

SPEAKER IDENTIFICATION USING A NONLINEAR SPEECH MODEL AND ANN

Lajish V.L¹, Sunil Kumar R.K² and Vivek P³

¹Department of Computer Science, University of Calicut,
Kerala - 673635, India.

¹lajish@uoc.ac.in, ²sunilcsirc@rediffmail.com,
³vivekpiritty@gmail.com

ABSTRACT

This paper introduces a nonlinear speech model for improved speaker identification. We modelled the speaker identity using Reconstructed Phase Space (RPS) of the speech signal and the Phase Space Point Distribution (PSPD) parameters. The PSPD parameters are extracted from five vowels uttered by different speakers. The speaker identification experiments are conducted based on the PSPD parameters using Feed Forward Multi Layer Perceptron (FFMLP). Overall performance of the Mel-frequency cepstral baseline system is compared with the proposed composite classifier system using both cepstral and PSPD features across ten different speakers. The experimental results indicate that the accuracy of the phase space approach by itself is below that of MFCC features and further it shows that the proposed approach in which PSPD features when used with MFCC, pitch and first formant frequency offers reasonable improvement in speaker identification accuracy of the system.

KEYWORDS

Speaker Identification, Nonlinear Systems, Reconstructed Phase Space, ANN

1. INTRODUCTION

The speaker recognition process is normally classified into two specific tasks namely speaker verification and speaker identification [1]. In verification, the goal is to determine from a voice sample if a person is whom he or she claims. It is the process of accepting or rejecting the identity claim of a speaker. In speaker identification, the goal is to determine which one of a group of known voices best matches the input voice sample [2]. It is the process of determining which registered speaker provides a given utterance. This technique makes it possible to use the speaker's voice to verify their identity and control access to services such as voice dialling, banking by telephone, telephone shopping, database access services, information services, voice mail, security control for confidential information areas, and remote access to computers [3]. Many speaker identifications systems are reported in literature [4] [5] [6] [7] [8]. The importance of modulation frequencies for speaker recognition and comparison of composite features have been studied in [9] and [10] respectively. Experimental evaluations of various features for robust speaker identification have also been carried out by many researchers [11] [12].

Source-filter models form the foundation of many speech processing applications such as speech coding, speech synthesis, speech recognition and speaker recognition technology. Usually, the DOI : 10.5121/ijait.2012.2502

filter is linear and based on linear prediction. The excitation for the linear filter is left undefined, modelled as noise, described by a simple pulse train or described by an entry from a large codebook [13]. While this approach has led to great advances in the last 30 years, it neglects structure known to be present in the speech signal. In practical applications, this manifests itself as an increase in the bit rate, less natural speech synthesis and an inferior ability to discriminate speech sounds. The replacement of the linear filter (or parts thereof) with nonlinear operators (models) should enable us to obtain an accurate description of the speech. This in turn may lead to better performance of practical speech processing applications. From a physics and mathematics viewpoint, in the traditional linear approach to speech modelling the true nonlinear physics of speech production are approximated via the standard assumptions of linear acoustics and 1D plane wave propagation of the sound in the vocal tract. Despite the limited technological success of the linear model in several applications, there is strong theoretical and experimental evidence for the existence of important nonlinear 3D fluid dynamics phenomena during the speech production that cannot be accounted for by the linear model. Examples of such phenomena include modulations of the speech airflow and turbulence [14].

A somewhat different approach for speech dynamics estimation based speaker recognition has been presented by Petry and Barone [15]. The approach is based on nonlinear chaos-theoretic approach called largest Lyapunov exponent and improvement on the cepstrum and delta-cepstrum [16]. The main arguments that show the evidences of nonlinearities in the speech signal are the excitation signal and the vocal-tract filter model are mutually independent. The airflow through the vocal tract is laminar and the wave equation is linear. Several speaker recognition systems are reported based on the non-linear properties of the speech signal. When only the nonlinear features are used to fully characterize a speaker, the results obtained are poor. Accuracy of only 6.49% using fractal dimension, 7.57% using entropy, 9.19% using largest Lyapunov exponents, and 19.19% combining all nonlinear features. The information provided by these parameters is not enough to discard other parameters. The same idea can explain the bad performance of the speaker identification system which uses only pitch, reaching accuracy of only 14.86% [17]. So, further investigations should be carried out in nonlinear point of angle to increase the recognition accuracy of the speaker recognition system. In this work we propose a novel speaker identification system based on non-linear properties of the speech signal and Artificial Neural Network (ANN). This paper is organized as three sessions. First Session deals with the speaker modelling using reconstructed phase space. Second session deals with the experimental results and the third session concludes the work.

2. SPEAKER MODELLING USING PHASE SPACE PARAMETERS

The phase space of a dynamical system is a mathematical space with orthogonal co-ordinate directions responding each of the variables needed to specify the instantaneous state of the system. Here phase space is the space in which all possible states of a system corresponding to one unique point. For mechanical systems, the phase space usually consists of all possible values of position and momentum variables. A plot of position and momentum variables as a function of time is called a phase diagram. The number of state variables determines the dimension of the system. These systems are typically treated as deterministic, i.e. if the state of the system at time t_0 is known, then the state of the system at any time t_1 is completely predictable.

Unfortunately, the entire state space of almost all real systems cannot be observed, if only one state variable is available. It would seem that accurate characterization of the system is impossible in this case, especially if the dimensionality and nonlinearity of the system are high. However, with the transformation on the observable variable known as a time-delay embedding, more information about the system is available than one might expect. Takens' theorem states that

under certain assumptions, phase space of a dynamical system can be reconstructed through the use of time-delayed version of the original scalar measurements [18]. This new state space is commonly referred to in the literature as a Reconstructed Phase Space (RPS), and has been proven to be topologically equivalent to the original phase space of the dynamical system, as if all the state variables of that system would have been measured simultaneously [19]. A reconstructed phase space can be exploited as a powerful signal processing domain, especially when the dynamical system of interest is nonlinear or even chaotic [20]. The RPS can be used to estimate the dynamical in-variants like attractor dimension, Kolmogorov Entropy etc. of the system [21]. Recently this approach has been taken in the realm of speech signal processing by many researchers [22] [23]. P Prajith has analysed nonlinear and chaotic nature of speech patterns and presented the use of Reconstructed State Space (RSS) for the recognition of Malayalam vowels [24]. V L Lajish studied State Space Point Distribution (SSPD) parameters derived from the gray scale based State Space Maps (SSM) of handwritten character patterns and used it for high speed Handwritten Character Recognition (HCR) applications [25]. N K Narayanan & V Kabeer has proposed SSM and SSPD of the gray scale images for the computer recognition of human faces [26]. In this work we used the nonlinear feature called Phase Space Point Distribution (PSPD) for speaker identification.

2.1. Phase Space of a Dynamical System

For a purely deterministic system, once its present state is fixed, the states at all future times are determined as well. Thus it will be important to establish a vector space called phase space or state space for the system, such that specifying a point in the space specifies the state of the system and vice versa. Then the information about the dynamics of the system can be obtained by studying the various features of the corresponding phase space distribution. The state of a particle moving in one dimension is specified by its position (x) and velocity (v). Its phase space is a plane. On the other hand a particle moving in three dimensions would have a six dimensional phase space with three position and three velocity directions. In phase space, momenta can be used instead of velocities.

2.2. Speaker Modelling Using Phase Space Point Distribution

In the case of speech signal, what we have is not a phase space object but a time series, only a sequence of scalar measurements. We therefore have to convert the observations into state vectors. This is the important problem of phase space reconstruction, which is technically solved by the method of time delay embedding. One of the profound results established in chaotic theory is the Takens' embedding theorem. Takens' theorem states that under certain assumptions, phase space of a dynamical system can be reconstructed through the time-delayed versions of the original scalar measurements. This new state space is commonly referred to in the literature as Reconstructed Phase Space (RPS), and has been proven to be topologically equivalent to the original phase space of the dynamical system.

Packard et al. [27] first proved the concept of phase space reconstruction in 1980. Soon after, Takens showed that a delay-coordinate mapping from a generic state space to a space of higher dimension preserves the topology [28]. This theorem provides important theoretical justification for the use of RPS's for system identification and pattern classification. Because the topology of the RPS is identical to the topology of the underlying system's phase space, we can expect the shape and density of the RPS attractor to provide valuable information of the system that generates a signal.

According to Takens' embedding theorem a reconstructed phase space can be produced for a measured state variable x_n , where $n = 1, 2, 3, 4, \dots, N$, via the method of delays by creating vectors given by

$$X_n = [x_n \quad x_{n+\tau} \quad x_{n+2\tau} \quad \dots \quad x_{n+(d-1)\tau}]$$

Where d is the embedding dimension and τ is the chosen time delay value. The row vector, X_n , defines the single point in the RPS. The row vectors then can be compiled into a matrix called a trajectory matrix to completely define the dynamics of the system and create a reconstructed phase space as

$$X_d = \begin{bmatrix} x_1 & x_{1+\tau} & x_{1+2\tau} & \dots & x_{1+(d-1)\tau} \\ x_2 & x_{2+\tau} & x_{2+2\tau} & \dots & x_{2+(d-1)\tau} \\ x_3 & x_{3+\tau} & x_{3+2\tau} & \dots & x_{3+(d-1)\tau} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_N & x_{N+\tau} & x_{N+2\tau} & \dots & x_{N+(d-1)\tau} \end{bmatrix}$$

A speech signal with amplitude values can be treated as a dynamical system with 1D dimensional time series data. Based on the above theory, this study investigates a method to model a Reconstructed Phase Space for vowels through the use of time delay versions of the original scalar measurements. Here trajectory matrix X_1 with embedding dimension $d = 2$ and time delay $\tau = 1$ and X_2 with embedding dimension $d = 3$ and time delay $\tau = 1$ are constructed by considering the speech amplitude values x_n as one dimensional time series data. The matrices X_1 and X_2 thus obtained are given below.

$$X_1 = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \\ x_3 & x_4 \\ \vdots & \vdots \\ x_N & x_{N+1} \end{bmatrix} \quad X_2 = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \\ x_3 & x_4 & x_5 \\ \vdots & \vdots & \vdots \\ x_N & x_{N+1} & x_{N+2} \end{bmatrix}$$

Next session describes the visual representation of the system dynamics by plotting the row vectors of the trajectory matrix constructed from speech signals.

2.3. Reconstructed Phase Space Distribution

The Reconstructed Phase Space Distribution plots (scatter graph) in two dimensions are constructed for five vowels /a/, /e/, /i/, /o/, and /u/ as follows. The normalized N sample values for the vowel sound is considered as the scalar time series x_n , where $n=1, 2, \dots, N$. For each speech signal, a trajectory matrix is formed with dimension $d=2$ and time delay $\tau=1$. Figure 1 shows the Reconstructed Phase Space (RPS) for vowel /a/ with $d=2$ and $\tau=1$.

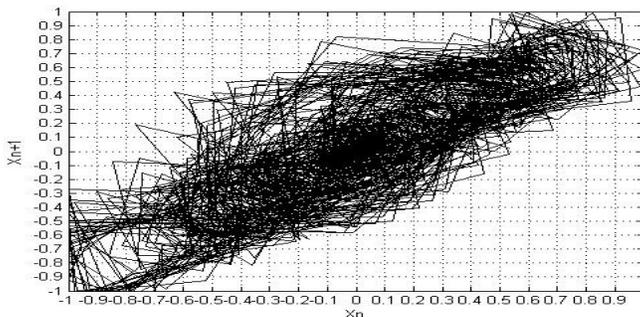


Figure 1. Reconstructed Phase Space (RPS) for the vowel /a/ with $d=2$ and $\tau=1$

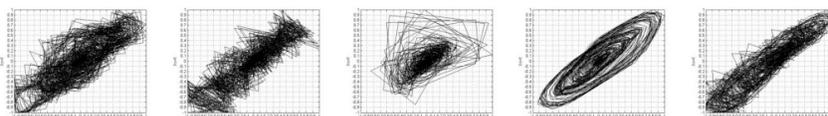


Figure 2. Reconstructed Phase Space (RPS) for five English vowels spoken by a single speaker.

We have analyzed the RPSs of a single vowel spoken by the same speaker and different speakers. The visual analysis of RPSs shows that the plots corresponding to the system dynamics of a particular vowel spoken by a single speaker are similar in nature and the RPSs of the same vowel spoken by different speakers differs considerably. This shows the credibility of the proposed features for speaker identification. Figure 3 shows the plots of RPSs of the vowel /a/ spoken by a single speaker. Figure 4 shows the plots of RPSs of the vowel /a/ spoken by five different speakers.



Figure 3. Reconstructed Phase Space (RPS) for five samples of the vowel /a/ spoken by a single speaker.

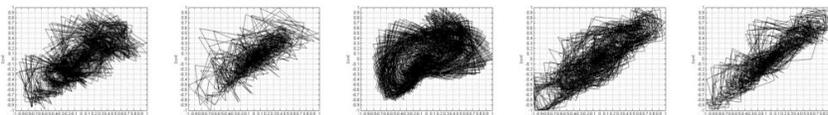


Figure 4. Reconstructed Phase Space (RPS) for vowel /a/ spoken by five different speakers.

Further a scatter plot named Reconstructed Phase Space Distribution is generated by plotting the row vectors of the above constructed trajectory matrix. (i.e. by plotting X_n versus X_{n+1}). Figure 5 shows the reconstructed phase space distribution plot for vowel /a/ with $d=2$ and $\tau=1$. Figure 6 shows the reconstructed phase space distribution plot for five different vowels spoken by a single speaker.

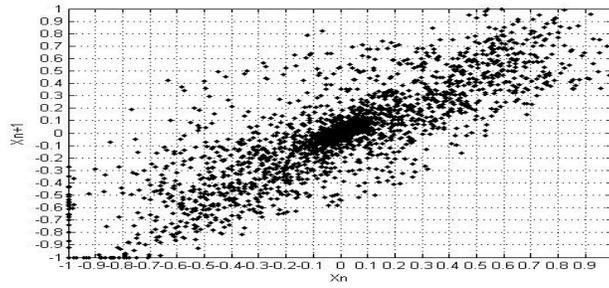


Figure 5. Scatter Plot for the vowel /a/ with $d=2$ and $\tau=1$

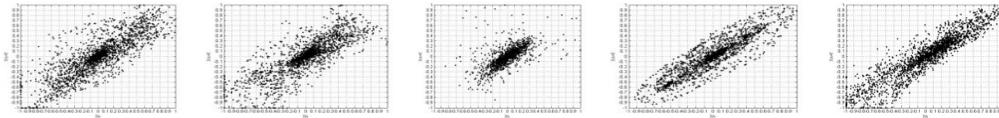


Figure 6. Scatter plots for 5 different vowels spoken by a single speaker

Figure 7 (a-e) shows the scatter plots of vowel /a/ spoken by five different speakers. Similar to RPSs the visual analysis of the scatter plots also shows that the plots corresponding to a particular vowel spoken by a single speaker are similar in nature and scatter plots of the same vowel spoken by different speakers differs considerably.

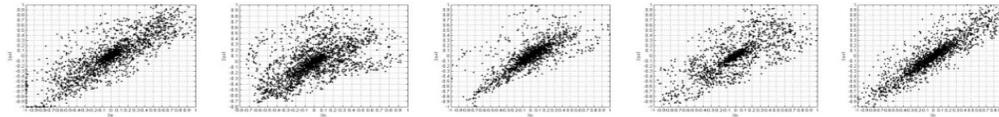


Figure 7. Scatter plot for vowel /a/ spoken by five different speakers

2.4. Phase Space Distribution Parameter

The complexity of RPS based approach due to the higher dimensionality of data can be reduced by extracting selected parameters from the reconstructed phase space. Here RPS is divided into grid with 400 uniform locations as indicated in Figure 5, and the number of phase space points distributed in each location is calculated. For this purpose, the RPS is divided into grid with 20 X 20 boxes. The box defined by coordinates (-1, 1), (-1, 0.9), (-0.9, 1) and (-0.9, 0.9) is taken as location 1. Box just right side to it is taken as location 2 and is extended towards X direction, with the last box with the coordinates (0.9, 1) (0.9, 0.9), (1, 0.9) and (1, 1) as location 20. This is repeated for the next row, taking the starting box as location 21 and repeated for all other rows. The reconstructed Phase Space Point Distribution (PSPD) parameter is calculated by estimating the number of Phase Space Points distributed in each location. This can be mathematically represented as follows. The Reconstructed Phase Space Distribution Parameter for location i in two dimensions can be defined as:

$$(PSPD)_i = \sum_{n=0}^N \delta([X_n, X_{n+\Gamma}], i)$$

$$\delta([X_n, X_{n+\Gamma}], i) = 1, \text{ if the phase space point defined by the row vector } [X_n, X_{n+\Gamma}] \text{ is in the location } i.$$

$$= 0, \text{ otherwise.}$$

Generally, PSPD for location i in d dimension can be defined as:

$$(PSPD)_i = \sum_{n=0}^N \delta([X_n, X_{n+\Gamma}, X_{n+2\Gamma}, \dots, X_{n+(d-1)\Gamma}], i)$$

Where,

$$\delta([X_n, X_{n+\Gamma}, X_{n+2\Gamma}, \dots, X_{n+(d-1)\Gamma}], i) = 1, \text{ if the phase space point defined by the row vector } [X_n, X_{n+\Gamma}, X_{n+2\Gamma}, \dots, X_{n+(d-1)\Gamma}] \text{ is in the location 'i'.$$

$$= 0, \text{ otherwise.}$$

Figure 8 shows the PSPD plot for vowel /a/ with $d=2$ and $\tau=1$. Figure 9 shows the feature vector graph plotted for the five samples of vowel /a/ of a single speaker. From the graph it is evident that the proposed feature vectors representing different samples of a particular vowel are similar for the same speaker and hence can be effectively used for speaker identification purpose. Next session describes the recognition experiments conducted using the proposed PSPD parameters and artificial neural networks.

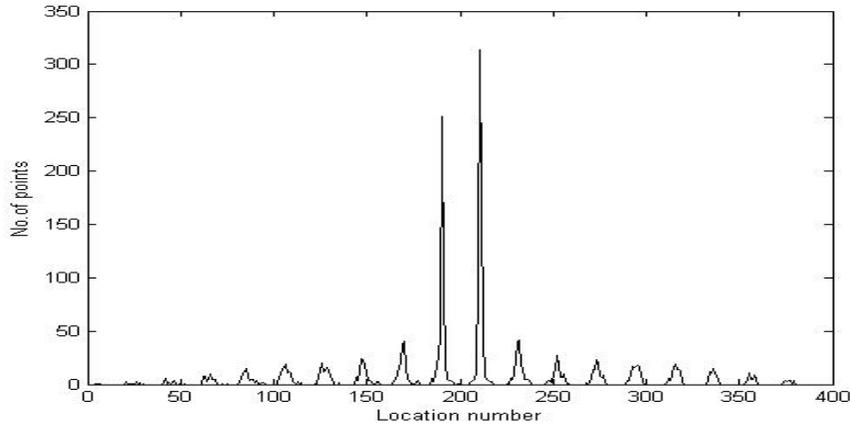


Figure 8. PSPD Plot for vowel /a/ with $d=2$ and $\tau=1$

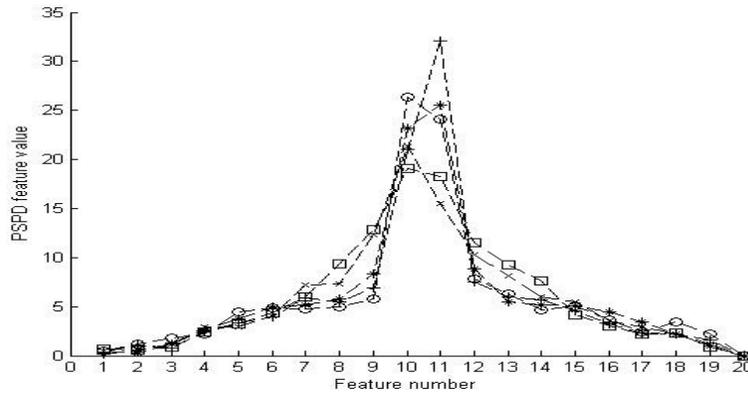


Figure 9. Feature vector graph plotted for the five samples of vowel /a/ of a single speaker.

3. EXPERIMENTAL RESULTS

Present study investigates the speaker identification capabilities of the proposed PSPD parameter and feed forward multilayer perceptron (FFMLP) based speaker identification system. The Reconstructed Phase Space (RPS) plots and its scatter plots obtained from five vowel samples of different users are thoroughly analysed. It is observed that the RPS patterns obtained for repeated samples of a particular vowel uttered by a single speaker are similar in nature and it varies from speaker to speaker. The PSPD parameters are extracted from five vowels uttered by 10 different speakers for the conduct of speaker identification experiments. Further, the multi layer feed forward neural network is simulated with the back propagation learning algorithm. A constant learning rate 0.01, is used (value of η was found to be optimum as 0.01 by trial and error method). The initial weights are obtained by generating random numbers ranging from 0.1 to 1. The number of input layer is fixed according to the feature vector size and the five output nodes representing five vowels. The recognition experiment is repeated by changing the number of hidden layers and number of nodes in each hidden layer for obtaining the successful architecture.

The network is trained using RPSDP feature and MFCC features extracted from the vowel samples of 10 speakers separately. Here we used a set of 50 samples each of five vowels spoken by 10 different speakers for iteratively computing the final weight matrix and disjoint set of vowels of same size from the database for identification purpose. The experiment is repeated by adding pitch and first formant frequency along with the MFCC feature and proposed PSPD parameters. The recognition accuracy obtained for ten different speakers based on above said features extracted from each of the five vowels using FFMLP classifier are tabulated in table 1. The experimental results indicate that although the accuracy of the phase space approach by itself is still currently below that of MFCC features, the combined feature approach is capable of considerably increasing the speaker identification accuracy.

Table 1. Speaker Recognition Results.

Parameters used	Size of the parameter	Speaker identification accuracy (%) (for each vowel)					Average (%)
		/a/	/e/	/i/	/o/	/u/	
MFCC	12	48.00	41.30	43.00	50.97	47.80	46.21
PSPD	20	36.00	26.00	26.00	38.00	32.00	31.60
PSPD + PITCH	21	40.92	33.00	31.00	46.00	39.00	37.98
PSPD + PITCH + FIRST FORMANT FREQUENCY	22	51.12	42.00	38.10	58.00	52.00	48.24
MFCC + PSPD + PITCH + FIRST FORMANT FREQUENCY	34	86.00	76.00	79.00	92.00	84.00	83.40

3. CONCLUSIONS

At present, the speaker identification systems produce acceptable results but it still lacks the necessary performance if they are to be used in commercial applications. In this work, we modelled the speaker identity based on the non-linear properties of the speech samples which are normally not considered in any of the conventional feature extraction methodologies. The speaker identification experiments are conducted based on the phase space distribution patterns derived from the Reconstructed Phase Space (RPS) of the speech signal named Phase Space Point Distribution (PSPD). The PSPD features obtained from five vowels are used for speaker identification purpose using the FFMLP. The experiment is repeated by taking different combination of PSPD, MFCC, pitch and first formant frequency. The experimental results indicate that the proposed phase space approach by itself is still below (31.60%) that of MFCC features (46.21%). The results further shows that the combined approach in which the PSPD features, when used with MFCC, pitch and first formant frequency, offers enormous improvement in speaker identification (on an average of 83.40%) accuracy.

REFERENCES

- [1] Campbell, J., (1997), "Speaker recognition: a tutorial", *Proceedings of the IEEE*, 85(9), pp. 1437–1462.
- [2] Douglas, A. Reynolds & Richard, C. Rose., (1995), "Robust text-independent speaker identification using Gaussian mixture speaker models", *Proceedings on speech and audio processing*, 3(1), pp 72–83.
- [3] Furui, S., (1997), "Recent advances in speaker recognition", *Pattern Recognition Letters*, 18 (9), pp. 859–872.
- [4] Atal B., (1972), "Automatic speaker recognition based on pitch contours", *Journal of the Acoustic Society of America*, 52 (6), pp. 1687–1697.
- [5] Atal B., (1974), "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification", *Journal of the Acoustic Society of America*, 55(6), pp. 1304–1312.
- [6] Bimbot, F., Magrin-Chagnolleau, I., & Mathan, L., (1995), "Second-order statistical measures for text-independent speaker identification", *Speech Communications*, 17, pp. 177–192.
- [7] Brunelli, R., & Falavigna, D., (1997), "Person identification using multiple cues", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(10), pp.955–966.
- [8] Farrell, K., Mammone, R., & Assaleh, K., (1994), "Speaker recognition using neural networks and conventional classifiers", *IEEE Trans. on Speech and Audio Processing*, 2(1), pp.194–205.
- [9] Vuuren, S., & Hermansky, H., (1998), "On the importance of components of the modulation spectrum for speaker verification", In *Proc. Int. Conf. on Spoken Language Processing (ICSLP 1998)* (Sydney, Australia), pp. 3205–3208.
- [10] Openshaw, J., Sun, Z., & Mason, J., (1993), "A comparison of composite features under degraded speech in speaker recognition", In *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 1993)* (Minneapolis, Minnesota, USA, 1993), pp. 27–30.
- [11] Reynolds, D., "Experimental evaluation of features for robust speaker identification", *IEEE Trans. on Speech and Audio Processing*, pp. 639–643.
- [12] Hardt, D., & Fellbaum, K., "Spectral subtraction and RASTA-filtering in text-dependent HMM-based speaker verification", In *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 1997)* (Munich, Germany), pp. 867–870.
- [13] Marcos Faundez-Zanuy, Unto Laine, Gernot Kubin, Stephen McLaughlin, Bastiaan Kleijn, Gerard Chollet, Bojan Petek & Amir Hussain (2005), "The COST-277 European Action: An Overview, Nonlinear Analyses and Algorithms for Speech Processing", *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, vol. 3817, pp 1-9.
- [14] H. M. Teager & S. M. Teager, (1989), "Evidence for Nonlinear Sound Production Mechanisms in the Vocal Tract, in *Speech Production and Speech Modelling*", W.J. Hardcastle and A. Marchal, Eds., NATO Advanced Study Institute Series D, vol.55.

- [15] Petry A., & Barone D., (2001), "Speaker identification using nonlinear dynamic features". *Chaos, Solitons and Fractals*, 13, pp.221–231.
- [16] Petry A., & Barone D., (2002), "Text-dependent speaker verification using Lyapunov exponents", In *Proc. Int. Conf. on Spoken Language Processing (ICSLP 2002)* (Denver, Colorado, USA), pp. 1321–1324.
- [17] Adriano Petry, Dante Augusto & Couto Barone, (2002), "Speaker identification using nonlinear dynamical features", *Chaos, Solitons & Fractals*, vol. 13(2), pp 221-231.
- [18] Takens, F., (1980), "Detecting strange attractors in turbulence", *Proc. Dynamical Systems and Turbulence*, Warwick, pp. 366-381.
- [19] Kubin, G. (1995), "Nonlinear speech processing in speech coding and synthesis", W. B. Kleijn and K. K. Paliwal, Eds.: Elsevier Science.
- [20] Kantz, H. & Schreiber, T., (2003), "Nonlinear time series analysis", Cambridge University Press, UK.
- [21] Abarbanel, H.D.I., (1996), "Analysis of observed Chaotic Data", Springer Verlag, New York.
- [22] Lindgren, A.C., Johnso, M.T., & Povinelli, R.J., (2003), "Speech Recognition using Reconstructed Phase Space features", *Proc. Int. Conf. on Acoustics, Speech and signal processing*, pp. 61-63, Hong Kong.
- [23] Michael T. Johnson, Richard J. Povinelli, Andrew C. Lindgren, Jinjin Ye, Xiaolin Liu & Kevin M. Indrebo, (2005), "Time-Domain Isolated Phoneme Classification Using Reconstructed Phase Spaces", *IEEE Trans. on Speech and Audio Processing*, vol.13(4), pp. 458-466.
- [24] P Prajith, (2008), "Investigations on the Applications of Dynamical Instabilities and Deterministic Chaos for Speech Signal Processing", PhD Thesis, Department of Physics, University of Calicut.
- [25] V L Lajish, (2007), "Adaptive Neuro – Fuzzy Inference Based Pattern Recognition Studies on Handwritten Character Images", PhD Thesis, University of Calicut.
- [26] N K Narayanan & V kabeer, (2010), "Face Recognition using Non-linear Feature Parameter and Artificial Neural Network", *International Journal of Computational Intelligent Systems*, vol 3(5), pp. 566 – 574.
- [27] Packard, N.H, Crutchfield, J.P, Farmer, J.D & Shaw, R.S, (1980), "Geometry from a time series", *Phys. Rev. Lett.*, vol. 45, pp.712-717.
- [28] Broomhead, D.S & King, J.P, (1986), "Extracting qualitative dynamics from experimental data", *Physica D*, vol.20, pp. 217-224.

AUTHORS

Dr.Lajish.V.L has been associated with University of Calicut, Kerala, INDIA as Head of the Department of Computer Science. He has worked as Scientist R&D in TCS Innovation Labs, Tata Consultancy Services Ltd. Mumbai, prior to joining the University. His prime areas of research include Digital speech and image processing, Pattern recognition algorithms and Indian language script technology solutions for mobile devices. He has more than thirty research publications in journals and peer-reviewed international and national conferences to his credit. After his masters in Computer Science from Vellore Institute of Technology, Dr.Lajish earned his Ph.D in Computer Science from University of Calicut in 2007. He is a senior life member of International Association of Computer Science and Information Technology.



Dr Sunilkumar R.K earned his Ph.D in Speech Signal Processing from University of Calicut, Kerala, India in 2004. He has published several research papers in national and international levels in the area of signal processing and artificial neural networks. His research interest includes speech signal processing and neural networks and active noise cancellation.



Vivek P earned his Masters in Computer Applications from the University of Calicut in 2010 and specialised in DSP System Design. He is an active research scholar working in the area of speech analytics under the supervision of Dr.Lajish.V.L

