

Marchiatura digitale di sequenze video stereoscopiche a disparità coerente

BENEDETTA BARBETTI,
MICHAELA SERVI

Relatori:
Alessandro Piva
Carlo Colombo

Correlatori:
Pasquale Ferrara
Francesca Uccheddu

Tesi Magistrale di Ingegneria Informatica
Università degli Studi di Firenze

10 Dicembre 2015

CONTESTO

Numerose applicazioni di elaborazione di immagini e video richiedono esplicite informazioni sulla **profondità** della scena. La **stereoscopia** permette di ottenere queste informazioni.

Campi applicativi

- Medicina
- Robotica
- Tracking
- Industria manifatturiera
- Cinema



STEREOSCOPIA

Tecnica di realizzazione e visione di immagini e filmati, atta a trasmettere una illusione di **tridimensionalità**



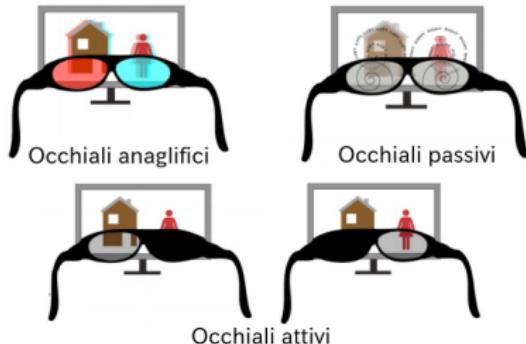
VIDEO STEREOSCOPICI

Il video stereoscopico è ottenuto inquadrando la stessa scena da
due **punti di vista diversi**

DISPOSITIVI DI RIPRESA E VISUALIZZAZIONE

Sistema di ripresa stereo

- Due telecamere sincronizzate
- Correttamente allineate
- Stessi parametri



Sistema di riproduzione

- **Attivo:** lenti sincronizzate con il televisore
- **Passivo:** lenti diversamente polarizzate
- **Anaglifico:** lenti passive con filtri di colore diverso

NECESSITÀ DI UNA PROTEZIONE

- Autenticazione
- Copyright
- Individuazione di falsificazioni e alterazioni



WATERMARKING

Inserimento di informazione aggiuntiva in contenuti multimediali in modo trasparente e robusto



nuove tecniche adatte alla diversità della coppia stereo

OBIETTIVO

Implementazione di un **sistema di marchiatura a disparità coerente** per video stereoscopici

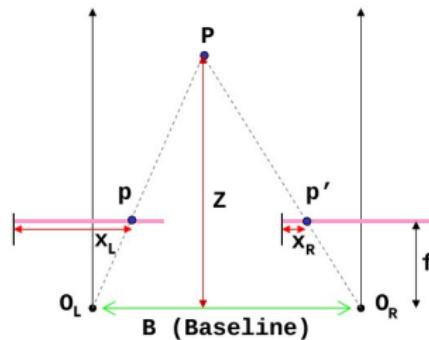
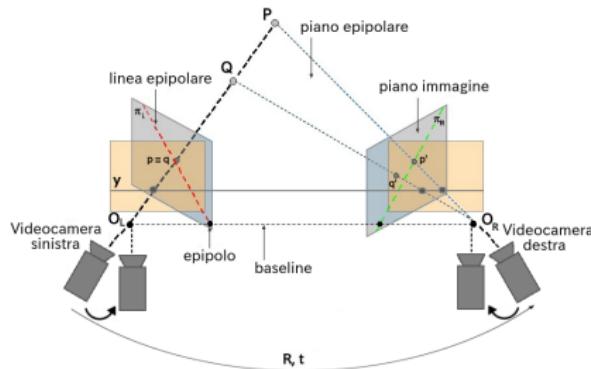
Due approcci:

- nel dominio spaziale
(Stato dell'Arte)
- nel dominio della frequenza

a disparità coerente

un punto fisico della scena deve
portare lo stesso campione di
watermark in entrambe le viste

BACKGROUND

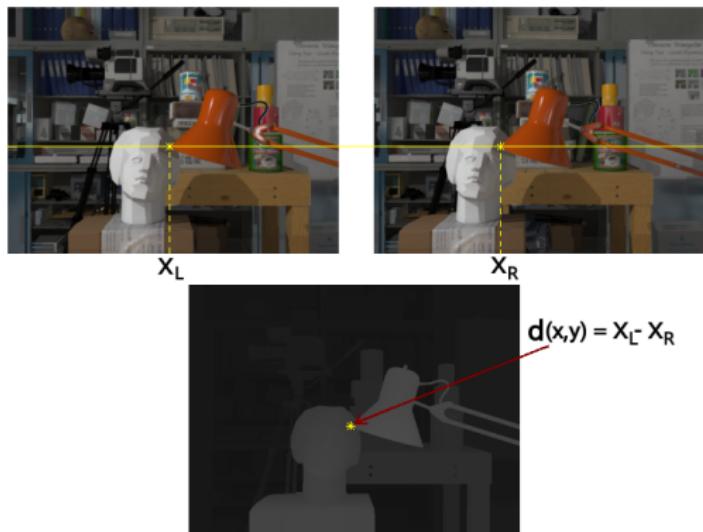


1. Calibrazione parametri intrisici ed estrinseci
2. Rettificazione
3. Calcolo delle corrispondenze
4. Computazione mappa di disparità

- Triangolazione:
- $$\frac{B}{Z} = \frac{(B+x_L)-x_R}{Z-f},$$
- $$Z = \frac{B \cdot f}{x_L - x_R} = \frac{B \cdot f}{d}$$
- $d = x_L - x_R$ è la **disparità**

MAPPA DI DISPARITÀ: CODIFICA

- Codificata come un'immagine in scala di grigi
- Punti più vicini alla telecamera sono più chiari e corrispondono a una disparità maggiore



MAPPA DI DISPARITÀ: COMPUTAZIONE

- **Metodi locali:** calcolare un valore di similarità (MSE, NCC..) all'interno di una finestra
- **Metodi globali:** minimizzare su tutta l'immagine una funzione di energia che racchiude le assunzioni di corrispondenza



In questa tesi è stato utilizzato il metodo globale di Kolmogorov and Zabih **Graph Cuts Stereo Matching Algorithm**

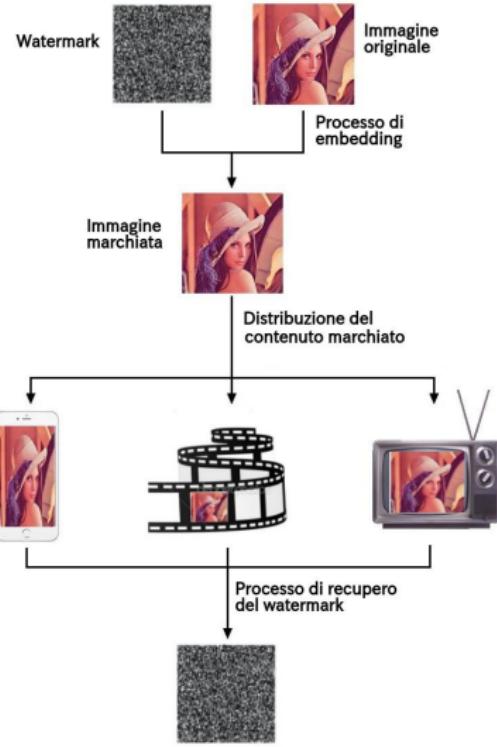
WATERMARKING

Processo di marchiatura

- Codifica di informazione nascosta in un contenuto originale
- Distribuzione del contenuto
- Ritrovamento del marchio

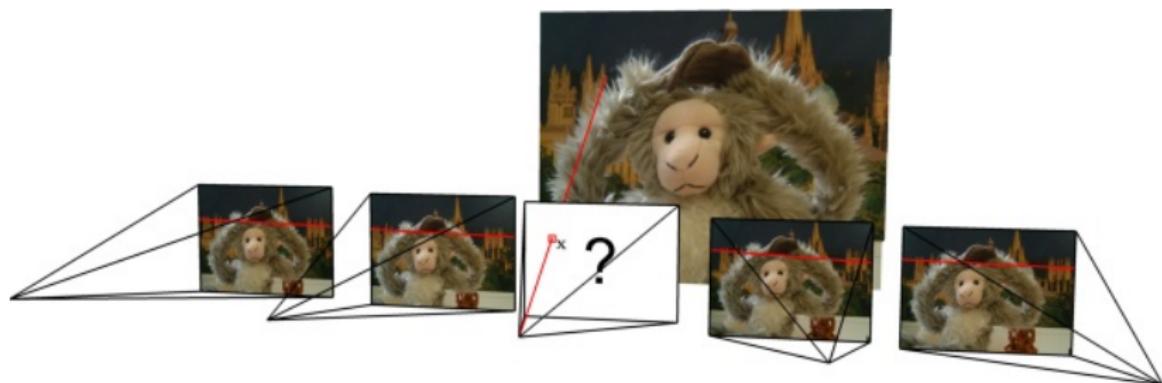
Proprietà

- Trasparenza
- Robustezza (compressione, view synthesis, distribuzione web)



VIEW SYNTHESIS

Dato un insieme di immagini della stessa scena ottenute da diversi punti di vista, una **nuova immagine** viene **creata** considerando una **camera virtuale** posizionata in un diverso punto dello spazio



METODI A DISPARITÀ COERENTE

- Marchiatura della **coppia stereo**
- **Punti corrispondenti** nelle due viste presentano lo **stesso** campione del **marchio**
- Il watermark viene modificato in base alla disparità prima di essere inserito nella vista destra
- Vantaggi:
 - Migliore **qualità visiva**
 - Maggior grado di **robustezza** contro attacchi di **view synthesis**

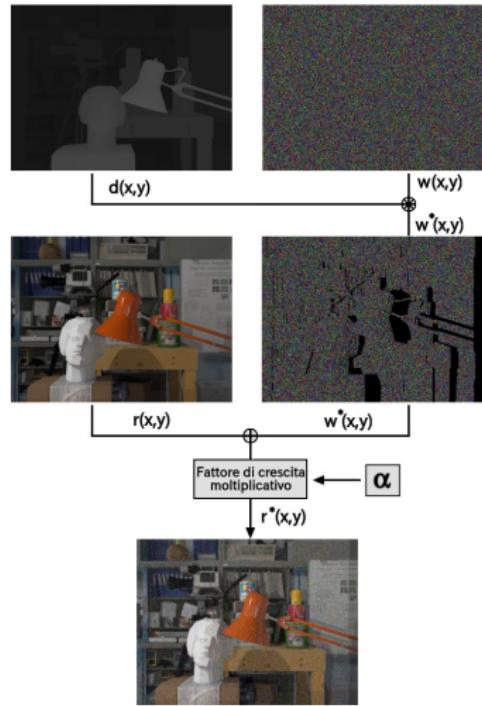
MARCHIATURA SPAZIALE A DISPARITÀ COERENTE

Codifica vista sinistra

- Metodo Spread Spectrum(**SS**):
 $I^w = I + \alpha w_K$
- Il marchio w_K segue $N(0, 1)$,
 $\alpha \in \{1, 3\}$

Codifica vista destra

- w_K deformato in base ai valori della disparità:
 $w_K^*(x, y) = w_K(x + d(x, y), y)$
- Inserito con SS: $r^w = r + \alpha w_K^*$



MARCHIATURA SPAZIALE A DISPARITÀ COERENTE

Decodifica vista sinistra

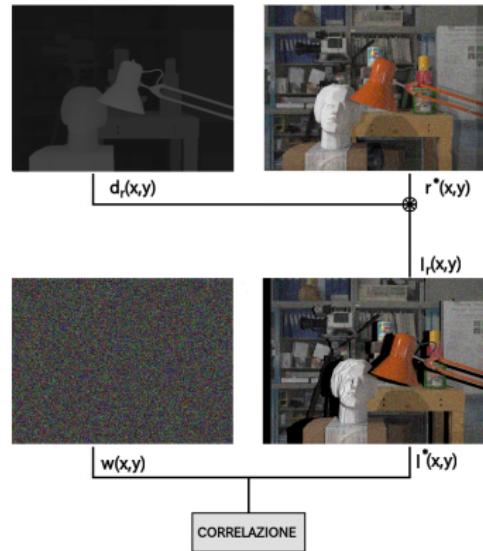
- Correlazione **vista sinistra e marchio originale**

Decodifica vista destra

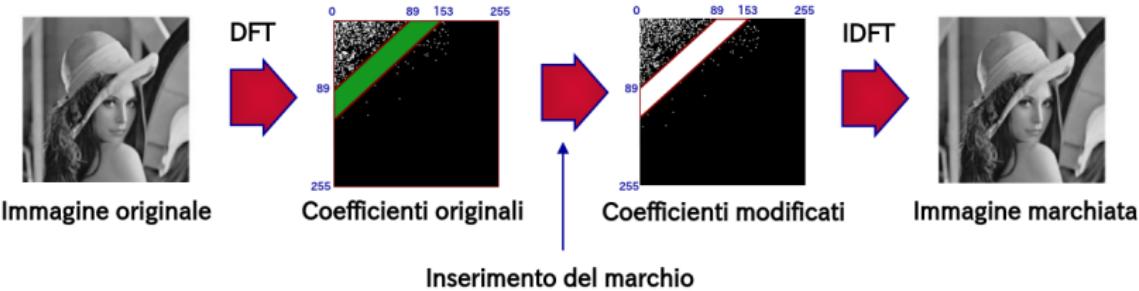
1. Correlazione **vista destra e marchio distorto**

oppure

2. Correlazione **vista sinistra ricostruita** (attraverso vista destra e disparità) e marchio originale



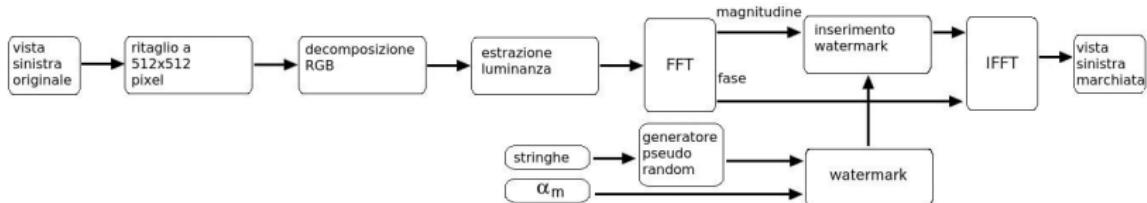
ALGORITMO DI MARCHIATURA NEL DOMINIO DELLA FREQUENZA



Vantaggi:

- Visibilità
- Invariante a traslazioni
- Resistente a degradazioni

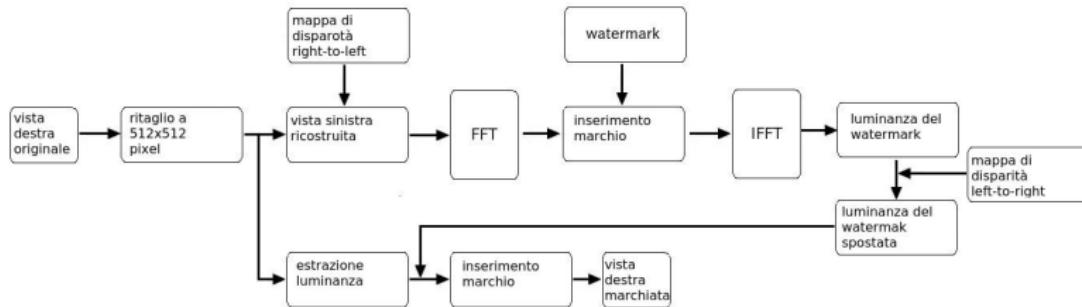
MARCHIATURA VISTA SINISTRA



Inserimento marchio:

- estrazione luminanza di I
- estrazione coefficienti a media frequenza
- alterazione coefficienti: $v'_{l_i} = v_{l_i} + \alpha w_i v_{l_i}$

MARCHIATURA VISTA DESTRA



Obiettivi:

- Pixel corrispondenti marchiati nello stesso modo;
- Marchio inserito spazialmente ma rilevabile in frequenza.

MARCHIATURA VISTA DESTRA

Idea:



Se consideriamo l'immagine \mathbf{I} scritta in funzione della sua trasformata \mathbf{L} :

$$\mathbf{I} = \frac{1}{MN} \sum \sum (|\mathbf{L}(u, v)|) \exp\{\phi(u, v)\} \exp\{-j2\pi(\frac{ux}{M} + \frac{vy}{N})\}$$

Dopo il processo di marchiatura:

$$\mathbf{I}_w = \frac{1}{MN} \sum \sum (|\mathbf{L}(u, v)| + \alpha |\mathbf{L}(u, v)| |w|) \exp\{j(\phi_L + \phi_w)\} \exp\{-j2\pi(\frac{ux}{M} + \frac{vy}{N})\}$$

MARCHIATURA VISTA DESTRA

Problemi:

- La procedura di detection si aspetta in ingresso un'immagine del tipo:

$$\mathbf{r}_w = \mathbf{r} + \frac{1}{MN} \sum \sum (\alpha |R(u, v)| |w| \exp\{j(\phi_R + \phi_w)\}) \exp\{-j2\pi(\frac{ux}{M} + \frac{vy}{N})\}$$

- L'alterazione estratta dalla vista sinistra è del tipo:

$$\alpha |L| |w| \exp\{j(\phi_L + \phi_w)\}$$

Soluzione:

- **modulo marchio:** modulo della vista sinistra ricostruita a partire dai pixel della vista destra;
- **fase marchio:** fase della vista sinistra sommata a quella del marchio di riferimento;

MARCHIATURA VISTA DESTRA

Il marchio inserito nella vista destra è quindi:

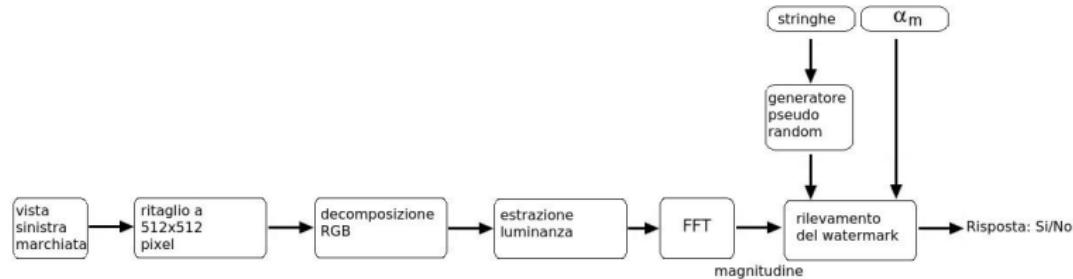
$$\alpha |R^{**}| |w| \exp\{j(\phi_L + \phi_w)\}$$

Le formule complete sono:

$$l_w = l + \frac{1}{MN} \sum \sum (\alpha |L(u, v)| |w| \exp\{j(\phi_L + \phi_w)\}) \exp\{-j2\pi(\frac{ux}{M} + \frac{vy}{N})\}$$

$$r_w = r + \frac{1}{MN} \sum \sum (\alpha |R(u, v)|^* |w| e^{\{j(\phi_L + \phi_w)\}*} e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})})$$

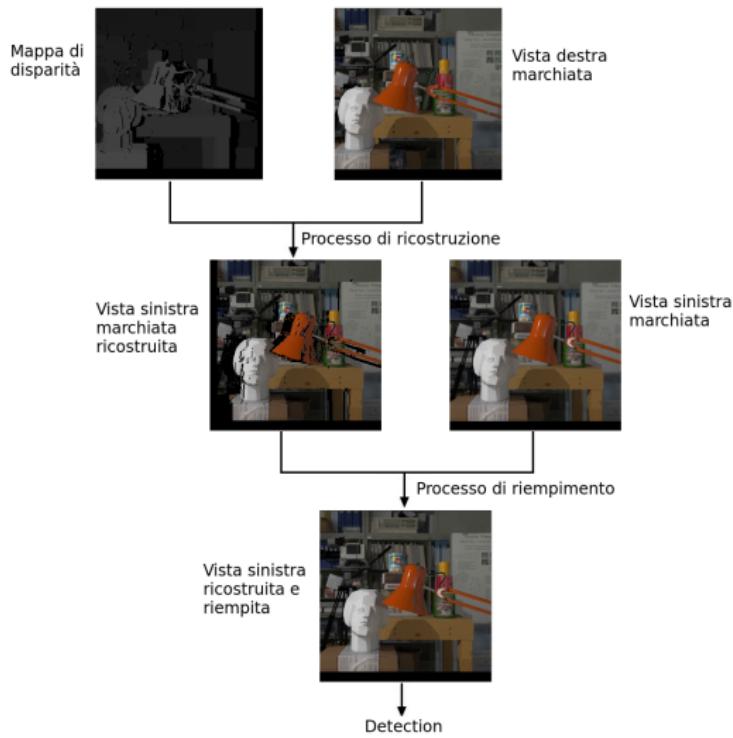
DETECTION VISTA SINISTRA



Detection basata su soglia:

- generazione marchio;
- calcolo likelihood $L(y)$ e soglia λ con il criterio di Neyman-Pearson;
- comparazione $L(y)$ con λ e decisione.

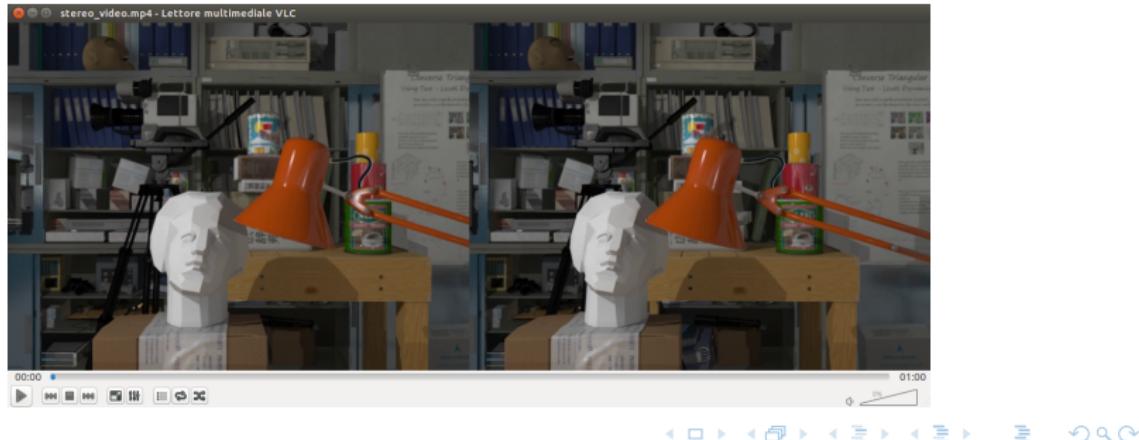
DETECTION VISTA DESTRA



RISULTATI SPERIMENTALI

Test effettuati su:

- sequenza video di 1 minuto creata dal **dataset Tsukuba**
- dimensioni frame
- frame rate 30
- group of picture 60



TEST DI ROBUSTEZZA

Compressione:

Si studia la degradazione del marchio dovuta alla compressione del video, in relazione alla potenza di inserimento.

Distribuzione web:

Si studia la resistenza del marchio in video che sono stati caricati su Youtube e successivamente scaricati.

Sintetizzazione di viste intermedie:

Si studia la risposta del marchio in caso di creazione di viste sintetiche a tre diverse distanze dalla vista sinistra.

COMPRESsIONE

potenza marchio	livello di compressione	Detection stereo
1	1	80%
1	15	40%
1	25	< 20%
1	30	< 20%
3	1	> 90%
3	15	> 90%
3	25	> 90%
3	30	> 90%

Statistica della detection spaziale

potenza marchio	livello di compressione	Detection stereo
0.3	1	100%
0.3	15	> 96%
0.3	25	36.6%
0.3	30	6.6%
0.6	1	100%
0.6	15	100%
0.6	25	86.6%
0.6	30	50%

Statistica della detection in frequenza

DISTRIBUZIONE WEB

potenza del marchio	Detection stereo
0.3	1
0.5	6
0.6	11

Statistica della detection in frequenza per video caricati su Youtube

livello di compressione	PSNR(dB)
15	46.0194
25	40.4861
YT	38.2039
30	37.5587

Valore medio di PSNR per diversi livelli di compressione

SINTETIZZAZIONE DI VISTE INTERMEDIIE



1/4



1/2



3/4



Detection
spaziale

>90%

>80%

>80%

Detection in
frequenza

100%

100%

>90%

TEST DI INVISIBILITÀ: MISURE DI QUALITÀ



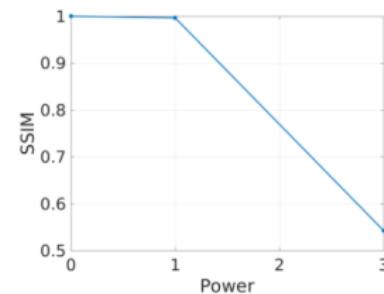
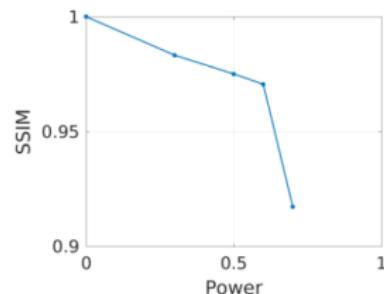
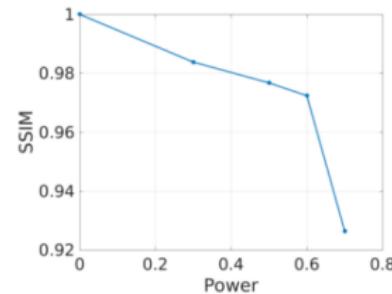
$$Q_{View}(x, y) = [I(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [\mathbf{S}_{View}(x', y')]^\gamma$$

$$Q_{Depth}(x, y) = [I(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [\mathbf{S}_{Depth}(x', y')]^\gamma$$

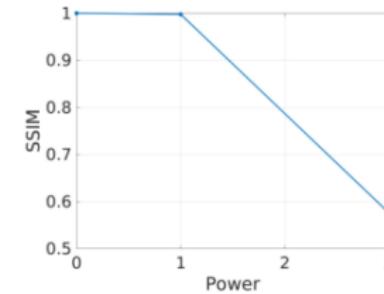
Metriche di Chamida et al.

- Basate sugli **edge** della mappa di disparità e della vista da esaminare
- Versione modificata di **SSIM**
- Metrica di tipo **Reduced-Reference**

TEST DI INVISIBILITÀ: MISURE DI QUALITÀ

vista
sinistravista
destra

metodo in frequenza



metodo spaziale

TEST DI INVISIBILITÀ: PSNR

- Si è misurato l'impatto visivo tramite PSNR tra i frame originali e marchiati;
- valori tipici per video marchiati sono tra 30 e 50 dB.

potenza del marchio	PSNR(dB)
0.3	46.0071
0.5	45.9505
0.6	45.9291

Valore medio di PSNR al variare della potenza del marchio

CONCLUSIONI

Nel dominio spaziale:

- fissiamo un degrado della qualità dell'immagine del 1%
- otteniamo la potenza del marchio pari a 1

CRF	potenza del marchio	True positive
1	1	70%
15	1	40%
25	1	< 20%
30	1	< 20%



CONCLUSIONI

Nel dominio della frequenza:

- fissiamo un degrado della qualità dell'immagine del 1%
- otteniamo la potenza del marchio pari a 0.3

CRF	potenza del marchio	detection stereo
1	0.3	100%
15	0.3	> 90%
25	0.3	36.6%
30	0.3	6.6%



CONCLUSIONI

Video stereo marchiato con potenza 0.3 compresso a fattore 15