

UNIVERSITÀ DEGLI STUDI DI PAVIA
FACOLTÀ DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

RICERCA DI CARATTERISTICHE
DI AMBIENTI STRADALI
IN IMMAGINI ALL'INFRAROSSO
PER APPLICAZIONI AUTOMOBILISTICHE

Relatore: Chiar.mo Prof. Virgilio Cantoni

Correlatore: Chiar.mo Prof. Alberto Broggi

Candidato: Pietro Cerri

ANNO ACCADEMICO 2001-02

Ai mie genitori

Indice

Introduzione	1
1 Analisi del problema	6
1.1 Tipologie di strada	7
1.2 I pali e le costruzioni	8
1.3 I segnali stradali	9
1.4 Le immagini all'infrarosso	11
1.5 La telecamera	14
2 Funzionalità sviluppate	17
2.1 Ricerca della strada e del punto di fuga	20
2.1.1 La trasformata di Hough	20
2.1.2 Le limitazioni e le operazioni successive	22
2.1.3 Le variabili di confidenza	26
2.1.4 Un metodo alternativo	31
2.2 Ricerca di Pali e di bordi di edifici	33
2.3 Ricerca di segnali stradali	37
3 Risultati e conclusioni	40
3.1 Risultati	40

3.1.1	La strada e il punto di fuga	41
3.1.2	I pali e gli edifici	49
3.1.3	Cartelli circolari	54
3.1.4	Immagini a varie temperature	58
3.1.5	Tempi di esecuzione	59
3.2	Conclusioni	61
3.3	Possibili sviluppi futuri	62
A	GOLD	64
B	Manuale utente	68
	Bibliografia	74

Elenco delle figure

1	Linee dell'orizzonte	3
1.1	Vari tipi di strade	8
1.2	Pedone con bordi verticali lunghi e continui	9
1.3	Segnali deformati e segnali di varie dimensioni	10
1.4	Pedone con bordi arrotondati	11
1.5	Stagioni diverse	13
1.6	Ombre	13
1.7	Zone dell'immagine inutilizzabili	14
1.8	Sfocatura durante le curve	15
2.1	I bordi	19
2.2	La trasformata di Hough	21
2.3	I voti nel dominio di Hough	23
2.4	Zone di ricerca e pendenza delle rette	24
2.5	Linee trovate dai metodi di Hough	25
2.6	Il risultato finale del metodo di ricerca della strada	28
2.7	Diagramma di flusso del metodo di ricerca del punto di fuga	30
2.8	Risultato del metodo alternativo di ricerca della strada	32
2.9	Bordi verticali	33
2.10	Le soglie per la ricerca dei pali	35

2.11	Soglia variabile per la ricerca dei pali	36
2.12	Risultato del metodo di ricerca dei pali	37
2.13	Risultato del metodo di ricerca dei segnali stradali	39
3.1	I vari tipi di strade riconosciute dal metodo	41
3.2	Una sola automobile parcheggiata	43
3.3	Strisce pedonali	44
3.4	Curva	45
3.5	Curva lieve	46
3.6	Metodo alternativo sulle curve	47
3.7	Metodo alternativo su tratti rettilinei	48
3.8	Risultati del metodo di ricerca dei pali	49
3.9	Sequenza di pali correttamente riconosciuta	50
3.10	Pali non correttamente riconosciuti	51
3.11	Confronto col metodo di ricerca dei pedoni	52
3.12	Pedone con bordi verticali lunghi e continui	53
3.13	Cartello identificato a intermittenza	54
3.14	Cartello di forma ellittica	55
3.15	Cartello in avvicinamento	55
3.16	Errori del metodo di ricerca dei segnali	56
3.17	Confronto col metodo di ricerca dei pedoni	57
3.18	Risultati con immagini a diverse temperature	58
3.19	Neve ai bordi della strada	59
A.1	Pannello di controllo di GOLD	66
B.1	Grafico dei tempi di esecuzione delle funzioni	71
B.2	Pannello di controllo delle funzioni	72

Elenco delle tabelle

3.1	Tempi di esecuzione	60
B.1	Funzione delle variabili del pannello di controllo	71

Introduzione

In questa tesi si sono sviluppate alcune funzionalità per studiare ambienti stradali, per lo più urbani, ed identificarne alcune caratteristiche. Il codice, in linguaggio C, è stato prodotto per un sistema composto da un'automobile fornita di una telecamera all'infrarosso e di un PC in grado di elaborare le immagini in arrivo o di salvarle su disco per poterle elaborare successivamente. La tesi è stata sviluppata nel laboratorio di visione dell'università di Parma ed è inserita all'interno del progetto ARGO [1] per lo sviluppo di un sistema in grado di condurre un veicolo stradale opportunamente modificato; in particolare la tesi si sviluppa nell'ambito del riconoscimento di pedoni.

La complessità di un tale progetto è evidente, si tratta di fornire ad un PC tutte le informazioni dell'ambiente stradale in cui la macchina si trova, sia esso una strada urbana o un'autostrada, e di fornire l'esatta posizione e natura degli ostacoli presenti. Successivamente il computer deve poter agire sulla guida dell'automobile e modificare adeguatamente velocità, direzione, e se necessario altri parametri, per adeguarsi alle informazioni fornite dalla telecamera o da altri dispositivi di acquisizione. Un progetto di questo tipo comporta evidenti problemi di sicurezza, sia per i viaggiatori che per i pedoni, che richiedono una particolare attenzione ad ogni particolare e ad ogni possibile evenienza ed un'attenta fase di test. I risultati in questo campo sono ancora altamente sperimentali e, nonostante gli sforzi della maggior parte delle case automobilistiche del mondo, non si sono ancora raggiunti dei risultati tali da essere proposti sul mercato, tant'è che non esistono ancora leggi che stabiliscano tutti gli aspetti legali

relativi alla responsabilità in caso di difetti del sistema o comportamenti non corretti.

Le immagini all'infrarosso non sono, come semplicisticamente si può credere, la soluzione ideale e totale di ogni problema di visione, infatti anche esse risentono di alcuni problemi irrimediabilmente legati al fenomeno fisico su cui si basa la tecnica. Un particolare problema che si è incontrato nel corso del lavoro è la pesante dipendenza delle immagini dall'illuminazione dell'ambiente ripreso, questo a causa della natura stessa della radiazione infrarossa, che lega la sua lunghezza d'onda alla temperatura dell'oggetto da cui viene emessa. I risultati sono stati verificati con sequenze di immagini riprese in diversi periodi dell'anno ed in diverse ore della giornata per assicurarsi che i programmi funzionino correttamente in qualsiasi condizione, sia con immagini fredde e con poco contrasto, sia con immagini calde con ombre molto evidenti.

L'ambiente stradale, in particolare quello urbano, può essere particolarmente complesso. Se in alcune strade di scorrimento gli elementi presenti nelle immagini sono pochi e piuttosto semplici da considerare, in zone più frequentate la presenza di pedoni, automobili parcheggiate, oggetti dell'arredamento stradale, rende particolarmente complessa l'elaborazione. Il programma è stato sviluppato tenendo conto anche della complessità delle immagini riprese, ed è stato testato su vari tipi di immagini.

Nel corso della tesi saranno usati più volte dei termini propri della prospettiva, non sempre però si darà a tali parole il senso che esse hanno nella teoria classica della prospettiva [5]. Quando si parlerà del punto di fuga si intenderà solo quello relativo alla strada che l'automobile sta percorrendo, dunque parlare di punto di fuga in immagini di curve risulta non corretto, benchè tale punto esista anche in queste immagini. Inoltre osservando i risultati si può notare che quasi sempre il punto di fuga anzichè appartenere all'orizzonte, come ci si aspetterebbe, si troverà un poco sopra; questa anomalia è causata da un errore di calibrazione della telecamera, che avviene prima delle elaborazioni del programma sviluppato, questo errore è in parte causato anche dalla particolare posizione della telecamera, che si trova molto in basso, e che

amplifica qualsiasi errore di calibrazione; nell'immagine (fig.1) sono mostrate le due linee: quella più chiara rappresenta l'orizzonte trovato col metodo di calibrazione di GOLD, quella più scura l'orizzonte trovato con il mio metodo.



Figura 1: Linee dell'orizzonte

Le funzionalità sviluppate sono principalmente tre:

- Ricerca del punto di fuga e della strada
- Ricerca dei pali e degli angoli verticali degli edifici
- Ricerca dei segnali stradali rotondi

Una sola funzione lavora direttamente sulle immagini fornite dalla telecamera, le altre si basano sull'individuazione dei contorni, che viene eseguita col metodo di Sobel, che fornisce, oltre al modulo del bordo, anche la fase, che diventa molto utile in tutte quelle applicazioni che necessitano la conoscenza dell'inclinazione del bordo.

Determinare la posizione della carreggiata risulta estremamente utile per capire la conformazione dell'ambiente stradale, e per qualsiasi successiva elaborazione; conoscere

la zona dell'immagine dove l'automobile si può spostare e dove i pedoni sono pericolosi è essenziale ai fini di evitare un incidente. Ciò che intendiamo trovare è una proiezione della strada lungo le sue linee laterali che vengono individuate nelle vicinanze dell'automobile, fino al punto di fuga, cioè un punto immaginario in cui si incontrano le linee formate dai bordi della strada, che ci appaiono incidenti, anziché parallele, a causa della prospettiva. Questo metodo, naturalmente, non funziona sulle curve, per questo si è provato ad inserire un metodo alternativo che non lavora sui bordi, ma sull'immagine principale, e cerca la strada come una zona con luminosità media simile.

Per trovare la conformazione della strada è stata utilizzata la trasformata di Hough; con questo metodo sono state trovate le linee che più probabilmente appartengono al bordo della strada, successivamente sono stati sperimentati vari metodi per scegliere quelle migliori ed il relativo punto di fuga.

Identificare pali e angoli verticali degli edifici risulta particolarmente utile nella ricerca dei pedoni perché permette di evitare che tali linee vengano considerate come pedoni a causa dal metodo di ricerca che favorisce gli oggetti con forte simmetria e con molti bordi verticali. Il nostro obiettivo principale è dunque identificare tutte le linee verticali che appartengono ad oggetti e non a persone, rientrano dunque in questa categoria pali, bordi di costruzioni, alberi, e tutti gli oggetti con bordi verticali lunghi, come cabine telefoniche e autobus. Un obiettivo, per altro secondario, sarebbe quello di distinguere tra tutti questi oggetti per avere una miglior conoscenza dell'ambiente.

Per la ricerca dei pali e dei bordi degli edifici si è utilizzato un metodo simile alla trasformata di Hough ma molto più veloce, , che lavorando sui bordi verticali, permette di considerare le linee verticali lunghe e continue, scandendo le colonne dell'immagine e non i pixel uno per uno.

La ricerca di segnali stradali non è legata tanto alla volontà di interpretare i segnali stessi, obiettivo difficilmente raggiungibile con una telecamera ad infrarossi, ma alla

necessità di distinguere i cartelli dai pedoni: i segnali rotondi infatti vengono spesso considerati come la testa di un pedone e inseriscono un errore difficilmente rilevabile dalla procedura di ricerca dei pedoni. In questo caso non si vuole generalizzare la ricerca a qualsiasi elemento quasi circolare che non appartenga alla figura del pedone perché il rischio di individuare anche le teste delle persone è troppo alto.

Per trovare i segnali stradali è stata utilizzata la trasformata di Hough generalizzata per la ricerca di forme circolari e alcuni accorgimenti per eliminare numerosi oggetti di forma quasi circolare che venivano considerati come segnali.

Il software è stato sviluppato all'interno di GOLD¹, una piattaforma per lo sviluppo di applicazioni per la visione realizzata grazie ad una collaborazione tra l'università di Parma e quella di Pavia.

Questa tesi è organizzata in 4 capitoli. Nel primo capitolo vengono analizzati e spiegati i problemi che si sono incontrati nel corso del lavoro, sia quelli legati all'ambiente stradale, sia quelli legati alle immagini all'infrarosso. Nel secondo capitolo vengono spiegate le funzionalità sviluppate e le tecniche utilizzate per raggiungere gli obiettivi prefissati. Nel terzo capitolo si trovano i risultati ottenuti con alcune riflessioni rispetto a quanto realizzato e ai possibili sviluppi futuri. Seguono due appendici, la prima presenta brevemente GOLD, la seconda è il manuale di utilizzo del pannello di controllo del software.

¹Acronimo di Generic Obstacle and Lane Detection

Capitolo 1

Analisi del problema

I problemi che si sono affrontati nello sviluppo della tesi sono piuttosto vari e complessi; essi riguardano sia la complessità e la diversità dei vari tipi di ambienti stradali, sia i problemi relativi alle immagini all'infrarosso, sia i problemi introdotti dalla telecamera a nostra disposizione. Di seguito sono esposti più chiaramente i problemi citati; nella presentazione si è volutamente invertito l'ordine con cui i problemi vengono affrontati dal software, per presentare per primi i problemi che hanno richiesto una maggior attenzione, cioè i problemi di visione in senso stretto, verranno quindi esposti nell'ordine:

- Problemi relativi alle immagini e alla loro complessità:
 - nella ricerca della strada,
 - nella ricerca di pali e bordi di edifici,
 - nella ricerca di segnali stradali;
- Problematiche dell'infrarosso;
- Difetti delle telecamera.

1.1 Tipologie di strada

Gli ambienti stradali esistenti sono molto numerosi e difficilmente riconducibili ad un unico o a più standard; l'esempio più immediato di questo fatto è dato dalla larghezza della carreggiata, che varia a seconda dell'importanza della strada che si percorre, ma anche per motivi costruttivi.

In questa tesi mi sono concentrato soprattutto su ambienti stradali urbani, la cui complessità e varietà è superiore rispetto alle strade extraurbane a causa del gran numero di oggetti che si possono incontrare e dei vari ambienti in cui le strade si trovano. Uno dei problemi maggiori che mi sono trovato ad affrontare è legato ai diversi modi in cui una strada può essere delimitata, questione molto importante visto che un obiettivo del lavoro è la ricerca della strada: segue un elenco dai più frequenti metodi di delimitazione che ho incontrato.

Tramite linee: presenta bordi dritti e netti facilmente individuabili (fig. 1.1.a).

Tramite Marciapiede: si tratta di una delimitazione abbastanza dritta ma non sempre netta (fig. 1.1.b).

Tramite macchine posteggiate: complessa a causa dei bordi frastagliati (fig. 1.1.c).

Nessuna delimitazione: la strada finisce nell'erba o nella terra, raro e piuttosto complesso.(fig. 1.1.d)

Naturalmente si possono trovare delimitazioni diverse sui due lati della strada e, a volte, anche sullo stesso lato, come ad esempio quando si ha un tratto con auto parcheggiate e un tratto senza auto. E' difficile trovare un solo modello che accomuni tutte questi tipi di strade, se non la forma che in tutti i casi risulta alla telecamera praticamente triangolare, con una base che occupa tutto il bordo inferiore dell'immagine e l'angolo superiore praticamente nel centro dell'immagine.

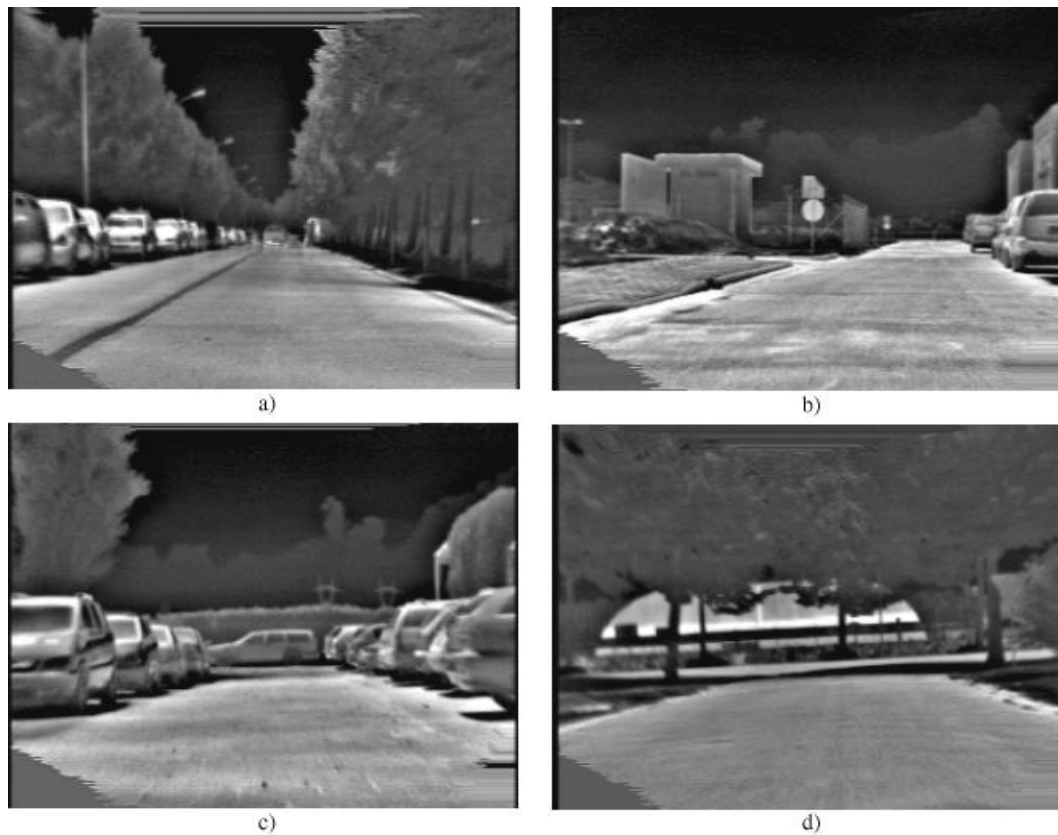


Figura 1.1: Vari tipi di strade
a) con linee, b) con marciapiede, c) con automobili, d) senza delimitazioni

Sulla strada si trova spesso del rumore dovuto alle imperfezioni della strada, che anche utilizzando un filtro anti-rumore difficilmente vengono eliminate, e alla presenza di segnaletica orizzontale (come i passaggi pedonali) che nel riconoscimento della strada sono degli elementi inattesi e di disturbo. Questo tipo di rumore, genera non pochi problemi, procurando il riconoscimento di linee di bordo inesistenti.

1.2 I pali e le costruzioni

La ricerca di pali e angoli di edifici non risulta particolarmente complessa, visto che si riduce all'individuazione di bordi verticali di una certa lunghezza; d'altra parte le

differenze tra i bordi dei pali e quelli degli edifici risultano con questo metodo scarse e non è certo facile distinguere tra i due oggetti, tale obiettivo è comunque secondario.

Considerare le linee lunghe pone alcuni problemi nell'individuare i pali corti, sia i pali più bassi, come quelli dei segnali stradali, sia quelli lontani.



Figura 1.2: Pedone con bordi verticali lunghi e continui

Il vero problema però si incontra quando ci sono pedoni che hanno dei bordi verticali piuttosto lunghi e continui, in questo caso i bordi del pedone, la schiena nella figura riportata (fig. 1.2), rischiano di essere considerati come pali, e questo è proprio l'opposto dell'obiettivo che ci siamo posti.

1.3 I segnali stradali

Se la ricerca di forme geometriche perfette (sintetiche o meno) risulta molto facile soprattutto in ambienti semplici, cercare delle forme non perfette in ambienti complessi è molto più complicato e richiede una grande attenzione alla scelta delle soglie



Figura 1.3: Segnali deformati e segnali di varie dimensioni

da utilizzare e alla definizione di metodi di appoggio, visto che il solo metodo della trasformata di Hough generalizzata rischia di fallire.

Il primo problema che si incontra nella ricerca dei cerchi è quello di avere delle circonferenze non perfette, sia per il bordo non regolare, causato da disomogeneità dello sfondo o da scarso contrasto con lo sfondo, sia, in primo luogo, dall'inquadratura non perfettamente frontale dei cartelli che induce delle deformazioni che tendono a trasformare i cerchi in ellissi. Inoltre non abbiamo informazioni sicure riguardo la dimensione dei cartelli: benché essa dovrebbe essere standard, ma con un po' di attenzione ci accorgiamo che questo non è sempre vero, nell'immagine varia a seconda della distanza. Anche l'altezza a cui sono posti i cartelli non è fissa, tant'è vero che possono esserci ad esempio cartelli sovrapposti sullo stesso palo, questo fatto comporta la necessità di espandere verticalmente la zona di ricerca dei segnali.

Se i problemi finora esposti spingerebbero ad utilizzare dei metodi meno precisi con soglie molto basse per identificare tutti i cartelli, vi sono problemi che spingono nel verso opposto. Nelle immagini di ambienti stradali compaiono molti oggetti di forma



Figura 1.4: Pedone con bordi arrotondati

quasi circolare che non sono cartelli e che dunque non devono essere considerati, basti pensare alle teste dei pedoni (fig. 1.4) o agli specchietti delle autovetture. Come già detto per il metodo di ricerca dei pali l'errore peggiore si verifica quando la testa di un pedone, che ha una forma quasi circolare, viene identificata come un cartello stradale: gli sforzi di affinamento della funzionalità vanno proprio in questa direzione.

1.4 Le immagini all'infrarosso

Le telecamere all'infrarosso riproducono sull'immagine la temperatura degli oggetti inquadrati analizzando le onde della radiazione infrarossa[2], nel nostro caso si tratta delle onde del lontano infrarosso (*FIR: Far InfraRed*) la cui lunghezza d'onda varia da $7\mu\text{m}$ a $20\mu\text{m}$. Si riescono dunque a distinguere oggetti di materiali diversi, di colori diversi, o comunque a temperature differenti; le immagini in scala di grigi su cui ho lavorato visualizzano queste differenze assegnando maggior luminosità (bianco) ad oggetti più caldi e minor luminosità (nero) ad oggetti freddi. Oggetti con colori chiari,

in particolare il bianchi, saranno riportati nell'immagine con tonalità scure perché i colori chiari riflettono quasi tutti i raggi solari e tendono a scaldarsi meno, il contrario accadrà ad oggetti scuri, ed in particolare neri. Osservando un'immagine qualsiasi che riporti una strada si nota come le linee bianche dei bordi siano più scure dell'asfalto nero, osserveremo quindi una sorta di negativo dell'immagine originale. Anche i diversi materiali si scaldano diversamente sotto l'illuminazione solare, oggetti metallici si scaldano molto e molto velocemente, mentre gli alberi non raggiungono mai temperature troppo elevate. La temperatura varia anche al variare dell'inclinazione con cui i raggi solari colpiscono un oggetto, più il bordo è perpendicolare, più verrà scaldato, infatti se si osservano attentamente gli edifici si può notare che i tetti sono più chiari dei muri. Esistono anche oggetti che non risentono dell'illuminazione solare, perché la loro temperatura è molto alta e perché sono essi stessi fonte di calore, rientrano in questa categoria le sigarette accese, che possono essere riconosciute come dei punti bianchi sull'immagine, e alcune zone delle automobili, come le marmitte, che raggiungono temperature elevatissime. Un caso a parte è rappresentato dai pedoni che pur avendo una propria temperatura, non elevata come nei casi precedenti, risentono comunque degli effetti del sole, infatti un pedone sotto il sole risulterà molto più chiaro di un pedone all'ombra.

Esiste una grande differenza tra immagini registrate con condizioni atmosferiche differenti o in stagioni differenti o in ore diverse della giornata a causa della diversa temperatura degli oggetti e della presenza di ombre.

Come risulta dall'esempio riportato (fig. 1.5) il contrasto tra gli oggetti varia drasticamente al variare delle stagioni, tanto che molti particolari evidenti nell'immagine più calda sono difficilmente localizzabili in quella fredda in cui gli oggetti assumono la temperatura dell'ambiente circostante; un fenomeno di questo tipo si verifica anche quando abbiamo degli oggetti che si trovano all'ombra.

Non sempre però immagini più calde e meglio illuminate sono migliori: nell'e-



Figura 1.5: Stagioni diverse
a) Inverno, b) Estate

sempio seguente (fig. 1.6) risultano evidenti nell'immagine estiva delle ombre che non introducono alcuna informazione utile, ripetendo semplicemente la forma degli oggetti che le proiettano, anzi possono introdurre degli errori, soprattutto nella ricerca dei bordi della strada.

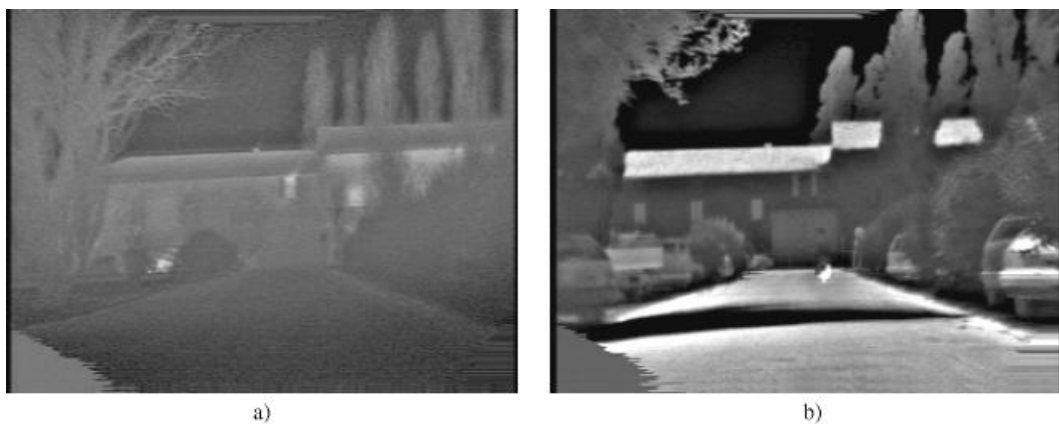


Figura 1.6: Ombre
a) Inverno, b) Estate

1.5 La telecamera

La telecamera con cui sono state acquisite le sequenze su cui si è lavorato, pur essendo di tipo professionale, ha alcuni difetti. La tecnologia delle telecamere ad infrarosso ha raggiunto degli ottimi risultati, ma i modelli più avanzati rimangono comunque confinati ad applicazioni militari; le telecamere in commercio, per quanto di buon livello, sono inferiori per prestazioni a quelle militari, fortunatamente però il loro prezzo è decisamente inferiore a quello degli apparati usati dall'esercito.

Analizzando le immagini si notano diverse imperfezioni. Vi sono alcune parti in cui non passa informazione: sui bordi, che sono neri, e nella zona in basso a sinistra, grigia scura. Vi sono inoltre delle zone in cui l'immagine è irrimediabilmente corrotta, come in basso a destra e in alto centralmente, dove, lungo la riga, viene riportato il colore della colonna più a sinistra (fig. 1.7); mentre i bordi laterali sono inseriti da imperfezioni nel sistema di acquisizione, le zone triangolari dove l'immagine è inutilizzabile sono causate da un difetto del sensore, che in quelle zone è irrimediabilmente rovinato.



Figura 1.7: Zone dell'immagine inutilizzabili

In tali zone dell'immagine è inutile lavorare, perché non abbiamo alcuna informazione sensata, anzi, spesso tali zone introducono un errore che può inficiare il buon funzionamento del programma, ad esempio i bordi neri laterali possono essere considerati come linee verticali. Mentre i bordi o la zona in alto non sarebbero stati comunque utilizzati, le zone laterali in basso potevano contenere informazioni utili per il riconoscimento dei bordi della strada, questo fatto introduce un'ulteriore complicazione alla ricerca della strada.

E' presente inoltre un altro difetto, quando l'immagine si sposta velocemente lateralmente, come nelle curve, le immagini risultano sfocate, questo a causa del tempo di apertura dell'obiettivo della telecamera.



Figura 1.8: Sfocatura durante le curve

Non certo vantaggiosa per la ricerca della strada é la posizione della telecamera, essa é infatti situata nella mascherina: questa posizione, molto adatta alla ricerca di ostacoli verticali come pedoni o pali, sfavorisce la ricerca di oggetti o forme sviluppate orizzontalmente come la strada che appaiono decisamente deformate e di cui vengono fornite poche informazioni: la strada occupa meno della metà dell'immagine; il lavoro

sarebbe molto più semplice con uno strumento di acquisizione posto più in alto. La stessa angolazione della telecamera, quasi parallela alla terra, crea dei problemi: in questa posizione è molto difficile riuscire a calibrare correttamente la telecamera, e perciò a definire la posizione dell'orizzonte, inoltre tutte le oscillazioni dell'automobile vengono purtroppo amplificate.

Capitolo 2

Funzionalità sviluppate

Per il raggiungimento degli obiettivi esposti nell'introduzione, tenendo conto dei problemi esposti nel capitolo precedente, sono state sviluppate diverse applicazioni, principalmente tre: una per la ricerca della strada, una per la ricerca dei pali e una per l'individuazione dei segnali stradali. La prima applicazione è di gran lunga la più importante perché fornisce informazioni generali sull'ambiente che possono essere utilizzate dalle altre, e vengono in effetti utilizzate dal metodo di ricerca dei pali.

Prima di elaborare le immagini è stato sempre applicato un filtro per eliminare il rumore, presente soprattutto sulla strada, e per rendere più nitidi i contorni: si è scelto di utilizzare un filtro mediano. Il filtro scandisce ogni pixel dell'immagine ed opera su una maschera 3 per 3 ordinando i 9 valori secondo la luminosità, e restituisce in uscita il valore del quinto elemento, cioè quello di mezzo. Questo filtro, per quanto piuttosto pesante, dà un buon risultato nell'eliminazione dei pixel fuori norma e dunque del rumore *salt & pepper* e nel mantenimento dei bordi nitidi, rendendo decisamente più semplici le operazioni successive. In realtà il filtro non elimina perfettamente il rumore presente sulla strada ed inoltre, uniformando alcune zone, può causare la formazione di bordi precedentemente inesistenti, creando in questo modo alcuni problemi; i casi di errore sono comunque molto rari e non pregiudicano eccessivamente gli ottimi risultati

del filtro.

Per la ricerca dei contorni è stato utilizzato il metodo di Sobel, questa scelta è giustificata dalla necessità di ottenere dei contorni chiari e di conoscere l'inclinazione dei bordi trovati: il filtro di Sobel utilizza infatti due maschere, una per la ricerca di bordi verticali e una per la ricerca di bordi orizzontali.

$$mod_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$mod_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Tramite la formula

$$modulo = \sqrt{mod_x^2 + mod_y^2}$$

dove mod_x e mod_y rappresentano i bordi verticali e orizzontali trovati, oppure, per velocizzare l'operazione, tramite la formula

$$modulo = |mod_x| + |mod_y|$$

si ottiene il modulo, cioè il contrasto del bordo, mentre applicando questa formula

$$fase = \frac{mod_x}{mod_y}$$

si ottiene la fase, cioè l'inclinazione del bordo. Per poter consultare velocemente questi valori si sono costruite due immagini, una per il modulo e l'altra per la fase, la prima è una matrice di interi, la seconda una matrice di valori in virgola mobile.

Per ottenere solo i bordi più interessanti, cioè quelli con maggior contrasto, si è utilizzata una soglia mobile in grado di selezionare sempre una certa percentuale di pixel dell'immagine. La soglia è stata calcolata scandendo tutta l'immagine del modulo e contando il numero di pixel presenti per ogni valore del modulo, tali numeri sono stati salvati in un array di 256 elementi; successivamente, partendo dall'ultimo elemento e avvicinandosi al primo, si sono sommati i numeri di pixel presenti fino ad arrivare ad un valore pari alla percentuale scelta e si è salvato l'indice dell'elemento raggiunto, tale indice rappresenta la soglia da utilizzare. I bordi con modulo superiore alla soglia sono stati salvati in un'immagine che è stata successivamente binarizzata per essere facilmente utilizzabile e visualizzabile (fig.2.1).

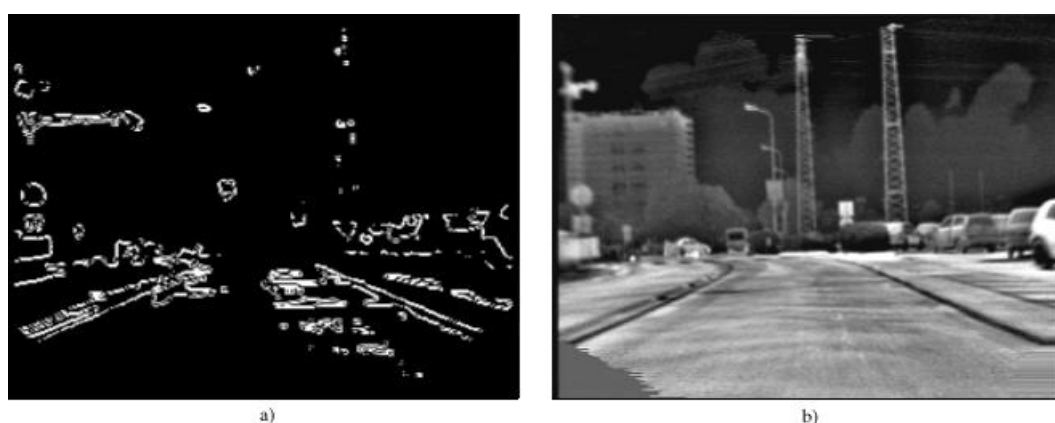


Figura 2.1: I bordi

Le zone di immagine inutilizzabili (vedi 1.5) sono state eliminate utilizzando una maschera per considerare solo le zone che contengono informazioni corrette; per poter applicare questo metodo sia alle immagini registrate con la telecamera attualmente in uso, sia con la telecamera utilizzata precedentemente, che aveva problemi simili ma in zone leggermente diverse, le dimensioni della maschera sono state mantenute variabili.

2.1 Ricerca della strada e del punto di fuga

Per risolvere problemi simili alla ricerca della carreggiata stradale sono presenti in letteratura diversi metodi, che utilizzano la ricerca di bordi o la ricerca di colori o pattern simili ad un modello. Entrambi i tipi di metodo sono stati provati, separatamente o insieme, ma, analizzando i risultati, si è deciso di privilegiare un metodo di ricerca dei bordi, che risente meno del problema delle ombre, e che ha dato buoni risultati anche in condizioni non ottimali; si è per altro mantenuto un controllo sulla ricerca della strada tramite luminosità dei pixel, ma solo come controllo per le imprecisioni del metodo scelto.

In seguito faremo uso di una variabile fornita da GOLD che misura la posizione dell'orizzonte sull'asse delle ordinate, come già spiegato nell'introduzione il punto di fuga si troverà al di sopra di questa linea di circa cinque pixel, su una linea che sarà da considerare il vero orizzonte.

2.1.1 La trasformata di Hough

Il metodo scelto è quello della trasformata di Hough che consente di trasportare le rette dell'immagine nel dominio di Hough che ha come assi il modulo della distanza della retta da un punto fisso ρ , nel nostro caso il punto in basso a sinistra, e l'inclinazione del segmento della distanza θ (fig.2.2).

Scandendo l'immagine, tramite alcuni algoritmi si può definire se un punto appartiene ad una certa retta, identificata da un ρ ed un θ , e votare per quella retta: una retta che ottiene molti voti corrisponderà ad un bordo dritto. A questo punto si potrà risalire dal dominio di Hough alla retta nell'immagine tramite la formula:

$$y = \frac{\rho}{\sin \theta} + \frac{x}{\tan \theta}$$

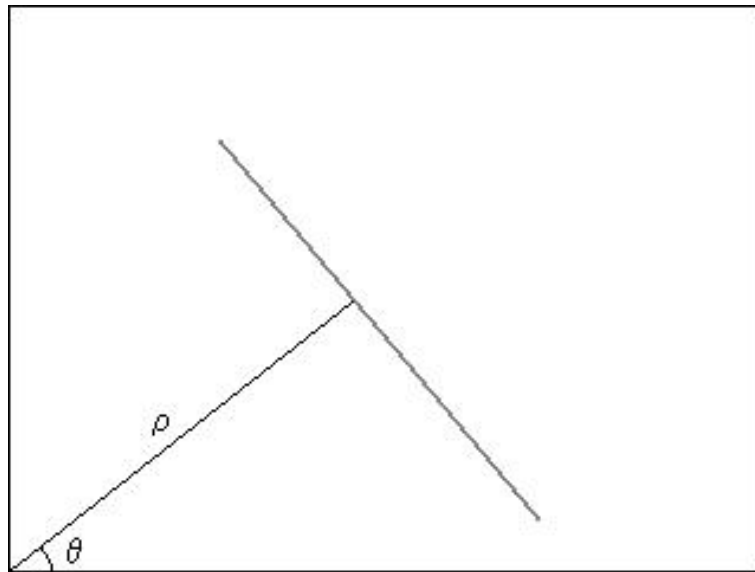


Figura 2.2: La trasformata di Hough

La trasformata di Hough, e tutte le applicazioni connesse, fanno un pesante utilizzo degli operatori trigonometrici, per questo si è scelto di salvare in un vettore i valori di seno coseno e tangente relativi ad ogni possibile valore assunto da θ . Particolare attenzione si è posta alla scelta della definizione del θ per non perdere in precisione ne' appesantire eccessivamente il programma; la soluzione adottata introduce un errore massimo di circa 0,9 gradi. Il θ può variare da $-\frac{\pi}{2}$, quando abbiamo una retta quasi orizzontale passante sotto il punto di riferimento, a π , che corrisponde ad una retta quasi verticale a sinistra del punto: i valori nel dominio di Hough variano da -52 a 105.

Come già ampiamente spiegato precedentemente (1.1) non sempre le linee da trovare erano nette e perfettamente dritte (condizioni ottimali per la ricerca di linee), per risolvere il problema dei bordi frastagliati o sfumati sono state sviluppate due diverse applicazioni della trasformata di Hough. Il primo metodo, quello più preciso, trovato un bordo calcola il ρ e il θ conoscendo la pendenza del bordo trovata con

Sobel, attraverso la formula

$$angolo = \arctan |fase|$$

si calcola l'angolo formato dalla distanza della retta dal punto fisso e dall'asse delle ascisse. Avendo utilizzato il modulo della fase si perde l'informazione sulla direzione dell'angolo che può però esser reperita considerando il segno della fase, infatti hanno fase negativa solo le rette che tagliano l'asse delle x a sinistra del punto fisso. Questo metodo funziona molto bene per le linee nette e dritte e richiede meno calcoli. Il metodo modificato invece quando trova un bordo cerca per ogni θ il relativo ρ e vota per ognuno dei valori trovati; per quanto più complesso e apparentemente meno preciso, questo metodo dà dei risultati discreti con bordi perfetti e ottimi risultati con bordi frastagliati o comunque non netti. Il ρ viene in entrambe le formule calcolato come

$$\rho = j * \cos \theta + (im_h - i) \sin \theta;$$

dove im_h rappresenta l'altezza dell'immagine, i e j rispettivamente la riga e la colonna dell'elemento dell'immagine considerato.

Esaminando i voti calcolati dai due metodi si notano alcune differenze: il metodo preciso registra voti più bassi e con una distribuzione puntiforme, il metodo modificato invece assegna molti voti ed ha una distribuzione dei voti sinusoidale, causata dalla votazione che viene eseguita per ogni θ (fig.2.3). È stato riportato solo il dominio di Hough relativo al metodo modificato perché più interessante e più facilmente visibile.

2.1.2 Le limitazioni e le operazioni successive

Questi metodi applicati senza limitazioni non risultavano comunque sufficienti, si è dunque scelto di affinarli limitando le zone di ricerca e la pendenza delle rette cercate.



Figura 2.3: I voti nel dominio di Hough

L'idea è che le informazioni relative al bordo della strada si trovano solo al di sotto del punto di fuga e le informazioni sul bordo destro si trovano a destra, mentre quelle per il bordo sinistro a sinistra; si è quindi limitata la ricerca alla zona sotto un ipotetico orizzonte, calcolato dai valori di calibrazione dell'immagine, e si è diviso a metà la zona per cercare i bordi destri a destra, e i bordi sinistri a sinistra. Si sono inoltre verificate, osservando immagini registrate in diverse posizioni della strada, le possibili pendenze del bordo e si è limitata la ricerca a tali pendenze, ovviamente differenti per i due lati. Queste limitazioni velocizzano notevolmente la ricerca e permettono una maggiore precisione.

Il dominio di Hough non viene utilizzato completamente, vengono utilizzate solo alcune zone, relative ai valori dei θ cercati, ma per semplicità di calcolo e di visualizzazione viene utilizzata una matrice delle dimensioni riportate precedentemente.

Entrambi i metodi vengono applicati sia per la ricerca dei bordi destri che per quella dei bordi sinistri, quest'ultima però richiede una doppia applicazione per poter considerare sia le linee che si trovano al di sopra del punto di riferimento, e che avranno

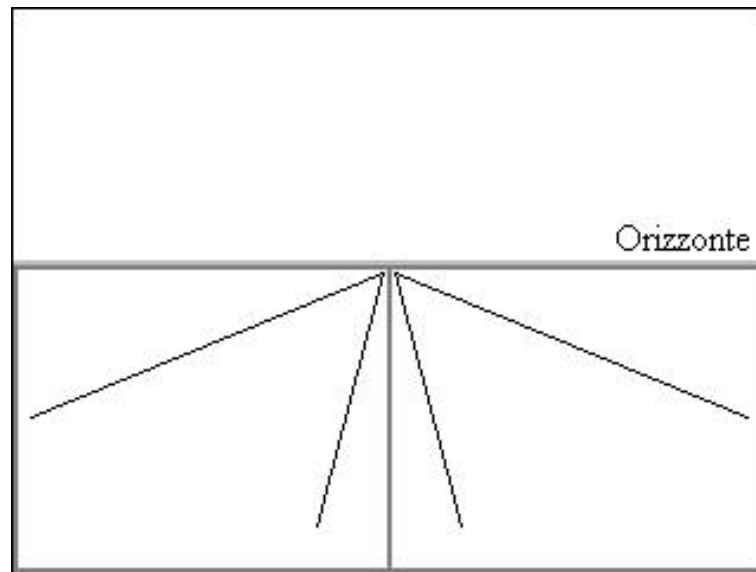


Figura 2.4: Zone di ricerca e pendenza delle rette

un θ superiore ai 90 gradi, sia le linee che si trovano sotto tale punto, e che avranno un θ negativo. Procedendo in questo modo ogni metodo viene applicato tre volte: una volta a destra e due volte a sinistra.

Per ogni metodo della trasformata di Hough vengono scelte due linee per lato, solo nel caso i voti nel dominio di Hough superino una certa soglia; in questo modo si trovano un massimo di quattro linee per lato, ma c'è la possibilità di non trovare nessuna linea. Nel caso non venga trovata nessuna linea da uno o da entrambi i lati non è possibile fare alcuna ulteriore operazione, nel caso vi sia almeno una linea da entrambe le parti si procede alla ricerca del punto di fuga. Le soglie utilizzate per la ricerca delle linee sono state calcolate empiricamente controllando i valori del dominio di Hough, esse variano al variare del lato, infatti possediamo solitamente molte più informazioni del lato destro rispetto a quelle del lato sinistro, e al variare del metodo, infatti il metodo di Hough più preciso raccoglie per ogni retta molti meno voti rispetto all'altro metodo. I punti che vengono considerati nel dominio di Hough, cioè le rette che hanno ottenuto più voti non possono però essere troppo simili, altrimenti i risultati

equivarebbero a quelli ottenuti considerare una solo punto, e quindi una sola retta, e si rischierebbe di considerare rette sbagliate; per questo motivo si è posta una differenza minima, sia per il ρ che per il θ entro la quale un solo punto può essere considerato.



Figura 2.5: Linee trovate dai metodi di Hough

Intersecando le quattro linee a destra con le quattro linee a sinistra si trovano sedici punti, tutti possibili punti di fuga; a questo punto si cerca la maggior concentrazione di punti nel raggio di dieci pixel, e si opera una media delle posizioni dei punti trovati; se la media è molto prossima ad una linea che si trova cinque pixel sopra all'orizzonte si prosegue, nel caso contrario si cerca tra tutti i punti il più vicino alla suddetta linea, se la distanza non supera una certa soglia si prosegue, altrimenti i dati raccolti vengono considerati non affidabili. Tutte queste operazioni possono essere eseguite sia con sedici punti, ma anche con un numero inferiore di punti, naturalmente sarà necessario avere almeno un punto.

Se osserviamo le linee trovate dai metodi di Hough (fig. 2.5) possiamo trovarne alcune praticamente coincidenti: questo è un ottimo segno, perché significa che entrambi i metodi hanno ottenuto molti voti per quella retta, e quasi sicuramente essa è un bordo

della strada. Alcune rette possono invece essere molto simili: anche questo risultato può essere letto come quello precedente in modo positivo, ma questo fatto può anche portare a delle imprecisioni causate dalle linee di delimitazione che non sono quasi mai rette monodimensionali, ma piuttosto strisce bidimensionali. Altre rette ancora possono trovarsi in zone inattese, queste sono quasi sempre causate dal rumore della strada.

2.1.3 Le variabili di confidenza

Anche nel caso non vengano trovate rette o le rette trovate vengano considerate non affidabili si vorrebbe che il metodo desse dei risultati, e allo stesso modo si vorrebbe che il metodo si accorgesse se vengono trovati punti decisamente poco probabili. Per raggiungere questo obiettivo sono state aggiunte delle variabili che controllano la confidenza di un punto trovato. La prima variabile, chiamata *confidenza* varia da 0 a 7, viene incrementata ogni volta che viene trovato un punto di fuga nella zona in cui ci aspettiamo di trovarlo, cioè abbastanza vicino al punto di fuga trovato precedentemente, e viene decrementata ogni volta che non viene trovata nessuna linea e ogni volta che si trova un punto di fuga anomalo. Se la confidenza è minore di 3 il punto trovato, se giudicato corretto, viene visualizzato ma, pur facendo aumentare la confidenza non viene considerato un punto di fuga molto affidabile, la posizione che assume viene ritenuta valida solo per il frame in cui viene trovata; in questo modo ci assicuriamo di non dare eccessiva importanza a punti dovuti a ombre o altri rumori, che variano rapidamente e che non riescono a far aumentare la confidenza abbastanza da farle superare la soglia necessaria per essere considerata affidabile; questo metodo può introdurre delle oscillazioni nelle linee disegnate, ma al massimo per la durata di 3 immagini, queste oscillazioni, che sono comunque controllate, non comportano un vero problema.

Se la confidenza è maggiore di 3 il punto viene considerato corretto e può essere

modificato solo da un punto che si trovi nel suo intorno, qualsiasi altro punto, come già detto, fa abbassare la confidenza; questa limitazione è dovuta alla constatazione che il punto di fuga varia molto lentamente e addirittura dovrebbe rimanere esattamente uguale su strade dritte, se viene trovato un punto decisamente lontano rispetto al punto precedente si può supporre che sia un punto sbagliato; evidentemente se vengono trovati una serie di punti al di fuori del raggio di dieci pixel la confidenza diminuirà fino ad arrivare ad un valore minore o uguale a 3 e verrà considerato e mostrato un nuovo punto di fuga, ciò avviene in prossimità di curve lievi o nel caso in cui l'automobile si sposti da un lato all'altro della strada. In questo caso viene introdotto un certo ritardo, che risulta però trascurabile, visto che non supera la durata di 3 immagini.

Se la confidenza raggiunge il valore massimo, cioè 7, si suppone di aver trovato un *punto sicuro*, e tale posizione viene salvata, nel contempo viene azzerato il valore dell'altra variabile, chiamata *incertezza*. La variabile *incertezza* varia da 0 a 5 e viene considerata solo nei casi in cui la confidenza sia a 0 o comunque non vi sia un punto di fuga da mostrare; se l'incertezza non è al suo valore massimo si suppone che il *punto sicuro* rappresenti una buona approssimazione del punto di fuga attuale, che non viene trovato a causa del rumore o di altri problemi dell'immagine. Questa funzione è molto utile salvo nei casi in cui, in effetti, la strada non debba essere trovata, come nel caso in cui ci si trovi in prossimità di un ostacolo che impedisca la vista della strada, in tale caso infatti la strada viene visualizzata anche se non dovrebbe esserlo, ma solo per pochi frame.

Il diagramma di flusso del metodo di ricerca del punto di fuga è riassunto per chiarezza. Esso si riferisce alla sola ricerca del punto di fuga partendo dalla conoscenza delle linee trovate con Hough ed arrivando alla localizzazione del punto di fuga più probabile (fig.2.7).

A questo punto rimangono da ricercare i bordi della strada, per trovarli utilizziamo ancora una volta le linee trovate dai metodi di Hough. Partendo dalla conoscenza dei

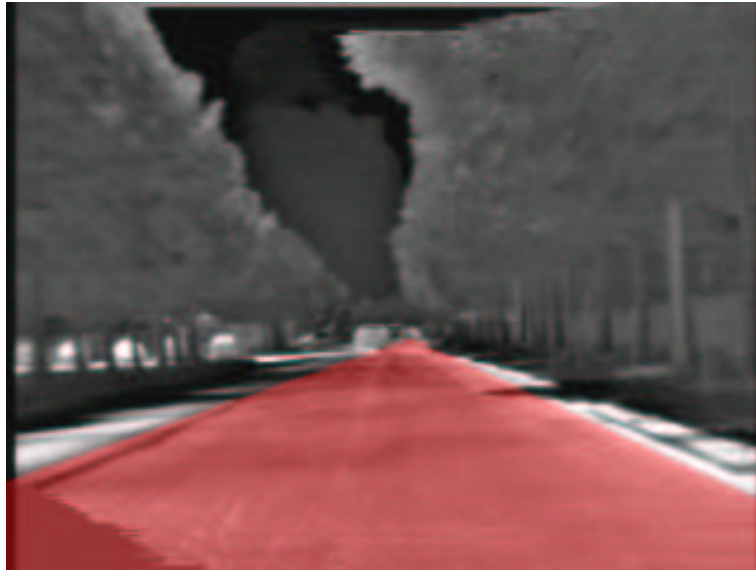


Figura 2.6: Il risultato finale del metodo di ricerca della strada

punti di intersezione tra le rette e il bordo sinistro dell'immagine possiamo utilizzare un metodo simile a quello utilizzato per trovare il punto di fuga che fa uso di due variabili di confidenza, una per il lato destro e una per il lato sinistro. Per prima cosa cerchiamo la maggior concentrazione di punti di intersezione col bordo, e a questo punto, dopo aver fatto una media, procediamo a confrontare con il bordo precedentemente trovato: se la variazione è inferiore ad una soglia o se la confidenza è inferiore ad un certo valore viene considerato il nuovo bordo, altrimenti si rimane con il bordo vecchio e la confidenza diminuisce. A differenza del metodo per la ricerca del punto di fuga i *punti sicuri* vengono assegnati anche quando la confidenza non è al valore massimo, e rimangono validi sempre. Questo permette di avere sempre un bordo destro e sinistro definiti, perché si è deciso di far in modo che la decisione di mostrare la strada sia presa solo dai informazioni provenienti dal metodo di ricerca del punto di fuga; in un solo caso la ricerca dei bordi blocca la visualizzazione dei risultati, cioè quando la strada risulta più stretta dell'autovettura, in un caso di questo tipo infatti il punto di fuga può essere stato calcolato correttamente, e solo il metodo di ricerca dei bordi si accorge

dell'errore.

La zona che viene trovata al termine di tutti questi procedimenti viene evidenziata con un colore rosso semitrasparente (fig.2.6).

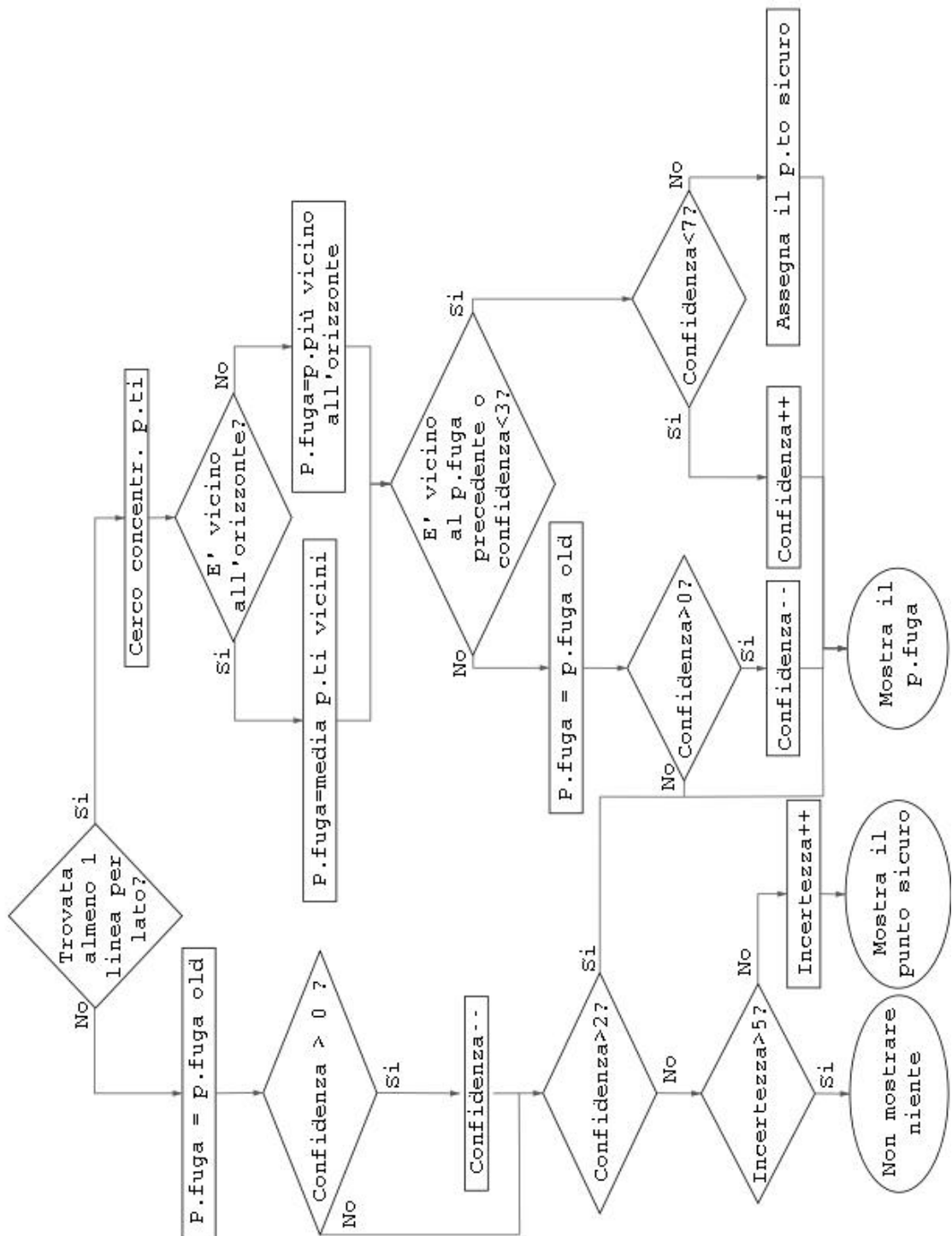


Figura 2.7: Diagramma di flusso del metodo di ricerca del punto di fuga

2.1.4 Un metodo alternativo

Il metodo esposto funziona molto bene su strade dritte o con curve lievi, ma, come ci attendavamo, trova i propri limiti in prossimità delle curve, per cercare di dare un risultato utile anche in queste situazioni si è deciso di inserire un metodo alternativo che dia buoni risultati in occasione dei malfunzionamenti del metodo principale; a questo secondo metodo è stata dedicata poca attenzione perché non rientra negli obiettivi principali che ci si era posti. Si è scelto di sfruttare la presunta uniformità di luminosità della strada; come già precedentemente accennato questa supposizione non sempre è verificata, in particolare in presenza di ombre o di linee sulla strada, come passaggi pedonali, ma questi casi sono rilevanti soprattutto con immagini calde. L'idea è di trovare la luminosità media della strada e di *espandere* la strada verificando che quadrati di due pixel di lato adiacenti alla strada trovata abbiano luminosità media simile a quella della strada; nel caso in cui il metodo basato su Hough funzioni si parte dalla strada trovata, in caso contrario si considera un rettangolo largo 120 pixel e alto 30 posto nella parte inferiore dell'immagine, zona in cui, salvo rarissimi casi, troviamo la strada.

Nel caso il metodo principale abbia funzionato si esegue un'operazione di distinzione tra i punti che effettivamente appartengono alla strada e quelli che invece non vi appartengono, evidenziando in questo modo le curve lontane o l'approssimarsi di un ostacolo.

L'immagine viene divisa in quadratini di 2 pixel di lato e viene creata una matrice con tanti elementi quanti sono i quadratini nell'immagine. Gli elementi della matrice possono avere 4 valori: 1 se la zona appartiene alla strada, -1 se non appartiene, 2 se è stata considerata strada dal metodo principale ma non da questo, 0 se la zona non è stata considerata. Per prima cosa vengono assegnati a 1 tutti gli elementi corrispondenti a quadrati appartenenti alla strada, nel caso il metodo principale abbia funzionato, o

appartenenti al rettangolo citato sopra, nel caso il metodo non abbia funzionato, in entrambi i casi viene calcolata la media della zona della strada.

Se il metodo principale ha funzionato si scansiona la matrice e per ogni quadratino appartenente alla strada si controlla che la sua media non differisca dalla media della strada per più di un certo valore, nel caso assegna all'elemento un valore pari a 2.

La parte inferiore dell'immagine, quella che si trova sotto l'orizzonte, viene scandita ed ogni volta che si incontra un elemento appartenente alla strada che ne ha vicino, cioè sopra, sotto, a destra o a sinistra, un altro non ancora considerato, si calcola la media del nuovo elemento: se la differenza tra la media della strada e quella dell'elemento è inferiore ad un certa soglia il valore dell'elemento viene impostato a 1, altrimenti a -1. Vengono effettuate scansioni finché non viene più trovato nessun quadratino appartenente alla strada.

I quadratini che hanno il valore pari a 1 vengono visualizzati in blu, quelli che hanno valore pari a 2 in verde, utilizzando colori semitrasparenti per fare in modo che l'immagine elaborata rimanga visibile(fig. 2.8).

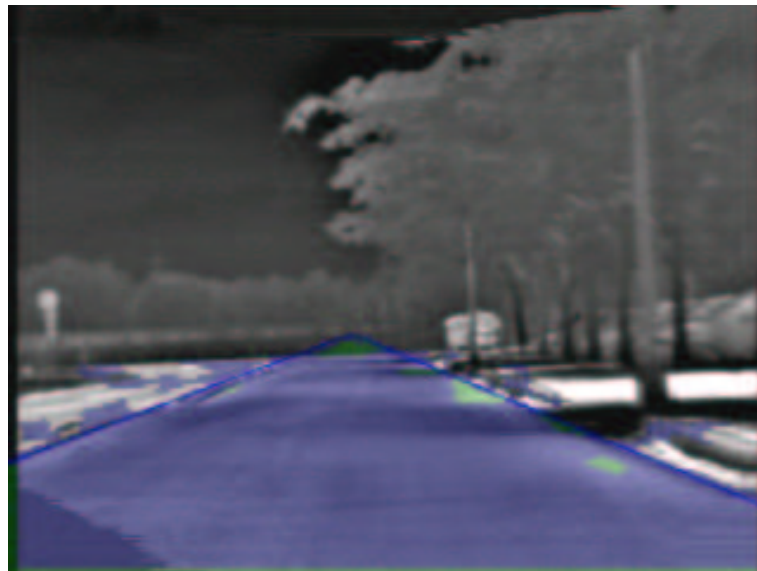


Figura 2.8: Risultato del metodo alternativo di ricerca della strada

2.2 Ricerca di Pali e di bordi di edifici

Pali ed edifici sono accomunati dai bordi laterali molto lunghi e verticali, per trovare tali elementi si è creata un'immagine contenente tutti i bordi verticali, o quasi verticali, trovati con il metodo di Sobel con modulo superiore ad una soglia molto bassa 2.9; mantenendo una soglia molto bassa sono stati considerati anche i pali che hanno poco contrasto con lo sfondo, si sono considerati anche i bordi quasi verticali per non perdere i pali leggermente storti. Nell'immagine si possono già notare quelli che sono i probabili pali e gli altri bordi causati da rumore o da oggetti piuttosto bassi.

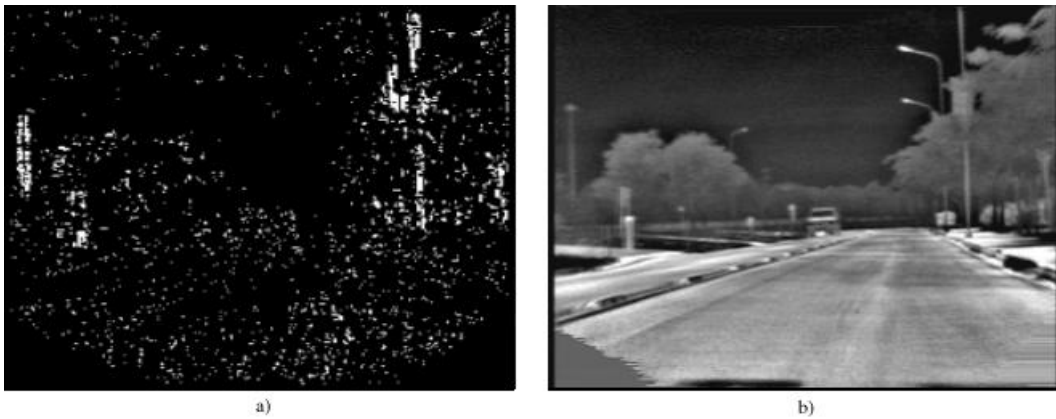


Figura 2.9: Bordi verticali

Si è scelto di non utilizzare il metodo di Hough per diversi motivi:

Per velocizzare il metodo: Hough è molto preciso ma molto pesante.

Perché ci servono solo i bordi verticali: è inutile utilizzare Hough solo per trovare delle linee verticali che possono essere facilmente trovate scandendo la colonna.

Perché Hough non da informazioni rispetto alla posizione dell'oggetto: sarebbe stato necessario considerare ancora una volta l'immagine originale per trovare l'inizio e la fine dei pali.

Il metodo scelto trova i suoi maggiori vantaggi nella semplicità e nella velocità di esecuzione.

Quello che ci interessa nell'immagine sono le colonne contenenti molti bordi verticali, ma questa informazione non è sufficiente ad identificare un palo, infatti si possono avere molte linee verticali corte causate da rumore o da oggetti con bordi verticali: dobbiamo anche considerare dei bordi continui. Quello che in realtà ci aspettiamo di trovare sono dei segmenti lunghi e continui, ma molto difficilmente li troveremo, infatti non sempre la pendenza dei bordi viene interpretata correttamente, a causa dello sfondo discontinuo, per questo troveremo dei bordi frastagliati o comunque non continui. Per ovviare a questo problema è stata inserita una variabile che indica la dimensione dei tratti senza bordi permessi all'interno delle rette.

Il metodo è molto semplice: per prima cosa si contano il numero di bordi verticali presenti per ogni colonna, questo valore viene salvato in un vettore, successivamente partendo dall'alto dell'immagine vengono cercate per ogni colonna delle linee continue e viene salvata in un vettore la lunghezza ed il punto di inizio della più lunga. Come già accennato, nella ricerca di linee continue si possono trovare dei buchi: ogni volta che si incontrano dei pixel senza bordi una variabile viene decrementata, fino a zero, ed ogni volta che si trovano dei pixel con bordi la stessa variabile viene incrementata, fino ad un massimo di venti; quando la variabile raggiunge il minimo si calcola la lunghezza della linea trovata e si continua cercando una linea più lunga.

Abbiamo a nostra disposizione due vettori che contengono, per ogni colonna dell'immagine, il numero di bordi verticali e la lunghezza del bordo più lungo, prima di utilizzare tali vettori viene operata una sorta di filtro mediano ad entrambi i vettori, viene cioè sostituito il valore contenuto nella posizione i con la media dei valori delle posizioni $i - 1$, i e $i + 1$. A questo punto possiamo verificare se i due valori raccolti per ogni colonna superano le due soglie definite (fig.2.10), in questo caso si cerca se esistono altre colonne vicine che superano le soglie e si formano dei quadrati che con-

tengono le linee più lunghe di colonne successive. Non ci si deve stupire che si creino dei quadrati e non solo delle linee, infatti la soglia bassa permette di trovare anche leggere sfumature verticali, che sono spesso presenti sui pali, e le linee non perfettamente verticali votano per colonne adiacenti.

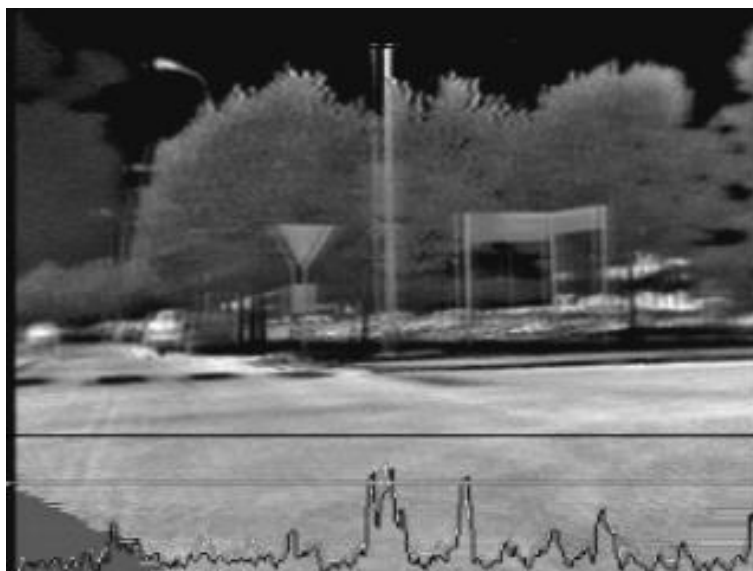


Figura 2.10: Le soglie per la ricerca dei pali

Utilizzando questo metodo si hanno dei buoni risultati con linee piuttosto lunghe, che corrispondono a pali vicini, ma risulta molto più difficile trovare dei pali lontani, che nell'immagine appaiono più corti, per questo motivo è stata introdotta una soglia che è massima ai bordi e diminuisce avvicinandosi alla colonna che contiene il punto di fuga, dove si dimezza (fig.2.11): solo i pali che hanno la base vicino al bordo della strada utilizzano la soglia modificata, tutti gli altri utilizzano la soglia fissa.

È stato provato anche un altro metodo, si era notato infatti che i pali e gli edifici potevano avere pochi bordi verticali ma molto concentrati, in questo caso il numero di bordi e la lunghezza della linea più lunga risultavano simili; si è dunque provato a considerare anche le colonne numero di pixel accesi e lunghezza simili e al di sopra di una certa soglia. Questo metodo funzionava piuttosto bene, trovava infatti un certo

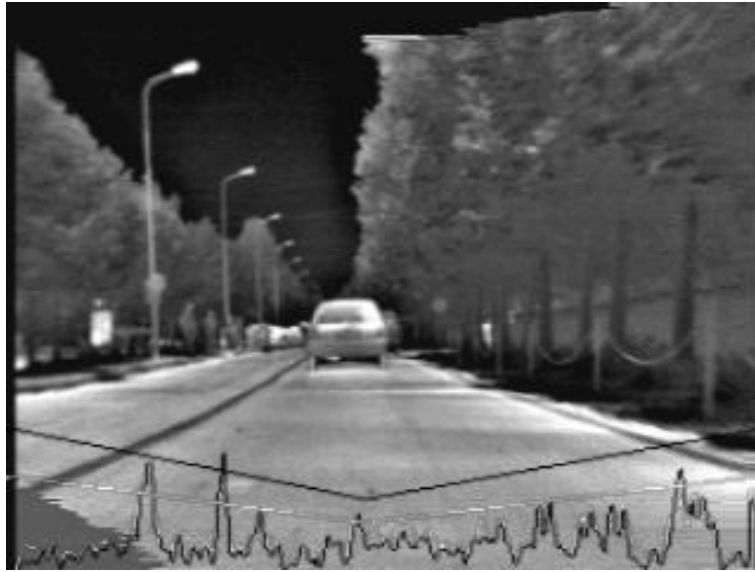


Figura 2.11: Soglia variabile per la ricerca dei pali

numero di pali che i metodi precedenti ignoravano, ma selezionava anche degli oggetti che non erano pali.

Distinguere tra pali ed edifici con questo metodo è piuttosto complicato, in quanto per semplificare la ricerca buona parte dell'informazione viene persa; è stato fatto un tentativo per trovare la distinzione, ma il tentativo è purtroppo fallito. La discriminante tra i due oggetti doveva essere la luminosità media, infatti il bordo di un edificio presenta sui due lati luminosità medie diverse, da una parte lo sfondo e dall'altra la costruzione; i pali invece presentano quasi sempre una certa luminosità e ai lati aree con luminosità diversa rispetto al palo, ma simile tra loro. Il metodo, in apparenza corretto, fallisce quando vengono trovati i bordi del palo e non il palo nella sua interezza, in questo caso infatti la luminosità media destra può essere molto diversa da quella sinistra.

I pali vengono visualizzati con dei quadrati azzurri semitrasparenti, vengono anche rappresentate le soglie, tramite rette, e i valori di ogni colonna tramite un diagramma (fig. 2.12).

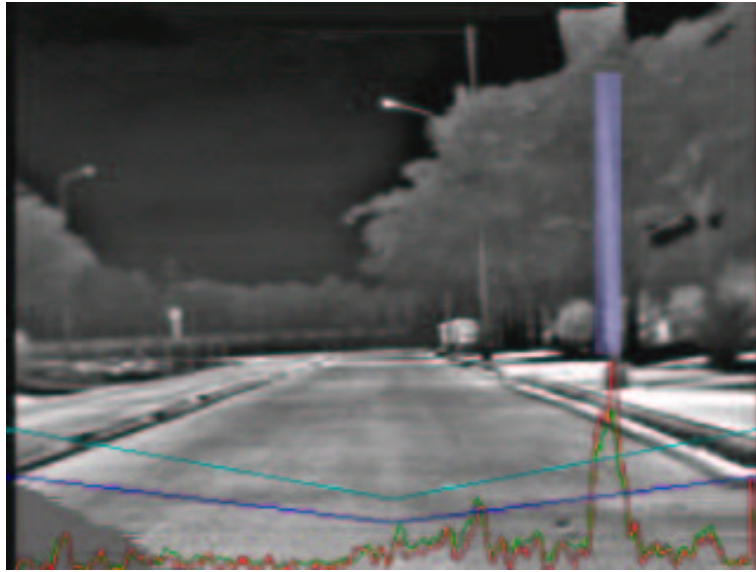


Figura 2.12: Risultato del metodo di ricerca dei pali

2.3 Ricerca di segnali stradali

Per la ricerca dei segnali stradali circolari si è scelto di utilizzare il metodo della trasformata di Hough generalizzata, che funziona molto bene per la ricerca di forme geometriche. Il metodo di hough generalizzato si utilizza per applicare hough ad una qualsiasi forma geometrica, semplice o complessa. Per utilizzarlo bisogna creare una r-table partendo da una rappresentazione dell'oggetto che si vuole cercare; per costruirla si usa un punto di riferimento nel centro dell'oggetto e si costruisce una tabella che riporta la posizione del centro rispetto ad ogni punto del contorno. Quando applichiamo il metodo facciamo votare ogni rilievo dell'immagine per ogni punto della r-table, orientandola secondo la direzione del contorno. Nel nostro caso, prendendo come punto di riferimento il centro del cerchio, la r-table sarà data da un punto distante quanto il raggio del cerchio, infatti tutti i punti del bordo sono equidistanti dal centro.

Le dimensioni dei segnali stradali sono pressoché standard, salvo rari casi, ma per poter trovare sia segnali lontani che segnali vicini, il raggio dei cerchi dovrà poter

variare da un minimo di 5 pixel ad un massimo di 15 pixel, in questo modo saranno considerati segnali da una distanza di cinque metri circa, fino a 70 metri circa, cioè tutti i segnali che possono indurre in errore nella ricerca dei pedoni.

Non sempre i segnali sono visti perfettamente di fronte dalla telecamera, perché l'automobile non è perfettamente allineata alla strada, perché il cartello si trova su una curva o perché il cartello stesso è storto: ricercare un cerchio perfetto in queste condizioni non darebbe alcun risultato. Si è pensato di modificare la ricerca in modo tale che il metodo desse buoni risultati anche con forme ellittiche, per questo si è deciso di modificare la *r-table* utilizzando, anziché un solo punto, tre punti allineati a distanze diverse.

Come già accennato la complessità della scena ripresa introduce molti problemi, infatti molti oggetti quasi circolari rischiano di essere considerati come cartelli; due metodi sono stati adottati per ovviare a questo problema. Il primo metodo considera il centro trovato e a partire da quel punto controlla i bordi presenti relativi a cerchi di varie dimensioni, cioè conta quanti pixel di bordo si trovano ad una certa distanza dal centro; se la percentuale di bordi presenti alla distanza pari al raggio del cerchio con cui abbiamo trovato il centro è maggiore delle altre percentuali, allora la forma viene considerata un cartello. Il secondo metodo invece cerca delle linee verticali al di sotto della circonferenza, ci aspettiamo infatti che un cartello sia sorretto da un palo; viene considerata una zona rettangolare di tre pixel di base e venti di altezza allineata sotto il centro del cartello stradale e si contano i bordi verticali presenti al suo interno, se il numero di bordi contati non supera una certa soglia la forma non viene considerata.

Per limitare la ricerca si era anche pensato di calcolare a priori l'altezza sull'orizzonte di un segnale al variare della distanza, in questo modo, conoscendo approssimativamente la dimensione di un cartello, si sarebbero potuti cercare cerchi di una certa dimensione solo ad una certa altezza; questa idea è stata abbandonata dopo aver notato il fatto che l'altezza dei cartelli è molto variabile e spesso capita anche che vi siano

cartelli sovrapposti; gli unici interventi possibili sono stati quelli di limitare la ricerca alla parte superiore dell'immagine e di alzare, all'aumentare del raggio, il bordo superiore della zona di ricerca, infatti mentre i cartelli lontani, di dimensioni ridotte, si trovano in un'area verticalmente ristretta, i cartelli più vicini possono essere in una fascia molto alta dell'immagine, ma anche piuttosto in basso.

I segnali vengono identificati da circonferenze rosse al cui interno viene riportato, sempre in rosso, il raggio del cerchio trovato (fig. 2.13).



Figura 2.13: Risultato del metodo di ricerca dei segnali stradali

Capitolo 3

Risultati e conclusioni

Questo capitolo contiene i risultati ottenuti col programma elaborato. Complessivamente i risultati sono buoni e possiamo dire che la procedura di riconoscimento della strada e del punto di fuga, nella quale maggior impegno è stato profuso, è quella che funziona meglio sia in condizioni ottimali che in condizioni complesse. Le ricerche di pali e di segnali stradali hanno dato dei risultati non banali ma non certo del livello raggiunto col metodo principale.

I possibili sviluppi futuri sono molti e si basano sia sull'affinamento dei metodi precedenti, che sull'appoggio a tecnologie e strumenti differenti, che certamente potrebbero apportare un notevole aiuto nella soluzione del problema.

3.1 Risultati

Descrivere i risultati di un programma di questo tipo è piuttosto complicato; non si può semplicemente dire il programma funziona oppure il programma non funziona, perché sarebbe estremamente riduttivo, il confronto con la realtà, tralasciando le approssimazioni tipiche dell'ambiente didattico, complica notevolmente le valutazioni. Di seguito riporterò alcune situazioni in cui il programma ha funzionato bene ed alcune

in cui i risultati non sono stati raggiunti, ma al di là della casistica e delle percentuali di successo, che per altro non sono presenti, è interessante domandarsi il motivo dei malfunzionamenti e domandarsi quali possibili soluzioni possono essere provate.

3.1.1 La strada e il punto di fuga

Il metodo di determinazione della strada e del punto di fuga ha dato buoni risultati praticamente in ogni situazione. Tutti i modelli di strada esaminati sono stati riconosciuti correttamente, anche se non sempre in modo perfetto, ne' in modo continuativo.

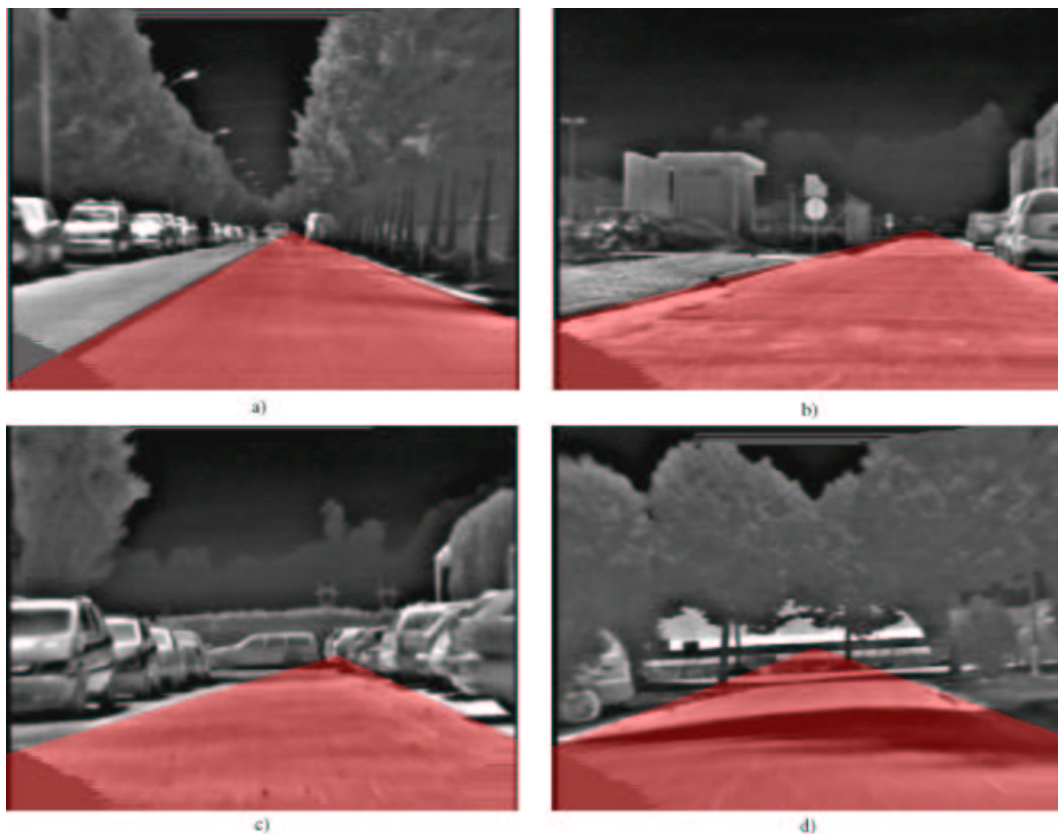


Figura 3.1: I vari tipi di strade riconosciute dal metodo

Le strade delimitate da linee vengono riconosciute abbastanza facilmente, come

è ovvio; in questo caso infatti funzionano molto bene entrambi i metodi di Hough, sia quello basato sulla direzione del bordo trovata con Sobel (preciso), sia quello che favorisce le linee con più bordi, qualsiasi angolazione abbiano tali bordi (modificato). Il punto di fuga e anche i bordi della strada vengono trovati in modo continuo e preciso (fig.3.1.a), questo caso è il più semplice, abbiamo infatti bordi dritti e molto marcati.

Le strade delimitate da marciapiedi sono più complesse, infatti i bordi non sono molto netti per quanto siano per lo più dritti, sono inoltre presenti più linee, quelle che seguono i 3 lati visibili del cordolo e quella eventuale dell'ombra, questo complica l'elaborazione, soprattutto col metodo di Hough che abbiamo chiamato impreciso. In realtà i risultati sono piuttosto buoni (fig.3.1.b), anche se alcune volte il cordolo del marciapiede viene considerato come strada, e altre volte no.

Le strade delimitate da autovetture sono decisamente difficili da elaborare, innanzitutto è impossibile considerare l'intera strada, si considera solo la parte libera da automobili, e questo è piuttosto ragionevole, inoltre il bordo fornito dalle macchine parcheggiate è molto frastagliato e diverso, ci possono essere auto molto calde e auto più fredde, il bordo può essere dovuto all'ombra o alla macchina stessa, le macchine non sono quasi mai perfettamente allineate e tra di loro e tra le automobili c'è una distanza variabile. Nonostante questo il metodo funziona piuttosto bene (fig.3.1.c), sempre che la distanza tra le automobili non sia eccessiva e non si veda il vero bordo della strada. Un errore in cui si può incappare è quello di considerare il bordo generato dalla linea delle portiere delle automobili e quindi di avere un risultato falsato, anche se di poco.

Le strade non delimitate sono le più complesse: il metodo di Hough che abbiamo definito preciso non funziona correttamente a causa dell'assenza di linee dritte. I risultati ottenuti variano a seconda di quanto è frastagliato il bordo, ma in ogni caso non sono troppo soddisfacenti (fig.3.1.d); fortunatamente questo caso è il più raro.

Il metodo non ha sempre i risultati sperati, in alcuni casi abbiamo dei fallimenti, o



Figura 3.2: Una sola automobile parcheggiata

perché la strada non viene trovata o perché viene trovata in modo sbagliato. Abbiamo un esempio molto chiaro di strada che non viene trovata quando si presenta un'unica automobile parcheggiata sul ciglio della strada (fig.3.2). In questo caso vengono trovate due linee diverse una generata dai bordi della macchina e una dai bordi della strada: il metodo si trova in difficoltà e non sempre riesce a funzionare correttamente, ammesso che esista una soluzione. Questo tipo di problema si può avere ogni volta che un ostacolo o un'ombra inseriscono una seconda possibile linea sull'immagine. Nel caso si tratti di un ostacolo (una macchina, un pedone o altro) sarebbe corretto restringere la corsia e considerare solo la parte di strada libera e percorribile, nel caso si tratti di ombre sarebbe corretto mantenere la forma originale della strada, quello che spesso capita invece è che la strada non viene semplicemente trovata.

Un problema simile si ha quando si incontrano sulla strada delle strisce pedonali, in questo caso abbiamo delle righe dritte e molto marcate (soprattutto d'estate o comunque sotto il sole) che possono trovarsi anche al centro della strada e che, a causa della prospettiva, appaiono oblique come i bordi della carreggiata. Capita spesso che



Figura 3.3: Strisce pedonali

tali linee raccolgano più voti del ciglio della strada e per questo vengano scelte come bordi dai metodi di Hough. Utilizzando le rette generate dalle strisce pedonali non si trova comunque un buon punto di fuga, e per questo il metodo non si accorge dell'errore, ma i bordi della strada trovati sono assolutamente sbagliati, si rischia di avere il modello di una strada molto più stretta rispetto alla realtà.

Nel caso siano delimitate da linee sia la strada che la corsia in cui si trova l'automobile si possono avere dei problemi: entrambe le linee vengono trovate e il metodo non riesce a scegliere quale delle due posizioni sia quella corretta, non esiste infatti alcun controllo per distinguere un tipo di linee da un altro tipo. I due bordi si alternano nella visualizzazione, si ottiene in questo modo un risultato pessimo, per nulla affidabile, fortunatamente dopo un breve periodo in cui il risultato è variabile, la situazione si stabilizza. In effetti non ci si era mai posti questo problema, che è stato riscontrato solo nelle verifiche finali; per continuità con le soluzioni adottate finora si dovrebbe considerare solo la corsia, cioè la parte di strada percorribile, ma non sempre il metodo privilegia questa scelta.

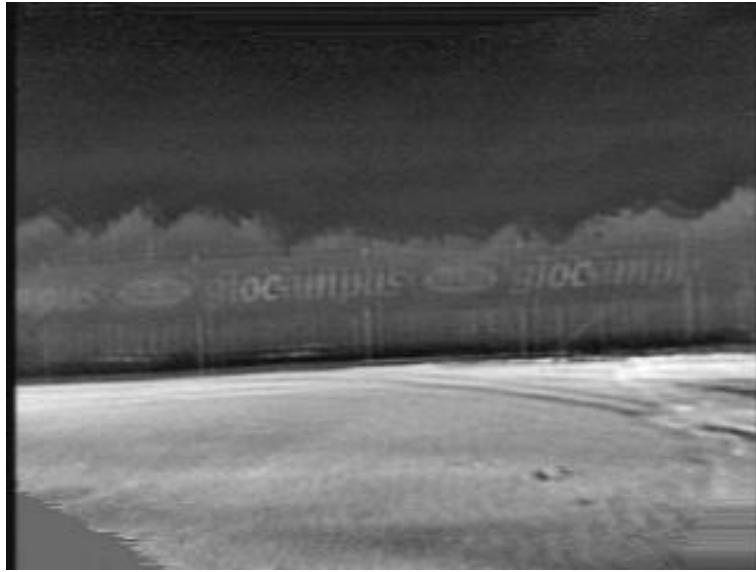


Figura 3.4: Curva

Sulle curve si sperimenta, come previsto, il vero limite del metodo; non si tratta di un vero e proprio malfunzionamento ma di una limitazione dovuta all'idea che ha generato il metodo, non si può infatti pensare di trovare correttamente la strada sulle curve se la cerchiamo mediante il riconoscimento di bordi dritti. Sulle curve il riconoscimento dovrebbe bloccarsi e non dare alcun risultato (fig.3.4), a volte, molto raramente, invece produce dei risultati, che non hanno alcun senso, e che sono causati dal rumore presente sulla strada.

Sulle curve lievi invece le cose funzionano meglio e possiamo dire, in generale, che il metodo funziona senza eccessivi problemi (fig.3.5), anche se, per i motivi esposti nel capitolo precedente (2.1.3) si ha un certo ritardo nel passaggio. Nell'immagine si nota come le linee si stiano adattando al cambiamento di direzione della strada e nel proseguire della scena le linee si spostano nella nuova posizione trovata.

Per cercare di avere dei risultati anche sulle curve e per cercare di perfezionare i risultati ottenuti sulle strade dritte utilizziamo il metodo che ho chiamato di *espansione*. Questo metodo secondario ha dato risultati poco continui e necessita di

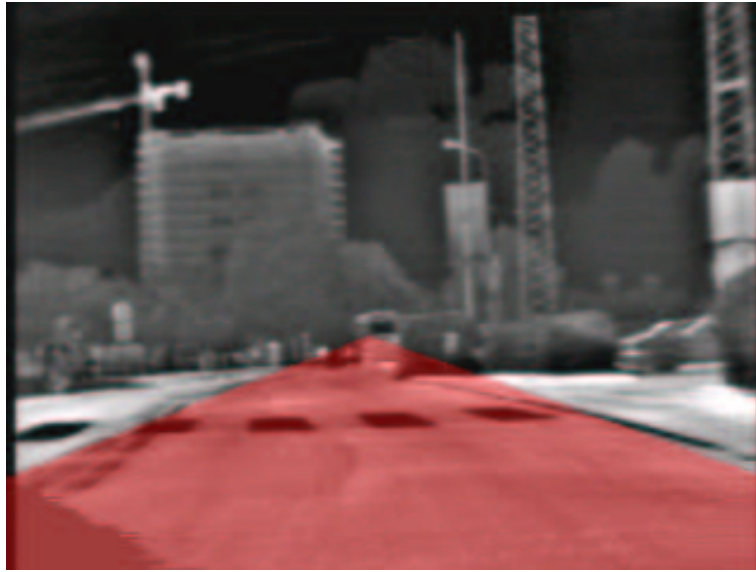


Figura 3.5: Curva lieve

una modifica alle variabili quasi ogni volta che si cambia sequenza. Il suo funzionamento sulle curve, cioè quando l'altro metodo è inutilizzabile, dà comunque delle informazioni non banali (fig.3.6).

L'utilità di conoscere la posizione della strada è già stata spiegata nell'introduzione, ma in particolare conoscere la posizione in curva diventa ancor più importante, visto che in tale momento si deve modificare la direzione di avanzamento dell'automobile. Come si nota dall'immagine i bordi trovati sono piuttosto frastagliati, ma comunque si può fornire un'idea della strada che approssima abbastanza bene la realtà.

Quando non ci trova in prossimità di una curva il metodo di espansione da delle informazioni interessanti soprattutto nell'individuazione di ostacoli sulla strada, purtroppo però risente pesantemente della presenza di ombre (fig.3.7), di segnaletica orizzontale e di rumore. Anche in questo caso i bordi sono frastagliati, vi sono addirittura delle zone di espansione che si trovano assolutamente al di fuori della strada; si nota bene come il rumore della strada e le ombre influiscano sul risultato. Le linee in blu rappresentano la strada trovata con il metodo principale, mentre in alcuni punti i bordi

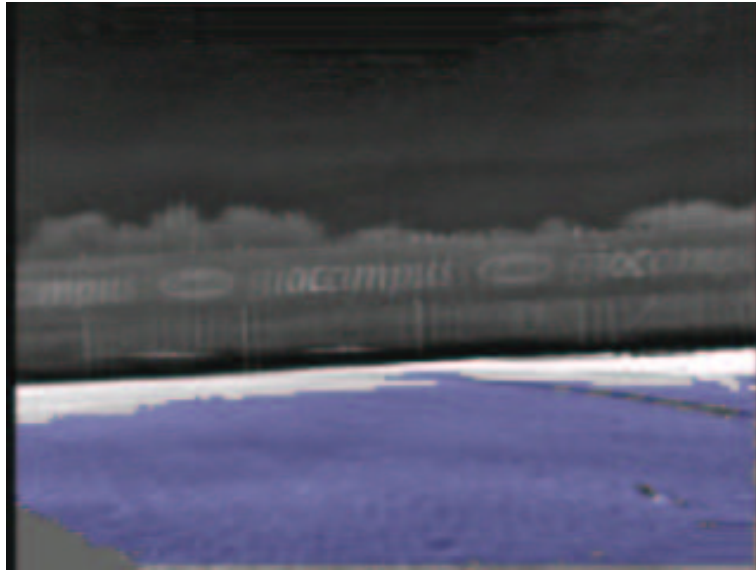


Figura 3.6: Metodo alternativo sulle curve

trovati coincidono, in altri sono decisamente diversi.

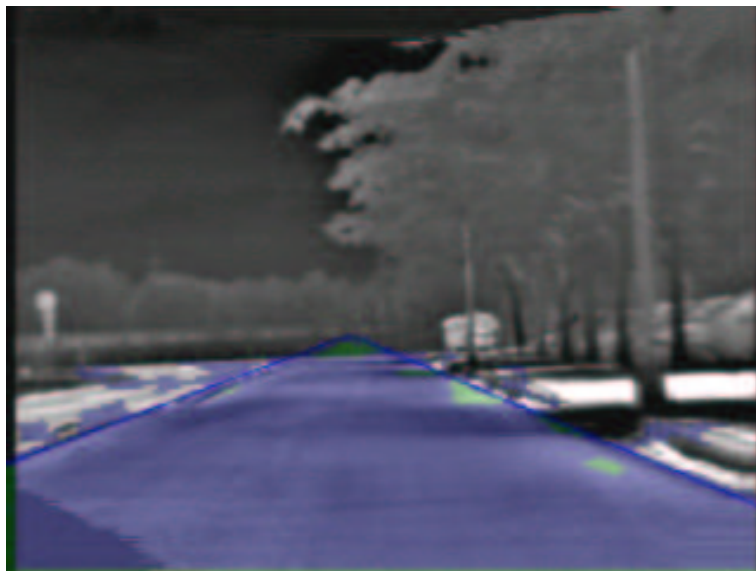


Figura 3.7: Metodo alternativo su tratti rettilinei

3.1.2 I pali e gli edifici

I risultati ottenuti con questo metodo appaiono soddisfacenti nell'individuazione di linee verticali lunghe, cioè di quello che pensiamo possano essere pali ed edifici. Aver mantenuto bassa la soglia sul modulo dei bordi ha inserito molto rumore, ma ha anche favorito l'individuazione di bordi e pali con uno sfondo non omogeneo, o comunque con poco contrasto.

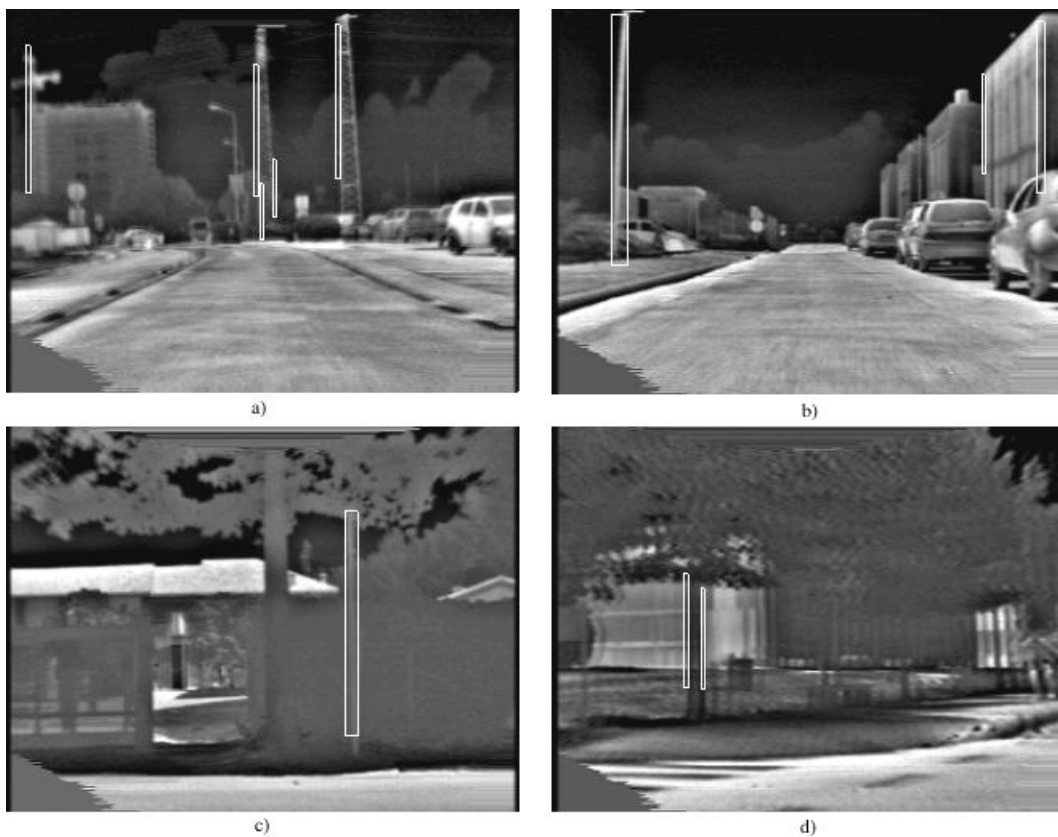


Figura 3.8: Risultati del metodo di ricerca dei pali

Come si può notare i pali della luce e gli angoli degli edifici vengono trovati abbastanza bene (fig.3.8). Se analizziamo la prima immagine (fig.3.8.a) possiamo notare come vengono presi anche degli oggetti che non sono propriamente pali, ma che hanno una forma molto simile, inoltre si può notare un altro problema, non sempre i pali

vengono considerati in tutta la loro altezza, spesso poi lo stesso bordo viene diviso in più parti. La seconda immagine (fig.3.8.b) mostra come il metodo funzioni bene anche con i bordi delle costruzioni e con i pali non perfettamente dritti; sulla sinistra dell'immagine troviamo un palo decisamente storto che viene individuato correttamente. I pali dei cartelli stradali vengono difficilmente individuati, in alcuni casi però possono essere trovati (fig.3.8.c), in questo caso il palo è molto vicino e visto lateralmente, la sua altezza è sufficiente. Un problema non superato riguarda la parte di palo che viene identificata, nell'immagine (fig.3.8.d) si distinguono chiaramente come vengono identificati i due bordi dell'albero, anziché l'albero intero; come già spiegato questa imperfezione genera problemi nelle successive possibili elaborazioni, ma l'entità di tali problemi è trascurabile.



Figura 3.9: Sequenza di pali correttamente riconosciuta

La ricerca di pali funziona bene anche col metodo della soglia variabile(vedi 2.2), si nota bene come la successione di pali non comporti problemi particolari nonostante l'evidente differenza nell'altezza dei due pali (fig.3.9); i pali più lontani ven-

gono riconosciuti con più difficoltà e con meno precisione, ma avvicinandosi i risultati migliorano decisamente.



Figura 3.10: Pali non correttamente riconosciuti

Un difetto, difficilmente correggibile utilizzando questa procedura, è quello che si nota considerando delle immagini di pali corti, come quelli dei segnali stradali; infatti il nostro metodo ricerca delle linee verticali lunghe, mentre questo tipo di pali generano delle linee piuttosto corte che non superano la soglia richiesta. I pali della segnaletica verticale molto raramente vengono trovati, per quanto vicini siano (fig.3.10), questo è un problema relativamente grave, infatti molto raramente tali pali generano errori nella ricerca dei pedoni, e se capita l'errore viene individuato dal metodo di ricerca di segnali circolari. (vedi 3.1.3).

Ma i risultati migliori che si possono apprezzare si osservano confrontando il metodo di ricerca dei pali con quello di ricerca dei pedoni (fig.3.11), infatti quest'ultimo spesso individua dei pali come pedoni. Ad un primo esame questo può non sembrare un problema, infatti sia pedoni che pali sono ostacoli da evitare, ma dobbiamo considerare il fatto che, mentre i pali sono fissi, i pedoni sono quasi sempre in movimento

e risulta evidente la differenza di attenzione che va posta ai due ostacoli, inoltre distinguere tra gli ostacoli presenti aiuta ad avere una maggior conoscenza dell'ambiente stradale. Il palo non si sovrappone completamente al pedone, la sovrapposizione si limita alla parte superiore del pedone, ma pedone e palo si trovano perfettamente sovrapposti verticalmente. L'informazione dell'errore nel riconoscimento del pedone deve essere a mio parere trasmessa soprattutto lungo le colonne, in modo tale che basti una sovrapposizione anche parziale affinché l'errore venga identificato.

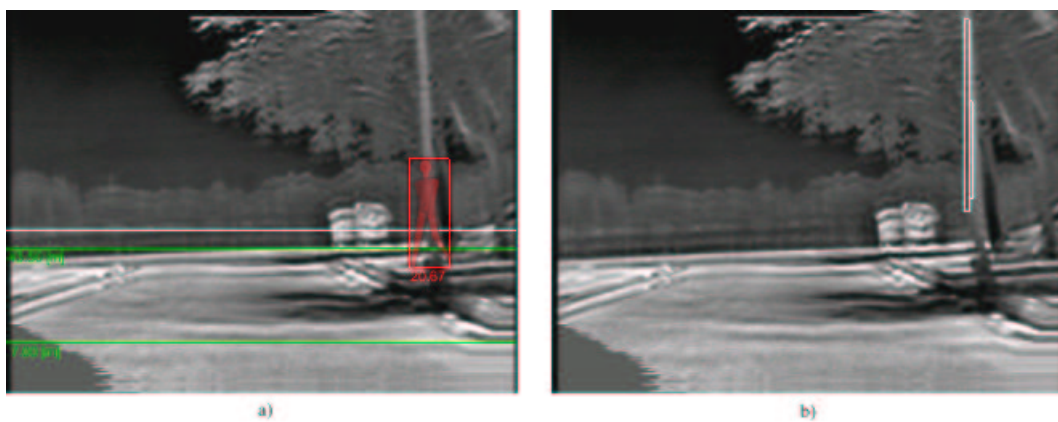


Figura 3.11: Confronto col metodo di ricerca dei pedoni

Per quanto riguarda il problema, cui già si accennava, dei pedoni con bordi verticali lunghi e continui, dobbiamo dire che il problema sussiste e non è del tutto risolto. In realtà le immagini in cui si ha questo tipo di problema sono veramente poche, si limitano a pedoni abbastanza vicini e molto dritti (fig.3.12), inoltre i bordi che vengono considerati come pali si trovano sempre nella zona laterale del bounding box e quindi, se l'informazione si trasmette verticalmente, tale errore non viene considerato.



Figura 3.12: Pedone con bordi verticali lunghi e continui

3.1.3 Cartelli circolari

La ricerca di segnali stradali circolari è quella che ha dato i risultati meno buoni, ciò nonostante possiamo dire di aver in buona parte raggiunto l'obiettivo prefissato. Il vero problema in questo caso è la complessità delle immagini che impone una difficile scelta delle soglie con cui lavorare, se da una parte infatti si vorrebbe tenere la soglia piuttosto alta per evitare di considerare come cartelli oggetti di forma quasi circolare, dall'altra si cerca di usare soglie piuttosto basse in modo da trovare tutti i cartelli, anche quelli leggermente deformati. Si è così raggiunto un buon compromesso che permette di non considerare la maggior parte degli oggetti con forme quasi circolari, tra cui anche i pedoni, ma che purtroppo identifica i cartelli stradali ad intermittenza (fig.3.13).



Figura 3.13: Cartello identificato a intermittenza

Come possiamo notare (fig.3.14) i cartelli, pur di forma ellittica vengono correttamente identificati, il problema che si incontra in questo caso è la difficile definizione del raggio, nella maggior parte dei casi viene considerato il raggio minore dell'ellisse, anche se questa non è forse la scelta migliore.

Inoltre possiamo notare che il metodo funziona bene per le diverse dimensioni considerate: avvicinandosi ad un cartello possiamo vedere come il cartello venga in modo



Figura 3.14: Cartello di forma ellittica

quasi continuativo identificato e la dimensione rilevata venga correttamente aggiornata di volta in volta (fig.3.15); riportare correttamente la dimensione del segnale è molto utile per poter valutare approssimativamente la distanza dell'automobile dal cartello.



Figura 3.15: Cartello in avvicinamento

Non tutti i cartelli vengono trovati, questo soprattutto a causa dello scarso contrasto tra il segnale e lo sfondo. Per la ricerca dei segnali si sono usati i bordi trovati con il metodo di Sobel che vengono utilizzati anche per il metodo di ricerca della strada e

del punto di fuga; i bordi trovati sono sogliaati con una soglia variabile che visualizza sempre una certa percentuale, in genere molto bassa, di pixel accessi su tutta l'immagine, per questo motivo i segnali con poco contrasto non vengono identificati e i segnali possono essere trovati ad intermittenza.

Vi sono poi alcune forme circolari che possono essere individuate come cartelli; ne esistono parecchie: specchietti retrovisori, scritte di cartelloni pubblicitari, parti di edifici, solo per citarne alcuni. Giustificare questi errori del metodo è a volte molto semplice, perché esistono veramente dei bordi circolari posti sopra linee verticali (fig.3.16.a), ma spesso diventa piuttosto complicato perché osservando l'immagine non si trovano forme circolari (fig.3.16.b). In questo caso conviene osservare i bordi trovati con Sobel e ci si può accorgere che sono semicircolari: a volte questo è sufficiente al metodo per identificare un cerchio. Fortunatamente, grazie agli accorgimenti inseriti, tali errori non si verificano con i pedoni.



Figura 3.16: Errori del metodo di ricerca dei segnali

Ancora una volta i risultati migliori si osservano confrontando il metodo con quello di ricerca dei pedoni (fig.3.17). La procedura di ricerca dei pedoni a volte considera i segnali circolari come teste di pedoni, ma visto che i segnali stradali vengono correttamente identificati questo problema può essere facilmente corretto e si possono sensibilmente affinare i risultati ottenuti.

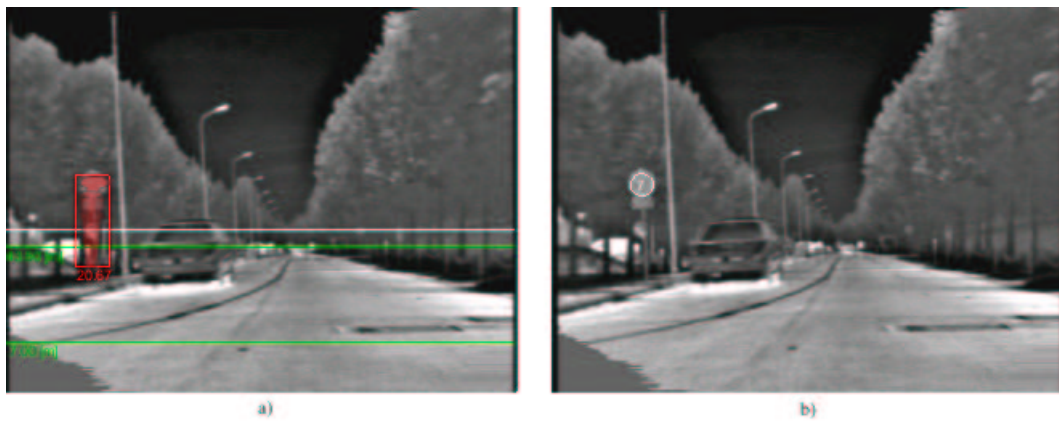


Figura 3.17: Confronto col metodo di ricerca dei pedoni

3.1.4 Immagini a varie temperature

Le immagini presentate finora sono state registrate nei mesi estivi e presentano per questo motivo un forte contrasto che rende più facile il compito dei vari metodi sviluppati; ciò nonostante il metodo funziona piuttosto bene anche con immagini registrate in mesi diversi (fig.3.18). La differenza tra le due immagini è evidente, ma regolando opportunamente le variabili si può ottenere un risultato buono anche con immagini più fredde, che purtroppo però non hanno risultati buoni come quelli delle immagini calde. L'immagine invernale è stata volutamente scelta tra quelle con il minor contrasto per evidenziare questo problema, in realtà immagini di questo tipo non sono così comuni.

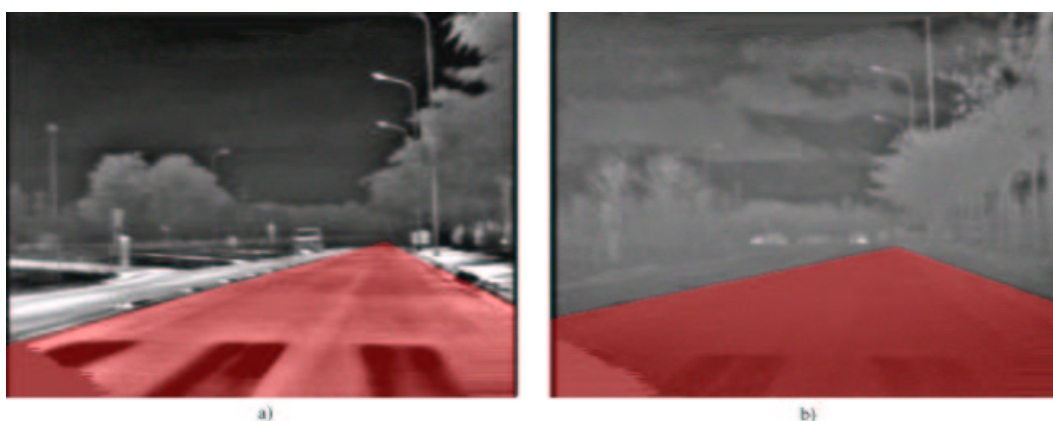


Figura 3.18: Risultati con immagini a diverse temperature

Dobbiamo aggiungere anche che la telecamera all'infrarosso non funziona con qualsiasi condizione atmosferica, la pioggia battente e la nebbia alterano notevolmente i risultati ottenibili da qualsiasi telecamera, in casi come questi l'unica soluzione è utilizzare un radar.

Una condizione particolare si verifica con la neve; se la strada è stata pulita si nota un forte contrasto con i bordi ricoperti di neve, che appaiono molto scuri (fig.3.19). Questa condizione è una condizione quasi ideale per il riconoscimento della strada.

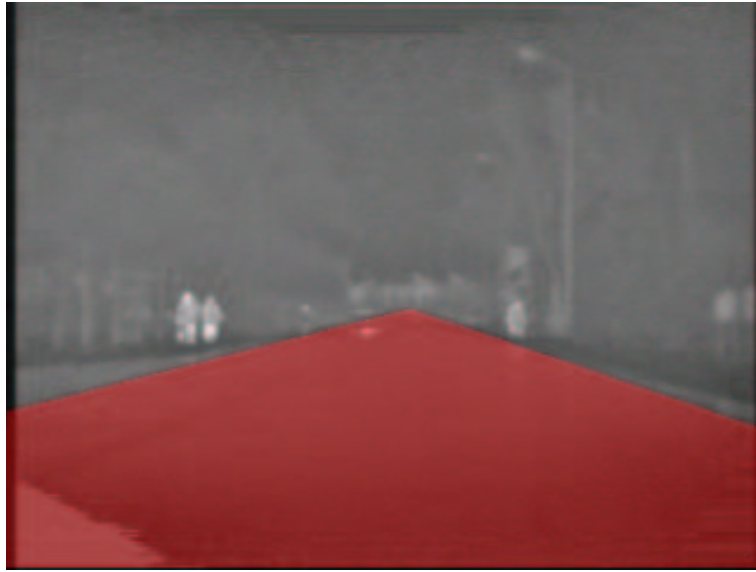


Figura 3.19: Neve ai bordi della strada

3.1.5 Tempi di esecuzione

Un altro importante metro di giudizio del funzionamento delle procedure sviluppate sono i tempi di esecuzione che devono essere necessariamente molto bassi per poter permettere l'utilizzo realtime delle funzioni. Purtroppo le funzioni sviluppate, sono piuttosto pesanti e richiedono un certo tempo, ma compilando con ottimizzazione per sfruttare a pieno le possibilità del PC si sono ottenuti dei risultati abbastanza buoni.

Nella tabella 3.1 sono riportati i tempi di esecuzione misurati in millisecondi di un test effettuato su di personal computer fornito di CPU AMD Athlon con frequenza di clock di 1,34 GHz con 256 MB di RAM DDR 266, utilizzando le immagini fornite da un imaged server montato su una macchina in rete locale.

Come si può notare vi sono alcune funzioni semplici, come la ricerca del punto di fuga e la ricerca di pali, che sono molto veloci, ma altre che richiedono più tempo.

Il filtro mediano ha un tempo di esecuzione piuttosto elevato, dovuto alla natura stessa del metodo che prevede per ogni pixel l'ordinamento di 9 valori; nonostante i tentativi velocizzare tale metodo non è stato possibile; si è anche provato ad eliminare il

Funzione	Min	Med	Max
Filtro mediano	22	26	30
Trasformata di Hough	32	39	48
Ricerca punto di fuga	6	6	11
Espansione strada	0	60	96
Ricerca pali	1	1	2
Ricerca segnali stradali	20	37	44
Totale	94	179	209

Tabella 3.1: Tempi di esecuzione

filtro ma si è verificato un peggioramento dei risultati e si è dunque deciso di mantenere il filtro nonostante il suo tempo di esecuzione.

La trasformata di Hough è la funzione chiave di tutto il progetto; all'interno del tempo di esecuzione è riportato anche quello di Sobel, che comunque ne rappresenta una parte praticamente trascurabile. Nonostante vengano applicate ad una parte dell'immagine e non all'immagine intera, bisogna tener conto che in totale le trasformate da fare sono 6 e il tempo di calcolo di una singola trasformata è piuttosto alto. L'unica soluzione per abbassare il tempo di esecuzione sarebbe rinunciare ad un tipo di trasformata, preferibilmente quella che abbiamo chiamato precisa, ma questo peggiorerebbe notevolmente i risultati ottenuti.

Un discorso simile vale per la ricerca dei segnali stradali, il calcolo della trasformata di Hough generalizzata è molto pesante e viene ripetuto per dieci volte, una volta per ogni raggio del cerchio cercato, è naturale che il tempo di esecuzione sia così alto, ma se proviamo, ad esempio, a cercare cerchi con raggi differenti in un'unica passata otteniamo risultati pessimi.

Il tempo di espansione della strada è il più variante; il motivo di questa variazione è duplice. Esiste una notevole differenza di operazioni da eseguire nel caso il metodo di ricerca del punto di fuga funzioni oppure no, nel primo caso l'espansione parte da un'area piuttosto grande e l'espansione, che richiede numerose scansioni dell'imma-

gine, è minima, nel secondo caso è necessaria un'espansione notevole per terminare il metodo. Inoltre a fare la differenza è senza dubbio il numero di scansioni dell'immagine necessarie: questo numero è difficile da calcolare e può variare molto anche in due immagini consecutive.

Il tempo totale di elaborazione si mantiene praticamente sempre al di sotto dei duecento millisecondi, che è un tempo relativamente buono per un'elaborazione di questo tipo. Dobbiamo però tenere conto che la prima esecuzione del programma richiederà un tempo un po' superiore a causa della necessità di inizializzare tutte le variabili, ed in particolare di calcolare il valore degli elementi dei vettori di seno, coseno e tangente.

3.2 Conclusioni

Per concludere è necessario verificare quanto i risultati raggiunti rispecchino gli obiettivi che ci si era posti all'inizio del lavoro.

L'obiettivo principale, cioè quello di trovare il punto di fuga e la strada è quello raggiunto in maniera più esaustiva: il metodo funziona bene praticamente con qualsiasi condizione, è abbastanza preciso e può essere applicato senza particolari modifiche. Tutti i tipi di strade possono essere individuate anche se con alcune imprecisioni, che possono essere causate dalle ombre, dal rumore della strada e da segnaletica orizzontale. La sua vera restrizione è sulle curve, ma questo era già stato previsto da principio. Il metodo di espansione della strada, per quanto funzionante, non da risultati di livello paragonabile al metodo principale, ma resta comunque un'alternativa possibile nei punti critici.

Il riconoscimento di pali e bordi di edifici è stato in buona parte attuato e trova nella velocità di esecuzione la sua maggior forza, infatti non necessita dell'applicazione della trasformata di Hough. La distinzione tra pali e angoli di costruzioni, che era un obiettivo iniziale, non è riuscita, ma, come già spiegato, questo è un problema

non centrale. Sono superabili anche gli errori introdotti dal riconoscimento di linee verticali sui bordi dei pedoni, come spiegato precedentemente. Se confrontato con la funzione di ricerca dei pedoni il metodo dà buoni risultati permettendo di riconoscere numerosi errori.

Per quanto riguarda i cartelli circolari possiamo dire di aver raggiunto un buon equilibrio tra l'identificazione di tutti i segnali e l'eliminazione di tutti gli oggetti di forma quasi circolare, ma l'obiettivo è raggiunto solo a metà a causa dell'intermittenza con cui vengono individuati i cartelli. Anche in questo caso i risultati nell'individuazione degli errori della funzione di riconoscimento dei pedoni sono buoni.

Nel complesso dunque si può dire che gli obiettivi siano raggiunti; senza dubbio sono possibili miglioramenti e sviluppi, che sono esposti nel paragrafo seguente.

3.3 Possibili sviluppi futuri

Gli sviluppi possibili del lavoro fatto sono diversi, alcuni sono legati ad un perfezionamento delle procedure sviluppate, altri invece implicano l'utilizzo di strumentazioni diverse dalla sola telecamera ad infrarossi.

Per migliorare l'identificazione della strada e della sua forma, in particolare nelle curve, sarebbe interessante sfruttare i segnali di controllo della vettura e verificare quando l'automobile sta curvando per poter interrompere il funzionamento del metodo principale di riconoscimento della strada ed inserire il metodo alternativo, d'altra parte quest'ultimo dovrebbe essere affinato; sono stati fatti dei tentativi in questo senso, ma hanno avuto scarso successo. In particolare si potrebbe cercare di uniformare i risultati ottenuti ad uno o più modelli di strada per semplificare l'elaborazione ed avere dei risultati più precisi. Si potrebbe inoltre bloccare il funzionamento di qualsiasi metodo quando un ostacolo si trova molto vicino al veicolo, ad esempio un'automobile se ci troviamo in coda, ciò è facilmente attuabile con l'utilizzo di un radar.

Sarebbe utile inserire una buona distinzione tra pali e angoli di edifici, questo può essere ottenuto in molti modi, in primo luogo affinando il metodo della luminosità spiegato nel capitolo precedente, magari cercando di individuare i pali nella loro interezza e non di individuare solo i bordi. Per identificare gli edifici si potrebbero cercare i bordi superiori, cioè i tetti, che danno quasi sempre un buon contrasto, tenendo conto di poter incontrare numerose forme diverse, che ci costringerebbero a considerare un'ampia gamma di modelli diversi. Inoltre un'ottima distinzione si potrebbe ottenere con un radar, che sarebbe in grado di fornire informazioni sull'area dell'ostacolo, che unite alle informazioni sui bordi ci darebbero una conoscenza molto particolareggiata dell'ambiente.

Inoltre è necessario raggiungere una maggior confidenza nel riconoscimento dei cartelli circolari, procedere al riconoscimento del tipo di segnale e determinarne il significato rappresenterebbe un enorme passo per un progetto di guida automatica: a mio parere il metodo più semplice per raggiungere tale obiettivo è passare dagli infrarossi al visibile e magari lavorare con una telecamera a colori, in questo modo l'informazione della forma potrebbe trovare conferma col colore.

Appendice A

GOLD

GOLD è una piattaforma per lo sviluppo di applicazioni per la visione che è dotata di molti strumenti indispensabili per rendere possibile o facilitare l'elaborazione delle immagini. Le principali funzioni e caratteristiche di GOLD sono:

- gestione dell'acquisizione di dati;
- possibilità di aggiungere nuove funzionalità;
- interfaccia grafica semplice;
- gestione delle finestre di output e della grafica;
- gestione di strumenti utili.

Le immagini possono essere acquisite in numerosi modi: si possono utilizzare delle directory contenenti una sequenza di immagini, anche compresse, oppure si possono acquisire da un imaged server, o anche real-time da un frame grabber a cui è collegata una telecamera; sono gestite immagini sia mono che stereo. Tutte le immagini acquisite devono essere calibrate, esistono strumenti appositi per eseguire questa operazione sia durante l'esecuzione che precedentemente. È possibile anche acquisire dati da strumenti diversi dalle telecamere, come un radar.

GOLD è quasi completamente modulare perciò l'inserimento di nuove funzionalità è estremamente semplice e naturale; le funzioni possono essere applicate in sequenza a quelle già presenti o in modo indipendente.

L'interfaccia grafica di GOLD è semplice e compatta e tramite l'uso di bottoni e pannelli supplementari facilita la corretta esecuzione di tutti i programmi.

I risultati ottenuti possono essere mostrati in modo molto chiaro ed elegante grazie alle numerose possibilità di GOLD che gestisce anche la creazione e la visualizzazione delle finestre di output e permette di salvare con estrema facilità i risultati delle elaborazioni.

All'interno della piattaforma sono sviluppati anche alcuni strumenti utili al controllo dei programmi sviluppati, come ad esempio il pannello di visualizzazione dei tempi di esecuzione.

Prima di eseguire GOLD è necessario settare alcune variabili contenute nel file di inizializzazione *gold.ini*. Modificando questo file è possibile scegliere su quali immagini lavorare e da dove reperire tali immagini, si può decidere quali funzionalità sarà possibile eseguire, e si possono fornire al sistema altre informazioni utili sulla sequenza da elaborare e sulle opzioni del programma.

Eseguendo GOLD viene visualizzato un pannello (fig.A.1) che è l'interfaccia grafica del programma.

Per cominciare la scansione di una sequenza di immagini è sufficiente schiacciare *start*, per terminarla schiacciare *stop*, il tasto *next* consente invece l'elaborazione del solo frame successivo a quello visualizzato.

La manopola in basso a sinistra gestisce la direzione e la velocità di scansione della sequenza; i due valori riportati indicano i frame di avanzamento per ogni esecuzione del programma e il tempo minimo che deve trascorrere tra la visualizzazione di due frame successivi. Normalmente il valore del primo elemento è impostato ad uno, per scandire la sequenza dall'inizio alla fine immagine per immagine, se si vuole bloccare

l'esecuzione su un frame è sufficiente settare il valore a zero, mentre se vogliamo spostarci nel verso opposto dobbiamo utilizzare un valore negativo.

I due *radio button* con valore zero servono per scegliere se eseguire o meno le funzionalità selezionate. Scegliendo il valore zero vengono visualizzate le immagini acquisite senza modifica, scegliendo il valore uno vengono mostrati i risultati delle funzioni selezionate tramite i bottoni, nel nostro caso *IR* e *Dummy*, cliccando sui bottoni sottostanti, che riportano la scritta *panel*, si accede ai pannelli di controllo della funzionalità.

Cliccando sul bottone *profiler* vengono visualizzati i tempi di esecuzione delle funzionalità selezionate (vedi anche fig.B.1).

Cliccando sul bottone *acquisition* è possibile calibrare le immagini acquisite. I valori di default vengono letti da un file di descrizione della sequenza (*DSC*), ma possono essere modificati durante l'esecuzione. I valori riportati rappresentano la posizione e l'inclinazione della telecamera. In particolare il valore *Left theta* viene utilizzato per il calcolo dell'orizzonte.

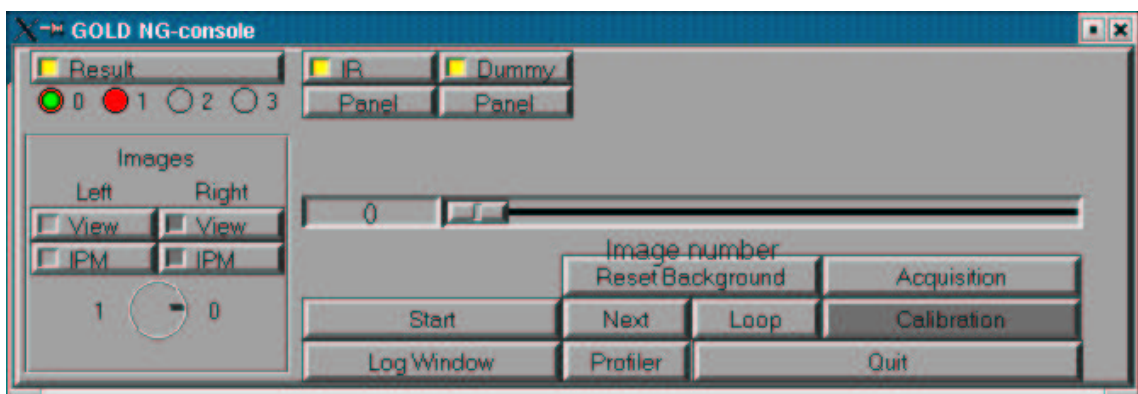


Figura A.1: Pannello di controllo di GOLD

Lavorare su di una piattaforma in continua modifica, come GOLD, porta alcuni vantaggi ed alcuni svantaggi. Si tratta di mantenere il proprio codice aggiornato alle

nuove versioni della piattaforma e dunque ad un lavoro di limatura e risistemazione ogni volta che una nuova versione viene caricata, ma d'altra parte ciò rende possibile sfruttare i vantaggi offerti dalle nuove funzionalità sviluppate. Due sono le funzionalità che hanno rivoluzionato GOLD nell'ultimo periodo, la prima è la stabilizzazione dell'immagine, la seconda l'aggiunta delle funzionalità grafiche delle librerie OpenGL.

Il tool di stabilizzazione viene applicato alle immagini in ingresso ed elimina con un'apposita procedura il movimento ondulatorio provocato da buche o imperfezioni del manto stradale per rendere disponibili all'utente una serie di immagini stabili, in cui cioè l'orizzonte è pressoché fisso. I vantaggi apportati da tale strumento al metodo di riconoscimento del punto di fuga e della strada sono stati notevoli, infatti la conoscenza dell'esatta ubicazione dell'orizzonte facilita notevolmente l'identificazione del punto di fuga.

Le librerie OpenGL invece hanno decisamente ampliato le possibilità nella visualizzazione dei risultati ottenuti, in particolare introducendo l'alpha blending, e hanno notevolmente velocizzato l'esecuzione della parte grafica del programma.

Appendice B

Manuale utente

Aperto il pannello di controllo di Dummy, come spiegato nell'appendice A, si possono configurare i valori di alcune variabili utilizzate nelle varie funzioni sviluppate. Il pannello si presenta come un quadrato con diversi bottoni ed alcune barre di scorrimento (fig.B.2) che servono per abilitare alcune funzioni e per regolare i valori di alcune variabili.

Il bottone *Show Dummy* serve per mostrare la finestra di output di Dummy, che contiene tutti i risultati delle elaborazioni eseguite.

Il bottone *Show Edges* mostra i bordi trovati col metodo di Sobel già soglia e binarizzati.

La barra *Edges %* regola la percentuale di pixel accesi nell'immagine dei bordi, conviene regolarla correttamente per avere dei buoni risultati; con immagini semplici si può lavorare bene con un valore del 5%, con immagini più complesse, soprattutto nella zona superiore dell'immagine, conviene utilizzare valori più alti, anche fino al 10%.

Il bottone *Show Lines* mostra le linee trovate con i metodi di Hough.

La barra *Search Multiplier* fa variare la soglia che regola l'individuazione delle linee nel dominio di Hough; se il metodo funziona normalmente tale valore può essere

lasciato a 1, come viene settato di default, ma se i bordi delle strade non vengono riconosciuti bene da Sobel, conviene abbassare il valore anche fino a 0,5.

Il bottone *Show Vanishing Point* mostra il punto di fuga e la strada trovata con il metodo principale di ricerca.

Il bottone *Show Road* mostra la strada trovata con il metodo secondario di ricerca, chiamato di espansione.

La barra *Road IN Threshold* regola la differenza massima accettabile per un quadratino all'interno della strada nel metodo di espansione; abbassare tale valore, com'è consigliato in immagini molto fredde, permette di avere una miglior identificazione degli ostacoli presenti sulla strada, il valore di default, 70, è consigliabile per immagini piuttosto calde.

La barra *Road OUT Threshold* regola la differenza massima affinché un quadratino possa essere considerato parte della strada; conviene abbassare tale valore per le immagini fredde in cui il contrasto è molto lieve, anche lavorando su immagini calde non conviene superare il valore di default, 40. La corretta assegnazione di queste variabili è molto importante, visto che il metodo risente molto della temperatura delle immagini, bisogna fare attenzione che l'espansione non sia eccessiva ma allo stesso tempo neanche troppo restrittiva. È conveniente che il valore della prima variabile venga mantenuto superiore a quello della seconda.

Il bottone *Show Horizon* mostra l'orizzonte su cui giace il punto di fuga e quello ideale.

Il bottone *Show Poles* mostra i pali e i bordi degli edifici trovati.

La barra *Vertical Threshold* regola la soglia sul numero minimo di pixel di linee verticali che deve essere presente in una colonna affinché questa venga considerata parte di un palo; è bene mantenere questo valore fisso perché abbassandolo si rischia di identificare erroneamente pedoni e altri oggetti ed alzandolo si possono perdere dei pali che altrimenti verrebbero individuati.

La barra *Length Treshold* regola la soglia sulla lunghezza minima di una linea verticale affinché questa venga considerata parte di un palo. Le indicazioni riportate per la variabile precedente valgono anche in questo caso.

La barra *Spot* seleziona il numero di pixel spenti che possono essere presenti in una linea verticale; conviene mantenere fisso tale valore che è stato fissato dopo molti tentativi a 10.

Il bottone *Show Circles* mostra i cartelli circolari trovati.

La barra *Circle Treshold* regola la soglia di voti nel dominio di Hough affinché un punto venga considerato il centro di un cartello. Alzare tale soglia serve per eliminare tutti i falsi veri presenti, abbassarla è utile per trovare tutti i cartelli, conviene sempre comunque non utilizzare valori inferiori a 1,5 e superiori a 3,5.

Il bottone *Show Profiler* mostra il profiler dei tempi di esecuzione; ci sono tre possibili visualizzazioni (fig.B.1):

Simboli che mostra una tabella che contiene il valore attuale, quello medio, quello massimo e quello minimo del tempo di esecuzione

Confronti che utilizza un istogramma per rappresentare i valori attuali

Variazioni che non è implementato

Il bottone *Show Vertical Edges* mostra i bordi verticali trovati.

Segue una tabella riassuntiva delle funzioni dei bottoni e delle barre di scorrimento (tab.B.1).

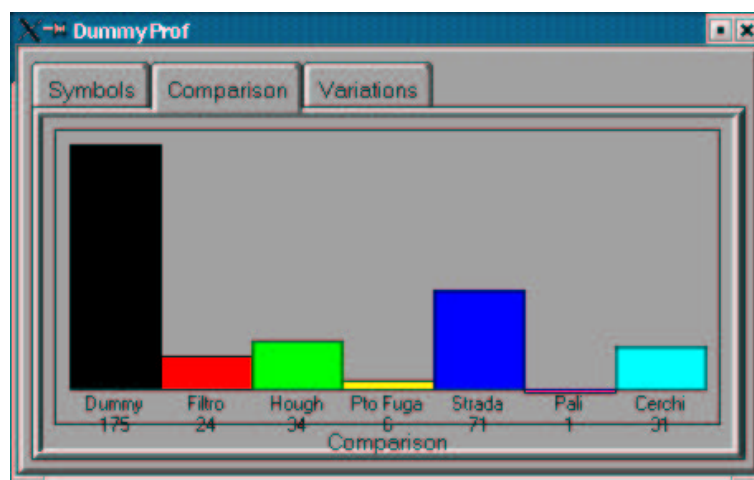


Figura B.1: Grafico dei tempi di esecuzione delle funzioni

Variabile	Funzione
Show Dummy	Mostra la finestra di output
Show Edges	Mostra i bordi trovati col metodo di Sobel
Edges %	Seleziona la percentuale di pixel accesi nell'immagine dei bordi
Show Lines	Mostra le linee trovate coi metodi di Hough
Search Multiplicator	Regola le soglie sui voti delle linee
Show Vanishing Point	Mostra il punto di fuga e la strada
Show Road	Mostra la strada ottenuta col metodo alternativo
Road IN Treshold	Regola la soglia per i quadrati di strada interni
Road OUT Treshold	Regola la soglia per i quadrati di strada esterni
Show Horizon	Mostra l'orizzonte trovato e l'orizzonte ideale
Show Poles	Mostra i pali e i bordi degli edifici trovati
Vertical Treshold	Regola la soglia sul numero di pixel di linee verticali
Length Treshold	Regola la soglia sulla lunghezza delle linee verticali
Spot	Seleziona il numero di pixel vuoti spenti in una linea
Show Circles	Mostra i cartelli rotondi trovati
Circle Treshold	Regola la soglia sulla ricerca dei cartelli
Show Profiler	Mostra la tabella dei tempi di esecuzione
Show Vertical Edges	Mostra i bordi verticali trovati

Tabella B.1: Funzione delle variabili del pannello di controllo

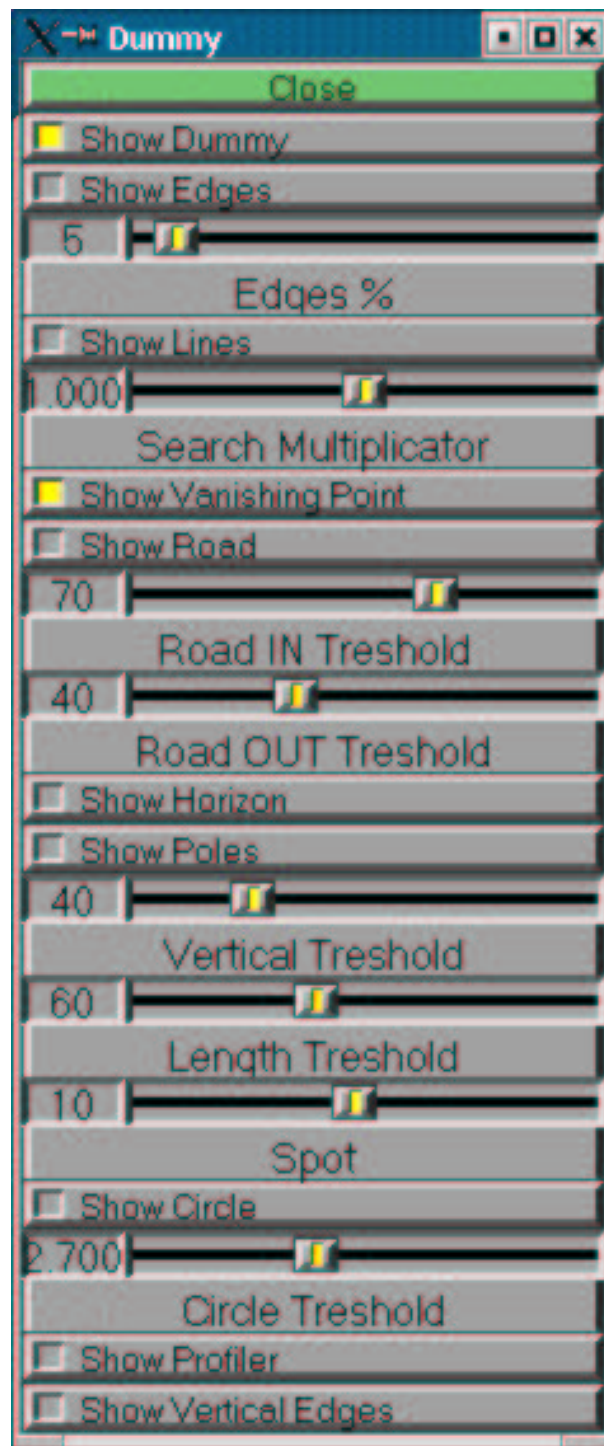


Figura B.2: Pannello di controllo delle funzioni

Ringraziamenti

Se dovessi ringraziare tutte le persone che mi hanno aiutato, in qualsiasi modo, nello sviluppo di questa tesi i ringraziamenti rischierebbero di essere più lunghi della tesi; mi limiterò a ricordare chi mi ha aiutato più direttamente.

Vorrei ringraziare innanzitutto il professor Alberto Broggi per avermi seguito con attenzione ed estrema gentilezza nel procedere della tesi ed il professor Virginio Cantoni che ha accettato molto gentilmente di farmi da relatore.

Non posso dimenticare l'intera staff del progetto ARGO, il professor Massimo Bertozzi e l'ingegner Alessandra Fascioli, che hanno seguito il procedere del lavoro aiutandomi qualora ne avessi bisogno.

Vorrei inoltre ricordare i compagni del laboratorio di visione di Parma: Paolo Grisleri, Marcello Carletti, Paolo Federici, Amos Tibaldi e Isabella Fedriga, tutti prodighi di savi consigli e pronti a risolvere i miei problemi sull'utilizzo di Linux, di GOLD, dei vari account e, nel caso, sulla grammatica inglese.

Una menzione speciale ai miei vecchi amici Dome, Fraka, Andrea e Al, i migliori compagni di università che qualsiasi studente possa trovare ;-)

Bibliografia

- [1] A. F. A. Broggi, M. Bertozzi. *ARGO Home Page*. <http://www.argo.ce.unipr.it>, dic. 2002.
- [2] S. Donati e A. Gilardini. *Tecniche infrarosso*. AEI Associazione Elettrotecnica ed Elettronica Italiana, set. 1997.
- [3] R. Katsuri e R. C. Fain. *Computer Vision: Advances and Applications*. IEEE Computer Society Press, Los Alamitos, California, 1991.
- [4] R. Katsuri e R. C. Fain. *Computer Vision: Principles*. IEEE Computer Society Press, Los Alamitos, California, 1991.
- [5] A. Marcolli. *Teoria del campo: corso di educazione alla visione*. Sansoni.
- [6] L. A. P. Barham, X. H. Zhang e M. Vache. The DEvelopment of a Driver Vision Support System using Far Infrared Technology: Progress to Date on the DARWIN Project. *IEEE Intelligent Vehicles Symposium 2000*, pp. 545–549, ott. 2000.
- [7] I. Pitas. *Digital Image Processing Algorithm*. Prentice Hall, Hemel Hempstead, Hertfordshire, UK, 1993.
- [8] D. Ranzato. *Localizzazione di autoveicoli in immagini all'infrarosso*. Tesi di laurea, Università degli Studi di Pavia - Facoltà di Ingegneria, 2001.

