Journal of Engg. Research **Vol. 4 No. (1) March 2016** pp. 1-20

تطوير جهاز الأتجاه المدعم (SVM) بإستخدام الطريقة الأمثل لتقنية الافتراس- الضحية وطريقة (هوك - جيف) للتعرف على المحادثة

***تينا ميتال و**ر.ك. شارما**

*قسم الإلكترونيات وهندسة الإتصالات' جامعة ثابار، باتيالا، البنجاب، الهند

**كلية الرياضيات وتطبيقات الحاسوب، جامعة ثابار، باتيالا، البنجاب، الهند

الخلاصة

إن التعرف التلقائي على المحادثات يتطلب تمثيل الكلمات بمجموعة سمات. حجم هذه المجموعات له أهمية كبيرة لأنه يؤثر على دقة التعرف ووقت التعرف ومتطلبات الذاكرة وكذلك تعقيد النموذج. جهاز دعم النواقل بالاتجاهات له تطبيق جيد في عملية التعرف على التخاطب. ولكن أداء هذا الجهاز يتأثر بمعاملاته. في هذا البحث تم تقديم تقنية مثلى مهجنة لتطوير القدرة التعليمية لجهاز الاتجاه المدعم لاختيار أفضل مجموعة سمات. هذه التقنيه المهجنة تكامل بين الطريقتين المذكورتين (الفريسة – المفترس وهوك – جيف). للتعامل مع متغيرات القرار ذات الطبيعة المركبة تم نمذجة طريقة مثلى ثنائية لتقنية (الفريسة – المفترس) ولزيادة البحث طبقت طريقة (هوك – جيف) أيضاً.

An improved SVM using predator prey optimization and Hooke-Jeeves method for speech recognition

Teena Mittal* and R. K. Sharma**

**Department of Electronics and Communication Engineering, Thapar University, Patiala, 147004, Punjab, India*

***School of Mathematics and Computer Applications, Thapar University, Patiala, 147004, Punjab, India, Email: rksharma@thapar.edu*

**Corresponding author: tnarang28@gmail.com*

ABSTRACT

For automatic speech recognition, speech signal is represented in terms of feature set. Size of the feature set is an important aspect as it affects recognition accuracy, computational time, memory requirement and complexity of the model. Support vector machine has good application prospects for speech recognition; nevertheless, performance of support vector machine is affected by its parameters. In this research work, a hybrid optimization technique is proposed to improve the learning ability of support vector machine and to select the most appropriate feature set. The hybrid technique integrates predator-prey optimization and Hooke-Jeeves method. To deal with mixed type of decision variables, binary predator-prey optimization technique has also been introduced. During the initial phase, search is performed by predator-prey optimization and to further exploit the search, Hooke-Jeeves method is applied. The proposed technique with support vector machine has been implemented to recognize TI-46 isolated word database in clean as well as noisy conditions and self-recorded Hindi numeral database. The experimental results obtained by proposed technique with support vector machine shows improved recognition rate. Furthermore ROC curve is analysed to verify sensitivity and specificity of results obtained by proposed technique with support vector machine.

Keywords: Feature selection; Hooke-Jeeves method; Predator prey optimization; Speech recognition; Support vector machine.

INTRODUCTION

Automatic speech recognition needs a preprocessing stage that extracts features, to obtain significant improvements in the recognition rate (Rabiner & Juang, 1993). Various signal processing techniques such as filter banks, linear prediction coding, cepstrum analysis (Rabiner, 1978), and Mel-Frequency Cepstral Coefficients

(MFCC) (Davis & Mermelstein, 1980) were addressed to extract the features. Since recognition rate and computation time are very much affected by the choice of feature set, feature selection algorithms are used to choose the most significant features. The algorithms for feature selection are classified as, wrapper method and filter method. The wrapper method generally achieves better recognition rate as compared to filter method because wrapper method utilizes the actual target learning algorithm whereas the filter method applies intrinsic data to select features while ignoring the target learning algorithm. However, these methods are computationally impractical because of high computational time and memory requirement. To overcome the disadvantages of the wrapper and filter methods, global search techniques are applied to select the most relevant feature set. Hua-chao *et al.* (2007) have proposed a feature selection and classification method for hyper-spectral images by integrating particle swarm optimization (PSO) algorithm with support vector machine (SVM). Liu *et al.* (2011) have designed a modified multi-swarm PSO for feature selection. To improve classification accuracy in binary problems, Sarafrazi & Nezamabadi-pour (2013) have proposed a hybrid system consisting of gravitational search algorithm with SVM. An ant colony optimization method with SVM is applied to improve the classification accuracy with appropriate feature subset by Huang (2009). Hu *et al.* (2015a) have applied firefly algorithm (FFA) to select proper input features for midterm interval load forecasting.

Support vector machine is a well-accepted pattern classification technique with different applications. The radial basis function (RBF) is the most commonly used kernel as it can classify multi-dimensional data and requires only two parameters (C and γ) to be set. Various optimization techniques have been proposed by several researchers to optimize SVM parameters. Xiang *et al.* (2006) have optimized SVM parameters by integration of PSO with the simulated annealing algorithm. Yan-bin *et al.* (2009) have applied modified PSO to optimize the key parameters of SVM. Bao *et al.* (2013) have optimized SVM parameters by applying the memetic algorithm based on PSO and pattern search. He *et al.* (2012) have applied PSO-SVM based technique to classify electronic nose data. Wei *et al.* (2011) have presented a method based on SVM and PSO for face recognition. İlhan & Tezel (2013) have developed genetic algorithm (GA) and SVM based method for parameters optimization to predict single nucleotide polymorphisms. Hu *et al.* (2013) have proposed a FFA based memetic algorithm to select the parameters of support vector regression forecasting model. Mandal *et al.* (2013) have presented a hybrid algorithm utilizing data filtering technique to forecast day-ahead electricity prices.

To improve the speech recognition rate, proper choice of SVM model parameters and most relevant feature subsets of speech is required (Zhao *et al.*, 2011). Shih-Wei *et al.* (2008) have developed PSO based approach for parameters determination of

SVM along with feature selection. A combination of discrete and continuous PSO (Huang & Dun, 2008) is applied to select the most relevant feature subset and to search SVM kernel parameters. Huang & Wang (2006) have presented a GA based approach to simultaneously optimize the parameters of SVM and feature subset for pattern classification. Hu *et al.* (2015b) have proposed a hybrid filter-wrapper approach for short-term load forecasting feature selection.

Although some of the global search techniques are applied to optimize feature set and to fine-tune the SVM parameters, these techniques still have some drawbacks while searching for the optimum solution. Omran *et al.* (2009) have discussed some of the deficiencies of GA and PSO. Sometimes, GA lacks the ability to produce better offspring and causes slow convergence near global optimum solution. Though PSO is one of the most promising global search techniques, it is possible that solution may converge to local solution if prey particles come close together during the search. To overcome this deficiency, many approaches have been proposed in the literature. One of the potential approaches is to include the predator particle with prey particles in a swarm as predator-prey optimization (PPO) model (Silva *et al.*, 2002). The predator attracts towards global best prey particle and prey particles try to escape from a predator that enhances the searching capability of the PPO. The main motive of predator and prey interaction is to maintain the balance between exploration and exploitation capability of PPO technique. Recently, Narang *et al.* (2014) applied PPO based technique to solve scheduling problem and concluded that the solutions obtained by PPO technique outperform the solutions obtained by PSO. Silva & Gonçalves (2013) have applied heterogeneous PSO for training of SVM. In yet another attempt, Silva & Gonçalves (2014) have investigated the use of scouting predator-prey optimizer to train SVM with non positive definite kernels. However, heterogeneous PSO and scouting PPO have not been applied to deal with mixed variables type problem.

Current research focuses on the integration of local and global search techniques. Harman & McMinn (2010) have discussed theoretical and empirical aspects of integration of local and global search techniques. They have concluded that such integration improves the convergence rate and computational time of global search techniques.

Researchers claim the superiority of an algorithm on the basis of test results. These claims need to be supported by evidence obtained from statistical tests. To compare the validity of different models, researchers have applied a number of statistical tests. Demšar (2006) has reviewed and recommended a set of non-parametric statistical tests and procedure to compare the classifier performance over multiple data sets. García & Herrera (2008) have given some statistical procedure to compare $n \times n$ classifiers. The symmetric mean absolute percentage error and mean absolute scaled error have extensively been used to evaluate the performance of different forecasting models

(Xiong *et al.*, 2013; Xiong *et al.*, 2014; Bao *et al.*, 2014a). The receiver operating characteristics (ROC) is a well recognized graphical technique used to evaluate the quality of classification models in various fields (Bradley, 1997).

The main contributions of this paper in the context of feature selection and to optimize SVM kernel parameters to improve speech recognition rate, are fourfold:

(1) A binary version of PPO technique is proposed.

(2) The integration of PPO, a global search technique with Hooke-Jeeves method, a local search technique is proposed.

(3) The proposed technique is tested on standard speech database TI-46 in clean as well as noisy environment for the recognition of isolated words. In addition, to further verify validity of the technique, it is tested on self recorded Hindi numerals speech database.

(4) The results obtained by proposed technique with SVM have also been validated using ROC curves and found satisfactory.

The remainder of the paper is organized into eight sections *i.e.,* feature extraction, support vector machine, predator prey optimization, Hooke-Jeeves method, development of proposed algorithm, experimental details, results and discussion, and conclusions.

FEATURE EXTRACTION

The feature extraction process converts the speech waveform to a set of features that are further explored in the classification process. The main aim of feature extraction is to extract attributes from the speech signal to differentiate between a wide set of distinct words. A wide range of techniques exists for representing the speech signal, such as linear prediction coding, MFCC, perceptual linear prediction, etc., out of which MFCC features are most commonly used. To extract the MFCCs, the speech sample passes through a hamming window to minimize the discontinuities of the signal. Then, discrete Fourier transform (DFT) is used to generate the Mel filter bank. After Mel frequency warping, a number of coefficients are obtained. Finally the inverse DFT is used for the Cepstral coefficients calculation (Rabiner, 1978). The steps involved in the MFCC feature extraction are shown in Figure 1. The features used in this work are MFCCs, their first derivatives (delta coefficients) and second derivatives (acceleration coefficients).

SUPPORT VECTOR MACHINE

Support vector machine, proposed by Vapnik (1995), is a binary classifier for pattern classification. For a binary classification, the training dataset is $\{(x_{i,} y_i), i = 1, 2, \ldots, n\}$, where the input space $x_i \in R^n$ and $y_i \in \{1,-1\}$ are the labels of the input space x_i and n denotes the number of data items in the training set. The hyper-plane is defined as $w.x_i + b = 0$, where x_i is a point on the separating hyper-plane, w is the normal vector to hyper-plane and b is bias value. The optimal hyper-plane to separate the two classes is obtained by minimizing the regularized training error,

$$E = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i$$

$$\text{subject to } y_i(\langle w, x_i\rangle + b) \geq 1 - \xi_i,\ \ \xi_i \geq 0,\ i = 1,2,\ldots,n \tag{1}$$

where $\langle,\rangle$ denotes the inner product, ξ_i is a slack variable, which defines the permitted misclassification error, C is the penalty coefficient and it specifies the trade-off between the empirical risk and the regularization term.

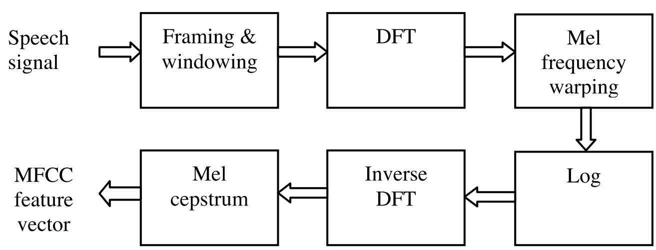

Fig. 1. Block diagram of MFCC feature extraction process

For simplification, Equation (1) can be transformed into the dual problem as:

$$max\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i, x_j)$$

$$\text{subject to } \sum_{i=1}^{n}\alpha_i y_i = 0,\ 0 \leq \alpha_i \leq C\ ,\ \gamma > 0 \tag{2}$$

where, α_i is the Lagrange multiplier and $K(x_i, x_j) = \Phi(x_i)\ \Phi(x_j)$ is the kernel function, which can map the data into a higher dimensional space through some nonlinear mapping $\Phi(x)$.

The kernel parameters should be carefully chosen as it implicitly defines the structure of the high dimensional feature space. RBF kernel has been undertaken in this work because it can classify multi-dimensional data and requires very few parameters to set. It is represented as:

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2}\|x_i - x_j\|^2\right) \tag{3}$$

A simplified form of this function is:

$$K(x_i, x_j) = \exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \gamma > 0, \text{ and } \gamma = \frac{1}{2\sigma^2} \tag{4}$$

Here, σ (or γ) is the kernel parameter.

PREDATOR PREY OPTIMIZATION

Kennedy & Eberhart (1995) have proposed continuous PSO and binary PSO (Kennedy & Eberhart, 1997) to deal with problems that have continuous and discrete decision variables, respectively. In binary PSO, velocity is a probability vector and it computes the likelihood of the binary variables having a value of one (Bao *et al.*, 2014b). In PPO algorithm, predator particle is also included in addition with prey particles. The influence of predator on preys is controlled by a factor *probability fear* (*pf*). The predator always tries to chase global best prey particle.

In continuous PPO, velocity of predator particle V_{Pi}^t is updated as:

$$V_{Pi}^{t+1} = C_4\left(Gbest_i^t - X_{Pi}^t\right), \; i = 1, 2, \ldots, n \tag{5}$$

where n is the number of dimensions; C_4 is a uniformly distributed scaled random number and it determines how quickly predator catches the global best prey particle; $Gbest_i^t$ is the global best prey position for i^{th} dimension, at t^{th} movement.

The prey velocity V_{il}^t is updated as:

$$V_{il}^{t+1} = \begin{cases} w \times V_{il}^t + C_1\left(Lbest_{il}^t - X_{il}^t\right) \times rand() + C_2\left(Gbest_i^t - X_{il}^t\right) \times rand() & pf < pf\max \\ w \times V_{il}^t + C_1\left(Lbest_{il}^t - X_{il}^t\right) \times rand() + C_2\left(Gbest_i^t - X_{il}^t\right) \times rand() + C_3 a_i \exp\left(-b_i d\right) & pf \geq pf\max \end{cases}$$

$$l = 1, 2, \ldots, NP; \; i = 1, 2, \ldots, n \tag{6}$$

where NP is prey particles in the swarm; C_1 and C_2 denote acceleration constant; C_3 is a uniformly distributed scaled random number and it controls the effect of predator particle on prey particles; parameter w is inertia weight and its value varies from 0.9 to 0.4 with movement; PPO parameters a_i and b_i affect the maximum amplitude of the predator influence over prey particles; $Lbest_{il}^t$ is local best prey position for i^{th} dimension of l^{th} particle at t^{th} movement; pf and $pf\max$ represent probability fear and maximum value of probability fear, respectively; d is Euclidean distance between predator and prey; *rand*() is a uniformly distributed random number.

For continuous PPO, positions of predator X_{Pi}^t and prey particles X_{il}^t are updated as:

$$X_{Pi}^{t+1} = X_{Pi}^t + V_{Pi}^t, \; i = 1, 2, \ldots, n \tag{7}$$

$$X_{il}^{t+1} = X_{il}^{t} + V_{il}^{t} \ , \ l = 1,2,\ldots,NP;\ i = 1,2,\ldots,n \tag{8}$$

In binary PPO, particles are restricted in the search area between zero and one, so the velocities of predator and prey particles are updated by sigmoid limiting transformation as:

$$S(Velocity) = \frac{1}{1 + \exp(-Velocity)} \tag{9}$$

The predator position is updated as:

$$X_{Pi} = \begin{Bmatrix} 1 & rand() < S(V_{Pi}) \\ 0 & otherwise \end{Bmatrix},\ i = 1,2,\ldots,n \tag{10}$$

The prey positions are updated as:

$$X_{il} = \begin{Bmatrix} 1 & rand() < S(V_{il}) \\ 0 & otherwise \end{Bmatrix},\ l = 1,2,\ldots,NP;\ i = 1,2,\ldots,n \tag{11}$$

HOOKE-JEEVES METHOD

Hooke-Jeeves method (Rao, 1996) is a direct search method and, therefore, it does not need any derivative information of fitness function and constraints. In this method, search is performed by sequential examination of trail solution by performing exploratory and pattern moves. Initially, exploratory move is executed to update the trail solution. If the updated and trail solutions are different then exploratory move is a success. The pattern move is performed by using two successive points obtained from exploratory move. The whole procedure is repeated until a termination criterion is met. The flow chart of Hooke-Jeeves method is shown in Figure 2.

DEVELOPMENT OF PROPOSED ALGORITHM

The proposed technique is applied to optimize the parameters of SVM kernel and to select the most relevant feature set. Initially global search technique 'PPO' is applied to explore the search area and to avoid possible local solution, local search 'Hooke-Jeeves' method is applied. The implementation of the proposed technique is described in the following sub-sections.

Particle representation

The RBF kernel parameters (C, γ) are continuous in nature and feature mask is of a binary nature. The particles are divided in two parts; first part consists of feature mask and the second part contains RBF kernel parameters. The feature mask is either 0 or 1; if the mask is 0, then corresponding feature will not be selected and if it is 1, then corresponding feature will be selected (Bao *et al.*, 2014a; Xiong *et al.*, 2015). Table 1 shows the representation of particle *l* with a dimension of $N_F + 2$.

Table 1. l^{th} prey particle representation

Input feature mask				C	γ
Discrete				Continuous	
$X_{1,l}$	--	--	$X_{N_F,l}$	$X_{N_F+1,l}$	$X_{N_F+2,l}$

where N_F represents number of features of speech; $X_{1,l}$ represents 1^{st} dimension of l^{th}decision variable.

Initialization of prey particles

The initial prey position and velocity representing feature mask are randomly generated as follows:

$$X_{il}^{0} = \begin{Bmatrix} 1 & if\ rand\,() \geq 0.5 \\ 0 & otherwise \end{Bmatrix}, \quad i = 1, 2, ..., N_F; \; l = 1, 2, ..., NP \tag{12}$$

$$V_{il}^{0} = \begin{Bmatrix} 1 & if\ rand\,() \geq 0.5 \\ 0 & otherwise \end{Bmatrix}, \quad i = 1, 2, ..., N_F; \; l = 1, 2, ..., NP \tag{13}$$

where NP is number of prey particles.

The prey position and velocity representing SVM kernel parameters (C and γ) are randomly initialized as:

$$X_{il}^{0} = X_i^{\min} + rand\,() \times \left(X_i^{\max} - X_i^{\min}\right), \quad i = N_F + 1, N_F + 2; \; l = 1, 2, ..., NP \tag{14}$$

$$V_{il}^{0} = V_i^{\min} + rand\,() \times \left(V_i^{\max} - V_i^{\min}\right), \quad i = N_F + 1, N_F + 2; \; l = 1, 2, ..., NP \tag{15}$$

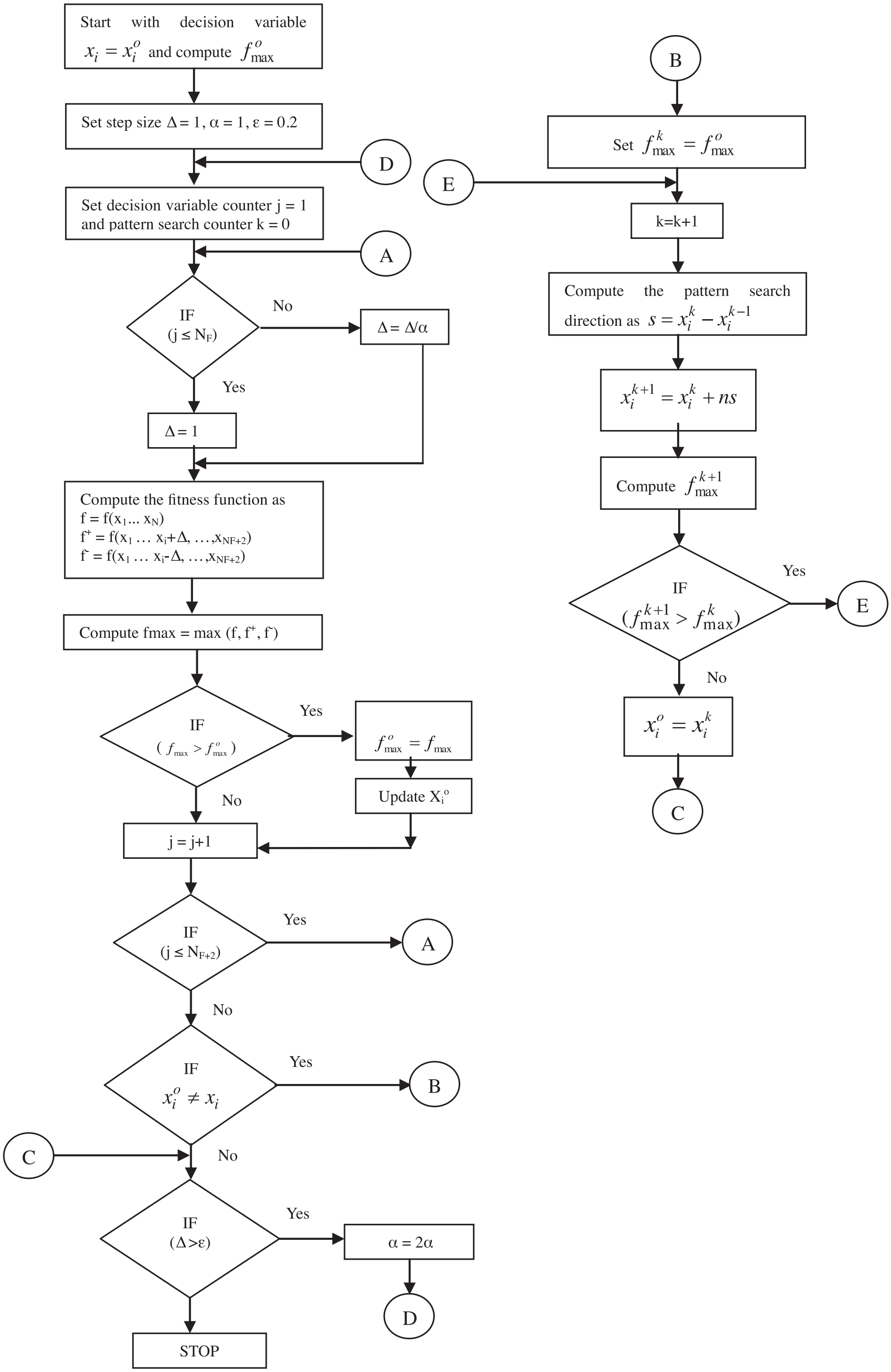

Fig. 2. Flow chart of Hooke-Jeeves method

Initialization of predator particle

The predator position and velocity representing feature mask are randomly generated as follows:

$$X_{pi}^{0} = \begin{Bmatrix} 1 & if\ rand() \geq 0.5 \\ 0 & otherwise \end{Bmatrix},\ i = 1, 2, ..., N_F \tag{16}$$

$$V_{pi}^{0} = \begin{Bmatrix} 1 & if\ rand() \geq 0.5 \\ 0 & otherwise \end{Bmatrix},\ i = 1,2, ..., N_F \tag{17}$$

The predator position and velocity representing SVM kernel parameters (C and γ) are randomly initialized as:

$$X_{Pi}^{0} = X_i^{\min} + rand\,()\left(X_i^{\max} - X_i^{\min}\right),\ i = N_F + 1, N_F + 2 \tag{18}$$

$$V_{Pi}^{0} = V^{\min} + rand\,()\left(V_i^{\max} - V_i^{\min}\right)\ ,\ i = N_F + 1, N_F + 2 \tag{19}$$

Fitness function

The two objectives of fitness function are classification accuracy and the feature set. There is a trade-off between these two objectives because accuracy of speech recognition is high with higher number of features and vice versa. Sarafrazi & Nezamabadi-pour (2013) have combined these objectives to achieve high accuracy with less number of features as:

$$Fit = W_1 \times F_1 + W_2 \times F_2 \tag{20}$$

where Fit is fitness function; W_1 represents the weight for recognition accuracy function F_1 and W_2 is the weight for feature set function F_2. The values of W_1 and W_2 are set to 1.

The recognition accuracy function F_1 is defined as:

$$F_1 = \frac{CR}{CR + IR} \times 100 \tag{21}$$

Here, CR represents the number of correct recognitions and IR represents the number of incorrect recognitions.

For the binary classifier, classes are divided into positive class and negative class. Thus, the classified test points can be divided into four categories that are represented in the confusion matrix. Table 2 represents the confusion matrix that illustrates relationship among these indices. The accuracy can also be represented in terms of indices of confusion matrix as:

$$F_1 = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \tag{22}$$

F_2 represents feature set function and is defined as:

$$F_2 = \left[1 - \frac{\sum_{k=1}^{N_F} f_k}{N_F} \right] \tag{23}$$

where f_k is the value of feature mask.

$$\text{with } f_k = \begin{cases} 1 & \textit{if feature is selected} \\ 0 & \textit{otherwise} \end{cases} \; . \tag{24}$$

Table 2. Confusion Matrix

	P	N
P	True positive (TP)	False Positive (FP)
N	False Negative (FN)	True Negative (TN)

For binary classifier, other performance measures are sensitivity and specificity. The sensitivity and specificity evaluate the discrimination capacity of classifier to classify between positive and negative class. The sensitivity and specificity are measured by true positive rate (TPR) and false positive rate (FPR), respectively. The TPR and FPR are computed as:

$$TPR = \frac{TP}{P} \; ; \; FPR = \frac{FP}{N} \tag{25}$$

where $P = TP + FN$ and $N = FP + TN$

Proposed algorithm for speech recognition

The main aim of this research work is to improve the recognition rate of TI-46 speech database and isolated Hindi words using SVM with RBF kernel. The feature set and SVM kernel parameters are randomly initialized as PPO particles as discussed in sub-section “initialization of prey particles” and “initialization of predator particle”. Prey particles are evaluated on the basis of fitness function as given in Equation (20) and local and global best solutions are updated. To further improve the quality of solution, local best solutions are optimized by Hooke-Jeeves method. The flow chart of proposed technique for speech recognition is presented in Figure 3.

EXPERIMENTAL DETAILS

In this section, the performance of the proposed technique with SVM has been evaluated on TI-46 database under clean and noisy conditions and self recorded Hindi digit database. The TI-46 database is standard isolated word speech corpus, having two subsets, TI-20 and TI-ALPHA. The TI-20 database consists of ten English digits and ten control words and the TI-ALPHA database consists of English alphabets. All samples have been recorded with a sampling frequency 12.5 kHz. The Hindi database consists of ten numerals recorded in a quiet room with sampling frequency 44.1 kHz. Table 3 shows the vocabulary used, number of instances and the number of male and female speakers in each database.

The experiments have been conducted by extracting MFCCs, delta and acceleration features for each utterance. The speech signal is given as input to a first-order finite impulse response high pass filter to spectrally flatten the signal. After this pre-emphasis, the silence region is removed from each utterance. The speech signal is divided into a fixed number of 40 frames each of 25ms with 50% superposition. After framing, the Hamming window is used for windowing as it introduces the least amount of distortion. MFCC feature vector that consists of 13 MFCCs and the corresponding 13 delta coefficients and 13 acceleration coefficients (total 39 coefficients) is extracted from each frame. The SVM technique with RBF kernel is implemented for classification. One-versus-all approach is used to construct the SVM classifier.

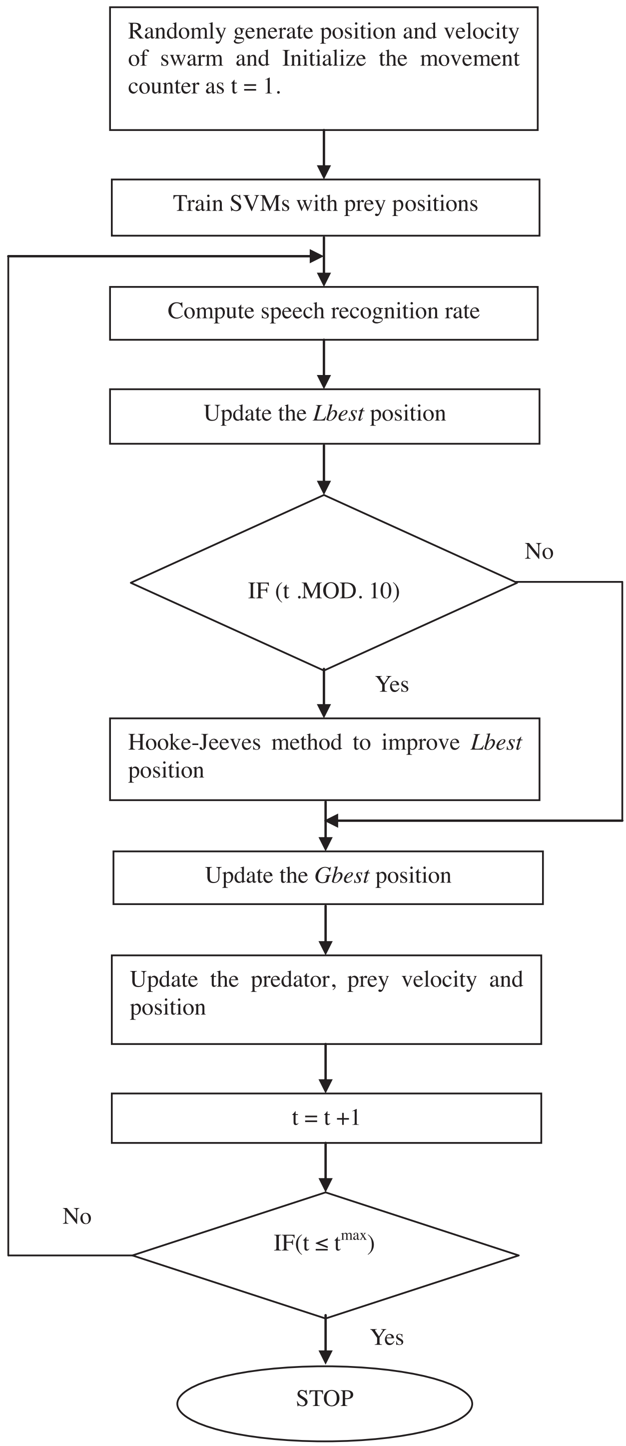

Fig. 3. Flowchart for proposed technique for speech recognition

Table 3. Databases

Database	Vocabulary	Number of instances	Number of speakers		Number of training samples	Number of testing samples
			Male	Female		
TI-20	Ten English digits and ten control words (NSIT speech disc 7-1.1, 1991)	26	8	8	3200	5120
TI-ALPHA	English alphabets “a” through “z”. (NSIT speech disc 7-1.1, 1991)	26	8	8	4160	6656
Hindi	Ten Hindi numerals, “shoonya”, “ek”, “do”, “teen”, “chaar”, “paanch”, “cheh”, “saat”, “aath”, and “nau”	20	2	2	400	400

RESULTS AND DISCUSSION

In order to test the performance of proposed technique, accuracy achieved by proposed technique with SVM model is compared with the results obtained by SVM model along with binary PSO, binary PSO and Hooke-Jeeves method, binary PPO and also with SVM model having default values of kernel parameters. The default value of kernel parameters are given in math library (MATLAB and Bioinformatics toolbox, 2009) and for RBF kernel, parameters are $C = 1, \sigma = 1$. The recognition rate (accuracy) obtained by these approaches is given in Table 4. The experimental results indicate that recognition rate achieved by proposed technique with SVM model is better than other techniques.

Table 4. Comparison of recognition rates (%)

Applied Technique / Dataset	Default value of SVM	Binary PSO with SVM	Binary PSO and Hooke – Jeeves with SVM	Binary PPO with SVM	Proposed Technique with SVM
TI-20	91.6	94.7	97.3	96.5	98.8
TI-ALPHA	86.2	90.1	94.0	91.7	96.6
Hindi digits	83.7	89.2	91.1	90.0	93.4

To investigate the effect of initial solution, 30 independent trials are given and the best, average and worst recognition rate is computed for Hindi digits database by applying various algorithm with SVM. The *standard deviation* (S.D.) is also computed to measure the variability of results. It is observed from Table 5 that the best, average and worst results of proposed technique with SVM are better than the results obtained by other techniques and also S.D. of results obtained by proposed technique is the least among its counterparts.

Table 5. The best, average and worst values of recognition rate (%)

Applied Technique	Best	Average	Worst	S.D.
Binary PSO with SVM	89.2	85.5	79.1	2.904
Binary PSO and Hooke – Jeeves with SVM	91.1	88.3	83.5	2.473
Binary PPO with SVM	90.0	86.4	82.5	2.841
Proposed Technique with SVM	93.4	92.0	88.5	1.912

To check the validity of the proposed technique, ROC curves have also been investigated. The ROC is a graphical plot between TPR and FPR that illustrates the performance of binary classifier. The TPR represents 'sensitivity' and FPR represents '1 - specificity'. Figure 4 shows the ROC curves plotted for Hindi digit database by applying various techniques. For Hindi digit database, 400 test samples and ten classes gives a total of 400 positive samples and 3600 negative samples. Based on these statistics, the ROC curves have been drawn by varying the threshold level for the techniques under consideration and the area under the curves (AUC) has also been computed and is given in Table 6. It can be observed that AUC of proposed technique with SVM is larger than the AUC for other techniques. So, the performance of proposed technique is better than the other considered techniques.

To evaluate the robustness of the proposed technique, experiments have been carried out in clean as well as noisy test samples. Noisy samples have been obtained by artificially adding white noise with different signal to noise ratio (SNR) (0, 5, 10, 15, 20, 30 and 40 dB) of TI-20 database. Figure 5 represents the recognition rate for different SNR achieved by various applied techniques with SVM model. It has been observed that recognition rate achieved by proposed technique with SVM model is higher than that of other techniques.

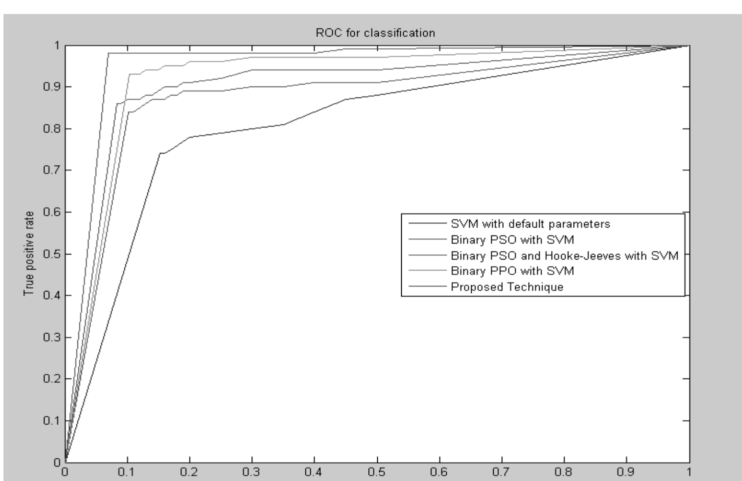

Fig. 4. ROC curves for different techniques for Hindi database

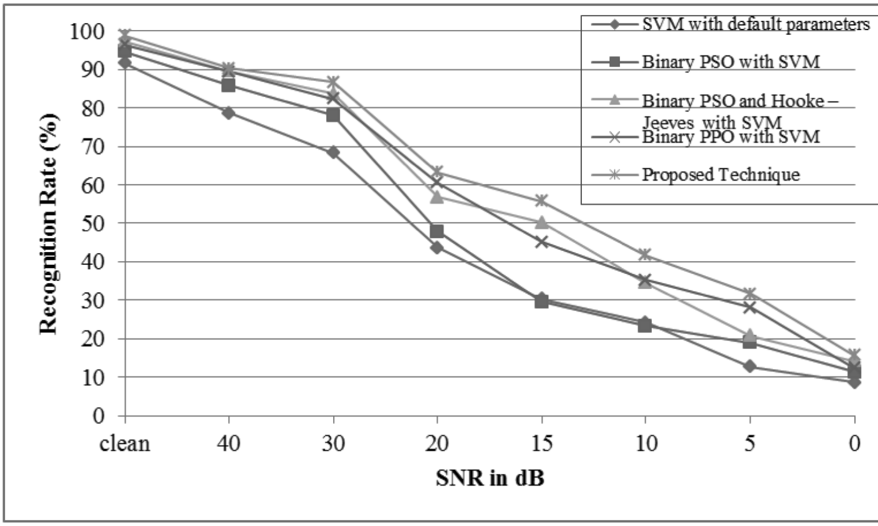

Fig. 5. Recognition rate obtained with various techniques under noisy conditions

Table 6. Area under the ROC curve for different techniques

Technique	Area under the curve
SVM with default parameters	0.8096
Binary PSO with SVM	0.8759
Binary PSO and Hooke-Jeeves with SVM	0.9045
Binary PPO with SVM	0.9166
Proposed Technique	0.9539

CONCLUSIONS

In this paper, a new hybrid optimization technique is proposed for searching the most relevant feature set of speech, and, optimal kernel parameters of SVM. This hybrid optimization technique is an integration of PPO technique and Hooke-Jeeves method. In this research work, binary PPO is proposed and successfully integrated with continuous PPO to deal with mixed type variables. The initial search is performed by PPO technique. The predator fear effect of PPO improves the search capability of algorithm. To further enhance the search capability, local search Hooke-Jeeves method is applied. Initially, the proposed technique with SVM model is tested on the TI-46 clean speech word database and then database with different range of SNR is also undertaken. The proposed technique has also been tested on Hindi speech database. The recognition rate achieved by proposed technique with SVM model is compared with the results obtained by SVM model along with binary PSO, binary PSO and Hooke-Jeeves method, binary PPO and with default values of SVM kernel. It has been seen that proposed technique with SVM model gives higher accuracy for all the databases used in this work. The statistical test, ROC has also been implemented and it has been observed that AUC for proposed technique is better than that of other techniques. The proposed technique can further be explored to solve other constraint optimization problems.

REFERENCES

Bao, Y., Hu, Z. & Xiong, T. 2013. A PSO and pattern search based memetic algorithm for SVMs parameters optimization. Neurocomputing 117 (Oct): 98-106.

Bao, Y., Xiong, T. & Hu, Z. 2014a. Multi-step-ahead time series prediction using multiple-output support vector regression. Neurocomputing 129 (April): 482-493.

Bao, Y., Xiong, T. & Hu, Z. 2014b. PSO-MISMO modeling strategy for multistep-ahead time series prediction. IEEE Transactions on Cybernetics 44 (5): 655-668.

Bradley, A.P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition 30 (7): 1145-1159.

Davis, S.V. & Mermelstein, P. 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Transactions on Acoustics Speech and Signal Processing 28 (4): 357-366.

Demšar, J. 2006. Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research 7(Jan): 1–30.

García, S. & Herrera, F. 2008. An extension on ‘‘statistical comparisons of classifiers over multiple data sets’’ for all pairwise comparisons. Journal of Machine Learning Research 9(Dec) 2677–2694.

Harman, M. & McMinn, P. 2010. A theoretical and empirical study of search-based testing: local, global, and hybrid search. IEEE Transactions of Software Engineering 36 (2): 226-247.

He, Q., Yan, J., Shen, Y., Bi, Y., Ye, G., Tian, F. & Wang, Z. 2012. Classification of electronic nose data in wound infection detection based on PSO-SVM combined with wavelet transform. Intelligent Automation and Soft Computing 18 (7): 967-979.

Hu, Z., Bao, Y. & Xiong, T. 2013. Electricity load forecasting using support vector regression with memetic algorithms. The Scientific World Journal 2013(292575).

Hu, Z., Bao, Y., Chiong, R. & Xiong, T. 2015a. Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection. Energy 84 (May): 419-431.

Hu, Z., Bao, Y., Xiong, T. & Chiong, R. 2015b. Hybrid filter-wrapper feature selection for short-term load forecasting. Engineering Applications of Artificial Intelligence 40 (April):17-27.

Hua-chao, Y., Shu-bi, Z., Ka-zhong, D. & Pei-Jun, D. 2007. Research into a feature selection method for hyperspectral imagery using PSO and SVM. Journal of China University Mining & Technology 17(4): 473-478.

Huang, C.L. 2009. ACO-based hybrid classification system with feature subset selection and model parameters optimization. Neurocomputing 73(1-3): 438-448.

Huang, C.L. & Dun, J.F. 2008. A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Applied Soft Computing 8(4):1381-1391.

Huang, C.L. & Wang, C.J. 2006. A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications 31(2): 231-240.

İlhan, İ. & Tezel, G. 2013. A genetic algorithm–support vector machine method with parameter optimization for selecting the tag SNPs. Journal of Biomedical Informatics 46 (2): 328-340.

Kennedy, J. & Eberhart, R.C. 1995. Particle swarm optimization. Proceedings of IEEE International Conference on Neural Network, Australia.

Kennedy, J. & Eberhart, R.C. 1997. A discrete binary version of the particle swarm algorithm. IEEE International Conference on Systems, Man, and Cybernetics, Washington, DC, USA.

Liu, Y., Wang, G., Chen, H., Dong, H., Zhu, X. & Wang, S. 2011. An improved particle swarm optimization for feature selection. Journal of Bionic Engineering 8 (2): 191-200.

Mandal, P., Haque, A.U., Meng, J., Srivastava, A.K. & Martinez, R. 2013. A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting. IEEE Transactions on Power Systems 28 (2): 1041-1051.

MATLAB version 7.8.0 and Bioinformatics Toolbox 3.3, **2009**. Natick, Massachusetts: The MathWorks Inc.

Narang, N., Dhillon, J.S. & Kothari, D.P. 2014. Scheduling short-term hydrothermal generation using predator prey optimization technique. Applied Soft Computing 21: 298-308.

NSIT speech disc 7-1.1, 1991. TI 46-Word speaker dependent isolated word corpus.

Omran, M.G.H., Engelbrecht, A.P. & Salman, A. 2009. Bare bones differential evolution. European Journal of Operational Research 196 (1):128-139.

Rabiner, L. 1978. Digital processing of speech signals. Prentice Hall.

Rabiner, L. & Juang, B. 1993. Fundamentals of speech recognition. Pearson Education.

Rao, S.S. 1996. Engineering optimization: theory and practice. John Wiley & Sons, New York.

Sarafrazi, S. & Nezamabadi-pour, H. 2013. Facing the classification of binary problems with GSA-SVM hybrid system. Mathematical and Computer Modelling 57 (1-2): 270-278.

Shih-Wei, L., Kuo-Ching, Y., Shih-Chieh, C. & Zne-Jung, L. 2008. Particle swarm optimization for parameter determination and feature selection of support vector machines. Expert Systems with Applications 35(4): 1817-1824.

Silva, A., Neves, A. & Costa, E. 2002. An empirical comparison of particle swarm and predator prey optimization. Proceedings of 13th Irish International Conference on Artificial Intelligence and Cognitive Science, Limerick, Ireland, Sep. 12-13, 24(64), 103-110.

Silva, A. & Gonçalves, T. 2013. Training support vector machines with an heterogeneous

particle swarm optimizer. Adaptive and Natural Computing Algorithms Lecture Notes in

Computer Science 78 (24): 100-109.

Silva, A. & Gonçalves T. 2014. Using a scouting predator-Prey optimizer to train support vector machines with non PSD Kernels. Nature Inspired Cooperative Strategies for Optimization Studies Computational Intelligence 512: 43-56.

Vapnik, V.N. 1995. The nature of statistical learning theory. Springer, New York.

Wei, J., Jian-qi, Z. & Xiang, Z. 2011. Face recognition method based on support vector machine and particle swarm optimization. Expert Systems with Applications 38(4): 4390-4393.

Xiang, L., Shang-dong, Y. & Jian-xun, Q. 2006. A new support vector machine optimized by improved particle swarm optimization and its application. Journal of Central South University of Technology 13(5): 567-571.

Xiong, T., Bao, Y. & Hu, Z. 2013. Beyond one-step-ahead forecasting: evaluation of alternative multistep-ahead forecasting models for crude oil price. Energy Economics 40 (November):405-415.

Xiong, T., Bao, Y. & Hu, Z. 2014. Does restraining end effect matter in EMD-based modeling framework for time series prediction? some experimental evidences. Neurocomputing. 123 (January):174-184.

Xiong, T., Bao, Y., Hu, Z. & Chiong, R. 2015. Forecasting interval time series using fully complex-valued RBF neural network with DPSO and PSO algorithm. Information Science 305 (June):77-92.

Yan-bin, L., Ning, Z. & Cun-bin, L. 2009. Support vector machine forecasting method improved by chaotic particle swarm optimization and its application. Journal of Central South University of Technology 16 (3): 478-481.

Zhao, M., Fu, C., Ji, L., Tang, K. & Zhou, M. 2011. Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. Expert Systems with Applications 38 (5): 5197-5204.

Submitted: 12/2/2015

Revised: 2/6/2015

Accepted: 7/6/2015