



Università degli Studi di Padova
Facoltà di Ingegneria
Corso di Laurea in Ingegneria dell'Informazione



MISURE DI CONFIDENZA PER ALGORITMI DI STEREO VISION

Relatore:
Prof. Pietro Zanuttigh
Correlatore:
Giulio Marin

Laureando:
Nicola Dal Lago

Anno Accademico: 2013/2014



Misure di Confidenza per Algoritmi di Stereo Vision

Visione stereo

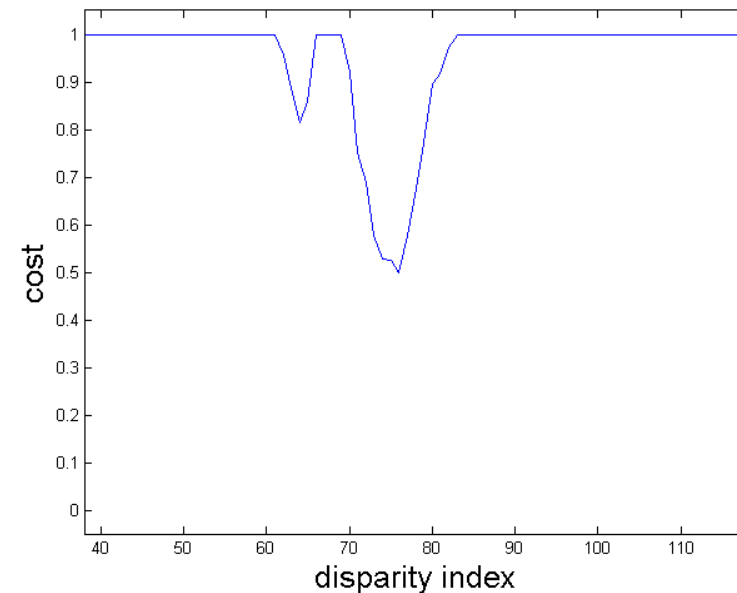
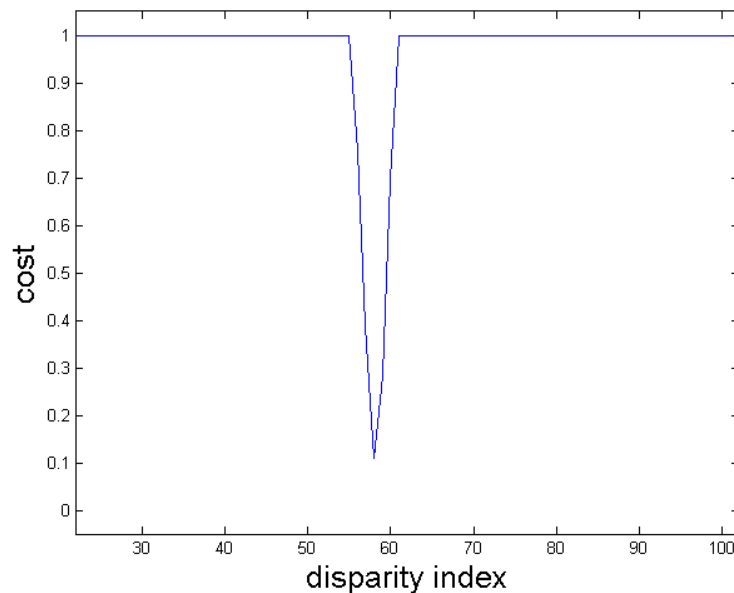


A partire dalle immagini delle due videocamere è possibile definire, per ogni pixel, una disparità. Questa disparità indica la distanza tra i due *punti coniugati* che sono proiezione dello stesso punto della scena. Da questa si calcola la *mappa di disparità*. Per facilitare l'algoritmo le immagini vengono prima *rettificate*, in modo che due punti coniugati si trovino nella stessa *retta epipolare*.



Funzione costo

A causa di molti fattori che possono influire negativamente nel corretto funzionamento dell'algoritmo per il calcolo della disparità (illuminazione, mancanza di trama, occlusioni, eccetera) la disparità per ciascun pixel può contenere errori. Si dà quindi a ciascun pixel un costo, che identifica quanto una determinata disparità sia corretta. L'algoritmo ne prende il minimo, ma non è detto sia sempre corretto.





Università degli studi di Padova

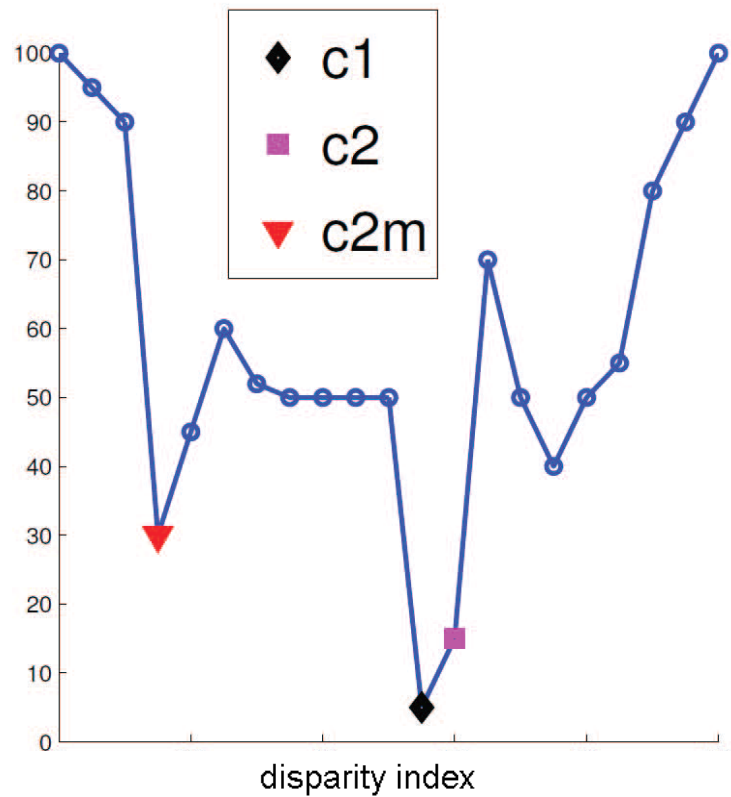
Tesi di Laurea

Misure di Confidenza per Algoritmi di Stereo Vision

Misure di confidenza

Vista la possibilità di errore, si rende necessario lo studio di varie tecniche per misurare la confidenza con la quale viene assegnato un determinato valore di disparità. Le misure che verranno utilizzate, sfrutteranno le caratteristiche locali e globali della curva, infine se ne combineranno alcune di esse.

I risultati verranno confrontati con la vera disparità per quella determinata scena, questa disparità, detta *ground truth*, viene fornita insieme alle immagini nei dataset presenti in <http://vision.middlebury.edu/stereo/data/>.



◆ c1: minimo della funzione costo;

■ c2: secondo minimo della funzione costo;

▼ c2m: secondo minimo locale più piccolo.

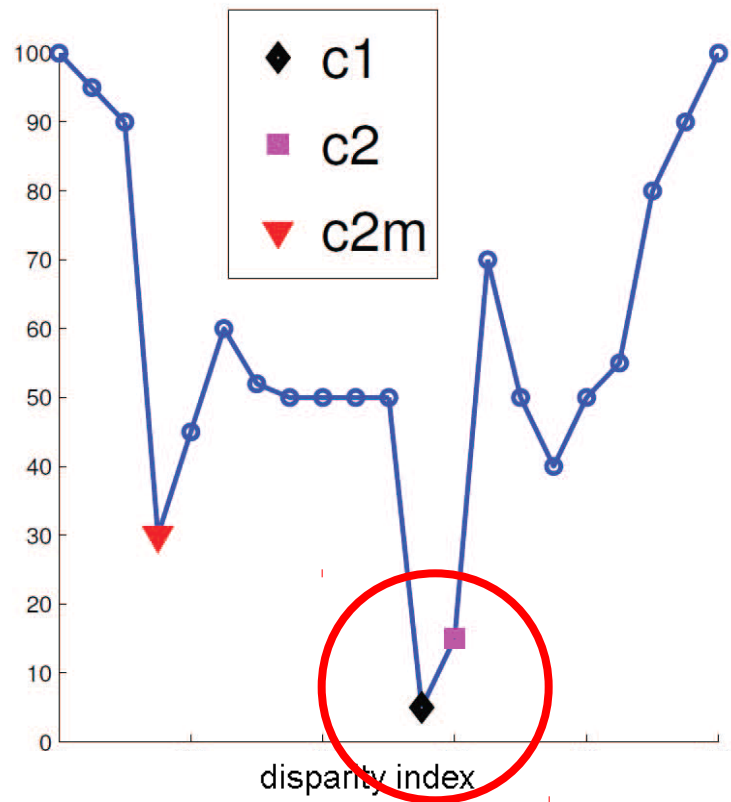


Università degli studi di Padova

Tesi di Laurea

Misure di Confidenza per Algoritmi di Stereo Vision

Proprietà locali della curva:



Curvature:

$$C_{CUR} = \frac{-2c(d_1) + c(d_1 - 1) + c(d_1 + 1)}{2}$$

Local Curve:

$$C_{LC} = \frac{\max\{c(d_1 - 1), c(d_1 + 1)\} - c_1}{\gamma}$$

In questo genere di metriche, si sfrutta il fatto che la forma della curva di costo intorno al minimo è indice di certezza nel confronto.

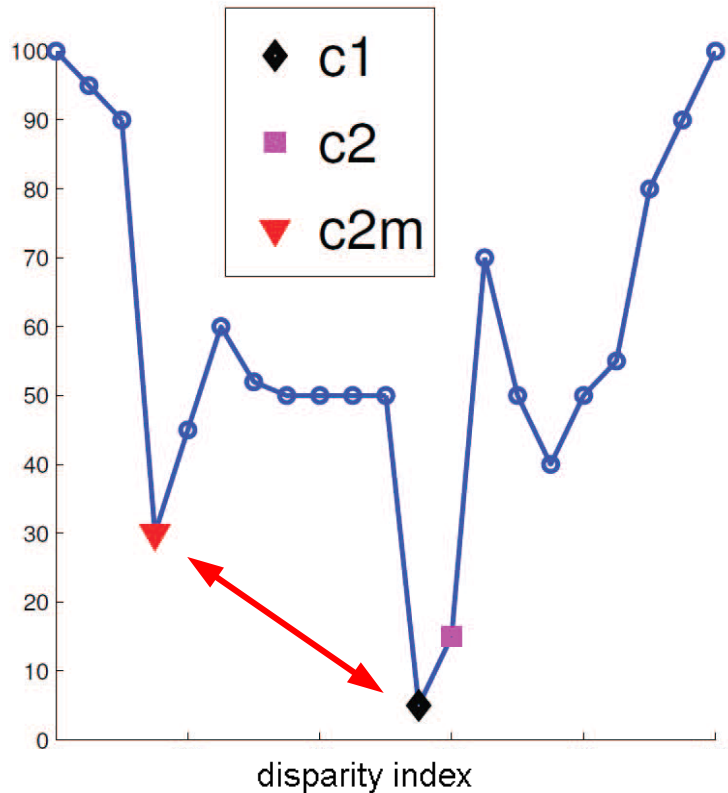


Università degli studi di Padova

Tesi di Laurea

Misure di Confidenza per Algoritmi di Stereo Vision

Minimo locale della curva:



Peak Ratio Naive: $C_{PKRN} = \frac{c_2 + \epsilon}{c_1 + \epsilon} - 1$

Maximum Margin: $C_{MMN} = c_2 - c_1$

Nonlinear Margin: $C_{NLM} = e^{\frac{c_2 - c_1}{2\sigma_{NLM}^2}} - 1$

Si basa sul concetto che la presenza di altri candidati è un'indicazione di incertezza, mentre la loro assenza di affidabilità.

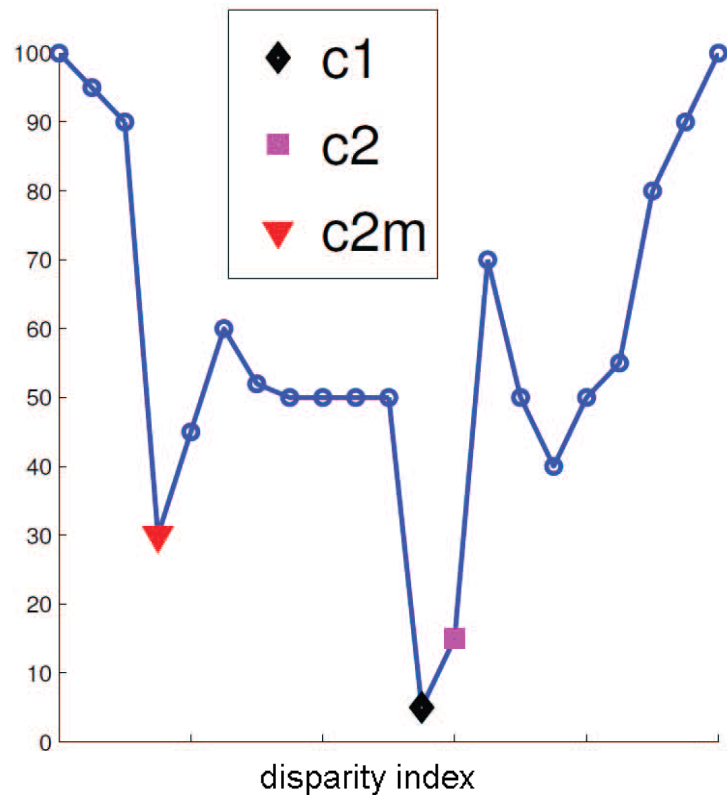


Università degli studi di Padova

Tesi di Laurea

Misure di Confidenza per Algoritmi di Stereo Vision

Intera curva:



Maximum Likelihood: $C_{MLM} = \frac{e^{-\frac{c_1}{2\sigma_{MLM}^2}}}{\sum_d e^{-\frac{c(d)}{2\sigma_{MLM}^2}}}$

Attainable Maximum Likelihood:

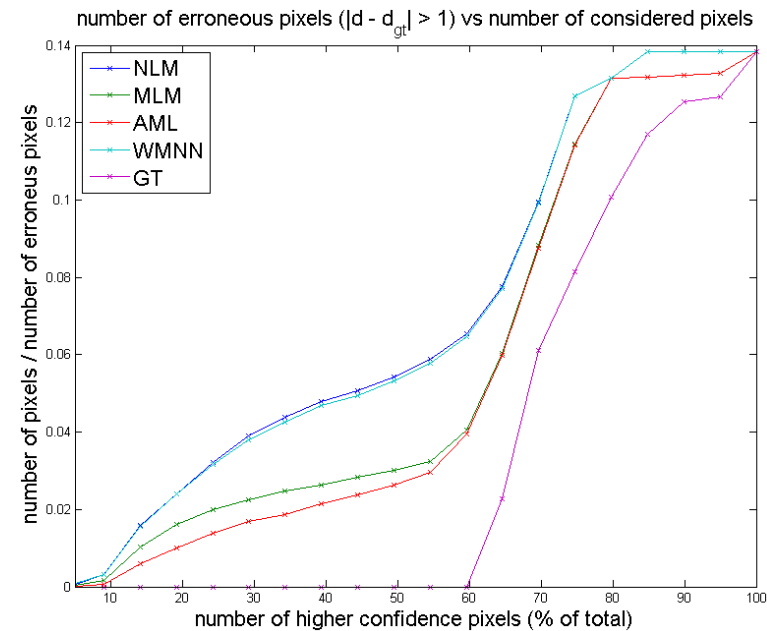
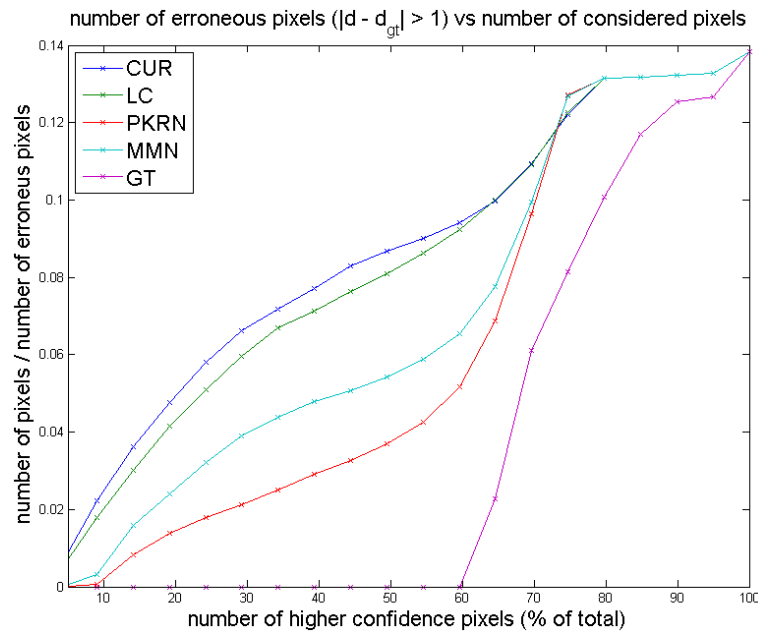
$$C_{AML} = \frac{1}{\sum_d e^{-\frac{(c(d)-c_1)^2}{2\sigma_{AML}^2}}}$$

Winner Margin Naive: $C_{WMNN} = \frac{c_2 - c_1}{\sum_d c(d)}$

Questi metodi convertono la funzione costo in una distribuzione di probabilità sulla disparità.



Risultati



Le tre misure che risultano essere le migliori sono, nell'ordine: AML, MLM, PKRN. Infatti le loro curve d'errore ordinate rispetto la confidenza, sono quelle che si avvicinano più di tutte alla curva ottenuta semplicemente ordinando l'errore in ordine crescente.



Misure di Confidenza per Algoritmi di Stereo Vision

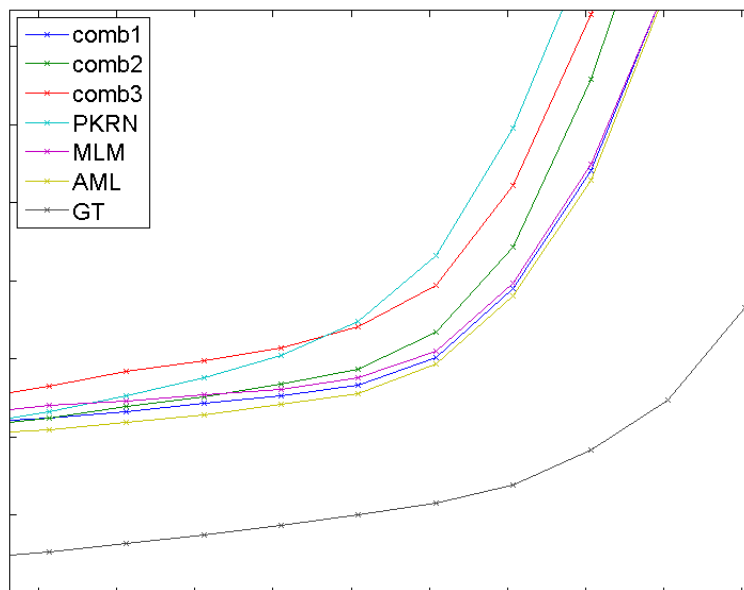
Combinazioni di misure

Considerando le metriche come probabilità e presupponendo che siano indipendenti, è possibile calcolarne una combinazione semplicemente moltiplicandole fra di loro. Le combinazioni scelte sono:

$C_{comb1} = C_{AML} \cdot C_{MLM}$ ovvero la combinazione delle due migliori;

$C_{comb2} = C_{AML} \cdot C_{MLM} \cdot C_{PKRN}$ le tre migliori;

$C_{comb3} = C_{NLM} \cdot C_{MLM}$ NLM ha dimostrato di essere una buona misura su alcuni dataset.



La migliore combinazione si è dimostrata essere la prima, riuscendo a migliorare MLM.



Conclusioni e sviluppi futuri

- La metrica singola migliore è **Attainable Maximun Likelihood**
 - La migliore combinazione è il prodotto tra **AML** e **MLM**
-

Questi risultati possono essere utilizzati per altri lavori, in quanto ci forniscono una indicazione di dove e di quanto l'algoritmo sbaglia:

- possono ottimizzare l'algoritmo stesso che calcola la disparità, in quanto è possibile stabilire in anticipo dove andrà a sbagliare e sfruttare l'informazione nei metodi di propagazione globale;
- essendo la visione stereo una tecnica non esente da errori, è possibile fonderne i risultati con altri sensori, ad esempio *tempo di volo*, e usare questa fusione per migliorare la precisione.