CSE 126/228F Lecture 2: April 3

Media Data Types & Operations

Images

Images are defined as matrices where each element represents a pixel (picture element). Each pixel represents color information for a specific location in the image. There can be only one color value per pixel. We know from optics that there are three primary colors that represent all colors: Red, Green, and Blue (RGB). A True Color image uses 24 bits to define each pixel value, 8 bits per primary color. True color images can therefore contain up to 2^{24} or ~16.7 million colors. The human eye can not distinguish more colors than this.

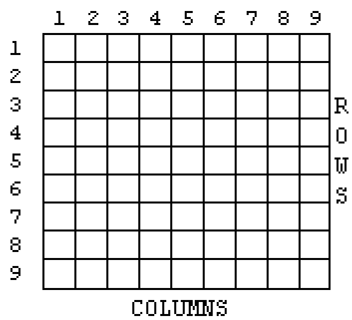


Image Formats

While RGB allows for a straightforward color representation, the human eye is more sensitive to intensity variations than color variations. Using this information, we can use more bits of the pixel to represent intensity information and less to represent color. Examples of this idea are:

- HSB: Hue, Saturation, and Brightness. In this format, Hue is the dominant color of the image. Saturation is a measure of how dominant that color is, and Brightness is intensity of the image.
- YUV: Y represents Intensity, U and V components represent chrominance (color). U is defined as Intensity (Y) - Blue (B), and V is defined as Intensity (Y) - Red (R). This is the format used by JPEG and MPEG.

Displaying Colors

To display colors, a computer must have a table to convert a color value into a signal sent to a display device. If a system can display 16.7 million colors, this table would exceed 16MB! There are two schemes commonly used to reduce the size of this table. One way is to simply use the highest-order bits of each color to create the new color. Although effective, the results are crude and require each application to use the same set of colors. The second method, a Color Look Up Table (CLUT), allows each application to define a table of colors. Then, each color to be displayed will map to the closest color in the CLUT. This allows an application to define a set of colors that best suits its needs and can provide much more accurate translations.

Pixel-Based Image Operations

- Editing
- Point Operations: Thresholding, color correction
- Filtering
- Compositioning: Combining two images into a single image.

Other Image Operations

- Displace
- Rotate
- Mirroring
- Scaling
- Skewing
- Warping: Take a surface and fit the image to it.

Graphics

Graphics tries to represent and manipulate images in a more intuitive manner. It uses models to represent image structures rather than storing the raw pixel data. An example of a model is an equation to represent a light source, which can be mathematically translated into changes in pixel intensity values such as reflections and shadows.

Model Types

- Geometric: 2D, 3D
- Solid Models: e.g.: a model to represent a sphere, cone, etc.
- Physical Models: e.g.: model object collision mathematically
- Empirical Models: use statistical observations to represent data. e.g.: Weather Forecasts

Models are converted into images by a process called *Rendering*.

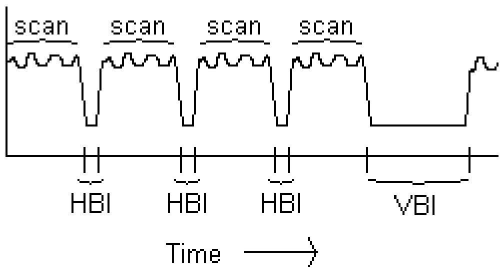
Graphics Operations

- Shading
- Light
- Texture Mapping

Animations

Most animations are still hand-drawn. They usually have about 15 drawings per second. For a one-hour animation, this is 54,000 drawings! The use of computer models can allow a designer to use a computer to animate an entire sequence or it can interpolate between images, reducing the number of actual drawings required to around 1 per second, or 3,600 drawings per hour.

Video



Video is still mostly used in analog form. The above diagram shows an analog video signal. A Video is drawn from left to right, top to bottom. Each scan is a single horizontal pass across the screen. This is followed by a Horizontal Blanking Interval, where the electron gun moves to the beginning of the next scan line. After every scan line has been drawn, a Vertical Blanking Interval allows the electron gun to move from the lower-right corner of the screen back to the upper-left, and the process begins again.

Analog Video Formats

- NTSC (National Television Systems Committee) has 525 lines of resolution with a 30Hz frequency rate. It is the most widely used standard.
- SECAM (Sequential Coleur Avec Memoire) is a frequency-modulated signal that has 625 lines of resolution and a 25Hz refresh rate. It is used in France and Eastern Europe.
- PAL (Phase Alternating Line) is similar to SECAM and is used in parts of Western Europe.

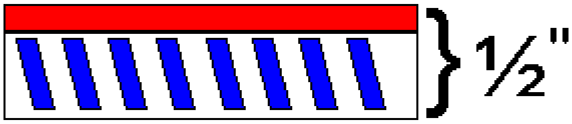
Because human vision is wider than tall, television is rectangular, its width is larger than its height. The ratio of width to height is called the *aspect ratio*. For standard TV, the aspect ratio is 4:3, giving a resolution of 700x525. HDTV has 1000 scan lines and an aspect ratio of 16:9, giving a resolution of 1778x1000.

Refresh Rates

The human eye can distinguish about 1/16 of a second. Despite this, some flicker can be seen even at 30 frames per second. In order to alleviate this problem, *Interlacing* is used. Interlacing draws odd scan lines in the first 1/60 of a second and then draws the even scan lines in the second 1/60 of a second, effectively converting a 30Hz signal to a 60Hz refresh rate while keeping the same bandwidth as the original signal.

Video Tape

VHS is the most common videotape standard. The magnetic patterns on the tape are structured as in the following diagram:



The Red strip is the audio and synchronization information. The blue strips are the video frames. The frames are diagonal because a spinning head creates them as the tape is moving.

Synchronization

An interesting problem is inherent in NTSC video. The advertised frame rate of NTSC video is 30 Hz, however due to a harmonic interference with the color carrier; the frame rate was dropped to 29.97 Hz, a 0.1% decrease in the frame rate. Because synchronization information is represented as hh:mm:ss:ff (hour:minute:second:frame#), this poses a serious synchronization problem. If we assume each frame is 1/30th of a second, then display time will drift away from presentation time. We can, however, correct the problem by doing the following: Drop the first two frame **numbers** (not the actual frames) of every minute not divisible by ten. This means that for a video sequence will look like the following:

```
Orig. Frame#: 0123..0123..0123..0123..0123..0123..0123..0123..0123..0123

Fixed Frame#: 2345..2345..2345..2345..2345..2345..2345..2345..2345..0123
```