Welcome to the course on:

DIGITAL VIDEO PROCESSING

CS 687.

Text Books:

1. **Handbook of Image and Video processing – Al Bovik (Alan C Bovik),** Academic Press, Second Edition, 2005.

2. **H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia – Iain E.G. Richardson**, Wiley, 2003

3. **Digital Image Sequence Processing, Compression, and Analysis – Todd R. Reed**, CRC Press, 2004.

4. **Digital Video Processing – A. Murat Tekalp,** Prentice Hall, 1995

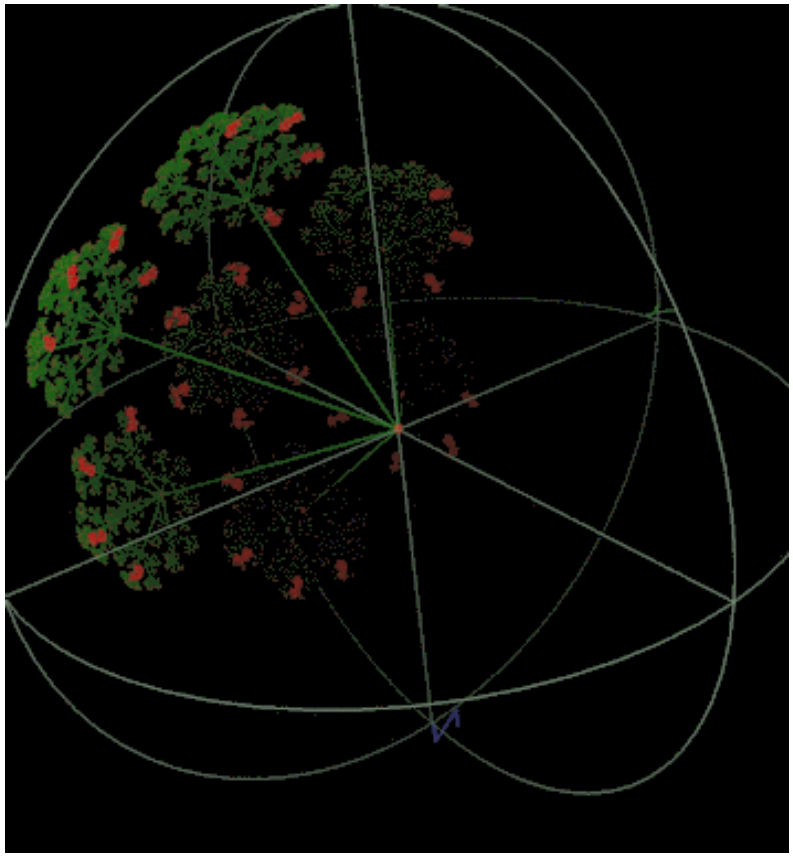
Reference:

IEEE Transactions of Circuits and Systems for Video technology.

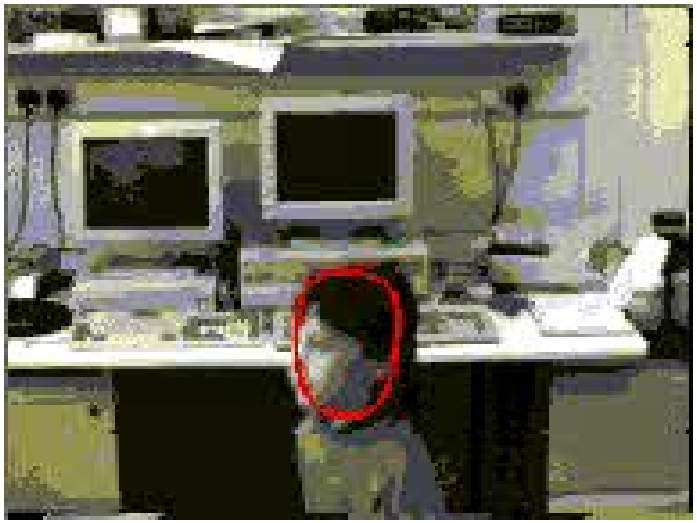
Key course contents:

- **Video Acquisition and Representation.**
- **Motion analysis.**
- **Video Object Tracking and segmentation**
- **Video Filtering**
- **Video coding, representation**
- **Content based Video retrieval**
- **Video based Rendering**

Real-world
VIDEO



Synthetic
Animation



Processed Video

VIDEO-BASED RENDERING

Augmented Reality Video

Processed Video

Original Video


0026

Digital Video processing tasks for various applications

Frame-Rate Conversion - An algorithm for motion-compensated interpolation. Important frame-rate conversions include 10 f/s to 20 f/s, 24 f/s to 30 f/s (film to TV), and 50 f/s to 60 f/s (PAL to NTSC).

Super-resolution - An algorithm that uses a number of low-resolution video frames to create a high resolution still frame. Also an algorithm can take low-resolution video and create enhanced-resolution video.

Mosaicing - An algorithm that takes a number of low resolution frames of a scene to generate a single large panoramic of the scene.

Deinterlacing - An algorithm for interlaced to progressive conversion. Possible applications include display of interlaced TV on a progressive computer monitor, generation of a high quality frame from an interlaced video of a still scene (e.g. for printing a picture from TV), and standards-conversion such as 480-line interlaced to 480-line progressive or 1080-line interlaced to 720-line progressive (important for Digital TV).

Video Stabilization - Handheld video cameras are often afflicted by unintentional camera movement (jerks). Stabilization process identifies this unintentional movement and compensates for it.

Segmentation - An algorithm to perform foreground/background segmentation of a scene. For example, separate the speaker in a video conference or news show. e.g. to identify the number of people in a scene or in a room.

Restoration and Noise Reduction - An algorithm to deblur a (blurred or out of focus) video signal and reduce other artifacts (e.g. motion picture film is often afflicted by scratches and salt-and-pepper noise).

Motion Estimation - Different algorithms exist for performing motion estimation. Algorithms may include different methods for gradient-based and/or block-based estimation, single-layer versus multi-layer estimation. Comparison criteria could include closeness to the true motion in the scene, performance when used for compression and complexity.

Video Watermarking - An algorithm to place an *invisible* watermark in a video, where the existence of the watermark can be identified by an associated watermark detection algorithm. In addition to being invisible and relatively easy to detect, the watermark should be difficult to remove. For example, adding noise or encoding/decoding the video should not destroy the watermark.

Object-Tracking - An algorithm to track an object in a video scene. Examples including tracking a soccer ball in a soccer game, cars driving down a street, or people moving in a room. The algorithm may be fully automatic, or semi-automatic where the user initializes the algorithm by telling it where the object(s) is (are) in the first frame.

Content-Based Representation and retrieval - An algorithm to identify and extract key-frames from a video sequence so that those frames may be used, for example, to provide a summary of the video or to enable efficient searching of the video.
Also represent motion trajectory, texture, color etc.

Video-based rendering

All video storage formats (encoding) are PCM based.

- **CCIR 601 used for broadcast stations**

- **MPEG-4 good for online distribution of large videos and video recorded to flash memory**

- **MPEG-2 used for DVDs and Super-VCDs**

- **MPEG-1 used for video CDs**

- **H.261**

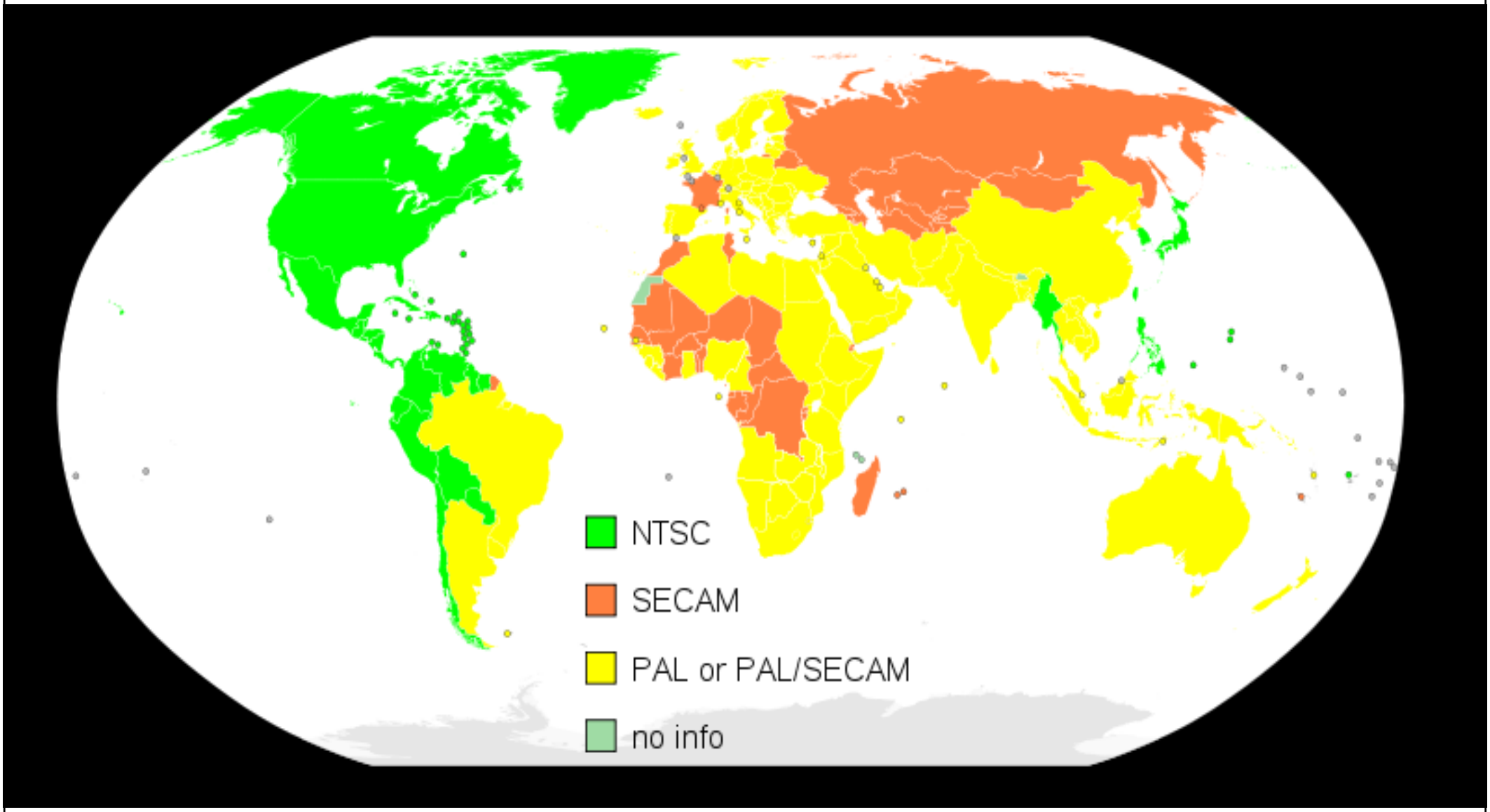
- **H.263**

- **H.264 also known as MPEG-4 Part 10, or as AVC (advanced Video Coding)**

- **Theora standardized but still in development. used for video over the internet (very low bitrate).**

Discs:

- **Video CD (VCD)**
- **DVD**
- **HD DVD**
- **Blu-ray Disc**
- **Sony ProDATA**


NTSC
SECAM
PAL or PAL/SECAM
no info

Editing sequence, of video clips:

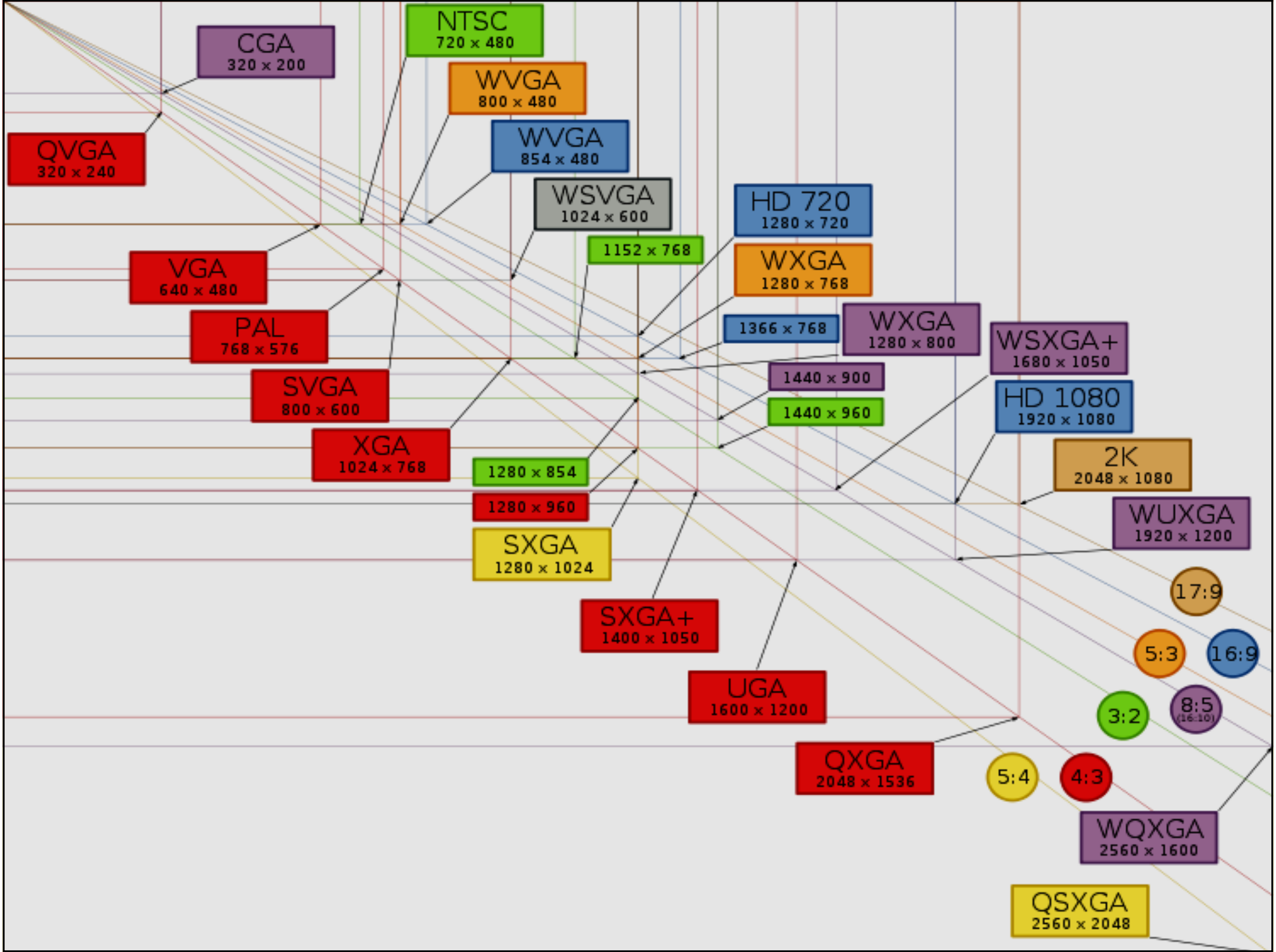
Avid's software and hardware is almost synonymous with the professional NLE market, but Apple's Final Cut Pro, Adobe Premiere, Sony Vegas and similar programs are also popular.

Many interfaces have been designed specifically to handle the requirements of uncompressed digital video (at roughly 400 Mbit/s):

- **Serial Digital Interface**
- **FireWire**
- **High-Definition Multimedia Interface**
- **Digital Visual Interface**
- **Unified Display Interface**
- **DisplayPort**
- **USB**

The interface for carrying MPEG-Transport compressed video: DVB-ASI (*Digital Video Broadcast Asynchronous Serial Interface*).

Compressed video is also carried using UDP-IP over Ethernet.
Two approaches exist for this:
- **Using RTP as a wrapper for video packets;**
- **1-7 MPEG Transport Packets are placed directly in the UDP packet**


CGA
320 × 200
NTSC
720 × 480
WVGA
800 × 480
QVGA
320 × 240
WVGA
854 × 480
WSVGA
1024 × 600
HD 720
1280 × 720
1152 × 768
WXGA
1280 × 768
VGA
640 × 480
1366 × 768
WXGA
1280 × 800
WSXGA+
1680 × 1050
PAL
768 × 576
1440 × 900
HD 1080
1920 × 1080
SVGA
800 × 600
1440 × 960
XGA
1024 × 768
2K
2048 × 1080
1280 × 854
1280 × 960
WUXGA
1920 × 1200
SXGA
1280 × 1024
17:9
SXGA+
1400 × 1050
5:3
16:9
UGA
1600 × 1200
8:5
(16:10)
3:2
QXGA
2048 × 1536
5:4
4:3
WQXGA
2560 × 1600
QSXGA
2560 × 2048

Streaming Media storage size (in megabytes)
= length (in seconds) * bit rate (in kbit/s) / (8 * 1024)

Real world example:

One hour of video encoded at 300 kbit/s (typical broadband video during 2005 and it's usually encoded in a 320×240 pixels window size) will be:

(3,600 s * 300,000 bit/s) / (8*1024*1024)
= 128 MB of storage.

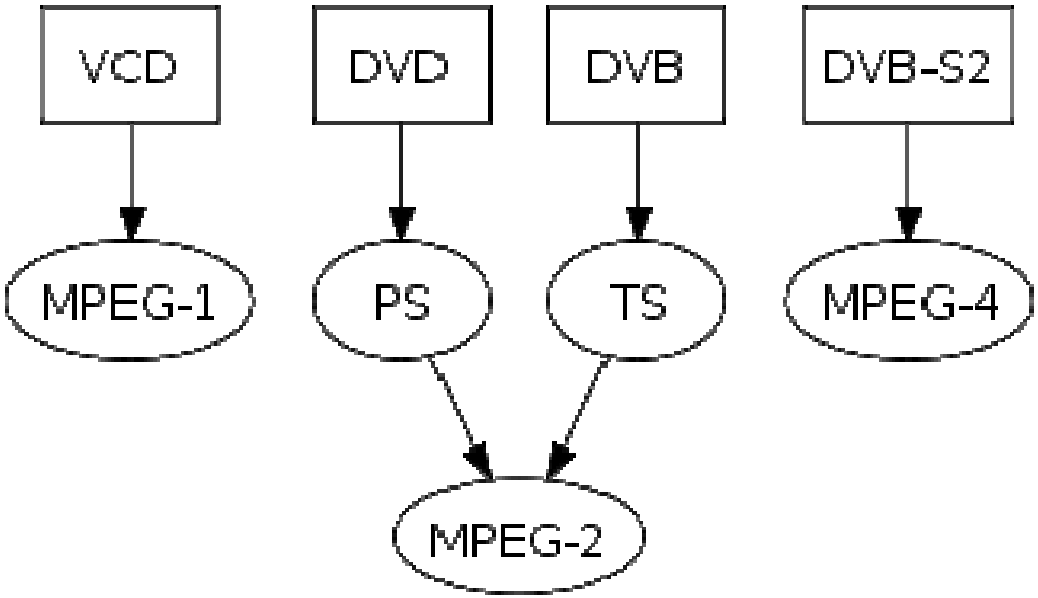
If this stream is viewed by 1,000 people at the same time using a Unicast protocol, you would need:

300 kbit/s * 1,000 = 300,000 kbit/s = 300 Mbit/s of bandwidth
Equivalent to around 125 GiB per hour.

Using a Multicast protocol the server sends out only a single stream that is common to all users.
That stream would only use 300 kbit/s of serving bandwidth.

Examples of Streaming Media systems (usually compressed):

- Flash media server
- Winamp and W-media
- Quicktime
- Peercast
- Freecast etc.



DVB – S2: Digital Video Broadcasting - Satellite - Second Generation;

DVB-SH: *Digital Video Broadcasting - Satellite services to Handhelds*;

DVB-T2: **Digital Video Broadcasting – Second Generation Terrestrial**;

TS is the format used to broadcast high-definition TV;

SONY – PS2, PS3;

List of digital video broadcast standards

ATSC family (*North America*)

ATSC (terrestrial/cable)
ATSC-M/H (mobile/handheld)

ISDB family (*Japan/Brazil*)

ISDB-S (satellite)
ISDB-T (terrestrial)
1seg (handheld)

ISDB-C (cable)
SBTVD (Brazil)

DVB family (*Europe*)

DVB-S (satellite)
DVB-S2

DVB-T (terrestrial)
DVB-T2

DVB-C (cable)
DVB-C2

DVB-H (handheld)
DVB-SH (satellite)

The *MPEG standards* consist of different Parts. Each part covers a certain aspect of the whole specification. The standards also specifies Profiles and Levels. Profiles are intended to define a set of tools that are available, and Levels define the range of appropriate values for the properties associated with them. MPEG has standardized the following compression formats and ancillary standards:

MPEG-1: is the first compression standard for audio and video. It was basically designed to allow moving pictures and sound to be encoded into the bitrate of a Compact Disc. To meet the low bit requirement, MPEG-1 downsamples the images, as well as using picture rates of only 24-30 Hz, resulting in a moderate quality. It includes the popular Layer 3 (MP3) audio compression format.

MPEG-2: Transport, video and audio standards for broadcast-quality television. MPEG-2 standard was considerably broader in scope and of wider appeal--supporting interlacing and high definition. MPEG-2 is considered important because it has been chosen as the compression scheme for over-the-air digital television ATSC, DVB and ISDB, digital satellite TV services like Dish Network, digital cable television signals, SVCD, and DVD.

MPEG-3: Developments in standardizing scalable and multi-resolution compression which would have become MPEG-3 were ready by the time MPEG-2 was to be standardized; hence, these were incorporated into MPEG-2 and as a result there is no MPEG-3 standard. MPEG-3 is not to be confused with MP3, which is MPEG-1 Audio Layer 3.

MPEG-4: MPEG-4 uses further coding tools with additional complexity to achieve higher compression factors than MPEG-2. In addition to more efficient coding of video, MPEG-4 moves closer to computer graphics applications. In more complex profiles, the MPEG-4 decoder effectively becomes a rendering processor and the compressed bitstream describes three-dimensional shapes and surface texture. MPEG-4 also provides Intellectual Property Management and Protection (IPMP) which provides the facility to use proprietary technologies to manage and protect content like digital rights management.

Several new higher-efficiency video standards (newer than MPEG-2 Video) are included (an alternative to MPEG-2 Video), notably:

MPEG-4 Part 2 (or Advanced Simple Profile) and
MPEG-4 Part 10 (or Advanced Video Coding or H.264). MPEG-4 Part 10 may be used on HD DVD and Blu-ray discs, along with VC-1 and MPEG-2.

In addition, the following standards, while not sequential advances to the video encoding standard as with MPEG-1 through MPEG-4, are referred to by similar notation:

MPEG-7: A multimedia content description standard.

MPEG-21: MPEG describes this standard as a multimedia framework.