INTERACTIVE CREATION OF CHINESE CHARACTERS

Dennis B. Ward, Xiaoning Wang, and W. A. Davis

ABSTRACT

This paper describes a program which allows the user to construct Chinese characters and text using interactive graphics. Input to the system comes from a terminal and an acoustic tablet, and the output is displayed on a color monitor.

Chinese characters can be built from a set of basic strokes. The program displays these strokes in a menu, and allows the user to select them in sequence and place them on the character under construction. Strokes can be placed in any position within the character, and a special linear stroke can also be altered in length or rotation angle. The order in which strokes have been added, which is important in Chinese characters, can be indicated by coloring the strokes in a sequence indicated by a series of color bars. Previously written strokes can be selected for alteration by indicating the stroke color or stroke type. Characters can be constructed and displayed in various sizes, with an overlayed grid being used as a guide. They can be displayed in a single colour to improve their appearance, and can be stored in a file for future retrieval. Stored characters may be used to construct Chinese text, which can also be stored and retrieved.

Cette publication décrit un programme permettant à l'usager de former des caractères et textes chinois en utilisant un graphisme interactif. L'entrée dans le système se fait par un terminal et une tablette acoustique, et la sortie est visualisée sur un moniteur en couleur.

Les caractères chinois sont construits à partir d'un ensemble de traits de base. Le programme affiche les traits sur un menu, et permet à l'usager de les selectionner en série et de les placer sur le caractère en construction. Les traits peuvent être placés dans n'importe quelle position à l'intérieur du caractère, et la longueur et l'angle de rotation d'un trait droit special peuvent aussi être modifies. L'ordre d'addition des traits, qui est important dans les caractères chinois, peut être connu èn colorant les traits dans une séquence indiquée par une série de barres en couleur. Des traits tracés préalablement peuvent être selectionnés pour modification en indiquant la couleur ou le type du trait. Des caractères de différentes tailles peuvent etre construits, en utilisant une grille guide en surimpression. Ils peuvent etre visualisés dans une seule couleur pour améliorer leur apparence, et etre stockés sur un fichier en vue d'un recouvrement ultérieur. Les caractères conservés peuvent être utilisés pour réaliser un texte chinois, qui peut aussi être stocké et recouvré.

本文摘述一个允许用户以交互式计算机图象显示方式构造汉字和汉字串的程序。输入信息经终端和图象映射装置进入系统，系统输出由彩色显光屏显示。

本系统中，汉字采由一组基本笔划组成。程序在荧光屏上显示该笔划集，其中一个通用的，直线型的笔划可在长度上任意变化并做旋转任意角度。用户可以任意选择笔划来构造任意的常用汉字。系统显示一个具有一定顺序的色谱，使构字顺序可以由构成该字的笔划的颜色来表明。已构成的汉字中所用的笔划可以通过指明其颜色或类型来再选择并加以修改或删除。汉字被写在一个特定的方格中，方格的尺寸可以变化，其中汉字的大小也随之变化。系统可以存贮和检索已构成的汉字。存贮在系统中的汉字可用来构造汉字串，系统也可以存贮和检索汉字串。

Graphics Interface '83

Introduction

The Chinese system of writing is exceedingly complex. More than 60,000 characters are known, of which some 3000 to 4000 are in general use. Computer processing of this vast array of characters is an important and serious problem, and there is a large literature in this area.

If the characters are viewed as individual entities, producing them is very difficult. However if the component strokes are analyzed, some simplifying factors are found. Firstly, the vast majority of characters can be produced using a relatively small number of strokes. Secondly, the more complex of these strokes are fixed in size and orientation.

These observations suggest that an interactive computer graphics system could be used to produce Chinese characters in a stroke oriented fashion, with strokes being selected from a menu. This paper will give a brief description of such a system which was developed at the University of Alberta[1]. A more complete description will appear elsewhere (1).

Stroke design

A central requirement in development of the system was the choice of a set of strokes. The set had to be sufficiently large to enable easy construction of characters, yet relatively small to reduce the complexity of the menu. The strokes had to be capable of being drawn quickly, and in a range of sizes.

Eventually a set of twenty strokes was chosen, and these may be seen in the stroke menu at the top of Figure 1. This set is not necessarily optimal – as time passed a number of other strokes were encountered which would make useful additions to the menu. The chosen set was, however, sufficient to produce reasonable versions of a large number of characters.

As can be seen in the menu, there are essentially two types of stroke. All but one of the strokes are of the "nonlinear" type. These strokes are as shown in the menu, and, except for overall changes in scale, nonlinear strokes cannot be changed in size or rotated. Each of these strokes was created as a pattern by editing a dot matrix, and the dot matrix for each stroke was then converted to a series of horizontal lines. The table of strokes is read in from a file each time the program is run. Changes of scale are achieved by multiplying line end point coordinates (which are relative to a stroke center) by the scale factor, and interpolating to fill lines in between.

The second type of stroke (appearing at bottom left of the stroke menu) is the "linear" stroke. This stroke can be rotated and changed in length, and is used to create any strokes which are not found in the standard set of nonlinear strokes. It is drawn as a long narrow box, filled in with horizontal lines.

System operation

The system runs on a PDP11/60 computer under the UNIX operating system, and is written in C. The display device is a NORPAK VDP. Input is taken primarily from an acoustic tablet, with a terminal being used to enter command confirmations and alphanumeric information.

Actions are selected by moving a flashing cursor, which is overlayed on the display screen. The acoustic tablet is mapped onto the screen, and the position of the cursor is controlled by the position of the acoustic tablet stylus. An action is performed when the cursor position has stabilized in an active portion of the screen. "Stabilization" of the cursor means that the variance of the cursor position remains small over a sequence of five cursor position measurements. The terminal bell is sounded to give some feedback each time an action has been selected.

System messages are displayed in a flashing white overlay at the bottom of the screen. They are used for a number of purposes, examples being prompts for terminal input, error messages, and signals that an action has been performed.

There are two modes of operation, "character" mode, in which characters are constructed by positioning strokes, and "text" mode, in which character displays are formatted. These will be further described in the following sections.

Character mode

The screen display for character mode is shown in Figure 1. It is divided into four regions – the stroke menu at the top left, the colour bars at top right, the work area at bottom left, and the command pads at the bottom right of the screen. The purpose of each portion of the screen will be explained below.

(1) Stroke menu

This part of the character mode display shows the strokes which can be used to construct Chinese characters. If the cursor is moved into one of the boxes surrounding a stroke, that stroke is selected as the "current stroke", and can be moved into the character being built in the work area. The stroke menu can also be used to reselect a previously added stroke for modification (see the discussion of the "reselect" command in the command pad description).

(2) Work area

Character construction takes place in the work area, but only within the template which occupies the central portion. All the rest of the work area outside the template is inactive.

The template is used to help the user position strokes accurately, and resembles an aid which is used by Chinese children when they are learning to draw characters. The current stroke (overlayed in brilliant white on the display) is positioned within the template by moving the cursor to the desired location. When a new current stroke

is selected, the previous stroke is written into the character at its present position, and in the appropriate colour (see the discussion of the colour bars below).

Linear strokes can be manipulated with the aid of the CCW, CW, LONGER and SHORTR commands, as described in the command pad discussion below. Strokes may be erased with the ERASE command.

(3) Command pads

Commands are selected by moving the cursor into an active command pad, which is identified by a brown colour. There are four basic types of commands, which are summarized below.

(a) current stroke manipulation commands

CW rotates a linear stroke clockwise
CCW rotates a linear stroke counterclockwise
LONGER lengthens a linear stroke
SHORTR shortens a linear stroke
ERASE erases the current stroke. If no new current stroke is selected, the erasure can be reversed by moving the cursor back into the template – the erased stroke will reappear at the new cursor position.
RELEASE releases the current stroke, which is written into the character in the appropriate colour.

(b) character manipulation commands

ROTATE 90 180 rotates the character being constructed by 90 (clockwise) or 180 degrees, whichever is selected. It only works for characters containing no nonlinear strokes. It is included in the command set because a large number of common Chinese characters are rotations of other characters.
SCALE + - increases or decreases the scale at which the current character is being drawn. There are four size levels which can be used.
SAVE saves the current character in a file. An identifying name for the character must be entered from the terminal.
RETRV retrieves a character which was previously saved. The character name is entered from the terminal. Any strokes which were previously in the template are wiped out.
RESTART erases the strokes currently in the template so a new character can be drawn. Confirmation from the terminal is required, and is given by entering "RETURN". Hitting any other key aborts the command.

(c) mode commands

GLOBAL shifts the system over to text mode. It requires confirmation similar to that for RESTART, described above.
QUIT halts the program. Confirmation from the keyboard is required, as for GLOBAL and RESTART.

(d) flag option commands

RESELECT ON OFF turns the reselect flag ON and OFF. Reselect is used to select a stroke already in the character to be the current stroke. This stroke can then be moved, modified or erased. With the reselect ON, moving the cursor into a stroke menu box will cause the first stroke of that type added to the character to become the current stroke, if such a stroke exists. Leaving the cursor in that box will cause the next instance of the stroke to become the current stroke. The reselect flag must be turned OFF before new strokes can be added to the character. (See the colour bars discussion for another method of reselecting a stroke.)
COLOR ON OFF changes the colour flag setting. When the colour flag is ON, the colour bars are active (the default system setting). If the colour flag is OFF the strokes are drawn in a default colour.
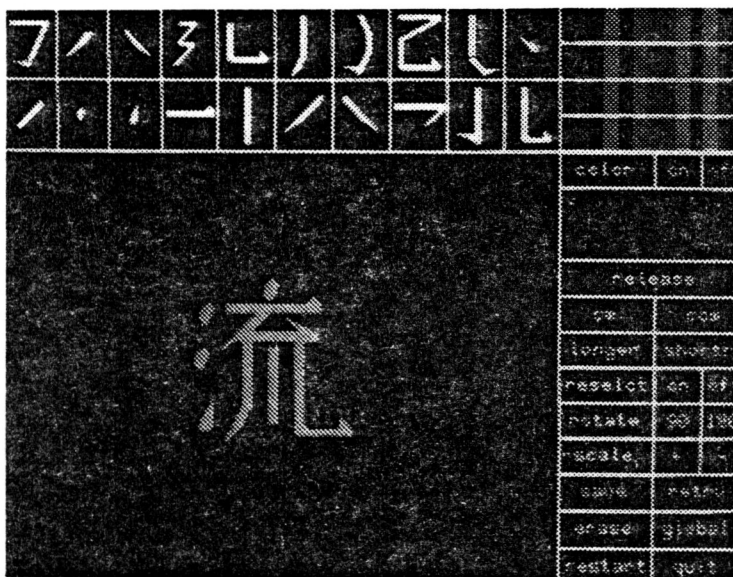

Figure 1 – Character mode display

(4) Colour bars

There are 52 colour bars, split into 4 levels of 13 bars each. They give the colours which are assigned to strokes as they are added to the character (assuming the colour flag is ON). The first colour assigned is at the bottom left, and the assignment goes on from left to right through the first level, then on to the second level, etc. The purpose of this is to identify the order in which strokes have been added, since this is important in correct Chinese character construction.

A further use of the colour bars is in reselection of a stroke to be modified. If the colour flag is ON, moving the cursor into the colour bar whose colour corresponds to the

desired stroke will make that stroke the current stroke. This method of reselecting a stroke is generally more convenient than use of the reselect flag commands.

If the colour flag is OFF, strokes are added in a default colour, and the reselect flag command must be used to pick strokes for modification.

Text mode

The text mode display is shown in Figure 2. It has two main regions – the work area to the left, and the command pads on the right. These are described in the following subsections.

(1) Work area

This is the area in which text is built up. The template indicates where the next character will be positioned. The template location is adjusted by moving the cursor within the work area, with the restriction that no part of the template is allowed to extend beyond the boundaries of the work area.

Characters can be added in various sizes, with the current size being indicated by the size of the template. One character can be selected as being the "current character", and can then be moved or replicated using commands described in the next subsection. The current character is displayed in the overlay in brilliant white, while other characters are displayed in the default color (typically a bright gold).

(2) Command pads

The text commands break into several classes, as described below.

(a) character commands

RETRIEVE prompts for the name of a character, which is retrieved from the character file and displayed at the current template position.
SELECT selects the character currently within the template to be the current character
MOVE moves the current character to the present template location
COPY puts a copy of the current character at the present template location
RELEASE releases the current character, which is written back into the display in the default colour.
ERASE erases the current character. Once erased, the character cannot be brought back.
SCALE + - changes the size of the template. The next character retrieved, moved or copied into the template will appear in a size corresponding to the template size. Two sizes are available, corresponding to the two smallest sizes in stroke mode.

(b) text commands

SAVE-WORKAREA saves the current display in a file. The file name is entered from the terminal.
RETR-WORKAREA retrieves a display which was previously stored in a file. The file name is entered from the terminal. The previous work

area contents are lost.
RESTART clears the work area. Confirmation must be given by hitting "RETURN" on the terminal. Any other keystroke will abort the command.

(c) mode commands

LOCAL switches the program to character mode. Confirmation is required from the keyboard, similar to RESTART.
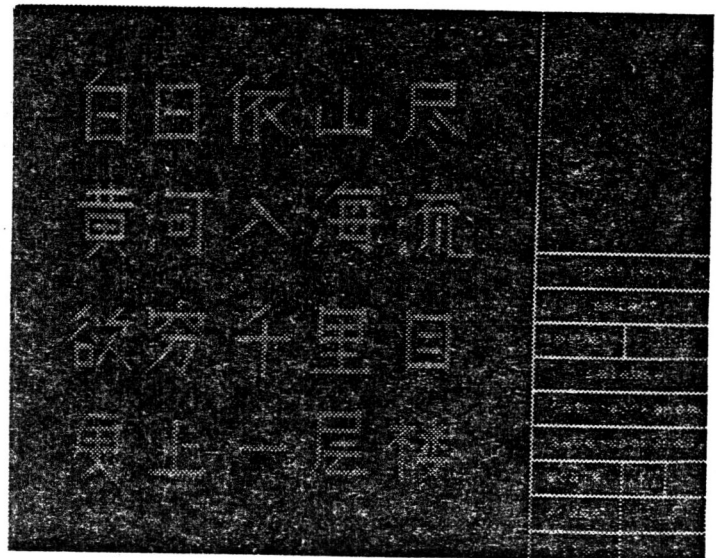QUIT halts the program. Confirmation from the keyboard is required.

Figure 2 – Text mode display

Results and Conclusions

The system is operational, and has been tried out by a number of people. It is quite easy to use, and produces Chinese characters of reasonable quality with a minimum of training and effort.

At present the package is used as an informal demonstration tool, showing what Chinese characters are like and how they are made. It is quite effective in this role, at least in part because the characters produced are quite attractive. This suggests that a similar approach might be used in connection with teaching Chinese. With a minimum of training a student would be able to produce attractive and recognizable characters, without having to master the use of a brush. Stored characters could be used for demonstrations of character construction and of Chinese text.

Such a system could likely be designed for a microcomputer which has reasonable graphics capabilities. The algorithms used to draw strokes are relatively simple and fast, and secondary storage requirements for characters and text would not be large, since the stroke data structures can be made very compact. Some simplification might be

necessary to reduce storage requirements, since the system currently uses about 64K of memory. The lower resolution of a personal computer screen would be overcome in character mode by making characters quite large, but would cause problems in text mode unless some sort of anti-aliasing algorithm (2) were used.

Production of Chinese characters for use in computer-based printing is another possible application. The approach described here would likely find its best application in production of characters for inherently low-resolution output devices, such as dot matrix printers. Creation of truly high quality Chinese characters would require the addition of many more stroke types. This is probably not worthwhile, as the greatest advantage of the current system is its simplicity. More sophisticated techniques (3,4) would likely be more effective.

We would like to thank John de Haan for his contributions to this work.

References

(1) Technical Report, University of Alberta Department of Computer Science. In preparation.

(2) Ni, G. and P. Tanner. "The application of anti-aliasing technique for displaying high quality Chinese characters," in Proc. 1982 International Conference of the Chinese-Language Computer Society, pp. 37-45, September 1982.

(3) Hsu, Y.-C. and R.T. Yeh. "A brush-pen Chinese character output system," in Proc. of 1982 International Conference of the Chinese-Language Computer Society, pp. 1-11, September 1982.

(4) Hobby, J. and G. Gu. "Using Metafont to design Chinese characters," in Proc. of 1982 International Conference of the Chinese-Language Computer Society, pp. 18-36, September 1982.