![NJDH New Jersey Digital Highway — Where History, Culture, and Learning Merge]

Digital Imaging:
A brief course on digitizing techniques

New Jersey Digital Highway
Orientation and Training

September 22, 2005 – Scholarly Communication Center

Topics to cover:

- Goals of Digital Preservation

- Concepts and Vocabulary

- Digital image standards for this project

- Hardware Recommendations

- Demo of a digital imaging platform

Experience not required

- You do not need to be an imaging expert to digitize.

  Excellent results can be achieved by following simple guidelines.

Digital Imaging Advantages

- **Easy manipulation**

- **Easy Storage and Retrieval**

- **Easy Access**

The Goal: Digitize to Preserve

- Conserve brittle, old, or delicate documents
- Digitize only once: capture as much data as possible.
- Reproduce and share freely with the public
- Increase object's usability, accessibility
- Transcend the object's physical limitations

Digitizing terms and concepts

- **Pixel:** A pixel is the smallest element of a digital image ("pixel" is short for "picture element").
- Also referred to in many cases as simply a "dot."

  Like the fibers of a fabric, individual pixels are too small to notice, but they "mesh" together to make up the whole item.
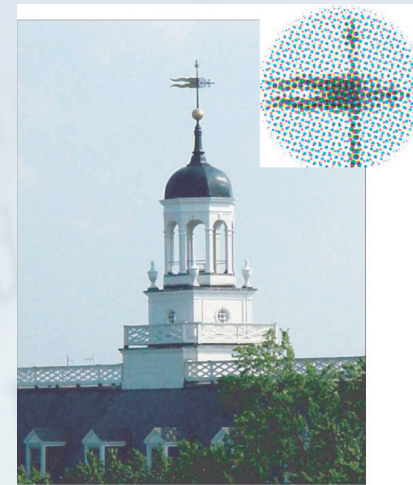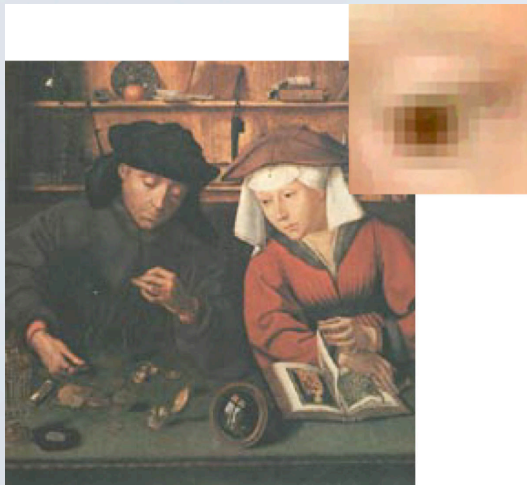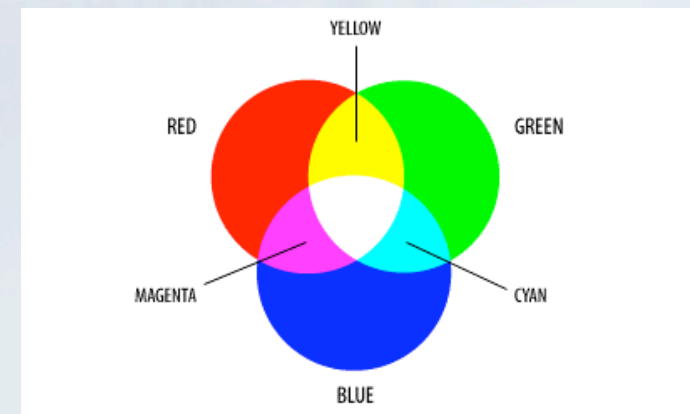
Digitizing terms and concepts

- **Resolution:** The number of pixels making up an image. For our purposes, resolution is measured in pixels per inch (ppi). Also called dots per inch (dpi).

- The higher an image's resolution, the more data it will contain about the image.

- High resolutions are needed to enlarge images beyond their original size.
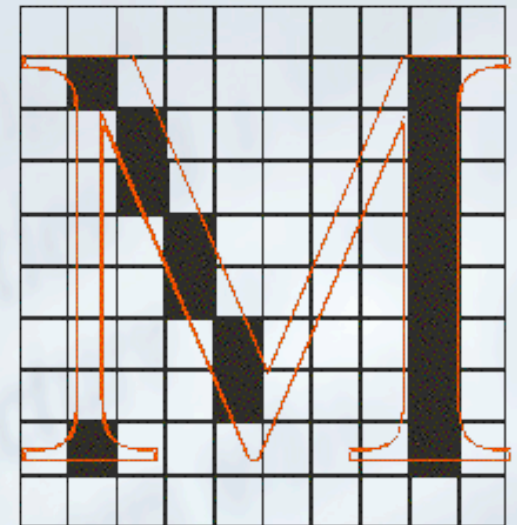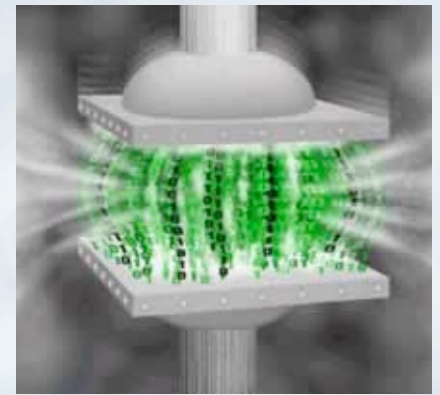
Digitizing terms and concepts

- **Color Space:** A computer model that describes a range of colors that are contained in an image

- *Described using primary colors.* Every color on a computer is a mixture of primary colors and described as percentages. Images with no color are expressed as percentages of Gray. And Black and White pixels can be only one of two colors: black, or white.

- *There is more than one color space* because no formula is perfect; one model may be able to describe colors than a different model may not have.

Digitizing terms and concepts

- **Data Compression:** A technique used to reduce the amount of space needed to store an image, or the time needed to send an image to another location (i.e. over the web).

  - Some compression methods are **lossy.** This means that a small amount of data is discarded in order to save space. This reduces image quality slightly, and the resulting image **will not** be the same as the original.

  - Some newer compression methods are lossless. This means that the best effort is made not to discard any data, and to ensure that the result image is **an exact copy** of the original.

Digitizing terms and concepts

- **File Format:** The scheme that is used to store the image data.
- Various formats exist:
  - **JPG Format** (for photos displayed on the web. Heavily compressed, sometimes very lossy.)
  - **GIF** Format (computer generated images, limited colors)
  - **TIF** Format (archival grade, no lossy compression, stores the most data)
    - We will use the TIF format to store archives of our images.

  - Additional, specialized formats are emerging (PDF, Djvu, Jpeg2000).
    - PDF and Djvu currently used for presentation.
    - Jpeg2000 being investigated for future use.

Standards for Imaging in NJDH

- Minimum requirements have been established.

- The standards ensure:
  - We can prepare for future technology advances
  - We can keep up with new displays and web standards
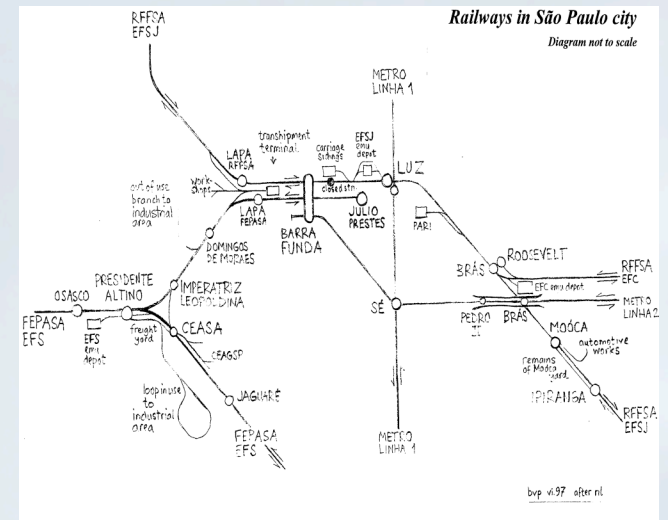 - We can easily migrate if standards must change.

Standards for Imaging in NJDH

- **Black and White Images**
  - Reserved for items that are in good condition
  - Object contains text and line diagrams **only**.

  - Minimum resolution:
    - 400 pixels per inch (ppi)

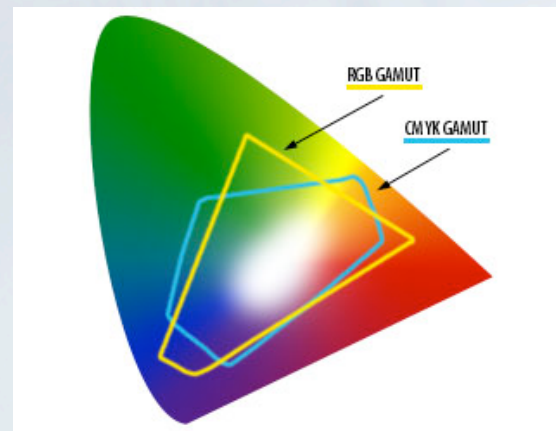# Standards for Imaging in NJDH

- **Grayscale images**
  - Photographs in good condition that do not contain *any* color information, other than shades of gray.

  - Minimum resolution: 600dpi.

# Standards for Imaging in NJDH

- **Color Images**
  - It's likely that most non-text items will be scanned in color.
  - Any item containing color information, including older documents of historical value.

  - Minimum resolution:
    - **600 dpi**

Standards for Imaging in NJDH

- **The 3,000 pixel rule**
  - Every image scanned *must* be at least 3,000 pixels in length or width.

  - If, even at the minimum dpi, the image is not at least 3,000 pixels on one axis,

    the resolution must be

    increased until this minimum

    is met.
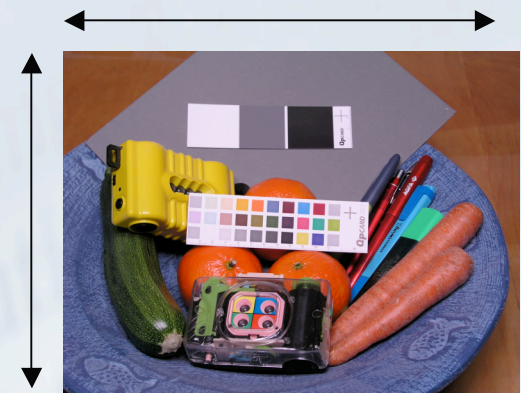
3,000 pixels

Getting Equipped

Hardware and Software
Recommendations

Hardware Recommendations

- **Considerations for hardware:**
  - Meeting the needs of your workspace:

    - Is portability required? (multiple "scan sites")
    - Fixed workstation doing multiple tasks?
    - Consider storage requirements

Hardware Recommendations

- **"What equipment should I use?"**
  - For a workstation (computer), up-to-date, existing equipment in many cases is fine.

    - Windows XP or Mac OS X
    - If PC: Pentium III Minimum, Pentium 4 or Intel Core Duo/AMD Athlon X2 series recommended
    - Large amount of RAM (1GB to 2GB or more)
    - Hard drive (at least 80GB free)
    - USB 2.0 capable, Firewire recommended
    - Should buy/upgrade to a CD/DVD Burner

Hardware Recommendations

- **"What equipment should I use?"**
  - If Buying New equipment:
    - Storage is a priority

      (lots of memory, large hard drive)
    - 1-2GB of RAM
    - At least 160GB hard drive (the larger the better)
    - Intel Core Duo processor for faster image handling
    - DVD Burner (Dual Layer recommended)

  - Even with these specs, a new computer can be purchased for around $1,000

Hardware Recommendations

- **"What equipment should I use?"**
  - Imaging Equipment
    - Flatbed scanner
      - Good for photographs, individual sheets of paper
      - Equipment of choice for most light-to-moderate scanning requirements
      - $Cheap!$
      - Drawbacks: labor-intensive if scanning large batches of documents.
    - Flatbed Scanner Requirements
      - 2400 ppi optical scanning resolution
        - Beware of "interpolated" or "software enhanced" resolution specs.

Hardware Recommendations

- **"What equipment should I use?"**
  - Imaging Equipment
    - Flatbed scanner
      - Typical price with these requirements: $99 - $200

Hardware Recommendations

- **"What equipment should I use?"**
  - Imaging Equipment
    - Sheet Feed or Document Scanners
      - Excellent choice if scanning very large volumes of letter or legal-sized documents, especially multi-page. Makes fast work of multipage documents.
      - Drawbacks: Not useful for photographs or "non-standard" sized objects.
      - Very pricey, but may be worth it in labor savings.
      - 600-1200 dpi resolution is acceptable.

Hardware Recommendations

- **"What equipment should I use?"**
  - Imaging Equipment
    - Sheet Feed or Document Scanners

      - Typical cost: $600 - $1,500

      - **NOT recommended for brittle objects!**

Hardware Recommendations
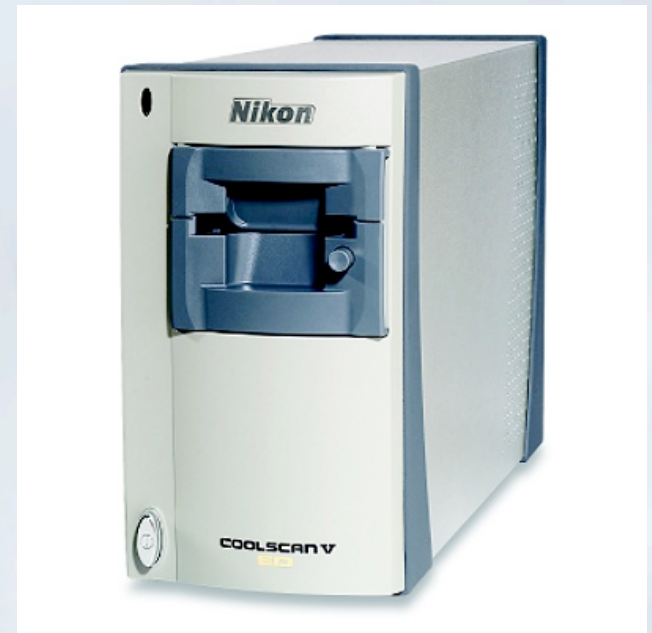
- **"What equipment should I use?"**
  - Imaging Equipment
    - Slide/Film Scanners
      - Virtually required if you have a significant number of slides or negatives that need to be digitized.
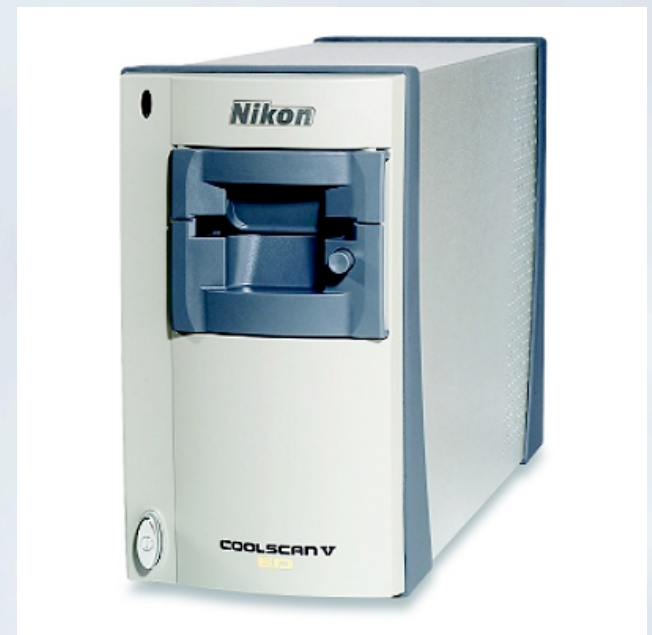   - Slide scanner "attachments" on flatbed scanners often will not do the job properly.
  - Requirements:
    - Capable of up to 3600 ppi resolution or higher

Hardware Recommendations

- **"What equipment should I use?"**
  - Imaging Equipment
    - Slide/Film Scanners
      - Typical cost: $600 - $2,000 depending on features.

Outsourcing to a Vendor

- Hiring a third-party company to digitize objects that require special handling.

- Often a requirement for large numbers of bound volumes, and large format items (maps, posters, items greater than letter size).

- Per-item cost can be very inexpensive if done in significant volume

- Vendor must adhere to minimum requirements (minimum resolutions, 3,000-pixel rule).

Software

- Image editing software (Adobe Photoshop, Corel Paint Shop Pro) is required to view and manipulate images, and create presentation copies.

- Costs:
  - Adobe Photoshop: $600 (institutional discounts available)
  - Paint Shop Pro: $195

Equipment Demo
Portable Imaging Platform

Questions or comments?

Technical Support Issues?
Please contact:

Isaiah Beard

Digital Projects Coordinator, NJDH

(732) 932-8573 ext 177

isaiah.beard@rutgers.edu