

Purging of silence for robust speaker identification in colossal database

P. Rama Koteswara Rao¹, Sunitha Ravi², Thotakura Haritha³

^{1,2}Department of ECE, NRI Institute of Technology, Pothavarappadu, Krishna Dt, India

³Department of ECE, P.V.P Siddhartha Institute of Technology, Vijayawada, A.P, India

Article Info

Article history:

Received Aug 18, 2020

Revised Dec 4, 2020

Accepted Jan 19, 2021

Keywords:

Dimensionality reduction

Large scale SVM

Multi-linear principle factor analysis

Un-pitched

Voiced

ABSTRACT

The aim of this work is to develop an effective speaker recognition system under noisy environments for large data sets. The important phases involved in typical identification systems are feature extraction, training and testing. During the feature extraction phase, the speaker-specific information is processed based on the characteristics of the voice signal. Effective methods have been proposed for the silence removal in order to achieve accurate recognition under noisy environments in this work. Pitch and Pitch-strength parameters are extracted as distinct features from the input speech spectrum. Multi-linear principle component analysis (MPCA) is utilized to minimize the complexity of the parameter matrix. Silence removal using zero crossing rate (ZCR) and endpoint detection algorithm (EDA) methods are applied on the source utterance during the feature extraction phase. These features are useful in later classification phase, where the identification is made on the basis of support vector machine (SVM) algorithms. Forward looking schostic (FOLOS) is the efficient large-scale SVM algorithm that has been employed for the effective classification among speakers. The evaluation findings indicate that the methods suggested increase the performance for large amounts of data in noise ecosystems.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

P. Rama Koteswara Rao
Department of Electronics and Communication Engineering
NRI Institute of Technology
Krishna Dt, Pothavarappadu, Andra Pradesh, India
Email: prkr74@gmail.com

1. INTRODUCTION

Ever-increasing data base size in real-world speaker recognition systems pose challenges such as large training time, poor response time and vast memory requirements [1-3]. Robustness and adaptability are the major aspects in real-world speaker recognition systems. From the previous work, it is observed that good results were achieved for the clean high-quality speech under matched conditions. However, under noisy environments and mismatched conditions, the performance of recognition system degrades significantly, much further by being acceptable. Therefore, sophistication is also a critical analysis in the identification of speakers. This motivates us to investigate new methods in different stages during the speaker recognition process. A typical speaker recognition system consists of mainly two stages: enrollment phase and classification phase [4, 5]. During the enrolment stage, speaker-specific information is extracted from the speech database in chronological mode. A cluster of such models tends to establish the speaker data base. An input speaker model is compared with the existing models in the database and then the results are expedited in the classification stage. In addition, features are extracted from input speech and transferred into a compact

representation which is relatively more discriminative and stable than the original signal [5-7]. This work proposes the experimental investigations for the various phases; feature extraction, noise elimination, dimensionality reduction and classification.

2. THE PROPOSED METHOD

In this research work, silence removal algorithms are employed to eliminate the back ground noise from the source utterance for the robust speaker recognition. The FOLOS SVM algorithm along with the MPCA dimensionality reduction technique is adopted for the classification of voices. Hence, the MPCA-FOLOS combination is employed in this work to improve the recognition accuracy under noisy environments for the large data sets as illustrated in Figure 1.

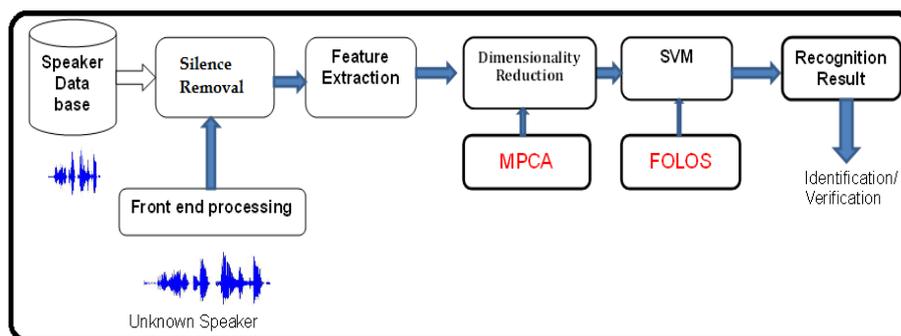


Figure 1. The speaker recognition process comprises of feature extraction, noise elimination, dimensionality reduction and voice classification

2.1. Feature extraction

The features extraction module collects a group of parameters which indicate the speaker-explicit data from the input signal. This data is the consequence of complex data processes at various phases of speech signal processing. Characteristics of speech signals are mainly differentiated by learning habits, vocal tract and vocal excitation of speakers. It is observed that the physiological based vocal tract features are more robust and less prone to mimicry [8-10]. Therefore, vocal tract related parameters like “mel frequency cepstral coefficients (MFCC)”, “linear predictive cepstral coding (LPCC)” and the other non-conventional features exhibit good results under clean and matched conditions [11]. These methods are developed by using linear predictive (LP) residual signals. However, the performance of these systems severely degrades under mismatched and noisy environments [12]. Therefore, it is required to derive a set of new features from the excitation source which are less prone to environmental noise. Excitation source characteristics such as pitch and pitch strength exhibit the physiological and behavioral aspects of the speaker. This work mainly focuses on feature extraction by accumulating characteristics from pitch and pitch strength parameters.

2.2. Noise removal

Voice activity detection algorithms are utilized to eliminate the environmental noise from the source articulation for the quality speaker identification. The noise removal comprises two main techniques: zero crossing rate and energy based speech utterance detector. In order to separate the pitched part of the speech signals, statistical nature of back ground is more significant.

In the speaker identification process, elimination of silent part is the crucial for robustness. Therefore, the preprocessing step involves the isolating of repeated information is important. The speech segment mainly comprises of unvoiced (U), silence (S) and voiced (V) events [13-15]. There is no voice in a silent zone, where sound waves don't even pulsate as in contingencies regions. This work utilizes the well-known silence removal methods.

2.3. Dimensionality reduction

This procedure minimizes dimensions of the feature matrix while holding the speaker discriminative data for the identification procedure. The size of information that is gathered from the input speech segments is very enormous since the fundamental qualities change gradually. Further, the identification procedure requires moderately little information to specify the attributes of speech.

In speaker recognition process, a large number of feature set demands vast memory and high processing time. Hence, it is necessary to minimize the complexity of the parameter matrices, where a high-dimensional space is transferred into a space of fewer dimensions. Apart from the benefit in computation, improvement in accuracy can be achieved under noisy environments [16]. The much more effective and being structured analysis used for the minimization for the experimental work is “multilinear principal component analysis” (MPCA) [17]. We have conducted several investigations and performed the comparative analysis with respect to population size. Since these methods are successful in large-scale recognition tasks, original contributions are carried out in this work with respect to large-scale speaker recognition.

2.4. Classification

Signal classification is a dynamic procedure for approving a given unknown individual in speaker identification process. The classification is commonly made by comparing with the existing data base. This process is generally partitioned into two sections namely: train and validate [18-20]. In training step, the artificial neural network is trained with the speaker-specific information. The validation is a procedure of calculations of coordinated score which is a proportion of closeness among applied and existing speaker models.

Verification is tested by comparing the estimated features to the speaker models. Support vector machine (SVM), is a successful discriminator and very effective method in recognition systems [21]. Specifically, SVM is operated in statistical learning procedures in order to attribute the risk minimization function. But, the main constraints of the SVM are computational complexity and poor performance at large data set conditions. Therefore, investigations have been carried out for the optimization of large-scale SVM Algorithms. An efficient SVM algorithm namely FOLOS has been derived and tested [22, 23].

3. SILENCE REMOVAL

In speech based applications, for removal of silence, two main techniques are widely adopted: zero crossing rate and energy based speech utterance detector. In order to separate the pitched part of the speech signals, statistical nature of back ground is more significant. The following section describes the important silence removal procedures.

3.1. STE algorithm

The short time energy (STE) algorithm comprises of three steps for the elimination or silence part and speech detection from the recorded speech utterance.

3.1.1. Pre-process

Extracted input signals are partitioned into frame blocks of size 18 ms by Hanning window technique.

3.1.2. Estimation of segments

The signal peaks of every block is determined analytically as,

$$P(i) = \sum_{n=1}^N x_i(n)^2 \quad (1)$$

Where, $i = 0, 1, 2, 3, \dots, N$ in which the sample count in every segment is denoted by the parameter N . The magnitude of n^{th} spike in i^{th} segment is represented by the function $x_i(n)$. Energy thresholds T_1 & T_2 are derived from the frames of signal. These levels are calculated by the following typical parameters.

$$\text{Enrg_Maxi} = \text{Maxi} (P(k)) \quad (2)$$

$$\text{Enrg_Mini} = \text{Mini} (P(k)) \quad (3)$$

$$T_1 = \text{Enrg_Mini} [5 + 2 \text{Log}_{10} (\text{Enrg_Maxi}/\text{Enrg_Mini})] \quad (4)$$

$$SL = \frac{\sum_k P(k)}{\sum_k 1} \quad (5)$$

Here k is perhaps the element of the segments with $P > T_1$

$$T_2 = T_1 + 0.25 (SL - T_1) \quad (6)$$

From the above equations, the maximum energy voice spike $Enrg_Maxi$, $Enrg_Mini$ specifies the least energy level and the mean energy level represents SL for above T_1 .

3.1.3. Silence elimination

The frames which haven't been recognized as "voice" by the above steps are marked as 'silence' and would be removed in signal utterance.

3.2. Zero crossing rate (ZCR) for Silence elimination

It is very difficult to record perfect clean speech. Therefore, it is obvious that some level of noise interferences the clean speech. As a result, there is a larger zero-crossing in silence speech patterns. The following three steps are employed for the silence elimination by zero-crossing segments.

3.2.1. Basic segmentation

The speech spectrum is divided to 10ms segments with the Hanning window

3.2.2. Zero cross rate estimation

In each frame, threshold point for the estimation is calculated as,

$$Z_{in} = \sum_{n=1}^{N-1} |\text{sgn}[x_m(n+1)] - \text{sgn}[x_m(n)]| \quad (7)$$

Where, m stands for the segment number, $x_m(n)$ represents an n^{th} frame in m^{th} segmen.

$$\text{sgn}(x_m(n)) = \begin{cases} 1 & \text{if } x_m(n) \geq 0 \\ 0 & \text{elsewhere} \end{cases} \quad (8)$$

3.2.3. Silence elimination

Unvoiced parts of speech are eliminated by taking higher rates of zero crossings in comparison with a threshold.

3.3. Detection and silence removal algorithm

The 'mean' μ calculation and the 'standard deviation' σ for 3100 segments of the voice signal that represents the background noise.

$$\mu = \frac{1}{3200} \sum_{i=1}^{3100} x(i) \quad (9)$$

$$\sigma = \sqrt{\frac{1}{2300} \sum_{i=1}^{3200} (x(i) - \mu)^2} \quad (10)$$

The deviation is calculated in the stored signal and categorized as voiced sample if $\frac{|x-\mu|}{\sigma} \leq 3$.

The speech signal is denoted as one and non-voiced segment as zero. The whole data is sampled in to other segments, each 10 ms long. Then the entire speech is transformed as series of 0 and 1 pattern. This first 100ms to 200ms of any speech signal represents back ground noise or silence portions. The back ground noise and silence portions are defined as white gaussian noise with the following frames.

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad (11)$$

First 3100 frames of the signals are used to calculate the parameters μ and σ as these samples are known as white noise. The possibility of data samples x satisfies,

$$P(|x - \mu| \leq \sigma) \text{ is } 0.68 \quad (12)$$

Hence, any sample x satisfies the condition $\frac{|x-\mu|}{\sigma} \leq 3$, then, it relates to background noise and the corresponding part can be eliminated.

4. DIMENSIONALITY REDUCTION

The dimensionality reduction technique MPCA can perform reduction by taking all tensor modes. Specifically, these projected tensors capture the variation present in the given feature set. More precisely, a selected set of features that varies in characteristics is identified [24]. It is obtained by selecting N_{sel} elements from each column of the feature matrix, where, N_{sel} to be a perfect square number. Therefore, N_T column vectors are selected of size $N_{sel} \times 1$. Each column vector is represented as $C_a^{(b)}$, where $0 \leq a \leq N_{sel} - 1$ and the respective matrix M is defined as,

$$M_{xy}^{(b)} = C_z^{(b)}; 0 \leq x \leq \sqrt{N_{sel}} - 1, 0 \leq y \leq \sqrt{N_{sel}} - 1 \quad (13)$$

$$z = x\sqrt{N_{sel}} + Y \quad (14)$$

From the matrix modified above, each b^{th} column vector is suggested by dimensionality reduction using MPCA. Y is an arbitrary constant. In the above process specified above, the distance matrix D is derived for each b^{th} matrix as,

$$D^{(b)} = M^{(b)} - \mu \quad (15)$$

$$\mu_{xy} = \frac{1}{N_T} \sum_{b=0}^{N_T-1} M_{xy}^{(b)} \quad (16)$$

In (16), μ represents the mean matrix, derived for $M^{(b)}$ and it is used for determining the distance matrix. To obtain the projection of matrix Ψ , tensors $T_1^{(b)}$ and $T_2^{(b)}$ are applied to distance matrix $D^{(b)}$:

$$\Psi = \sum_{b=0}^{N_T-1} T^{(b)} (T^{(b)})^T \quad (17)$$

The specific form for calculation is denoted in (17), and it is required to determine the tensors $T_1^{(b)}$ and $T_2^{(b)}$. This projection matrix is subjected to the generalized eigenvector problem.

5. RESULTS AND DISCUSSION

5.1. Experimentation setup

5.1.1. Speaker database

The proposed speaker recognition techniques are implemented in the signal processing research tool, MATLAB. These techniques are tested by using the speech database taken from Neurosciences Institute. In experimental setup, all 640 speakers (446 males and 194 females) are used for the training and testing [25]. During the second stage, a set consists of 240 speakers (males-142 and females-98), were chosen arbitrarily in datasets. In learning stage, every speaker having clean speech is trained where as testing is done separately on speech data from NIST database with channel noise and white noise [26]. The experiments are conducted by preparing database by adding white noise to the clean speech of database at different SNRs. This is dynamically generated by MATLAB toolbox.

5.1.2. Feature extraction

In the feature extraction stage, twenty five speech samples are extracted from every user in to prepare speaker-specific features. Twenty speech samples are employed for the training process, and the other samples are used for the testing of recognition. The speech signals are normally sampled at 16 KHz and establishes as windows of size 25ms. Each frame comprises of FFT-based 256 power spectrum vectors to develop feature vectors. Next stage, these feature vectors are applied for dimensionality reduction to develop an optimized feature set.

5.2. Results

The performances of the proposed technique are evaluated on the sentences that are taken from NIST database. The robustness is measured by using noisy database, intentionally made by inserting white gaussian noise to the original database at different levels like signal to noise ratios 10, 20 and 30 in dBs. That blue areas in the sound performance signal reflect the parts of the voice whereas green areas reflect the neglected silence zones as seen in Figures 2-4. From the experimental results, it is identified that enhancement in SNRs improves that performance significantly. The improvement of our technique is similar

to that of sentences have been used from clean noise free data base. But, the higher SNRs improve the effectiveness. At the SNR of 10dB, the standard techniques eliminate major portion of the voice signal as silent, but the proposed method removes only the non-voiced parts.

The results obtained from the experiments that are conducted after silence removal step show that the new feature vectors exhibit reasonable improvement over original values from Tables 1 and 2. Specifically, with the inclusion of background noise at 30 dB SNR, the accuracy is improved by 2.68% as compared to the silence removal step by using MPCA method. Identically, the noisy conditions reduced upto 1.82% for the “channel noise” at the signal to noise ratio of 30 decibels. From the improved results, it is found that the existence of back ground noisy conditions, MPCA generates more effective results. The evaluation of the silence removal is summarized in Tables 1 and 2.

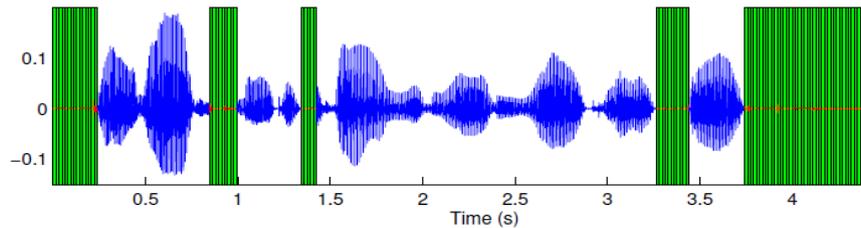


Figure 2. Silence removal by STE algorithm with white gaussian noise at 30 decibels

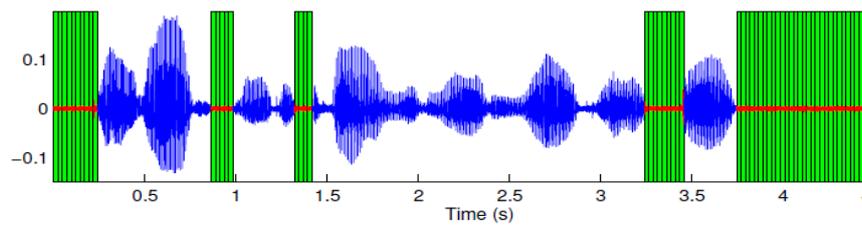


Figure 3. Silence removal by zero crossing rate (ZCR) algorithm at 30 decibels for white gaussian noise

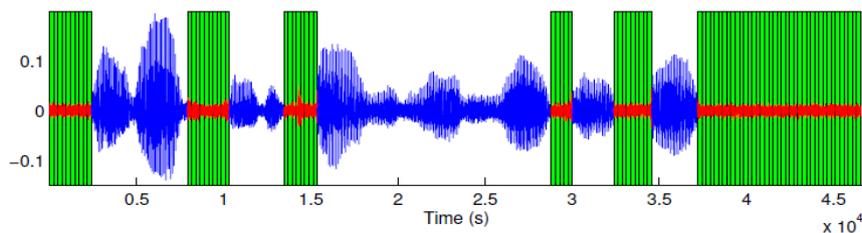


Figure 4. Silence removal by end point detection algorithm at 30 decibels for white gaussian noise

Table 1. Speaker recognition rate degraded by additive gaussian noise

Method	Noise at 30 decibels	Noise at 20 decibels	Noise at 10 decibels
STE algorithm	96.02%	94.3%	93.25%
ZCR algorithm	98.5%	96.9%	95.05%
End point detection algorithm	97.5%	95.6%	94.34%
Recursive least squares (RLS) algorithm	94.87%	92.74%	91.55%
log minimum mean square error logMMSE algorithm	95.23%	92.41%	91.33%

Table 2. Speaker recognition rate degraded by additive channel noise

Method	Noise at 30 decibels	Noise at 20 decibels	Noise at 10 decibels
STE algorithm	97.02%	95.35%	94.25%
ZCR algorithm	98.92%	96.65%	94.45%
End point detection algorithm	98.25%	95.17%	94.12%
Recursive least squares (RLS) algorithm	97.45%	94.35%	93.15%
log minimum mean square error logMMSE algorithm	96.24%	95.13%	92.45%

5.2.1. Population vs accuracy

This experimental setup uses different population sizes (150, 300, 450, 600, 750, 900, 1050, and 1200). Accuracy is calculated by performing speaker identification assessments on voices picked at random mostly from 1500 voice database. Figure 5 and 6 reflect the effect of population size with accuracy under noisy conditions. The accuracy with the application of MPCA feature extraction is improved.

Further investigations are carried out in order to evaluate the efficiency of MPCA with FOLOS SVM algorithm method. The recognition system is tested on noisy database provided by NIST database having gaussian and channel noises. Tests have been performed using 1200 voices (males 712 and females 488 females) mostly from NIST dataset. Every speaker's classification is performed on about 6 sentences of 28 seconds clean voices.

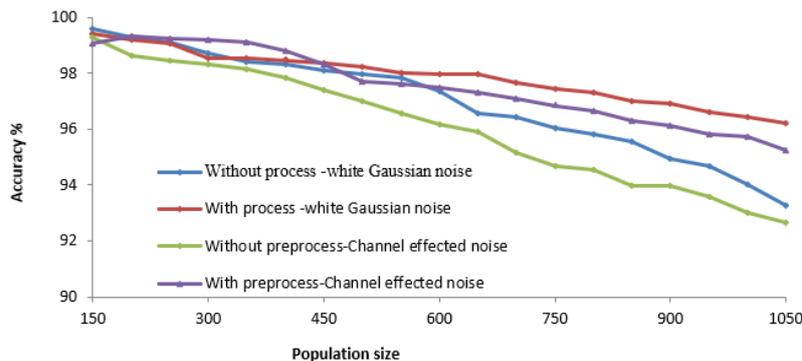


Figure 6. Speaker detection levels for additive white gaussian noise and channel noises signal to noise ratios of 10 decibels

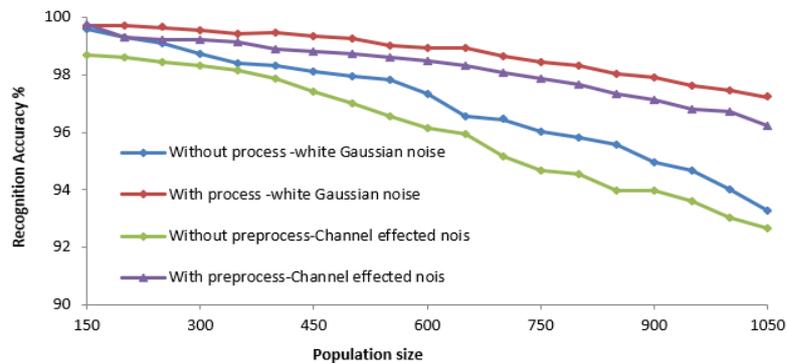


Figure 6. Speaker detection levels for additive white gaussian noise and channel noises signal to noise ratios of 30 decibels

Results from studies with gaussian noise at signal to noise ratio of 30 decibels increased by 2.76 percent (without application of algorithms) but 3.26 percent (with application of algorithms) after employing MPFA transition and forward looking schostic SVM algorithm. Alternatively, the rise in channel affected noise at 30 decibels is 1.88% and 1.95% accordingly. According to the above data, the existence of noise influences output method but is much more prevalent.

6. CONCLUSION

The results achieved from the experiments indicate that the MPCA transformation improves the accuracy when compared to the conventional features. Though, the reduction is 50% in computational complexity, the MPCA-FOLOS outperforms the state-of-art speaker recognition techniques. Compared to other techniques, our suggested methodologies improve more the overall performance of the robust speaker recognition system at noisy conditions. Though, the different noise contents present in the system the improvement is significant. As explained in the above mentioned sections, there are considerable constraints

during learning and testing stages, it does not contribute to significant increase. If the overall response of the recognition system is taken, the procedure of the candidate selection represented by FOLOS is the optimal set of feature vectors. The proposed technique is experimentally proved and also tested that the performance is improved than the existing technologies. Also it is identified that the complexity of SVM is reduced significantly for the large number of speakers in candidate selection step.

REFERENCES

- [1] Wenlu Ma., and Han Liu., "Classification method based on the deep structure and least squares support vector machine," *Electronics Letters*, vol. 56, no. 11, pp. 538–541, 2020.
- [2] J. Dass. Narawane, Y., Mahapatra, R. N., and Sarin, V "Distributed Training of Support Vector Machine on a Multiple-FPGA System," *IEEE Transactions on Computers*, vol. 69, no. 7, pp. 1015–1026, 2020.
- [3] H. Wang., Pan, X., and Xu, Y., "Simultaneous Safe Feature and Sample Elimination for Sparse Support Vector Regression," *IEEE Transactions on Signal Processing*, vol. 67, no. 17, pp. 4043–4054, 2019.
- [4] J. Jia., Zhou, S., Yin, Y., Wu, B., Chen, W., Meng, F., and Wang, Y., "Inferring Emotions from Large-Scale Internet Voice Data Clean," *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp.1853–1866, 2019.
- [5] W. Zuo., Wu, X., Lin, L., Zhang, L., and Yang, M. H., "Learning Support Correlation Filters for Visual Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 5, pp.1158–1172, 2019.
- [6] Zhu, F., Gao, J., Xu, C., Yang, J., and Tao, D., "On Selecting Effective Patterns for Fast Support Vector Regression Training," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3610–3622, 2018.
- [7] Shen, X. J., Mu, L., Li, Z., Wu, H. X., Gou, J. P., and Chen, X., "Large-scale support vector machine classification with redundant data reduction," *Neurocomputing*, vol. 172, pp. 189–197, 2016.
- [8] H. Joachims., "Making large-scale support vector machine learning practical," *Advances in learning*, pp. 169–184, 1999.
- [9] Himaja, D., Padmaja, T. M., and Krishna, P. R., "Over sample Based Large Scale Support Vector Machine for Online Class Imbalance Problem," *International Conference on Big Data Analytics*, 2018, pp. 348–362.
- [10] Ho, C. H., and Lin, C. J., "Large-scale Linear Support Vector Regression," *Journal of Machine Learning Research*, vol. 13, pp. 3323–3348, 2012.
- [11] S. Sra., "Efficient Large Scale Linear Programming Support Vector Machines," *European Conference on Machine Learning*, 2006, pp. 767–774.
- [12] Baek, J., and Kim, E., "A new support vector machine with an optimal additive kernel," *Neurocomputing*, vol. 12, pp. 279–299, 2019.
- [13] Ghaddar, B., and Naoum-Sawaya, J., "High dimensional data classification and feature selection using support vector machines," *European Journal of Operational Research*, vol. 26, no. 3, pp. 993–1004, 2018
- [14] A. Sarveniazi., "An Actual Survey of Dimensionality Reduction," *American Journal of Computational Mathematics*, vol. 4, no. 2, pp. 55–72, 2014.
- [15] Cheng, Z., and Lu, Z., "A Novel Efficient Feature Dimensionality Reduction Method and its Application in Engineering," *Hindawi Complexity*, Article ID 2879640, 2018.
- [16] P. R. K. Rao, "Random Forest Algorithm with a Half-Voting and Weighted Decision Trees for Interior Pedestrian Tracking," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 6971–6976, 2019.
- [17] Rao, P. R. K., and Rao, Y. S., "Dimensionality Reduction Techniques and SVM Algorithms for Large Population Speaker Identification," *International Journal of Signal Processing Systems*, vol. 4, no. 2, pp. 86–96, Apr. 2016.
- [18] Khine, K. L., and Nyunt, T. S., "Predictive geospatial analytics using principal component regression," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 3, pp. 2651–2658, Jun. 2020.
- [19] Vásquez, J. L., Vásquez, J., Briceño, J. C., Castillo, E., and Travieso, C. M., "Feature selection of RAPD haplotypes for identifying Peach Palm (*Bactris gasipaes*) landraces using SVM," *Wseas Transactions on Computers*, vol. 9, no. 3, pp. 205–214, Mar. 2010.
- [20] Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., and Saeed, J., "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," *Journal of Applied Science and Technology Trends*, vol. 01, no. 2, pp. 56–70, 2020.
- [21] Lerner, B., Yeshaya, J., and Koushnir, L., "On the Classification of a Small Imbalanced Cytogenetic Image Database," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 2, 2007.
- [22] Liu, R., Lin, Z., Su, Z., and Tang, K., "Feature extraction by learning Lorentzian metric tensor and its extensions," *Pattern Recognition*, vol. 43, pp. 3298–3306, 2010.
- [23] Ijaia, "New feature selection model-based ensemble rule classifiers method for dataset classification," *International Journal of Artificial Intelligence and Applications*, vol. 8, no. 2, pp. 17–43, Mar. 2017.
- [24] Rao, K. S., and Sarkar, S., "Robust Speaker Recognition in Noisy Environments," Springer eBook., 2014.
- [25] Chougule, S. V., and Chavan, M. S., "Robust Spectral Features for Automatic Speaker Recognition in Mismatch Condition," *Procedia Computer Science*, vol. 58, pp. 272–279, 2015.
- [26] Mak, M. W., and Chien, J. T., "Machine Learning for Speaker Recognition," Cambridge University Press, 2019.

BIOGRAPHIES OF AUTHORS

P. Rama Koteswara Rao is professor in the Department of ECE, NRI Institute of Technology, Pothavarappadu, Krishna Dt, AP, India. He obtained his Doctorate in Signal Processing from the College of Engineering, Andhra University, Visakhapatnam. He has twenty-six years of professorial experience and five years of industrial experience. He is specialized in speech signal processing, pattern detection, image processing, and embedded devices. He published more than 25 articles and presented many research papers at international conferences. He published two patents for India's Intellectual Property. He's an ISOI, ISTE member.



Sunitha Ravi is a Professor in Department of ECE, NRI Institute of Technoogy, Pothavarappadu, Agiripalli Mandal, Krishna District, Andhra Pradesh, India. She has submitted her thesis in Image Processing at K.L. Educational Society, Deemed to be University, Guntur District, Andhra Pradesh. She has twenty three years of Professiona experience. She is specialized in Signal and Image Processing and Embedded system. She Published 10 articles and published one patent.



T. Haritha is a Professor in Department of ECE, P.V.P. Institute of Technology, Vijyawada, Krishna Dt, Andhra Pradesh, India. She was awarded her Doctorate in Mobile Communications from Gitam University, Visakhapatnam. Twenty years of Professiona experience. She is specialized in mobile communications, image processing and embedded devices. She published neary 10 articles and presented papers at International Conferences. She published one patent and is member of ISTE.