SELECTION OF THE MOST LEGIBLE MATRIX CHARACTERS IN COMPUTER-MAN COMMUNICATIONS

Ching Y. Suen* and Michael G. Strobel†
*Department of Computer Science, Concordia University, Montréal
†Department of Psychology, University of Montréal, Montréal

Abstract

An experiment was conducted to determine the most legible set of 5 × 7 matrix characters. The characters studied were compiled from an extensive survey of over 30 common computer output systems. A PDP-12 computer was used to store the 120 matrices of 36 alphanumeric symbols and present them to 15 subjects in random order using a precision display unit. Only one stimulus was generated at a time. The duration of display of the matrix was increased until the subject could label it correctly. The experimental data yielded a set of dot-matrices which offers the highest legibility for communications between the computer and human readers.

Résumé

La lisibilité optimale d'un ensemble de caractères formés par des éléments d'une matrice 5 × 7 était investiguée par une expérience avec 15 sujets. On a étudié les caractères provenant de 30 systèmes d'affichage. Les 120 matrices de 36 symboles alphanumériques étaient emmagasinées et présentées par un ordinateur PDP-12 dans un ordre aléatoire sur l'écran, un symbole à la fois. La durée de la présentation a été augmentée jusqu'au point où le sujet pouvait identifier correctement le symbole. Le résultat est un ensemble des caractères qui sont le plus facilement reconnaissables pour être employés dans la communication homme-machine.

INTRODUCTION

In recent years, matrix characters, especially the 5 x 7 font, have widespread use in computer output systems such as CRT's, thermo-printers and high-speed printers (McLaughlin, 1973). This is due to the increasing usage of digital devices, the flexibility of point and graphical presentation and the ready availability of matrix character generators in IC chips. As matrix characters become more widely used, their legibility become imperative. Legible characters do not only provide better reading, they also play a vital role in applications such as aerospace operations where penalties for common misreading errors could be disasterous. However, due to the quantization process, matrix characters are less legible than their printed counterparts. One can easily detect this by comparing one's speed when reading materials printed in these two fonts. Indeed, the importance of the legibility of matrix characters has been investigated by several researchers (e.g., Huddleston, 1971; Ellis et al, 1974; and Riley, 1976).

This paper describes an experiment conducted to determine the most legible set of matrix characters for computer -man communications. In this investigation, an extensive survey of over 30 computer output terminals was conducted. Alphanumeric character sets produced by leading manufacturers of ROM matrices, matrix printers, CRT's and other computer terminals were examined. A few symbols were also invented by the authors. Altogether, 122 different models of the 36 alphanumeric matrix symbols were assembled. They comprised 78 models of letters and 44 models of numerals. Thus, on the average, a symbol was represented by 3.36 different shapes. For example, the letter A had 4 forms as shown below.

```
  *        ***       ***       ***
 * *      *   *     *   *     ** **
*   *     *   *     *   *     *   *
*   *     *****     *   *     *   *
*****     *   *     *****     *****
*   *     *   *     *   *     *   *
*   *     *   *     *   *     *   *
```

METHOD

A PDP-12 computer was programmed to present the matrix characters one at a time on a VR-12 precision display unit. The stimulus pattern appeared as a 5 x 7 dot matrix measuring 2.54 cm (1 in) high. Subjects were seated at approximately 48 cm (19 in) in front of the screen so that the stimulus image subtended an angle of 3^{o} at the fovea.

Luminosity of the dot matrix was adjusted to perceptual threshold at 1 ms duration, in a fluorescent lit environment of 10 to 12 ft-candle illuminance.

Fifteen college students, 9 men and 6 women participated as experimental subjects. Their age ranged from 19 to 35, with a median of 28. All had natural or corrected 20/20 vision.

The subjects were informed of the purpose of the experiment and they were urged to attempt to label the stimulus pattern even in cases when identification was uncertain. Prior to the test, five characters were presented as training stimuli to accustom the subjects to the experimental procedure and to focus their attention to the active area of the screen.

Following a random sequence, the experimenter entered the characters from a teletype. The verbal responses of the subjects were also recorded by the experimenter on the teletype in order to avoid distractions and possible typing mistakes made by the subjects.

During the first trial of the experiment, the complete set of stimuli was presented at 1 ms duration. Correct responses were indicated by the bell of the teletype. Misidentified patterns were presented again at increased durations until all stimuli were recognized without error. The durations used in subsequent trials were: 1, 5, 25, 100, 500, 1000, 4000 ms.

RESULTS AND DISCUSSIONS

Errors made by the subjects over all stimulus durations were tallied. Visual inspection of the scores showed marked differences of performance among the subjects. The range of total errors made by the men extended from 5 to 61 with a mean of 27.2. The number of errors made by the women ranged between 18 and 186, with a mean of 79.2. A t-test showed that the difference in performance between male and female subjects was statistically significant at the 0.05 level. This might be attributed to the level at which the subjects were familiar with matrix characters used in computer outputs.

Several patterns were recognized on the first trial, i.e., no error was ever made by any subject at the 1 ms duration during the first trial: one A, one D, one J, one L, one P, one T, one Z and one 4, two 3's, two 4's, one 5, one 6, and four 7's. The following patterns were recognized without error on the second trial, still at 1 ms duration: all patterns of A, B, D, E, H, I, K, L, O, P, R, T, U, X, Y, and the numerals 1 to 9 with the exception of 6. These patterns constitute the set which is least difficult to design. Table 1 shows the patterns selected for an alphanumeric display of high legibility. When variations of the same character were found equally legible during the experiment, the one which was judged to be most pleasing was selected. Within this set of patterns, M, S, V, W and particularly Q seem to be the most difficult ones to identify and a special design effort is required. The numeral Ø caused a special problem to the subjects. Some subjects were not familiar with the discriminating slash, while others were influenced by the FORTRAN convention to slash the letter O.

The patterns which gave rise to more than 10% of wrong identifications could be classified into two categories: those which were predominantly confused with another symbol as shown in Table 2; and those which were hard to label at all, resulting in a wide spread of different responses as shown in Table 3. In both tables, the frequency of misidentifications is shown in brackets. Note that one pattern of V falls in both categories. Note also that the two designs for the letter Q appear in the list of unacceptably high recognition errors.

As this experiment was progressing, the matrices were also tested by using distance and information measurements (Suen et al, 1976). The results show that there are some discrepancies between the optimal set selected in this study and the optimum set selected by the theoretical approach of distance and information measurements. A thorough investigation of the problem is underway and the findings will be reported in the future.

REFERENCES

1. Alphanumeric Displays, Auerbach Publishers, New York, 1971.
2. McLaughlin, R. A. "Alphanumeric display terminal survey," Datamation, 71-92, Nov. 1973.
3. Huddleston, H. F. "An evaluation of alphanumerics for a 5 x 7 matrix display," Proc. Conf. on Displays, IEEE Publication 80, 145-147, Sept. 1971.
4. Ellis, B., Burrell, G. J., Wharf, J. H., Hawkins, T. D. and Peters, D. V. "The performance in illuminances up to 80,000 lux of a light emitting diode display having a 3 mm character height," Proc. SID, vol. 15, 150-160, 1974.
5. Riley, T. M. "Dot-matrix alphanumeric identification as a function of font and discrete element degradation," Proc. SID International Sym., 110-111, 1976.
6. Suen, C. Y., Shiau, C., Shinghal, R. and Kwan, C. C. "Reliable recognition of handprinted characters," Proc. Joint Workshop on Pattern Recognition and Artificial Intelligence, 98-102, June 1976.

ACKNOWLEDGEMENTS

The authors would like to thank various computer terminal manufacturers, too numerous to be identified here, for providing data on their matrix character sets.

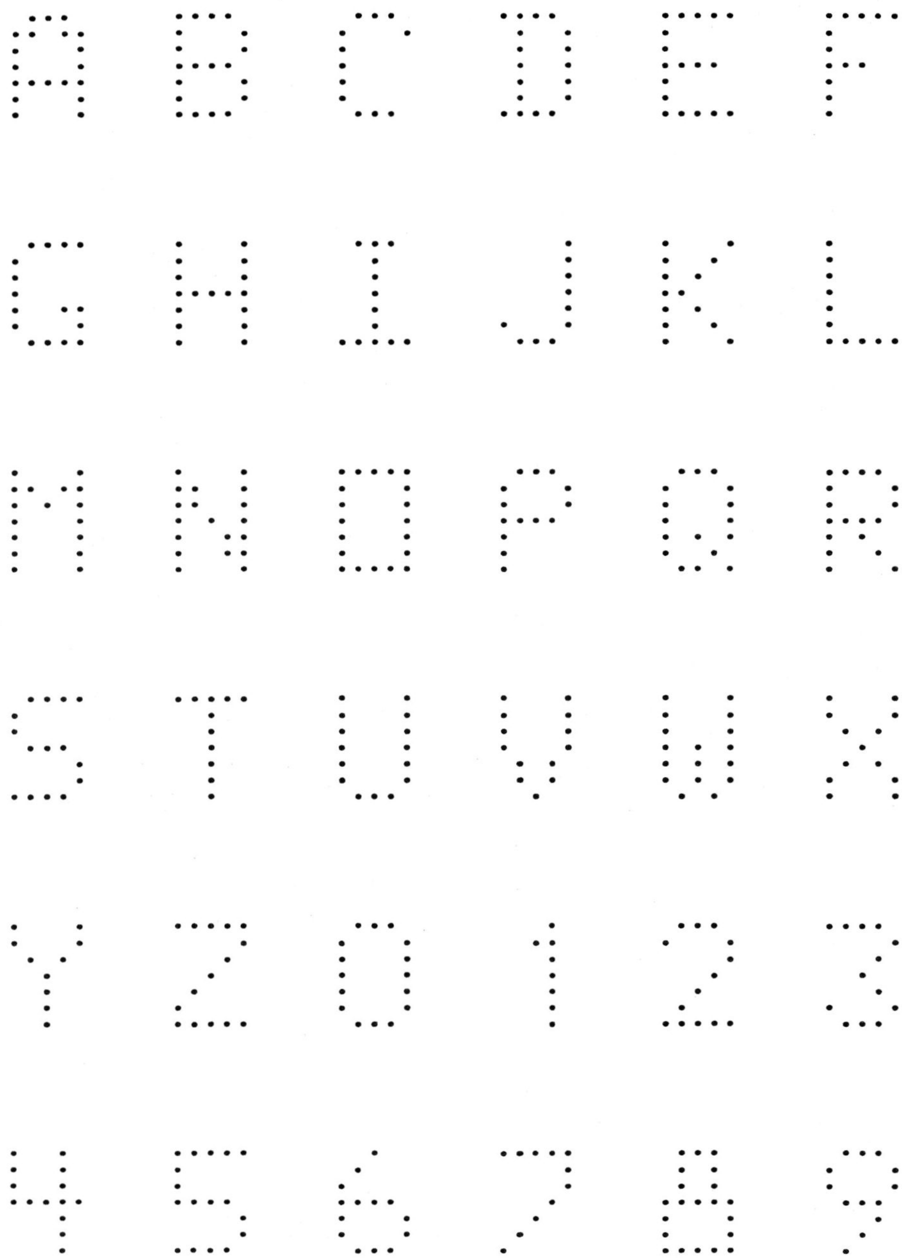


Table 1. Alphanumeric patterns of high legibility.

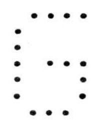
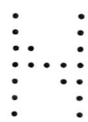
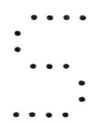
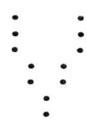
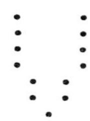
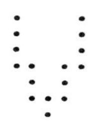
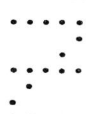
Pattern	Error				
	G (8)	E	?		
	6 (7)				
	H (23)	W (3)			
	5 (9)	E			
	Y (20)	H	?		
	Y (10)	U (2)	?		
	U (10)	Y (2)	?		
	Y (12)	W (2)	? (3)	H	D
	2 (9)	5 (2)	? (3)	S	E

Table 2. Consistent or dominant confusions.
Comment: Frequency in brackets, ? = unidentified.

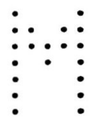
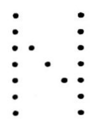
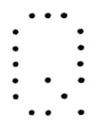
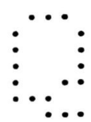
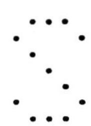
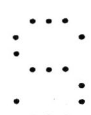
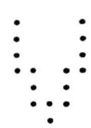
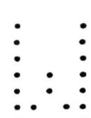
Pattern	Error				
	H (5)	W (2)	F	?	
	H (5)	W (4)	? (2)	T	
	K (3)	? (3)	G (2)	Ø (2)	O
	? (7)	K (2)	Ø (2)	D	E
	8 (9)	3 (2)	5	X	?
	8 (5)	9 (2)	W (2)	H	D
	Y (12)	? (3)	W (2)	H	D
	H (4)	? (3)	I		

Table 3. Inconsistent or scattered confusions.
Comment: Frequency in brackets, ? = unidentified.