LBSC 690: Information Technology
Lecture 10
Multimedia and Web Integration

William Webber
CIS, University of Maryland

Spring semester, 2012

Today's lecture

Three parts:

- Digitally representing and compression images, video, and sound (old multimedia)
- Integrated diverse web services into a web site (new multimedia)
- Critically considering benefits of multimedia (for curmudgeons)

Raster versus vector graphics

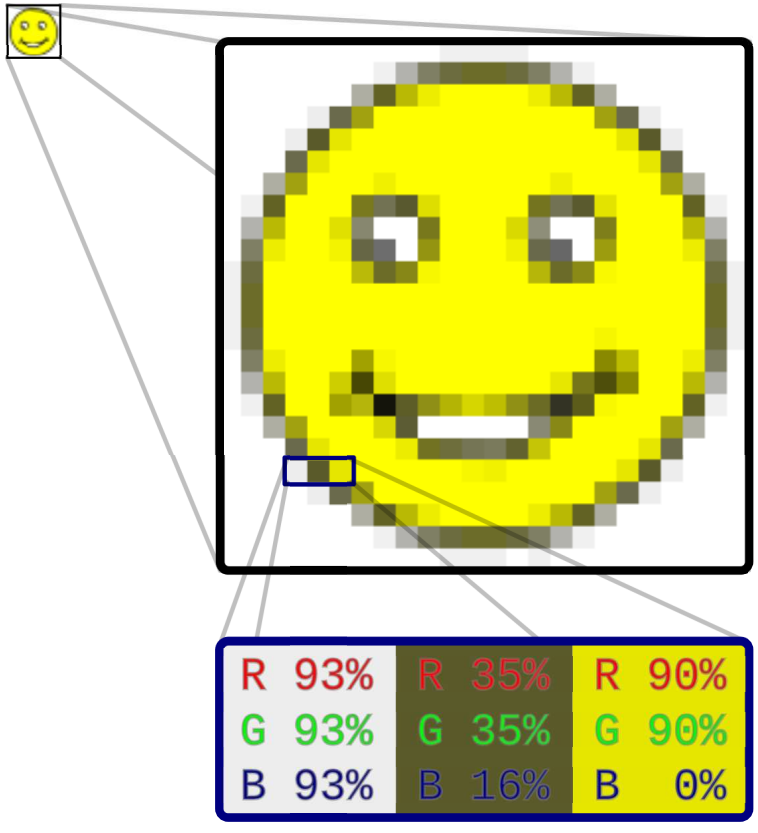
Two basic ways of representing digital graphics:

Raster picture represented as matrix (array) of dots (pixels) (png, jpeg, gif)

Vector picture represented as set of lines and shapes (ps, pdf, svg)

Note: most physical displays are raster graphics, so vector graphics must be converted to raster to display.

Raster graphics



- Divide (square) image into $h \times w$ matrix of pixels:
  - h pixels high
  - w pixels wide
- For each pixel, hold n-bit value
- . . . representing 2^n different colors

[0] Image: Wikipedia

Color depth

Color depth refers to number of bits per pixel (or pits per channel, or number of distinct colors)

Black and white one bit per pixel (1 = white, 0 = black)

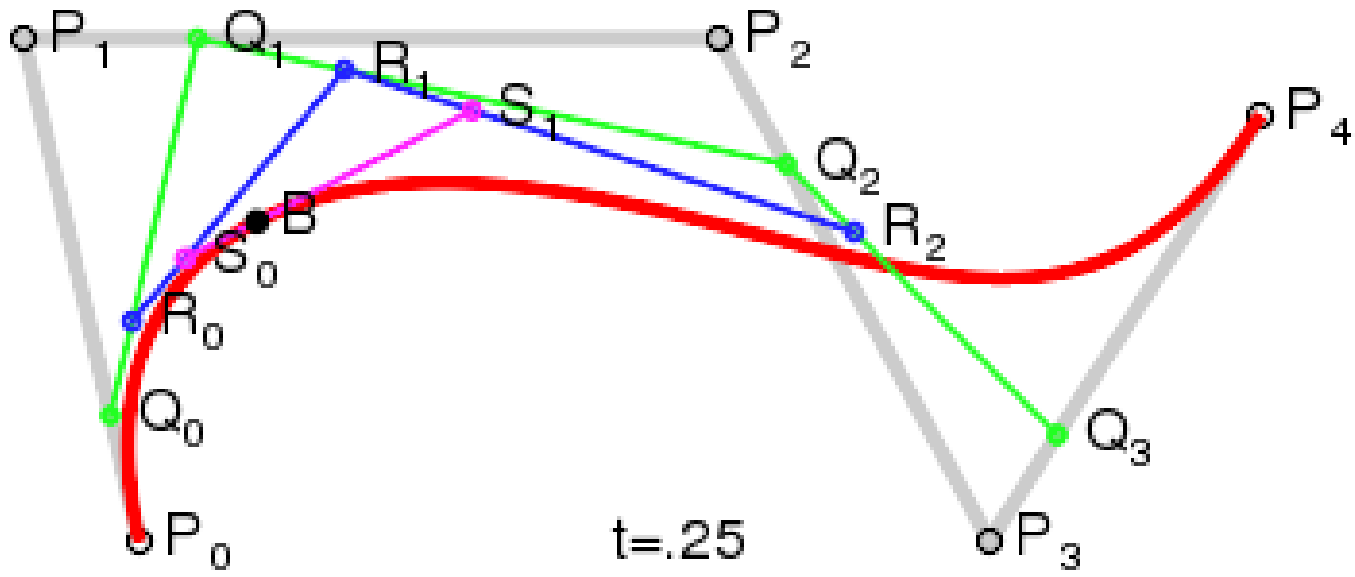
Grayscale single channel, n bits, for 2^n shades of grey

RGB three channels, n bits each, for 2^{3n} colors

- 8 bits per channel or 24 bits overall for 16 million colors called *true color*

RGBA add alpha channel: degree of transparency/opaqueness

Vector graphics



- Drawing represented as series of lines, curves, filled shapes
- Each shape has an “equation” behind it
- For 2-d graphics, a fundamental equation is the Bezier curve
- For 3-d graphics, a fundamental equation is polygon

[0]Image: Wikipedia

Vector vs. raster

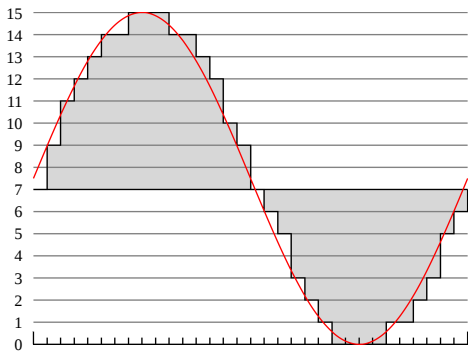
- Vector graphics can be arbitrarily scaled without loss of resolution
  - raster graphics have finite resolution, pixelate on scaling
- But vector graphics can only be computer(+human) generated
  - you can’t take a vector graphic photo
  - our AI not sufficient to figure out equations behind physical representation of real world
- Also, vector graphics can only approximate real shapes, shades humans may wish to draw

Summary: use vector graphics where you can, and stay in vector graphic format for as long as you can!

Taking photos: analogue $\rightarrow$ digital

- Taking photos an analogue to digital conversion:
  - We convert the continuous visual field . . .
  - into a pixelized representation
- Two choices on the digital end:
  - The resolution or number of pixels to capture (measured in megapixels), in 2d
  - The fidelity or color depth to capture at each point, in bits per pixel (not normally described with digital cameras)
- Beyond some resolution and depth, the human eye can't detect the difference
- But, quality has as much to do with the analogue to digital conversion (sensitivity of color detector, quality of optics, degree of noise). Otherwise, you're just getting a very high-fidelity representation of noise!

Audio



- Sound is a (compound) (continuous) wave in a medium (particularly, air)
- We convert (for each stereo channel) by:
  - sampling at certain points in time
  - measuring strength (sound pressure) of wave, to an integer value (2^n for n bits), at each point of time (quantization)
- Both sampling rate and quantization effect fidelity of signal
- Note similarity / differences with 2d image analogue $\rightarrow$ digital conversion

[0]Image: Wikipedia

Video

- Video can be represented as ordered series of 2d images
- Additional choice is number of frames per second (FPS)
- Human eye can detect below 12 fps, but strained by above
- Film projectors now often have 72 fps, but each frame repeated three times, for 24 distinct images per second
- Hz of your monitor is the refresh rate ($\approx$ FPS), with 60Hz common

Compression: why

- Images, video, sound, can take up a lot of space raw
  - How many bytes would a true-color image from a 14MP digital camera take up?
- But can be very effectively compressed.
  - JPEG version of photo can be less than 10% of size of naive representation.

Compression: how

The basic idea in compression is to find repetitive (redundant) information, and represent it more concisely.

- In image:
  - Assign codes only to colors actually used
  - Assign shorter codes to frequent colors
  - Represent run of n pixels of color k not as $kkkkk \ldots k$, but nk
- In video (additional to image):
  - Record the difference between a previous frame and the current frame
- In sound:
  - Record the change in sound pressure from previous level
  - Predict future changes from past ones

Lossless vs. lossy compression

Relative to “naive” (full, raw) representation

Lossless no information (fidelity) is lost; raw representation can be perfectly recreated (GIF, TIFF, PNG; WAV)

Lossy information (fidelity) is lost; raw representation can only be approximately recreated (JPEG; MP3; MPEG)

How of lossy compression

- Simple lossy compression: just reduced sampling rate / resolution / color depth / quantization globally
- More advanced: reduce fidelity locally (e.g. in most red area, give more of color map to red shades than green)
- Gives higher compression rates than lossless
- . . . particularly for types of file that have less raw redundancy (e.g. photos vs. line drawings)

Why of lossy compression

Why can we do this?

- Human eye, mind, ear can (partially) recreate approximated sound
- Analogue $\rightarrow$ digital was lossy to begin with!

Lossless compression of non-analogue data (e.g. text) doesn't (generally) make sense.

Integrating web services in your site

- Much dynamic, networked functionality and content can now be embedded directly on web page
- . . . without need for backend server support (on your server)

Architecture: simple embeds

- Simplest architecture embeds pagelet from other site in <iframe>
- For example, youtube videos are embedded in this way.
- But little or no
  - customizability or programmability by embedder
  - interactivity with rest of page

Architecture: Javascript

Richer embeds use Javascript

- Include link to (generally obfuscated) Javascript library (= set of functions) from service provider in your web page (downloaded by client's browser)
- Library sets up, manipulates on-page widget using DOM calls
- Call library functions from Javascript in your page to initialized, run, and interact with embedded widget
- Uses AJAX to communicate with back-end service (theirs, not yours)

Twitter

- Twitter provides a Twitter widget that is embedded inside your page.
  `http://terpconnect.umd.edu/~wew/twitter.html`
- Twitter also provides a web page where you can generate the HTML / Javascript to cut-and-paste for your widget
  `https://twitter.com/about/resources/widgets`
- HTML / Javascript then cut-and-past to website (view source)

Google Maps

- Google Maps widget a more complicated interface, provide much greater flexibility and programmability. `http://terpconnect.umd.edu/~wew/gm.html`
- Google provides example code that can be cut and pasted and then modified – as long as you understand basic Javascript (which you do!) (Also “wizards” for creating cut-and-paste)
- Interaction of Google Map library with DOM to create widget is a little more visible. (View source)
- The API is very extensive; we’re scarcely scratching the surface here.
- Note: the number of calls without a registration key is limited!

Multimedia and comprehension

Rockwell and Singleton, “The Effect of the Modality of Presentation of Streaming Multimedia on Information Acquisition”, `Media Psychology`, 9:179-191, 2007

- 132 subjects (student volunteers).
- Subjects watched presentation on Mali, Africa.
- Three presentation modalities:
  1. Text only (powerpoint presentation)
  2. Text with audio (presenting material almost word-for-word)
  3. Text with audio and video (video of presenter)

Multimedia and comprehension (cont.)

- Students were given 10-question quiz on their understanding of the material.
- Mean number of correct answers per presentation medium:
  - Text only: 7.04.
  - Text-audio: 6.37.
  - Text-audio-video: 5.98.
- Students also asked to rate presentation on various factors on (scale 1 to 5).
- Text-only version significantly preferred to multimedia (audio not significantly different); e.g.:
  - “The presentation was educational”: 4.24 / 4.00 / 3.73
  - “The presentation was interesting”: 3.47 / 3.00 / 2.84

Hypertext and comprehension

Niederhauser et al, “The Influence of Cognitive Load on Learning from Hypertext”, `J. Educational Computing Research`, 23(3): 237–255, 2000.

- 46 subjects (student volunteers).
- Parallel articles on theories of learning:

Constructivist: knowledge of outside world is an interpretation

Behaviorist: outside world is objectively knowable through (scientifically-directed) experience

- Hypertext links allowed students to click between corresponding article sections
- Students could choose to use or not use hypertext features

Hypertext and comprehension (cont.)

- Use of hypertext features measured
- Students tested on knowledge after reading
- Use of hypertext led to large and significant decrease in students' comprehension of text
- . . . even after controlling for reading comprehension, domain knowledge, etc..

Further reading

Popular treatment of the “book reading mind” vs. the “internet mind”: