2009 10th International Conference on Document Analysis and Recognition

Lexicon-based Word Recognition Using Support Vector Machine and Hidden Markov Model

Ahmad[1], A.R, Viard-Gaudin[2], C., Khalid[3], M,
[1]*Universiti Tenaga Nasional, Malaysia*
[2]*Universite de Nantes, France*
[3]*Universiti Teknologi Malaysia, Malaysia*
emails: abdrahim@uniten.edu.my, marzuki@utmkl.utm.my, christian.viard-gaudin@univ-nantes.fr

Abstract

Hybrid of Neural Network (NN) and Hidden Markov Model (HMM) has been popular in word recognition, taking advantage of NN discriminative property and HMM representational capability. However, NN does not guarantee good generalization due to Empirical Risk minimization (ERM) principle that it uses. In our work, we focus on online word recognition using the support vector machine (SVM) for character recognition. SVM's use of structural risk minimization (SRM) principle has allowed simultaneous optimization of representational and discriminative capability of the character recognizer. We evaluated SVM in isolated character recognition environment using IRONOFF and UNIPEN character database. We then demonstrate the practical issues in using SVM within a hybrid setting with HMM for word recognition by testing the hybrid system on the IRONOFF word database and obtained commendable results.

1. Introduction

There are two domains of handwriting recognition; online and offline which are differentiated by the nature of their input signals. Offline recognition uses static representation of a digitized document such as check, form, mail or technical documents. On the other hand, online recognition relies on information acquired during the production of the handwriting such as the coordinates of the writing trajectory, pen pressure and time. Mobile communication systems such as Personal Digital Assistant (PDA), electronic pad and smart-phone have online handwriting recognition capability integrated in them. It is important to try to improve the accuracy of the recognition engine in the devices while trying to constrain space for parameter storage and improving processing speed.

Handwriting recognition can be by character, word or sentence. A character recognizer is simple. It only needs to be trained with sample characters from the alphabets used in the language. Word recognizers are complex if they are general purpose but are simpler if it is based on specific lexicon. They can either be a holistic or segmentation based recognizer. In holistic recognizer, words are recognized as a whole and each word is modeled individually. In segmentation-based recognition (SegRec), character recognizer is used to recognize many hypothesis characters made up of combination of the smaller over segmented slices. The word recognizer combines the most suitable combination of hypothesis that gives the highest word score among all words in the lexicon. The best word score eventually determine the ultimate segmentation points of the word.

There are two ways to segment words; (a) Output Space Segmentation (OUTSEG) and (b) Input Space Segmentation (INSEG). OUTSEG approach allows segmentation points to be decided at the output space. Initially, the word signal is segmented implicitly into uniform size entities which can be overlapping, that is smaller or equal to character. Then, the recognition is carried out at the output space by associating groups of these smaller entities to form a particular character in a word. INSEG, on the other hand requires segmentation points to be decided explicitly by using spatial information in the word. All possible segmentation points are determined and cuts are made at these points. Then the cuts are combined into character hypotheses and passed to character recognizer. These character hypotheses can represent part of a character, a full character, a few characters, or part of a character combined with part of another characters.

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.248

The recognition process involves selecting only the character hypotheses that represents actual character signals forming a particular word. The observation probability for each character hypothesis is normally obtained by using a neural network (NN). The final step in handwriting recognition is to compare the handwritten word that is being recognized with the reference patterns to determine their similarity to decide which pattern or model best represents the word being recognized. SegRec word recognition is essentially a best path problem that incorporates character classification scores, segmentation information and possibly the language model.

Character classification score can be pure probability values or negative log probabilities. The overall score for a path in the recognition graph is given by the product of the character score of the arcs traversed. The probability of a given word is given by summing over all possible ways of character combinations to produce that word. The most probable word is the recognition result. The forward algorithm described in HMM is an efficient dynamic programming (DP) technique to compute the above sum. The Viterbi algorithm which picks the best single path in the graph as the recognition is often used to approximate the forward algorithm, for computational reasons. This means replacing the sum in equation with the largest term to make this approximation.

The main objective of this work is to investigate the usage of Support Vector Machine (SVM) for character recognition in place of NN within a hybrid of SVM/HMM system. The motivation for it is the work by Ganapathiraju [11] in speech recognition (SR) system using a similar method. SVM's usage in handwriting has been demonstrated by some researchers. However most reported work was just in character recognition. In this work, SVM is first used to train an online character recognizer using IRONOFF and UNIPEN character databases and the results compared with other methods. This has been reported in earlier work [9].

The character recognizer is then retrained with characters cut from IRONOFF word database and used in a hybrid SVM/HMM word recognizer. The word recognizer is used to find new segmentation points in a word and generate a new set of characters. The SVM character recognizer is then retrained with the new character database and again is used to improve the segmentation into characters. This is repeated until no more improvements can be made. The final word recognizer obtained can be used to recognize words based on medium sized lexicon in French and English language in the IRONOFF word database. The system were tested and performed fairly satisfactorily; with more than 98% recognition rate on a small English language lexicon.

This paper is organized as follows: Section 2 introduces SVM and some usages of SVM in handwriting reported in the literature. We describe our hybrid SVM/SVM word recognizer in section 3 followed with the description of the databases and experiments conducted to verify the system in the following section. Experiments results are discussed in section 5. Section 6 contains a summary that concludes this paper.

2. SVM in handwriting recognition

We first discuss SVM briefly and then move on to highlight some of the ways SVM have been used in handwriting. Our observation is that so far, in all cases reported, SVM have been more popularly used in character recognition rather than in word recognition.

2.1 SVM Theory

SVM has been used in recent years as an alternative to NN. SVM, unlike NN, takes into account learning examples as well as structural behavior. It achieves better generalization due to structural risk minimization (SRM). SVM formulation approximates SRM principle by maximizing the margin of separation.

Consider a set of l linearly separable data and its class $\{(x_1,y_1), (x_2,y_2),\dots, (x_l, y_l)\}$ where $x_i \in R^d$ and $y_i \in \{\pm 1\}$, the SVM classifier is:

$$f(x) = sgn\ (w.x + b) \qquad (1)$$

where w and b are parameters that maximize the margin with respect to the two classes. The classifier, expressed with input and output vectors information is as follows:

$$f(x) = sgn(\sum_{i=1}^{l} \alpha_i\, y_i (x_i.x) + b) \qquad (2)$$

where the parameters α_i for each corresponding input vectors needs to be found in order to find the maximal margin classifier. The α_i values are mostly zero and those inputs with non-zero α_i are called the support vectors and they contribute strongly towards the decision function. If the data set is not linear, the decision function used is:

$$f(x) = \mathrm{sgn}(\sum_{i=1}^{N} \alpha_i\, y_i K(x_i.x) + b) \qquad (3)$$

where a kernel K is used in the mapping of the non-linear input space into a linear space. The maximal margin classifier is found in the linear space. There are a few possible kernels that can be chosen; Linear, Polynomial, Radial basis function and Hyperbolic tangent. SVM training involves solving a convex quadratic programming (QP) problem with equality and inequality constraints obtained by the objective of margin maximization. The solution solves for the nonzero parameters α_i's in the formulation and extracts the support vectors corresponding to it.

2.2 Usages of SVM in handwriting recognition

SVM have been used in handwriting recognition in a number of ways, among others; as a standalone recognizer in a fixed feature based character recognition system, as a replacement of HMM in a sequence processing based character recognizer or as a final decider in the final output of a handwriting recognition system. Bahlmann [1] describes an approach for on-line character recognition that combines dynamic time warping (DTW) and SVM by establishing a new kernel called the Gaussian DTW (GDTW) kernel. Bahlmann compared his SVM based approach to an HMM based technique using the UNIPEN character database and showed that he achieve comparable results.

Camastra [2] describes a cursive character recognizer as a module in an offline cursive word recognition system based on a segmentation and recognition approach. The character classification is done by using SVM and a Neural Gas. The Neural Gas is used to verify whether lower and upper case version of a certain letter can be joined in a single class or not. Once this is done for every letter, the character recognition is performed by SVM.

Many researches use different methods for combining multiple classifiers to compensate the weakness of one classifier, by the strength of the other classifiers. Bellili [3] uses a combination of multilayer perceptron (MLP) neural network and SVM classifiers. The SVMs are used to improve the performances of an MLP based digit recognizer. The hybrid SVM/MLP architecture is based on the idea that the correct digit class of the recognizer almost systematically belongs to the two maximum MLP outputs and that some pairs of digit classes constitute the majority of the recognizer errors. Specialized local SVMs are introduced to detect the correct class among these two classification hypotheses. The hybrid MLP-SVM recognizer achieves a recognition rate of 98.1%, for real mail zipcode digits recognition task, a performance better than several classifiers reported in earlier researches.

SVM was also used in non-Roman handwriting recognition such as in Thai language, Arabic and Devanagari/Telugu. Sanguansat [4] uses HMM and SVM with score space kernel which is a generalized Fisher kernel for online Thai handwritten character recognition. HMM is first used for multi-classification, and then SVMs are applied to resolve any uncertainty remaining after the first-pass HMM-based recognizer. Chakravarthy [5] uses SVM for online handwritten character recognition for Indian scripts. A number of separate feature vector combinations were used and compared. The standard gaussian kernel is used for training.

In summary, a search in the literature for online word recognition shows that a hybrid word recognizer using SVM and HMM has not been attempted.

3. Hybrid SVM/HMM word recognizer

The general layout of the architecture for our handwritten word recognition system is depicted in figure 1. It is a SegRec recognizer and lexicon-driven which is similar in structure to the works by Tay [6] and Caillault [7]. Table 1 compares our system with the two. The main difference which is the centre of our thesis is the use of an SVM in place of the ANN and TDNN for the character recognizers. The training was done at the character level; the reason being that normally, SVM training does not involve correcting gradients like ANN but requires quadratic optimization. [1]

Table 1 Comparison of work

Author	Segment. method	Domain	Character Recognizer	Training method
Tay [6]	INSEG	Off-line	ANN	Character and word level
Caillault [7]	OUTSEG	On-line	TDNN	Word level
Our work	INSEG	On-line	SVM with RBF kernel	Character level

The input word to the recognizer is preprocessed and over segmented into slices. Slices are combined to generate character hypothesis which will be recognized by the trained character SVM [9]. The HMM Viterbi algorithm with the help of a lexicon is used to obtain the N-best list of the words from the lexicon.

[1] It is observed that Telstra Australia has patented a gradient-based SVM training method for SVM recently. See [8]

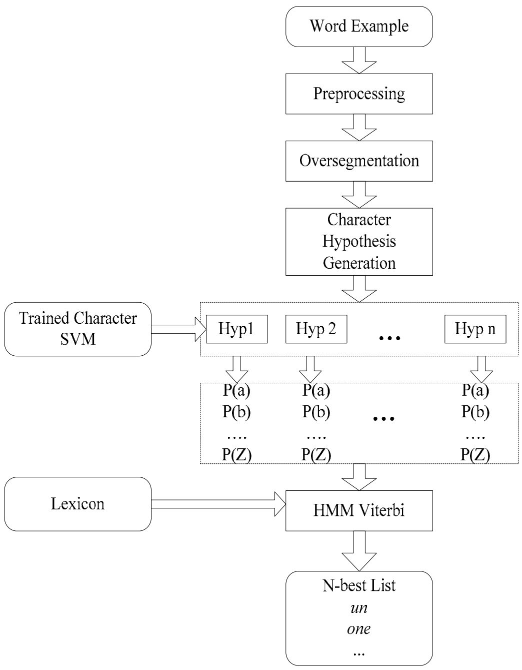

Figure 1 Architecture of the word recognition system

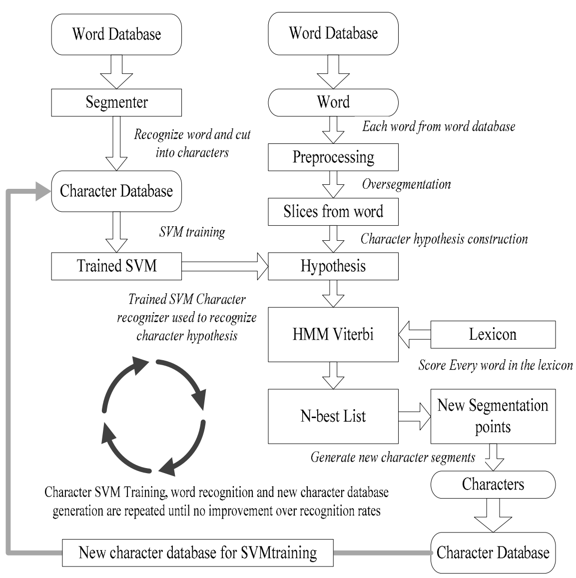

Figure 2 Training process for the word recognizer

3.1 Overview of Hybrid SVM/HMM Training

The training process for the word recognizer is given in Figure 2. The aim of the training is to optimize the character SVM by using the set of characters which are the best segmented from the word as a result of the recognition. The quality of the word recognizer relies on the quality of the SVM character recognizer, which in turn relies on the segmentation made on the words into characters. Initial training uses the bootstrap SVM character recognizer. The objective in training of the system is then to try to improve further the initial bootstrap SVM character recognizer by improving the segmentation of the input word while recognizing it.

In Figure 2, it can also be seen that the front-end involves preprocessing, segmentation, character hypothesis generation from the segments and feature extraction of the hypothesis. These are followed by the recognition portion of the system which involves the recognition of each character hypothesis by using the SVM character recognizer and the use of Viterbi algorithm or dynamic programming algorithm to find the word score for each word in the lexicon given the input word signal.

The initial hybrid SVM/HMM word recognition system uses a character recognizer trained using the set of characters which were cut from words in the IRONOFF word databases. This is necessary as isolated characters are different from characters cut from words. Initially, we used a commercial handwriting recognizer to cut the word database into our initial isolated character database. For testing and comparison purposes, we have trained separately the character SVMs for characters obtained from cheque words, characters obtained from French words and characters obtained from English words. These SVMs were used in separate hybrid word recognizers catering for the respective word types for testing purposes.

4. Databases and Experiments

We only make use of IRONOFF [10] online word database for training and testing the hybrid word recognizer. The database is limited to the lexicon that consists of 197 words from the English and French language. 30 of the French words are French cheque words. To evaluate our hybrid word recognition system, we use the complete word database to measure the percentage of correctly recognized words from the database (those words having true class in position 1). We also compare the performance of the word recognizer by comparing the percentage position of the true class in certain top position other than 1, between 2 to n. For example, we can take a word as "correctly recognized" if the true class is in the top 10 position and compare recognition performances based on this.

5. Results and discussions

The baseline system is the word recognizer using the character SVM trained with the characters segmented from the word with the commercial recognizer. We performed recognition on the individual word databases and regenerate new character database using the segmentation points generated during the recognition. We initially obtained the word recognition rates as shown in Table 2. The result shows that Top(1) recognition rate is not very high, where all except English word databases gives below 98% recognition rate. However, the Top(10) recognition rate of almost or above 99%, indicates that although the recognizers made some errors in word recognition, they are still able to recognize well within the Top(10) positions.

Table 1 Initial word recognition

Database	Lexicon size	Recognition rate (%)			
		Top(1)	Top(2)	Top(3)	Top(10)
English	26	98.77%	99.44%	99.50%	100%
Cheque	30	76.71%	91.64%	95.71%	99.99%
French	171	63.25%	77.90%	85.15%	98.86%
All words	197	64.53%	79.05%	86.20%	99.91%

To demonstrate further effectiveness of the system, we retrain the individual character SVM for the English word database. The character SVMs recognition rate, the number of support vectors for each character SVM and the word recognition rate of the word recognizer based on the new character SVM are given in Table 3.

Table 2 Improvements in English word recognizer

Iteration	Character SVM Rec. Rate (%)	nSV	Word Rec.Rate Top(1)
Baseline	80.46	8591	98.77%
First iteration	74.37	8449	98.49%
Second iteration	74.27	8296	98.83%
Third iteration	74.32	8349	98.99%

As observed from the table, the performance of the new character SVM were less than the baseline character recognizer but it did not change the word recognition rate too much. The word recognition rates were within 1% above and below the baseline word recognition rate. The number of support vectors is also generally lower than the baseline character recognizer which means the SVM model is getting smaller in size. As we performed word recognition and resegmentation repeatedly, the word recognition rate converges to around 98.9%.

6. Conclusion

We have demonstrated the viability of using a hybrid of SVM and HMM in word recognition. Although the result of recognition using IRONOFF word database is not very promising for the bigger lexicon, it shows that the method is feasible especially for smaller English language lexicon.

References

[1] BAHLMANN, C. H., B. BURKHARDT, H. (2002) Online handwriting recognition with support vector machines - a kernel approach. Eighth International Workshop on Frontiers in Handwriting Recognition, 2002.

[2] CAMASTRA, F. (2007) A SVM-based cursive character recognizer. Pattern Recognition, 40, 3721-3727.

[3] BELLILI, A., GILLOUX, M., GALLINARI, P. (2000) An Hybrid MLP-SVM Handwritten Digit Recognizer.

[4] SANGUANSAT, P., ASDORNWISED, W., JITAPUNKUL, S. (2004) Online Thai handwritten character recognition using hidden Markov models and support vector machines. International Symposium on Communications and Information Technologies, ISCIT 2004. Japan.

[5] CHAKRAVARTHY, V. S., CHANDRASEKHAR, C. (2007) Online Handwritten Character Recognition Systems for Devanagari, Telugu and Kannada. Online Handwritten Character Recognition Consortium. Madras, India, Indian Institute of Technology.

[6] TAY, Y. H., KHALID, M., YUSOF, R., VIARD-GAUDIN, C. (2003) Offline Cursive Handwriting Recognition System based on Hybrid Markov Model and Neural Network. IEEE Int'l Symp on Computational Intelligence in Robotics and Automation (CIRA-2003) Kobe, Japan.

[7] CAILLAULT, E., VIARD-GAUDIN, C. (2006) Using Segmentation Constraints in an Implicit Segmentation Scheme for On-line Word Recognition. 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2006) La Baule, France.

[8] KOWALCZYK, A. (2001) A Gradient Based Training Method for A Support Vector Machine. IN LTD., T. R. D. M. P. (Ed. Patent Office, Australia, Canberra. G06F 15/18 ed. Australia.

[9] Ahmad, A. R., Khalid, M., Viard-Gaudin, C., Poisson, E. (2004). Online Handwriting Recognition Using Support Vector Machine. IEEE Region 10 Conference TENCON 2004. . Chiangmai Thailand. 1: 311- 314.

[10] VIARD-GAUDIN, C., LALLICAN, P.M., KNERR, S., AND BINTER, P. (1999) The IRESTE On/Off (IRONOFF) Dual Handwriting Database. International Conference on Document Analysis and Recognition. Bangalore.

[11] GANAPATHIRAJU, A. H., J.E.; PICONE, J. (2004) Applications of support vector machines to speech recognition. IEEE Transactions on Signal Processing, 52, 2348 - 2355