A Hybrid Recognition & Speech Synthesis System for Handwritten Punjabi Words

Dinesh Kumar
Assistant Prof.and Head
Department of Information Technology
DAVIET, Jalandhar
erdineshk@gmail.com

Neeta Rana*
Student – M.Tech
Department of Computer Science
DAVIET, Jalandhar
neetarana@ymail.com

Abstract: - In this paper, a Handwriting Recognition and Speech Synthesis System is presented. Here, the recognition is of Online Handwritten Punjabi Words. A handwritten word is represented as a sequence of strokes. Support vector machines have been used for constructing the stroke recognition engine. Our proposed approach is database independent. It can recognize any Punjabi word based on the style of writing. To make it more interactive the concept of speech synthesis is also added in this proposed system. After recognition the proposed system speaks the recognized word. For speech synthesis, the technique used is CTTS (Concatenative Text-to-Speech).

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

I. INTRODUCTION

Handwriting Recognition plays an important role in the modern world. It can solve more complex problems and make human job easier. 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 of two types: Offline handwriting Recognition, Online handwriting Recognition. 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.[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]

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]

II. RELATED WORK

Lots of Research is available in Indian script recognition. Till now very less research is done for recognizing handwritten Punjabi words. The research for Punjabi Language is valuable for recognizing only 2460 Punjabi words that are already stored in the database of that system [14]. The proposed method of this system is called small line segments and it is based on idea of chain code and elastic matching techniques. Here recognition is done in two stages: First stage recognizes the stroke and character is evaluated in next stage on the basis of recognized strokes. Researchers have utilized many different approaches for both segmentation and recognition tasks of word

recognition. Vast work is done to recognize other languages like English, Tamil, Bengali, Persian, Urdu, Hindi, Bengali, Thai etc. In the fiels of Speech Synthesis, the researchers have used various techniques like Concatenative synthesis, Formant synthesis, Articulatory synthesis, HMM-based synthesis, Sinewave synthesis. Detail description of previous research is written below:

A. Research on Handwriting Recognition Neural Network:

Using Neural Network method online handwritten English characters are recognized [4][20]. An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by biological nervous systems, such as the brain. ANNs, like people, learn by example. An Artificial Neural Network is an adaptive, most often nonlinear system that learns to perform a function (an input/output map) from data. Neural network recognizers learn from an initial image training set. The trained network then makes the character identifications.

[a] ***Support Vector Machine:*** Support Vector Machine technique is used for recognizing online handwritten English Word[1], online Thai characters[2], online handwritten Devnagari & Telugu Words[7], Handwritten numerals of Kannada Script[15].

[b] ***Dempster -Shafer Theory:*** Dempster- Shafer theory is used to recognize handwritten Telugu Characters [16].This is a mathematical theory of evidence. It allows one to combine evidence from different sources and arrive at a degree of belief (represented by a belief function) that takes into account all the available evidence.

[c] ***Hidden Markov Model:*** Online Handwritten Tamil & Bangle Words are recognition using the concept of Hidden Markov Models [9][12]. HMMs have intrinsic properties that make them attractive for handwriting recognition. HMM use statistical algorithms that can automatically extract knowledge from training patterns, in contrast to knowledge edge-based approaches.

[d] ***Octal Graphs Technique:*** The proposed approach ventures a solution for offline handwritten recognition,

which converts the letter written into an octal graph, by representing each pixel of the given character as a node of a graph. Each node has eight fields so termed as octal graph. The graph tries to represent the basic form of a letter independent of the style of writing. Using the weights of the graphs and by the appropriate feature matching with the predefined characters, the written characters are recognized.[10]

[e]   Genetic Algorithms: The genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. The basic idea of genetic algorithm comes from the fact that it can be used as an excellent means of combining various styles of writing a character and generating new styles [17]. Human beings are able to recognize a new character because of their power to visualize parts of the known styles into the unknown character. This is the basic idea behind recognition using genetic algorithm.

B.   Research on Speech Synthesis

[a]   Concatenative synthesis: Concatenative synthesis is based on the concatenation (or stringing together) of segments of recorded speech. Generally, concatenative synthesis produces the most natural-sounding synthesized speech. However, differences between natural variations in speech and the nature of the automated techniques for segmenting the waveforms sometimes result in audible glitches in the output.[22] There are three main sub-types of concatenative synthesis: Unit selection synthesis, Diphone synthesis, Domain-specific synthesis. Combined benefit of Concatenative and Statistical Synthesis techniques are taken to develop Hybrid Text-to-Speech System.

[b]   Duration Modeling Technique: This technique is used for Hindi Text-to-Speech Synthesis [3]. The primary goal of this technique is to model the duration pattern of natural speech, considering various features that affect the pattern.

[c]   Hidden Markov Model (HMM): HMM based speech synthesis system developed in Nagoya Institute of Technology (Nitech-HTS) for a competition of text-to-speech synthesis systems using the same speech databases, named Blizzard Challenge 2005[5]. HMM-based synthesis is a synthesis method based on hidden Markov models, also called Statistical Parametric Synthesis. In this system, the frequency spectrum (vocal tract), fundamental frequency (vocal source), and duration (prosody) of speech are modeled simultaneously by HMMs. Speech waveforms are generated from HMMs themselves based on the maximum likelihood criterion.[22]

[d]   Corpus-Based Synthesis: This technique is used to synthesis Slovenian language [6].

III.   DESCRIPTION OF PROPOSED SYSTEM

The proposed system consists of two pahases. First phase is Online Handwritten Punjabi Word Recognition. Second phase is Speech synthesis of recognized word.

The technique used for developing recognizer is Support Vector Machine and for speech synthesis the technique used is CTTS (Concatenative Text-to-Speech). Support vector machines (SVMs) are a set of related

supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. After completion of First phase next phase is Speech Synthesis system. For developing proposed Speech Synthesis system, we used CTTS (Concatenative Text-to-Speech) synthesis technique. Database containing Punjabi Language's wavefiles is created. Finally we have integrated these two proposed systems. Here, output of first phase becomes input for next phase. Firstlly, it recognizes the word then it speaks the recognized word.

A.   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

B.   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 linespace 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.

C.   Feature Extraction & Stroke Classification

For this two approaches are used:

[a]   Single Recognition Engine Approach

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 classifed using SVMs. Kernel and associated parameters are experimentally determined. One-against-the-rest strategy is implemented for multiclass classification of strokes.

[b]   Multiple SVM Engines for Stroke Recognition

Multiple SVM engines have been explored for enhancing the recognition accuracy. Strokes are preclassified into two categories based on a threshold set for curve length. An SVM-based engine is constructed for each stroke category. Strokes with curvelength below the threshold are classified as small strokes. Small strokes are subjected to normalization and smoothing as mentioned in smaller number of points.

[c]   Character Identification from Strokes

A character consists of one or more strokes. All possible strokes in the script are identified from the dataset and the characters are expressed in terms of strokes. This results in the formulation of rules defining possible stroke combinations. There can be multiple rules for a character, which reflect the different ways in which it can be written.These rules are used to determine the character from

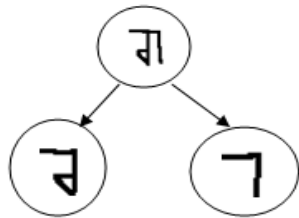a set of recognized strokes. The rule for writing the character "ਗ" is shown in Figure 1.

Figure 1(Rules for representing Punjabi Character "g")

After recognition of charater words are recognisation. Word recognition criteria are shown in Figure 2.
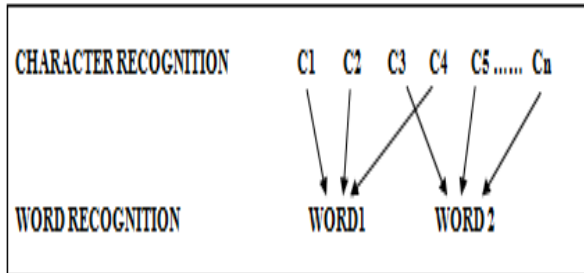

Figure 2 (Recognition Steps)

Second phase of this proposed system is Speech Synthesis. The speech synthesis is of recognized word. Database containing the wavefiles of Punjabi words is created. This speech generation method is concatenative TTS (CTTS). In this approach, speech is generated by concatenating the best compatible segments according to certain concatenation rules. Speech generated by this approach inherently possesses natural quality. However, its quality depends on the size of the recorded database, as high-quality CTTS needs an extensive database. Block Diagram of proposed system is shown in Figure 3.
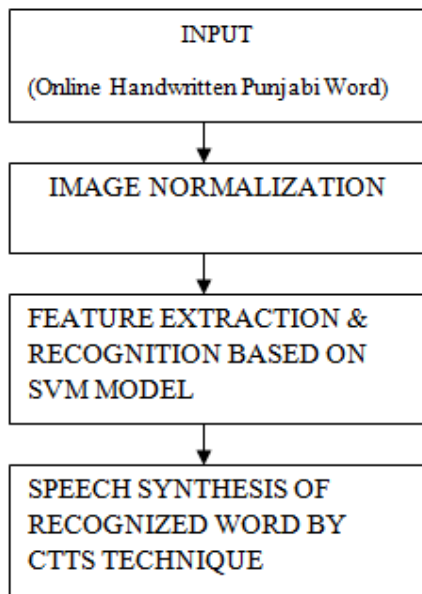

Figure 3 (Block Diagram of proposed System)

IV. 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]

A. *Character Set For Punjabi*

[a] *Consonants*

Punjabi language consists of 41 consonants. Consonants list of Punjabi language is written in Table 1.

Table 1 (List of Characters)

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

[b] *Vowels*

There are nine vowel phonemes in Punjabi. They are vowels making only one sound. All consonants use the vowel. Table2 show the vowels.

Table 2 (List of Vowels)

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

[c]   *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)

V.   WORKING OF RECOGNIZER & SPEECH SYNTHESIS SYSTEM

A.   Input

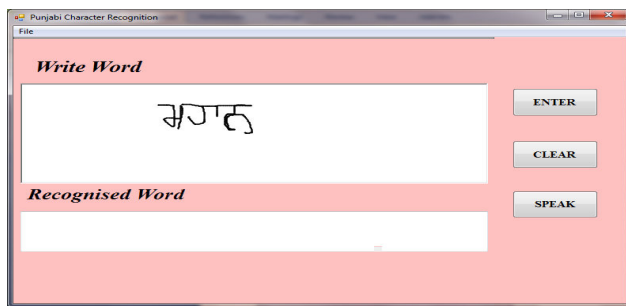Using any pointing device the user will write a Punjabi Word under "Write word" option.

Figure 4 (Input to the System)

Here, in Figure 4 the word written by the user is "`mhwn`".

B.   Processing

For processing, three controls are taken in the proposed system: Enter, Clear, Speak (see Figure 4). Enter is used to send handwritten word for processing. Clear is used to clear previous entry. Speak is used to Speak the recognized Word. Here the processing for recognizing handwritten word is by SVM & Speech synthesis is by CTTS.

C.   Output

After processing the result will be displayed under "Recognized Word" option. As shown in Figure 5.
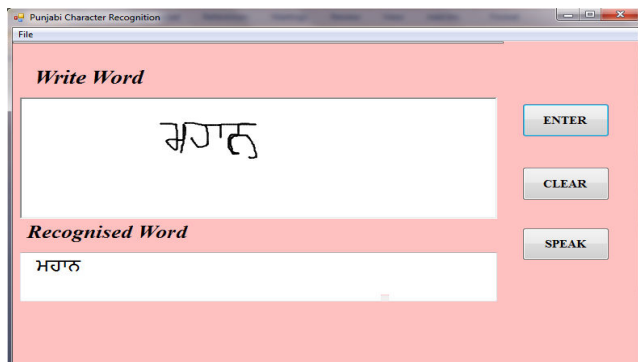

Figure 5 (Output of the Proposed System)

VI.   CONCLUSION

Till now the appropriate evaluation of the proposed system is not done. But we can evaluate our system based on Recognition accuracy & writer dependency and naturalness of speech.

Our proposed system is totally dependent on training sets. There are two or three ways of writing each Punjabi character. Here, in this proposed system we had tried to cover maximum writing styles. It will give maximum writer independency. By increasing the database of wavefiles the naturalness in speech can be achieved. Our proposed system works for recognizing online handwritten words of Punjabi. Furthur this concept can be applied to recognize Punjabi sentences.

VII.   REFERENCES

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

[2] Parinya Sangnansat, Widhyakom Asdornwised, and Somchai Jitapunkul "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, 2004.

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

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

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

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

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

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

[9] Bharath A, Sriganesh Madhvanath "Hidden Markov Models for Online Handwritten Tamil Word Recognition" 9th International Conference on

Document Analysis and Recognition (ICDAR 2007), Curitiba, Brazil, pp 23-26, 2007.

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

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

[12] S. K. Parui, K. Guin, U. Bhattacharya and B. B. Chaudhuri "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, 2008.

[13] Tabassam Nawaz, Syed Ammar Hassan Shah Naqvi, Habib Ur Rehman & Anoshia Faiz "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, 2009.

[14] Anuj Sharma, R. Kumar and R.K. Sharma "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, 2009.

[15] S.V.Rajashekararadhya, P. Vanaja Ranjan "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, 2009.

[16] T.Sitamahalakshmi Dr. A.Vinay Babu, M.Jagdeesh "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, 2010.

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

[18] Sophea PRUM, Muriel Visani, Jean-Marc Ogier "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, 2010.

[19] Bharath A. and Sriganesh Madhvanath "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, 2010.

[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 "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, 2010.

[22] www.wikipedia.com