Print Quality Evaluation of a Large Number of Data

Issei Yamasaki* and Taizo Iijima**

1. Introduction

The method for print quality evaluation described in "Printing Specifications for Optical Character Recognition" [1] which is recommended by ISO seems to be too microscopic and has little reproducibility of measured values. A method for quantitative evaluation of printed images has been proposed [2]. This paper discusses the feasibility of the method described in the previous paper [2] using 'model' data and actual data. Please refer to previous paper [2] for details of evaluation method.

2. 'Model' data

'Model' data are patterns with artificial noises (cf. Fig. 1). Artificial noises are added manually to the standard pattern with nominal strokewidth, considering mean darkness and the centroid of each 'model' pattern are equal to those of the standard pattern with nominal strokewidth. Especially, the patterns T-1, T-2, and T-3 are equally distorted each other from the standard pattern with nominal strokewidth.

Fig. 1 'Model data (Characters with artificial noises).

3. Sampling Interval and Classes of Strokewidths of Standard Pattern

To determine classes of strokewidths which should be prepared for the purpose of estimating the similarity when the darkness of input pattern coincides with that of standard pattern, the following

---

This paper first appeared in Japanese in Joho－Shori (Journal of the Information Processing Society of Japan), Vol. 13, No. 8 (1972), pp. 525～532.

* Information Science Division, The Electrotechnical Laboratory
** Faculty of Electrical Engineering, Tokyo Institute of Technology

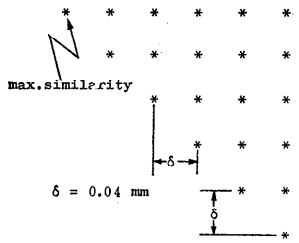

Fig. 2. Relatively translated positions of standard pattern and an input pattern

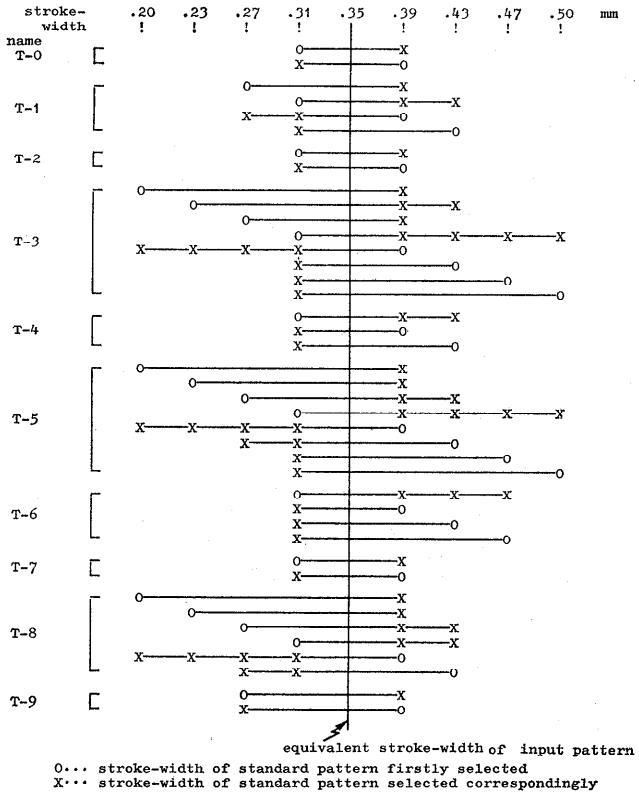

Fig. 3. Usable strokewidth at average estimation error under one per cent and the maximum estimation error under five per cent.

experiment has been made with a computer program using 'model' data. The similarity between an input pattern and the standard pattern at

equal mean darkness is estimated. The input pattern and the standard pattern are relatively translated previously as shown in Fig. 2, and strokewidths of standard patterns prepared for estimation are changed. The average (with respect to relative positional translation) of the estimation is less than or equal to 1.0 per cent and the maximum error is less than 5.0 per cent, if the increment of strokewidth prepared for standard pattern is 0.08 mm (cf. Fig. 3).

Considering a stroke is thickened or thinned symmetrically with respect to the center-line of the stroke, the sampling interval should be 0.04 mm at most for the purpose of preparing standard patterns which have strokewidth every increment of 0.08 mm.

4. Interval of Calculating Similarity for Estimating True Maximum Sililarity and Its Position

For the purpose of determining the interval —— estimation interval d —— of calculating similarity for estimating true maximum similarity and its position, the following experiment is made with a computer program using 'model' data. 'Model' data are previously translated relatively as shown in Fig. 4, and the value of true maximum similarity and its position are estimated. Results are shown in Table 1. Table 1 results follows: Firstly, estimation error of true maximum similarity in any translated positions does not change significantly, even if the estimation interval is varied from 0.08 to 0.20 mm. Secondly, the estimation error of the position of true maximum similarity, especially in the case of translation on diagonal (① and ③

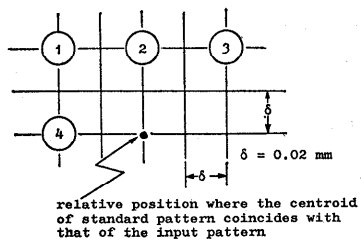

Fig. 4. Relative positions (① ~ ④) of a standard pattern and an input pattern previously translated.

in Fig. 4), is approximately ten times larger, if the estimation interval is 0.12 mm or longer. Therefore, it is reasonable to take 0.08 mm as the estimation interval d.

Table 1. Estimation Errors of the Maximum Similarity and Its Position in the Case of Varying the Strokewidths of Standard Pattern

TRANSLATED POSITION		ESTIMATION ERRORS											
	ESTIMATING DISTANCE (GRID POINT)	SIMILARITY				POSITION-X				POSITION-Y			
		4	6	8	10	4	6	8	10	4	6	8	10
①	MEAN	4.19	4.59	4.39	4.41	0.66	4.42	3.97	3.75	0.61	3.34	7.07	5.27
	MAX.	7.44	7.87	7.87	7.79	2.35	9.44	9.18	8.48	2.20	10.08	36.70	20.07
	MIN.	1.50	1.73	1.37	1.84	0.00	0.11	0.14	0.23	0.00	0.57	0.55	0.89
②	MEAN	1.57	1.79	2.01	1.88	0.78	0.53	0.45	0.33	0.00	2.20	5.76	4.69
	MAX.	4.29	4.29	4.21	3.98	2.56	1.52	1.43	0.96	0.00	6.94	38.03	21.81
	MIN.	0.00	0.07	0.02	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.11
③	MEAN	4.03	4.33	4.05	4.00	0.66	2.42	1.86	1.57	0.63	2.71	6.40	5.71
	MAX.	7.28	7.89	7.85	7.62	2.38	9.44	9.18	6.88	2.30	6.01	36.70	20.07
	MIN.	1.50	1.77	1.37	1.80	0.00	0.11	0.14	0.23	0.00	0.64	0.55	0.89
④	MEAN	2.53	2.83	2.87	2.85	0.00	2.62	2.36	2.08	0.91	0.83	0.75	0.65
	MAX.	4.63	4.63	4.42	4.49	0.00	8.77	8.54	6.24	2.38	2.31	2.27	2.25
	MIN.	0.65	0.77	0.77	0.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

(%)

5. Evaluation of Characters Printed with a Line Printer

5.1 Acquisition of Data

The method for evaluation of print quality of a character pattern has been tested with a computer program using actual data. Actual data are OCR-A numerals printed with a line printer. Sample sheet (cf. Fig. 5) is scanned by a CRT flying spot scanner, and video signal is sampled every interval of 0.04 mm in horizontal and vertical directions, and the sampled signal is converted into sixty-four-level digital signal. Total number of data is one hundred and fifty.

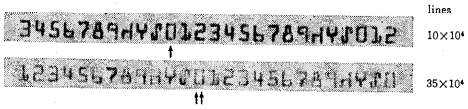

Fig. 5. Test sheet samples.

5.2 Threshold value

Shape of character, and the quantity of spots and voids are varied with the level of threshold. The threshold value θ, taking the peak value of print contrast signal PC_{peak} in the domain R as a standard,

is determined as follow:

$$\theta \equiv \begin{cases} 0.3 & , \quad 0 < PC\beta_{peak} \leqq 0.525 \\ PC\beta_{peak} / 1.75 & , \quad 0.525 < PC\beta_{peak} \leqq 1 . \end{cases}$$

Example of two-valued black and white character pattern are shown in Fig. 6.

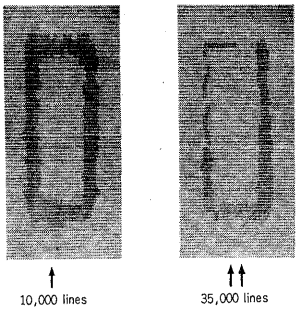

Fig. 6. Examples of two-valued black and white character patterns.

5.3 Procedure of Evaluation

The outline of the evaluation procedure is shown in Fig. 7.

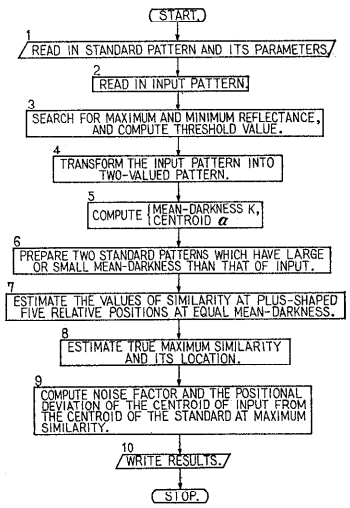

Fig. 7. Simplified procedure of print quality evaluation.

6. Results and Conclusion

Noise factor ε^2 and positional deviation D which expresses the

distance between centroids of input and standard patterns at true maximum similarity, are shown in Fig. 8. The average (with respect to all of characters printed on the same line), and values for characters '0' and '7', for example, are shown. Variations of the peak value PCS_{peak} of print contrast signal in the domain R, mean darkness K, and threshold value θ with printed lines are shown in Fig. 9.

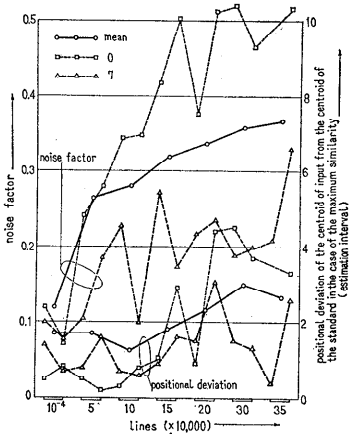

Fig. 8. Result of evaluation for OCR-A numerals.

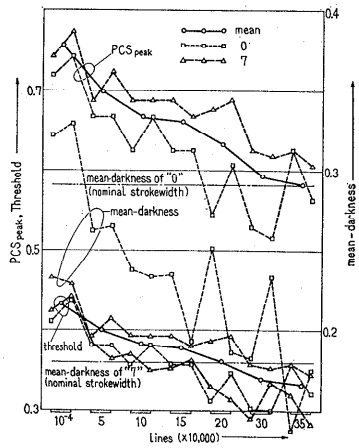

Fig. 9. Variation of PCS_{peak}, threshold, and mean darkness with printed line.

Values of ε^2, D, PCS_{peak}, K, and positional deviation d serve for the purpose of quantitative print quality evaluation.

References

1] ISO RECOMMENDATION R 1831: "Printing Specifications for Optical Character Recognition" (1971)

2] Yamasaki, I. and Iijima, T.: "A Method for Print Quality Evaluation of a Large Number of Data", Information Processing in Japan, 12, pp. 119-125 (1972)