Esperienze di Elaborazione di Immagini in Tempo Reale su Veicoli in Movimento

Alberto Broggi
Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Parma
I-43100 Parma, Italy
broggi @ CE.UniPR.IT

Sommario

In questo lavoro viene presentato il sistema integrato sul veicolo mobile Argo presso il Dipartimento di Ingegneria dell'Informazione di Parma. Argo e` un veicolo sperimentale sviluppato nell'ambito del Progetto Finalizzato Trasporti 2 del CNR, rivolto allo studio di tecniche e tecnologie innovative per l'aumento della sicurezza stradale. Sul veicolo sono state installate due telecamere stereo che inviano le immagini a sistemi di elaborazione specializzati: vengono considerate due diverse soluzioni, una basata su hardware dedicato (PAPRICA) e una su hardware commerciale (PC Pentium MMX).
L'elaborazione consiste nel determinare la corsia di marcia e la posizione di eventuali ostacoli. I sistemi di output sono basati su avvertimenti acustici, ottici (sia su monitor che su pannello di controllo a led), ed è in fase di studio anche l'integrazione di un attuatore che consenta al veicolo di intervenire in situazioni di pericolo o anche di raggiungere una guida completamente automatizzata.

1. Introduzione

La capacità di percepire l'ambiente circostante è essenziale per il miglioramento della sicurezza della guida di autoveicoli. La visione artificiale rappresenta un utile mezzo per percepire l'ambiente circostante senza alterarlo grazie all'utilizzo di telecamere (*sensori passivi*); infatti altri sensori, come laser, rilevatori di distanza, sensori tattili o acustici, permettono l'acquisizione di dati riguardanti l'ambiente esterno in maniera anche più diretta, però hanno lo svantaggio di introdurre nell'ambiente stesso stimoli o segnali che possono portare a problemi di "inquinamento" (*sensori attivi*). In molti campi di applicazione la presenza di questi disturbi può indurre problemi di non semplice risoluzione. In particolare, nel caso della navigazione in ambienti non controllati, in cui è coinvolto un numero elevato di veicoli, l'utilizzo di sensori attivi porterebbe a interferenze e a un inquinamento inaccettabile dell'ambiente, compromettendo potenzialmente il funzionamento dei sistemi che utilizzano le misure effettuate. L'utilizzo di tecniche di visione artificiale è quindi basilare nel progetto e nello sviluppo di veicoli con maggior grado di sicurezza.

Tuttavia la guida automatica di veicoli basata su visione in ambienti esterni presenta alcuni problemi, intrinseci all'uso della visione:

- a differenza degli ambienti interni (laboratori, officine,…), l'ambiente esterno non può essere fortemente strutturato. Infatti le strade per il transito dei veicoli presentano un ampia varietà di tipologie e il dotarle di strutture di supporto alla visione artificiale risulta proibitivo dal punto di vista dei costi. Un veicolo autonomo deve quindi essere in grado di basarsi esclusivamente sulla segnaletica esistente, senza richiedere nessun'altra infrastruttura;
- l'elaborazione delle immagini acquisite deve essere sufficientemente robusta rispetto alle variazioni della luminosità e del contrasto della scena. Situazioni come ombre e scarsa visibilità (foschia, pioggia) devono essere previste; inoltre il sistema deve poter affrontare anche cambi repentini delle condizioni di luminosità come, ad esempio, il passaggio tra zone soleggiate e zone in ombra, oppure anche l'ingresso e l'uscita da gallerie, ecc.
- data la criticità dell'applicazione, è necessario disporre di elevate capacità di elaborazione: la potenza di calcolo richiesta cresce con la velocità del veicolo, inoltre l'uso di telecamere comporta l'elaborazione di grandi quantità di dati. Si rende quindi necessario investigare sia l'utilizzo di architetture dedicate che lo sviluppo di opportuni algoritmi per ottenere elaborazioni in tempo reale. Tuttavia è evidente che il successo dei sistemi sviluppati non risiede solamente nella loro efficacia ma è anche strettamente legato al loro costo commerciale. È quindi necessario utilizzare sistemi standard e già diffusi oppure sviluppare soluzioni dedicate ma a basso costo.

Nonostante la complessità del problema, la visione artificiale è stata utilizzata con successo in un ampio numero di applicazioni e con differenti finalità, come, ad esempio: l'individuazione della corsia di marcia, l'inseguimento di altri veicoli, il controllo dei sorpassi e tecniche di parcheggio automatizzate. Sono peraltro necessari ulteriori sviluppi e indagini per portare queste soluzioni in prodotti disponibili sulle vetture commerciali. I principali problemi per raggiungere questo obiettivo sono:

- il costo,
- l'efficacia e l'affidabilità del sistema,
- l'integrazione con il sistema veicolo e
- l'interazione con il guidatore.

Nel seguito saranno analizzate le due funzionalità di base: l'individuazione della corsia di marcia e l'individuazione degli ostacoli.

1.1 Individuazione della Corsia di Marcia

Nella maggior parte dei prototipi di veicoli autonomi, il controllo della traiettoria è effettuato calcolando la posizione del veicolo rispetto alla strada; in seguito, grazie ad opportuni attuatori, il veicolo viene mantenuto in una posizione corretta. Anche se alcuni prototipi sono stati progettati per funzionare su strade non strutturate o addirittura fuori strada [38][34], la maggioranza dei sistemi sviluppati si basa sul riconoscimento della segnaletica orizzontale, come, ad esempio, le linee che delimitano le corsie di marcia.
L'individuazione della segnaletica orizzontale richiede di affrontare alcuni problemi, come:

- le ombre proiettate da alberi o altri oggetti;
- le occlusioni della segnaletica da parte di altri veicoli o oggetti che possono trovarsi sul piano stradale; e
- le discontinuità della segnaletica stessa.

Tuttavia questo approccio permette di ridursi all'analisi di singole immagini senza dover analizzare sequenze di immagini come nel caso del flusso ottico [22][25][26] o più immagini acquisite concorrentemente come nel caso della visione stereoscopica [6][7][8]. Inoltre alcune ipotesi possono velocizzare e/o aiutare l'intero processo:

- invece di analizzare l'intera immagine, ci si può limitare ad analizzare specifiche regioni dove è più probabile trovare la segnaletica d'interesse (zona di attenzione) [24];
- il sistema può ipotizzare che le linee di corsia abbiano determinate caratteristiche [24][30][39], come: una forma particolare, siano parallele, varino con continuità, ...
- si può sfruttare la conoscenza della calibrazione del sistema di visione e ipotizzare inoltre che la geometria della strada sia nota a priori, per esempio, che la strada sia piatta senza dossi o cunette [3][36].

Inoltre gli approcci utilizzati per il riconoscimento della corsia di marcia possono essere ulteriormente suddivisi in due classi: quelli guidati dalle aspettative, in cui, per esempio, un modello ipotetico viene deformato fino a coincidere con la situazione reale [14][23][29][36]; e quelli guidati dai dati basati sull'estrazione di specifiche caratteristiche. Quest'ultimo approccio è preferibile, in quanto permette di limitare l'analisi delle immagini alla sola estrazione delle caratteristiche di interesse, e quindi richiede una minor potenza computazionale.

1.2 Localizzazione di Ostacoli

In alcuni sistemi sviluppati l'individuazione di ostacoli è basata sul riconoscimento ed è quindi limitata all'individuazione di specifici modelli aventi una ben precisa forma [26], sagoma [5] o colore [41]. Anche in questo caso ci si può limitare all'analisi di singole

immagini, ma chiaramente l'approccio fallisce nelle situazioni in cui gli ostacoli non corrispondono al modello utilizzato.
Se, più generalmente, si intende come ostacolo un qualsiasi oggetto che ostruisce il percorso del veicolo [3], l'individuazione degli ostacoli consiste nella determinazione dello spazio libero di fronte al veicolo, ovvero la zona in cui il veicolo può muoversi senza pericolo. Per affrontare questo problema si possono usare due tecniche differenti, ma entrambe basate sull'analisi di più immagini:

- il flusso ottico, basato sull'analisi di immagini acquisite in istanti diversi;
- la visione stereoscopica, basata su immagini acquisite da angolazioni differenti.

Queste tecniche quindi richiedono una grande potenza di calcolo, e inoltre devono essere implementate in maniera da sopportare sia i movimenti del veicolo, che le inevitabili imprecisioni nella calibrazione del sistema. Inoltre, nel caso del flusso ottico, gli ostacoli sono individuati *indirettamente* dall'analisi del campo di velocità nelle immagini; mentre nel caso della visione stereoscopica, l'individuazione degli ostacoli avviene in modo *diretto* e non è limitata dalla velocità del veicolo e degli ostacoli.
Nella sezione seguente sono analizzati gli algoritmi per l'individuazione della corsia di marcia e per la localizzazione di generici ostacoli.

2. Algoritmi di Elaborazione delle Immagini

I paragrafi seguenti presentano il sistema GOLD (Generic Obstacle and Lane Detection) [4][3][1], le cui funzionalità sono:

- individuazione della corsia di marcia (Lane Detection);
- localizzazione degli ostacoli (Obstacle Detection).

Entrambe queste funzionalità condividono la parte iniziale dell'elaborazione, basata su una trasformazione geometrica dell'immagine acquisita, mirata a omogeneizzare il contenuto informativo dei pixel. Tale trasformazione, descritta nel prossimo paragrafo, è nota come *Inverse Perspective Mapping* IPM [32], ed è stata utilizzata con successo sia per l'analisi di immagini stereoscopiche che per l'individuazione di ostacoli, ma può semplificare significativamente anche l'individuazione della segnaletica orizzontale.

2.1 Rimozione dell'Effetto Prospettico (IPM)

L'*Inverse Perspective Mapping* (IPM) è una tecnica ampiamente utilizzata [40][31][37][33][35] che permette di rimuovere l'effetto prospettico dalle immagini acquisite nel caso in cui si conoscano a-priori la conformazione della scena (come ad esempio nell'ipotesi di strada piatta) e i parametri del sistema di visione. La procedura di rimozione dell'effetto prospettico campiona l'immagine acquisita mappandone

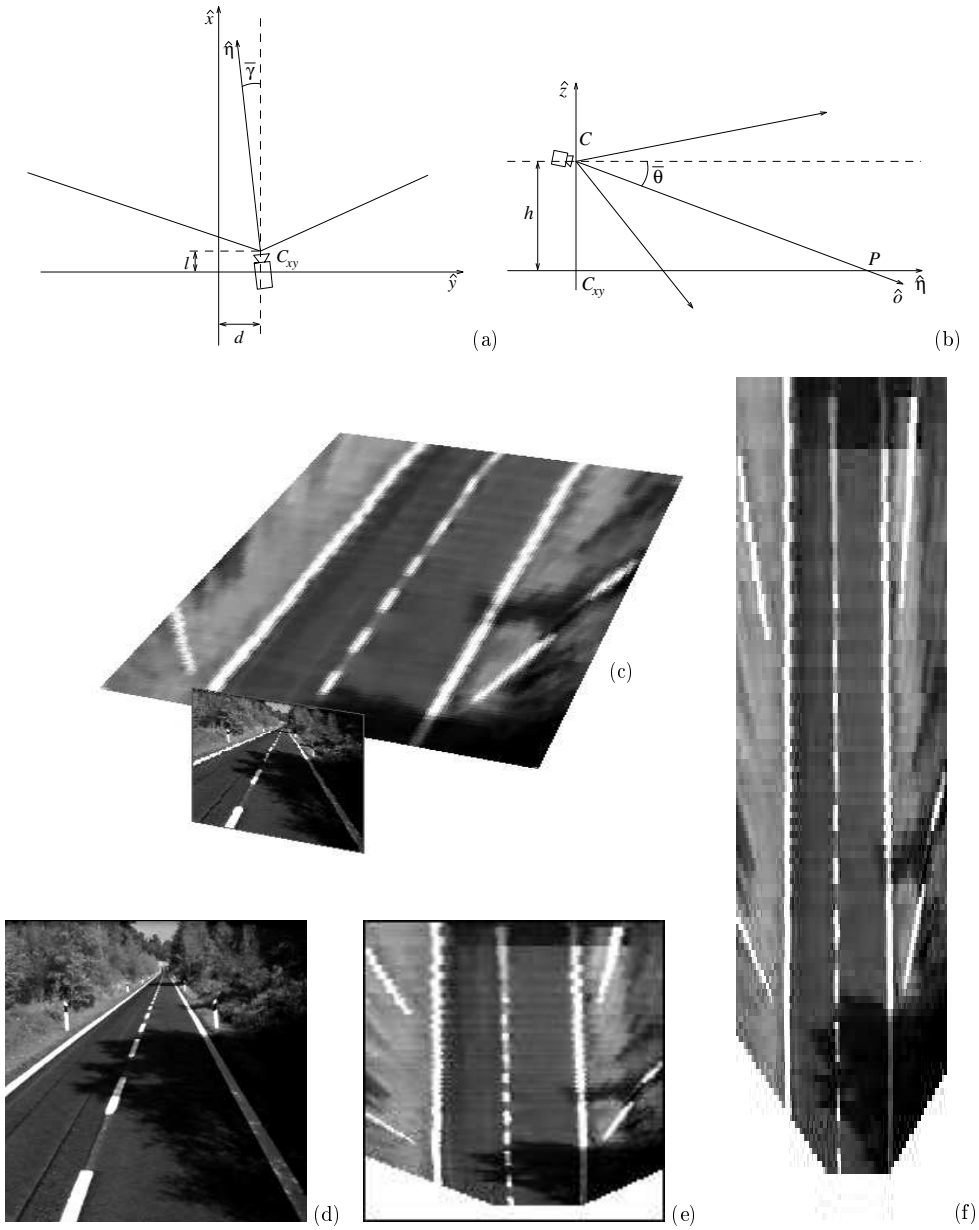

Figura 1: La rimozione dell'effetto prospettico: (a) e (b) i parametri necessari, (c) la relazione mondo tridimensionale - spazio immagine, (d) l'immagine acquisita, (e) l'immagine riorganizzata e (f) immagine riorganizzata con rapporto tra larghezza e lunghezza unitario.

ogni pixel in una differente posizione creando un array bidimensionale di pixel; l'immagine risultante (*immagine riorganizzata*) rappresenta una vista dall'alto della porzione di strada di fronte al veicolo (figura 1.e).
Definiamo i due spazi euclidei:

- $W \equiv \{(x,y,z)\} \in E^3$ è lo spazio tridimensionale che rappresenta il mondo reale (spazio del mondo reale), mentre
- $I \equiv \{(u,v)\} \in E^2$ rappresenta lo spazio bidimensionale dell'immagine dove la scena tridimensionale è proiettata (spazio immagine).

Partendo dall'ipotesi di strada piatta di fronte al veicolo, l'immagine acquisita dalla telecamera appartiene allo spazio *I*, mentre l'immagine riorganizzata è definita come il piano z=0 dello spazio tridimensionale *W*, quindi il processo di riorganizzazione consiste nel proiettare l'immagine acquisita sul piano $z = 0$ (vedi figura 1).
Per rimuovere l'effetto prospettico è necessario conoscere i seguenti parametri (figure 1.a e 1.b):

- punto di osservazione: la posizione della telecamera $C \equiv (l,d,h) \in W$;
- angolo di osservazione: la posizione dell'asse ottico è determinata da due angoli:
  - ✓ γ_c: l'angolo formato dalla proiezione dell'asse ottico sul piano $z=0$ (xy) e dall'asse x;
  - ✓ θ_c: l'angolo formato dalla proiezione dell'asse ottico sul piano yz e dall'asse z;
- apertura: l'apertura dell'obbiettivo della telecamera (2α);
- risoluzione: la risoluzione del sistema di visione ($n \times n$).

La trasformazione $g: I \rightarrow W$ può essere definita come:

$$\begin{cases} x(u,v) = h \times ctg\left[(\theta_c - \alpha) + u \times \dfrac{2\alpha}{n-1}\right] \times \cos\left[(\gamma_c - \alpha) + v \times \dfrac{2\alpha}{n-1}\right] + l \\ y(u,v) = h \times ctg\left[(\theta_c - \alpha) + u \times \dfrac{2\alpha}{n-1}\right] \times sin\left[(\gamma_c - \alpha) + v \times \dfrac{2\alpha}{n-1}\right] + d \\ z(u,v) = 0 \end{cases}$$

2.1.1 Rimozione dell'Effetto Prospettico da coppie di Immagini Stereoscopiche

Nel caso di visione stereoscopica, la rimozione dell'effetto prospettico permette di ottenere coppie di immagini riorganizzate [4]. L'utilizzo di appropriati parametri di riorganizzazione nel caso in cui sia valida l'ipotesi di strada piatta permette di far

coincidere i due risultati. In questo caso le due immagini riorganizzate sono identiche e due punti di esse aventi le stesse coordinate rappresentano lo stesso punto del piano stradale e sono quindi punti omologhi.

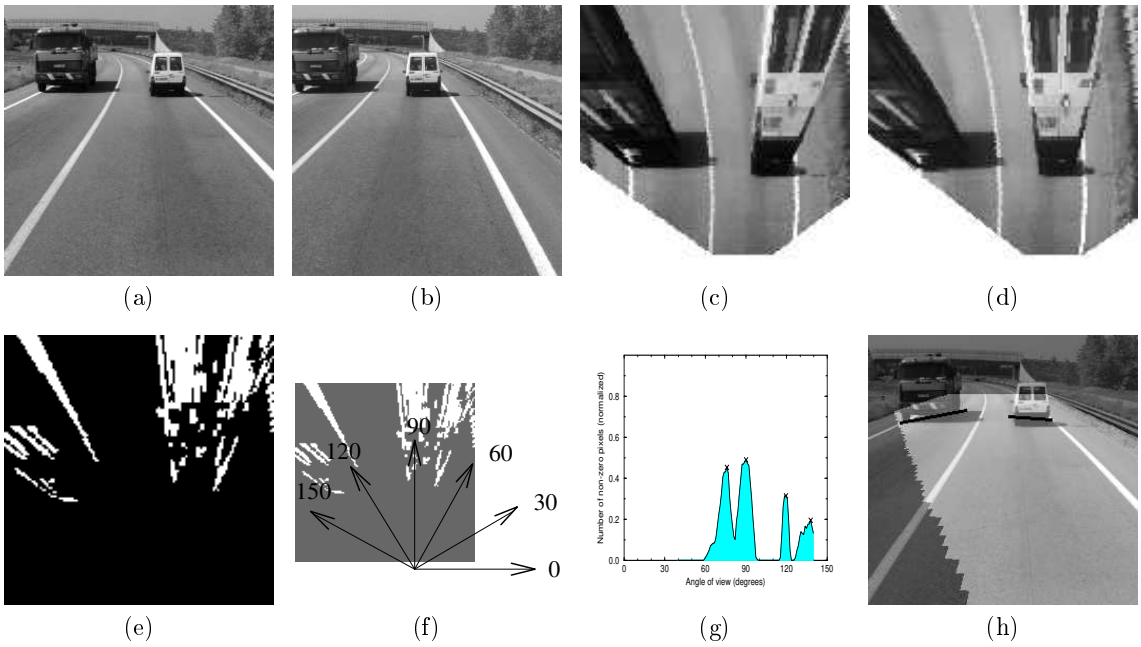

Figura 2: Rimozione dell'effetto prospettico e visione stereoscopica: (a) e (b) le due immagini acquisite, (c) e (d) le due immagini riorganizzate, (e) l'immagine differenza, (f) la posizione del fuoco e il calcolo dell'istogramma polare, (g) l'istogramma polare e (h) il risultato finale dell'elaborazione, in cui due tratti individuano la base degli ostacoli

2.2 Individuazione degli Ostacoli

La presenza di uno o più ostacoli sul piano stradale di fronte al veicolo rappresenta una deviazione significativa dall'ipotesi di strada piatta; in questo caso la rimozione dell'effetto prospettico da immagini stereoscopiche porta a produrre due immagini riorganizzate che differiscono (vedi figura 2).

La presenza di ostacoli sul piano stradale può quindi essere rilevata calcolando e analizzando le differenze tra le due immagini riorganizzate. Quindi, l'elaborazione di basso livello delle immagini si riduce al calcolo della differenza tra le immagini riorganizzate e alla sua binarizzazione grazie ad una soglia [3]. Inoltre la successiva analisi del risultato (*immagine differenza*) permette di calcolare sia la distanza che l'ingombro degli ostacoli.

Infatti, grazie al fatto che le due telecamere del sistema di visione inquadrano un ostacolo da differenti punti di vista, i due bordi verticali di un ostacolo vengono trasformati nell'immagine differenza in una coppia di triangoli (vedi figura 2.e). Naturalmente, data la non omogeneità cromatica e l'irregolarità dei bordi di un generico ostacolo, i triangoli generati possono essere frastagliati e non necessariamente disgiunti, ma pur tuttavia sempre riconoscibili. Grazie a ciò l'individuazione degli ostacoli si riduce all'individuazione di coppie di questi triangoli.
L'immagine differenza viene scandita radialmente lungo le rette passanti per un punto chiamato fuoco, per ogni angolo di scansione il numero di punti sopra soglia viene riportato in un diagramma chiamato istogramma polare (figura 2.g). Nel caso in cui il fuoco sia posizionato tra le proiezioni delle telecamere sul piano stradale, l'istogramma polare presenta dei picchi significativi in corrispondenza dei triangoli generati dalla presenza degli ostacoli. Inoltre alcune caratteristiche dei picchi come l'ampiezza o la larghezza permettono di raggruppare quelli corrispondenti allo stesso ostacolo.
Dato che la posizione di un picco nell'istogramma polare determina l'angolo di vista di un bordo verticale dell'ostacolo il raggruppamento dei picchi corrispondenti permette di calcolare l'ingombro di un ostacolo.
Un'ulteriore analisi dell'immagine differenza lungo le direzioni individuate dai picchi permette di individuare il vertice inferiore dei triangoli; questo rappresenta il punto di contatto tra la strada e l'ostacolo e quindi la sua individuazione rende possibile determinare la distanza dell'ostacolo.
La figura 3 presenta differenti situazioni operative con uno o più ostacoli di fronte al veicolo; una linea che codifica sia la posizione che la distanza dell'ostacolo individuato permette di valutare la correttezza dei risultati, si noti che gli ostacoli nella parte sinistra dell'immagine sono individuati solo parzialmente dato che sono solo parzialmente visti dalla telecamera destra del sistema di visione.

2.3 Individuazione della Corsia di Marcia

Nell'immagine riorganizzata le linee delimitanti le corsie di marcia vengono trasformate in linee chiare su fondo scuro pressoché verticali e di ampiezza costante. Quindi un pixel appartenente alle linee della corsia di marcia è più chiaro rispetto i suoi vicini lungo l'asse orizzontale. Di conseguenza la prima fase dell'individuazione della corsia di marcia è basata sulla ricerca di transizioni scuro-chiaro-scuro lungo l'asse orizzontale; ciascun pixel è confrontato con i suoi vicini orizzontali ad una data distanza e una nuova immagine che codifica il risultato viene prodotta. Dato che differenti condizioni di illuminazione come, ad esempio, la presenza di ombre rendono non costante la luminosità della segnaletica, viene applicato un ulteriore filtraggio che uniforma la luminosità delle linee verticali [10].
L'immagine risultante viene binarizzata e analizzata al fine di identificare catene di pixel sopra-soglia che vengono approssimate con spezzate. Un ulteriore fase di unione e approssimazione delle spezzate permette di colmare eventuali lacune nelle linee individuate.

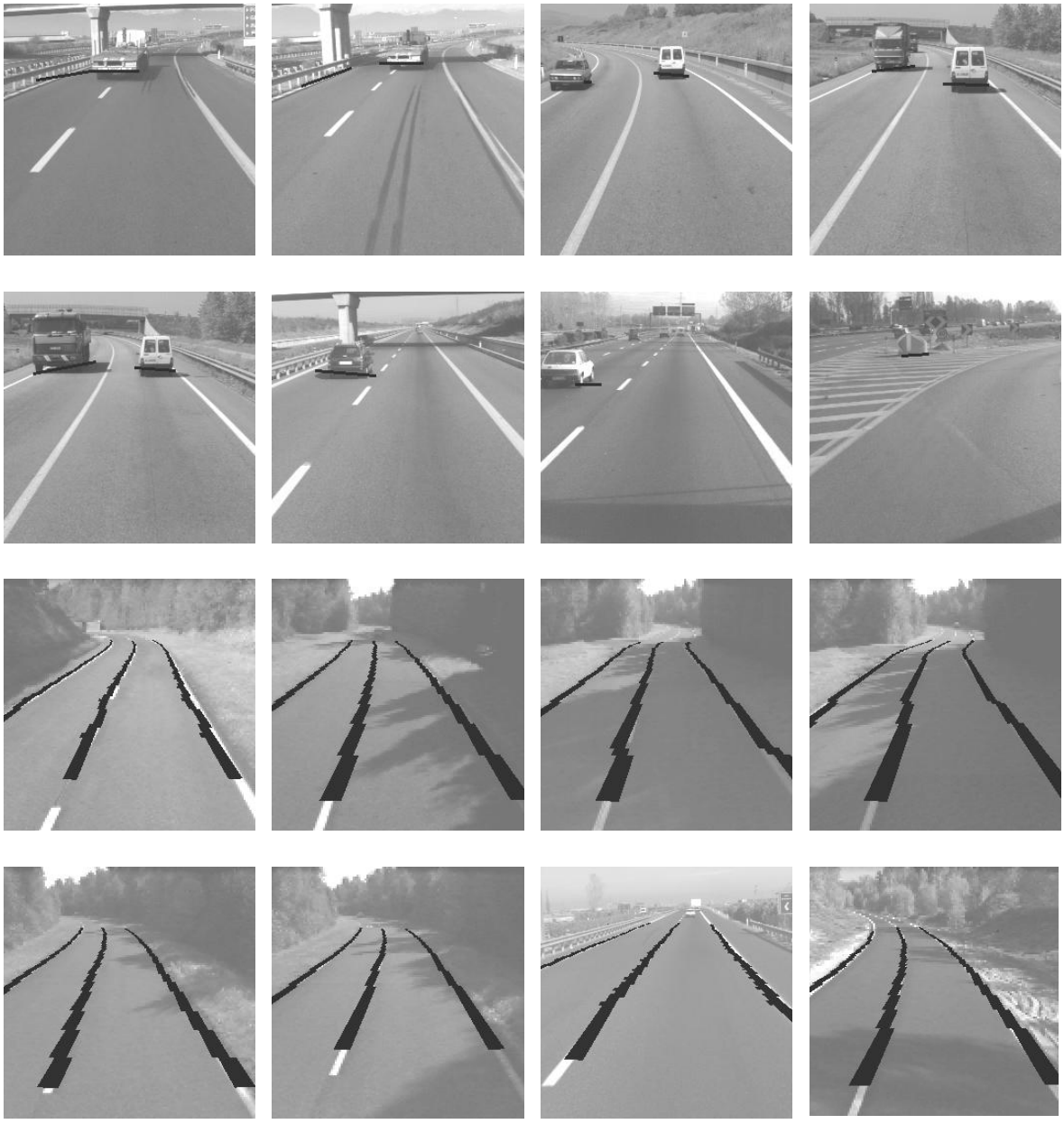

Figura 3: Risultati della localizzazione degli ostacoli e dell'individuazione della corsia di marcia ottenuti in differenti condizioni ambientali

Il sistema è stato configurato ipotizzando che la strada sia a due corsie; le spezzate precedentemente calcolate vengono confrontate tra di loro e raggruppate seguendo criteri di parallelismo. Ad ogni gruppo viene assegnato un punteggio legato ad alcuni parametri, fra cui: il parallelismo, la larghezza delle corsie e la regolarità delle linee stradali. Il gruppo che ottiene il punteggio più alto individua la corretta geometria delle corsia di marcia e, grazie all'ipotesi di strada piatta, permette di determinare la posizione del veicolo rispetto alla stessa.
La figura 3 presenta alcuni dei risultati ottenuti in differenti condizioni di illuminazione, nel caso di ombre e con ostacoli.

3. Architettura di Calcolo del Sistema

In questo specifico campo di applicazione il tempo di risposta del sistema rappresenta uno degli aspetti maggiormente critici poiché influisce direttamente sulla massima velocità consentita al veicolo. La scelta del sistema di calcolo è un quindi un punto chiave del progetto [20].

3.1 Hardware Specializzato

Nella prima versione del sistema GOLD è stato possibile ottenere un'elevata velocità di elaborazione grazie all'utilizzo di PAPRICA [18][17][19], un sistema hardware dedicato, sviluppato appositamente per essere utilizzato a bordo di un veicolo mobile come acceleratore hardware per compiti di elaborazione in tempo reale di immagini provenienti da telecamera. Lo sviluppo hardware dell'architettura PAPRICA è stato curato dal Dipartimento di Elettronica del Politecnico di Torino, mentre l'ambiente di programmazione [12] è stato sviluppato dall'unità del Dipartimento di Ingegneria dell'Informazione dell'Università di Parma.
PAPRICA è un'architettura SIMD massivamente parallela composta da 256 processori elementari e dotata di un insieme di caratteristiche studiate espressamente allo scopo di velocizzare l'esecuzione delle funzionalità di assistenza alla guida del sistema GOLD: l'integrazione diretta sulla scheda di un dispositivo di acquisizione delle immagini, un paradigma di elaborazione morfologico fondamentale per l'elaborazione di basso livello delle immagini, un meccanismo di virtualizzazione dei processori, un supporto hardware per la riorganizzazione efficiente delle immagini. Più precisamente il sistema PAPRICA consente di:

- acquisire e memorizzare nella Memoria Immagine coppie di immagini stereoscopiche a 256 toni di grigio con risoluzione 512x512 pixel a 25Hz o 512x256 pixel a 50Hz;

- elaborare in parallelo immagini mediante operatori morfologici;
- applicare trasformazioni geometriche alle immagini grazie ad un modulo hardware specifico [21] (questa caratteristica è di fondamentale importanza per la rimozione dell'effetto prospettico, realizzata in meno di 3ms su un'immagine 512x512 e producendo un'immagine 128x128);
- visualizzare i risultati intermedi e finali dell'elaborazione su un monitor connesso alla Memoria Immagine.

L'analisi dei principali colli di bottiglia del sistema ha consentito di sviluppare una nuova versione dell'architettura PAPRICA, attualmente in fabbricazione [27][16][2][28].
Il nucleo del nuovo sistema PAPRICA-3 è una architettura dedicata di tipo SIMD basata su un array lineare di Q processori elementari a singolo bit. L'array dei processori è connesso sia ad una Memoria Immagine esterna che ad un dispositivo di acquisizione delle immagini per mezzo di un bus dati bidirezionale a Q bit, perciò ogni operazione trasferisce un array completo di Q pixel in una volta, 1 bit per pixel in caso di trasferimenti in/da memoria e 8 bit per pixel in caso di acquisizione o output di immagini. Su singolo chip verranno integrati 32 processori e l'architettura sarà espandibile modularmente in modo che la dimensione dell'array dei processori coincida con la larghezza delle immagini in ingresso. Questa soluzione riduce il problema del meccanismo di virtualizzazione dei processori che si è dimostrato un aspetto critico del progetto. L'array dei processori è in grado di elaborare un'intera riga dell'immagine in ogni ciclo di clock.
Una rete di interconnessione dei processori riconfigurabile [12] è stata inclusa per supportare l'IPM: essa consente di riorganizzare una riga dell'immagine durante l'acquisizione della riga seguente [11].

3.2 Hardware Commerciale

Attualmente è in fase di valutazione come sistema di elaborazione anche un'architettura di tipo general-purpose: un processore Pentium 200 MMX standard che, grazie alla sua elevata potenza computazionale, consenta di soddisfare a costi contenuti le specifiche di tempo-reale.
In particolare la tecnologia MMX permette di aumentare le prestazioni del software utilizzando una tecnica di tipo SIMD: molteplici dati possono essere elaborati in parallelo mediante una singola istruzione. Le nuove istruzioni general-purpose supportate dalla tecnologia MMX realizzano operazioni aritmetiche e logiche in parallelo su molteplici dati impacchettati in quantità a 64 bit. Queste istruzioni, quindi, aumentano le prestazioni di applicazioni basate su algoritmi computazionalmente pesanti che realizzano operazioni locali ricorrenti su dati elementari. In particolare nell'elaborazione di immagini a toni di grigio l'informazione è rappresentata sotto forma di quantità a 8 bit (pixel) e quindi un'istruzione può operare su 64/8=8 pixel contemporaneamente. L'utilizzo di tali caratteristiche direttamente a livello di linguaggio assembly sembra preliminarmente consentire di raggiungere un aumento delle prestazioni di un fattore circa 10 nei casi di

differenza tra due immagini e operazioni morfologiche rispetto al codice generato da un compilatore C standard (GNU GCC 2.7.2).
La disponibilità di sistemi di elaborazione a basso costo ed elevate prestazioni come il Pentium facilita la fase di sperimentazione ma non soddisfa, al momento attuale, le specifiche di costo, consumi e dimensioni imposte dalla specifica applicazione. Una soluzione traducibile in termini di prodotto richiede a tutt'oggi architetture VLSI specifiche simili all'architettura PAPRICA o che utilizzano soluzioni miste analogico-digitali.

4. Il Veicolo Sperimentale ARGO

Nell'ambito del Progetto Finalizzato Trasporti 2 è stato acquisito un veicolo Lancia Thema 2000 al fine di sperimentare sul campo gli algoritmi e le architetture per la navigazione basata su visione sviluppati nel corso di questi anni.
La disponibilità di tale veicolo sperimentale, denominato ARGO[1][9], ha consentito di sviluppare, sperimentare e mettere a punto svariati approcci alla navigazione autonoma, tuttavia una particolare enfasi è stata data all'implementazione delle funzionalità di localizzazione di ostacoli e di determinazione della corsia di marcia, che rappresentano funzionalità basilari per la guida automatica dei veicoli [4].
Il veicolo ARGO, è stato attrezzato con un sistema di visione stereoscopico per l'acquisizione in tempo reale delle immagini stradali. Due diversi sistemi di elaborazione, discussi nel capitolo 3, sono stati integrati sul veicolo: l'architettura massivamente parallela PAPRICA e un sistema commerciale basato su PC.
L'aumento della sicurezza stradale si può raggiungere in tre differenti modalità:

- ad un primo livello l'azione del sistema si esplica nel fornire segnalazioni di pericolo al conducente;
- un intervento attivo del sistema è invece richiesto qualora si desideri non solo la segnalazione di situazioni pericolose, ma anche il recupero da tali situazioni (secondo livello);
- il terzo livello di sicurezza consiste nella guida completamente automatizzata.

[1] Il nome ARGO deriva dal nome di due distinti miti collegati alla navigazione e alla visione. Il primo è Argo, la prima nave della storia avente un nome, utilizzata da Giasone e da un gruppo di 50 eroi mitologici, gli Argonauti, per navigare in tutto il mare Mediterraneo alla ricerca del vello d'oro. Siccome la nave Argo era equipaggiata con tutti gli strumenti e le attrezzature necessari per il governo e la guida della nave in mare aperto, attraverso i forti venti e gli enormi flutti, essa può rappresentare l'archetipo del perfetto mezzo di navigazione. Il secondo mito è Argo Πανόπτης (colui che tutto vede), il dio dai molti occhi con il dono della visione a 360 gradi. Poiché Argo aveva cento occhi e poteva farne riposare alcuni, mentre teneva desti i rimanenti, egli era un assai abile guardiano, perciò fu incaricato dalla dea Era di fare la guardia alla vacca in cui Io (la sacerotessa di Era) era stata trasformata. Travestito da pastore, Hermes riuscì ad ottenere la fiducia di Argo e lo fece addormentare. Una volta che Argo fu addormentato, Hermes lo uccise. In ricordo di Argo, Era prese i suoi occhi e li dispose sulla coda del suo uccello preferito, il pavone.

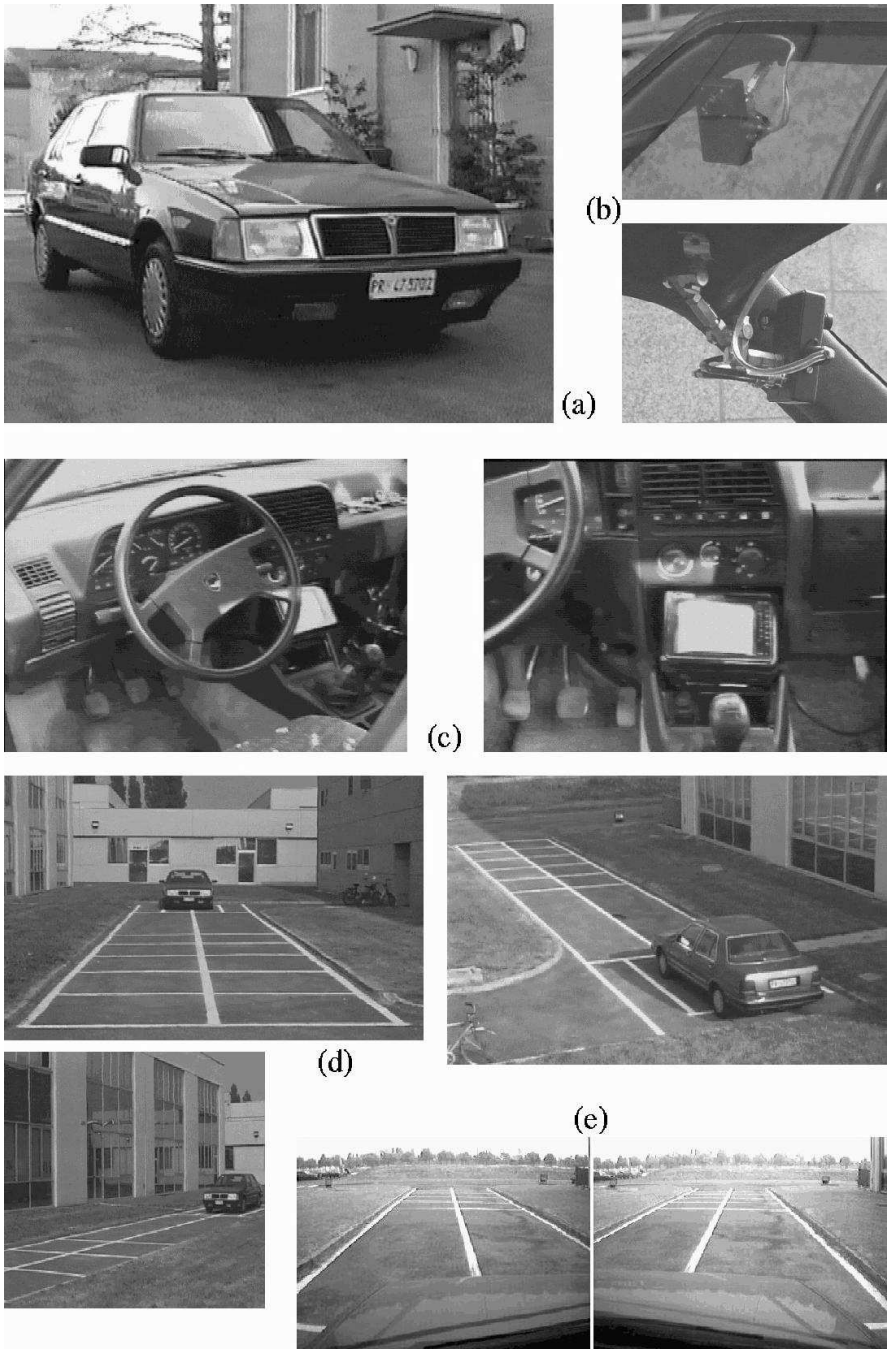

Figura 4: (a) il veicolo sperimentale ARGO, (b) il sistema di visione, (c) l'interno del veicolo in cui è visibile il monitor di output, (d) lo stradello attrezzato per la calibrazione ed (e) la coppia di immagini stereoscopiche acquisite da ARGO in tale stradello

Correntemente sul veicolo sperimentale ARGO il risultato dell'elaborazione è fornito al guidatore mediante un monitor, un pannello a led e mediante segnalazioni acustiche; l'intervento sulla sicurezza è perciò di primo livello. Tuttavia si prevede nell'immediato futuro di integrare un attuatore sullo sterzo per controllare in modo automatico la guida, intervenendo sulla sicurezza a livello più alto.

4.1 Sistema di Visione

Indipendentemente dal livello di sicurezza scelto, la funzionalità principale che viene richiesta è la determinazione delle condizioni dell'ambiente circostante. Sul veicolo sperimentale ARGO è installato un sistema di visione stereoscopico costituito da una coppia di telecamere in grado di acquisire immagini a 256 tonalità di grigio tra loro sincronizzate. Si è optato per telecamere su scheda a basso costo di dimensioni $3{,}2\ \text{cm} \times 3{,}2\ \text{cm}$, con una distanza focale di $6.0\ \text{mm}$, una risoluzione di 360 linee e sincronizzabili dall'esterno.

4.1.1 Configurazione del Sistema di Visione

Le telecamere sono poste all'interno dell'abitacolo, sulla parte più alta del parabrezza e sono posizionate alla massima distanza tra loro in modo tale da aumentare le differenze tra le coppie di immagini acquisite simultaneamente. Le telecamere sono orientate in modo tale da avere assi ottici paralleli e da inquadrare la scena anche al di sopra dell'orizzonte così da poter gestire anche le situazioni di strada non piana, anche se il fatto di inquadrare una porzione di cielo può rappresentare un problema per la luminosità dell'immagine: in condizioni di eccessivo contrasto il sensore può portare all'acquisizione di immagini eccessivamente saturate.

4.2 Sistema di Acquisizione

Ognuno dei due sistemi di elaborazione sperimentati a bordo del veicolo dispone di un dispositivo di acquisizione delle immagini.
Il sistema PAPRICA include una scheda frame-grabber piggy-back in grado di acquisire contemporaneamente da due telecamere sincronizzate una coppia di immagini di 512×512 pixel, che vengono memorizzate direttamente nella Memoria Immagine.
Il sistema commerciale basato su PC è invece dotato di una scheda PCI Matrox capace di acquisire tre immagini di 768x576 pixel contemporaneamente. Le immagini vengono memorizzate direttamente nella memoria principale del calcolatore grazie all'utilizzo del DMA. In entrambe le configurazioni l'acquisizione delle immagini avviene in tempo reale, a 25 frame per secondo nel caso di frame complete o a 50 field per secondo nel caso di acquisizione a singolo field.

4.3 Sistema di Elaborazione

Come si è detto al capitolo 3 sono attualmente in fase di valutazione le prestazioni di due differenti sistemi di elaborazione, entrambi sperimentati a bordo del veicolo ARGO: un sistema special-purpose massivamente parallelo, PAPRICA, ed un sistema commerciale standard consistente in un PC con processore Pentium 200 MMX.

4.4 Output

Il risultato dell'elaborazione è fornito al guidatore attraverso una serie di dispositivi di output installati a bordo del veicolo: un monitor, un pannello di controllo basato su led e un dispositivo per segnalazioni acustiche.
Il monitor può essere pilotato dalla scheda frame-grabber del sistema PAPRICA, che genera direttamente il segnale composito, come pure dal PC grazie ad una apposita scheda di conversione VGA/PAL. Sul monitor viene visualizzata l'immagine acquisita dalla

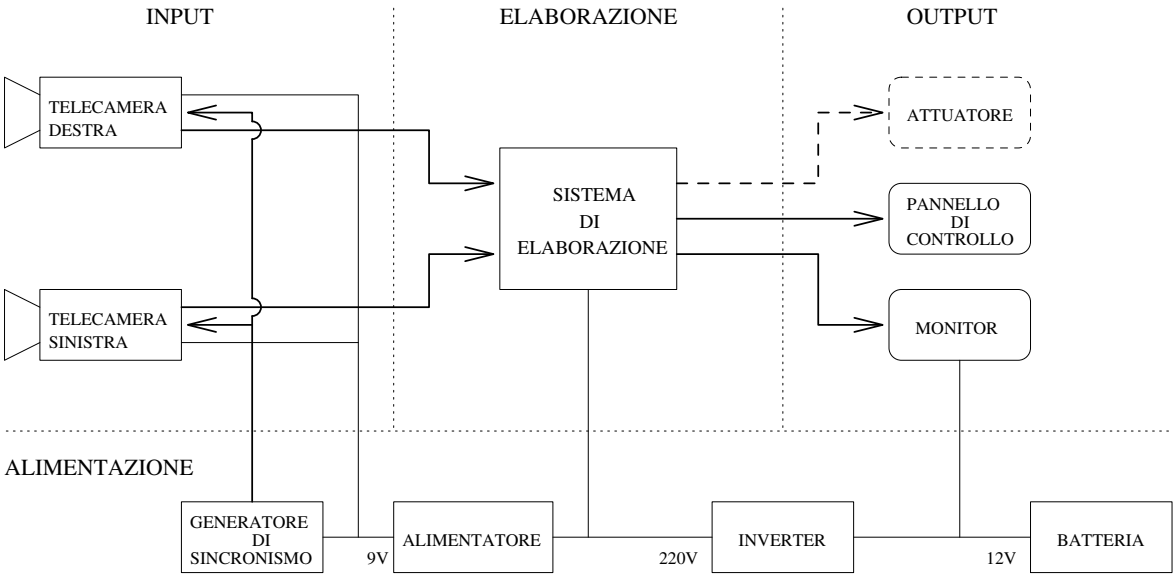

Figura 5: Schema dell'intero sistema di elaborazione installato su ARGO: sono visibili le parti di acquisizione, elaborazione ed output dei dati, nonché i dispositivi di alimentazione

telecamera sinistra a toni di grigio con sovrapposti i bordi strada in colore verde e tratti di colore rosso che codificano la posizione, la distanza e la dimensione degli ostacoli.
Il pannello di controllo a led, che come il monitor può essere pilotato da entrambi i sistemi di calcolo, è utilizzato esclusivamente per la funzionalità di inseguimento della corsia di

marcia: i led luminosi codificano la posizione del veicolo rispetto alla carreggiata e la rotazione da impartire allo sterzo.
È attualmente allo studio la possibilità di integrare un attuatore che permetta al sistema di intervenire sullo sterzo controllando la traiettoria del veicolo.

4.5 Ambiente di Calibrazione

Per la calibrazione del sistema di visione è stato predisposto uno stradello attrezzato con un'opportuna griglia di $4\text{m} \times 20\text{m}$ dipinta sul terreno. Una volta acquisite due viste stereoscopiche della griglia, considerando le coordinate nelle immagini di punti della griglia opportunamente scelti è possibile ricavare i parametri delle telecamere come l'angolo di apertura, l'orientamento degli assi ottici, l'altezza da terra delle telecamere e la loro distanza.
La calibrazione viene effettuata prima della partenza e impiega circa 30 secondi.

5. Conclusioni

In questo lavoro sono stati descritti i principali risultati ottenuti nel corso degli ultimi anni dall'Unità Operativa del Dipartimento di Ingegneria dell'Informazione dell'Università di Parma riguardo al progetto mirato all'incremento della sicurezza stradale.
Dopo alcuni anni di studio e valutazione degli approcci migliori, dei sistemi di elaborazione più efficienti, degli algoritmi più efficaci e dei sistemi più promettenti di interfaccia con il guidatore, è stato acquisito il veicolo sperimentale ARGO, una Lancia Thema 2000, su cui sono stati implementate le funzionalità di individuazione della corsia di marcia e localizzazione di ostacoli. Le particolari soluzioni adottate hanno reso il sistema molto robusto rispetto ai movimenti del veicolo (rollìo e beccheggio) e agli errori di calibrazione del sistema di visione. Durante i test dell'intero sistema si sono valutate anche le prestazioni delle due funzionalità. L'individuazione della corsia di marcia, che può essere effettuata solo su strade attrezzate (cioè con linee di demarcazione visibili), ha portato ad una percentuale di esatta individuazione in circa il 90 % dei casi, mentre la funzionalità di localizzazione di ostacoli, per il particolare modo con cui è stata implementata, è risultata essere ancora più robusta. Infatti, qualunque ostacolo viene localizzato con successo (gli unici problemi sono dati dalla determinazione della distanza di oggetti vicini o di oggetti comunque piccoli a grande distanza); a causa dei movimenti del veicolo, però, può succedere che la posizione reciproca delle telecamere non sia più nota; questo causa l'individuazione di un ostacolo anche quando questo non esiste.
In ogni caso, però, l'analisi di sequenze di immagini al posto di immagini singole, prevista per l'immediato futuro, porterà a diminuire il problema causato da falsi positivi; l'utilizzo della alta correlazione temporale tra le immagini di una sequenza consentirà anche di ottenere una maggior robustezza dei risultati per la determinazione della corsia di marcia.
Inoltre, è ora in fase di sperimentazione una estensione del sistema al fine di eliminare l'assunzione di strada piana: un ulteriore processo consentirà di modificare gradualmente il modello orizzontale della strada, prevedendo anche la possibilità che essa contenga parti

non piane come dossi, cunette, salite o discese. Questo porterà ovviamente ad un appesantimento dell'elaborazione, ma le tempistiche attuali (circa 10 immagini elaborate al secondo) permettono sicuramente di rallentare il processo allo scopo di incrementarne l'efficacia.
Infine è attualmente in fase di studio un attuatore da integrare sullo sterzo per dare al veicolo la possibilità di operare a livelli di sicurezza maggiori: in particolare per recuperare autonomamente in situazioni di pericolo e per raggiungere la guida completamente automatizzata.

Bibliografia

[1] M. Bertozzi and A. Broggi. Real-Time Lane and Obstacle Detection on the GOLD System. In I. Masaky, editor, *Proceedings IEEE Intelligent Vehicles'96*, pages 213—218, Tokyo, Japan, September 18-20 1996.

[2] M. Bertozzi and A. Broggi. An effective environment for fast prototyping of real-time applications on PAPRICA-3. *Journal of Systems Architecture*, 1997.

[3] M. Bertozzi and A. Broggi. GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. *IEEE Transactions on Image Processing*, 1997.

[4] M. Bertozzi and A. Broggi. Vision-based Vehicle Guidance. *IEEE Computer*, 30(7):49—55, July 1997.

[5] M. Bertozzi, A. Broggi, and S. Castelluccio. A Real-Time Oriented System for Vehicle Detection. *Journal of Systems Architecture*, 43(1-5):317—325, March 1997.

[6] M. Bertozzi, A. Broggi, D. Colla, and A. Fascioli. Sensing of Automotive Environments Using Stereo Vision. In *Proceedings ISATA - International Symposium on Automotive Technology and Automation*, Florence, Italy, 1997.

[7] M. Bertozzi, A. Broggi, and A. Fascioli. A Stereo Vision System for Real-Time Automotive Obstacle Detection. In *Proceedings ICIP - Third IEEE International Conference on Image Processing*, Lausanne, CH, September 16-19 1996.

[8] M. Bertozzi, A. Broggi, and A. Fascioli. Real-Time Obstacle Detection using Stereo Vision. In *Proceedings EUSIPCO-96 - VIII European Signal Processing Conference*, Trieste, Italy, September 10-13 1996.

[9] M. Bertozzi, A. Broggi, and A. Fascioli. Obstacle and Lane Detection on the ARGO Autonomous Vehicle. In *Proceedings IEEE Intelligent Transportation Systems Conference'97*, Boston, USA, November 1997. IEEE Computer Society.

[10] A. Broggi. Robust Real-Time Lane and Road Detection in Critical Shadow Conditions. In *Proceedings IEEE International Symposium on Computer Vision*, Coral Gables, Florida, November 19-21 1995.

[11] A. Broggi. The Evolution of a Massively Parallel Vision System for Real-Time Automotive Image Processing. In *Proceedings IEEE IPPS'96 - International Parallel Processing Symposium*, pages 724—728, Honolulu, HI, 1996. IEEE

Computer Society.

[12] A. Broggi. The PAPRICA-3 Programming and Simulation Environment. In *Proceedings MPCS-96 - IEEE International Conference on Massively Parallel Computing Systems*, Ischia, Italy, May 6-9 1996.

[13] A. Broggi. Global Communications on a Linear Array Architecture. *Journal of Parallel Algorithms and Applications*, 11(1-2):27—43, 1997.

[14] A. Broggi and S. Bertè. A Morphological Model-Driven Approach to Real-Time Road Boundary Detection for Vision-Based Automotive Systems. In *Proceedings Second IEEE Workshop on Applications of Computer Vision*, pages 73—90, Sarasota, FL, December, 5-7 1994.

[15] A. Broggi and S. Bertè. Vision-Based Road Detection in Automotive Systems: a Real-Time Expectation-Driven Approach. *Journal of Artificial Intelligence Research*, 3:325—348, December 1995.

[16] A. Broggi, G. Conte, G. Burzio, L. Lavagno, F. Gregoretti, C. Sansoè, and L. M. Reyneri. PAPRICA-3: A Real-Time Morphological Image Processor. In *Proceedings ICIP - First IEEE International Conference on Image Processing*, pages 654—658, Austin, TX, November 13-16 1994.

[17] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, R. Passerone, and L. M. Reyneri. Design and Implementation of the PAPRICA Parallel Architecture. *The Journal of VLSI Signal Processing*, 1997.

[18] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, and L. M. Reyneri. The PAPRICA Massively Parallel Processor. In *Proceedings MPCS - IEEE International Conference on Massively Parallel Computing Systems*, pages 16–30, Ischia, Italy, May 2-6 1994.

[19] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, and L. M. Reyneri. The Evolution of the PAPRICA System. *Integrated Computer-Aided Engineering Journal - Special Issue on Massively Parallel Computing*, 4(2):114–136, 1997.

[20] A. Broggi and F. Gregoretti. Editorial: Special-Issue on "Special-Purpose Architectures for Real-Time Imaging," A. Broggi and F. Gregoretti guest-editors. *Real-Time Imaging Journal*, 2(6):329—330, December 1996.

[21] A. Broggi, F. Gregoretti, and C. Sansoè. A Smart Sensor for the Real-Time Detection of Road Markings through the Removal of the Perspective Effect. In *Proceedings ISATA - International Symposium on Automotive Technology and Automation*, pages 425—432, Stuttgart, Germany, September, 18-22 1995.

[22] S. Carlsson and J. O. Eklundh. Object detection using model-based prediction and motion parallax. In *Proceedings European Conference on Computer Vision*, volume 427, pages 297—306, Antibes, France, April 1990. LNCS, Springer Verlag.

[23] J. Crisman and C. Thorpe. SCARF: A Color Vision System that Tracks Roads and Intersections. *IEEE Transactions on Robotics and Automation*, 9(1):49–58, February 1993.

[24] E. D. Dickmans and B. D. Mysliwetz. Recursive 3-D Road and Relative Ego-State

Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:199—213, May 1992.

[25] W. Enkelmann. Obstacle detection by evaluation of optical flow field from image sequences. In *Proceedings European Conference on Computer Vision*, volume 427, pages 134—138, Antibes, France, April 1990. LNCS, Springer Verlag.

[26] A. Giachetti, M. Campani, R. Sanni, and A. Succi. The Recovery of Optical Flow for Intelligent Cruise Control. In *Proceedings IEEE Intelligent Vehicles'94*, pages 91—95, Paris, October 1994.

[27] F. Gregoretti, R. Passerone, C. Sansoè, and A. Broggi. The PAPRICA-3 Parallel Processor. In *Proceedings MIPRO'96 Symposium*, Opatija, Croatia, May 20-24 1996.

[28] F. Gregoretti, L. M. Reyneri, C. Sansoè, A. Broggi, and G. Conte. The PAPRICA SIMD array: critical reviews and perspectives. In L. Dadda and B. Wah, editors, *Proceedings ASAP'93 - IEEE Computer Society International Conference on Application Specific Array Processors*, pages 309—320, Venezia, Italy, October 25-27 1993.

[29] K. Kluge and S. Lakshmanan. A Deformable Template Approach to Lane Detection. In I. Masaky, editor, *Proceedings IEEE Intelligent Vehicles'95*, pages 54—59, Detroit, September 25-26 1995.

[30] D. Koller, J. Malik, Q.-T. Luong, and J. Weber. An integrated stereo-based approach to automatic vehicle guidance. In *Proceedings of the Fifth ICCV*, pages 12—20, Boston, 1995.

[31] J. Kosecka, R. Bajcsy, and M. Mintz. Control of Visually Guided Behaviors. In C. Brown and D. Terzopoulos, editors, *Real-time Computer Vision*. Cambridge University Press, 1994.

[32] H. A. Mallot, H. H. Bülthoff, J. J. Little, and S. Bohrer. Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics*, 64:177—185, 1991.

[33] L. Matthies. Stereo vision for planetary rovers: stochastic modeling to near real-time implementation. *International Journal of Computer Vision*, 8:71—91, 1992.

[34] L. E. Mettala and O. Firschein. Reconnaissance, surveillance and target acquisition research for the unmanned ground vehicle program. In *Proceedings International Conference on Intelligent Autonomous Systems*, pages 257—279, 1993.

[35] M. Ohzora, T. Ozaki, S. Sasaki, M. Yoshida, and Y. Hiratsuka. Video-rate image processing system for an autonomous personal vehicle system. In *Proceedings IAPR Workshop on Machine Vision and Applications*, pages 389—392, Tokyo, 1990.

[36] D. Pomerlau and T. Jochem. Rapidly Adapting Machine Vision for Automated Vehicle Steering. *IEEE Expert*, 11(2), April 1996.

[37] D. A. Pomerleau. RALPH: Rapidly Adapting Lateral Position Handler. In I. Masaky, editor, *Proceedings IEEE Intelligent Vehicles'95*, pages 506—511, Detroit,

September 25-26 1995.

[38] H. Schneidermann and J. S. Albus. Progress and prospects for vision-based automatic driving. In *Proceedings 1993 annual meeting of IVHS AMERICA*, pages 569—575, 1993.

[39] J. Weber, D. Koller, Q.-T. Luong, and J. Malik. New results in stereo-based automatic vehicle guidance. In I. Masaky, editor, *Proceedings IEEE Intelligent Vehicles'95*, pages 530—535, Detroit, 1995.

[40] Y. Zheng, D. G. Jones, S. A. Billings, J. E. W. Mayhew, and J. P. Frisby. SWITCHER: a stereo algorithm for ground plane obstacle detection. *Image and Vision Computing*, 8:57—62, 1990.

[41] T. Zielke, M. Brauckmann, and W. von Seelen. Intensity and Edge-based Symmetry Detection with an Application to Car-Following. *CVGIP: Image Understanding*, 58:177—190, 1993.

[42] F. Thomanek, E. D. Dickmanns, and D. Dickmanns. Multiple Object Recognition and Scene Interpretation for Autonomous Road Vehicle Guidance. In *Proceedings IEEE Intelligent Vehicles'94*, pages 231–236, Paris, October 1994.