Michael G. Noblett,[1] *B.S.*

Image Processing and Statistical Analysis as an Aid in the Comparison of Typewritten Impressions

REFERENCE: Noblett, M. G., "**Image Processing and Statistical Analysis as an Aid in the Comparison of Typewritten Impressions,**" *Journal of Forensic Sciences,* JFSCA, Vol. 32, No. 4, July 1987, pp. 963-974.

ABSTRACT: A device is being evaluated which images and statistically evaluates typewritten impressions on paper. The device can differentiate between typeface designed and manufactured by one company and typeface copied from the same design but manufactured by a different company. The device can make precise and detailed measurements of typewritten impressions while reducing the effects of the ribbon and paper to provide a valid and reliable image of the individual typewritten characters.

KEYWORDS: questioned documents, typewriters, comparative analysis, digital image processing

The concept of physical measurement of typewritten impressions is certainly not novel to document examiners. The first test performed on a typed document is to measure the horizontal spacing. Beyond that, however, measurements of the individual characters do not provide a great deal of additional information, due in part to our inability to define and quantify what we see. Unless the measured differences are grossly inconsistent, very little can be said as to their significance. Even relatively modern optical devices such as the Projectina and other optical comparison devices face problems of measurement and comparison because the device is able to look at only one impression (or pair of impressions) at a time.

Discussion

Figure 1 shows a single lowercase "t" at ×35 magnification. The inherent problem in measurements of typewriter impressions on paper is clear in this figure. The document examiner, looking at any single typewritten impression, observes the shape of the character and any gross abnormalities such as bent and broken serifs, misalignment, or off-foot striking. However, as magnification is increased to a point at which these same abnormalities should be visible in the fine structure of the impression (slightly bent serif, for instance), the abnormalities are masked by the interplay of the typewriter ribbon, type font, and paper. The edge of the character becomes ragged and ill-defined; the information we seek is lost in this ragged edge. If we think of this ragged edge as noise and the information we seek as signal, we are faced with the same problem that has plagued scientific instruments throughout history: How can we reduce the noise while we preserve and amplify the signal?

Received for publication 12 Sept. 1986; revised manuscript received 16 Oct. 1986; accepted for publication 20 Oct. 1986.

[1]Special agent, Federal Bureau of Investigation Laboratory Division, Washington, DC.

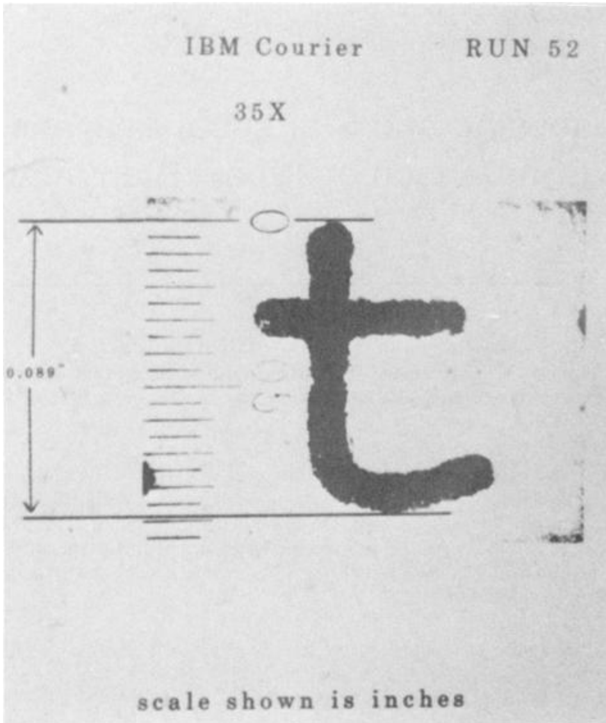

FIG. 1—*X35 magnification of a typewritten impression showing the ragged edge created by the ribbon and paper.*

A more detailed look at the document problem shows that we may have the capability of solving many of our problems with the same methods used by the electronic and electrical engineering community. If we consider the three components which make up our typewritten impression, we find that one, the element (ball, wheel, thimble, and so forth), is clearly the signal we are trying to measure. The other two, the ribbon (cloth or film) and the paper, generate the noise we are trying to suppress. Since paper consists of a random distribution of fibers and the ribbon contributes a random distribution of effects,[2] averaging these effects should eliminate or at least significantly reduce them.

The technique employed by the electrical engineering disciplines is signal processing. Since our problem is not really a signal as much as an image, application of some principles of digital image processing was a logical starting point.

Analysis of the problem in the Document Section of the FBI Laboratory indicated that there were certain requirements that were absolute. The first was that the method chosen had to reduce or eliminate the noise effects. The second was that the method had to maintain the spatial relationships which existed in the original character (a lowercase "t" still had to look like a lowercase "t" when the processing was finished). And third, the relative gray scale density in the character had to be preserved (if there was a void in the middle of the lowercase "t" caused by the typing element that void still had to be there when the processing was finished).

[2]The cloth pattern observed in the typewritten impression is not always in the same place, and the plastic film does not always deposit the carbon with the same edge abnormalities.

With these absolute requirements in mind, a system was designed which used the techniques of digital image processing to aid the document examiner in the comparison of typewritten impressions. This system is currently being evaluated in the FBI Laboratory, and initial results are encouraging.

The system can be described as a very accurate ruler which can be instructed to perform certain measurements on typewritten characters and display the results of these measurements both as numbers which can be evaluated statistically and as a picture which can be evaluated subjectively by the document examiner.

The initial analysis of the problem and approach also suggested some practical considerations. For instance, the system averages similar individual characters to produce an average or typical character. The individual measurements are statistically processed to produce the needed data, that is, average width and its variance. It is both logical and statistically correct that the more individual characters which go into this average, the more accurately the average will represent the character it purports to be. Likewise, any statistical evaluation will be more accurate because the statistics will be based on a larger population. Since most typewritten documents received as questioned items in the FBI Laboratory are relatively short, and since numerous repetitions of the character are required for the statistics to be valid, special emphasis should be placed on evaluating letters which occur frequently in the English language.

Instrumentation

The actual measurement of the impressions is done using a Prime computer configured with 2 Mbyte of core memory and disc storage of 400 Mbyte. Output from the Prime computer is communicated to a standard terminal as well as a line printer. Information about the typewritten impressions is entered into the computer automatically from the questioned document. The questioned document is mounted on a drum and held flat by vacuum. The diameter of the drum is 7.6732 in. (19.4899 cm) and the circumference is 24.106 in. (61.229 cm); it rotates at 300 rpm. As the drum is rotating, a photomultiplier traverses the drum axially at a scan density of 600 lines per inch. A reading is taken from the photomultiplier every 6.944 μs (a sampling frequency of 72 kHz) to produce a picture element of 0.0016 in. (0.0041 cm) in diameter. This picture element is input to the computer and is further characterized by assigning a value of how dark the picture element is in one of 256 shades of gray from absolute black to absolute white. Scanning a document 8½ by 11 in. (21.59 by 27.94 cm) produces 33 660 000 picture elements and, since each picture element requires 1 byte of gray scale information, storage of this picture takes 33.6 Mbytes of space on a magnetic disc.

After the document is stored in computer memory, the information is acted upon by well documented and reliable algorithms which start stripping the characters one by one for individual analysis.

Figure 2 illustrates a typical stripped character. The dots represent spacings on the actual document of 1.66 mil (0.04 mm). The vertical and horizontal uneveness is introduced by a standard printer being employed as a graphics device. A true graphics printout is planned for the future and would not exhibit the apparent anomaly. Alphabet letters are utilized to signify the grey level of each picture element. In order not to overly clutter the presentation, each letter represents five adjacent shades of grey. The dots denote the lowest 6 shades of grey, lowercase "a" to "y," the next 125 shades, and uppercase "A" to "Y," the highest 125 shades (256 in all). An operator settable analog threshold is employed to suppress dirt effects without compromising character integrity. In Fig. 2 that threshold is 90, which causes all levels below that value to become dots. This is a somewhat confusing display which requires getting used to, but there appears to be no better way of portraying shades of grey to the required resolution. The highly nonlinear human eye tends to suppress small variations in grey scale in the more usual type of presentation, but computer processing stresses them to

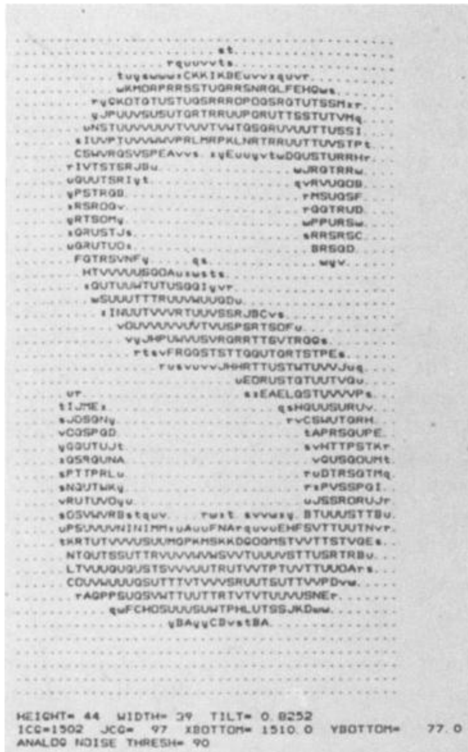

FIG. 2—*A computer printout of a single character. The arrows indicate areas where the random effect of the ribbon and paper occur. The letters which make up the character represent gray scale with darker areas indicated by uppercase letters higher in the alphabet.*

provide the document examiner with information heretofore unavailable. Below the pictorial presentation of the letter are several numbers which are the first measurements made on an individual impression.

Character width is defined as the number of picture elements contained within and including the left and rightmost extremities of the character—regardless of the grey level of those bounding picture elements. Width is ascertained after the analog threshold is applied. Similarly, height considers the top and bottommost picture elements. Character tilt is derived with a moment of inertia regimen and subtracts out the page tilt.

The character center of gravity is computed to provide leading and trailing pitch data to the centers of gravity of the preceding and succeeding characters. This method is used in preference to intercharacter space measurements, since the latter is very sensitive to stroke thickness variations. There is a serious question as to whether such pitch information is meaningful, except when large populations are involved. Lastly, the character's vertical displacement from the datum of the line of text based on its center of gravity is computed.

The character represented by Fig. 2 also contains the random noise effects discussed earlier and shows how these noise effects tend to obscure the actual shape of the character represented.

The method chosen to provide a noise-free character is to average all like characters. Figure 3 represents an average character where ten symbols we know as "s" existed on the document. These characters were aligned relative to their centers of gravity, and all of their

FIG. 3—*A computer printout of an average character made up of ten characters one of which was shown in Fig. 2. Note that the two areas marked in Fig. 2 do not appear. The numbers associated with width, height, and so forth, are expressed in picture elements. Each picture element is 0.001 66 in. (0.004 22 cm).*

individuals picture elements were summed maintaining both the value of the grey level in the picture element and its location. The individual sums were then divided by ten to obtain the fairly clear image shown in Fig. 3. Note that the two "bumps" on the upper and middle portions of Fig. 2 have disappeared in Fig. 3. This indicates that those two bumps were an artifact of that one "s" only and not associated with the other nine "s"s in the document.

The measurement data associated with the individual characters (height, width, and so forth) are statistically combined to form the mean and standard deviation. The presentation of this data which appears above the letter in Fig. 3 is termed the attribute table; one attribute table is required for each symbol which the computer processes. In addition to the statistical data, several housekeeping functions are recorded. A file name is appended to the character as well as a reference number. The date and time of the operation is recorded. The total number of characters found on the document is shown (in this case, 697 characters), and the analog noise threshold which was described previously is recorded. The table also shows the total number of like categories found (in this case, 88) and finally the maximum number of categories which the system is able to store. Specific to this particular computer run, on 14 March 1985 at 3:30 p.m., a document described as IBM Scribe was read into the computer. The document contained a total of 697 characters, and these characters were further grouped into 88 categories based on similar shapes. Of the 88 categories found, one looks like Fig. 3 and this single category has a population of 10 characters.

Since the device makes fine and accurate measurements, it seemed logical that a good test to evaluate its ability to provide useful forensic science information would be to look at the problem of differentiating lookalike type styles. These type styles often copy IBM and are available on most machines and in most of the popular single element configurations (ball, print wheel, thimble, and so forth).

Figure 4 shows text prepared using three different but lookalike single elements. A careful study of these three samples will show differences which are real and demonstrable. In actual practice, however, the document examiner seldom has the luxury of having all the characters on the element appear on the questioned document.

Test Results

Nine documents were prepared which bear numerous repetitions of the letters E, T, O, A, N, S, R, which together represent approximately 58% of the letters which appear printed in the English language, and the numerals 0 through 9.

Three documents were prepared on an IBM Selectic II equipped with an IBM Courier twelve-point element. Three documents were prepared on the same machine using an element manufactured by DSG (Philadelphia, PA) and called Courier 12. Three were prepared on a Royal typewriter using an element manufactured by CSA (Neuchatel, Switzerland), also called Courier 12.

All nine documents were scanned on the laboratory instrument and measurements taken of the individual characters. From these measurements an average character was created. For this paper, the measurements used were height, width, and tilt. The average and how much the individual measurements varied from that average were calculated and printed.

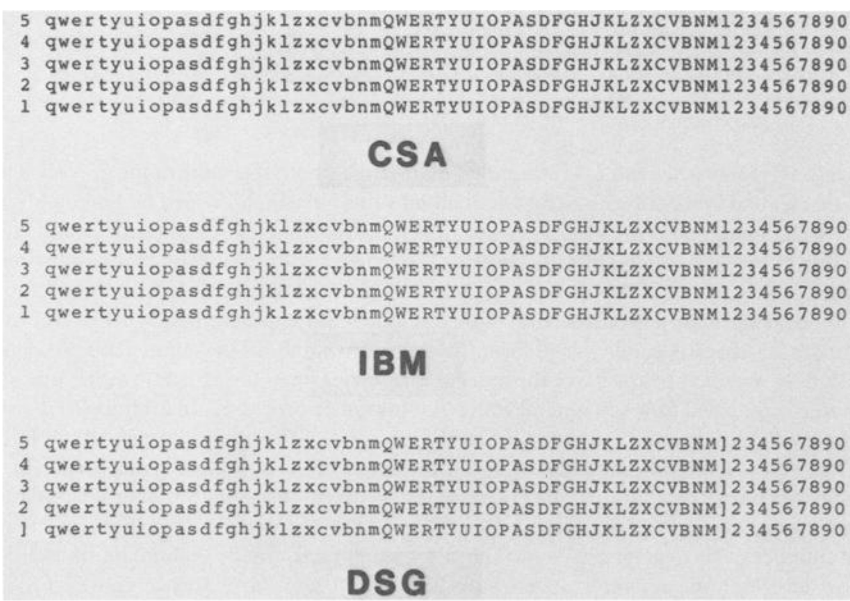

FIG. 4—*Five lines of typewritten text prepared using single elements designated courier manufactured by three different companies.*

A confidence interval was calculated for these data. The confidence interval describes the range (in real space) around the average which the measurements will have. The confidence level chosen was 99% and is derived from the student t test as follows:

$$u = \bar{X} \pm t_{\alpha} \frac{s}{\sqrt{n}}$$

where

$\bar{X}$ = average
s = standard deviation,
n = frequency, and
t_{α} = table value associated with the chosen confidence level.

For example, on Run 52, the average height of the lowercase "t" was measured at 54.34 units (0.0902 in. or 0.2291 cm) with the standard deviation of 0.6708. On Run 52, the lowercase "t" appeared 20 times. Thus,

$$\begin{aligned} \bar{X} &= 54.35 \\ s &= 0.6708 \\ n &= 20 \\ t_{\alpha} &= 2.539 \text{ (from the literature)} \\ u &= 54.35 \pm 2.539\ 0.6708/\sqrt{20} \\ &= 54.35 \pm 0.381 \\ &= 53.97 \text{ to } 54.73 \\ &= 0.0895 \text{ to } 0.0908 \text{ in. (1 unit} = 0.001\ 66 \text{ in.)} \end{aligned}$$

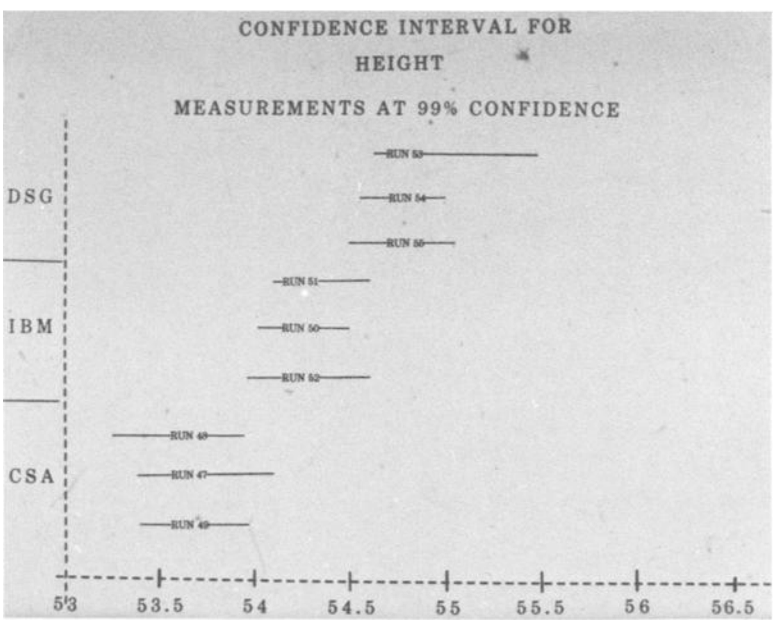

FIG. 5—*Confidence intervals calculated from nine tests (three each from the three manufacturers) measuring the height of the lowercase. "t." The numbers along the horizontal axis are picture elements. Each picture element is 0.001 66 in. (0.004 22 cm).*

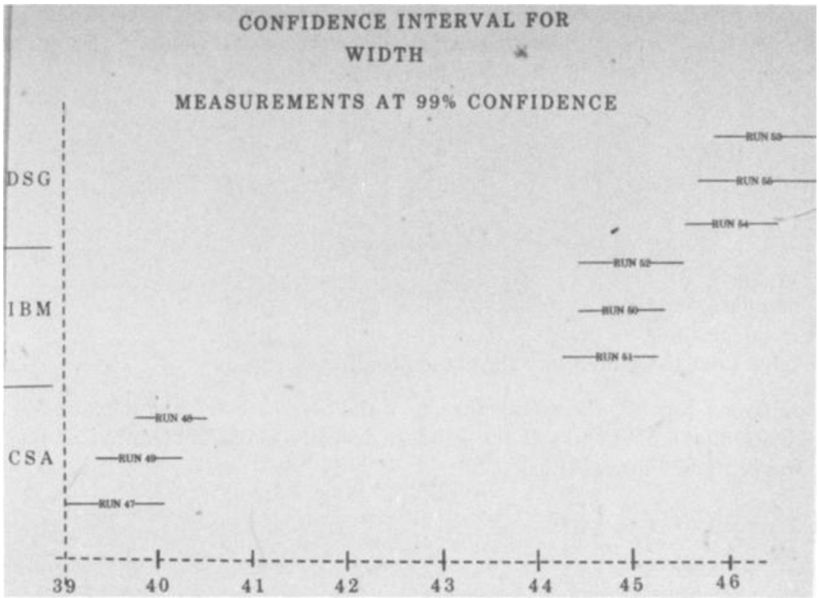

FIG. 6—*Confidence intervals calculated from nine tests (three each from the three manufacturers) measuring the width of the lowercase "t." The numbers along the horizontal axis are picture elements. Each picture element is 0.001 66 in. (0.004 22 cm).*

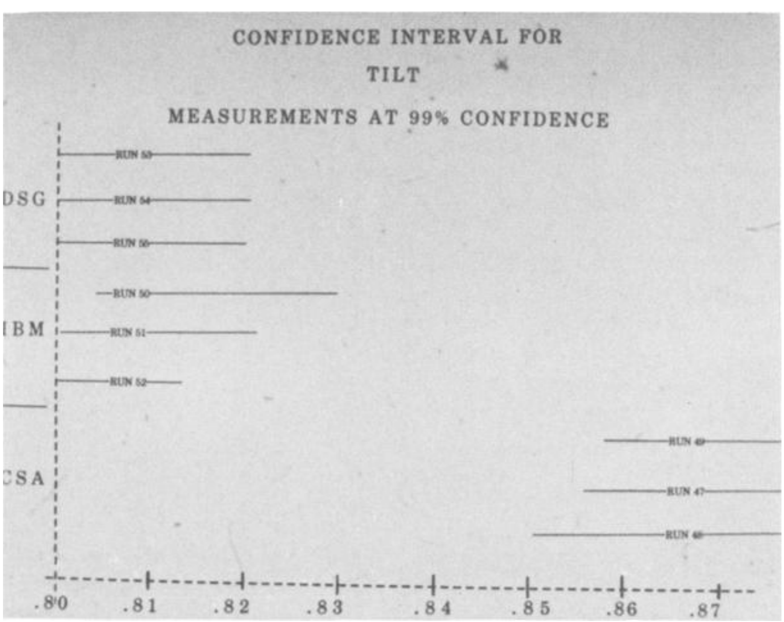

FIG. 7—*Confidence intervals calculated from nine tests (three each from the three manufacturers) measuring the tilt of the lowercase "t." The numbers along the horizontal axis are radians.*

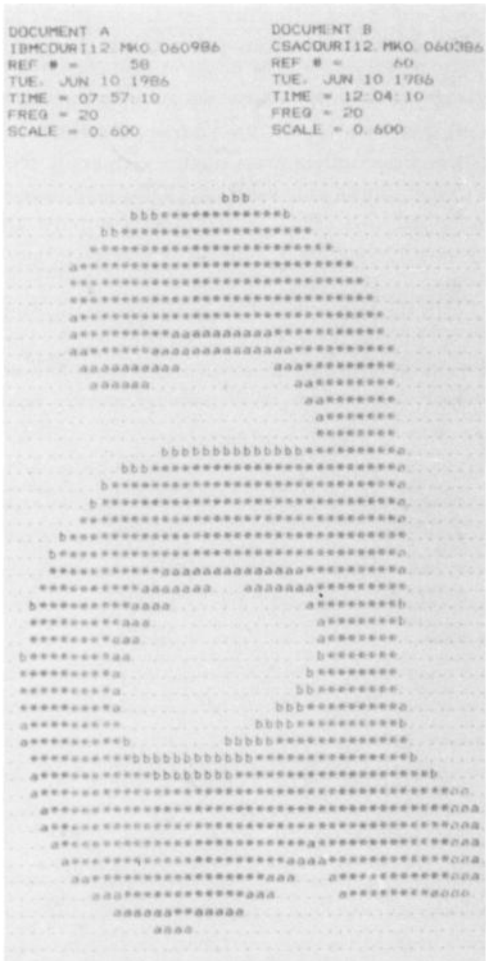

FIG. 8—*A computer printout of an average character made up of 20 individual lowercase* "a"*s prepared using an IBM Courier single element.*

In other words, we can say with a 99% confidence that the lowercase "t" represented by Run 52 will measure to be between 0.0895 and 0.0908 in. (0.2273 and 0.2306 cm) or will vary only plus or minus 0.0006 in. (0.0015 cm) from the average of 0.0902 in. (0.2291 cm).

Confidence intervals calculated from all nine runs measuring these three characteristics (height, width, and tilt) are shown. The numbers along the horizontal axis for height and width are expressed in picture elements; each picture element is 0.001 66 in. (0.004 22 cm).

As can be seen from Figs. 5, 6, and 7, the three type styles cluster well among themselves. Likewise, they separate when compared with their lookalikes.

The confidence intervals shown are valid and represent data symbolically. However, in operation, this is not the method of choice because it can look at only one set of data at a time. Operationally, the instrument will use techniques of multiple analysis of variance to determine significant differences between two sets of data from two exemplars.

Although the preceding data is compelling as an argument that there is a significant difference between the impressions based on the measurements of the lowercase "t," it is not immediately apparent just where this difference lies. The method chosen to show this difference in a more graphic form is a simple overlay and subtraction.

Figure 8 is the average character formed from 20 lowercase "a"'s typed on a machine bearing an IBM Courier 12 element. Figure 9 is the average character formed from 20 lowercase "a"'s typed on a machine bearing a CSA Courier 12 element.

The subtraction program first looks at every picture element to see if it has a grey value or is white. If the picture element has a grey value, the value is replaced by the letter "a." This image of the character shaped with "a"s is stored in the computer and filed as (in this case) IBM Courier 12. Likewise, the document we want to compare against this IBM Courier is examined for the presence of grey scale values, and those picture elements in this second document are replaced with the letter "b." This image of the character shaped with "b"s is stored in the computer and filed as (in this case) CSA Courier 12. These two images are then placed one over top of each other in the same way the average character was generated and

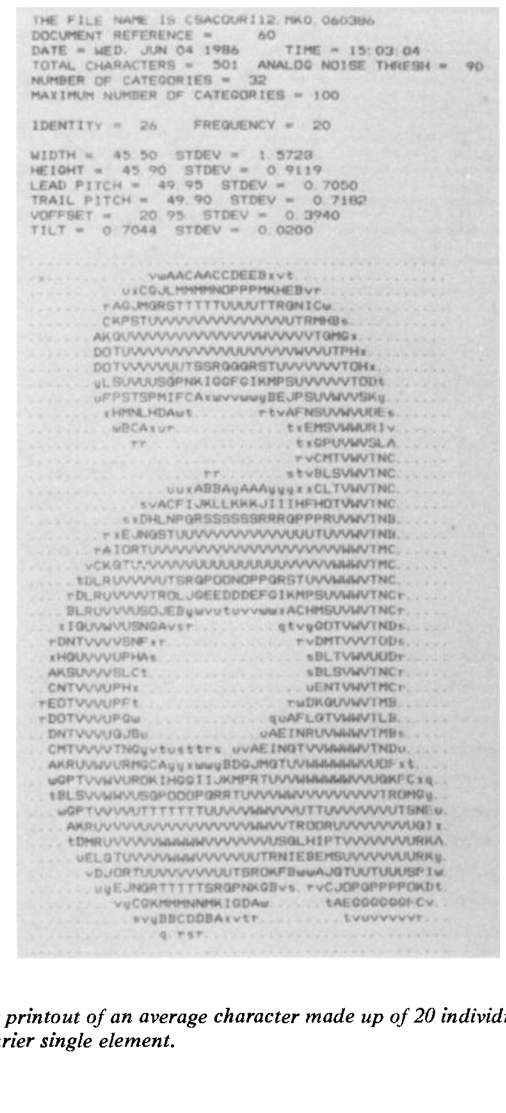

FIG. 9—*A computer printout of an average character made up of 20 individual lowercase "a"s prepared using a CSA Courier single element.*

the two images are compared. Where the picture element contains both an "a" and a "b," the picture element is replaced with an asterisk, indicating that a value was found in this location in both the IBM and CSA documents. If an "a" is found in a picture element with no corresponding "b," an "a" is printed. Likewise, if a "b" is found in a picture element with no corresponding "a," a "b" is printed. Using these rules, a third image is generated which shows the differences. This image is shown as Fig. 10.

Conclusion

The ability of this machine to make and record measurements of typewritten impressions in a laboratory environment has been demonstrated, and the value of this information is clear. What remains is a lengthy process of evaluating the ability of the machine to produce usable data in the real world environment in which document examiners operate.

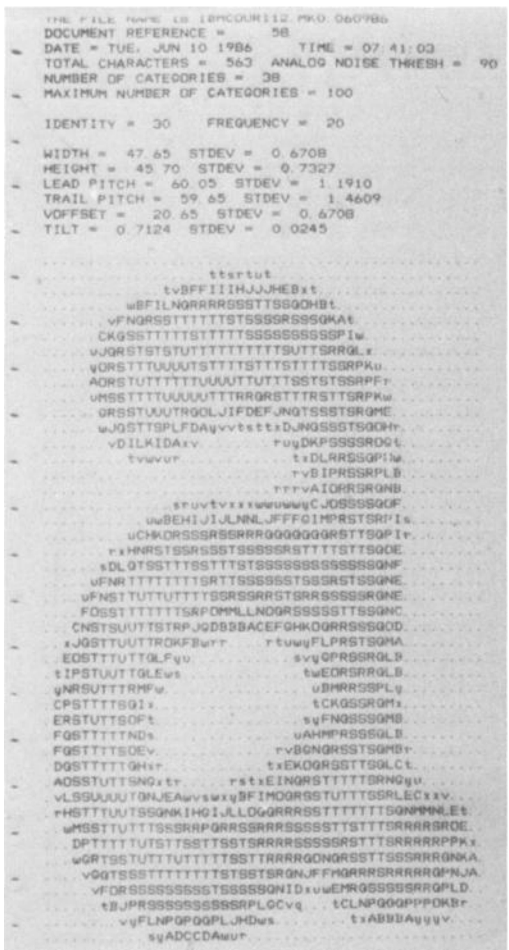

FIG. 10—*A computer printout showing where the differences occur between the character in Fig. 8 and the character in Fig. 9.*

Bibliography

Alder, H. L. and Poessier, E. B., *Introduction to Probability and Statistics*, W. H. Freeman, San Francisco, 1972.

Anthony, A. T., "Comparison of Modern Typestyles," *Journal of Forensic Sciences*, Vol. 31, No. 2, April 1986, pp. 710–718.

Hoel, P. G., *Elementary Statistics*, Wiley, New York, 1971.

Schlang, A., "Typewriter Impression Comparison Study Considerations," presented at the International Symposium on Document Examination, Quantico, VA, July 1985.

Suen, C. Y., "N-Gram Statistics For Natural Language Understanding And Text Processing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. Pami-1, No. 2, April 1979, pp. 164–172.

Address requests for reprints or additional information to
Michael G. Noblett
FBI Laboratory
Washington, DC 20535