J.King Saud Univ., Vol.7, *Comp. & Inf. Sci.*, pp. 1-11 (A.H 1415/1995).

An Intelligent Arabic Text Editor for the Motor Impaired

A. Abdenbi Mandar

College of Computer and Information Sciences,
King Saud University, P.O.Box 51178 ,Riyadh ,11543, Saudi Arabia

(Received 19 October 1992; accepted for publication 21 June 1993)

Abstract. We describe a dynamic array editing method of Arabic alphanumeric alphabet with the aid of an on-off sensor controlled by a motor impaired person. The main work presented in this paper is the development of an intelligent editor that adapts itself to the user. Two modes are possible with this editor: the first, the normal mode intended for handicapped users with normal intellectual abilities; the second, learning mode is to be used by handicapped users with low intellectual abilities. Both modes are based on statistical data on the Arabic language collected from various texts for the first mode and from a limited and almost finite language for the second.

Introduction

A number of computerized communication aids for the motor impaired are currently available in which the handicapped communicates with the system without the keyboard [1,2]. The method used in general is as follows: a set of items is displayed on a tablet or a screen and scanned in closed loop. The user can then select the data item required which can be a character or a word. Usually the selection is done using a push button or an On-Off sensor adapted to the handicapped user [3,4,5].

Among the available aids, one can find editors designed to be used by persons with more than one handicap [6]. In this paper we present an intelligent Arabic text editor for the motor impaired with or without a low mental handicap [6].

We describe a dynamic array edition method of alphanumeric alphabet with the aid of an On-Off sensor controlled by the impaired. The matrix is scanned and when a character is selected, a new matrix is presented to the user with its elements arranged by decreasing probability. The first character is the most probable after the selected one.

The system can operate in two modes:

- *Normal user mode (NUM)* : This mode is intended for handicapped users with normal ntellectual ability. In this case, a table of statistical data collected previously from lifferent kinds of texts is stored in the system and used by the editor.

- *Learning mode (LM)* : This mode is intended for handicapped users with intellectual ability lower than normal. In this case, the user usually has a limited vocabulary, so the tatistical data is built from the text he creates.

This method allows for considerable increase in the man-machine communication peed compared to methods using a fixed matrix.

In this paper we present first the principle of the method, then we give the tatistical matrix of the Arabic texts, and finally we illustrate the method by giving an xample in the NUM.

Method Principle

) Statistics

Various types of texts, were used to build the statistics matrix. Its elements are lescribed below:

	$a_{01,01}$	$a_{01,02}$	$a_{01,41}$
	$a_{41,01}$	$a_{41,02}$	$a_{41,41}$

Forty- two Arabic characters are needed to represent any Arabic text.

Fig.1. Statistics matrix

Each element ($a_{i,j}$) represents the frequency of appearance of the character j after haracter i.

The above matrix is represented for the alphabetical characters only but it can be xtended to punctuation marks (space, period, comma, etc..).

b) Dynamic matrix vectors

From the above matrix the dynamic matrix vectors are built such that each vector Vxi = xi1 xi2 xi3 xin where xi1 is the most probable letter after xi and xin is the least probable.

$$P(xi_1/xi > P(xi_2/xi > \ldots . \quad \ldots > P(xi_n/xi \qquad (1)$$

These vectors will be displayed one at a time in a form of an array. During the editing, this array is scanned and after the selection of character xj, the next array displayed is represented by the vector Vxj. This process is described bellow.

The first array displayed when the system starts is represented by vector V[space]. Then after the selection of character xi, the corresponding array is displayed:

xi_{01} xi_{02} xi_{03} ... ...
xi_{08} xi_{09} xi_{10} ... ... This array represents the vector V_{xi} in two dimensions
xi_{21} xi_{22} xi_{23} ... ...

During editing the text created is displayed and the user has the option either to store it on the system disk or print it.The selection is done using an on-off sensor. Different sensors can be used with the system depending on the nature of handicap. As explained before, the system can operate in two modes:

- **NUM** or normal user mode
- **LM** or learning mode

Normal User Mode

The statistics matrix and the dynamic matrix vectors are determined from various texts on different subjects such as science, sports, religion, etc... The dynamic matrix vectors are determined once and cannot be changed. This mode is intended for handicapped users with normal intellectual abilities. Figure 2 describes the system and shows that the statistics matrix is determined at state (2) and the system remain in state (3), the dynamic matrix editor.

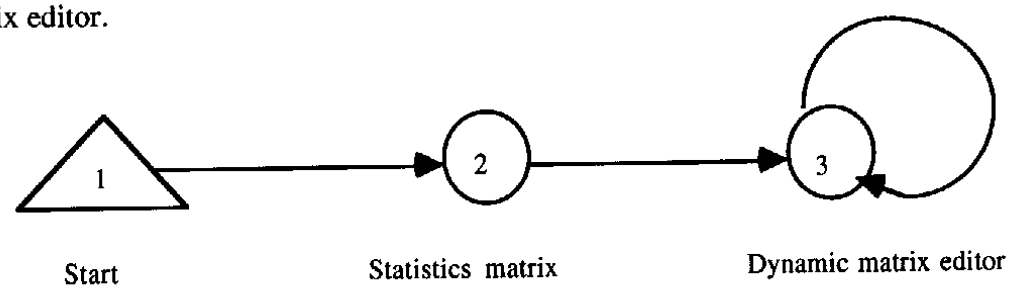

Fig. 2. Normal user mode

Learning Mode

This mode is intended for users with limited intellectual ability. The system for this mode is described in Fig.3. As shown an initial matrix represented by one dynamic matrix vector is used for the first editing. The initial matrix has its elements ordered alphabetically.

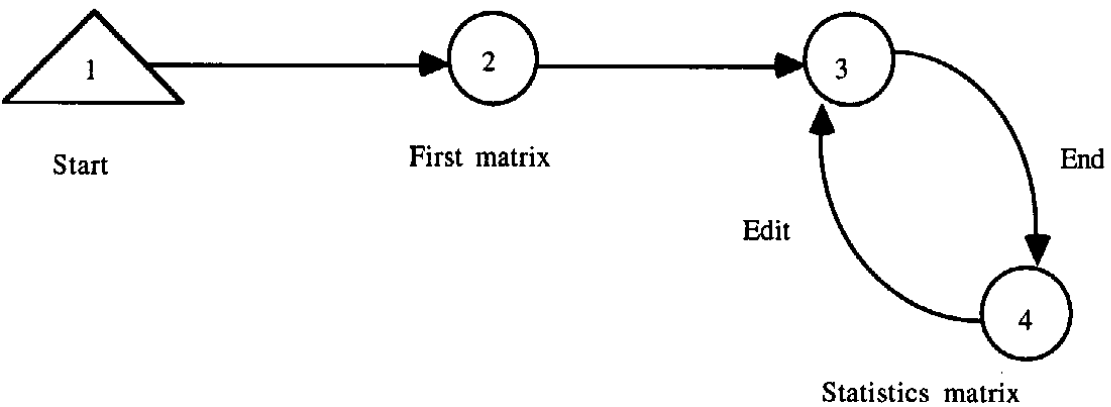

Fig. 3. Learning mode

From the text edited a statistics matrix is generated which will be used for the next edition. From each text edited the elements of the statistics matrix are recalculated.

In both cases-**NUM** and **LM** once the user selects the desired character or a punctuation mark which is appended to the end of the message, a new matrix is displayed depending on the selected character. When the composed text reaches a given length it can be either printed or stored on the system disk. Figure 4 gives an overview of the system.

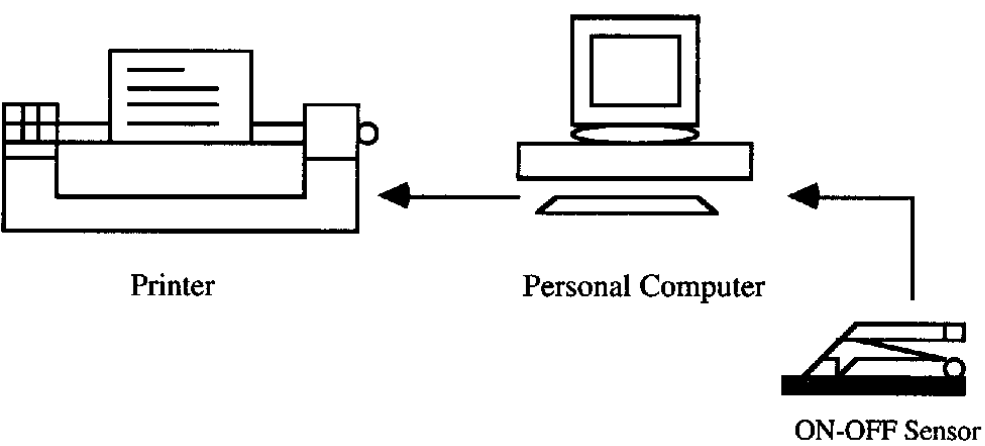

Fig. 4. The system hardware

Cursor speed control

The user can control the length of time that the cursor stays at one character of the array before moving to the next. This feature allows the user to increase his throughput as his skill increases.

Statistics Matrix

Appendix 1 shows ten rows of the statistics matrix built from a text of about 16,000 characters. A longer text would give a more accurate matrix and hence a faster editor. A zero means either the character of the column is never placed after the character of the corresponding row or that the text is not long enough. The sign "_" represents the space.

The corresponding dynamic matrix vectors built from the matrix described above are shown in Appendix 2.

Editing Speed

Let C be the number of characters scanned before the selection of a given character. C is inversely proportional to the speed of characters selection.

To compose the sentence: بسم الله الرحمن الرحيم using the proposed method we found C= 42. If we compare this number to the one that could be obtained from a matrix in which the characters are ordered by decreasing probability of appearance (we suppose that such a matrix is the best one using a static method) we find that C= 63.

We compared the cost of some randomly chosen words and we can conclude that our method increases the speed of editing by at least 50 % [7,8].

Conclusion

In this paper we proposed a method for Arabic text editing for the motor impaired. The results of our study for a normal user are described. Our statistics were collected from a short Arabic text which is not very reliable. However, we found that this method increases the speed of editing by at least 50 % compared to a static method.

In many cases of handicap the learning mode is more suitable. In this case we believe that the speed of editing can be much higher than the one used in the Normal User Mode.

Appendix 1

Some vectors of the statistics matrix obtained from a short text of approximately 16,000 characters.

ص ض ظ ط ع غ ف ق ك لآ لأ ء آ أ ؤ ا ئ ا ب ة ت ث ج ح خ د ذ
ش س ز ر - ي ى و ه ن م لآ لا

ا	45	0	0	2	0	26	2	51	3	75	4	8
	36	5	49	9	82	7	40	7	21	12	13	2
	37	2	21	10	14	906	7	25	4	138	47	127
	29	23	0	12	298							
ت	0	1	9	1	0	0	30	21	2	10	0	7
	24	9	15	0	19	4	13	14	4	10	5	0
	33	7	9	32	6	13	0	0	0	0	25	26
	27	19	39	9	114							
ج	0	0	1	0	0	0	35	5	0	7	0	0
	0	0		15	2	13	6	8	0	1	0	0
	0	7	0	0	0	0	6	0	0	0	1	16
	14	21	16	1	12	15						
	1	9	6	0	0	1	66	76	31	9	1	14
	10	2	12	0	2	0	19	2	2	18	0	0
	6	5	19	10	32	0	0	0	0	0	6	14
	27	31	70	159								
	0	0	0	0	0	0	56	16	13	12	0	3
	0	0	37	3	46	6	9	5	6	24	3	4
	0	0	3	10	2	100	0	0	0	5	20	32
	9	29	6	21	63							

ف	0	0	3	0	12	1	22	0	5	16	0	1
	3	0	2	2	30	0	11	1	1	9	1	4
	15	0	1	7	29	5	0	0	1	3	4	4
	16	21	99	35	42							
ل	0	0	0	0	0	5	2	26	17	74	7	36
	53	22	34	43	26	6	55	30	28	7	9	2
	81	32	35	52	64	59	0	0	0	9	236	44
	87	58	103	69	184							
م	0	0	0	10	0	1	128	11	22	29	17	14
	20	6	18	2	23	3	66	5	6	8	2	0
	48	3	20	12	9	30	0	1	0	3	9	173
	26	21	7	63	278							
و	5	0	22	0	4	2	219	25	11	23	3	23
	8	8	31	8	48	13	15	2	10	16	9	3
	21	3	24	22	3	80	0	0	0	13	59	102
	27	3	10	42	59							
-	0	7	110	0	48	0	1015	173	3	94	9	44
	62	21	29	12	33	6	28	10	34	9	7	8
	148	8	179	60	80	135	0	0	2	23	282	47
	83	421	146	86								

Appendix 2

Some vectors built from the statistics matrix shown in the Appendix 1

vectors	Elements of the vectors										
ا	ل	–	لآ	ن	ر	ت	ب	د	م	ء	س
	ع	ح	ه	ئ	لأ	و	ص	ف	ك	ط	ض
	ي	ق	ذ	ج	ز	ش	لا	خ	ث	لا	ة
	ؤ	ا	ظ	غ	آ	أ	ا	ى			
ت	–	ى	ع	ق	ا	ه	ن	م	ح	ب	ر
	و	د	ش	س	ل	ت	ض	أ	خ	ف	ي
	ج	غ	ك	ط	ز	ص	ة	لأ	ؤ	ء	ا
	ئ	ث	ذ	ظ	لأ	لأ	لا	لآ			
ج	ا	ه	م	و	د	–	ن	ر	ي	س	ت
	ع	ز	ل	ب	ذ	أ	ص	لأ	ى	ء	آ
	ؤ	ا	ئ	ة	ث	ج	ح	خ	ش	ض	ط
	ظ	غ	ف	ق	ك	لآ	لأ	لا			
ر	–	ب	ي	ا	ك	ة	ى	ه	و	س	ف
	ض	ج	ن	د	ح	ق	أ	ت	أ	ع	م
	غ	خ	ر	ش	ص	ء	ئ	ث	ؤ	ا	ذ
	ز	ط	ظ	ل	لآ	لأ	لا	لآ			
ع	ل	–	ا	ر	د	ن	و	ض	ي	م	ب
	ة	ت	ق	س	ه	ز	ص	ى	ش	لآ	ظ
	ج	ذ	ط	ف	ك	ء	أ	أ	ؤ	ا	ئ
	ث	ح	خ	ع	غ	لآ	لأ	لا			

ف	ى	–	ي	ر	ك	ا	و	ت	ه	ع	ا
	س	ض	ق	ة	ل	ظ	م	ن	أ	ح	لآ
	د	ذ	ئ	ج	ش	ص	ط	ف	لا	ء	آ
	ؤ	ب	ث	خ	ز	غ	لآ	لآ			
ل	م	–	ى	ه	ع	ت	ي	ك	ل	و	س
	ح	ق	ن	ذ	ج	ف	د	غ	ش	ص	ب
	ر	خ	ة	ط	لآ	ث	ض	ز	ئ	ا	ظ
	ء	آ	أ	ؤ	ا	لآ	لا	لا			
م	–	ن	ا	س	ي	ع	ل	ت	ه	ر	ة
	و	ح	ف	د	ث	ج	ق	ب	ؤ	ك	م
	ض	ى	خ	ص	ش	ز	غ	لآ	ذ	ط	ئ
	لآ	ء	آ	أ	ا	ظ	لآ	لا			
و	ا	ن	ل	م	–	ر	ي	د	ه	ب	ف
	ت	ج	أ	ق	ع	ض	س	ز	لآ	ة	ص
	ى	ط	ح	خ	ذ	ء	ا	ث	ظ	غ	ك
	و	ئ	ش	آ	ؤ	لآ	لا	لا			
–	ا	و	م	ف	ب	ع	ي	ل	أ	ت	–
	ه	ك	ح	ق	ا	ن	ج	ص	ر	د	س
	لآ	خ	ذ	ش	ث	ض	ظ	غ	آ	ط	ز
	ة	لا	ى	ء	ؤ	ئ	لآ	لا			

References

[1] Levinsky, J. L."The Keyboardless APPLE". *Proceedings of the Johns Hopkins First National Search for Applications of Personal Computing to Aid the Handicapped.* IEEE, Oct 1981.

[2] Young,C. "Versa-scan Keyboardless Input System for Personal Computer Use by Severly Disabled Persons".*Proceedings of the Johns Hopkins First National Search for Applications of Personal Computing to Aid the Handicapped.* IEEE, Oct 1981.

[3] Mandar, A.M. and Dubus, J.P. "Nouvelle technique de communication a l'usage des infirmes moteurs cerebraux par edition de matrice dynamique." Innovation et technologie en biologie et medecine. 5, No 6,(1984).

[4] Cannon,T.W. "An Optically Actuated Keyboard System". *Proceedings of the John Hopkins. First National Search for Applications of Personal Computing to Aid the Handicapped.* IEEE, Oct 1981.

[5] McFerland, W.D."A Communication Aid for the Non-Oral Severely Disabled". *Proceedings of the John Hopkins. First National Search for Applications of Personal Computing to Aid the Handicapped.* IEEE, Oct 1981 .

[6] Alghamdy, S. A. "Dynamic Array Edition for Severely Physically Disabled Persons". *Final Project Repor*t, Computer Engineering Department, King Saud University, Jan 1987.

[7] Aljabri, A. and Mirza, A. "Characterization of Written Arabic as an Information Source". *Proceeding of Third Saudi Engineering Conference.* 1992, pp. 904.

[8] Al Kadi, M."Arabic Cryptography Features". *Proceedings of Third Saudi Engineering Conference.* 1992, pp. 910.

محرر ذكي للنصوص العربية لاستخدام الأشخاص المعاقين حركيًا

عبدرب النبي منظر

قسم علوم الحاسب، كلية علوم الحاسب والمعلومات، جامعة الملك سعود، الرياض، المملكة العربية السعودية

ملخص البحث. تصف هذه الورقة طريقة تحرير للألفبائية الرقمية العربية باستخدام مصفوفة حركية، وذلك بمساعدة نبيطة تحسس من نوع وصل / قطع يتم التحكم فيها من قبل الشخص المعاق حركيا. والجزء الرئيس من العمل الموصوف في هذه الورقة هو تطوير محرر ذكي يكيف نفسه مع مستخدمه. وهناك أسلوبان للعمل يمكن اتباع أي منهما مع هذا المحرر: الأسلوب المعتاد المتبع مع المستخدمين المعاقين الذين يتمتعون بقدرات عقلية عادية، والأسلوب الثاني هو أسلوب التعلم الذي يمكن استخدامه مع المعاقين الذين لهم قدرات عقلية منخفضة. وكلا الأسلوبين يستند إلى بيانات إحصائية عن اللغة العربية تم جمعها من نصوص عديدة فيما يخص الأسلوب الأول، أما فيما يخص الأسلوب الثاني فكانت النصوص اللغوية محدودة.