# How to Evaluate a Scanner

By Derek Jenkins

**Introduction:**
This paper is an add-on to my other white paper about scanner evaluation. In this document I will be a little more focused on the interpretation of the truth behind the numbers quoted or inferred in today's advertising and actual evaluation of images. I am not saying anyone is directly misrepresenting data, but I think the average inexperienced person assumes a lot and that can get them in trouble.

To most people the individual statements like resolution, camera size, scan speed, bit depth and output types, all blur into one, but they are individual facts that work together to make a scanner. With the evaluation of each fact and image more information about the scanner is uncovered. This paper is written as a guide to the evaluation of a scanner and its images.

**The Mythical Scanner:**
To make this easy I am going to use the specifications from a new scanner that just came on the market. It does not matter who made it or what the model is as this is an exercise for you to learn from. For simplicity I am going to call this the DTS 1000C from "Desk Top Scanners Inc". And the advertising material says:-

| | |
|---|---|
| Type: | Desktop color book scanner |
| Light Source: | 2 x 28W fluorescent lamps |
| Size: | 18x24 inches |
| Document Thickness: | 2 inches |
| Image Sensor: | 10.5 Megapixel Area Array |
| Resolution: | 200 / 240 / 300 / 400 / 600dpi |
| Scan Mode: | 24bit color, 8bit grey scale, 1bit Black & white |
| Scan Speed: | 9.9 seconds |
| Auto Book Correction: | Curve text line correction |
| Interface: | USB 2.0 with cable less than 6ft. |
| Recommended PC: | Intel Core 2 Duo 2.33Ghz or faster<br>2GB memory and 256meg Video card |
| Accessories: | Glass plate, Angle book holder & manual switch |

Let us look at each statement and dissect what it says and does not say.

# How to Evaluate a Scanner

By Derek Jenkins

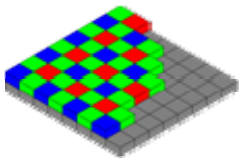
**Type: Desktop color book scanner for face up originals**
This does not tell us anything. It's a scanner, it is for books and the books are scanned face up.

**Light Source: 2 x 28W fluorescent lamps**
This is not important. All it tells us is that scanner has integral lights. Most scanners do. Some do not. It makes no difference in the scanners performance.

**Size: 5.5x8.5 to 18x24 inches**
This is an important fact. This is the maximum and minimum areas that can be scanned. I don't know why we have a minimum as software cropping should take care of this. This should be investigated further but should not affect scanner operation.

**Document Thickness: 2 inches**
This tells us the depth of field that the camera can handle. Or in other terms the camera can only focus over a range of 2 inches. This is enough for most books but what if you scan a tight book that when open causes a big page curve. This could be more than 2 inches.

**Sensor Size: 10.5 Megapixels Area Array**



This tells us that the scanner is using a rectangular area array the same as a Sony or Canon camera. It most likely has a Bayer filtered array with 3.5 Meg of Red, Green and Blue pixels. With this type of CCD the software/hardware in the camera mixing the RGB pixels together to form a single color point. Being a 10.5 mega pixel area array I estimate that is will scanning about 2592x3888 pixels

**Resolution: 200 / 240 / 300 / 400 / 600dpi**
Now we get to some of the really grey areas. What is resolution? Do they mean scan resolution or do they mean output resolution. Is it always true optical resolution?

This then goes into the Modular Transfer Function or MTF of the scanner. In simple terms this is how the scanner (Camera/Lens/Electronics) resolves (sees) a dot pattern. If you CAN NOT can see the number of transitions from black to white in 1 inch then you are not able to scan that resolution.

# How to Evaluate a Scanner

By Derek Jenkins

In the literature for the Mythical scanner it says ***"Clean, crisp resolution. You will have 24-bit color, 8-bit grayscale, 1-bit B&W imaging for superb reproduction of complex originals -- with output resolution up to 600 dpi."*** Let us break this down and analyze what this for us.

***"Clean, crisp resolution"*** Well this is marketing words and has no value.

***"You will have 24-bit color"*** That is correct as the scanner has a color camera.

***"8-bit grayscale"*** Most color systems can translate to grayscale. Some can even scan in native grayscale.

***"1-bit B&W imaging"*** This is produced by taking the Color (or maybe GS) image and thresholding it to Black and White. Depending on the algorithm used, this could be a fixed threshold or a dynamic threshold. Depending on the capability of the software, this can do a good job or a bad one. The test really comes when a color image with a picture is involved and how well the algorithm converts the image.

***"For superb reproduction of complex originals"*** Again another marketing statements with no facts, unless we want to assume that the scanner will provide for superb reproduction of complex originals in Black & White. I would not assume this. I would test it.

***"With output resolution up to 600 dpi"*** This is the only part that is important. The key word is **OUTPUT**. It does not say that it is scanning at 600dpi it says that it can **output** 600dpi. So unless it scans at 600dpi the only way it can possibly get that high is to scaling from its TRUE OPTICAL RESOLUTION or scan resolution.

Let us explore this further. From above we estimated that the camera is 2592x3888 pixels or about 10.5meg in size. I am going to round this to 2700x3600 pixels to make the calculations easier. If the scanner scans the whole area in one go then it scans the 18 inches by 24 inches at 150dpi.

If I wanted to output a 300dpi image of the 18x24 inch scanning area I would have to scale the image **UP** by a factor of two. If I wanted 600dpi then I would have to scale it four times.

# How to Evaluate a Scanner

By Derek Jenkins

Generally scaling up any significant amount in both X and Y dimensions does no good whatsoever. All you have is a big image, taking up more disk space. You do not have any more details. You cannot read finer print. You have gained nothing.

On the assumption that the scanner scans at only 150dpi then any output above 150dpi is a scaled output. But the image quality is still only that of the original 150dpi image. ** See the scan examples later in quality evaluation.

We must consider that the manufacturer may zoom in on the images to increase the resolution. This is valid and normal, but then you are only scanning a smaller area at the higher resolution. If I zoom to 300dpi then I can scan only 9x12 inches or if I zoom to 600dpi then I can only scan 4.5x6 inches.

With zooming you can trade off scan area with resolution. The higher the resolution, the smaller the scan area. The lower the resolution, the large the scan area.

**Scan Mode: 24bit color, 8bit grey scale, 1bit Black & white**

This tells us what the scanner can do. It can output 24bit color, 8 bit grayscale and 1 bit black & white. Well, if this is an area array with a Bayer filter then it always scans in color and converts to GS or black and white.

**Scan Speed: 9.9 seconds 17x22 inch color 300dpi**

In the literature it says that a "17x22 inch 300dpi area can be scanned in 9.9 seconds". This quantifies things a little, but leaves out a lot of information.

- What is the scan speed at 200dpi and 600dpi? Is more or less?
- What is the scan speed if I output Black & White?
- What is the speed if I de-skew the page or run curve correction?
- Is this the speed of the scan only?
- What about the time it takes to write to disk?

**Auto Book Correction: Curve text line correction**

This is a nice feature. Most book scanners have it. How well does it work? Does it do better on a single or double page spread? How does it deal with the book cover and the feathering of the pages when the book is open?

# How to Evaluate a Scanner

By Derek Jenkins

**Interface: USB 2.0 with cable less than 6ft.**

A simple interface. Not much to be said except that it has limits. 400 Megabits / sec is 50 Megabytes / second on the USB bus. Practical throughput may be as low as 18 MB / sec. This seems fast but what if you are scanning the full 18x24 area at 600 dpi color. This is 18x24x600x600x3 pixels or 455MB of image, if the raw image comes from the camera electronics. It would take about 25 seconds to transfer the image to the PC and then at least 11 seconds to write it to disk. This does not include any overhead to de-skew or curve correct. That is 36 seconds.

From above it states that it takes 9.9 seconds for a 300dpi scan. If that is the "Cycle time" from scan to scan then we know that you can do 6 scan / min. If it is only the SCAN time and not the SAVE time then it will be slower.

From experience 6 scan / min is slow and steady. 12 scan / min is hard work on the operator but can be done. Practical speeds sit between the two numbers.

**PC Required: Intel Core 2 Duo 2.33Ghz or faster**
**2GB memory and 256meg Video card**

Nothing special here. A very basic PC but it could be considered as an additional $1000 expense to make the system operational.

**Accessories: Glass plate, Angle book holder & manual switch**

Again, nothing special. It is nice they have these but don't forget that they are additional expenses to make you system operational.

**Quality Evaluation:**

We have read the advertising literature. We have asked the sales person for a quote. We have insured that ALL the items were included:-

1. Hardware
2. Software
3. Accessories
4. PC
5. Shipping
6. Installation & training
7. Warranty and Support

But it would be a good idea to try the scanner in a similar environment to what we will be using it. Your peers are a great source but so is trying the scanner your self. A trade show is a great place to test equipment. Many manufacturers don't

# How to Evaluate a Scanner

By Derek Jenkins

like this as it puts them on the spot. I have even heard of one that declines doing samples because “the equipment got damaged on the way to the show”. Yes, this does happen, but not at every show.

Have a test plan. Have test targets and have a USB stick/CD/DVD to transfer the images onto.

**The Test Plan.**

1. Scan an 8.5x11 inch test target and save it as a RAW TIFF image.
2. Scan it at 200, 300, 400 & 600 dpi.
3. Scan it in GS & Color.
    a. Time how long it take between one scan and the next, especially at 200 & 300 dpi.
4. Scan the test target along with a color target to make an 11x17 inch area.
    a. If you don’t have a large target you can simulate one by using multiple targets or moving the one target you have. The goal is to simulate scanning the full scan area by placing the target in each corner of the scan area.
5. Scan some targets and save them in high quality JPEG. Don’t forget to time the scan to scan for cycle time.
6. Now scan an actual book or document from your collection. Scan at 200 or 300 dpi, scan in Color, GS or BW. Turn on the recommended settings and time yourself over a group of pages. This should be as close to real life as you can make it.

**Evaluation – Read the other White Paper**

The first white paper I wrote is specifically targeted at “How to Evaluate a Scanner”. This document is a follow up, more user friendly version, showing the “How to” of that evaluation. Most of this can be done with the scans from above and a sharp eye.

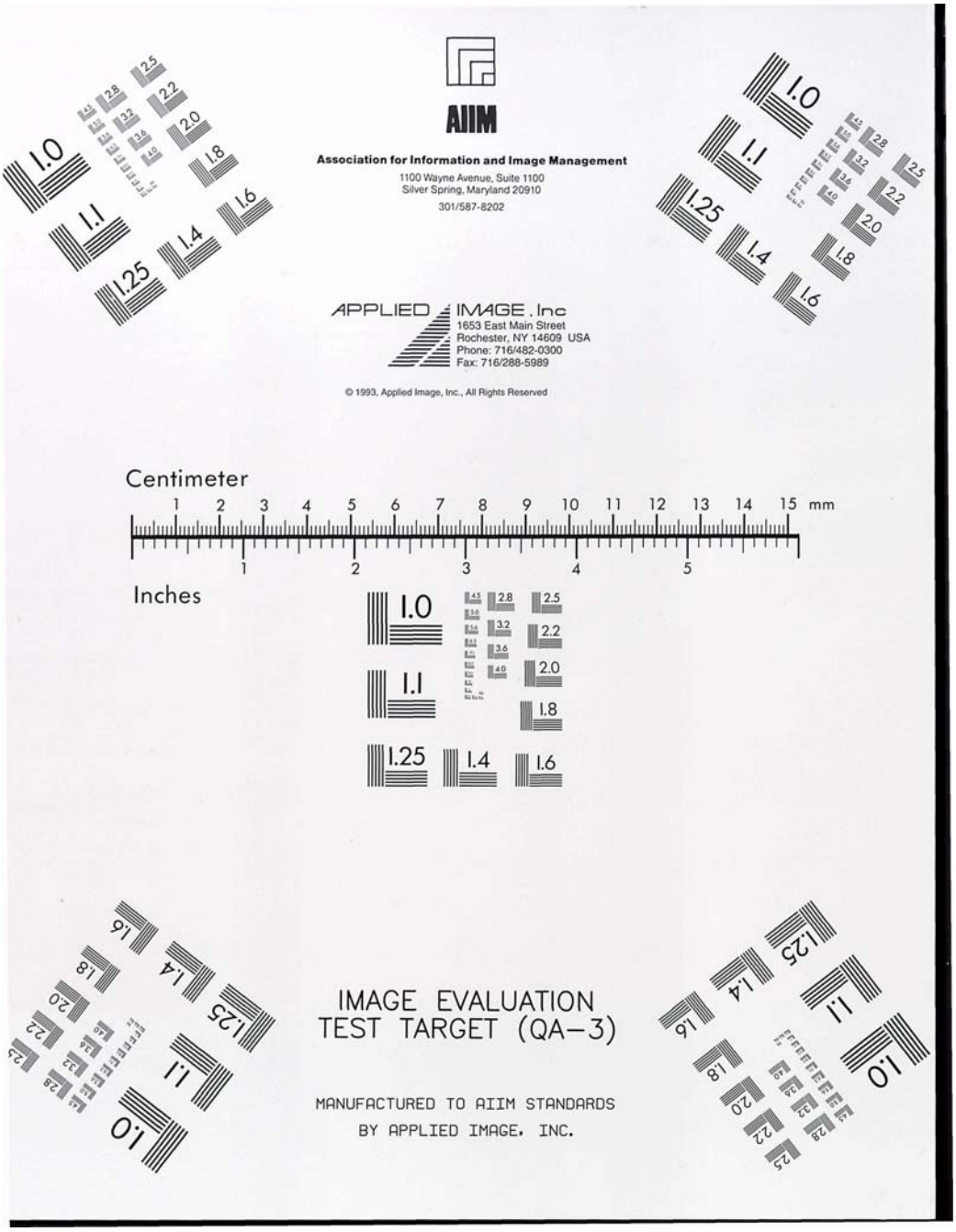
To help me, I asked a reseller of the “Mythical Scanner” to scan some targets for evaluation. I sent him the targets and a list of the scans I wanted. Try to be present when the scans are done. In this case I was not able to be there. Anyway, my evaluation is based on the images he sent back.

I will break down the evaluation into Speed, Flexibility and Quality. Speed and Flexibility are subjective to your needs and budget. Image Quality should not be sacrificed as you are only going to scan this once. Therefore, I will walk through a simple quality review to show how I would look at the scanner.

# How to Evaluate a Scanner

By Derek Jenkins

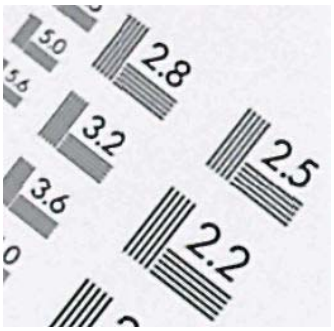
This target was scanned at 300, 400 & 600 dpi. I will review these three resolutions and compare them to another scanner. As the images are large I will cut a small segment to demonstrate what I am looking at. At first glance the image looks OK but when dissected flaws start showing up.

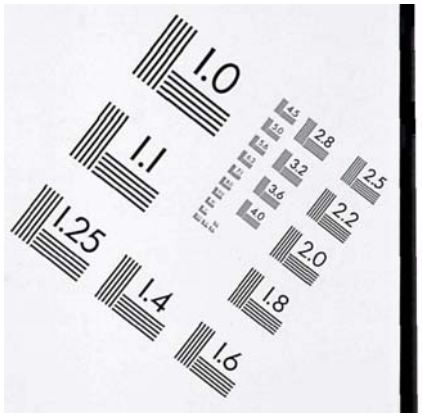


# How to Evaluate a Scanner

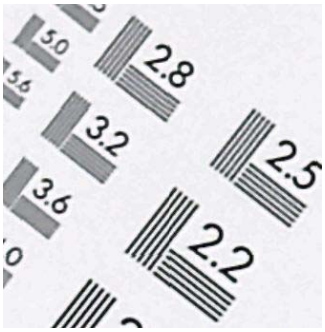
By Derek Jenkins

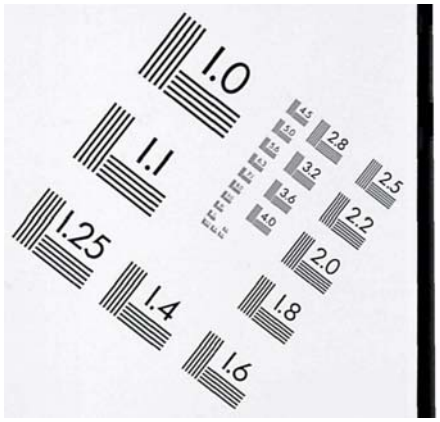
300dpi test Scan. Read 2.5 maybe 2.8



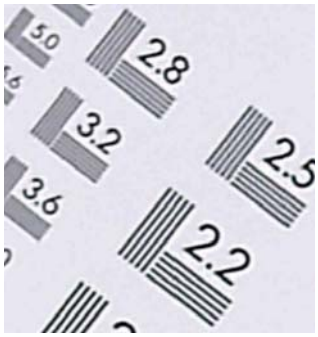



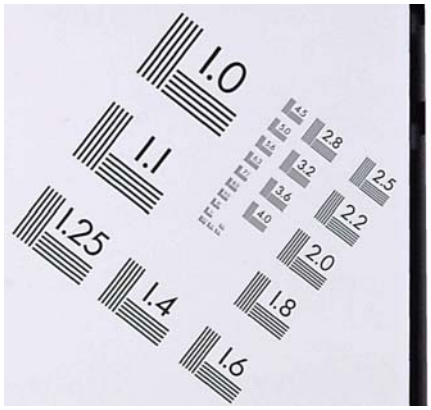

400dpi test Scan. Read 2.5 maybe 2.8




600dpi test Scan. Read 2.8 maybe 2.8



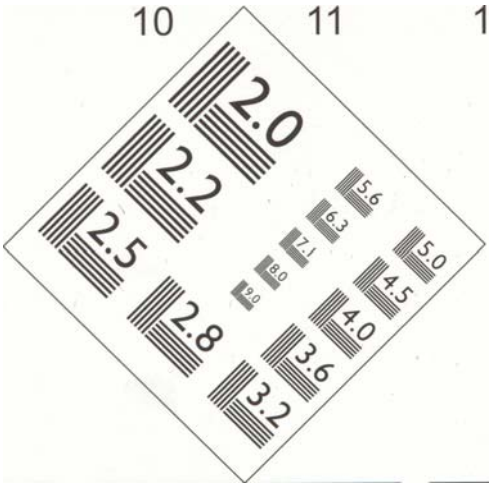



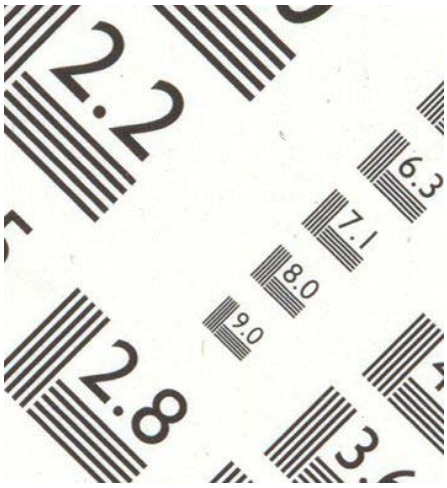
# How to Evaluate a Scanner

By Derek Jenkins

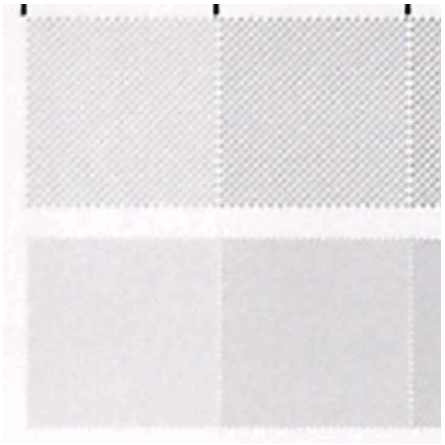
The first thing to notice on the images is that the same target number can be read on all three resolutions. This says that the maximum amount of data was achieved at less than 300dpi. If I had a 200dpi scan I expect it to have the same results. This scanner is either a low resolution scanner or the camera and optics are so poor that it cannot achieve a higher MTF (clarity/focus) on a scan to see more details. It may scale to generate 400 or 600dpi output but it is NOT a 400 or 600 dpi scanner.

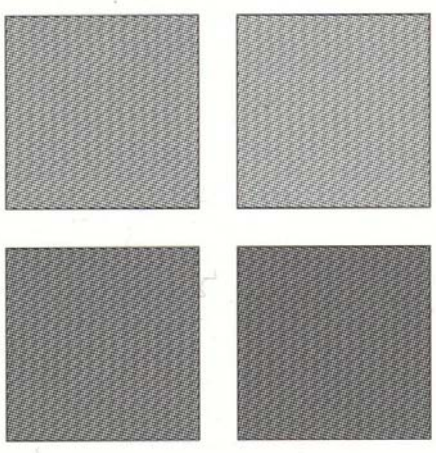
Compare this to another scanner. At 1:1 I can read 8.0 or maybe 9.0 off this 600 dpi scan. It is a little difficult to read here but is clear as day on the original.




Here on the left is another 600dpi snippet of the dot mesh pattern. You can just see the pattern on the top squares, the bottom ones just blur together. You can't see the B/W transitions. You can't see the image features. The 600dpi image in the right has a fine dot matrix that is seen very clearly.

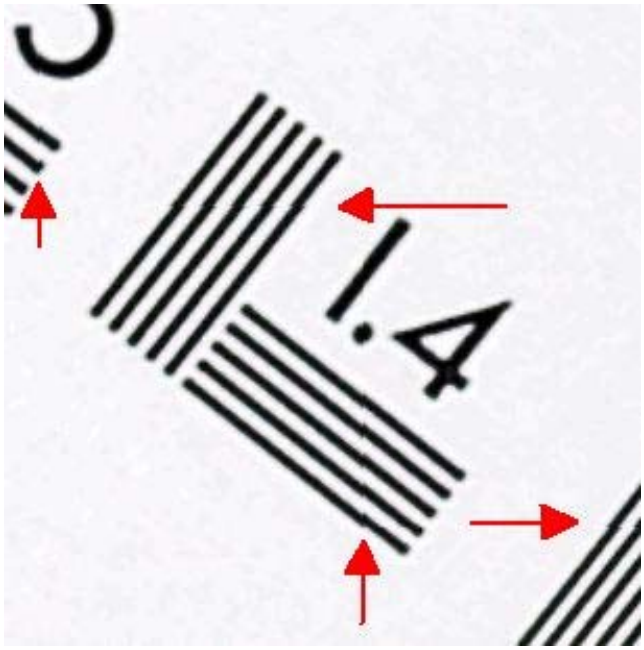




# How to Evaluate a Scanner

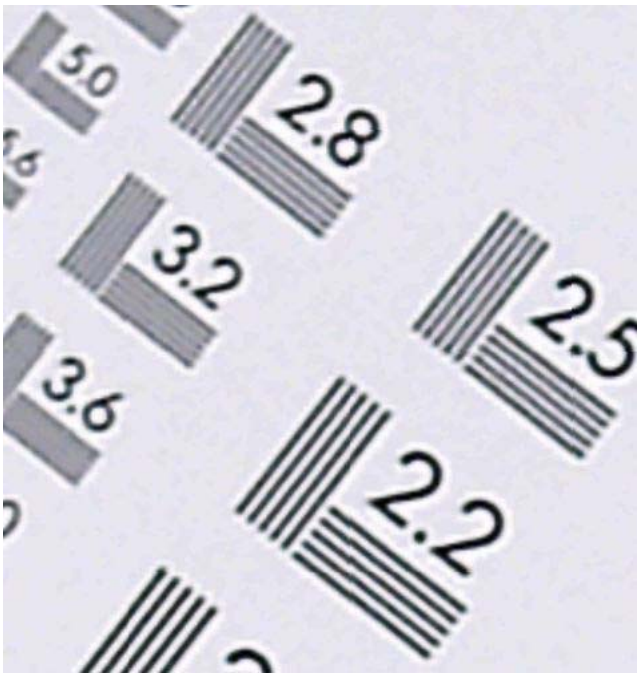
By Derek Jenkins

While looking at the images I noticed two other features that need pointing out. The first I think is the result of a de-skew or even stitching. The other is the result of very heavy use of a sharpen filter.




My guess here is that the image was de-skewed using a very simple shift both horizontal & vertical. This is not the technique used on better equipment.

It could also be the result of scanning the image in patches and stitching it back together again. Either way this is not acceptable.




Around each character is a white halo. The halo is the result of a very heavy sharpening to boost the clarity of the B/W transition. In Color or G/S scanning a very low filter is most effective. However, for B&W scanning the background is washed out and only the black is left so a heavier filter can improve the quality of a cheaper scanner.

# How to Evaluate a Scanner

By Derek Jenkins

**Summary:**
Scanner and image evaluation is a very, very broad topic. I have tried to stay with the core points that can be looked at by most people. There is so much more that I have not even touched on the complex topics like color, lighting, color temperatures and optical systems. I have also not covered the image distortion and its required mathematical correction that is inherent in some scanner designs.

Selection of a scanner is always going to be a compromise. Many time budgets dictate what is purchased. Please keep in mind that you cannot buy a cheap scanner and scan for high quality even though the salesman will tell you it can do it.

This outline has been produced as a source of information for you. I have tried to be as generic as possible and not slanted in any direction. I am very willing to spend time and discuss this with you to help you produce good evaluation criteria.

If you remember a few things from this presentation they should be:-

- **Don’t believe the salesman.**
- **Test and evaluate subjectively.**
- **Don’t Panic**

# 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 has been involved in the microfilm digitizing and document scanning industry since its inception. Coming from a systems software and R+D background Derek allied his new company with the fledgling scanner manufacturer called SunRise Imaging® and participated in the development of a new microfilm scanner. Over the years SunRise® became the leader in the film scanning industry and IImage Retrieval was its largest and most experienced reseller and repair facility worldwide.

In 2002, IImage Retrieval severed ties with SunRise® and its principles formed another company called nextScan. They developed a new line of microfilm scanners and applications that set new quality and performance standards in the film scanning market.

Driven by IIRI's customer's requests, they expanded its products to include the i2s line of rare book scanners. The introduction in the US of the DigiBook™ archival scanners was heralded by the Ransom Center at the UT Austin digitizing their pristine two volume edition of the Gutenberg Bible. Since then, many other establishments have purchased scanners to digitize their special collections and archives including the Library of Congress (LC), NARA, the University of Pittsburgh and the University of Florida. The University of Florida at Gainesville has seven of the smaller i2s CopiBook systems scanning a massive collection of Caribbean newspapers. The LDS Church also has eight CopiBooks installed at various sites across the United States.

Recently we installed DigiBook scanners in Cairo and another Arabic county with the LC, as part of the World Digital Library project.

The i2s line of high resolution scanners were designed for the archival industry but also fill a very flexible roll scanning any large flat objects from museum displays, tapestries and pictures to maps drawing or newspapers.

**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