
Informatica Personale

Contributi digitali per rappresentare immagini e suoni: Parte I

Dante Picca

Anno Accademico 2009-2010

Premessa

Nella prima parte del corso abbiamo visto che un elaboratore elettronico deve essere in grado di trattare quattro tipi di informazioni:

Numeri

Caratteri

Immagini

Suoni.

Successivamente abbiamo visto come gli elaboratori elettronici trattano i primi tipi di informazioni.

Adesso il focus sarà concentrato sui restanti due.

Il formato di un file

Per il trattamento dei testi, come visto, ogni lettera è trasformata in un numero utilizzando la tabella ASCII.

Poiché i bit che costituiscono un file sono collegati alle informazioni in esso contenute, ad ogni bit è associato uno specifico significato, che dipende dal tipo di contenuto conservato nel file. Poiché i bit di un file possono rappresentare molti contenuti diversi, è necessario definire una convenzione che associ ad ogni bit un significato.

Il **formato di un file** è la convenzione usata per leggere, scrivere ed interpretare i contenuti del file.

Nel seguito vedremo come sono conservate le immagini ed i suoni nei file e perché esistono formati diversi.

Che cos'è una immagine

Che cos'è una immagine ?

Raffigurazione grafica, fotografica o plastica di qualcosa o qualcuno.

Da "Dizionario De Mauro della lingua italiana"

Forma esteriore degli oggetti corporei, in quanto viene percepita attraverso il senso della vista ..., o rimane impressa in una lastra o pellicola o carta fotografica.

Da "Il vocabolario Treccani"

Che cos'è una immagine

Nel caso di *immagini analogiche* (ad esempio le fotografie), sulla pellicola vengono posti dei materiali fotosensibili che alterano il loro stato quando colpiti dalla luce. L'immagine rappresentata sulla carta fotografica è ottenuta per analogia con la quantità di luce che ha impresso i diversi punti della pellicola durante la fase dell'esposizione (lo scatto).

Nel caso di *immagini digitali*, i diversi colori che rappresentano l'immagine sono memorizzati come numeri. In fase di acquisizione delle immagini in formato digitale (macchina fotografica digitale, scanner, ...) ad ogni colore è associato un numero. In fase di visualizzazione (monitor, stampante, ...) ad ogni numero è fatto corrispondere un colore.

Le immagini digitali

Pensiamo al processo di acquisizione di una immagine usando una macchina fotografica digitale. Come si fa a rappresentare la natura con un insieme di numeri (e quindi in formato digitale)?

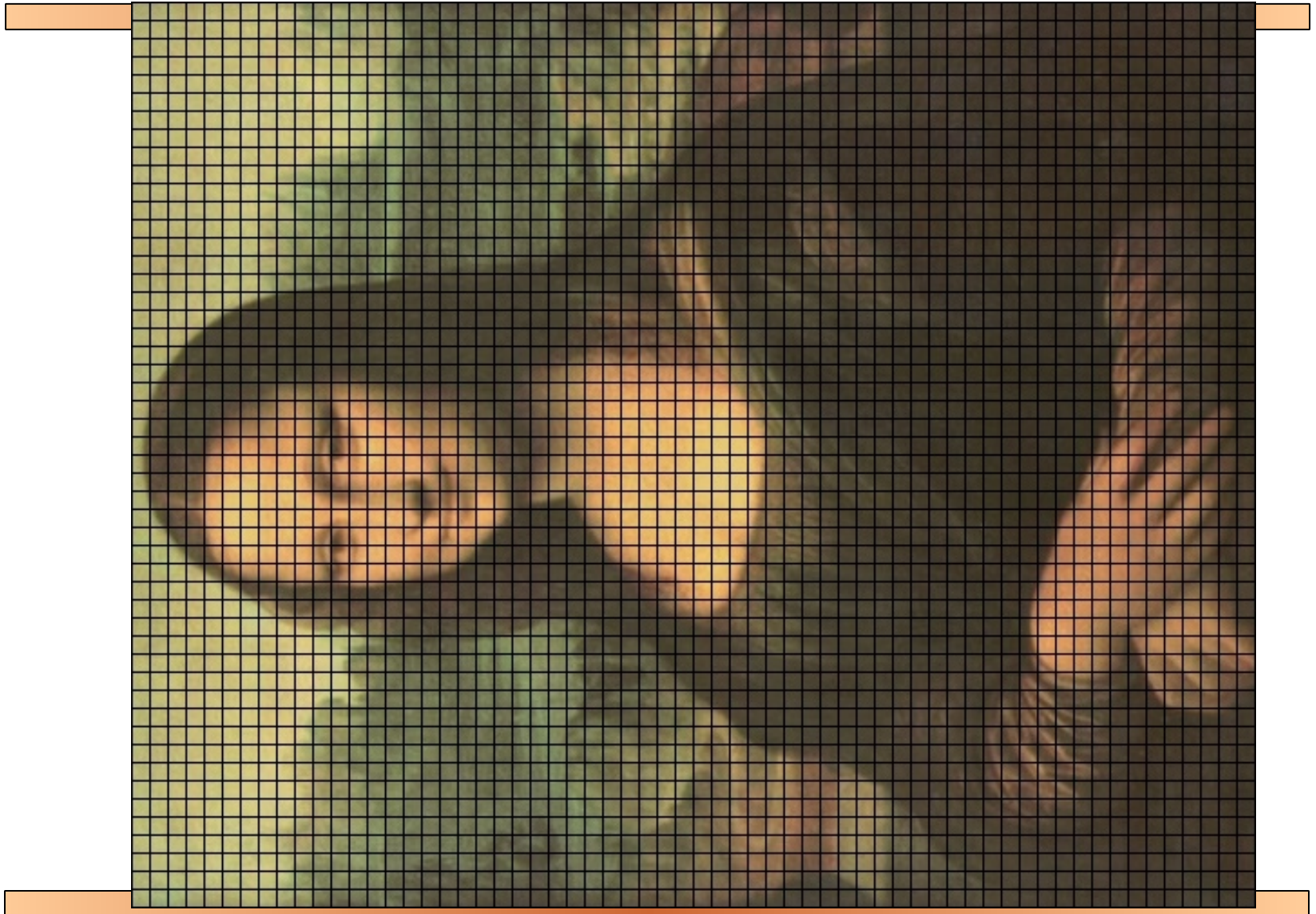
Per comprendere questo processo, detto di digitalizzazione, è necessario introdurre due concetti: il *campionamento spaziale* e la *quantizzazione cromatica*.

Il campionamento spaziale

Per **campionamento spaziale** si intende l'operazione con cui una immagine continua viene trasformata in un insieme discreto di rettangoli più o meno grandi, i **pixel**. Una immagine viene così rappresentata come una matrice di pixel. Per ogni immagine deve essere definito quante righe e quante colonne di pixel la compongono (ad ex. 640x480).

Nell'esempio seguente ad ogni pixel è associata una porzione dell'immagine con molti colori diversi. In realtà la periferica utilizzata per digitalizzare l'immagine assegna, per ogni pixel, un colore solo che corrisponde alla media dei colori presenti nel pixel stesso. Il pixel è, quindi, l'elemento minimo di informazione relativo ad una immagine, poiché ad ogni pixel è associato un colore.

1 pixel



La quantizzazione cromatica

Sappiamo che i colori possono assumere infiniti valori, poiché infinite sono le sfumature. Ovviamente non è possibile rappresentare in formato digitale tutti questi colori, poiché significherebbe associare ad ogni colore un numero compreso tra zero ed infinito. D'altronde memorizzare infiniti colori sarebbe anche inutile perché:

- l'occhio umano non è in grado di percepire la differenza tra due colori troppo simili;
- le periferiche di output hanno una capacità di rappresentazione dei colori molto più limitata della capacità di percezione dell'occhio umano.

La quantizzazione cromatica

Per tale motivo, dopo il campionamento spaziale, è opportuno effettuare una **quantizzazione cromatica**.

Con quantizzazione cromatica si intende l'operazione con la quale viene assegnato, ad ogni pixel, un numero - compreso in un opportuno intervallo - che ne identifica il colore.

Per ogni pixel, quindi, viene prima calcolato il colore medio presente nel rettangolo e, successivamente, a tale colore viene assegnato un numero compreso in un intervallo finito di possibili valori.

La quantizzazione cromatica

Per ogni immagine viene definito il numero di colori diversi possono essere rappresentati al suo interno. Tale valore viene indicato come **profondità di colore** e si esprime in bit.

<i>Profondità di colore (bit)</i>	<i>Numero di colori diversi</i>
1	2
2	4
4	16
8	256
24	16.777.216

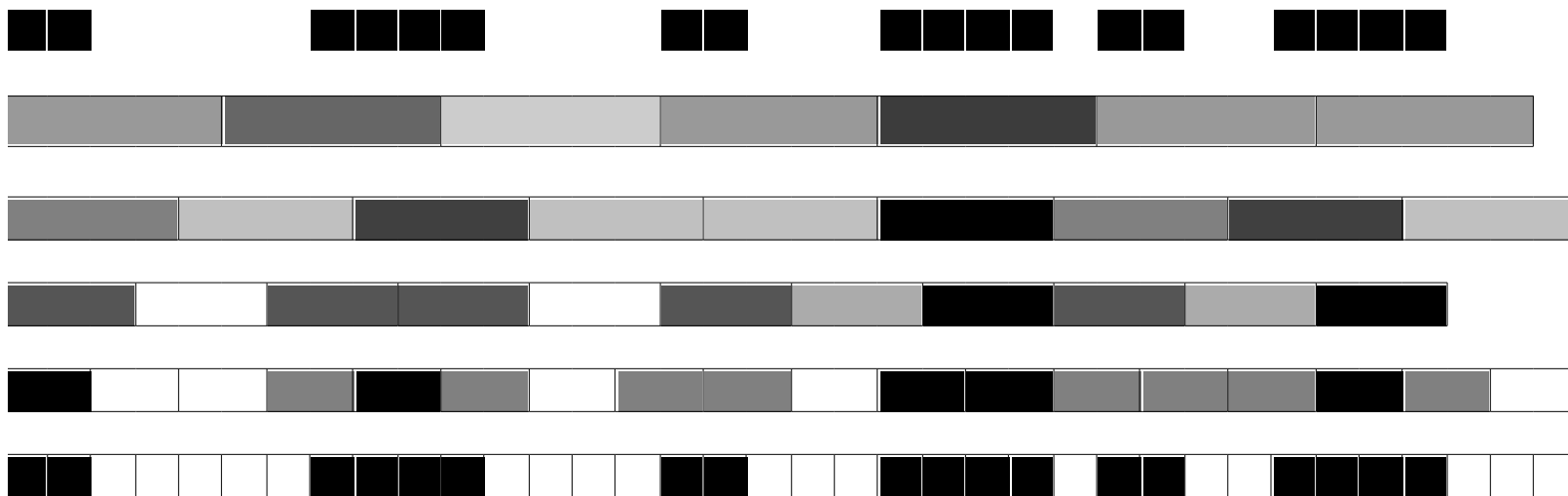
Risoluzione spaziale

Quanti pixel servono per visualizzare bene una immagine ? Questa domanda è mal posta, poiché non esiste un valore assoluto, ma tale valore dipende da cosa vogliamo distinguere nell'immagine.

Per *risoluzione spaziale* si intende lo spazio minimo tra due oggetti per cui, all'interno dell'immagine, essi sono ancora visibili come due entità separate. In una immagine digitale più fitti sono i pixel, ovvero quanto più grande è la dimensione della matrice, più la risoluzione è alta.

Quando si acquisisce una immagine (ad esempio allo scanner), è necessario specificare la risoluzione spaziale con cui si intende acquisire l'immagine. Questa è generalmente espressa in dpi (Dot Per Inch).

Risoluzione spaziale



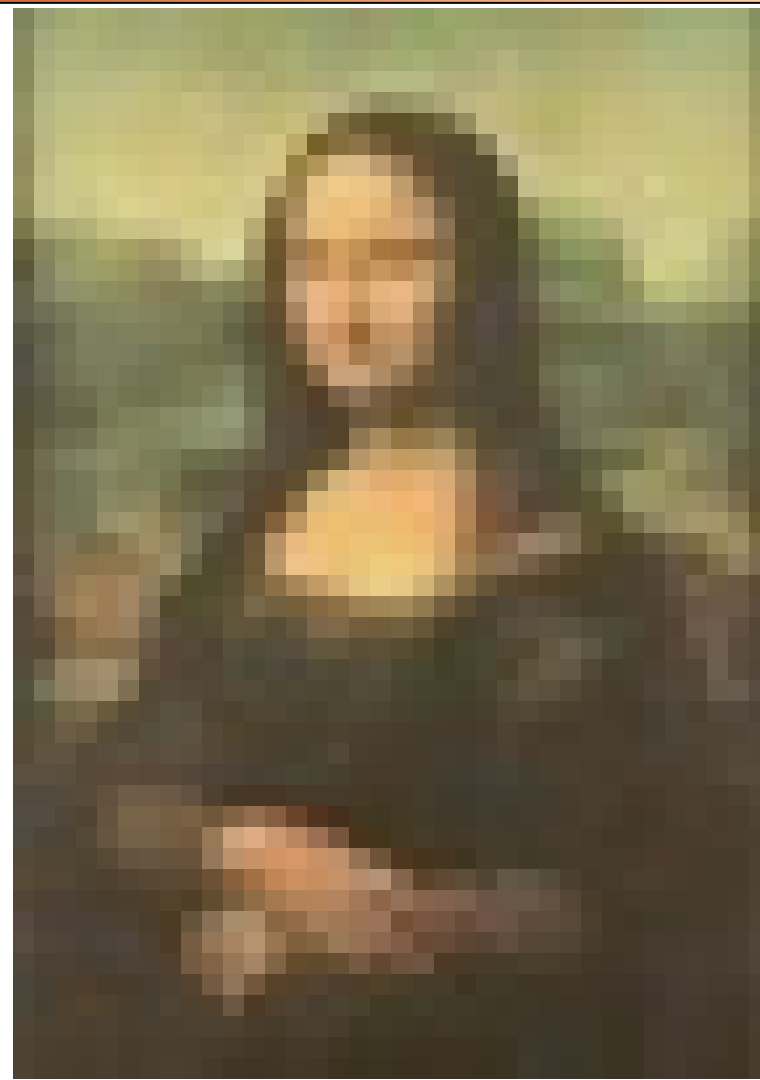
Risoluzione spaziale

A parità di colori rappresentati, tanto più è alta la risoluzione spaziale tanto più peserà il file che la contiene. Supponendo di utilizzare una profondità di colore di 3 byte, si ottiene la seguente tabella per il peso dell'immagine.

<i>Dimensione</i>	<i>Peso (in byte)</i>
100x100	30k
200x200	120k
640x480	922k
1024x768	2.359k

Non sempre, però, è necessaria una elevata risoluzione spaziale per riuscire a capire il contenuto di una immagine: il nostro cervello è un eccellente supporto.

Risoluzione spaziale



Risoluzione spaziale: le diapositive

72 dpi

1200 dpi

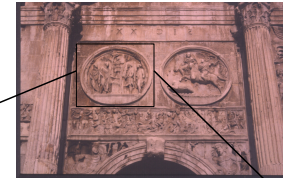


21kB



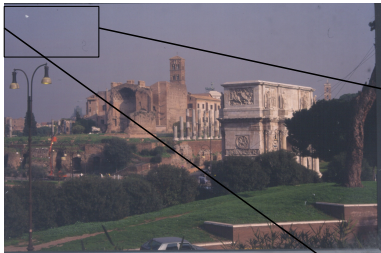
2kB

5,3MB



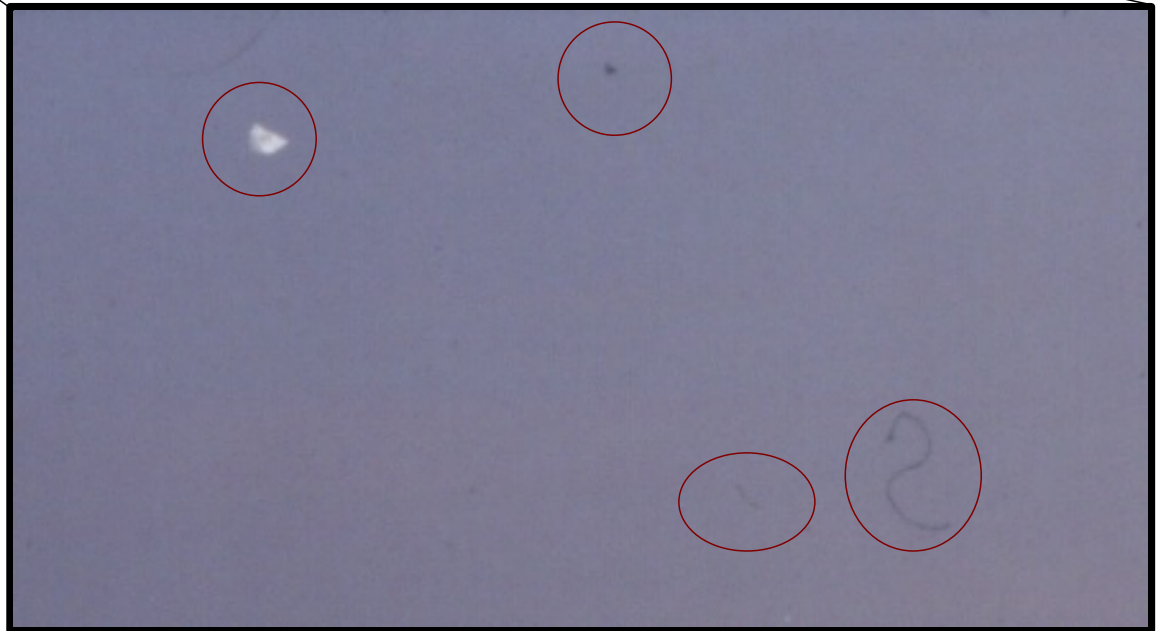
300kB

Risoluzione spaziale: dettagli inutili

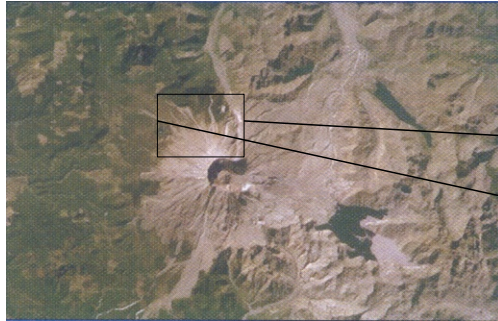


5,4MB

1200 dpi



Risoluzione spaziale: scansione di un libro



500KB

200 dpi



Risoluzione cromatica

Anche la domanda di quanti colori servono per visualizzare correttamente una immagine è mal posta, perché dipende da ciò che vogliamo distinguere nell'immagine.

Per *risoluzione cromatica* si intende la minima differenza di colore per cui, all'interno dell'immagine, essi sono ancora visibili come due entità separate. In una immagine digitale maggiore è la profondità di colore, più la risoluzione è alta.

A causa del funzionamento dei nostri occhi, la risoluzione cromatica dipende molto dalle sfumature di colore che si intendono separare. Per tale motivo, definire un valore ottimale è difficile.

Anche in questo caso, a parità di risoluzione spaziale, al crescere della risoluzione cromatica aumenta il peso del file.

Risoluzione cromatica: colori ottimizzati

24 bit



1 bit



2 bit



3 bit



Risoluzione cromatica



8 bit (256 colori)
opportunamente scelti



8 bit (256 colori) non
opportunamente scelti

Risoluzione cromatica

16M



256



256
ott.

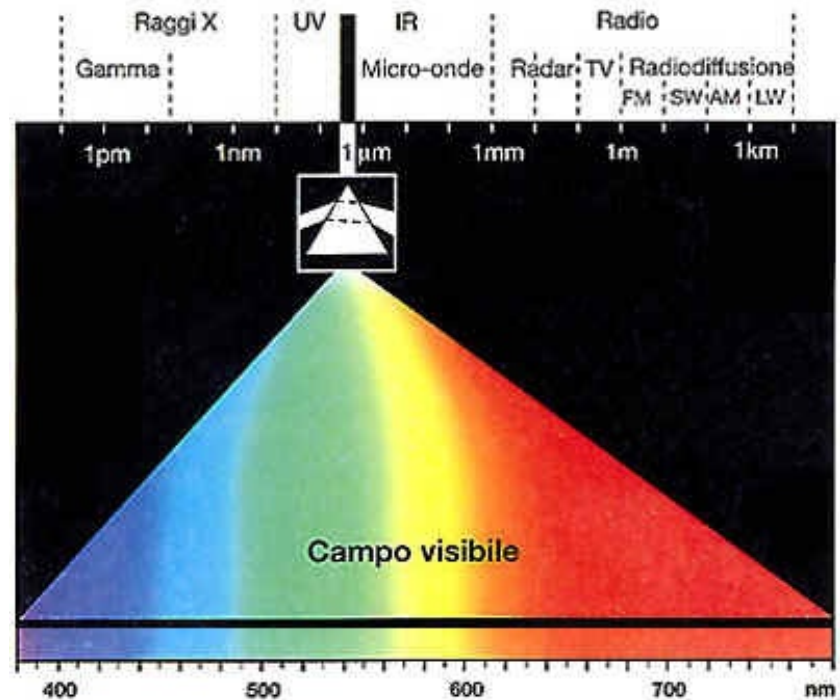


8



La rappresentazione dei colori

In che modo sono rappresentati i colori in formato digitale? Sappiamo che, ad ogni colore che compone lo spettro della luce bianca, corrisponde una lunghezza d'onda distinta. Tale valore potrebbe essere utilizzato come valore digitale da associare al colore.



La rappresentazione dei colori

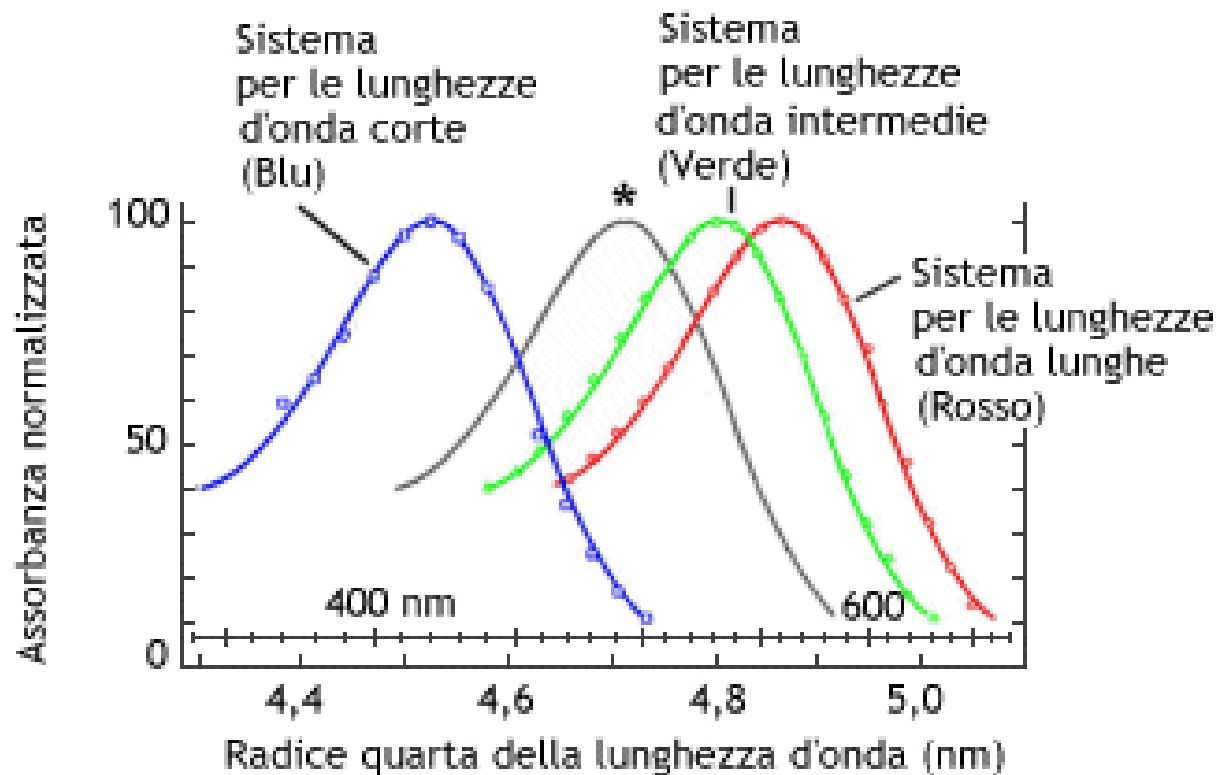
Questa idea, per quanto semplice, non è facile da realizzare. Il metodo più utilizzato per rappresentare i colori si basa sul funzionamento dell'occhio umano.

Sulla nostra retina sono presenti due tipi di rivelatori: i coni ed i bastoncelli.

Mentre i bastoncelli sono adatti alla visione in condizioni di scarsa luminosità (ad esempio di notte), i coni sono adatti alla visione in condizioni di buona luminosità.

La rappresentazione dei colori

Esiste un solo tipo di bastoncelli, mentre esistono tre tipi diversi di coni, caratterizzati da una diversa risposta in funzione della lunghezza d'onda della luce incidente.



Il sistema tricromatico RGB

Noi percepiamo i colori come somma degli stimoli ricevuti dal cervello per i tre tipi di coni.

Nella retina ci sono circa 130 milioni di fotoricettori!

Poiché i tre tipi di coni sono particolarmente sensibili al rosso (Red), al verde (Green) ed al blu (Blu), questo sistema di rappresentazione dei colori si chiama **tricromatico RGB**.

Questo sistema si basa sulla **sintesi additiva** dei colori, cioè il nostro cervello associa ad un punto un colore che è dato dalla somma di opportune quantità dei tre colori primari. Tali quantità sono definite dallo stimolo ricevuto dai coni.

Questo è il sistema utilizzato anche nelle televisioni, nei monitor e nei proiettori.

Il sistema tricromatico RGB



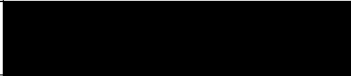
Il sistema tricromatico RGB

In formato digitale i colori sono generalmente rappresentati in formato RGB, cioè specificando quanto rosso, quanto verde e quanto blu compongono il colore di ogni pixel dell'immagine.

Le quantità di rosso di verde e di blu per ogni pixel, sono definite ciascuna da un valore compreso tra 0 e 255 (un byte) e, quindi, l'immagine ha 24 bit di profondità (che corrisponde a 16.777.216 colori diversi).

Tutti i colori che vediamo in televisione o sul monitor sono rappresentati come un numero composto dalla quantità di rosso, di verde e di blu.

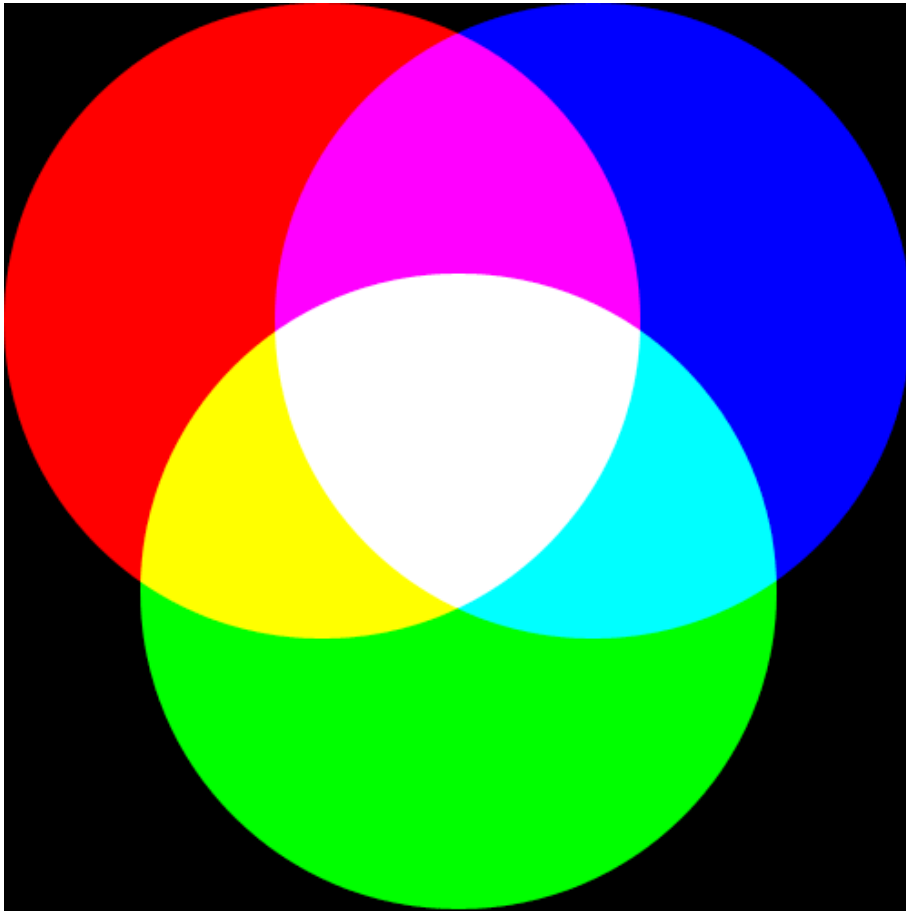
Il sistema tricromatico RGB

Colore	Rosso	Verde	Blu	
Nero	0	0	0	
Bianco	255	255	255	
Rosso	255	0	0	
Verde	0	255	0	
Blu	0	0	255	
Ciano	0	255	255	
Magenta	255	0	255	
Giallo	255	255	0	
Marrone	153	102	51	
Rosso scuro	153	40	76	

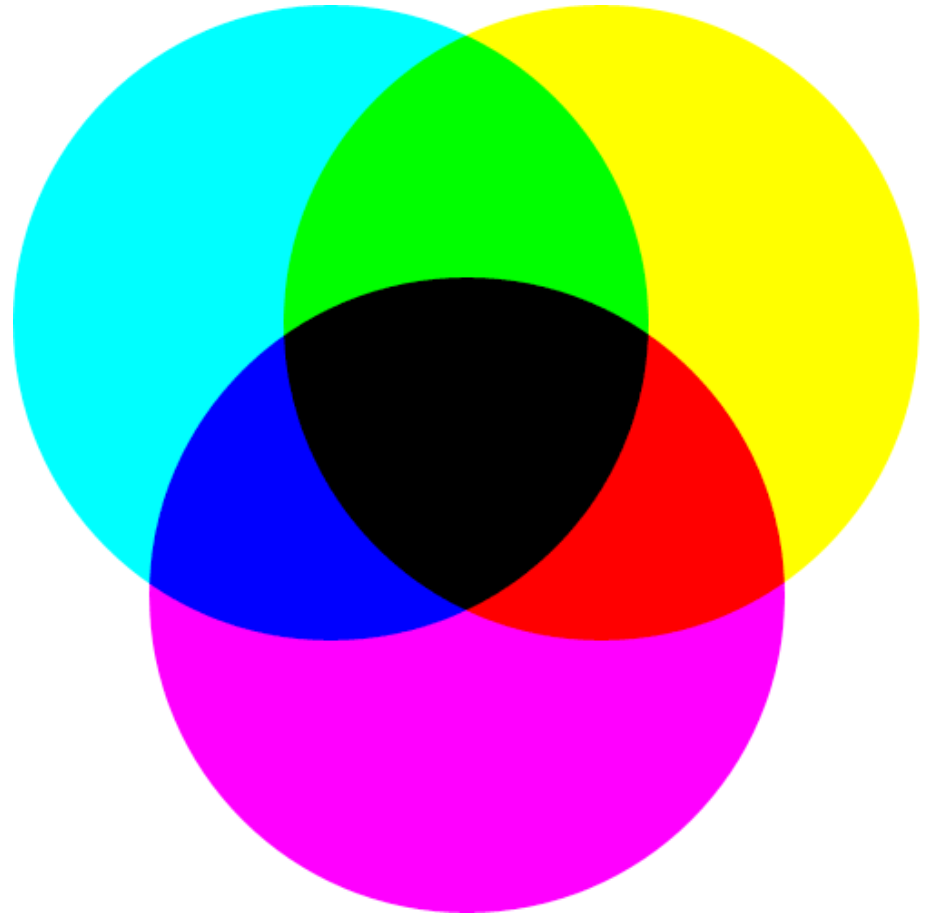
Il sistema tricromatico CMY

Nella stampa, l'inchiostro invece di emettere luce (come i cristalli del monitor) la assorbe, e non si può quindi utilizzare una sintesi additiva: è necessario utilizzare la sintesi sottrattiva. Per tale motivo, invece del sistema tricromatico additivo RGB, si usa il **sistema tricromatico sottrattivo CMY**, dove i colori primari sono il ciano (Ciano), il magenta (Magenta) ed il giallo (Yellow). La stampante, dosando opportunamente gli inchiostri, è in grado di riprodurre su carta i diversi colori visibili sul monitor. Ciano, magenta e giallo sono detti i colori complementari del rosso del verde e del blu.

Sistemi tricromatici RGB e CMY

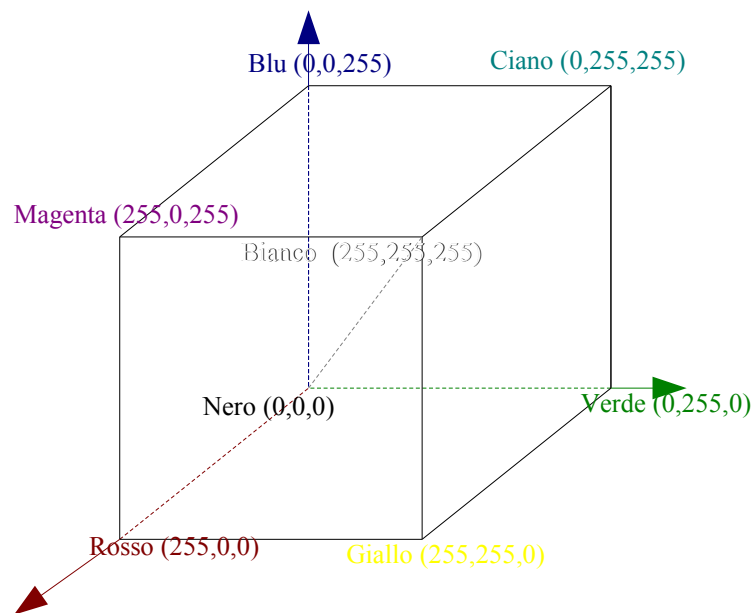


RGB

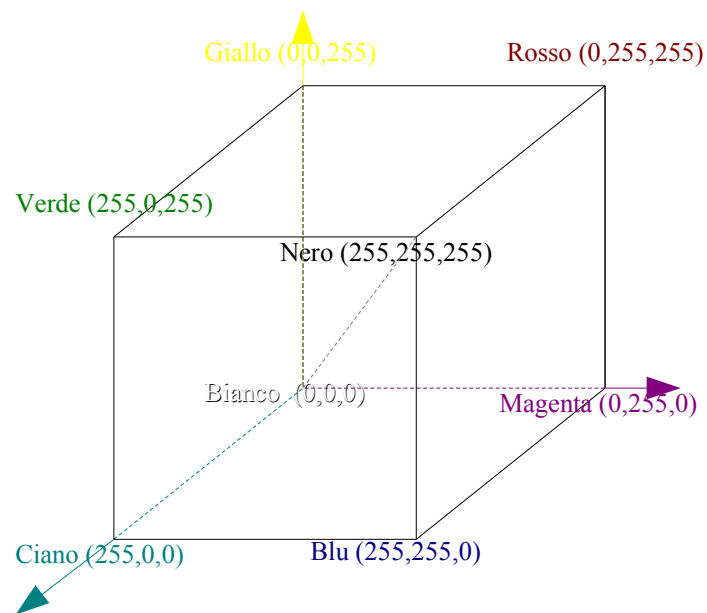


CMY

Sistemi tricromatici RGB e CMY



RGB



CMY

Sistema tricromatici RGB e CMY

Se un colore ha valori nel sistema RGB:

rosso = r

verde = v

blu = b

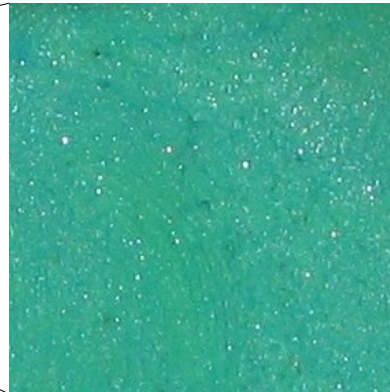
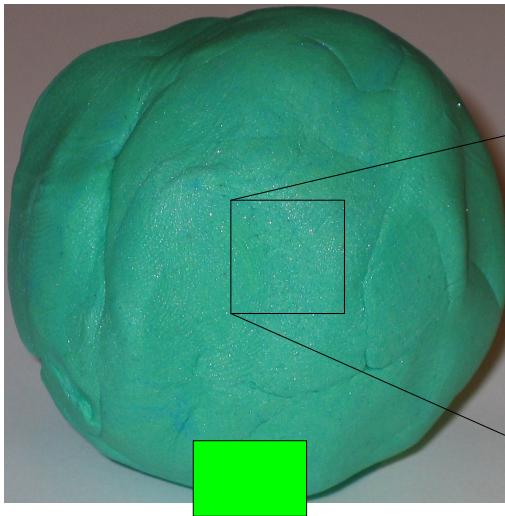
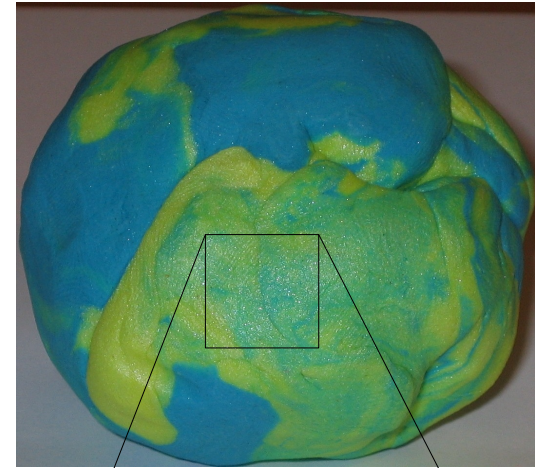
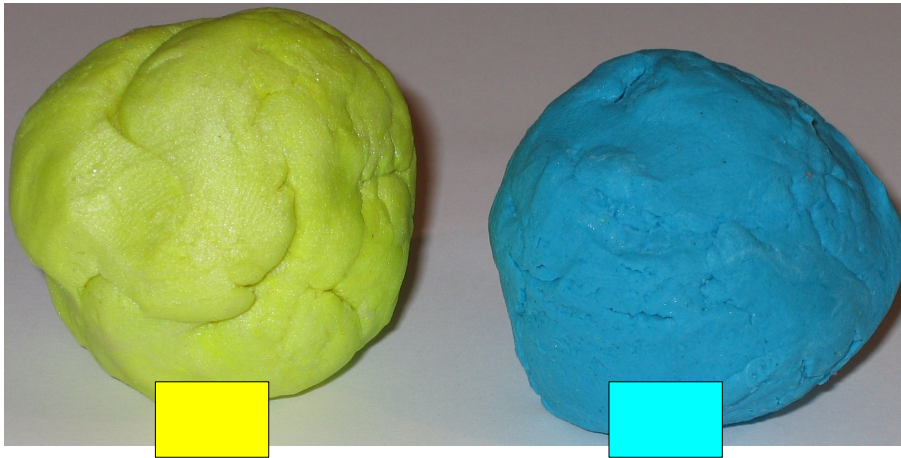
nel sistema CMY avrà valori:

ciano=255-r

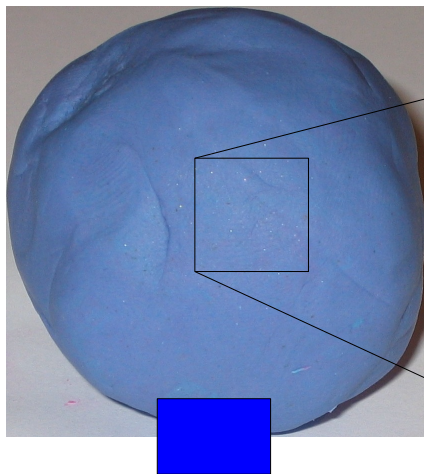
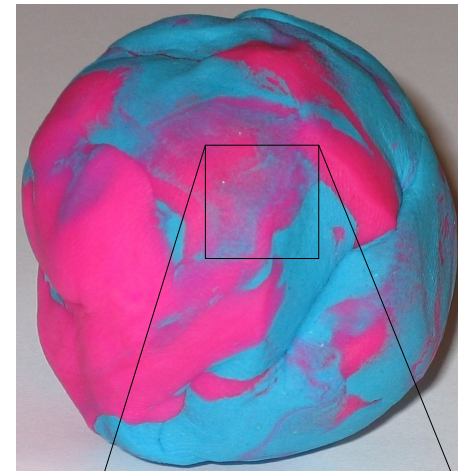
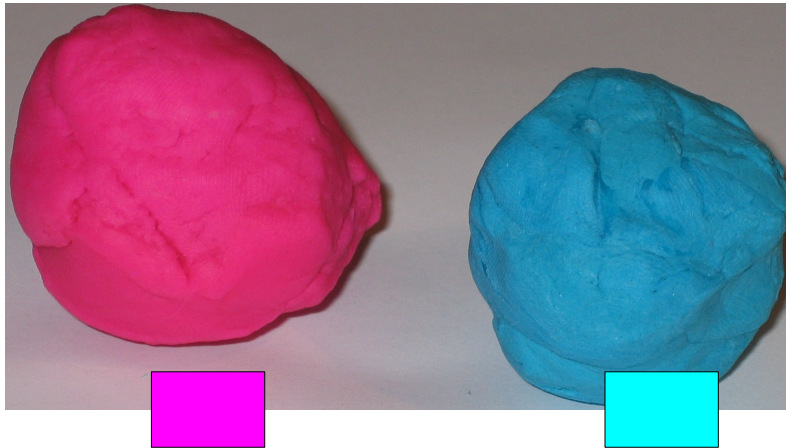
magenta=255-v

giallo=255-b

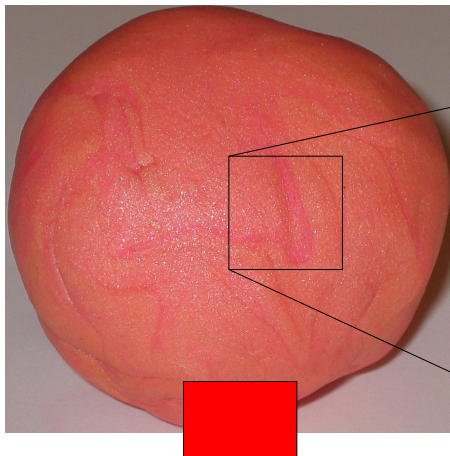
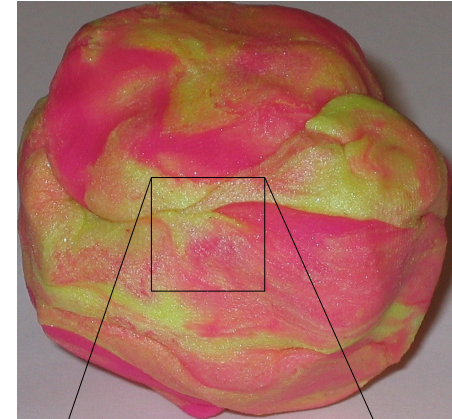
Il formato tricromatico CMY



Il formato tricromatico CMY



Il formato tricromatico CMY



Il formato BMP

Sulla base di quanto abbiamo imparato, un file contenente una immagine può essere archiviato con un formato molto semplice: per ogni pixel vengono indicati i 3 byte corrispondenti al rosso, al verde ed al blu.

Ancora una volta, come già avvenuto con la tabella ASCII, abbiamo trasformato una qualche entità (prima le lettere, ora i colori) in numeri.

Il dispositivo di input (ad esempio lo scanner) deve memorizzare per ogni punto le quantità di rosso, verde e blu. Il computer, quando invia una immagine al monitor, deve leggere queste quantità ed inviare gli opportuni comandi alla periferica.

Il formato BMP

Questo formato di conservazione delle immagini è detto **bitmap** ed i file così archiviati hanno estensione **.bmp**. Ovviamente, nel file bitmap devono essere inserite anche altre informazioni necessarie alla corretta visualizzazione delle immagini, come il numero di pixel in una riga, la risoluzione spaziale utilizzata o la profondità di colore.

Il problema fondamentale di questo formato è la dimensione del file: una immagine a tutto schermo salvata in formato bitmap pesa circa 2,5M Byte. Tale difetto rende questo formato utilizzabile solo per un ristretto insieme di applicazioni, ma lo rende inutilizzabile ad esempio per il web.

Il formato GIF

Nel 1987, con l'obiettivo di definire un formato per la conservazione delle immagini che fosse sufficientemente leggero da poter essere trasferito via rete, fu sviluppato il formato **GIF (Graphic Interchange Format)**. Le idee alla base del formato gif sono molto semplici:

- le schede video di allora non erano in grado di visualizzare molti colori (già 256 era un buon numero). L'idea fu di estrarre per ogni immagine, dai 16 milioni di colori possibili, un sottoinsieme di 256 colori che – meglio degli altri – rappresentano i colori presenti nell'immagine. Tali colori possono essere rappresentati con un singolo byte, riducendo la dimensione del file di un fattore 3.

Il formato GIF

- poiché l'immagine contiene solo 256 colori, è facile avere sequenze di pixel simili in punti diversi dell'immagine: è possibile effettuare una compressione dei dati, con tecniche non molto diverse da quelle che utilizziamo per gli altri file (ad esempio con winzip).

Il risultato è un file di dimensioni contenute che ben si adatta a conservare immagini artificiali (cioè realizzate completamente al computer), come ad esempio pulsanti, banner o icone.

Il formato PNG

La restrizione a 256 colori visibili contemporaneamente, nonché la presenza di alcuni brevetti sugli algoritmi alla base della compressione dei dati nel formato GIF, hanno spinto la comunità a sviluppare un nuovo formato più adatto alle mutate esigenze (ad esempio lo sviluppo dell'internet e le schede video sempre più potenti).

Per rispondere alle mutate esigenze, nel 1995 fu introdotto il formato **PNG (Portable Network Graphics)**. Oltre alla possibilità di visualizzare immagini sino a 48 bit (281.474.976.710.656 colori diversi) ed essere libero da brevetti, il formato PNG contiene molte altre caratteristiche che lo rendono adatto alla conservazione delle immagini digitali.

Il formato PNG

Utilizzando questo formato è possibile conservare immagini con qualità (e quantità di informazioni) superiore rispetto al formato bitmap, pur mantenendo le dimensioni del file contenute.

Tali caratteristiche rendono questo formato adatto all'archiviazione di immagini per le quali è richiesta una alta qualità. Questo formato è anche adatto ad essere utilizzato nei siti web.

Sfortunatamente, l'algoritmo utilizzato per la compressione dei dati nel formato PNG dà i migliori risultati di compressione con immagini non fotorealistiche.

Formato lossless e lossy

I formati visti sino a questo momento sono detti formati **lossless** (senza perdita di informazioni), poiché gli algoritmi di compressione conservano tutti i dati presenti nell'immagine originale.

Una volta definita la risoluzione e la profondità di colore, gli algoritmi lossless cercano di comprimere i dati (riducendo quindi lo spazio dell'immagine) conservando tutte le informazioni presenti nella immagine iniziale.

Questi algoritmi sono simili, per funzionamento, a quelli usati da tanti programmi (ad esempio winzip) per comprimere un generico file.

Formato lossless e lossy

Nel caso delle immagini, non tutte le informazioni sono indispensabili, poiché il nostro cervello non percepisce alcune sfumature o lievi variazioni di colore: eliminare questi dati non comporta alcuna perdita apparente di informazioni da parte del fruitore.

L'applicazione di un metodo di compressione dati **lossy** (con perdita di informazioni) comporta perdita di qualità. Decomprimendo un file compresso con un algoritmo lossy la copia ottenuta sarà peggiore dell'originale, ma in genere abbastanza simile da non comportare sostanziale perdita di informazione per il fruitore. Ciò è possibile poiché i metodi di compressione lossy tendono a scartare le informazioni poco percepibili, archiviando solo quelle essenziali.

Il formato JPEG

Il formato lossy più usato per la conservazione delle immagini è il formato **JPEG (Joint Photographic Experts Group)**.

L'idea dell'algoritmo alla base di questo formato è quella di eliminare dall'immagine tutte le informazioni che il nostro cervello non è in grado di percepire. Come risultato si ottiene un file molto più leggero dell'originale e che archivia una immagine molto simile (per noi) all'originale.

L'algoritmo di compressione utilizzato - basato sullo studio della percezione umana rispetto agli stimoli visivi - è particolarmente adatto alla memorizzazione di fotografie e, più in generale, di immagini fotorealistiche.

Il formato JPEG

In fase di salvataggio dell'immagine compressa con il formato JPEG, l'utente può decidere la qualità dell'immagine archiviata: minore sarà la qualità dell'immagine, più leggero sarà il file che la contiene.

La differenza tra un file conservato con un formato lossy e quello conservato con un file lossless si evidenzia quando si effettuano degli zoom per visualizzare alcuni dettagli dell'immagine.

Le specifiche del formato JPEG prevedono una versione lossless dell'algoritmo che, però, è scarsamente utilizzata.

Il formato JPEG

100%
160k



50%
33k



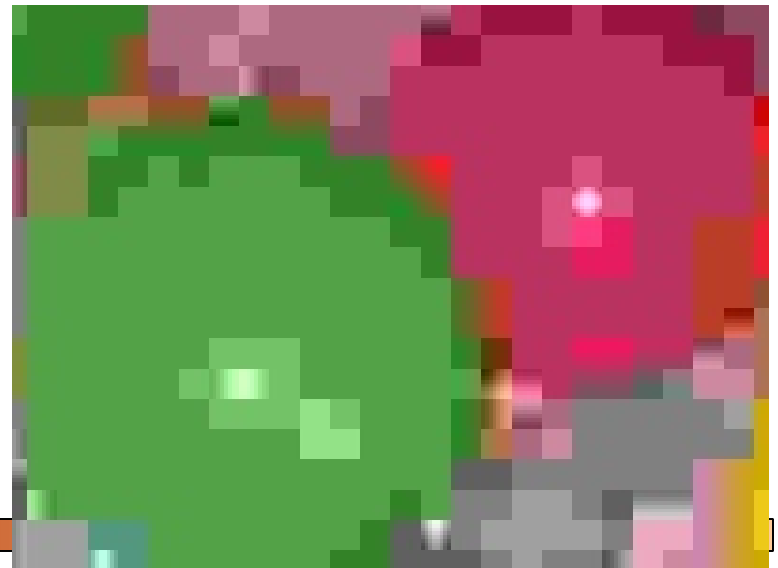
30%
24k



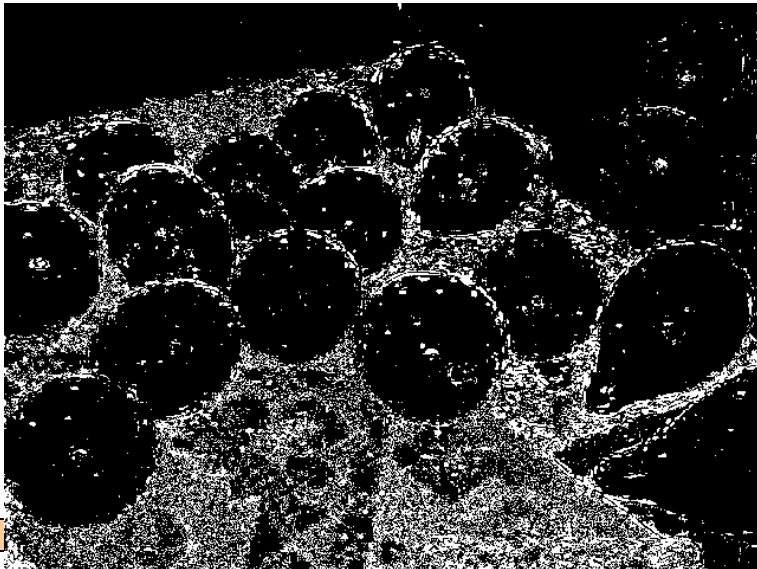
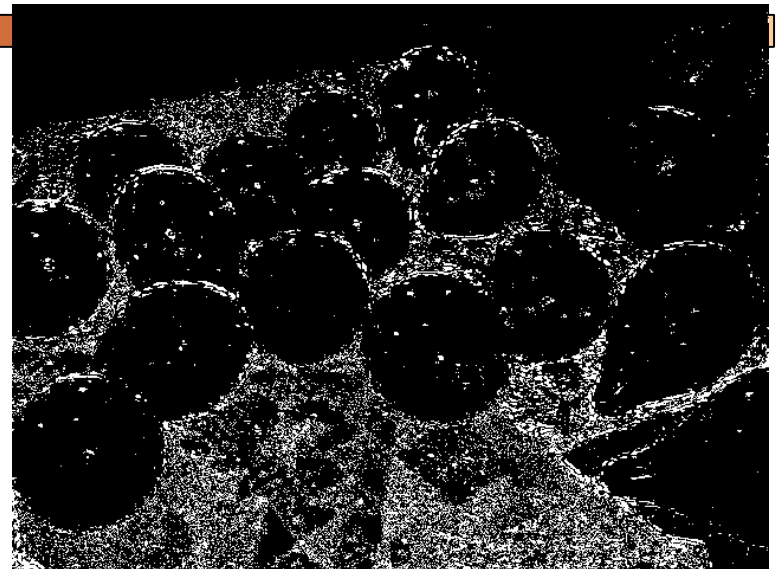
1%
6k



Il formato JPEG



Il formato JPEG



Il formato JPEG

Gif: 2,67 kB



Jpeg: 14,2 kB



Formati di tipo raster o vettoriale

Tutti i formati visti sino ad ora formalizzano come archiviare i singoli pixel che compongono una immagine. Nessuno di questi formati specifica come archiviare la rappresentazione di un oggetto.

L'immagine di una linea nera tracciata sullo schermo sarà così rappresentata da una sequenza di pixel bianchi con, ogni tanto, un pixel nero.

Oltre a questa tipologia di formati (detta **raster**), ne esiste una ulteriore detta **vettoriale**. In questo caso le informazioni vengono rappresentate attraverso un insieme di primitive matematiche che sono punti, linee, curve e poligoni opportunamente colorati.

Formati di tipo raster o vettoriale

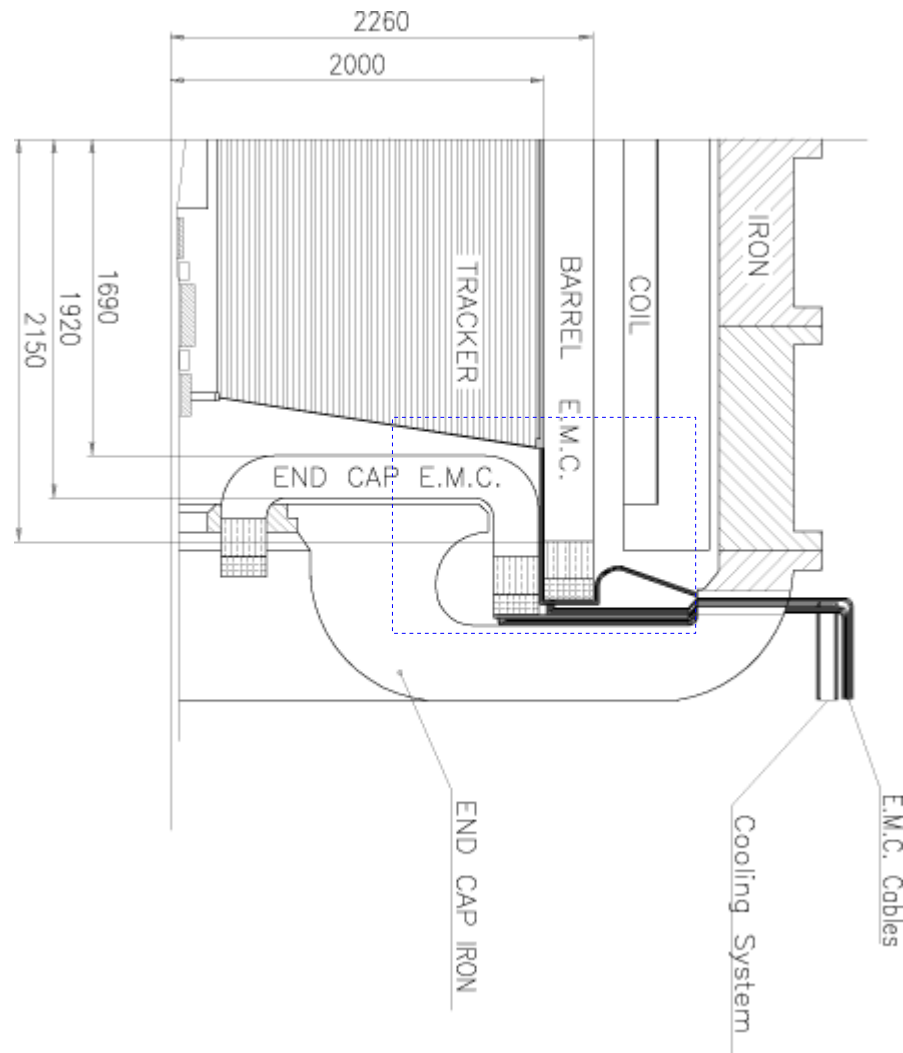
I formati vettoriali trovano una naturale applicazione nell'editoria, nell'architettura, nell'ingegneria e nella grafica tridimensionale.

I vantaggi di un formato digitale sono:

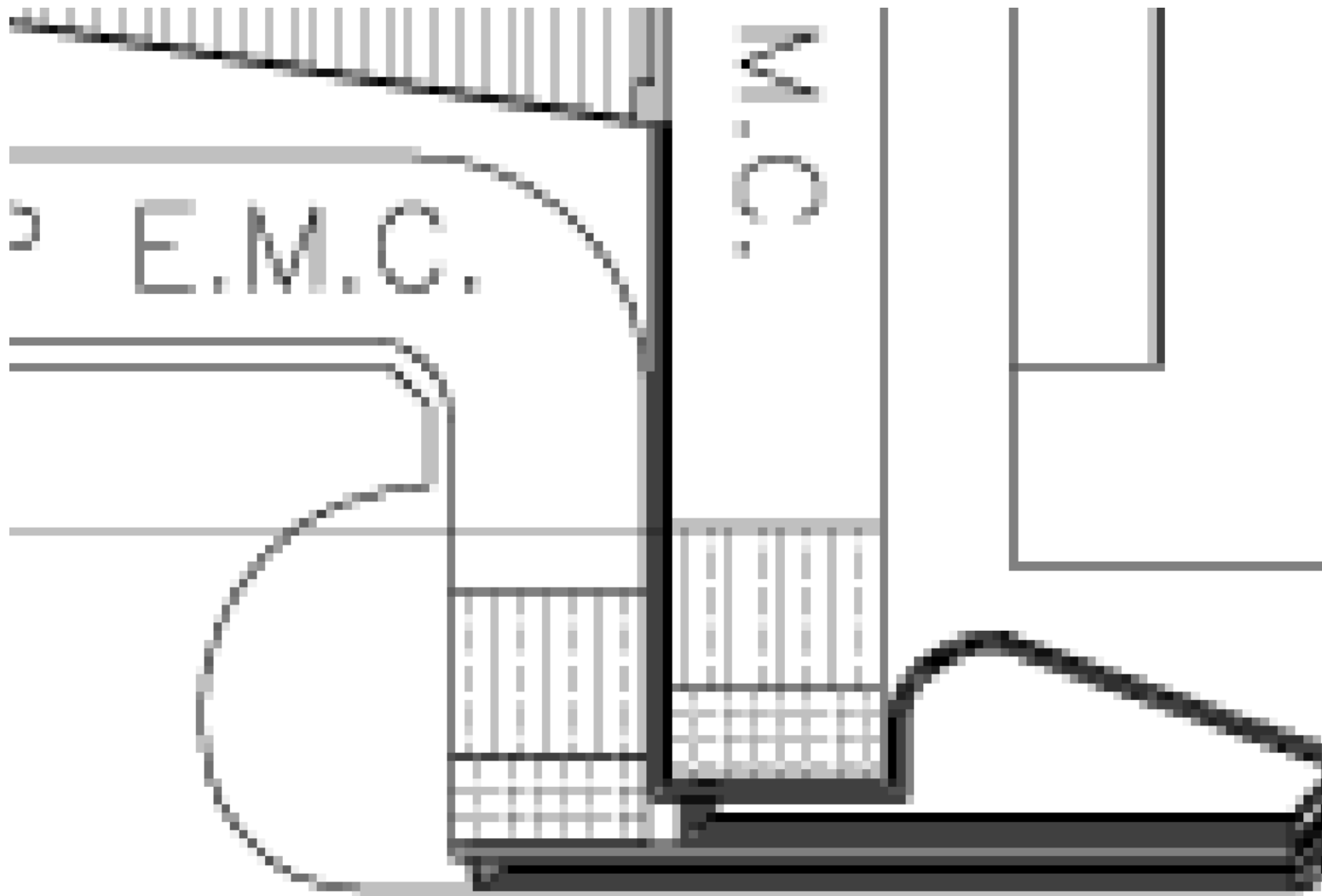
- completa modificabilità del contenuto delle immagini (ad esempio editando un testo di un documento)
- sulle primitive possono essere applicati algoritmi (ad esempio per effettuare delle rotazioni o dei cambi di prospettiva)
- effettuando uno zoom la qualità del dettaglio non degrada.

Sfortunatamente una fotografia archiviata in formato vettoriale ha un peso molto elevato e richiede una grande potenza di calcolo per poter essere elaborata.

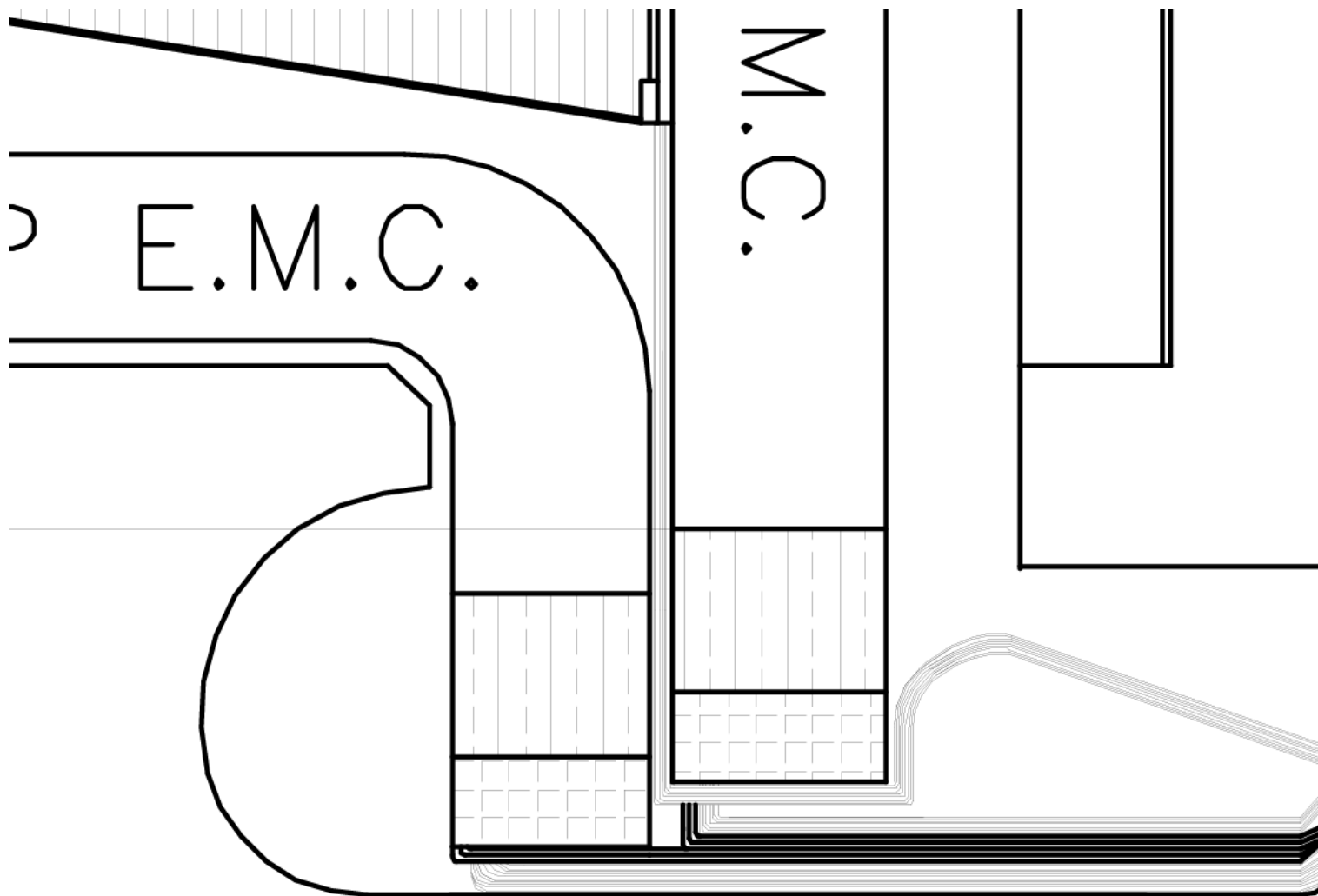
Formati di tipo vettoriale: immagine di partenza



Formati di tipo vettoriale: zoom su immagine raster



Formati di tipo vettoriale: zoom su immagine vettoriale



Il formato JPEG 2000

Il formato JPEG 2000 è una evoluzione del formato JPEG nato per superare alcune delle limitazioni emerse nel suo utilizzo.

Le nuove caratteristiche del formato JPEG 2000 sono:

- Codifica lossless e lossy in un solo algoritmo
- Compressione lossy più efficiente (minor peso del file a parità di qualità percepita dall'utente)
- Robustezza agli errori di trasmissione
- Codifica per regioni di interesse (in una stessa immagine posso avere zone diverse compresse con qualità diverse)
- Compressione lossy senza perdita di qualità cromatica
- Trasmissione progressiva in funzione dell'accuratezza e della risoluzione.

Il formato TIFF

Il TIFF (Tagged Image File Format) è un formato di archiviazione delle immagini di tipo raster molto diffuso in ambito istituzionale per la sua versatilità.

Rispetto ai formati visti in precedenza il TIFF è molto flessibile consentendo di scegliere, al momento dell'archiviazione, tra più algoritmi diversi. Ad esempio:

- l'algoritmo CCITT, utilizzato per l'archiviazione dei fax e delle immagini in bianco e nero
- l'algoritmo JPEG per le immagini foto realistiche
- un algoritmo lossless di qualità inferiore al png

Il formato TIFF

Il TIFF è largamente utilizzato per lo scambio di immagini fra stampanti e scanner perché permette di specificare numerose indicazioni aggiuntive ad esempio le tabelle di calibrazione del colore o di conservare immagini secondo le specifiche RGB, CMYK ed altre.

Un file TIFF può contenere più immagini divise: si possono inserire in un unico file tutte le pagine che compongono un fax o un documento, ovvero gestire più livelli di una singola pagina: i layer.

Queste caratteristiche rendono il formato TIFF adatto per attività di archiviazione dei documenti o per la memorizzazione di grandi quantità di immagini.

Il formato PS

Il PS (PostScript) più che un formato per la memorizzazione delle immagini è un linguaggio di programmazione della pagina, particolarmente adatto alla descrizione di testi ed immagini.

Inizialmente il PostScript era usato come linguaggio per il controllo delle stampanti.

Il PostScript consente di descrivere pagine di testo e grafica in modo indipendente dalla risoluzione e dal dispositivo di visualizzazione.

Un file postscript può quindi essere visualizzato o stampato alla massima risoluzione consentita su una qualsiasi piattaforma compatibile.

Il formato PS

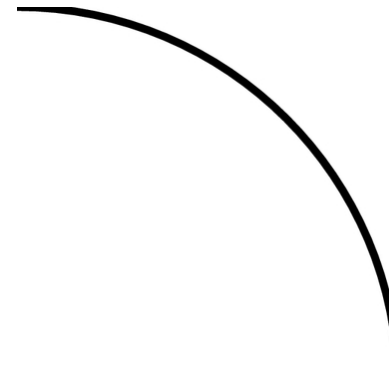
Il PostScript consente di conservare le informazione sulle immagini in formato vettoriale.

Scrive
Hello world!
in courier 20

```
%!PS
/Courier findfont
20 scalefont
setfont
72 500 moveto
(Hello world!) show
showpage
```

disegna un arco con centro nella posizione $x=100$, $y=150$, con raggio di 50 pun

```
100 150 50 0 90 arc
```



Il formato PDF

Il formato PDF (Portable Document Format) è un formato per la conservazione di immagini e testi basato su un linguaggio di descrizione di pagina. Eredita molte delle proprietà del PostScript ma i file pdf sono molto più leggeri e veloci da gestire.

Il formato PDF è probabilmente tra i più noti, ma possiede alcune caratteristiche meno note e molto importanti.

- è standard ISO;
- implementa i layer;
- consente di visualizzare le pagine mentre le altre sono ancora in download;
- è aperto: chiunque può sviluppare un software che scrive o legge file pdf senza dover pagare la Adobe;

Il formato PDF

- consente di memorizzare al suo interno immagini raster (lossless o lossy) o vettoriali, definendo la qualità;
- un file pdf può essere realizzato per massimizzare la fruizione dei suoi contenuti dalle persone diversamente abili;
- all'interno di un file pdf è possibile inserire dei form da riempire e dei programmi scritti con un linguaggio di programmazione simile al javascript.
- il contenuto di un file pdf può essere cifrato o possono essere implementati i DRM: in questo modo la lettura, la modifica, la stampa o la copia del contenuto possono essere abilitate o disabilitate ai diversi gruppi di utenti.

Il video

Sulla base di quanto visto, immaginare un formato per archiviare i video dovrebbe essere molto semplice: è sufficiente visualizzare una sequenza di immagini bitmap per avere un video di eccellente qualità.

In realtà questo non è possibile: infatti il nostro cervello per non percepire che sta vedendo una sequenza di immagini fisse, ma percepire un movimento continuo ha bisogno almeno di 20 fotogrammi al secondo.

La disponibilità di memoria e di banda passante per trasmettere un video composto da 20 immagini bitmap al secondo sarebbe troppo grande per i nostri scopi.

Il video

Vediamo un caso pratico: la codifica PAL utilizzata in Europa per la televisione prevede un video di dimensione 720x576 con 25 immagini (frame) al secondo. Questo significa che, per ogni secondo di video sono necessari:

$720(\text{pixel}/(\text{frame} \cdot \text{riga})) \cdot 576(\text{righe}) \cdot 24(\text{bit}/\text{pixel}) \cdot 25(\text{frame}/\text{secondo}) = 249\text{Mbit}.$

Una banda passante di 250Mbit/s è troppo grande non solo per la rete internet, ma anche per gli altri dispositivi infodomestici. La velocità con cui lettore preleva dati da un CD audio è di 1,4Mbit/s. Nel caso di dati su disco DVD questo valore sale a 5Mbit/s: comunque 50 volte più piccolo.

Anche se riuscissimo a leggere dal DVD a 250Mbit/s, il DVD sarebbe troppo piccolo per contenere un film: al più potrebbe archiviare poche decine di secondi.

Per superare questo problema si può sostituire alla sequenza di immagini bitmap una sequenza di immagini jpeg. Questa scelta, per quanto vada in una direzione giusta, sfortunatamente non consente di risolvere il problema: da 250Mbit/s si passa a circa 30Mbit/s.

Si potrebbe ridurre la qualità delle singole immagini jpeg, ma esistono altre tecniche che consentono di ottenere un risultato migliore.

Un ulteriore miglioramento dell'algoritmo di codifica dei video consiste nel memorizzare, per ogni fotogramma, solo le differenze con quello che lo ha preceduto. In questo modo è possibile non memorizzare nuovamente informazioni già presenti nel file. Ovviamente ogni tanto è necessario memorizzare un fotogramma completo, altrimenti la qualità si degrada.

Oggi quasi tutti i formati utilizzati per la conservazione dei video consentono di definire il rapporto tra il peso del file e la qualità di fruizione percepita dall'utente sulla base di alcuni parametri come: numero di frame al secondo, qualità dei singoli frame conservati, frequenza di memorizzazione di un frame completo o di altri frame di controllo.
