

# Corso di **Sistemi interattivi**

## **Lezione 6. Colore, percezione e teoria delle immagini**

Prof. Rudy Melli (rudymelli@ababrera.it)

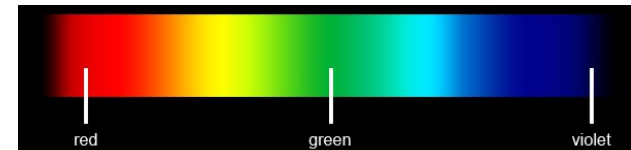
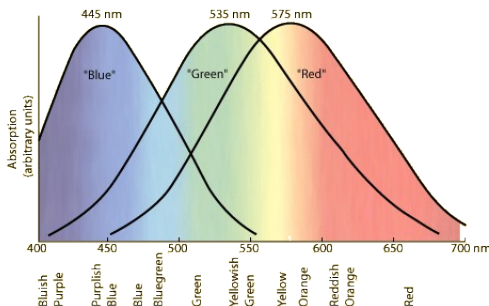
[www.vision-e.it/si](http://www.vision-e.it/si)

**ACCADEMIA DI BELLE ARTI DI BRERA**  
**Anno accademico 2018/2019**

# La visione umana

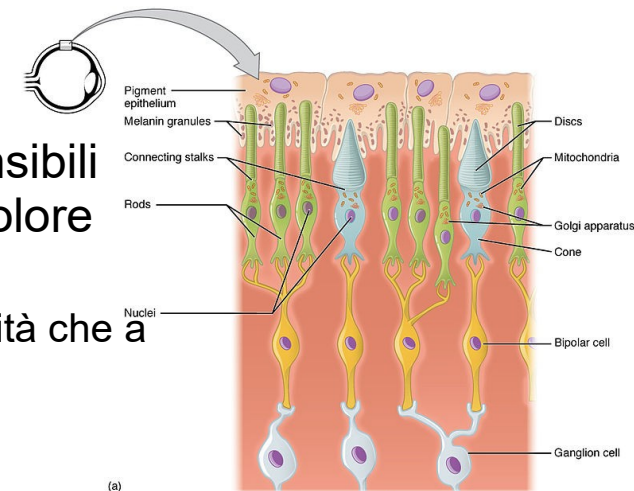
❏ L'occhio umano è sensibile alle onde elettromagnetiche di lunghezza d'onda tra i 380nm a 780nm che costituiscono la luce visibile (o spettro visibile)

- Il colore è una lunghezza d'onda espressa in nanometri
- Lunghezze d'onda diverse vengono interpretate come colori diversi
- Ogni oggetto assorbe alcune lunghezze d'onda riflettendo le altre che stabiliscono il colore. Un oggetto se riflette solo il rosso apparirà rosso, se assorbe tutte le lunghezze apparirà nero, se le riflette tutte bianco



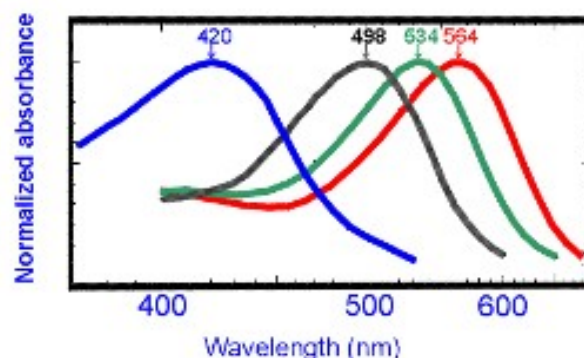
❏ La retina possiede circa 120 milioni **bastoncelli** sensibili alla luce acromatici e 6 milioni di **coni** sensibili al colore concentrati nella fovea

- Maggiore sensibilità al contrasto ed ai cambi di luminosità che a quelli cromatici



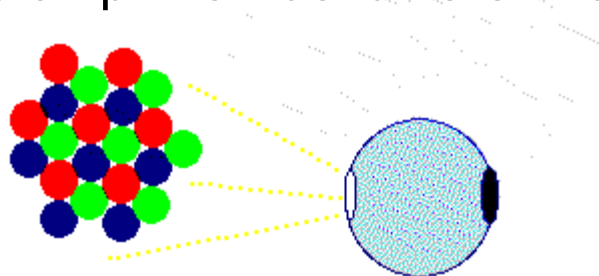
# La visione umana

- Esistono 3 tipi distinti di coni composti da proteine diverse che assorbono solo alcune frequenze cromatiche: rosso, verde e blu
  - Lo spettro di assorbimento (o range di assorbimento) di ogni cono è ampiamente sovrapposto agli altri

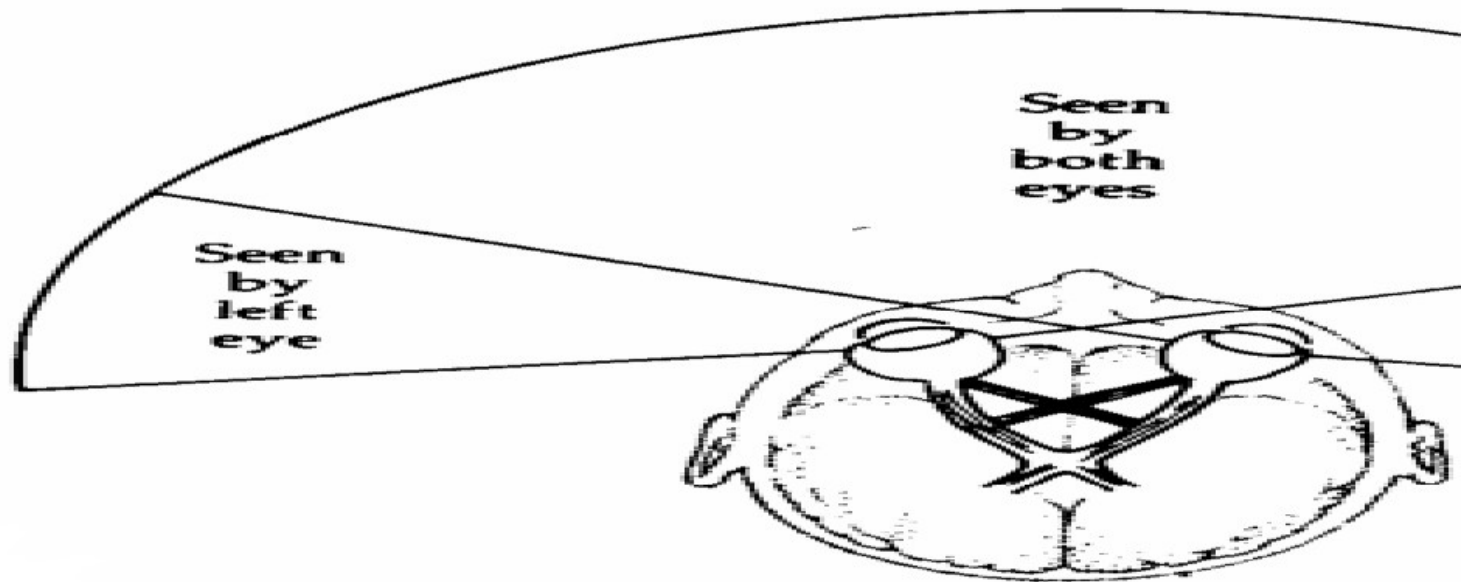


After Bowmaker & Dartnall, 1980

- Il nostro occhio percepisce il colore (o tinta, hue in inglese) come miscela di radiazioni luminose di qualsiasi lunghezza d'onda, combinando i tre colori primari dei differenti coni. Questa proprietà è detta tricromia



# Visione binoculare



**La visione binoculare per  
stima della distanza  
(terza dimensione)**

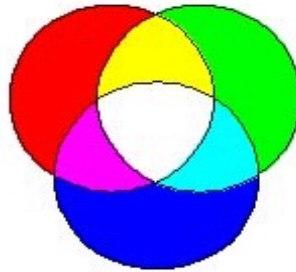
---

# QUANTIFICARE IL COLORE

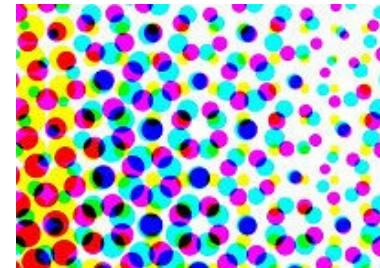
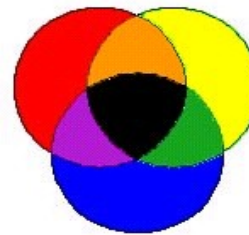
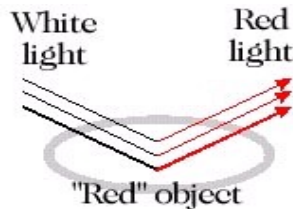
- ◆ Il nostro occhio, a differenza dell'orecchio, non è in grado di effettuare un'analisi spettrale, ma riporta una sensazione risultante dalla combinazione di tutte le lunghezze d'onda visibili
- ◆ E' inoltre impossibile comunicare la sensazione corrispondente ad un certo stimolo. Quello che possiamo comunicare è che due stimoli diversi producono la stessa sensazione.
- ◆ Possiamo definire tre tipi di colori:
- ◆ **colori spettrali** (prodotti da radiazioni contenenti una stretta fascia di lunghezze d'onda), **magenta** o porpora (radiazioni contenenti una combinazione degli estremi dello spettro visibile, cioè blu-viola e rosso-giallo), **bianco** (radiazioni a spettro continuo)
- ◆ La teoria di Young-Helmholtz (1801) ipotizza che per riprodurre la sensazione di un colore si possano combinare opportunamente tre radiazioni primarie (sistemi *tricromici*)
- ◆ La prima conferma sperimentale ottenuta da esperimenti sull'occhio si ebbe nel 1964

# RIPRODUZIONE DEL COLORE

- ◆ Un colore può essere riprodotto con tre tecniche fondamentali in grado di riprodurre lo stesso stimolo:
- ◆ **sintesi additiva:** miscela di luci di differente cromaticità



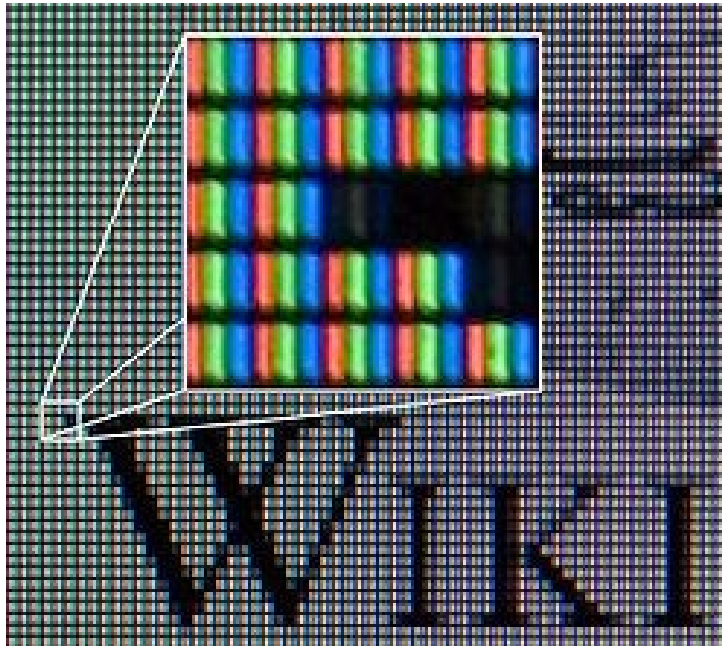
- ◆ **sintesi sottrattiva:** miscela o sovrapposizione di mezzi che assorbono diversamente alcune componenti spettrali



- ◆ **integrazione spaziale:** punti diversamente colorati e vicini che nell'occhio vanno a sovrapporsi producendo un risultato analogo alla sintesi additiva (questo processo si ha nei monitor e nella stampa di inchiostri mediante retinatura)

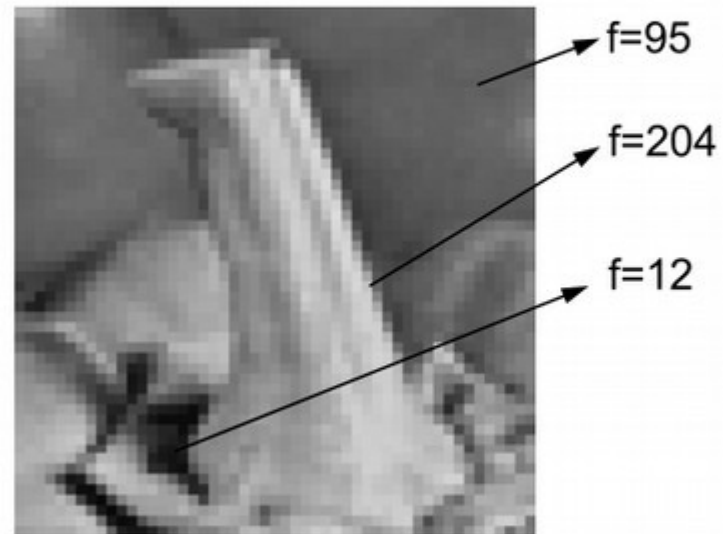
# Integrazione spaziale

- ❑ Punti diversamente colorati e vicini che nell'occhio vanno a sovrapporsi producendo un risultato analogo alla sintesi additiva
- ❑ Questo processo si ha nei monitor e nella stampa di inchiostri tramite retinatura



# Cos'è un'immagine digitale

Un'**immagine digitale** può essere vista come una funzione bi-dimensionale  $f(x,y)$ , dove  $f$  rappresenta l'**intensità** o **livello di grigio** dell'immagine in quel punto: 0 rappresenta il **nero**, 255 il **bianco**

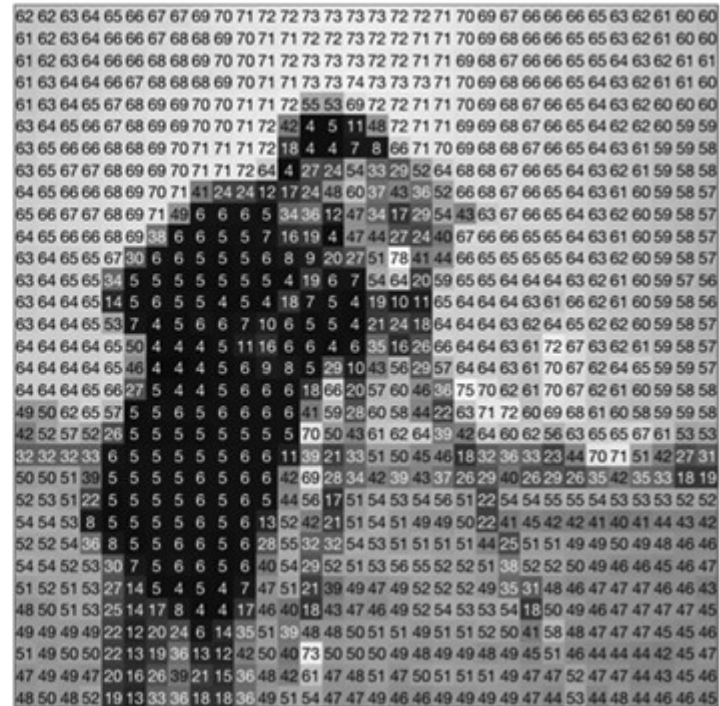


Un'*immagine* è quindi una **matrice** di elementi chiamati **pixels** (picture elements).



# Le immagini digitali

Cosa vede l'occhio umano e cosa vede un computer



# Immagine ed informazioni spaziali

L'immagine è la rappresentazione matriciale della camera del mondo reale

Ogni pixel ha un'informazione spaziale che dipende dalla camera e dai suoi parametri, dalla lente

rappresentazione in spazio 2D di una distribuzione 3D



*Different or similar height?  
The perspective problem*

---

# Video

Sequenza temporale di immagini (frame) normalmente campionata o trasmessa con un intervallo di tempo costante  $\Delta t$  detto **framerate** e misurato in *fps* (frame per secondo)

## Risoluzione

Vi sono 3 tipi di risoluzione:

1) **Spaziale (W,H)**

- Dimensione in pixel dell'immagine, larghezza W per altezza H

2) **Cromatica nbpp**

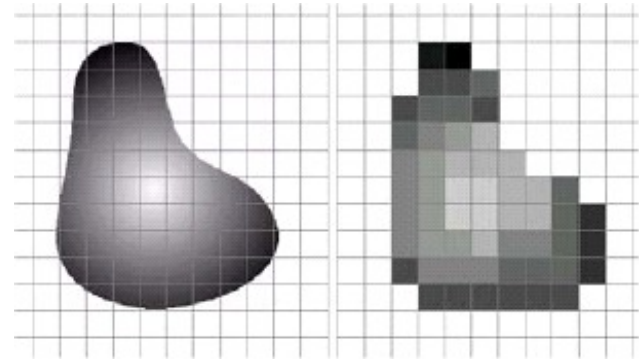
- Numero di bit per pixel, cioè con quanti colori o valori di luminosità posso rappresentare il singolo pixel. Solitamente rappresentata con un numero multiplo di 8 (8, 16, 24)

3) **Temporale fps**

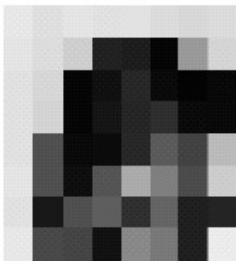
- Vale solo per i video ed è la frequenza di riproduzione, cioè il numero di frame al secondo

# Risoluzione

- ◆ *La risoluzione spaziale* dovrebbe essere legata al piu' piccolo dettaglio di interesse (dpi)
- ◆ In realta' dipende
- ◆ dal sistema di acquisizione
- ◆ Dalla vicinanza dell'oggetto alla camera
- ◆ dal sistema di visualizzazione



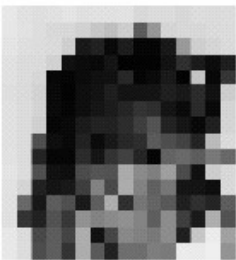
N = 8



N = 64



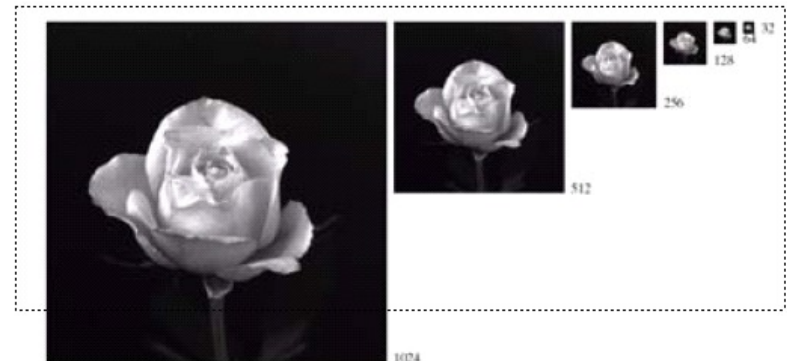
N = 16



N = 128



Stesse dimensioni dell'immagine, varia il numero di pixel.



Stessa dimensione del pixel, varia la dimensione dell'immagine

# Tipi di immagine

 Le immagini sono principalmente:

- A colori (24bit per pixel, 3 canali, 8 bit per canale: RGB)
- In scala di grigi (1 canale, 8 bit per canale)
- Binarie (1 bit per canale, bianco e nero). **ATTENZIONE:** un immagine a 24bit per canale, può contenere anche immagini in scala di grigi o in bianco e nero, infatti con molti dati posso anche rappresentarne pochi (basta mettere in R, G e B lo stesso valore per rappresentare tutti i livelli di grigio).

Un pixel a scala di grigi in RGB si rappresenta inserendo il medesimo valore nelle 3 componenti:


se  $p_{BN} = (128)$  (valore luminosità del pixel  $p$ ) in RGB diventa  $p_{RGB} = (128, 128, 128)$

# Immagine in bit

## Un'immagine è descritta da

- Il numero di canali  $nc$  : un'immagine a scala di grigi ha 1 canale (il canale luminosità), un'immagine RGB ha 3 canali (rosso, verde e blu)
- Il numero di bit per canale  $bpc$  (del pixel): 1, 8, 16, 32. Cioè con quanti bit rappresento l'informazione legata al singolo canale
- Il numero di bit per pixel totale  $bpp$  (1, 8, 16, 24, 32) che è il risultato del numero di bit per canale moltiplicato per il numero di canali

$$bpp = bpc * nc$$

 Un'immagine a scala di grigi può essere salvata in formato 24bit oppure 8bit. Nel primo caso sarà un'immagine a 3 canali e per ogni pixel, le 3 componenti RGB conterranno il medesimo valore.

---

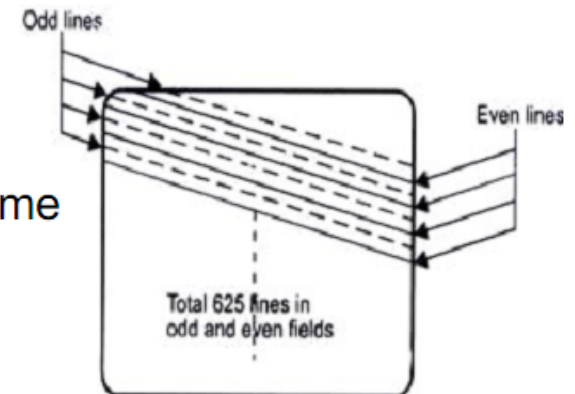
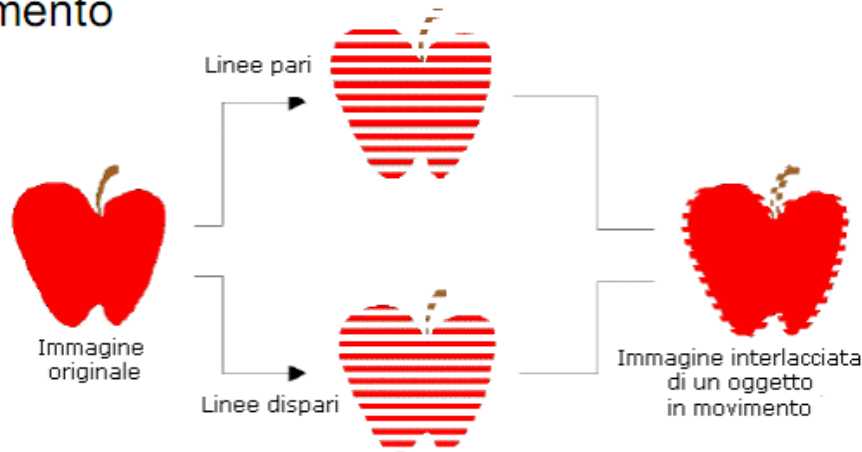
# Occupazione di memoria

- Un immagine rgb (3 canali) di dimensione  $n \times m$  NON compressa occupa esattamente:  
$$n * m * 3 \text{ bytes}$$
  - Ad esempio un immagine rgb 640x480 occupa:  
 $640 * 480 * 3 = 921600 \text{ byte} \rightarrow \sim 1 \text{ MByte}$
- Quando un'immagine compressa viene letta da disco per essere visualizzata viene copiata in memoria RAM l'intera immagine decompressa con un'occupazione pari alla formula sopra



## Interlacciamento

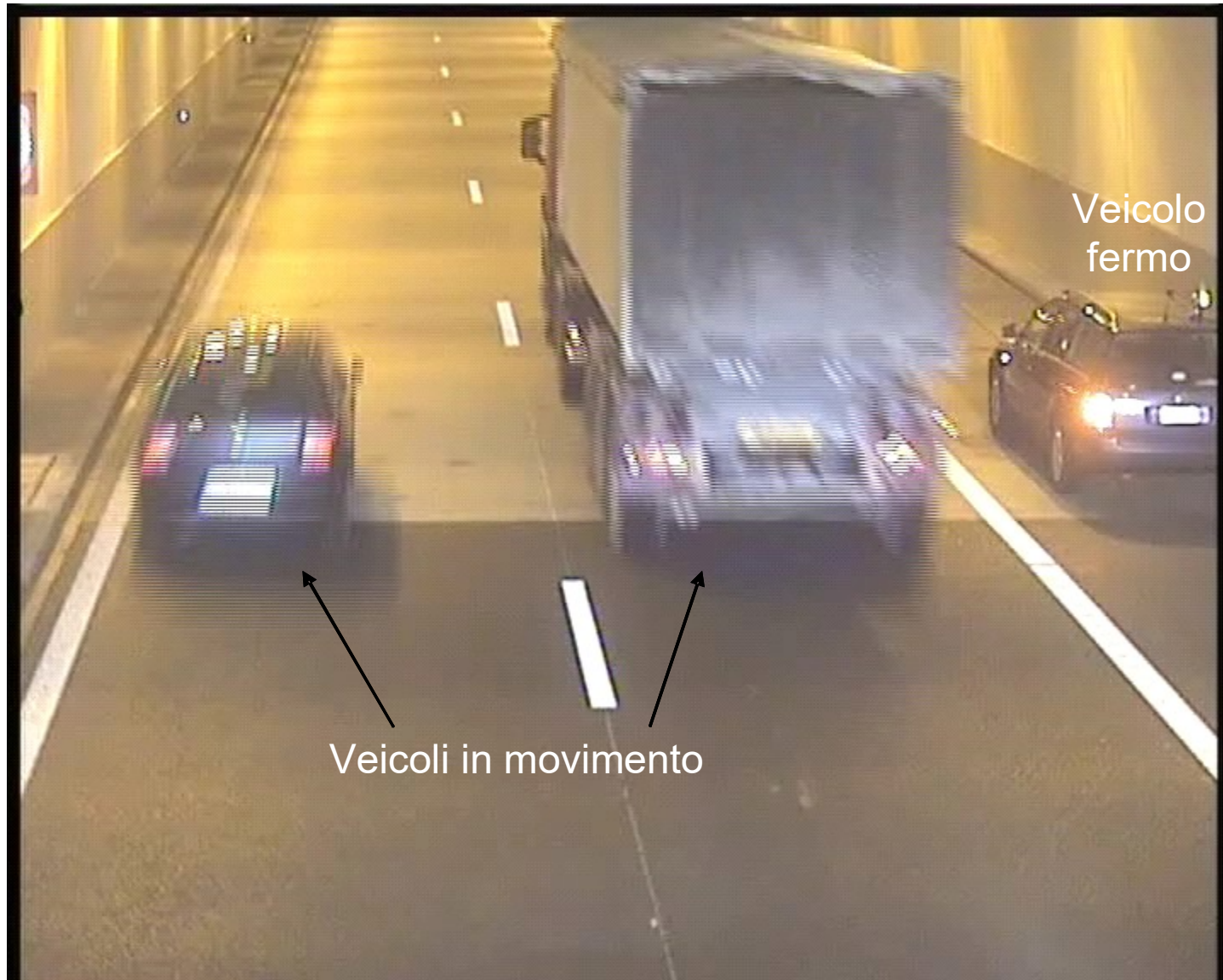
- ◆ Introdotto nel 1934 per aumentare la fluidità di visione preservando la definizione (in linee) → la tecnologia di allora non permetteva di aumentare la velocità di acquisizione
- ◆ Framerate invariato → introduzione di effetto blur motion
- ◆ Ogni frame è separato in 2 campi
  - Le righe pari del campo al tempo  $t$ , quindi
  - Le righe dispari del campo al tempo  $t + \Delta t$
- ◆ In ricezione i 2 campi vengono uniti per formare il frame
  - Problemi di deinterlacciamento per immagini in movimento



- ◆ I segnali di sincronismo discriminano tra tipo di righe ed allineano il frame in modo che non vi siano spostamenti orizzontali o verticali dell'immagine

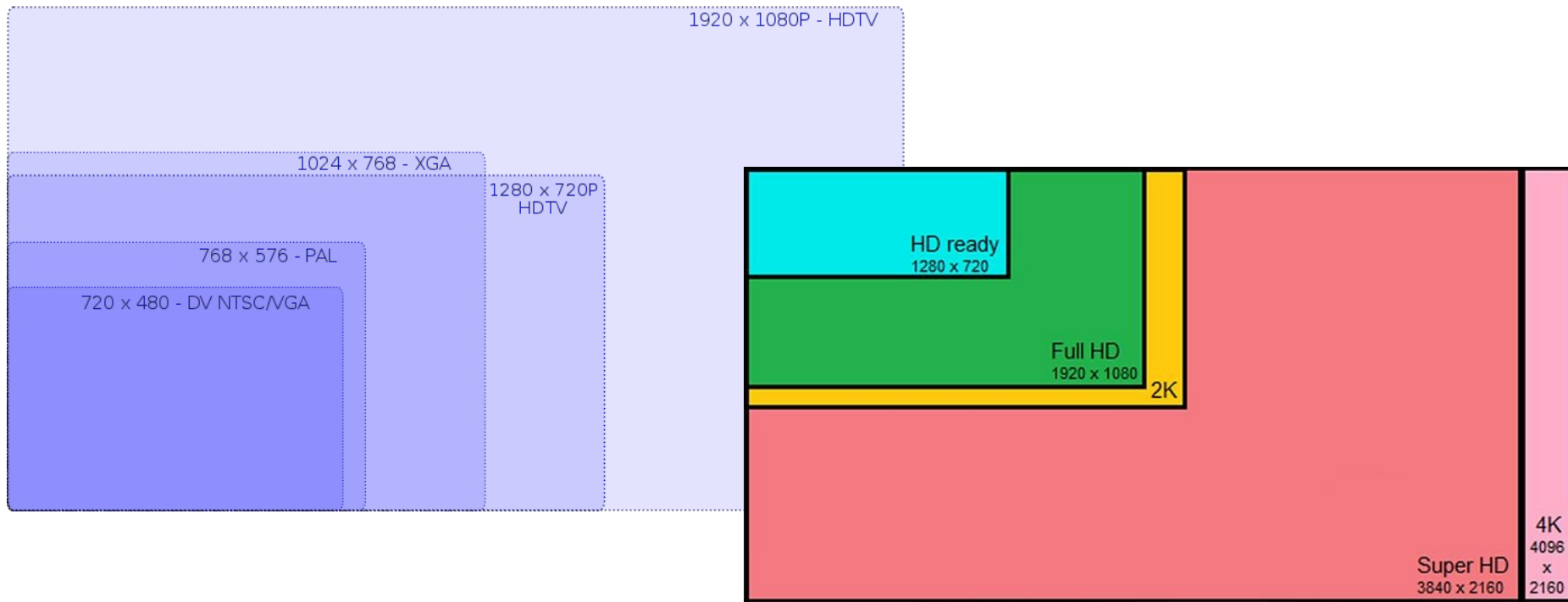


## Esempio di frame interlacciato



# Formati video

- ▶ DVD -> 720x576, 704x576 PAL – 720x480, 704x480 NTSC
- ▶ HDTV (720p) 1280x720
- ▶ Full HD (1080p) 1920x1080
- ▶ SuperHD o QuadHD (2160p) 3840x2160
- ▶ La *i* dopo la risoluzione verticale (720i, 1080i) indica che la sorgente video è interlacciata.



# Formati Immagini

NOME	COMPRESSIONE	DESCRIZIONE	Note
BMP	Non compresso	8 e 24bit, nessuna perdita di informazione	Formato nativo di windows
GIF	Compressione a 256 livelli	Memorizza la palette dell'immagine in modo da minimizzare lo spazio occupato. E' a perdita di informazione	Ottimo per immagini con pochi colori o testo bianco e nero
JPG	Compressione lossy	A perdita di informazione. Ha un'eccezionale livello di compressione	Ideale per immagini ad alta definizione
PNG	Compressione loss-less	Senza perdita di informazione	Immagini ad alta definizione di cui si vuole mantenere tutte le informazioni

# Formati Video

**AVI NON è un formato di compressione ma un contenitore di stream audio/video che possono essere compressi o no**

NOME	COMPRESSIONE	DESCRIZIONE	Note
Raw	Non compresso	Formato di dati raw, non compresso	
Divx/ Xvid	Alta compressione	Compressione basata sul formato MPEG4 in grado di ridurre sensibilmente l'occupazione di spazio preservando la qualità.	Da utilizzare per archiviare grandi quantità di ore video
Xvid	Alta compressione	Con le stesse caratteristiche del Divx ma Open source	Ottima qualità
H264/ Mpeg4/ H265	Alta compressione	Codec recente (H265) dalla compressione più performante (minor spazio, H265 circa il 40% in meno di H264) a parità di qualità	Ottima qualità. Tempi di compressione più lunghi di xvid

**NB: Durante la creazione di una patch che lavora su un filmato, tenete conto della compressione del file, più un file è compresso più il PC dovrà lavorare per decomprimerlo per visualizzarlo, sacrificando le prestazioni della vostra installazione.**

---

## Ridurre un filmato

**La dimensione di un filmato può risultare molto oneroso per il computer, in particolar modo con computer non recenti.**

**La caratteristica principale da ridurre è la risoluzione, cioè la dimensione del frame:**

- **Nel caso di acquisizione ed elaborazione in tempo reale da sorgente video esterna si consiglia di non superare il formato VGA (640x480)**
- **Nel caso di riproduzione di filmati si consiglia di non superare il formato VGA o HD720p per impedire scatti o interruzioni**

**ATTENZIONE: la dimensione in byte di un filmato non influiscono direttamente sulle prestazioni per l'acquisizione o l'elaborazione**

- **un video di pochi MB FullHD è più oneroso computazionalmente di un video VGA più grande in termini di MB**

## Esercizi

- ❏ Aprire un filmato e scrivere a video il framerate usando la funzione *text()*
- ❏ Visualizzarlo lo stesso filmato visualizzando il framerate ma in fullscreen
- ❏ Applicare al filmato un filtro con la sottofunzione *filter*  
(ad es. *m.filter(THRESHOLD, 0.5);*)
- ❏ Alla pressione del mouse o del tasto SPAZIO, salvare un frame della finestra di rendering con la funzione *saveFrame()* in vari formati: jpg, png, gif, ...
- ❏ Salvare il contenuto della finestra di rendering come video usando la libreria VideoExport