open.michigan

Author(s): Paul Conway, 2008-2011.

License: Unless otherwise noted, this material is made available under the terms of the Creative Commons Creative Commons Attribution - Non-Commercial - Share Alike 3.0 License: http://creativecommons.org/licenses/by-nc-sa/3.0/

We have reviewed this material in accordance with U.S. Copyright Law and have tried to maximize your ability to use, share, and adapt it. The citation key on the following slide provides information about how you may share and adapt this material.

Copyright holders of content included in this material should contact **open.michigan@umich.edu** with any questions, corrections, or clarification regarding the use of content.

For more information about **how to cite** these materials visit http://open.umich.edu/education/about/terms-of-use.

Any **medical information** in this material is intended to inform and educate and is **not a tool for self-diagnosis** or a replacement for medical evaluation, advice, diagnosis or treatment by a healthcare professional. Please speak to your physician if you have questions about your medical condition.

Viewer discretion is advised: Some medical content is graphic and may not be suitable for all viewers.





Citation Key

for more information see: http://open.umich.edu/wiki/CitationPolicy

les L Chara L Adapt								
Jse + Share + Adapt								
{ Content the copyright holder, author, or law permits you to use, share and adapt. }								
Ø PD-GOV	Public Domain – Government: Works that are produced by the U.S. Government. (17 USC § 105)							
Ø PD-EXP	Public Domain – Expired: Works that are no longer protected due to an expired copyright term.							
Ø PD-SELF	Public Domain – Self Dedicated: Works that a copyright holder has dedicated to the public domain.							
(G)) 2ERO	Creative Commons – Zero Waiver							
(C) 8Y	Creative Commons – Attribution License							
(C) BY SA	Creative Commons – Attribution Share Alike License							
(O) BY-NC	Creative Commons – Attribution Noncommercial License							
(O) BY-NC-SA	Creative Commons – Attribution Noncommercial Share Alike License							
GNU-FDL	GNU – Free Documentation License							

Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

Public Domain – Ineligible: Works that are ineligible for copyright protection in the U.S. (17 USC § 102(b)) *laws in your jurisdiction may differ

{ Content Open.Michigan has used under a Fair Use determination. }



2

Fair Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (17 USC § 107) *laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should do your own independent analysis to determine whether or not your use will be Fair.

SI 675 Digitization for Preservation

Week 2 – Image Digitization and Guidelines

What is a bitmap?
Representation and rendering
Guidelines
Transformations

"Our culture wants both to multiply its media and erase all traces of mediation..." Bolter & Grusin, 1996

Code as Map: Digital Bitmaps

A bit-mapped image is a digital picture made up of rows of pixels in a grid.

- Resolution
- Dynamic range
- Pixel size and shape



5

First Digital Image (176 x 176 pixels) Russell Kirsch, NIST, 1957 http://www.nist.gov/public_affairs/techbeat/tb2007_0524.htm#image

SI 675 Digitization for Preservation Winter 2011

Spatial Resolution – Pixel Grid



Adapted from Draco2008, Flickr, <u>http://www.flickr.com/photos/draco2008/4571175927/sizes/z/in/photostream/,</u> <u>http://creativecommons.org/licenses/by/2.0/</u>

Pattern and Symbol in Code



Coded pattern:

3 mm letter "e" at 600 dpi scanned from negative microfilm

4,900 bits vs. 8 bits

Coded symbol: e = 01100101 (ASCII code 101)

The number of possible colors or shades of gray that can be included in a particular image.

- Bitonal: sampled image level for each pixel is rounded to 0 (black) or 1 (white). One bit per pixel.
- 8-bit gray: sampled image level for each pixel is rounded to one of 256 values. 8 bits per pixel.
- 24-bit color: the R, G, and B image values for each pixel are rounded to one of 256 values and encoded as 8 bytes.

- Real Resolution (proportion surface detected)
- Addressable Resolution (claimed surface coverage)
- Real resolution depends upon engineering.
- Synthetic resolution inflates addressable resolution with software.
- Beware of scanner manufacture claims. Undertake rigorous testing and benchmarking.

Digital imaging uses a hierarchy of coding systems to represent an object.

- Data (strings of 0s and 1s)
- File format (organized data)
- Image headers (digital data about one image file)
- Metadata (digital data about digital objects)

Visual Representation

- "... representation is always of something or someone, by something or someone, to someone." [Mitchell]
 - Intentionality [camera | scanner]
 - Processes [darkroom | Photoshop]
 - Materiality [content | artifact]



SI 675 Digitization for Preservation Winter 2011

- "Our culture wants both to multiply its media and erase all traces of mediation..." (Bolter & Grusin 1996).
 - Technology transparency is goal
 - Hypermediacy: presence of old in new gives a sense of multiplicity
 - Technologies of rendering (making visible) are tools for mediating mediation
- "No text is an innocent by-product...but rather a constructed product," perhaps unconsciously [Cook 2001]

Summary of Digital Concepts

- Digital imaging is representation
- Communication requires shared symbols
- Bitmaps are pictures without intelligence
- Meaning through metadata coding or text conversion

Guidelines for Digitization



National Archives (2004)



Library of Congress (2006)





North Carolina (2007)

Colorado (2008)





Specific Research Questions

How is consensus represented in "best practices" for photograph digitization?

- HI: evidence of community
 - tightly focused influences
- H2: convergence over time
 - increase synthesis; single document
- H3: experience + capabilities raise quality
 - significant rise in specifications

Research Methods

- Definition / ID of target guidelines
 - review of 95 potential documents
 - criteria (expressive, specific, photos)
- Network analysis of 17 documents
 - extract personal and organizational names
 - search for patterns of relationships
- Quality normalization
 - method for comparing specifications
 - BQI = resolution + depth

Findings: Influences of Documents



Findings: Influences of Organizations



Findings: Influences of Experts



Findings: Experiments to Synthesis

20



SI 675 Digitization for Preservation Winter 2011

- NARA 04: "We recommend doing minor post-scan adjustment to optimize image quality and bring all images to a common rendition." (p. 31)
- □ LoC 2006: "The Library wishes to provide researchers with a reproduction of the entire original item." (p. 15)
- NC 2007: "Access images may be of varying quality and are generally manipulated for better display upon the screen or page." (sec. 4.2)
- Colorado 2008: "As a rule, the key to quality imaging is not to capture at the highest resolution possible, but to scan at a level that matches the informational content of the original." (p. 22)

Metamorfoze Preservation Imaging (2007)

National Library of the Netherlands and the Nationaal Archief

- Emphasis on Technical Quality Criteria enforced at the point of scanning.
 - Tonal reproduction
 - Noise
 - Signal-to-noise ratio
 - Illumination
 - Color cast and accuracy
 - Detail reproduction (resolution)
 - Errors (distortion, mis-registration, artifacts, defects)

Federal Agencies Digitization Guidelines

- I2-agency collaborative (images and audiovisual)
- Digital still images, version 1.0 2009
 - □ NARA (2004) guidelines as point of departure
 - Gap Analysis (research priorities)
 - Content Categories/Digitization Objectives
 - Digital Imaging Framework (quality definitions)
 - Tiff Image Metadata guidelines
 - Glossary and format information

http://www.digitizationguidelines.gov/



- Scanner setup and calibration
- Raw scan data adjustment (archival master)
- Post-scan manipulation (mediation)
- Creation of derivative versions (access)

Common Technical Decisions

- Assign color profile (screen or print)
- 2. Adjust/Correct color
- 3. Adjust/Correct tone (histogram)
- 4. Crop /deskew
- 5. Reverse polarity (negative to positive)
- 6. Apply sharpen mask
- 7. Remove scanner and film effects
- 8. Resize for screen display (store master)

Adjust Color

Item as it appears today

Item adjusted for aging





Champions of America, 1865. Baseball and Jackie Robinson. (American Memory).

Adjust Tone

Balanced Tone Values



Shadow Values Highlighted



Nebraska State Historical Society, Digital Imaging Laboratory, History Revealed, 1998.



Union Pacific Workers, 1866.

Army Teamsters, 1864.



Denver Public Library, Western History Photograph Collection



Ohio Historical Society, African-American Experience Collection

SI 675 Digitization for Preservation Winter 2011



Polarity Shift

Copy Negative



Original Photograph



@ PO-EXP

Julia Williams Wadsworth, 1930. Slave Narratives, American Memory, American Memory.

Sizing - Change in Scale

Original Photo – ca. 3" x 5"



Cropped Negative – 3" x 5"



Carnegie Museum of Art, Teenie Harris Archive, 1915-1970+

1921

SI 675 Digitization for Preservation Winter 2011

1927

Effects of Intermediate



Portrait of Marian Anderson, 1940. By Carl van Vechten Copy negative, reversed.

Creative Americans. American Memory, Library of Congress.



Altered Context

Henry Bedford, ex-slave. 1930.



Ø PD-GOV

African-American Experience in Ohio Ohio Historical Society (American Memory)

Hattie Ann Nettles, 1937.



Slave narratives from the Federal Writers Project American Memory, Library of Congress [accompanies 2-page story]

Pictorial Rendering Intent – Four Ways

- The photographic image is rendered. (As Is)
 - Match appearance of original
- The original appearance of the photograph is rendered. (As Was)
 - Reverse aging (color fading, physical damage)
- The photographer's goal is rendered. (As Desired)
 - Adjust for errors in exposure or processing
- The original scene is rendered. (As Seen)
 - Remove effects of intermediate

Adapted from: Frey & Reilly. (1999, 2006). *Digital Imaging for Photographic Collections,* pp. 28-29.



As Is (1)



Denver Public Library, Western History Photograph Collections

As Is (2)



Baseball Card Collections, American Memory, Library of Congress.

As Was



Baseball Card Collections, American Memory, Library of Congress.

SI 675 Digitization for Preservation Winter 2011

As Desired



PD-EXP

Ignacio, head chief of Utes , 1870. Denver Public Library, Western History Photograph Collections [X-30722]

As Seen



Ø PO-EXP

Bull Run, Va. View of Bull Run, 1860-65. Library of Congress, Civil War glass plate negative collection. <u>http://hdl.loc.gov/loc.pnp/cwpb.00618</u>



Use Case Studies: Research Context

- Library of Congress / Prints and Photographs
 - United States based
- Criteria
 - Significant use [18 months]
 - Tangible product
 - Primarily LC content used
- Seven deep case studies
 - Orientation toward visual artifacts
 - Influences on decisions to use
 - Content, characteristics, context



PRINTS & PHOTOGRAPHS ONLINE CATALOG (PPOC)



http://www.loc.gov/pictures/

Participants | Projects | Products

- PI. Civil War Stereographs (Brady, Sullivan, Gardner)
- P2. 1870s Turkestan albums
- P3. Lewis Hine children (NCLC)
- P4. Depression era roots music (Ben Shahn)
- P5. Russell Lee biography
- P6. FSA photographers (Jack Delano)
- P7. Baseball history (Bain Collection)

Qualitative Methods of Investigation

- Phase one telephone interviews
- Collection and project analysis
- Phase two in-person interviews
 - In-depth
 - Semi-structured
 - On-site

	Beyon	d Retrieva	al: Phase T	wo Inte	rview Data	
	Recording			Transcript		
Participant	Hours	Minutes	Total Min	Pages	Paragraphs	Words
1	3	27	207	60	618	26,047
2	2	28	148	42	367	19,538
3	2	44	164	50	417	24,356
4	2	46	166	46	404	21,139
5	2	21	141	41	404	18,344
6	2	18	138	44	466	19,064
7	1	24	84	26	288	10,768
TOTAL	17	28	1048	309	2,964	139,256
Mean			150	44	423	19,894

Findings: From Status to "Modes of Seeing"

Discovering [P1, P4, P5, P7]

- Evidence for something never seen or noted
- Community-based validation
- Ultra high resolution and depth
- Tools for manipulation

Storytelling [P1, P2, P3, P5, P6]

- Centerpieces for intellectual puzzles
- Composition and emotion
- Reject cropping; physical manipulation

Landscaping [P2, P4, P5]

- Windows on historical space and time
- Archival context beyond the image
- Technical issues arise in publication only



[Conway, 2011]

Discovering [P4, P7]





Shahn. Jeeter Gentry, Elmer Thompson and Fiddlin' Bill Hensley, Asheville, North Carolina http://hdl.loc.gov/loc.pnp/fsa.8a17159



Unknown. Weldon Wyckoff, Philadelphia AL (baseball), 1913 http://www.loc.gov/pictures/collection/ggbain/item/2001704361/

Storytelling [P3, P6]





Hine: of the spinners in Whitnel Cotton Mfg. Co. N.C., 1908 P3: Cora Lee Griffin, Whitnel, North Carolina. http://hdl.loc.gov/loc.pnp/nclc.01555 Ø PD-GOV **P6 P5 STORYTELLING P3 P4 P1 P2** DISCOVERING **P4 P7**

Delano: In the convict camp in Greene County, Georgia, 1941. http://hdl.loc.gov/loc.pnp/fsa.8c29075

Discovering and Storytelling [P1]

45



Landscaping [P4]



SI 675 Digitization for Preservation Winter 2011

46

Landscaping and Storytelling [P5]



Ø PD-GOV



Lee. Hair tonic salesman, New York City, 1936 http://hdl.loc.gov/loc.pnp/fsa.8a21070



Lee. Cafe. Mogollon, New Mexico http://hdl.loc.gov/loc.pnp/fsa.8b25484



Discovering and Landscaping [P2]





Unknown. Steamboat, Turkestan Album, 1872 http://www.loc.gov/rr/print/coll/287_turkestan.html



Missing Intersection of D + S + L





Implications: Building & Using

- Masters, derivatives, and versions archives
- Full intellectual mapping from source to derivative
- Transparency of visual editing
- One size does not fit all
- Beyond search to navigation





Jack Delano View of the old sea town, Stonington, CT, 1940. http://hdl.loc.gov/loc.pnp/pp.print

Further Research Needed

Prevalence of the various modes

- Survey research with many users
- Extend to other collections
- Explore intersections
- Move toward causality
 - Hypotheses on technical char.
 - Measure impact of digitization
- Remediation's wake
 - What remains of the source?
 - Preserving authenticity of the digital products



Materiality Lost (Sassoon 2004)

- Photographs (and other sources) are technological and cultural objects.
- Those who control the technologies also control the image and its reproduction (just as printers/publishers control the text and its reproduction)
- Three important features of the photographic source:
 - Photograph as object (materiality)
 - Concept of the original (context)
 - Photographic meaning (content)
- Where once materiality and meaning were bound up in a complex, synergistic and symbiotic relationship, the resultant digital object is an ephemeral ghost whose materiality is at best intangible." (p. 199)

Image Literacy

- From image specifications to technology performance
- Image literacy (n): The ability to read, interpret and use generally accepted imaging results, to handle the corresponding performance information, to express ideas and opinions, to make decisions and solve related problems.
- A taxonomy of image performance evaluation is at the heart of image literacy.
- High precision + calibration to adjust for inaccuracies in technologies
- Image literacy lends familiarity with basic technology, characteristics, and evaluation methodologies of digital imaging.

Williams/Burns [2009]. "Preparing for the Image Literate Decade"

References - Conway

- Conway, P. (1999). "Digital Technology Made Simpler." In Preservation of Library & Archival Materials, 3rd Edition. Edited by Sherelyn Ogden. Andover, MA: Northeast Document Conservation Center.
- Conway, P. (2008) "Best Practices for Digitizing Photographs: A Network Analysis of Influences." Proceedings of IS&T's Archiving 2008, Imaging Science & Technology, Berne, June 24-27.
- Conway, P. (2009). "The Image and the Expert User." Proceedings of IS&T's Archiving 2009, Imaging Science & Technology, Arlington, VA, May 4-7, pp. 142-50.
- Conway, P. (2009). "Building Meaning in Digitized Photographs." Proceedings of the Chicago Colloquium on Digital Humanities and Computer Science 1.
- Conway, P. (2010). "Modes of Seeing: Digitized Photographic Archives and the Experienced User." American Archivist 73 (Fall).
- Conway, P. (2011). "Fields of Vision: Toward a New Theory of Visual Literacy for Digitized Archival Photographs." Archivaria 72.

References - Other

- Bolter & Grusin. [1996] "Remediation." Configurations 4. <u>http://muse.jhu.edu/journals/configurations/v004/4.3bolter.html</u>
- Peterson. [2005] Introduction to Basic Measures of a Digital Image for Pictorial Collections. Library of Congress. <u>http://digitalarchive.oclc.org/da/ViewObjectMain.jsp?</u>

fileid=0000070519:000006286369®id=13040

- Mitchell, WJT. [1990] "Representation." in *Critical Terms for Literary Study*, edited by Frank Lentricchia and Thomas McLaughlin. Univ. of Chicago Press.
- Cook, T. [2001] "Archival Science and Postmodernism: New Formulations for Old Concepts," Archival Science 1 (2001): 7.
- Metamorfoze Preservation Imaging Guidelines DRAFT. National Library of the Netherlands. [2007] June. <u>http://www.metamorfoze.nl/publicaties/richtlijnen/guidelinespijune07.pdf</u>
- Frey and Reilly. [1999] Digital Imaging for Photographic Collections: Foundations for Technical Standards.
- Sassoon. [2004] "Photographic Materiality in the Age of Digital Reproduction." In Photographs Objects Histories: On the Materiality of Images, edited by Elizabeth Edwards, pp. 186-202,217-18.
- Williams/Burns. [2009] "Preparing for the Image Literate Decade." In IS&T Archiving 2009 Proceedings.

Thank you!

Paul Conway

Associate Professor School of Information University of Michigan www.si.umich.edu

