Speech Synthesis System for Online Handwritten Punjabi Word: An Implementation of SVM & Concatenative TTS

Dinesh Kumar
Assistant Prof. & Head
Department of Information Technology
DAVIET, Jalandhar

Neeta Rana
M.Tech
Department of Computer Science
DAVIET, Jalandhar

ABSTRACT
This research covers two phases: recognition & speech synthesis. The main aim of this research was to prepare a system which speaks the handwritten Punjabi word. Till now, the research for Punjabi word recognition is limited to 2460 Punjabi characters only (i.e. only for words available in database). In our proposed system, technique used for recognition is Support Vector Machine & for speech synthesis technique used is CTTS (Concatenative Text-to-Speech). For recognition, the proposed approach is database independent. But for speech synthesis, the proposed approach is database dependent.

KEYWORDS
Handwritten Word Recognition, Offline handwriting Recognition, Online handwriting Recognition, Support Vector Machine, Speech Synthesis, Text- to-Speech Synthesis System, Concatenative Text-to-Speech

1. INTRODUCTION
A very relevant present-day field of natural interface research is hand writing recognition technology. Handwriting Recognition is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices. [22] It is improving the interface between man and machine in many applications. Handwriting recognition is comparatively difficult, as different people have different handwriting styles.
Handwriting Recognition is of two types:

- Offline handwriting Recognition(Offline: Data is collected before the processing start)
- Online handwriting Recognition(Online: Data is captured as it is written)

On-line handwriting recognition involves the automatic conversion of text as it is written on a special digitizer or PDA, where a sensor picks up the pen-tip movements as well as pen-up/pen-down switching. That kind of data is known as digital ink and can be regarded as a dynamic representation of handwriting. [22]

Off-line handwriting recognition involves the automatic conversion of text in an image into letter codes which are usable within computer and text-processing applications. [22]

Handwriting recognition inherited a number of technologies from Optical Character Recognition (OCR). OCR engines are used for machine printed text and ICR for hand "printed" text. Till date there is no OCR/ICR for handwriting Recognition.
For recognition the techniques used are: Neural Network, Support Vector Machine, Dempster-Shafer Theory, Hidden Markov Model, Octal Graphs Technique, Genetic Algorithms, and Elastic Matching Method.

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a **speech synthesizer**, and can be implemented in software or hardware. A **text-to-speech (TTS)** system converts normal language text into speech [22]. There are various technique used for Text-to-speech synthesis techniques. Some of them are: Concatenative synthesis, Formant synthesis, Articulatory synthesis, HMM-based synthesis, Sine wave synthesis.

2. PUNJABI LANGUAGE
Punjabi is an Indo-Aryan language spoken by inhabitants of the historical Punjab region (north western India and in Pakistan). Punjabi is called also Gurmukhi (in India) or Shahmukhi (in Pakistan). Gurmukhi means "from the mouth of the Guru". It was devised during the 16th century by Guru Nanak, the first Sikh guru, and popularized by Guru Angad, the second Sikh guru.
According to the Ethnologue 2005 estimate, there are 88 million native speakers of the Punjabi language, which makes it approximately the 12th most widely spoken language in the world. According to the 2008 Census of Pakistan, there are 76,335,300 native Punjabi speakers in Pakistan and according to the 2001 Census of India; there are 29,102,477 Punjabi speakers in India. [22].

2.1 CHARACTER SET FOR PUNJABI
2.1.1 Vowels (ਲਗਾ ਮਾਤਰਾ)
There are nine vowel phonemes in Punjabi. They are vowels making only one sound. All consonants use the vowel. Table 1 shows the vowels.

Table 1 (List of Vowels)

ਿ (Sihari)	ੀ (Bihari)	ਾ (Kanna)
੍ (Lavan)	ੁ (Dulavan)	ੂ (Aunkar)
ੇ (Dulonkar)	ੋ (Hora)	ੌ (kanora)

1.1.2 Consonants (ਵਿਅੰਜਨ)

Punjabi language consists of 41 consonants. Consonants list of Punjabi language is written in Table 2.

Table 2 (List of Characters)

ਅ (ura)	ਆ (aira)	ਇ (iri)	ਸ (sassa)	ਹ (haha)
ਖ (kakka)	ਖ਼ (khakha)	ਗ (gaga)	ਘ (ghaga)	ਙ (nanna)
ਛ (chacha)	ਛ਼ (shasha)	ਜ (jaja)	ਝ (jhaja)	ਞ (nainna)
ਠ (tainka)	ਠ (thatha)	ਡ (dadda)	ਢ (dhadda)	ਯ (naanna)
ਥ (tatta)	ਥ (thattha)	ਦ (dadda)	ਧ (dhadhha)	ਃ (nana)
ਫ (pappa)	ਫ਼ (fafa)	ਬ (baba)	ਭ (bhabha)	ਂ (mamma)
ਯ (yaya)	ਰ (rara)	ਲ਼ (lala)	ਵ (vavva)	ੜ (rarha)
ਸ਼ sassha	ਖ਼ khakhha	ਗ਼ gaggha	ਜ਼ jajjha	ਫ਼ faffha	ਲ਼ lallha

2.1.3 Auxiliary Signs

It serves to add a nasal sound to a particular vowel. These signs are represented in Table 3.

Table 3 (Auxiliary Signs used in Punjabi)

ੰ (Tippi)	ੱ (Adak)	ਂ (Bindi)

3. OVERVIEW OF PROPOSED SYSTEM

The proposed system consists of two phases. First phase is Online Handwritten Punjabi Word Recognition. Second phase is Speech synthesis of recognized word. As already described, the technique used for developing recognizer is **Support Vector Machine** and for speech synthesis the technique used is **CTTS (Concatenative Text-to-Speech).**

Block diagram for proposed system is shown below.

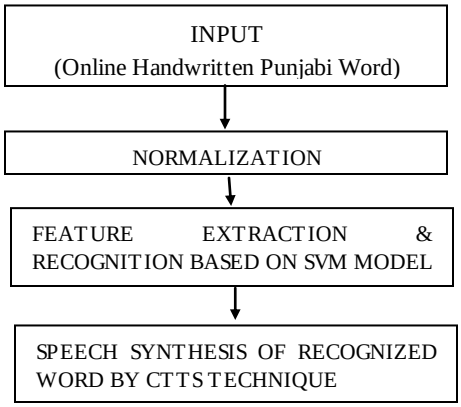

Figure 1(Block Diagram of Proposed System)

3.1 Word Recognition

Recognition Process is described below.

3.1.1 Stroke Representation of Character

Each character is represented as a combination of strokes. A stroke is defined as the trajectory traced by the pen from a pen-down event to a pen-up event and is represented using the data captured as the stroke is written. The number of points collected varies with the stroke and the speed of writing. Strokes used to represent Vowels, Auxiliary Signs & Punjabi characters are shown in table 4 & 5.

Table 4 (List of Strokes to represent Punjabi Vowels & Auxiliary Signs)

Punjabi Characters	Strokes	Strokes	Punjabi Characters	Strokes
ਾ	ੁ		ੰ	ੱ
ਿ	ੂ		ੱ	ਂ, ੱ
ੀ	ੇ	੍, ੁ	ਿ	ੀ
ੁ	ੋ		ੀ	ੂ
ੂ	ੌ		ਂ	ੇ
ੇ	ੇ, ਾ			

Table 5 (List of Strokes to represent Punjabi Characters)

Punjabi Characters	Strokes	Strokes	Punjabi Characters	Strokes	Strokes
ੳ	ੳ		ਬ	ਬ	੫,—,—
ਅ	ਅ		ਭ	ਭ	
ੲ	ੲ	ੲ,—	ਮ	ਮ	੫,—
ਸ	ਸ	ਸ,—	ਨ	ਨ	੮,੭
ਹ	ਹ		ਪ	ਪ	
ਕ	ਕ		ਫ	ਫ	
ਖ	ਖ	੫,—	ਬ	ਬ	ੀ,੬
ਗ	ਗ	ਗ,ੀ	ਭ	ਭ	
ਘ	ਘ		ਮ	ਮ	
ਙ	ਙ		ਯ	ਯ	ੀ,ੴ
ਚ	ਚ	ੀ,ੲ	ਰ	ਰ	
ਛ	ਛ	ੀ,ਓ	ਲ	ਲ	ੲ,—	ੲ,੭
ਜ	ਜ	ੀ,ੀ	ਵ	ਵ	
ਝ	ਝ	੨,ੀ	ੜ	ੜ	੩,।
ਞ	ਞ	ੲ,—	ਛ	ਛ	.
ਟ	ਟ		ਸ਼	ਸ਼	.
ਠ	ਠ	੦,—	ਖ਼	ਖ਼	.
ਡ	ਡ		ਗ਼	ਗ਼	.	ਗ,ੀ
ਦ	ਦ		ਜ਼	ਜ਼	.
ਣ	ਣ	ੲ,੮,	ਲ਼	ਲ਼	—,.	ੲ,੭,.
ਤ	ਤ					

3.1.2. Normalization

Some characters can be written using more than one stroke with some strokes extending above or below the main part of the character. The main part of a character gives a measure of the line space and character size used by the writer. Generally, there will be at least one stroke in the main part of the character from which this information is obtained. This main stroke, taken to be the largest stroke that occurs in the main part of the character, is identified and the character is scaled down by its height value. In cases where it is not possible to identify the main stroke, the height of the character is used for normalization.

3.1.3 Feature Extraction

Each stroke is represented as an n-dimensional feature vector depending on the choice of the number of points for stroke representation. The features chosen to represent the curve are the co-ordinates of points in the preprocessed stroke. The stroke is then classified using SVMs. Kernel and associated parameters are experimentally determined. One-against-the-rest strategy is implemented for multiclass classification of strokes.

Figure 2 represent the feature extraction of Punjabi character "ੳ". Here, F stands for **Feature**.

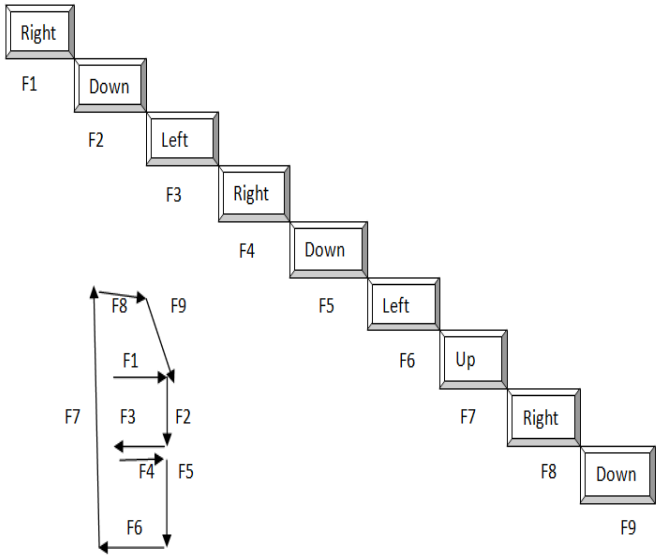

Figure 2 (Feature Extraction of Punjabi character "ੳ")

3.1.4 Recognition of Punjabi Word

Here, the recognition process consists of three steps: 1.Stroke Identification. 2. Making character from recognized strokes. 3. Concatenate character to get word. Figure No.3 is describing the whole process.

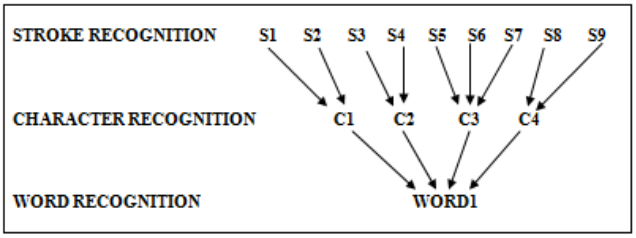

Figure 3 (Recognition Steps)

Recognition process of word "ਜੀਵਨ" is shown in Figure 4.

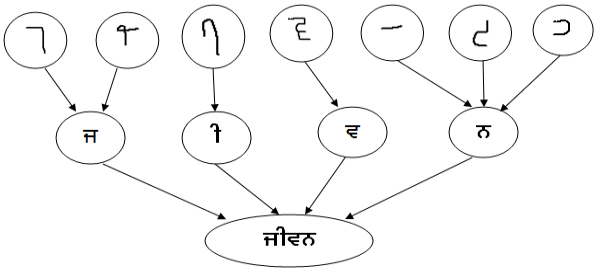

Figure 4 (Recognition of word "ਜੀਵਨ")

3.2 Text-to-Speech Synthesis:

Second phase of this proposed system is Speech Synthesis. The speech synthesis is of recognized word. This phase of proposed system is database dependent. Till now, Size of our database is 1000 sound files. The technique used is Concatenative Text to

speech synthesis. Concatenative synthesis uses a large database of source sounds, segmented into *units*, and a *unit selection* algorithm that finds the sequence of units that match best the sound or phrase to be synthesised, called the *target*. The selection is performed according to the *descriptors* of the units, which are characteristics extracted from the source sounds, or higher level descriptors attributed to them.

The selected units can then be transformed to fully match the target specification, and are concatenated. However, if the database is sufficiently large, the probability is high that a matching unit will be found, so the need to apply transformations, which always degrade sound quality, is reduced. As, this technique is used to speak text. But in our proposed system, we are applying this technique to speak only one word. So, the proposed system used this technique to search the recognized word from the database and play the related sound file. If recognized word is not available in the database then this technique concatenate the character sound files to speak this word. The whole speech synthesis process is shown in Figure 3.

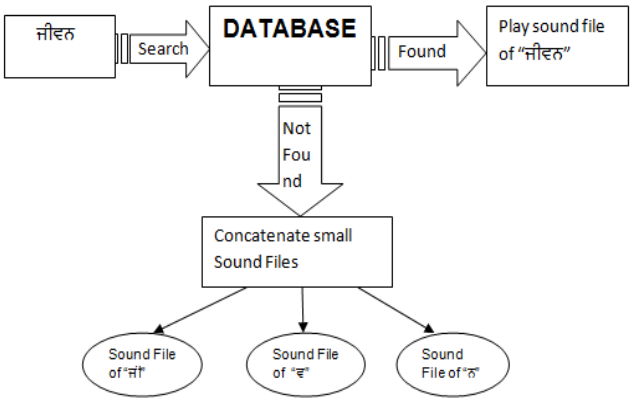

Figure 3 (Speech Synthesis Process)

Finally, the speech synthesis system is integrated with recognition system.

4. EVALUATION

Handwriting Recognition is evaluated based on Accuracy & Writer Dependency. On the basis of survey done on 15 writers the average accuracy of character recognition is 90% and average accuracy of word recognition is 80%. The table no. 6 is showing the recognition rate of each Punjabi character. It's mainly varies due to style of writing.

Table 6 (Recognition Rate)

Character	Recognition Accuracy(%age)	Character	Recognition Accuracy(%age)
ੳ	98%	ਫ	93%
ਅ	97%	ਬ	96%
ੲ	95%	ਭ	90%
ਸ	98%	ਮ	98%
ਹ	98%	ਯ	95%
ਕ	90%	ਰ	98%
ਖ	92%	ਲ	92%
ਗ	92%	ਵ	95%
ਘ	90%	ੜ	90%
ਙ	93%	ਫ਼	87%
ਚ	98%	ਸ਼	88%
ਛ	95%	ਖ਼	87%
ਜ	98%	ਗ਼	80%
ਝ	92%	ਜ਼	80%
ਞ	90%	ਲ਼	80%
ਟ	98%	ਾ	95%
ਠ	94%	ਿ	94%
ਡ	92%	ੀ	93%
ਦ	92%	ੁ	92%
ਨ	93%	ੂ	92%
ਤ	95%	ੇ	92%
ਥ	93%	ੈ	93%
ਢ	92%	ੋ	93%
ਧ	93%	ੌ	95%
ਨ	95%	੍	95%
ਪ	93%	ੰ	90%

It's manually tested that the average word recognition accuracy of proposed system is 80%. There are one or more ways of writing each Punjabi character. Here, in this proposed system we had tried to cover maximum writing styles. So, through this system we have achieved average 75% of writer independency.

Our system is writer dependent. It depends on the style of writing. If one or more writers have same style of writing Punjabi character then the system will recognize in similar way. In this system, we try to cover maximum style of writing the Punjabi characters. If the writer style matches with available training set then the word or character is recognized otherwise not recognized or give wrong result.

Speech synthesis evaluation is based on Naturalness. If the word is present in database then the naturalness is 100%. If the word is not present then the naturalness is less (approx. 75%).

5. CONCLUSION

As the recognizing accuracy of proposed system is 70%. More work is there to improve recognizing rate of Punjabi words. As the style of writing some Punjabi character are same. In that case the ambiguity problem arises. We have to make some concepts to remove the ambiguity problem. To get best speech synthesis rate, database of system should be large. So, there is scope to increase the database of proposed system.

6. REFERENCES

[1] Abdul Rahim Ahmad, Christian Viard-Gaudin, Marzuki Khalid, Emilie Poisson (2004) "Online Handwriting Recognition using Support Vector Machine" TENCON 2004 IEEE Region 10 Conference, Vol. No. 1 **,** pp 311-314.

[2] Parinya Sangnansat, Widhyakom Asdornwised, and Somchai Jitapunkul (2004) "Online Thai Handwritten Character Recognition Using Hidden Markov Models and Support Vector Machines" International Symposium on Communications and Information Technologies (ISCIT2004) Sappom, Japan, pp-492-497.

[3] N. Sridhar Krishna, Partha Pratim Talukdar, Kalika Bali, A.G. Ramakrishnan (2004) "Duration Modeling for Hindi Text-to-Speech Synthesis System" INTERSPEECH-2004, pp 789-792.

[4] Muhammad Faisal Zafar, Dzulkifli Mohamad, and Razib M. Othman (2005) "On-line Handwritten Character Recognition: An Implementation of Counterpropagation Neural Net" World Academy of Science, Engineering and Technology vol 10, pp-232-237.

[5] Heiga Zen and Tomoki Toda (2005) "An Overview of Nitech HMM-based Speech Synthesis System for Blizzard Challenge 2005" INTERSPEECH2005, pp 93-96.

[6] Jerneja Zganec Gros, Ales Mihelic, Nikola Pavesic, Mario Zganec, Stanislav Gruden (2005) "Slovenian Text-to-Speech Synthesis for Speech User Interfaces" World Academy of Science, Engineering and Technology 11 2005, pp 1-5.

[7] H. Swethalakshmi1, Anitha Jayaraman, V. Srinivasa Chakravarthy, C. Chandra Sekhar (2006) "Online Handwritten Character Recognition of Devanagari and Telugu Characters using Support Vector Machines" Tenth International Workshop on Frontiers in Handwriting Recognition.

[8] Ujjwal Bhattacharya, Bikash K. Gupta and Swapan K. Parui(2007) "Direction Code Based Features for Recognition of Online Handwritten Characters of Bangla" Ninth international conference in ICDAR, Vol.1, pp 58-62.

[9] Bharath A, Sriganesh Madhvanath(2007) "Hidden Markov Models for Online Handwritten Tamil Word Recognition" 9th International Conference on Document Analysis and Recognition (ICDAR 2007), Curitiba, Brazil, pp 23-26.

[10] R. Jagadeesh Kannan, R. Prabhakar(2008) "An Improved Handwritten Tamil Character Recognition System using Octal Graph" Journal of Computer Science Vol 4, Issue 7: pp 509-516.

[11] Sara Izadi, Mehdi Haji, Ching Y. Suen (2008) "A New Segmentation Algorithm for Online Handwritten Word Recognition in Persian Script" ICFHR, pp 1140-1142.

[12] S. K. Parui, K. Guin, U. Bhattacharya and B. B. Chaudhuri (2008) "Online Handwritten Bangla Character Recognition Using HMM" Proceedings of 19th ICPR published by IEEE Computer Society Press, Digital Object Identifier 10.1109/ICPR.2008.4761835, pp 1-4.

[13] Tabassam Nawaz, Syed Ammar Hassan Shah Naqvi, Habib ur Rehman & Anoshia Faiz (2009) "Optical Character Recognition System for Urdu (Naskh Font) Using Pattern Matching Technique" International Journal of Image Processing, (IJIP) Vol no. 3 : Issue no.3, pp-92-104.

[14] Anuj Sharma, R. Kumar and R.K. Sharma (2009) "Recognizing Online Handwritten Gurmukhi Characters using Comparison of Small Line Segments" International Journal of Computer Theory and Engineering, Vol. 1, issue No.2, pp-131-135.

[15] S.V.Rajashekararadhya, P. Vanaja Ranjan (2009) "Support Vector Machine based Handwritten Numeral Recognition of Kannada Script" 2009 IEEE International Advance Computing Conference (IACC 2009) Patiala, Print ISBN 978-1-4244-2927-1 India, pp 381-386.

[16] T.Sitamahalakshmi Dr. A.Vinay Babu, M.Jagdeesh (2010) "Character Recognition Using Dempster-Shafer Theory Combing Different Distance Measurement Methods" International Journal of Engineering Science and Technology Vol. No. 2, Issue No. 5, pp 1177-1184.

[17] Rahul Kala, Harsh Vazirani, Anupam Shukla and Ritu Tiwari(2010) "Offline Handwriting Recognition using Genetic Algorithm" IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1,

[18] Sophea PRUM, Muriel Visani, Jean-Marc Ogier(2010) "On-line Handwriting word recognition using a bi-character model" 2010 International Conference on Pattern Recognition, Digital object identifier 10.1109/ICPR.2010.62, pp 2700-2703.

[19] Bharath A. and Sriganesh Madhvanath (2010) "On the Significance of Stroke Size and Position for Online Handwritten Devanagari Word Recognition: An Empirical Study" 2010 International Conference on Pattern Recognition. Digital object identifier:10.1109/IVPR.2010.501, pp 2033-2036.

[20] Walid A. Salameh, Mohammed A. Otair "Online Handwritten Character Recognition Using an Optical Backpropagation Neural Network" Issues in Informing Science and Information Technology.

[21] Stas Tiomkinyz, David Malahy, Slava Shechtmanz, and Zvi Konsz(2010) "A Hybrid Text-to-Speech System that Combines Concatenative and Statistical Synthesis Units" in IEEE transactions on Audio, Speech and Language Processing, Vol PP, Issue 99, pp 1-25.