

Andrea Venturi
0000184706

Elaborato di Visione Artificiale
A.A. 2005/2006

SVOLGIMENTO DEL PROGETTO

Primo approccio

I passi base su cui ho fondato la strategia sono:

1. Scomposizione dell'immagine nelle componenti rosso e blu che compongono i cartelli stradali.
2. Eliminazione del rumore.
3. Etichettatura delle componenti connesse nello spazio rosso e blu.
4. Calcolo ed estrazione delle caratteristiche delle varie forme geometriche.
5. Classificazione dei cartelli stradali fissando delle regole ed analizzando i dati ottenuti dal punto precedente.

Secondo approccio

1. Segmentazione

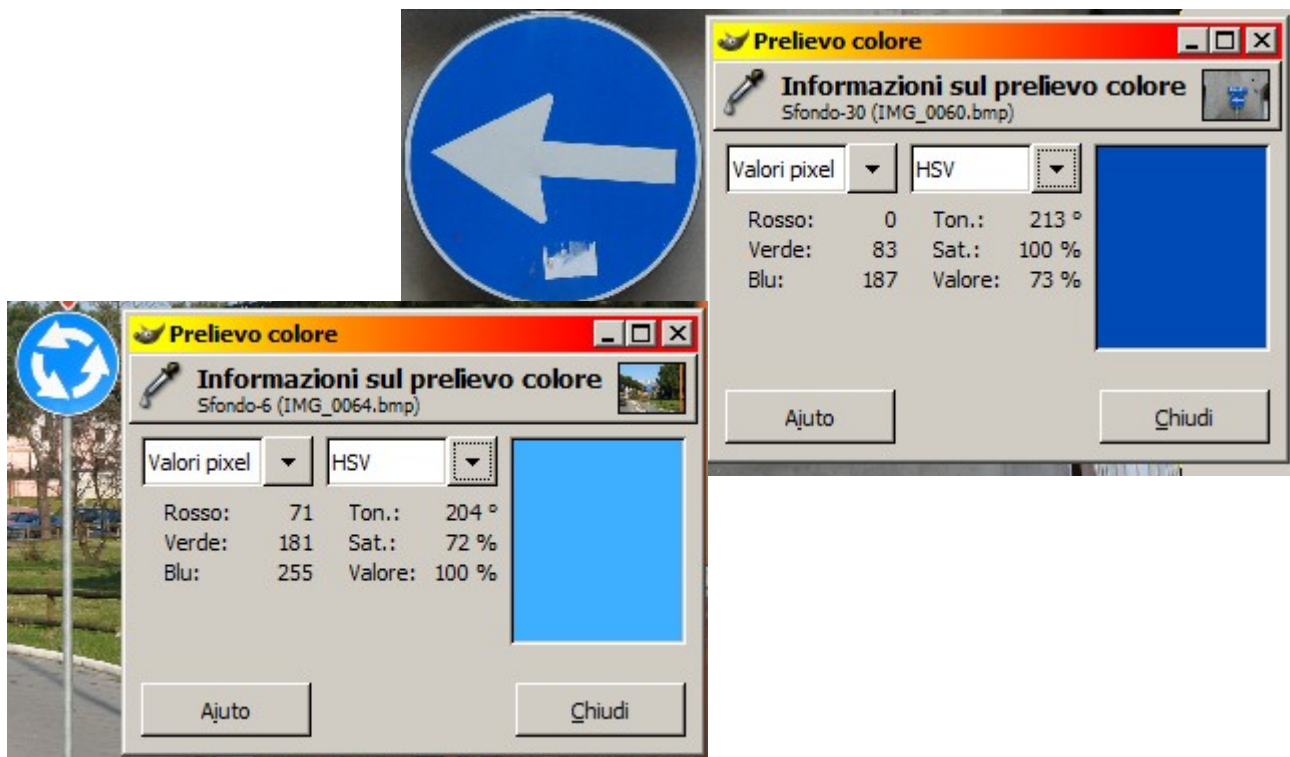
Lo scopo del primo passo è suddividere l'immagine nelle due componenti che interessano, ovvero il rosso e il blu, per poi eliminare i pixel che hanno un colore non riconducibile a quello di un cartello.

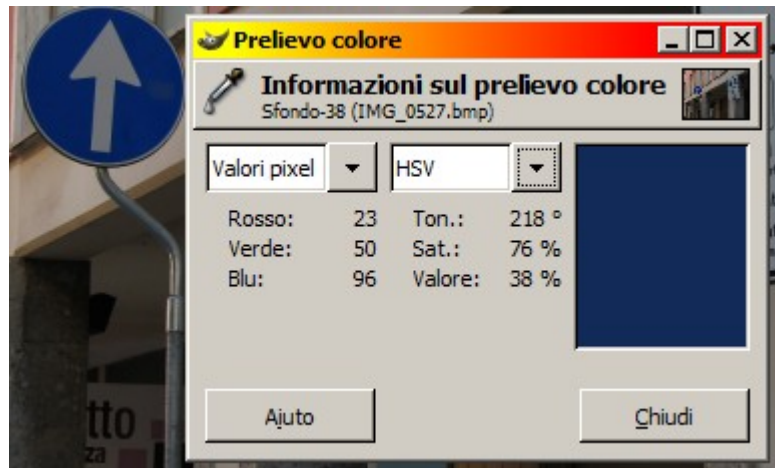
Come è noto, ogni colore è ottenuto dalla miscelazione dei colori base, quindi è necessario definire un range di valori per cui un pixel può essere ammissibile.

Lo standard informatico è RGB e anche gli strumenti messi a disposizione permettono di ricavare solo queste informazioni.

Tuttavia l'utilizzo di questo spazio, seppur ottimo per la generazione dei colori, non è altrettanto buono per catalogarli così come li percepisce l'occhio umano. Fortunatamente esistono altri spazi più adatti allo scopo come ad esempio il modello HSI.

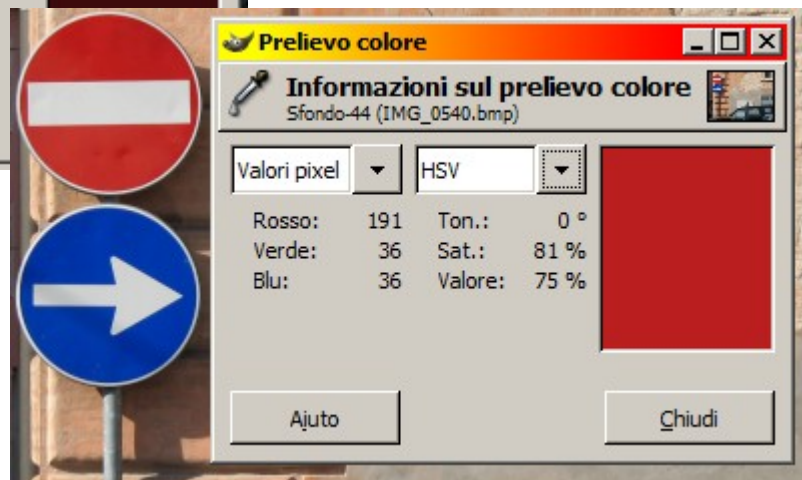
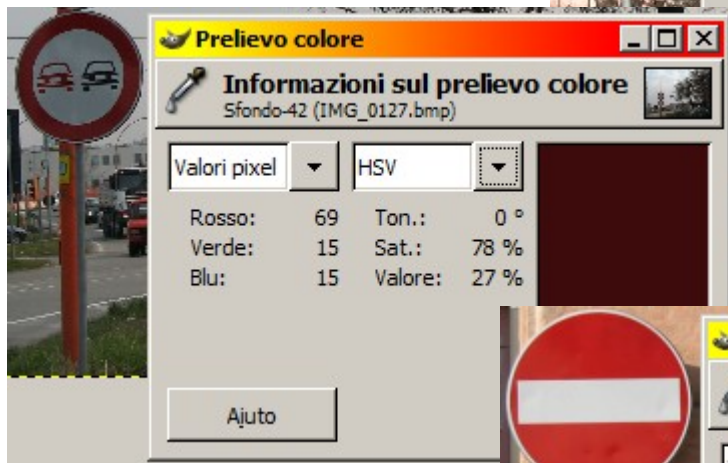
Alcuni esempi pratici:





Come si può subito notare, i valori delle componenti RGB hanno ben poco in comune nei tre casi, mentre per quanto riguarda HSI, il tono e la saturazione si possono inserire in un range ristretto.

La stessa situazione si verifica per il rosso:



Quindi un possibile algoritmo per la segmentazione del colore sarà così:

Data in input un'immagine a colori:

estrai le componenti R,G,B in 3 vettori;

per ogni pixel dell'immagine:

calcola le componenti H (tono) ed S (saturazione);

se H ed S rispettano le soglie per il rosso:

metti 1 nella corrispondente posizione nel vettore R;

metti 0 nella corrispondente posizione nel vettore B;

se H ed S rispettano le soglie per il blu:

metti 1 nella corrispondente posizione nel vettore B;

metti 0 nella corrispondente posizione nel vettore R;

altrimenti

metti 0 nella corrispondente posizione nel vettore B;

metti 0 nella corrispondente posizione nel vettore R;

azzerare il vettore G;

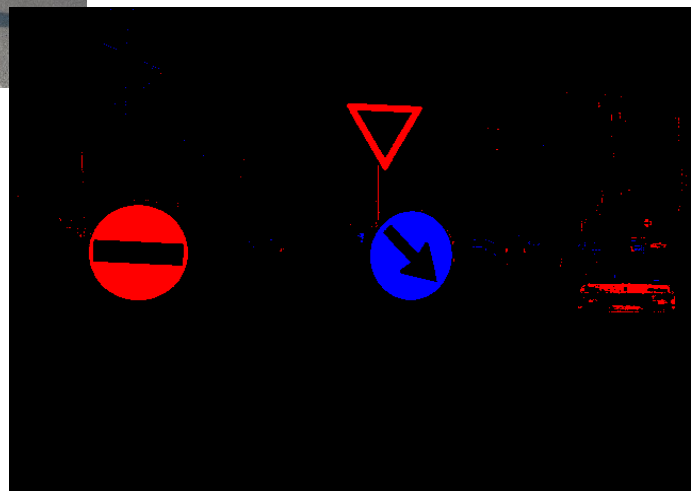
crea una nuova immagine;

setta come componenti della nuova immagine i vettori R,G,B;

restituisce l'immagine;

Fine.

Esempio:



2. Eliminazione del rumore

A prima vista può sembrare un problema minore, che si possa risolvere con

tecniche non troppo elaborate, se si pensa al rumore come qualche pixel o gruppo di pixel sparso qua e là nell'immagine.

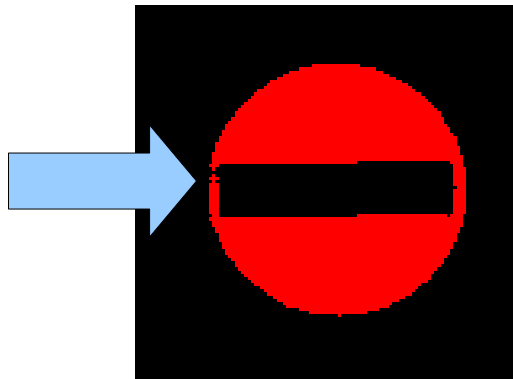
Però affrontando un ampio numero di casi ci si può presto accorgere che non è elementare la soluzione.

In fase di progettazione la prima tecnica sperimentata è stato l'utilizzo degli operatori morfologici, che hanno fornito dei risultati discutibili.

Effettuando una Opening con un elemento strutturante di questo tipo

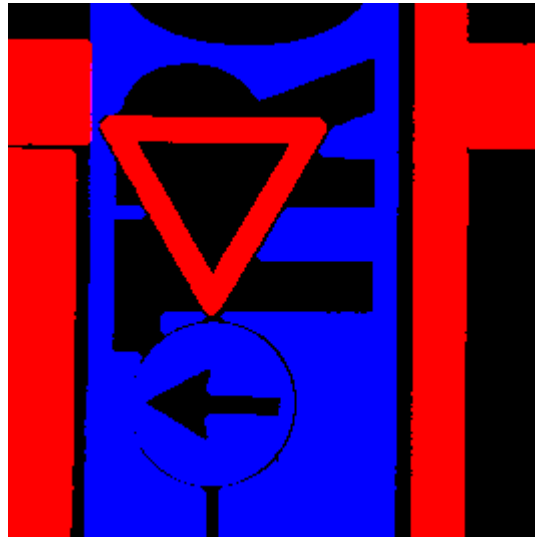
```
0 1 0
1 1 1
0 1 0
```

quindi di forma vagamente circolare e dimensioni minime si possono ottenere dei problemi con i cartelli di piccole dimensioni oppure con quelli di controsenso:



Si può notare la discontinuità nella forma.

Si può correre ai ripari effettuando una Dilation, ma per un problema che si risolve, uno nuovo se ne crea:



In questo caso il cartello si salda con lo sfondo.

Altre alternative potrebbero essere:

- usare l'Erosion come prima operazione, ma anche questo porterebbe a una eliminazione di parti piccole ma essenziali;
- la Closing potrebbe unire erroneamente il cartello con lo sfondo;
- aumentare la dimensione dell'elemento strutturante peggiorerebbe ulteriormente le cose.

La soluzione prevede l'eliminazione del rumore con una tecnica completamente differente e verrà attuata nella fase di etichettamento delle componenti connesse.

3. Etichettatura delle componenti connesse

Questa operazione è di fondamentale importanza.

Data in input l'immagine segmentata, ricavata dal primo punto, e il colore delle componenti che si vogliono etichettare (R o B), con una prima scansione diretta (quindi da sinistra a destra, dal basso all'alto) si andranno ad analizzare i seguenti pixel:

P3	P	
P2	P1	P0

Se nessuno dei 4 è etichettato, significa che P è il primo pixel di una nuova regione, quindi gli si assegna una nuova etichetta, ad esempio un numero incrementale.

Invece se due pixel hanno etichette diverse, significa che P è il punto di collegamento, tra due aree, quindi in un altro vettore si annota l'equivalenza tra le due etichette, e al pixel P si assegna, ad esempio, l'etichetta col valore maggiore.

Ultimo caso, se un solo pixel ha etichetta, oppure due o più hanno etichetta uguale, P assumerà la stessa etichetta.

Nel secondo passo, si andrà ad eseguire la fusione delle etichette equivalenti.

Con una prima scansione si scorre il vettore delle equivalenze, per ogni equivalenza, si mantiene il valore di etichetta maggiore.

La seconda scansione serve a sostituire, nel vettore delle etichette, le equivalenze, secondo quanto riportato nel rispettivo array.

A questo punto, interviene la fase di eliminazione del rumore, come preannunciato.

Per ogni etichetta presente, si esegue un conteggio del numero di pixel con quella etichetta e se questo valore è inferiore a una certa soglia, quella zona viene eliminata.

E' chiaro che questa tecnica non garantisce la totale affidabilità, ma durante i test è risultata migliore rispetto agli operatori di morfologia.

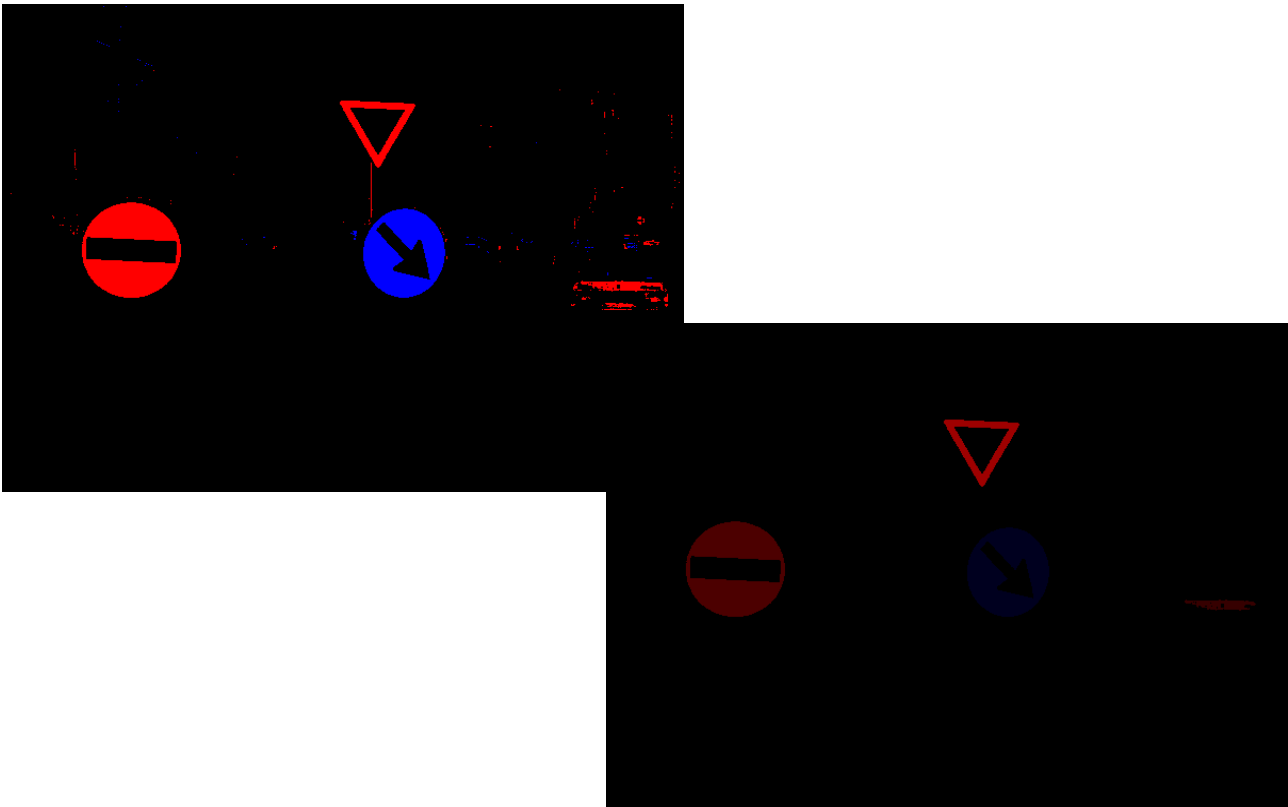
Impostare la soglia richiede una certa attenzione, perché è necessario valutare quale margine di errore si vuole mantenere.

Tenere una soglia troppo alta significa rischiare di eliminare i cartelli in secondo piano o comunque sulla media distanza; nel caso opposto, ovvero con una soglia bassa, si consente a molte zone di rumore di giungere alla fase di estrazione delle caratteristiche, con un conseguente decremento delle prestazioni.

La soglia utilizzata nel progetto filtra via i cartelli posti a una distanza superiore agli 8-10 metri in immagini con risoluzione 640x480 (inoltre dipende dal tipo:

un obbligo e un controsenso hanno maggiore area di un divieto o una precedenza).

Alcuni risultati:



4. Estrazione delle caratteristiche

Le caratteristiche ritenute discriminanti e che vengono estratte sono: colore, area, perimetro, centro dell'oggetto, raggio, alcune distanze del contorno dal centro, il rapporto tra l'area della figura e l'area del rettangolo in cui è inscritta, l'area sotto e sopra al centro.

Significato e metodologia di estrazione delle misure (data in input l'immagine ottenuta al passo precedente):

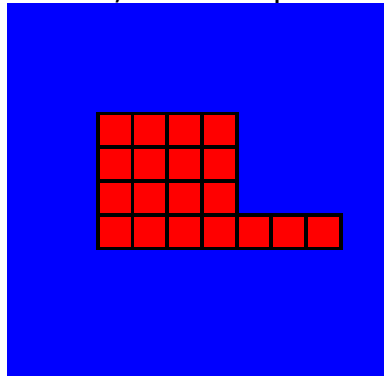
- con area si intende il numero di pixel che costituiscono la figura, quindi tramite una scansione dell'immagine, dentro un vettore avente tante posizioni quante possono essere le tonalità in scala di grigi (255) si contano i pixel per ogni tonalità.
- il perimetro è il numero di pixel componenti il bordo; si parte con una scansione, quando viene trovato un pixel diverso dal background si inizia a calcolare il perimetro della figura trovata in questo modo: si procede in senso antiorario e il pixel successivo si cerca in base allo spostamento immediatamente precedente, ovvero:

3	2	1
4	P	0
5	6	7

Se si proviene da destra, il prossimo pixel si cerca dalla posizione 6 e in senso antiorario fino alla 5; se si proviene da alto-destra si parte dalla posizione 7, se si proviene da sinistra dalla posizione 2 e così via. Quindi lo spostamento avviene con metrica D8.

Quando si esamina una figura, i pixel sono considerati di bordo se: ancora non sono stati visitati (se ne terrà traccia in un vettore), sono dello stesso colore della figura che si sta processando, hanno un pixel di background vicino con metrica D4.

Durante lo scorrimento del contorno si può verificare un particolare evento spiacevole a cui è stato posto un rimedio su misura: è il caso in cui una figura abbia una forma insolita, ad esempio:



in questo caso, si incorrerebbe nell'errore di incastrarsi nell'appendice di destra, perché non si troverebbe più un successore.

La soluzione prevede, ad ogni spostamento, il salvataggio delle coordinate del pixel di provenienza e della posizione da cui cercare il pixel successivo.

Quindi, quando si incorre in questo caso, non avendo trovato un pixel del contorno tra quelli adiacenti e nemmeno avendo concluso il calcolo del perimetro (il punto di partenza non è adiacente), si torna sui propri passi, finché non si trova un nuovo pixel del contorno.

Quando si termina il calcolo del perimetro di una figura, si ripristinano le coordinate da cui si era partiti, e si continua la scansione dell'immagine finché non si trova un pixel di un colore non ancora incontrato.

- il centro dell'oggetto viene calcolato come differenza delle coordinate minima e massima sui due assi. Questi quattro valori si ricavano durante la scansione per il calcolo del perimetro:

$$X_{\text{centro}} = X_{\text{min}} + (X_{\text{max}} - X_{\text{min}}) / 2$$

$$Y_{\text{centro}} = Y_{\text{min}} + (Y_{\text{max}} - Y_{\text{min}}) / 2$$

- il raggio è calcolato come media tra le semidifferenze delle coordinate minima e massima sui due assi:

$$\text{Raggio} = (X_{\text{max}} - X_{\text{min}}) + (Y_{\text{max}} - Y_{\text{min}}) / 4$$

- le distanze del contorno dal centro sono dei campionamenti effettuati sul bordo della figura. Utilizzando l'algoritmo per il calcolo del perimetro modificato, ovvero mantenendo la struttura che permette di costeggiare l'oggetto e scegliendo un intervallo di campionamento, si calcolano le distanze dal centro ad ogni intervallo.
- il calcolo dell'area situata sotto e sopra il centro della figura sarà utile per determinare il verso di un triangolo.
- il rapporto tra l'area della figura e l'area del rettangolo in cui è inscritta, è

un altro indicatore sulla forma dell'oggetto. L'area del rettangolo viene calcolata come il prodotto tra le differenze delle coordinate minima e massima sui due assi:

$$R = \text{area} / ((X_{\max} - X_{\min}) * (Y_{\max} - Y_{\min}))$$

5. Classificazione

Si è quindi giunti alla fase finale. Il compito svolto da questo passo è di classificare i cartelli stradali, secondo opportune regole e fissando delle soglie.

- La prima caratteristica discriminante, senza alcun dubbio, è il colore.
- Come secondo si è considerato il rapporto tra il quadrato del perimetro e la sua area: questo rapporto ha valori caratteristici a seconda dei casi:
 - divieto generico
 - divieto di sosta
 - controsenso
 - stop
 - precedenza e pericolo
 - obbligo con una freccia
 - obbligo con due frecce
 - rotonda
- controllo sulla rotondità: si confronta il raggio con le distanze dal centro nei punti di campionamento. Se tutti i valori sono uguali, a meno di una tolleranza, significa che la figura è circolare.
- nel caso in cui si sia stabilito che la figura è triangolare, è necessario conoscerne l'orientamento. Se l'area sottostante il centro è maggiore di quella sovrastante significa che la base del triangolo è in basso e viceversa.
- il rapporto tra l'area della figura e l'area del rettangolo in cui è inscritta assume valori specifici per i casi sopraelencati. E' stato utilizzato per riconoscere i triangoli e gli ottagoni. Questo valore è valido anche per gli altri segnali, ma non è stato utilizzato per non rendere troppo selettivo il riconoscimento.

Ecco i valori rilevati sul Training Set, su cui sono state scelte le soglie:

DIVIETO GENERICO			
IMMAGINE	AREA	PERIMETRO	P ² /A
IMG_0082	3572	284	22,58
IMG_0085	1482	183	22,60
IMG_0117	5426	350	22,58
IMG_0125	315	83	21,87
IMG_0127	1317	175	23,25
IMG_0515	17161	620	22,40
IMG_0523	6030	379	23,82
IMG_0524	5271	344	22,45
IMG_0538	2751	258	24,20
IMG_0548	8924	448	22,49

DIVIETO DI SOSTA			
IMMAGINE	AREA	PERIMETRO	P ² /A
IMG_0082	4279	288	19,38
IMG_0111	3466	259	19,35
IMG_0118	2483	212	18,10
IMG_0489	1126	147	19,19
IMG_0491	4798	305	19,39
IMG_0493	5852	298	15,17
IMG_0497	2219	205	18,94
IMG_0498	4737	298	18,75
IMG_0528	1675	179	19,13
IMG_0528	1476	168	19,12
IMG_0538	829	129	20,07

CONTROSENSO				
IMMAGINE	AREA	PERIMETRO	P ² /A	Rapporto
IMG_0067	3930	230	13,46	0,61
IMG_0077	5108	265	13,75	0,60
IMG_0078	7645	326	13,90	0,58
IMG_0088	3589	225	14,11	0,58
IMG_0089	747	107	15,33	0,59
IMG_0102	1913	163	13,89	0,58
IMG_0106	7198	319	14,14	0,57
IMG_0540	4402	249	14,08	0,57
IMG_0541	5431	277	14,13	0,57

PERICOLO				
IMMAGINE	AREA	PERIMETRO	P ² /A	Rapporto
IMG_0530	4310	351	28,58	0,30
IMG_0537	1433	210	30,77	0,28

PRECEDENZA				
IMMAGINE	AREA	PERIMETRO	P ² /A	Rapporto
IMG_0064	1228	186	28,17	0,30
IMG_0066	1765	226	28,94	0,30
IMG_0071	471	116	28,57	0,30
IMG_0077	1326	197	29,27	0,29
IMG_0091	3264	302	27,94	0,31
IMG_0511	8249	477	27,58	0,30



STOP				
IMMAGINE	AREA	PERIMETRO	P ² /A	Rapporto
IMG_0502	2006	159	12,60	0,64

OBBLIGO con una freccia			
IMMAGINE	AREA	PERIMETRO	P ² /A
IMG_0060	12527	401	12,84
IMG_0071	3557	212	12,64
IMG_0077	4154	230	12,73
IMG_0091	4054	225	12,49
IMG_0094	2529	182	13,10
IMG_0097	5050	252	12,58
IMG_0102	2421	174	12,51
IMG_0502	2057	161	12,60
IMG_0511	9367	341	12,41
IMG_0524	9081	348	13,34
IMG_0527	3639	217	12,94
IMG_0540	4792	245	12,53

OBBLIGO rotonda			
IMMAGINE	AREA	PERIMETRO	P ² /A
IMG_0064	1274	139	15,17

OBBLIGO pedoni			
IMMAGINE	AREA	PERIMETRO	P ² /A
IMG_0495	7975	322	13,00

RISULTATI DEI TEST








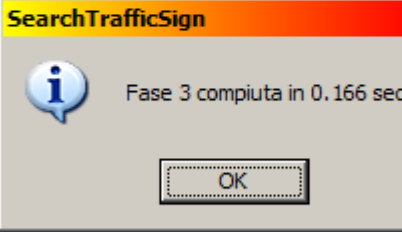

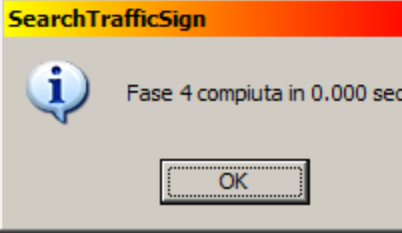


SearchTrafficSign	SearchTrafficSign
 Confronti completati in 14.774 s	 Confronti completati in 13.213 s
Immagini:	Immagini:
IMG_0060.bmp: THR =1.000 FHR=0.000	IMG_0061.bmp: THR =1.000 FHR=0.000
IMG_0064.bmp: THR =1.000 FHR=0.000	IMG_0065.bmp: THR =1.000 FHR=0.000
IMG_0066.bmp: THR =1.000 FHR=0.000	IMG_0068.bmp: THR =0.500 FHR=0.000
IMG_0067.bmp: THR =1.000 FHR=0.000	IMG_0070.bmp: THR =0.000 FHR=0.000
IMG_0071.bmp: THR =1.000 FHR=0.000	IMG_0074.bmp: THR =1.000 FHR=0.000
IMG_0077.bmp: THR =1.000 FHR=0.000	IMG_0098.bmp: THR =1.000 FHR=0.000
IMG_0078.bmp: THR =1.000 FHR=0.000	IMG_0103.bmp: THR =1.000 FHR=0.000
IMG_0082.bmp: THR =1.000 FHR=0.000	IMG_0105.bmp: THR =1.000 FHR=0.000
IMG_0085.bmp: THR =1.000 FHR=0.000	IMG_0107.bmp: THR =1.000 FHR=0.000
IMG_0088.bmp: THR =1.000 FHR=0.000	IMG_0108.bmp: THR =1.000 FHR=0.000
IMG_0089.bmp: THR =1.000 FHR=0.000	IMG_0110.bmp: THR =0.667 FHR=0.000
IMG_0091.bmp: THR =1.000 FHR=0.000	IMG_0112.bmp: THR =1.000 FHR=0.000
IMG_0094.bmp: THR =1.000 FHR=0.000	IMG_0113.bmp: THR =0.500 FHR=0.000
IMG_0097.bmp: THR =1.000 FHR=0.000	IMG_0119.bmp: THR =0.500 FHR=0.000
IMG_0102.bmp: THR =1.000 FHR=0.000	IMG_0121.bmp: THR =1.000 FHR=0.000
IMG_0106.bmp: THR =1.000 FHR=0.000	IMG_0122.bmp: THR =1.000 FHR=0.000
IMG_0111.bmp: THR =1.000 FHR=0.000	IMG_0126.bmp: THR =1.000 FHR=0.000
IMG_0117.bmp: THR =1.000 FHR=0.000	IMG_0494.bmp: THR =0.500 FHR=0.000
IMG_0118.bmp: THR =1.000 FHR=0.000	IMG_0501.bmp: THR =1.000 FHR=0.000
IMG_0125.bmp: THR =1.000 FHR=0.000	IMG_0505.bmp: THR =1.000 FHR=0.000
IMG_0127.bmp: THR =1.000 FHR=0.000	IMG_0507.bmp: THR =1.000 FHR=0.000
IMG_0489.bmp: THR =1.000 FHR=0.000	IMG_0512.bmp: THR =1.000 FHR=0.000
IMG_0491.bmp: THR =1.000 FHR=0.000	IMG_0516.bmp: THR =0.000 FHR=0.000
IMG_0493.bmp: THR =0.000 FHR=0.000	IMG_0518.bmp: THR =1.000 FHR=0.000
IMG_0495.bmp: THR =1.000 FHR=0.000	IMG_0520.bmp: THR =1.000 FHR=0.000
IMG_0497.bmp: THR =1.000 FHR=0.000	IMG_0522.bmp: THR =1.000 FHR=0.000
IMG_0498.bmp: THR =1.000 FHR=0.000	IMG_0526.bmp: THR =1.000 FHR=0.000
IMG_0502.bmp: THR =1.000 FHR=0.000	IMG_0529.bmp: THR =1.000 FHR=0.000
IMG_0511.bmp: THR =1.000 FHR=0.000	IMG_0531.bmp: THR =1.000 FHR=0.000
IMG_0515.bmp: THR =1.000 FHR=0.000	IMG_0532.bmp: THR =1.000 FHR=0.000
IMG_0523.bmp: THR =1.000 FHR=0.000	IMG_0534.bmp: THR =1.000 FHR=0.000
IMG_0524.bmp: THR =1.000 FHR=0.000	IMG_0535.bmp: THR =0.667 FHR=0.000
IMG_0527.bmp: THR =1.000 FHR=0.000	IMG_0539.bmp: THR =1.000 FHR=0.000
IMG_0528.bmp: THR =1.000 FHR=0.000	IMG_0542.bmp: THR =1.000 FHR=0.000
IMG_0530.bmp: THR =1.000 FHR=0.000	IMG_0545.bmp: THR =1.000 FHR=0.000
IMG_0537.bmp: THR =1.000 FHR=0.000	IMG_0549.bmp: THR =1.000 FHR=0.000
IMG_0538.bmp: THR =1.000 FHR=0.000	
IMG_0540.bmp: THR =1.000 FHR=0.000	Media
IMG_0541.bmp: THR =1.000 FHR=0.000	
IMG_0548.bmp: THR =1.000 FHR=0.000	Tempo = 0.367 THR=0.870 FHR=0.000
Media	<input type="button" value="OK"/>
Tempo = 0.369 THR=0.975 FHR=0.000	
<input type="button" value="OK"/>	

COMMENTI SUI TEMPI DI CALCOLO

I test sono stati effettuati su un notebook con queste caratteristiche: processore AMD Athlon 64 bit 3200 (frequenza 2000Mhz) con 1 Gb di RAM DDR, sistema operativo Windows XP Home Edition Service Pack 2.

Come si può osservare dai risultati riportati sopra, l'elaborazione di una immagine in formato bitmap (l'unico formato consentito), con risoluzione 640x480, impiega tra gli 0.3 e 0.4 secondi. I tempi subiscono notevoli variazioni utilizzando immagini con altre risoluzioni.

Ecco alcuni tempi impiegati dalle varie fasi del programma:

le tre immagini hanno diverse caratteristiche:

- la prima ha un solo cartello, uno sfondo di facile eliminazione e pochi disturbi;
- la seconda ha tre cartelli, uno sfondo di facile eliminazione e qualche disturbo;
- la terza ha due cartelli, ma uno sfondo con elementi di colori simili a quelli dei segnali e di grandi dimensioni.

La fase 1, ovvero la segmentazione, è indipendente dal contenuto dell'immagine, se non lo si era intuito dallo pseudo codice inserito precedentemente, balza subito all'occhio osservando i tempi.

La fase 2, cioè l'etichettamento delle componenti connesse e l'eliminazione del rumore, dipende largamente dal contenuto. Ciò che influenza l'elaborazione è la proporzione tra background e foreground. Più esteso è lo sfondo, più veloce sarà questa fase (lo sfondo non viene etichettato).

La fase 3, ossia l'estrazione delle caratteristiche, dipende fortemente dal contenuto. Ogni forma (sia cartelli che rumore non eliminato) che giunge a questa fase, verrà analizzata per il rilevamento delle feature, quindi si avrà un aumento del tempo di calcolo all'aumentare degli elementi.

La fase 4, ovvero la classificazione, seppur dipendente dal numero di figure rilevate, è velocissima, in quanto si riduce al confronto tra i parametri rilevati e le soglie per il riconoscimento.

COMMENTI E MIGLIORAMENTI

Come si può notare dai risultati ottenuti si ha un'altissima percentuale di cartelli riconosciuti sia nel Training che nel Test Set.

Da ulteriori prove condotte su immagini nuove, si è osservato che questa percentuale tende a diminuire, comunque mantenendosi su livelli più che discreti.

In ogni caso il punto di forza del progetto è il bassissimo tasso di falsi positivi.

Senza dubbio non è stato fatto il massimo possibile, sia per ragioni di tempo che di conoscenze possedute.

Alcuni miglioramenti potrebbero essere:

- nel calcolo del perimetro, quando si procede in diagonale, sommare $\sqrt{2}$ anziché 1.
- Utilizzare una tecnica di eliminazione del rumore meno "brutale".
- Creare un metodo per rigenerare i contorni di un oggetto (ad esempio: quando sul contorno della forma si presenta una concavità che potrebbe essere dovuta a una imprecisa segmentazione del colore, riempirla prima di calcolare il perimetro).
- Utilizzare dei filtri per evidenziare meglio i contorni.
- Migliorare la segmentazione del colore. Renderla adattativa alle condizioni di luce.
- Estrarre delle feature ancora più discriminati.

Ritengo la difficoltà di questo elaborato medio-alta se confrontata al livello medio degli elaborati di altri Corsi.

Nonostante abbia affrontato "in solitaria" il progetto, non mi sono mai trovato in vicoli ciechi, seppur abbia incontrato ostacoli impegnativi.

L'impegno speso si può quantificare in circa 15 giorni con una media di 6 ore al giorno. Date le dimensioni del progetto credo che anche in coppia i tempi totali si sarebbero accorciati di poco, la sostanziale differenza si sarebbe trovata nella qualità.

Probabilmente i limiti della soluzione presentata sono dovuti al fatto che è stata sfornata da una sola mente, quindi può presentare parti discutibili.

Senza dubbio è l'elaborato che mi ha dato più soddisfazioni.