Information Hiding (Streganografia e Watermarking)

Indice

- Introduzione a Strumenti per l'Information Hiding
- Steganografia
  - Un po' di storia
  - Modelli steganografici
  - Immagini come contenitori di informazioni
  - Steganalisi
- Watermarking
  - Introduzione e applicazioni
  - Caratteristiche
  - Tecniche di watermarking e attacchi

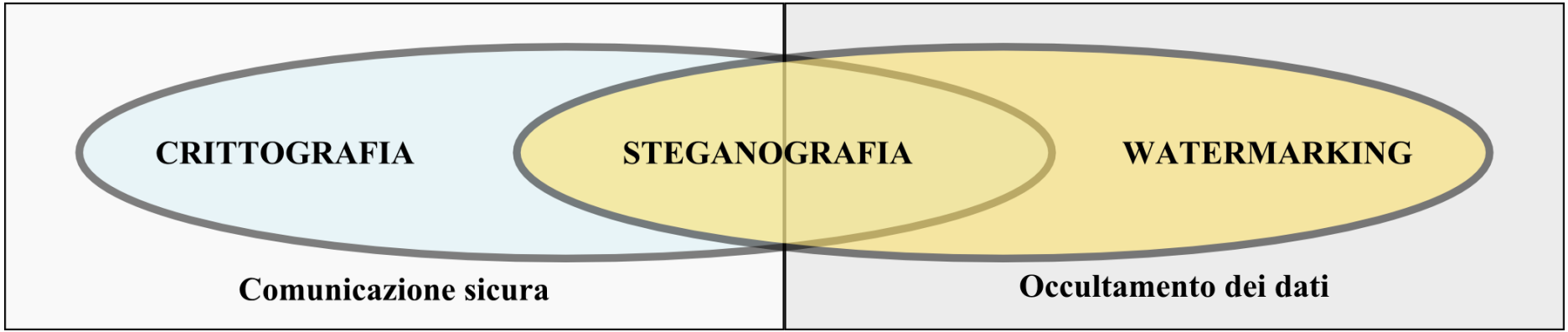
Riservatezza dell'Informazione

- La circolazione e la condivisione delle informazioni è importante così come il **desiderio** o la **necessità** di mantenere alcune informazioni **riservate**
- L'approccio più utilizzato per rendere **privata** la conversazione è quello di rendere il messaggio **incomprensibile** a chi non è il destinatario
- L'avvento del digitale ha portato anche al settore della riservatezza delle informazioni nuovi paradigmi di implementazione di teorie e tecniche già note.

Tre tecniche

- **Crittografia:** si basa sulla codifica dei messaggi mediante appositi algoritmi di cifratura che lo rendono incomprensibile a chi non è a conoscenza dei relativi sistemi di decodifica.
- **Steganografia:** nasconde l'esistenza stessa del messaggio, includendolo in un mezzo "neutrale" e garantendo quindi la segretezza della comunicazione stessa.
- **Watermarking:** (filigranatura) inserisce opportune informazioni (spesso nascoste) in testi, immagini o video, per segnalarne l'originalità o il proprietario.

Sicurezza e Occultamento



Crittografia

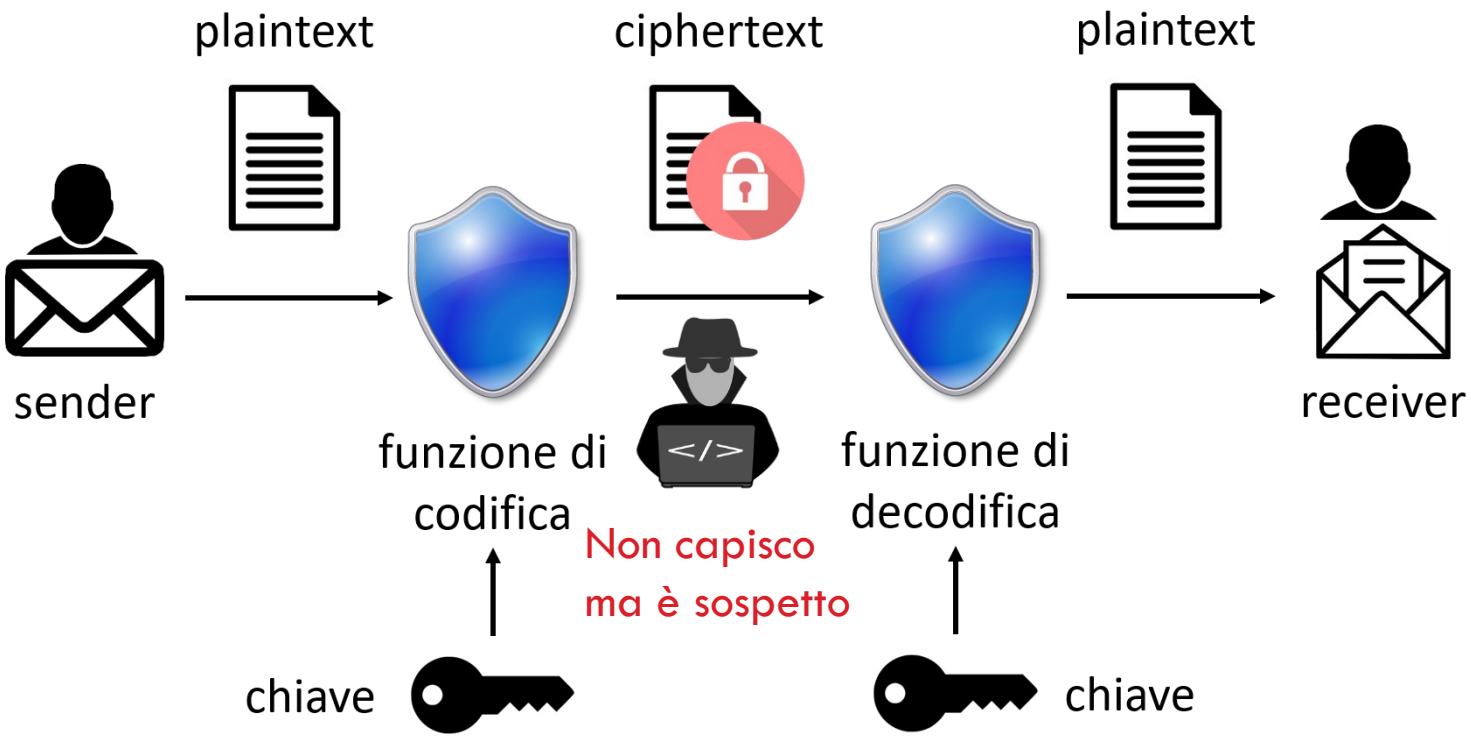
- La crittografia è la scienza che studia le tecniche e le metodologie per cifrare (codificare) un testo in chiaro al fine di produrre un testo cifrato comprensibile solo ad un destinatario legittimo

- Il ricevente deve possedere l'informazione sufficiente (*chiave*) per decifrare il testo cifrato, recuperando così il testo in chiaro.

La crittografia ha lo scopo di nascondere il **contenuto** di un messaggio, ma cambiarsi testi cifrati (incomprensibili) genera sospetti.

Fallisce quando il contenuto del ciphertext trasmesso viene decrittato

Crittografia



Steganografia

- La steganografia è l’arte di nascondere un messaggio (che si vuole rimanga segreto) all'interno di un altro messaggio contenitore che si vuole sia pubblico e non sospetto
- Il contenitore può anche avere un aspetto totalmente diverso dal messaggio segreto e deve essere in grado di nascondere la stessa esistenza della comunicazione.

La steganografia fallisce se il tentativo di trasmettere materiale confidenziale viene scoperto, anche se non si è in grado di decrittarne il contenuto

Sistema Steganografico

Watermarking

- Il *watermarking* (letteralmente filigranatura) è usato per inserire opportune informazioni in un segnale, in particolare su file multimediali, eventualmente per segnalarne l'originalità o per indicare il titolare dei diritti di proprietà.
- Un *watermark*, proprio come le filigrane delle banconote, deve essere visibile solo in certe condizioni – per esempio dopo l'applicazione di opportuni algoritmi.

Steganografia nella Storia

- Nelle sue Storie, Erodoto racconta che
  - Istieo, intorno al 440 a.C., rasò la testa del suo schiavo più fidato e la tatuò con un messaggio per Aristagora di Mileto che scomparì appena gli furono ricresciuti i capelli.
  - Demerato, un Greco alla corte del re di Persia Serse:
    - Avvisò Sparta di un imminente attacco, rimuovendo la cera da una coppia di tavolette di legno, incidendo un messaggio e successivamente ricoprendolo con la cera.
    - Quando il messaggio sotto la cera giunse a destinazione, nessuno immaginò la sua esistenza, finché Gorgo, moglie di Leonida, ebbe una premonizione e fece grattare via la cera.

Inchiostri Invisibili

- Gli antichi romani usavano scrivere fra le righe di un testo utilizzando un inchiostro fatto con sostanze naturali come il succo di limone, l'aceto o il latte. Il messaggio nascosto diventava visibile una volta che il testo veniva avvicinato ad una fonte di calore.
- Lo scienziato Giambattista Della Porta (XVI secolo) spiegò come comunicare tramite un uovo sodo, preparando un inchiostro con 30 grammi di allume in mezzo litro d'aceto, e usandolo per scrivere sul guscio.

Microdot

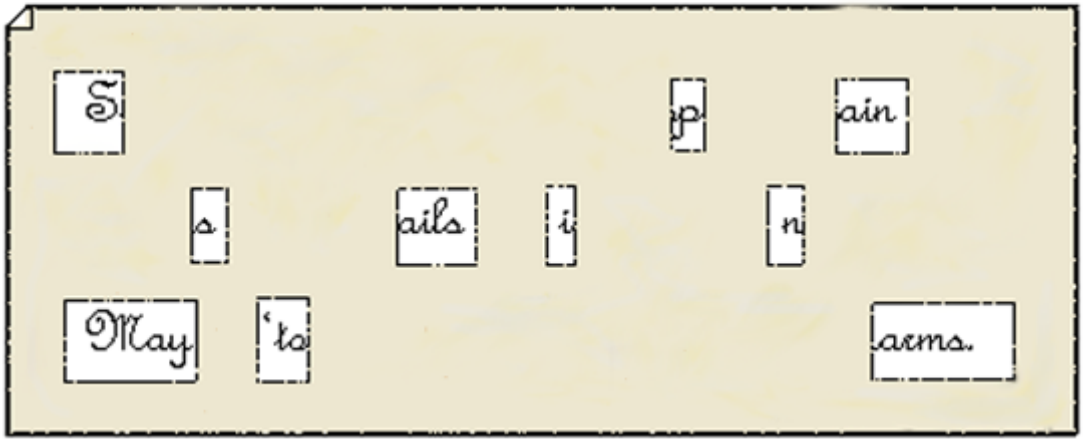
- Fotografie della dimensione di un punto dattiloscritto che, una volta sviluppate e ingrandite, possono diventare pagine stampate di buona qualità.
- Il primo microdot fu scoperto dall'FBI solo nel 1941, grazie a una soffiata

Griglie di Cardano

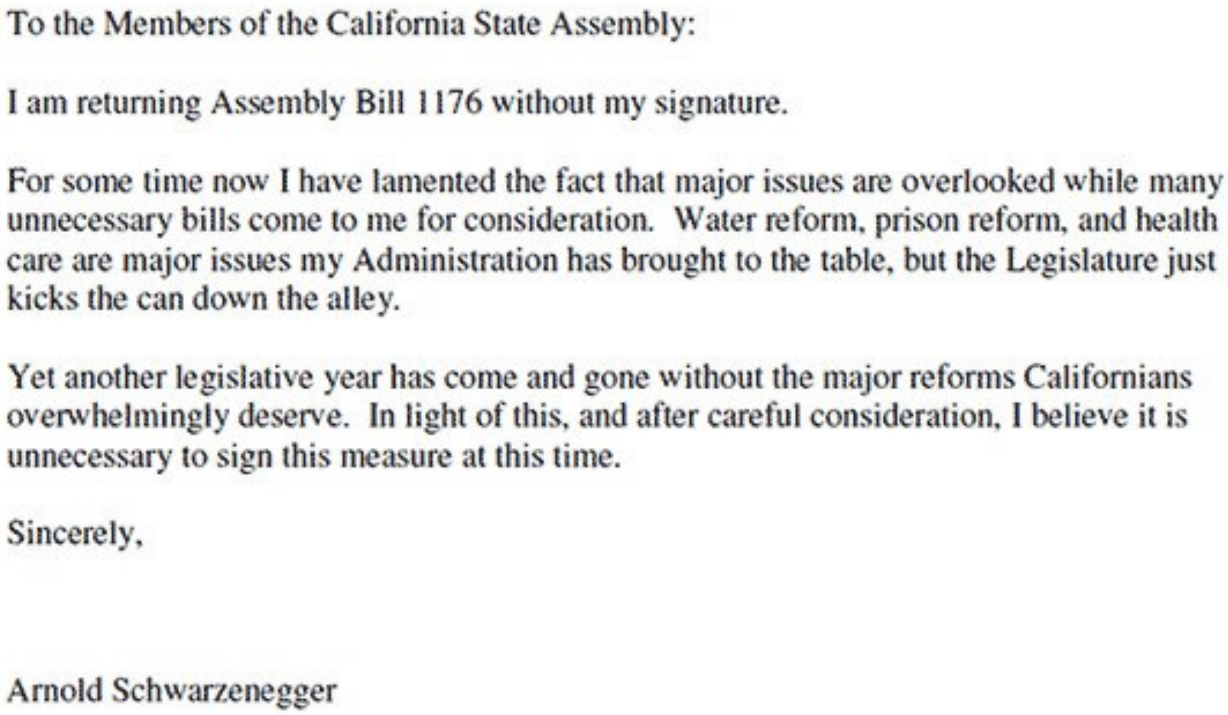
- Le griglie di Cardano (1501 – 1626) erano fogli di materiale rigido nei quali venivano ritagliati fori rettangolari a intervalli irregolari.
- Il messaggio segreto veniva scritto nei buchi (ciascun buco poteva contenere una o più lettere), dopodiché si toglieva la griglia e si cercava di completare la scrittura del resto del foglio in modo da ottenere un messaggio di senso compiuto, il quale poi veniva inviato a destinazione.
- Applicando sul foglio una copia esatta della griglia originaria, era possibile leggere il messaggio nascosto

Esempio di Griglia di Cardano

Sir John regards you well and spekes again that
all as rightly 'vails him is yours now and ever.
May he 'tone for past d'lays with many charms.



Acrostico



To the Members of the California State Assembly:

I am returning Assembly Bill 1176 without my signature.

For some time now I have lamented the fact that major issues are overlooked while many unnecessary bills come to me for consideration. Water reform, prison reform, and health care are major issues my Administration has brought to the table, but the Legislature just kicks the can down the alley.

Yet another legislative year has come and gone without the major reforms Californians overwhelmingly deserve. In light of this, and after careful consideration, I believe it is unnecessary to sign this measure at this time.

Sincerely,

Arnold Schwarzenegger

L'acrostico (dal greco ákros, «estremo» e stíchos, «verso») è un componimento poetico all'interno del quale le sillabe o le lettere iniziali di ciascun verso formano in verticale una parola, un nome, una frase.

Cifrari Nulli

- Un cifrario nullo (null cypher) è un cifrario di occultamento in cui il testo in chiaro è mescolato con una grande quantità di materiale non cifrato.

- Esempio: Mettendo insieme la prima lettera di una parola su tre del seguente testo si ottine"Wikipedia" come messaggio nascosto:

It's important **w**e allow anyone **i**nterested in gaining **k**nowledge accesst o **i**nformation which is **p**ublished freely. There **e**xists a website **d**evoted to this **i**dea, and you **a**re on it!

Steganalisi

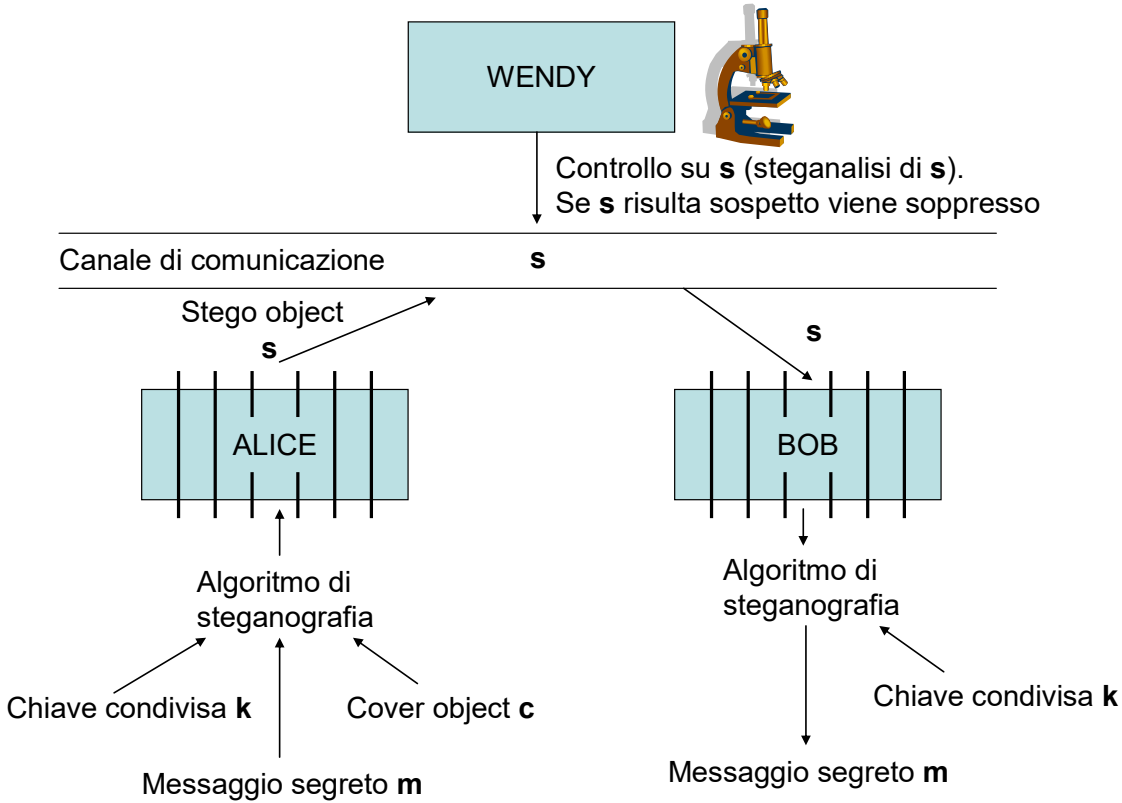
- La steganalisi è l'insieme dei metodi e delle tecniche capaci di attaccare un sistema steganografico.

- Così come la crittoanalisi ha il compito di svelare l'output cifrato di un sistema crittografico, lo scopo della steganalisi è quello di etichettare con certezza un oggetto come sospetto (contenente, cioè, al suo interno un messaggio occulto) o come innocuo (privo di informazioni nascoste).

"Se il fine della crittoanalisi è quello di rivelare il dato, quello della steganalisi il fine è quello di scoprirne la presenza"

Il problema dei prigionieri

La moderna formulazione degli studi steganografici si deve al lavoro di G.J. Simmons che nel 1984 propose il *problema dei prigionieri*



Steganalisi

- Wendy, oltre a potere esaminare tutti i messaggi che i due prigionieri si cambiano, può assumere un ruolo attivo oppure passivo.
  - Ruolo passivo se si limita a esaminare il messaggio che i due prigionieri si scambiano.
  - Ruolo attivo se può alterare i messaggi scambiati

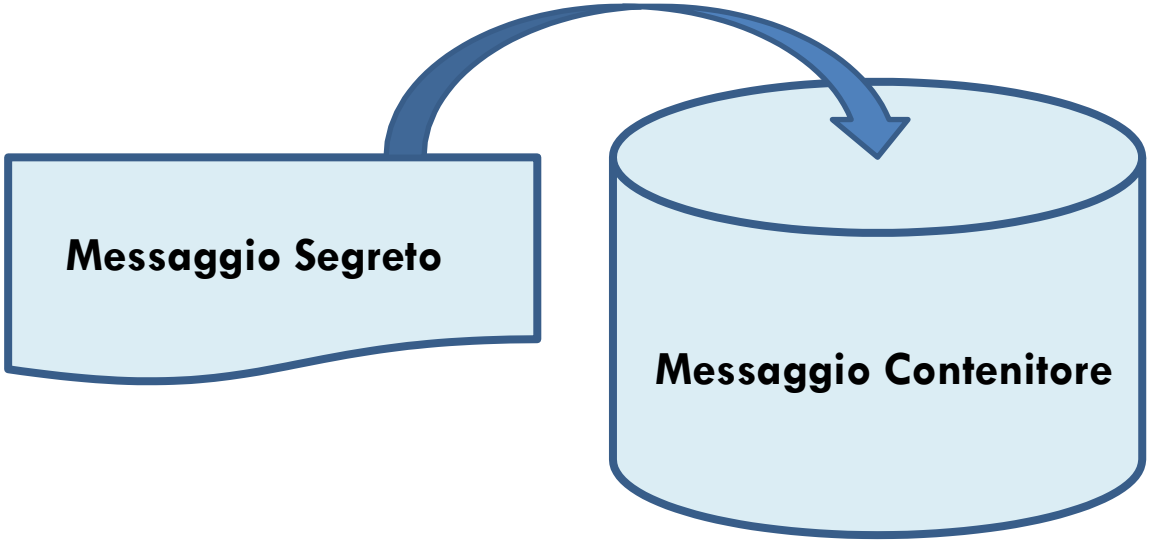
Modelli Steganografici: Contenitori

- Lo schema di base della steganografia presuppone l'esistenza di due messaggi:
  - Messaggio *segreto*
  - Messaggio *contenitore*
- In base a come si gestisce il contenitore si parla di:
  - Steganografia *iniettiva*
  - Steganografia *generativa*

Steganografia iniettiva

- La steganografia *iniettiva* (la più utilizzata) consente di inserire il messaggio segreto in un messaggio contenitore esistente modificandolo in modo tale da contenere il messaggio segreto, e risultare, al livello al quale viene percepito dai sensi umani, praticamente indistinguibile dall'originale.

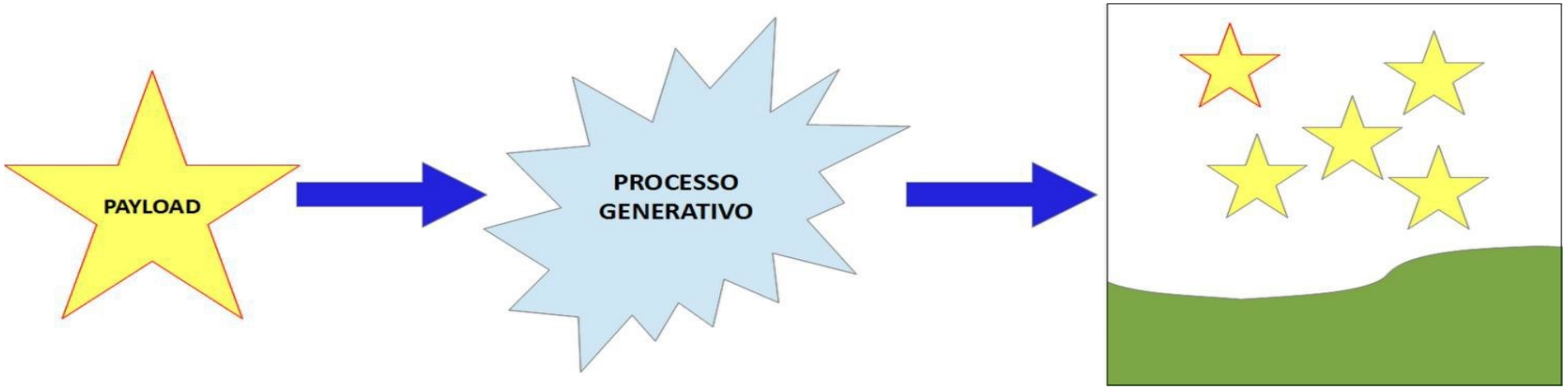
Steganografia Iniettiva



Steganografia generativa

- La steganografia *generativa* consente di generare, a partire dal messaggio segreto, un messaggio contenitore atto a nascondere nel migliore dei modi quel messaggio segreto.

Steganografia Generativa



Modelli Steganografici

- L'immagine designata a contenere il messaggio conviene detta immagine **cover** (contenitore**)**.
- Il messaggio viene detto **payload** (carico)
- Il risultato dell'inserzione del paylod nella cover image, viene chiamato **stego-image** (immagine stego).

```
Stego-image = F (G(cover), H(Payload))
```

dove*:*

- *F* è la funzione steganografica
- *G* è una funzione che elabora l'immagine **cover**
- *H* è una funzione che elabora il messaggio da inserire, potrebbe essere una funzione criptografica o anche la funzione identità.

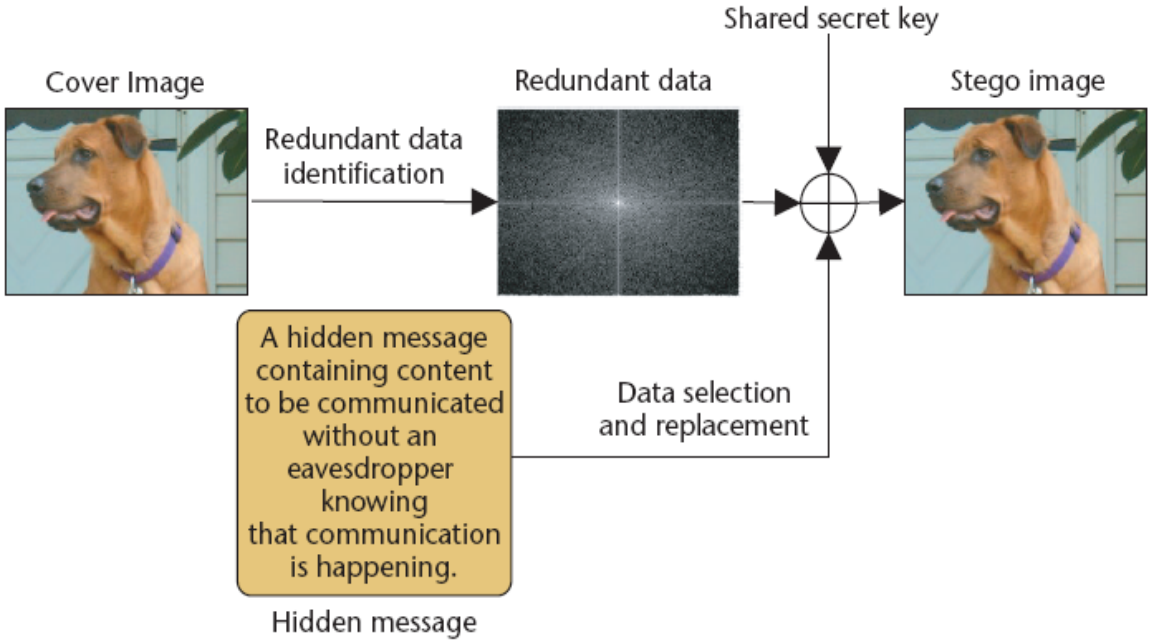
Modelli Steganografici: Tecniche

- Una classificazione alternativa a quella basata sui contenitori, si basa invece sulle tecniche utilizzate per l'aggiunta di informazione nel contenitore:
  - steganografia sostitutiva
  - steganografia selettiva
  - steganografia costruttiva

Steganografia sostitutiva

- L’obiettivo è sostituire un elemento di scarsa importanza del cover medium con un elemento del messaggio segreto che si vuole nascondere.
- Le tecniche sostitutive sono le più utilizzate e hanno una diffusione tale che, spesso, con il termine steganografia si fa implicito riferimento a esse.
- Il principale difetto di queste tecniche risiede nella possibile alterazione delle caratteristiche statistiche del rumore presente nel contenitore.

Steganografia sostitutiva



N. Provos and P. Honeyman, *Hide and Seek: An Introduction to Steganography*, IEEE Security & Privacy, 2003.

Steganografia selettiva

- L'idea di base è di scegliere il supporto a seconda del messaggio da occultare procedendo per tentativi, ripetendo una stessa procedura fintanto che il risultato non soddisfi una certa condizione.
- Il mezzo sopravvissuto al processo di selezione contiene effettivamente l'informazione segreta, ma è un contenitore "naturale".
- Il problema di questa tecnica è che è molto dispendiosa rispetto alla quantità di informazione che è possibile nascondere.

Steganografia selettiva

Secret Bits	1110011
Cover Text	Česká republika jekrásné místo
Steganographic Text	Česká republik a jekrásné místo 1 1 1 0 0 1 1

S. Khan et al. “Czech Text Steganography Method by Selective Hiding Technique” Proceedings of the World Congress on Engineering 2015 Vol I WCE 2015, July 1-3, 2015, London, U.K.

Il bit "zero" è memorizzato nelle lettere non accentate e il bit "uno" è memorizzato in quelle accentate (con annotazione aggiuntiva nella rappresentazione ASCII, usando caratteri di estensione non visibili al lettore ma solo al computer).

Steganografia costruttiva

- L’obiettivo è sostituire il rumore presente nel medium utilizzato con l'informazione segreta opportunamente modificata in modo da imitare le caratteristiche statistiche del rumore originale.
- Non è facile costruire un modello del rumore ed è possibile che qualcuno, in grado di disporre di migliori risorse, riesca a costruire un modello più accurato e a distinguere tra il rumore originale e il suo sostituto.
- Se il modello del rumore utilizzato dovesse cadere nelle mani dell’avversario, questi potrebbe utilizzarlo per controllare che un messaggio sia conforme ad esso.

Scelta del cover

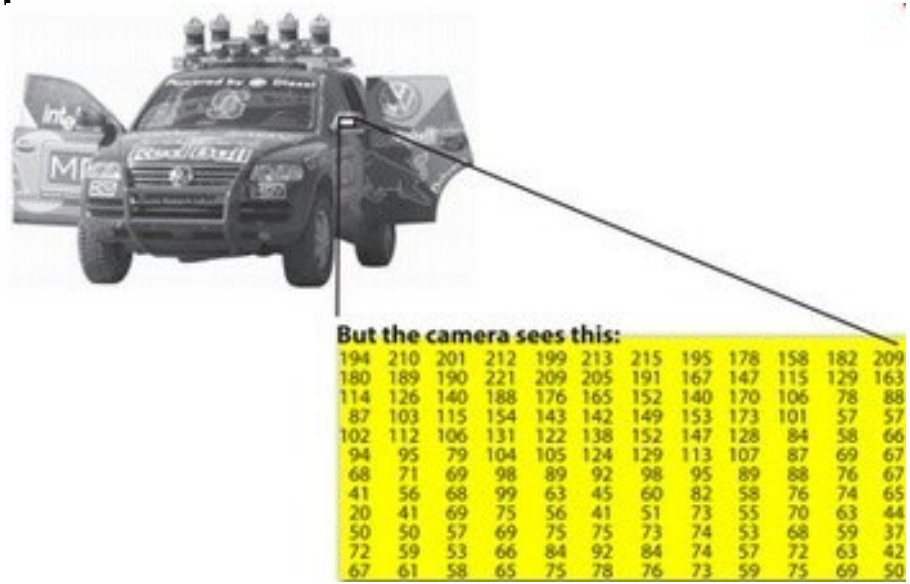
- La steganografia ha la necessità di utilizzare cover object che siano possibilmente inediti e mai usati in precedenza:
- Non c’è niente di più facile che scattare una foto, per ottenere un cover medium inedito
- Le cover più utilizzate sono le immagini
- La tecnica più utilizza è quella del Least Significant Bit (LSB)

Esempi di cover

- I dati multimediali (audio e video) sono eccellenti contenitori: infatti, a seguito della digitalizzazione, essi contengono del rumore di quantizzazione che fornisce lo spazio di manovra necessario all'inserimento dei dati
- La compressione con perdita (lossy) è in grado di introdurre ulteriori quantità di rumore.
- La tecnica di steganografia iniettiva su immagini sicuramente più diffusa è quella che si basa sulla modifica del bit meno significativo (LSB).

Immagini come cover

- Una immagine digitale è una matrice di pixel
- Il termine pixel deriva da picture element
- Il pixel contiene l'informazione relativa alla rappresentazione della realtà che è stata catturata tramite uno scanner, una macchina)

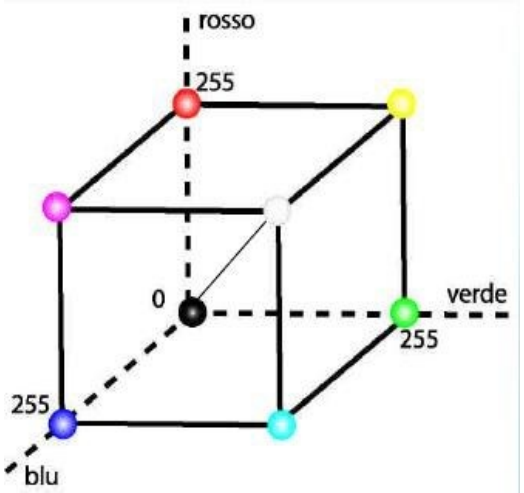


Immagini come cover

Le immagini digitali a colori sono array di valori

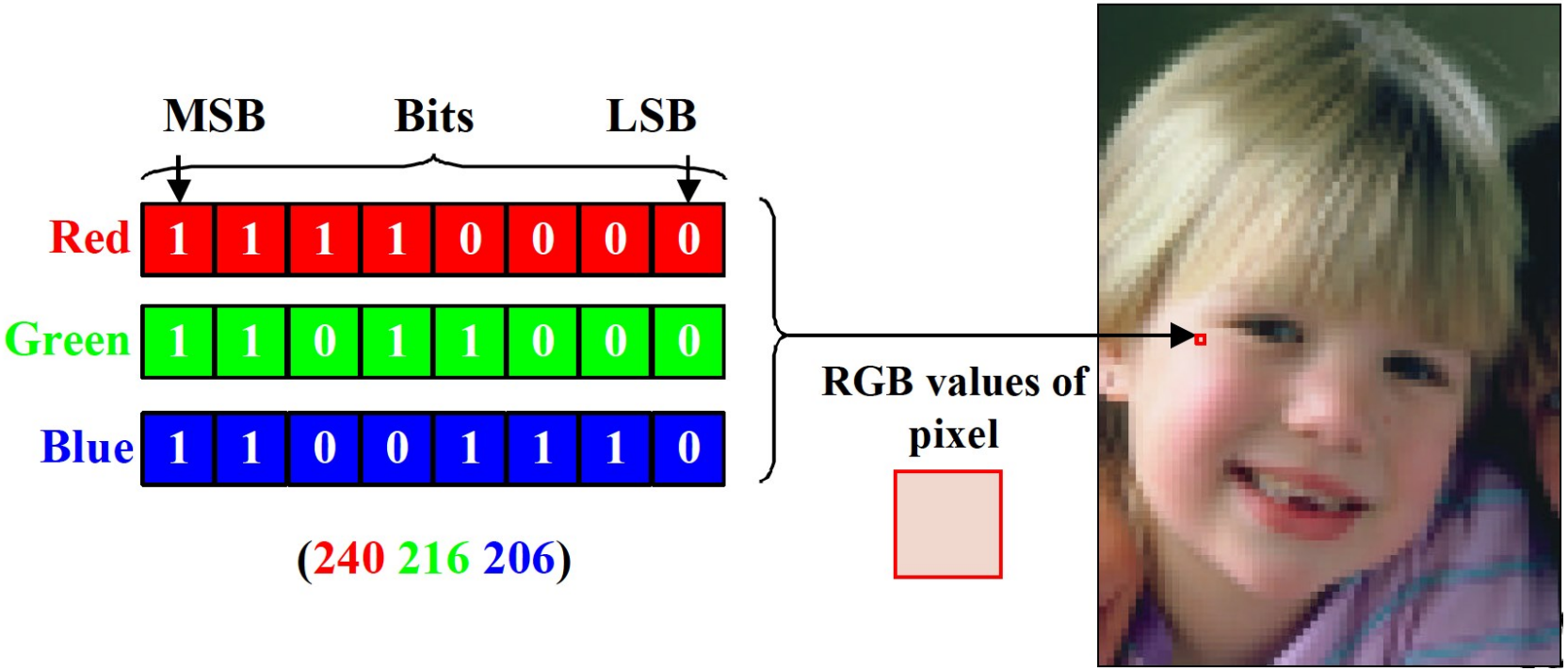
di pixel nel dominio spaziale (Molti diversi formati)

Nei personal computer:

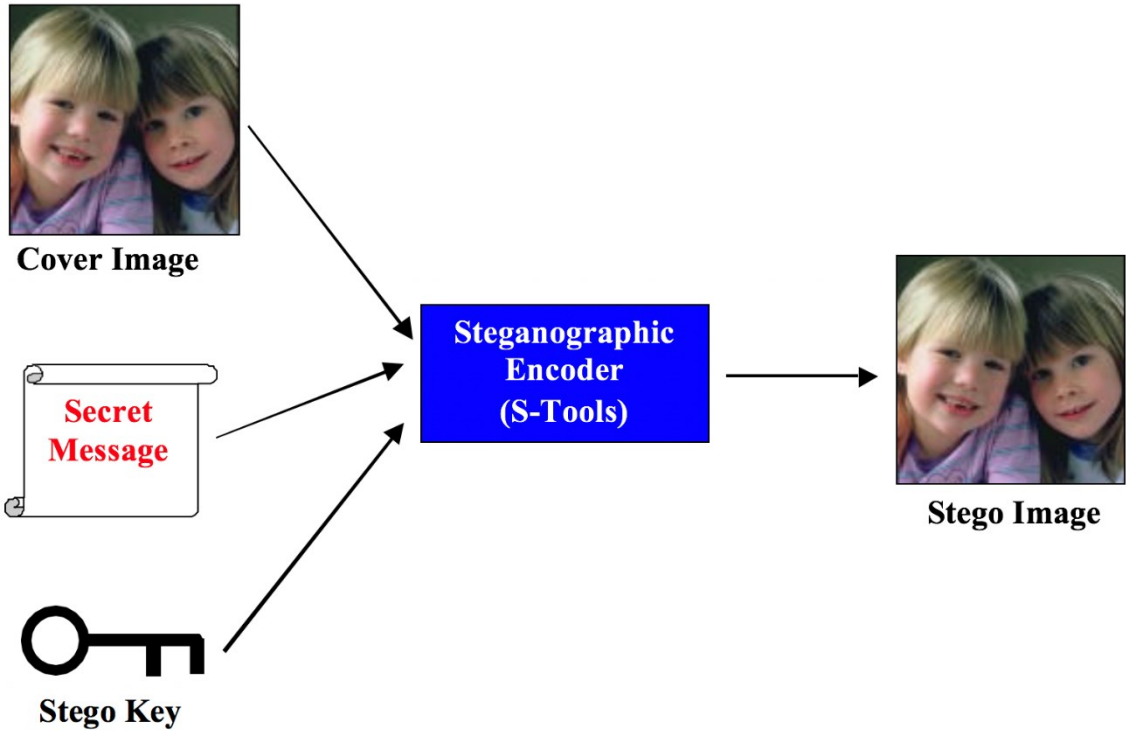


- Lo spazio di colore comune è RGB (Red, Green, Blue).
- Nelle immagini con 24 bit/pixel viene assegnato un colore ad ogni posizione di pixel
- Ogni componente di colore ha un valore di 8 bit compreso tra 0 e 255
- (0,0,0) corrisponde al nero più scuro;
- (255,255,255) corrisponde al bianco più chiaro.

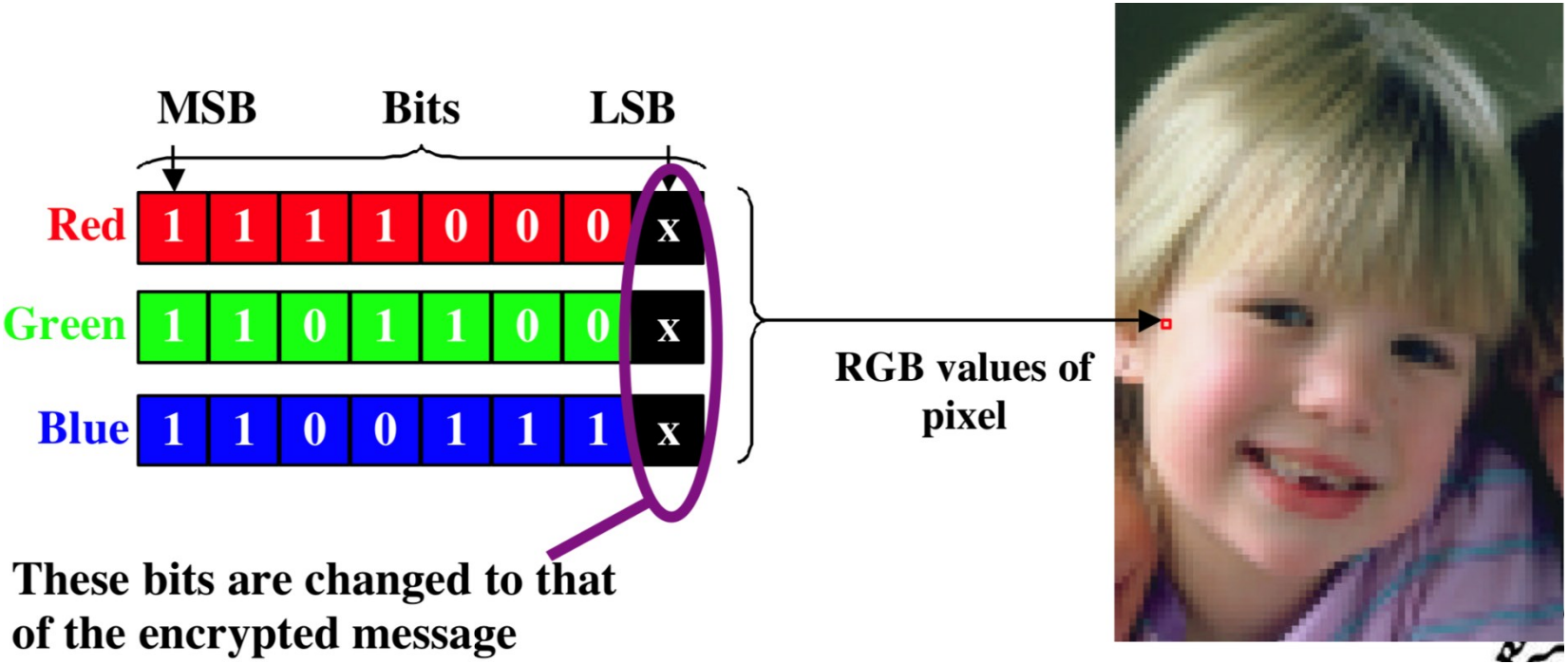
RGB 24 bit/pixel



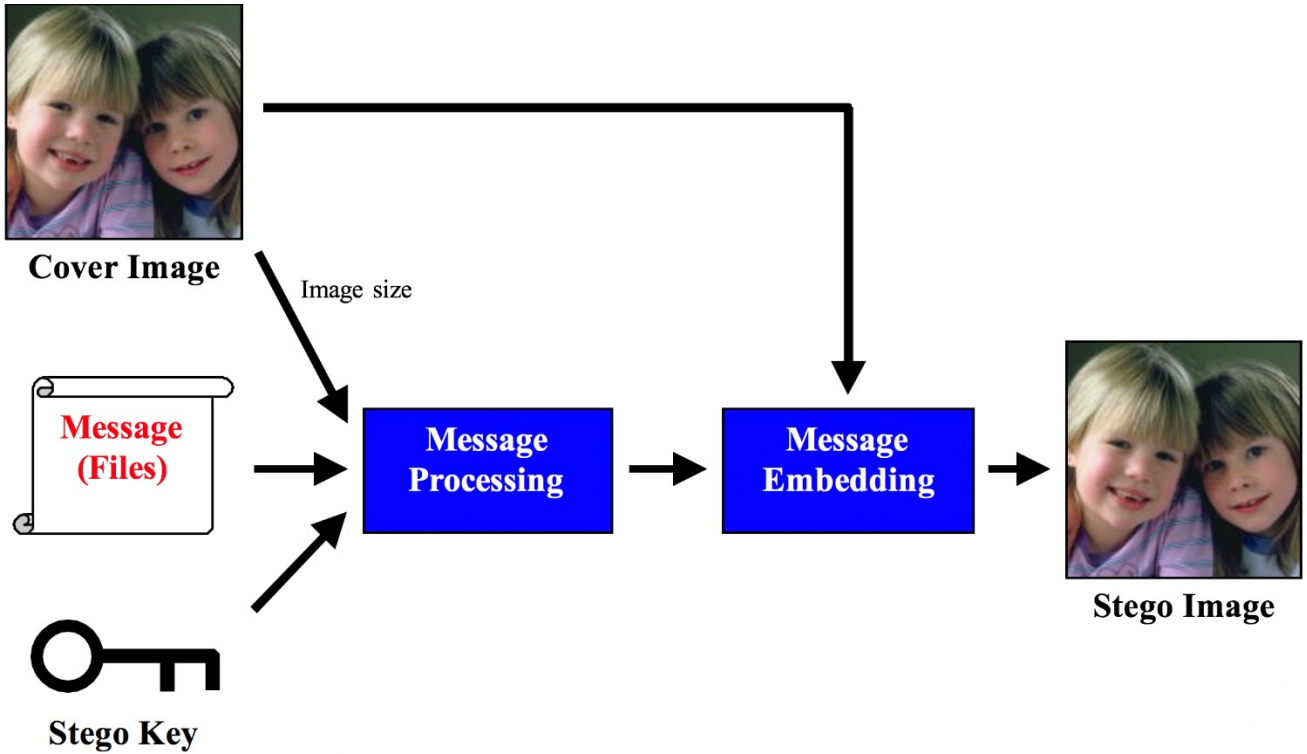
Cifratura Steganografica



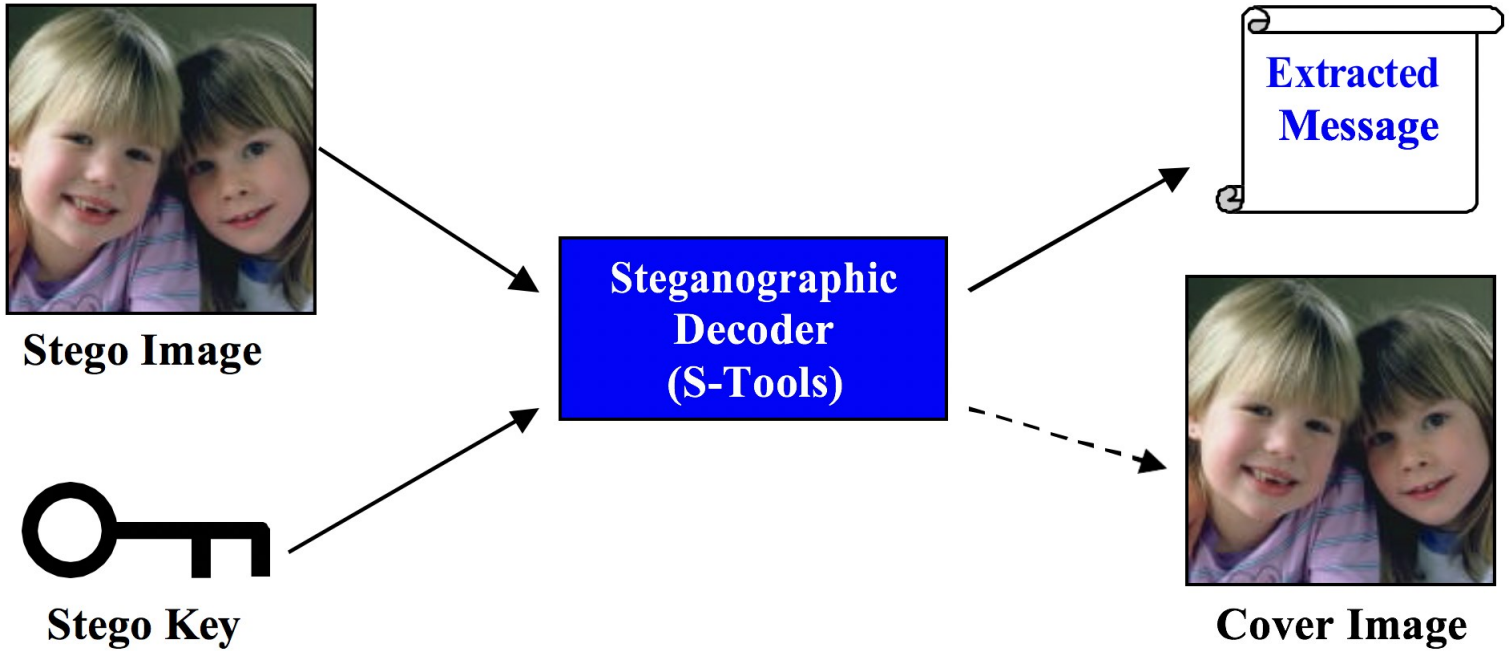
Cifratura Sostitutiva



Cifratura Stenografica con Crittografia



Decifratura Stenografica



Steganalisi

Esistono tools che consentono di individuare, estrarre e/o distruggere messaggi nascosti all'interno di cover sospetti.

- 2Mosaic, StirMark Benchmark: rimuovono messaggi da immagini.
  - http://www.petitcolas.net/fabien/watermarking/2mosaic/index.html
  - http://www.petitcolas.net/fabien/watermarking/stirmark/
- StegDetect: attacca file steganografati utilizzando gli attacchi statistici.
  - http://www.outguess.org/detection.php
- StegBreak: individua ed estrae messaggi segreti inseriti da vari software.
  - http://manpages.ubuntu.com/manpages/hardy/man1/stegbreak.1.html

Watermarking Digitale

- Insieme di tecniche e metodi per l'inclusione di informazioni in un file multimediale, che possono essere poi rilevate o estratte per avere informazioni sulla sua origine e/o provenienza.

Watermarking Digitale

- I watermark, possono essere evidenti per l'utente del file o latenti (nascoste all'interno del file)
- Quando il watermark è nascosto, il watermarking è una forma particolare di steganografia
- Un documento con un watermark è ancora accessibile, ma marcato in modo permanente.

Crittografia vs Watermarking

Crittografia

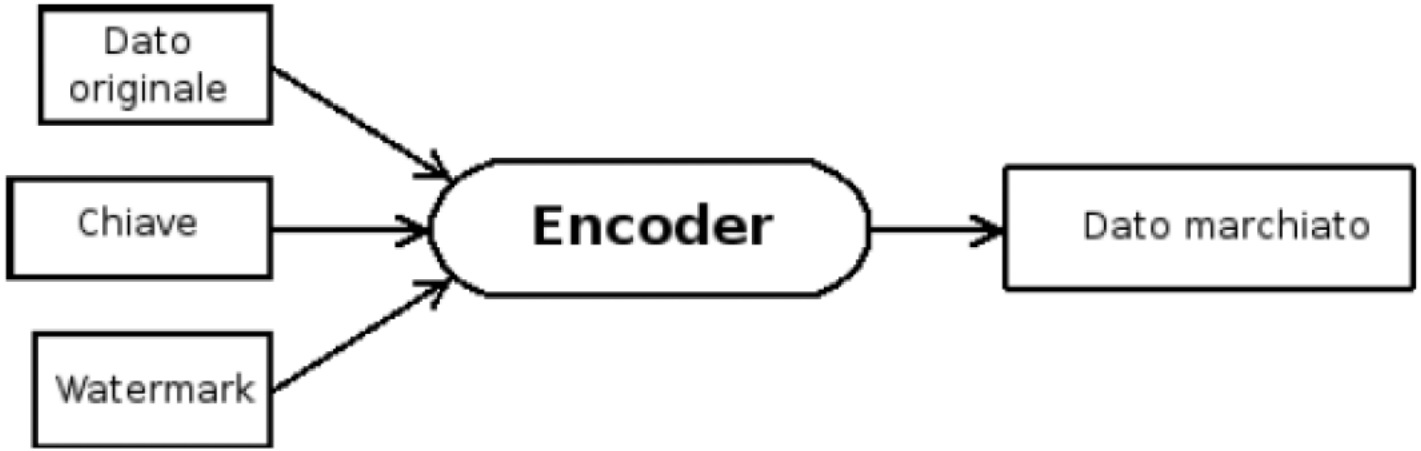
- implica la trasformazione dei documenti in modo che il loro contenuto non sia visibile senza una chiave per decifrare.

Watermarking

- lascia il file originale (quasi) intatto e riconoscibile

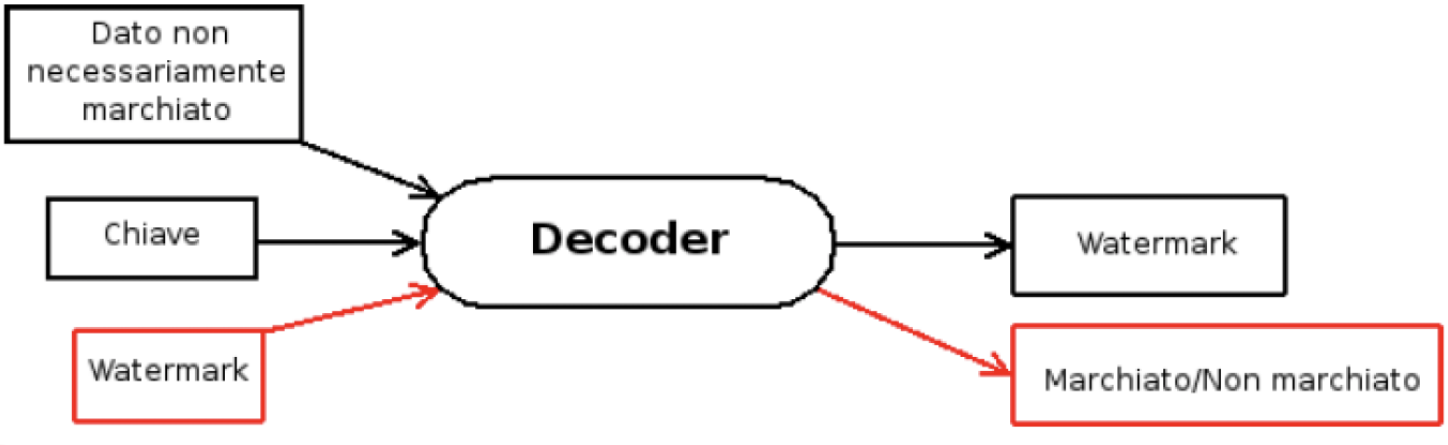
Esempi di uso

Inserimento del Watermark



Esempi di uso

Estrazione e Rilevamento
del Watermark



Obiettivi del Watermarking

- Rendere manifesto a tutti gli utenti chi sia il legittimo proprietario del documento (nel caso in cui il marchio sia visibile);
- Dimostrare l'originalità di un documento non contraffatto;
- Evitare la distribuzione di copie non autorizzate;
- Marcare alcune caratteristiche specifiche del documento;
- Segnare il percorso di vendita del documento, utilizzando un marchio differente per ciascun acquirente.

Caratteristiche

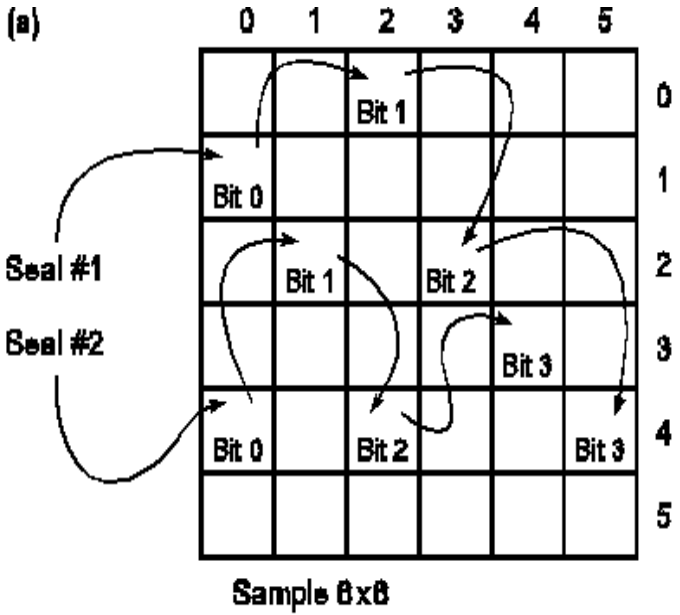
- Poco invadente: abbastanza invisibile da non degradare la qualità dei dati e da impedire che venga trovato e cancellato.
- Sufficientemente Visibile: visibile da scoraggiare il furto.
- Facilmente rilevabile: il proprietario dei dati o un'autorità di controllo indipendente deve poterlo individuare facilmente.
- Univoco: il suo recupero dovrebbe identificare senza ambiguità il proprietario dei dati.
- Molteplice: deve essere possibile generare diversi watermark
- Robusto: difficile da rimuovere da chi volesse contraffare il copyright dei dati.

Watermarking di Immagini

La posizioni di alcuni pixel e un checksum sono il watermark:

- Si scelgono i 7 bit più significativi di 8 diversi pixel.
- I segmenti vengono concatenati e si ottiene un checksum di 56 bit.
- I bit del checksum vengono scritti nell'ultimo bit di 56 pixel scelti in modo casuale.

Watermarking di immagini: un esempio



(b)

0	1	2	3	4	5	
55	73	71 71	123	123	205	0
120 121	123	70	72	147	199	1
130	123 123	67	66 66	73	123	2
140	133	120	72	70 70	117	3
158 159	142	123 122	123	69	71 70	4
195	178	150	112	67	70	5

71 ← Original pixel value
70 ← Pixel value after embedding a checksum bit

Vantaggi e svantaggi del checksum

- Semplice e veloce:
  - L'inserimento del checksum cambia (in media) solo la metà del numero di pixel; la distorsione visiva è limitata.
  - Si possono avere più watermark, purché non si sovrappongano.
- Fragile:
  - L'intero watermark può essere rimosso mettendo a zero tutti i bit meno significativi
  - Non funziona in caso di compressione con perdita di informazione.

Approfondimenti: Cifrari Nulli

Un cifrario nullo (null cypher) è un cifrario di occultamento in cui il testo in chiaro è mescolato con una grande quantità di materiale non cifrato.

- Il testo che segue fu realmente inviato da una spia tedesca durante la seconda guerra mondiale:

```
Apparently neutral's protest is thoroughly discounted and
ignored.Isman hard hit. Blockade issue affects pretext for embargo on
by-products, ejecting suets and vegetable oils.
```

- Considerando in sequenza la seconda lettera di ogni parola, si ottiene il messaggio:

```
Pershing sails from NY June 1
```

Approfondimenti:
La Steganographia di Tritemio

- Il primo testo stampato di steganografia fu scritto dall'umanista e teologo tedesco Johannes von Heidenberg (detto *Trithemius*), tra il 1499 e il 1500.
- Il testo, intitolato *Steganographia*, circolò a lungo in forma manoscritta e fu stampato solo nel 1606. Poco dopo venne iscritto nell'*Index Librorum Prohibitorum* in quanto "pericoloso e colmo di superstizioni" e la Chiesa ritenne quei manoscritti troppo espliciti, e condannò le copie esistenti ad essere bruciate.
- *Il libro III co*ntiene messaggi cifrati nascosti all'interno di quella che è apparentemente un'opera di magia.
- Dopo quasi 500 anni questi crittogrammi sono stati rilevati e risolti. (Dal 1606 si sapeva che cifrari simili erano presenti nei Libri I e II).
- Un dilemma è stato risolto solo nel 1998 da Jim Reeds [J. Reeds, "Solved: The Ciphers in. Book III of Trithemius's Steganographia", Cryptologia, 1998]
- Di conseguenza *Steganographia* non può più essere considerato come uno dei principali trattati demoniaci dei primi tempi moderni, ma si rivela senza ambiguità come il primo trattamento della crittografia in Europa.

Approfondimento: Stenografia sostitutiva

- Si basa sull'osservazione che la maggior parte dei canali di comunicazione (linee telefoniche, trasmissioni radio, ecc.) trasmettono segnali che sono sempre accompagnati da qualche tipo di **rumore**.
- Questo rumore può essere sostituito da un **segnale** (il **messaggio segreto**) che è stato trasformato in modo tale che, a meno di conoscere una chiave segreta, è indistinguibile dal rumore vero e proprio, e quindi può essere trasmesso senza destare sospetti.
- La tecnica impiegata è concettualmente molto semplice, consiste nel sostituire i **bit meno significativi** (LSB least significant bit) dei file digitalizzati con i bit che costituiscono il messaggio segreto.
- Quello che succede quindi è che il file contenitore risultante, dopo un'iniezione steganografica, si presenta in tutto e per tutto **simile** all'originale, con differenze difficilmente percettibili e quindi, a meno di confronti approfonditi con il file originale (non effettuabili ad occhio nudo) è difficile dire se le eventuali perdite di qualità siano da imputare al rumore od alla presenza di un messaggio segreto.

Approfondimento: Stenografia selettiva

- Ha un valore puramente teorico e viene raramente utilizzata e mirano a scegliere il supporto a seconda del messaggio da occultare. Viene effettuata una selezione dei supporti a disposizione o si usano semplici algoritmi per generarli, procedendo per tentativi finché non vengono rispettate particolari condizioni.
- Un esempio di tecnica selettiva è quella di impostare una funzione *hash* che controlli la parità dei bit del file digitale contenitore, assumendo il valore 1 se il cover contiene un numero di bit uguali ad 1 dispari, e il valore 0 se ne contiene un numero pari. A questo punto, volendo codificare il bit 0, acquisendo il file binario si controlla se il numero di bit uguali ad 1 sia pari, in caso affermativo si è trovato un file adatto a contenere l'informazione codificata, altrimenti si procede con l'acquisire un altro file digitale.
- Ha il pregio di risultare praticamente impossibile da identificare, in quanto il supporto, nonostante contenga effettivamente un messaggio segreto, non risulta assolutamente modificato.
- Ha il difetto di rivelarsi una soluzione alquanto dispendiosa in termini di tempo e insoddisfacente dal punto di vista dell'esiguo quantitativo di dati segreti che permette di celare. Infatti, dovendo aumentare il numero di bit da nascondere, aumenta anche il tempo per il reperimento o la generazione del supporto, il quale è costretto a soddisfare più vincoli contemporaneamente.

Approfondimento: Stenografia Costruttiva

- Opera più o meno come la steganografia sostitutiva, con la differenza che nel modificare il file contenitore si tiene conto di un modello di rumore, nel senso che si tenta di sostituire il rumore presente con il messaggio segreto nel rispetto delle caratteristiche statistiche del rumore originale.
- Secondo questa concezione, un buon sistema steganografico dovrebbe basarsi su un modello del rumore e adattare i parametri dei suoi algoritmi di codifica in modo tale che il falso rumore contenente il messaggio segreto sia il più possibile conforme al modello di partenza.
- Però non è facile costruire un modello accurato del rumore. La costruzione di un modello del genere richiede grossi sforzi ed è probabile che qualcuno, in grado di disporre di maggior tempo e di risorse migliori, riesca a costruire un modello più accurato riuscendo ancora a distinguere tra il rumore originale e un sostituto.
- Inoltre, se il modello del rumore utilizzato dal metodo steganografico dovesse cadere nelle mani del "nemico" egli lo potrebbe analizzare per cercarne possibili difetti e quindi utilizzare proprio il modello stesso per controllare che un messaggio sia conforme ad esso. Così, il modello, oltre ad essere parte integrante del sistema steganografico, fornirebbe involontariamente uno strumento di attacco estremamente efficace proprio contro lo stesso sistema.

Approfondimento: Scelta del cover

- La steganografia ha la necessità di utilizzare cover object che siano possibilmente inediti e mai usati in precedenza:
- E’ improbabile comporre di volta in volta nuovi brani musicali per inviare messaggi (essi risulterebbero comunque sospetti)
- E’ difficile modificare semplicemente il testo di pagine html (in tal modo si produrrebbero errori ortografici e frasi senza senso)
- E’ pericoloso utilizzare gli spazi in file testuali o alcuni tag dei file html (un siffatto modello di rumore può essere replicato se si conosce l’algoritmo di steganografia adoperato).
- Malelingue sostengono che negli anni ’80 l’ex primo ministro inglese Margaret Thatcher, preoccupata per la fuga di notizie riservate, lasciate filtrare alla stampa da suoi non troppo fedeli collaboratori, fece programmare i loro word processor in modo che il nome dello scrivente fosse codificato nella spaziatura delle parole

Approfondimento: Immagini

- Una immagine digitale è una matrice di pixel; questo termine deriva da *picture element*
- Il pixel contiene l'informazione relativa alla rappresentazione della realtà che è stata catturata tramite uno scanner, una macchina fotografica o un frame grabber (per i video)
- Ogni pixel contiene un quantità di informazione che può essere espressa in **bit** (binary digit).
  - il numero di bit riservati per ogni pixel viene denominato **profondità di colore**
  - Data la profondità di colore N, il numero di possibili tonalità per una immagine digitale è 2^N
  - La dimensione dell'immagine è rappresentata dal numero dei pixel che la compongono
- Per esprimere la dimensione si usa il formato **WxH**
  - dove W indica il numero di pixel orizzontali e
  - H il numero di pixel verticali
  - Esempio: 640x480 pixel
- Con il termine risoluzione si indica la densità dei pixel in relazione alla dimensione del supporto di visualizzazione (per esempio, un foglio di carta o uno schermo)
- La risoluzione si esprime comunemente in pixel per inch (ppi) o dot per inch (dpi)

Approfondimento: Immagini

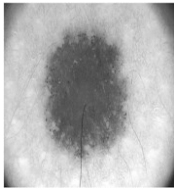
24 inch - grandezza dello schermo

Dimensioni
1920×1080
cioè Full HD

→ Risoluzione
91,79 ppi

Dimensioni
3840×2160
cioè Ultra HD (4k)

→ Risoluzione
183,58 ppi



- Ogni pixel contiene un quantità di informazione che può essere espressa in **bit** (binary digit).
- il numero di bit riservati per ogni pixel viene denominato **profondità di colore**
- Data la profondità di colore N, il numero di possibili tonalità per una immagine digitale è 2^N
- Esempio N = 1 ⇒ 2 tonalità, N = 4 ⇒ 16 tonalità, N = 8 ⇒ 256 tonalità

Approfondimento:
Formati file immagini

- Esistono tanti formati per file che rappresentano immagini o video. Noi abbiamo visto .bmp (bit map) a 24 bit, ci sono anche .bmp a 8 bit ma anche altri formati tipo .gif, .jpg, jpeg, .pict, .png e tanti altri. Ovv.iamente l'algoritmo per iniettare il messaggio nel contenitore cambia da formato a formato
- Va poi considerato che esistono formati **Lossless** (compressione dati senza perdita) ma anche formati **Lossy** (compressione con perdita) come .jpeg. In questo secondo caso non è possibile iniettare come visto per .bmp. Ad esempio, se venisse iniettato un messaggio m in un file bitmap e dopo questo venisse convertito in .jpeg, le informazioni andrebbero inevitabilmente perse.
- La compressione .jpeg, infatti preserva le caratteristiche visive dell'immagine ma non che l'esatta informazione contenuta nella sequenza di pixel, non sarebbe quindi possibile risalire al file bitmap originario ed estrarre m.

Approfondimento: Immagini

- L'occupazione di memoria è data dal prodotto tra la dimensione dell'immagine e la profondità di colore del singolo pixel
  - Occupazione = (Dimensione) x (Profondità di colore)
  - Esempio: Una immagine a colori (RGB **raw**) 640x480 occupa in memoria 9830400 bit pari a circa 1.23 MB
- Per limitare l'occupazione di memoria è possibile comprimere l'immagine ad esempio tramite **Differential Encoding:**
  - Invece di salvare in memoria il valore assoluto per ogni pixel, si può tenere traccia della differenza tra un valore e quello precedente (riducendo così il range)
  - Esempio: se per digitalizzare un valore si necessita di 12 bit, mentre la differenza richiede soli 3 bit, allora posso risparmiare fino al 75% della memoria.
- La compressione può essere *lossless* o *lossy*
  - Compressione lossless: **reversibile,** e.g., file PNG e file ZIP
  - Compressione lossy: **ricostruzione approssimata**, dove maggiore è la compressione, maggiore è l'errore, e.g., file JPEG e file MP3
- **Transform Encoding:** L' occhio umano è meno sensibile alle alte frequenze spaziali:
  - Se l'ampiezza di una componente ad alta frequenza cade sotto una certa soglia, l'occhio NON LA RILEVA
  - La quantizzazione può essere meno accurata alle alte frequenze e ossono essere utilizzati meno bit.

Approfondimento:
Applicazioni del Watermarking

Prova di Proprietà

- A utilizza una chiave privata per generare un watermark e la inserisce nel document
- A rende l'immagine disponibile al pubblico
- B afferma di essere il proprietario dell'immagine derivata dall'immagine pubblica
- A produce l'originale non marcato e dimostra la presenza del watermark di A

Impronte digitali

- Utilizzate per evitare la duplicazione e la distribuzione non autorizzata.
- Un watermark distinto (un'impronta digitale) è incorporata in ogni copia dei dati.
- Se vengono trovate copie non autorizzate, l'origine della copia può essere determinata recuperando l'impronta digitale.

Approfondimento:
Applicazioni del Watermarking

Autenticazione e verifica dell'integrità

- I Watermark possono essere utilizzati per rilevare anche il minimo cambiamento nel documento.
- Una chiave unica associata alla fonte viene utilizzata per creare la filigrana che viene poi incorporata nel documento.
- La chiave viene utilizzata per estrarre il watermak; l'integrità del documento è verificata sulla base dell'integrità di questo.

Etichettatura del Contenuto

- I bit aggiunti contengono annotazioni, che forniscono ulteriori informazioni sui dati.
- Le fotocamere digitali annotano le immagini con ora e data di quando la fotografia è stata scattata.
- Le macchine di imaging medico annotano le immagini con il nome del paziente.

Approfondimento: Watermarking deboli e forti

I watermark possono essere classificati a seconda di alcune loro proprietà, che dipendono dallo scopo con cui sono stati inseriti all'interno del documento.

Un watermark può essere visibile o invisibile.

- Watermark visibile: utilizzato per codificare informazioni che devono essere rese pubbliche all'utente finale.
- Watermark invisibile: utilizzato in quei contesti in cui il proprietario legittimo vuole garantirsi i diritti d'autore

Resistenza agli attacchi:

- Un watermark fragile può essere facilmente attaccato, distrutto e reso irriconoscibile da quasi ogni tipo di manipolazione dei dati.
- Un watermark robusto deve resistere alle più comuni operazioni e trasformazioni sui dati

Approfondimento: Watermarking deboli e forti

Resistenza agli attacchi:

- Un watermark fragile può essere facilmente attaccato, distrutto e reso irriconoscibile da quasi ogni tipo di manipolazione dei dati. Si applica quando occorre dimostrare che una informazione non è più quella originale anche se alterata in minima parte.
- Un watermark robusto deve resistere alle più comuni operazioni e trasformazioni sui dati, in quanto è utilizzato quando la proprietà del documento deve essere provata o garantita. Si applica quando occorre dimostrare l'origine di una informazione, anche quando essa viene fortemente distorta e/o manipolata?

Esempio: Usando uno strumento grafico come photoshop è possibile in una decina di minuti modificare una immagine.

- Se occorre dimostrare che l'immagine, anche se modificata, è la stessa allora uso un watermark robusto.
- Se occorre dimostrare che l'immagine, anche se minimamente modificata, non è più autentica allora uso un watermark fragile