How to Evaluate a Scanner

By Derek Jenkins

Introduction:
The main focus of this paper is to provide an introduction to the complex issues of evaluating and choosing a Book or Microfilm scanner for use in a library, conversion bureau or archival work place. There is no right or wrong answer, but being educated allows you to make a more informed decision. This report is primarily aimed for the non-technical user, but does introduce some technical issues that will be covered as simply as possible. For more detailed and technical information on evaluating scanners, IImage Retrieval is available to assist along with experts like Don Williams (ex-Kodak), the staff at NARA and the archival staff at the Library of Congress.

The Key Metrics:
Choosing a scanner to scan books or microfilmed documents can be a daunting task, especially when the industry has dozens of scanners available from a few thousand to over a hundred thousand dollars. All of these systems can scan, but you have to look at if they will meet your project objectives, your budget, and if the scanner is within your staff's capability. In many ways the budget drives the selection, but buying the wrong machine could prove to be a waste of money and cause the project to be both late and with unacceptable image quality.

This paper strives to give you a look into how to properly evaluate a scanner in as unbiased a method as possible. Below is a list of evaluation criteria in order of most to least importance. However, the points are interrelated so selection is commonly a compromise among all the points.

Image Quality: This is the main topic in the paper and the most important factor when buying a scanner. If you can not produce a good image then there's little point in scanning.

Reliability: The first year may be covered by a warranty but frequent failures will force you to purchase a maintenance contract just to be able to use the machine. IIRI has seen cases where customers bought poor quality scanners and ended up spending 50% of their time repairing the scanner and buying extended contracts that cost more money than if they had bought a better quality scanner to begin with.

Performance: If the time it takes from one scan to the next is too long then scanning more than a few pages or frames at a time is not practical. At the other end of the spectrum, there's little point in buying a scanner that can do 200 pages per hour when your staff can only scan 30 pages per hour.

Functionality: Knowing what you will be using the scanner for will help define the functionality required in the system. Things to consider are if you are scanning on demand or are converting in volume as well as if you want a finished product from the scanner or if you will post process it into the final formats.

How to Evaluate a Scanner

By Derek Jenkins

Cost: Cost is a tradeoff between image quality, reliability, performance, and functionality and this paper will try to help you determine the best balance of those metrics. Ultimately though, only you can decide what scanner is best for you.

How to Evaluate a Scanner

By Derek Jenkins

Background:

1. At the basic level, all scans are either color or grayscale. The associated hardware and software defines what image is finally delivered you.

2. All scanners capture an uncompressed data stream that is viewed in two dimensions (length and height).

3. Scanners DO NOT scan in PDF, JPG or TIFF. These are the output formats from the software that is running the scanner.

4. You can only evaluate the quality of a scanner by looking at an uncompressed image. You can not evaluate the quality by looking at an image that has been compressed as the image is degraded during compression.

5. Images from a "similar" model should never be used to evaluate a scanner. Different models use different cameras, electronics, lighting and scanning methods.

Hardware and Image Formats:

After scanning, images normally arrive in the memory of the PC in either 8 bit greyscale or 24 bit color depending on the camera being used. Sometimes images arrive in different color bits, but 8 bit and 24 bit are the most common. From this point on, any output images are a ***derivative*** of the original scanned image as they have gone through some form of processing.

For instance, a bi-tonal image is made by converting an image from its original greyscale or color format to a 1 bit mono-chrome representation. This can be done in many ways, perhaps with a simple static thresholding or with a more sophisticated dynamic threshold, but at the end of the day it is still a software mathematical algorithm that produces the output. This algorithm has nothing to do with the quality of the scan that was originally captured by the scanner.

Compression:

Once in the PC's memory, the captured image is then normally compressed for storage. The most common method for a bi-tonal image is called CCITT-G4 compression. This "compressed" image is then wrapped in a TIFF header that contains basic image information like width, height and bit depth. Hence the common term of TIFF-G4 or sometime Group 4 or even just TIFF.

Some of the typical compression schemes for color and grayscale images are JPEG or JPEG-2000. They take an image, compress its size with a mathematical algorithm and store it on disk. In most cases these formats achieve their savings in size by introducing a slight "loss" in the image quality. The higher the loss ratio is, the smaller the file size but at the expense of increased image degradation.

How to Evaluate a Scanner

By Derek Jenkins

Image Quality Evaluation:
All compressed image formats are images that have been manipulated and degraded to achieve a smaller storage foot print. Therefore, these images are not valid to be used for scanner quality evaluation. **The only true representation of the scanner's image quality is an uncompressed color or greyscale image that is typically stored in a RAW or TIFF format.**

For evaluation, it is important to consider scanning some of your proposed work. However, this still provides a very subjective and limited review. Later projects may have an entirely different type of material or film stock. Testing with one project in mind may not show inherent scanner weaknesses that could be an issue later on. A better method is to scan a known entity like a standard scanner target. A target simulates the basic components of all images to allow evaluation of a scanner's quality. (*See Applied Image at the end of this paper.)* These targets are available in many different layouts and most will do an excellent job of demonstrating the strengths and weaknesses of the scanner. Working with this known entity allows removal of the personal feelings about the documents and brings the evaluation to a more technical level.

Don Williams, an Image Quality and Calibration Consultant and formerly a research imaging scientist in the Imaging Science Division of Eastman Kodak has been working with NARA and the Library of Congress to produce an improved mathematical method to evaluate both film and book scanners. This process started with the SFT targets that have evolved into the SFT II targets and now a new version specifically for book scanners will soon be available. Considering the type of work that you are doing this may be little overkill, but it is the ultimate method for scanner evaluation.

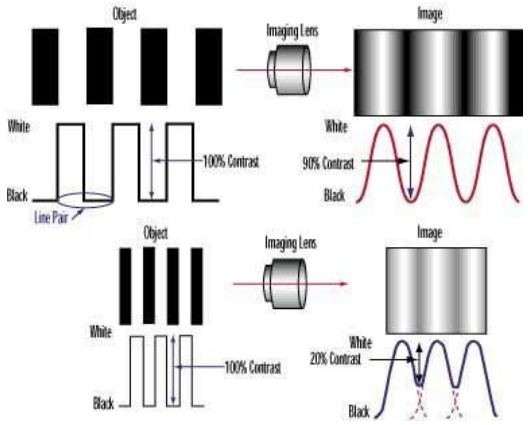
Without going to extremes, scanning any of the commercially available film or paper targets will allow you to evaluate many of the basic functions of a machine. However, you must insure that the images being evaluated were produced on the exact model with the exact lens, lighting and camera that you are considering purchasing. As stated earlier, allowing samples from "something similar" will make the comparison totally invalid.

A scanner consists of multiple components like camera, interface boards, lens, lighting and a transport that, when used together, produce an image. One of the evaluations of a scanning system is to view the sharpness or focus of an image before any enhancements. The finer the viewable details are, the better the scanner. In optical terms this is called the MTF or Modular Transfer Function.

How to Evaluate a Scanner

By Derek Jenkins

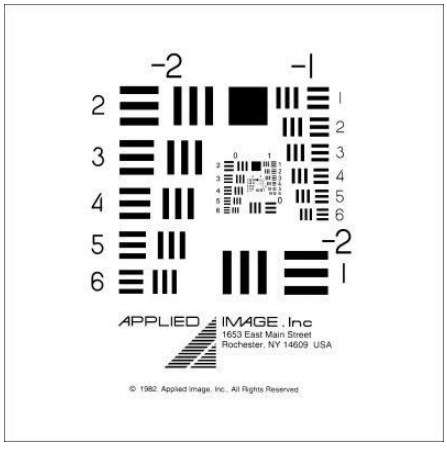
In the example below, taken from the Edmund Optics technical support Web site, an image with of black and white lines is evaluated. While the frequency of lines in low the black to white contrast can be seen. As the frequency of the lines increases the edges turn into a blur and definition is lost.



Quantitatively measuring MTF takes a special program but a simple evaluation can be made by looking at certain parts of a scanned target.

When evaluating images they should all be viewed on the same monitor at a 1:1 ratio or 100% zoom. Other levels of zoom scale the image up or down and do not provide a true representation of the original scan.

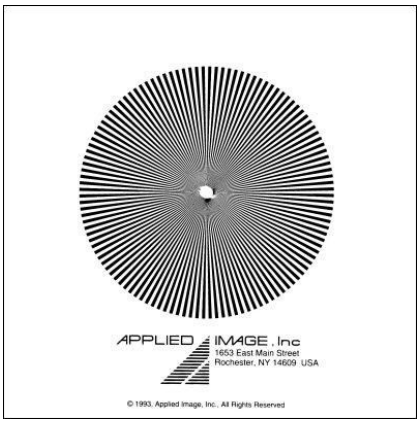
Line Pair per Millimeter (LPPM) Target. These are read by looking for the highest numbered line pair on the target that you can distinctly see the black and white lines. The higher the number the better the image *at the resolution scanned*. Don't compare 200dpi on one machine with 300dpi from another.



Fan Wedge Target. This is a fan shaped series of lines that are thicker at the bottom and thinner at the top. This is used like the LPPM and you look for the highest number where you can see the distinct difference between the black and white lines. The higher quality scanner will have the higher line delineation. In addition to the simple numeric reading you can also compare the straightness of the target lines.

How to Evaluate a Scanner

By Derek Jenkins



Pestrecov Star Sector Target. This is a circle of lines, thicker on the outside and thinner at the center. This is evaluated by looking to see how close to the center you can still see the distinct black & white lines. Again, the better scanner will show distinct lines closer to the center.

Text Size Targets. This is the alphabet in upper and lower case printed in ever smaller type fonts. The evaluation is performed by comparing which lines of text can be read. The best scanner will render the smallest fonts clear and legible. This is very important when scanning books. Footnotes, especially, may require a scanner that can scan smaller fonts than some scanners can provide.

Straight line Test. This test is valid for both book and film scanners. It is evaluation of a diagonal line scanned on a book scanner or filmed and then scanned on a microfilm scanner. The very simple evaluation is to review how straight this line is. A significant wave distortion in the line indicates major issues with the scanner.

Color Separation Test. This test is normally only used on a color book scanner and tells a great deal about the scanner as well as being easy to see. All you need to do is to look at any part of a target with a solid black vertical or horizontal edge. This could be lines or text or anything that has a good transition from black to white. Look at the line at 100% and see if you can see any color shadow or halo. Then zoom to 200%, 300%, 400% or more looking to see if this color shadow is seen. Look on both sides of the line or character. Color halos at the default 100% reveal a low end scanner or a bad array. Average quality scanners will not show a color halo until 400% or so. The best scanners will not show any haloing at all.

Full Area Scan. In this also a test for book scanners, the maximum area size is fully utilized. A target is placed in each corner of the area and one is also placed in the center of the scan area. Evaluate the images for optical distortion. The poorer quality the scanner the more differences will be viewed between each target's results.

Image target evaluation must be done at different resolutions. Some scanners like the i2s Digibook and the nextScan Eclipse have variable optical system. This means that the operator moves the camera and refocuses the lenses to achieve different resolutions. Other scanners like the Book-eye and the Wick & Wilson RS 300 have a fixed optical resolution. This means that the camera and lens are fixed at a given resolution, like 300dpi. Then image is then scaled up or down to achieve the desired resolution. Scaling down can be an acceptable image as it is a derivative of the original but scaling up

How to Evaluate a Scanner

By Derek Jenkins

involves interpolation and adding pixels that were not in the original scan. This is not acceptable if precise image representation is required.

How to Evaluate a Scanner

By Derek Jenkins

Color Evaluation:
There are many more steps involved in performing a true color evaluation that just scanning a target, and the scope of color evaluation is not covered in this paper. If a scanner has the ability to use an ICC profile, and most book scanners can, then color variances can be corrected. While scanning a GretagMacbeth® color target is still useful, true color evaluation is only possible under very controlled circumstances with special calibration tools. This requires the scanner and monitor both be calibrated and an ICC profile generated which ensures that the monitor is reproducing a perfect replica of the scanned original. You would only be able to confirm that the monitor is doing this perfect replication if you view the original document in a special calibrated light box. This detail is beyond what most people can practically do. Entities like NARA or OCLC that have goals for the Pristine Archival scanning buy the equipment and search out the expertise and training to do this type of "full" color evaluation. As long as the scanner produces decent eye discernable color and has the ability to use ICC profiles then this generally is all that you should need in most cases.

Reliability:
If the scanner is constantly failing and needs frequent maintenance then it will not be available for the task it was intended. These ongoing repair costs will effectively increase the purchase price of the scanner so what appeared a good buy suddenly becomes a burden to the operational budget.

Performance:
A manufacturer's documentation may state that is takes one second to scan an image or that it scans at a rate of 300 pages per min. Check carefully what they mean by that.

> **Book Scanners:**
> There are several elements to the action of scanning. A scanner may take just 1 second to do the scan but it also may take additional time to set the exposure, correct the curvature and save the image to disk. The time that should be looked at is not the scan time but the full cycle time that takes all these steps into account. The cycle time is the time from the start of a scan to the time you can start the next scan, **repeatedly**. Depending on your goals (the number of materials you have to scan and the amount of time you have to complete a project), this cycle time could be too long and then scanning more than a few pages at a time is not practical.
>
> Looking at this from another angle, if you are scanning a book using a glass plate, it's important to consider how fast can the operator release the glass, turn the page, close the glass and then press the scan button. If that time is less than the scanner's save time then the operator we will be waiting for the scanner. If the operator's actions take longer than the save time then the scanner is waiting for the operator.

How to Evaluate a Scanner

By Derek Jenkins

Performance is not everything, but it is important. Ideally, IIRI would prefer that the scanner is waiting on the operator to finish turning the page than the operator waiting for the scanner to finish saving. Running at 12 scans per min or about 24 pages per min is about as fast as an operator can work for any extended time.

Microfilm Scanners:
These scanners are normally rated in pages per minute but the size of pages changes depending on where you are in the world. For example, in the US they use 8.5 x 11 inches in Europe they use A4 which is 210mm x 297mm or 8.27 x 11.69 inches. So a scanner in the US rated at 300 pages is the same as a European one rated at 310 pages.

But now let's consider the full system. You may be able to scan at 300ppm but you need to look at if you can process, compress and output those images at the same rate. If you can't then the scanner has to be slowed down. If you have to slow it down then you don't need a 300ppm scanner. This is especially evident when scanning newspapers as the images are very large and the normally stored as uncompressed (RAW) TIFF images. This means that the system is trying to write to disk between 50 and 70 megabytes of data per minute which is the upper limit of the average disk drive. So the scanner has to slow down to a rate that can be saved to disk and practically you get between 40 and 50ppm.

Now let's look at film handling as part of the equation. To set a base line a 100ft roll of 24x film scanned at 200dpi takes about 7 minutes to scan. Then the roll has to be rewound, dismounted and the next roll loaded, tested, index data entered and then the scanner started. Assuming that this takes 3 minutes then we have a 10 min per roll production time, **for the mythical perfect roll of film** or 6 rolls per hour. This is 42 minutes of scanning and 18 minutes of "scanner management".

From direct experience, one operator can run 2 scanners and keep them going; but it is hard work.

Now let's consider the same production environment with a 500ppm scanner. The roll takes about 5 minutes to scan. You still have the 3 minutes of overhead making the cycle time 8 minutes. This works out to be 7.5 rolls per hour or 37 minutes scanning time and 23 minutes management.

Technically this will work with one scanner but two would be nearly impossible. And as soon as the rhythm of scan/mount/scan/mount is disrupted then a scanner goes idle and the productivity is reduced. Also consider the practical issues of interruptions for telephone, bathroom, lunch, etc. and so a faster scanner is not necessarily better.

How to Evaluate a Scanner

By Derek Jenkins

How to Evaluate a Scanner

By Derek Jenkins

Functionality Questions to Ask:

Book Scanning

1. If the majority of your books and documents are 8.5x11 inches or 11x17 inches do you really need to have a scanner that will scan the occasional 22x34 inch map?

2. Do you need a book cradle or is most of your work flat?

3. Do you need a glass platen to hold the work?

4. Do you need a vacuum table?

5. Do you want the scanner to just scan and then you will post process all images later with other software like Photoshop®, or BookRestorer® or some other software program?

6. Do you want the scanner to deliver finished work at the expense of speed?

7. Do you want the scanner to be portable?

8. Are you scanning for pristine archival work?

9. Do you have operators that are capable of operating a complex optical system?

Microfilm/fiche Scanning

1. Are you doing volume production scanning or are you doing on demand scanning?

2. Will a trained operator be using the scanner or will it be open to the public?

3. Do you want to post process the images as they are scanned or will you post process them in using other tools later?

4. Are you scanning for archival work or is this for simple monochrome retrieval?

5. Do you want the scanner very portable?

6. Do you have the infrastructure in place to manage a complete production system?

Cost:

This is a complex topic as on one hand, you need a scanner to do the work required while on the other hand, you may have budget/funding constraints. IIRI sees two steps to evaluating the cost of a system: The first is focusing in on the type and volume of work you

How to Evaluate a Scanner

By Derek Jenkins

are going to do. The second is the capability of the system. Common to both is the reliability, training and support offered by the vendor.

Below are several questions that are common and different to book and microform scanners.

Common Questions:

1. What is the warranty period?

2. What is the cost of an extended warranty?

3. Is support onsite or telephonic?

4. Does the system being purchase need a PC or monitor or is it included?

5. Is freight included in the quote or is it billed later? Note: a lot of the equipment is made overseas so freight can be high.

6. Is the scanner's capability being exaggerated so that you will buy it? Should you actually be looking at that next model up to meet your goals even if it does cost more and may not be in your current budget?

7. If the price seems too good to be true is this the end of a model's production life? Will it be replaced with an improved version soon?

8. Is the model a new model with no proven track record?

9. What happens if the scanner is delivered and it is DOA?

10. Is the scanner a true optical resolution scanner or does it scale up or down to achieve the requested resolution.

Book Scanners

1. Does the scanner have a book cradle?

2. Is it mechanical or electric?

3. Does the system have a glass?

4. How does a color scanner scan grayscale? Does the scanner do grayscale at a native level or does it generate the grayscale from a color scan?

How to Evaluate a Scanner

By Derek Jenkins

5. Does the system have operator exchangeable cameras for higher (or lower) resolutions?

Microfilm/fiche Scanners

1. Does the film scanner scan both 16mm and 35mm film?
2. Does the fiche scanner scan mixed mode, 16mm & 35mm, fiche?
3. Can it handle 1000ft reels?
4. How fast is the rewind speed?
5. How long does it take to setup?
6. What level of training is required to operate the scanner?
7. What is the cycle time when scanning jacket fiche? Cycle time is the time from the start of one scan to the start of the next scan.
8. How does the scanner handle 96 frame COM? How does it handle 297 frame COM?
9. How does it handle the different resolutions & reduction ratios?
10. Is the scanner a true optical resolution scanner?

Any purchase is a compromise of quality, performance, cost and functionality; before you buy, think of what compromises you are prepared to take for the project that you are taking on.

Suggestions:

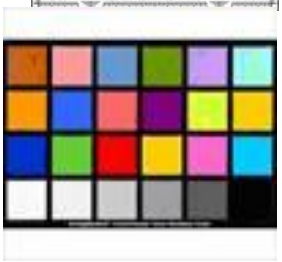
- Verify the information given in the advertising literature or provided by the salesperson.
- Check all the numbers in a real world setting like a trade show.
- Sit down at the scanner and use it like you would if you had purchased it.

How to Evaluate a Scanner

By Derek Jenkins

- Be wary of statements like “the scanner was damaged in shipment” if the performance is not as advertised.
- Ask for multiple references and call them all.
- Try and find a user with the scanner that was not a reference since a company is not likely to give out the name of a site that had problems.
- Ask the references probing questions to explore and compare how you are planning to use the scanner.
- Ask the references about the scanner’s reliability and the post sales support.
- Ask the salesperson for samples to be done on a machine that is exactly like the one you are considering. If possible, be present when the samples are done. Maybe try to do them at a trade show.
- **For a book scanner**
  - Ask for targets like the Kodak TL5003 “Old man target” to be scanned in both Color and grayscale and at different resolutions and save them as uncompressed TIFF images.

  - Ask for a scan of the GreTag color patch target at different resolutions and saved as uncompressed TIFF.



  - Scan an area the full size that can be handled by the system. Scan it multiple times moving a target to each corner.
- **For a Microfilm scanner**
  - Obtain a standard target like the Kodak Old Man Target. Have it filmed 2400 times on a silver negative and processed by a reputable company.
  - Send a 20ft segment to the scanner manufacturer and ask them to scan it at 200, 300 and 400dpi in grayscale and monochrome.
  - If possible be present to see the scanning and then take the images away with you.
  - Scan a clear strip of 16mm film and also 35mm film and look at the exposure from one side of the film to the other.

How to Evaluate a Scanner

By Derek Jenkins

- If a document is 8.5x11 at 200dpi then the image should be very close to 1700 pixels x 2200 pixels. Also the ratio between the width and the height should be very close to 1.29 (2200/1700=1.294).

All this information will let you know the capabilities and limitations of the scanner.

Summary:
Scanner and image evaluation is a broad topic; this paper has tried to stay with the core points that can be looked at by most people. If you are interested in learning more, other topics to look at are lighting, color temperatures, optical systems and image distortion. IIRI has tried to be as generic as possible and not slanted in any direction; if you have further inquiries, we are willing to spend time and discuss this with you to help you produce good evaluation criteria.

How to Evaluate a Scanner

By Derek Jenkins

Company Background:
IImage Retrieval, Inc. (IIRI) was founded in 1990 by CEO and President, Derek Jenkins and VP, Greg English, in Dallas, Texas. Early on, IIRI aligned itself with SunRise Imaging, a microfilm scanner manufacturer, and participated in the development of new scanners. SunRise quickly became the leader in the film scanning industry with IImage Retrieval as its largest and most experienced reseller and repair facility worldwide.

In 1995, IIRI opened an office in London, England which provides the same services as the US office, including microfilm, rare book and special collections scanning. In 2001, driven by customer requests, Derek expanded IImage Retrieval Inc.'s product line to include the rare book scanners produced by i2s including the DigiBook and eScan lines. These lines of high resolution, color or monochrome, planetary scanners expanded the film scanning capabilities IIRI currently had. In particular, the DigiBook line of high resolution scanners were initially designed for the archival industry but their flexible design easily allowed them to scan large flat materials such as maps, drawings, and newspapers in addition to archival projects.

In 2002, Derek and Greg founded a new company, NextScan, with the purpose of designing and developing a new generations of more technologically advanced microfilm scanners. When the new NextScan microfilm scanners hit the market, they quickly outperformed any other scanner, proving themselves to be years ahead of the competition. Today, NextScan is still a leader in microfilm scanning.

From the company office in a northern suburb of Dallas, TX, IImage Retrieval has grown from a small service bureau and software development company to the premier reseller of book, archival and microfilm scanner equipment in the US. Providing sales, installation, training, service and parts distribution for i2s and NextScan product lines as well custom software development for the imaging industry, IImage Retrieval is one of the premiere scanning companies worldwide.

Reference:
Applied Image, Inc www.AppliedImage.com Luke Hobson (585)482-0300
Image Calibration Targets

Image Science Associates, LLC Don Williams (315)573-4782
Image Quality and Color Calibration Consultant
Don-Williams@rochester.rr.com

Edmund Optics www.edmundoptics.com
Optical equipment supplier