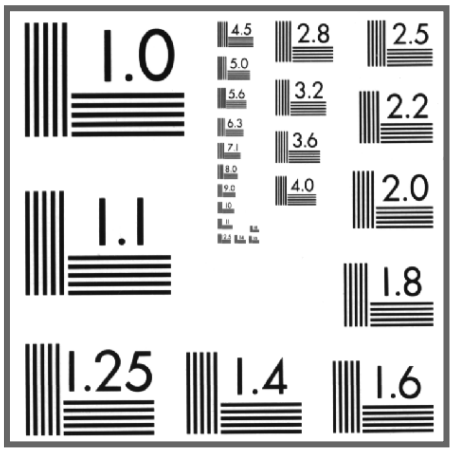
Best Practices and Recommendations for digital images to microfilm

Robert Breslawski, Worldwide ImageLink Equipment and Media Product Manager, Eastman Kodak Company, Graphic Communications, Document Imaging.

Introduction:

The practice of recording images of valued documents and pictures to microfilm started in the early 19th century with Jonathon Dancer creating the first photomicrographs and Rene Prudent Patrice Dagron receiving the first ever patent for a simple microfilm viewer in 1859[1]. This allowed for both a "capture" and "retrieval" mechanism for the practice of document photomicrography we call microfilming. The first practical application of commercial microfilm came about with the 1925 patent of the "Checkograph" filming unit for capturing check images allowing banks to create permanent images of all records and avoid transcription errors. The development of 35 mm cameras in the early 1930s allowed preservation filming to begin in the academic and archival institutions. Microfilm was used extensively in several wars starting with the Franco Prussian war of 1870-1871 and as late as World War II using microfilm to airmail "V-Mail" to avoid shipping heavy paper mail in merchant marine vessels vulnerable to submarine attack. The perceived value of condensing information by optically filming documents, engineering drawings, and cartographic materials to microfilm continued to grow after WWII. Growth continued at a higher rate as active retrieval systems allowed microfilm to provide valuable information capture, presentation and storage in a cost effective manner. With the founding of the National Micrographics Association (NMA) in 1943 practices to manufacture microfilm, supplies and equipment started to be standardized in conjunction with organizations such as the National Bureau of Standards (later to become National Institute of Standards (NIST), National Archives of the US (NARA), Library of Congress (LOC), American National Standards Institute (ANSI), manufacturers such as Eastman Kodak Company, 3M, Dupont, IBM and many others. NMA has since become the Association for Image and Information Management (AIIM) and with the modern nomenclature of Enterprise Content Management it is now called simply AIIM. However, NMA took on the particular task of writing standards and best practices for production of imaged and processed microforms. Of particular note is ANSI/AIIM MS23-2004 Practice for Production, Inspection, and Quality Assurance of First-Generation, Silver Microforms of Documents. This document was first published 1981 by combining two previous documents NMA-MS 110 and NMA MS-104 related to operational procedures (1974) and quality control and inspection (1972). As it relates to processes, procedures and quality control this document is still considered the "bible" of modern micrographics. In the late 1960s and beyond we saw the beginning of Computer Output Microfilm (COM) as a means to capture computer-generated data (typically ASCII) directly on to film rather than first printing to paper and then filming.

Recognizing the growth of the COM industry NMA produced Recommended Practice for Alphanumeric Computer-Output Microforms - Operational Practices for Inspection and Quality Control. These practices and documents all share the common thread of using optical techniques for capture onto microfilm. The critical test element used in measuring quality in all these systems is based on the NIST/NBS 1010A & ANSI/ISO test chart #2 shown below.



The key features of this target are 5 sets of equally spaced lines and spaces that can be filmed at any optical reduction ratio and by simply examining the resultant image one can determine camera/film system. The resolution is determined by multiplying the smallest readable element by the reduction ratio to determine line pairs per mm resolution. NMA/AIIM went on further to define a term called Quality Index, which again refers to this target and allows the user to determine what is "Fair", "Good" or "High" quality.

In summary, there is a good body of standardized reference materials for measuring quality and ensuring good quality control for traditional optical microfilming.

A New Technology is born.

In the late 1980s companies such as Anacomp Inc., Agfa Gevaert and later on Micrographic Technology Corporation (now Global Information Distributors (GID)) introduced all points addressable imaging devises that no longer used fixed formslides and thus the ANSI/AIIM standard for quality control of COM recorders was no longer applicable. With the cooperation of manufacturers and users AIIM developed AIIM MS62-1999 Recommended Practice for COM Recording Systems Having an Internal Electronic Forms Generating

System - Operational Practices for Inspection and Quality Control. This recommended practice describes a digital target which when called out by the software will write an image intended to verify the performance of the imaging device. This target is independent of the particular COM manufacturer as it is written from digital data only. The primary application of these devices at that time was still outputting Enterprise Report Management (ERM) data.

With the advent in the 1980s document scanners document imaging allowed users to convert paper documents to electronic imaging increasing communicability of information. However many commercial, government and academic institutions still prefer/require permanent records to be on microfilm. The need to store this information precipitated the development of many digital media types and systems. Experts quickly realized that a significant portion of this information would be subject to digital obsolescence if not preserved carefully. Continued media format, computer operating system changes and changing program applications have left much information irretrievable.

Another New Technology is born.

In 1995 Kodak introduced the Kodak Digital Science Document Archive Writer Model 4800. The Kodak Document Archive Writer offered rapid conversion of scanned documents to LE 500 (minimum life expectancy 500 years) media called "archive storage media". It was introduced into the standard document filming application as a replacement for traditional optical filming and the first question to be asked was of course how does one quality control this device? A simple answer would be to use a digitally created target such as that described in AIIM MS 62 referenced above but users associate that recommended practice with COM rather than document imaging. Kodak chose to introduce its own test target to test all elements of the writer and included such a target in the software to do so. The software test target image should be considered the same as the printer test image that one is familiar with in office printers. An example is shown below:

The key to this target is that one tests the printer only and not the scanner that may have produced an image to be printed. The printer is part of the system and each element of the system must be tested independently. The Kodak Document Archive Writer and others that have been introduced more recently should all be considered part of a system and independently controlled from the document scanners and retrieval devices associated with capturing, storing and retrieving images.

Imaging System Quality Control.

For any system intending to preserve digital documents and images to microfilm one must consider the following components of the system:

1. Document Scanner Quality Control
2. Film Writer Quality Control

Each element of the system should be critiqued and controlled individually and ultimately as a whole by choosing the right test processes. Unlike traditional optical filming systems where capture and recording were combined in digital systems the capture and recording (writing) steps can be and are separated. Verification of quality and recording of results along each step of the way are critical components of a full imaging system. Problems occur when improper targets are chosen for any part of the system.

Document Scanner Settings and Quality Control

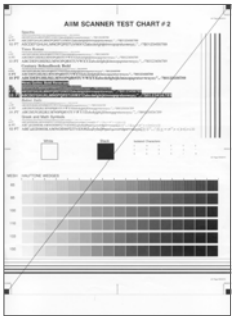
Critical factors in producing quality images from document scanners include the following:

1. Calibration (as required)
2. Brightness/Contrast
3. Resolution settings (DPI)
4. Thresholding (when converting to black and white).
5. Mechanical and operational issues.

Let's talk about each individual step of the factors above.

Calibration:

Calibration sets the parameters so that the image sensors know what is truly white and what is truly black. For color or grayscale scanners this is still important so that proper tone scale or image hue will be captured. The need for calibration varies from scanner to scanner between manufacturers and from model to model from any given manufacturer. For some calibration is factory set and no user intervention is required. For others daily calibration may be needed and in some cases the user interface will report a need for calibration to the operator.



Brightness/Contrast:

Brightness and contrast settings determine how bold or weak character or picture character will be. This setting may be somewhat dependent on the intended use of the digital image. If an image is to be used to create a permanent record on microfilm or for optical character recognition (OCR) one may want to set the brightness so that the characters are a little bolder compared to when the image will be used just for viewing on a monitor. This is because some loss may occur during the film writing and since OCR converts to bitonal during recognition slightly bolder characters can yield better results.

Resolution:

Dots per inch (DPI) are a key factor in scanning documents. DPI basically determines the number of discrete sampling points a scanner uses to capture an image. The more the better for image quality but file size becomes larger. Again the choice of scanning resolution becomes dependent on the application. For common office documents 200 DPI is frequently considered adequate but for documents with fine print or forms featuring descriptors in 4pt type size one may want to consider 300 or 400 dpi or sometimes even 600 dpi. For the purposes of film writing 300 dpi is usually optimum.

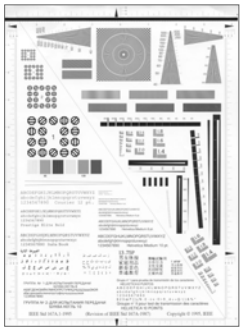
Thresholding:

Scanner sensor arrays are typically considered to be "8 bit" grayscale producing 256 levels of gray when capturing an image. For storage purposes these are frequently converted to Bitonal or binary images to reduce storage space, improve contrast and prepare the images for many film writers. The midpoint of threshold is 128 but typically a setting of about 160 to 180 will produce optimum results. As in the case of brightness control there maybe application specific needs that will affect the set point for thresholding. Slightly bolder characters typically have less loss in reproduction later on especially if converting from digital to analog (film writing) and analog to digital (film scanning).

Mechanical and operational issues:

Mechanical issues include monitoring the performance of the document scanner using standardized or reference test targets on a daily basis and keeping performance records. Image quality can be affected by a number of factors including normal wear of transport mechanisms, added features such as imprinters and ordinary dirt and dust build-up. Cleaning frequency may well depend on the type and number of documents being scanned. Keep in mind that "dirt is the enemy".

Typical document scanner quality control targets are shown below:



These targets maybe used for both quantitative and qualitative control of the document scanners. These targets are to be used for controlling the document scanner and not intended to then be written to film as a quantitative control factor for the film writers. The resolution target on the left is described in ISO 12653-1:2000 Electronic imaging -- Test target for the black-and-white scanning of office documents -- Part 1: Characteristics. This target can be used for quantitative quality control of the writer. The target on the right is described in ANSI/AIIM MS44-1988 (R1993) Recommended Practice for Quality Control of Image Scanners. This target can be used for qualitative control of the document scanner and later on can be written to film for comparison purposes.

Film Writer Quality Control

In 2008 the Property Records Industry Association (PRIA) published a paper called Recording Electronic Images on Roll Microfilm A Best Practices White Paper[2].

This document addresses many practical issues related to recording critical property records onto black and white microfilm for preservation purposes.

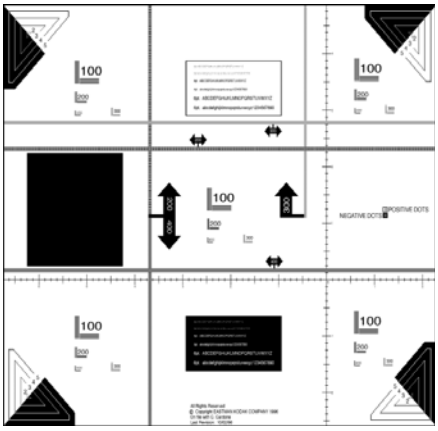
The intent of this section is to describe in detail the specific quality control of film writers. As mentioned previously the critical factors are:

1. Set the writer up at the correct density/contrast levels. (It may be necessary to consult the manufacturer's guidelines.
2. Utilize a "canned" digital target to test the writer only. Do NOT use a scanned in resolution target.
3. Utilize scanned in test targets for comparative purposes relevant to your application.
4. Use standard film processing control as outlined in ANSI/AIIM MS23-1998 Standard Recommended Practice - Production, Inspection, and Quality Assurance of First-Generation, Silver Microforms of Documents.

Test Targets for Writer Quality Control.

As mentioned before it is common for desktop printer manufacturers to supply a digital target or utilize a Windows Printer Test when assessing performance of a printer. The same is true with a film writer as in reality it is not a camera it is really a printer. Several vendors now supply film writers and in most cases they will supply a digital target such as the Kodak Document Archive Writer shown below:

This target is not intended for printing onto paper and then scanned for writing to the writer. Rather it is written directly



from the Quality Control Software built into the application software. Regardless of the manufacturer they will typically supply a similar target.

Those users desiring to utilize a standardized target can also follow the direction of AIIM MS62-1999 Recommended Practice for COM Recording Systems Having an Internal Electronic Forms Generating System - Operational Practices for Inspection and Quality Control. This standard has been adopted by ISO and is now available as ISO 14648-1:2001 - Micrographics -- Quality control of COM that generate images using a single internal display system -- Part 1: Characteristics. A sample is shown below:

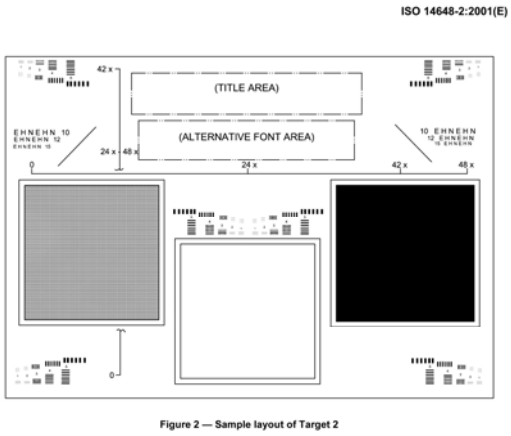


Figure 2 — Sample layout of Target 2

The key elements of either of these test targets are that they are digitally created and recorded without any reduction or scaling ratios applied. The targets test the individual elements of the imaging device and ensure that character fonts can be suitably resolved and that there are no significant mechanical aberrations. While most of these modern film writers have few moving parts they still have film transport mechanisms requiring high precision and consistent operation over time. The test element has for example ladder patterns of varying periodic frequency from one dot on, one dot off all the way up to 4 dots on and 4 dots off. These targets provide confidence that the writer is performing consistently over time.

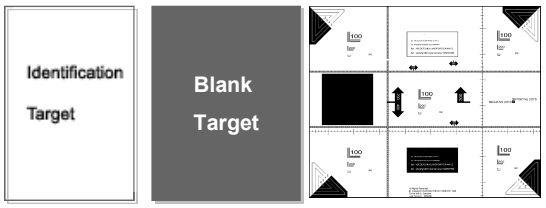
Optical versus Digital Targets:

Many people think they can scan a standard optical resolution target such as ISO Test Chart #2 and then write it out onto the microfilm and "read" the resolution pattern as done traditionally with a microfilm camera. This is not possible due to misalignment of the scanner sensor arrays with the fixed spatial frequency of the test pattern. The misalignment of sensor and test pattern frequency is called digital aliasing[3]. For a discussion and examples of this issue see AIIM TR26 Resolution as It Relates to Photographic and Electronic Imaging.

Quality Control Procedures:

As in ANSI/AIIM MS 23 it is recommended that the writer quality control target be included with each roll of film along with standard targeting such as identification, density, operator declarations. A reference scan of a document scanner target and a production target should also be included. Consistent inclusion of each of these targets on every roll serves to provide a trail of evidence if film quality becomes an issue during inspection or at any later time.

A typical film header including required test targets might look as follows:



Continued on next line

Summary:

The number of industry providers for preserving digital documents and images continues to grow. These provide valuable tools for the permanent preservation of valuable documents as long as adequate care is taken during the capture and recording processes. The steps outlined above and discussed in the related documents are critical to achieving the desired goals.

References:

[1]. http://www.srlf.ucla.edu/exhibit/text/default.htm

[2]. PRIA: Recording Electronic Images on Roll Microfilm A Best Practices White Paper

[3]. AIIM TR26 Resolution as It Relates to Photographic and Electronic Imaging

Contact Information:

Robert Breslawski, Worldwide ImageLink Equipment and Media Product Manager, Eastman Kodak Company, Graphic Communications, Document Imaging. Phone: 585-477-6771 or email: Robert.Breslawski@kodak.com.