile trarre vantaggio nell'utilizzo di questi sensori.

Ovviamente in queste applicazioni la risoluzione spazio variante offerta dal sensore retinico non costituisce un problema anzi, il compromesso tra la risoluzione offerta e il peso ridotto dei byte dà un supporto di qualità per l'obiettivo preposto a queste applicazioni.

Consideriamo ad esempio il problema della teleconferenza. In questo caso è evidente che la riduzione dei dati che transitano attraverso la rete è il fattore più importante da minimizzare mentre la risoluzione spazio variante evidenziata a video è comunque comoda, in quanto si può focalizzare la telecamera sulle caratteristiche salienti dell'individuo (quali la bocca) non essendo utile un alta risoluzione anche nella periferia dell'immagine ovvero in zone lontane dalle caratteristiche salienti dell'individuo sopra menzionate.

Parametri di interesse

Come citato nel paper [5] vi sono dei parametri interessanti che permettono di evidenziare la riduzione di peso tra due immagini generate da un sensore retinico e uno tradizionale. Consideriamo innanzitutto la dimensione del sensore, e la

dimensione del pixel più piccolo. Di fatto quest'ultimo determina la risoluzione in quanto questa è definita come il rapporto tra l'area in mm^2 e il numero di pixel occupanti tale area. Si vede infatti dalla Figura 22 come si evidenzia la differenza di risoluzione tra due immagini di dimensione uguale.

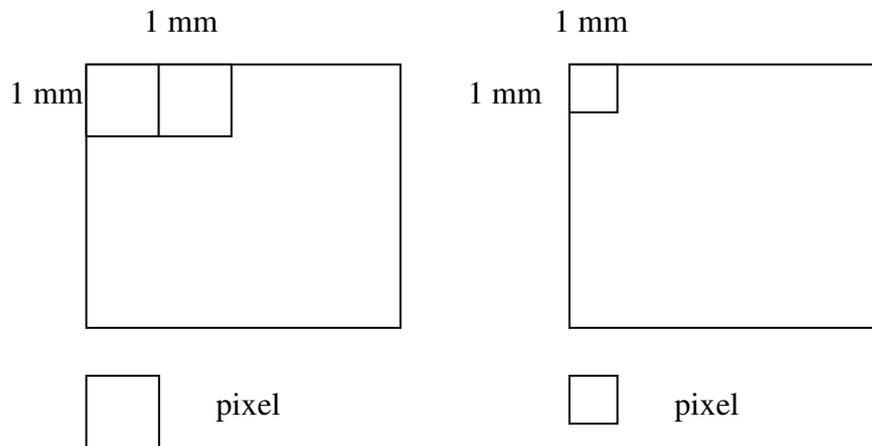


Figura 22. A sinistra la risoluzione minore, a destra la risoluzione maggiore

Il numero totale dei pixel di un'immagine è direttamente collegato all'ammontare dell'informazione acquisita: nei sensori tradizionali questo parametro è fissato dalla particolare tecnologia usata. Evidenziamo altri due parametri molto importanti chiamati sempre in [2] parametri R e Q. Il primo è definito come il rapporto tra la dimensione del pixel più grande e la dimensione del pixel più piccolo. Il parametro Q è definito come il rapporto tra la dimensione del sensore e la dimensione del pixel più piccolo. Nel primo prototipo di sensore retinico (costruito agli inizi degli anni 90) il parametro R era uguale a 13.7 ma ovviamente era limitato dalla tecnologia dell'epoca. Nei sensori tradizionali il parametro R è uguale a 1 perché la dimensione del pixel più piccolo e di quello più grande coincidono. Il parametro Q è molto importante poiché rappresenta la dimensione di un'immagine prodotta da un sensore tradizionale mantenendo lo stesso campo visivo e la stessa risoluzione del corrispondente sensore retinico. Se ad esempio il parametro Q fosse uguale a 200 significa che, qualora si volesse riprodurre lo stesso quantitativo di informazione prodotto da un'immagine log-polare, la dimensione dell'immagine generata da un sensore tradizionale con risoluzione costante dovrebbe avere una dimensione di 200×200 pixel.

I parametri di interesse dell'ultima realizzazione del sensore [5] vengono riassunti nella tabella 3.

PIXEL PERIFERICI	PIXEL NELLA FOVEA	PIXEL TOTALI	R	Q	DIMENSIONE SENSORE
27720	5473	33193	17	1100	7.1 mm

Tabella 3. Parametri caratteristici del sensore di ultima generazione

In questo caso il parametro Q è stato notevolmente aumentato. Un immagine derivata da un sensore tradizionale dovrebbe avere una dimensione di 1100×1100 per produrre lo stesso aumentare di informazione derivata da un sensore di tipo retinico! Quindi la riduzione di peso in bit è notevole; considerando infatti un immagine log-polare in scala di grigio si ha che ogni pixel si assegnano 8 bit e quindi segue facilmente la riduzione complessiva del peso in bit:

$$\text{IMMAGINE LOG POLAR} \text{ -----} \rightarrow (27720 + 5473) \times 8 = 265544 \text{ bit}$$

$$\text{IMMAGINE TRADIZIONALE} \text{ ----} \rightarrow (1100 \times 1100) \times 8 = 9680000 \text{ bit}$$

La riduzione è più di un fattore 30.

Per comprendere maggiormente l'utilità di questa riduzione si può fare una considerazione simile ed opposta che predilige l'aspetto qualitativo del problema. Disponendo di un determinato numero di pixel uguale sia per l'immagine log-polare acquisita tramite la telecamera dotata del sensore retinico, che per l'immagine a risoluzione costante (tradizionale), acquisita tramite una telecamera con un sensore tradizionale, premettendo come al solito di usare la massima risoluzione uguale per entrambi i sensori e lo stesso campo visivo si perviene a due immagini diverse rappresentate in Figura 23. Si può notare che la perdita di informazione che si ha nell'immagine tradizionale è notevole rispetto all'immagine log-polare che pur presenta una risoluzione decrescente dal centro verso la periferia.



Da tradizionale a retinico..



Remapping cartesiano



Figura 23. In alto la simulazione di un immagine ottenuta tramite un sensore tradizionale con lo stesso numero di pixel utilizzato nell'immagine in basso, ottenuta con il sensore retinico.

2.5. CALCOLO DEL PESO DELL'IMMAGINE LOG POLARE

Nel paragrafo precedente si è voluto evidenziare come l'immagine log-polare presenti una dimensione ridotta, se confrontata con un'immagine acquisita tramite un sensore tradizionale con la medesima risoluzione. Nella trattazione che segue si vuole stimare il numero dei fotorecettori del sensore retinico (e quindi il numero di pixel totali dell'immagine log-polare) partendo da informazioni note quali la dimensione del sensore e la risoluzione massima di quest'ultimo; confrontando ancora una volta il risultato con il numero di pixel dell'immagine acquisita con il sensore tradizionale aventi la medesima dimensione e risoluzione del sensore retinico.

Schematizziamo semplicemente il sensore log-polare come evidenziato nella Figura 24. A destra un equivalente sensore tradizionale, dotato cioè della stessa dimensione del sensore retinico e della stessa risoluzione massima. ρ_0 è il raggio del cerchio comprendente la fovea, in cui i pixel interni hanno tutti la stessa dimensione e vengono disposti in modo che man mano che ci si avvicina al centro del sensore il loro numero diminuisce. Nella fovea quindi non viene mantenuta la caratteristica del sensore di spazio varianza, perché nel suo interno i fotorecettori hanno tutti la stessa dimensione. Fuori da questa zona, negli anelli successivi i pixel variano la dimensione ma il numero di questi attorno agli anelli rimane costante. La risoluzione di entrambi i sensori è la stessa: il pixel di dimensione inferiore nella fovea del sensore log-polare ha la medesima dimensione del pixel nel sensore tradizionale. Inoltre come detto, anche la dimensione dei due sensori coincide. Quindi confrontiamo i due sensori a parità di risoluzione, campo visivo e dimensione come fatto in precedenza.

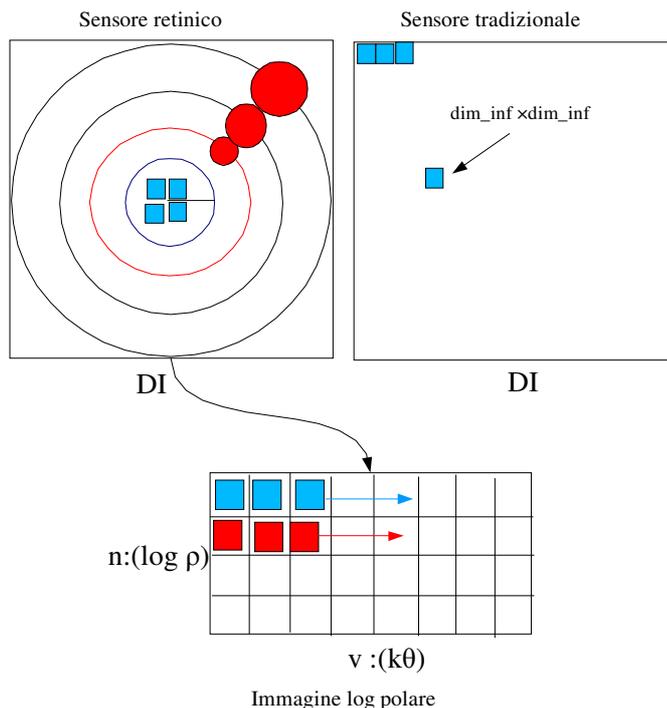


Figura 24. La figura esemplifica come i pixel corrispondenti ai fotorecettori nel sensore retinico vengano mappati nel piano log-polare specificando che in questo caso la risoluzione del sensore retinico e del sensore tradizionale è la medesima.

Il numero totale di pixel all'interno della fovea viene calcolato considerando l'area di essa e dividendo per l'area del singolo pixel al suo interno, che possiede la dimensione inferiore.

Quindi il numero di pixel della fovea è circa uguale a:

$$\text{Num_fovea} \sim \pi \rho_f^2 / \text{dim_inf}^2 \quad (2.5.1)$$

Il passo successivo è quello di calcolare il numero totale dei pixel esterni alla regione di fovea dove il sensore conserva la caratteristica peculiare di spazio varianza. Per fare questo calcoleremo il numero di anelli e il numero di fotorecettori che si addensano per anello.

Consideriamo le usuali posizioni:

$$n_i = \log_a (\rho_i / \rho_0) \quad (2.5.2)$$

$$v_i = k \times \theta_i \quad (2.5.3)$$

dove n_i , v_i rappresentano le coordinate discretizzate dell'immagine log-polare. La base a del logaritmo rappresenta un fattore di scala che caratterizza la risoluzione

radiale mentre k caratterizza la risoluzione angolare; ρ_0 è il raggio della fovea. Tramite la posizione (3.4) possiamo ricavare il numero massimo di anelli N :

$$N = \log_a (\rho_{\max} / \rho_0) \quad (2.5.4)$$

basandoci sul raggio massimo ρ_{\max} dell'anello più esterno dove sono collocati i fotorecettori (schematizzati come cerchi) corrispondenti ai pixel di dimensione maggiore. La base a del logaritmo può essere calcolata tramite la posizione:

$$a = e^{\ln(\rho_{\max} / \rho_0) / N} \quad (2.5.5)$$

Ora indichiamo con D_{\max} la distanza tra il cerchio della fovea e l'anello più esterno. Inoltre chiamiamo DI (come nella Figura 24) la dimensione del sensore. Per D_{\max} valgono le posizioni:

$$D_{\max} = a^N \rho_f - \rho_f \cong (DI / 2 - \rho_f) \quad (2.5.6; 2.5.7)$$

dove si è approssimata la grandezza D_{\max} con la differenza tra $DI/2$ e ρ_f .

A questo punto calcoliamo il numero massimo di fotorecettori per anello.

Il numero totale di anelli N dipende dal numero di cerchi(fotorecettori) e dalla dimensione del sensore. Ora consideriamo la Figura 25.

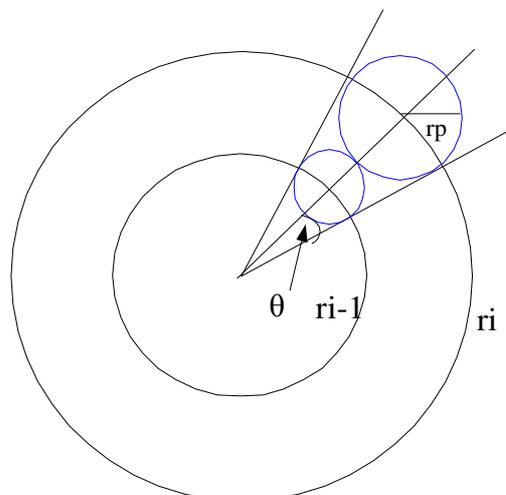


Figura 25. Il fotorecettore individua una zona angolare.

Per semplicità si sono disegnati soltanto due anelli, l'anello r_i e l'anello r_{i-1} . Si è evidenziato in blu un cerchio corrispondente ad un pixel sull'anello più esterno. Come al solito indichiamo con r_i ed r_{i-1} il valore del raggio dell'anello più esterno e il valore del raggio più interno rispettivamente.

Il valore del raggio r_p del cerchio in blu sull'anello può essere calcolato mediante la posizione:

$$r_p = r_i \sin \theta \quad (2.5.8)$$

Si nota facilmente dalla figura che il numero di cerchi M per anello è uguale al rapporto tra π e θ . Detto questo si ricava facilmente una relazione tra r_i e r_{i-1} .

$$r_i(1 - \sin \theta) = r_{i-1}(1 + \sin \theta) \quad (2.5.9)$$

Sappiamo che (raggio esterno) $r_N = a^N \rho_0$ quindi riscriviamo la (2.5.9) come segue e ricaviamo il valore a in un altro modo rispetto a (2.5.5):

$$\begin{aligned} a^N \rho_0 - a^N \rho_0 \sin \theta &= a^{N-1} \rho_0 + a^{N-1} \rho_0 \sin \theta \Rightarrow a^N / a^{N-1} = (1 + \sin \theta) / (1 - \sin \theta) \\ \Rightarrow a &= (1 + \sin \theta) / (1 - \sin \theta) \end{aligned} \quad (2.5.10)$$

Quindi sfruttando ancora la posizione (3.9) con $i=N$ risulta che:

$$\begin{aligned} r_N &= a r_{N-1} = \\ &a(a(r_{N-2})) = \\ &a(a(a(r_{N-3}))) = \\ \cdot & \quad \cdot a^N r_0 = [(1 + \sin \theta) / (1 - \sin \theta)]^N r_0 \end{aligned} \quad (2.5.11)$$

e inoltre adoperando correttamente la definizione di logaritmo, con la base $a = (1 + \sin \theta) / (1 - \sin \theta)$, rivediamo la 3.4:

$$\log_a r_N / r_0 = N; \quad (2.5.12)$$

Poiché per r_N vale la relazione:

$$r_N(1+\sin\theta) \sim DI/2 \quad (2.5.13)$$

si ha che:

$$N = \log_a (DI/2) / [(1+\sin\theta)r_f] \quad (2.5.14)$$

Possiamo quindi ora ricavare θ dalla (2.5.14) così che

$$\begin{aligned} \theta &= \sin^{-1}[(a^k - 2r_f) / 2r_f] \\ ; k &= \log_a (DI - N + 1) \end{aligned} \quad (2.5.15)$$

e da θ si ricava V come: $V = \pi/\theta$

Abbiamo quindi ricavato una serie di equazioni (2.5.6), (2.5.7), (2.5.13), (2.5.15) che compongono un sistema di 4 equazioni nelle incognite desiderate: a , ρ_{max} , ma soprattutto N , V . Questi due ultimi valori moltiplicati tra loro danno un'indicazione dei pixel parziali del sensore che sommati a quelli interni alla regione di fovea danno il numero di pixel totali dell'immagine log-polare.

Il sistema a cui si è pervenuti è qui riassunto:

$$\left\{ \begin{aligned} a &= (1 + \sin \theta) / (1 - \sin \theta) \\ (DI/2 - \rho_f) &= \rho_f (a^N - 1) \\ r_N(1 + \sin \theta) &= DI/2 \\ \theta &= \sin^{-1}[(a^k - 2\rho_f) / 2\rho_f]; \text{ con } k = \log_a(DI - N) \end{aligned} \right.$$

, il numero di pixel totali del sensore retinico a cui si perviene: $N \times V + \text{num.fovea}$, mentre il numero di pixel totali del sensore tradizionale è: $DI^2 / \text{dim_inf}^2$

Con questa ultima trattazione ci si è ricondotti al calcolo più generale del valore di N e V che determina il numero di pixel dell'immagine log-polare prodotta dal sensore retinico e il numero di pixel dell'immagine prodotta dal sensore tradizionale, partendo unicamente da parametri noti quali la dimensione del sensore e la

risoluzione massima, (parametri coincidenti ad entrambi i sensori) e conoscendo il valore ρ_f del raggio del cerchio della fovea. Punto cardine rimane sempre la notevole riduzione di peso dell'immagine generata dal sensore retinico.

3. GLI INVARIANTI GEOMETRICI NELLE IMMAGINI RETINICHE

3.1 INTRODUZIONE.

Nel Capitolo precedente si sono illustrate le caratteristiche principali delle immagini log polari derivate dalla particolare struttura del sensore retinico. Le immagini log polari presentano anche altre proprietà molto importanti. Esse evidenziano particolari invarianti geometrici ricavabili sotto opportune condizioni. Prima di evidenziare gli invarianti nelle immagini retiniche cominceremo nel verificare gli invarianti nelle immagini prospettiche acquisite tramite una telecamera tradizionale. Cominceremo innanzitutto a spiegare come avviene il processo di formazione dell'immagine acquisita tramite telecamera, illustrando anche quali modelli di telecamera e quindi di proiezione verranno usati.

Il processo di formazione dell'immagine⁵ di una scena tridimensionale acquisita tramite una telecamera implica una selezione tra i raggi luminosi riflessi o trasmessi dagli oggetti presenti nella scena che vengono proiettati nel piano di formazione dell'immagine: lo scopo del processo di selezione dei raggi è quello di produrre un'immagine che sia una buona rappresentazione della scena tridimensionale. Per ottenere questo risultato di selezione il modello che viene utilizzato è il modello "pinhole camera", dove si schematizza la telecamera con una scatola chiusa con un foro praticato su di un lato. Tra tutti i raggi luminosi propagati da ogni singolo punto dello spazio tridimensionale, solo quelli direzionati attraverso il foro entrano nella scatola. Possiamo verificare questo ultimo passaggio con la Figura 26.

⁵Argomento ampiamente trattato in [8], da cui sono tratte anche le immagini

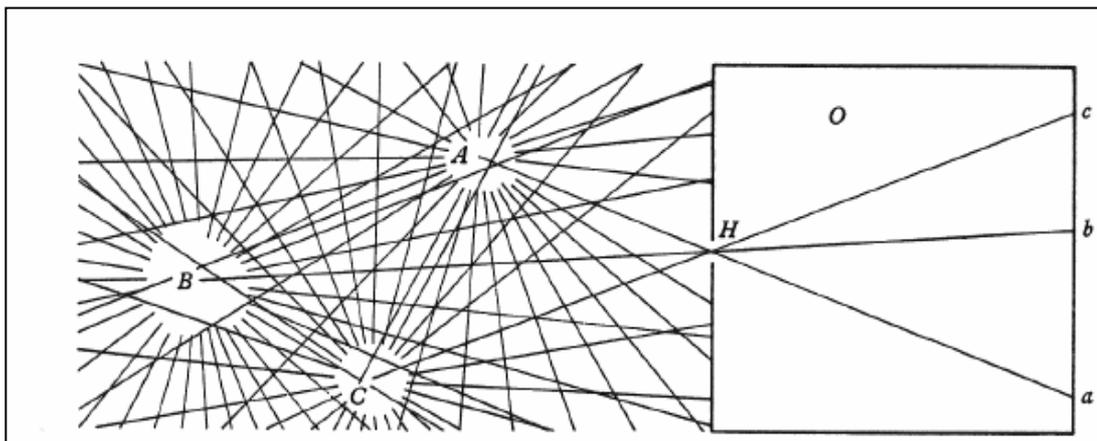


Figura 26. Attraverso il foro passano solo i raggi direzionati su esso e si proiettano su a, b,c

Ad ogni punto dell'oggetto 3D corrisponde nello spazio immagine una piccola elisse di luce come contributo all'immagine del punto dell'oggetto. L'insieme di tutte le elissi prodotte da tutti i punti dello spazio genera l'immagine. Più piccolo è il foro più alta è la risoluzione dell'immagine, ma non è possibile ridurre eccessivamente la dimensione del foro poiché così facendo diminuirebbe di conseguenza anche la quantità di luce raccolta. Dovremmo quindi, a scapito della risoluzione, costituire il foro di dimensione elevata. In pratica viene sostituito al foro una lente, essa permette di concentrare tutta la luce proveniente da una zona più ampia, che focalizza sulla superficie di formazione dell'immagine. La lente però, introduce nell'immagine una distorsione.

Il modello 'pinhole camera' realizza la proiezione prospettica, l'immagine è una proiezione bidimensionale del mondo tridimensionale su di un piano. Un punto nello spazio tridimensionale viene proiettato, attraverso una retta detta raggio di proiezione passante per il punto di proiezione (il pinhole), sul piano di proiezione bidimensionale. Questo passaggio viene esemplificato dalla Figura 27.

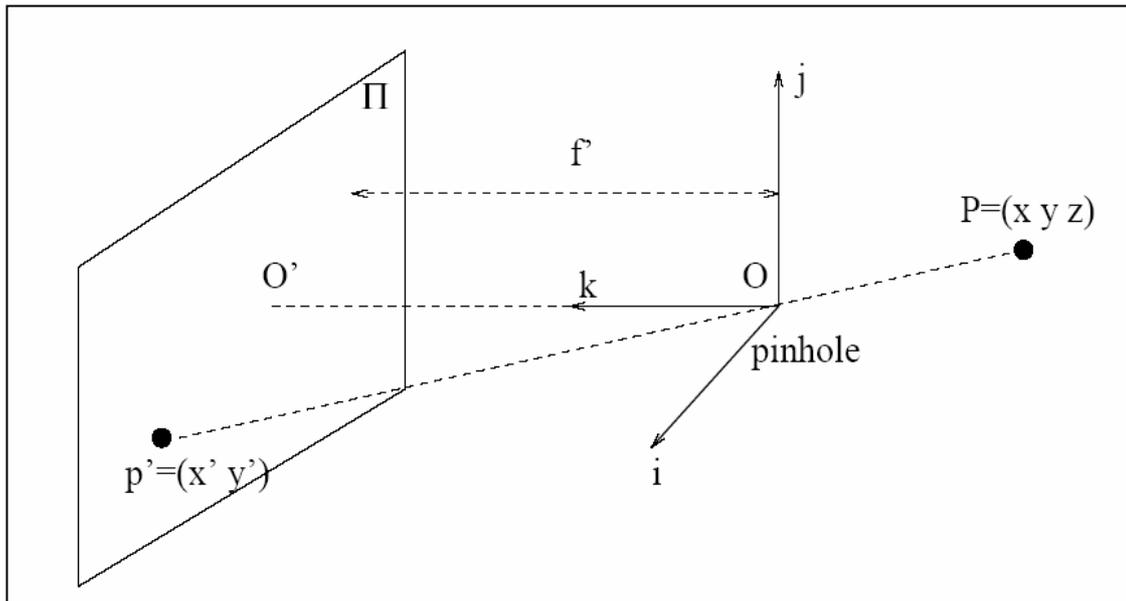


Figura 27. La proiezione prospettica.

La Figura 27 evidenzia il piano di proiezione (piano immagine) Π , il sistema di riferimento della telecamera; le direzioni degli assi sono identificati dai versori i, j, k .

Il versore k identifica la direzione dell'asse ottico della telecamera.

La distanza focale f rappresenta la distanza tra il pinhole ed il piano immagine.

Il punto $P(x, y, z)$ viene proiettato attraverso la retta di proiezione nel punto $P'(x', y')$ appartenente al piano immagine

Le equazioni che regolano questa trasformazione sono:

$$x' = f' x / z$$

$$y' = f' y / z$$

Oltre alla proiezione prospettica sono stati proposti altri modelli geometrici per modellare il processo di proiezione di un punto nella scena 3D nel piano immagine. Evidenziamo ad esempio il modello 'weak perspective' [9] e il modello di proiezione ortografica.

Nel primo dei due modelli sopra citati si assume che la profondità della scena sia piccola rispetto alla distanza media tra la telecamera e la scena inquadrata.

$$x' = f' x / z_0$$

$$y' = f' y / z_0$$

Esemplifichiamo il modello "weak perspective" con la Figura 28.

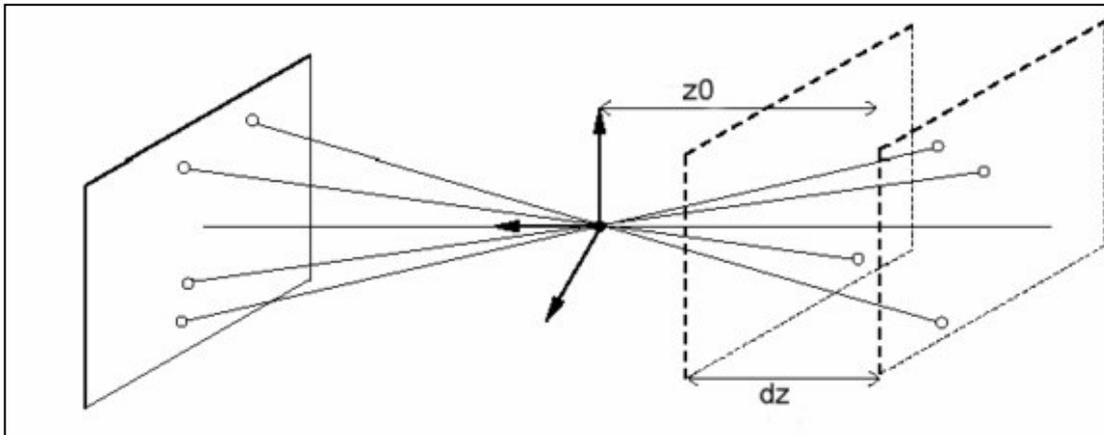


Figura 28. Il modello "weak perspective", $dz \ll z_0$

La proiezione ortografica invece, è un'approssimazione valida nel caso in cui la distanza tra la scena inquadrata ed il piano immagine sia molto grande. In questo caso le rette di proiezione di ogni punto dell'oggetto della scena inquadrata sul piano immagine sono parallele tra loro e perpendicolari al piano immagine. In questo caso le equazioni della trasformazione si semplificano:

$$x' = x$$

$$y' = y$$

Esemplifichiamo la proiezione ortografica con la Figura 29.

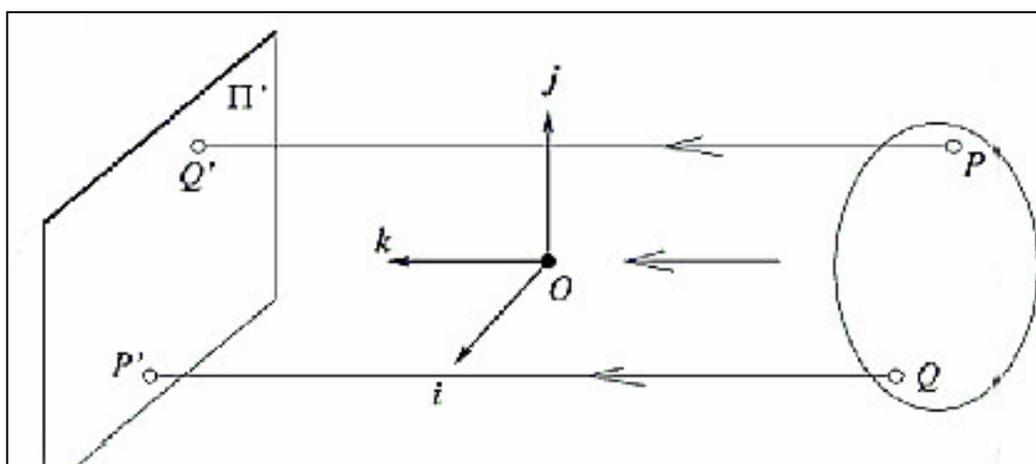


Figura 29. La proiezione ortografica.

Nella maggior parte dei casi pratici il modello che interpreta meglio la realtà è il modello “pinhole” di proiezione prospettica con cui possiamo rappresentare con maggiore fedeltà la scena 3D.

3.2 GLI INVARIANTI NELLE TRASFORMAZIONI GEOMETRICHE

Gli invarianti sono proprietà geometriche che non si modificano in seguito ad una trasformazione geometrica. Una branca della geometria ha come oggetto lo studio di queste proprietà, definendo così specifiche trasformazioni che hanno determinate proprietà invarianti.⁶

L'esempio più generale è quello della cosiddetta “trasformazione affine”: essa è una applicazione biettiva T fra due piani Π e Π' che fa corrispondere al punto P di coordinate X e Y il punto P' di coordinate x e y secondo la formula:

$$\begin{aligned}X &= ax + by + e \\ Y &= cx + dy + f\end{aligned}$$

dove i coefficienti a, b, c, d sono numeri reali. L'applicazione è biettiva se $ad-bc \neq 0$.

Questo tipo di trasformazione ha varie proprietà invarianti:

- la conservazione del parallelismo fra rette
- la conservazione dell'incidenza fra rette
- la conservazione del rapporto fra segmenti paralleli

Una proprietà che in questa particolare classe di trasformazioni non è invariante è la forma della figura.

⁶ Nozioni tratte da [10],[11], [12]e [13]

3.3 INVARIANZE GEOMETRICHE IN IMMAGINI PROSPETTICHE

Consideriamo ora la visione di un'immagine catturata con una telecamera convenzionale,⁷ se viene modificata la posizione della telecamera o similmente la posizione degli oggetti inquadrati, alcune proprietà geometriche degli oggetti restano invariate mentre altre si modificano, quindi inducono un particolare tipo di trasformazione a seconda del tipo di proprietà invarianti che si manifestano. Al fine di determinare gli invarianti si deve specificare anticipatamente il modello di telecamera utilizzato che realizza un determinato tipo di prospettiva. Il più semplice modello geometrico di telecamera che possiamo considerare è quello identificato dalla proiezione ortografica trattato nel paragrafo precedente, in questo caso non si ha la distorsione dell'immagine dovuta alla prospettiva in quanto come detto, le rette che proiettano nel piano immagine i vari punti degli oggetti nella scena tridimensionale, sono paralleli fra loro. Il modello che, come detto, viene praticamente usato è il modello di proiezione prospettica, che introduce nell'immagine la distorsione dovuta alla prospettiva.

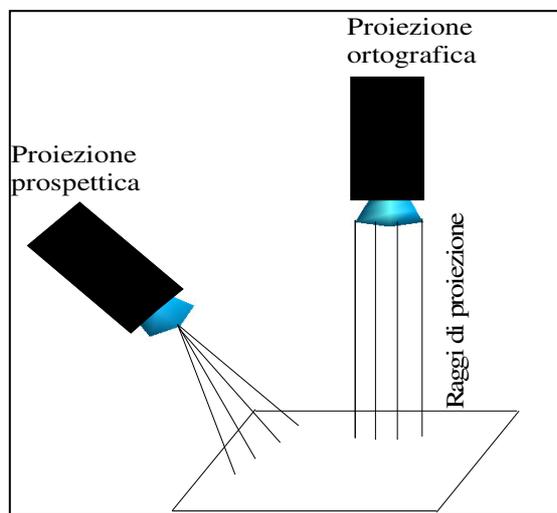


Figura 30. Esempi di 2 telecamere che realizzano i 2 tipi di modelli prospettico e ortografico.

Utilizzando il modello di proiezione ortografica possiamo ricavare vari invarianti nell'immagine inquadrata: la lunghezza dei segmenti, l'area delle figure, il

⁷ Acquisisce immagini in formato rettangolare e non retinico

parallelismo fra rette, la conservazione dell'ampiezza degli angoli; in pratica questi invarianti delineano una particolare classe di trasformazione affine quali le isometrie⁸. In un isometria viene mantenuta anche la forma della figura.

Utilizzando invece il modello di telecamera “weak perspective” si verifica la proprietà che le linee parallele nello spazio rimangono parallele anche nel piano immagine.

Nel caso di telecamera affine le quantità che si possono considerare invarianti sono:

- parallelismo tra rette
- Rapporto di lunghezze lungo rette parallele o collineari
- Rapporto delle aree

Consideriamo ora gli invarianti sotto proiezioni prospettiche. Questo tipo di invarianti sono molto importanti in quanto usati in algoritmi di geometria computazionale. Alcuni esempi di invarianti sotto questo tipo di proiezione sono:

- La concorrenza⁹
- La collinearità¹⁰
- La condizione di tangenza
- La presenza di cuspidi

ma l' invariante più importante, utilizzato soprattutto per risolvere problemi di geometria computazionale in ambito di visione artificiale, è il cross-rapporto: considerando il fascio di rette uscente dal punto O (Figura 31) si ha che:

$$\frac{AD \times BC}{BD \times AC} = \frac{ad \times bc}{bd \times ac}$$

⁸ Particolari isometrie sono le traslazioni e le rotazioni

⁹ Tre rette o segmenti sono concorrenti se passano per lo stesso punto

¹⁰ Punti collineari appartengono alla stessa retta

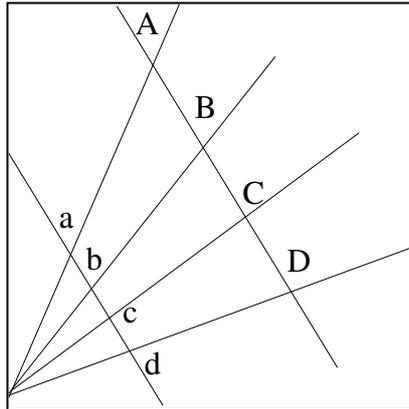


Figura 31. Cross Rapporto, invariante proiettivo non elementare

3.4 INVARIANZE GEOMETRICHE IN IMMAGINI RETINICHE

Si è visto quindi che la distorsione nelle immagini è un fenomeno intrinseco introdotto dal tipo di meccanismo visivo e utilizzando i vari modelli della telecamera questo è attenuato o incrementato, comportando quindi variazioni delle proprietà geometriche dell'immagine più o meno ampie. Nelle immagini acquisite dalla telecamera dotata del sensore retinico, una variazione della posizione degli oggetti nell'immagine, rileva una notevole variazione della geometria dovuta all'intrinseco formato log-polare con cui esse vengono memorizzate. Chiariamo ulteriormente questo aspetto molto importante. Notiamo ad esempio l'immagine sottostante in cui è stato visualizzato un pattern piano raffigurante dei cerchi concentrici, il cui centro viene a coincidere con la fovea del sensore; in seguito ad uno spostamento del pattern, ortogonale all'asse ottico della telecamera, le forme riprodotte nell'immagine log-polare cambiano rispetto all'immagine log-polare originale. Infatti in seguito allo spostamento, i cerchi non possiedono più la simmetria radiale rispetto alla fovea e i pixel che identificano i cerchi vanno ad interessare i fotorecettori del sensore più in periferia, e vengono mappati in zone diverse del piano log-polare.

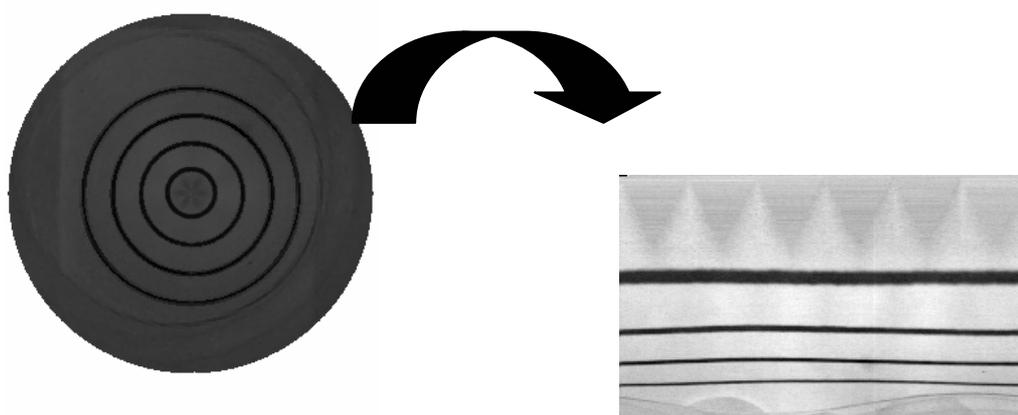


Figura 32. Pattern con cerchi concentrici e corrispondente immagine log-polare. I cerchi del pattern vengono trasformati in righe.

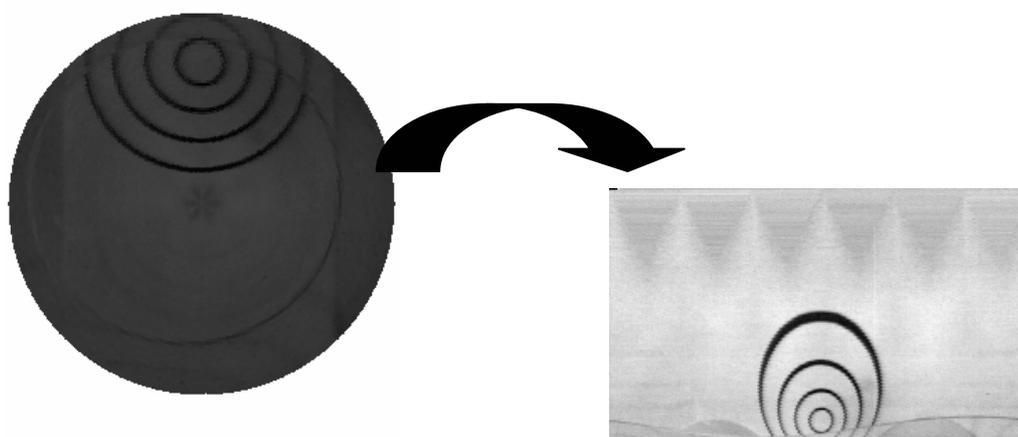


Figura 33. Pattern circolare dopo lo spostamento lungo la direzione ortogonale all'asse ottico e corrispondente immagine log-polare. Le righe della Figura 29 nell'immagine log-polare compaiono ora come delle curve.

Per ricavare gli invarianti nelle immagini retiniche considereremo sempre una telecamera inquadrante un piano ortogonale all'asse ottico della stessa e concentreremo l'attenzione prevalentemente sulla distorsione delle forme visualizzate, dovuto al fenomeno di spazio varianza intrinseco del sensore.

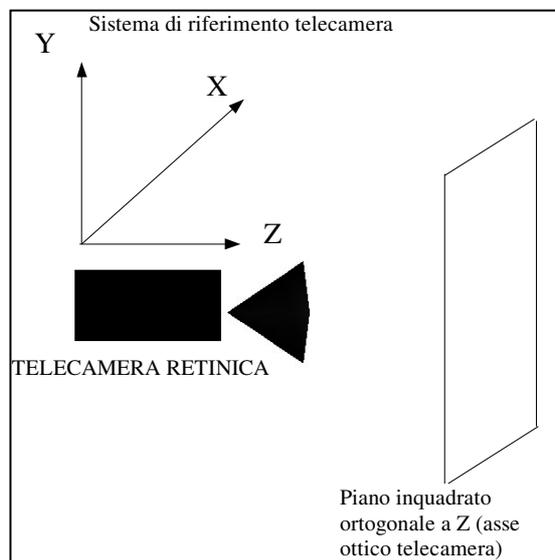


Figura 34. Modello di proiezione usato per l'acquisizione delle immagini con la telecamera retinica. L'asse ottico è ortogonale al piano inquadrato.

Utilizzando questo modello si considerano gli oggetti dell'immagine semplici forme geometriche e se ne studiano gli invarianti.

Esaminiamo inizialmente degli invarianti per semplici forme del piano log-polare che si manifestano sotto opportune condizioni. Prendiamo in esempio un segmento AP nel piano cartesiano, questo trasformato nel piano log-polare apparirà come una curva.

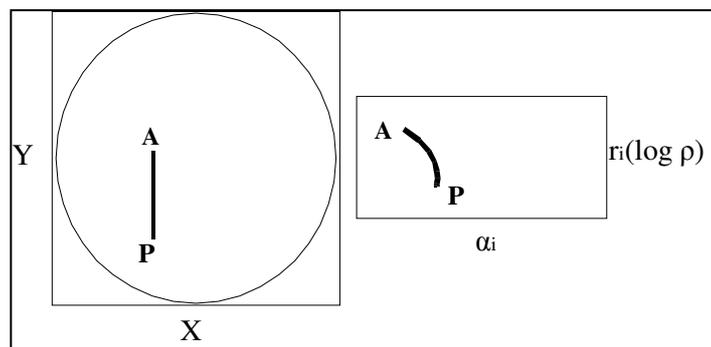


Figura 35. Il segmento AP viene trasformato nel piano log-polare in una curva

Deriviamo preliminarmente alcune considerazioni analitiche che saranno utili successivamente. Le equazioni sono costruite riferendosi al paper [14]. Queste relazioni dimostrano come si può costruire incrementalmente una curva nel piano log-polare che corrisponde ad un segmento nel piano cartesiano. Si evidenzia che due o più forme costruite in questo modo (corrispondenti a due o più segmenti nel piano cartesiano) sono uguali sotto opportune condizioni che verificheremo. Queste relazioni sono valide per pixel del sensore nella parte spazio variante e non per i pixel che identificano la fovea del sensore. Infatti come detto, nella fovea non vi è il

fenomeno di spazio varianza. Come al solito, consideriamo i due piani di lavoro Cartesiano e log-polare i cui punti di appartenenza rispettivamente sono:

$$X_c = [X, Y]' \text{ e } X_{lp} = [r, \alpha]$$

Che sono legati dalla relazione (in forma semplificata)

$$X_{lp} = [\frac{1}{2} \log_u (x^2 + y^2), \arctan (y/x)]'$$

e da quella inversa

$$X_c = U' [\cos \alpha, \sin \alpha]' = \rho [\cos \alpha, \sin \alpha]'$$

Ora consideriamo una retta nel piano cartesiano di equazione $y = mx + q$; tramite un cambio di coordinate possiamo calcolare l'espressione della retta in coordinate log polari mediante la posizione:

$$r = r_0 - \log \cos(\alpha - \alpha_0)$$

Un segmento nel piano cartesiano può essere definito tramite due punti p_1 e p_2 , o tramite un punto p , la lunghezza l e l'angolo γ (che chiameremo angolo di tangenza) come appare in modo chiaro nella Figura 36.

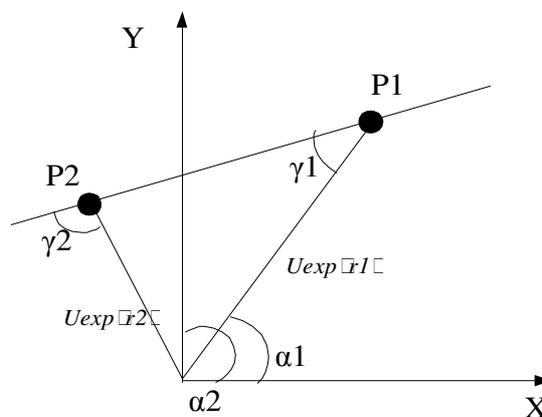


Figura 36. Segmento P1P2 nel piano cartesiano (X,Y)

Le linee con α costante nel piano log-polare corrispondono a linee radiali con angolo polare costante nel piano cartesiano. Dalla Figura si ricava la relazione tra γ e α :

$$\gamma_2 - \gamma_1 = \alpha_2 - \alpha_1 = \Delta\alpha \quad (1)$$

grazie poi al teorema di Eulero possiamo dedurre la nuova relazione:

$$\sin \gamma_2 = U^{r_2 - r_1} \sin \gamma_1 = U^{\Delta r} \sin \gamma_1 \quad (2)$$

possiamo quindi costruire un punto r_2 nel piano log-polare per un dato angolo α_2 attraverso la relazione:

$$r_2 = r_1 + \log u(\sin \gamma_2 / \sin \gamma_1) \quad (3)$$

a questo punto è noto che nell'immagine log-polare l'angolo α varia discretamente in quanto i cerchi disposti sugli anelli del sensore coprono una zona angolare determinata, come si può meglio notare dalla Figura 37.

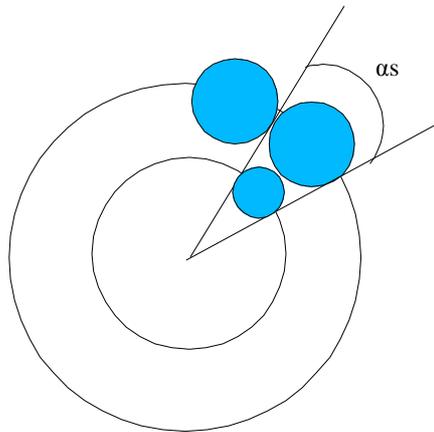


Figura 37. I due fotorecettori vicini coprono una zona angolare, i valori degli angoli nel piano log-polare sono discretizzati

Possiamo così definire un intervallo $\Delta\alpha = \alpha_s = \alpha_{i+1} - \alpha_i$. Per pixel vicini tra loro possiamo quindi riscrivere in una diversa forma la relazione (1) che diventa:

$$\gamma_{i+1} = \gamma_i + \alpha_s \quad (4)$$

mentre l'equazione (3) che rappresenta la relazione tra due punti r_i in coordinate log polari vicini tra loro diventa semplicemente:

$$r_{i+1} = r_i + \log_u(\sin(\gamma_i + \alpha_s) / \sin \gamma_i) \quad (5)$$

Con quest'ultima relazione si dimostra come costruire incrementalmente la curva nel piano log-polare partendo da un punto (r_i, α_i) . La curva corrisponde ad un segmento nel piano cartesiano il cui punto iniziale corrispondente possiede un angolo di tangenza γ_i .

Si può notare che la relazione (5) dipende soltanto dall'angolo γ_i ma non dal particolare segmento AP, infatti i parametri U (base del logaritmo) e l'intervallo α_s (risoluzione angolare) sono parametri intrinseci del sensore.

Quindi grazie a questa considerazione finale possiamo trarre questa conclusione. Consideriamo un segmento AP nel piano cartesiano per cui operiamo una traslazione radiale, l'angolo γ_i si conserva, come si verifica dalla Figura 38; quindi la curva corrispondente al segmento AP prima della traslazione sarà uguale alla curva corrispondente al segmento AP traslato., In questo caso quindi la traslazione del segmento nel piano cartesiano si traduce in una traslazione della curva nel piano log-polare che conserva la forma.

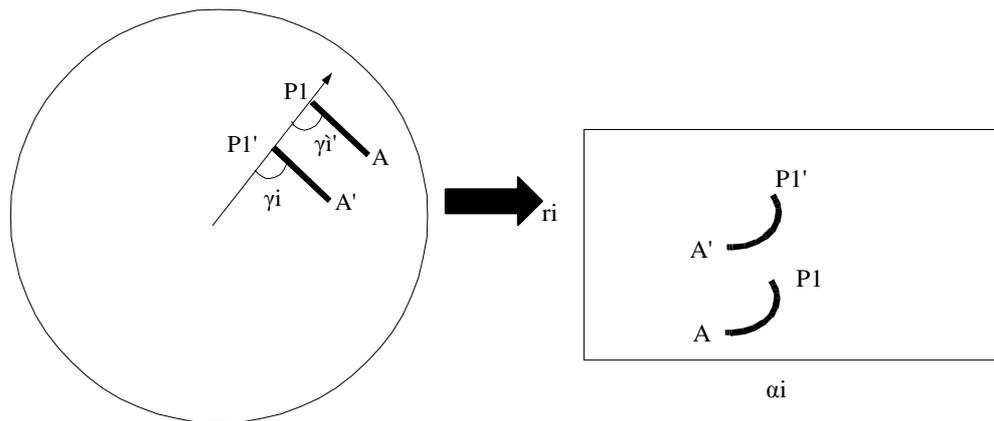


Figura 38. Una volta traslato il segmento AP1 conserva l'angolo di tangenza quindi le forme corrispondenti nel piano log-polare sono uguali e evidenziano anch'esse una traslazione.

Se operiamo una traslazione del segmento AP1 come nella Figura 39, l'angolo di tangenza γ_i si conserva nel modulo ma non nel segno. In questo caso la curva A'P'

corrispondente al segmento A'P1' in seguito alla traslazione viene anche ruotata oltre che traslata: risulta simmetrica alla curva prima della traslazione.

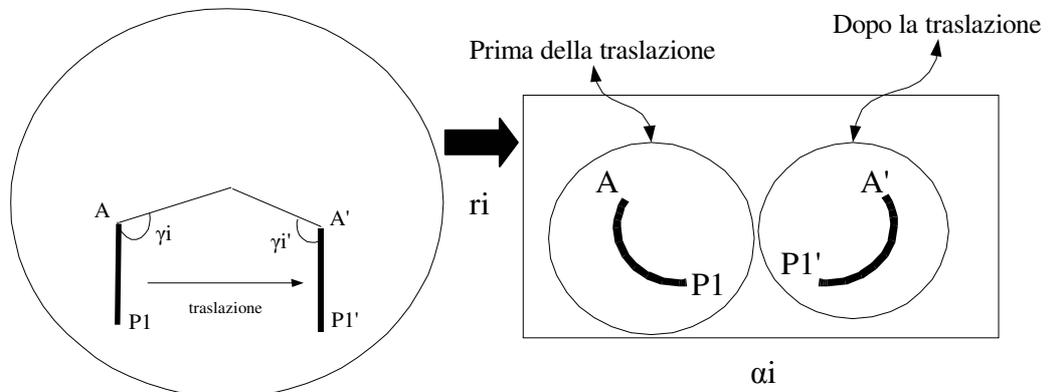


Figura 39. In questo caso vi è una semplice traslazione del segmento ma gli angoli di tangenza sono uguali in modulo e non nel segno. La curva dopo la traslazione risulta simmetrica alla curva prima della traslazione

Con la trattazione operata fino ad ora si è voluto far notare come determinate curve nell'immagine log-polare corrispondenti a segmenti in un'immagine convenzionale in seguito a opportune traslazioni conservino la forma; puntiamo ora l'attenzione su delle caratteristiche generali invarianti dell'immagine log-polare che possono essere maggiormente sfruttati in algoritmi di elaborazione di immagini acquisite tramite il sensore retinico.

Le immagini log polari presentano invarianti in seguito a scalature dell'immagine e in seguito a rotazioni della stessa sotto opportune condizioni. Questi invarianti sono verificati se e solo se la telecamera si muove lungo l'asse ottico (dando luogo alla scalatura dell'immagine) o se essa ruota sempre attorno all'asse ottico. L'asse ottico nelle acquisizioni effettuate è perpendicolare al piano su cui giace l'immagine catturata, come già evidenziato con la Figura 34.

Vediamo inizialmente gli invarianti in seguito a rotazione nell'immagine. Operando una rotazione di un oggetto che conservi la distanza dal centro dell'immagine in cui è centrato l'asse ottico della telecamera, questo corrisponde nel piano log-polare ad una semplice traslazione lungo una riga dell'oggetto stesso, che conserva la forma e la dimensione.

La Figura 40 esemplifica questo aspetto, mentre la Figura 41 verifica con un immagine reale acquisita con la telecamera retinica questo tipo di invariante.

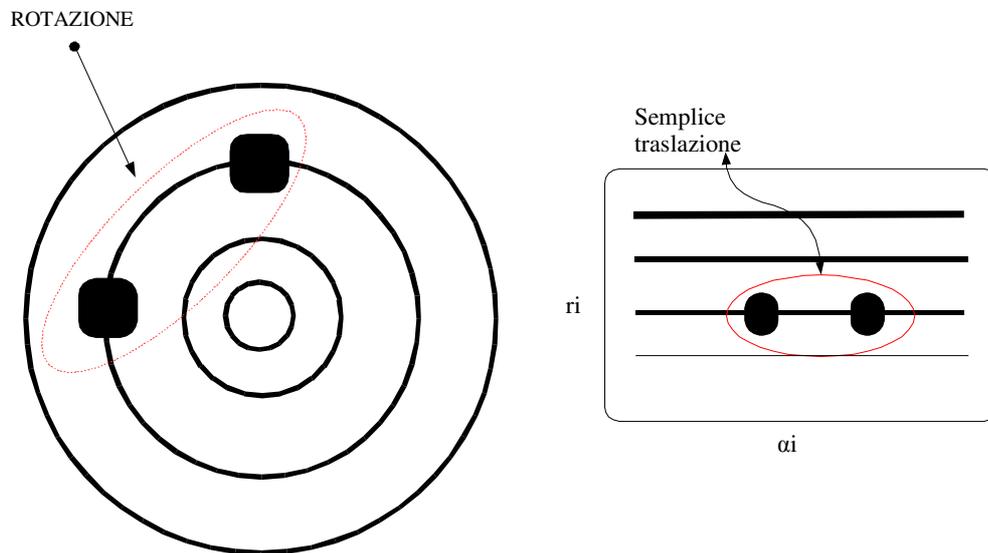


Figura 40. La rotazione dell'oggetto nel piano cartesiano si traduce in una traslazione dello stesso nell'immagine log-polare conservandone la forma e la dimensione.

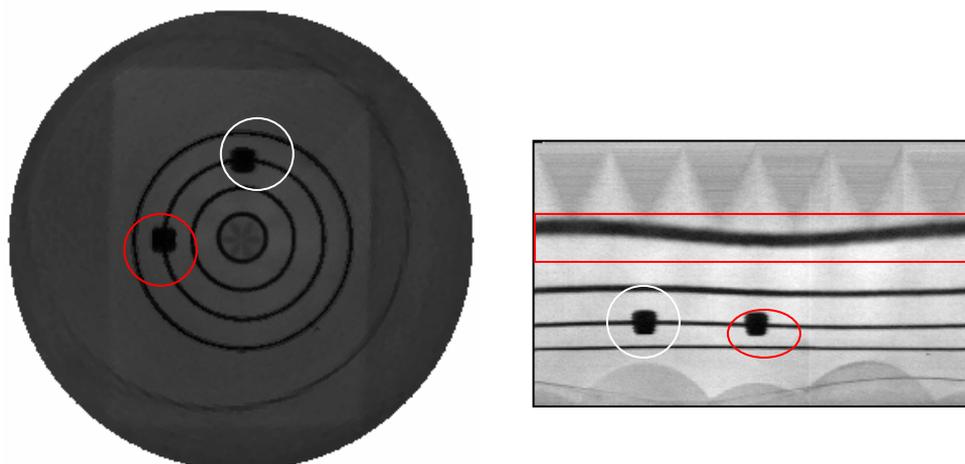


Figura 41. L'oggetto cerchiato in rosso in seguito alla rotazione trasla lungo la riga dell'immagine lo-polare.

Possiamo verificare invece dall'immagine 42 come i cerchi che vengono trasformati nelle righe dell'immagine log-polare non conservino la dimensione,

infatti come osserviamo, le prime righe dell'immagine log-polare corrispondenti ai cerchi più interni dell'immagine cartesiana sono più grosse delle ultime.

Passiamo ora agli invarianti in seguito a scalature nell'immagine. La scalatura nell'immagine può avvenire spostando la telecamera lungo l'asse ottico parallelo al piano inquadrato, o viceversa spostando il piano inquadrato verso la telecamera., sempre lungo l'asse ottico. Operando questi movimenti, consideriamo ora un oggetto visualizzato (o una forma) centrato con l'asse ottico della telecamera. La dimensione dell'oggetto nell'immagine cartesiana varia: aumenta o diminuisce a seconda che allontaniamo o avviciniamo la telecamera. Nell'immagine log-polare invece, questa scalatura si traduce in una semplice traslazione lungo le colonne, mantenendo così la dimensione dell'oggetto costante, e quindi invariante. Cambiamenti di dimensione nell'immagine cartesiana visualizzata, si traducono quindi in semplici traslazioni nella corrispondente immagine log-polare.

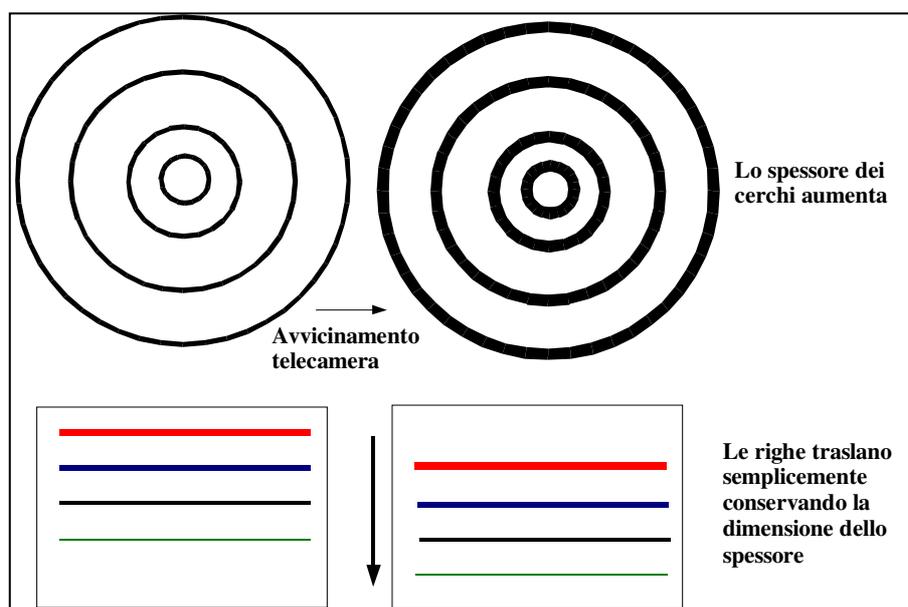


Figura 42. In seguito all'avvicinamento della telecamera, la scalatura dei cerchi si traduce in una semplice traslazione delle righe nell'immagine log-polare

Come evidenziato inizialmente perché questo invariante esista, l'oggetto o la forma visualizzata deve essere centrato (come nel caso dei cerchi) con l'asse ottico della telecamera.

Possiamo invece verificare che un oggetto che si sposta radialmente nell'immagine non presenta invarianti, possiamo solo dire che questa traslazione radiale,

nell'immagine cartesiana, si traduce in uno spostamento dell'oggetto lungo la colonna dell'immagine log-polare, mentre la forma e la dimensione dell'oggetto variano, come esemplificato nella Figura 43.

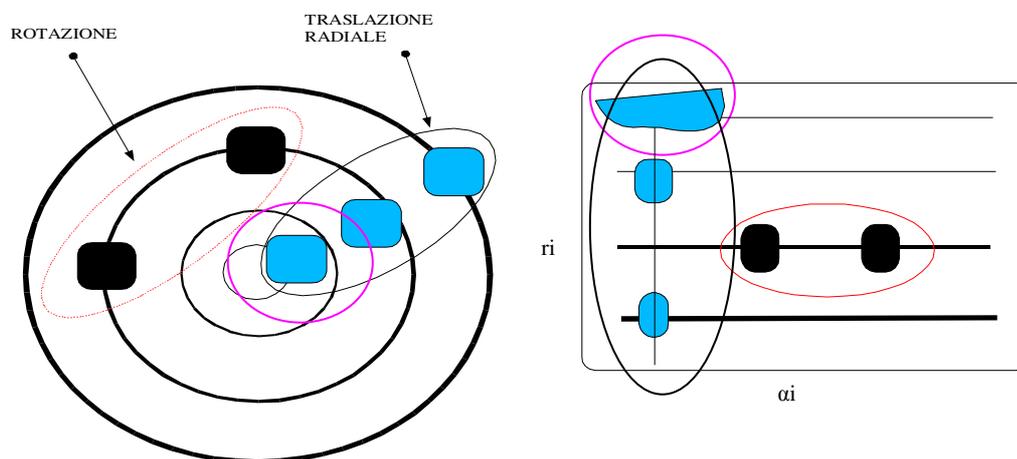


Figura 43. L'oggetto cerchiato in fucsia essendo identificato dai pixel nella fovea appare nel piano log-polare appare deformato nella forma.

Sempre notando la Figura 43 possiamo effettuare anche un'altra considerazione. L'oggetto che giace nel cerchio più interno dell'immagine cartesiana gode di maggiore risoluzione mentre quando viene mappato nell'immagine log-polare i pixel che lo identificano vengono amplificati nella dimensione (come specificato nel Capitolo1) e questo effetto provoca un aumento nella grandezza dell'oggetto in questione oltre che una deformazione. La situazione degenera quando l'oggetto visualizzato è nella fovea del sensore e in questo caso mappato nelle primissime righe del piano log-polare appare totalmente sformato.

3.5 INVARIANZE GEOMETRICHE IN IMMAGINI OMNIDIREZIONALI RETINICHE

L'acquisizione delle immagini retiniche omnidirezionali è effettuata tramite la telecamera retinica abbinata ad uno specchio il cui asse di simmetria coincide con l'asse ottico della telecamera. Sfruttando lo specchio è possibile avere un campo visivo di 360° avendo quindi una visione della scena più ampia. Consideriamo

innanzitutto un immagine omnidirezionale retinica acquisita visualizzando un pattern disposto su di un piano ortogonale all'asse ottico della telecamera e quindi all'asse di simmetria dello specchio. Come precisato nella Figura numero 44.

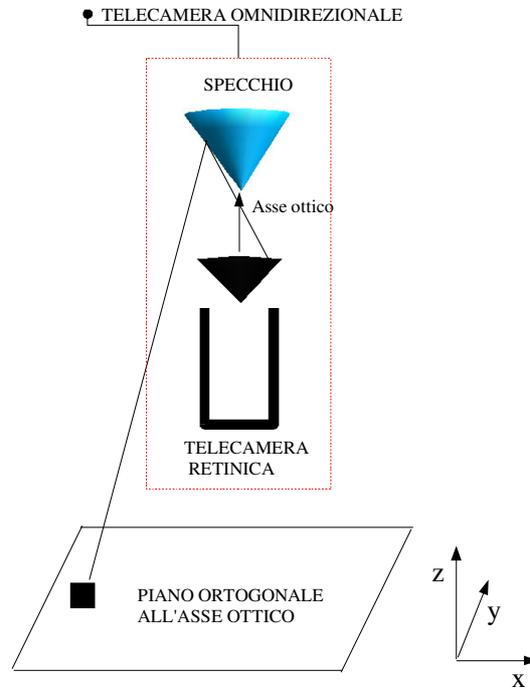


Figura 44. Complesso con telecamera omnidirezionale e piano ortogonale all'asse ottico visualizzato

Considerando questo tipo di immagini log polari si osserva costruendo dei pattern opportuni un comportamento simile alle immagini retiniche (non omnidirezionali) log polari. Infatti prendendo come esempio il pattern già considerato dei cerchi concentrici, si può notare che se i cerchi sono centrati sull'asse ottico della telecamera, questi sono ancora delle righe parallele nell'immagine log-polare omnidirezionale come era per le immagini log polari prospettiche con lo stesso pattern. Anche in questo caso quindi, esaminando un oggetto che compie una rotazione lungo un cerchio, corrisponderà una traslazione di esso lungo una riga dell'immagine omnidirezionale log-polare, conservando la forma e la dimensione dell'oggetto. L'oggetto che trasla radialmente invece, non possiede questo invariante, nella traslazione lungo la colonna dell'immagine log-polare, come è stato precedentemente evidenziato con le immagini log polari non omnidirezionali.

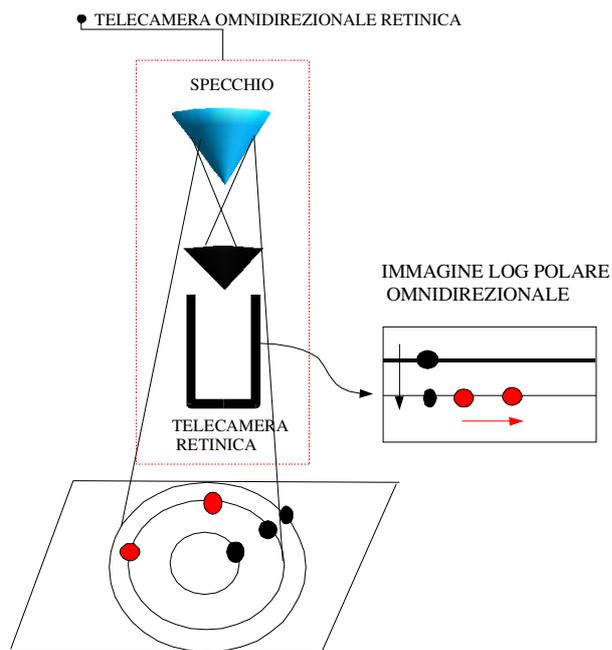


Figura 45. L'immagine omnidirezionale retinica in questo caso ha lo stesso comportamento dell'immagine retinica. Un movimento radiale o angolare si traduce in una traslazione dell'oggetto lungo la colonna o la riga. Anche in questo caso per la rotazione dell'oggetto rimane invariante la forma e la dimensione mentre per la traslazione radiale questo invariante non si verifica.

Più in generale si può dire che l'immagine omnidirezionale log-polare della scena attorno alla telecamera compare già "srotolata". Puntualizziamo meglio questo aspetto. Un immagine omnidirezionale acquisita con un sensore tradizionale appare come nella Figura 46.

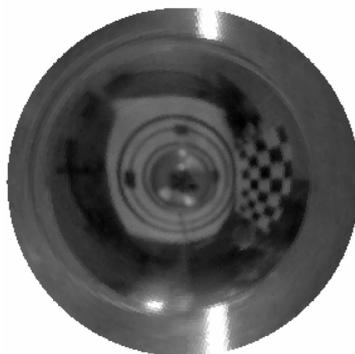


Figura 46. L'immagine omnidirezionale acquisita tramite un visore omnidirezionale con sensore tradizionale.

Per rendere la scena piu' interpretabile l'immagine necessita di un ulteriore elaborazione che la renda "srotolata" e quindi rettangolare, il cosiddetto "unwarping" dell'immagine omnidirezionale.

La stessa immagine acquisita tramite il visore omnidirezionale dotato di sensore retinico invece appare già "unwarp" così come evidenziato dalla Figura 47.

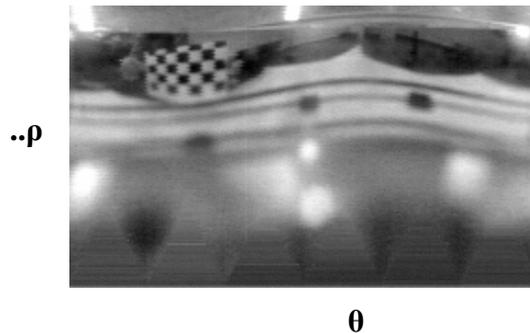


Figura 47. L'immagine omnidirezionale log-polare appare già "unwarp" e non necessita di ulteriore elaborazione.

Questo risultato è un ovvia conseguenza del fatto che l'immagine acquisita dal sensore retinico subisce già in fase di mappaggio dei pixel nel piano log-polare un naturale "unwarping" visto che l'angolo θ corrisponde all'azimuth mentre con il valore di ρ si identifica la distanza dal centro della telecamera. Questo quindi è un ulteriore vantaggio introdotto dall'uso del sensore retinico nel visore omnidirezionale. Ovviamente l'immagine omnidirezionale log-polare appare con una risoluzione minore rispetto alla corrispondente immagine non omnidirezionale dato il maggior campo visivo della telecamera.

I risultati ottenuti fino ad ora analiticamente saranno verificati con l'acquisizione di immagini reali, utilizzando opportuni pattern creati, nel paragrafo successivo.

3.6 PROSPETTO RIASSUNTIVO DELLE TRASFORMAZIONI LOG POLARI

Si è visto quindi che l'immagine log-polare prodotta dal sensore retinico e la stessa immagine omnidirezionale retinica presentano delle caratteristiche invarianti, che possono essere sfruttate in sede di sviluppo di particolari algoritmi di elaborazione di immagini per facilitarne l'esecuzione o per diminuirne la

complessità. Gli invarianti maggiormente utilizzabili per questo scopo rimangono le invarianze in seguito a scalature o rotazioni dell'immagine.

Qui di seguito viene presentato uno schema che riproduce le immagini cartesiane risultanti dalla visualizzazione di opportuni pattern e le corrispondenti trasformazioni log polari, inoltre vengono presentate anche immagini omnidirezionali log polari con le corrispondenti immagini omnidirezionali cartesiane.

L'acquisizione delle immagini è stata effettuata all'interno del laboratorio tramite la telecamera retinica Giotto gentilmente concessa allo IAS LAB dal prof. Giulio Sandini del dipartimento di Ingegneria Informatica di Genova. Per l'acquisizione delle immagini retiniche è stato utilizzato un cavalletto su cui poggiare la telecamera. Sono stati stampati vari pattern su fogli A4 e disposti su di un piano ortogonale all'asse ottico della telecamera e anche paralleli all'asse ottico.

Sono state effettuate due tipi di acquisizione con due tipi di parametri di distanza utilizzati che sono riportati qui di seguito.

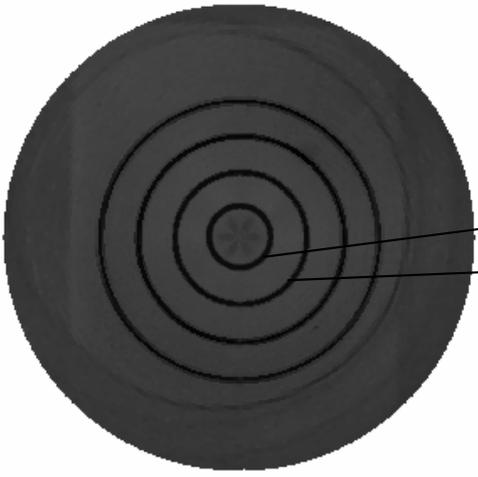
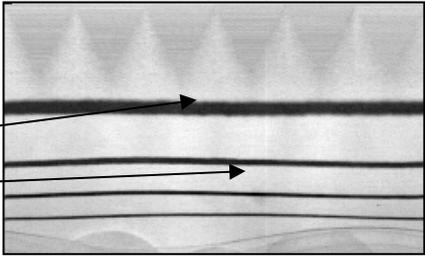
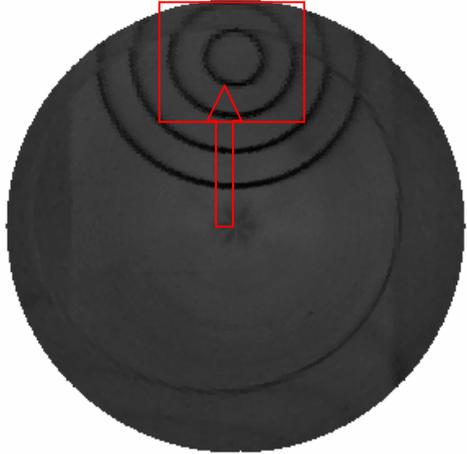
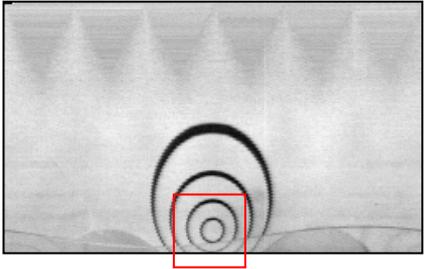
1° ACQUISIZIONE	2° ACQUISIZIONE
Distanza dall'obiettivo 21 cm	Distanza dall'obiettivo 32 cm
Altezza telecamera da terra 77.5 cm	Altezza telecamera da terra 77.5 cm

Le immagini poi sono state salvate in formato bitmap con l'uso del software di acquisizione appositamente creato.

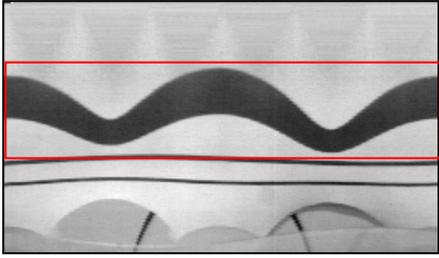
Per l'acquisizione delle immagini omnidirezionali retiniche sono stati stampati dei pattern su dei fogli molti più grandi di dimensione 75 cm * 150 cm; disposti su di un piano ortogonale all'asse ottico della telecamera e all'asse di simmetria dello specchio. I parametri di distanza vengono riportati di seguito.

SPECCHIO UTILIZZATO	Specchio iperbolico, raggio 3 cm
Distanza specchio- obiettivo	7 cm
Altezza telecamera da terra	18 cm
Raggio del primo cerchio	17 cm
Raggio del secondo cerchio	28 cm
Raggio del terzo cerchio	38 cm

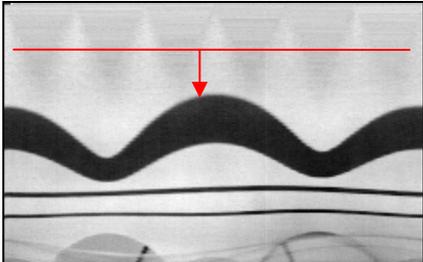
RISULTATI DELLE ACQUISIZIONI IMMAGINI RETINICHE

RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Il pattern di base è composto da cerchi concentrici, il centro di esse è in foveazione ovvero la fovea del sensore coincide col centro dei cerchi. Si può notare che i cerchi vengono mappati sulle righe dell'immagine log-polare. Inoltre i cerchi dell'immagine cartesiana hanno lo stesso spessore mentre quando sono mappati nell'immagine log-polare come righe presentano spessori crescenti dalla prima riga verso l'ultima in quanto passano da una zona dell'immagine a minor risoluzione ad una zona a maggiore risoluzione.</p>	
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	

In questo caso il pattern precedente è stato spostato di 4 cm in alto rispetto alla fovea (ortogonalmente all'asse ottico) quindi non essendovi la simmetria radiale precedente l'immagine log-polare ora appare diversa: non c'è più la corrispondenza cerchi- righe, ora l'immagine log polare appare come una serie di semicerchi e possiamo notare che solo i due più interni(evidenziati) si chiudono a cerchio.

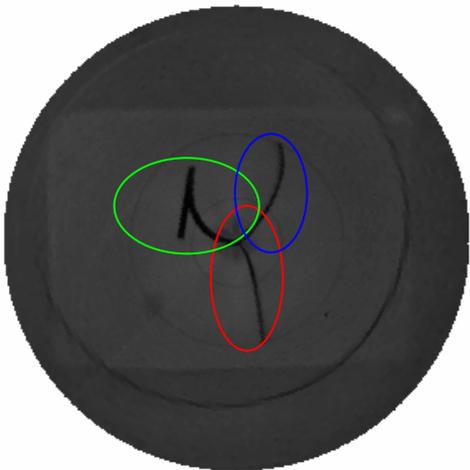
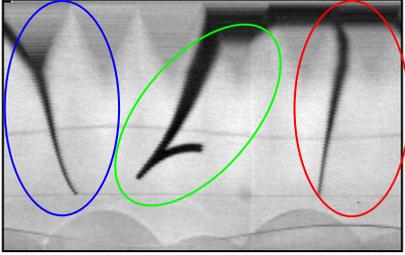
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	

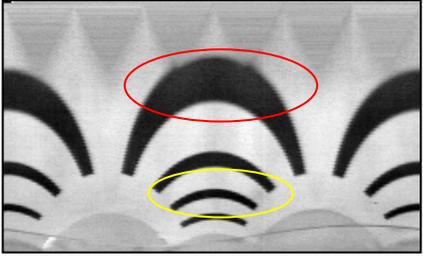
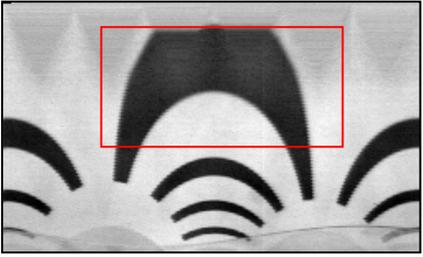
Il pattern è composto sempre da cerchi concentrici e da un'ellisse centrata con l'asse ottico della telecamera. La zona evidenziata nel rettangolo rosso è l'ellisse mappata nell'immagine log-polare.

RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	

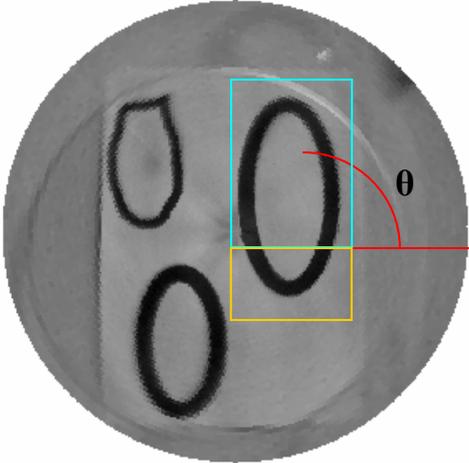
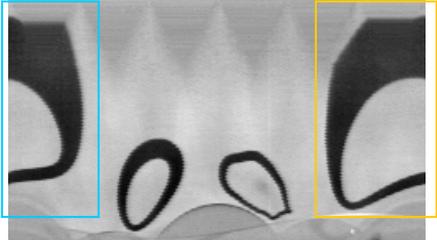
In questo caso il pattern è lo stesso del precedente ma la telecamera è stata spostata in avanti di 15 cm. Sia la forma nell'immagine log-polare rappresentante l'ellisse cartesiana, che le due righe rappresentanti i due cerchi concentrici, sono traslati verso

il basso. La linea rossa evidenzia l'entità della traslazione. In seguito alla scalatura dell'immagine cartesiana (in questo caso un ingrandimento), che quindi aumenta lo spessore dei cerchi, nell'immagine log-polare le righe corrispondenti ai cerchi non aumentano nella dimensione ma traslano semplicemente, questo grazie alla proprietà di invarianza in scalature per le immagini log polari.

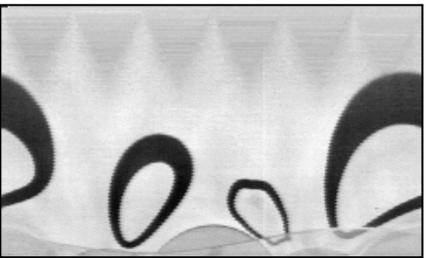
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Esempio di mapping log-polare di alcune traiettorie curve che si dipartono dal centro. Non essendovi nessun tipo particolare di simmetria le linee curve vengono mappate ancora in linee curve nell'immagine log-polare. Visto che le curve nell'immagine cartesiana partono dal centro dirette verso la periferia dell'immagine, nella corrispondente immagine log-polare queste partono dalle prime righe (identificate da valori bassi del raggio) e si dirigono verso la parte bassa dell'immagine log-polare (identificata da alti valori del raggio).</p>	

RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>In questo pattern vengono evidenziate delle semplici righe parallele tra loro; nell'immagine log-polare queste appaiono come semicurve di dimensione variabile a seconda della zona (a minor o maggior risoluzione) che occupano. Ad esempio le righe centrali dell'immagine vengono maggiormente deformate nell'immagine log polare, poiché sono mappate nelle prime righe che presentano minore risoluzione.</p>	
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Tra questo pattern e il precedente vi è stato un' avvicinamento della telecamera di 10 cm ed uno spostamento dell'asse ottico. In questo caso possiamo notare che la riga</p>	

centrale essendo visualizzata in parte dalla fovea provoca un mapping nell'immagine log polare di una curva molto più deformata rispetto al pattern precedente.

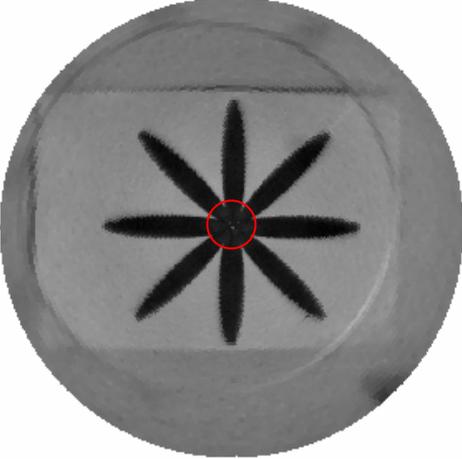
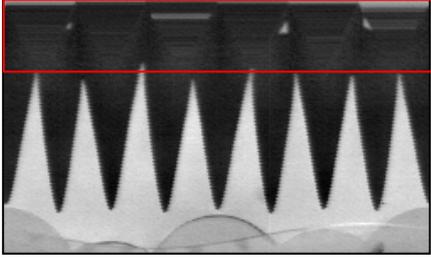
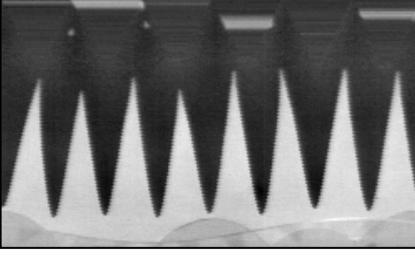
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	 <p data-bbox="935 835 1018 902">$0 < \theta < 90$ gradi</p> <p data-bbox="1246 835 1358 902">θ prossimi a 360°</p>

Il pattern evidenzia tre linee chiuse spostate dalla fovea e di dimensione diversa. Possiamo notare come la linea chiusa più grande venga spezzata nella rappresentazione log-polare in due parti. Essa infatti attraversa la linea evidenziata in rosso dove parte la misura dell'angolo θ (rappresentante le colonne nell'immagine log-polare). Quindi una parte di essa verrà mappata nell'immagine log-polare nelle prime colonne a cui corrispondono gli angoli da 0 a 90 gradi, mentre l'altra parte verrà mappata nelle ultime colonne a cui corrispondono gli angoli prossimi a 360° .

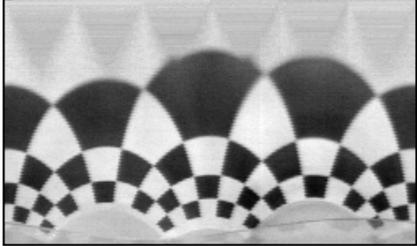
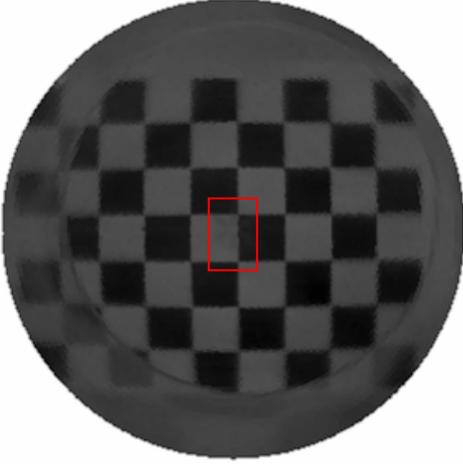
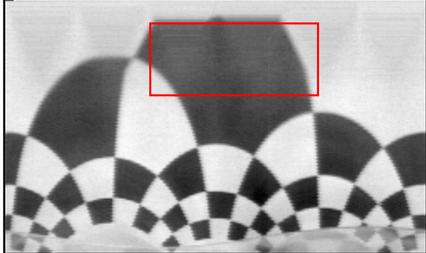
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	

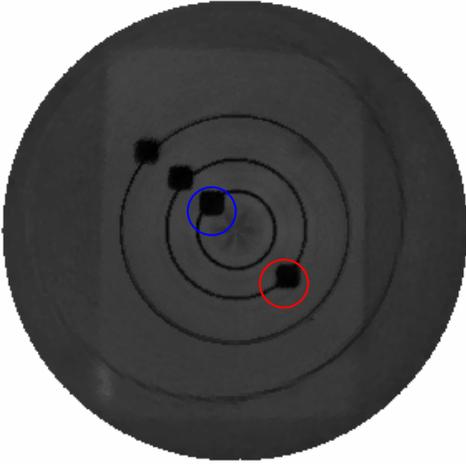
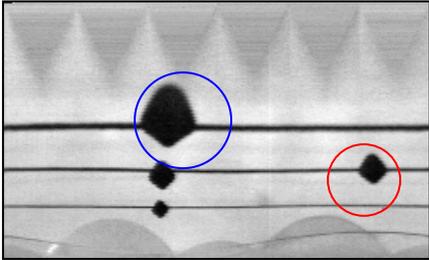
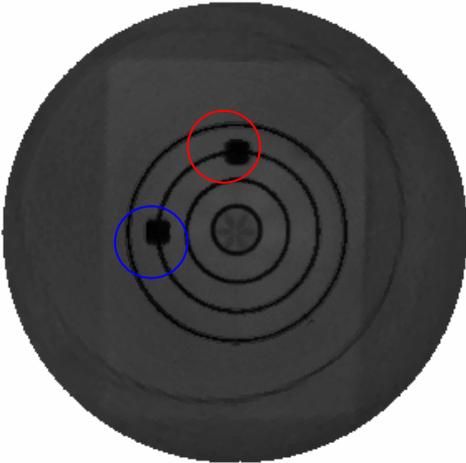
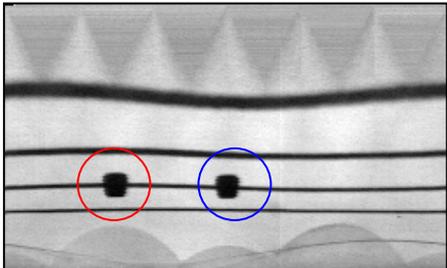
La telecamera è stata spostata in avanti verso il pattern di 10 cm, in questo caso però gli oggetti non sono centrati con l'asse ottico della telecamera e quindi la proprietà di

invarianza alle scalature dell'immagine non può essere verificata.

RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>I petali del fiore vengono mappati lungo le colonne dell'immagine log-polare essendovi la simmetria radiale. La parte centrale del pattern cerchiato essendo in foveazione (centrata nella fovea) viene mappata lungo le prime righe dell'immagine log-polare evidenziate. Si può notare che i petali che nell'immagine cartesiana sono formati da linee curve per effetto del mapping log-polare vengono raddrizzati nell'immagine log-polare apparendo come triangoli.</p>	
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Anche per la visualizzazione di questo pattern vi è stato lo spostamento della</p>	

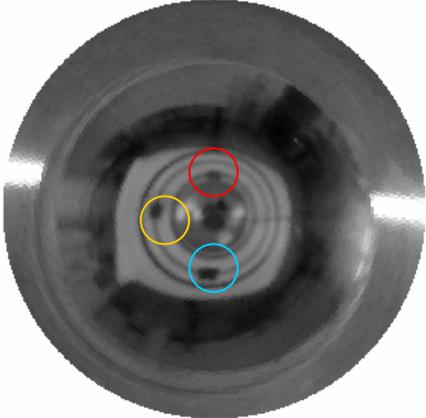
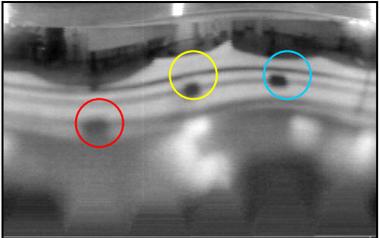
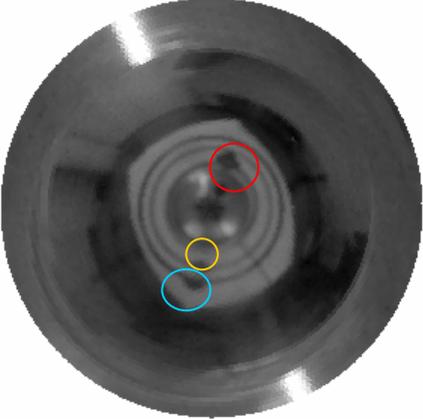
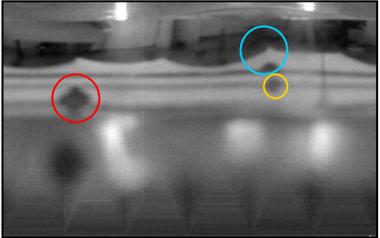
telecamera di 10 cm.

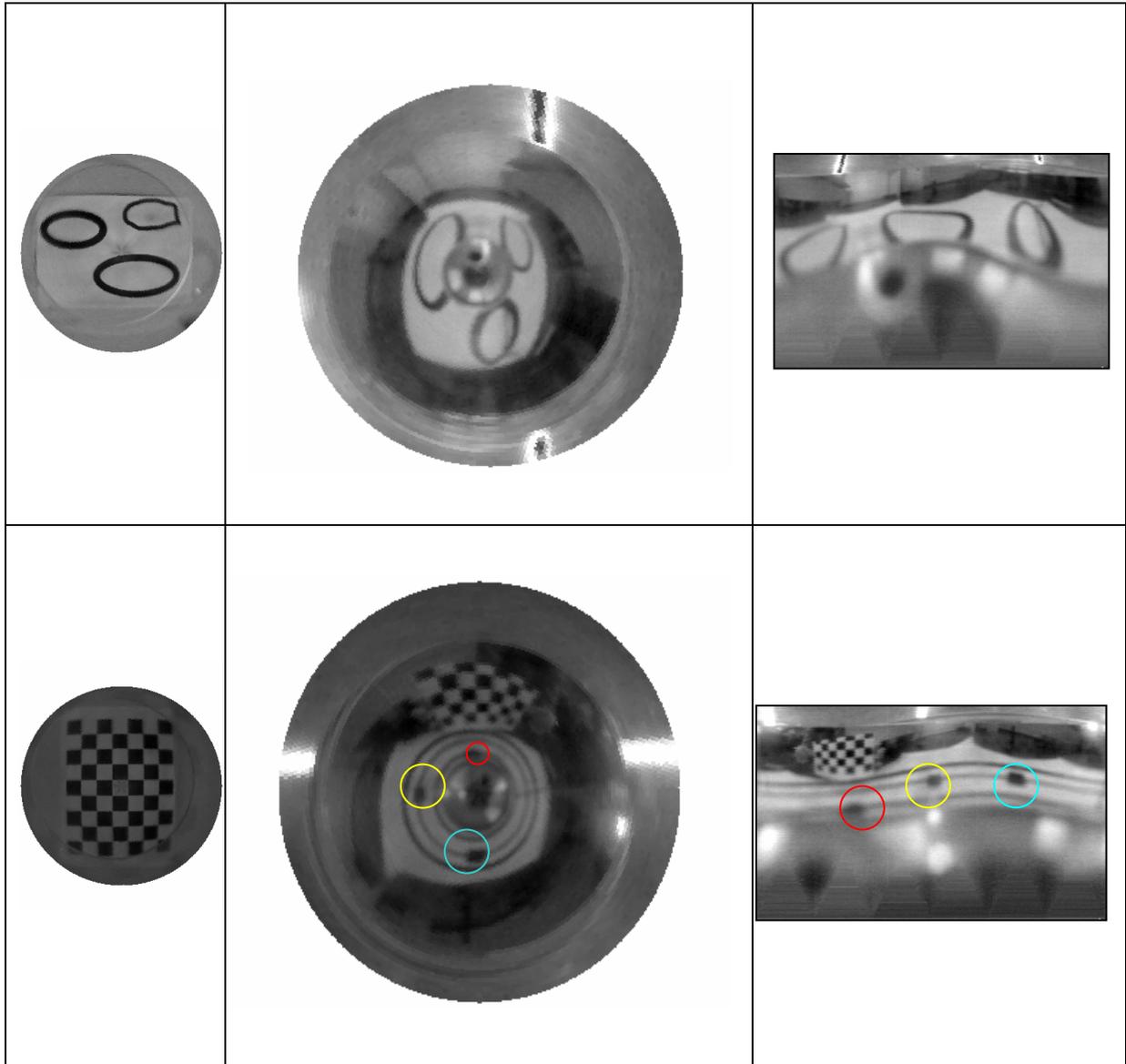
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>I quadrati centrali della scacchiera nell'immagine catesiana vengono deformati maggiormente nell'immagine log-polare, perché si trovano in una zona a maggiore risoluzione dell'immagine cartesiana. Essi vengono mappati nelle prime righe dell'immagine log-polare che presentano invece una risoluzione minore.</p>	
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Telecamera spostata in avanti di 10 cm. Il quadrato centrale viene molto deformato, in quanto parte di esso viene visualizzato dalla fovea del sensore e quindi viene mappato nelle prime righe dell'immagine log polare dove, come al solito, si ha una deformazione maggiore.</p>	

RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Nel primo pattern viene evidenziato un movimento radiale di un oggetto passante per la fovea, anche in questo caso l'oggetto posto nel cerchio più interno dell'immagine cartesiana appare più grande nell'immagine log-polare, si evidenzia una traslazione dell'oggetto lungo le righe dell'immagine log polare, ma né la dimensione né la forma si conservano. Per l'oggetto cerchiato in rosso si ha invece uno spostamento di colonna rispetto agli altri tre oggetti in quanto esso ha un angolo θ maggiore.</p>	
RAPPRESENTAZIONE CARTESIANA	RAPPRESENTAZIONE LOG-POLARE
	
<p>Il secondo pattern evidenzia lo spostamento angolare lungo un cerchio, mappato nello spostamento lungo una riga nell'immagine log-polare. In questo caso l'oggetto</p>	

conserva la forma e la dimensione grazie alla proprietà di invarianza alle rotazioni delle immagini log polari

IMMAGINI OMNIDIREZIONALI RETINICHE

PATTERN PIANO	OMNIDIREZIONALE CARTESIANA	OMNIDIREZIONALE LOG-POLARE
		
		



I pattern visualizzati nelle prime colonne sono stati posti in un piano ortogonale all'asse ottico della telecamera con centro di simmetria coincidente con l'asse ottico. In questo caso le immagini log polari omnidirezionali presentano la medesima struttura di quelle log-polari retiniche.

Nell'ultimo pattern il piano è disposto parallelamente all'asse ottico della telecamera e compare nell'immagine log-polare con la sua forma originale anche se con minore risoluzione.

Importante è notare che nelle immagini log polari omnidirezionali acquisite più di un terzo delle righe sono occupate dall'immagine riflessa nello specchio della telecamera stessa poiché esso ha vertice sferico.

In questo Capitolo si sono evidenziati gli invarianti geometrici presenti nelle immagini prospettiche, retiniche ed omnidirezionali retiniche. Inoltre è stata riportata una serie di immagini acquisite con la telecamera retinica per verificare gli invarianti nelle immagini log polari e per evidenziare altre proprietà strutturali delle immagini log-polari. Si è visto in particolare che nelle immagini prospettiche l'invariante più importante è il cross rapporto, invariante fondamentale nella "computer vision", in quanto questo invariante viene molto utilizzato in algoritmi di visione computazionale usando telecamere sfruttanti i sensori tradizionali. Per le immagini log polari si sono evidenziati gli invarianti a scalature e a rotazioni che uniti al pregio della dimensione ridotta dell'immagine log polare possono costituire un potente supporto per numerose applicazioni nell'ambito della visione computazionale. Si è precisato poi che le immagini omnidirezionali retiniche sotto particolari condizioni, hanno la medesima struttura di quelle log-polari. Offrendo un campo visivo maggiore e non necessitando di successive elaborazioni per l'"unwarping", le immagini omnidirezionali retiniche costituiscono un grande vantaggio nell'uso della telecamera omnidirezionale retinica oggetto di questa tesi.

4. EDGE DETECTION SU IMMAGINI LOG-POLARI

4.1 INTRODUZIONE

L'edge detection di un immagine è il processo di individuazione delle linee di bordo o contorni all'interno di essa. Questo tipo di elaborazione è molto importante in quanto permette di giungere al rilevamento di zone dell'immagine che possono essere utili in svariate applicazioni. Un esempio è la navigazione di robot mobili. In questa applicazione un robot dotato di telecamera deve muoversi all'interno di un ambiente, evitare ostacoli e riconoscere determinati oggetti. La rilevazione dei contorni delle immagini in questo particolare caso può essere utile. L'edge detection è stata maggiormente studiata in modo approfondito per quanto riguarda immagini in toni di grigio, come le immagini acquisite dalla telecamera retinica in nostra dotazione.

L'essenza degli algoritmi di edge detection risiede nella individuazione delle discontinuità esistenti nell'immagine tra i valori dei pixel, che in taluni casi possono evidenziare dei bordi nell'immagine elaborata.

Esistono in letteratura tre tipologie di strumenti analitici che conducono al rilevamento dei bordi: gli operatori differenziali, il calcolo statistico e l'analisi in frequenza. Gli algoritmi più utilizzati appartengono alla prima classe, gli altri due si basano rispettivamente sulla ricerca dei punti ad alta varianza ed ad alta frequenza.

L'algoritmo più noto e diffuso nel campo dell' "image processing" in immagini mono canale è sicuramente l'algoritmo di Canny [15]. L'idea base di Canny [15] è quella di abbinare alla fase della differenziazione dell'immagine, un'operazione di smoothing al fine di rilevare le brusche variazioni di luminosità dei pixel. Lo smoothing è ottenuto attraverso la convoluzione con un filtro gaussiano. Il filtro gaussiano serve per eliminare la presenza di minimi locali nella funzione di intensità dei pixel. Definendo le funzioni gaussiane e le corrispondenti derivate prime con le seguenti posizioni:

$$G_{\sigma}(x) = (1 / \sqrt{2\pi\sigma}) \times e^{-x^2/2\sigma^2} \longrightarrow G'_{\sigma, x} = -x / \sqrt{2\pi\sigma^3} \times e^{-x^2/2\sigma^2}$$

$$G_{\sigma}(y) = (1 / \sqrt{2\pi\sigma}) \times e^{-y^2/2\sigma^2} \longrightarrow G'_{\sigma, y} = -y / \sqrt{2\pi\sigma^3} \times e^{-y^2/2\sigma^2}$$

il primo passo dell'algoritmo è quello di compiere la convoluzione dell'immagine $I(x,y)$ con i filtri:

$$f_v(x, y) = G'_{\sigma, x} G_{\sigma}(y)$$

$$f_h(x, y) = G'_{\sigma, y} G_{\sigma}(x)$$

$$R_v(x, y) = I(x, y) * f_v(x, y)$$

$$R_h(x, y) = I(x, y) * f_h(x, y)$$

si determina poi il valore assoluto del gradiente dell'immagine così filtrata:

$$| R(x, y) | = | R_v(x, y)^2 + R_h(x, y)^2 |$$

I passi finali dell'algoritmo sono i cosiddetti “non maxima suppression” e “hysteresis thresholding”. La fase “non maxima suppression” consiste nell'assottigliamento dei bordi fino a renderli spessi un solo pixel, mediante l'eliminazione di picchi situati nell'intorno dei massimi locali. In particolare si scandisce l'immagine pixel per pixel e si analizza il valore del modulo del gradiente (precedentemente calcolato) per ogni pixel. Se il modulo del gradiente è nullo, il pixel è marcato come non appartenente al bordo; se invece non è nullo, si determina la direzione ortogonale al contorno, e per i pixel adiacenti che sono situati lungo questa direzione si verifica che essi non abbiano il modulo del gradiente superiore. Se ciò avviene il pixel è marcato come non appartenente al bordo, se invece i moduli dei gradienti calcolati per i pixel adiacenti sono inferiori, allora il pixel viene marcato come edge.

L'ultima fase è come detto la fase di “hysteresis thresholding”, o sogliatura ad isteresi. Infatti dopo la fase di “non maximal suppression” abbiamo ancora la presenza di alcuni massimi locali creati dal rumore. Operando però una sogliatura con un singolo valore di soglia si presentano tali problemi. Se prendiamo una soglia troppo bassa, in modo da prendere i bordi veri, ma molto deboli, anche i massimi

locali rumorosi vengono selezionati. Se i valori dei massimi locali autentici lungo un contorno fluttuano al di sotto e al di sopra di tale soglia, avremo come risultato un bordo spezzettato. Ecco perché la semplice sogliatura non è una buona soluzione, mentre la soluzione ideale è quella di fissare due soglie T_h e T_l .

Il risultato dopo la sogliatura ad isteresi è un'immagine con tutti i contorni spessi un pixel e ripuliti dal rumore.

4.2. SPACE VARIANT EDGE DETECTION

Come visto nel Capitolo 1, è stata implementata una funzione all'interno del software di visualizzazione delle immagini log polari, che opera un edge detection basato su l'algoritmo di Canny. L'edge detection viene implementato sia per le immagini log-polari sia per le immagini cartesiane. Per queste ultime verrà implementato l'edge detection spazio variante, che ora analizzeremo

Utilizzando il filtro di Canny con una gaussiana che ha una larghezza predefinita, vengono rilevati solo gradini (discontinuità tra i pixel) con una certa pendenza, però poiché l'immagine cartesiana presenta una risoluzione spazio variante essa presenterà degli edge che hanno pendenze diverse. Come al solito l'aumento di risoluzione si ha dalla periferia al centro dell'immagine dove la risoluzione è maggiore; utilizzando Canny con una varianza (σ) stretta, è nella zona centrale dell'immagine dove sarà individuato un numero maggiore di Edge, mentre nella periferia dell'immagini può avvenire che l'individuazione degli edge fallisca. Aumentando la varianza si aumenta la larghezza della gaussiana e quindi la dimensione del filtro: con un filtro più ampio si individueranno variazioni graduali dell'intensità su un'area più vasta e perciò viene aumentata la raffinatezza di ricerca degli edge; viceversa un filtro con dimensione minore, e quindi con varianza minore, provvederà a rilevare molte piccole variazioni in dettaglio e quindi la probabilità di individuare gli edge sarà maggiore.

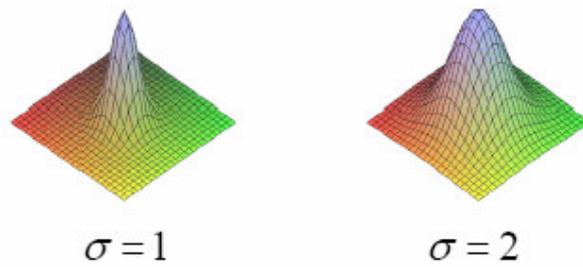


Figura 48. Immagine di Gaussiane tridimensionali con due valori di varianza

Diminuendo la varianza però ovvieremo al problema di ricerca degli edge nella periferia ma nel centro dell'immagine vi è la possibilità di rilevare edge indesiderati.

Quindi a causa di queste considerazioni la soluzione ideale è stata quella di creare un filtro variabile, che si adatti alla zona di immagine considerata variandone opportunamente la varianza. La parte centrale dell'immagine verrà convoluta con filtro determinato da un particolare valore di sigma per poi aumentare linearmente verso la periferia dove la varianza assumerà il valore massimo. L'ideale sarebbe applicare ad ogni anello una varianza diversa visto che la risoluzione varia al variare degli anelli. In pratica la dimensione del filtro (e quindi la varianza) non viene variata per ogni anello, ma l'immagine viene divisa in 21 zone come evidenziato dalla Figura 50, (approssimando gli anelli nell'immagine come cornici rettangolari), partendo da un valore della varianza (σ) uguale a 1.1 per la parte centrale dell'immagine e aumentando sigma nel passaggio da una zona alla successiva di un fattore K uguale a 0.05.

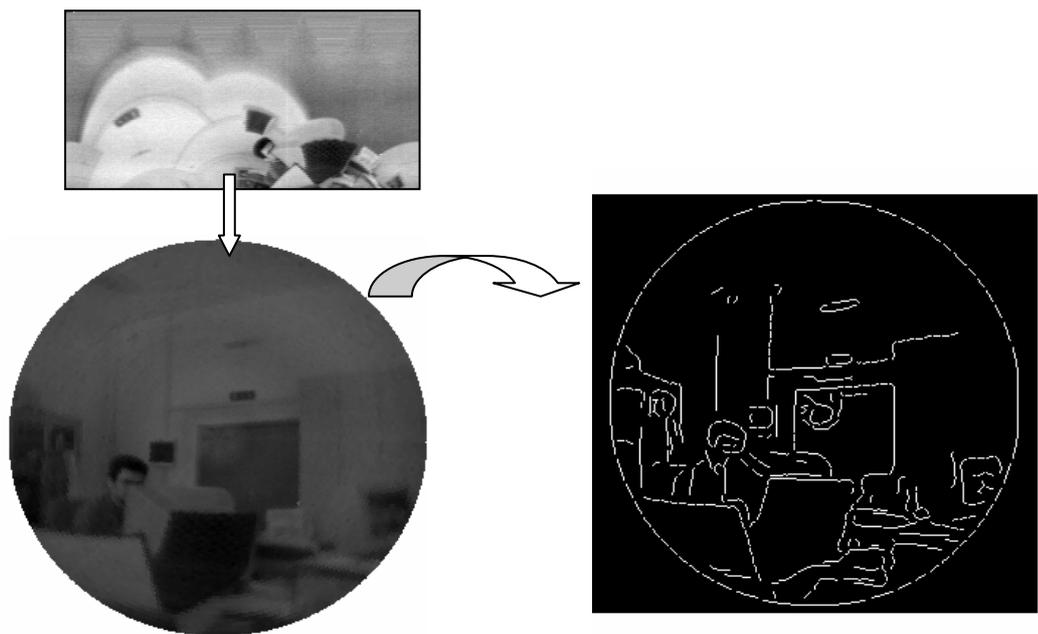
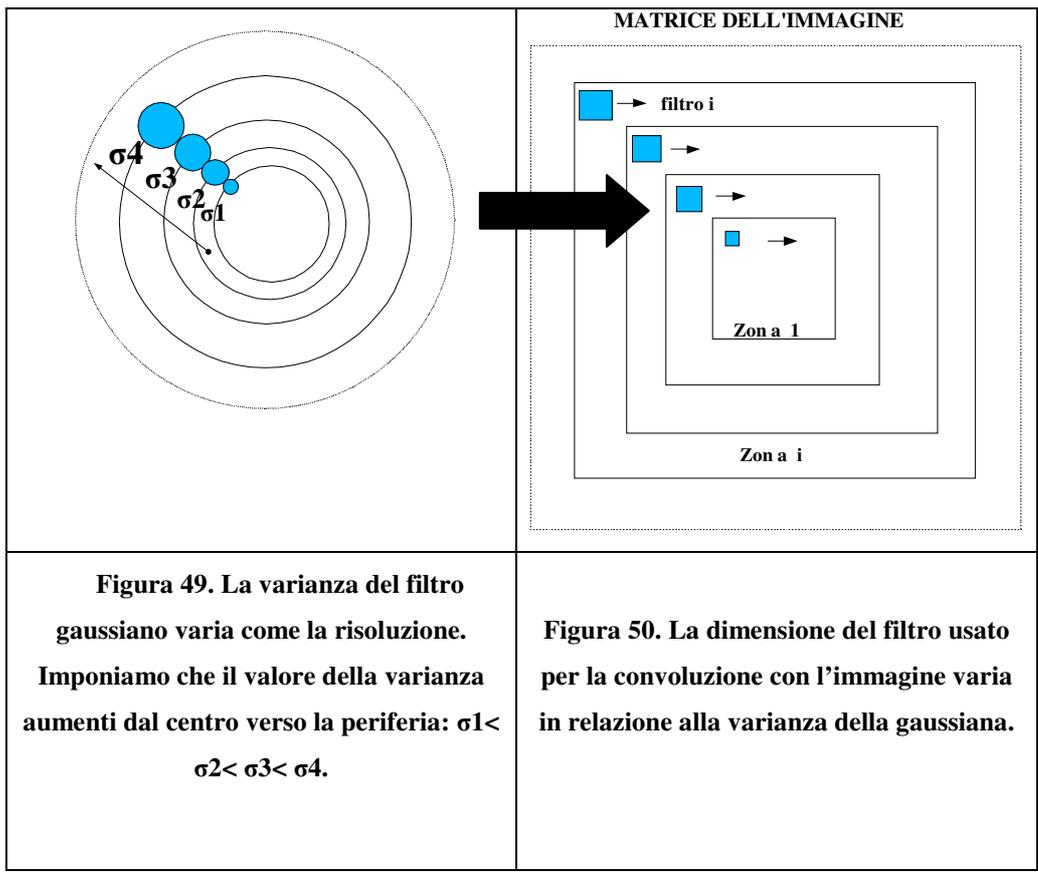


Figura 51. Esempio di space variant edge detection su immagine cartesiana rimappata da log-polare

Dopo aver applicato lo space variant edge detection sulle immagini cartesiane si è testato anche l'edge detection (sempre con il filtro di Canny ma non spazio variante) direttamente sulle immagini log polari rimappando poi il risultato dell'edge nella corrispondente immagine cartesiana, al fine di confrontare i due risultati. Possiamo visualizzare un esempio nella Figura 52 e in seguito evidenzieremo i problemi che emergono da questo tipo di azione.

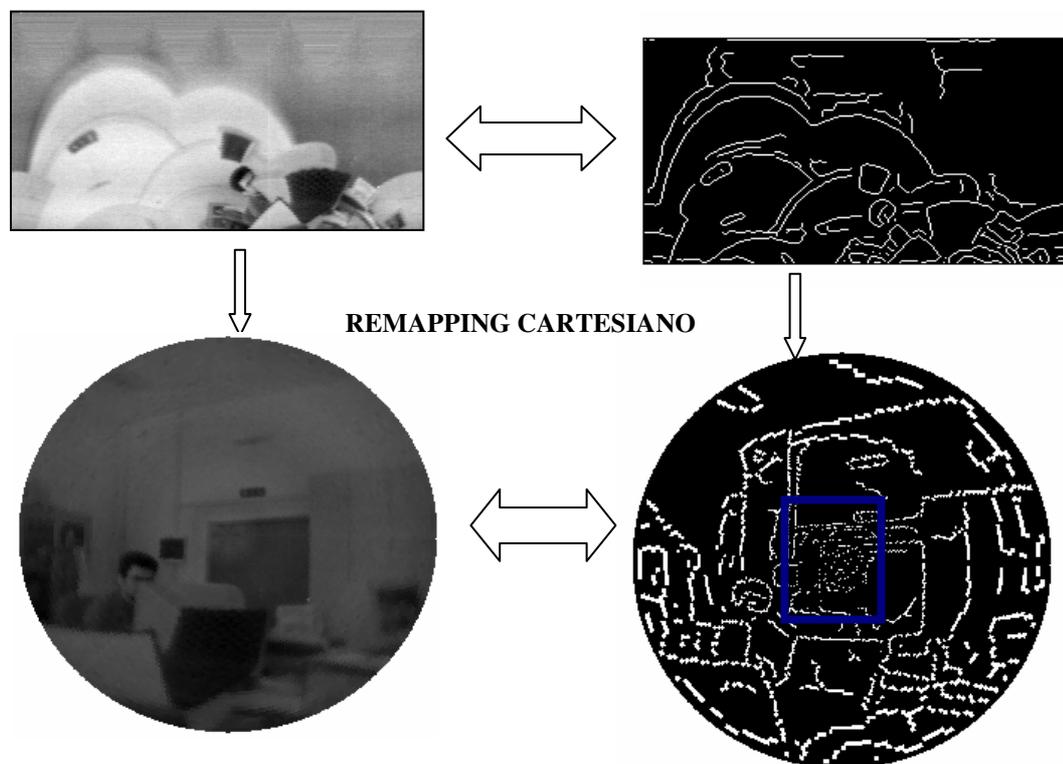


Figura 52. L'edge detection in questo caso viene effettuato sull'immagine log-polare e l'immagine con gli edge viene rimappata nella corrispondente cartesiana.

In questo ultimo esempio però si nota che gli edge nella parte più lontana dal centro dell'immagine hanno uno spessore maggiore di un pixel; questo perché i pixel appartenenti agli edge calcolati nell'immagine log-polare, quando vengono rimappati nell'immagine cartesiana vengono replicati più volte, se appartenenti ad una zona dell'immagine dove la risoluzione è minore. Nella parte centrale dell'immagine

invece avviene il fenomeno inverso, perché i pixel appartenenti agli edge della fovea, vengono rimappati con pochi pixel, e si presentano come bordi isolati non connessi, come si evidenzia nella parte riquadrata in blu della Figura 52.

Quindi possiamo affermare che in questo caso, l'edge detection spazio variante applicato all'immagine cartesiana da un migliore risultato rispetto al edge detection applicato in quest'ultimo caso, come si può capire dal confronto delle due immagini di Figura 51 e di Figura 52.

Al fine di formare un'immagine migliore della precedente ottenuta con il secondo tipo di edge detection (sull'immagine log-polare), si elabora l'immagine cartesiana rimappata da log-polare diminuendo lo spessore degli edge in periferia e collegando i pixel isolati nella fovea. Per espletare questo tipo di elaborazione viene usata una funzione di Matlab atta ad elaborare le immagini in bianco e nero. Essa è la funzione "Bwmorph". Può essere utilizzata con vari parametri e a seconda del tipo di parametro usato il tipo di elaborazione sull'immagine varia. Si è usata la funzione con il parametro "thin" per assottigliare gli edge presenti nella periferia dell'immagine, mentre si è usata la funzione con il parametro "bridge" per connettere il più possibile gli edge nella zona centrale dell'immagine.

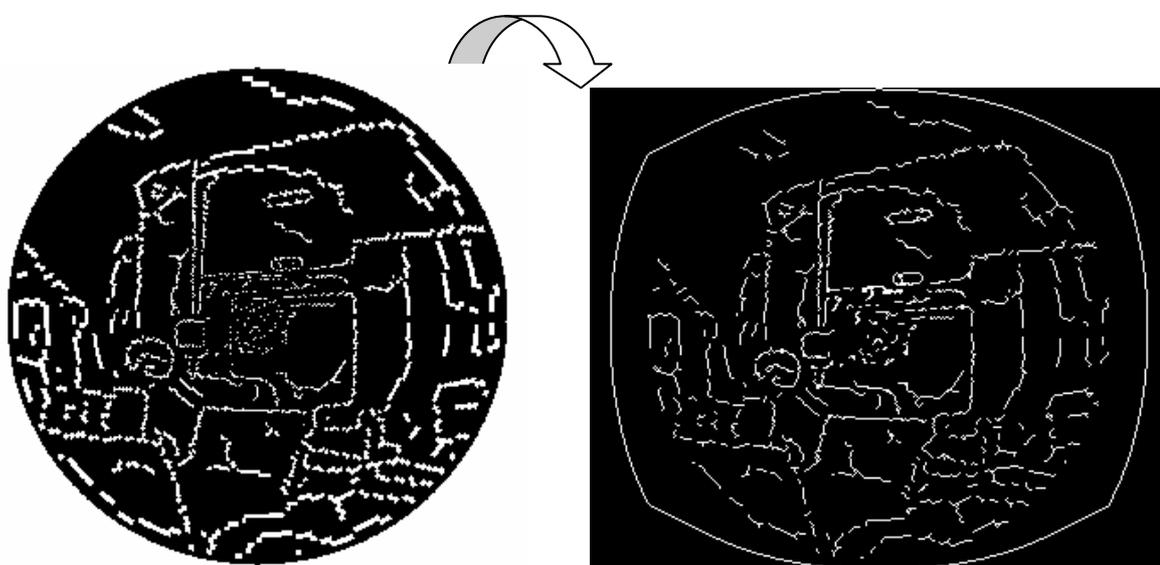


Figura 53. Risultato dell'operazione di 'morphing' dell'immagine degli edge cartesiana rimappata dall'immagine degli edge log-polare.

L'immagine degli edge che abbiamo ottenuto dopo l'elaborazione è sicuramente migliore di quella antecedente l'elaborazione. Se però confrontiamo quest'ultimo risultato con l'immagine degli edge (di Figura 54) ottenuta con l'edge detection

spazio variante possiamo verificare che l'edge detection sull'immagine cartesiana rimane la soluzione migliore, in quanto permette di ricavare edge più delineati.

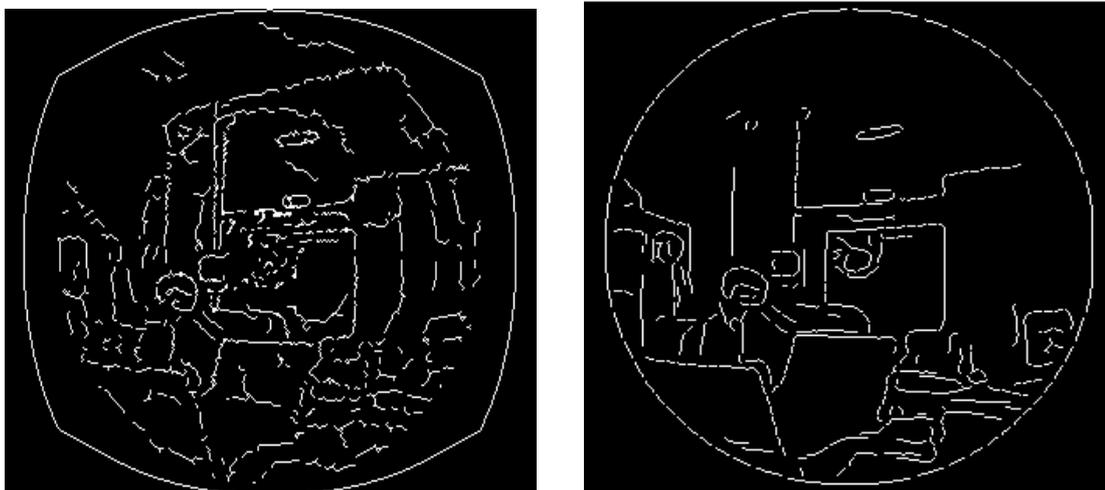


Figura 54. A sinistra l'immagine degli edge su immagine log polare con successivo remapping ed elaborazione. A destra l'immagine degli edge ottenuta con edge detection spazio variante.

5.SPECCHIO DEDICATO A TELECAMERA RETINICA

5.1 INTRODUZIONE

Nella progettazione degli specchi a simmetria rotazionale per telecamere omnidirezionali alcuni importanti parametri da analizzare sono l'errore assoluto e l'errore relativo che si compie nel misurare la distanza incognita agli oggetti visualizzati e il fattore di distorsione delle immagini acquisite che è introdotto dallo specchio. Scegliendo opportunamente determinati parametri di progettazione possiamo ottimizzare gli aspetti sopra citati.

Utilizzando il profilo dello specchio omnidirezionale proposto da Marchese e Sorrenti [16] è possibile eliminare la distorsione delle immagini acquisite dalla telecamera omnidirezionale. Lo specchio viene chiamato Isometrico in quanto le immagini degli oggetti attorno al robot sono indipendenti dalla distanza da essa e inoltre l'errore assoluto nelle misure delle distanze è costante. Nella soluzione di progettazione proposta invece in [17] la prima parte del profilo dello specchio fornisce un errore relativo costante nel calcolo delle distanze.

L'università di Praga invece (con supervisor Tomas Pajdla and Hadas Roth) ha progettato e costruito uno specchio omnidirezionale per la telecamera retinica che permette di acquisire immagini omnidirezionali retiniche aventi risoluzione angolare costante.

Anche in questa tesi, come ultimo lavoro, si vuole progettare uno specchio appositamente creato per la telecamera retinica Cmos che, viste le differenze sostanziali rispetto ai normali sensori, necessita di appositi accorgimenti in fase di progettazione. Ulteriore obiettivo della progettazione sarà quello di ottenere un profilo che determini un errore relativo costante.

5.2 ALGORITMO PER LA CREAZIONE DEL PROFILO DELLO SPECCHIO

Nella progettazione dello specchio è stata utilizzata una tecnica già sviluppata in [3] e in [4]: data una generica funzione $f : \mathcal{R}^2 \rightarrow \mathcal{R}^2$, si può sintetizzare uno specchio che proietti ciascun punto (X, Y) appartenente ad un regione limitata del piano di riferimento in un punto (x, y) sul piano del CMOS, in modo tale che

$(x, y)=f(X, Y)$. Data la simmetria rotazionale dello specchio, in realtà interessa sintetizzare un profilo radiale che proietti i punti a distanza D ($0 \leq D \leq D_{MAX}$) sul piano di riferimento in punti a distanza d ($0 \leq d \leq d_{MAX}$) sul piano del CMOS, dal centro del sensore. La proiezione segue una relazione $d=f^*(D)$, dove f^* è una generica funzione $f^*: [0:D_{max}] \rightarrow [0:d_{max}]$.

La soluzione esatta di questi problemi porta alla risoluzione di un'equazione differenziale; sempre [16] propone una soluzione basata su un' approssimazione locale lineare del profilo dello specchio, ottenuta sostituendo in ogni punto il profilo stesso con la sua tangente. L' algoritmo presentato in [16] verrà adattato considerando il fatto che si vorrà mappare nel centro dell' immagine una circonferenza di raggio D_{min} in modo da eliminare il corpo del robot. In pratica la funzione f^* avrà un dominio del tipo $[D_{MIN}, D_{MAX}]$ e il codominio $[0, d_{MAX}]$ ricordando che $f^*(D_{MIN})=0$.

Per i dettagli dell' algoritmi riferirsi a [3] o [4].

Il passo più importante è ora scegliere la funzione f^* che tenga conto delle nostre esigenze.

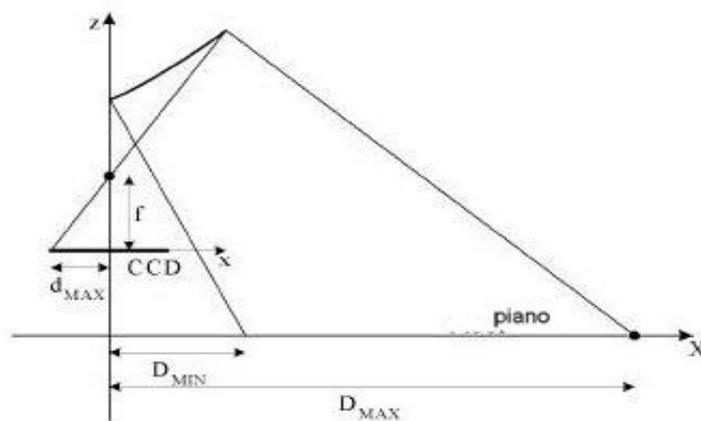


Figura 55. Schema che evidenzia come le distanze $[D_{MIN}, D_{MAX}]$ vengano mappate nel ccd in $[0, d_{MAX}]$.

5.3 CARATTERISTICHE PROGETTUALI

Come visto nel Capitolo 1, il sensore retinico presenta una distribuzione dei pixel spazio variante, e inoltre la dimensione dei fotorecettori disposti sugli anelli aumenta

radialmente. Il fenomeno di spazio varianza è assente nella fovea, in cui i fotorecettori mantengono la stessa dimensione e la spaziatura fra essi è costante.

La misura delle distanze di oggetti che si trovano attorno alla telecamera è un aspetto cruciale che trova applicazione ad esempio nel campo della “robot navigation”, in cui un robot mobile che si sposta su un piano deve riconoscere la posizione dei vari oggetti che si trovano nelle vicinanze di esso, al fine di avvicinarsi a loro o di evitarli (nel caso degli ostacoli). Per conoscere la posizione di un oggetto è sufficiente ottenere il valore dell’azimuth e della distanza D dell’oggetto dal robot. Data la simmetria rotazionale dello specchio possiamo calcolare l’azimuth sapendo le coordinate (x, y) della proiezione P' del punto P sul piano del sensore con la semplice formula:

$$\phi = \arctan(y/x)$$

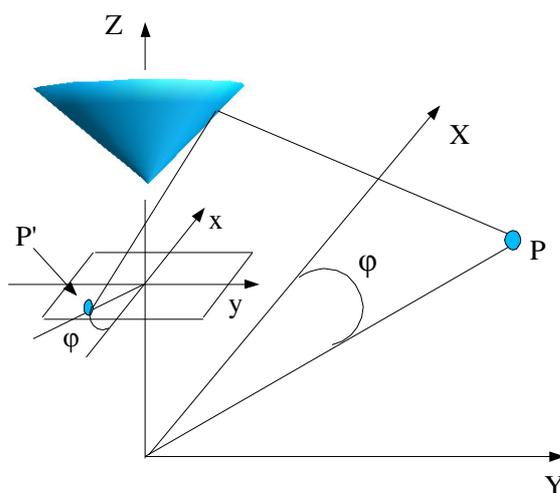


Figura 56. Conoscendo la distanza D e il valore dell’azimuth θ si può calcolare la posizione del punto P

Per ricavare la distanza D ci si basa sulla conoscenza della funzione f^* che mappa un punto ad una distanza D dal centro del robot, in un punto a distanza d dal centro del sensore. Abbiamo quindi $d = f^*(D)$, e la funzione inversa $D = f^{*(-1)}(d)$ dove ovviamente $d = (x^2+y^2)^{1/2}$ con (x, y) rappresentanti come detto le coordinate della proiezione P' del punto P sul sensore.

Poiché il pixel del sensore ha una dimensione finita A possiamo ricavare soltanto la posizione (x, y) del suo centro C , approssimando così la misura effettiva della

proiezione P' con il valore Pc. Per un determinato pixel quindi avremo un insieme di valori (x, y) discretizzati nel valore del centro del pixel C. Questo comporta un insieme di distanze $d = (x^2+y^2)^{1/2}$ tali che (x, y) siano appartenenti all'insieme A che copre la dimensione del pixel. Questo comporta che la funzione inversa $f^{*(-1)}(d)$ produca un insieme di valori Di rappresentante l'insieme di valori discreti, rappresentanti le distanze di un oggetto.

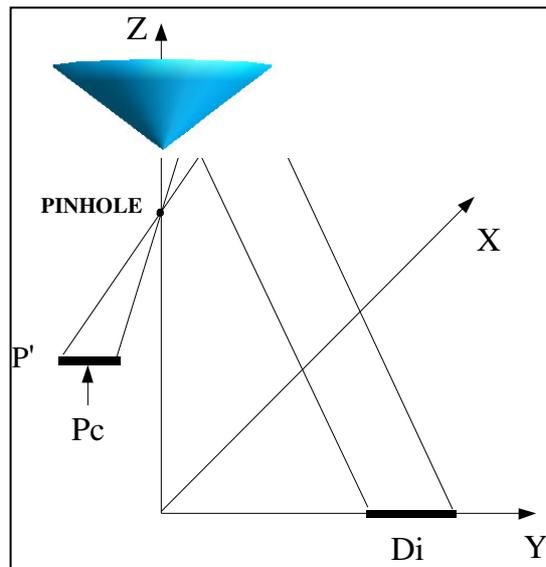


Figura 57. Avendo il pixel P' un' area finita il suo valore viene discretizzato nel centro Pc e produce un insieme di valori Di nel piano tramite l'applicazione della funzione inversa $f^{*(-1)}(d)$.

Ovviamente assoceremo al pixel Pc a distanza d dal centro del sensore soltanto una delle possibili distanze Di sul piano reale; al fine di minimizzare l'errore assoluto sceglieremo il valore medio di tale insieme.

Nel sensore retinico i fotorecettori (pixel) esterni alla fovea aumentano di dimensione con la distanza d dal centro; la distanza tra i centri dei fotorecettori (pixel) consecutivi varia come già specificato precedentemente secondo una legge del tipo: $\rho_0(a^{n_i} - a^{n_{i-1}})$ dove come al solito ρ_0 è il raggio del cerchio più interno, a la base del logaritmo, n_i e n_{i-1} gli anelli consecutivi dove sono disposti i vari fotorecettori. La Figura 58 evidenzia quanto detto e pone anche in evidenza come i raggi si riflettano dallo specchio ai centri dei pixel.

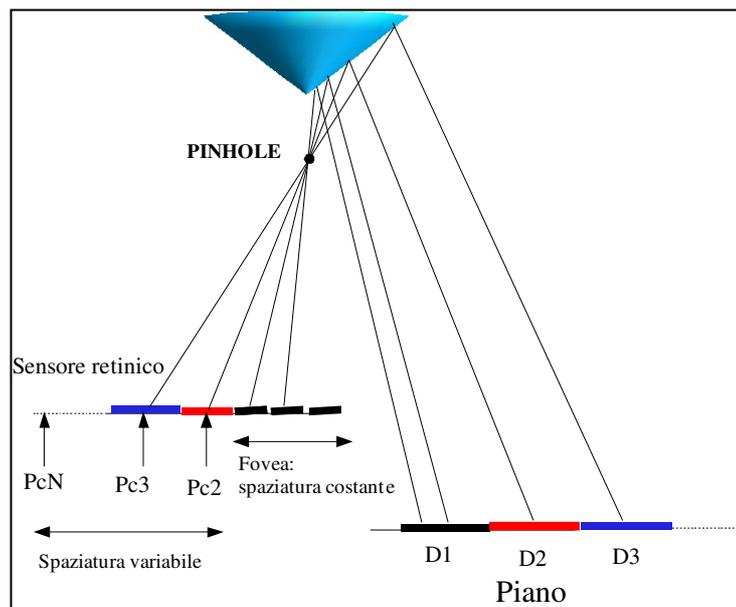


Figura 58. Il sensore retinico ha i fotorecettori non equamente spazati fuori della fovea mentre nella fovea la spaziatura è costante.

La scelta di progettare uno specchio che comporti un errore relativo costante piuttosto che un errore assoluto costante è motivata dal fatto che un errore assoluto costante comporta un errore relativo molto alto nelle vicinanze del robot, come si può notare nella Figura 59 che rappresenta l'errore assoluto e quello relativo nello specchio Isometrico proposto da Marchese e Sorrenti [16]

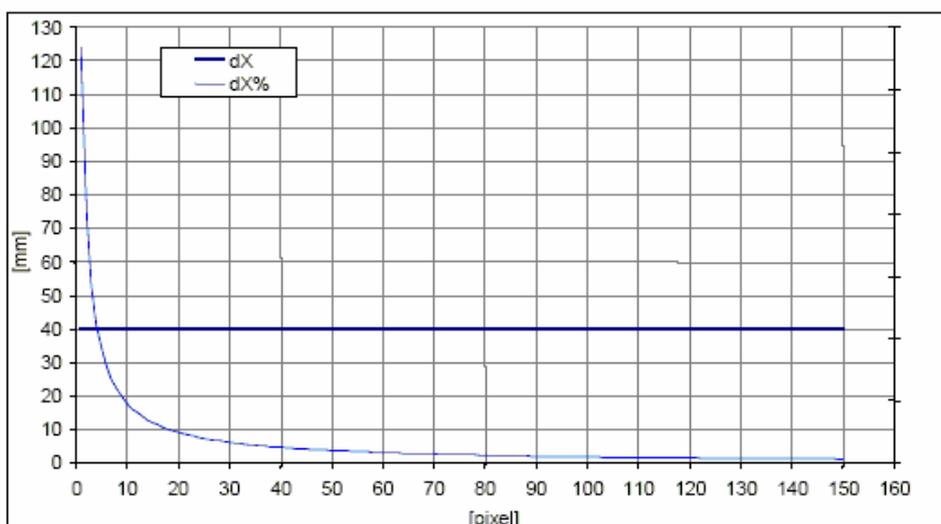


Figura 59. Andamento dell'errore assoluto (dX) ed errore relativo(dX%)

Vogliamo avere maggiore precisione nelle vicinanze del robot in quanto è nelle vicinanze di questo che abbiamo, data la struttura del nostro sensore, maggiore

risoluzione e dove possiamo analizzare gli oggetti visualizzati in maniera più dettagliata.

Un'ulteriore caratteristica molto rilevante dello specchio progettato è la curvatura di questo in corrispondenza del vertice. Gli specchi conici con discontinuità nel vertice permettono di mappare un'intera circonferenza C1 nel centro dell'immagine, mentre specchi conici con vertice sferico (senza discontinuità) mappano nel centro dell'immagine tutti i punti ad una distanza $D \neq 0$ prossimi al centro del robot. Oltre a questo aspetto questi due tipi di specchi uniscono altri pregi e difetti specificati nella tabella 4 come indicato sia in [16] che in [17].

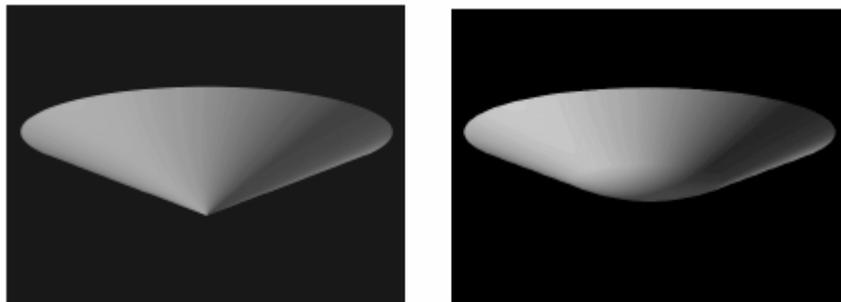


Figura 60. Profilo di specchio conico e specchio conico con vertice sferico

SPECCHIO	VANTAGGI	SVANTAGGI
CONICO	E' possibile eliminare dall'immagine il corpo del robot scegliendone opportunamente i parametri geometrici	Distanza massima misurabile molto ridotta. Scarsa risoluzione nella misura di distanze piccole.
CONICO CON VERTICE SFERICO	Buona risoluzione nella misura di piccole distanze. Errore relativo basso nelle misure di distanze sul piano di gioco	Parte centrale dell'immagine inutilizzabile perche' occupata dal corpo del robot Distanza massima misurabile molto ridotta.

Tabella 4. Confronto tra i due tipi di specchi analizzati.

Data la struttura del nostro sensore non possiamo permetterci di sprecare la parte centrale dell'immagine con informazione inutile quale potrebbe essere il corpo del robot, visto che essa corrisponde alla fovea che deve essere sfruttata essendo la zona

a maggior risoluzione. Opteremo quindi per la soluzione dello specchio con discontinuità nel vertice.

Prima di progettare lo specchio per il sensore retinico progetteremo altre due tipologie di specchi che permetteranno di capire maggiormente come strutturare gli insiemi 'Di' e 'di' per lo specchio retinico e per i quali valuteremo l'influenza della struttura di questi ultimi insiemi nella risoluzione dell'immagine riflessa dallo specchio. Le simulazioni degli specchi progettati verranno eseguite in ambiente POV- RAY, il programma di ray tracing che ci permette di effettuare le simulazioni. Inoltre verrà utilizzato ambiente virtuale il campo di calcio di Robocup.

5.4 SPECCHIO CON SUDDIVISIONE DEGLI INSIEMI NEL PIANO LINEARE

In questo specchio la funzione $f^*(D)=d$ viene ricavata costruendo una tabella che associa ad ogni distanza sul piano la distanza del centro del pixel dal centro del sensore con lo stesso procedimento seguito in [16] e [17]. In questo primo specchio progettato consideriamo un sensore tradizionale. L'obiettivo come già detto è quello di commettere un errore relativo (nel caso peggiore) costante, e comunque sempre al di sotto dei una percentuale stabilita.. Vedremo dopo la costruzione degli insiemi il significato di "caso peggiore". Si deve dividere il piano in opportuni insiemi in cui il centro rappresenta l'approssimazione della distanza D che corrisponde alla distanza del pixel sul sensore. Gli insiemi nel piano verranno creati in questo modo:

$$D_1 = D_{MIN} + \alpha D_{MIN}$$

$$D_2 = D_1 + \alpha D_1$$

.....

$$D_i = D_{i-1} + \alpha D_{i-1}$$

dove la costante α rappresenta l'errore relativo che si commette. Gli insiemi dei centri dei pixel sul sensore sono costruiti in questo modo:

$$d_1 = (0.0075 / 2)$$

.....

$$d_i = d_{i-1} + (0.0075 / 2)$$

puntualizzando che la costante 0.0075 corrisponde alla dimensione del pixel ovvero alla particolare risoluzione del sensore considerata.

Quindi per come è costruito lo specchio un punto P ad una distanza D dal centro del robot con $D_{i-1} < D < D_i$ verrà proiettato in un pixel che dista d_{i-1} o d_i dal centro del sensore. Nel caso peggiore si verifica che $D = D_{i-1}$ ed il punto P è proiettato in un pixel che dista d_i ; quindi l'errore assoluto sarà uguale a $D_i - D_{i-1} = \alpha D_{i-1} = \alpha D$ e l'errore relativo pari ad $\alpha D / D = \alpha$, costante. Il caso peggiore si ha quindi, quando un punto ad una distanza D nel piano viene mappato in un centro del pixel errato, successivo o precedente al centro del pixel che lo identifica correttamente.

Dopo avere costruito gli insiemi sul piano e sul CMOS per completare la progettazione sono stati scelti i parametri essenziali per la creazione del profilo dello specchio. Essi vengono riportati nella tabella sottostante.

Altezza telecamera da terra	100 cm
Distanza specchio-telecamera	10 cm
α	0.02
Dmin (distanza minima visibile)	30 cm
Dmax (distanza massima visibile)	596 cm

Tabella 5. Parametri progettuali e visivi dello specchio

Di seguito si riporta il profilo dello specchio generato con Matlab e lo specchio simulato con POV-RAY.

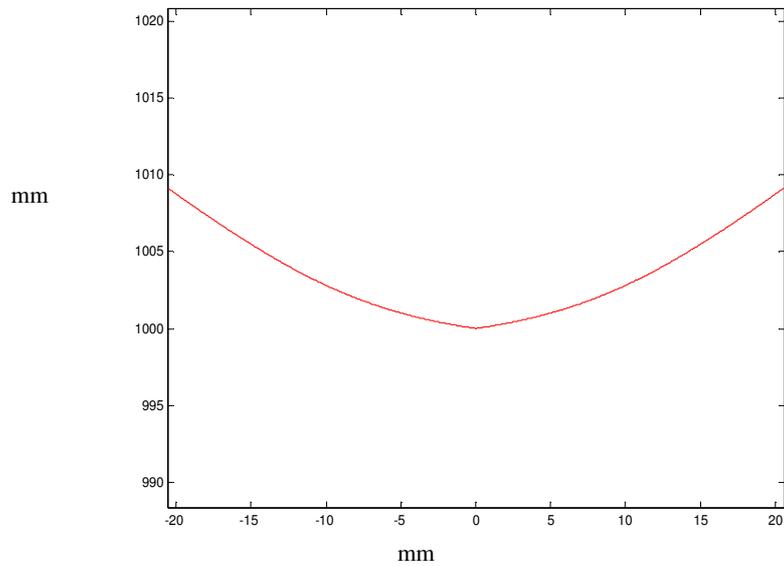


Figura 61. Profilo dello specchio creato.



Figura 62. Immagine dello specchio in ambiente POV RAY

Al fine di verificare la corretta misurazione delle distanze ed il conseguente errore relativo si simula la visione di un campo di gioco con il posizionamento di una serie di sfere cui verrà calcolata la distanza. La Figura 63 e la tabella 6 riportano i risultati ottenuti.

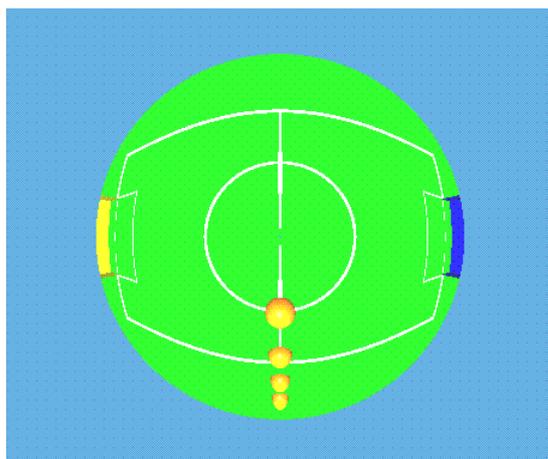


Figura 63. Visione del campo di gioco e di oggetti posti a diverse distanze dal centro del robot.

Distanza (cm)	Oggetto	Pixel dell'oggetto	Distanza stimata (cm)	Errore assoluto (cm)	Errore relativo [%]
80		50	79.164	0.836	1.04
180		92	181.86	1.86	1.03
280		114	281.15	1.15	0.4

Tabella 6. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco.

La risoluzione offerta da questo specchio ovviamente decresce in maniera lineare visto come abbiamo suddiviso gli insiemi nel piano. E' interessante notare che l'errore relativo nelle simulazioni si mantiene sotto il 2%, valore che avevamo detto essere l'errore relativo costante nel caso peggiore che volevamo ottenere.

5.5 SPECCHIO CON SUDDIVISIONE DEGLI INSIEMI NEL PIANO ESPONENZIALE.

Il passo successivo è quello di creare un nuovo specchio con la suddivisione degli insiemi D_i sul piano con legge esponenziale secondo una determinata base b . La suddivisione sarà quindi di questo tipo:

$$D_1 = k * b^0$$

$$D_2 = k * b^1$$

.....

$$D_i = k * b^i$$

mentre la suddivisione degli insiemi 'di' nel sensore sarà la stessa dello specchio precedentemente progettato.

Anche in questo caso per come è costruito lo specchio un punto P ad una distanza D dal centro del robot con $D_{i-1} < D < D_i$ verrà proiettato in un pixel che dista d_{i-1} o d_i dal centro del sensore. Nel caso peggiore $D=D_{i-1}$ ed il punto P è proiettato in un pixel che dista d_i ; quindi l'errore relativo sarà uguale a $D_i - D_{i-1} / D_{i-1} = (k * b^i - k * b^{i-1}) / k * b^{i-1} = (b^i - b^{i-1}) / b^{i-1} = b - 1$.

Vengono ora riportati come precedentemente fatto i parametri scelti per la progettazione.

Altezza telecamera da terra	100 cm
Distanza specchio-telecamera	10 cm
b	1.03
Dmin (distanza minima visibile)	20 cm
Dmax (distanza massima visibile)	1735 cm

Tabella 7. Parametri progettuali e visivi dello specchio

Di seguito si riporta il profilo dello specchio generato con Matlab e lo specchio simulato con POV RAY.

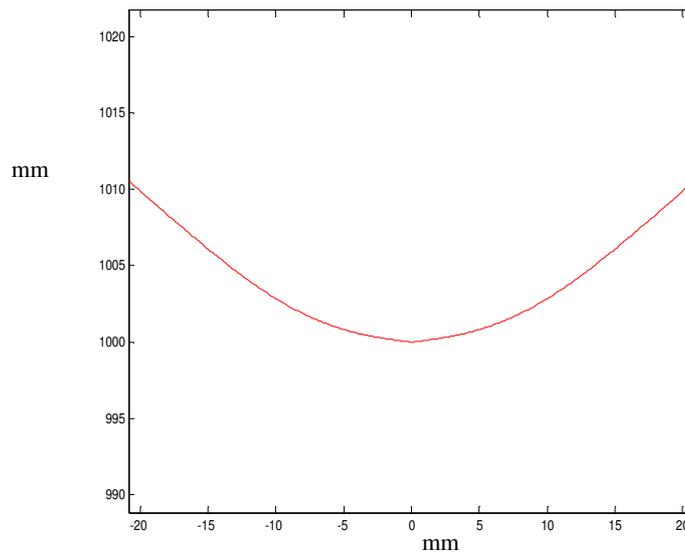


Figura 64. Profilo dello specchio creato.

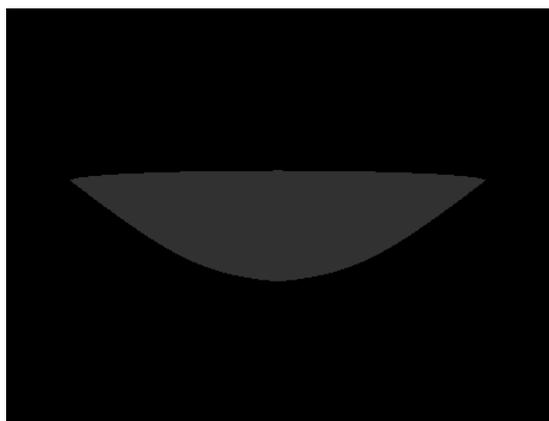


Figura 65. Immagine dello specchio in ambiente POV RAY

In questo tipo di specchio ovviamente la risoluzione dell'immagine diminuirà esponenzialmente dal centro verso la periferia della stessa.

Anche ora al fine di verificare la corretta misurazione delle distanze ed il conseguente errore relativo si simula la visione di un campo di gioco con il posizionamento di una serie di sfere cui verrà calcolata la distanza. La Figura 66 e la tabella 8 riportano i risultati ottenuti.

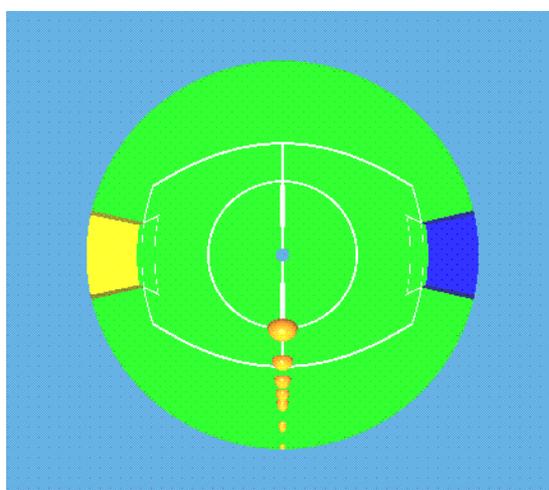


Figura 66. Visione del campo di gioco e di oggetti posti a diverse distanze dal centro del robot.

Distanza (cm)	Oggetto	Pixel dell'oggetto	Distanza stimata (cm)	Errore assoluto (cm)	Errore relativo [%]
80		48	80.238	0.238	0.29
180		76	183.86	3.86	2
800		125	781.35	18.5	2.31

Tabella 8. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco.

Anche per questo specchio possiamo notare che nella misura delle distanze incognite degli oggetti sul piano l'errore relativo si mantiene sotto il 3%, valore prefissato essere l'errore relativo massimo commettabile. La risoluzione però ora varia in maniera esponenziale dal centro dell'immagine verso la periferia visto il tipo di suddivisione fatto per gli insiemi nel piano.

5.6 SPECCHIO PER SENSORE RETINICO A ERRORE RELATIVO COSTANTE

Nella progettazione dello specchio per il sensore retinico dovremo tenere conto che la spaziatura tra i pixel è variabile per i pixel esterni alla fovea, mentre è costante per i pixel interni ad essa, come evidenziato dalla Figura 58. Quindi dovremo costruire gli insiemi 'di' di conseguenza seguendo la struttura del sensore. Per i centri dei pixel nella fovea avremo una suddivisione degli insiemi del tipo:

$$d_1 = (0.0065 / 2)$$

.....

$$d_{\max \text{ fovea}} = d_{i-1} + (0.0065 / 2)$$

dove la costante di valore 0.0065 è la dimensione del pixel più piccolo del sensore retinico.

Mentre per gli insiemi 'di' relativi ai centri dei pixel esterni alla fovea la suddivisione sarà del tipo:

$$d_{\max \text{ fovea} + 1} = \rho_0 \times a^{\max \text{ fovea} + 1}$$

$$d_{\max \text{ fovea} + 2} = \rho_0 \times a^{\max \text{ fovea} + 2}$$

.....

$$d_i = \rho_0 \times a^{ni}$$

$$d_N = \rho_0 \times a^N$$

Per quanto riguarda la suddivisione degli insiemi Di nel piano seguiremo una strada simile: per i primi n insiemi Di nel piano che sono proiettati nei centri dei pixel della fovea faremo questa suddivisione:

$$D_1 = D_{\min} + kD_{\min} \text{-----} > \text{corrispondente } _ a _ d_1$$

$$D_2 = D_1 + kD_1 \text{-----} > \text{corrispondente } _ a _ d_2$$

$$D_i = D_{i-1} + kD_{i-1} \text{-----} > \text{corrispondente_a_} d_i$$

$$D_{\max_fovea} = D_{\max_fovea-1} + kD_{\max_fovea-1} \text{-----} > \text{corrispondente_a_} d_{\max_fovea}$$

dove con max_fovea si indicano il numero degli anelli nella fovea e con k il valore dell'errore relativo che vogliamo commettere. Per gli insiemi Di nel piano che sono proiettati nei centri dei pixel esterni alla fovea avremo invece questa suddivisione:

$$D_{\max_fovea+1} = +Q \times b^{\max_fovea+1}$$

.....

$$D_j = Q \times b^j$$

dove Q è una costante che permette di variare la distanza massima rappresentabile e b la base dell'esponenziale. Vengono inoltre ora riportati come precedentemente fatto i parametri scelti per la progettazione.

Altezza telecamera da terra	100 cm
Distanza specchio-telecamera	10 cm
b	1.03
Dmin (distanza minima visibile)	10 cm
Dmax(distanza massima visibile)=	400 cm
Dmax_fovea	12 cm
Q	23

Tabella 9. Parametri progettuali e visivi dello specchio.

Di seguito si riporta come prima il profilo dello specchio generato con Matlab e lo specchio simulato con POV RAY.

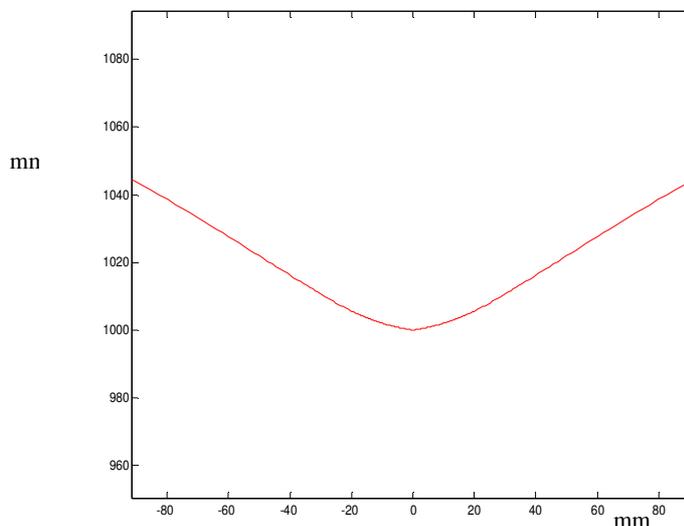


Figura 67. Profilo dello specchio retinico.

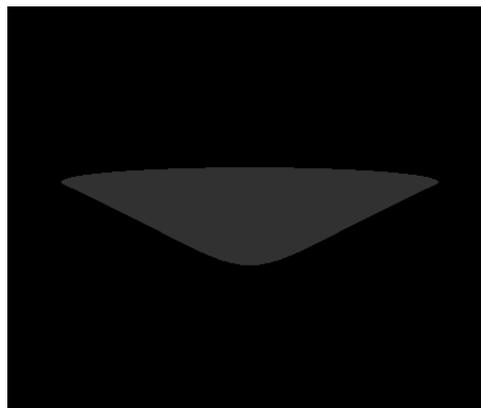


Figura 68. Immagine dello specchio retinico in ambiente POV RAY.

A questo punto dopo aver progettato lo specchio, per poter simulare in ambiente POV RAY la visione dell'immagine generata dal sensore retinico è necessario trasformare l'immagine renderizzata in un'immagine cartesiana spazio variante. Inoltre per riuscire a calcolare le distanze basandoci sui pixel dell'immagine generata dovremo operare alcuni accorgimenti. Poiché la tabella degli insiemi Di che abbiamo creato rappresenta le proiezioni sui centri dei pixel del sensore retinico, scegliendo un determinato pixel nell'immagine non vi sarebbe corrispondenza con il pixel del sensore retinico. Questo perché nell'immagine generata da POV-RAY i pixel sono considerati tutti della stessa dimensione.

Nell'algoritmo di progettazione dello specchio viene costituita una seconda lookup table dove viene associato ad ogni centro del pixel del sensore retinico un insieme di pixel a spaziatura costante, aventi tutti la dimensione minima uguale alla dimensione del pixel più piccolo del sensore retinico. Così per ogni pixel di grandezza variabile del sensore retinico verrà corrisposto un insieme di pixel (simulati) di dimensione costante di un sensore normale. Riusciremo così a capire nell'immagine generata da POV RAY che un determinato pixel corrisponde ad un pixel più grande del sensore retinico.

. In particolare nel calcolo delle distanze potremo identificare la posizione del pixel desiderato nell'immagine cartesiana, calcolare a quale insieme esso appartenga della lookup table e associarlo al centro del pixel del sensore retinico che ci interessa per risalire alla distanza stimata. La Figura 69 esemplifica quanto detto fino ad ora. Per esempio se la posizione del pixel che ci interessa calcolare è localizzata nell'area cerchiata in giallo (in questo caso il Pix3) tramite la lookup table creata possiamo risalire al pixel del sensore retinico (il pixel numero 2) e quindi verificare la distanza stimata nella tabella contenente gli insiemi Di.

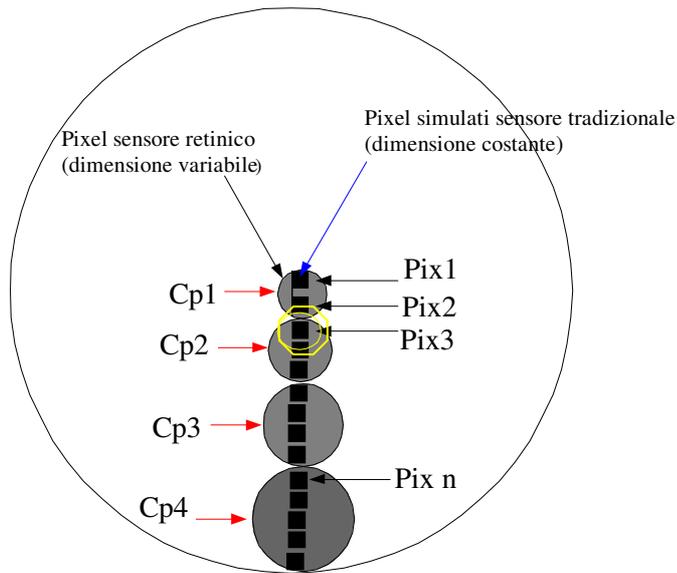


Figura 69. Schema di come vengono simulati i pixel di dimensione costante nell'immagine POV RAY calcolati in base al pixel del sensore retinico.

Verifichiamo ora quanto detto creando l'immagine della visione del campo di gioco da parte del sensore, con alcuni oggetti disposti su di esso; successivamente creiamo l'immagine cartesiana spazio variante e infine effettuiamo il calcolo delle distanze.



Figura 70. Visione del campo di gioco e di alcuni oggetti posati a diverse distanze su di esso ottenuta con uno specchio spazio variante e un sensore tradizionale simulato con POV-RAY.

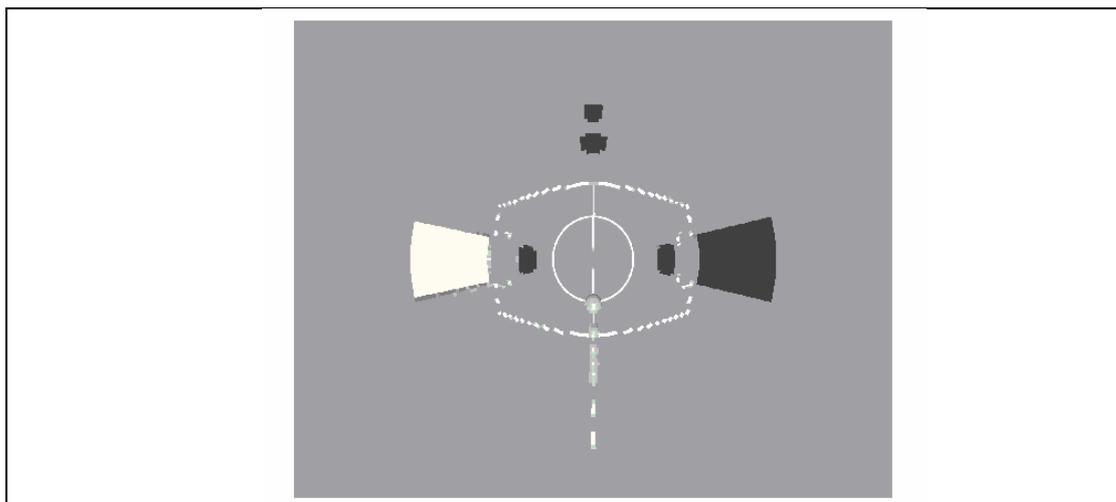


Figura 71. Immagine del campo di gioco trasformata nella corrispondente cartesiana spazio variante.

Distanza Oggetto	Pixel simulato dell'oggetto	Pixel del sensore retinico	Distanza stimata (cm)	Errore assoluto (cm)	Errore relativo [%]
180	139	101	183.19	3.19	1.7
280	181	112	281.27	1.27	0.4
340	240	119	377.41	3.59	1

Tabella 10. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco

Anche in quest' ultimo caso l'errore relativo permane sotto il 3%, valore stabilito come la percentuale di errore relativo massimo che abbiamo imposto essere costante.

Quindi con lo specchio progettato possiamo ricavare la distanza degli oggetti attorno alla telecamera ;questo tramite un semplice accesso ad una lookup table dove sono memorizzati i valori delle distanze sul piano relativi ai pixel nell'immagine. Inoltre l'errore relativo commesso nel calcolo delle distanze sarà al massimo costante, anche se come abbiamo visto dalle simulazioni questo si mantiene su un valore molto piu' ridotto del valore che abbiamo imposto preliminarmente. Uno svantaggio che si presenta è invece la risoluzione non costante dell'immagine omnidirezionale acquisita; infatti questa diminuisce esponenzialmente dal centro dell'immagine verso la periferia .

Le considerazioni e le simulazioni sono state effettuate sull'immagine cartesiana e per il calcolo delle distanze si sono dovuti operare gli accorgimenti specificati al fine di costituire la lookup table. Nella realtà comunque, avendo automaticamente l'immagine log-polare il calcolo della distanza oltre che il calcolo dell'azimuth è molto più diretto. In quanto per calcolare la distanza D dell'oggetto interessato nell'immagine log-polare si visiona direttamente il valore della riga a cui corrisponde il pixel relativo all'oggetto nell'immagine, e si accede direttamente alla tabella delle distanze senza passare per la lookup table. Per il calcolo dell'azimuth invece, si visiona il valore della colonna del pixel dell'immagine log-polare e sapendo la risoluzione angolare minima si perviene all'azimuth. Per esempio prendiamo l'immagine omnidirezionale retinica della figura sottostante (tra le immagini acquisite del prospetto riassuntivo del Capitolo 2):

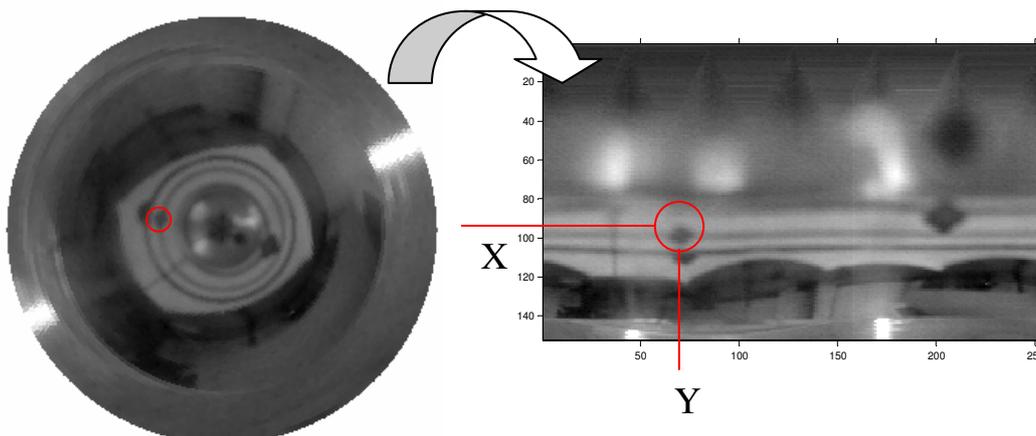


Figura 72. L'oggetto cui vogliamo calcolare la posizione è quello cerchiato, per risalire alla distanza dal centro valutiamo la riga dell'immagine a cui corrisponde il pixel dell'oggetto, e accediamo alla tabella delle distanze D_i ; mentre per l'azimuth valutiamo la colonna di appartenenza del pixel e moltiplichiamo per la risoluzione angolare.

Volendo calcolare la posizione dell'oggetto cerchiato basterà considerare il pixel dell'immagine corrispondente log-polare, valutando per la distanza la riga e accedendo alla tabella delle distanze dal centro D_i ; mentre per l'azimuth valutando la colonna e moltiplicando questa per la risoluzione angolare del sensore, sottraendo poi l'angolo θ corrispondente al quadrante più vicino.

CONCLUSIONI

Nella presente tesi si è analizzato inizialmente il funzionamento e la struttura del sensore retinico, abbinato ad uno specchio omnidirezionale. L'ottimo compromesso tra risoluzione, campo visivo e peso dell'immagine introdotto da questo nuovo tipo di sensore studiato, è un vantaggio per l'utilizzatore in un vasto numero di applicazioni, in cui è cruciale conseguire un alleggerimento del peso dell'informazione elaborata, senza rinunciare a dati rilevanti per il raggiungimento dell'obiettivo preposto al tipo di applicativo. Ecco che il sensore retinico insieme al visore omnidirezionale si propone come un interessante alternativa agli usuali sensori, non solo in ambito di progetti specializzati nella robotica mobile, ma anche in applicazioni più comuni quali la videosorveglianza o la videoconferenza, dove lo scambio di dati attraverso una rete e l'elaborazione degli stessi deve essere molto veloce senza dover comportare uno spreco eccessivo di risorse.

Dopo aver appreso la struttura del sensore si sono esaminate le proprietà delle immagini log polari, acquisite tramite la telecamera retinica.. Sono state effettuate un gran numero di acquisizioni in modo da mettere in evidenza la struttura intrinseca delle immagini log polari. Sono stati ricavati alcuni invarianti geometrici delle immagini retiniche, sfruttabili in algoritmi di visione computazionale. Alcune volte sono emersi dei problemi in fase di acquisizione delle immagini in quanto il software di acquisizione fornito insieme alla telecamera non era molto stabile, e non aveva un adeguata gestione degli errori. Inoltre avendo la possibilità di acquisire solo lo stream di dati identificanti le immagini, si è provveduto a implementare un software per la gestione e il salvataggio su disco (in un opportuno formato) delle immagine log polari acquisite tramite la telecamera retinica. Nell'acquisizione delle immagini retiniche omnidirezionali i problemi riscontrati sono dovuti al fatto che si è usato un supporto per lo specchio omnidirezionale che non permetteva un perfetto allineamento dell'asse di simmetria dello specchio con l'asse ottico della telecamera retinica. Essendo questa ultima molto sensibile a questo tipo di parametro le immagini omnidirezionali retiniche acquisite non erano precise. Sarebbe stato meglio costruire un supporto che garantisse un allineamento più accurato dell'asse ottico

della telecamera e dell'asse di simmetria dello specchio per garantire una precisione nelle immagini acquisite superiore.

Inoltre si è implementato un particolare tipo di Edge Detection, basato sul filtro di Canny, testandolo sulle immagini cartesiane spazio variante. L'edge detection sviluppato permette di trovare i bordi dell'immagine cartesiana spazio variante in maniera più precisa rispetto ad un normale edge detection, in quanto il filtro usato si adatta nelle varie zone dell'immagine cartesiana che presentano risoluzione variabile. E' stato testato anche l'edge detection direttamente sulle immagini log-polari ma i risultati non sono stati buoni, in quanto i bordi evidenziati nella periferia dell'immagine erano troppo spessi, mentre nel centro dell'immagine si presentavano spezzettati.

Infine come ultima fase della tesi si è progettato uno specchio omnidirezionale per la telecamera retinica. L'algoritmo di progettazione dello specchio permette la localizzazione degli oggetti visualizzati attorno alla telecamera tramite un semplice accesso ad una tabella di corrispondenze: pixel nell'immagine - posizione dell'oggetto nel piano. Inoltre il profilo dello specchio creato permette la misura delle distanze degli oggetti nel piano commettendo al più un errore relativo costante. Lo specchio omnidirezionale progettato abbinato alla telecamera retinica fornisce un ottimo supporto in applicazioni di navigazione di robot mobili, i quali usando il sensore di visione omnidirezionale retiniche, possono localizzare gli oggetti attorno a loro e operare numerose azioni all'interno dell'ambiente di navigazione, come ad esempio evitare gli ostacoli, afferrare degli oggetti, o semplicemente raccogliere informazioni della scena visualizzata.

Gli sviluppi futuri di questo lavoro di tesi sono molti. Innanzitutto si dovrebbe migliorare il software di acquisizione delle immagini retiniche al fine di renderlo più stabile; inoltre visto che da questa tesi si sono ricavate le proprietà più importanti e gli invarianti delle immagini log polari, si potranno studiare ed implementare numerosi algoritmi, nell'ambito della visione computazionale, che sfruttino queste caratteristiche. Inoltre si potrà realizzare fisicamente lo specchio progettato, in modo da dotare un robot mobile del sensore di visione omnidirezionale e quindi di verificare con dei test reali le potenzialità di questo sensore.

APPENDICE

SOFTWARE DI ACQUISIZIONE

La telecamera utilizzata in laboratorio era corredata da un software di acquisizione dell'immagini ma non permetteva di salvare le immagini log polari e quelle cartesiane su disco fisso. Era possibile salvare su disco fisso solo il file di acquisizione con formato *.lps. Questo file contiene uno stream di caratteri rappresentante la serie di immagini acquisite. Il singolo carattere identifica il valore a 8 bit del pixel di ogni immagine log-polare visualizzata. A questo punto è stato creato un software apposito per caricare il file di acquisizione e salvare le immagini log polari oltre a quelle cartesiane dopo essere state opportunamente rimappate. Ulteriori funzioni sono stati create per l'elaborazione delle immagini log polari. Nella pagina successiva viene presentata una panoramica dell'interfaccia grafica del software creato. Nella tabella successiva vengono riassunte le specifiche funzioni aggiuntive espletate.

FUNZIONE	DESCRIZIONE
ACQUISIZIONE	Carica il file .lps e crea le immagini log polari e le immagini cartesiane.
CARICA IMMAGINI	Carica le immagini da disco visualizzandole nelle due liste
VISUALIZZAZIONE	Per visualizzare le immagini selezionarle dalla lista; se è abilitata l'opzione di edge detection oltre alle immagini verrà visualizzata anche la corrispondente immagine con gli edge.
SIMULAZIONE	Questa funzione permette di simulare l'immagine che sarebbe prodotta da un sensore tradizionale a parità di risoluzione e parità di campo visivo.
ZOOM FOVEA	Può essere visualizzata solamente la parte dell'immagine cartesiana corrispondente alla fovea, dove vi è maggiore risoluzione.
SALVA	Permette di salvare le immagini selezionate in formato bitmap.

Tabella 11. Funzioni abilitate dal software di acquisizione delle immagini.

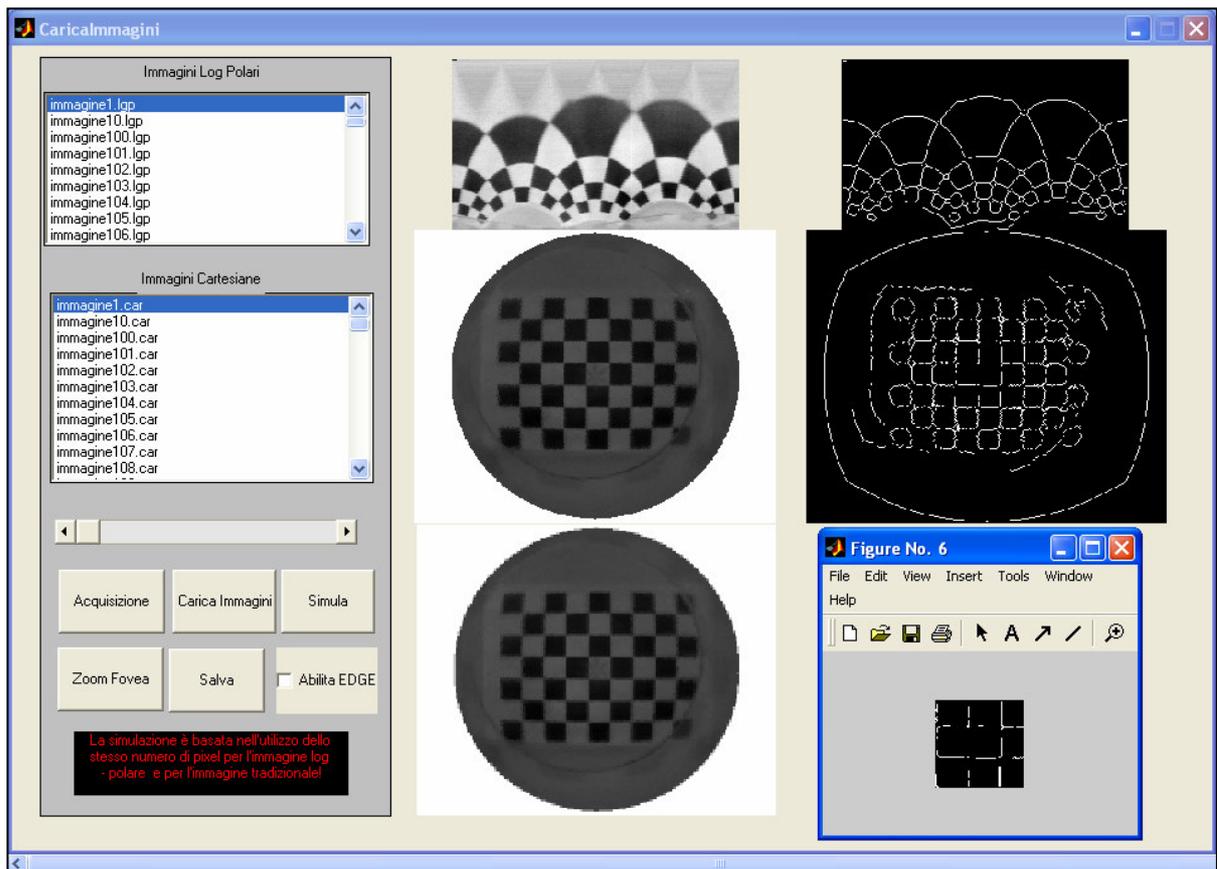


Figura 73. Interfaccia grafica del software di acquisizione delle immagini.

INDICE DELLE FIGURE

Figura 1. Struttura del sensore tradizionale. I fotorecettori hanno tutti la stessa dimensione.....	4
Figura 2. Struttura della telecamera omnidirezionale retinica.....	5
Figura 3. Spaccato dell'occhio. Immagine tratta da [1].....	8
Figura 4. Sezione dell'occhio. Immagine tratta da [2]	9
Figura 5. A. Sezioni istologiche della retina umana. Da 1 a 9 gli strati della retina con sclera e corioide. B.Fovea	11
Figura 6. Schema specifico funzionale dei coni e dei bastoncelli.	13
Figura 7. Flusso delle informazioni attraverso le cellule nervose sino al nervo ottico. Immagine tratta da [2].....	14
Figura 8. Dagli occhi della mosca al dettaglio del singolo Ommatide.	17
Figura 9. A sinistra, sezione trasversale di ommatidi, a destra sezione longitudinale della formazione retinulare.	17
Figura 10. Layout del sensore. La densità dei fotorecettori è notevolissima al centro.	19
Figura 11. Esempio Distribuzione dei fotorecettori. 64 anelli per 48 cerchi (fotorecettori).	19
Figura 12. Layout della fovea nel primo sensore retinico realizzato con tecnologia Cmos.	20
Figura 13. Schematizzazione degli anelli nel sensore e dei fotorecettori disposti sugli anelli. Notare la diminuzione della dimensione dei fotorecettori dall'anello piu' esterno verso il centro, e la costanza della dimensione all'interno della fovea.	21
Figura 14. Layout della fovea dell'ultimo sensore Cmos realizzato. I pixel hanno dimensione costante	21
Figura 15. Immagine log-polare acquisita con la telecamera retinica.	23
Figura 16. L'immagine log-polare trasformata in cartesiana appare ora facilmente interpretabile.	23
Figura 17. Viene rappresentato il mapping di un pixel dal piano cartesiano a quello log-polare	25

Figura 18. Per rappresentare l'informazione prodotta dal fotorecettore del sensore, viene replicato n volte un pixel nell'immagine cartesiana. Possiamo notare la transizione sensore- immagine log-polare- immagine cartesiana.....	26
Figura 19. A sinistra l' immagine log-polare, a destra l'immagine cartesiana rimappata.	26
Figura 20. In questa immagine si nota che i quadrati che vanno ad interessare la fovea del sensore in sede di mappaggio log-polare compaiono più grandi e deformati perché si collocano nelle prime righe dove la risoluzione è minore.	27
Figura 21. Mapping dei pixel nella fovea nell'immagine log-polare.	27
Figura 22. A sinistra la risoluzione minore, a destra la risoluzione maggiore	29
Figura 23. In alto la simulazione di un immagine ottenuta tramite un sensore tradizionale con lo stesso numero di pixel utilizzato nell'immagine in basso, ottenuta con il sensore retinico.	31
Figura 24. La figura esemplifica come i pixel corrispondenti ai fotorecettori nel sensore retinico vengano mappati nel piano log-polare specificando che in questo caso la risoluzione del sensore retinico e del sensore tradizionale è la medesima.	33
Figura 25. Il fotorecettore individua una zona angolare.....	34
Figura 26. Attraverso il foro passano solo i raggi direzionati su esso e si proiettano su a, b,c.....	38
Figura 27. La proiezione prospettica.	39
Figura 28. Il modello "weak perspective", $dz \ll z_0$	40
Figura 29. La proiezione ortografica.	40
Figura 30. Esempi di 2 telecamere che realizzano i 2 tipi di modelli prospettico e ortografico.....	42
Figura 31. Cross Rapporto, invariante proiettivo non elementare.....	44
Figura 32. Pattern con cerchi concentrici e corrispondente immagine log-polare. I cerchi del pattern vengono trasformati in righe.	45
Figura 33. Pattern circolare dopo lo spostamento lungo la direzione ortogonale all'asse ottico e corrispondente immagine log-polare. Le righe della Figura 29 nell'immagine log-polare compaiono ora come delle curve.	45
Figura 34. Modello di proiezione usato per l'acquisizione delle immagini con la telecamera retinica. L'asse ottico è ortogonale al piano inquadrato.....	46
Figura 35. Il segmento AP viene trasformato nel piano log-polare in una curva.....	46

Figura 36. Segmento P1P2 nel piano cartesiano (X,Y)	47
Figura 37. I due fotorecettori vicini coprono una zona angolare, i valori degli angoli nel piano log-polare sono discretizzati	48
Figura 38. Una volta traslato il segmento AP1 conserva l'angolo di tangenza quindi le forme corrispondenti nel piano log-polare sono uguali e evidenziano anch'esse una traslazione.....	49
Figura 39. In questo caso vi è una semplice traslazione del segmento ma gli angoli di tangenza sono uguali in modulo e non nel segno. La curva dopo la traslazione risulta simmetrica alla curva prima della traslazione.....	50
Figura 40. La rotazione dell'oggetto nel piano cartesiano si traduce in una traslazione dello stesso nell'immagine log-polare conservandone la forma e la dimensione.	51
Figura 41. L'oggetto cerchiato in rosso in seguito alla rotazione trasla lungo la riga dell'immagine lo-polare.....	51
Figura 42. In seguito all'avvicinamento della telecamera, la scalatura dei cerchi si traduce in una semplice traslazione delle righe nell'immagine log-polare	52
Figura 43. L'oggetto cerchiato in fucsia essendo identificato dai pixel nella fovea appare nel piano log-polare appare deformato nella forma.	53
Figura 44. Complesso con telecamera omnidirezionale e piano ortogonale all'asse ottico visualizzato	54
Figura 45. L'immagine omnidirezionale retinica in questo caso ha lo stesso comportamento dell'immagine retinica. Un movimento radiale o angolare si traduce in una traslazione dell'oggetto lungo la colonna o la riga. Anche in questo caso per la rotazione dell'oggetto rimane invariante la forma e la dimensione mentre per la traslazione radiale questo invariante non si verifica.	55
Figura 46. L'immagine omnidirezionale acquisita tramite un visore omnidirezionale con sensore tradizionale.....	55
Figura 47. L'immagine omnidirezionale log-polare appare già "unwarp" e non necessita di ulteriore elaborazione.....	56
Figura 48. Immagine di Gaussiane tridimensionali con 2 valori di varianza	72
Figura 49. La varianza del filtro gaussiano varia come la risoluzione. Imponiamo che il valore della varianza aumenti dal centro verso la periferia: $\sigma_1 < \sigma_2 < \sigma_3 < \sigma_4$	73

Figura 50. La dimensione del filtro usato per la convoluzione con l'immagine varia in relazione alla varianza della gaussiana.	73
Figura 51. Esempio di space variant edge detection su immagine cartesiana rimappata da log-polare	73
Figura 52. L'edge detection in questo caso viene effettuato sull'immagine log-polare e l'immagine con gli edge viene rimappata nella corrispondente cartesiana. ...	74
Figura 53. Risultato dell'operazione di 'morphing' dell'immagine degli edge cartesiana rimappata dall'immagine degli edge log-polare.	75
Figura 54. A sinistra l'immagine degli edge su immagine log polare con successivo remapping ed elaborazione. A destra l'immagine degli edge ottenuta con edge detection spazio variante.....	76
Figura 55. Schema che evidenzia come le distanze [DMIN, DMAX] vengano mappate nel ccd in [0,dMAX].	78
Figura 56. Conoscendo la distanza D e il valore dell'azimuth θ si può calcolare la posizione del punto P.....	79
Figura 57. Avendo il pixel P' un' area finita il suo valore viene discretizzato nel centro Pc e produce un insieme di valori Di nel piano tramite l'applicazione della funzione inversa $f^{*-1}(d)$	80
Figura 58. Il sensore retinico ha i fotorecettori non equamente spaziatu fuori della fovea mentre nella fovea la spaziatura è costante.....	81
Figura 59. Andamento dell'errore assoluto (dX) ed errore relativo(dX%)	81
Figura 60. Profilo di specchio conico e specchio conico con vertice sferico	82
Figura 61. Profilo dello specchio creato.	85
Figura 62. Immagine dello specchio in ambiente POV RAY.....	85
Figura 63. Visione del campo di gioco e di oggetti posti a diverse distanze dal centro del robot.	86
Figura 64. Profilo dello specchio creato.	87
Figura 65. Immagine dello specchio in ambiente POV RAY.....	88
Figura 66. Visione del campo di gioco e di oggetti posti a diverse distanze dal centro del robot.	88
Figura 67. Profilo dello specchio retinico.....	90
Figura 68. Immagine dello specchio retinico in ambiente POV RAY.	91
Figura 69. Schema di come vengono simulati i pixel di dimensione costante nell'immagine POV RAY calcolati in base al pixel del sensore retinico.	92

Figura 70. Visione del campo di gioco e di alcuni oggetti posati a diverse distanze su di esso ottenuta con uno specchio spazio variante e un sensore tradizionale simulato con POV-RAY.	92
Figura 71. Immagine del campo di gioco trasformata nella corrispondente cartesiana spazio variante.	93
Figura 72. L'oggetto cui vogliamo calcolare la posizione è quello cerchiato, per risalire alla distanza dal centro valutiamo la riga dell'immagine a cui corrisponde il pixel dell'oggetto, e accediamo alla tabella delle distanze Di; mentre per l'azimuth valutiamo la colonna di appartenenza del pixel e moltiplichiamo per la risoluzione angolare.	94
Figura 73. Interfaccia grafica del software di acquisizione delle immagini.	98

INDICE DELLE TABELLE

Tabella 1. Caratteristiche peculiari dei coni e dei bastoncelli	13
Tabella 2. Dettagli tecnici del sensore retinico e della telecamera Giotto.....	22
Tabella 3. Parametri caratteristici del sensore di ultima generazione.....	30
Tabella 4. Confronto tra i due tipi di specchi analizzati.	82
Tabella 5. Parametri progettuali e visivi dello specchio	84
Tabella 6. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco...	86
Tabella 7. Parametri progettuali e visivi dello specchio	87
Tabella 8. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco...	88
Tabella 9. Parametri progettuali e visivi dello specchio.	90
Tabella 10. Risultati della misura di distanze di 3 oggetti posti sul campo di gioco.	93
Tabella 11. Funzioni abilitate dal software di acquisizione delle immagini.	97

BIBLIOGRAFIA

- [1] “L’occhio, i grandi temi della medicina”, Fabbri Editore.
- [2] G. Chiarugi, “Istituzioni di anatomia dell’uomo”, Vallardi Editore.
- [3] Georgii A. e Mazokhin-Porshnyakov, “Insect Vision”.
- [4] James Missler e Farhad A. Kamangar, “A neural network for pursuit tracking inspired by the fly visual system“. *Neural Networks*, 8 (1995),no 3, 463-480.
- [5] Giulio Sandini e Giorgio Metta, “Retina like sensors: motivations, technology and application. (2002).
- [6] Consorzio Unitek, “Giotto, retina like camera”.
- [7] Lars Schwabe, Klaus Obermayer, “Modeling the adaptive visual system: a survey of principled approaches”. *Neural Networks*, 16 (2003). 1353-1371
- [8] Michael Brady e Barrow H., “Computer Vision”. Prentice-Hall, Englewood Cliffs, 1982.
- [9] E. Trucco and A. Verri. “Introductory Techniques for 3-D Computer Vision”. Prentice Hall, Inc., Upper Saddle River, New Jersey 07458, 1998.
- [10] Isaac M. Jaglom, “Trasformazioni geometriche: le isometrie”. *Matematica moderna*, Zanichelli,1983.
- [11] Eugenio Bertini, “Complementi di geometria proiettiva”, Zanichelli 1927.
- [12] Mayer Frane, “Rapporto sui progressi della geometria proiettiva nell’ultimo 4° secolo”.
- [13] L. Mundy, Andrew Zisserman, “Invariance in computer vision”, The Mit Press, Cambridge, Massachusetts, London, England, 1992.
- [14],Konrad Schindler, “Construction and Detection of straight lines, distances, and circles in log polar images”, *Computer Graphics and Vision*, TU, Graz Inffeldgasse 16, from *Omnivis 2004*. P.Sturm, T.Svoboda and S.Teller (eds.)
- [15] J. Canny A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 8, No. 6, Nov 1986.
- [16] Marchese F., Sorrenti D.G. Omni-Directional Vision with a Multi-Part Mirror. 4th International Workshop on RoboCup, pp. 289-298, 2000.

[17] E. Menegatti, F. Nori, E. Pagello, C. Pellizzari, and D. Spagnoli. Designing an omnidirectional vision system for a goalkeeper robot. In A. Birk, S. Coradeschi, and S. Tadokoro, editors, RoboCup-2001: Robot Soccer World Cup V., pages pp. 78–87. Springer, 2002.

[18] Tomhas Pajdla and Hadas Roth, Panoramic imaging with SVAVISCA camera - simulations and reality, Ctu-cmp-2000-16, Center for machine perception - Czech Technical University, October 2000.

Y. Yagi, “Map based Navigation for a mobile robot with Omnidirectional Image Sensor” COPIES “IEEE Trans. On robotics and automaton, vol. 11, no. 5, October 1995.

V. Srinivasan, “A New Class of Mirrors for Wide-Angle Imaging” Mandyam Centre for Visual Sciences, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra, ACT 2601, Australia

Stefan Gachter, “Mirror Design for an Omnidirectional Camera with a Uniform Cylindrical Projection Using the SVAVISCA Sensor” OMNIVIEWS – Omnidirectional Visual System FP5 RTD – FET Project No: IST–1999–29017 The technical description of the deliverable D2:”Mirror implementation”.

Ales Ude, Christopher G. Atkeson, Gordon Cheng, “Combining peripheral and foveal humanoid vision to detect, pursue, recognize and act”, In: Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Las Vegas, USA, October 2003, pp. 2173-2178

F.L. Lim*, S. Venkatesh, G.A.W. West, “Resolution consideration in spatially variant sensors”, Curtin University of Technology, Department of Computer Science, GPO Box U/987. Perth 6001. Australia

A.M. Wallace *, D.J. McLaren, “Gradient detection in discrete log-polar images”, Department of Computing and Electrical Engineering, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK

Yoshio Matsumoto, Kazunori Ikeda, Masayuki Inaba and Hirochika Inoue, "Visual Navigation using Omnidirectional View Sequence", Graduate School of Information Science, Nara Institute of Science and Technology 8916-5, Takayamacho, Ikoma-city, Nara, 630-0101, Japan

R. Manzotti, A. Gasteratos, G. Metta, and G. Sandini, "Disparity Estimation on Log-Polar Images and Vergence Control", Laboratory for Integrated Advanced Robotics (LIRA-Lab), Department of Communications Computer and System Science, University of Genoa, Viale Causa 13, Genoa I-16145, Italy

RINGRAZIAMENTI

Vorrei dedicare questa tesi principalmente alla mia famiglia, a mio padre Bruno, a mia madre Lucia e a mia sorella Cinzia che mi hanno appoggiato e mi hanno permesso di completare questo mio lungo percorso di studi. Vorrei ringraziare inoltre la mia ragazza Claudia che mi ha sostenuto moralmente in questi ultimi anni di università. Ringrazio infine il dr. Emanuele Menegatti ed il prof. Pagello Enrico che mi hanno permesso di sviluppare questa tesi e di lavorare all'interno dei laboratori IAS LAB, oltre al dr. Enrico Grisan, il dr. Alvisè Lastra e al dr. Fabio Berton (dell'università di Genova) per il prezioso aiuto datomi.