Scanning and Images Part II

Instructor Mike Garrambone

Interesting Genealogy Topics

Scanning and Images Part II: Color, Images, OCR, & Tricks

Mr. Michael W. Garrambone

Family History Jamboree
Saturday, 28 Mar 2026
Centerville, Ohio

Scanning & Images II 1

1

What this is About

Everyone wants images in their genealogy documents

Great Uncle Bob

- Family Photos
- Documents
- Paper Narratives
- Maps and Charts
- Graphic Treasures

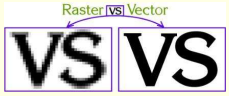
The next level about what you need to know to capture and keep those images

Scanning & Images II 2

2

Agenda

MG?



- Introduction
- Scanning a Photo
- B/W and Color
- Image Ideas
- Fixing Images
- Image Types
- Format Types

Scanning & Images II 3

3

Recall: Scanning is

Pixel Grabbing

Scanning: Using a device (a scanner) to detect and capture the information, arrangement, color, text characters or images and storing this information in a computer graphic or file of some format

Image

Scanner

Computer File

Scanning & Images II 4

4

Let's Scan a Photo

3" wide x 4" tall

Cute Little Tara Pooh

- Scan information
  - Scan at 300 dots per inch (dpi)
  - Create a JPG file
  - Scan in color
  - File name is TaraPooh 300.jpg

Pixel = Picture Element, one dot on a display screen

Scanning & Images II 5

5

Scanning a Photo

Photo on the scanner glass

Using Photoshop Elements scanning software here

Image is 3" wide x 4" tall

Scanning tells the computer how big the image really is

- Result
  - (3x300) x (4x300)
  - Width is 900 & Height is 1200 pixels
  - Total number of pixels is 1,080,000 (1.08 MP)
  - Each pixel needs color information

Scanning & Images II 6

6

Scanning and Images Part II

Instructor Mike Garrambone

Color Information

Color makes an image more interesting

1. Black and white
   - Like a printing press
   - Makes the smallest file size

2. Shades of black and white (gray scale)
   - Eight shades
   - 16, 32, or more shades

3. Color
   - Eight color
   - 16, 32, 64, 128, 256, or millions

Scanning & Images II 7

7

Talk about Black & White

Today's Scanners can capture Black & White and Color images

Simple Black ■ and □ White takes only one computer bit to store each pixel's information

1 bit for Black/White

1	0

- Eight shades of Black and White takes three computer bits to store each pixel's information

Simple 8 shades for gray scale

3 bits for 8 shades

7	To	0	
0	0	0	White
•			
•			
•			
1	1	1	Black

Scanning & Images II 8

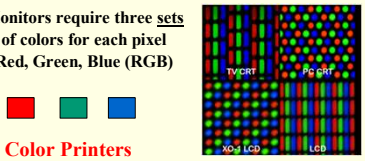
8

Color on Monitors & Printers

Computer Monitors

Light

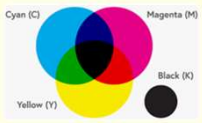
Monitors require three sets of colors for each pixel Red, Green, Blue (RGB)

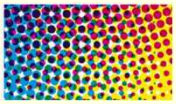


Color Printers

Printers require four sets of colors for each dot

Ink





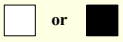
Scanning & Images II 9

9

B/W & Color File Examples

(8 bits = 1 byte) (1 million bytes = 1 MB)

Black and White



- Let's look at pure B/W (No shades)

Uses 1 bit per pixel

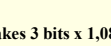
Takes 1 bit x 1,080,000 pixels for photo

Takes 1,080,000 bytes, or 1,080 Kilobytes (KB), or 1.08 Megabytes (MB) for the photo

Eight Colors

- Let's look at 8 colors

Uses 3 bits per pixel



000 = 0	100 = 4
001 = 1	101 = 5
010 = 2	110 = 6
011 = 3	111 = 7

Takes 3 bits x 1,080,000 pixels for photo

Takes 3,240,000 bytes, or 3,240 KB, or 3.24 MB for the photo

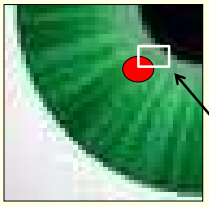
Scanning & Images II 10

10

The Truth: Scanners Only Sample Information

Density: The number of dots in the space of an inch (both directions)

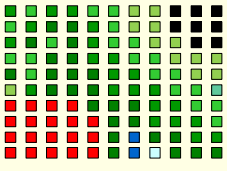
Subject being scanned



This Section Only

300 dots per inch

Sample taken

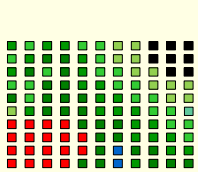


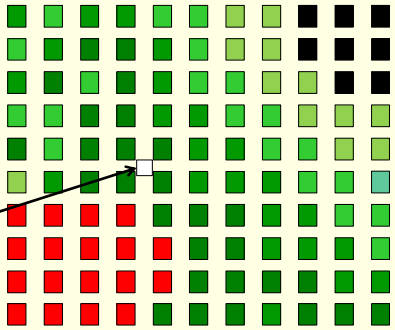
300 dots per inch

Scanning & Images II 11

11

Stretching an Image





- Pixels can be stretched
- But there is a limit

Pixelation

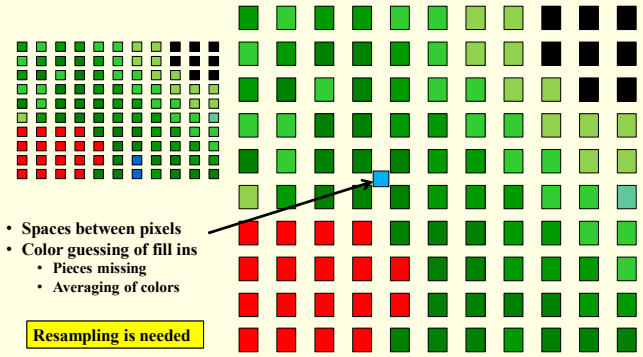
Scanning & Images II 12

12

Scanning and Images Part II

Instructor Mike Garrambone

Effects of "too much" Stretching



- Spaces between pixels
- Color guessing of fill ins
  - Pieces missing
  - Averaging of colors

Resampling is needed

Scanning & Images II 13

13

Images

Images enhance our genealogy

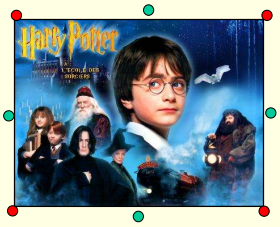
- **IMAGE**: a visual representation of something: as *(1)*: a likeness of an object produced on a photographic material *(2)*: a picture produced on an electronic display (as a television or computer screen)

Scanning & Images II 14

14

Aspect Ratio

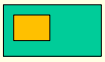
Every image has shape called its Aspect Ratio



The Aspect Ratio
Matches the length of the image to the width
Shows the numbers in reduced form

Width: 1024 pixels	1024	4
Height: 768 pixels	768	3

Monitors also have Aspect Ratio
They are normally 4 : 3 or 16: 9
My monitor is 1920 x 1200, or 8 x 5



Scanning & Images II 15

15

Framing in PowerPoint

Just some examples

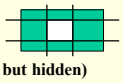
Scanning & Images II 16

16

Four PowerPoint Cropping Secrets

Density: The number of dots within an inch (both directions) = DPI

1. You can easily crop an image in PowerPoint
   This removes the cropped parts from view (still there, but hidden)

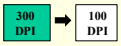


Keep just the white

2. You can put any part of image back to the original from the crop
   This allows you to fix mistakes!

3. You can discard the edges of the cropped image
   Reduces the file size and the briefing size as well

4. You can change (e.g., reduce) the image density in PowerPoint
   PowerPoint can take a big file image and reduce it's DPI

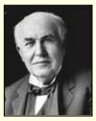


Scanning & Images II 17

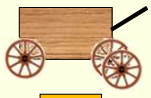
17

Easy PowerPoint Image Duplication

Copy an image with (Ctrl C) then duplicate with (Ctrl D)



So
This becomes this



Many wagon wheels

Place the first one and the others will fall in the path with more Ctrl Ds

Scanning & Images II 18

18

Scanning and Images Part II

Instructor Mike Garrambone

Recall the Four Popular Image Types

Each image type has different characteristics

- **Raster**: Made of various dots
- **Vector**: Made of various lines
- **Character**: Made of a standard type-sets (font) and size (points)

Letters, Symbols, Numbers

- **Compound**: Combinations of the above that looks like this

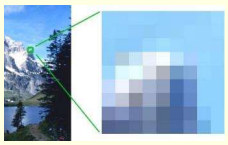
19

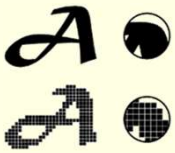
Raster (Bitmap Images)

Raster Images are made from dots

Original photo

Blown up looks like boxes





- BMP (Bitmap)
- TIFF (Tagged Image File Format)
- JPEG (Joint Photographic Expert Group)
- GIF (Graphics Interchange Form)
- PSD (Photoshop Document)
- TGA (Truevision Graphics Adapter)
- PNG (Portable Network Graphic)
- MNG (Multiple-image Network Graphics)

20

File Compression

Compression: Reducing the size of image to take less storage

1. You can change from one file format to another

A BMP becomes a smaller TIFF becomes a smaller JPG

BMP file
4, 571 KB

TIFF file
860 KB

JPG file
240 KB

2. You can use the PPT compression button

21

File Compression Example

The data for an image is kept in the image file

- Compression
  - Can reduce the size of the file
  - Can either keep all the image information, or you can lose some information
  - LossLess compression (all info is kept - BMP, PNG, GIF)
  - Lossy compression (some info is lost - JPG, HAM, JBIG)

Example Black and White file (letters used in place of binary code)

WWWWWWWWWWWWBWWWWWWWWWWWWBBBWWWWWWWWWWWWWWW
WWWWWWWWBWWWWWWWWWWWWWW

This changes to this (with no loss of info)

12W1B12W3B24W1B14W

Same file info but compressed and no information loss

22

Some Vocabulary

Color Words:

- **Transparency**: Ability to see through part of the image and show background
- **Alpha Channel**: An area where parts of the image can be transparent
- **Color Depth**: The number of colors in the image (more depth-increase size)
- **Color Palette**: A list of the colors available in image
- **Gamma Correction**: A way to correct the image lightness or darkness on devices

File Words:

- **Compression**: A way to make the file size smaller
- **Lossless Compression**: File smaller, but no information lost
- **Lossy Compression**: File smaller, but image quality is reduced
- **Interlaced**: Send files on the internet where the image builds over time

23

BMP Files

BMP Basic Thoughts:

- Two-dimensional, universal to windows
- Any size, any height, any width, any density
- Editable by pixel – easy to recolor sets of pixels

Advantages:

- No compression, keeps quality
- May be only choice on your scanner
- One bit to 24 bit color

Summary:
Can be beaten by Tiff, JPG, PNG

Disadvantages

- Files are very large
- Not support by web browsers
- Hard to stretch to larger image

24

Scanning and Images Part II

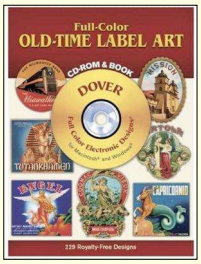
Instructor Mike Garrambone

TIFF (Tagged Image File Format)

TIFF Basic Thoughts:
- Two-dimensional, universal to windows
- Most widely supported format (MAC/Windows)
- Editable by pixel – easy to recolor sets of pixels

Advantages:
- Optional compression, keeps quality
- May be only choice on your scanner
- Can use any color depth

Disadvantages
- Not support by web browsers
- Files are typically very large
- Many different type of TIFF files



Summary: Industry standard, beats BMP

Scanning & Images II 25

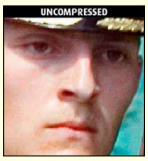
25

JPG (Joint Photographic Experts Group)

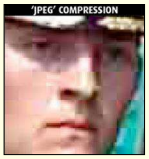
JPG Basic Thoughts:
- Two-dimensional, universal to windows
- Most widely supported format
- Platform independent, Internet standard

Advantages:
- Superior and variable compression
- Interlacing for progressive building
- 24 bit color depth

Disadvantages
- Compression loses information
- Edit and resave causes degradation
- Not suitable for simple pictures



Summary:
Industry standard, Internet standard



Scanning & Images II 26

26

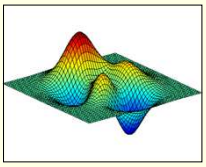
GIF (Graphic Interchange Group)

GIF Basic Thoughts:
- Two-dimensional, universal to windows
- Most widely supported, Internet standard
- Can create animated images

Advantages:
- Lossless compression and transparency
- Interlacing for progressive development
- Supports single or multiple image data

Disadvantages
- Only 256 Colors
- Lossless compression inferior to JPG
- Limited transparency (in Alpha Channel)

Summary:
Low color, small files, animation



Scanning & Images II 27

27

PNG (Portable Network Graphic)

PNG Basic Thoughts:
- Two-dimensional, universal to windows
- Most widely supported Internet standard
- Can have any color depth

Advantages:
- High level lossless compression
- Alpha Channel transparency
- Gamma correction and interlacing

Disadvantages
- Lack of support in older browsers
- Provides less compression than JPG
- No support for multi-image, animation like GIF

Summary:
Full color, small files, transparency

Scanning & Images II 28

28

Vector Images

Images are made by putting together lots of lines

Lines

Easy to stretch lines without loosing image

Great for precision
Great for graphic plotters

- CGM (Computer Graphics Metafile)
- SVG (Scalable Vector Graphics)
- AI (Adobe Illustrator)
- CDR (CorelDraw)
- GEM (Graphics Environment Manager)

Scanning & Images II 29

29

TXT or ASCII Character files

Each character has a symbol-code

- Character Files
  - .TXT (Text) --- DOS text files
    - Works with windows
  - .DOCX (Word Files)
    - Great for word processing

1212 Cat Street ← You see this

1 2 1 2 sp C a t sp S t r e e t

49,50,49,50,32,67,141,116,32,83,116,114,101,101,116

Computer sees this ↑

Example ASCII Code

Binary	Oct	Dec	Hex	Glyph
010 0000	040	32	20	space
010 0001	041	33	21	!
010 0010	042	34	22	"
010 0011	043	35	23	#
010 0100	044	36	24	$
010 0101	045	37	25	%
0110000	060	48	30	0
011 0001	061	49	31	1
011 0010	062	50	32	2
100 0001	101	65	41	A
100 0010	102	66	42	B
100 0011	103	67	43	C
110 0001	141	97	61	a
110 0010	142	98	62	b
110 0011	143	99	63	c

Scanning & Images II

30

Scanning and Images Part II

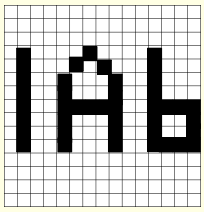
Instructor Mike Garrambone

Optical Character Reader (OCR)

Your scanner is told to look for characters

- OCR (Optical Character Reader)
  - Scans a document and captures characters
  - Dots are gone and only characters remain

Text at 300 dpi is best
Times New Roman Font is Best
Be ready to edit and format



1 A b

Note. OCR is not always perfect

31

I Scanned this document as a JPG Image

```
**********************************************************************************
ITALY, AGRIGENTO, NARO - CIVIL REGISTRATION

Naro (Agrigento).  Ufficio dello stato civile.                    +--------------+
   Stato civile, 1821-1865. -- Salt Lake City : Filmati dalla     |EUROPE        | I
   Genealogical Society of Utah, 1988. -- in 20 bobine di         |FILM AREA     |
   microfilm ; 16 mm.                                             +--------------+

   Microfilm dei registri originali nell'Archivio di Stato, Agrigento.
   Microfilm di alta riduzione (42x).  Adoperate una macchina di alto
      ingrandimento.
   Births, marriages, banns, deaths.
   Include indici parziali.

   Nati, morti, notificazioni, ---------------------------------- 1468162        I
      matrimoni  1821-1823
   Nati, morti, notificazioni, ---------------------------------- 1518044
      matrimoni  1823-1825
   Nati, morti, notificazioni, ---------------------------------- 1518045
      matrimoni  1825-1828
   Nati, morti, notificazioni, ---------------------------------- 1518046        I
      matrimoni  1828-1830, 1829
   Matrimoni    1829 ------------------------------------------- 1546587        I
      Nati, morti  1830
      Nati, morti, notificazioni,
      matrimoni    1831
   Nati, morti, notificazioni, ---------------------------------- 1468217
      matrimoni    1832-1833
      Nati         1834
   Nati, morti  1834-1835 -------------------------------------- 1468218
      Notificazioni, matrimoni  1834-1835
      Nati                      1836
```

See next Slide

32

Then "OCR it" into a White Text File

ITALY, AGRIGENTO, NARO - CIVIL REGISTRATION

Naro (Agrigento). Ufficio dello stato civile.
Stato civile, 1821-1865. -- Salt Lake City : Filmati dalla Genealogical Society of Utah,. 1988. In 20 bobine di microfilm: 16 mm

20 Rolls

Partial Indices (nice)

Microfilm dei registri originali nell' Archivio de Stato, Agrigento.
Mircofilm di alta riduzione (42x) Adoperate una macchina de alto ingrandimento
Births, Marriages, Banns, Deaths.
Include indici parziali

Film Order Numbers

Contents

Nati, morti, notificazioni- - - - - - - - - - - - - - - 1468162
matrimoni 1821 -1823
Nati, morti, notificazioni- - - - - - - - - - - - - - - 1468163
matrimoni 1823-1825

33

Portable Document Format (PDF)

Your scanner can create single page or multi-page PDFs

- PDF files are independent of application software, hardware, and operating systems
- PDF files capture a complete description of a fixed-layout flat document, including text, fonts, graphics, and other information needed to display it

The PDF combines three technologies:

1. Uses PostScript for generating the layout and graphics
2. The font tables are embedded and travel within the file
3. It stores everything into a single file, compressed where needed

http://en.wikipedia.org/wiki/Portable_Document_Format

34

Summary

Today we talked about:

- Photo Scanning and File Size
- B/W and Color Examples
- Stretching and Aspect Ratio
- Some Nifty PowerPoint Tools
- Reviewed Image Types
- Reviewed File formats

GIF File Image

You now know more that the average

35

Point of Contact Information

Mr. Michael W. Garrambone

: Home 937-233-3255
: 4138 Quail Bush Drive
Dayton, Ohio 45424
: mgarrambone@aol.com

36

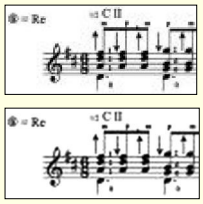
Scanning and Images Part II

Instructor **Mike Garrambone**

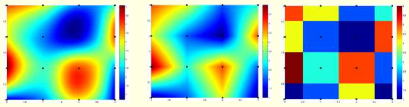
Six Resampling Methods

Scaling up using nearest neighbor

Scaling up using 2 SAL method



Scaling up using Lanczos method



Bicubic | Bilinear | Nearest Neighbor

Pixelation

37

More about Images

Every image has size, shape, and density

Document (Image) Size:
Width: 14.22 inches
Height: 10.667 inches
Resolution: 72 pixels/inch

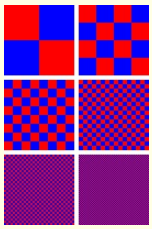
Total number of Pixels (2.25 M)
Width: 1024 pixels
Height: 768 pixels

☑ Scale Size
(Affects proportion when image is resized)
☑ Constrained Proportions
(Constrains aspect ratio)
☑ Resample Image
(Interpolate pixel information)

Interpolation methods: Bilinear, Bicubic, Bicubic Smoother, Bicubic Sharper, Nearest Neighbor

38

Two Visual Techniques



Dithering and Halftone

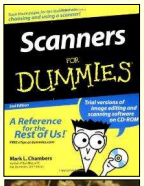
- **Dithering:** A a technique used in computer graphics to create the illusion of color depth in images with a limited color palette (color quantization)
- A dithered image, makes colors not available in the palette by approximating them by a diffusion of colored pixels from within the available palette
- The human eye perceives the diffusion as a mixture of the colors within it
- Dithered images, particularly those with relatively few colors, can often be distinguished by a characteristic graininess, or speckled appearance

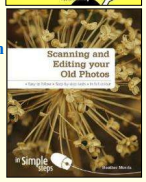
Halftone: The reprographic technique that simulates continuous tone imagery through the use of dots, varying either in size, shape, or spacing

39

References

- Scanning 101 A few tips by Wayne Fulton
http://www.scantips.com/
- Hewlett-Packard Official Scanner Handbook
Busch, Krzywicki, & Burden, 1999
- Image File Formats
http://en.wikipedia.org/wiki/Image_file_formats#CGM
- Image Scanner
http://en.wikipedia.org/wiki/Image_scanner
- Optical Character Recognition
http://en.wikipedia.org/wiki/Optical_character_recognition
- Laser Printing
http://en.wikipedia.org/wiki/Laser_printing
- Scanners for Dummies
- Scanning and Editing Your Old Photos
- Official Scanner Handbook





40