Information Technology, J. Moneta (editor)

OPTICALLY WEIGHTED DOT-MATRIX FARSI AND ARABIC NUMERALS

BEHROOZ PARHAMI

Arya-Mehr University of Technology
Tehran, Iran

Abstract: Representation of digits by optically weighted symbols is a way of presenting information in combined digital and analog form. Such symbols can be used to display three-dimensional patterns on conventional printers, with the third dimension visualized by the darkness of printed page areas. In this paper, a general framework for the design of optically weighted numerals is established and several examples are given to illustrate the general concepts. It is shown that Farsi and Arabic numerals lend themselves very well to optically weighted representation and design examples are given for both decimal and hexa-decimal digits.

Index Terms: Arabic output, character set design, Farsi output, numeric output, optical distance, optically weighted numerals, output devices, printer output, three-dimensional output.

INTRODUCTION

Numerals with area coverage (darkness) proportional to the digit values provide an interesting way of representing information in combined digital and analog form. Such optically weighted numerals can be used to represent three-dimensional patterns on conventional printers; the darkness of printed page areas makes the third dimension visible, provided it varies in a limited range. This is particularly helpful for extracting macro information from large volumes of numerical data. If the optically weighted numerals resemble the conventional ones, an easily readable printout will result which also provides accurate digital data when desired.

Optically weighted numerals have been used before (Figure 1). The purpose of this paper is to provide a general framework for the design of such character sets and to apply the proposed guidelines to the design of several sets of optically weighted Farsi and Arabic numerals [3].

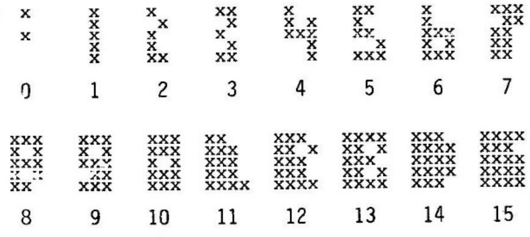

Figure 1. Optically Weighted Representation of Hexa-Decimal Digits [1].

Figure 2 shows the decimal Farsi and Arabic numerals. It appears that Farsi and Arabic numerals are quite suitable for optically weighted representation because of the following properties:

1. The normal representation of zero covers the smallest area among all numerals.

2. The alternative representations of Figure 2 provide some flexibility, in the sense that the more convenient representation may be used in each case.

3. Except possibly for '9', the numerals may be easily rewritten such that the area coverage increases with the value represented.

Farsi and Arabic Representation	.	١	٢	٣	٤	٥	٦	٧	٨	٩
Alternative Farsi Representation					۴	۵	۶			
Digit Value	0	1	2	3	4	5	6	7	8	9

Figure 2. The Ten Decimal Farsi and Arabic Numerals.

Let us first discuss, in relatively general terms, some design considerations for optically weighted numerals and then proceed with specific designs for Farsi and Arabic languages.

GENERAL DESIGN CONSIDERATIONS

Let us assume that we wish to design optically weighted representations for radix-r digits (0, 1,2,...,r-1). Each digit will be represented by an m by n matrix (m rows and n columns) of black and white squares, with the number of black squares increasing with the digit value x. The number of black squares will be taken as a linear function of x to create the best visual effect. That is

$$f(x) = ax + b \quad (1)$$

where a and b are positive, non-zero integers. Therefore, we must have:

$$mn \geq r \quad (2)$$

Since the height of each numeral is usually greater than its width, we also require that:

$$m > n \quad (3)$$

To provide the best contrast in the printed page, it is desirable that b be as small and $a(r-1)+b$ as large as possible. This can be achieved by selecting:

$$a = \lceil mn / (r-1) \rceil - 1 \quad (4)$$

Then, b can be selected in the following range:

$$1 \leq b \leq mn - a(r-1) \quad (5)$$

The best choice for b is probably somewhere near the middle of its allowable range, since too small a value will reduce the darkness of the representation for r-1 and too large a value will limit our ability to provide a distinctive shape for the representation of r-1.

Once a and b have been selected, the area for each of the r numerals is known from (1) and the next step is to select dot patterns with the given areas in such a way that the shape of each pattern resembles that of the corresponding numeral.

In addition to ease of recognition, another criterion that can be used to evaluate a given set of symbols is their distinguishability, measured by the distance between pairs of symbols. Given two symbols, we define their <u>(optical) distance</u> as the minimum number of positions (squares) in which they differ when both symbols are shifted horizontally and vertically to occupy every possible position within the m by n frame. The shifting is required since the shapes of symbols are important in their recognition and not their positions within the matrix.

To verify that the minimum distance in a set of symbols is no less than d, we need only compare each symbol with the next $\lfloor d/a - 1 \rfloor$ numerals in the sequence, as the following symbols are certain to satisfy the condition because of the difference between covered areas being at least d. To determine the distance between two symbols c_1 and c_2, a simple procedure is to superimpose them (without rotation) such that a maximum number of black squares overlap. The number of non-overlapping black squares in both symbols is the distance between c_1 and c_2, which is denoted by $d(c_1,c_2)$. As an example, the minimum distance for the numeral set in Figure 1 is d = 3.

It can be shown that for numeral sets of the type considered here, the distance d that can be achieved is bounded by the following inequality:

$$d \leq a + 2\min[(b-1),(mn - ar + a - b)] \quad (6)$$

This is easily provable by considering the worst-case conditions for d(0,1) and d(r-2,r-1).

<u>APPLICATION TO FARSI AND ARABIC NUMERALS</u>

We now apply the concepts presented so far to the design of several sets of Farsi and Arabic numerals (Figure 2). Assuming r = 10, Inequalities (2) and (3) show that the smallest possible matrix has m = 4 and n = 3. We will then have a = 1 and $1 \leq b \leq 3$. Selecting b = 1 will result in f(0) = 1 and f(1) = 2, causing d(0,1) = 1, regardless of the symbol designs. This is clearly undesirable. Selecting b = 3, on the other hand, will result in f(9) = 12. Therfore, the representation for '9' covers the entire matrix which, in addition to having no distinctive shape, results in d(8,9) = 1. The above distances can also be found from Inequality (6).

Hence, we select b = 2. Inequality (6) indicates that the minimum distance can be no more than three. In fact, a distance of three can be achieved but not with easily recognizable numerals (Figure 3). If a smaller distance is acceptable, somewhat more readable numerals can be designed. However, a larger matrix appears to be necessary for easily readable numerals and, therefore, we will not elaborate on this case any further.

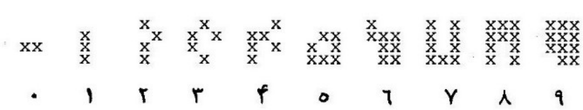

Figure 3. Distance-3 Decimal Farsi and Arabic Numerals in a 4 by 3 Matrix.

Let us now try m = 5 and n = 4. We have a = 2 and $1 \leq b \leq 2$. We choose b = 1. By Inequality (6), the minimum distance can be no more than 2. Figure 4 shows a possible numeral set with d = 2. One desirable property of the symbol set of Figure 4, which makes the numerals more readable, is that all the black squares are solidly connected to each other, while in Figure 3 some squares are touching only at the corners.

Now consider an even larger matrix with m = 7 and n = 5. We have a = 3 and $1 \leq b \leq 8$. Experimenting with different values of b shows that b = 4 and b = 3 are the most suitable choices. Figure 5 shows a possible set of numerals for b = 4. Removing the black squares marked by an 'o' from Figure 5, we obtain a possible set of numerals for b = 3. Both sets have a minimum distance of d = 5 and appear to be equal with respect to readability. Also, in both cases, the alternative representations for '4', '5', and '6' are more readable than those in the first row of Figure 5.

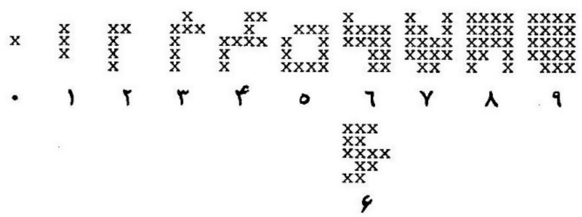

Figure 4. Decimal Farsi and Arabic Numerals in a 5 by 4 Matrix.

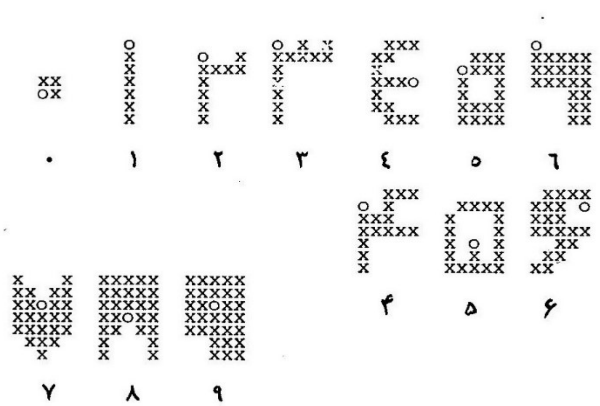

Figure 5. Decimal Farsi and Arabic Numerals in a 7 by 5 Matrix.

Hexa-decimal digits offer a wider range of values for the darkness of the printed characters. We, therefore, consider sets of optically weighted hexa-decimal numerals next.

The hexa-decimal numerals of Figure 1 do not satisfy the conditions stated in the previous section. A somewhat more readable set can be designed by following the stated guidelines. Taking the same values for m and n (m = 5, n = 4), we find a = 1 and 1 $\leq$ b $\leq$ 5 (r = 16). Choosing b = 3, we can construct the numeral set shown in Figure 6. Alternatively, we can use letters 'A' through 'F' for digits '10' through '15' (Figure 6). In either case, the minimum distance is d=3.

Farsi and Arabic hexa-decimal numerals can be represented similarly in a 5 by 4 matrix. Figure 7 shows a possible design for Farsi hexa-decimal numerals with a = 1 and b = 3. For this set, the minimum distance is d = 3.

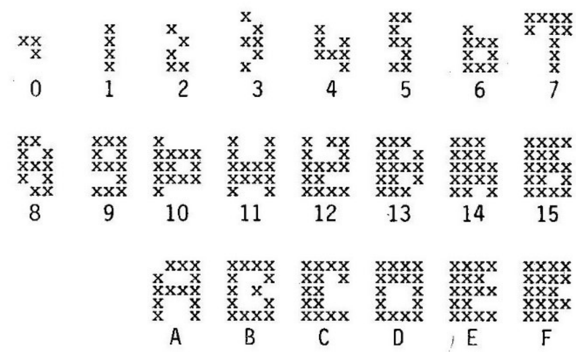

Figure 6. Hexa-Decimal Numerals Represented in a 5 by 4 Matrix.

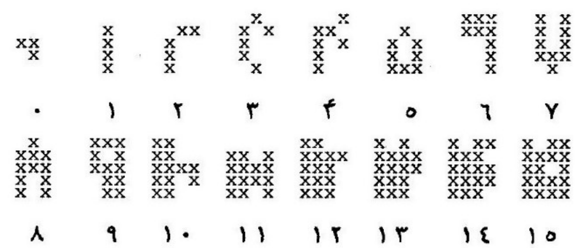

Figure 7. Farsi and Arabic Hexa-Decimal Numerals in a 5 by 4 Matrix.

Figure 8 shows Farsi and Arabic hexa-decimal numerals in a 7 by 5 matrix. Here, we have a = 2 and 1 $\leq$ b $\leq$ 5. Selecting b = 3, we can construct a numeral set with a minimum distance of d = 4, as shown in Figure 8 (two representations for the digit '6' are given on the second row). This set is easily readable partly because all black squares in each symbol are solidly connected to each other.

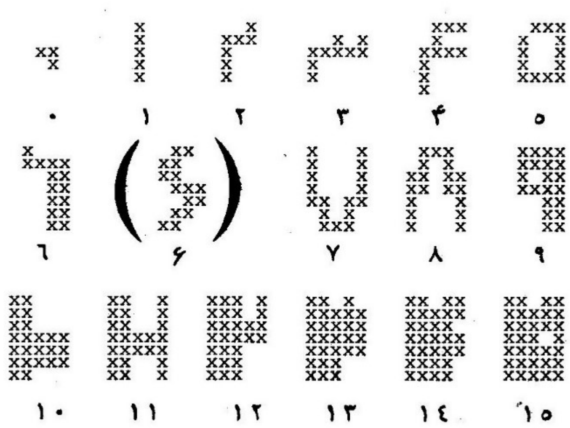

Figure 8. Hexa-Decimal Farsi and Arabic Numerals in a 7 by 5 Matrix.

CONCLUSION

In this paper, we have presented some general guidelines for the design of optically weighted radix-r numerals and demonstrated the general approach through specific examples. It was shown that Farsi and Arabic numerals lend themselves very well to optically weighted representation and several examples of the many possible designs were given. This result is a pleasant contrast to the numerous difficulties in conventional Farsi and Arabic input and output [2], [5], [6] and in the design of line-segment displays [4]. The same techniques are applicable to larger matrices and will undoubtedly result in much more readable symbols.

We have only discussed dot-matrix patterns. It is also possible to construct numerals with smooth edges so that the area they cover increases linearly with the values represented. For this purpose, and also for dealing with large dot-matrix patterns, automated design aids are clearly needed for experimenting with the many possible design choices.

REFERENCES

[1] "Display Technique Presents Data in Four-Dimensional Analog/Digital Patterns," Computer, Vol. 6, No. 6, pp. 43-44, June 1973.

[2] Mavaddat, F. and B. Parhami, "Informatics in Iran: Problems and Prospects," Proceedings of the International Conference on Computer Applications in Developing Countries, Bangkok, Thailand, August 1977, pp. 121-133.

[3] Parhami, B., "Representation of Digits by Optically Weighted Dot-Matrix Characters," Computer Systems Laboratory, Arya-Mehr University of Technology, Technical Report CSL-75-006, May 1975.

[4] Parhami, B., "Low-Cost Output Displays for Microcomputer Applications," Proceedings of the Second India Symposium on Computer Architecture and System Design, New Delhi, India, November 1976, pp. 111-119.

[5] Parhami, B., "On the Use of Farsi and Arabic Languages in Computer-Based Information Systems," Presented to the Symposium on the Use of Indian Languages in Computer-Based Information Systems, New Delhi, India, April 1978.

[6] Parhami, B. and F. Mavaddat, "Computers and the Farsi Language: A Survey of Problem Areas," Information Processing 77 (Proceedings of IFIP Congress, Toronto, Canada, August 1977), North-Holland, Amsterdam, 1977, pp. 673-676.