



Generation of Personalized Handwriting Written in Different Language – A Technical Review

Preeti P. Bhatt¹, Jarul Munshi², Juhi Patel²

¹Assistant Professor, ²Student

Babu Madhav Institute of Information Technology,
Uka Tarsadia University, Bardoli, Gujarat, India

ABSTRACT

Digitization of handwritten document is important for better processing and storage. But people wish to have their own personal handwriting as font because Personal font style gives a personal touch and make the document more eye-catching. Font generation is a process of generation of a handwritten document-based writing characteristic which is specific to user's writing style. In this paper comparative study of font generation techniques has been presented for different languages such as English, Bangla, and Chinese. We also compare different methods like RNN, a Macro structure model, LSTM etc. which were used for font generation in different languages. Different parameters have been considered for regenerating someone's font like character glyph, character space and connection style.

Keywords: Font Generation, Glyph synthesis, RNN, LSTM

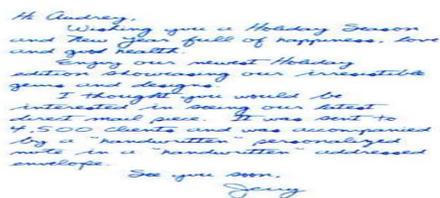
INTRODUCTION

People of personalized era desire to have their own personal font so that documents can be displayed as if

handwritten by their own hands. If you use personal style font than it makes communication more attractive. Personal handwriting fonts will also attract the eye. We can create personal touch while inviting by card in own writing, writing books in personal writing style. Even we can create logo, notes and header using attractive personal writing which creates good impacts on reader minds.

It is thought that different language calligraphy can mirror one's emotion and persona. Many authors have been long looking for personal fonts so that they can not only present their idea but also show their personality and feelings to readers.

Font generation is a process of generation of the handwritten document-based writing characteristic which is specific to user's writing style. Regenerating someone's handwriting by applying different method, this process is known as font generation. We demonstrate font generation with one example here.



*Hi Audrey,
Wishing you a Holiday Season
and New Year full of happiness, love
and good health.
Enjoy our newest Holiday
season slowasing our irresistible
game and sleazins.
I thought you would be
interested in seeing our latest
direct mail piece. It was sent to
4,300 clients and was accompanied
by a "handwritten" personalized
note in a "handwritten" addressed
envelope. See you soon,
Jarul*

Process &
Generate



	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
10	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
20	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
30	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
40	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
50	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
60	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
70	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

Figure 1

Personal calligraphy creation had been developed for many languages such as Chinese [1][2], English[3], Bangla[4]. Also, many tools had been developed by font engineers. For example, FontForge, Calligraphy both are an open source font editor.

For generating font in any language we require parameters. They are,

- Character glyph,
- Connection style,
- Character space,
- Cursiveness.

Visual representation of the character is known as character glyph. In some case, the character might be represented by more than one glyph. Example letter 'J' is the character while a particular sketch of that character is known as a glyph. A style using which characters are connected with each other that style is known as connected style. Amount of space between each individual character is known as character space. A handwriting in which letters are usually connected in a slanted and flowing manner is known as cursiveness.

Challenges:

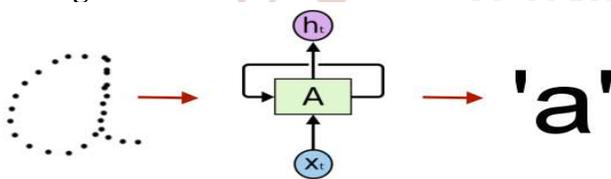


Figure 2

In figure 2 series of pen, points represent the letter 'a'. It is quite easy to generate character 'a' from series of pen points letter. A much more difficult to reverse the process, i.e. to train a model that takes the letter 'a' as an input and produces a series of points that we can connect to make the letter 'a'.

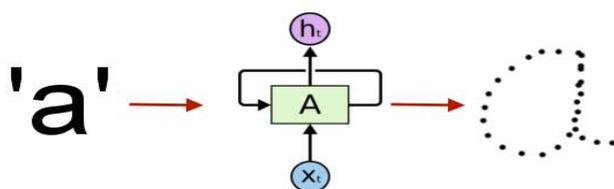


Figure 3

Figure 2 represents font recognition and figure 3 represents font generation. Font generation is quite difficult compare to font recognition so that we choose this topic.

Literature review:

For generating Chinese character two phases are proposed, pre-processing phase and production phase. In the pre-processing phase sizes and positions of components of every Chinese character is label manually. In the production phase for creating personal handwriting, the user had to handwrite the small subset of the Chinese character. The system can track every stroke, distinguish and extract components from the user's handwriting character by character. Different methods are explained for it such as component analysis, positions and sizes of components in Chinese character, component extraction and glyph synthesis [1].

Whole Chinese character first splits into multiple constituents parts such as strokes and radicals. After that algorithm analysis and learns the characteristics of character handwriting styles both defined in Chinese national font standard and those exhibited in a person's own handwriting records. Parametric character shape representation method is used for characterizing the shape difference between components. A method calculates stroke similarity according to the feature points of a stroke. The topological structure of Chinese character, automatic extracting character components task are used for generating Chinese character [2].

By predicting one data point at a time, one can generate complex sequences with a long-term structure using long-short-term memory recurrent neural networks. The approach is demonstrated for text and online handwriting. It is then extended to condition its predictions on a text sequence. Novel CNN that allows a recurrent network to condition its estimates on a supporting annotation sequence [3].

Specific writing has macro structures like neighbouring word roughness, line orientation, margin orientation, and interline spacing etc. as well as micro properties like inter-character transitions and individual character shapes. Bangla writing is more complex in nature because of the same, a grouping of stored character database and stroke synthesis by a spline with paint brushing is used. [4].

ISI graphy tool for the Android platform has been developed to collect data on the handheld device. It can collect and stores handwriting samples on large-scale written in any language of distinct users. A tool

is used for collecting online handwriting sample database. In a single device, this application supports multiple user accounts. To provide security SQLite database is used. Data are temporarily stored in buffer memory. This tool does not depend on the particular script. The advantage of this tool is that it can send the collected data directly via electronic mail [5].

Generation of character shapes done based on Bayesian networks, in which experimental results of digit shows more natural character shapes than HMMs. Character shapes are generated from specified texts by searching the most feasible input point sequences. Representing random variables and their dependencies by the probabilistic graph is acknowledged as a Bayesian network. Its node represents random variables and its arcs represent dependencies between random variables. Also overall shapes, relationships between graphemes, strokes, and points are more natural [6].

An approach for analysing simple, as well as complex movements such as cursive handwriting, describe in this paper. Cursive calligraphy is described as the superimposition of vital strokes with an elliptic form that results from the algebraic summation of Beta

velocity profiles. Each stroke is totally described by a set of ten parameters that characterize the movement both in the kinematics and the static domains. The beta-elliptic model can be applied not only to French handwriting but also to Arabic handwriting words [7].

Font Forge is an open source program which allows the creation and modification of fonts in many standard formats. It can automatically guess at widths for characters and even produce kerning pairs without human intervention. It has a scripting language which allows batch processing of many fonts at once. Font Forge has support for Apple's font formats. It can read and generate Apple font files both on and off a Macintosh. Font Forge allows you to manipulate bitmap fonts as well as outline fonts [8].

Paper describe not only about font recognition but also about font generation specific to author's writing style. For recognition different methods like converting to gray scale, thresholding, and smoothing, etc method are uses. For generating font RNN and LSTM is use. For handwriting recognition 98.2% accuracy was achieved [9].

Comparative Study:

Reference	Language	Classifier	Features
[1] [2]	Chinese	<ul style="list-style-type: none"> • Parametric Character Shape Representation • Novel Algorithmic method 	<ul style="list-style-type: none"> • For characterize the shape difference between components in a standard Chinese character and their counterparts in a personal handwritten character sample parametric character shape representation method is use. • Novel algorithmic method first split Chinese character into multiple constituent parts and then algorithm analyses and learns the characteristics of character.
[3][6][7] [9]	English	<ul style="list-style-type: none"> • RNN • LSTM 	<ul style="list-style-type: none"> • RNN allow generation of arbitrary lengths of recurrent network and also ideal for text and speech analysis. • LSTM store previous input for long and also it was design for remove long term dependency.
[4][8]	Bangla	<ul style="list-style-type: none"> • Macro Structure Model 	<ul style="list-style-type: none"> • Work on combination of stored character database and stroke synthesis because Bangla language is complex.

Table 1

Reference	Classifier	Advantage	Disadvantage
[3][6][2]	RNN	Allows us to generate arbitrary Sequence.	Unable to store previous inputs for long, this reduces the system's stability.
[3][7][1]	LSTM	Better at storing and accessing information than RNNs.	It makes predictions difficult and confusing. It is more costly to calculate the network output and apply back propagation.
[4][5]	Macro structure model	Less time consuming	High error rate compare to RNN

Table 2

Conclusion and future work:

This paper demonstrated about font generation in different language and its parameters to synthesis the character. Also, we discuss challenges to generate font. Shown comparative study of font generation in different languages like Chinese, English, Bangla, and methods like RNN, macrostructure model, and LSTM.

For many languages such as Chinese, English font generation is done but for the Gujarati language, there is no implementation yet so that we can implement font generation for the Gujarati language in future. Based on a comparative study we can try the combination of any two methods or more like RNN and Macrostructure model to generate font for Indian script.

References:

- Lin, Jeng-Wei, et al. "Complete font generation of Chinese characters in personal handwriting style." *Computing and Communications Conference (IPCCC), 2015 IEEE 34th International Performance*. IEEE, 2015.
- Liu, Peng, Songhua Xu, and Shujin Lin. "Automatic generation of personalized Chinese handwriting characters." *Digital Home (ICDH), 2012 Fourth International Conference on*. IEEE, 2012.
- Graves, Alex. "Generating sequences with recurrent neural networks." *arXiv preprint arXiv: 1308.0850* (2013).
- Chaudhuri, Bidyut B., and Abhisek Kundu. "Synthesis of individual handwriting in bangla script." *Proceedings of the ICFHR* (2008).
- Das, Arindam, and Ujjwal Bhattacharya. "ISI graphy: A tool for online handwriting sample database generation." *Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on*. IEEE, 2013.
- Choi, Hyunil, Sung-Jung Cho, and Jin Hyung Kim. "Generation of handwritten characters with bayesian network based on-line handwriting recognizers." *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*. IEEE, 2003.
- Bezine, Hala, Adel M. Alimi, and Nasser Sherkat. "Generation and analysis of handwriting script with the beta-elliptic model." *Frontiers in Handwriting Recognition, 2004. IWFHR-9 2004. Ninth International Workshop on*. IEEE, 2004.
- Williams, George. "Font creation with Font Forge." *Euro TEX 2003 Proceedings, TUG boat 24.3* (2003): 531-544.
- R. Shah et.al, Handwriting Recognition, Learning and Generation, "Research project.2016