Research Institute for Future Media Computing
未来媒体技术与研究所

Institute of Computer Vision
计算机视觉研究所

多媒体系统导论
Fundamentals of Multimedia System

授课教师：文嘉俊
邮箱：wenjiajun@szu.edu.cn

2024年春季课程

Outline of Lecture 03

- Color Science
  - Light and Spectra
  - Human Vision
  - Image Formation
  - Color-Matching Functions
  - CIE Chromaticity Diagram
  - Out-of-Gamut Colors
  - Color Coordinate Schemes
- Color Models in Image
- Color Models in Video
- Experiments

Color Science

◆ Human Vision

- These spectral sensitivity functions (光谱灵敏度函数，或视锥函数) are usually denoted by a vector function $\boldsymbol{q}(\lambda)$, with components

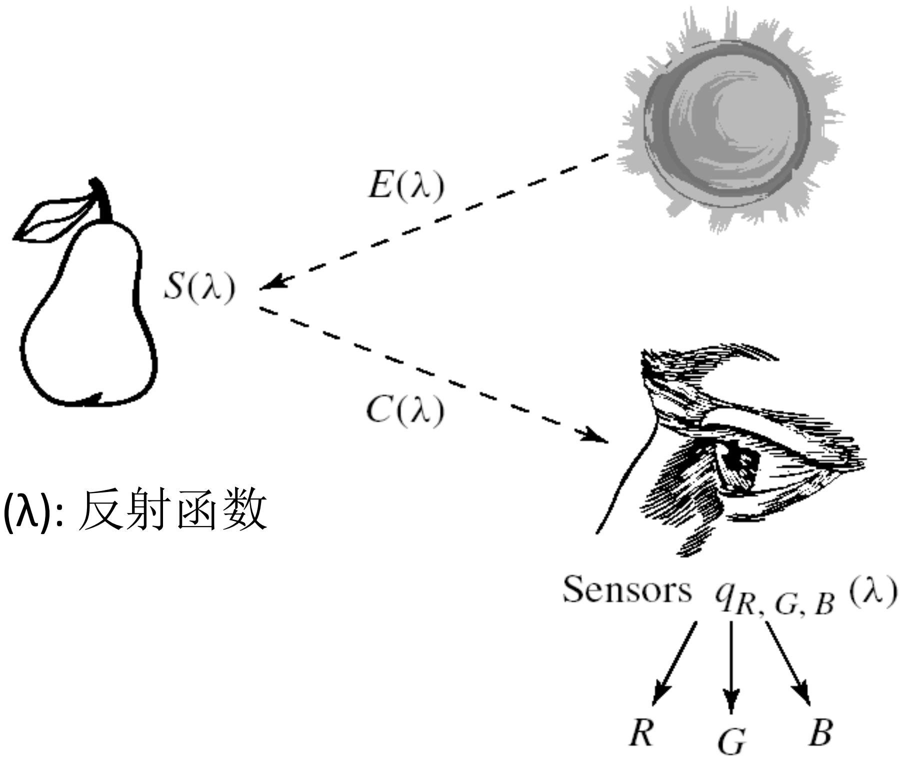
$$\boldsymbol{q}(\lambda) = \left(q_R(\lambda), q_G(\lambda), q_B(\lambda)\right)^T.$$

- The response in each color channel in the eye is proportional to the number of neurons firing（神经元激发）.
- We can succinctly write down this idea in the form of an integral:

$$\begin{aligned} R &= \int E(\lambda)\, q_R(\lambda)\, d\lambda \\ G &= \int E(\lambda)\, q_G(\lambda)\, d\lambda \\ B &= \int E(\lambda)\, q_B(\lambda)\, d\lambda \end{aligned}$$

Color Science

- **Image Formation**
  - The equations that take into account the image formation model are:
  - *color signal* $C(\lambda) = E(\lambda)S(\lambda)$



$$
\begin{aligned}
R &= \int E(\lambda)\, S(\lambda)\, q_R(\lambda)\, d\lambda \\
G &= \int E(\lambda)\, S(\lambda)\, q_G(\lambda)\, d\lambda \\
B &= \int E(\lambda)\, S(\lambda)\, q_B(\lambda)\, d\lambda .
\end{aligned}
$$

S(λ): 反射函数

Color Science

◆ CIE Chromaticity Diagram

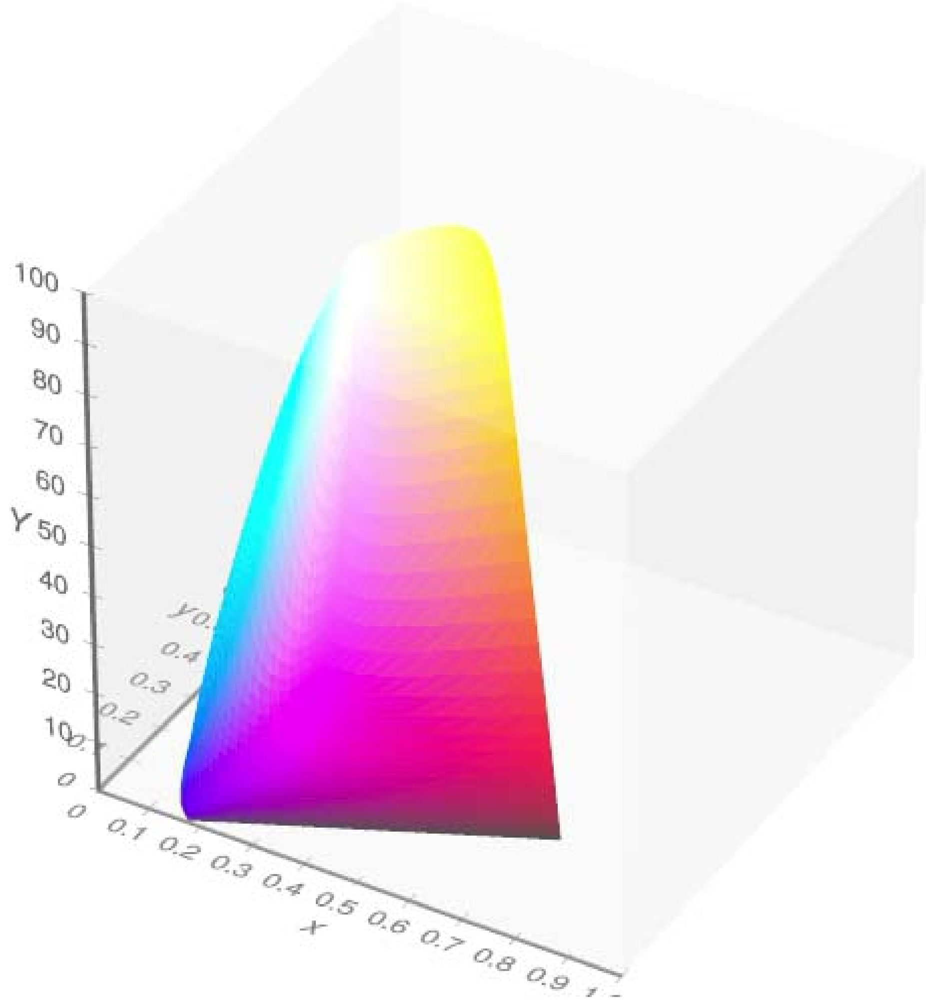
- A color is the set of **tristimulus values(三色值)** X, Y, Z defined

$$
\begin{aligned}
X &= \int E(\lambda)\, \bar{x}(\lambda)\, d\lambda \\
Y &= \int E(\lambda)\, \bar{y}(\lambda)\, d\lambda \\
Z &= \int E(\lambda)\, \bar{z}(\lambda)\, d\lambda
\end{aligned}
$$

- Y is called the ***luminance***
- All color information and transforms are tied to these special values

$$
\begin{aligned}
x &= X/(X+Y+Z) \\
y &= Y/(X+Y+Z) \\
z &= Z/(X+Y+Z)
\end{aligned}
$$

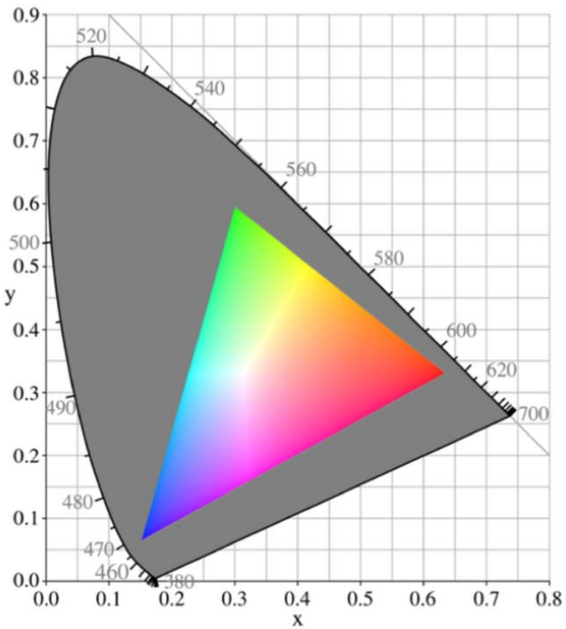
Normalization: insensitive to light intensity

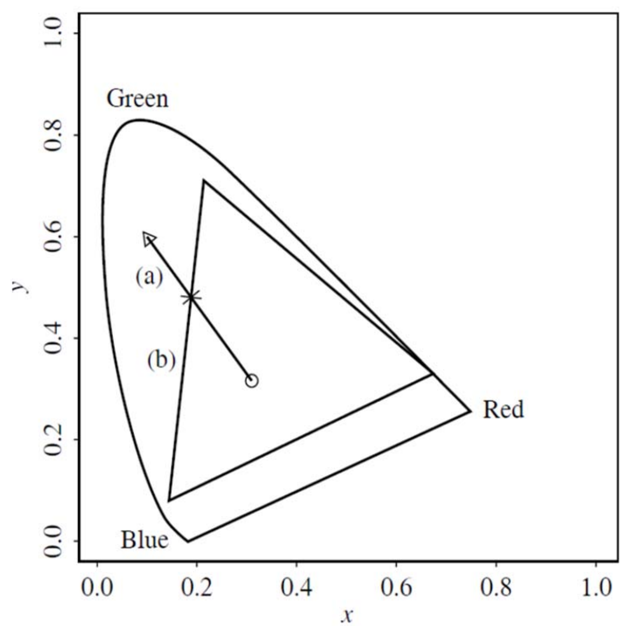


Color Science

◆ Out-of-Gamut Colors

- **Gamut**(色彩空间，色域)：refers to the subset of colors which can be accurately represented in a given circumstance, such as within a given **color space** or by a certain **output device**.
- The out-of-gamut (超色域) color show by a triangle is **approximated** by the intersection of (a) the line from that color to the white point with (b) the boundary of the device color gamut.

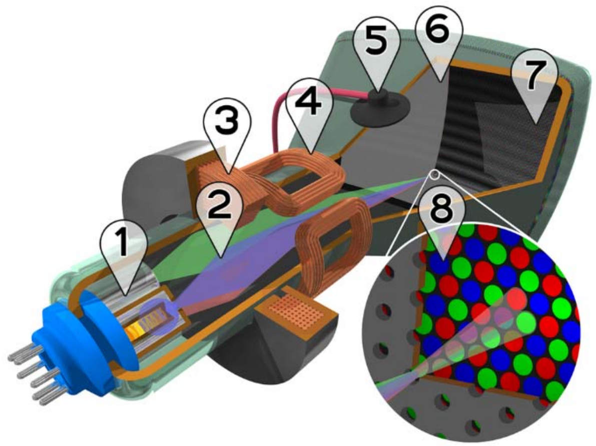




Color Models in Image

- **Gamma Correction**
  - The RGB numbers in an image file are converted back to analog, and drive the electron guns in the cathode ray tube (CRT, 阴极射线管).
  - The light emitted is actually roughly proportional to the voltage raised to a power; this power is called "gamma," γ.
  - In order to make images look right, make a gamma correction.

$$R \rightarrow R' = R^{1/\gamma} \Rightarrow (R')^{\gamma} \rightarrow R,$$



1. Three electron emitters
2. Electron beams
3. Focusing coils
4. Deflection coils
5. Connection for final anodes
6. Mask for separating beams
7. Phosphor layer
8. Close-up of the phosphor-coated inner side of the screen

Color Models in Video

◆ Color Model

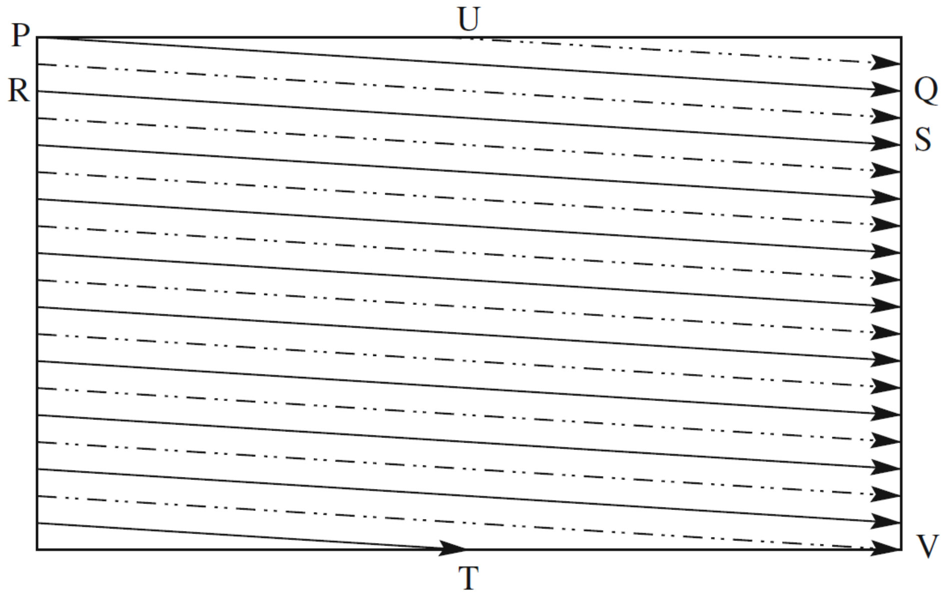
模型	应用领域	Comments
XYZ	理论实验	CIE, 三色值
Lab	纺织、制造	CIE
RGB	电子设备	CIE、三基色、加模型
sRGB	网页	平衡显示和色彩感知
HSV/HIS/HSL	艺术	色调、饱和度、亮度
CMYK	印刷工业	减模型
YUV	彩色视频	欧洲、PAL
YIQ	彩色视频	北美、NTSC
YCbCr	数字视频	视觉特征，低存储

Outline of Lecture 04

- Analog Video
- Digital Video
  - Chroma Subsampling
  - CCIR and ITU-R Standards for Digital Video
  - High-Definition TV
  - Ultra High Definition TV
- Video Display Interfaces
  - Analog Display Interfaces
  - Digital Display Interfaces
- 3D Video and TV

Analog Video

- Traditional TV programs were sent and received as an analog signal.
- The brightness of video signal is a monotonic function of voltage
- An **analog signal** $f(t)$ samples a time-varying **image**.
  - **Progressive scanning**: traces a complete image (frame) line by line at regular intervals.
  - **Interlaced scanning**: the odd-numbered lines are traced first, then the even-numbered lines.



Analog Video

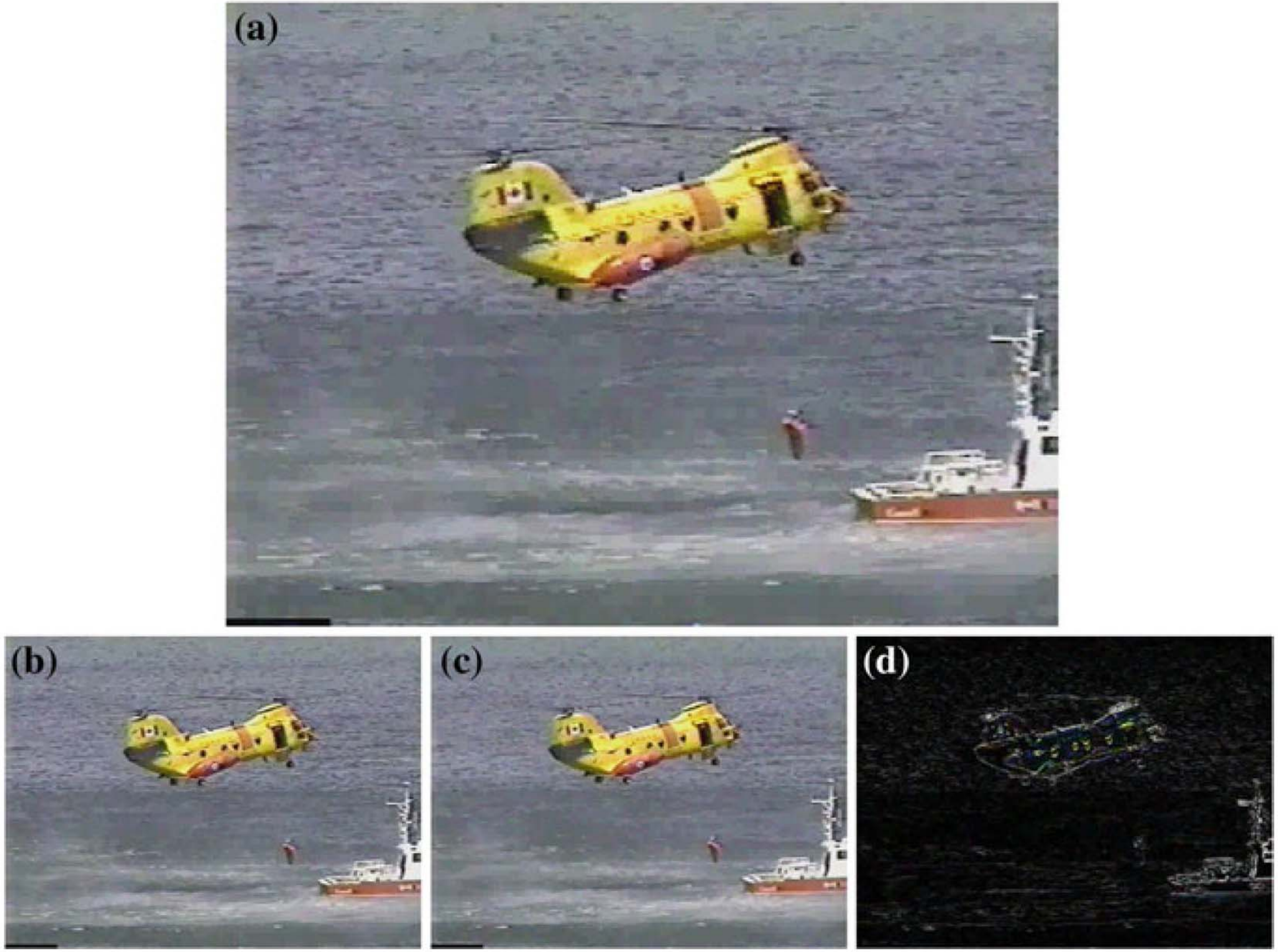
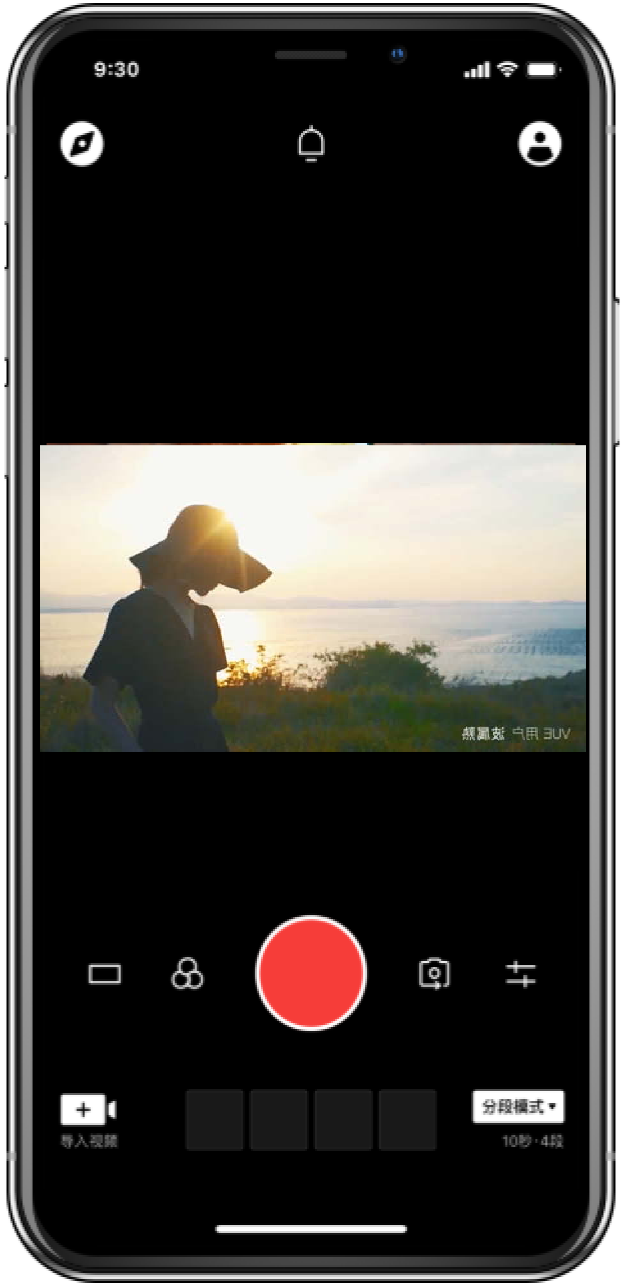


Fig. 5.2 Interlaced scan produces two fields for each frame: **a** the video frame; **b** Field 1; **c** Field 2; **d** difference of fields

Digital Video

- The advantages of digital representation for video:
  - Video can be stored on digital devices or in memory, **ready to be processed** (noise removal, cut and paste, etc.), and integrated to various multimedia applications;
  - **Direct access** is possible, which makes nonlinear video editing achievable as a simple, rather than a complex task;
  - **Repeated recording** does not degrade image quality;
  - Ease of **encryption** and better tolerance to channel noise.



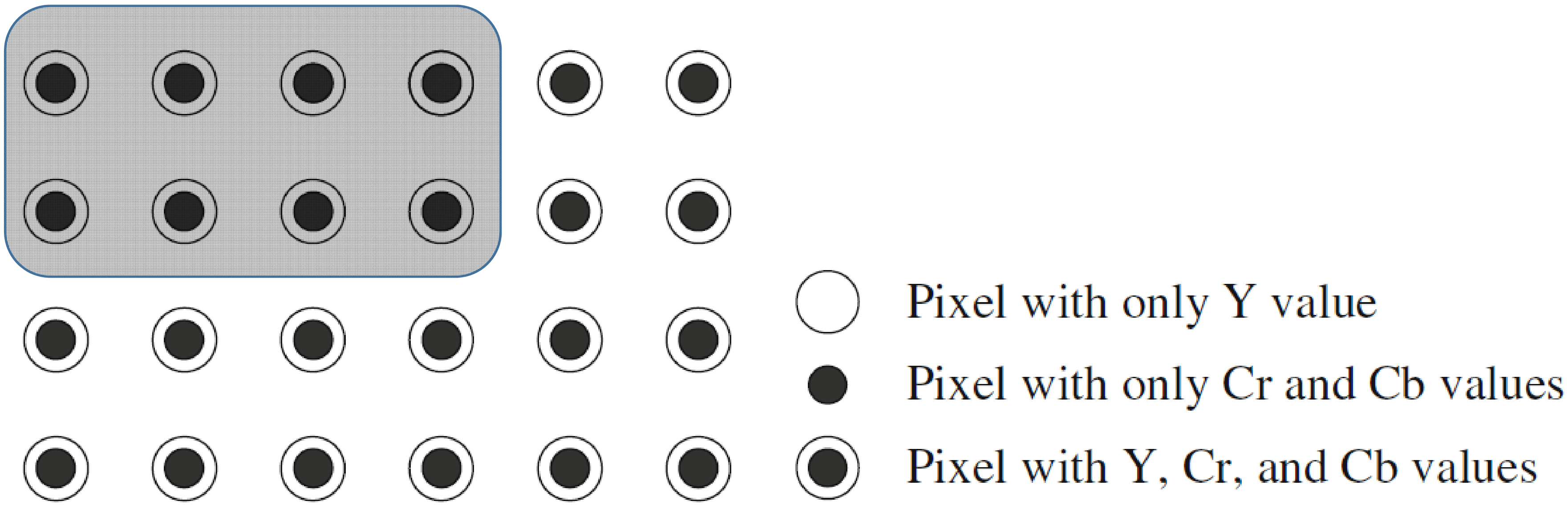
http://vue.video/

Digital Video

- Chroma Subsampling
  - Since humans see color with much less spatial resolution than they see black and white, it makes sense to “**decimate**(削弱)” the **chrominance signal**.
  - RGB signals are first converted into color component space. The usual color space is **YCbCr**.
  - Interesting (but not necessarily informative!) names have arisen to label the different schemes used.
  - To begin with, numbers are given stating how many pixel values, per four original pixels, are actually sent: 4:4:4, 4:2:2, 4:1:1, 4:2:0.

Digital Video

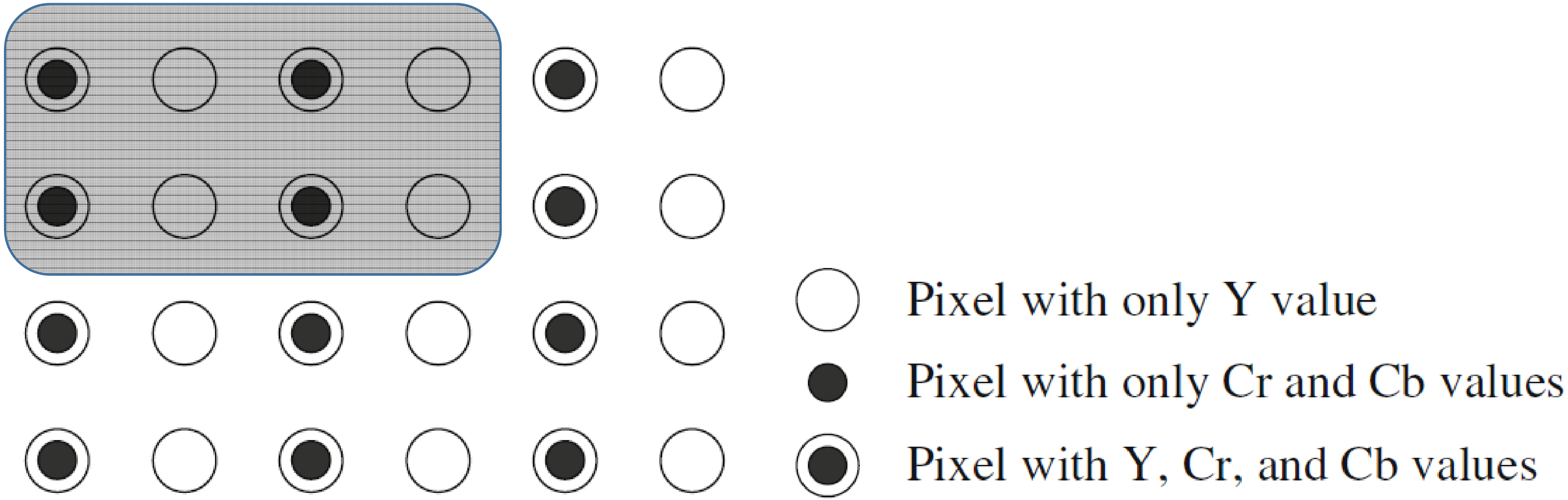
- Chroma Subsampling – 4:4:4
  - The chroma subsampling scheme “4:4:4” indicates that no chroma subsampling is used.
  - Each pixel's *Y*, *Cb* and *Cr* values are transmitted, 4 for each of *Y*, *Cb*, *Cr*.
  - 3 Bytes for each pixel.



4:4:4

Digital Video

- Chroma Subsampling – 4:2:2
  - The scheme "4:2:2" indicates *horizontal subsampling* of the *Cb*, *Cr* signals by a factor of 2.
  - That is, of four pixels horizontally labelled as 0 to 3, all four *Y*s are sent, and every two *Cb*'s and two *Cr*'s are sent, as (*Cb*0, *Y*0)(*Cr*0, *Y*1)(*Cb*2, *Y*2)(*Cr*2, *Y*3).

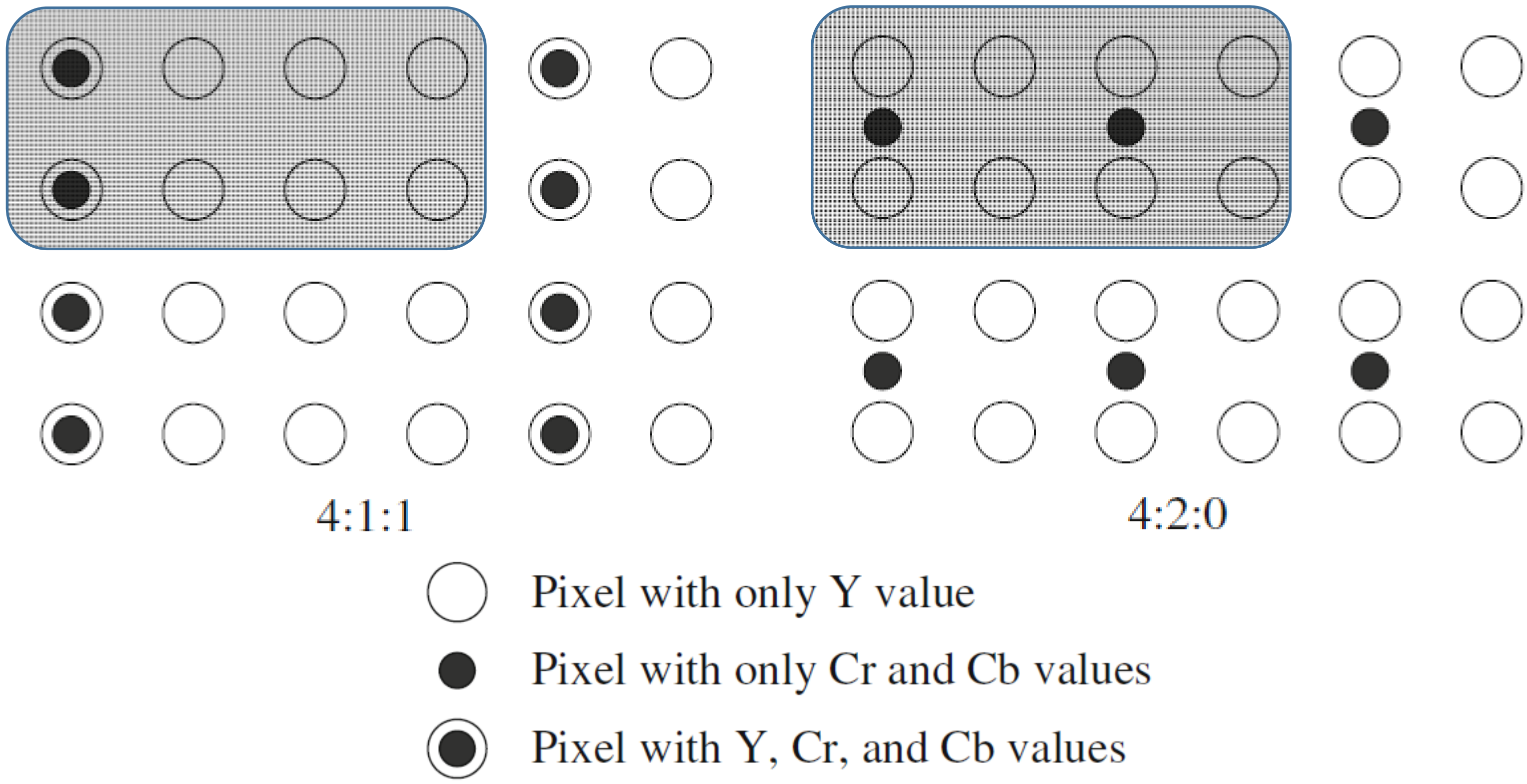


4:2:2

Digital Video

◆ Chroma Subsampling

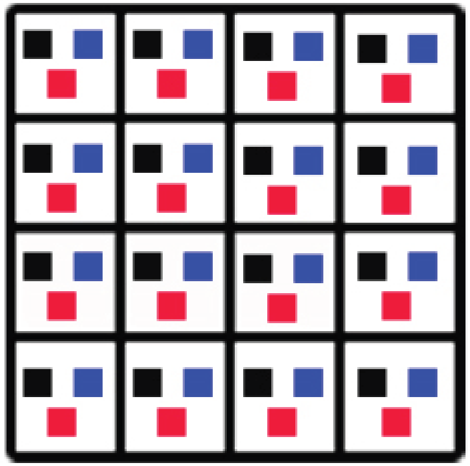
- The scheme "4:1:1" indicates *horizontal subsampling* of signals by a factor of 4.
- The scheme "4:2:0" indicates *vertical and horizontal subsampling* of signals by a factor of 2.



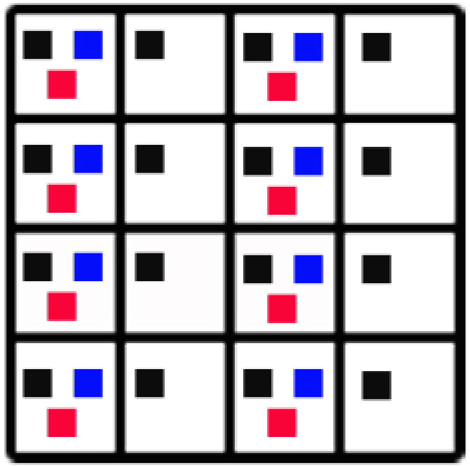
Digital Video

- Chroma Subsampling

3 Bytes



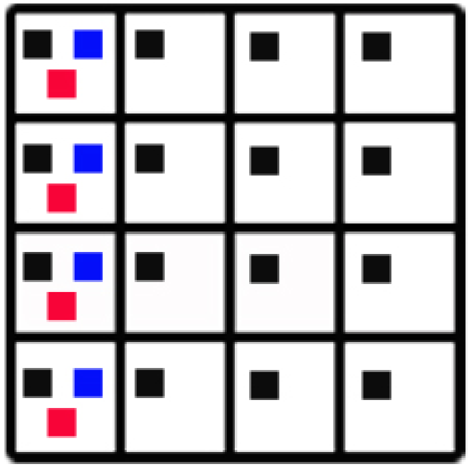
4:4:4



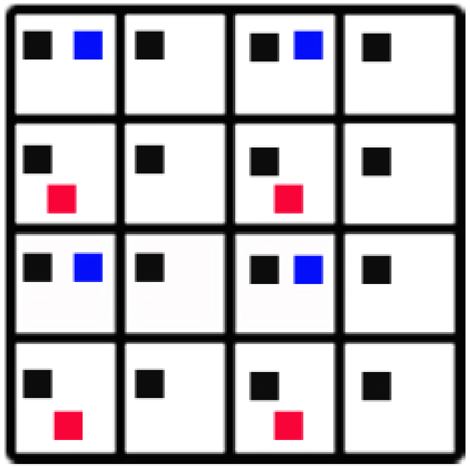
4:2:2

2 Bytes

1.5 Bytes



4:1:1



4:2:0

1.5 Bytes

Digital Video

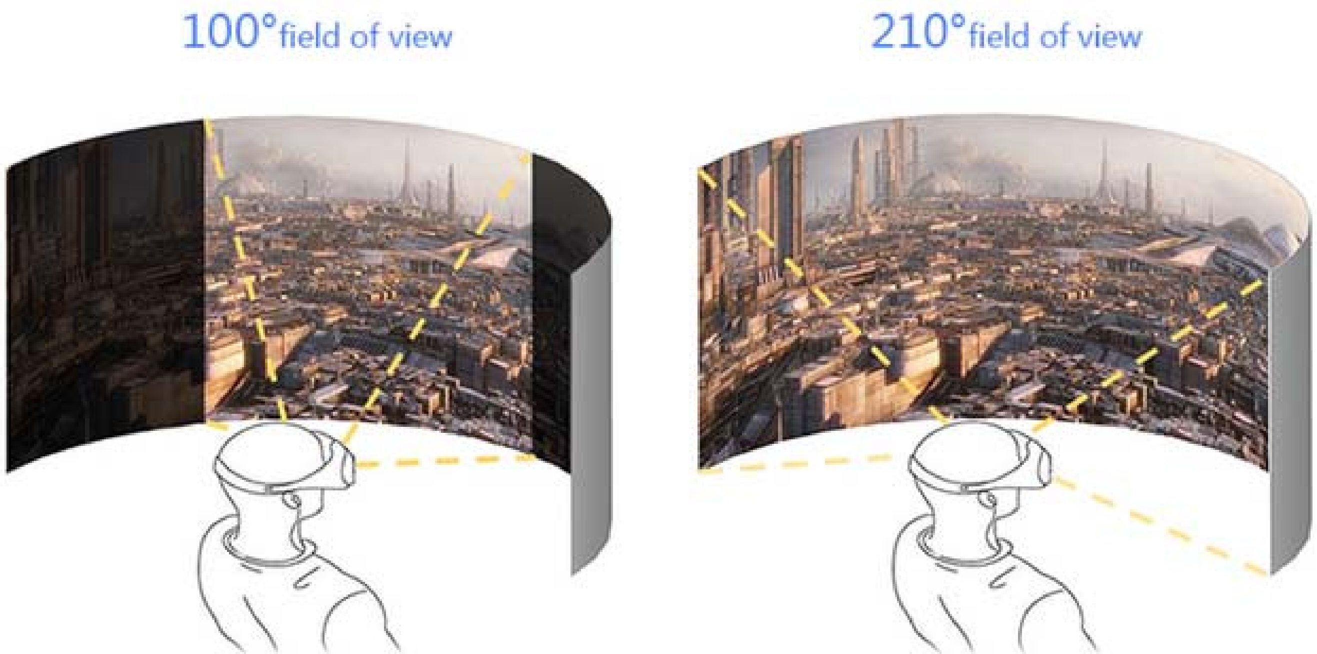
- Consultative Committee for International Radio： CCIR-601
- International Telecommunication Union： ITU-R Rec. 601

Table 5.3 ITU-R digital video specifications

	Rec. 601 525/60 NTSC	Rec. 601 625/50 PAL/SECAM	CIF	QCIF
Luminance resolution	720×480	720×576	352×288	176×144
Chrominance resolution	360×480	360×576	176×144	88×72
Color subsampling	4:2:2	4:2:2	4:2:0	4:2:0
Aspect ratio	4:3	4:3	4:3	4:3
Fields/sec	60	50	30	30
Interlaced	Yes	Yes	No	No

Digital Video

- High-Definition TV
  - Increase the **visual field**, especially its width.
  - **Progressive scan**, avoiding serrated（锯齿） edges to moving objects



Digital Video

- High-Definition TV
  - Increase the **visual field**, especially its width.
  - **Progressive scan**, avoiding serrated edges （锯齿） to moving objects

Table 5.4 Advanced digital TV formats supported by ATSC

Number of active pixels per line	Number of active lines	Aspect ratio	Picture rate
1,920	1,080	16:9	60P 60I 30P 24P
1,280	720	16:9	60P 30P 24P
720	480	16:9 or 4:3	60P 60I 30P 24P
640	480	4:3	60P 60I 30P 24P

  - The bandwidth for 1080p 30fps (4:2:0 format) HDTV video at 1s ?
  - 1920x1080x30x1.5 Bytes = 93MB

Digital Video

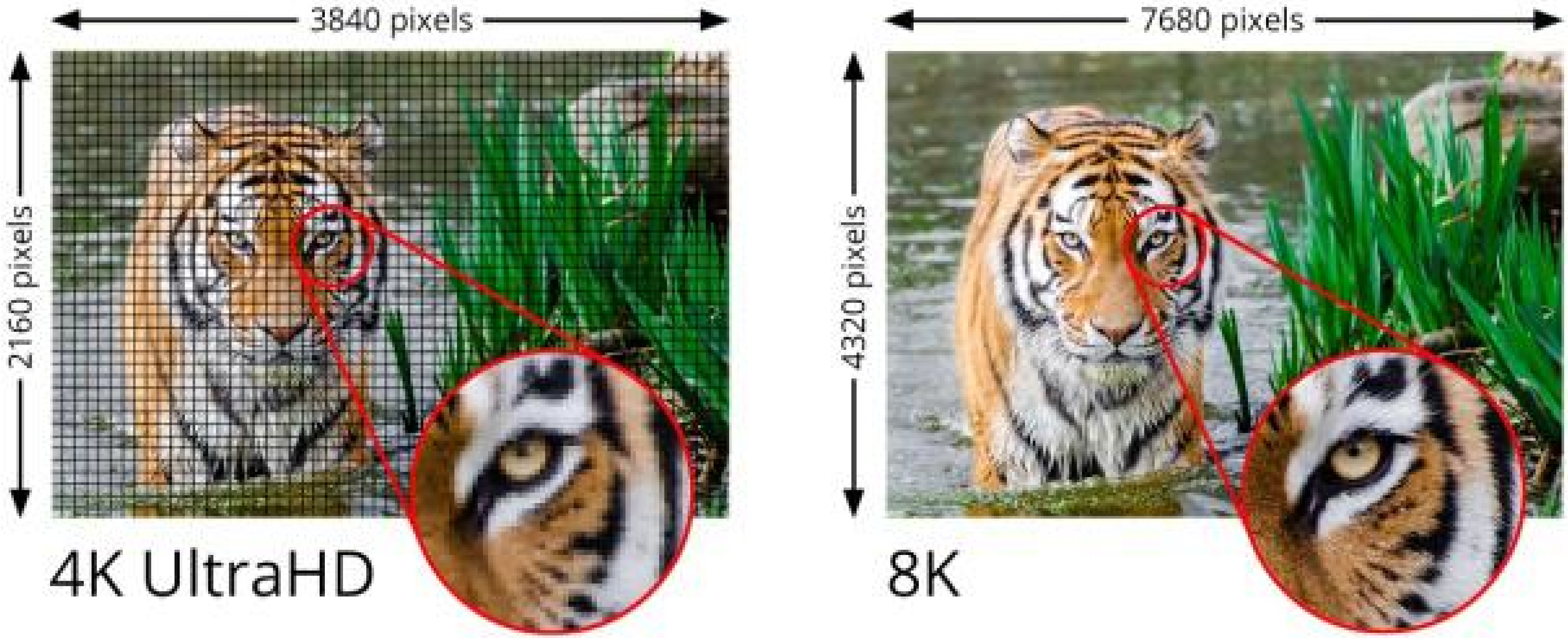
- Ultra High Definition TV (UHDTV)
  - **4K** UHDTV: 2160P (3840×2160)
  - **8K** UHDTV: 4320P (7680 × 4320)

  Aspect ratio: 16:9, Bit-depth: up to 12 bits,

  Frame Rate: up to 120

Digital Video

- Ultra High Definition TV (UHDTV)



Outline of Lecture 04

- Analog Video
- Digital Video
  - Chroma Subsampling
  - CCIR and ITU-R Standards for Digital Video
  - High-Definition TV
  - Ultra High Definition TV
- Video Display Interfaces
  - Analog Display Interfaces
  - Digital Display Interfaces
- 3D Video and TV

Video Display Interfaces

- Analog Display Interfaces

  - Component Video

  a) Three separate video signals for the red, green, and blue image planes

  b) It gives the best color reproduction, since there is no “crosstalk” between the three different channels.

  c) Requires more bandwidth and good synchronization of the three components.

Video Display Interfaces

- Analog Display Interfaces

  - Composite Video

  a) Color ("chrominance") and intensity ("luminance") signals are mixed into a single carrier wave.

  b) Chrominance is a composite of two color components (I and Q, or U and V).

  c) Used by broadcast color TV. **Interference** between the luminance and chrominance signals.

Video Display Interfaces

◆ Analog Display Interfaces

- S-Video

a) S-video (separated video, or super-video, e.g., in S-VHS) uses two wires: one for luminance and another for a composite chrominance signal.

b) There is less crosstalk between the color information and the crucial grayscale information.

Video Display Interfaces

◆ Analog Display Interfaces

- Video Graphics Array (VGA)

a) First introduced by IBM in 1987. It has since been widely used in the computer industry

b) The VGA video signals are based on analog component RGBHV(red, green, blue, horizontal sync, vertical sync).

c) VGA: 640x480 (70Hz)->QXGA: 2048x1536 (85Hz).

Video Display Interfaces

- Digital Display Interfaces

  - Digital Visual Interface (DVI)

  a) It was developed for transferring digital video signals, particularly from a computer's video card to a monitor.

  b) It carries uncompressed digital video and support DVI-D, DVI-A (analog), or DVI-I (Both), and VGA.

  c) Its transmission format is based on PanelLink, a high-speed serial link technology using transition minimized differential signaling (TMDS).

Video Display Interfaces

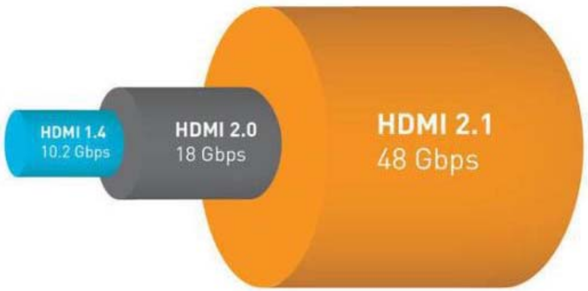
- Digital Display Interfaces

  - High-Definition Multimedia Interface (HDMI)

  a) Backward-compatible with DVI. It has been widely used in the consumer market since 2002.

  b) HDMI supports both *RGB* and *YCbCr* 4:4:4 or 4:2:2.

  c) HDMI supports digital audio.



版本	2.1版	2.0版
画质	3D视频 支持8K，7680x4320分辨率 支持4K/120Hz、8K/60Hz 静态HDR(HDR静态数据) 动态HDR(HDR动态数据) 可变刷新率(VRR) 快速媒体切换(QMS) 自动低延模式(ALLM)	3D视频 支持4K，3840x2160分辨率 支持4K/60Hz 静态HDR(HDR静态数据)
音质	支持多音频流 32个音频通道 自动音视频同步 1536kHz音频采样率 加强音频回传通道(eARC)	支持多音频流 32个音频通道 自动音视频同步 1536kHz音频采样率
速率	48Gbps总带宽 快速帧传输(QFT)	18Gbps总带宽

Video Display Interfaces

◆ Digital Display Interfaces

- DisplayPort

a) DisplayPort is a digital display interface developed by VESA, starting from 2006.

b) It is the first display interface that uses **packetized data** transmission, like the Internet or Ethernet .

c) 2019, DP2.0, support 8K, even up to 10K.

d) It is royalty-free. make DisplayPort a strong competitor to HDMI in the consumer electronics market

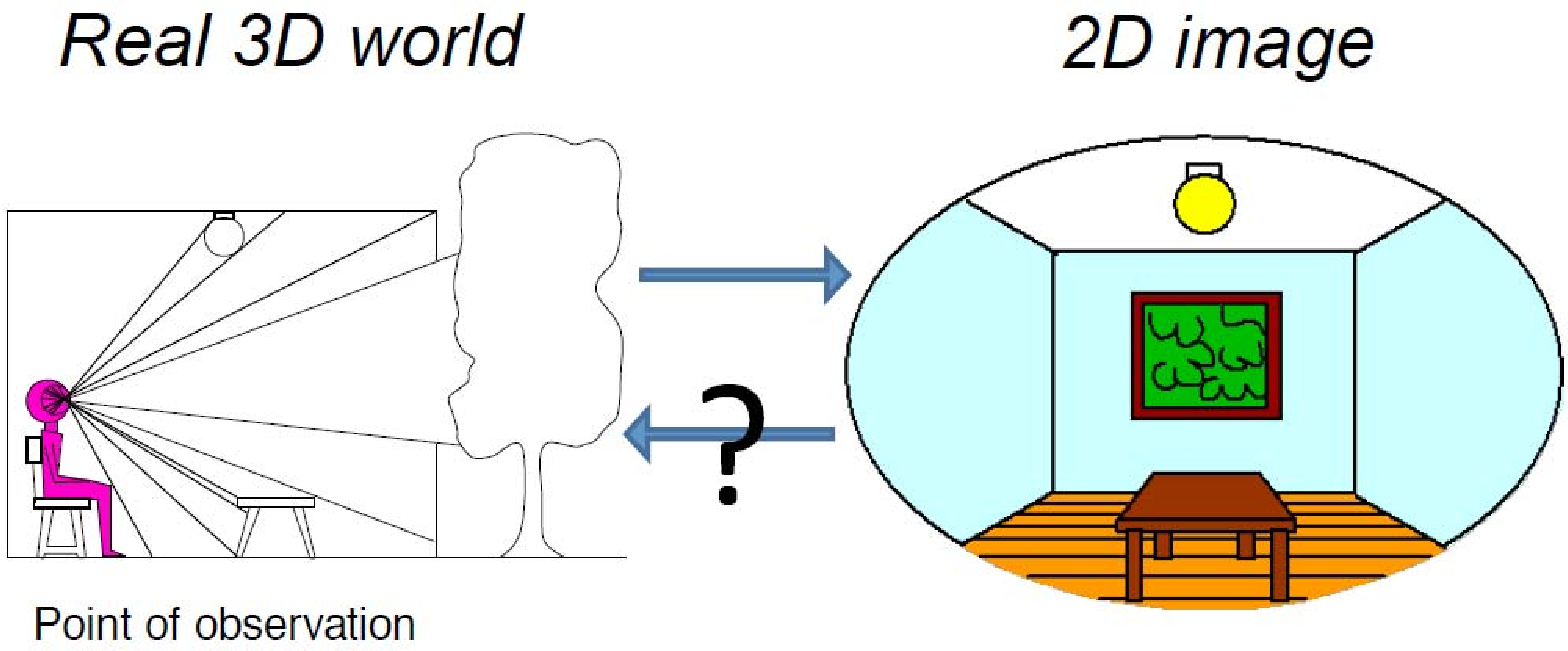
https://baijiahao.baidu.com/s?id=1641627122287252962&wfr=spider&for=pc

Outline of Lecture 04

- Analog Video
- Digital Video
  - Chroma Subsampling
  - CCIR and ITU-R Standards for Digital Video
  - High-Definition TV
  - Ultra High Definition TV
- Video Display Interfaces
  - Analog Display Interfaces
  - Digital Display Interfaces
- 3D Video and TV

3D Video and TV

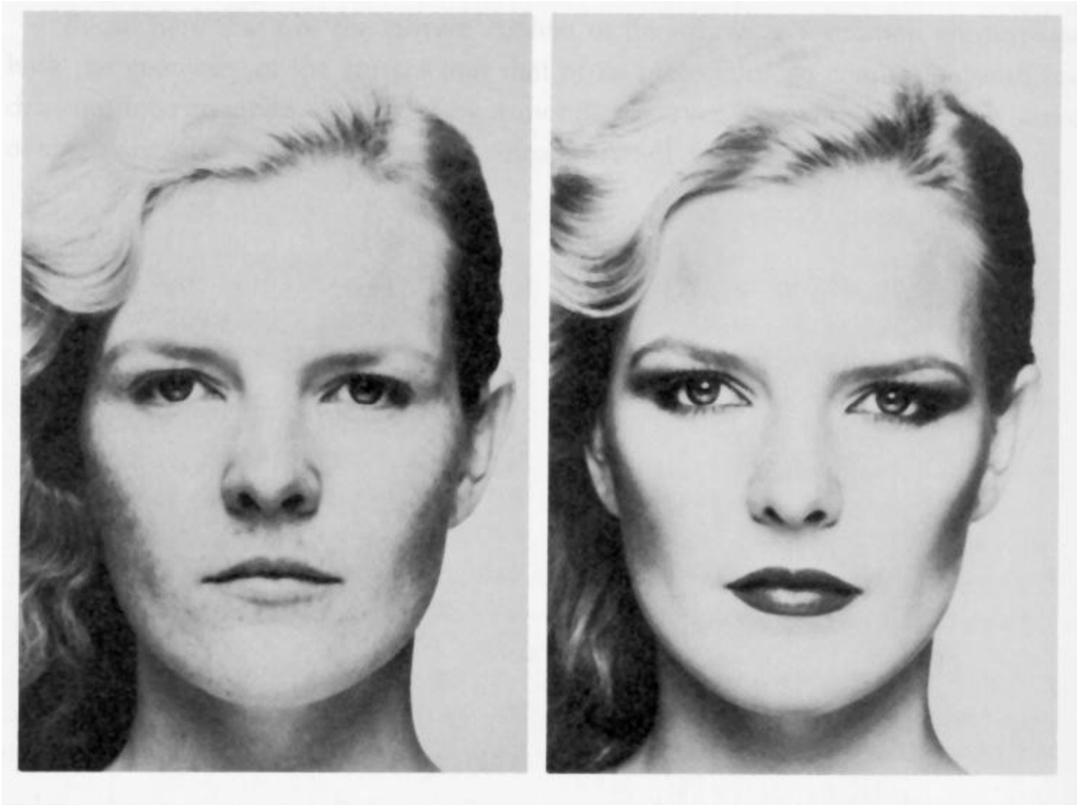
- Cues for 3D Percept
  - **How does the human eye automatically extract 3D geometric information from an image?**



3D Video and TV

- Cues for 3D Percept
  - **How does the human eye automatically extract 3D geometric information from an image?**
  - **Monocular Cues**

a) Shading



3D Video and TV

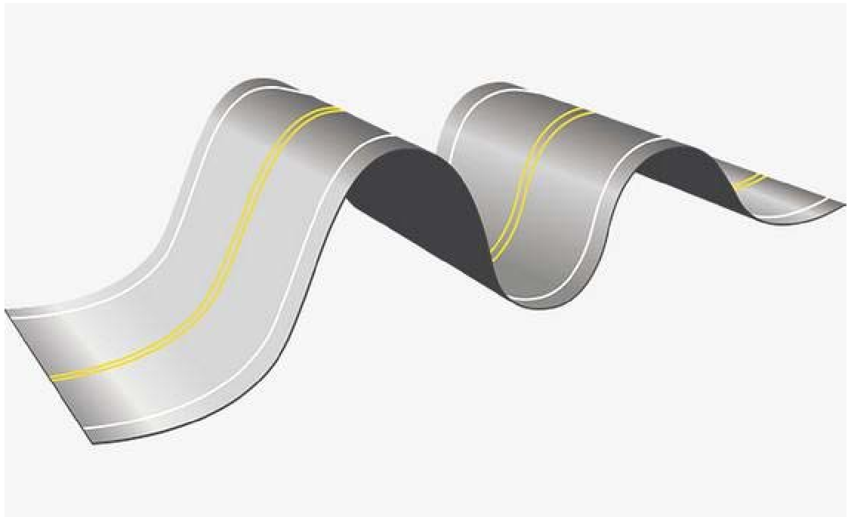
◆ Cues for 3D Percept

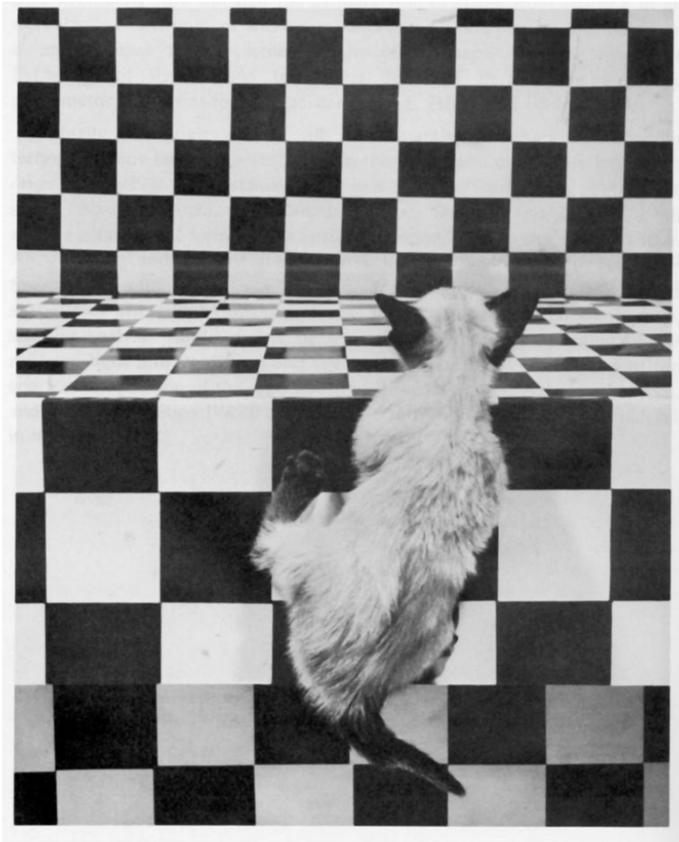
- **How does the human eye automatically extract 3D geometric information from an image?**

- **Monocular Cues**

a) Shading

b) Texture





Borrowed from Prof Fei-Fei Li, Standford Vision Lab

3D Video and TV

- Cues for 3D Percept
  - **How does the human eye automatically extract 3D geometric information from an image?**
  - **Monocular Cues**

a) Shading

b) Texture

c) Focusing

3D Video and TV

- Cues for 3D Percept
  - **How does the human eye automatically extract 3D geometric information from an image?**
  - **Monocular Cues**

a) Shading
b) Texture
c) Focusing
d) Motion

Borrowed from Prof Fei-Fei Li, Standford Vision Lab

3D Video and TV

- Cues for 3D Percept
  - **How does the human eye automatically extract 3D geometric information from an image?**
  - **Monocular Cues**

a) Shading

b) Texture

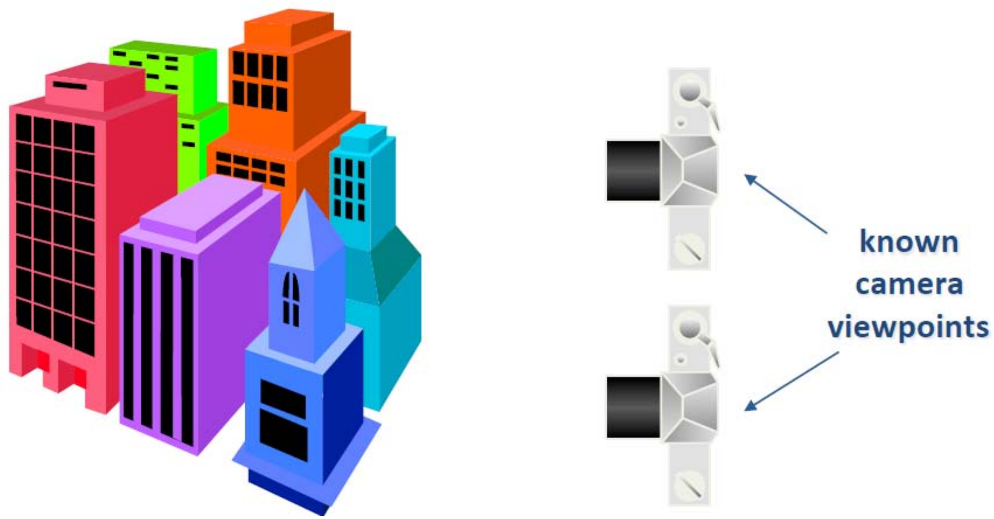
c) Focusing

d) Motion

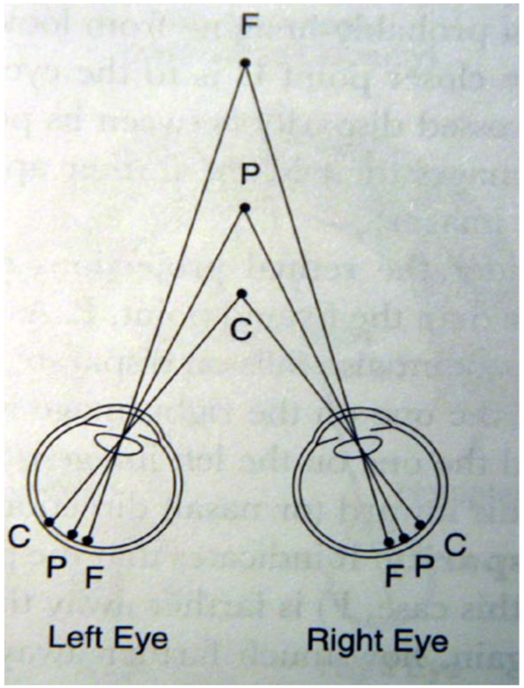
e) Others: Occlusion, Perspective scaling, Relative size, Haze, etc.

3D Video and TV

- Cues for 3D Percept
  - **Binocular Cues**
  - Our left and right eyes are separated by a small distance -- *interocular distance* (瞳孔距离)
  - Images of objects are shifted horizontally – *disparity* (视差). It is dependent on the object's distance from the eyes – *depth.*



Borrowed from Prof Fei-Fei Li, Standford Vision Lab



3D Video and TV

- Cues for 3D Percept
  - **Binocular Cues**
  - Our left and right eyes are separated by a small distance -- *interocular distance* (瞳孔距离)
  - Images of objects are shifted horizontally – disparity (视差). It is dependent on the object's distance from the eyes – depth.

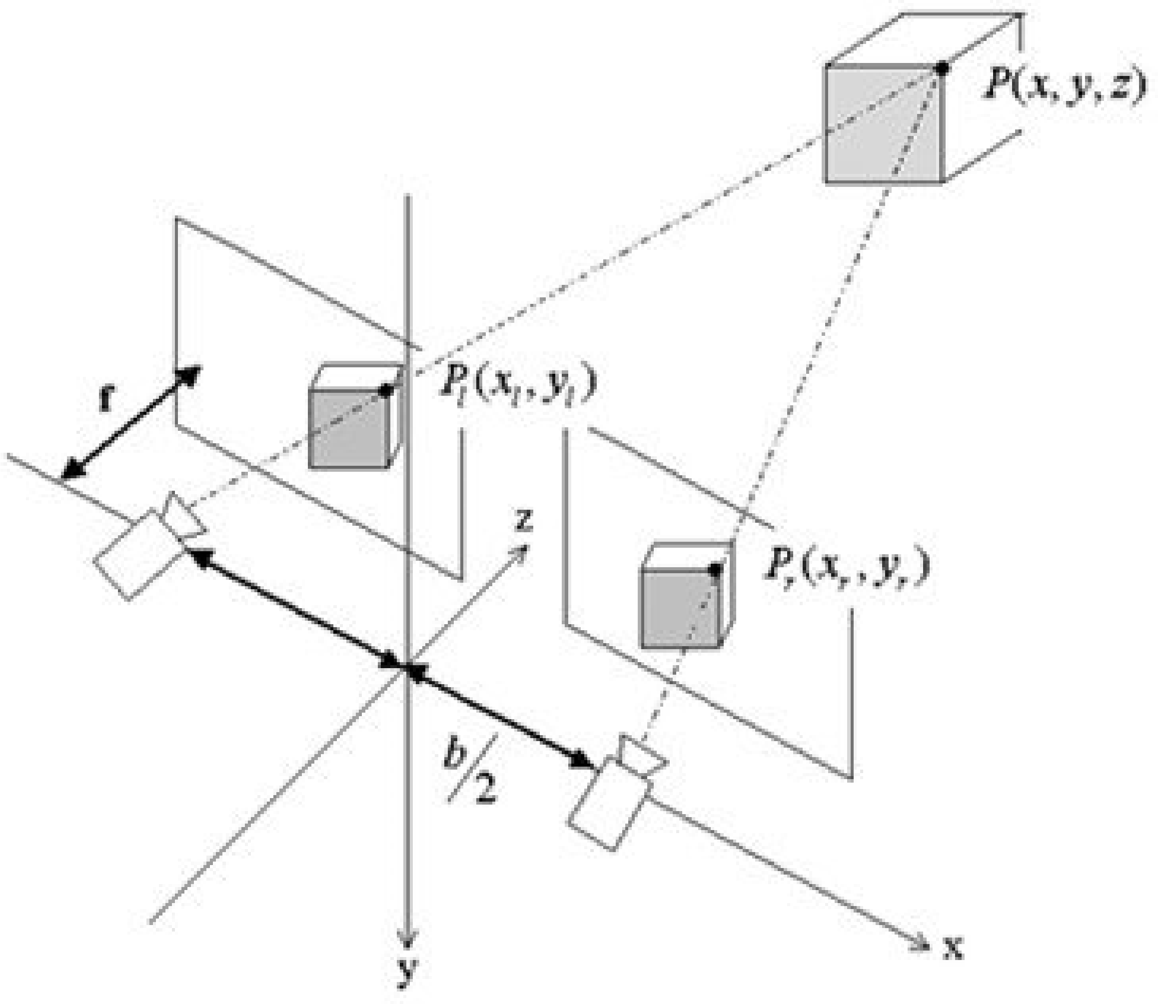
Left

Right

Combined

3D Video and TV

- 3DCamera Model

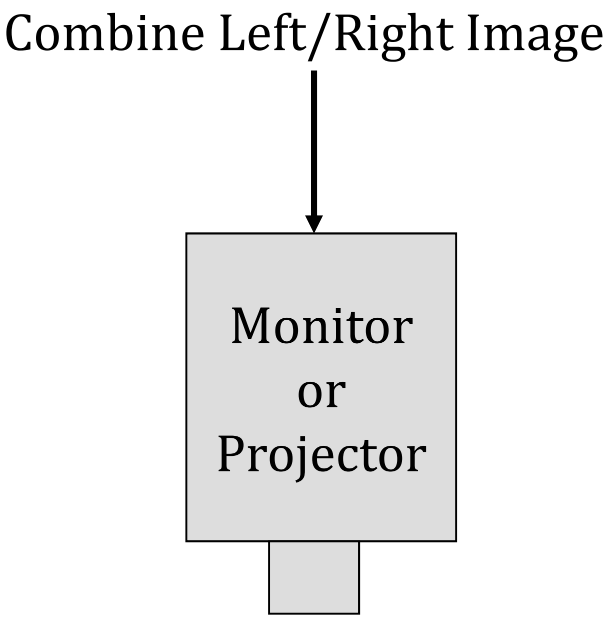


3D Video and TV

- 3DMovie and TV Based on Stereo Vision

  - Passive

  a) Colored Glasses



3D Video and TV

- 3DMovie and TV Based on Stereo Vision

  - Passive

  a) Colored Glasses

  b) Circularly Polarized Glasses (圆形偏光)

3D Video and TV

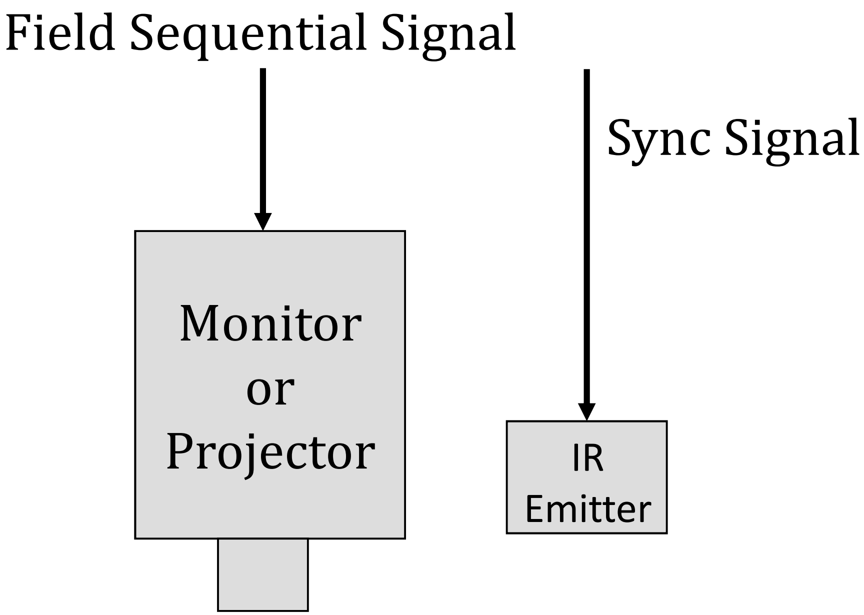
- 3DMovie and TV Based on Stereo Vision

  - **Passive**

    a) Colored Glasses

    b) Circularly Polarized Glasses (圆形偏光)

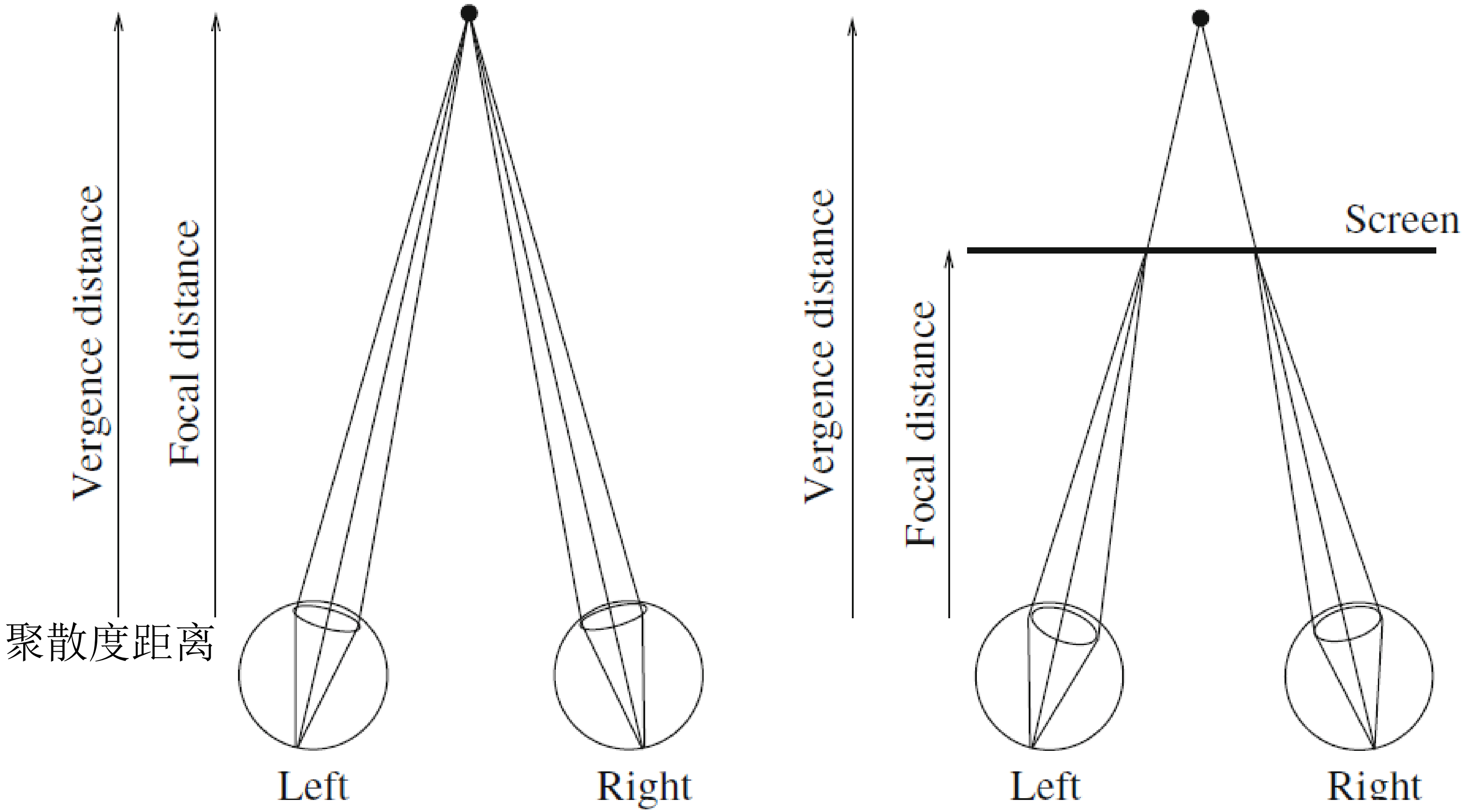
  - **Active** - TV with Shutter Glasses



LCD Shutter Glasses

3D Video and TV

- 3DMovie and TV Based on Stereo Vision
  - **Autostereoscopic** (Glasses-Free) Display Devices

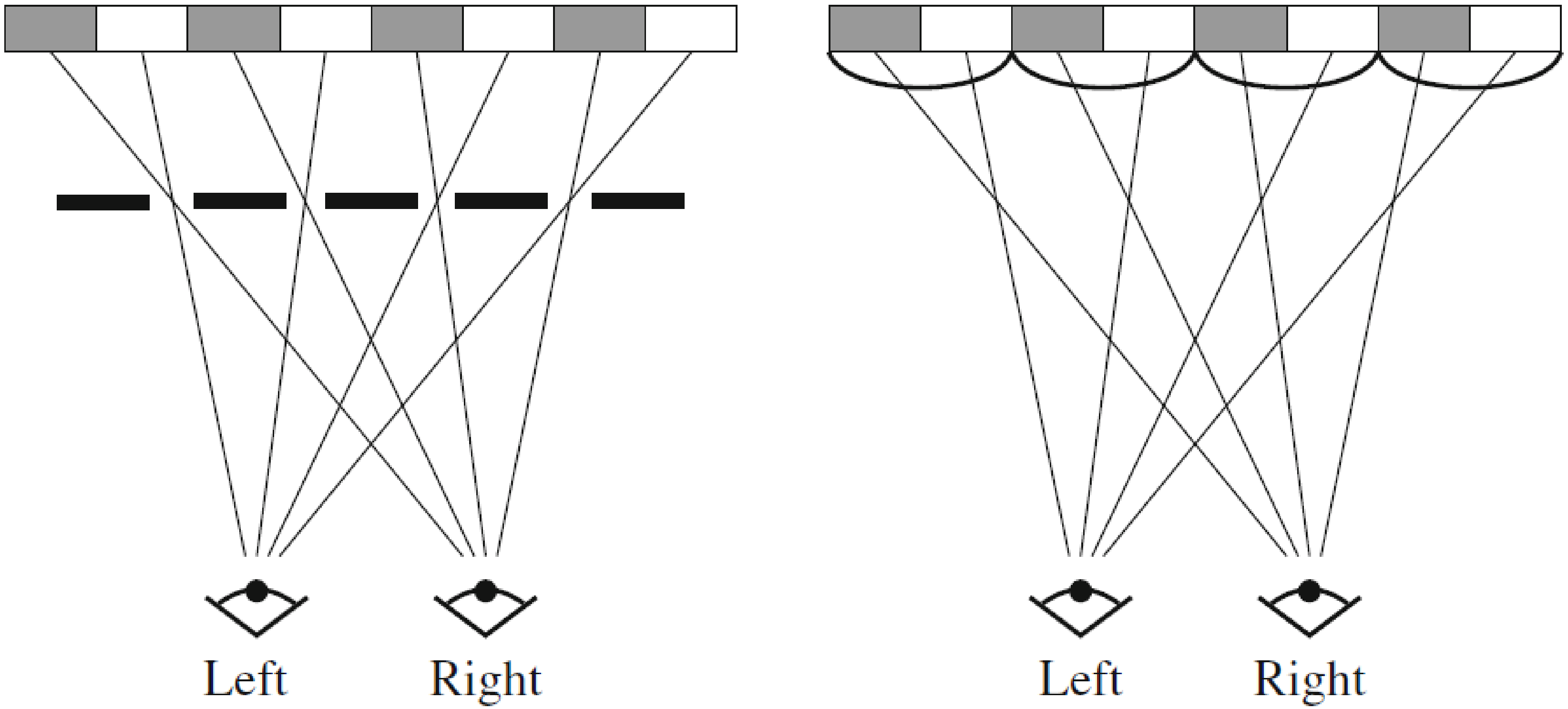


视觉辐辏调节冲突

The Vergence-Accommodation Conflict. **a)** Real World and **b)** 3D Display

3D Video and TV

- 3DMovie and TV Based on Stereo Vision
  - **Autostereoscopic** (Glasses-Free) Display Devices



Autostereoscopic display devices. **a)** Parallax Barrier and **b)** Lenticular Lens

自由立体显示设备 视差屏障技术 柱状透镜

Experiments and Class Assignments

- Experiments: Chroma Subsampling

  - *demo_chromesubsampling.m*

- Class Assignments

  1、隔行扫描视频的优点是什么？它存在哪些问题?

  2、数字视频采用色度二次采样，目的是什么？为什么是可行的?