Experimental Study on Distinguishing the Laser Printed and Copied Documents by Using ImageXpress

Yuchen Wang[1], Yong Wang[2], Zhongyuan Ji[1], Xibin Song[3]

[1]Key Laboratory of Evidence-Identifying in Universities of Shandong, Shandong University of Political Science and Law, Jinan, Shandong Province, China

[2]Chongqing Public Security Department of Chengdu Railway Public Security Bureau, Chongqing, China

[3]Technical Squadron of the Criminal Police Unit of the Lixia Area Sub-bureau of Jinan Municipal Public Security Bureau, Jinan, Shandong Province, China

Keywords: printed and copied documents; quantification; laser printer-copier; distinguish; ImageXpress

Abstract: In this paper, ImageXpress, a printing quality inspection instrument commonly used abroad, is introduced into the examination of copied and printed documents. Taking the laser printed and copied documents as experimental materials, this paper explores the principle of ImageXpress, analyzes the feasibility of printing and copying documents, and factors affecting the inspection and appraisal results. Experiments are carried out to fully demonstrate the feasibility and practicability of this instrument in the field of document examination. Further, in order to fully illustrate the feasibility of the instrument, on the basis of previous studies, documents printed and copied by the HP LaserJet Pro M1136 laser printer-copier are taken as samples to measure the relevant parameters of the printed and copied documents under the guidance of the international standard ISO/IEC 13660. The data samples are used to distinguish printed and copied documents, so as to explore the criteria for distinguishing printed and copied documents, and to provide data to support future research.

1. Introduction

At present, representative digital image analysis instruments are ProScan 2000 produced by the company of BST Pro Mark in Germany, GX100 and GX200 developed by the company of VIPER in the United States, PrintVision/9000NT, PrintVision/Genesis and PrintVision/Apollo produced by the Advanced Vision Technology Ltd. company (AVT) in Israel, as well as the CP2000 central control system developed by the enterprise of Heidelberger in Germany. The ImageXpress Image Expert System (IX) used in this paper is an image analysis tool developed by the ImageXpress Inc. in the United States. It seamlessly links various system analysis programs through the same session protocol shared by different systems. It can analyze the usage of papers, toner cartridges and powdered inks, as well as the printing machine and other printing parameters.

The analysis system can measure all features described in the ISO13660 international standard. It integrates the measurement of color and density through the X-Rite instrument (X-Rite is an image analysis instrument which integrates illuminometer, optical densitometer, densitometer and colorimeter), the laser analysis system for topological analysis and the human eye recognition system. It also integrates a variety of digital image measurement algorithms based on mathematical analysis, such as the gray level measurement algorithm system, the connectivity measurement algorithm system, the graphic area measurement algorithm system, the roughness measurement algorithm system and other digital image analysis algorithms to meet the needs of different image analysis.

2. The Influence of ROI Setting on Measurement Results

This section explores the influence of the size and location of ROI on the measurement results.

  DOI: 10.25236/iwmecs.2019.007

The sample is an A4 paper document filled with the Chinese character of “Zang”. The characters are in Song typeface; the word size is three. The document is printed by the HP LaserJet Pro M1136 when the toner cartridge is half used. The Chinese character with the coordinates of (13, 14) is taken out.

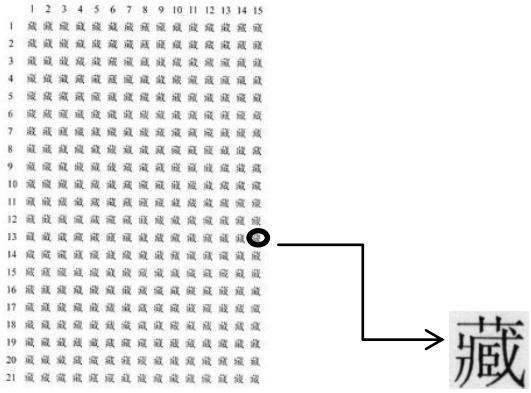


Figure 1. The Chinese character of “Zang” with the coordinates of (13, 14)

2.1 The influence of ROI size on measurement results

Changing the size of ROI means changing the size of the image and text contained in ROI. There are two main situations. One is that the image and text in the changed area have polar changes (abbreviated as case 1); the other is that the image and text in the increased area have no polar changes (abbreviated as case 2). The measurement threshold is 130; other measurement parameters are set automatically by the system.

2.1.1 Gray-level feature measurement

After changing the size of ROI, the gray-level characteristic values of several characters are measured, and the size of ROI is increased successively. For case one, the measurement schematic diagram and measurement values are as following.

Table 1. Measurement of gray values after changing the size of ROI in case 1

No.	1	2	3	4
ROI (area)	16.002	17.121	20.747	22.944
Gray Avg.	200.76	209.136	212.911	214.483
Profile StDev	34.92	36.507	37.41	35.602
Profile Gray Min	78.644	80.301	88.064	96.18
Profile Gray Max	226.204	228.301	229.789	230.198
Profile Gray Avg.	167.22	171.202	177.152	182.083

Figure 2. Diagram for the measurement of gray level after changing the size of ROI

The measurement of case 2: ROI is the same as the case 1. To simplify the measurement process, the polarity of case 1 in measurement is changed (dark pixels are measured).

Table 2. Measurement of gray values after changing the size of ROI in case 2

No.	1	2	3	4
ROI size (area)	16.002	17.121	20.74	22.944
Gray Avg.	45.76	45.76	45.76	45.76
Profile Stdev	33.33	35.25	34.438	35.792
Profile Gray Min	75.69	80.264	84.653	89.417
Profile Gray Max	224.027	225.026	228.276	228.409
Profile Gray Avg.	164.31	168.601	173.714	176.290

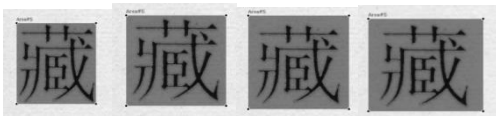


Figure 3. Diagram for the measurement of gray level after changing the size of ROI

2.1.2 Roughness feature measurement

The roughness of the edges of several strokes of characters is measured after changing ROI size (Rise Time features are selected). The size of the first ROI is set as the minimum, and then the sizes of ROI are increased successively. The measurement figures and values are as following.

Table 3. Measurement of toughness features after changing ROI sizes

No.	1	2	3	4
ROI size (area)	0.187	0.26	0.315	0.419
Rise Time	14	18.55	22.579	32.167

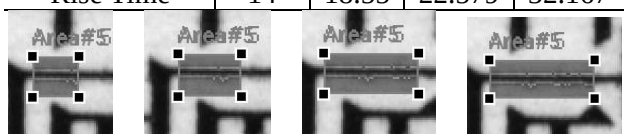


Figure 4. Diagram for the measurement of toughness after changing the size of ROI

In addition, the connectivity of strokes edges and other special features are measured after changing the size of ROI. The results show that with the changes of ROI sizes, the gray values, the levels of edge roughness and the connectivity of strokes edges are consistent with above measurement. Accordingly, for case 1, after changing the ROI size, the measured values of characters in the same position on the same scanned image are inconsistent; for case 2, after changing the ROI size, the measured values of characters in the same position on the same scanned image are changed.

2.1.3 Reasons and countermeasures to reduce the impact

With the changes of ROI size, there are different trends of measurement values in the two cases. The reason is that with the changes of ROI size, the image and text information contained in ROI will increase. In IX system, the measurement of many characteristic quantities is the process of calculating the average value of quantitative information in the whole ROI. The system calculates all the pixels that meet the threshold value and polarity, resulting in changes in measured values. Therefore, it is suggested that the ROI should be fixed in the actual measurement process.

2.2 The influence of ROI position on measurement results

The so-called change of ROI position means that in the measurement, the position of image and text is changed for the same size of ROI. The changing of ROI position can also be divided into two situations: one is that there are polar changes of image and text in the added area (hereinafter referred to as case 1), in which the measured values will change with the change of ROI position; the other is that there is no polar change of image and text in the added area (hereinafter referred to as case 2). Now the second situation is surveyed and explored. Samples are consistent with the previous section, and the polarity is set as dark pixels. The measurements show that:

Table 4. Measurement of gray level characteristic after changing ROI position

No.	1	2	3	4
Gray Avg.	45.76	45.76	45.76	45.76
Profile Stdev	35.251	35.406	35.126	35.527
Profile Gray Min	83.904	83.737	83.722	84.056
Profile Gray Max	228.349	228.131	225.657	228.409
Profile Gray Avg.	172.025	171.833	171.899	172.290

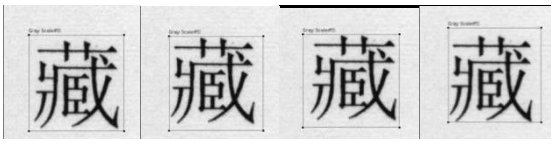


Figure 5. Diagram for the measurement of gray level after changing ROI position

The situation is consistent with the conclusion in the previous section, so it is suggested that the measurement position should be fixed in the actual measurement process by looking for a reference object or by using the Location in the Toolbox of IX system.

Figure 6. Location in the Toolbox can fix the measurement position

3. The Effect of Scanning on Measurement Results

Because of the projection and reflection principles of prism and gap, the scanner can transform the image, and then digitize the image through CCD element. Due to the different qualities and accuracy of scanners, scanners with relatively low accuracy (especially low resolution) and without data correction after scanning are liable to cause the information loss and coding errors of the original manuscripts, which will affect the measurement.

3.1 Analysis on the information error of images acquired by the scanner

This paper only involves the examination of black and white images; it only discusses the geometric errors, and does not analyze the errors caused by color distortion. There are many factors that cause geometric errors, such as external factors, the scanning direction, the scanning process of the scanner, the internal structure of the scanning system of the scanner, the usage degree of parts in the imaging system and the status of the image sensor. The following two methods of measurement and statistics are used in this paper.

(1) Length measurement statistics means to measure the scanning image of the same object in horizontal or vertical direction. If there are scanning errors, the length of the measured image will be inconsistent in multiple scan cases.

(2) Proportional measurement statistics means to measure horizontal and vertical measurements of the scanning image of the same object. If there are scanning errors, the horizontal and vertical scales of measurements will be inconsistent.

3.2 The influence of warm-up time of the scanner on the scanned image

Scanners' warm-up time refers to the period between the scanner running on power and the sTable and accurate measurement of the image. In order to find out the warm-up time of the scanner, the following experiments are carried out. The experimental sample material filled with the character of "one" produced by HP LaserJet Pro M1136 is shown as following.

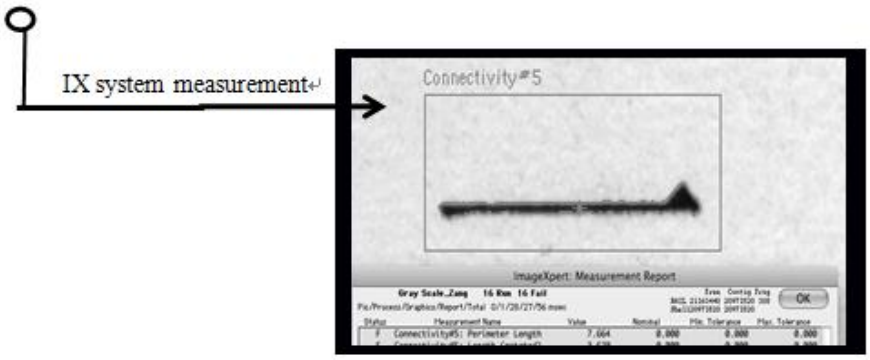


Figure 7. The measurement sample for the scanner's warm-up time

The document is scanned with a scanner at 800dpi in different time intervals; the perimeter and length of the word are measured through connectivity characteristics. The following measurement data are obtained, in which the scanning time 0 represents the actual length of the printed image and text.

Table 5. Measurement of warm-up time before scanning

Warm-up time (min)	1	2	3	5	10	15	20
Measuring length	3.598	3.598	3.598	3.598	3.598	3.620	3.598
perimeter	7.561	7.617	7.664	7.733	7.664	7.637	7.52

(attached Table)

Warm-up time (min)	25	30	35	40	45	50
Measuring length	3.598	3.598	3.598	3.598	3.598	3.598
perimeter	7.576	7.564	7.561	7.571	7.626	7.626

The measurement results of scanning the Chinese character “one” at different time after the starting up of the scanner show that the measured length of the word “one” basically remains unchanged after the starting up of the scanner, but there are also small fluctuations within the first 20 minutes; the measured perimeter of the word “one” varies greatly after the starting up of the scanner. Meanwhile, the shape of the data graph is changing slightly. After 20 minutes, the fluctuation of the length of the perimeter slows down obviously.

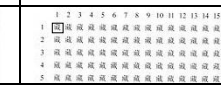
3.3 The influence of different scanning resolutions on feature measurements

ISO13660 Digital Quality Printing Standard stipulates that the minimum scanning resolution should be 600 dpi when analyzing digital images. In order to meet different scanning requirements, in this paper, the scanning system is set up with six different scanning resolutions: 50 dpi, 100 dpi, 200 dpi, 400 dpi, 800 dpi, 1200 DPI and 2400 dpi. But for IX system, using high scanning resolution also brings about defects such as long scanning time, less scanning contents, large scanning image and difficult measurement. In order to make up the shortcomings of high scanning resolution, the following experiments are carried out.

3.3.1 Scanning time measurement

Now we measure the scanning time needed for the same scanning area with a scanning resolution of more than 400 dpi. The following data can be obtained.

Table 6. Measurement of scanning time, resolution and maximum scanning area

Resolution (dpi)	400	800	1200	2400
Scanning time (s)	15	60	125	450
Maximum scanning area				

Then the scanning time of the same size scanning area is measured with a scanning resolution of more than 400 dpi. The scanning area is a Chinese character “Zang” as shown in the following figure. The following data can be obtained.

Table 7. Scanning time and resolution measurement

Resolution (dpi)	400	800	1200	2400
Scanning time (s)	5	6	9	30

It can be seen that the scanning resolution is directly proportional to the scanning time and image quality, and inversely proportional to the size of the scanning area. In order to improve the timeliness of actual research and case handling, we need to balance factors of scanning resolution, scanning time and scanning range to find a suiTable scanning resolution. According to the recommendation of IS013660 standard, an appropriate resolution should be found from the range of 800 dpi to 1200 dpi. It is also necessary to test the relationship between random error and scanning

resolution.

3.3.2 Relationship between random error and scanning resolution

According to the method proposed by Quan-fa Zhang and his colleges in determining the relationship between random error and resolution, the following measurement experiments are carried out. A circular image is printed with HP LaserJet Pro M1136 and scanned for five times at 800 dpi, 1200 dpi and 2400 dpi resolution respectively. The radius length of the scanned image in the horizontal direction (represented by the letter of a) and the vertical direction (represented by the letter of b) are measured and the standard deviation is calculated as following.

The following data are obtained through measurement.

Table 8. The measurement and analysis of random error

Scanning resolution	800		1200	
	a	b	a	b
The first scan	5.355	5.292	8.806	7.916
The second scan	5.355	5.292	8.806	7.916
The third scan	5.355	5.292	8.806	7.916
The fourth scan	5.355	5.292	8.806	7.916
The fifth scan	5.355	5.292	8.806	7.916
mean square error	0	0	0	0

According to the mean square error calculated from the measurement data, it can be seen that the scanning does not affect the length of a distance. According to the measured values of a, b and a/b, it can be seen that the measurement values are different with the same resolution, but the differences are very small and do not have great impacts on the measurement. Therefore, we should try our best to select the same direction of the characters or strokes, or find the same character or the same part of the character in measurement.

4. Conclusions

Through above experiments about factors influence the measurement of scanned characters, following conclusions can be drawn.

1) The scanning system equipped in the IX system is a scanning system with high precision.

2) The scanning environment, the scanning resolution, random errors, mechanical errors and internal construction errors have little influence on the measurement, which can be neglected. The measured data does not need to be corrected.

3) In order to obtain data with high precision and high stability, the following settings are recommended. The 800 dpi resolution should be selected for scanning; the 1200 dpi can be used when measuring the edge of strokes; the scanning is recommended to be carried out after 20 minutes after the scanner is opened; strokes in the same Chinese character or in the same direction should be measured.

Acknowledgements

In this paper, the research was sponsored by the Foundation for School-level Scientific Research Projects of Shandong University of Political Science and Law. The fruit was achieved by the Academic Innovation Team of Young Teachers Engaging in Forensic Science and Evidence Research, Shandong University of Political Science and Law.

References

[1] Karunasekera S A, Kingsbury N G, A distortion measure for image artifacts based on human visual sensitivity [J]. IEEE Int Conf. Acoust, Speech and Signal Processing, 5 (1994) 117-120.

[2] Chou C H, Li Y C, A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile [J]. IEEE Trans. Circuits and Systems for Video Tech, (1995) 467-476.

[3] Wang Z, Bovik A C. A universal image quality index [J]. IEEE SignalProcessing Letters, 9(2002) 81-84.

[4] Webster A A, Jones C T, M H Pinson, et al, An objective video quality assessment system based on human perception [C]. SPIE, 1913(1993) 15-26.