TECHNICAL FEATURES OF A HIGH RESOLUTION FACSIMILE SYSTEM

Reed O. Slaugh*

Abstract: Rapicom, Inc. has developed a system to transmit press ready film for remote printing of high quality images including 150 line screen color magazine separations.

The major components of the system will be described with particular detail given to those areas which are unique or where the system has an advantage over existing technologies.

The input uses a drum scanner with a helium neon scanning spot and a focused detector to avoid some of the effects of spot flair.

The output is also a drum scanner which uses a helium argon laser to expose film and has multiple resolution capability from 600 scan lines per inch to 1200 scan lines per inch.

The transmission system uses data compression with an adaptive run length compression algorithm. The algorithm will be compared to Huffman and other algorithms.

The communications system uses selective block repeat synchronous transmission. The system has the capability for full communications error control when transmitting to multiple receivers over a wide variety of communications circuits. It can transmit simulatiously to locations on three different continents through multiple satellite circuits.

*Rapicom, Inc.

Technical Features of a High Resolution Facsimile System

Before I get into the main part of this paper, I would like to introduce very briefly my company. Rapicom is wholely owned by Ricoh of Japan. We are the world leaders in high speed digital compressed facsimile. The company was founded in the Silicon Valley on the San Francisco peninsula in 1966 and was originally named Dacom. We were the original developers of the sub-one-minute document facsimile and are the industry leaders with over 25,000 units in the field under the Dacom, Rapicom, Rapifax, Kalle and Ricoh labels. We also developed the first high resolution digital compressed publications facsimile system in 1970, originally for The Wall Street Journal, and we currently have over 200 systems in the field. The system that I will discuss today is a major enhancement of this system for use in transmitting full magazine pages, including the color separations, with Time Magazine as the first installation in 1980.

A dictionary definition of facsimile is "an exact copy", or the process of transmitting an exact copy via a communications circuit.

Many of the systems currently under discussion use the concept of generation of information within a closed system, then outputting the information in the final format, the "computer-to-plate" concept.

Facsimile takes an image that is already in the final format and transmits this image to other locations with the received image "an exact copy" of the original.

One of the main advantages of facsimile is that it is very insensitive to the source of the input material.

If we use a news magazine as an example, the copy within the magazine may come from many different sources. Advertising is generated by various ad agencies, color news photos are generated by one or more color houses, black and white photos may be generated in house, and color photo covers and color artwork covers each by other sources. It is not unusual for a magazine to have 10 or more different sources for the images that go into the final product.

Facsimile is ideal in such a situation because it can take as input a plate ready image from whatever source and transmit the image to remote terminals located throughout the world. The preparation of the input image for transmission is ideally no different than the preparation of the image for exposure to a printing plate.

Quality

What goes into making "an exact image"?

With facsimile, as with most things in life, the copy is never exactly like the original, but we try to make the quality of the copy so good that it very difficult to tell the difference.

What constitutes acceptable quality?

One definition I have heard used is that quality is acceptable when the average reader can not tell the difference between a fax and a non-fax copy.

Another definition that I have heard is that quality is acceptable when the art director with a microscope can not tell the difference between a fax and a non-fax copy.

The latter level of quality is probably never attainable. The level generally used is that the publication production crew have to look closely to tell which is fax and which is non-fax.

I would now like to examine some of the elements that go into producing a quality facsimile image.

Resolution

The first factor of quality is resolution. Newsmagazine editoral color halftones are generally in the 133 line screen range, with covers in the 150 line screen range.

One of the prime considerations in transmitting halftones is the elimination of moire patterns. The normal moire encountered in color separations is caused by the interference of the screen for one color with the screen for another color.

In facsimile, moire patterns are generated as "beat" frequencies or harmonics between the frequency of the halftone screen and the scan frequency. If the frequency of this harmonic is low enough, it can become very visible as a form of moire.

General facsimile experience has indicated that to avoid moire completely, the scan resolution must be at least 10 times the halftone screen density. Thus a 65 line screen halftone would require a scan resolution of 650 lines per inch. This would also indicate that 133 line screen halftones would require 1330 lines per inch sampling resolution, and 150 line screen would require 1500 lines per inch. Rapicom has developed techniques which reduce these resolution requirements, as I will explain later.

Another quality consideration is the accurate scan resolution of the smallest printable elements - the very small highlight and shadow dots.

A somewhat arbitrary rationale for picking this resolution would be that the smallest dot of interest should be scanned with four picture elements or pels. The smallest dot should be a two pel by two pel square. If we use this criteria to resolve 3% dots of 133 line screen, the resolution needed is about 1500 lines per inch.

Scanning System

Another factor that becomes critical in this area is the problem of actually being able to resolve small element areas in scanning reflective copy.

You are probably aware that transmission densitometers can easily be used to obtain an accurate indication of halftone dot size or percentage. Reflection densitometers, however, can not be used to obtain accurate dot sizes from reflective copy. The graph shows typical errors found in using a reflection densitometer to measure dot size.

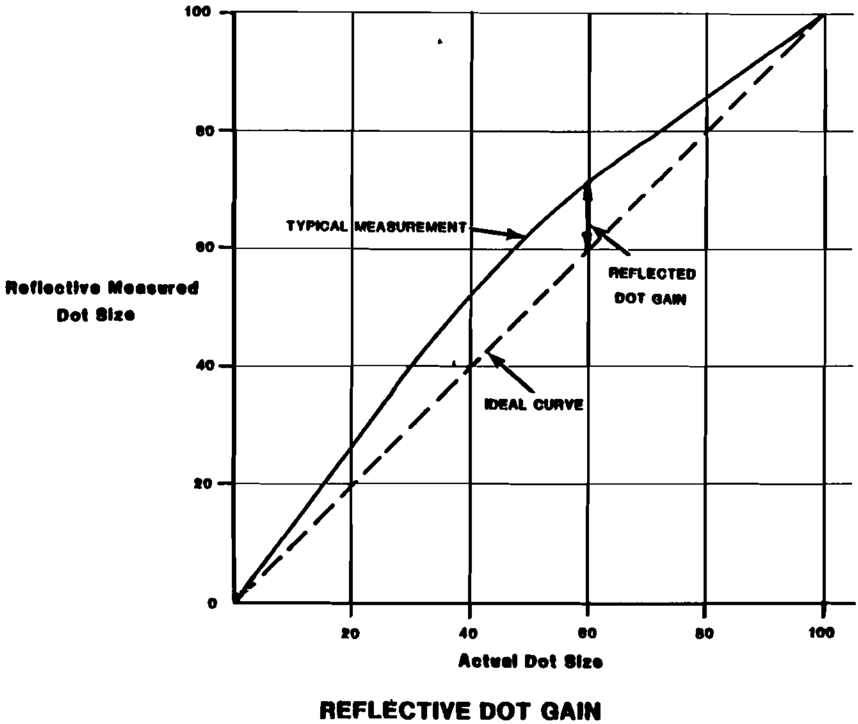

Figure 1. Reflective density curve.

The same physical properties that make it difficult for accurate reflective dot size measurements, make accurate scanning of reflective copy difficult.

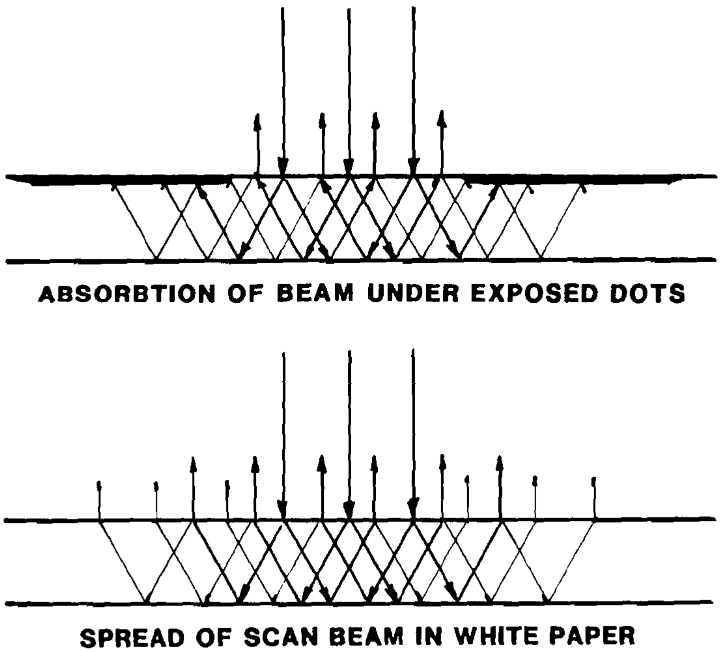

Figure 2. Scan beam reflection.

A simple explanation of the problem is that when one is scanning an all white area, the emulsion acts like a small reflecting waveguide for the light beam. The scan beam enters the emulsion, bounces around, then reflects back some distance from the point of entry. This increases the effective size of the scan beam. This spot growth or flair is further complicated by the fact that it does not occur in small white areas such as those in small shadow dots. In the small white dots which are surrounded by black areas, the beam enters the emulsion then bounces under the black exposed areas and is absorbed.

The net result is that the reflected light that a scanner sees in a small white area is much less than in an all white area. The amplitude in a 5 percent dot can be reduced by 50 percent.

The technique that we have used to minimize the effects of spot flair is to use a dual focus scanning system. The system uses a focused helium neon laser spot to scan the image and detects the image with a detector which is also in a focus plane. By having the detector in a focal plane, only the exact area of the scan spot is seen by the detector and the flair area of the spot is not detected. This is a technique which is not available in the current flat bed scanners.

Because of the need for high accuracy and the requirement of a focused detector, the optical system is implemented in a drum scanner. To enable the customer to make speed/quality trade-offs, the system has multiple resolution capability. In the standard configuration, these resolutions are 1200 and 1000 scan lines per inch, each with a sample resolution of 1400 samples per inch. There are also 1000, 800, and 600 lines per inch, each with 1000 samples per inch.

The next area where specialized circuitry is used is in the video sampling process. As indicated earlier, moire is a potential problem with facsimile. Sampling moire is caused by periodic errors in the video sampling process. We have found that this moire can be significantly reduced by an error correction algorithm in the sampling process.

When an analog video is sampled and changed into a black and white digital signal, errors will always result. The circuitry used in the Rapicom system analyzes each error and uses the result to correct the next sample. The example shows a typical video waveform for halftone dots near 50 percent. The high amplitude represents white, and the low amplitude black, but there are gray areas in between as the signal changes from black to white. The normal method of sampling and digitizing the signal is to establish a threshold about half way between black and white, with all signals above called white and those below, black.

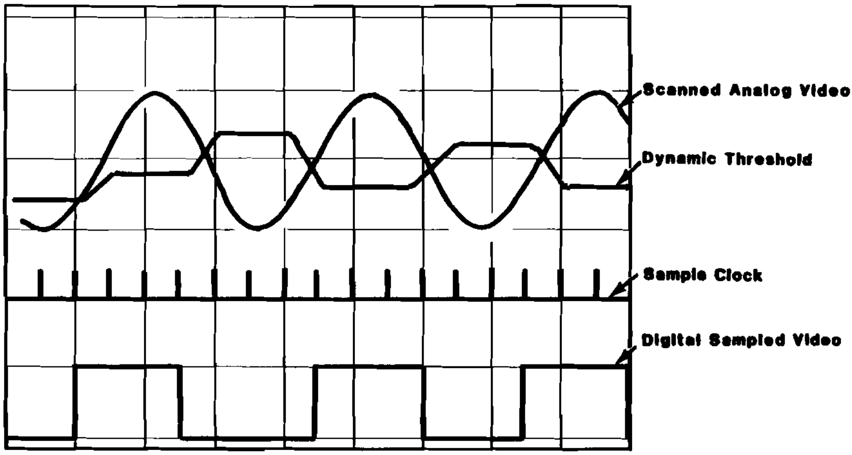

Figure 3. Video waveforms.

A digital system can only change state at discrete points in time represented by the sample clock. Because of sampling, there is always an error between the unsampled video and the sampled video output. The system uses this error to adjust the threshold point to correct for the error at the next sample. We have found this system to be very effective, not only in reducing moire, but also in enabling the system to reproduce each dot in the proper size, especially in the critical highlight, shadow and midtone areas.

Data Compression

The next factor necessary for quality facsimile is efficient transmission of the facsimile information. This is where data compression is very important. The role of data compression is to improve the speed, the efficiency, or cost; and the reliablity of moving information from one point to another and to do so without seriously degrading this information.

In order to understand the need for data compression I would like to look briefly at the amount of data required to transmit a facsimile document.

I have already discussed the high resolution needed for scanning half tone images. One problem with this is that as the resolution increases, so does the information to be transmitted, by a square factor.

For example, if we scan an 8 1/2 x 11 inch letter at 100 lines per inch in each direction, we generate 935,000 picture elements, or pels. If the resolution doubles to 200 lines per inch, the number of pels quadruples to 3,740,000. A newspaper page scanned at 1000 lines per inch generates a million bits per square inch or 360 million bits per page. The problem is even greater when we want to increase the resolution to that needed for magazine reproduction. For the four separations needed for a color magazine page scanned at 1200 by 1400 lines per inch, the number of pels generated is 806,000,000, almost a billion bits. If we want to transmit the newspaper page in one minute, this requires a transmission circuit capable of 6 million bits per second. There are several ways of reducing this requirement.

One way is to lower the scan resolution. Our optical system and dynamic threshold allow the resolution to be lower than would otherwise be required, but high quality still requires high resolution.

Sophisticated modulation techniques allow facsimile and other systems to transmit high bit rates in small analog channels, which reduces the time necessary to transmit a document. There is a requirement for even further reduction, and this requirement is filled by data compression.

The purpose of data compression is to reduce the number of bits that have to be transmitted.

The basic concept used in data compression is that adjacent areas are likely the same. A white area is very likely surrounded by white, a black element is probably connected to other black elements. The important features are how big the white and black areas are, and the exact location of the transition point from black to white.

```
00000000000000000000000000000000   32W
00000000000111111111111100000000   11W, 13B,  8W
00000000001111111111111110000000   10W, 15B,  7W
00000000001111111111111110000000   10W, 15B,  7W
00000000001111111111111100000000   10W, 14B,  8W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001110000000000000000000   10W,  3B, 19W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111111111000000000000   10W, 10B, 12W
00000000001111111111100000000000   10W, 11B, 11W
00000000001111111111100000000000   10W, 11B, 11W
00000000001111111111100000000000   10W, 11B, 11W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111000000000000000000   10W,  4B, 18W
00000000001110000000000000000000   10W,  3B, 19W
00000000001110000000000000000000   10W,  3B, 19W
00000000001111000000000000000000   10W,  4B, 18W
00000000001111111111111100000000   10W, 14B,  8W
00000000001111111111111110000000   10W, 15B,  7W
00000000001111111111111110000000   10W, 15B,  7W
00000000000011111110000000000000   12W,  7B, 13W
00000000000000000000000000000000   32W
```

```
 32 W    32 Pels
100000    6 Bits

 11 W    13 B    8 W       32 Pels
001011,001101,001000    18 Bits
```

896 Pels Scanned and 480 Bits Transmitted

Compression Ratio = 1.87

Figure 4. Data Compression

To illustrate the conçept of data compression, I have a typewritten capital E that has been scanned horizontally at 240 lines per inch. In this computer output we are representing each picture element (pel). Zeros are used to represent white elements and ones are use to represent black elements.

You may notice as you look at this representation, that there is a large amount of redundant information. In a facsimile compression system, the objective is to eliminate the redundant information and code the image in such a way that you can transmit just enough information so that the receiver can figure out how to plot the original image.

The common way of doing this is to count elements or run lengths of elements, and then transmit only the count or address of the position point where the image changes from one color to another.

In the illustration, the run lengths for each scan line are shown at the right of the line. We can then see that the numerical columns on the right contain the same information as the picture element representation on the left. We can also see an apparent compression, in that the run lengths on the right take less space than the image on the left takes.

The next step is to take the numerical run length information and convert this information to something that is efficiently handled by digital machines which only recognise ones and zeros. This is shown at the bottom of the illustration.

The first line of 32 white can be represented by binary 100000, a 6 bit code which gives a compression of about 5 to 1.

The second line has three different run lengths. If we use 6 bits to represent each run length, this requires 18 bits, or a compression ratio of less than 2 to 1. In fact if we use this same scheme for the entire letter, we have reduced the amount of information by only 1.87 to 1. This indicates that we need to be more efficient than just using straight binary coding.

The next illustration shows what happens with halftone. It represents 5 percent 100 line screen halftone dots scanned at 1000 lines per inch.

```
+++++++++++++++++++++  21W
+++0+++++++++0+++++++   3W, 1B, 9W, 1B, 7W
++000+++++++000++++++   2W, 3B, 7W, 3B, 6W
+++0+++++++++0+++++++   3W, 1B, 9W, 1B, 7W
+++++++++++++++++++++  21W
+++++++++++++++++++++  21W
++++++++0+++++++++0++   8W, 1B, 9W, 1B, 2W
+++++++000+++++++000+   7W, 3B, 7W, 3B, 1W
++++++++0+++++++++0++   8W, 1B, 9W, 1B, 2W
+++++++++++++++++++++  21W
+++++++++++++++++++++  21W
+++0+++++++++0+++++++   3W, 1B, 9W, 1B, 7W
++000+++++++000++++++   2W, 3B, 7W, 3B, 6W
+++0+++++++++0+++++++   3W, 1B, 9W, 1B, 7W
+++++++++++++++++++++  21W
+++++++++++++++++++++  21W
++++++++0+++++++++0++   8W, 1B, 9W, 1B, 2W
+++++++000+++++++000+   7W, 3B, 7W, 3B, 1W
++++++++0+++++++++0++   8W, 1B, 9W, 1B, 2W
+++++++++++++++++++++  21W
```

```
  21 W    21 Pels
 10101     5 Bits

   3W     1B     9W     1B     7W    21 Pels
 00011,00001,01001,00001,00111  25 Bits
```

420 Pels Scanned and 340 Bits Transmitted

Compression Ratio = 1.24

Figure 5. Halftone compression.

Here the information content, or the number of transitions, is greater and the numeric codes take almost as much space as the picture element representation.

If you look at the binary coding for the second scan line, you see that we use 25 binary bits to code 21 scanned elements, which gives a compression ratio of less than one.

Now, if we are going to make the compression system worthwhile, we must do something to increase the efficiency of our coding scheme. The major problem in the illustration I am using is that we are using constant length codes. To code a run length of 21 we need 5 bits. But if we use 5 bits to code a run length of one element, the 4 leading zeros are severely reducing our efficiency. However, if we use different code lengths for different run lengths, we must do it in a manner that the reconstruction machine can follow what the compressor is doing.

The most widely used method is accomplish this is to use statistical run length codes. Here one does a statistical analysis of the various run lengths in a type of document. If for example, the most common run length were 4 elements, this length would be given the shortest code, and if 5 were the next most common run lengh length then it would be assigned the next shortest code.

The international compression standard for document facsimile is a statistical code known as a Huffman code.

Statistical codes work well if all the information on the documents you wish to transmit is similar, as it is in typewritten documents.

We find that this scheme is not as efficent in compressing newpaper and magazine images. The problem here is that the pictures and graphics have characteristics which are different from the characteristics of the text. If we base our code on the characteristics of text, it will not be efficient for half tones. If we use halftone characteristics as the basis for our code statistics, the code will not be efficient for text.

At Rapicom, we use an adaptive coding scheme that very rapidly adapts itself to efficiently code each type of information that our systems scan. The systems use short codes in areas of fine detail, such as halftone, and use long codes in large open areas such as headlines.

This adaptive coding significantly improves efficiency. But there is still a need for further improvement, so we look for other areas to gain efficiency.

If we return to our illustration, you can observe that as we go from scan line to scan line, much of the time adjacent lines are very similar.

The more complex compression schemes take advantage of this similarity between adjacent lines with coding multiple lines. The publications facsimile system developed by Rapicom codes three adjacent lines at one time. This results in a system that gives an overall compression ratio of 14 to 1 measured over a very large average of newspaper pages including halftone and text. It represents the most efficient system currently available.

There is also a multiple line version of the Huffman code which has the acronym READ which again improves the efficiency of the Huffman code at the cost of complexity.

There is continuing effort to develop even more efficient compression systems. One of the most promising will potentially draw together several different concepts currently used in the graphic art industry.

To explain this, we can consider OCR (Optical Character Recognition) as a form of facsimile. If we return to our scanned letter "E", I will try to show the similarity.

In OCR documents are scanned the same as in facsimile. The OCR algorithms then block out each character and compares each with a stored library of characters to find the best match, thus "recognizing" the "character".

There is a compression concept being worked on by several companies called either Adaptive Character Recognition (ACR), or by the more general term, Adaptive Symbol Recognition (ASR).

In ACR or ASR, an image is scanned, and characters or symbols are blocked out, much the same as in OCR. However, in ACR, as each character or symbol is scanned, the system builds its own adaptive library of characters or symbols.

As each symbol is blocked, it is compared with the symbol library. If a match is found, a code for the symbol is transmitted. If no match is found, a new symbol is added to the library and the new code plus all the information necessary to describe the symbol is transmitted to the receive site. In this manner, the system can quickly build a library that matches the particular characteristics of the document being scanned.

This system appears to have good potential with text only input. Those companies doing research in this area talk about compression ratios of 40 to 1 or better.

The concept becomes much more complex when the technique is applied to graphics and halftone. There has been research done in this area which appears promising, but is still in the research phase.

To review, the basic concept used in data compression is run-length coding. The efficiency of the codes can be improved by using statistical run length codes, such as Huffman codes, or by using adaptive codes. The efficiency of each method can further be improved by operating on multiple lines, at the cost of added complexity. One of the most interesting concepts that is still primarily in the research phase is Adaptive Symbol Recognition, which merges the concepts of run length encoding and OCR techniques.

The compression technique used in the Rapicom Multiple Resolution Magazine facsimile system is the most efficient for this type of imagery currently in use.

Communications

The next area of quality that I would like to discuss is the communications system.

The general design guideline used in developing the Rapicom compression system was that it should be able to operate over almost any type of circuit and that any communications errors would not affect the quality of the output image. We have demonstrated the effectiveness of our design on circuits throughout the world. The ability to operate reliably over almost any circuit is one of the main reasons the system has been so successful. The system is designed to work with the long delays encountered with satellite transmission, with sufficient storage for about 100,000 bits in the propagation path at any time. This will accommodate a two way link over three different satellites at a 56 kilobit transmission rate.

The systems currently installed operate over a wide variety of circuits throughout the world. One of the most extensive in terms of distance covered is the system for Sing Tao newspapers of Hong Kong. Sing Tao currently transmits from Hong Kong to receive sites in San Fransicso, New York, here in Toronto, and in London, and will add this year a receive site in Sydney. The system can transmit to all sites simultaneously. The circuits used in this system are standard satellite voice channels with modems operating at 9600 bits per second. At this rate, the system transmits full newspaper pages in about 25 minutes.

In Europe, with systems in almost every country, the standard circuit used is the 48 kilobaud group band circuit. This is an analog circuit which is normally used by 12 voice channels. One of our first customers was the International Herald Tribune which transmits at 72 kilobits on the group circuit.

Another example is the N.M.P.P. group in France which uses multiple systems to transmit for 12 different newspapers.

For facsimile, high speed digital bandwidth compression modems can be used on the group band, with the fastest transmission rate currently 153.6 kilobits per second. An example of this is the London Financial Times system which transmits pages from London to Frankfurt in less than 3 minutes.

In the United States, the system has been used for years with satellite transmission for Dow Jones and The Wall Street Journal. Time Magazine uses the system for color transmission via satellite from New York to Chicago, Los Angeles, and Hong Kong. Starting this fall, Gannett will launch a new nationwide newspaper called USA TODAY, using the system to transmit with dedicated satellite earth stations from Washington, D.C. to about 15 receive sites throughout the United States.

In each of these applications, the system uses an effective communications error control subsystem. The compressed video data is sent in long blocks of synchronous bits. At the end of each block, a check code is attached. At each receive site, the check code is used to detect if any errors have occurred in the transmission process. If errors are detected, the receiver sends a message back to the transmit site, asking that the block in error be retransmitted. This retransmission process insures that only good data is reconstructed and plotted. The Gannett system has a subsystem that will enable simultaneous transmission to up to 32 different sites with full communications error control.

Recording System

The last area of quality that I will briefly discuss is the film recorder. The system uses an air-cooled argon ion laser to plot the image onto film. The spot is gaussian in shape. When gaussian elements are convolved together to form larger areas, errors are made, particularly in the very small dot areas. It would be a proper topic of a separate paper to discuss methods of optimizing this spot size and minimizing errors.

I will now just say that the effect of these errors has been analyzed and the system has circuitry to electronically correct the size of the very small dot areas. The net result is that over the range of interest, there is extremely good correlation between the picture element area and the exact area exposed.

To conclude, I would like to show a few slides of the equipment I have been discussing. The basic drum scanner and recorder are manufactured by Muirhead, Limited. Rapicom has modified the basic units for high resolution requirements. In particular we completely replaced the optical system with the laser based optical system.

The compression system is not limited to drum scanners for input and output, though the drums give higher quality currently. The system can use a variety of input/output devices. We currently have systems installed using the Muirhead flat-bed scanners. An interesting side benefit of compression is that one can send from a flat bed scanner to a drum recorder, or vice versa.

I suppose that many of you have read a magazine or newspaper that had been transmitted by a Rapicom facsimile system and never realized that it was a facsimile copy. This is in accordance with our goal - to make our facsimile copy "an exact image" of the original.

Thank you.