Intensità percepita

è una f.ne logaritmica della intensità incidente nell'occhio

L'immagine si può pensare come una f.ne $f(x,y)$ con x e y coordinate spaziali

$$f(x,y) = i(x,y)\,r(x,y) \qquad 0 < i < \infty \quad 0 < r < 1$$

i: luce incidente   r: luce riflessa

Passaggio da continuo a discreto

con campionamento e quantizzazione

PIXEL

L'immagine può essere espressa come una matrice di pixel. val pixel: [0, 255]

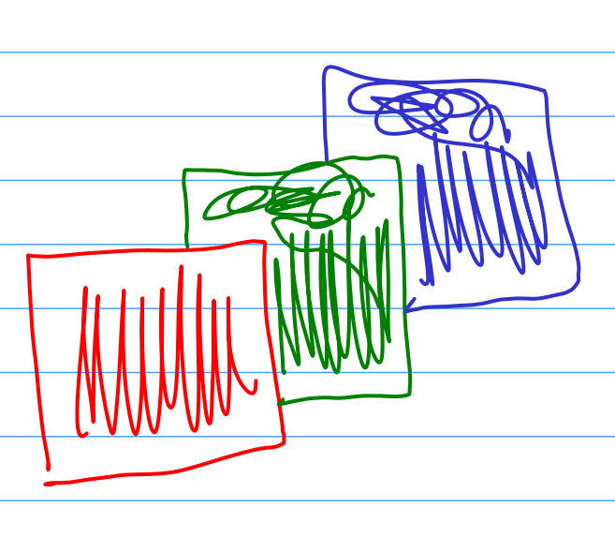
B/W: 1 bit per pixel

Scala di grigi: 8 bit per pixel

Colori: 3 canali da 8 bit ciascuno

Immagini RGB

Ogni immagine è scomponibile nelle sue 3 componenti rosso, blu e verde



Immagini vettoriali e raster

- predilige le forme chiuse, buono per il ridimensionamento
- raster non è ottimo per il ridimensionamento

Vettoriale: pesante computazionalmente, non buono per immagini reali

Raster: migliore per immagini reali

Vettoriale: si occupa di colorare da pixel x a pixel y dello stesso colore

Raster: si occupa di colorare ogni pixel

PSNR

fattore di qualità rispetto a una immagine reale (max ≃ 52)

Operazioni

Prodotto

Nell'image processing si tende a fare il prodotto punto a punto di elementi corrispondenti.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} ax & yb \\ cz & dw \end{bmatrix}$$

Operazioni affini

Si "spostano" le immagini

A ogni operazione è associata una matrice (detta affine)

- Identità $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ $x = v$, $y = w$
- Scaling $\begin{vmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{vmatrix}$ $x = c_x v$, $y = c_y w$
- Rotation $\begin{vmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$ $x = v\cos\theta - w\sin\theta$, $y = v\sin\theta + w\cos\theta$
- Translation $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{vmatrix}$ $x = v + t_x$, $y = w + t_y$
- Vertical Shear $\begin{vmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ $x = v + s_v w$, $y = w$
- Horizontal Shear $\begin{vmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ $x = v$, $y = w + s_h v$

Forward Mapping

(v, w) pixel input

(x, y) pixel output

T matrice affine

$$[x, y, 1] = [v, w, 1] * T$$

Inverse Mapping

$$[v, w, 1] = [x, y, 1] * \mathrm{inv}(T)$$

Utile per colmare i "buchi" lasciati dal forward mapping

Interpolazione

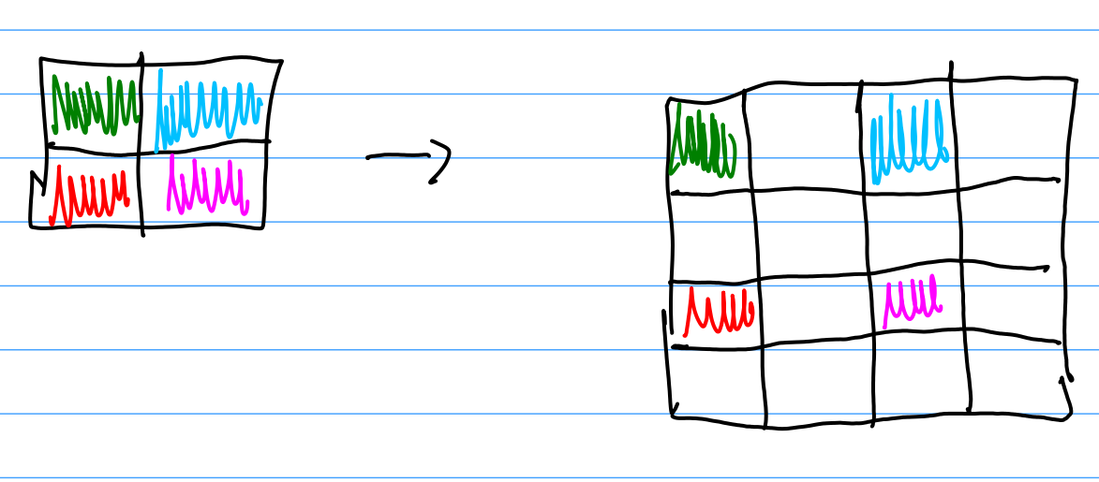
Da fare in seguito a una operazione affine per "colmare" i vuoti

Effettua una stima dei valori ignoti

Zooming in

(2x) $n \times n \to 2n \times 2n$

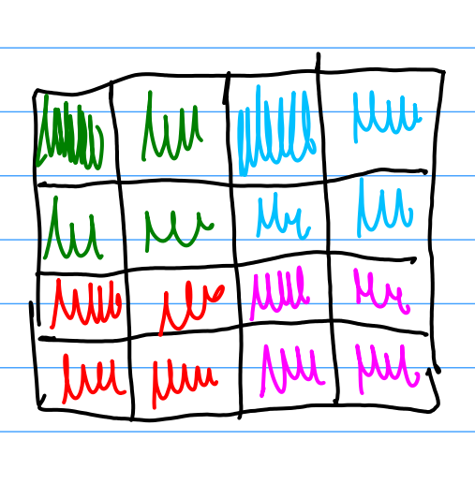
Occorre stimare i valori delle zone vuote



Va fatta una stima per riempirli

Tipi di interpolazione

Replication: pixel vicini vengono replicati. Viene l'immagine a quadri



Bilinear $\underbrace{V(x,y)}_{\text{Valore del pixel in posizione }(x,y)} = ax + by + cxy + d$

Dobbiamo scegliere a, b, c e d in base ai valori dei 4 pixel vicini.

$$V(2,2) = 2a + 2b + 4c + d$$

$$\begin{cases} 100 = V(1,1) = a + b + c + d \\ 150 = V(1,3) = a + 3b + 3c + d \\ 50 = V(3,1) = 3a + b + 3c + d \\ 200 = V(3,3) = 3a + 3b + 9c + d \end{cases}$$

Ricaviamo i valori a, b, c e d

V(2,2) ora è calcolabile

L'immagine bilineare viene un po' più sfocata

BICUBIC

Utilizza i 16 pixel più vicini

$$V(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij}\, x^i y^i$$

I 16 coefficienti sono determinati da un sistema a 16 equazioni → Dipendono dai pixel vicini.

Interpolazione ai bordi

2 possibilità:

- Non fare nulla
- Interpolare coi valori dei pixel vicini anche se di minor numero

Zooming out

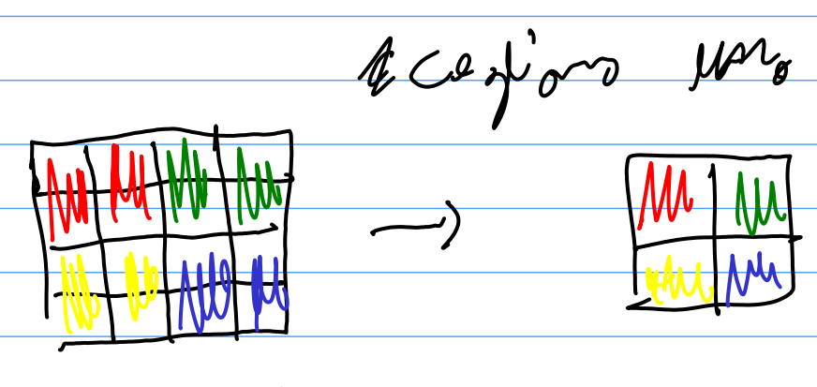
(0,5)

$n \times n \to n/2 \times n/2$

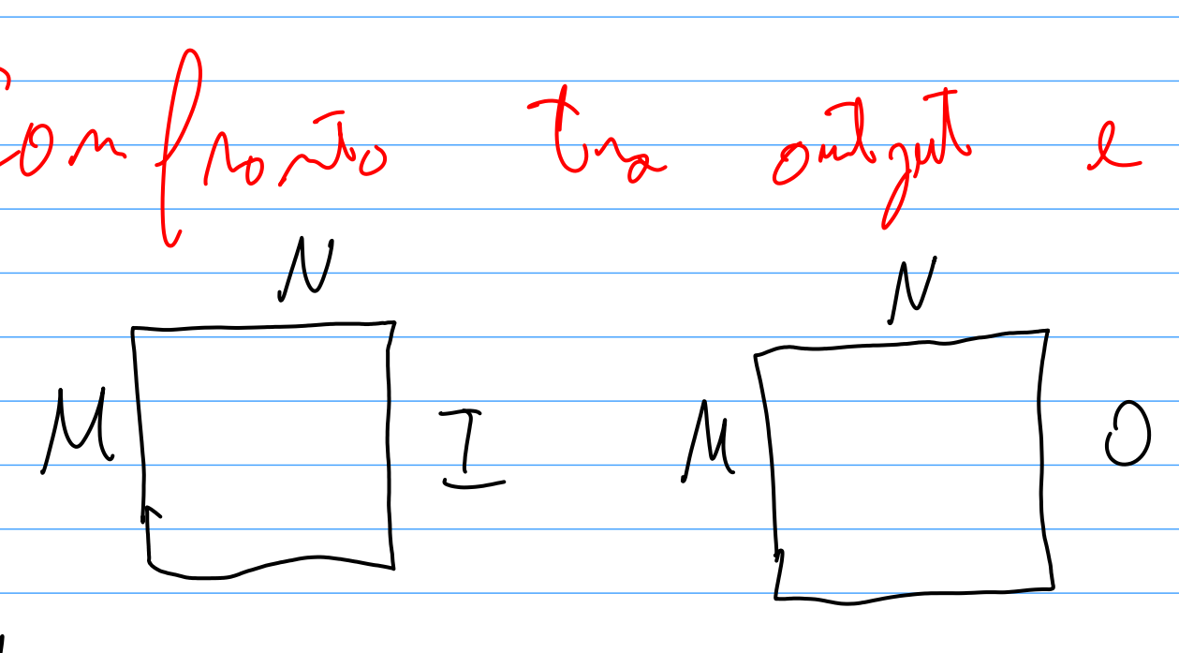
Due strade

scegliamo uno solo dei 4 bit

- Decimazione
- Media dei 4 pixel



Confronto tra output e "ideale"



MSE: errore quadratico medio

$$MSE = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \left(I_{xy} - O_{xy}\right)^2$$

più è basso e meglio è

PSNR: altro parametro di qualità

$$PSNR = -10 \log_{10} \frac{MSE}{S^2}$$

S: max valore del pixel (di solito 255)

più alto è e meglio è

Valore max più alto (sperimentalmente non riesce mai ad andare oltre 52)

35: soglia minima desiderata del PSNR

Il PSNR non è il miglior parametro per valutare la qualità di un algoritmo di interpolazione ma è il più diffuso