Information Technology
Inside and Outside
- David Cyganski & John A. Orr

III. Graphics and Visual Information

5. From the Real World to Images and Video

Hoon-Jae Lee
hjlee@dongseo.ac.kr
http://cg.dongseo.ac.kr/~hjlee

5. From the Real World to Images and Video

- Objectives:
  - the various ways in which an image represents the real world, and the ways in which it is never a perfect representation;
  - how images are formed, optically, photographically, and electronically;
  - how the quality of images is measured and expressed;
  - how images that were never visible in the real world (such as radar, medical ultrasonics imaging, and so forth) may be created by computer;
  - how images are represented in computers by binary numbers;
  - how color information is expressed and stored;
  - how the human eye works, including its color discrimination and stereo vision capability; and
  - how it is possible to represent continuous motion with still images.

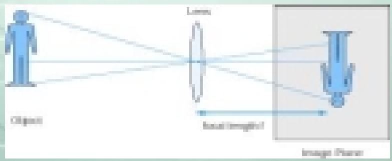
5.1 Introduction

5.1 Introduction
- ``a picture is worth ten thousand words(**百聞 不如一見**)"
  - a visual image conveys **a lot of information** at once.
- An ***image*** is a representation (usually two-dimensional) of objects in the real world

5.2 Images: Information without Words or Numbers
- **Images** play a fundamental role in the **representation**, **storage**, and **transmission of important information** throughout our professional and personal lives.
- In many professions, including **publishing**, **art**, **film making**, **architecture, and medicine**, it is crucial to be able to represent and manipulate information in image form.
- Furthermore, with the development of **multimedia technology** and **virtual reality**, **many other professions** are beginning to explore the power of representing information in visual form.

5.3 Cameras and Image Formation

- The film-based **camera** is over 150 years old.
- The essential components of this system are: the ***object*** *or* ***scene*** to be imaged, the ***lens***, and the ***image recording medium*** (retina of the eye, film, or other device).
- **Figure 5.1:**The operation of an imaging system based upon projection by a lens of a scene onto an imaging plane.

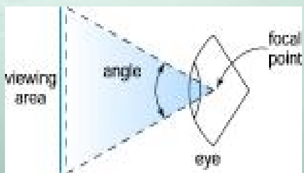


5.3 Cameras and Image Formation(2)

- The image recording medium is usually located in a plane parallel to the lens, known as the ***image plane***
- The image is **inverted**; this is usually of no consequence because the **display device** may easily correct this condition.
- The resulting image represents a ***projection*** from the **three-dimensional object** world to the **two-dimensional image** world.
  - **Two eyes → brains** take these two images and merge them to recreate three-dimensional images in our brain.
- The ***focal length*** specifies the distance from the lens to the image plane
  - 35~50 mm focal length is considered ``**normal**''
  - 28 mm focal length is ``**wide angle**,'' and
  - 135 mm focal length is ``**telephoto**.'
- The process of **reducing the dimensionality of the information** (from three dimensions to two in photography) is referred to as **projection** and is fundamentally a mathematical concept.

5.4 Human Visual Discrimination and Acuity

- The ability of the eye to resolve fine detail : ``**lines per degree of visual arc**''.
  - an image is brought closer to our eye,we can resolve more detail.
  - In bright illumination, the adult, visually impaired human can resolve approximately **60** lines per degree of visual arc.
  - **visual arc** mean the angle covered by the area being viewed at the apparent focal point of the eye, as shown in Figure5.2.
- **Figure 5.2:**Depiction of the geometry of the viewing angle of the eye.



5.4 Human Visual Discrimination and Acuity(2)

- Human can resolve approximately **60 lines per degree of visual arc**.
- The eye will discern this black-white transition as a line. Hence, one line requires **two strings of pixels**.
- The human eye can individually discern **120 pixels per subtended degree of visual arc**, or **60 lines per degree**
- **550 dpi**(dots per inch) in printer would be sufficient to fool any eye into thinking a picture was rendered without pixelization if the paper were held at distances of a foot or greater.
  - This explains the long-term popularity of **600** dots per inch (dpi) laser and ink jet printers.

5.5 Other Types of Image Formation

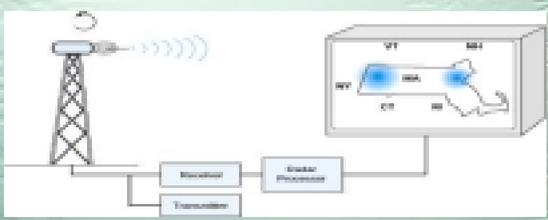
- Lens-based **cameras**
- Other devices - **radar, sonar, X-rays,** and **tomography** (``CAT scans") - differ from traditional cameras in two ways:
  - (1) the **type of energy** used to form the image (instead of visible light, radio,sound waves, X-rays, or radio emissions of nuclei under the influence of a magnetic field are used); and
  - (2) the **geometry of the system** that relates the locations of the objects in the real world (three-dimensional) to the image world (two-dimensional).

5.5 Other Types of Image Formation (2)

- ``**Radar**'‘(``**Radio Detection and Ranging**,'‘) is fundamentally different in several ways from normal photography:
  - The **type of energy** used to form the image (radio waves vs. light waves);
  - The fact that **the illumination must be supplied by the imaging system**, rather than the surrounding ambient conditions. That is, cameras and human eyes operate with visible light; radar, however, must supply its own ``illumination" using radio waves;
  - The **geometry of the image** (based on polar coordinates rather than rectangular coordinates). The image is formed by rays emanating from the center of the image, corresponding to the radar location; and
  - The fact that **the radar site (camera) is located in the image plane** rather than perpendicular to the image plane and some distance away.

5.5 Other Types of Image Formation (3)

- **Figure 5.3:Diagram of a radar system**. The antenna is like a rotating searchlight,sending rays of radio energy at successive angles around the compass. Wherever the searchlight beam strikes an object, energy bounces back to the radar, and is plotted on the image at the appropriate angle and distance from the antenna. The antenna is located at the center of the radar image.



5. 6 Converting Images to Bits

- All image processing is now performed digitally.

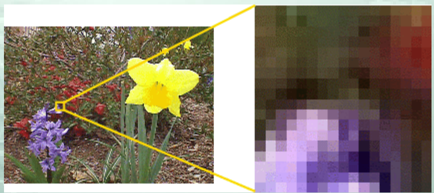
5.6.1 From Continuous Information to a Discrete Representation

- *Continuous* information → *Discrete* information : *analog to digital conversion(A/D conversion)*
  - *Only discrete information can be perfectly represented using binary digits.*
  - *black-and-white photograph*
  - *brightness*
  - *a finite number of small picture elements, or pixels*
  - **Q) 1 frame=512 x 512 images with gray level=256(8-bit),**
    # of pixels ?   512 x 512 = 262,144 = 2^18 pixels/frame
    # of bits/sec in 30 frames/sec? 512 x 512 x 256 x 30
    = 2,013,265,920 = 2^31 bit/sec

5. 6 Converting Images to Bits(2)

5.6.2 Pixels: A Matter of Spatial Resolution

- **Figure :**A picture of a flower bed, represented as a collection of tiny pixels and a close up view of a small section of that flower. The close-up view reveals the pixels which were small enough to not be noticed before.

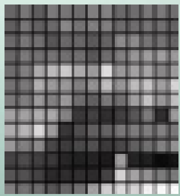


- ***pixelization, sampling, scanning, or spatial quantization*** : The process of breaking continuous image into a grid of pixels

5. 6 Converting Images to Bits(3)

5.6.2 Pixels: A Matter of Spatial Resolution (2)

- The definition of ``**pleasing**'' is driven by the use to which the picture will be put.
  - To the Internet user, the idea of pleasing often centers around the idea of conveying the content with a **minimum of download time.**
  - **requiring as few pixels as make the picture recognizable**
- **Figure 5-4:**A 13 x 13 rectangular grid of pixels with black borders clearly delineating pixel boundaries.



5. 6 Converting Images to Bits(4)

5.6.2 Pixels: A Matter of Spatial Resolution (3)

Figure 5.6:A 256 by 256 pixel image.

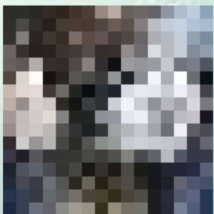
Figure 5.7:A 128 by 128 pixel image.

5. 6 Converting Images to Bits(5)

5.6.2 Pixels: A Matter of Spatial Resolution (3)

Figure 5.8:A 64 by 64 pixel image.

Figure 5.9:A 16 by 16 pixel image.



5. 6 Converting Images to Bits(6)

5. 6.3 Shades of Gray

- **256 brightness levels** (recall that a binary number with 8 bit positions may take on 256 different values, or equivalently that 2 raised to the power 8 equals 256

Figure 5.10:A 6-bit (64 gray levels) image.

Figure 5.12:A 1-bit (black and white only) image.

Figure 5.11:A 3-bit (8 gray levels) image.

5. 6 Converting Images to Bits(7)

5. 6.3 Shades of Gray

- **256 brightness levels** :*quantization*
  - 6-bit word means that a total of $\mathbf{2^6=64}$ possible gray levels
  - codes ranging from black (000000) to white (111111).
  - 256 x 256 ***at the same spatial resolution***, but with 6 bits (for 64 gray levels), **3 bits** (8 gray levels), and **1 bit** (2 gray levels) used to represent and store each pixel.
  - the effect of differing gray-scale resolutions on image quality is clearly discernible, ***but is of a different nature than the effect of differing spatial resolutions*.**
  - 32 gray levels (5 bits) → the image can be stored using 64x64x32=20,480 bits, or 2560 bytes, or 2.5 kB.
  - the image with **256 x 256 pixels, and six bits per pixel** for a total of 64 gray levels, was determined to be ***good enough*** for an image of this size in an application such as this textbook.

5. 6 Converting Images to Bits(8)

5.6.4 Color Representation

- How can we represent ***color images*** in binary form?
- ***We can represent a color with three numbers indicating the amounts of red, green, and blue light*** that combine to produce that color. This system for specifying colors is known as the **RGB system**.
  - For example, **10 units of red, green, and blue** will form **white** of a certain intensity.
  - If we increase this to **20 units of red, green, and blue**, we will still have a **white** light, but it will be **more intense**.
  - The impaired human has **three kinds of cells** in the eye that are sensitive to different ranges of wavelengths of light and are used to distinguish color.
  - The three values for **red, green, and blue** content in an RGB can produce a response by the eye like that of any other color because our eyes can only interpret a color from the three responses or the respective cells.

5. 6 Converting Images to Bits(9)

5.6.4 Color Representation(2)

- **Standard** color televisions use tight clusters of **red,green, and blue** color sources to create the illusion of other colors.
- **To digitize a color image:**
  - First, spatially quantizing the image into **pixels**, as we did for black-and-white imagery.
  - Then, determining the **RGB representation for each pixel**. (determine the amount of red, green, and blue needed to represent the color at the pixel's location.)
  - Finally, **digitizing** these three numbers, to represent each value by a binary number of a predefined length.
  - For example, **3 bits** for each color value → represent **2^3 =8 different intensity levels of red, of green, and of blue**.→ a total storage of ***9 bits* per pixel**--three bits for each of the three colors. → **8 x 8 x 8=512 different possible color combinations**.

5. 6 Converting Images to Bits(10)

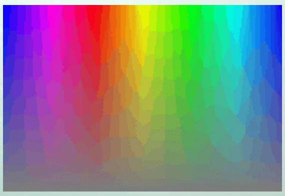
5.6.4 Color Representation(3)

- **HLS (hue, luminance, and saturation → 색조,휘도,채도)**
  - The ***hue*** of a pixel represents where its pure color component falls on a scale that extends across the visible light spectrum, from red to violet.
  - The ***luminance*** of a pixel represents how bright or dark the pixel is.
  - The ***saturation*** represents how ``pure" the color is; that is, how much it is or is not diluted by the addition of white, with 100% indicating no dilution with white.

5. 6 Converting Images to Bits(11)

5.6.4 Color Representation(4)

- **Figure 5.13:**Diagram showing the colors corresponding to the full range of hue (along horizontal axis) and saturation values (0% saturation at the bottom and 100% saturation at the top).



5. 6 Converting Images to Bits(12)

5.6.5 Color Discrimination

- **Saturation** : The average person can discern about **100 saturated colors** from each other.
- **Both luminance and hue** : one can discern about 6,000 variations of color intensity.
- In practice, **8 bits per color or 24 bits per pixel** are general used for full color representation

Figure 5.14: An image with 24 bit color (8 bits each for red, green and blue.)

Figure 5.15: An image with 16 bit color resolution.

Figure 5.16: An image with 8 bit color resolution.

5. 6 Converting Images to Bits(13)

5.6.5 Color Discrimination(2)

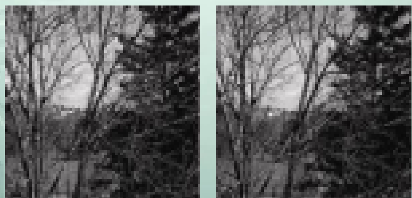
Figure 5.17: An image with 4 bit color resolution.

Figure 5.18: An image with 4 bit color resolution in which a **dithering process** has been used to obtain a better representation on average.

5.7 Binocular Vision and 3D displays

- **Binocular (or stereo, or 3D) vision** provides us with additional visual information referred to as ***depth perception***.
  - the scene look more ``real" to the viewer
  - some specific information as to the location of objects in the front-to-back direction.

Figure 5.19: Example of images produced by **binocular (stereo) vision** with camera separation equivalent to the spacing of the human eyes. Some people can cross their eyes in such a way as to fuse these two pictures in their minds into single stereo scopic scene. The role of stereo displays is to produce this without effort, talent, or eye strain.



5.7 Binocular Vision and 3D displays(2)

- **Figure 5.20:** Stereo scopic shutter glasses are shown being used to view a 3-D display of terrain. The operation of the electronic shutters in this case is synchronized with images shown on the computer screen via an infra-red link (note the small box on top of the computer monitor) between the computer and the glasses.

5.8 From Images to Video

5.8.1 Human Visual persistence

- If you look at a well-lit scene and then close your eyes, you will notice that the image can still be sensed for some time after the eyes close. This is due to the amount of time that the retina retains some of the information with which it has been stimulated. This phenomenon, which places limits on how fast our visual system can react to changes, is known as ***visual persistence*** or ***visual latency***.
  ➔ Take **advantage** of this to develop techniques for digital video systems
  ➔ This period, on average, is about **50 milliseconds, or 1/20 second**.
  ➔ The average human visual system can only take in about **20 different images per second** before they begin to blur together.
- No flashing is evident at above **50 flashes per second**, and perception even for the brightest of lights disappears for rates above **80 flashes per second ➔ 60 Hz electrical system** used in the United States causes electric lights to flicker at a rate of **120 times/sec.**

5.8 From Images to Video(2)

5.8.1 Human Visual persistence

- **The frame rate in the motion picture industry** to provide a more pleasing **24 frames per second** to provide an improved illusion of continuous motion.
  ➔ Again the phenomenon was addressed by having the shutter open and close, this time **twice** for each single motion of the film, producing a **48-flashes-per-second presentation**.
- **Television**, interestingly enough, displays **30 new images per second**, but suffers from the same flash phenomenon if simply presented.
  ➔ The way this is accomplished is that **60 times per second**, *every other line or raster is changed*.

5.8 From Images to Video(3)

5.8.2 Adding Up the Bits

- **One hour's** worth of music stored on a **compact disc** would require storage of over **5 billion bits(608 MB)** of information
- TV : 512 x 512 pixels → 3 bits per color per pixel, for a total of 9 bits per pixel → 60 frames per second → 3600 seconds = **500 billion bits per hour**
- Francis Ford Coppola's *The Godfather*, at over 3 hours,would require nearly **191 GB**--over 191 billion bytes--of memory using this approach.