CS 414 – Multimedia Systems Design

Lecture 5 – Digital Video Representation

Klara Nahrstedt

Spring 2014

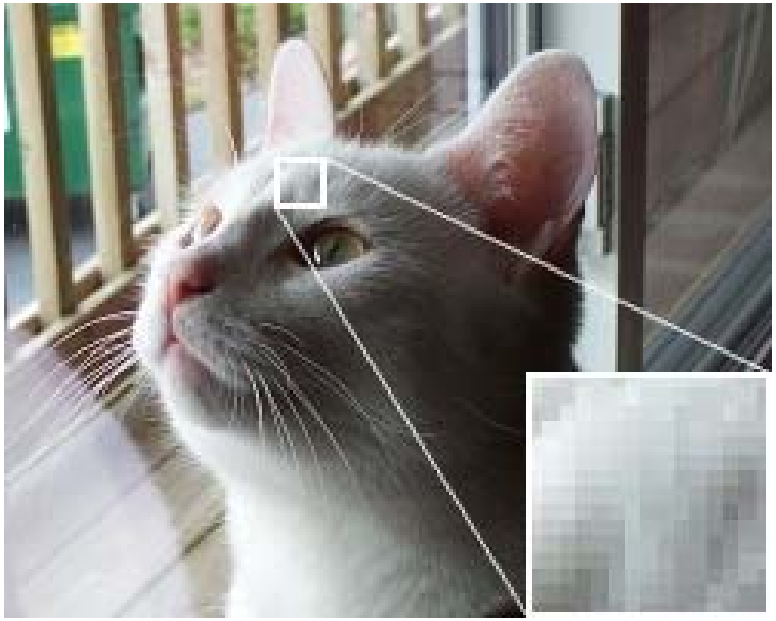
Administrative

- MP1 will be out (February 2)
  - Deadline of MP1 is **February 19 (Wednesday), 5pm**
  - You can have 2 bonus day if needed (just keep in mind that you can have totally 3 bonus days for all three MPs)
  - Submit via compass
- **MP1 discussion during Lecture on February 7 (Friday)**

Today Introduced Concepts

- Digital Image Representation
  - Quantization, Color Issues, Image Size
- Video – Additional Visual Perception Dimensions
  - Resolution, Brightness, Temporal Resolution
- Television
  - Analog, Digital
  - NTSC, HDTV, …

Color Quantization
Example of 24 bit RGB Image



24-bit Color Monitor

Image Representation Example

24 bit RGB Representation (uncompressed)

128	135	166	138	190	132
129	255	105	189	167	190
229	213	134	111	138	187

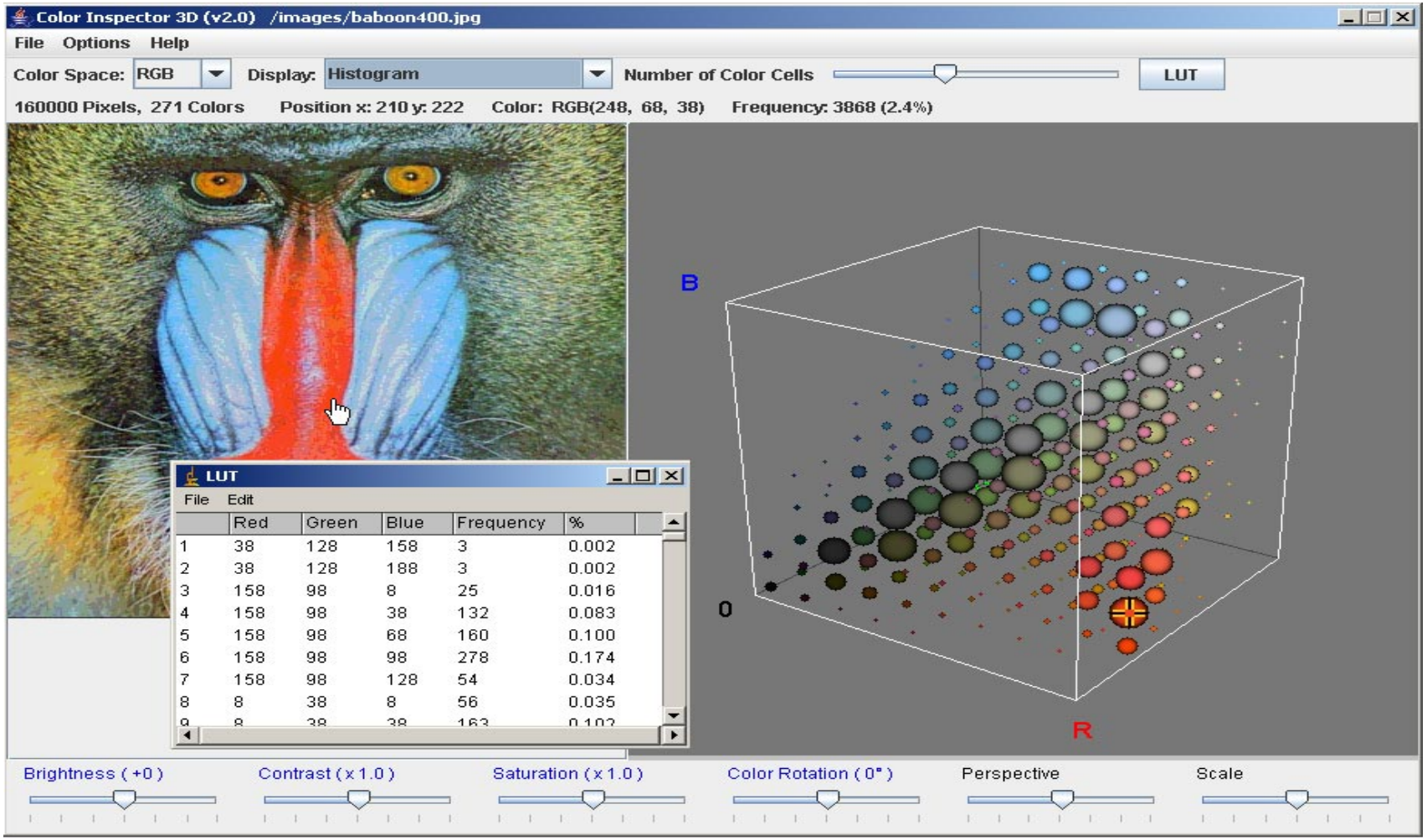
128	138
129	189
229	111

135	190
255	167
213	138

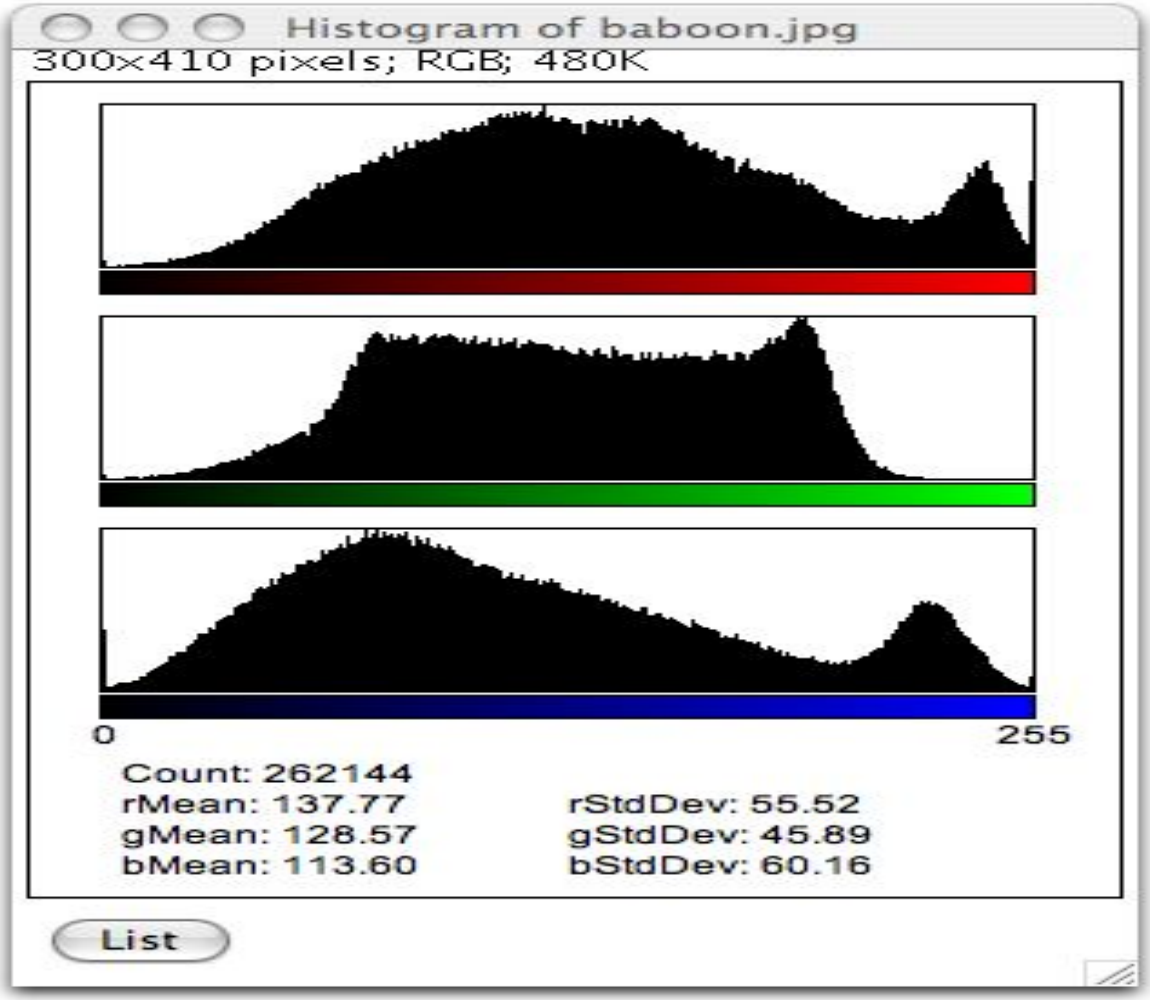
166	132
105	190
134	187

Color Planes

Image Properties (Color)



Color Histogram



Spatial and Frequency Domains

- Spatial domain
  - refers to planar region of intensity values at time t

- Frequency domain
  - think of each color plane as a sinusoidal function of changing intensity values
  - refers to organizing pixels according to their changing intensity (frequency)

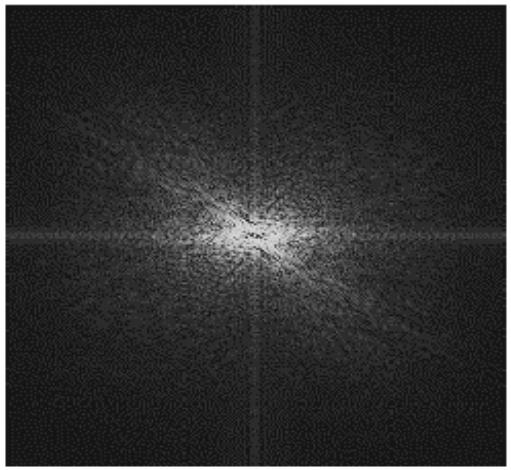


Image Size (in Bits)

- Image Size = Height x Width X Bits/pixel

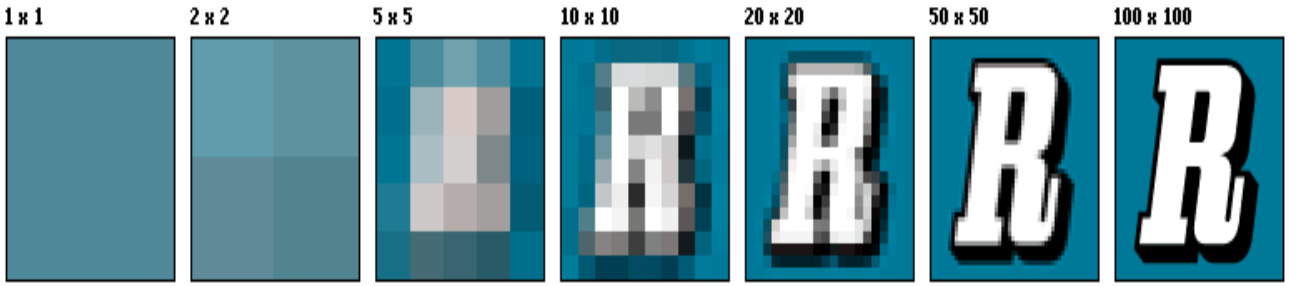
- Example:
  - Consider image 320x240 pixels with 8 bits per pixel
  - Image takes storage 7680 x 8 bits = 61440 bits or 7680 bytes

What is 2D Video?

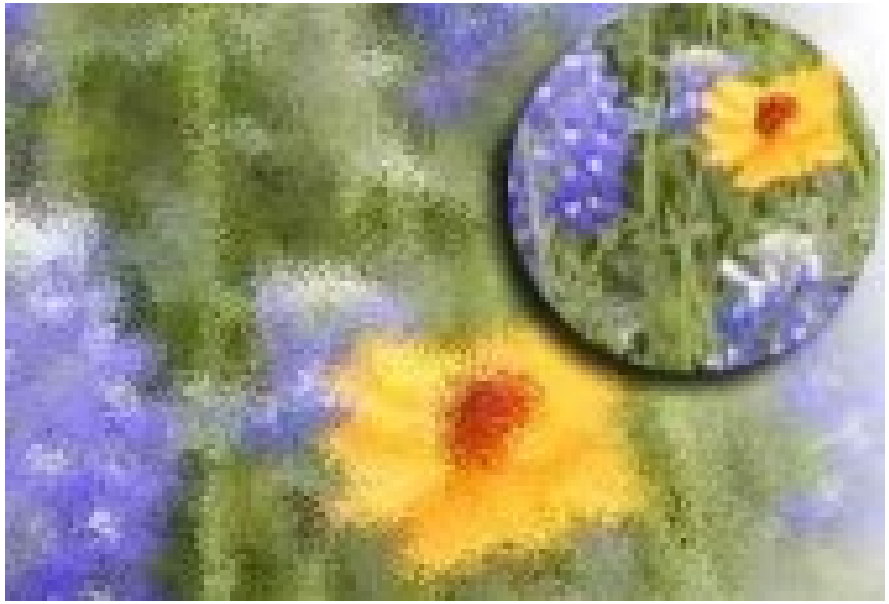
- 300 image frames

Visual Perception: Resolution and Brightness

- **Visual Resolution**
  (depends on:)
  - Image size
  - Viewing distance



- **Brightness**
  - **Perception of brightness is higher than perception of color**
  - Different perception of primary colors
    - Relative brightness: green:red:blue=
      59%:30%:11%

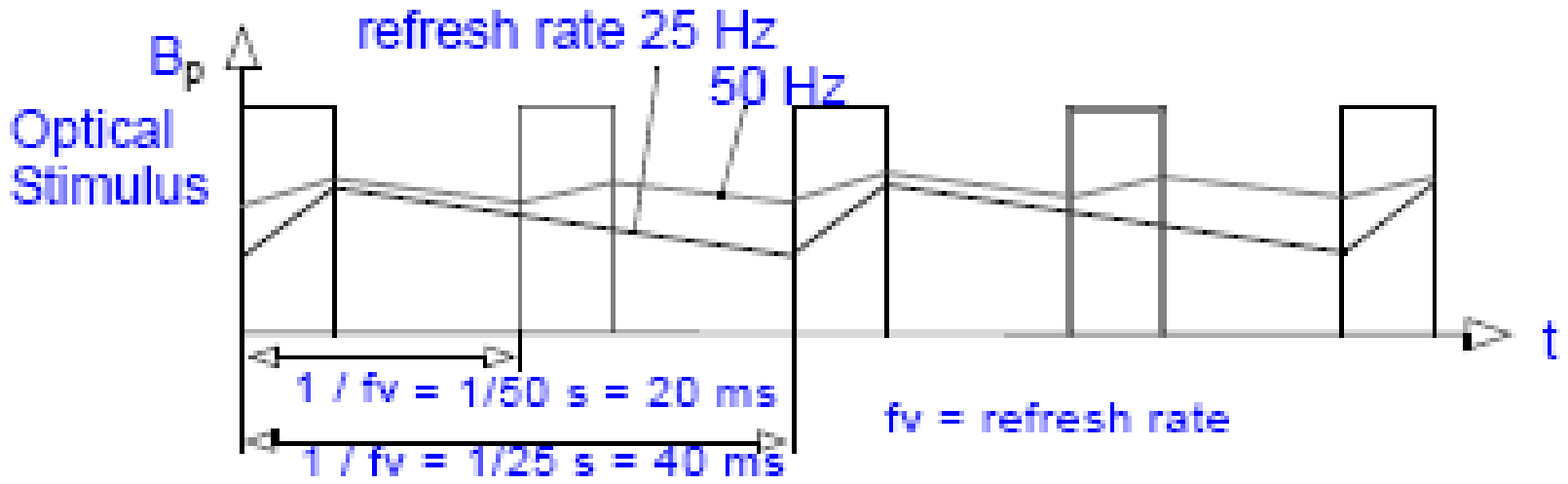


Source: wikipedia

Visual Perception: Resolution and Brightness

Visual Perception: Temporal Resolution

- *Effects caused by inertia of human eye*
- *Perception of 16 frames/second as continuous sequence*
- *Special Effect: Flicker*



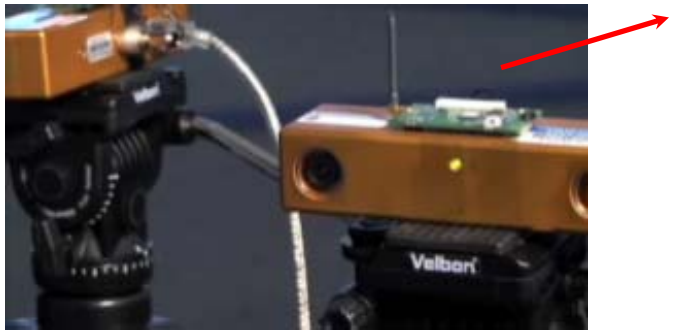
Temporal Resolution

- **Flicker**
  - Perceived if frame rate or refresh rate of screen too low (<50Hz)
  - Especially in large bright areas
- **Higher refresh rate requires:**
  - Higher scanning frequency
  - Higher bandwidth

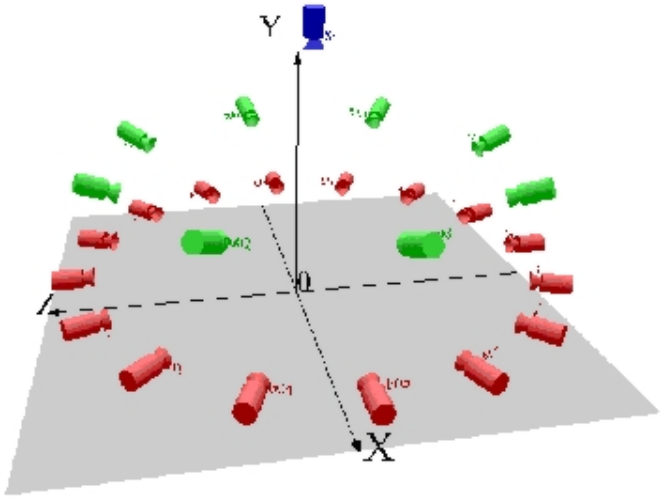
Visual Perception Influence

- **Viewing distance**
- **Display ratio**
  - width/height – 4/3 for conventional TV
  - Width/height – 16/9 for HDTV
- Number of details still visible
- **Intensity** (luminance)

3D Video



- Stereo video and Free-viewpoint video

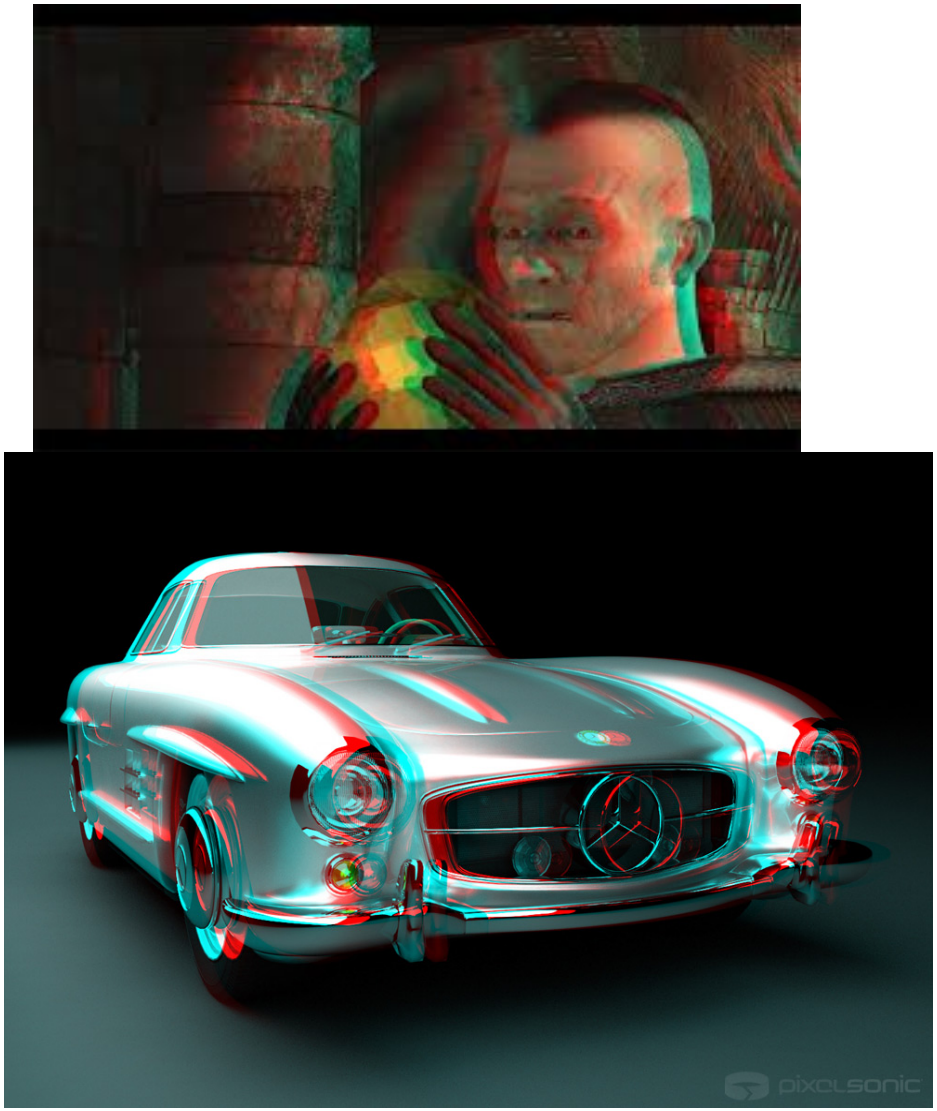


Stereo 3D Image

- Most stereoscopic methods present two offset images separately to the left and right eye of the viewer.
- These two-dimensional images are then combined in the brain to give the depth perception
- Visual requirements for 3D video
  - Simultaneous perception
  - Fusion (binocular ‘single’ vision)
  - Stereopsis (impression of depth)

3D Image

- Binocular viewing of scene creates
  - Two slightly different images of the scene in the two eyes due to the eyes' different positions on the head
  - These differences, referred to as binocular disparity, provide information that the brain can use to calculate depth in the visual scene, providing depth perception

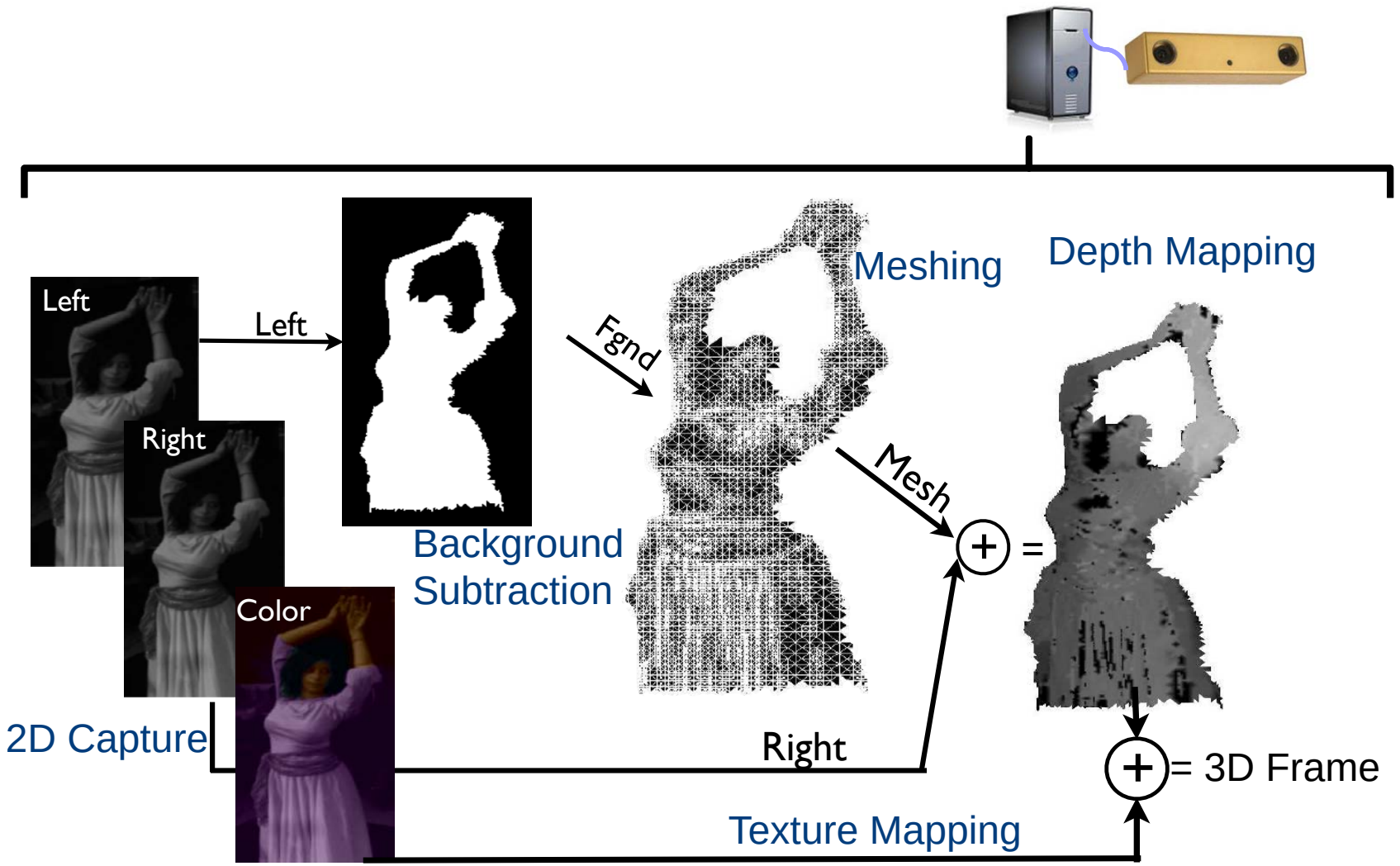


Stereoscopic Image
http://www.pixelsonic.com/2011/04/mercedes-300-sl-stereoscopic/

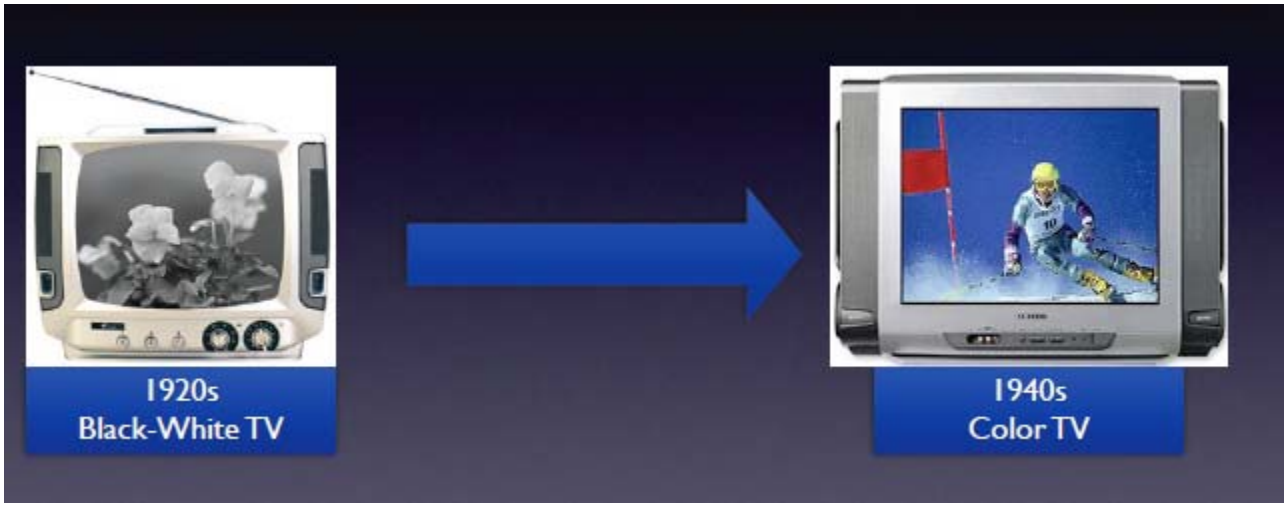
3D Image/Video - Depth Perception

- Depth perception
  - Visual ability to perceive world in 3D and the distance of an object
  - Depth sensation is corresponding term for animals (it is not known whether they perceive depth in the same subjective way that humans do)
- Depth cues
  - Binocular cues that are based on receipt of sensory information in 3D from both eyes
  - Monocular cues that can be represented in just 2D and observed (depth) with just one eye.

3D Teleimmersive Video



Television History (Analog)

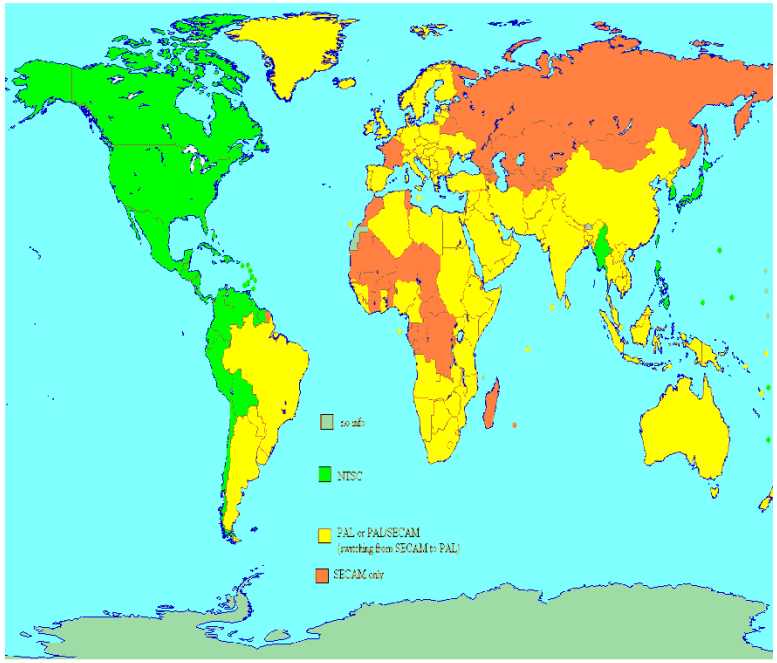


- 1927, Hoover made a speech in Washington while viewers in NY could see, hear him

- AT&T Bell Labs had the first "television"
  - 18 fps, 2 x 3 inch screen, 2500 pixels

Analog Television Concepts

- Production (capture)
  - 2D
  - structured formats
- Representation and Transmission
  - popular formats include NTSC, PAL, SECAM
- Re-construction
  - scanning
  - display issues (refresh rates, temporal resolution)
  - relies on principles of human visual system

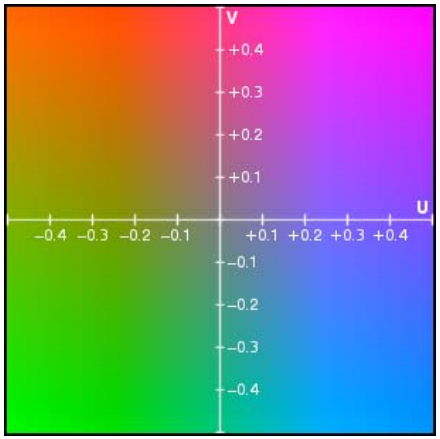


Color Space: YUV

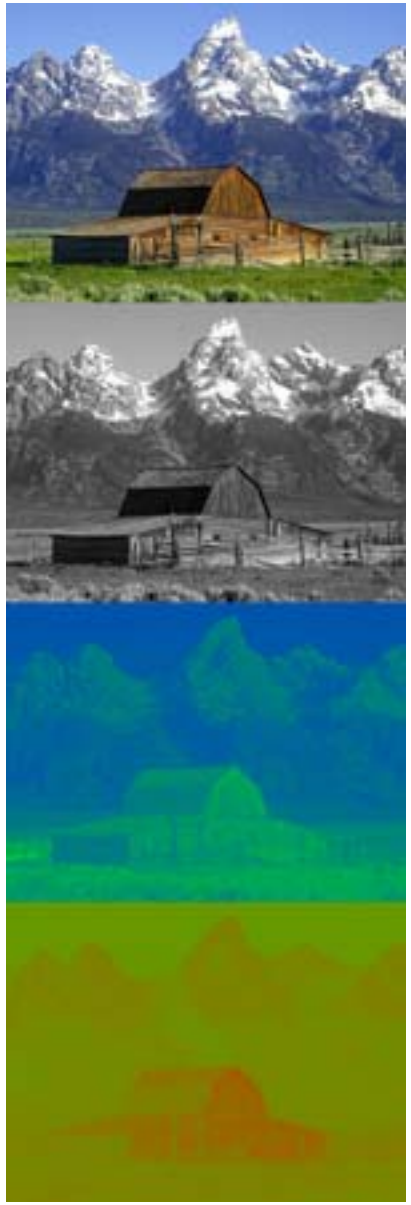
- PAL video standard
  - Y is luminance
  - UV are chrominance

- YUV from RGB

  Y = .299R + .587G + .114B
  U = 0.492 (B - Y)
  V = 0.877 (R - Y)



U-V plane at Y=0.5

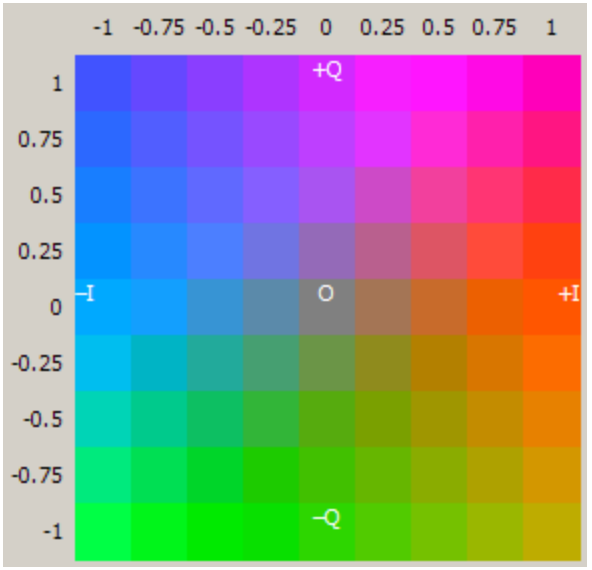


Source: wikipedia

YIQ (NTSC)

- YIQ from RGB

$$Y = .299R + .587G + .114B$$
$$I = .74 (R - Y) - .27 (B - Y)$$
$$Q = 0.48 (R - Y) + 0.41 (B - Y)$$

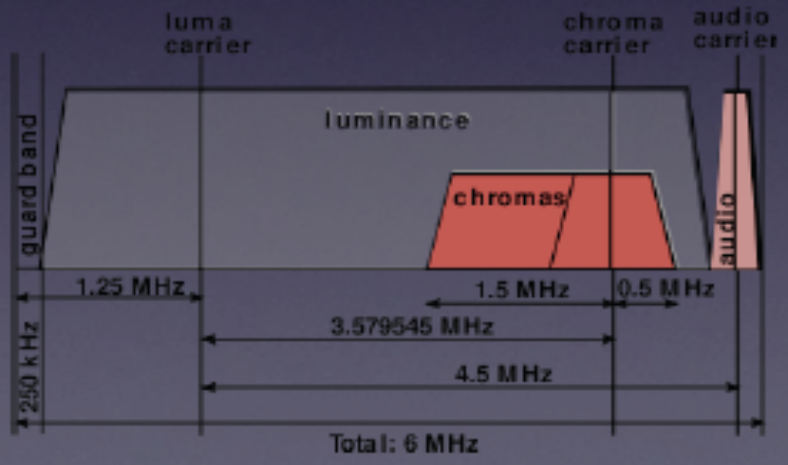


YIQ with Y=0.5

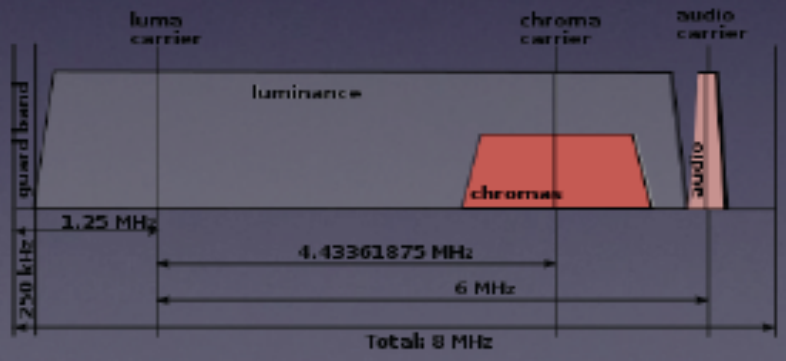
Source: wikipedia

Video Representations

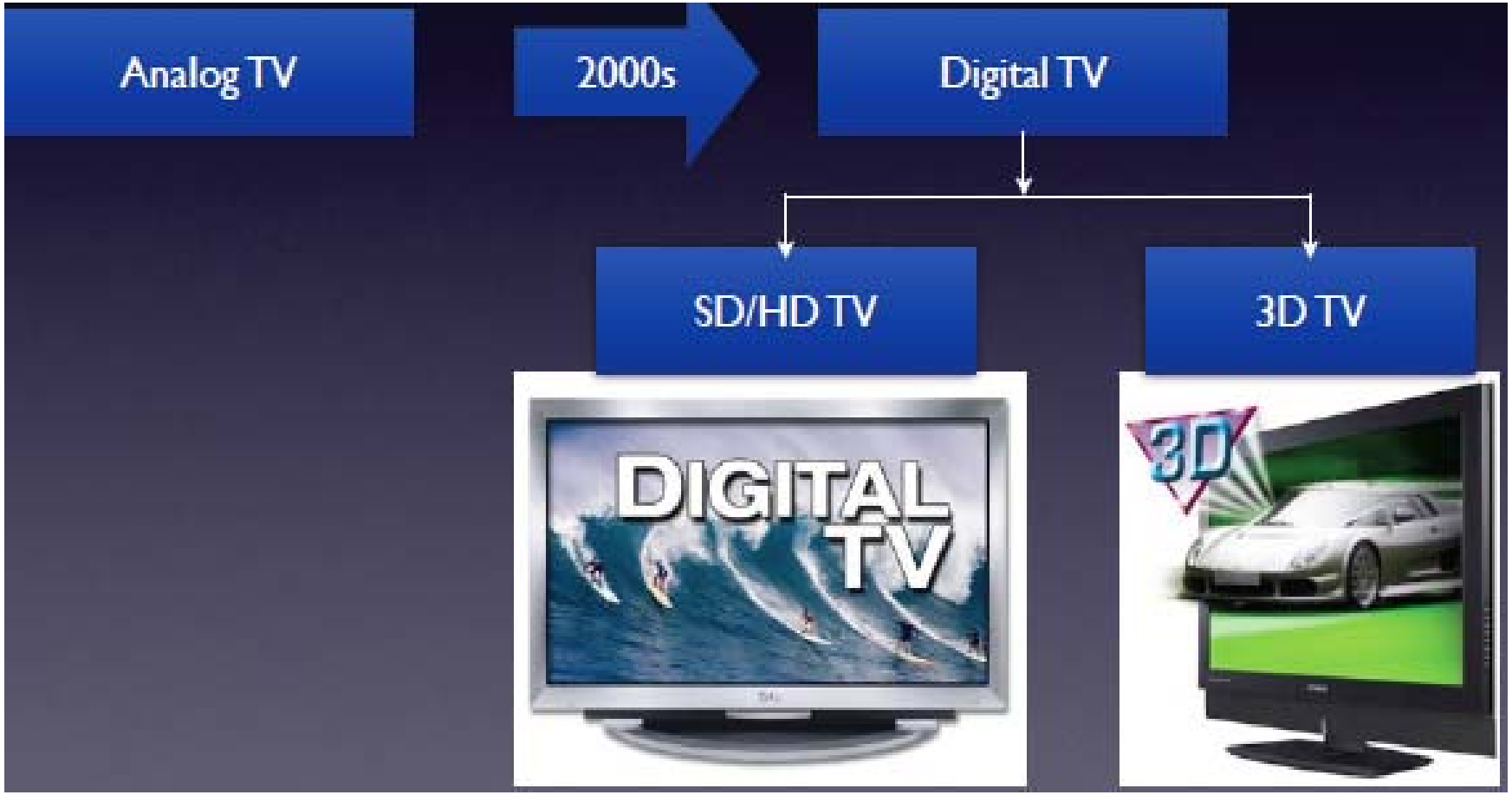
- NTSC (USA, Japan, etc)
  - 525 lines, 59.94Hz
  - Interlaced, 480 lines visible, 29.97 fps (480i)



- PAL (Europe, China, etc)
  - 625 lines, 50Hz
  - Interlaced, 576 lines visible, 25 fps (576i)



TV History



HDTV (Digital)

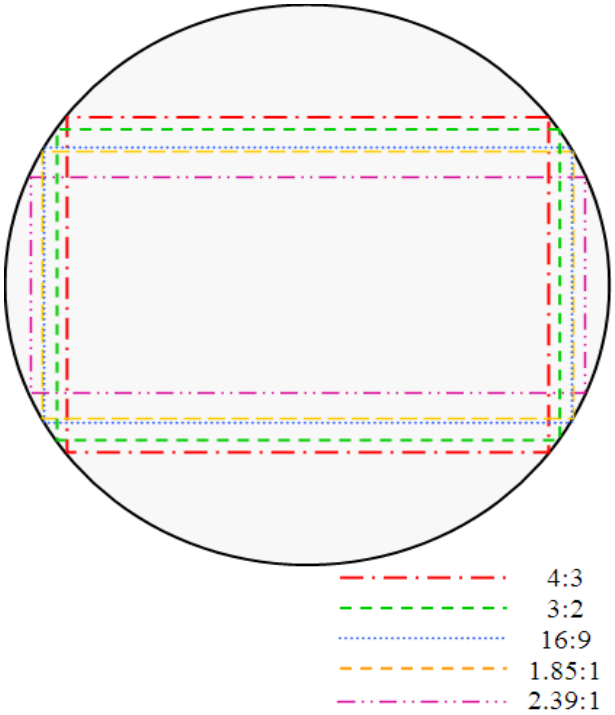
- Resolutions:
  - 1920x1080 (1080p) – Standard HD (HDTV)
  - 2160p, …
  - 4096x2304 (4096p) – 4K High HD
- Frame rate:
  - HDTV - 50 or 60 frames per second
  - HDTV – 120 fps

HDTV

- Interlaced (i) and/or progressive (p) formats
  - Conventional TVs – use interlaced formats
  - Computer displays (LCDs) – use progressive scanning
- MPEG-2 compressed streams
- In Europe (Germany) – MPEG-4 compressed streams

Aspect Ratio and Refresh Rate

- Aspect ratio
  - Conventional TV is 4:3 (1.33)
  - HDTV is 16:9 (2.11)
  - Cinema uses 1.85:1 or 2.35:1
- Frame Rate
  - NTSC is 60Hz interlaced (actually 59.94Hz)
  - PAL/SECAM is 50Hz interlaced
  - Cinema is 24Hz non-interlaced



Source: wikipedia

Resolution

	Broadcast Resolution	Aspect Ratio	Pixel Aspect Ratio	Display Resolution
480i (SD NTSC)	704x480	4:3	10:11	640x480
480i (SD NTSC)	704x480	16:9	40:33	854x480
576i (SD PAL)	704x576	4:3	12:11	768x576
576i (SD PAL)	704x576	16:9	16:11	1024x576
720p (HD)	1280x720	16:9	1:1	1280x720
1080p (Full HD)	1920x1080	16:9	1:1	1920x1080

Digital Video and TV

- Bit rate: amount of information stored per unit time (second) of a recording
- Color Coding: YCrCb
  - Subset of YUV that scales and shifts the chrominance values into range 0..1

$$Y = 0.299R + 0.587G + .114B$$
$$Cr = ((B-Y)/2) + 0.5$$
$$Cb = ((R-Y)/1.6) + 0.5$$

Digital Video and TV

- Color space compression
  - YUV444
    - 24 bits per pixel
  - YUV422
    - 16 bits/pixel
  - YUV411
    - 12 bits/pixel

Digital Video and TV

- DVD video
  - Since 1997
  - Resolution and frame rate
    - 704x480 at 29.97 fps
    - 704x576 at 25 fps
  - Bitrate: 9.8 Mbps

Digital Video and TV

- Blu-ray video
  - since 2006
  - Resolution and frame rate
    - 1920i (@59.94 fps) – interlaced
    - 1920p (@24 fps) – progressive
    - ….
  - Bitrate : 40 Mbps

3DTV

- Refresh rate no less than 120 Hz
- Synchronized shutter glasses to enable different views for different eyes

Summary

- Digitization of Video Signals
  - Composite Coding
  - Component Coding
- Digital Television (DTV)
  - DVB (Digital Video Broadcast)
    - Satellite connections, CATV networks – best suited for DTV
    - DVB-S – for satellites (also DVB-S2)
    - DVB-C – for CATV

SMPTE Time Codes

- Society of Motion Picture and Television Engineers defines time codes for video
  - HH:MM:SS:FF
  - 01:12:59:16 represents number of pictures corresponding to 1hour, 12 minutes, 59 seconds, 16 frames
    - If we consider 30 fps, then 59 seconds represent 59*30 frames, 12 minutes represent 12*60*30 frames and 1 hour represents 1*60*60*30 frames.
- For NTSC, SMPTE uses a 30 *drop frame* code
  - increment as if using 30 fps, when really NTSC has only 29.97fps
  - defines rules to remove the difference error