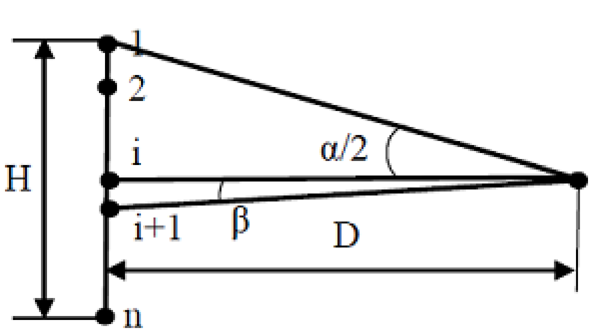
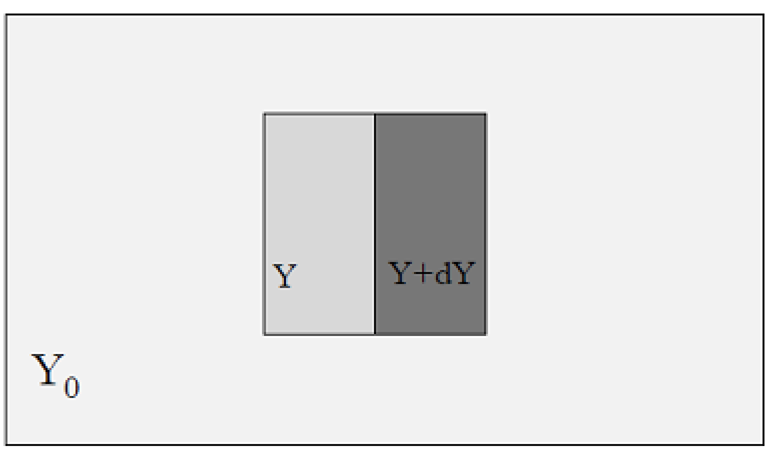
Curs 1

1. Enumerate the limitations of the human vision systems

Visual acuity (in eng. also “acuteness” or “clearness” of vision)
“measures the capacity of the human eye to distinguish spatial details at a given distance”
e.g. different objects and shapes of same light intensity
e.g of a test : smaller and smaller letters on a standard eye chart
α – whole screen viewing angle
β – separation angle
D – visualisation angle
H – image/video height
FROM EXPERIMENTAL CONCLUSIONS
-> D/H = 4 … 5
-> β ~ 1’ (1/60 of a degree)

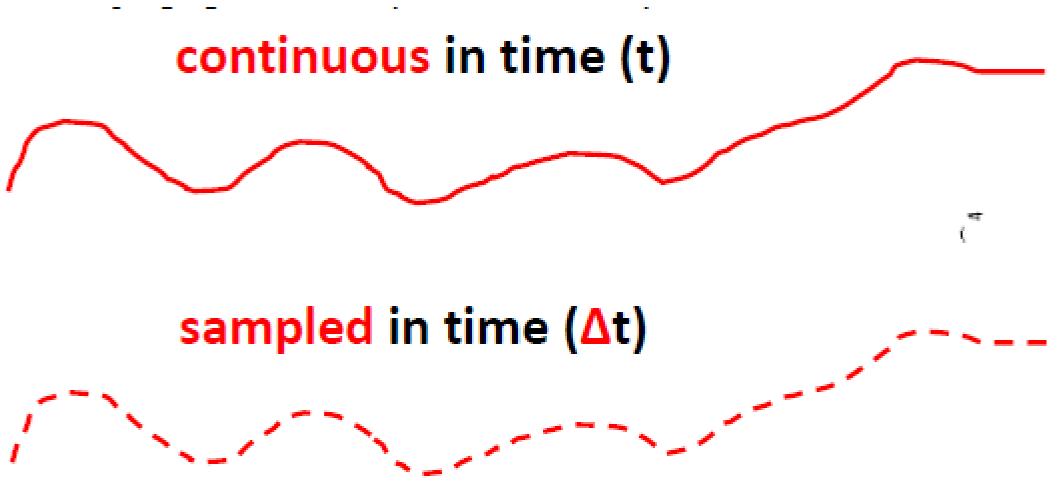


Contrast sensitivity
“measures the capacity of the human eye to distinguish details of different objects from their background based on their difference in luminance and/or colour”



Persistance of vision
“measures the capacity of the human eye to retain the “impression” made by the image even after the image is gone”linked to the perception of motionand image flickering

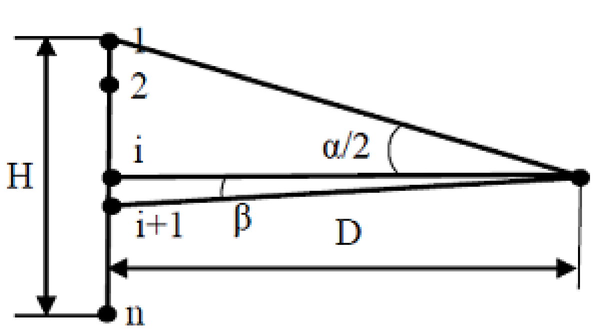
2.The sampling algorithm used by electronic vision systems



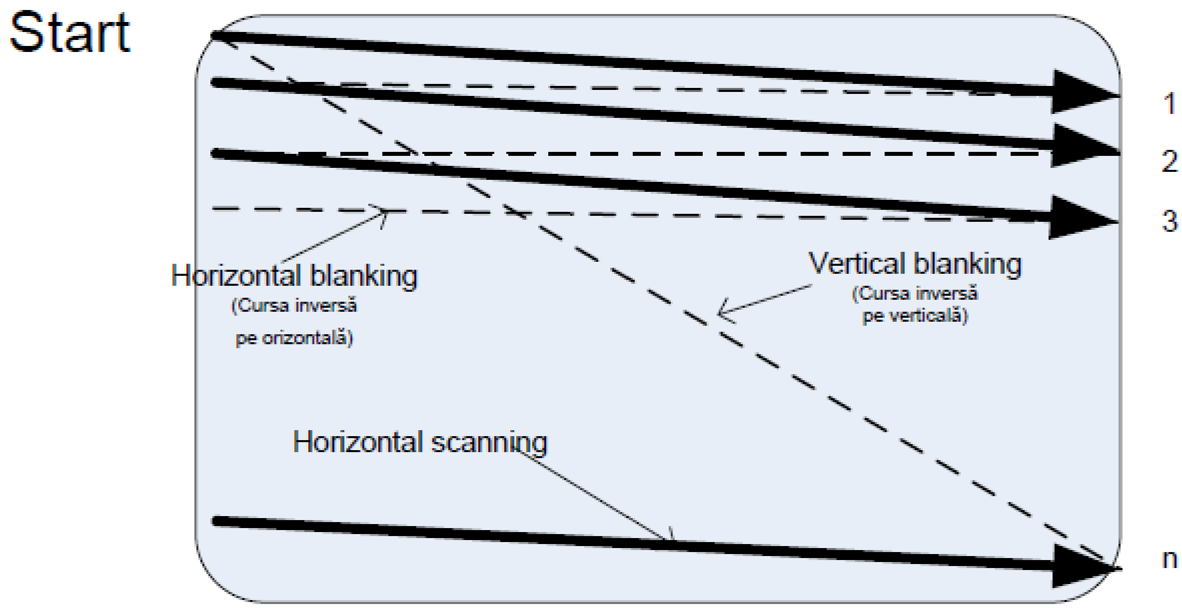
1 sample in time at (t) = 1 image (2D or 3D)
1 sample in time at (t) = 1 frame (in TV&videoengineering)
1 sample in 2D space at (y and x) = 1 pixel (in TV&videoengineering !)
this is DIGITAL TV and video
Frame rate = number of frames per second
Frame rate a.k.a. “vertical frequency” (in TV)
It is a frequency -> it is measured in Hz
Frame rate ~ 25 Hz (minimum)
Number of lines

$$n = \frac{2 \cdot arctan \frac{H}{2 \cdot D}}{\beta} \approx 440 \div 660 \; lines$$

where : FROM EXPERIMENTAL CONCLUSIONS
-> D/H = 4 … 5
-> β ~ 1’ (1/60 of a degree)

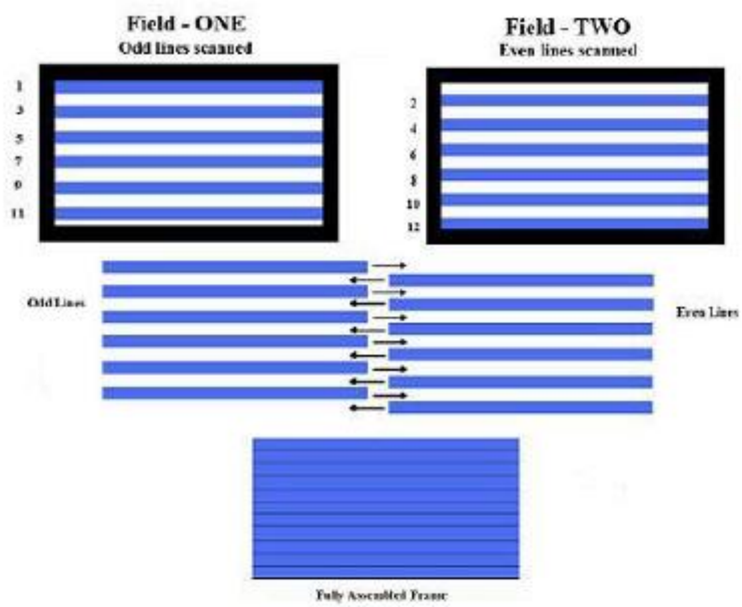


How are the lines ordered ? -> from technical availability
scanning = methods of “exploring the screen” and separate it into lines (it is a process)
progressive scanning = one line after the other



interlaced scanning = 1 frame divided into 2 fields

Even numbered lines go into FIELD1
Odd numbered lines go into FIELD2
-> number of lines should be ODD!



Line issues
-> number of lines should be ODD!
(depending on how much bandwidth is available)
-> number of lines should be more than ~ 440
(depending on what is considered “optimal” for the D/H ratio

Curs 2

1.The basic components of the TV signal

(Din LABORATOR)
Semnalul de televiziune, sau cum se mai numeşte „semnalul complex de televiziune” conţine mai multe componente. În primul rând avem semnalul electric care poartă informaţia privitoare la strălucirile imaginii. Acest semnal se numeşte „semnal video” sau semnal de imagine.

În afară de aceasta, sunt necesare semnale care să blocheze fasciculul de electroni care realizează analiza, respectiv sinteza imaginii, în timpul întoarcerilor fasciculului de electroni. Aceste semnale se numesc semnale de stingere şi sunt două semnale distincte: unul care asigură stingerea pe verticală BV (semnal de stingere V) şi unul pe orizontală BH (semnal de stingere H).

Mai este necesară asigurarea sincronizării între emisie şi recepţie. Este strict necesar, pentru redarea unei imagini cu aceeaşi configuraţie geometrică cu cea iniţială, ca generatoarele de baleiaj din receptor să aibă aceeaşi frecvenţă cu cele de la emiţător. Semnalele folosite în acest scop sunt numite semnale de sincronizare. Ca şi în cazul semnalelor de stingere, vom avea semnale de sincronizare pe orizontală – „semnale sincro

H” (SH), respectiv semnale de sincronizare pe verticală – „semnale sincro V” (SV). Ansamblul acestor semnale constituie semnalul complex de sincronizare.

2. How the TV signal is constructed

Although over 50 years old, the standard television signal is still one of the most common way to transmit an image. Figure 23-9 shows how the television signal appears on an oscilloscope.

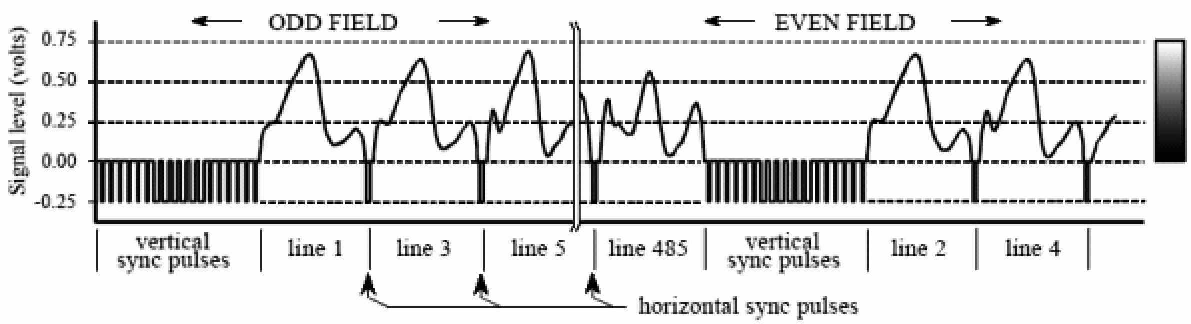


FIGURE 23-9
Composite video. The NTSC video signal consists of 30 complete frames (images) per second, with each frame containing 480 to 486 lines of video. Each frame is broken into two fields, one containing the odd lines and the other containing the even lines. Each field starts with a group of vertical sync pulses, followed by successive lines of video information separated by horizontal sync pulses. (The horizontal axis of this figure is not drawn to scale).

This is called composite video, meaning that there are vertical and horizontal synchronization (sync) pulses mixed with the actual picture information. These pulses are used in the television receiver to synchronize the vertical and horizontal deflection circuits to match the video being displayed. Each second of standard video contains 30 complete images, commonly called frames. A video engineer would say that each frame contains 525 lines, the television jargon for what programmers call rows. This number is a little deceptive because only 480 to 486 of these lines contain video information; the remaining 39 to 45 lines are reserved for sync pulses to keep the television's circuits synchronized with the video signal.

Standard television uses an interlaced format to reduce flicker in the displayed image. This means that all the odd lines of each frame are transmitted first, followed by the even lines. The group of odd lines is called the odd field, and the group of even lines is called the even field.

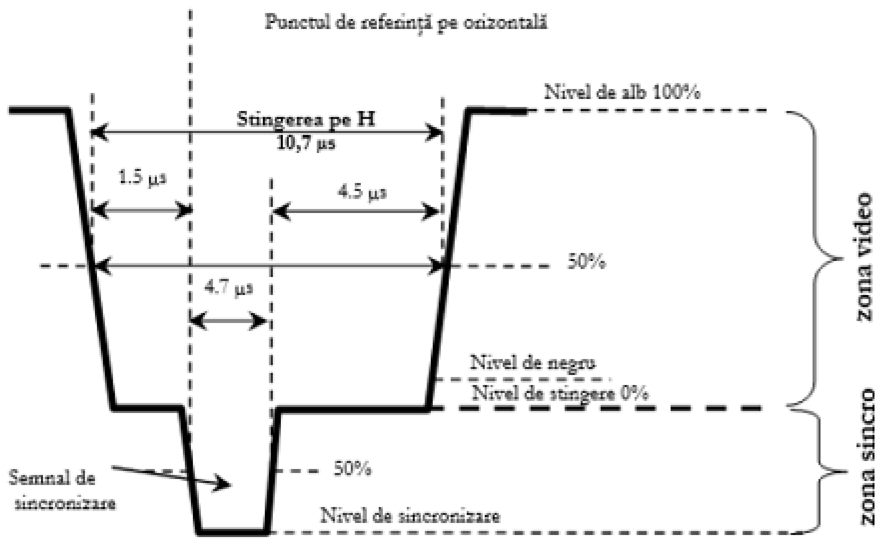
Since each frame consists of two fields, the video signal transmits 60 fields per second. Each field starts with a complex series of vertical sync pulses lasting 1.3 milliseconds. This is followed by either the even or odd lines of video. Each line lasts for 63.5 microseconds, including a 10.2 microsecond horizontal sync pulse, separating one line from the next. Within each line, the analog voltage corresponds to the grayscale of the image, with brighter values being in the direction away from the sync pulses. This places the sync pulses beyond the black range. In video jargon, the sync pulses are said to be blacker than black.

3. The basic elements of TV signal processing

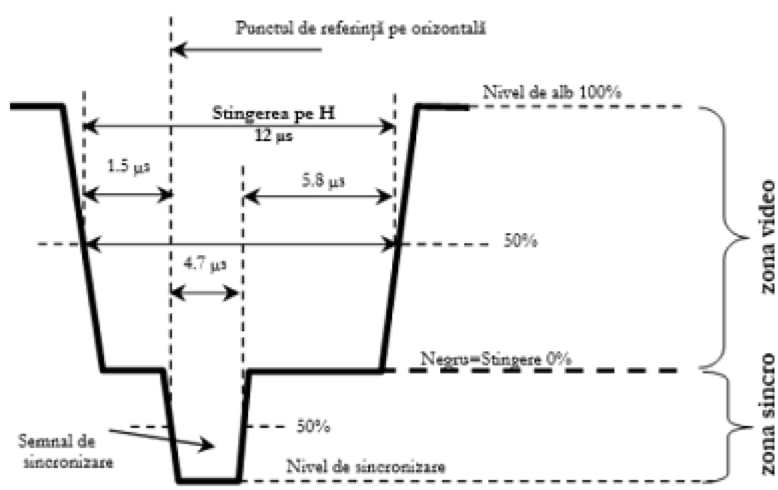
The hardware used for analog-to-digital conversion of video signals is called a frame grabber. This is usually in the form of an electronics card that plugs into a computer, and connects to a camera through a coaxial cable. Upon command from software, the frame grabber waits for the beginning of the next frame, as indicated by the vertical sync pulses. During the following two fields, each line of video is sampled many times, typically 512, 640 or 720 samples per line, at 8 bits per sample. These samples are stored in memory as one row of the digital image.

Standard values for the TV signal

The 525/60 standard



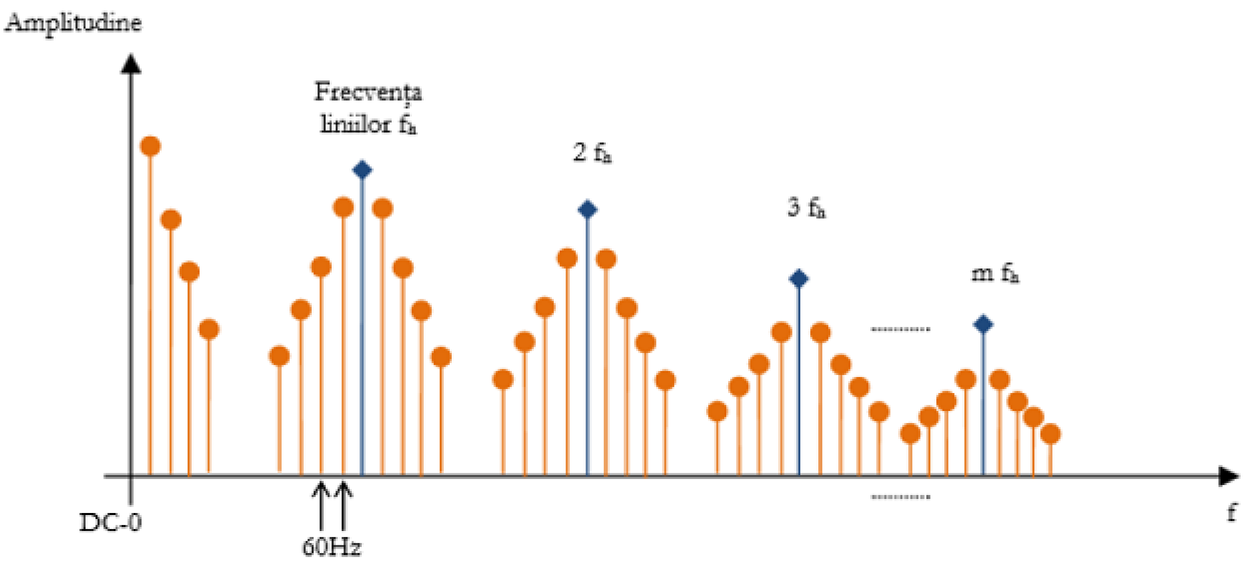
The 625/50 standard



Spectrum

the set of harmonic components in a television signal. The width and structure of the spectrum depend on the scanning parameters and content of the image being transmitted. The lower limit of a television signal spectrum with line-by-line scanning is assumed to be equal to the frame frequency or, with interlaced scanning, to the field frequency. A DC component, characterizing the average brightness of the image, is usually not present as such in a television signal. The upper limit of the spectrum fmax is determined from the transmission conditions of the major harmonic components for alternating black and white image elements along a line; fmax = 1/2KnpZ2, where K is a constant (usually 0.6–0.9), n is the frame frequency, ρ is the aspect ratio of the frame (ratio of width to height), and Z is the number of lines. With the television standards adopted in the USSR, n = 25 per sec, Z = 625, p = 4/3; when K = 0.9, fmax = 6 megahertz (MHz).

The television signal spectrum for a stationary black-and-white image, as well as the signal spectrum for the brightness of a stationary color image, is discrete and consists of separate groups of spectral lines, which are formed by harmonics of the line frequency fine and by side lines. In each group, the most intense harmonic is fnne. When objects move and the content of the transmitted images changes, side bands having a continuous spectrum appear around the discrete spectral lines; the width of a band usually does not exceed several hertz.



C3.1 Identify the basic components of light

In physical terms, light constitutes a small section in the range of **electromagnetic radiation**, extending in wavelength from about 400 to 700 nanometers.

$$\text{Wavelength} = \frac{\text{Speed of light}}{\text{Frequency}}$$

Visible Light

400 nm	Ultraviolet
450 nm	Violet
500 nm	Blue
550 nm	Green
600 nm	Yellow
	Orange
650 nm	
700 nm	
750 nm	Red
800 nm	Infrared

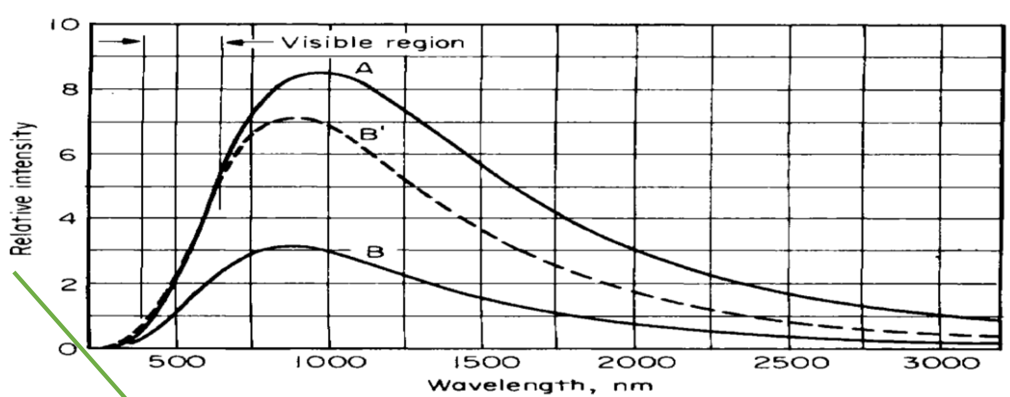
Under ideal conditions, the human visual system can detect:

- Wavelength differences of 10^–8 cm(=> **sensitivity of 10^-8cm**)
- Intensity differences as little as **1 %**

When a beam of light traveling in air falls upon a glass surface at an angle, it is *refracted* or bent. **The amount of refraction depends upon the wavelength,** its variation with wavelength being known as *dispersion*. Similarly, when the beam, traveling in glass, emerges into air, it is refracted (with dispersion). A glass prism provides a refracting system of this type. Because different wavelengths are refracted by different amounts, an incident white beam is split up into several beams corresponding to the many wavelengths contained in the composite white beam. This is how the spectrum is obtained.

- **Spectrum**

A plot of the power distribution of a source of light is indicative of the watts radiated at each wavelength per nanometer of wavelength.



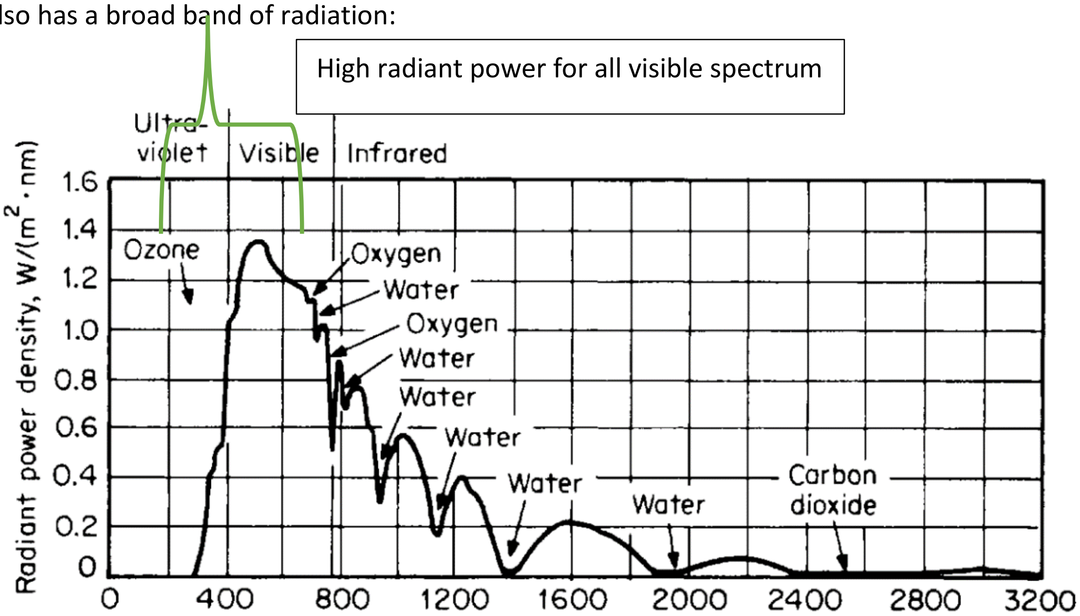
Radiant Intensity is the radiant flux leaving a point on the source, per unit solid angle of space surrounding the point. Unit: watts per steridian(asta ar veni: unghi solid,este fara dimensiune si fara prea mare importanta zic eu), W sr-1. (definitie de pe slide).

- A,B,C – 3 cazuri, suprafete din materiale si dimensiune diferite.

White light perception

A very broad band extending throughout the visible spectrum is perceived as white light. (asta e si intuitiv din moment ce lumina e o combinatie de unde electromagnetice cu wavelength diferite ce se intind pe tot spectrul vizibil 400-700 nm).

Daylight also has a broad band of radiation:



Individual narrow bands of wavelengths of light are seen as strongly colored elements. Increasingly broader bandwidths retain the appearance of color, but with decreasing purity, as if white light had been added to them.

A sensation of white light can also be induced by light sources that do not have a uniform energy distribution. Among these is fluorescent lighting, which exhibits sharp peaks of energy through the visible spectrum.

Notiuni de pe slide (un indicator bun imi pare unitatea de masura)

Irradiance is the radiant flux incident on a receiving surface from all directions, per unit area of surface. Unit: **W m-2.**

Spectroradiometry: All the properties of the radiant flux depend on the wavelength of the radiation. The prefix spectral is added when the wavelength dependency is being described. Thus, the spectral irradiance is the irradiance at a given wavelength, per unit wavelength interval. The irradiance within a given waveband is the integral of the spectral irradiance with respect to wavelength (8). Unit: **W m-2 nm-1.**

Note: **Spectroradiometers** are devices designed to measure the spectral power distribution of a source. Spectroradiometers typically take measurements of spectral irradiance and spectral radiance.

Absorptance is the fraction of the incident flux that is absorbed by a medium. **Reflectance** and **Transmittance** are equivalent terms for the fractions that are reflected or transmitted.

Radiant Flux is the amount of radiation coming from a source per unit time. Unit: **watt, W.**

Radiance is the radiant flux emitted by a unit area of a source or scattered by a unit area of a surface. Unit**: W m-2 sr-1.**

C3.2 Understand how human and machine perceive color

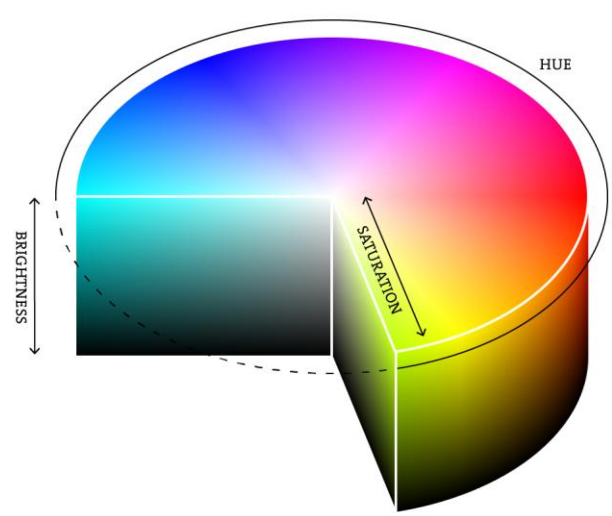
The color sensation associated with a light stimulus can be described in terms of three characteristics:

- **Hue** - is the actual color. It is measured in angular degrees counter-clockwise around the cone starting and ending at red = 0 or 360 (so yellow = 60, green = 120, etc.).

- **Saturation** - is the purity of the color, measured in percent from the center of the cone (0) to the surface (100). At 0% saturation, hue is meaningless.

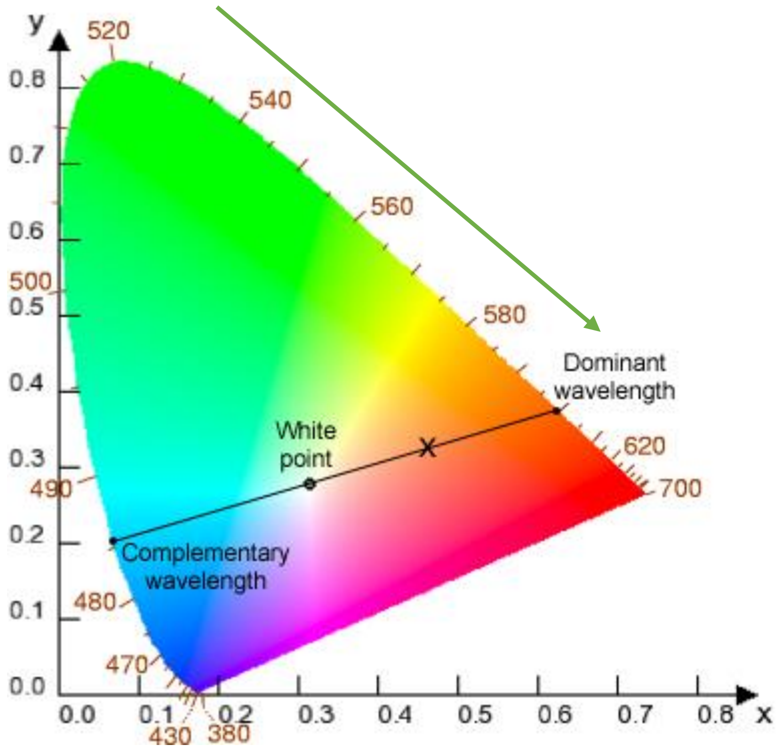
  - Saturation pertains to the strength of the hue. Spectrum colors are highly saturated. White and grays have no hue and, therefore, have zero saturation. Pastel colors have low or intermediate saturation

- **Brightness** - is measured in percent from black (0) to white (100). At 0% brightness, both hue and saturation are meaningless.
  - Brightness pertains to the intensity of the stimulation. If a stimulus has high intensity, regardless of its hue, it is said to be brig



The psychophysical analogs of hue, saturation, and brightness are :

- **Dominant wavelength**



- Intuitively, the dominant wavelength of "x" corresponds to the primary hue of "x".

- **Excitation purity**
- **Luminance**

By using definitions and standard response functions, which have received international acceptance through the International Commission on Illumination, the dominant wavelength, purity, and luminance of any stimulus of known spectral energy distribution can be determined by simple computations. Although roughly analogous to their psychophysical counterparts, the psychological attributes of hue, saturation, and brightness pertain to observer responses to light stimuli and are not subject to calculation. These sensation characteristics—as applied to any given stimulus—depend in part on other visual stimuli in the field of view and upon the immediately preceding stimulations

Color sensations arise directly from the action of light on the eye. They are normally associated, however, with objects in the field of view from which the light comes. The objects themselves are therefore said to have color. Object colors may be described in terms of their hues and saturations, such as with light stimuli. The intensity aspect is usually referred to in terms of lightness, rather than brightness. The psychophysical analogs of lightness are *luminous reflectance* for reflecting objects and *luminous transmittance* for transmitting objects.

At low levels of illumination, objects may differ from one another in their lightness appearances, but give rise to no sensation of hue or saturation. All objects appear as different shades of gray. Vision at low levels of illuminatio [...] is c [...] *vision,* which takes place at higher lev [...] 2 c [...] values for photopic and scotopic vision.(daca va intereseaza continuarea tabelului sa stiti ca nu sunteti normali)

High level of illumination

Low level of illumination

Wavelength, nm	Photopic Vision	Scotopic Vision
390	0.00012	0.0022
400	0.0004	0.0093
410	0.0012	0.0348
420	0.0040	0.0966
430	0.0116	0.1998
440	0.023	0.3281
450	0.038	0.4550
460	0.060	0.5670
470	0.091	0.6760
480	0.139	0.7930

Only the rods of the retina are involved in scotopic vision; cones play no part. Because the *fovea centralis* is free of rods, scotopic vision takes place outside the fovea. The visual acuity of scotopic vision is low compared with photopic vision. (chestiuni de incultura generala).

Course 4-Television Standards

1) List main standardization bodies regulating the TV & video industry

Early standards were developed by equipment manufacturers under the banner of Radio Manufacturers Association(RMA),then the Radio, Electronic and Television Manufacturers Association(RETMA) and later the Electronic Industries Association(EIA).

Nowadays the main standardization bodies in the TV and video industry are:

-ITU-International Telecommunication Union

-ISO-International Standardization Organization

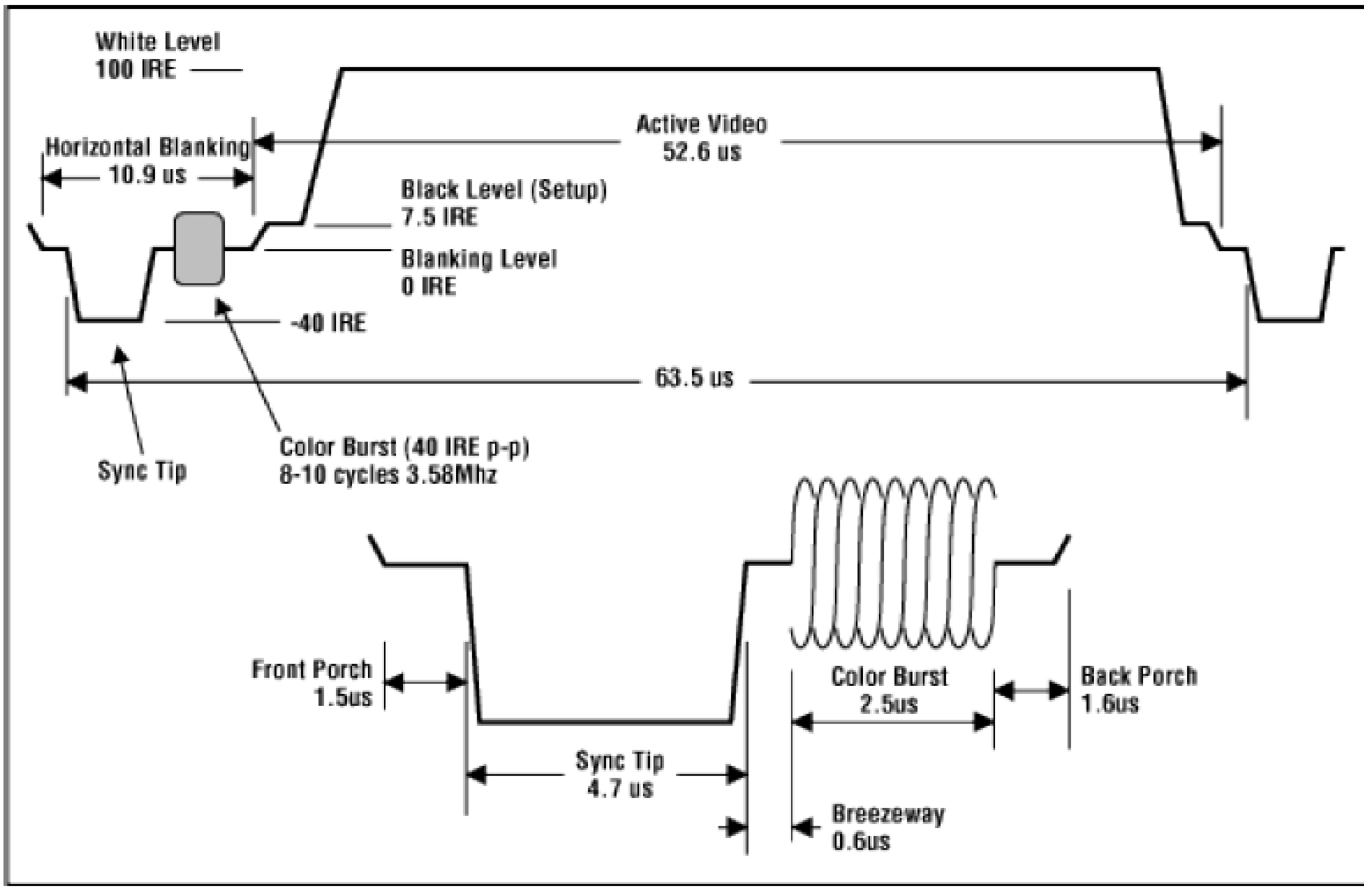
-IEC-International Electrotechnical Commission

2) Correlate TV signal properties with values and standardized specifications

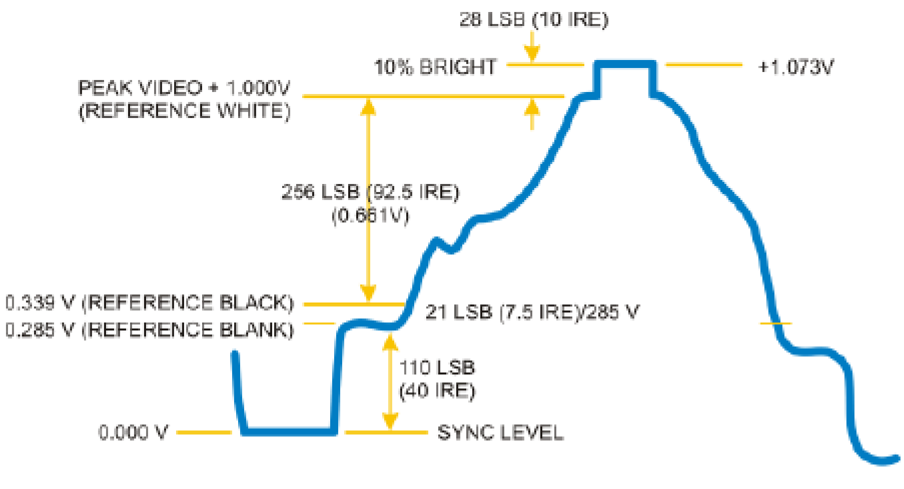
Broadcasters using analog television systems encode their signal using NTSC(National Television Systems Committee),PAL(Phase Alternation Line rate) or SECAM(Séquentiel

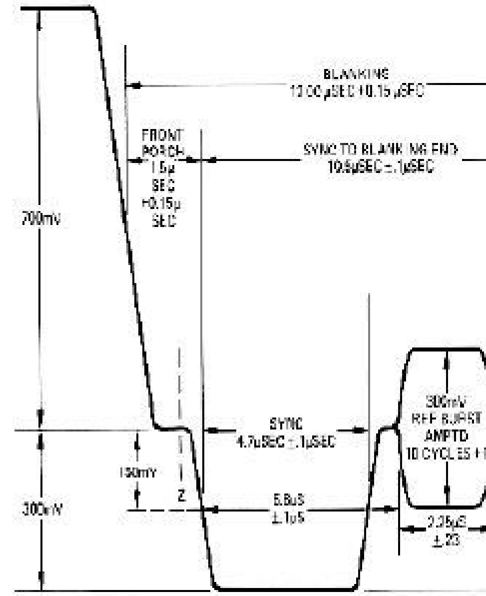
Couleur Avec Mémoire) analog encoding and then use RF modulation to modulate this signal onto a Very high frequency (VHF) or Ultra high frequency (UHF) carrier. PAL's color encoding is similar to the NTSC systems. SECAM, though, uses a different modulation approach than PAL or NTSC.

NTSC



PAL


28 LSB (10 IRE)
10% BRIGHT
+1.073V
PEAK VIDEO + 1.000V
(REFERENCE WHITE)
256 LSB (92.5 IRE)
(0.661V)
0.339 V (REFERENCE BLACK)
0.285 V (REFERENCE BLANK)
21 LSB (7.5 IRE)/285 V
110 LSB
(40 IRE)
0.000 V
SYNC LEVEL


BLANKING
FRONT
PORCH
SYNC
4.7uSEC ±.1μSEC
700mV
300mV
Z
REF BURST
AMPTD

Signal duration/frequency comparison

	NTSC M	PAL B,G,H	PAL N	PAL M	SECAM B,G,H	SECAM D,K,K',L
Lines/Fields	525/60	625/50	625/50	525/60	625/50	625/50
Horizontal Frequency	15.734 kHz	15.625 kHz	15.625 kHz	15.750 kHz	15.625 kHz	15.625 kHz
Vertical Frequency	60 Hz	50 Hz	50 Hz	60 Hz	50 Hz	50 Hz
Color Subcarrier Frequency	3.579545 MHz	4.433618 MHz	3.582056 MHz	3.575611 MHz		
Video Bandwidth	4.2 MHz	5.0 MHz	4.2 MHz	4.2 MHz	5.0 MHz	6.0 MHz
Sound Carrier	4.5 MHz	5.5 MHz	4.5 MHz	4.5 MHz	5.5 MHz	6.5 MHz

C5. Digital Television

1) Digital TV signal generation

Digital television (DTV) is the transmission of audio and video by digitally processed and multiplexed signal, in contrast to the totally analog and channel separated signals used by analog television. Digital TV can support more than one program in the same channel bandwidth. It is an innovative service that represents the first significant evolution in television technology since color television in the 1950s.

Digital Video Broadcasting (DVB) uses coded orthogonal frequency-division multiplexing (OFDM) modulation and supports hierarchical transmission. This standard has been adopted in Europe, Australia and New Zealand.

2) Comparison between analog & digital

DTV has several advantages over analog TV, the most significant being that digital channels take up less bandwidth, and the bandwidth needs are continuously variable, at a corresponding reduction in image quality depending on the level of compression as well as the resolution of the transmitted image. This means that digital broadcasters can provide more digital channels in the same space, provide high-definition television service, or provide other non-television services such as multimedia or interactivity. DTV also permits special services such as multiplexing (more than one program on the same channel), electronic program guides and additional languages (spoken or subtitled). The sale of non-television services may provide an additional revenue source.

Digital and analog signals react to interference differently. For example, common problems with analog television include ghosting of images, noise from weak signals, and many other potential problems which degrade the quality of the image and sound, although the program material may still be watchable. With digital television, the audio and video must be synchronized digitally, so reception of the digital signal must be very nearly complete; otherwise, neither audio nor video will be usable. Short of this complete failure, "blocky" video is seen when the digital signal experiences interference.

Analog TV started off with monophonic sound, and later evolved to stereophonic sound with two independent audio signal channels. DTV will allow up to 5 audio signal channels plus a sub-woofer bass channel, with broadcasts similar in quality to movie theaters and DVDs.

3) Main DTV standards

Several regions of the world are in different stages of adaptation and are implementing different broadcasting standards. Below are the different widely used digital television broadcasting standards (DTB):

- Digital Video Broadcasting (DVB) uses coded orthogonal frequency-division multiplexing (OFDM) modulation and supports hierarchical transmission. This standard has been adopted in Europe, Australia and New Zealand.
    - DVB systems distribute data using a variety of approaches, including:
        - Satellite: DVB-S, DVB-S2 and DVB-SH
            - DVB-SMATV for distribution via SMATV
        - Cable: DVB-C, DVB-C2
        - Terrestrial television: DVB-T, DVB-T2

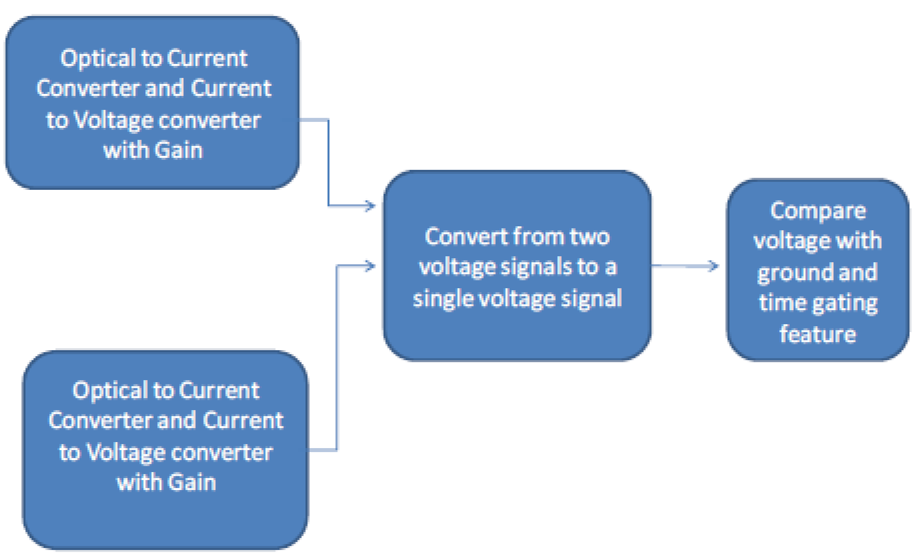
- Digital terrestrial television for handhelds: DVB-H, DVB-SH
- Microwave: using DTT (DVB-MT), the MMDS (DVB-MC), and/or MVDS standards (DVB-MS)

- Advanced Television System Committee (ATSC);
- Integrated Services Digital Broadcasting (ISDB);
- Digital Terrestrial Multimedia Broadcasting (DTMB);
- Digital Multimedia Broadcasting (DMB).

4) DTV signal's properties correlated with standard specifications

COURSE 6 - Video sensors

Optical to Electrical Converter Block Diagram



What is a sensor?

Digital SLR cameras capture their images on a silicon semiconductor referred to as a digital sensor. This sensor is composed of an array of photosensitive diodes called photosites that capture photons (subatomic light particles) and converts them to electrons, much like solar panels convert light to energy. This build up of electrons in each photosite is converted to a voltage which in turn is converted to digital data as a picture element or 'pixel'. These pixels are then relayed in consecutive order and stored as an image on the camera's memory as a file. These files can then be viewed on the camera in the LCD screen, or uploaded to a computer where they can also be viewed or manipulated with imaging software.

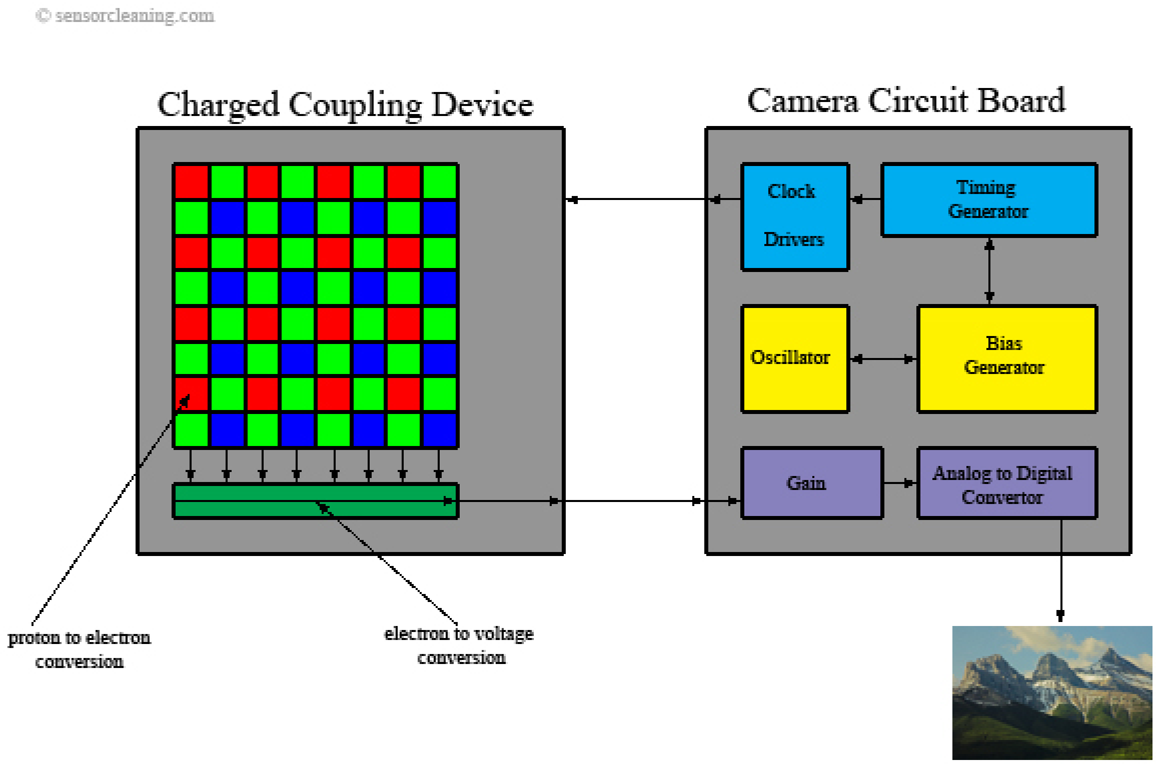
Sensor types

In the digital camera world there are basically two types of digital sensors: CCD and CMOS, with some new and exciting technologies on the horizon.

Charged Coupling Device (CCD) sensors were invented in 1969. Today CCD sensors are used in a multitude of devices from scanners to telescopes.

CCD sensors derive their name from how the charge is read after an image is captured. Utilizing a special manufacturing process, the sensor is able to transport the built up charge across itself without compromising the image quality. The first row of the array is read into an output register, which in turn is fed into an amplifier and an analog to digital converter. After the first row has been read, it is dumped from the read out register and the next row of the array is read into the register. The charges of each row are 'coupled' so as each row moves down and out, the successive rows follow in turn. The digital data is then stored as a file that can be viewed and manipulated.

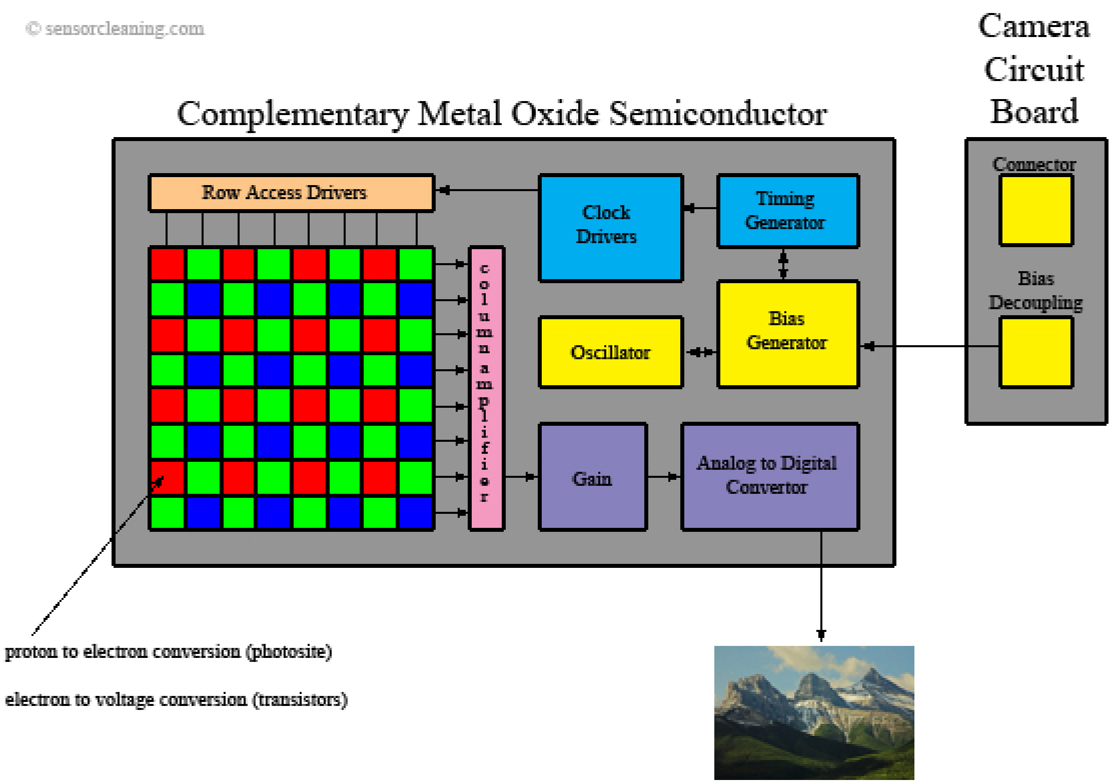
CCDs have been used for many imaging applications because the have relatively low noise compared to other sensors. Although the CCD achieves nearly 100% fill factor (the percentage of of active area to total pixel area) it has several drawbacks. Because CCDs use a charge transfer bucket brigade, 100% charge transfer efficiency is from one pixel to another is imperative during readout since charge loss accumulates across the readout path. This can be achieved but requires slow frame rates. Additionally, because the whole chip is essentially a capacitor, a high capacitive driving system is needed to perform integration and readout which also prohibits high frame rates. Fabrication also presents a drawback of the CCDs. CCD fabrication requires a specialized fabrication process and has not been successfully integrated with other CMOS electronics despite decades of attempts.



Until recently, CCD sensors have been predominant in digital SLR's because of their high quality/ low noise images and their maturity, having been produced for over thirty years. But new manufacturing methods of an old technology have led to inroads for it to surpass CCD sensors.

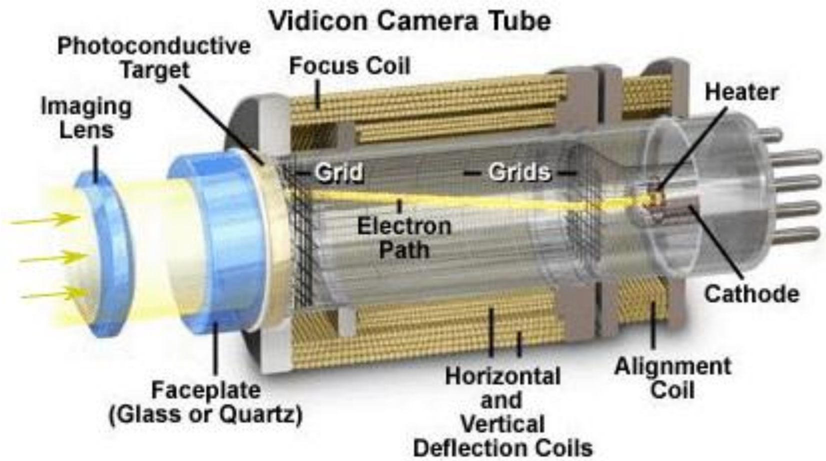
CMOS (Complimentary Metal Oxide Semiconductor) sensors or Active Pixel sensors were first discussed in length in 1992. 1993 through to 1995, JPL developed much of the technology which would be implemented into the CMOS sensor.

CMOS sensors derive their name from the way they are manufactured. They are cut from a CMOS wafer which is cheaper to produce than a CCD wafer, provides less power consumption, and also allow for more involved circuitry alongside of the photosite array. Each photosite in the CMOS sensor has three or more transistors which has its benefits and its drawbacks. The transistors allow for processing to be done right at the photosite, and each pixel/photosite can be accessed independently. Because the transistors occupy space on the array, some of the incoming light hits the transistors and not the photosites, which leads to picture noise. CMOS sensors also function at a very low gain which may contribute to noise.



Basic parameters for CMOS and CCD sensors:

- fill factor
- Pixel well capacity
- Quantum efciency
- dynamic range
- bit depth
- photon shot noise
- readout noise
- reset noise
- dark current noise
- power supply noise
- electro-magnetically coupled radiation



A "Vidicon Camera Tube" or "Pick up Tube" was used as a replacement of CCD's (Charged coupled device) for converting an optical image into an electrical signal. It has an ability to store intensity information over the entire photosensitive area simultaneously on photoconductive surface.

Cathode ray tube: It is a vacuum tube through which the beam of electrons are focused. Electrons are nothing but cathode rays.

Color tubes use three different phosphors which emit red, green, and blue light respectively. They are packed together in stripes (as inaperture grille designs) or clusters called "triads" (as in shadow mask CRTs).[18] Color CRTs have three electron guns, one for each primary color, arranged either in a straight line or in an equilateral triangular configuration (the guns are usually constructed as a single unit). (The triangular configuration is often called "delta-gun", based on its relation to the shape of the Greek letter delta.) A grille or mask absorbs the electrons that would otherwise hit the wrong phosphor.[19] A shadow mask tube uses a metal plate with tiny holes, placed so that the electron beam only illuminates the correct phosphors on the face of the tube;[18] the holes are tapered so that the electrons that strike the inside of any hole will be reflected back, if they are not absorbed (e.g. due to local charge accumulation), instead of bouncing through the hole to strike a random (wrong) spot on the screen. Another type of color CRT uses an aperture grilleof tensioned vertical wires to achieve the same result.

Photoconductive Target: It is a thin layer deposited on a transperent metal film. The target which is placed on the inner surface of the faceplate consists of three layers:

1) The transperent target electrode deposited directly onto the faceplate.
2) A dielectric layer.
3) A photoelectric layer facing on the electron gun that is located toward the back of the camera tube.

[to be continued]

Course 7

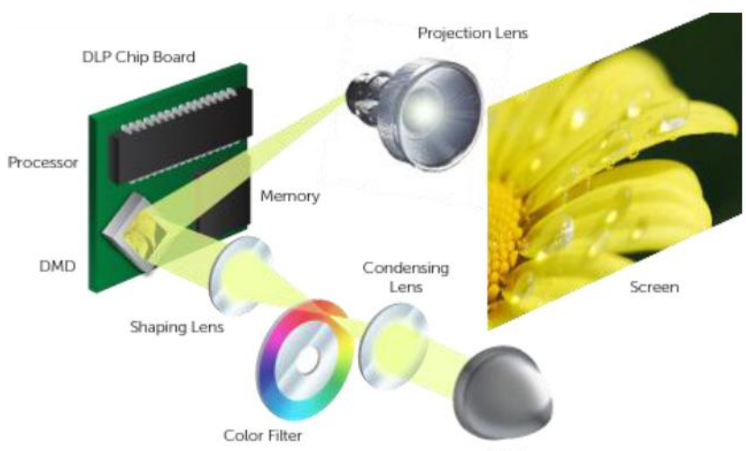
Electro-optical Conversion Process

At the heart of the module that converts RF signals to light is a laser diode. The basic principle is direct modulation of the incoming RF signal onto the output of the laser diode

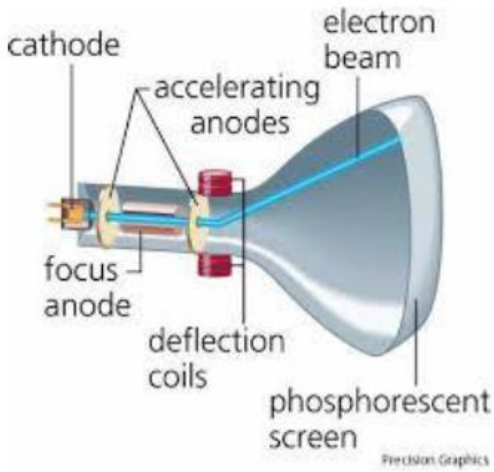
The RF input signal directly modulates the laser diode bias current about the optimal DC working point, sometimes referred to as the quiescent point, which is typically 40mA. Modulation gains range from 0.02 to 0.2mW/mA and a monitor photodiode maintains the stability of the fixed operating point of the laser.

Main technologies available

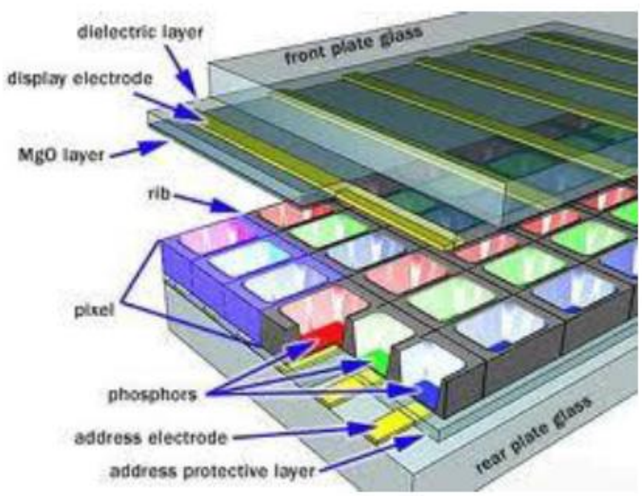
1.The video projector

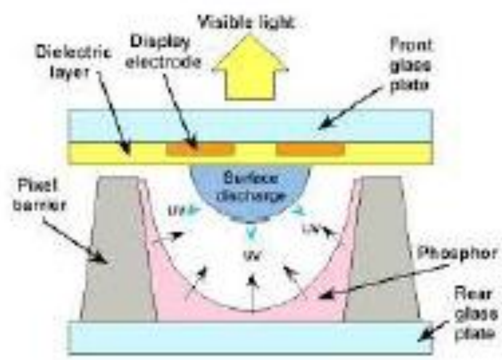


2.Cathodic Ray Tube

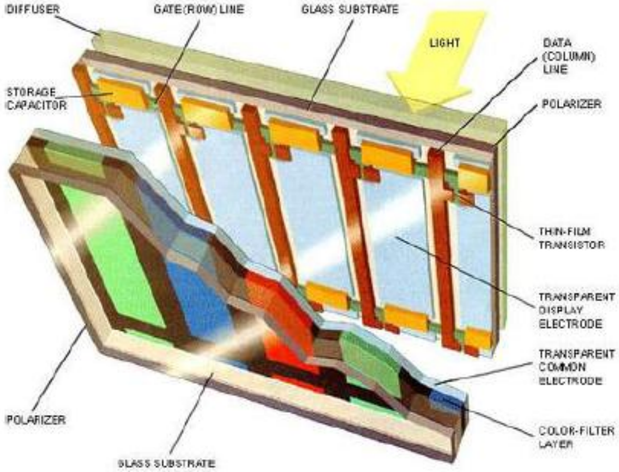


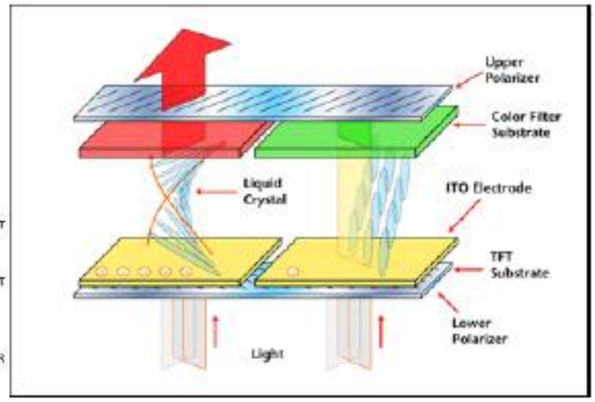
3. PLASMA screen



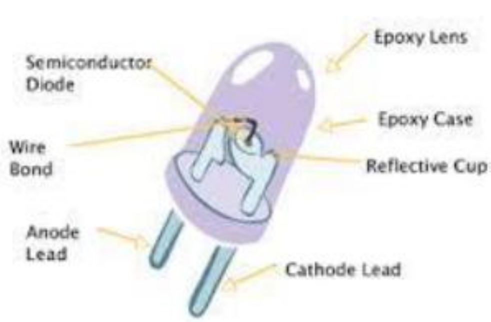


4. LCD screen





5. LED screen



LCD screens

The particular type of liquid used in displays is known as nematic, for its elongated rod-shaped molecules. These crystals have the property that when placed into a container with the surfaces slightly grooved, the nematic molecules will naturally align themselves with the pattern of the grooves. If an electric field is applied, the molecular alignment will shift to be more or less parallel to the field

The alignment of the crystal molecules is capable of rotating the light polarization.

Pixels are addressed by:

-Passive-matrix LCD architectures-slow and have a tendency to smear moving object, still widely used because it has low costs

-Active-matrix LCD architectures-done by adding transistor switches=>response speed and contrast ratio are greatly improved, but adding the transistors increases the cost(almost doubles)

LCD can be produced in transmission mode, or they can be made I a reflective mode by placing a mirror below the display. Ambient illumination passing through the display is reflected by the mirror.

Advantage: no power is required for illumination

Color in LCDs requires separate pixels for red green and blue, with color filters to produce the color. Thus , color LCDs have 3 times the pixels in the horizontal direction.

The angle of view of LCDs is limited because of their use of polarized light. Most LCDs are usable only over a field of view of about 30 deg,which is adequate for single-user application, but not for multiple users

LCD panels are difficult to fabricate in larger sizes

PLASMA

Emissive display that uses the radiation produced by an electronic discharge in a gas, such as xenon. In a color display, the plasma emits ultra violet rays that excite normal color phosphors.

The structure of a one type of color PDP is a sandwich of 2 glass plates that contain a mixture of gasses that are excited locally by electric fields. When sufficient voltage is applied to a pair of electrodes, a discharge occurs at their intersection. Ultraviolet energy is emitted which excited the color phosphor and creates a spotlight at the color of the phosphor. The PDP has a digital input and performs its own \DAC.

There are 2 general ways to limit the discharge current—the DC method where an external resistor is used, and the AC method where an internal capacitor is used(the later approach is more popular).

Some of the design considerations of PDP are :

They are not suited for interlaced scanning because there is too much flicker.

` The angle of view is very wide for PDP making them suitable for almost any viewing situation.

Present PDPs have not achieved luminous efficiency but it is expected to improve .

They require more operating power than even CRTs, of the equivalent brightness.

CURS 8

8.1. Understand basic components of the TV signal spectrum (nu sunt sigur ca asta trebuie)

Pentru a putea determina spectrul semnalului video trebuie să vedem care este frecvenţa maximă respectiv frecvenţa sa minimă a semnalului.
Frecvenţa video minima: cea mai joasă frecvenţă a unei componente este cea care corespunde frecvenţei de explorare pe verticală f_V=50 Hz pentru sistemul 625/50 (sau 60Hz pentru 525/60). Această frecvenţă este considerată ca **frecvenţă limită inferioară** sau frecvenţă video minimă.
Pentru determinarea frecvenţei video maxime trebuie luate în considerare cele mai mici detalii ce pot fi reproduse în sistemul de televiziune. Acestea au dimensiuni egale cu elementul de explorare, adică cu secţiunea fasciculului de electroni. Frecvenţa video maximă corespunde numărului maxim de perioade descrise de semnal într-o secundă. Pentru a găsi această valoare, se determină numărul de puncte pe care le conţine imaginea, apoi la câte puncte corespunde o perioadă completă a semnalului şi, în sfârşit, câte imagini se transmit într-o secundă. Pe scurt se poate scrie formula:

f max=Nr.max puncte/imagine * nr. Imagini/sec * ½= N* fk* ½

Dacă considerăm un relief de potenţial corespunzător unei imagini formate din elemente albe şi negre (tabla de şah) cu dimensiunea elementului de explorare, la două elemente alăturate le va corespunde o perioadă, de aici şi înmulţirea cu factorul ½ din expresia ecuaţiei de mai sus. Considerând o imagine care are un număr n de linii pe verticală şi formată din pătrate foarte mici negre şi albe, numărul de elemente de imagine pe orizontală va fi H/V*n. În consecinţă, numărul maxim de puncte de pe imagine va fi:
N=H/V*n*n
În cazul standardului 625/50, unde n=625 linii, H/V=4/3 şi f_k=25 Hz, rezultă deci:

f max=6,5MHz

8.4. List main approaches used in digital video compression

The need for compression in a digital television system is apparent from the fact that the bit rate required to represent an HDTV signal in uncompressed digital form is about 1 Gbits/s and that required to represent a standard-definition television signal is about 200 Mbits/s, while the bit rate that can reliably be transmitted within a standard 6 MHz television channel is about 19 Mbits/s. This implies a need for about a 50:1 or greater compression ratio for HDTV and

10:1 or greater for standard definition.

The ATSC (Advanced Television Systems Committee) Digital Television Standard specifies video compression using a combination of compression techniques. For reasons of compatibility these compression algorithms have been selected to conform to the specifications of MPEG-2, which is a flexible internationally accepted collection of compression algorithms.

Figure 5.2.1 shows the overall flow of signals in the ATSC DTV system. Video signals presented to the system are first digitized (if not already in digital signal form) and sent to the encoder for compression; the compressed data then are transmitted over a communications channel. On being received, the possibly error-corrupted compressed signal is decompressed in the decoder, and reconstructed for display.

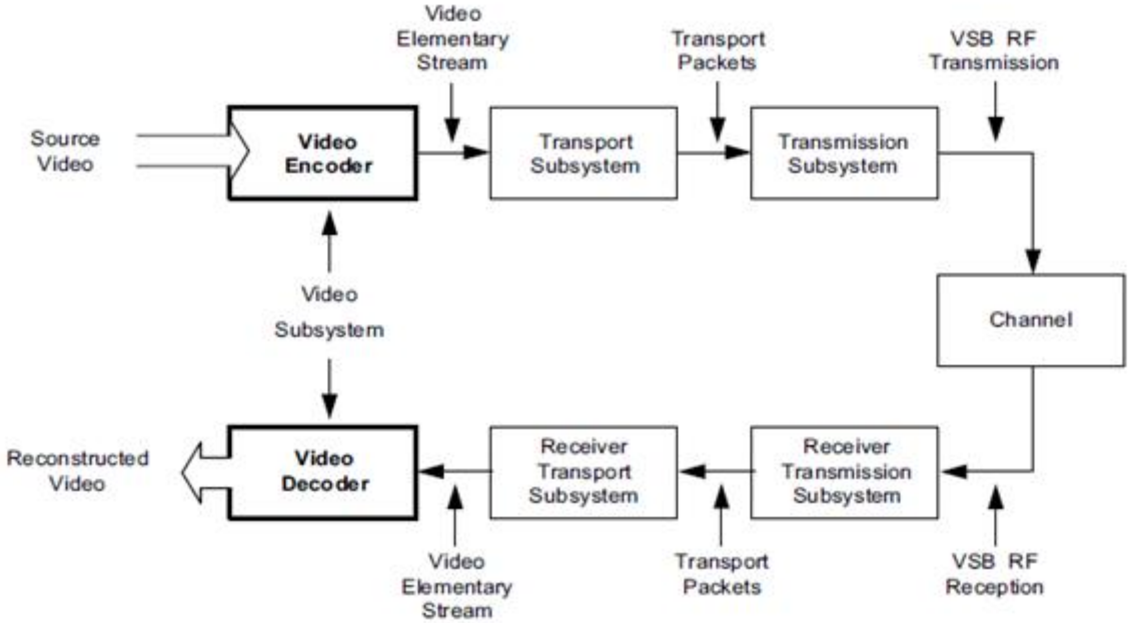


Figure 5.2.1 Video coding in relation to the DTV system.

There could be different sampling rates depending on the video format:
For the 1080-line format, with 1125 total lines per frame and 2200 total samples per line, the sampling frequency is 74.25 MHz for the 30.00 frames per second (fps) frame rate [1]. For the 720-line format, with 750 total lines per frame and 1650 total samples per line, the sampling frequency
is 74.25 MHz for the 60.00 fps frame rate. For the 480-line format using 704 pixels, with 525 total lines per frame and 858 total samples per line, the sampling frequency is 13.5 MHz for the 59.94 Hz field rate. Note that both 59.94 fps and 60.00 fps are acceptable as frame or field rates for the system.

Film mode: material originated at 24 frames per second, such as a film, is typically converted to 30 or 60 frame-per-second video for broadcast [1]. In the case of 30 fps interlaced television, this means that each four frames of film are converted to ten fields, or five frames of video. In the case of 60 fps progressive-scan television, each four frames of film are converted into ten frames of video.

Color Component Separation and Processing:
The input video source to the ATSC DTV video compression system is in the form of RGB components
matrixed into luminance (Y) and chrominance (Cb and Cr) components. The luminance component represents the intensity, or blackand-white picture, while the chrominance components contain color information. While the original RGB components are highly correlated with each other; the resulting Y, Cb, and Cr signals have less correlation and are thus easier to code efficiently.
In the Y, Cb, Cr color space, most of the high frequencies are concentrated in the Y component; the human visual system is less sensitive to high frequencies in the chrominance components than to high frequencies in the luminance component.
Number of lines encoded:
The video coding system requires that the coded picture area has a number of lines that is a multiple of 32 for an interlaced format, and a multiple of 16 for a non-interlaced format. This means that for encoding the 1080-line format, a coder must actually deal with 1088 lines ($1088 = 32 \times 34$).
The extra eight lines are always the last eight lines of the encoded image. These dummy lines do not carry useful information, but add little to the data required for transmission.

Course 9

Q2) Describe the television workflow

The television workflow is based on several stages:

1. Planning - covers the workflow up to the point of shooting and sets out the different conventions and practices that need to be adopted right at the start of the process.

2. RUSHES MANAGEMENT/Capturing - covers the capture and handling of content on location or in a studio up to the point of rushes archive and management.

3. POST-PRODUCTION - covers the workflow from the ‘ingest’ of material into the edit for low or high resolution editing through to the completion of the master.

4. DELIVERY - covers the production of masters for delivery to broadcasters, clients or audiences.

Q3) Enumerate basic equipment used by the TV workflow

- Digital mixer
- TV Monitors with AV input
- Oscilloscope
- Signal generator
- Video cameras
- Computer
- Cables for AV connection and control
- DVD player

Course 10

Q2) Compare main connectors for the analog video signal

1. TRS connector

This type is a common family of connector typically used for analog signals, primarily audio. It is cylindrical in shape, typically with two, three or four contacts. Three-contact versions are known as TRS connectors, where T stands for "tip", R stands for "ring" and S stands for "sleeve".

2. RCA connector

RCA Plugs for composite video (yellow) and stereo audio (white and red)

An RCA connector, sometimes called a phono connector or cinch connector, is a type of electrical connector commonly used to carry audio and video signals. The connectors are also sometimes casually referred to as A/V jacks.

RCA connectors began to replace the older quarter-inch phone (TRS) connectors.

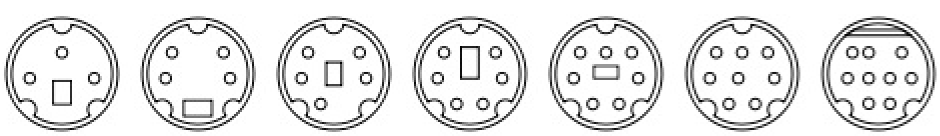
One problem with the RCA connector is that, when connecting the male into the female, the inner 'hot' (signal) connection is made before the 'cold' (ground) connection has been guaranteed; this often produces a loud buzz. Another problem with the RCA connectors is that each signal requires its own plug.

3. BNC connector

Male 50 ohm BNC connector

The BNC connector is a miniature quick connect/disconnect radio frequency connector used for coaxial cable. BNC connectors are made to match the characteristic impedance of cable at either 50 ohms or 75 ohms. They are usually applied for frequencies below 4 GHz and voltages below 500 volts.

4. S-video (mini-DIN) connector



Separate Video, commonly termed S-Video, Super-Video and Y/C, is a signaling standard for standard definition video. By separating the black-and-white and coloring signals, it achieves better image quality than composite video, but has lower color resolution than component video.

The mini-DIN connectors are a family of multi-pin electrical connectors used in a variety of applications

Mini-DIN connectors are 9.5 mm in diameter and come in seven patterns, with the number of pins from three to nine.

5. SCART connector

A male SCART connector (21-pin)

SCART is a 21-pin connector for connecting audio-visual (AV) equipment. The signals carried by SCART include both composite and RGB (with composite synchronization) video, stereo audio input/output and digital signaling.

6. TV aerial plug

TV aerial plug

Antenna plugs are male antenna connectors that fit into female antenna sockets.

Antenna sockets are female antenna connectors that have slots or holes which accept the pins or blades of antenna plugs inserted into them and deliver or receive TV signal to or from the plugs. Sockets are usually designed to reject any plug which is not built to the same standard.

Q3) Compare main connectors and protocols used for the digital video signal

1. FireWire/IEEE 1394

IEEE 1394 is an interface standard for a serial bus for high-speed communications and isochronous real-time data transfer.FireWire can connect up to 63 peripherals in a tree or daisy-chain topology.

It has a maxim bit rate of 3200Mbit/s.

2. DVI/ Digital Visual Interface

The interface is designed to transmit uncompressed digital video and can be configured to support multiple modes such as DVI-D (digital only) , DVI-A (analog only), or DVI-I (digital and analog). The DVI specification is compatible with the VGA interface.

•Maximum bit rate 3.96Gbit/s (single),

•Maximum resolution of the image 2560×1600@60 / 3840×2400@33

3. HDMI / High-Definition Multimedia Interface

HDMI (High-Definition Multimedia Interface) is a proprietary audio/video interface for transferring uncompressed video data and compressed or uncompressed digital audio data from an HDMI-compliant source device, such as a display controller, to a compatible computer monitor, video projector, digital television, or digital audio device.

•Maximum bit rate 18 Gigabit/s (HDMI v2.0)
•Maximum resolution of the image 4096×2160 @ 24

4. Display Port

Interface is primarily used to connect a video source to a display device such as a computer monitor, though it can also be used to carry audio, USB, and other forms of data.

It was designed it to replace VGA, DVI, and FPD-Link. DisplayPort is backward compatible with VGA and DVI.

5. SDI

Serial digital interface (SDI) is a family of digital video interfaces. Additional SDI standards have been introduced to support increasing video resolutions (HD, UHD and beyond), frame rates, stereoscopic (3D) video, and color depth.

SDI standards are used for transmission of uncompressed, unencrypted digital video signals within television facilities. They can also be used for packetized data.

Table of SDI standards:

Standard	Name	Introduced	Bitrates	Example video formats
SMPTE 259M	SD-SDI	1989[2]	270 Mbit/s, 360 Mbit/s, 143 Mbit/s, and 177 Mbit/s	480i, 576i
SMPTE 344M	ED-SDI		540 Mbit/s	480p, 576p
SMPTE 292M	HD-SDI	1998[2]	1.485 Gbit/s, and 1.485/1.001 Gbit/s	720p, 1080i
SMPTE 372M	Dual Link HD-SDI	2002[2]	2.970 Gbit/s, and 2.970/1.001 Gbit/s	1080p
SMPTE 424M	3G-SDI	2006[2]	2.970 Gbit/s, and 2.970/1.001 Gbit/s	1080p
SMPTE ST-2081*	6G UHD-SDI		6 Gbit/s	4Kp30
SMPTE ST-2082*	12G UHD-SDI		12 Gbit/s	4Kp60

COURSE 11 – 3D

Basic components of 3D spatial visual perception

Your visual system takes two-dimensional images projected onto your two retinas and uses these images to reconstruct a three-dimensional perception of the world around you. To perceive the depth in a visual scene, your visual cortex relies on two kinds of information: **the information that your binocular vision** provides by integrating the two slightly different images from your two eyes (stereoscopic vision -a), and **the information that your monocular vision** provides from the image perceived by each eye separately.

a) The 3D spatial visual perception is based on the ability of the human eyes of stereoscopic vision. This means that you have to use both eyes.When you focus on an object, each of your eyes has a slightly different view of the object. Your left eye tends to see a little more of the left side of the object, while your right eye sees a little more of the right side. Your brain automatically uses this information, plus the angle your eyes have to turn to focus on the object, to supply you with an estimate of the distance of the object. **(2 views of the object, your brain makes 1 image of them)**

b) But even with monocular vision, you can receive an impression of depth, because your brain deduces it from several indicators. These indicators are:

- Interposition (Whenever one object hides your view of another object either partly or completely, your brain deduces that the hidden object is farther away, simply because the other object is in front of it), (also, the shadows matter here)
- Object size is another depth indicator (When you don't know the exact size of two objects, but you do know that they are identical, and one of

them projects a smaller image on your retina, you interpret it as being farther away.

- Parallax movement occurs when when you are in motion yourself, and objects that are at different distances from you appear to be moving at different speeds. The farther away these objects are, the smaller the parallax movement.
- Linear perspective is the well known phenomenon in which parallel lines appear to converge as they recede into the distance and appear to meet at a point on the horizon. If there are two objects, and one of them is closer to this point, your visual system assumes that this object is farther away.

o

One important source of ambiguity for the visual system is that the world is three-dimensional, but the images that it projects onto your retina are two-dimensional. Hence differing objects, depending on their distance and orientation, may occupy the same amount of surface area on your retina. Your brain therefore becomes confused, and tries to use other indicators to clarify the situation. Two such indicators are your own past experience with the object in question and the experience of the human species, which is encoded in your genes.

So => when we want to make a 3D movie, we have to build a complex image, videostream 2 componenets (image for the left and the right eyes too), then to separate them and filter out the undesired components before reaching the eye (the left comp for right eyes and vice versa)

List main technologies used for 3D video rendering

1. **Color space: anaglyph technologies -passive**

The films had two different layers. These layers with two different colours, typically red and blue, were superimposed in a way to counterbalance each other. On screen, the two predominantly red and blue images were projected by using a single projector.

The audience were given 3D glasses with one red lens and the other blue. Since the colored lens of the glasses filter out any color but their own, the left eye, which has a red lens, only sees the red shades and the right eye, which has the blue lens, sees the blue shades. **This enabled the two retinas to form two different images and hence the optical illusion of depth was created**.

However, the colour filtering by the lenses distorted the final colour and many in the audience watching a 3-D film complained of headaches and nausea. The picture quality was also low and also monochrome pictures could be created this way.

2. Polarisation -passive

A polarised light wave vibrates on only one plane. The light produced by the sun is unpolarised, meaning it is made up of light waves vibrating on many different planes. It can however be transformed into polarised light using a polarising filter. A polarising filter has tiny parallel lines etched into it, this means it will only let light vibrating on a particular plane through.

As with old fashioned 3D, the film is **recorded using two camera lenses sat side by side**. **But in the cinema, the two reels of film are projected through different polarised filters**. So images destined for viewers' left eyes are polarised on a horizontal plane, whereas images destined for their right eyes are polarised on a vertical plane.

Cinema goers' glasses use the same polarising filters to separate out the two images again, giving each eye sees a slightly different perspective and fooling the brain.

3. The active glasses technology -active

Active 3D uses battery-operated shutter glasses that do as their name describes: they rapidly shutter open and closed. This, in theory, means the information meant for your left eye is blocked from your right eye by a closed (opaque) shutter. Then the information meant for your right eye is blocked from your left. And repeating this so rapidly that the interruptions do not interfere with the perceived fusion of the two images into a single 3D image. All that's required of the TV is the capability to refresh fast enough so each eye gets at least 60 frames per second.

Modern active shutter 3D systems generally use liquid crystal shutter. Each eye's glass contains a liquid crystal layer which has the property of becoming opaque when voltage is applied, being otherwise transparent. The glasses are controlled by a timing signal that allows the glasses to alternately block one eye, and then the other, in synchronization with the refresh rate of the screen. The timing synchronization to the video equipment may be achieved via a wired signal, or wirelessly by either an infrared radio frequency(Bluetooth) transmitter.

Active 3D can be found on plasma, LCD, LED LCD, and all front and rear projectors for the home.

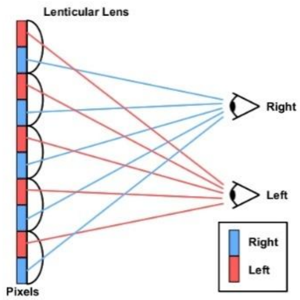
4. Autostereoscopic 3D

The main idea in it is that the viewer does not have to wear a special glass for this.

Stereo parallax: When viewing a scene in real life, an observer sees a different image - with each eye. Movement parallax: When he moves his head, the viewer sees different images. An autostereoscopic 3D display provides a different image to each slot, producing both stereo and movement parallax with a small number of views. (few viewers).

2 main solutions:

Lenticular lenses

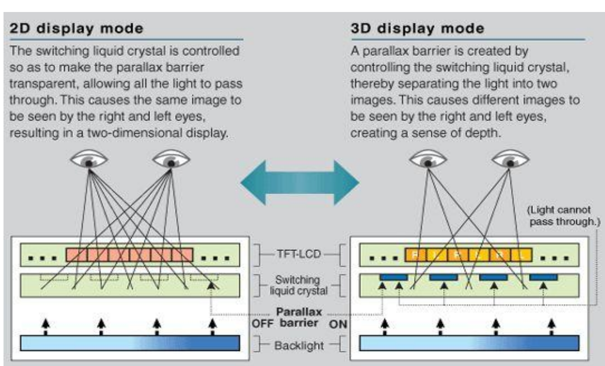


The less popular of the two autostereoscopic models involves the use of lenticules, which are tiny cylindrical plastic lenses. These lenticules are pasted in an array on a transparent sheet, which is then stuck on the display surface of the LCD screen. So when the viewer sees an image, it is magnified by the cylindrical lens.

However, lenticular lenses technology is heavily dependant on where you are sitting. It requires a very specific ‘sweet spot’ for getting the 3D effect, and straying even a bit to either side will make the TV’s images seem distorted.

Parallax barrier

The parallax barrier is a fine grating of liquid crystal placed in front of the screen, with slits in it that correspond to certain columns of pixels of the TFT screen. These positions are carved so as to transmit alternating images to each eye of the viewer, who is again sitting in an optimal ‘sweet spot’. When a slight voltage is applied to the parallax barrier, its slits direct light from each image slightly differently to the left and right eye; again creating an illusion of depth and thus a 3D image in the brain.



Important: these technologies are OK for a limited number of viewers (you are somehow tracked and have to sit in that certain position!)

Compare main approaches used in 3D video rendering

Active vs passive techniques

<u>Advantages of active:</u>

- **Unlike red/cyan color filter (anaglyph) 3D, shutter glasses are colour neutral, enabling 3D viewing in the full colour spectrum.**
- Unlike in a Polarized 3D system, where the screen resolution is halved when the images are combined, the active shutter system retains full resolution by combining the images over time.

Disadvantages of active:

- **Flicker can be noticed except at very high refresh rates,** as each eye is effectively receiving only half of the monitor's actual refresh rate.
- LC shutter glasses are shutting out light half of the time; moreover, **they are slightly dark even when letting light through, because they are polarized**.
- **Frame rate has to be double that of a non-3D, anaglyph, or polarized 3D systems to get an equivalent result.** All equipment in the chain has to be able to process frames at double rate;
- more expensive than anaglyph and polarized 3D glasses.
- **From brand to brand, shutter glasses use different synchronization methods and protocols.**

Course 12:

List main technologies used for 3D video acquisition

3D video acquisition consists of two main parts: capturing and post-processing. Setups with 2-3 cameras are typically used for capturing 3D video. Problems and challenges with multi-camera systems are temporal synchronization, geometrical calibration, and color balance between the
individual cameras. Furthermore depth sensor enhanced camera setups are rarely used, because of the limited spatial resolution and depth range of available sensors. Future development in this field might enable direct capture of depth enhanced 3D video.

For coding and display purposes, the captured sequences need to be converted from the production into the transport format by post-processing. Regarding the video-only formats minor adjustments, like color correction, subsampling or color format conversion, might be necessary, while more complex algorithms, like rectification and depth estimation, are required for the depth-enhanced formats. Various algorithm for estimating depth or disparity maps from 2-view [2], 3-view (within MPEG), and N-view video [3] have been developed, but they are still error-prone and can be highly complex.Further advancement of depth estimation algorithms is expected, as highly accurate depth maps are mandatory for the success of depth-enhanced 3D video formats.

Acquisition can occur from a multitude of methods including 2D images, acquired sensor data and on site sensors.

Acquisition from 2D images

3D data acquisition and object reconstruction can be performed using stereo image pairs. Stereo photogrammetry or photogrammetry based on a block of overlapped images is the primary approach for 3D mapping and object reconstruction using 2D images.

A semi-automatic method for acquiring 3D topologically structured data from 2D aerial stereo images has been presented by Sisi Zlatanova.[9] Software used for 3D data acquisition using 2D images include e.g. Autodesk 123D Catch, ENSAIS Engineering College TIPHON (Traitement d'Image et PHOtogrammétrie Numérique).[10] CyberCity 3D Modeler, ORPHEUS, ...

Acquisition from acquired sensor data[edit]

Semi-Automatic building extraction from LIDAR Data and High-Resolution Images is also a possibility. Ex: The extracted building outlines are then simplified using an orthogonal algorithm to obtain better cartographic quality. Watershed analysis can be conducted to extract the ridgelines of building roofs. The ridgelines as well as slope information are used to classify the buildings per type. The buildings are then reconstructed using three parametric building models (flat, gabled, hipped).[15]

Acquisition from on-site sensors

LIDAR and other terrestrial laser scanning technology[16] offers the fastest, automated way to collect height or distance information. LIDAR or laser for height measurement of buildings is becoming very promising.[17] Using laser scans and images taken from ground level and a bird's-eye perspective, Fruh and Zakhor present an approach to automatically create textured 3D city models. This approach involves registering and merging the detailed facade models with a complementary airborne model. Finally, the two models are merged with different resolutions to obtain a 3D model.

Slide 4: (intelegerea diferentei intre pasiv si activ, cred):