Tecniche di Digital Watermarking

Alfredo De Santis

Dipartimento di Informatica
Università di Salerno

ads@dia.unisa.it

http://www.dia.unisa.it/professori/ads

Settembre 2019

Outline

- Definizione di Watermarking
  - Motivazioni
  - Applicazioni
- Tipologie di Watermarking
- Proprietà del Watermarking
  - Percettibilità
  - Robustezza
- Schemi di Watermarking
- Watermarking di Dati Multimediali
  - Immagini
  - Audio

Definizione

- Il **watermark** è essenzialmente una sorta di «*filigrana*»
  - In **ambito informatico** si parla di **digital watermark**, che tipicamente è costituito da una **sequenza di bit**
    - Un logo, una stringa, etc.
- Il **watermarking** è invece la **tecnica** che permette:
  - L'**inserimento** (***embedding***) del watermark all'interno dei dati
  - La **rilevazione** e/o **estrazione** (***detection*** o ***extraction***) del watermark stesso, dai dati in cui esso è stato inserito
- Applicabile su: immagini, video, audio, dati testuali, etc.

Motivazioni

- Il digital watermark è **principalmente utilizzato** per
  - Copy-protection
  - Copyright-protection
- Mediante il digital watermark è possibile
  - Rendere identificabile il legittimo proprietario dei dati
  - Dimostrare l'originalità di dati non contraffatti
  - Evitare la distribuzione di copie non autorizzate
  - Inserire all'interno del dato informazioni ad esso relate
  - Etc.

Cenni Storici

- **1826** - Tecnica del **dandy roll**
  - Ideata da **John Marshall**

  - Durante la fabbricazione della carta, veniva impresso su di essa un marchio (**mark**) mediante un timbro metallico rivestito d'acqua (**water**)

Tipologie di Watermarking

- Diverse tipologie di watermaking sono state introdotte in letteratura
- Tali tipologie sono caratterizzate **in base al modo in cui il watermark può essere rilevato e/o estratto**, da parte di una componente detta **decodificatore** (o decoder)
- Le **principali tipologie** sono le seguenti
  - Blind Watermarking
  - Semi-Blind Watermarking
  - Non-Blind Watermarking

Tipologie di Watermarking

- Diverse tipologie di watermaking sono state introdotte in letteratura
- Tali tipologie sono caratterizzate **in base al modo in cui il watermark può essere rilevato e/o estratto**, da parte di una componente detta **decodificatore** (o decoder)
- Le **principali tipologie** sono le seguenti
  - **Blind Watermarking**
  - Semi-Blind Watermarking
  - Non-Blind Watermarking

Blind Watermarking

- Sia **M** un dato (video, immagine, etc.) o un documento e sia **M'** la versione di **M** in cui è stato inserito il watermark **w**
  - **M'** è detto dato marchiato (o marcato)
- È possibile rilevare e/o estrarre il watermark **w** direttamente da **M'**
  - **Non è necessario quindi il dato originale**
- Le tecniche di blind watermarking richiedono generalmente **fasi di progettazione ed implementazione più complesse**

Tipologie di Watermarking

- Diverse tipologie di watermaking sono state introdotte in letteratura
- Tali tipologie sono caratterizzate **in base al modo in cui il watermark può essere rilevato e/o estratto**, da parte di una componente detta **decodificatore** (o decoder)
- Le **principali tipologie** sono le seguenti
  - Blind Watermarking
  - **Semi-Blind Watermarking**
  - Non-Blind Watermarking

Semi-Blind Watermarking

- Le tecniche di semi-blind watermarking necessitano di alcune informazioni **per poter rilevare e/o estrarre** il watermark
  - Generalmente viene fornito al decodificatore il watermark stesso
    - In tal modo è possibile individuare se il watermark è presente o meno, per cui viene **restituito un valore booleano**: ***true*** **o** ***false***

Tipologie di Watermarking

- Diverse tipologie di watermaking sono state introdotte in letteratura
- Tali tipologie sono caratterizzate **in base al modo in cui il watermark può essere rilevato e/o estratto**, da parte di una componente detta **decodificatore** (o decoder)
- Le principali tipologie sono le seguenti
  - Blind Watermarking
  - Semi-Blind Watermarking
  - **Non-Blind Watermarking**

Non-Blind Watermarking

- Un watermark inserito mediante uno schema di non-blind watermarking può essere **rilevato e/o estratto a patto che il decodificatore sia in possesso del dato originale**
- Questa tipologia di watermarking è in generale di **più facile progettazione ed implementazione**
  - Tuttavia, la necessità di avere il dato originale, ne **limita gli ambiti applicativi**

Percettibilità del Watermark

- Il watermark può essere anche **caratterizzato** in base alla sua **percettibilità**
  - **Visibile**
    - Il watermark è percepibile all'utente
  - **Invisibile**
    - Il watermark non è percepibile all'utente
    - Per valutare la presenza del watermark è necessario un algoritmo di estrazione e/o rilevazione

Visibile

Invisibile

Robustezza del Watermark – 1/3

- Un'ulteriore caratterizzazione del watermark viene fatta in base alla sua **robustezza**
  - Per robustezza si intende la **capacità del watermark di «resistere» ad eventuali alterazioni**
    - Involontarie e/o maliziose
- In base a tale caratterizzazione vengono definite due categorie di watermark
  - Fragile
  - Robusto

Robustezza del Watermark – 1/3

- Un'ulteriore caratterizzazione del watermark viene fatta in base alla sua **robustezza**
  - Per robustezza si intende la **capacità del watermark di «resistere» ad eventuali alterazioni**
    - Involontarie e/o maliziose
- In base a tale caratterizzazione vengono definite due categorie di watermark
  - **Fragile**
  - Robusto

Robustezza del Watermark – 2/3

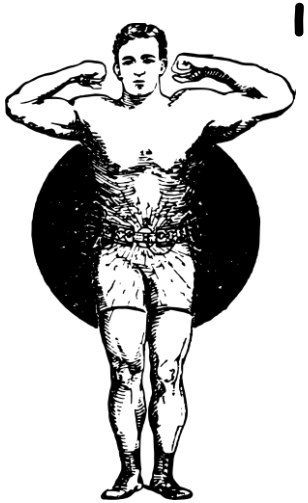
- Un watermark **fragile** è
  - Altamente «sensibile» alle alterazioni
    - Anche semplici modifiche al dato in cui il watermark è incluso, possono far fallire il processo di estrazione e/o rilevazione dello stesso
  - Utile per **rilevare manomissioni**
    - Se il dato viene manipolato (manomesso), il watemark può non essere rilevato
      - Indicando quindi una manomissione

Robustezza del Watermark – 1/3

- Un'ulteriore caratterizzazione del watermark viene fatta in base alla sua **robustezza**
  - Per robustezza si intende la **capacità del watermark di «resistere» ad eventuali alterazioni**
    - Involontarie e/o maliziose
- In base a tale caratterizzazione vengono definite due categorie di watermark
  - Fragile
  - **Robusto**

Robustezza del Watermark – 3/3

- Un watermark **robusto**
  - Deve essere **difficile da rimuovere**
  - Deve essere **resistente a manipolazioni** (involontarie o maliziose) del dato in cui esso è inserito
  - Può essere utilizzato per **trasportare informazioni**
    - Tali informazioni potrebbero essere correlate ai dati stessi, ad es., informazioni sul copyright, etc.



Schema di Watermarking

- Uno **schema di watermarking** definisce
  - Come il watermark viene **inserito** (o **immerso**) nel dato
    - Fase di inserimento (embedding)
  - Come il watermark viene **rilevato** e/o **estratto** dal dato marchiato e successivamente reso disponibile
    - Fase di estrazione e/o rilevazione (extraction e/o detection)

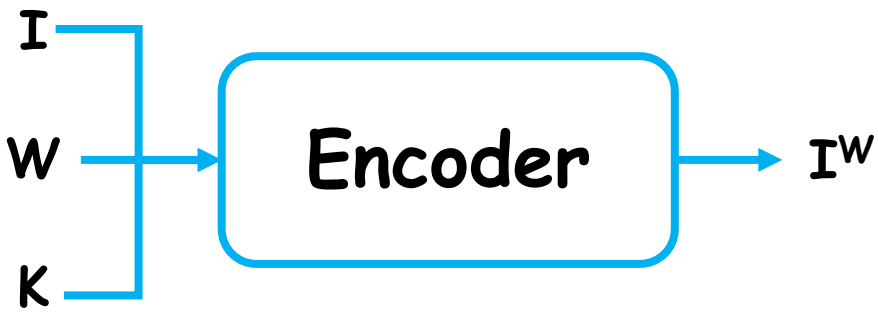
Fasi di uno Schema di Watermarking

- Schema generale
  - Inserimento (embedding)
  - Estrazione e/o rilevazione

Fasi di uno Schema di Watermarking

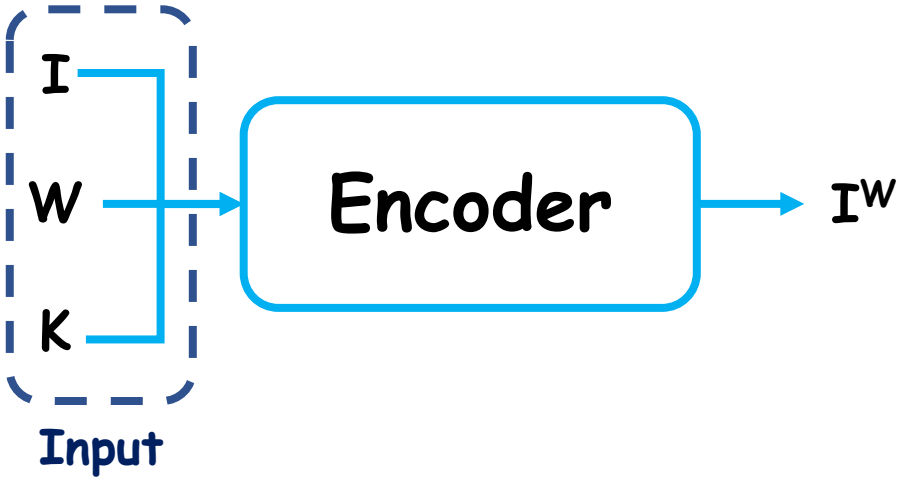
- Schema generale
  - **Inserimento (embedding)**
  - Estrazione e/o rilevazione

Schema generale di inserimento



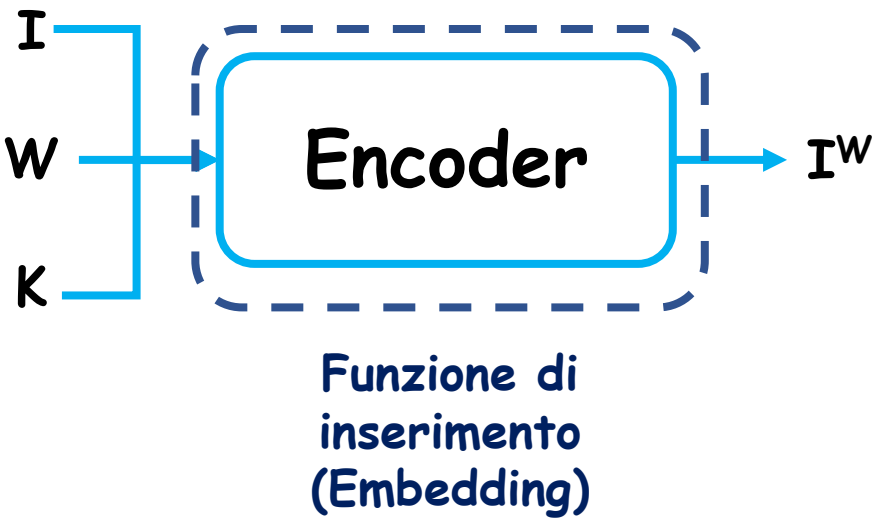
Schema generale di inserimento

- **Input**
  - **I** – Dato di Input
  - **W** – Watermark
  - **K** – Chiave



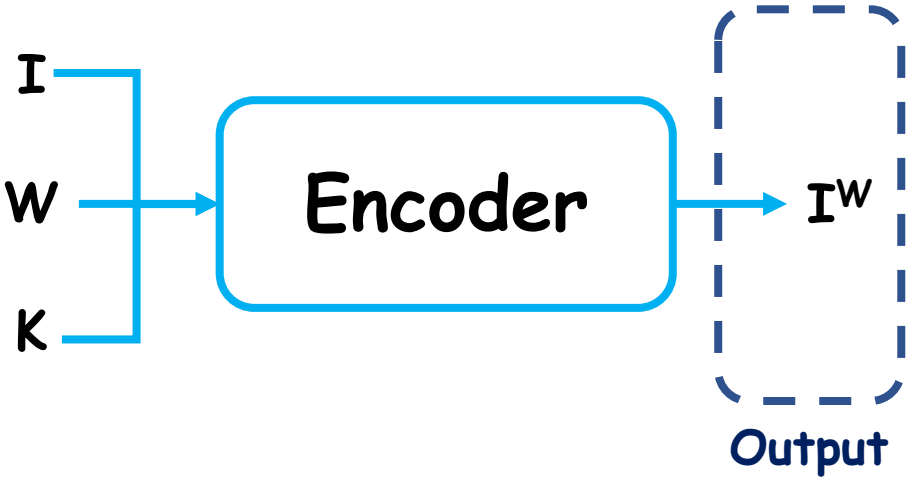
Schema generale di inserimento

- **Input**
  - **I** – Dato di Input
  - **W** – Watermark
  - **K** – Chiave



Schema generale di inserimento

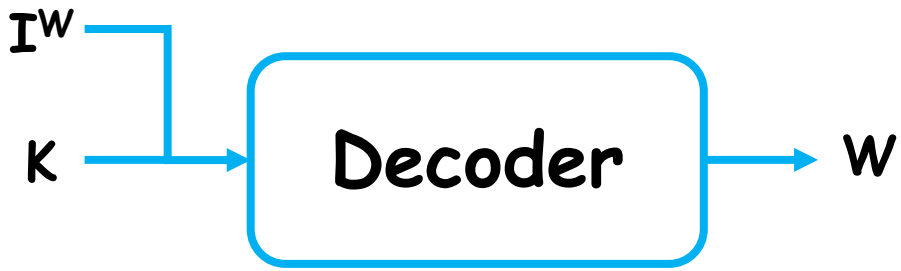
- **Input**
  - **I** – Dato di Input
  - **W** – Watermark
  - **K** – Chiave
- **Output**
  - I^W – Dato marchiato



Fasi di uno Schema di Watermarking

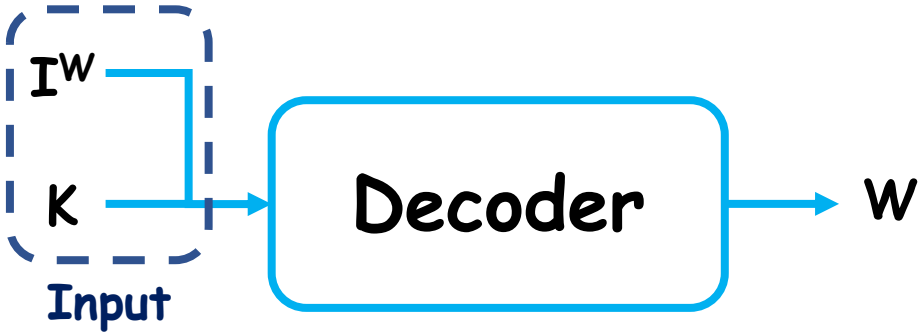
- Schema generale
  - Inserimento (embedding)
  - **Estrazione e/o rilevazione**

Schema generale di estrazione/rilevazione



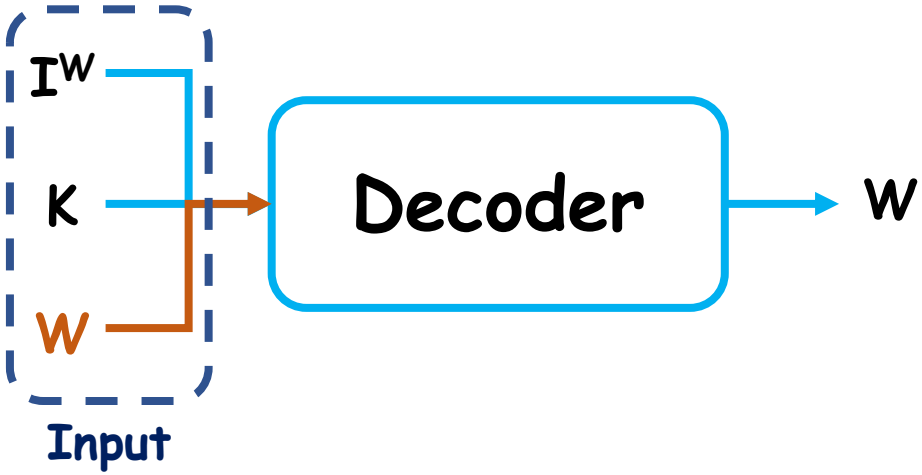
Schema generale di estrazione/rilevazione

- **Input**
  - I^W – Dato di Input
  - **K** – Chiave



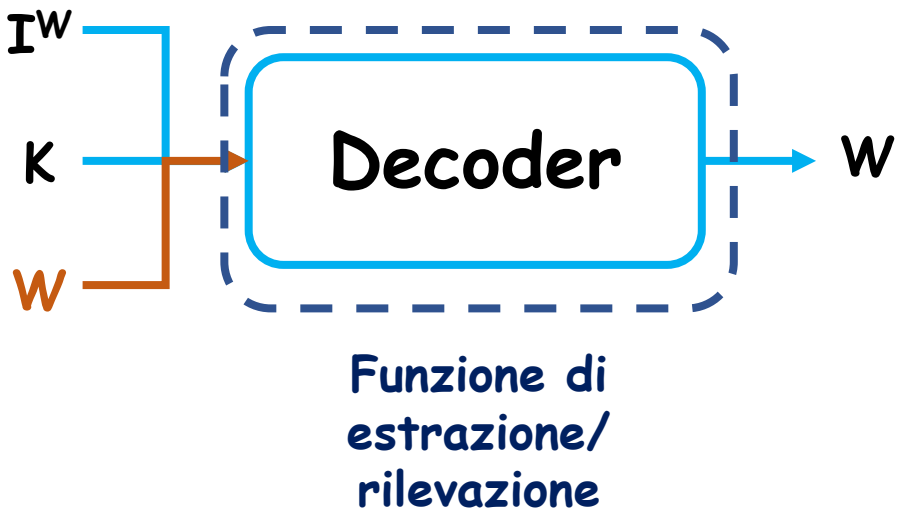
Schema generale di estrazione/rilevazione

- **Input**
  - I^W – Dato di Input
  - K – Chiave
- **Input Opzionali**
  - W – Watermark



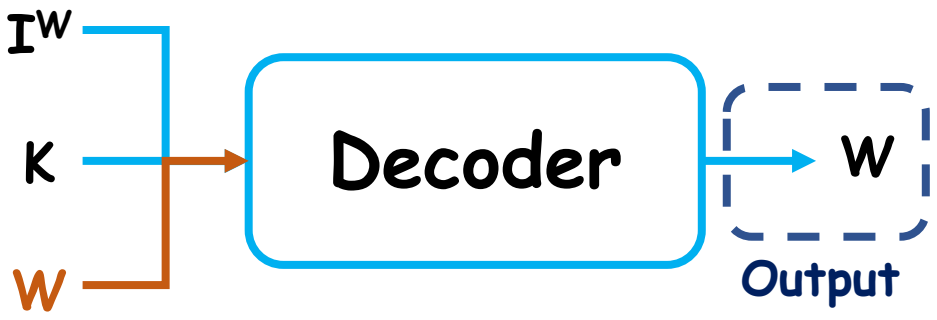
Schema generale di estrazione/rilevazione

- **Input**
  - I^W – Dato di Input
  - K – Chiave
- **Input Opzionali**
  - W – Watermark



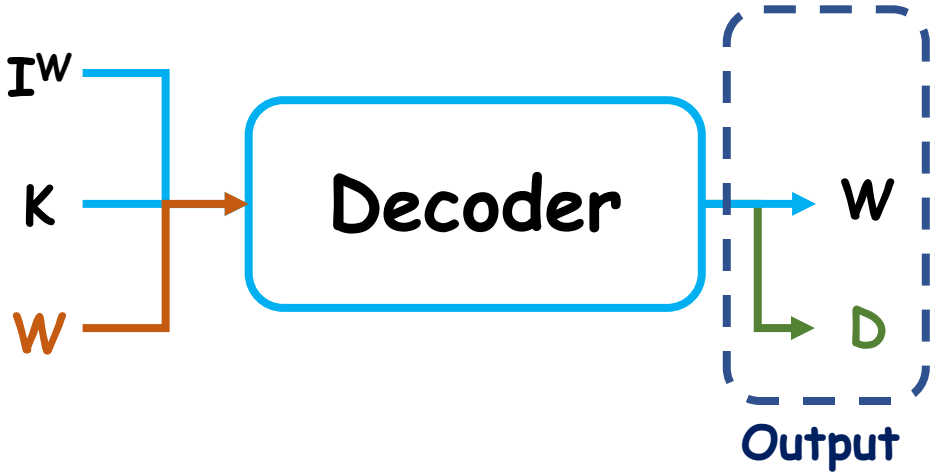
Schema generale di estrazione/rilevazione

- **Input**
  - I^W – Dato di Input
  - K – Chiave
- **Input Opzionali**
  - W – Watermark
- **Output**
  - W – Watermark



Schema generale di estrazione/rilevazione

- **Input**
  - I^W – Dato di Input
  - K – Chiave
- **Input Opzionali**
  - W – Watermark
- **Output**
  - W – Watermark
- **Output Opzionali**
  - D – Detected
    - True o false



Watermarking di Dati Multimediali

- Watermarking su
  - Immagini

  - Audio

Watermarking di Dati Multimediali

- Watermarking su
  - Immagini

  - Audio

Watermarking di Immagini

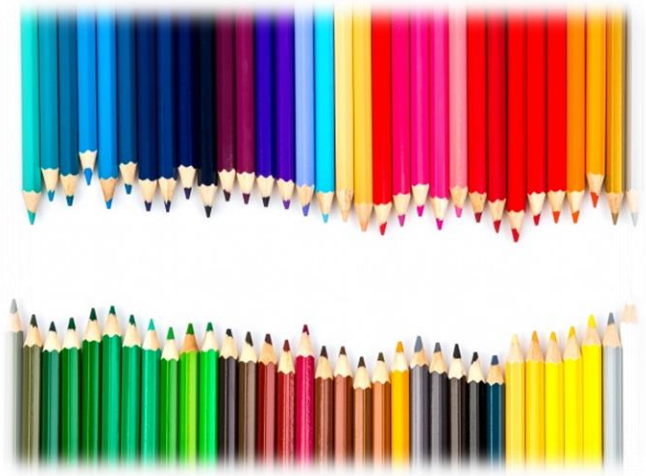
- Le tecniche di watermarking su immagini operano principalmente in **due domini**

Watermarking di Immagini

- Le tecniche di watermarking su immagini operano principalmente in due domini

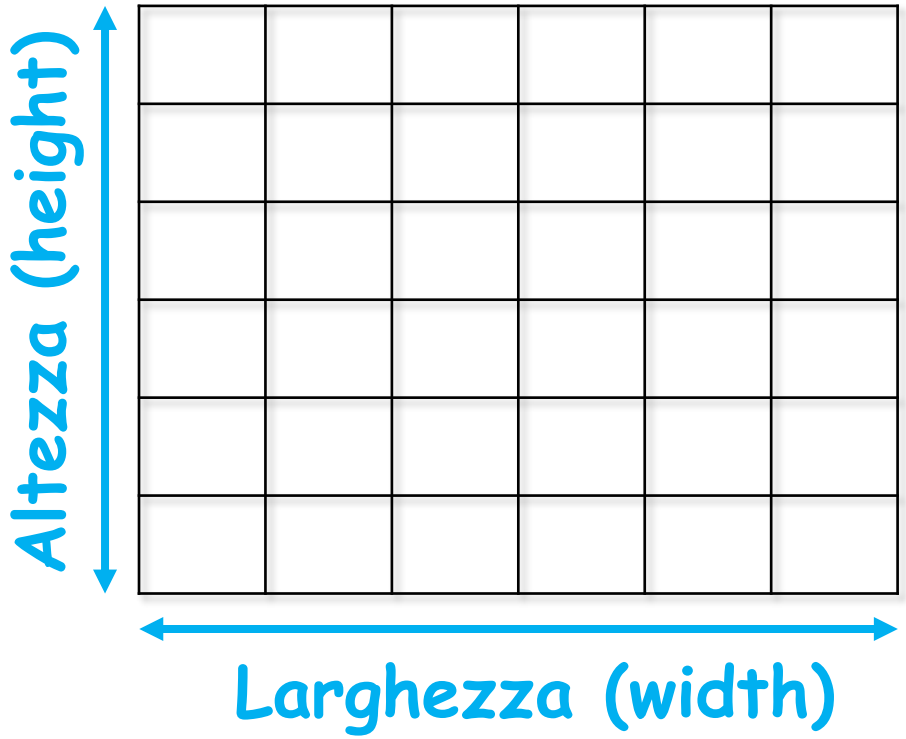
Watermarking di Immagini
Spazio dei Colori

- Le tecniche di image watermarking che operano nel dominio spaziale **effettuano delle elaborazioni sullo spazio dei colori di un'immagine**



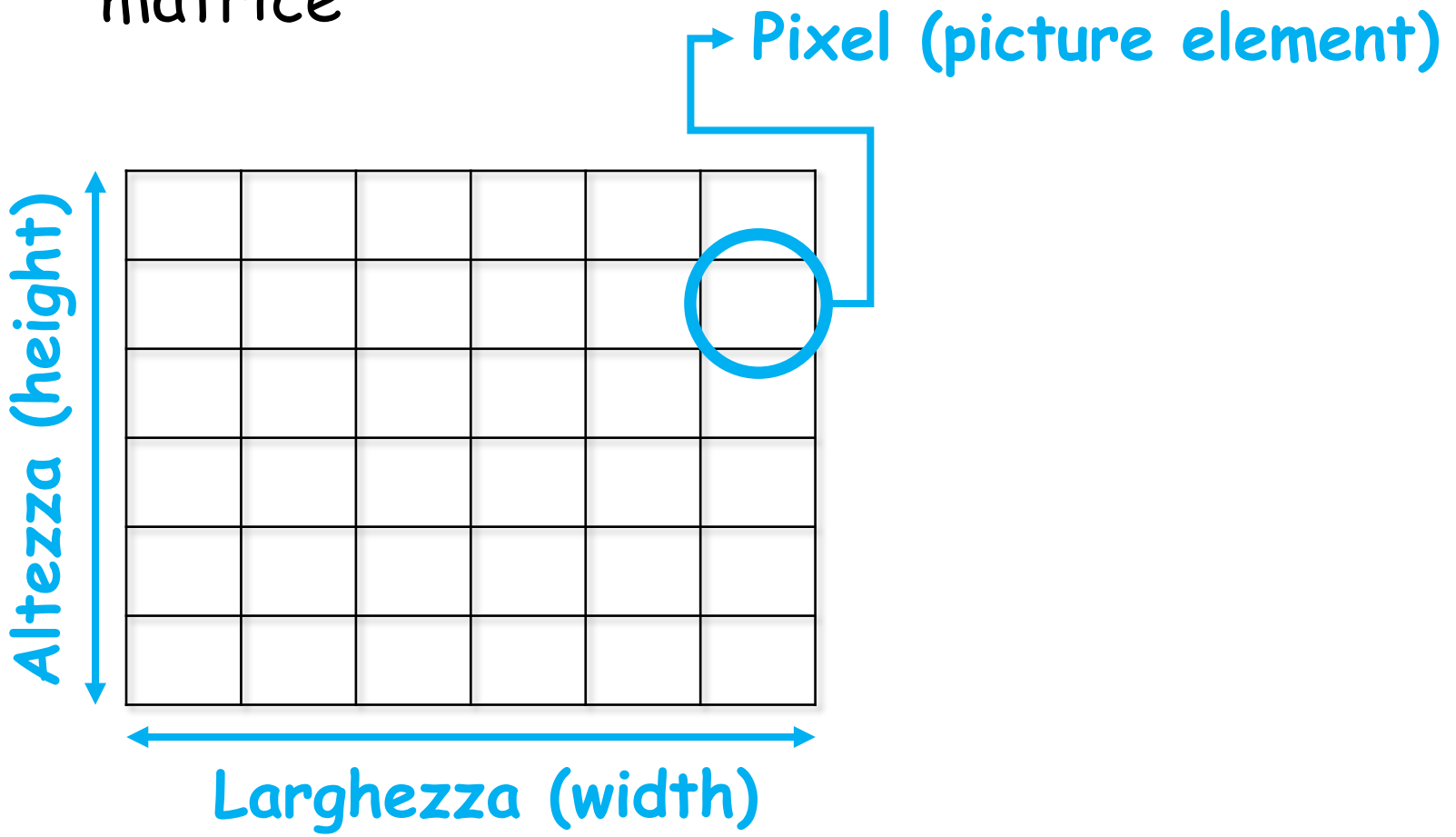
Watermarking di Immagini
Spazio dei Colori

- Un'immagine ***I*** può essere vista come una matrice



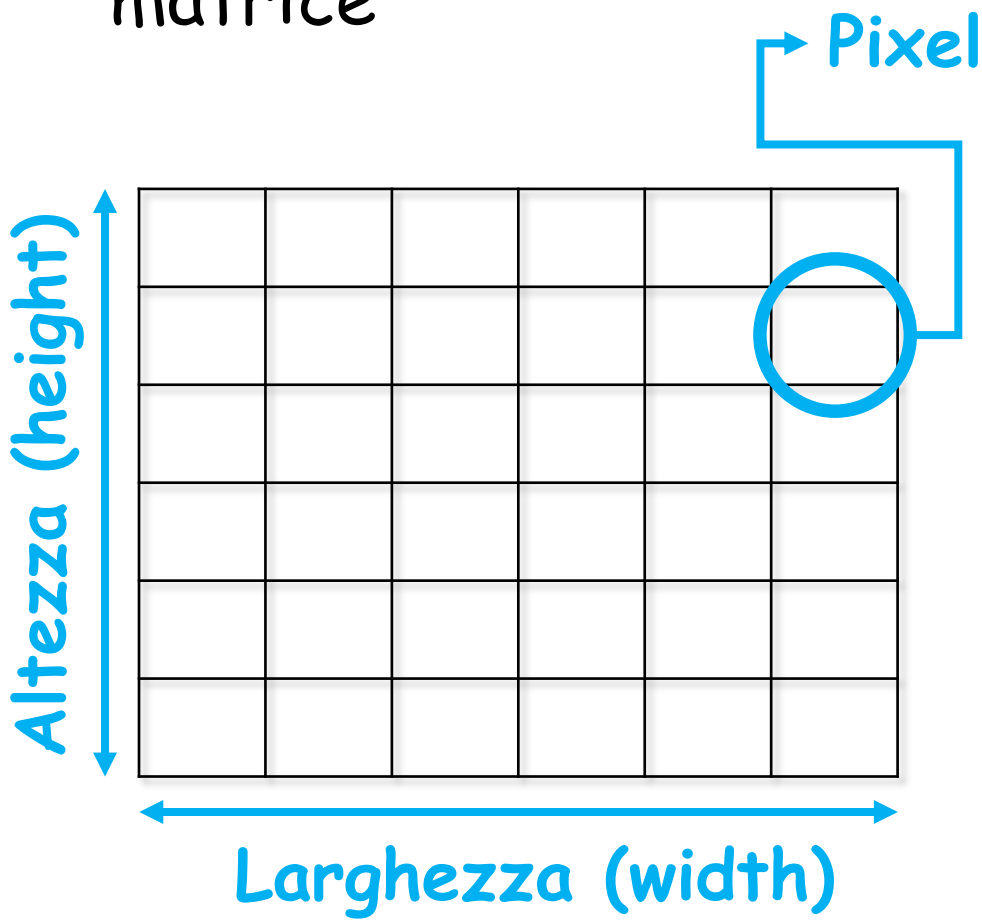
Watermarking di Immagini
Spazio dei Colori

- Un'immagine ***I*** può essere vista come una matrice



Watermarking di Immagini
Spazio dei Colori

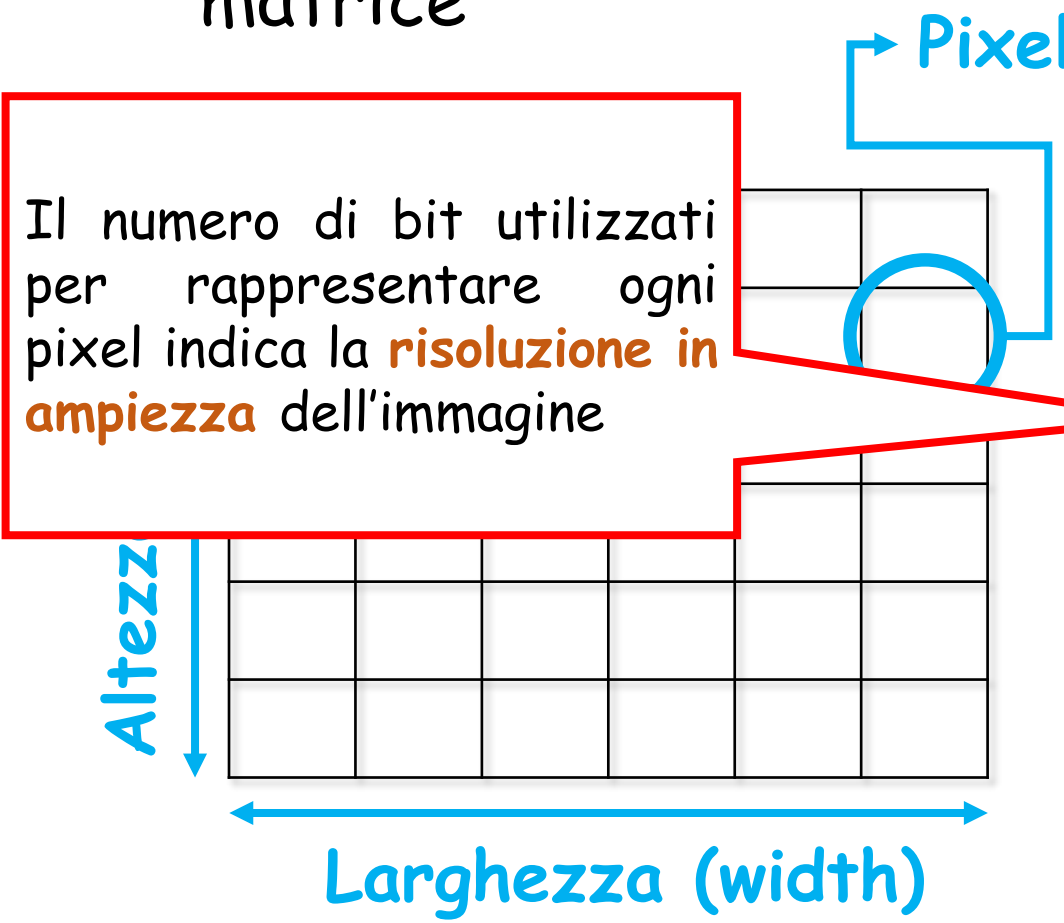
- Un'immagine ***I*** può essere vista come una matrice



- Un pixel è associato ad una stringa binaria
  - Può essere rappresentato mediante 8, 16, 24, 32 bit
  - Contiene informazioni sul colore o la tonalità di grigio in quel punto specifico dell'immagine

Watermarking di Immagini
Spazio dei Colori

- Un'immagine ***I*** può essere vista come una matrice



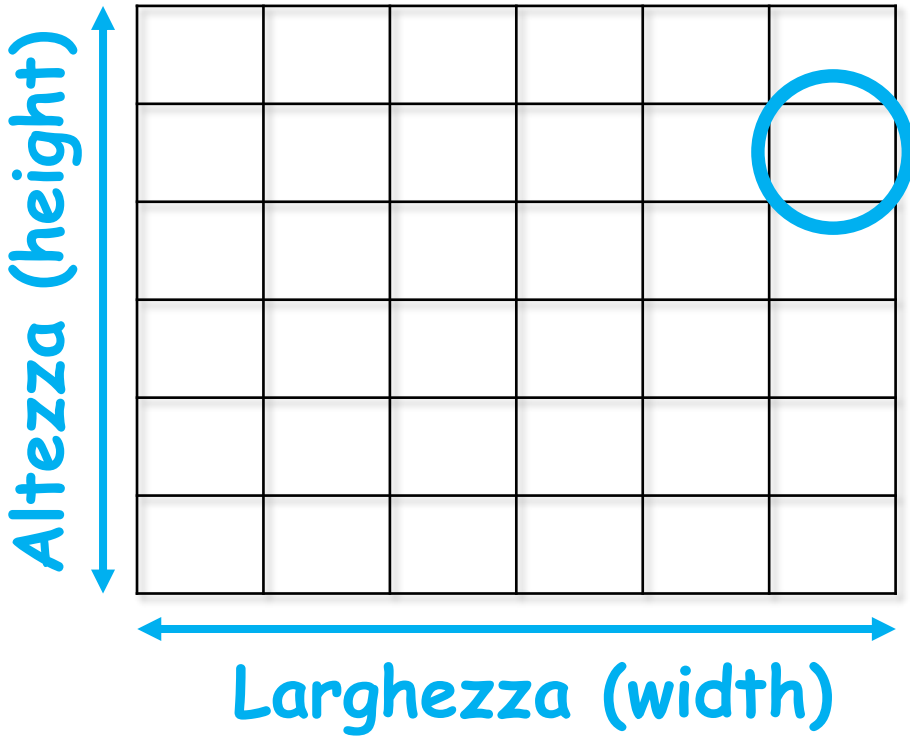
- Un pixel è associato ad una stringa binaria
  - Può essere rappresentato mediante 8, 16, 24, 32 bit
  - Contiene informazioni sul colore o la tonalità di grigio in quel punto specifico dell'immagine

Watermarking di Immagini

La **risoluzione** di un'immagine è data dal numero di pixel dell'immagine stessa

- Larghezza (**width**) × Altezza (**height**)

una

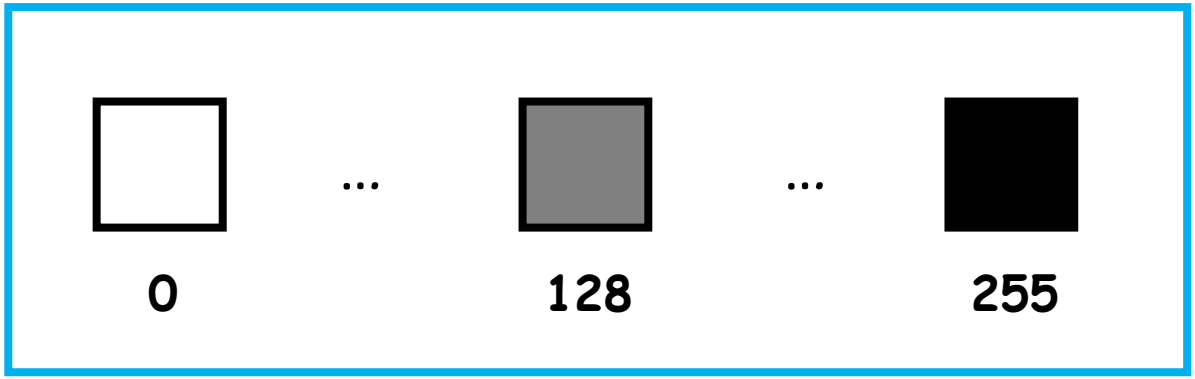


- Un pixel è associato ad una stringa binaria
  - Può essere rappresentato mediante 8, 16, 24, 32 bit
  - Contiene informazioni sul colore o la tonalità di grigio in quel punto specifico dell'immagine

Watermarking di Immagini
Spazio dei Colori

Anatomia di un pixel

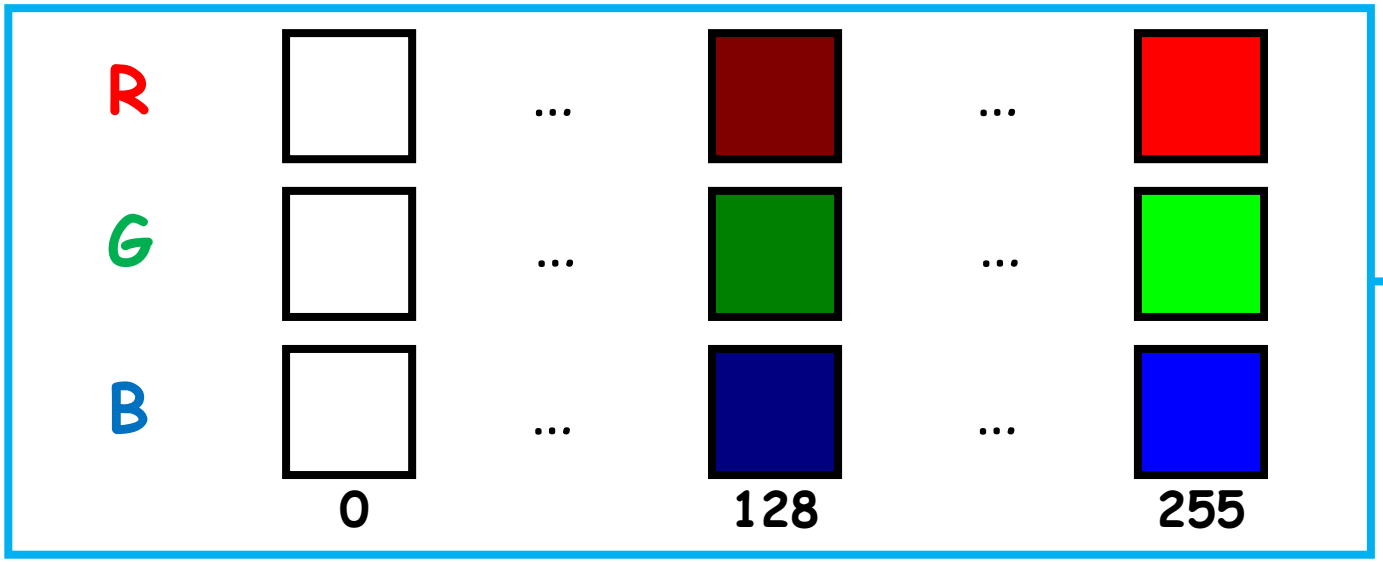
- Nelle immagini a toni di grigio un pixel viene generalmente rappresentato mediante 8 bit
  - **2^8 = 256** tonalità di grigio
    - Valori da **0** (*bianco*) a **255** (*nero*)

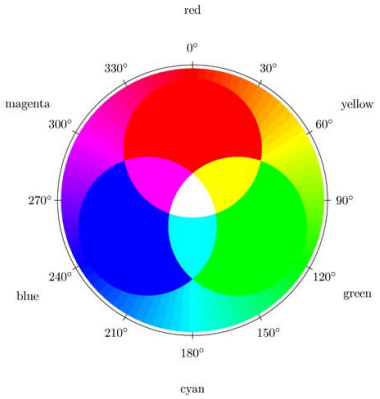


Watermarking di Immagini
Spazio dei Colori

Anatomia di un pixel

- Nelle immagini a colori
  - Il colore viene espresso mediante la combinazione di tre componenti (Modello RGB)
    - Rosso (Red), Verde (Green) e Blu (Blue)
    - Ogni componente varia in modo indipendente





Watermarking di Immagini
Spazio dei Colori

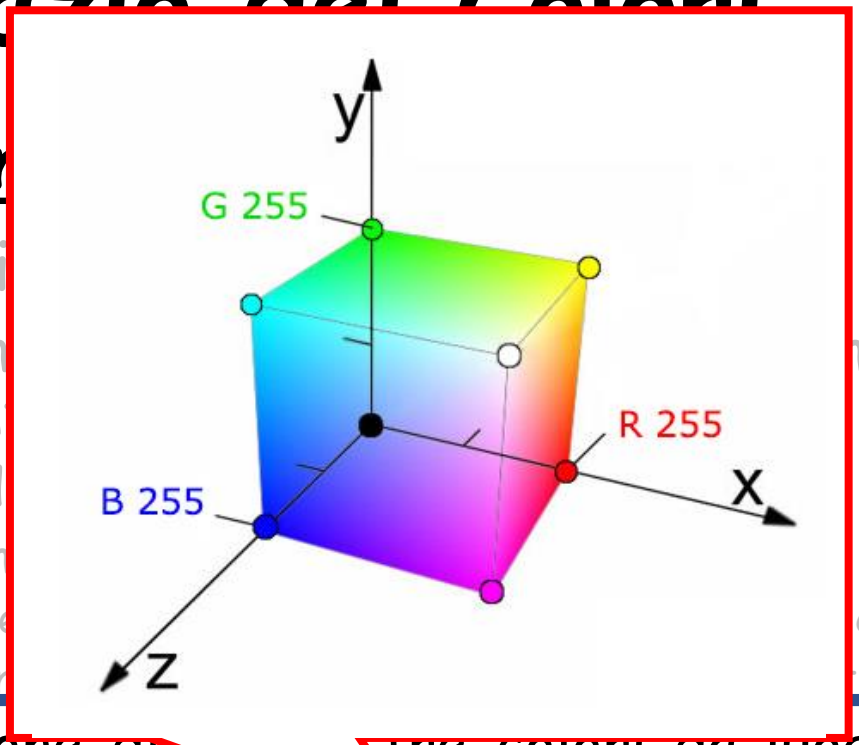
Anatomia di un pixel

- Nelle immagini a colori
  - Il colore viene espresso mediante la combinazione di tre componenti (Modello **RGB**)
    - Rosso (**Red**), Verde (**Green**) e Blu (**Blue**)
    - Ogni componente varia in modo indipendente
      - In genere vengono usati 8 bit per rappresentare un colore
      - **Quindi un pixel verrà rappresentato da 24 bit (8 bit × 3)**
  - La combinazione di questi tre colori da luogo ad uno spazio tridimensionale (una dimensione per colore) noto come spazio dei colori RGB

Watermarking di Immagini
Spazio dei Colori

Ar

- Nelle immagini
  - Il colore vien … nazione di tre componen…
    - Rosso (**Red**…
    - Ogni compon…
      - In genere … colore
      - Quindi ur… × 3)



- La combinazione di … tre colori dà luogo ad uno spazio tridimensionale (… dimensione per colore) noto come **spazio dei colori RGB**

Watermarking di Immagini
Spazio dei Colori

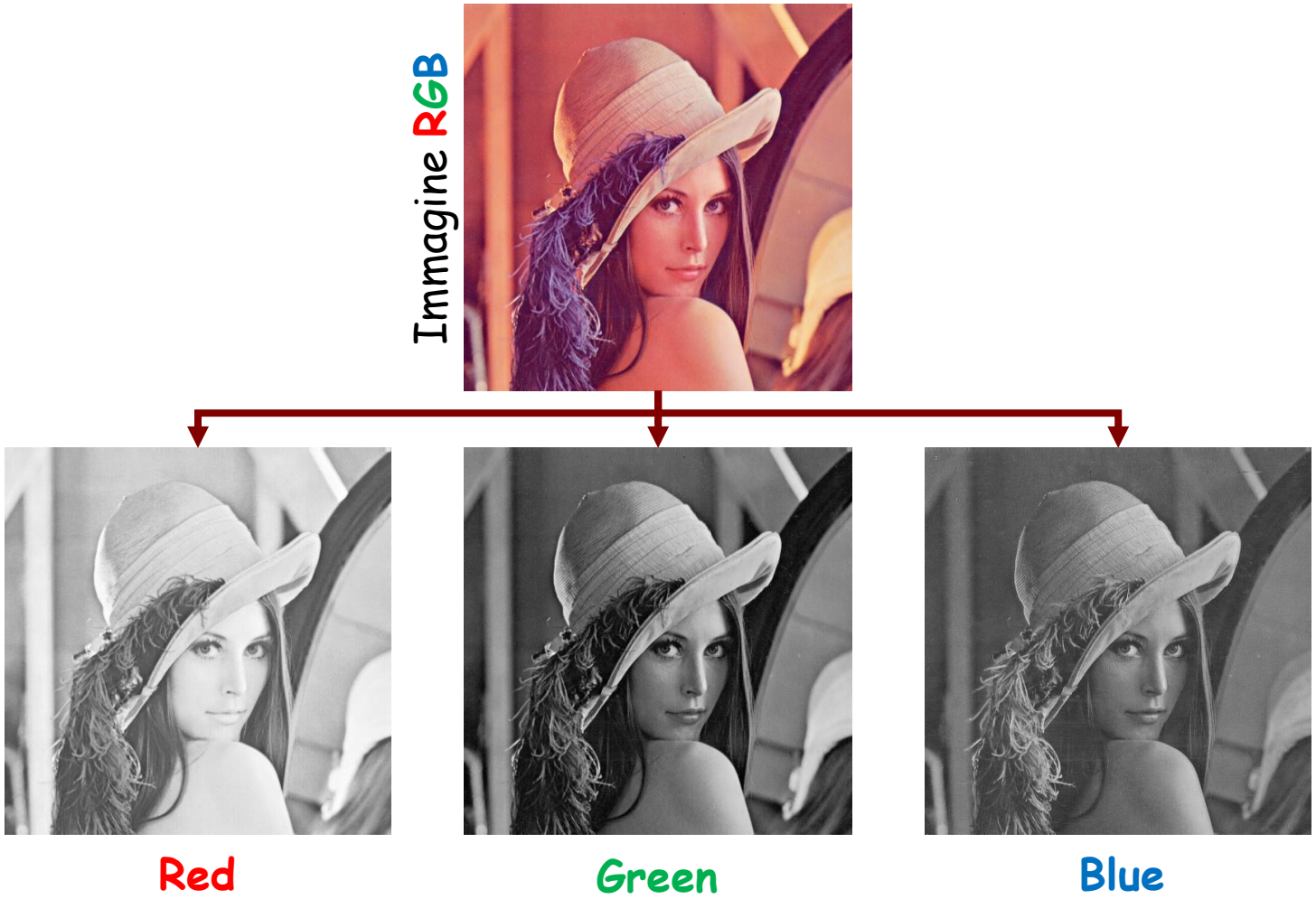
Definizione

- Uno spazio dei colori si definisce mediante un **modello matematico**, al quale viene associata una **funzione di mappatura**
  - Funzione *RGB*

- ***Esempio***
  - *RGB*(0, 0, 0) → Bianco
  - *RGB*(255, 255, 255) → Nero

Watermarking di Immagini
Spazio dei Colori

Immagine RGB scomposta nelle tre componenti: Red, Green, Blue



Watermarking di Immagini
Spazio dei Colori

- Un altro spazio dei colori tridimensionale è quello denominato **YUV**
- Tale spazio è caratterizzato da tre componenti
  - **Y** che indica la **luminanza**
    - *Intensità della luce espressa in scala di grigio (luminosità)*
  - **U** che indica la **tonalità**
    - *Presenza del colore (rosso, verde, giallo, etc.)*
  - **V** che indica la **saturazione**
    - *Descrive la vividezza del colore (molto forte, quasi bianco, etc.)*
- In tale spazio, le due componenti **U** e **V** sono rappresentative della colorazione dell'immagine e sono dette componenti di crominanza

Watermarking di Immagini
Spazio dei Colori

- È possibile passare dallo spazio dei colori **RGB** a quello **YUV** e viceversa
- La luminanza di un pixel, rappresentato mediante RGB, può essere calcolata attraverso la somma pesata dei tre colori
  - **Y** = 30% di R + 60% di G + 10% di B
    = 0.3 * R + 0.6 * G + 0.1 * B
- La crominanza di un pixel, rappresentato mediante RGB, può essere definita come
  - **U** = R – Y
  - **V** = B – Y

Watermarking di Immagini
Spazio dei Colori

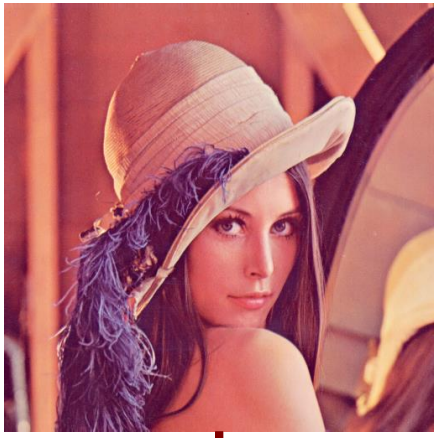
- È possibile passare dallo spazio dei colori **RGB** a quello **YUV** e viceversa
- La lumi di un pixel, rappresentato mediante RGB calcolato attraverso la somma

- Lo spazio dei colori **YUV** è molto utile nell'ambito della **compressione delle immagini**
  - Dal momento che **l'occhio umano ha meno sensibilità sulle componenti di crominanza** è possibile non considerarle

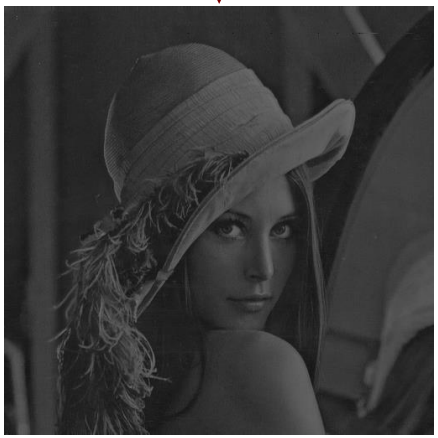
Watermarking di Immagini
Spazio dei Colori

Immagine **RGB** convertita nello spazio dei colori **YUV** e scomposta nelle tre componenti: **Y**, **U**, **V**

Immagine **RGB**



Y



U

V

Watermarking di Immagini **Dominio Spaziale**

- Le tecniche di image watermarking nel dominio spaziale **operano direttamente sui valori dei pixel che compongono l'immagine**
  - Le **operazioni sono generalmente semplici** e poco onerose dal punto di vista della complessità computazionale
    - Utili per implementazione su dispositivi con hardware limitato
  - Un watermark inserito nel dominio spaziale risulta generalmente poco robusto
    - Fragile

Watermarking di Immagini **Dominio Spaziale**

- Una tecnica di image watermarking nel dominio spaziale è conosciuta come **LSB** (**L**ess **S**ignificant **B**it)
  - Il watermark viene inserito modellando tutti (o una parte) i valori dei pixel
    - Effettuando operazioni sul bit meno significativo
- Il watermark inserito con tale tecnica risulta essere
  - **Poco oneroso** dal punto di vista computazionale
    - Sono necessarie solo poche operazioni bit-a-bit
  - **Fragile**
    - Una manipolazione dell'immagine che modifica (ad es. settando a 0) il bit meno significativo di ogni pixel, elimina il watermark

Watermarking di Immagini
Dominio Spaziale – Schema Generale

Fase di inserimento (embedding) – 1/2

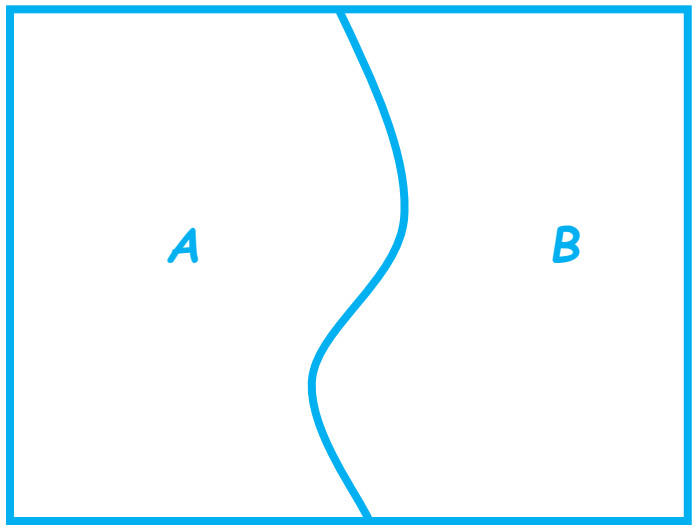
- Siano **A** e **B** due partizioni dell'immagine **I**
  - **A** e **B** contengono i valori dei pixel
  - I valori dei pixel sono uniformemente distribuiti in **A** e **B**
  - La dimensione delle due partizioni risulta essere simile
    - $|A| \approx |B|$
  - Il partizionamento viene effettuato mediante una chiave segreta

Watermarking di Immagini
Dominio Spaziale – Schema Generale

Fase di inserimento (embedding) – 1/2

- Siano **A** e **B** due partizioni dell'immagine **I**
  - **A** e **B** co[illegible]ng
  - I valori de[illegible]buiti in **A** e **B**
  - La dimension[illegible]a essere simile
    - $|A| \approx |B|$
  - Il partizionar[illegible]ante una chiave segret[illegible]

Esempio di Partizionamento



Watermarking di Immagini
Dominio Spaziale – Schema Generale

Fase di inserimento (embedding) – 2/2

- Il valore del watermark **w** (intero)
  - Viene aggiunto al valore di tutti i pixel della partizione ***A***
  - Viene sottratto al valore di tutti i pixel della partizione **B**
  - **w** è un intero piccolo abbastanza
    - La sua addizione/sottrazione non provoca all'immagine una degradazione percettibile
      - Il watermark risulta essere quindi **invisibile**

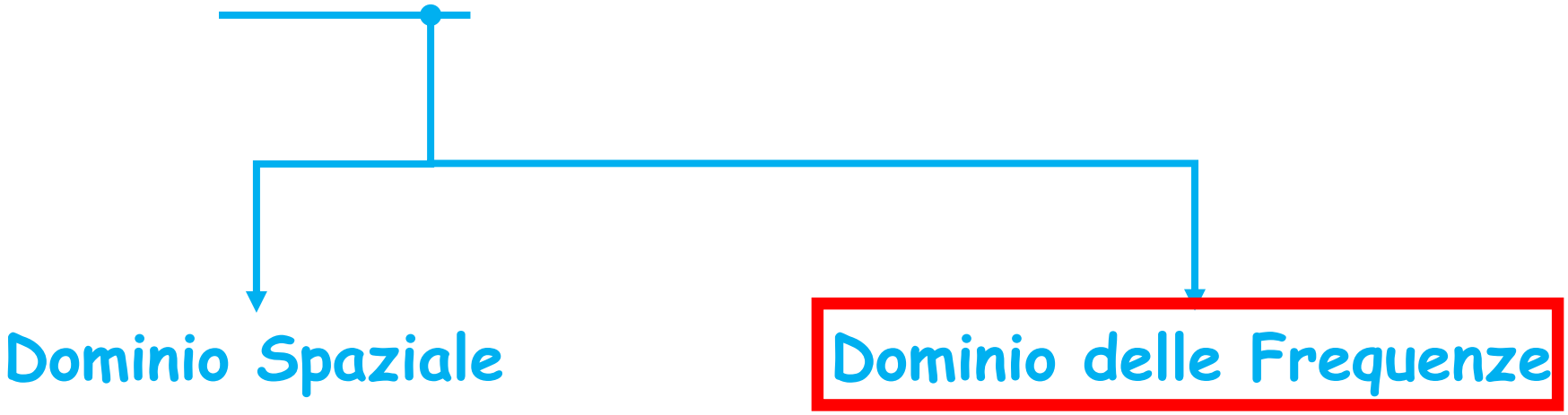
Watermarking di Immagini
Dominio Spaziale – Schema Generale

Fase di rilevazione

- Mediante la medesima chiave segreta utilizzata per l'embedding, vengono ridefinite le partizioni **A** e **B** dell'immagine **I**
- Ciascuna partizione contiene i medesimi pixel selezionati nella fase di embedding
  - Viene calcolata la media dei valori di tutti i pixel appartenenti alla partizione **A** ed alla partizione **B**
  - Siano **A'** e **B'** rispettivamente la media della partizione **A** e della partizione **B**
    - La differenza sarà prossima a uno dei seguenti valori
      - **2 * w**, se il watermark **w** è presente
      - **0**, altrimenti

Watermarking di Immagini

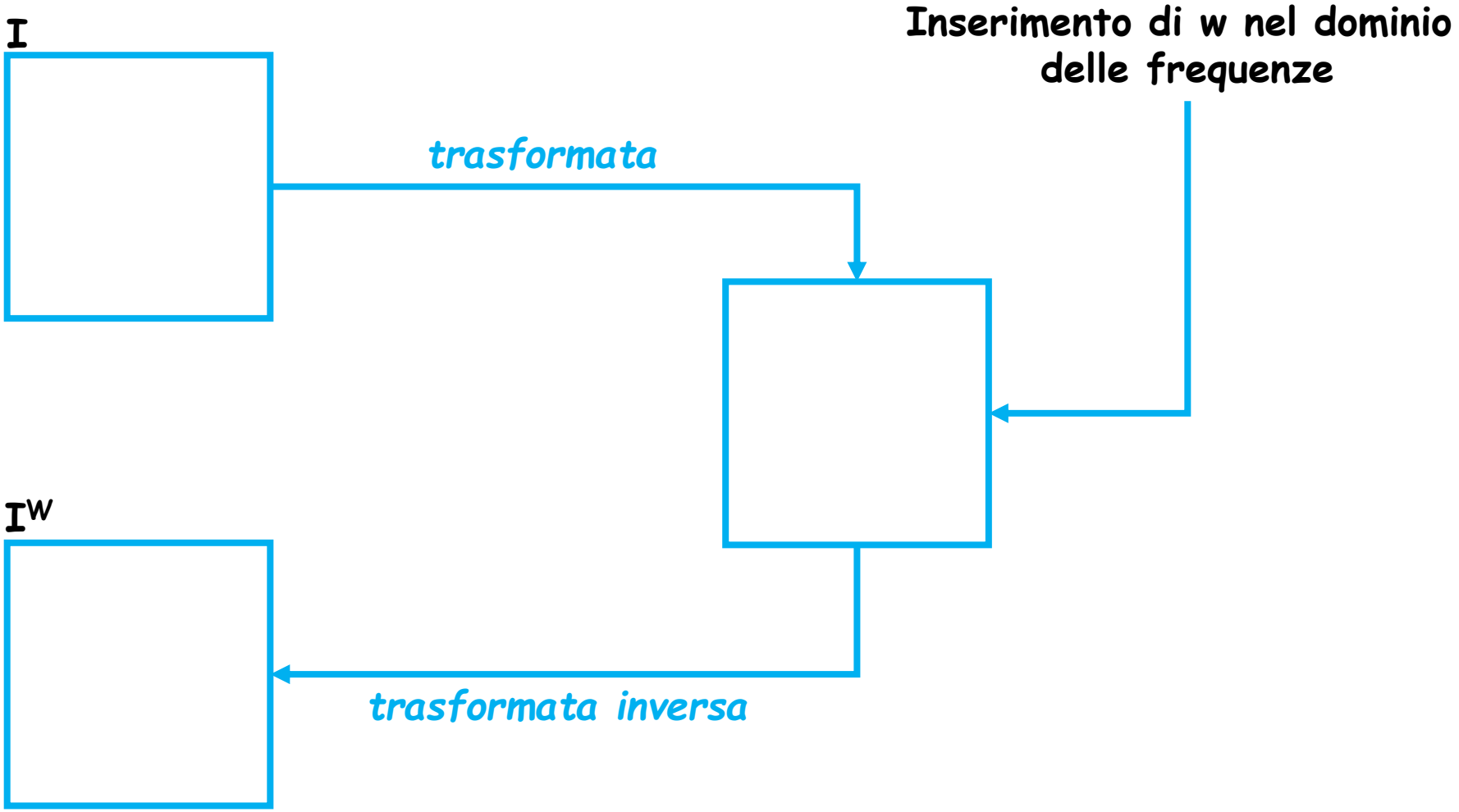
- Le tecniche di watermarking su immagini operano principalmente in due domini



Watermarking di Immagini **Dominio delle Frequenze**

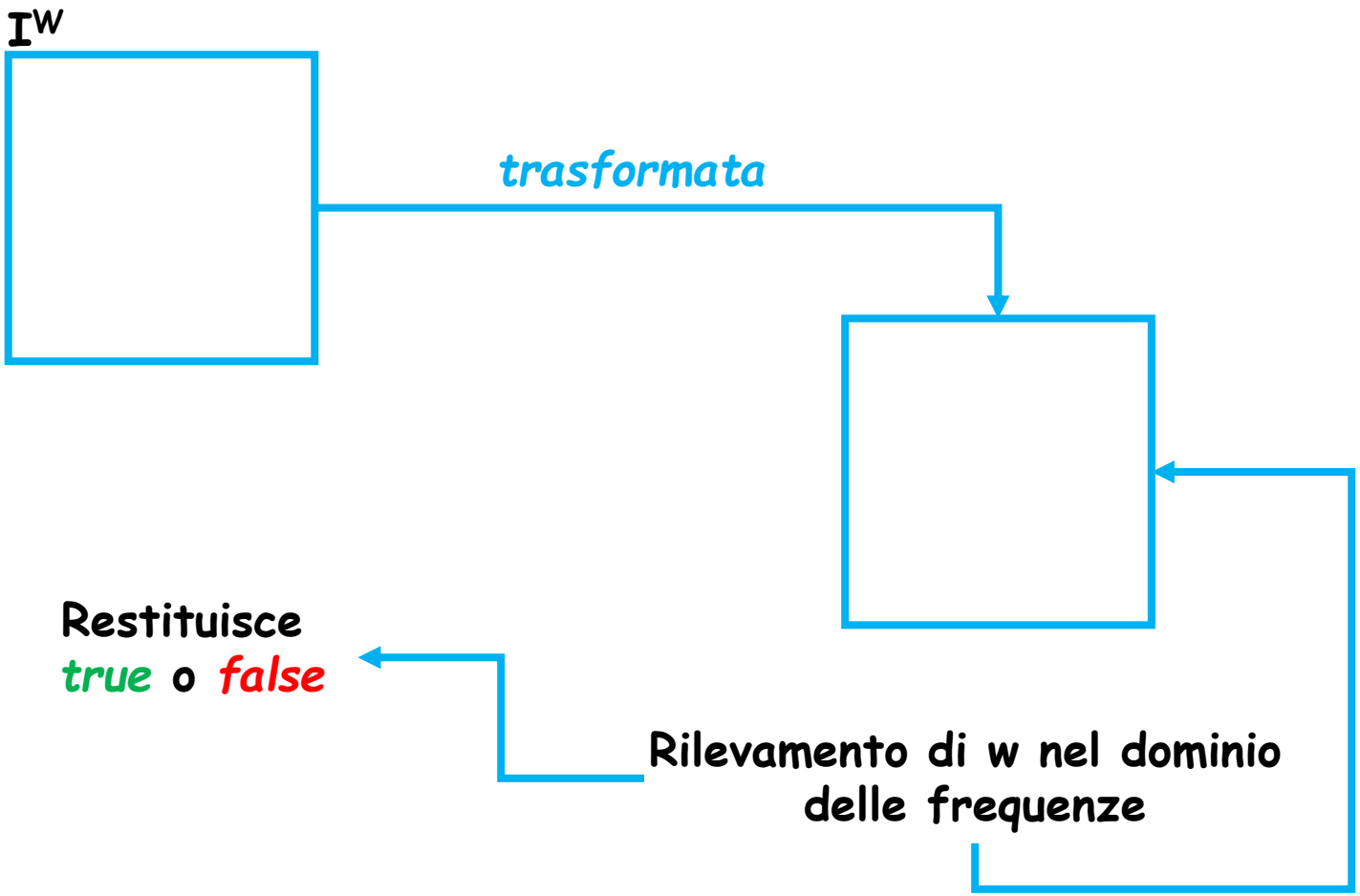
- Le tecniche di image watermarking che operano nel **dominio delle frequenze effettuano i seguenti passi**
  1. Applicano una specifica trasformata all'immagine
  2. Effettuano le operazioni relative all'inserimento o alla rilevazione del watermark
  3. Applicano la trasformata inversa
     - Ottenendo così l'immagine marcata

Watermarking di Immagini
Dominio delle Frequenze



Schema di Inserimento

Watermarking di Immagini
Dominio delle Frequenze



Schema di Rilevamento

Watermarking di Immagini
Dominio delle Frequenze

- Le trasformate più utilizzate sono
  - **DCT** (Discrete Cosine Transform)
  - **DWT** (Discrete Wavelet Transform)
  - **DFT** (Discrete Fourier Transform)

Watermarking di Immagini
Dominio delle Frequenze

DCT (Discrete Cosine Transform)

Definizione Matematica

- La DCT è una funzione lineare invertibile nel dominio dei numeri reali

$$\mathrm{DCT}(x_1, x_1, \dots, x_N) = X_1, X_2, \dots, X_N$$

- Gli N dati di input, $x_1, x_2, \dots, x_N \in R^N$, vengono trasformati nel modo seguente

$$X_k = \sum_{n=0}^{N} x_n \cos\left[\frac{\pi}{n}\left(n + \frac{1}{2}\right)k\right] \qquad k = 0, 1, \dots, N$$

Watermarking di Immagini
Dominio delle Frequenze

DCT (Discrete Cosine Transform)

Motivazioni

- I dati di output della DCT ($X_1, X_2, \ldots, X_N \in R^N$) costituiscono una nuova rappresentazione dei dati di input
  - Tale rappresentazione **rende l'applicazione di alcune operazioni più efficiente e agevole**
  - **I valori dei dati di input vengono compattati e concentrati nei primi dati di output**

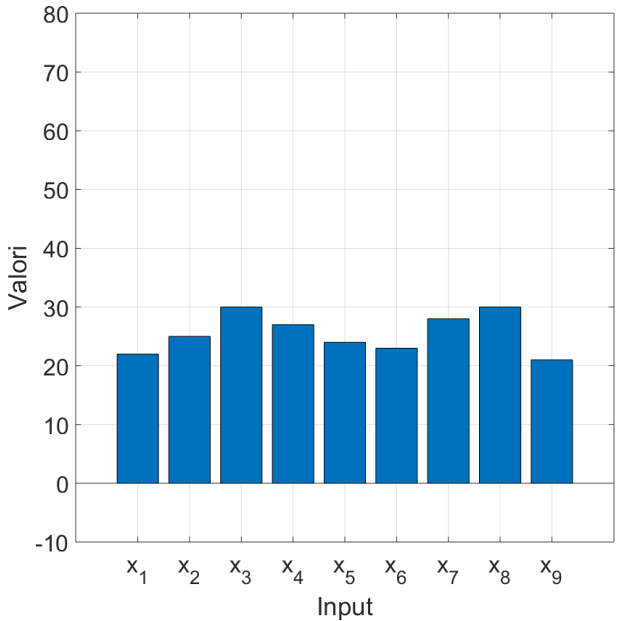
Watermarking di Immagini
Dominio delle Frequenze
DCT (Discrete Cosine Transform)
Esempio 1

Input

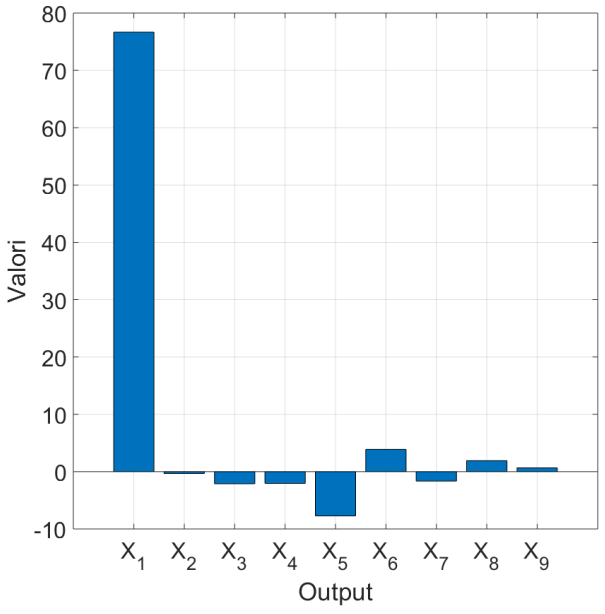
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
22	25	30	27	24	23	28	30	21

Output (Applicazione della DCT)

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
76,7	-0,3	-2,1	-2,0	-7,7	3,9	-1,7	1,9	0,7

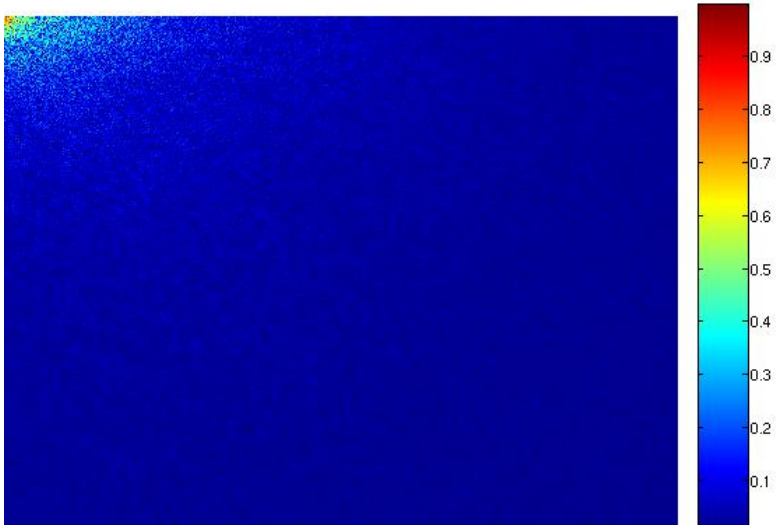


DCT



Watermarking di Immagini
Dominio delle Frequenze
DCT (Discrete Cosine Transform)
Esempio 2

Immagine Originale



Applicazione della DCT

Watermarking di Immagini
Dominio delle Frequenze

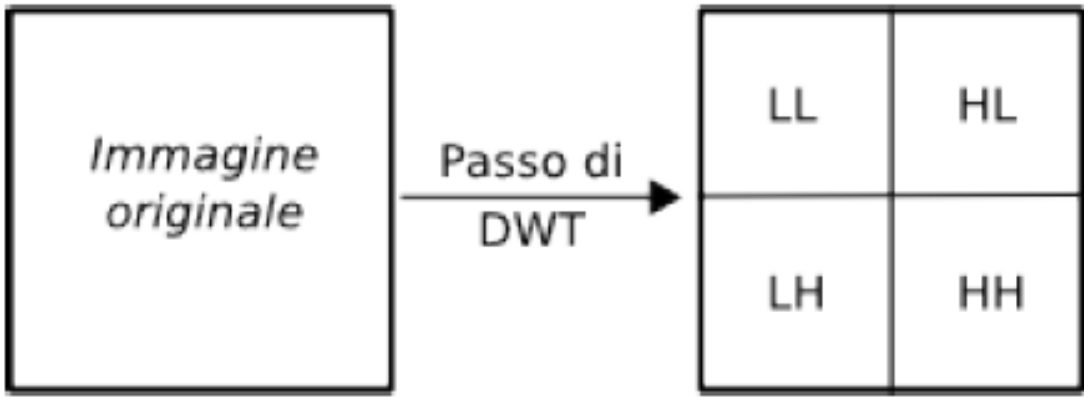
DWT (Discrete Wavelet Transform)

- La DWT è una trasformata invertibile
  - Calcolando l'inversa della trasformata DWT è possibile ottenere l'informazione iniziale
- Mediante la DWT le componenti di un segnale (l'immagine) vengono scomposte in
  - Componenti ad alta frequenza
  - Componenti a bassa frequenza

Watermarking di Immagini
Dominio delle Frequenze

DWT (Discrete Wavelet Transform)

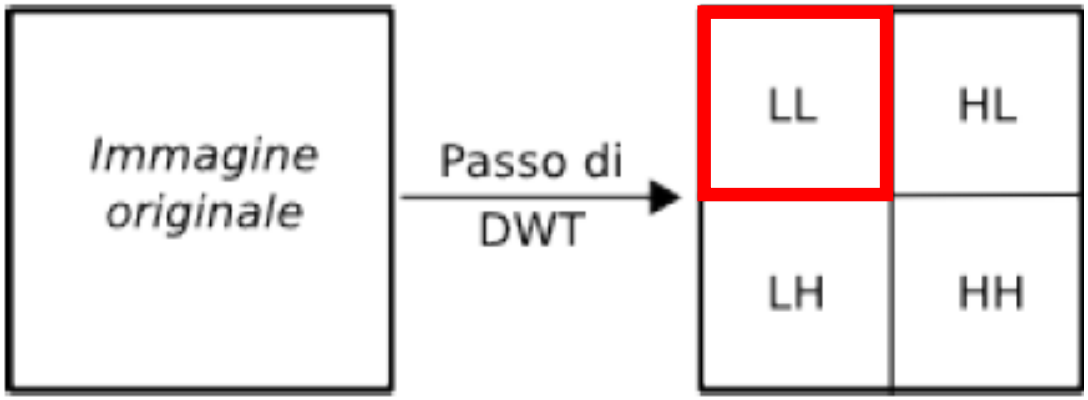
- Applicando la DWT, l'immagine viene suddivisa in quattro sottobande
  - **LL, HL, LH e HH**



Watermarking di Immagini
Dominio delle Frequenze

DWT (Discrete Wavelet Transform)

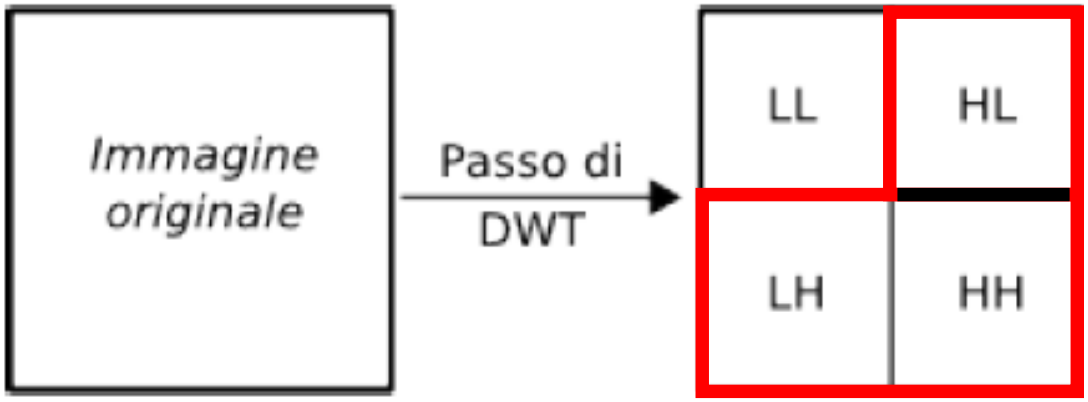
La sottobanda **LL** contiene le **basse frequenze**



Watermarking di Immagini
Dominio delle Frequenze

DWT (Discrete Wavelet Transform)

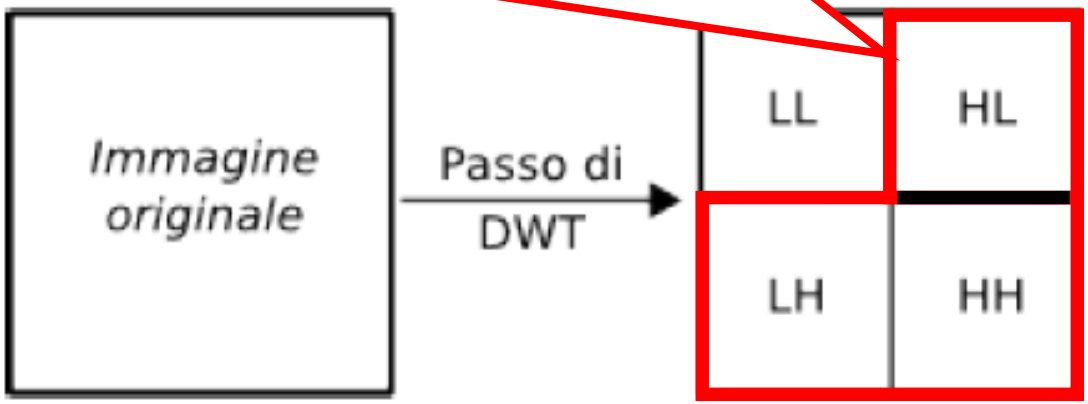
Le sottobande **HL**, **LH** e **HH** contengono le **alte frequenze**, **meno percepibili dall'occhio umano**



Watermarking di Immagini
Dominio delle Frequenze

DWT (Discrete Wavelet Transform)

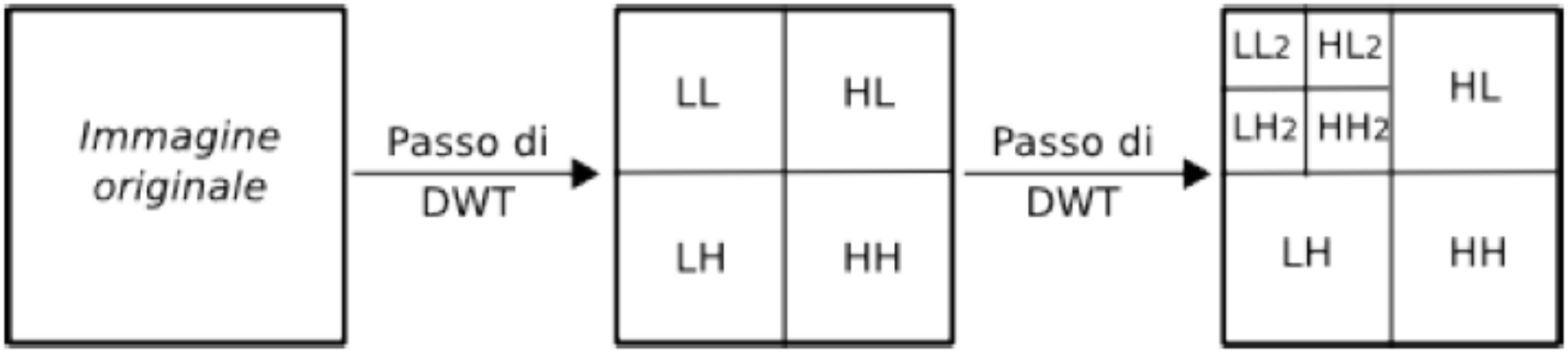
- Nelle sottobande delle basse frequenze operano gli algoritimi di watermarking
  - In tal modo il watermark non risulterà percepibile



Watermarking di Immagini
Dominio delle Frequenze

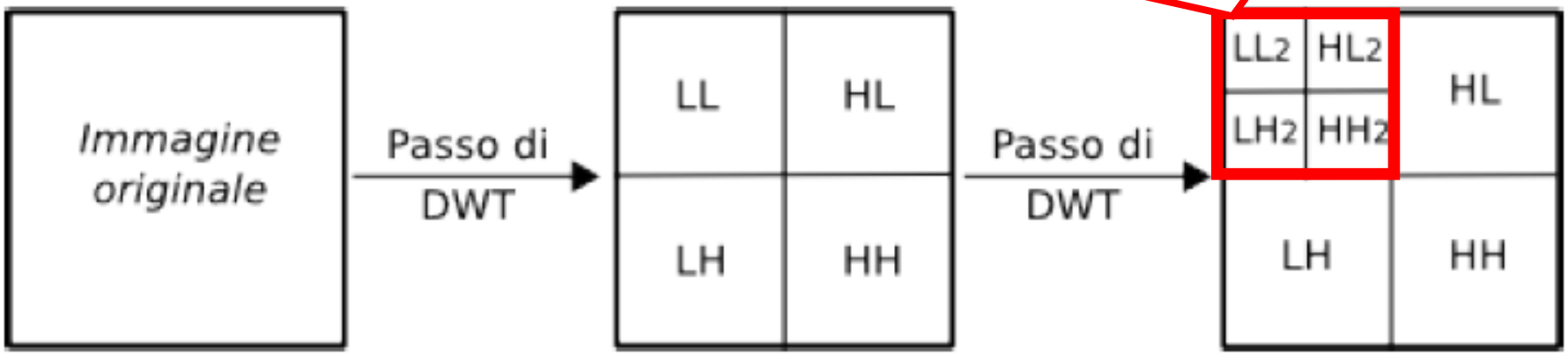
DWT (Discrete Wavelet Transform)

- La sottobanda **LL** può essere ulteriormente scomposta mediante l'applicazione di un ulteriore passo della DWT



Watermarking di Immagini
Dominio delle Frequenze

- L'applicazione del secondo passo della DWT sulla sottobanda **LL** genera quattro ulteriori sottobande, denominate **LL2**, **HL2**, **LH2** e **HH2**
- La sottobanda **LL2** sarà quella contenente le basse frequenze



Watermarking di Immagini
Dominio delle Frequenze

DWT (Discrete Wavelet Transform)

Esempio

Immagine Originale

Applicazione
della DWT

Applicazione della DWT
sulla sotto-banda **LL**

Watermarking di Immagini
Dominio delle Frequenze

DFT (Discrete Fourier Transform)

Definizione Matematica

- La DFT è definita nel dominio dei numeri complessi
  - Trasforma l'input $x_0, x_1, x_2, \dots, x_{N-1}$, composto da N numeri complessi, in una successione di N numeri complessi $X_0, X_1, X_2, \dots, X_{N-1}$

$$X_k = \sum_{n=0}^{N-1} x_n e^{-ik\left(\frac{2\pi}{N}\right)n}, k = 0, 1, 2, \dots, N-1$$

Watermarking di Immagini
Dominio delle Frequenze

DFT (Discrete Fourier Transform)

Esempio

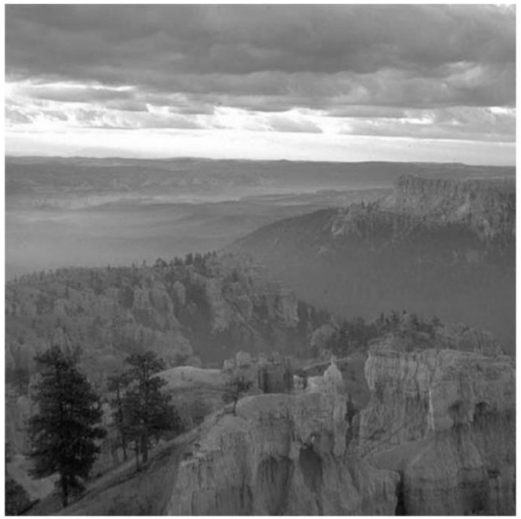
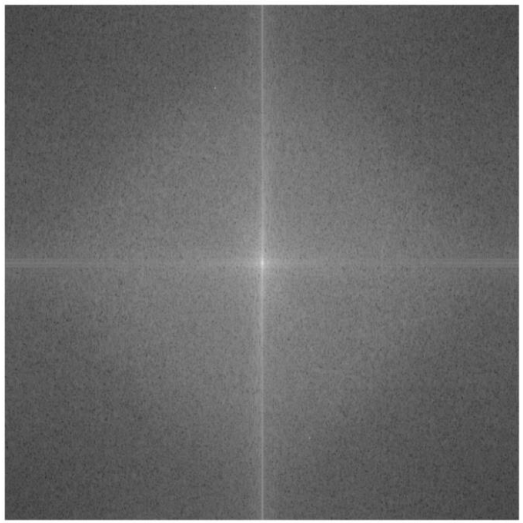


Immagine Originale



Applicazione
della DFT

Possibili Attacchi

- Partendo dall'immagine marcata è possibile **effettuare diversi attacchi**, al fine di **rendere il watermark non rilevabile** o **solo parzialmente rilevabile**
- I principali attacchi sono
  - Compressione Lossy
  - Distorsioni Geometriche
  - Operazioni di Signal Processing
  - Rewatermarking

Possibili Attacchi

- Partendo dall'immagine marcata è possibile **effettuare diversi attacchi**, al fine di **rendere il watermark non rilevabile** o **solo parzialmente rilevabile**
- I principali attacchi sono
  - **Compressione Lossy**
  - Distorsioni Geometriche
  - Operazioni di Signal Processing
  - Rewatermarking

Possibili Attacchi **Compressione Lossy**

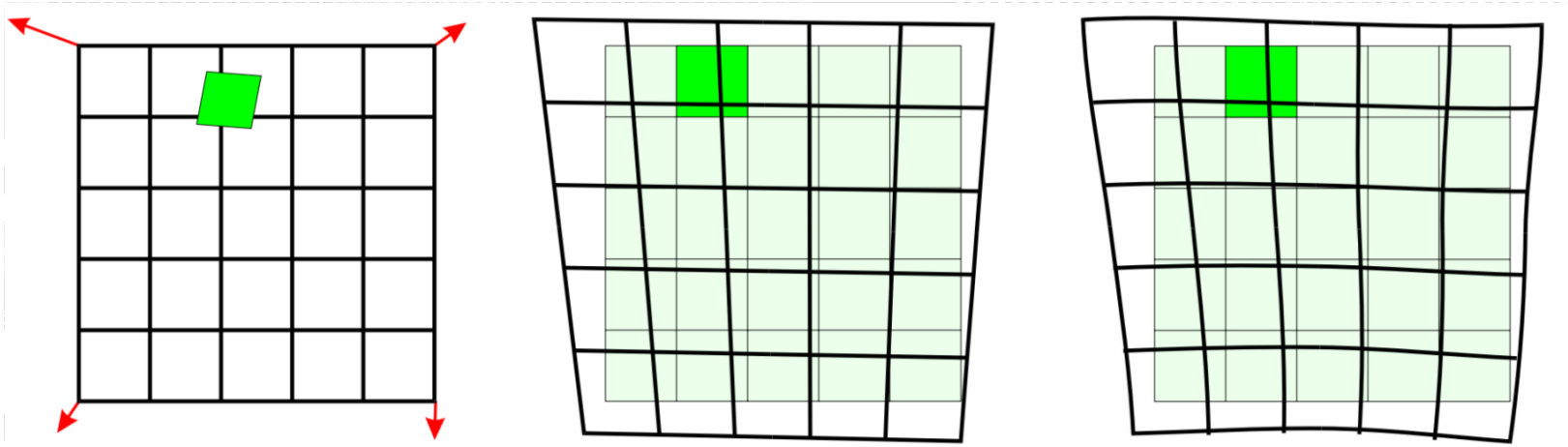
- Molti schemi per la compressione lossy (JPEG, JPEG2000, etc.) eliminano le informazioni meno percepibili dall'uomo
  - Allo scopo di ridurre le dimensioni dell'immagine
- Solitamente il watermark viene nascosto proprio nelle informazioni meno percepibili
  - La perdita di tali informazioni può alterare o far perdere il watermark
  - Gli schemi di watermarking nel dominio delle frequenze sono solitamente robusti contro attacchi basati sulla compressione lossy

Possibili Attacchi

- Partendo dall'immagine marcata è possibile **effettuare diversi attacchi**, al fine di **rendere il watermark non rilevabile** o **solo parzialmente rilevabile**
- I principali attacchi sono
  - Compressione Lossy
  - **Distorsioni Geometriche**
  - Operazioni di Signal Processing
  - Rewatermarking

Possibili Attacchi
Distorsioni Geometriche

- Un altro possibile attacco agli schemi di watermarking è rappresentato dalle **distorsioni geometriche**
- Tali distorsioni possono essere di diverso tipo
  - Ritaglio (Crop)
  - Rotazione
  - Shifting
  - Allungamento (Stretching)



distorsioni geometriche

- Tali distorsioni possono essere di diverso tipo
  - Ritaglio (Crop)
  - Rotazione
  - Shifting
  - Allungamento (Stretching)

Possibili Attacchi

- Partendo dall'immagine marcata è possibile **effettuare diversi attacchi**, al fine di **rendere il watermark non rilevabile** o **solo parzialmente rilevabile**
- I principali attacchi sono
  - Compressione Lossy
  - Distorsioni Geometriche
  - **Operazioni di Signal Processing**
  - Rewatermarking

Possibili Attacchi
Operazioni di Signal Processing

- Attacchi basati sulle operazioni di signal processing
  - Aggiunta di un offset costante al valore dei pixel
  - Aggiunta di rumore all'immagine
  - Aggiunta di disturbo all'immagine
  - Applicazione di filtri
  - Etc.
- Un altro attacco è basato sulla conversione Digitale/Analogico (D/A) → Analogico/Digitale (A/D)
  - Viene effettuata una stampa dell'immagine (conversione D/A) e poi una digitalizzazione di tale stampa, mediante l'acquisizione da scanner (conversione A/D)
  - Nel processo fra la conversione D/A → A/D vengono perse informazioni (e potenzialmente anche il watermark)

Possibili Attacchi

- Partendo dall'immagine marcata è possibile **effettuare diversi attacchi**, al fine di **rendere il watermark non rilevabile** o **solo parzialmente rilevabile**
- I principali attacchi sono
  - Compressione Lossy
  - Distorsioni Geometriche
  - Operazioni di Signal Processing
  - **Rewatermarking**

Possibili Attacchi **Rewatermarking**

- Con il termine rewatermarking si intende l'**applicazione di un ulteriore watermark** ad un'immagine già marcata
- Gli obiettivi principali di tale attacco sono
  - «Sovrascrivere» il watermark precedentemente inserito
  - Ingannare l'algoritmo di estrazione e/o rilevazione

Watermarking di Dati Multimediali

- Watermarking su
  - Immagini

  - Audio

Watermarking di Audio

- Le tecniche di watermarking che operano su audio digitale si basano principalmente sulla caratteristica del mascheramento uditivo dell'apparato uditivo umano

- **Mascheramento Uditivo**
  - Un segnale audio debolmente percettibile, **diviene non percettibile** in presenza di un forte segnale audio di frequenza leggermente superiore

Watermarking di Audio

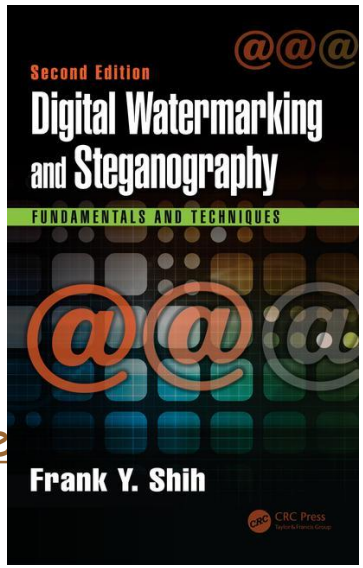
- **Ad esempio**, assumendo la presenza di un segnale principale, a una data frequenza, e di un segnale secondario, di livello più basso e ad una frequenza leggermente inferiore, quest'ultimo segnale è "mascherato" dal primo e l'orecchio umano non è in grado di percepirlo
- È possibile effettuare l'embedding di un watermark operando nel dominio
  - Temporale
  - Delle frequenze

Watermarking di Audio
Possibili Attacchi

- Attacchi in ambiente digitale
  - Compressione Lossy
  - Audio Processing
    - Denoising, Equalizzazione, Ricampionamento, etc.
  - Ritaglio (Cropping)
- Attacchi in ambiente analogico
  - Conversione da D/A → A/D
  - Rumore di sottofondo
  - Modifica dell'ampiezza del segnale

Bibliografia

- Digital Watermarking and Steganography: Fundamentals and Techniques (Second Edition), Frank Y. Shih, 2017, CRC Press.
  - Capitolo 1: da pagina 1 a pagina 6
  - Capitolo 2: da pagina 9 a pagina 14
  - Capitolo 3: da pagina 15 a pagina 20 (Solo definizioni di Least-Significant-Bit Substitution, Discrete Fourier Transform, Discrete Cosine Transform, Discrete Wavelet Transform)
  - Capitolo 4: da pagina 35 a pagina 41 (Sezione 4.2.2 esclusa)
- http://printer-blogarchive.yale.edu/blog/2012/09/14/archives-yale-shie dandy-roll



Ulteriori Riferimenti Bibliografici

- Cox, Ingemar J., et al. Digital watermarking. Vol. 1558607145. San Francisco: Morgan Kaufmann, 2002.
- Cox, Ingemar, et al. Digital watermarking and steganography. Morgan Kaufmann, 2007.
- Shih, Frank Y. Digital watermarking and steganography: fundamentals and techniques. CRC Press, 2017.
- Seitz, Juergen. Digital watermarking for digital media. IGI Global, 2005.
- Katzenbeisser, Stefan, and Fabien Petitcolas. Information hiding techniques for steganography and digital watermarking. Artech house, 2000.
- Arnold, Michael Konrad, Martin Schmucker, and Stephen D. Wolthusen. Techniques and applications of digital watermarking and content protection. Artech House, 2002.
- Mintzer, Fred, Gordon W. Braudaway, and Alan E. Bell. "Opportunities for watermarking standards." Communications of the ACM 41.7 (1998): 57-64.
- Cox, Ingemar J., et al. "Secure spread spectrum watermarking for multimedia." IEEE transactions on image processing 6.12 (1997): 1673-1687.
- Dugad, R., Ratakonda, K., Ahuja, N., A new wavelet-based scheme for watermarking images. Proceedings of IEEE ICIP 1998 (Vol. 2, pp. 419-423)