OCLC Online Computer Library Center

Microfilm Selection for Digitization

Kelly Barrall,
Manager Quality Assurance
&
Christine Guenther
Sr. Digital Project Manager

OCLC Preservation Service Centers

NEH May 17, 2005

CG

Introduction

NDNP microfilm selection parameters: Technical criteria

- (intellectual criteria)
  (completeness of title run)
- Text quality
- Low reduction ratio (lower than 20x)
- Resolution (5.0 or higher)
- Narrow range of density readings

CG

1. The quality of original text and microfilm capture. Poorly prepared original material, no matter how well microfilmed, yields poor results. Microfilm of bound material may have page curvature, gutter shadows, or out of focus pages that influence digital image quality. Preference in selection should be given to titles on higher quality microfilm.
2. The reduction ratio used when microfilming the original newspaper. This ratio directly influences image quality and OCR results. The lower the reduction ratio (below 20x) the better. (If the reduction ratio is too high to allow scanning at 400 dpi, tests on sample images should be performed to determine if a lower resolution (e.g., 300 dpi) provides acceptable confidence levels in OCR text.)
3. The camera master negative microfilm duplicated for scanning should have resolution test patterns readable at 5.0 or higher. For camera master microfilm without resolution test charts, resolution should be estimated by comparison to film with resolution test charts and original material.
4. Variations in density within images and between exposures. Such variations require adjustment of scanning parameters within a reel. Density readings should follow current standards, but the range should ideally be narrower than the standards allow (e.g. .90-1.20). Best results are obtained from microfilm with variations in density readings of no more than 0.2 within an image and between exposures.

USNP guidelines

- www.loc.gov/preserv/usnpspecs.html (current specifications)

CG

Highlights:

•Focus on text quality

•Quality index of 8.0 or above

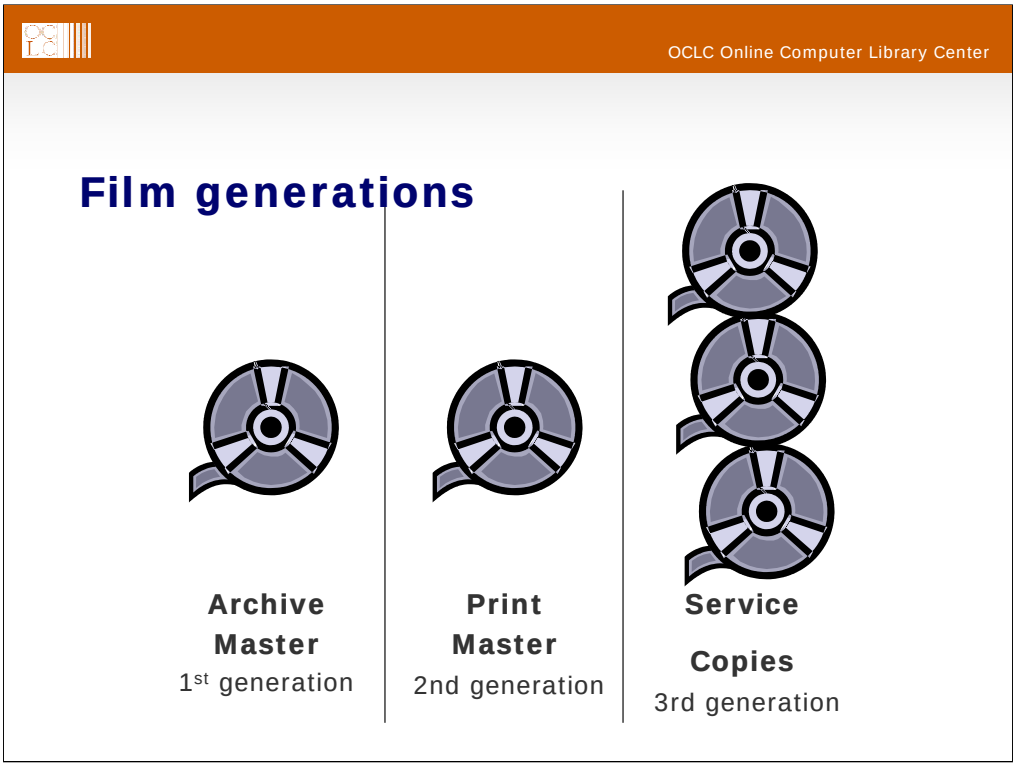
•High resolution camera

•Reduction genrally between 16x and 20x.

•The institution and/or microfilming agency shall produce, in accordance with referenced standards and specifications, for each newspaper title:
• one set of silver negative first-generation preservation master microfilms, which shall be used only to produce:
• one set of second-generation silver negative intermediate print masters, from which will be produced:
• X number of positive service copies

•ISO Test Chart No. 2 shall be filmed immediately preceding the unit target and text portion of each reel and immediately following the text portion of each reel.

•Background densities on the camera negative shall be no lower than .80 and no higher than 1.20, with an aim point of 1.0 for newspapers of average text quality.



KB

Archive Master:

Off the camera

Negative

Should never be touched, goes to storage (NUS, etc)

Print Master (surrogate of the master, by Preservation Standards):

First copy – closest to the resolution of the archive master (loss of resolution with each generation)

Also negative

Used to “Print” or “scan” --- generation of choice for scanning

Some institutions choose to go directly to 2nd generation in positive polarity.

Hence, the 2nd generation to be used for NDNP projects should be the Print master negative, but as an exception positive reels might be included to complete a run.

What's better: Negative or Positive?

- Negative: Usually less noise, better contrast for text
- Positive (usually 3rd generation): Lower resolution, more scratched

Check: 2nd generation positives have emulsion on the top when right-reading

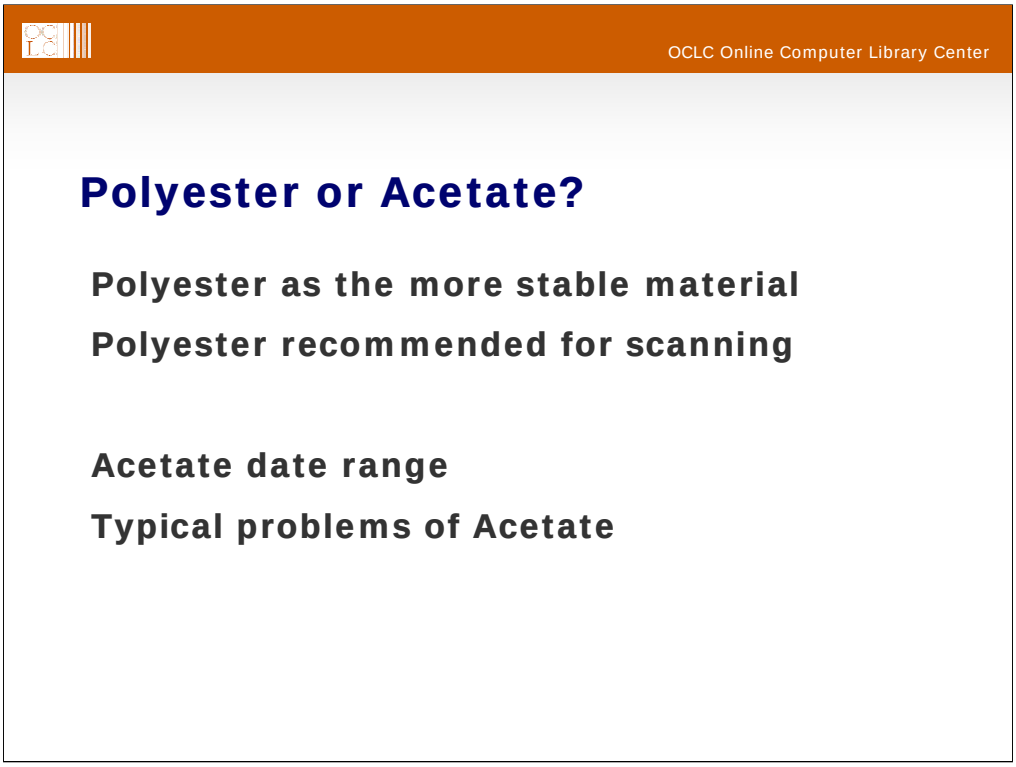
KB

Negative (Inverting the polarity during scanning):

Background noise is less prevalent

Easier to eliminate scratches.

Positive: Scratches and noise are more difficult to eliminate.



KB

Polyester is durable and stable

Recommendation: Duplicate Acetate reels to Polyester before scanning – Acetate is too risky to handle, might tear/break…

Acetate date range 1923-1970 and until 1980.

Acetate problems deterioration, warping, curling, buckling, brittle, redox and blisters appear on the film.

To determine acetate - Hold a wound roll of film up to the light and examine the side of the roll. If it is opaque (light does not pass through), the film is likely acetate. Unwind several inches of film. Acetate film tends to curl. Last resort – attempt to tear a portion of leader. If it tears, it's acetate.

Overview: Key selection criteria

- DENSITY
- RESOLUTION & FOCUS
- REDUCTION RATIO
- PHYSICAL CONDITION

- QUANTITATIVE ANALYSIS

KB

Density

- Definition
- Optimum Density Range
- Density and how it affects focus
- Density variance and how it affects OCR

KB

Density is the measurement of light able to pass through film. The density of microfilm is measured in two ways, d.min and d.max. D.max is measured on the background of the exposed area. D.min is measured on the unexposed area of the film.

The d.max density range should be between .90 and 1.20 with an aim point of 1.0 for material with an average text quality. A lower density range between .90-1.0 is preferred for older pages with faint or broken text.

If the density is too light or too dark, contrast is compromised and focus problems become more apparent. This happens when the density of the background of the material is underexposed and the text becomes faint or similar in density, for example: the curves of a cursive letter S may be lost. Or in the case when it is overexposed the letters start to fill in, for example the inside of a lower case e.

Density variance is measured in two ways, the density throughout the reel from frame to frame and the density on any given frame or page. The density throughout a reel should not vary by more than .20 readings on the densitometer. Uniform density, even illumination on the page or light balance, is just as important as the density range. When the illumination of the page is uneven, text will be lost. Density variance caused by poor light balance is typically a problem in the corners.

ADD WORDING ABOUT SCANNER SETTING . For most microfilm scanners the threshold is set to the entire reel, frames that vary in density may be compromised. This could lead to more attended scanning and higher prices.

Resolution & Focus

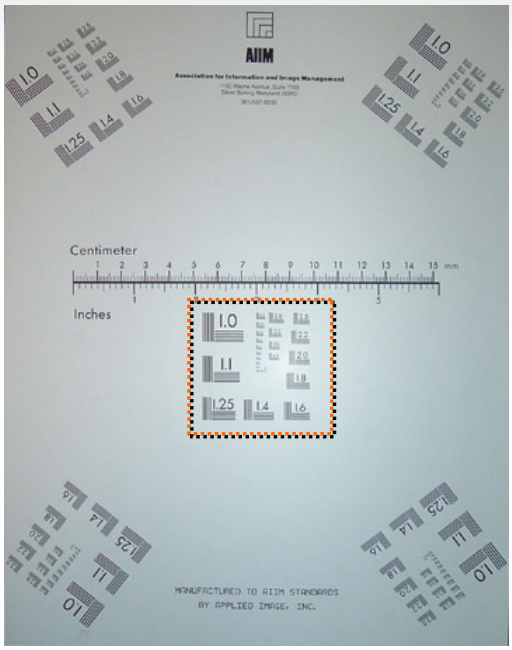
- Definition
- Technical targets / test charts / line pairs
- Poor resolution – fuzzy text – poor OCR

Hint: if no resolution target, use microscope and compare text quality with similar reel that has the test charts

KB

Resolution is a test that measures whether the film captured the fine detail or clarity of the document.

Quality index is measured in low, medium and high quality. Quality index is the relationship between the size of printed text and the patterns on a resolution target. It was developed to assist in determining an appropriate reduction for microfilm. For example: if the measurement of the lower case e, best representing the majority of the text is 1.75mm, the material should be filmed using a reduction that is guaranteed to produce a reading of a 6.3 on the test patterns.



Resolution Test Chart

KB

Reduction Ratio

- Definition
- Recommendation: below 20x
- Formula for missing reduction ratio

KB

Reduction is the relationship between the size of the original document and the size of the microfilmed image. Generally referred to with an “X”. 12X represents the reduction ratio 12:1. Example: 12X or 12:1 means the size of the image on film will be 1/12th the actual size of the original document.

Reduction should always be identified on preservation quality microfilm. In the event it is not, you can use the following formula to determine the reduction. Size of the original divided by the size of the frame = reduction.

Physical condition: 3 reasons for "poor film"

- Poor original
- Poor filming
- Poor condition of the film

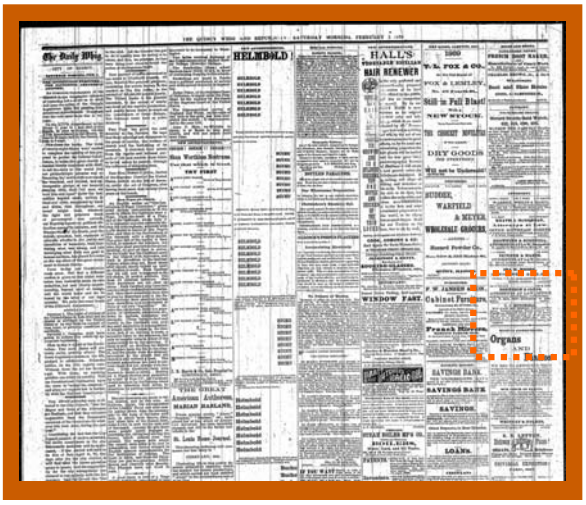
CG

Poor original **(original poorly printed, poor paper, poorly prepared for filming: pages missing, tight binding causing shadow…)**

Poor filming **(focus/resolution, missing technical targets, uneven lighting, …)**

Poor condition of the film **(scratches, weak splices, brittle acetate, warping….)**

Example: Heavily scratched film

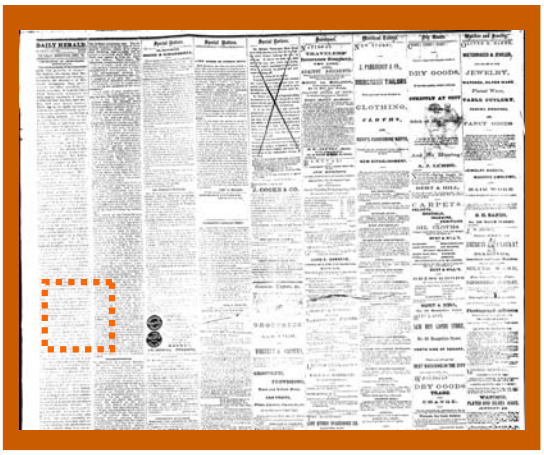


Line that separates columns

Heavy scratches

CG

Example: Uneven lighting – lost text



Lost text

CG

Quantitative Analysis

- Image count:
  - Technical targets
  - Bibliographic targets (incl. Irregularities)
  - Newspaper Pages
  - ISE (intentional second exposures)
  - Duplicates (unintentional)
- Plus: scanning resolution target

CG

"How many images do you really have?"

- multiplier for your per image costs

Scanning resolution

- Pixel array of the scanner is limiting factor: it determines the maximum file size that can be captured.
- Knowing the maximum file size and the page dimensions, the maximum ppi will be calculated.
- If paper is very large, filmed at high reduction ratio, a 400 ppi scan will not be possible / exceed the scanner array.

CG

Examples…

8000 pixel scanner array

Divided by 400 ppi

= inches (20inches) in height (maximum capture size at 400ppi.)

Questions?

Kelly Barrall/Christine Guenther
1-800-773 7222