Creating Rich and Long-Lasting Digital Image Collections

Franziska Frey
School of Print Media
Rochester Institute of Technology

Setting Goals

- Best managed conversion projects have clear goals
- Tendency to dive right into technology

Imaging Purposes

To protect vulnerable originals from use	67%
To produce printed reproductions	77%
To make collection accessible over the Internet	86%
To include in a collection management system	86%
To document conservation treatment	58%
Other	28%

Goal of Digitization

- The digital image is used as a reference
- The digital image is used for print reproduction
- The digital image represents a "replacement" of the original
- The digital image is used for conservation documentation

Ultimate Goal

- Be able to render the images captured today on tomorrow's systems
  - Quality
  - Digital Preservation

How are Digital Libraries Evaluated?

- Almost no research on implications of image quality
- User interfaces and usability in terms of finding the right image have been evaluated
- Why this gap?
  - Do users know what they can demand in terms of image quality?
  - Visual literacy

artimaging.rit.edu

Experimentation

- Define quality criteria based on objective and subjective metrics
- Develop a method to connect objective, measurable image quality to subjective image quality as perceived by the observers
- Benchmark current quality

Psychophysical Testing: Hard Copy Experiments

Softcopy Experiments

Digitization Strategies Involve Decisions

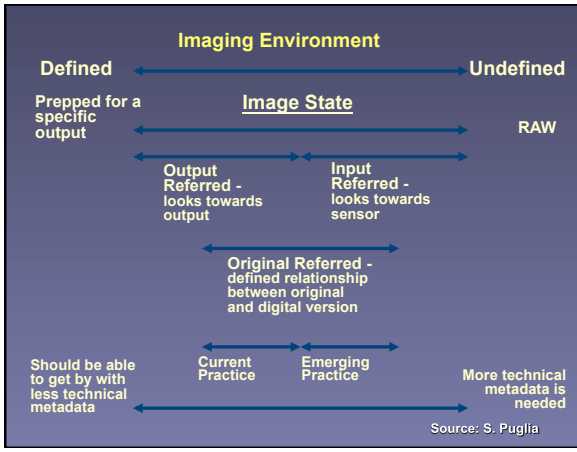
- Key imaging decisions have to be made
  - Intent: what are the significant properties of the source to be copied, then to be preserved?
  - However, these decisions have also to be taken in conventional duplication

Decision Making

- Decisions need to be based on good knowledge of the specific topic
- Do not hesitate to consult with outside sources!

Consequences of Decisions Vis-à-vis Preservation

- Resolution
  - As size increases (e.g., decisions to capture and keep 48 bit, high resolution files), management overhead increases
  - This holds true especially if the storage unit bills per MB or GB per year



Techniques to Preserve Images (*after Chapman*)

- Phase 1—Production
- Phase 2—Appraisal
- Phase 3—Deposit
- Phase 4—Archiving and Preservation
- Phase 5—Discovery and Delivery

Phase 1 — Production

- Imaging does matter
  - Steve Puglia: "We feel that the managed environment needs to be extended beyond the digital repository and forwarded in time to include the digitization process…." (*IS&T Archiving Conference, 2008*)
- Formats do matter
- Documentation does matter

Conventional Photography

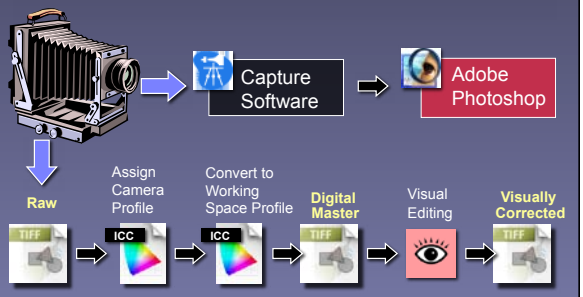
- Behavior of photographic systems was well defined (closed systems)
- System design by scientists (at companies like Kodak) took care of many factors
- Reproductions options were limited and professional photographers considered the output media when capturing the original scene
- After photograph was taken, considerable effort required to make modifications

Digital Photography

- Digital photography systems offer many advantages, primarily in terms of flexibility and convenience
- Increase in flexibility carries a high price
  - "Nature" no longer takes care of some of the variables

updig.org

Workflow Overview



Fine Art Reproduction Workflows

Workflow Process General Function	Specific Workflow Process Steps and Considerations	Additional Steps and Considerations
1. Image capture	Objective targets used	Lighting set up used to illuminate the artwork including polarization
		Camera calibration
		Flat-fielding
2. Proofing and image file preparation	Monitor Calibration	
	Working color space	Screen background used for file viewing
	Viewing environment	Physical image size on the screen
	Sharpening	Image orientation
	Resolution and file size	
3. Image delivery	File format	Image layers for documentation of image processing conducted
	ICC color management	
	Delivery media	
	Guide prints and proofs	
4. Image archiving	Archiving protocol	Proper handling and storage of guide prints
	Metadata	
	Image naming	

Conservation of Documents

- Preparation for scanning and the actual scanning step are well suited for conservation treatment and re-housing
- Improper handling is the greatest cause of damage to library and archive materials

Handling Issues

- Large, clean area to allow for safe handling of materials during digitization
- Plan enough time to prepare documents for scanning

Handling Issues—Training

- Training plan needed for scanning technicians
- "Before-the-job training" and continuous guidance

Standards and Best Practices

- In recent years many fields have found that standards and best practices are necessary to be able to easily exchange information

Standardization Needed For:

- Imaging
- File formats
- Metadata
- Text
- Scientific data
- …

www.idealliance.org

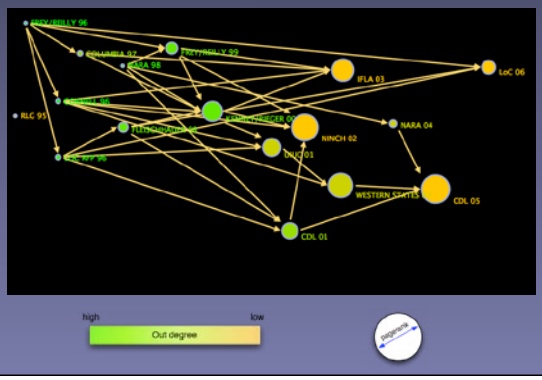
Best Practice

- Best practices can be defined as the most efficient (least amount of effort) and effective (best results) way of accomplishing a task, based on repeatable procedures that have proven themselves over time for large numbers of people.

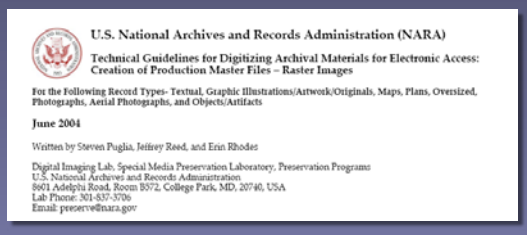
Guideline

- A guideline is any document that aims to streamline particular processes according to a set routine. By definition, following a guideline is never mandatory.

Influence and Accumulation of Knowledge Over Time (Paul Conway)



www.archives.gov/research/arc/techguide-raster-june2004.pdf



U.S. National Archives and Records Administration (NARA)

Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images

For the Following Record Types- Textual, Graphic Illustrations/Artwork/Originals, Maps, Plans, Oversized, Photographs, Aerial Photographs, and Objects/Artifacts

June 2004

Written by Steven Puglia, Jeffrey Reed, and Erin Rhodes

Digital Imaging Lab, Special Media Preservation Laboratory, Preservation Programs
U.S. National Archives and Records Administration
8601 Adelphi Road, Room B372, College Park, MD, 20740, USA
Lab Phone: 301-837-3706
Email: preserve@nara.gov

www.digitizationguidelines.gov

Other Digitization Guidelines

- Library of Congress
  - http://www.loc.gov/rr/print/tp/DgtlMastersSamplSpecsSelctdRcmndFinal7_2004.pdf
  - http://memory.loc.gov/ammem/formats.html
- California Digital Library
  - http://www.cdlib.org/news/pdf/CDLImageStd-2001.pdf

www.digitalpreservation.gov/formats

Basic Approaches to Quality Control

- Subjective visual inspection
- Objective evaluation done in software on digital files

Monitor Viewing Conditions

- ISO 3664—Viewing Conditions for Graphic Technology and Photography (under revision)
  - Defines "environment" in scanning area
  - Ensures consistency of visual quality control steps

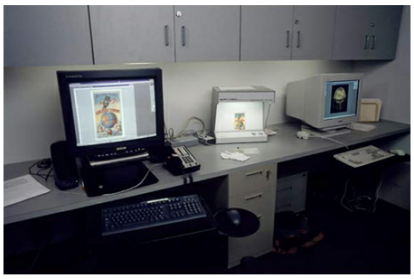
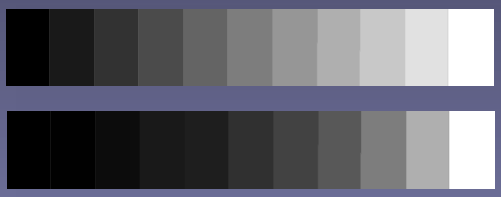


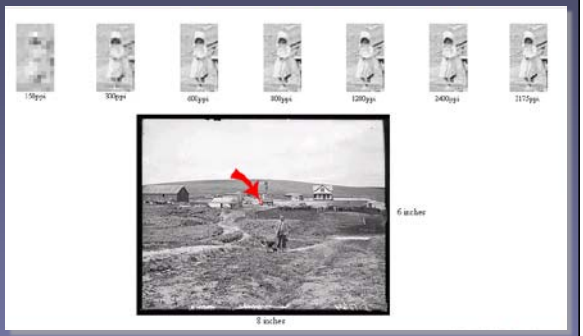
Image processing area including Barco monitor, GTI viewing booth, and DELL workstation with calibrated monitor.

Tone Reproduction



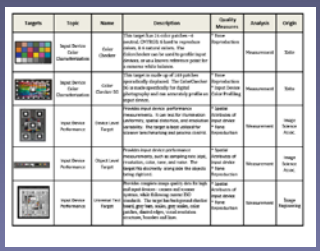
Spatial Resolution

- No magic formula available
- Adapt to the usage
- Keep future developments in mind, e.g., high resolution monitors
  - Watch what is being developed and not what is being used at this time



Objective Targets

- Input targets—output targets

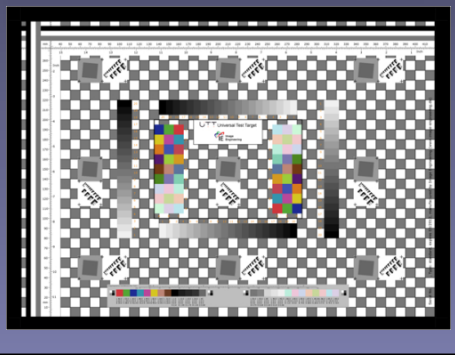


Targets

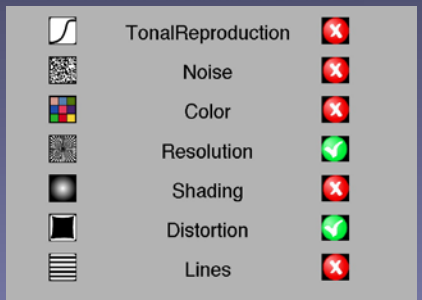
- "Test targets represent known values from an object or in a digital file"
- General uses of test targets:
  - Calibration of a device
  - Characterization & performance testing of a device
  - Process control & color reproduction of a system
  - Resolution & addressability, gray balance and tone reproduction of a device

Test Targets

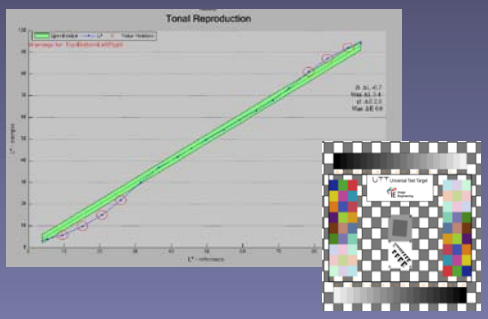
Universal Test Target



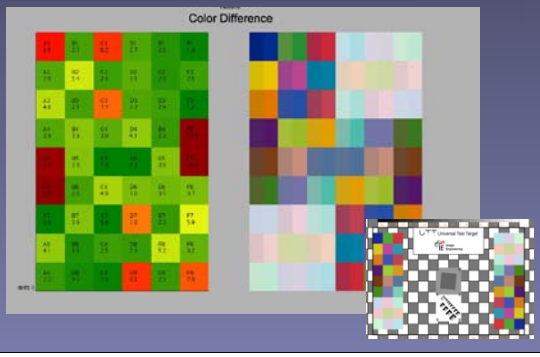
Results from the UTT



Results from the UTT



Results from the UTT



Ultimate Goal

- Be able to render the images captured today on tomorrow's systems
  - Quality
  - Digital Preservation

Responsibility

- "Digital preservation will only happen if organisations and individuals accept responsibility for it.
- Acceptance of responsibility should be explicitly and responsibly declared….."

UNESCO, "Guidelines for the Preservation of Digital Heritage," 2003

Stewardship

- Long-term management of heritage materials (digital objects) through collaboration, throughout all phases of object life cycle.
  - Rights holders
  - Collection managers
  - Repository/preservation staff
  - Centers of expertise (researchers, scientists)
  - Auditors
  - Content users and their communities

Managing Risk

- Security and access control
  - Preventing unauthorized use, tampering or theft
  - Protecting rights holders
- Data obsolescence
  - Media incompatible with players
  - Formats
- Functional obsolescence
  - Formats incompatible with user needs
- Fiscal obsolescence

Summary

- Image quality matters
- Storage is not synonymous with digital preservation, and storage is neither free nor cheap
- Stewardship and digital preservation require active oversight of content, technologies, and user expectations
- Distributed, but shared expertise centers and tools will be essential to managing costs

Thanks to Stephen Chapman
Harvard Law Library

School of Print Media
Rochester Institute of Technology
69 Lomb Memorial Drive
Rochester, NY 14623-5604
Phone: 585-475-2712
fsfpph@rit.edu