ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Fast Sound Verification Using Support Vector Machine and Particle Swarm Optimization Algorithms

[I]Thaer M. Taha, [II]Hazem M. El-Bakry, [III]Amal Ibrahim, [IV]Samir Abd-Elrazik, [V]Magdi Z. Rashad
[I,III,V]Dept. of Computer Science, [II,IV]Dept. of Information Systems
[I,II,III,IV]Faculty of Computer and Information Sciences, Mansoura University, Mansoura, EGYPT

Abstract
The sound verification has been defined as during the consideration of the talking term as an input sound and conformity it with the already in listed vote on based on the several criteria. In this paper, the proposed model will be explained. It's dependent on the combination between two algorithms (SVM-PSO), where particle swarm optimization (PSO) with support vector machine (SVM), the proposed approach uses the class label of the input sample as the actual output. The main goal of my idea is for reduction in time. To evaluate the effectiveness of the proposed approach, SVM based approach is compared with the combination of (SVM) and (PSO) classification algorithm. The experimental results show that, both the sound verification and training time of the suggested combination algorithms are better than that of the separately (SVM) algorithm. Furthermore, it has been demonstrated that these algorithm converges are faster than the traditional gradient regular. Moreover, a percentage improvement in the time and accuracy is shown compared to the regular algorithm.

Keyword
Support vector machine (SVM), particle swarm optimization (PSO), artificial neural network (ANN), Automatic speech recognition (ASR), sound verification, and machine learning (ML).

I. Introduction

The sound signal is the fastest and the most natural method of communication between humans. This fact has motivated researchers to think of sound as a fast and efficient method of interaction between human and machine [1]. Speech recognition features a nice potential, to become a very important issue of interaction between the human and computer within the close to future. The success of a sound verification system not solely should verify the characteristics exist in input mode during a purpose in time, however additionally has input mode changing over time [2]. Sound verification technologies Well-established profit several rehabilitation and medicine engineering industries within the development of health or help devices for paraplegics and therefore the disabled, wherever sound verification is employed as a patient-machine interface that Transfers
Management commands from patient to the machine
And Provides feedback from the machine to the patient [3].
Sound verification problem is a sub branch of pattern recognition. Some popular techniques to tackle this problem are:

- Artificial neural networks (ANN),
- Dynamic time warping (DTW), and
- Hidden Markov modeling (HMM) [4].

In general, some of the most common problems faced by the researcher in sound verification are:

- Contrast Speaker: in this case it is clear exactly the same word differently by different people due to the age or sex of the speaker or the anatomical differences, and the speed of sound, the emotional state of the speaker and the difference between dialects.
- Background noise: the close noise surroundings ideology will add noise to the signal. Even the speaker himself will add noise by the approach they speak.
- Above segmental aspects: the impact of intonation and place pressure on the sections. These aspects have an effect on one word pronunciation
- Continuing nature of the sound: When we speak, and rarely there is a break between words. Sound most likely be continuous without interruption of the vote. This makes it very difficult to detect individual words.
- Other external factors are: the position of the microphone in respect to the speaker, the direction of the microphone and many others, and sound. Overall, the process of recognition is split in many different components as shown in **Figure1**.

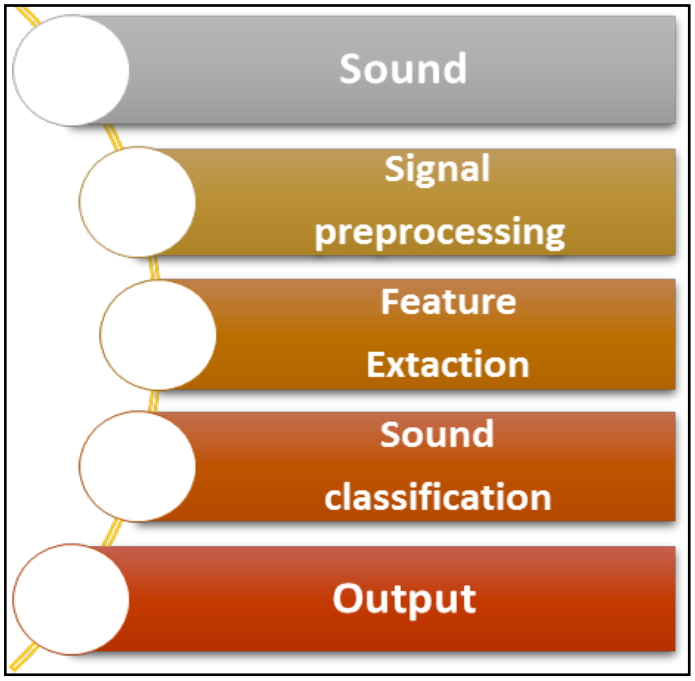

Fig 1: Sound recognition process. [5]

a) Stage one (sound), which consists of the acoustic setting additionally to the transduction instrumentation (microphone, electronic equipment and AD- converter) will have a really severe impact on the illustration of sound was born. As an example, you'll have an extra impact ensuing from the supplementary noise or echo chamber.
b) Stage two (signal preprocessing), is intended to deal with these problems, as well as the derivation of acoustic representation both good at separating layers of the speech sounds and effectiveness

in literacy and sources that are related to the difference.
c) Stage three (feature extraction), you must be able to extract certain features of the sound signal processed in advance. It can be done with a variety of techniques such as cestrum and spectral analysis.
d) Stage four (sound classification), is trying to box in the extraction and classification features respect to the audio input to appropriate the best voice in the 'vocabulary group' known and represents this as a way out [5]. Sound verification (SI):- Also known as automatic speech recognition (ASR) is to improve the process or the development of voice recognition or sound by the idea raised by the researcher through a certain adjustment that may occur either on a specific algorithm or change sports symbols or hybrid two algorithms to give best all of the above and more have in improving sound verification results in either shorten the time or improvement in resolution or improvement in performance, which may provide an accurate verification of authorized sound signals. Many applications are developed based on automatic speech recognition (ASR) such as text to sound system, public address system, mobile, and personal communication. In the recent work, the enhancement of sound verification has been studied using the technique of artificial neural networks. The preprocessing consists of the process of digitization of the analog signal. This process includes the sampling, quantization and then the digitalized of the analog signal that can be recognized from the computer. So, the sounds that are entered at the algorithm are analog and cannot be recognized. So, it is needed to do preprocessing the conversion of the signal to be recognized from the computer [6].

II. Related work

The Overall Performance average MLP 96%, RBF 98.69%. [2]. The performance is dropped to about 30% and 16% recognition accuracy when the SNRs decreased to 4 dB and −2 dB respectively. [3]. that the authors have reported recognition rates of 98% and 99.5%, [4]. Performs very well with MFCCs as input having more than 90% successful classification rate [5]. The overall speech recognition accuracy value obtained by this experiment is 97.14%. [6]. in these works, to unravel the optimum path to cut back time for sound verification looking out draw back, within the search area. It has been shown that mistreatment PSO with a technique causes a substantial reduction in sound verification time whereas maintaining the system accuracy. In distinction, mistreatment direct ANN while not clipping isn't a promising approach. PSO has been not accustomed to solve several drawback - complete optimization cause. Convergence proportion detour has developed from 88.5 to 95.06 %. [7]. Some of the verification of the automatic systems of speech (ASR) may not recognize the sound signals are accurate, they might face some difficulties and suffer from any of these challenges, such as sound signals to reduce noise lack of accuracy, and extract tankers appropriate feature, and choose the feature, and methods of recognition sounds [8]. Mostly choose the appropriate feature and proper recognition [9]. As a result, improved technology and particle swarm optimization are working to provide more accurate results in improved results or sound verification in any shorten the time or improvement in resolution or improvement in performance [10]. In the standard framework for enhancement of sound corrupted by additive noise, a digitalized noisy signal is considered to be a containing clean sound and noise. This signal is usually transformed into the time-frequency domain, where modeling of sound and noise is easier [11]. In this paper, the enhancing sound verification performance are being investigated useful techniques by using particle swarm optimization (PSO) algorithm context to provide additional information to the recognizer. Show that MMSE estimators employing only the super Gaussian noise prior (GS+SN) reduce the word error rates (WER) under the presence of noise a 10%-25%. [12]. This can be done through different ways. The directly recognition is a process of which is talking on the basis of features of sound signal of the speaker. Substantial increase in overall word accuracy simply from (59%) to (73%) [13]. The advantages of sound verification in include some of best of the existing features such as input sound are easy to do because it does not demand a specialized skill as does typing or push button procedures. When the user is not constant information can be input even or doing other action through the eyes, hands, Terminal ear, legs or since a mike or cell phone is being utilized input unit. The calculated accuracy and error rates according to formulas are from 82% to 90%. [14]. Most sound verification systems can be classified according to the following categories:
i. Speaker Dependent versus Speaker Independent
ii. Isolated versus Continuous
Key word based versus Sub word unit based [15]. Ask a targeted support device (SVM) in 1995. From a strict mathematical theory, and explained and implemented that small sample can increase the accuracy pattern classification methods and data installation. Successfully used in SVM sound verification, text recognition, and determine the error and other field research [16]. In this section, the proposed was described the PSO-SVMA system of classification. This study aims at improving the algorithm (PSO) through the use (SVM) of road classification seeded by detecting Subset of the best features of the information and the best estimate Regularization parameter values for the nucleus of SVM model [17]. The machine learning (ML) includes a modern techniques such as support vector machine (SVM), the structure of that depends on the applied mathematics learning, the problem of over-fitting could be resolved that, it has outstanding generalization capability within the situation of minor sample, native best resolution low convergence and magnitude relation endured in ANN. On the otherwise, the benefit of SVM is wedged in order to choose occasion SVM parameters. This study gives the best accuracy level that in 77%. [18]. The experiment result shows excellent trade with profit 139.5% while the stock gains about 80.6%. [21]. The best average recognition rate of 94.8% is achieved using the DCT (50x50) feature vector and the PSO-based feature selection algorithm. [22]. Voice recognition accuracy for each user should exceed 80% over 10 trials under normal conditions. [23]. Notice that the delay constraint is met while the area constraint becomes violated. The final area is 59% and the final delay is 21.99%. [24]. SVM is the best classification task for sound signals as it gives highest accuracy rate. The classification system of SVM is a supervised that minimizes an upper bound on its expected error. The algorithm SVM can be attributed to the inherent nature of best performance. The technique k-NN implementation least in comparison to the other ML algorithm. Complexity wise, k-NN takes extremely long time to classify audio signals whereas SVM takes only few seconds. The results of the experiments performed on sound signals classification task using Support Vector Machine (SVM) and k Nearest Neighbor (k-NN) classifiers. Comparative Accuracy (k-NN, 74.6%) and (SVM, 90%). Speed Performance in second (k-NN, 5) and (SVM, 2). [25]
When the speaker pauses shortly between each word that is called isolated sound, while when speech continuous the person talk

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

in uninterrupted words and likely long stream with very a few breaks between in or maybe not. Systems area unit isolated sound verification explicit to make. Continuous speech in words spoken on the opposite hand area unit subjected to the co-articulation impact, within which the pronunciation of a word is changed by the words encompassing it.

In this work, located the study between the voices of males and females, consisting of two groups, the first group which contains 40 voice for males, also the second group contain 40 voice for females, the test vote has been dependent on the vote free Site, database has been used called (urban sound) it consisting of 225 voted for males and 225 for females voice from [19].

III. PSO algorithm

The particle swarm optimization is comprised of a collection of particles that move around search space influenced by their own best past location of the whole swarm or a close neighbor each iteration a particles velocity is updated using:

Vi(t+1)=vi(t)+(c1*rand()*(pi best-pi(t)))+(c2*rand()*(Pgbest-pi(t)))

Where vi(t+1) is the new velocity for the i^{th} particle c_1 and c_2 are the weighting coefficients for the personal best and global best positions respectively, Pi(t) is the ith particles position at time t.Pibest is the ith particles best known position, and ***P***gbest is the best position known to the swarm. The rand () function

generates an uniformly random variable ∈ [0, 1] variants on this update equation consider best positions within a particles local neighborhood at time t.

A particles position is updated using:

$P_i(t+1) = P_i(t) + v_i(t)$

Algorithm (below) provides a pseudocode listing of the particle swarm optimization algorithm for minimizing a cost function. [20], [21], [22], and [24].

IV. The proposed algorithm

In the proposed method, the goal is to reduction in time it takes to identify the speaker of samples previously recorded wave. A primary focus on accuracy and speed. The proposed method is implemented using MATLAB.

Let N be the particle number of particle swarm.

A. Initialization

1. Initialize the population of N particles with random positions and velocities on D dimensions in the solution space.
2. Set the velocity vectors v_i (i=1,2,…N) to zero.
3. For each position $p_i \in R^{d+2}$ of the particle P_i (i=1,2,…N) from the swarm, train the SVM classifier and compute the fitness function value.

B. Particle swarm search

1. Detect the best global position $\boldsymbol{P_i}$ in the swarm which shows the minimal value of the fitness function value over all explored trajectories.
2. Update the speed and position of each particle.
3. For each candidate particle $\boldsymbol{P_i}$, train the SVM classifier and compute the fitness function f(i).
4. Update the best position of each particle if its current position has a smaller fitness function.

C. Convergence

1. If the maximum iterations times have reached, then exit, else return to 2.1

D. SVM-PSO Training and classification

1. Select the best global position of the particle swarm and train the SVM with the detected feature subset modeled with the optimized parameters C and σ.
2. Make SVM classification based on the trained classifier.[26], [27].

E. Combination between SVM-PSO algorithms

1. prepocessing, it working on convert the sounds from analog signals to digital signals.
2. Feature extraction, by using (Discrete Cosine Transform) technique is extrcted the feature through (fast fourier transform).
3. In k-fold cross-validation, the training data is randomly split into k mutually exclusive subsets (the folds) of approximately equal size.
4. The proposed work will be clarified in this paper, the new model depends on the combination between SVM-PSO algorithms for classification, this discuss originally explored is intended at improving sound verification through reduction in time by using SVM-PSO instead of SVM classifier by discovering the part of best powerful features and choose the shorter time, The PSO algorithm depends enhanced structure in experience. The proposed technique of SVM-PSO its part of the machine learning and contain two techniques through using PSO for improve the parameters of SVM. The begins work PSO with queries for the optimal particle iterative and n-randomly chosen particles. symbolizes an applicant solution for every particle is a dimensional vector. Each selection solution to assess its efficiency through the cross validation technique is called built SVM classifier. PSO algorithm controls the selection of possible subsets that lead to best forecast accuracy. The uses of the maximum suitable particle in algorithm is to lead to the setting up of n-candidate particle. so, on the model, all following individual of chosen particles fits best than precedent. This process involves on until the efficiency of SVM convergence. PSO is used to detection maximum feature partial by searching for the better feature mixtures as they fly within the matter area from the stomach database. An action clarifies SVM classification is as follows during in figure **2 (a)**, and combination proposed SVM-PSO is as follows during **figure 2 (b)**.

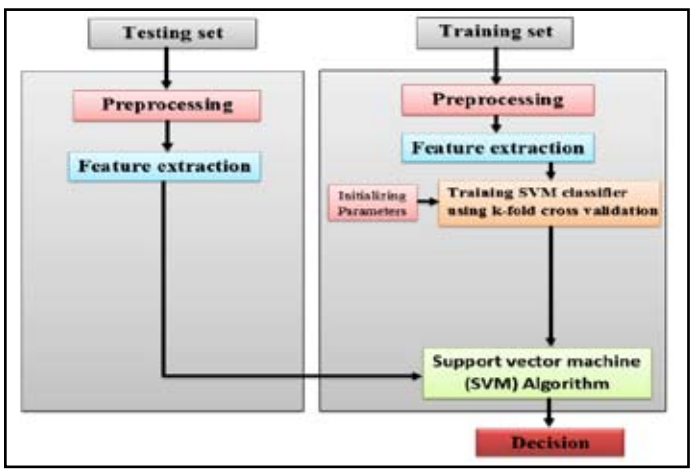

Fig. 2 (a): SVM for sound classification.

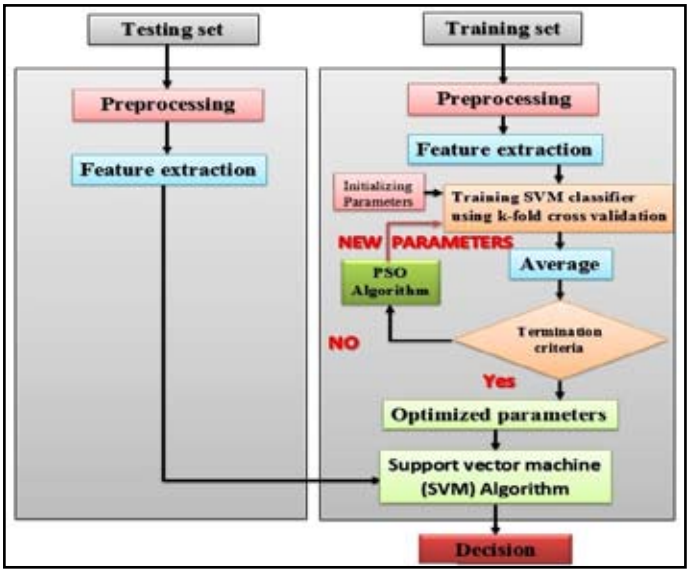

Fig. 2 (b): combination between PSO and SVM for sound classification.

5. Then, the fitness functions and the local best the fitness of each particle will be evaluated and global best parameters will be compared. Once completed, the value of the fitness task converges depends on the velocity and position of each particle will be updated. After converges, the global best particle in the swarm is fed to SVM classifier for training. Finally, The sounds were being tested on SVM separately and were being trained on SVM-PSO.
6. The results of tested sounds in proposed model guide that using combination between SVM-PSO algorithms give us the fast sound verification, reduction in time, and sound classification, best than the use of separately SVM algorithm.

In order to validate the algorithm, which recently worked to edit them and to ensure the validity of the final results, the small sample shows the best results that seem to consist of 18 voice of the total votes which, the practical side application has been worked on both composed of 40 voice test for men and 40 voice test for women before and after optimization, a simple simulation has been implemented using MATLAB software.

Here, some examples are presented about the best practical application of the results that explain their work and have a group of voices as a sound performance for the algorithm performance evaluation and compared with some traditional optimization algorithms swarm of particles, and then, offer the first metrics used sound performance at that time, provide examples below. In the chart below, the reader can see clearly the time difference between the standard algorithm and combination algorithms through shortening the time to recognize the many different voices as found in the following example in **figure 3**.

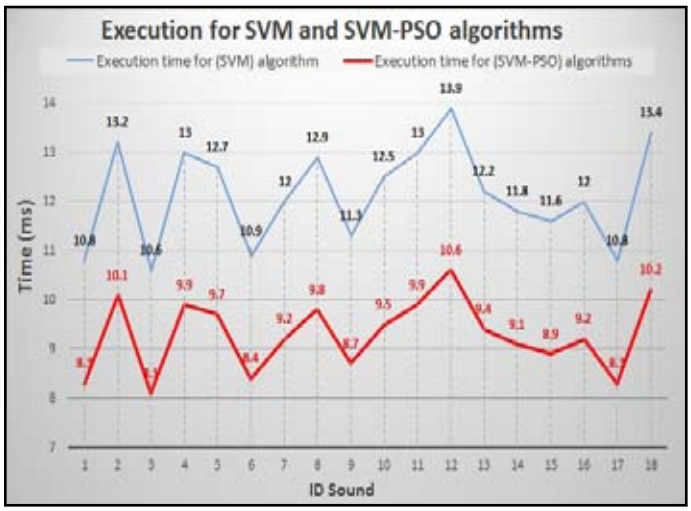

Fig. 3: The results for testing SVM separately and (SVM-PSO) algorithms.

V. Experiment Results:

To investigate the performance of the proposed approach, practical experimental has been conducted on real voices. To illustrate the effectiveness of the improved algorithm, the consumed time before and after optimization is estimated. **Tables 1**and **2** show the speakers sounds.

Table I. Result for testing males sounds (SVM) and (PSO) algorithms (ms)

Classifier	Tested sounds	Time of test (ms)	Total reduction time (ms)
SVM	40 voice	393.2	91.8
SVM - PSO	40 voice	301.4	

Table II. Result for testing females sounds (SVM) and (PSO) algorithms (ms)

Classifier	Tested sounds	Time of test (ms)	Total reduction time (ms)
SVM	40 voice	394.9	92.5
SVM - PSO	40 voice	302.4	

The first table contains voices of males while the second one contains female's voices. In addition, the recognition time in seconds is given. Furthermore, the amount of shortening time for these voices is listed. In the next figures, the table's results will be shown clearly through of statistical data for both tables by **figure 4** for males sound and **figure 5** for female's sounds.

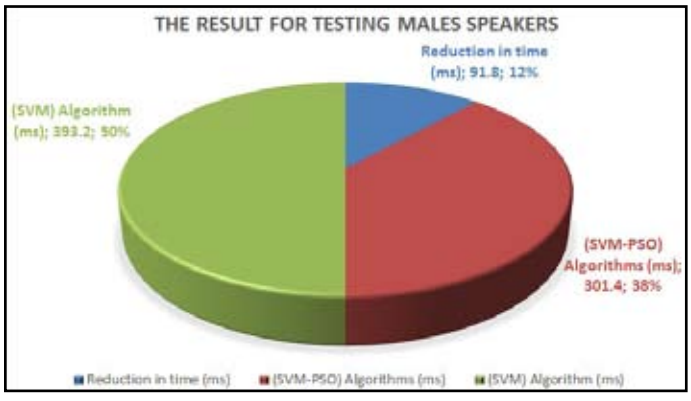

Fig 4: description test voice for male's speakers.

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

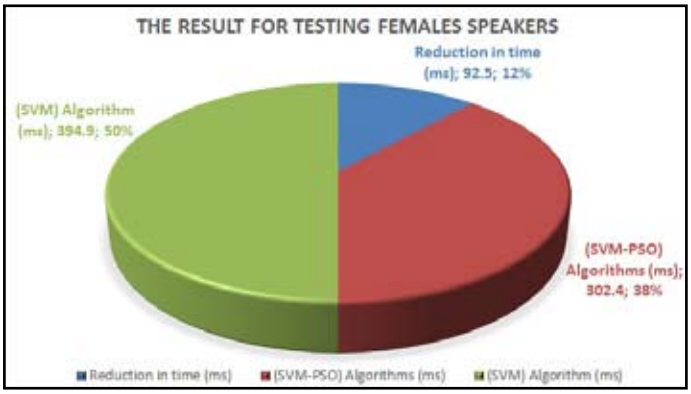

Fig. 5: description test voice for female's speakers.

Calculate the rate has been revealed to how that proposed work improves the sound verification, consider: [6].

Reduce in time = (reduction time / original time)*100=
= (184.3/788.1)*100=23.5%.

For further clarification, the best results have been incorporated that appeared to have consisting of 18 votes, 9 for males and 9 for females, from the practical application of an excel program and sample results. Samples are divided into two groups and each group has 9 votes so that the reader has a clear vision on the chart and the difference between every sound before the improvement after amendment the algorithm and using the same sound as they showed us here in this flow chart in the first Experiment one **(6: a)** for males, and Experiment two **(6: b)** for females.

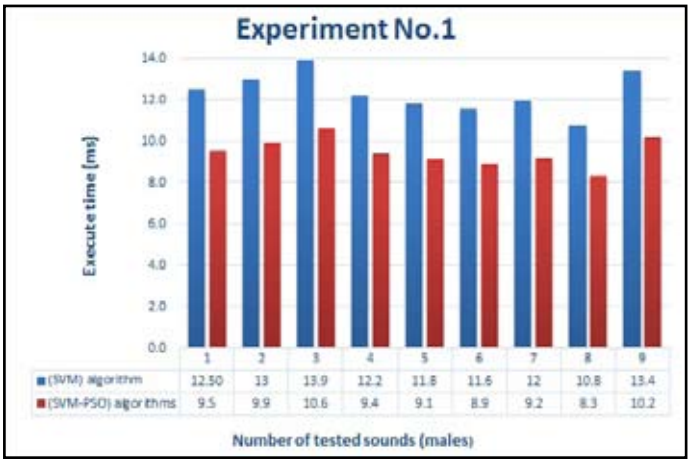

Fig. 6(a): Experimental results for sound recognition of the first group.

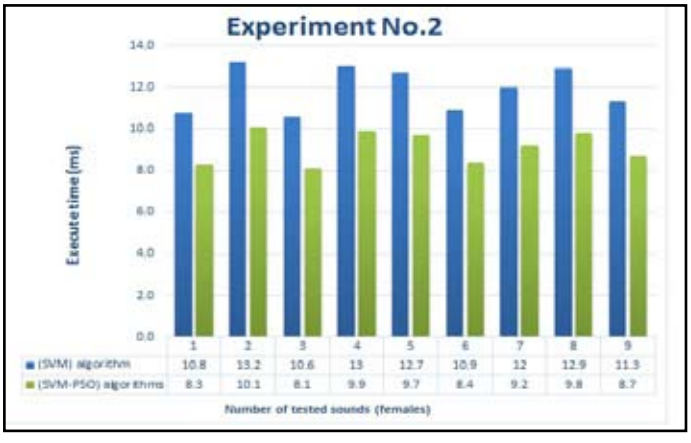

Fig. 6(b): Experimental results for sound recognition of the second group.

VI. Conclusion

This paper has been developed through insert some primary for combination between two algorithms, the first one Support Vector Machine and the second Particle swarm optimization (SVM-PSO). [18]. This strategic model for fast sound verification has been presented. By using this speaker verification technology many uses have been achieved. This technology helps physically challenged skilled persons Sound verification, most effective, reliable and customary medium to speak in real time systems. This has been done by combining (SVM) and (PSO) algorithms for classification rule to boost the performance based mostly sound verification system. The square measure numerous applications of sound is still to be far away from reality simply because of the lack of economical and reliable noise removal mechanism and techniques for conserving or up the comprehensibility for the sound signals. During this paper a try has been stepped towards mensuration the methodologies for soft computing based mostly sound verification techniques for speech sweetening in multimedia system applications. Through this review it's found that improvement technique used wide for performance improvement of sound verification in multimedia system applications. What is more, the given approach may be utilized in multimedia system applications. Primarily, we have a tendency to conjointly mention the normally used enhance performance throughout manual time at a rate of 23.5%.

References

[1]. *M. El-Ayadi, M. S. Kamel, F. Kaaray,"Survey on speech emotion recognition: Features, classification schemes and databases, Pattern Recognition, 44 (2011), pp. 572–587.*

[2]. *uojiang,"A Novel Isolated Speech Recognition Method based on Neural Network, 17 (2011) 264-269.*

[3]. *Y. Chan, P. C. Yong, S. Nordholm, C. K. F. Yiu, H. K. Lam, "A hybrid noise suppression filter for accuracy enhancement of commercial speech recognizers in varying noisy conditions, Applied Soft Computing, 14 (2014) 132-139.*

[4]. *G. Dede, M.H. Sazlı,"Speech recognition with artificial neural networks, Digital Signal Processing, (2010) 763–768.*

[5]. *G. T. Tsenov, V. M. Mladenov,"Speech Recognition Using Neural Networks", Senior Member, IEEE, 10 (2010) 181-186.*

[6]. *Selvaraj, and B. Ganesan,"Enhancing Speech Recognition Using Improved Particle Swarm Optimization Based Hidden Markov Model", Hindawi Publishing Corporation, 270576 (2014) 1-10.*

[7]. *N. Najkar, F. Razzazi, and H. Sameti,"An evolutionary decoding method for HMM-based continuous speech recognition systems using particle swarm optimization", Springer-Verlag London, 17 (2014) 327–339.*

[8]. *B. Singh, N. Kapur, and P. Kaur, "Speech recognition with hidden Markov model: a review," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 3, (2012), pp. 400–403.*

[9]. *Gupta and D. S, Wadhwa,"Speech feature extraction and recognition using genetic algorithm," International Journal of Emerging Technology and Advanced Engineering, vol. 4, no. 1, (2014), pp.363–369.*

[10]. *H. Hu and S. Zahorian, "An experimental comparison of fundamental frequency tracking algorithms," The Journal of Acoustical Society of America, vol. 132, (2012) pp. 2092–2098.*

[11]. J. S. Casals and V. Zaiats, "Speech Enhancement for Automatic Speech Recognition Using Complex Gaussian Mixture Priors for Noise and Speech" Springer-Verlag Berlin Heidelberg vol. 5933 (2010) pp. 60–67.
[12]. J. Cai, R. C. Miller, and S. Seneff,"Enhancing Speech Recognition in Fast-Paced Educational Games using Contextual Cues" , MIT Computer Science and Artificial Intelligence Laboratory 32 Vassar Street, Cambridge, Massachusetts 02139, USA, (2011), pp. 1-6.
[13]. S. Kaur, E. G. Kaur, "Enhancement of Speech Recognition Algorithm Using DCT and Inverse Wave Transformation" Journal of Engineering Research and Applications, Vol. 3, Issue 6, Nov-Dec (2013), pp.749-754.
[14]. S. K. Sarangi, and G. Saha, "A Novel Approach in Feature Level for Robust Text-Independent Speaker Identification system higher", IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction, December,27-29.(2012) p.p.1-5.
[15]. S. K. Gaikwad, B. W. Gawali, and P. Yannawar, "A Review on Speech Recognition Technique", International Journal of Computer Applications (0975 – 8887), Vol.10– No.3, November (2010).
[16]. C. yongqi, "LS_SVM Parameters Selection Based on Hybrid Complex Particle Swarm Optimization", Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Hainan University, vol.17,(2012), pp. 706 – 710.
[17]. F. Ardjani, K. Sadouni, M. Benyettou, "Optimization of SVM MultiClass by Particle Swarm (PSO-SVM)", IEEE. Vol, 978-1-4244-6977-2/10/ (2010), pp.1-4.
[18]. A. S. H. Basaria, B. Hussina, I. G. P. Anantaa, J. Zeniarjab , "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization ", Elsevier Ltd, Procedia Engineering vol,53 (2013).pp.453 – 462.
[19]. http://www.freesfx.co.uk.
[20]. http://www.cleveralgorithms.com/natureinspired/swarm/pso.html.
[21]. Q. Wen, Z. Yang, Y. Song, P. Jia,"Automatic stock decision support system based on box theory and SVM algorithm", Expert Systems with Applications 37 (2010) p.p1015–1022.
[22]. R. M. Ramadan, R. F. Abdel – Kader," Face Recognition Using Particle Swarm Optimization-Based Selected Features", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 2, No. 2, June (2009), p.p51-66
[23]. A. Upadhyay, "Enhancement of Speech Recognition Using Neural Networks Using Clustering of Vocabulary", Journal of Data & Network Security www.ijdnsonline.com Volume 2 No.1, Feb10, (2013), p.p121-125.
[24]. M. B. Abdel Halim, and S. E. –D. Habib, "Particle Swarm Optimization for HW/SW Partitioning", Aleksandar Lazinica (Ed.), www.intechopen.com, ISBN: 7619-48-0, (2009), p.p. 978-953. P. Mahan, G. Singh, "Comparative
[25]. Analysis of Machine Learning Algorithms for Audio Signals Classification", IJCSNS International Journal of Computer Science and Network Security, VOL.15 No.6, June (2015), pp.49.
[26]. R. Wang, Z. Hu, L. Chen, J. Xiong, "An Approach on Feature Selection of Cascaded Support Vector Machines with Particle Swarm Optimization Algorithm", Springer-Verlag Berlin Heidelberg, LNEE 140, Vol.2, (2012) .pp. 689–700.
[27]. N. Najkar, F. Razzazi, and H. Sameti, "A novel approach to hmm based speech recognition systems using particle swarm optimization. Math Compute Model 52(11-12): (2010), p.p1910–1920.