Multimedia Hardware

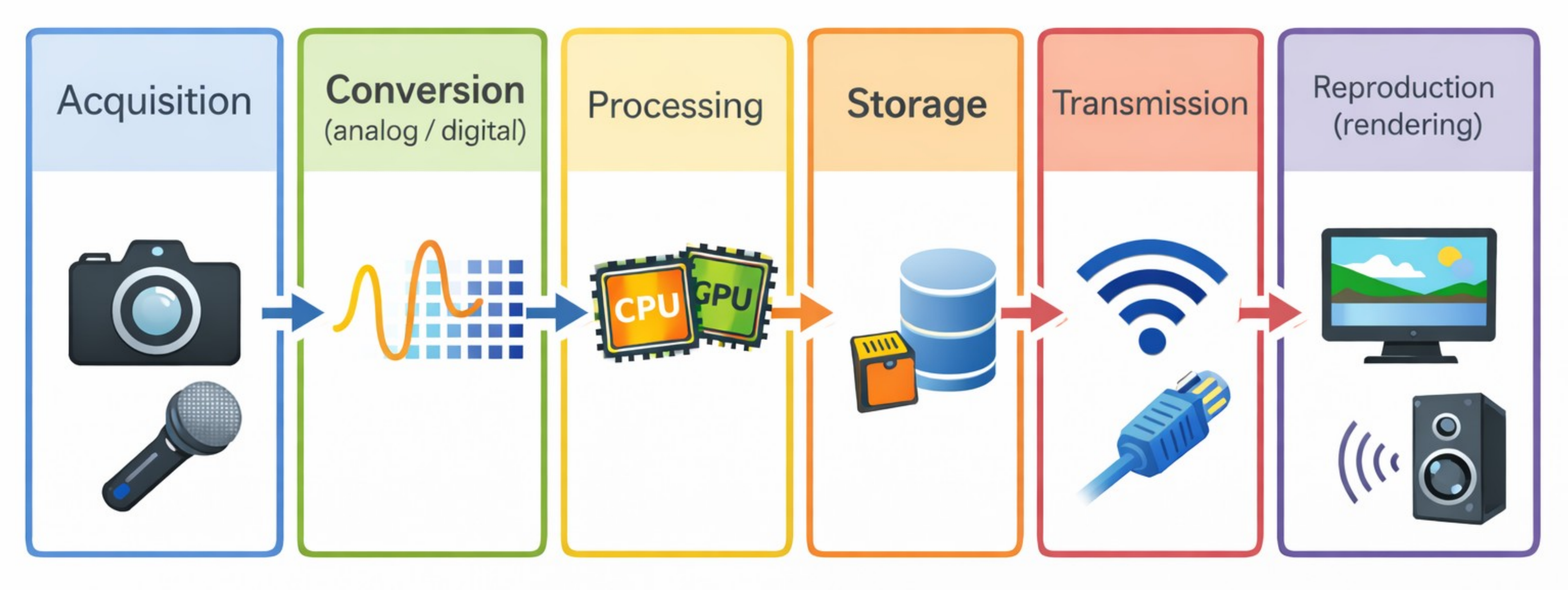
Why Hardware Matters?

- Multimedia systems process continuous signals (image, video, audio)
- Hardware determines
  - what quality is achievable
  - what is real-time
  - what formats and codecs are practical

Overview

- Multimedia pipeline
- Image/video acquisition and reproduction
  - Camera
  - Projector / Display
- Audio acquisition and presentation
  - Microphone
  - Speakers
- Processing hardware

The Multimedia Pipeline



Hardware Constraints

- Throughput (data rate)
- Latency (delay)
- Precision (bit depth, noise)
- Power & thermal limits
- Cost & integration

Signal Acquisition

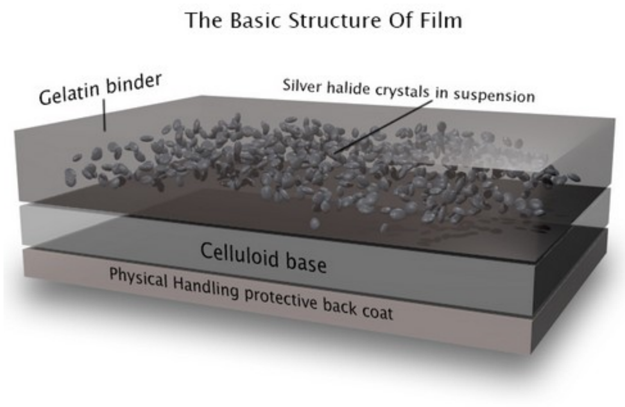
- Images: light → photons → electrons → voltage
- Audio: pressure waves → membrane → voltage
- Capture is:
  - noisy
  - limited in resolution
  - limited in dynamic range
- Implicit filtering determined by hardware

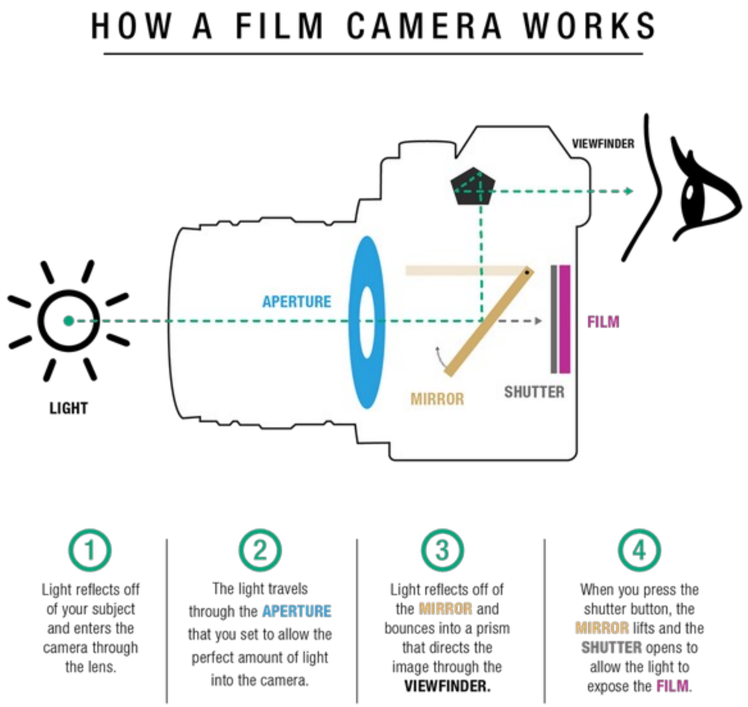
Signal Reproduction

- Inverse of signal acquisition: digital → physical
- Requires actuators rather than sensors
- Image/video: light emission or modulation
- Audio: air pressure change generation
- Hardware limitations
  - finite resolution
  - non-linear response
  - limited dynamic range
- Perceived quality defined by hardware

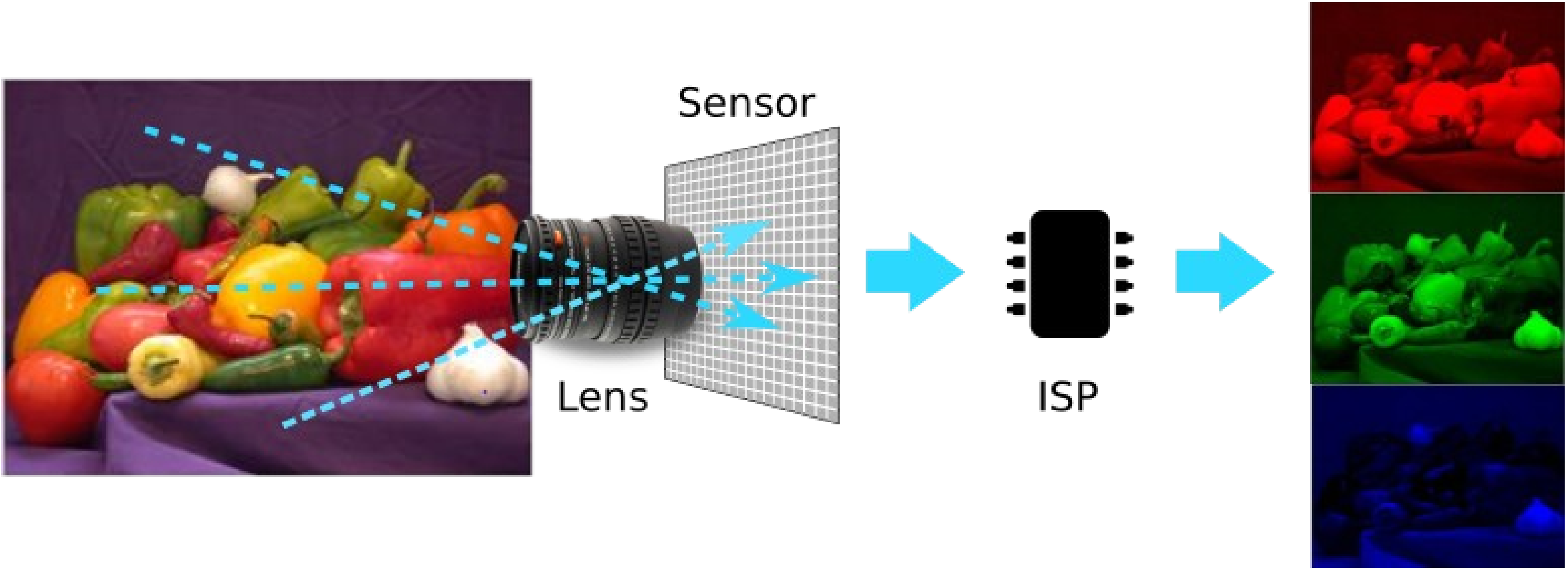
Analog image formation

- Light is focused onto photosensitive film
  - Film coated with silver halide crystals
  - Exposure causes chemical change





Digital camera



Lens

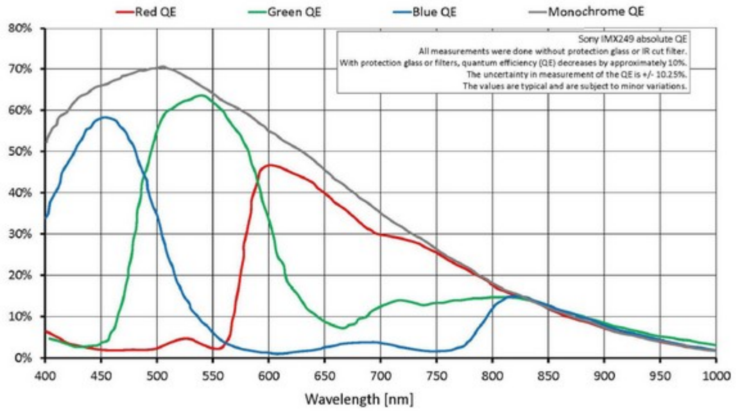
- Lens focuses light onto the sensor
  - focal length (field of view)
  - aperture (light intake)
- Trade-offs:
  - larger aperture - more light, less depth of field
  - compact lenses - more distortion
- Mechanical vs fixed-focus lenses
- Strong constraints in mobile devices

Optical distortion

- Chromatic aberration
- Vignetting
- Radial distortion

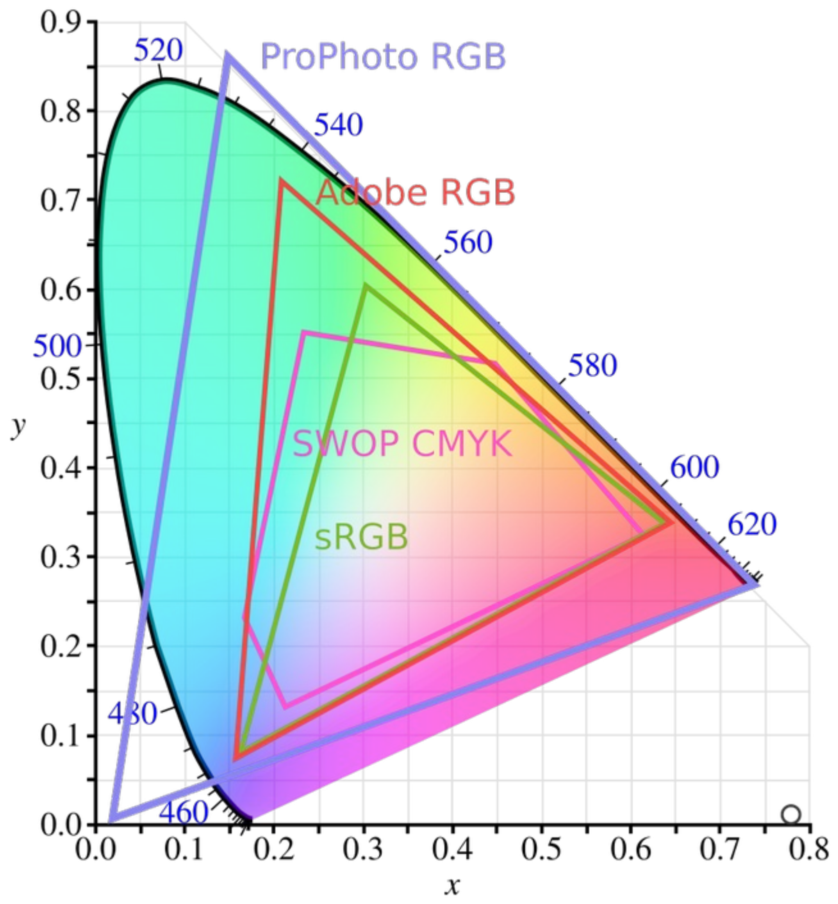
Sensor

- Converts incoming photons into electrical charge (photoelectric effect)
- Sensor response is spectral, not “true color”
- Color sensing uses overlapping spectral responses implemented as a hardware approximation
- Each pixel measures intensity, color is inferred
- Sensor characteristics
  - noise
  - dynamic range
  - low-light performance



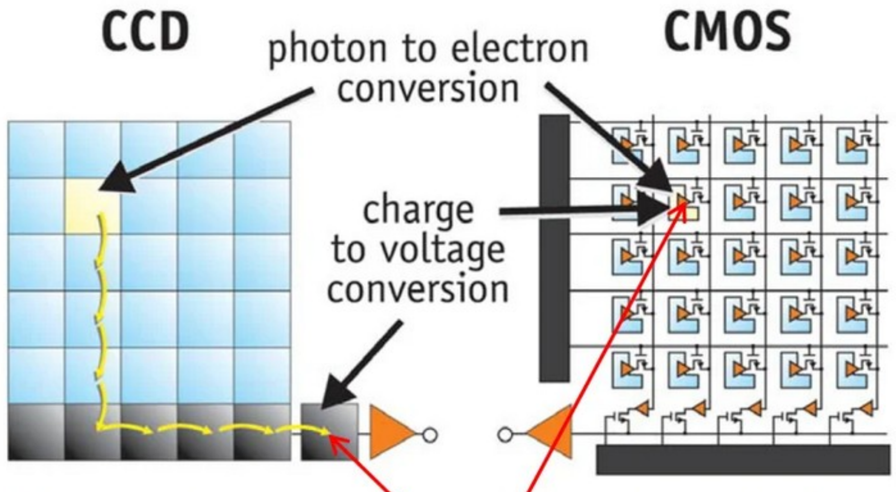
Digital Color

- Full spectrum reproduction is unnecessary
  - Reliance on human metamerism
  - Displays and cameras use color primaries
- Implemented using device-specific color primaries and calibration tables
  - Exact color reproduction requires calibration
  - Different color gamut



Sensor Technologies

- Charge Coupled Device (CCD)
  - high image quality (historically)
  - slower readout
  - higher power
- Complementary Metal Oxide Semiconductor (CMOS)
  - per-pixel or per-row readout
  - fast, low power
  - dominant in modern devices



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

Sensor Layout

- Color Filter Array (CFA)
  - One channel per pixel
  - Other values interpolated - demosaicing
- Foveon X3
  - Absorption depth
  - Vertically stacked photodiodes

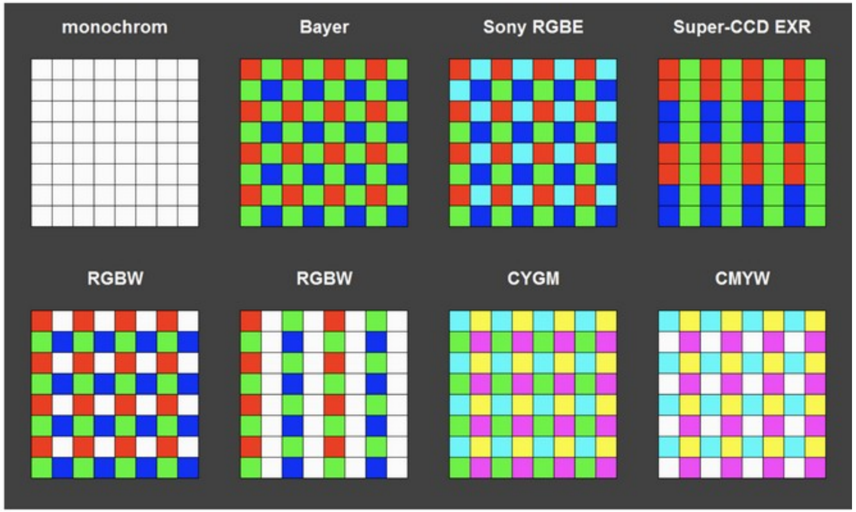
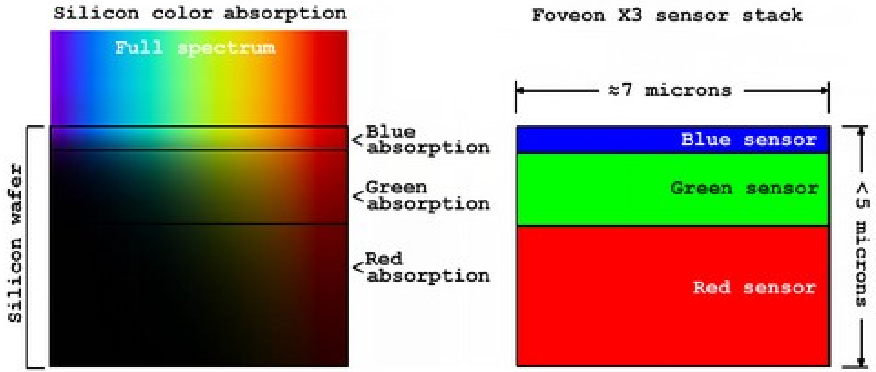


Image Courtesy: Frank Klemm



Shutter

- Rolling shutter
  - rows sampled at different times
  - motion distortions (“jello effect”)
- Global shutter
  - all pixels sampled simultaneously
  - better for fast motion and vision tasks
- Hardware trade-off: cost, complexity, noise

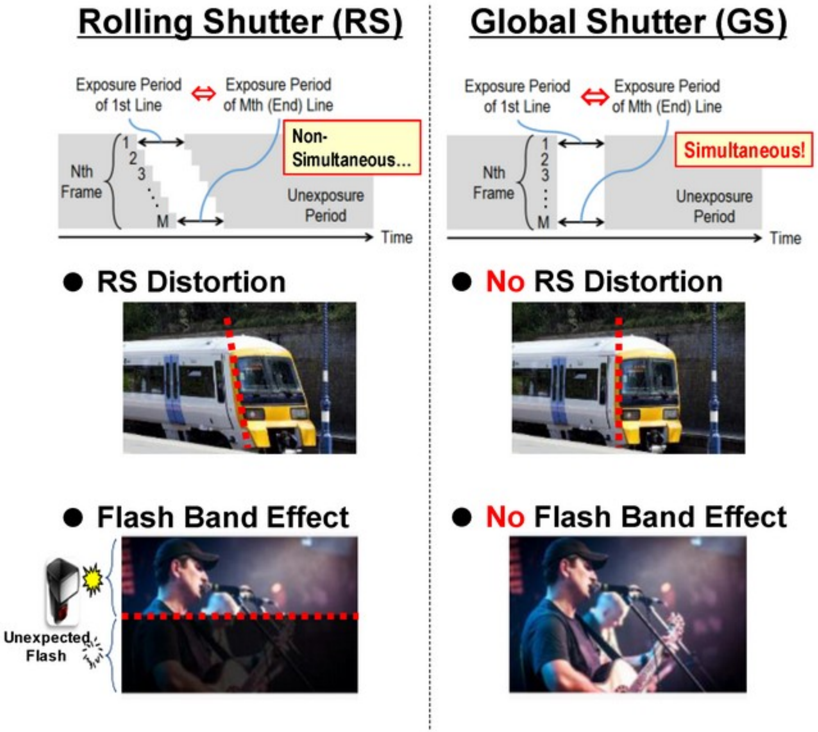
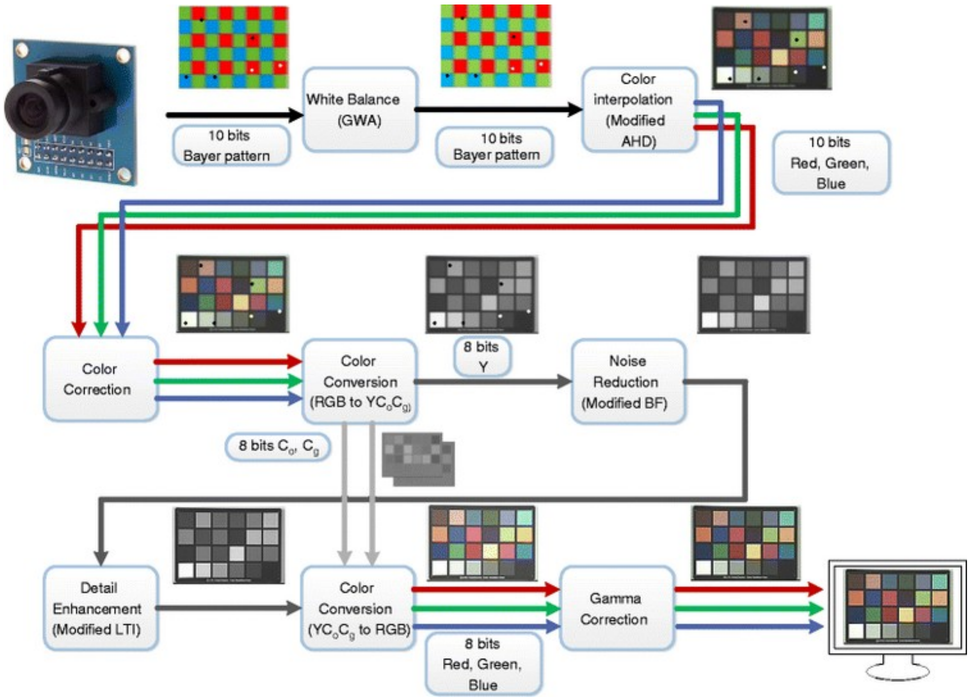


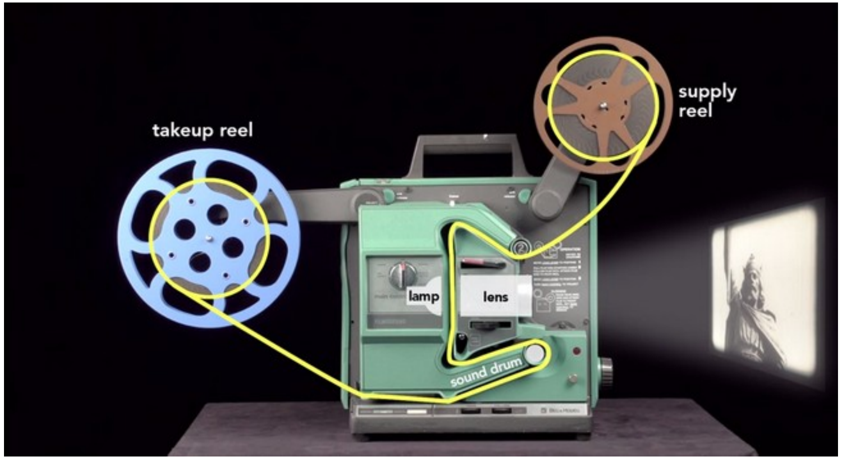
Image Signal Processor

- Dedicated fixed-function hardware, before CPU/GPU access
  - gain control
  - demosaicing
  - noise suppression
  - tone mapping
- Highly device-specific



Analog image reproduction

- Backlight is filtered by the film
- For motion, a film reel is physically moved
  - Film stops
  - Light passes through frame
  - Film moves to next frame
  - Light source is continuous

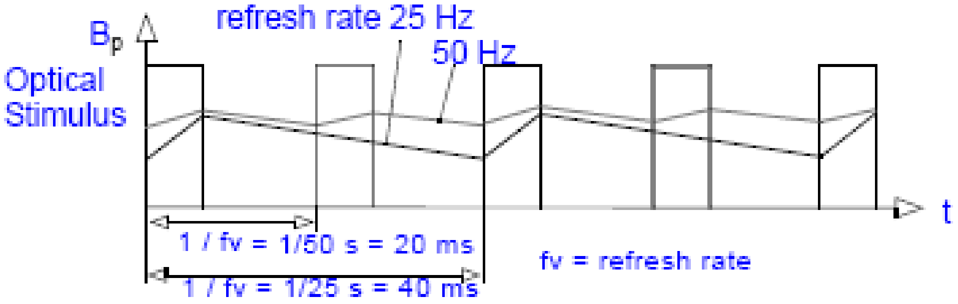


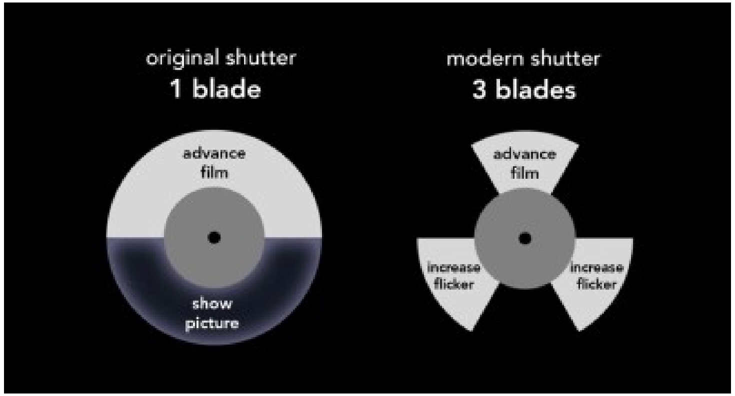
Temporal resolution

- Human perception system (eye+brain) can perceive about 10 - 12 images per second as separate images.
- Display refresh and shutter design must satisfy these limits

Flickering

- Shutter closed during film movement
- High illumination change - flickering
- Darker display - higher shutter frame rate



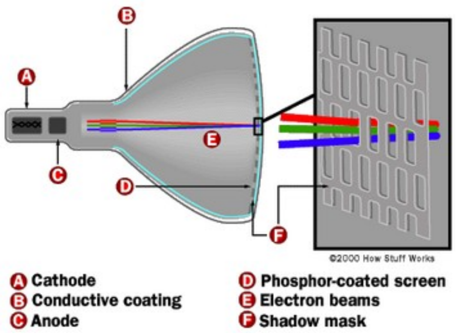


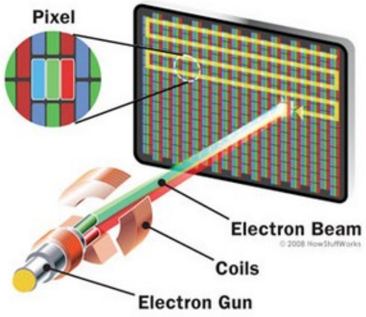
Analog color

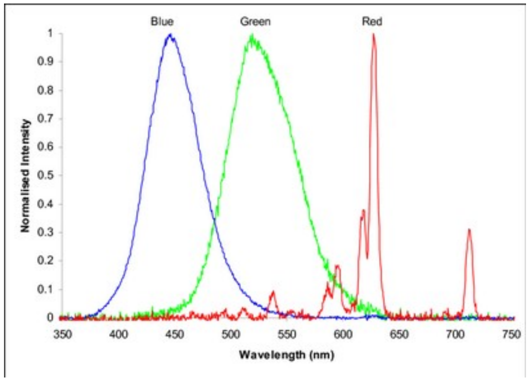
- Composite video
  - Suitable for lower bandwidth
  - Luminosity and color mixed in common signal
  - Channel cross-talk
  - Analog TV: NTSC, PAL, SECAM
- Component video
  - Separate signals for color channels
  - Better image reproduction, no cross-talk
  - BNC, RCA, VGA

Cathode displays

- Electron beam traversing matrix of fluorescent particles
- When a particle (pixel) is hit, it briefly glows
- To maintain realism, the beam has to refresh the screen fast enough (refresh frequency)
- CRT monitors (<50Hz flickering) (~100Hz no flickering)

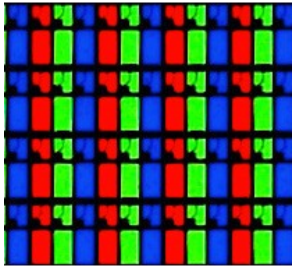


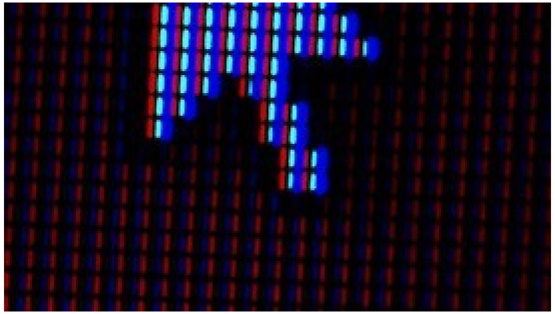


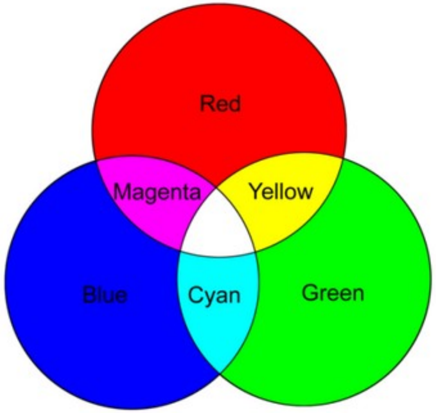


Digital color

- Additive color model
  - Starting with “black”
  - Additively mix color
- Devices
  - Monitors
  - Television
  - Projectors

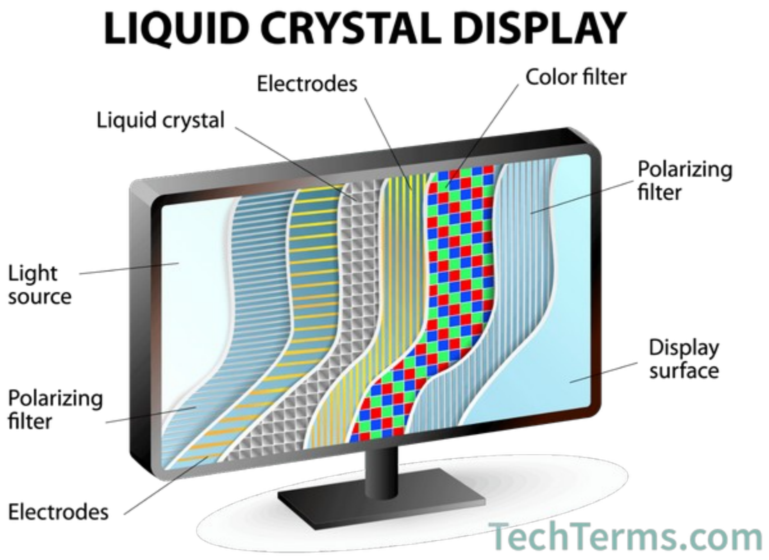


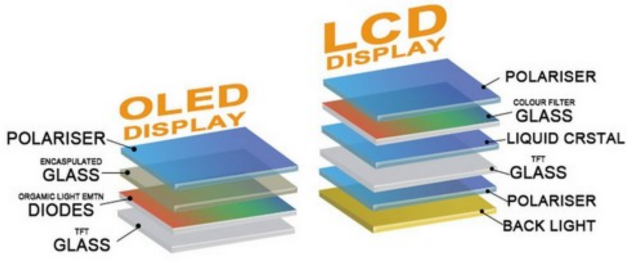




Flat panel displays

- Liquid crystal (LCD)
  - Fluorescent backlight
- Light-emitting diode (LED)
  - LED backlights
  - Still LCD
- Organic LED (OLED)
  - OLEDs for individual pixels
- *(Plasma)*





3D video

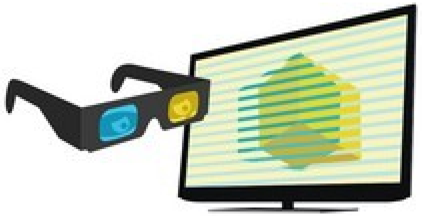
- Stereoscopic photography
  - Two images at two positions (50 mm to 75 mm apart)
  - Impression of a third dimension
- Video technology
  - Wearable technology
  - Autostereoscopy

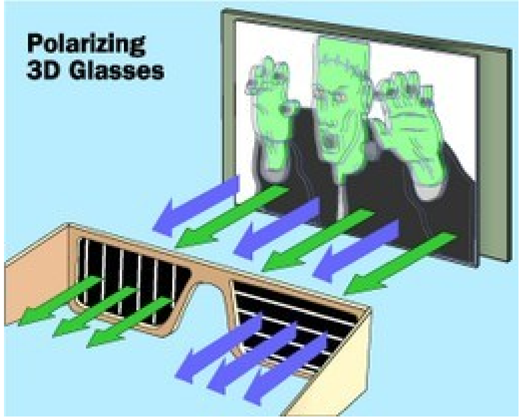
3D Video Acquisition

- Requires capturing multiple viewpoints
  - dual-camera rigs
  - multi-camera arrays
- Key constraints
  - camera synchronization
  - geometric calibration

Wearable 3D video technology

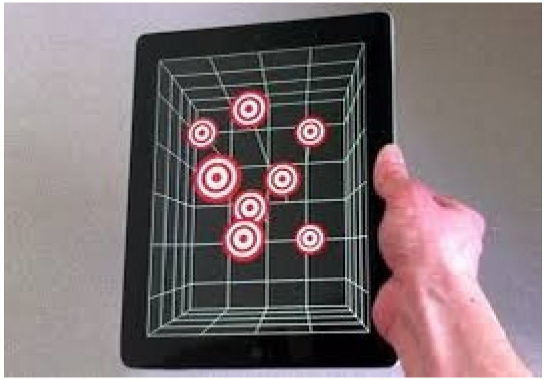
- Anaglyph glasses (passive)
- Polarized glasses (passive)
- Binocular HMD (active)
- Active shutter glasses (active)

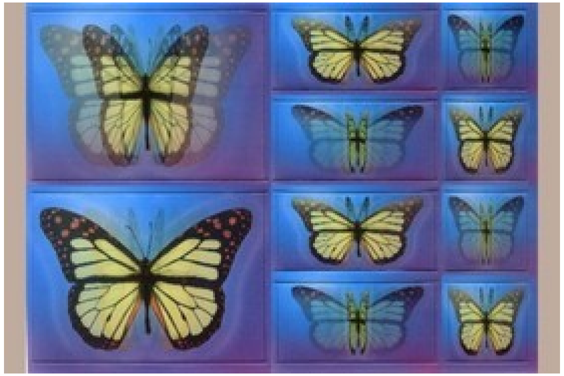


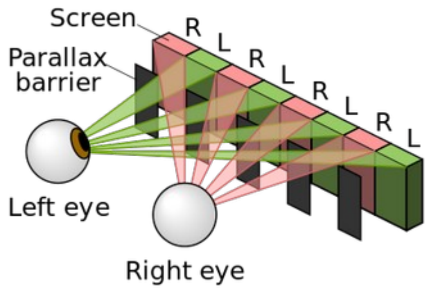


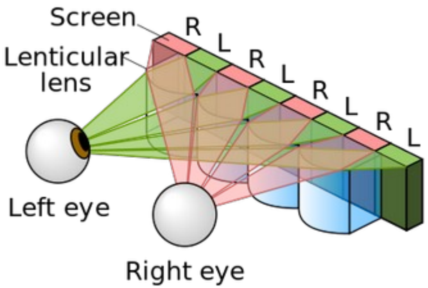
Autostereoscopy (Glassesless 3D)

- Eye/head tracking (active)
- Parallax barrier (passive)
- Lenticular lens (passive)









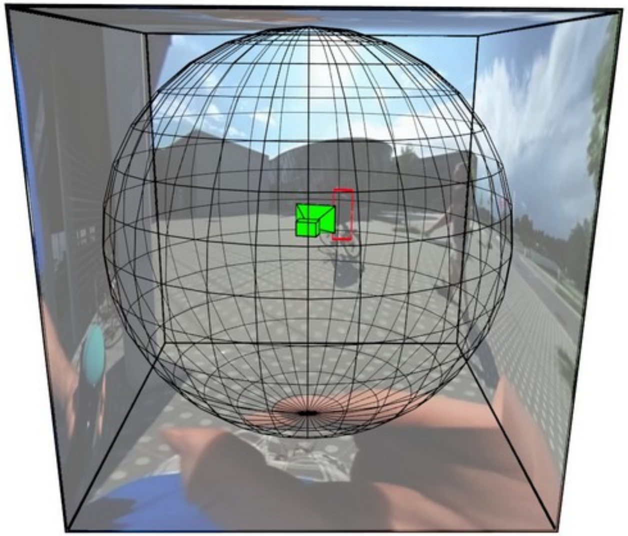
Author: Cmglee

Problems with 3D video

- Technology
  - Resolution
  - Frame rate
  - Cross-talk
- Usefulness
  - Limited value
- Cost
- Health
  - Motion sickness
  - Headaches
  - Nausea
  - Disorientation

Omnidirectional video

- Single camera origin
- Projection sphere
- Interactive experience
  - Panoramas (aesthetic)
  - Live entertainment - BBC
  - Sports and tourism
  - Consumer-created (travel)

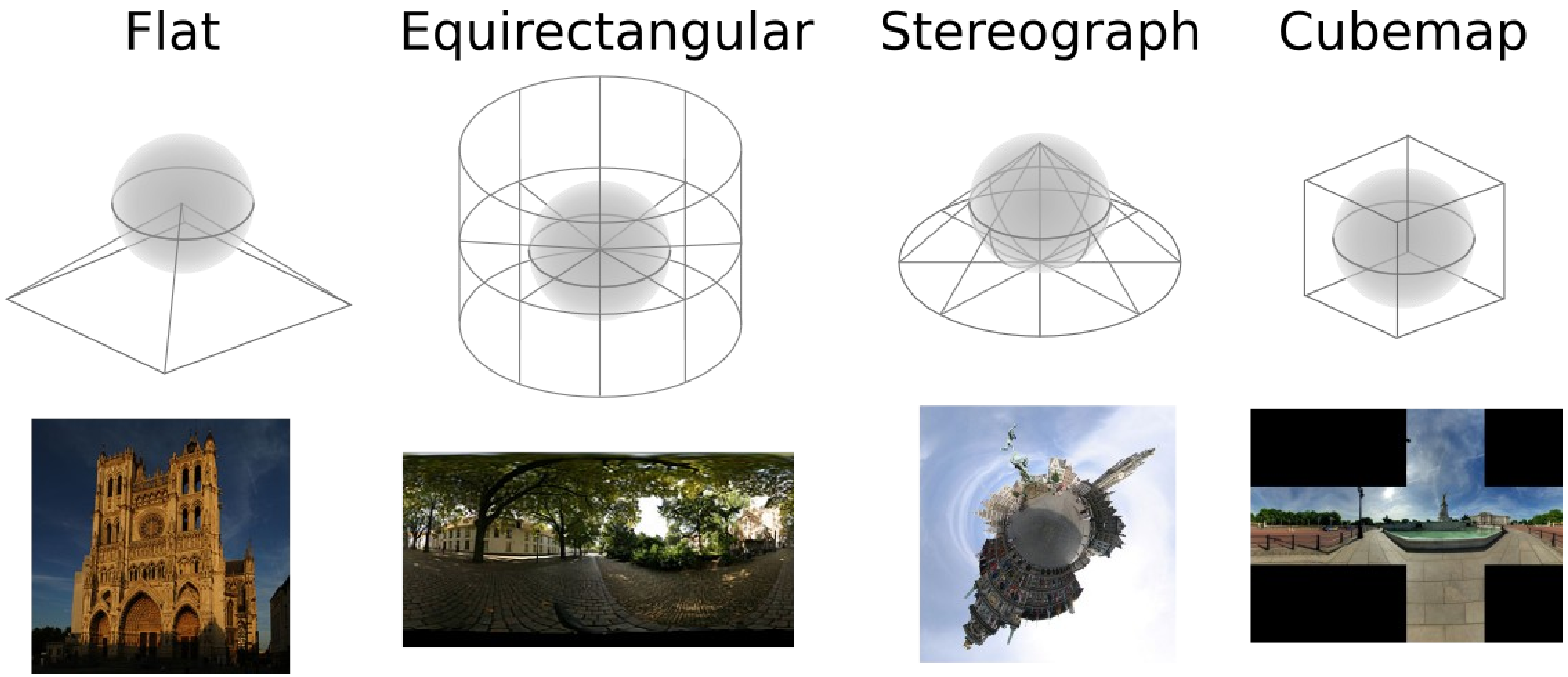


Omnidirectional video acquisition

- Multi-camera
  - Wide-lens (less cameras, low resolution)
  - Narrow-lens (more cameras, high resolution)
- Calibration (factory, post-processing)

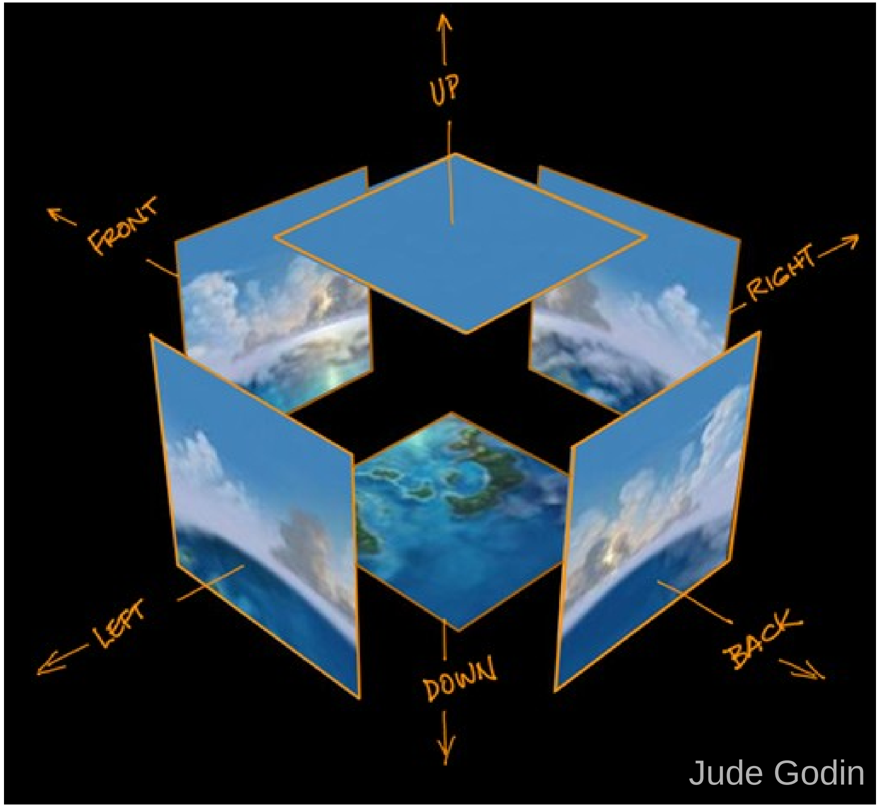



Re-projection approaches



Novel view generation

- Rendering
  - UV mapping
  - Cubemap
- Limitation
  - View angle
  - Resolution

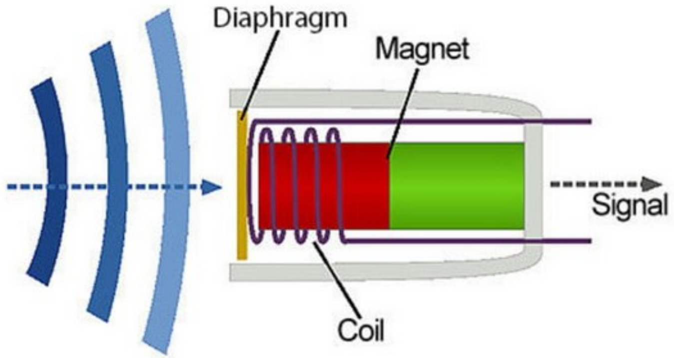


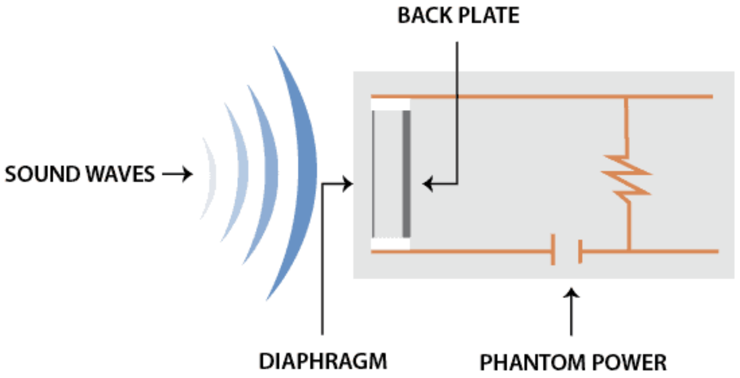
Audio Acquisition

- Microphone converts pressure waves to electrical signals
  - microphone
  - pre-amplifier
  - anti-alias filter (low-pass)
  - Analog-to-Digital (ADC)
- Hardware quality strongly affects signal-to-noise ratio
  - Noise introduced before digitization is irreversible
  - ADC clock stability directly impacts timing and noise

Microphone Types

- Dynamic microphone
  - Robust, passive
  - Induction based
  - Low sensitivity
- Condenser microphone
  - Capacitance based
  - High sensitivity
  - Require power (phantom / bias)
- MEMS microphones
  - Very small condenser microphone
  - consistent manufacturing
  - Dominant in phones, wearables



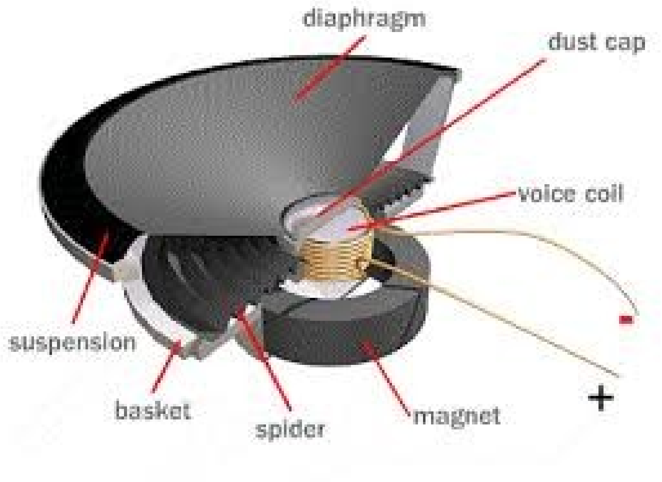


Audio Reproduction

- Digital samples converted to analog signals (Digital-to-Analog - DAC)
- Amplifier drives speaker or headphones
- Hardware constraints
  - size limits low-frequency reproduction
  - enclosure affects sound quality

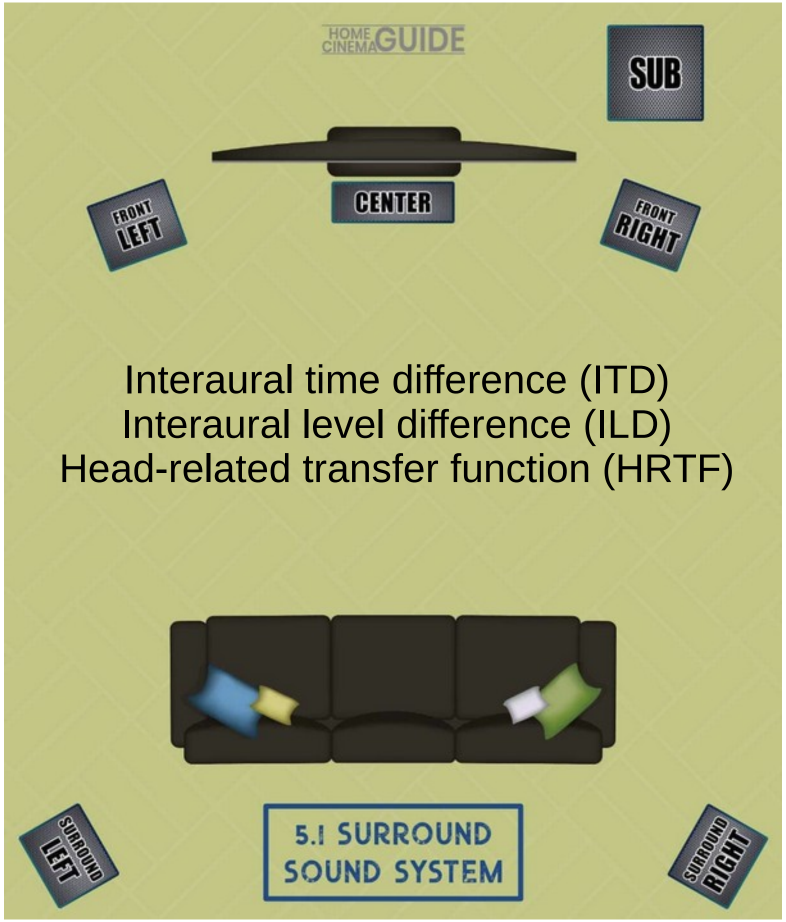
Speakers and Headphones

- Speakers
  - enclosure size limits bass
  - efficiency vs size trade-off
- Headphones
  - easier low-frequency reproduction
  - less power needed
- Amplifier limitations
  - clipping
  - distortion
  - power consumption



Spatial Audio

- Stereo
  - left / right channel separation
  - Limited spatial cues
- Beyond stereo
  - Surround sound (5.1, 7.1)
  - Object-based audio (Dolby Atmos)
- Hardware challenges
  - Speaker placement
  - Synchronization
  - Calibration

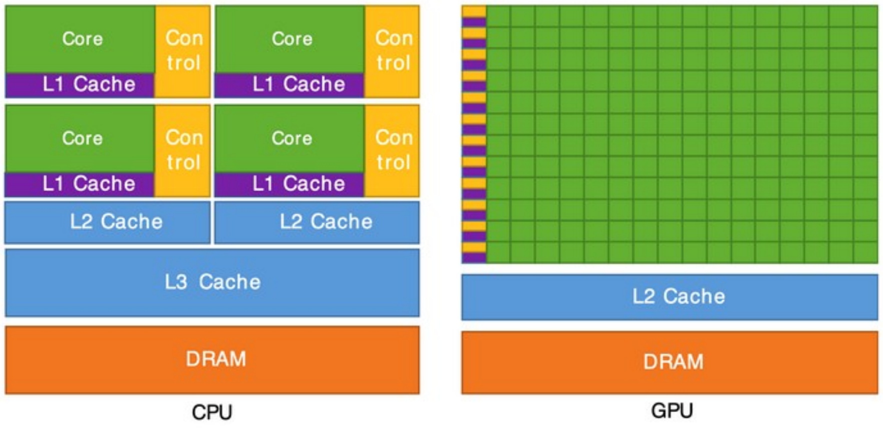


Processing for Multimedia

- Multimedia data
  - high-volume (images, video, multi-channel audio)
  - continuous (real-time streams)
  - time-dependent (deadlines matter)
- Real-time constraints (dropped frame / audio glitch)
- Processing quality is bounded
  - available compute
  - memory bandwidth
  - power budget

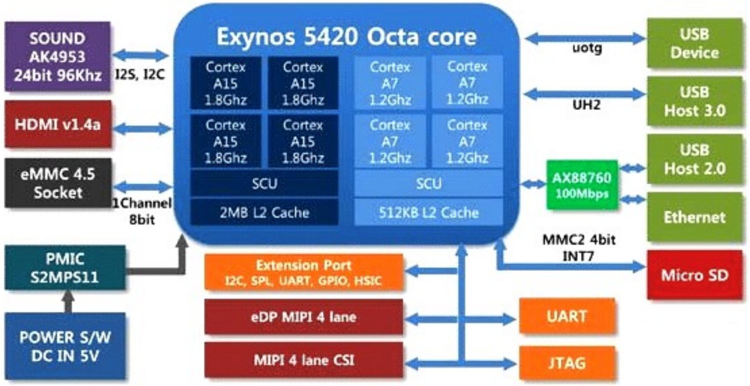
Parallel processing

- Multimedia workloads are naturally parallel
  - per-pixel image operations
  - block-based video decoding
  - audio filtering
- Hardware parallelism
  - SIMD instructions
  - GPU cores
  - hardware accelerators
- Key constraint is memory bandwidth



Heterogeneous systems

- Modern SoCs combine:
  - CPU + GPU + DSP + accelerators
- Software schedules work across units
- Performance comes from coordination, not one processor
- Asynchronous execution → buffering → latency

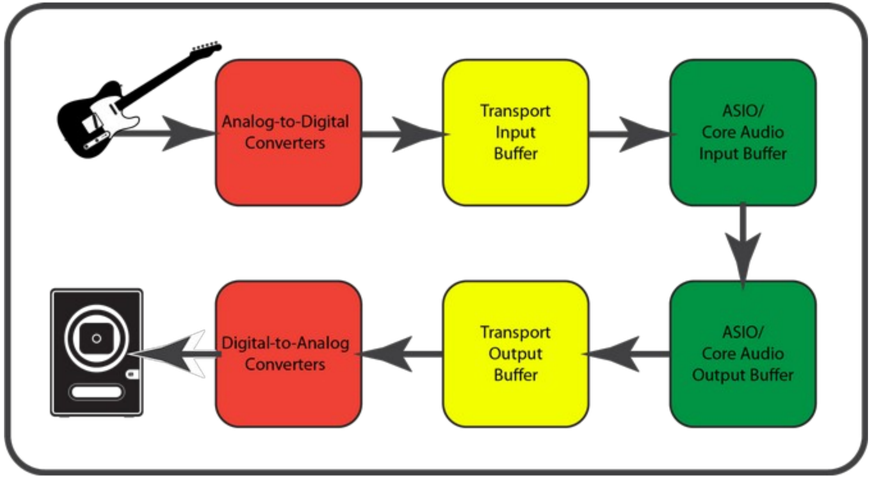


Throughput

- Amount of data processed per unit time
- How much can the system handle continuously?
  - Can we decode 4K video at 60 fps?
  - Can we stream 1.4 Mbps audio without dropouts?
- Limited by weakest link
  - Processing capacity
  - Memory/IO bandwidth

Latency

- Time delay between input and output
- How long does it take before output appears?
  - Delay between speaking and hearing your voice
  - Delay between camera capture and display
- Latency is additive
  - Buffering
  - Processing, movement
  - Synchronization, scheduling



Throughput and Latency

- Throughput: how much data per second
- Latency: how long until output appears
- High throughput ≠ low latency
- Multimedia systems must balance both
  - larger buffer - higher latency
  - smaller buffer - dropouts

Codecs and Standards

- Hardware often supports specific codecs only
  - Fast implementation of bottleneck operations
  - Energy efficient
  - Limited flexibility (dedicated units)
- The role of standards
  - Interoperability
  - Long hardware lifespan
- Hardware support impacts codec adoption

Summary

- Multimedia systems are constrained by hardware
- Acquisition and reproduction define quality bounds
- Processing relies on heterogeneous hardware
- Fixed-function units dominate efficiency
- Understanding hardware explains system-level design choices