Scanning of Analog Materials for Preservation, Archiving, and Distribution

Steve D. Bowman, Ph.D., P.E., P.G.
Geologic Hazards Program Manager

Various Scanner Hardware

Scanning Resolutions

Choosing the right scanning resolution is critical to ensure end-users can effectively use the data, and to reduce the need for re-scanning and additional handling.

Slide Enlargement	Scanning Resolution (dpi)							
	1000	2000	3000	4000	5000	6000	7000	8000
1x (1.34”H by 0.91”V)	1000	2000	3000	4000	5000	6000	7000	8000
2x (2.68”H by 1.82”V)	500	1000	1500	2000	2500	3000	3500	4000
3x (4.02”H by 2.73”V)	333	667	1000	1333	1667	2000	2333	2667
4x (5.36”H by 3.64”V)	250	500	750	1000	1250	1500	1750	2000
5x (6.70”H by 4.55”V)	200	400	600	800	1000	1200	1400	1600
6x (8.04”H by 5.46”V)	167	333	500	667	833	1000	1167	1333
7x (9.38”H by 6.37”V)	143	286	429	571	714	857	1000	1142
8x (10.72”H by 7.28”V)	125	250	375	500	625	750	875	1000
9x (12.06”H by 8.91”V)	111	222	333	444	556	667	778	889

Make sure you are using optical resolution, it is hardware dependent, need to verify the scanner specifications.

Scanning Bit-Depth

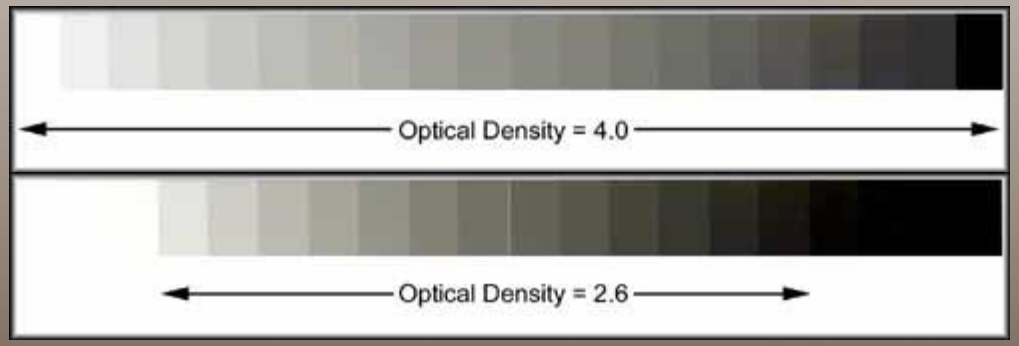
Choosing the right scanning bit-depth is critical to ensure end-users can effectively use the data, and to reduce the need for re-scanning and additional handling.

Bits per Pixel	Bits per Channel				Colors or Grey Levels
	1 Channel Greyscale	3 Channel RGB			
		Red	Green	Blue	
1	1	--	--	--	2
8	8	3	3	2	256
12	12	4	4	4	4096
16	16	5	5	6	65,536
24	24	8	8	8	16.8 million
36	36	12	12	12	68.7 billion
48	48	16	16	16	2.8E14

Scanner Optical Density or Dynamic Range: Why it Matters

Optical density/dynamic range (D_{max}) is the ability of the scanner to measure the range of color tones (range from black to white in a grayscale image, or ranges of red, blue, and green tones in a color image).

Non-professional scanners will typically have a low dynamic range. Document oriented scanners will also have a lower dynamic range than transparency/film scanners, as film has a much larger possible dynamic range than paper printed products.



Good film scanner

Poor film scanner

Color Spaces and Profiles

- Gamut or range of reproducible colors, defined by standards, such as AdobeRGB (1998), sRGB, etc. and helps to reproduce accurate color on a variety of devices. Choosing a color space is required if color calibration is used. The AdobeRGB color space was developed to encompass most reproducible colors and humanly visible, where the sRGB color space was developed mainly for computer displays and consumer devices, and can encompass fewer reproducible colors. As a result, the AdobeRGB color space should be used whenever possible for color images.

- Suitable scanner software should be used that allows for scanner color calibration using targets (reflective and transparent) to develop ICC color profile files to retain consistent color across a range of devices (scanners, printers, monitors). The Gray Gamma 2.2 color space is the most common for grayscale images.

Scanner Hardware

- A high-quality scanner is required for archival use.
- Most consumer scanners to do not have the appropriate optics and sensor(s) for archival resolutions, optical density, and low geometric and optical distortion. Lack the ability to control focus.

Make and Model	Max Optical Resolution (dpi)	Bit Depth (bits/pixel)[1]		Optical Density (Dmax)	Max Scan Size (in)	Max Transparent Size (in)	Reliability (MCBF)	Sensor[2]	Light Source[3]	Scan Speed[4]
		Color	Grey							
Epson										
V850	6400x9600	48 I/E	16 I/E	4.0	8.5x11.7	Yes, 8x10	100,000	CCD	LED	
V560	6400	48 I/E	16 I/E	3.4	8.5x11.7	Yes, 2.7x9.5	30,000	CCD	LED	
V9	4800	48I/24E	16I/8E	?	8.5x11.7	No	10,000	CIS	LED	
11000	2400	48 I/E	16 I/E	3.8	12.2x17.2	Yes, 12.2x16.5	100,000	CCD	Cold FL	
Canon										
9000F Mark II	4800, 9600 (film)	48 I/E	48I 16-8E	?	8.5x11.7	?	?	12 line CCD	White LED	7 sec. Color
DR-F120	600	24[5]	8[5]	?				1 line CIS CMOS	RGB LED	10/18 ppm 20/36 ppm

Scanner Driver Software

- Many scanner manufacturer drivers are not suitable for high-quality scanning.
- Drivers may not be available for current operating systems, as many scanners were discontinued years ago. Few high-quality scanners still made.

- Silverfast scanner software – can control many scanners from Canon, Epson, Microtek, Nikon, and others.
  - Automatic IT8 color calibration of scanners.
  - Printer color calibration (ICC color profile for specific scanner and paper).
  - Multiple scan image enhancements:
    - Multi-exposure / High Dynamic Range (HDR) – reduce shadowing, etc.
    - Infrared dust and scratch removal – transparencies (films/slides) on select scanners only.
    - Film grain reduction
    - Noise reduction
  - Focus control (rare for most scanner software).
  - Develop scanning workflows for consistent scanning.

Output File Formats

- No compression or lossless compression file formats, such as TIFF and PDF should be an available option in the scanner driver software. If any post-scan processing is to be performed, lossy compression formats, such as JPEG, should not be used, due to compression degradation each time a file is saved.

- TIF – universal image file format; lossless, but can contain compressed (LZW, JPEG, or deflate) data; extensive metadata possible, including geolocation.
- PDF – universal document file format; lossless or compressed (JPEG, JPEG2000, deflate); extensive metadata possible using Adobe's XMP system.

- Compression
  - Lossless – compressed data can be uncompressed exactly to the original.
  - Lossy – compressed data when uncompressed does not match the original due to data loss, also results in data quality reduction. Every save degrades the data quality.
    - JPEG2000 can be a good compression format for relatively small file sizes; however, does require computational power to compress and uncompress.