



Feature Extraction Techniques for Speech Processing: A Review

Mohammed Arif Mazumder¹, Rosalina Abdul Salam^{1,2}

¹Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM), 71800 Nilai, Negeri Sembilan Malaysia, arif.marias@gmail.com

²Islamic Science Institute, Universiti Sains Islam Malaysia (USIM), 71800 Nilai, Negeri Sembilan, Malaysia, rosalina@usim.edu.my

ABSTRACT

In digital signal processing, speech processing is one of the areas that is used in many type of applications. It is one of an intensive field of research. The major criterion for good speech processing system is the selection of feature extraction technique, which plays a major role in achieving higher accuracy. In this paper, most commonly used techniques for feature extraction such as Linear Predictive Coefficient (LPC), Mel Frequency Cepstral Coefficient (MFCC), Perceptual Linear Prediction (PLP), Relative Spectral Perceptual Linear Prediction (RASTA-PLP) and Wavelet Transform (WT) are presented. Comparisons that highlight the strengths and the weaknesses of these techniques are also presented. Studies show that feature extraction techniques are mainly selected based on the requirement of the applications. Wavelet transform outperform other techniques for the analysis of non-stationary signals in audio signal. Enhanced Wavelet transform technique is a way forward and studies can be focused on its coefficients. Hybrid methods can be further explored to increase the performance in speech processing. A number of hybrid methods were reviewed, and studies show that Mel-Frequency Cepstral Coefficients (MFCC) provide better results for speech processing applications with standard coefficient for classification.

Key words: Linear Predictive Coefficient (LPC); Mel Frequency Cepstral Coefficient (MFCC); Perceptual Linear Prediction (PLP); Relative Spectral Perceptual Linear Prediction (RASTA-PLP); Wavelet Transform (WT)

1. INTRODUCTION

Speech processing involves a huge amount of signal data. In speech processing the speech signal is very crucial. This can later affect the classification and recognition stages. The nature of speech signal is a non-stationary signal. This make it more complex during the feature extraction stage. Dimensionality reduction is very important to ensure minimum or zero data lost during the feature extraction stage. The time domain waveform of a speech signal give us auditory information of this non-stationary signal. This waveform of the speech describes minimum information of a speech signal. Feature extraction in speech processing is very crucial especially for accuracy and performance. Currently,

many existing methods that are available for feature extraction in speech processing. The most commonly used are such as Linear Predictive Coefficient (LPC), Perceptual Linear Prediction (PLP), Mel Frequency Cepstral Coefficient (MFCC), Relative Spectral Perceptual Linear Prediction (RASTA-PLP) and Wavelet Transform (WT). These methods are explained and discussed in this paper. Comparative studies for these methods are provided. Studies show that methods are selected based on its applications. In recent years, hybrid methods are also introduced, and it shows that in most cases hybrid methods outperformed single methods. However, suitable methods are selected based on the domain. In the next following sections, the overview of what features extractions will be explained, it will be followed by the comparative studies of the most commonly used methods. Then the hybrid feature extractions methods and their major properties will be discussed.

2. FEATURE EXTRACTION

In a speech processing the process of extracting important information from a speech signal and reducing noise and unwanted information is called feature extraction. Basic operation of feature extraction involves spectral analysis, parametric transformation and statistical modeling [1]. The output is a parameter vector [2]. However, it is normal to lose useful information while removing unnecessary information [3]. Feature extraction involves the process of converting the speech signal into digital form [4]. This basic step of feature extraction is shown in Figure 1.

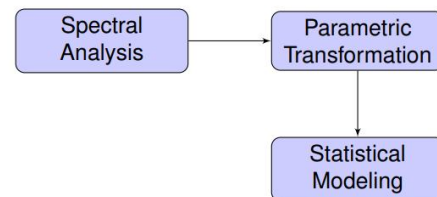


Figure 1: Basic Operation of Feature Extraction [1]

Spectral Analysis is the first stage of speech analysis and it includes spectro temporal analysis of signal [1]. In Parametric Transforms, two fundamental operations that are differentiation and concatenation stage are applied to create signal parameters from signal measurements [1]. Signal parameters were generated from few underlying multivariate random processes and this happen in Statistical Modeling stage.

3. FEATURE EXTRACTION TECHNIQUES

Speech signal can be retrieved directly from the digitized waveform [5]. Large data of speech signal requires suitable and reliable feature extraction techniques. This can improve the performance and computationally more effective. It will remove various source of information, such as whether the sound is voiced or unvoiced, that is whether speech are affected by noise or not [6].

3.1 Linear Predictive Coding (LPC)

In Linear Predictive Coding (LPC) analysis, a speech sample approximately combines past speech samples linearly. LPC is a frame based analysis of the speech signal [7]. LPC feature extraction process are shown in Figure 2. Adjacent frames in input speech signal are separated and is framed blocked into frames of samples. In order to minimize the signal discontinuities each individual frame is windowed [8]. This is followed by auto correlating each frame of windowed signal and then it converts each frame of autocorrelations into LPC parameter set by using Durbins method [8]. The LPC features vector were then created.

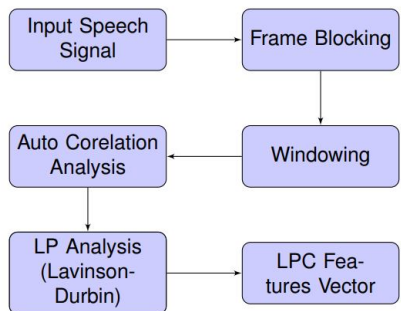


Figure 2: Linear Predictive Coding (LPC) Feature Extraction Process [6]

3.2 Mel Frequency Cepstral Coefficient (MFCC)

The Mel-frequency Cepstrum Coefficient (MFCC) technique is mainly used to create the fingerprint of the sound files. MFCC feature extraction process is shown in Figure 3.

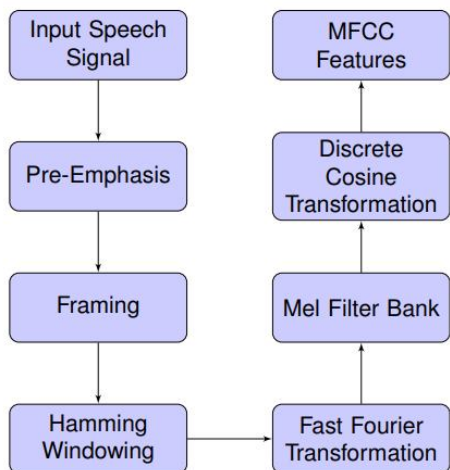


Figure 3: Mel Frequency Cepstral Coefficient (MFCC) Feature Extraction Process [9]

First input speech signal is divided into overlapping frames. Windowing is applied and then it is subjected to fast Fourier Transform. In the next step the frequency domain signal is converted to Mel frequency scale. Then the log Mel scale spectrum is converted to time domain using Discrete Cosine Transform (DCT) [9]. The result of the conversion is called Mel Frequency Cepstrum Coefficient. MFCC mainly concentrates on the static characteristics of a signal.

3.3 Perceptual Linear Prediction (PLP)

The Perceptual Linear Prediction basically discards irrelevant information such as noise and not similar to human voice. PLP is very similar to LPC but PLP is close to human voice system. The process of PLP is shown in Figure 4 [10].

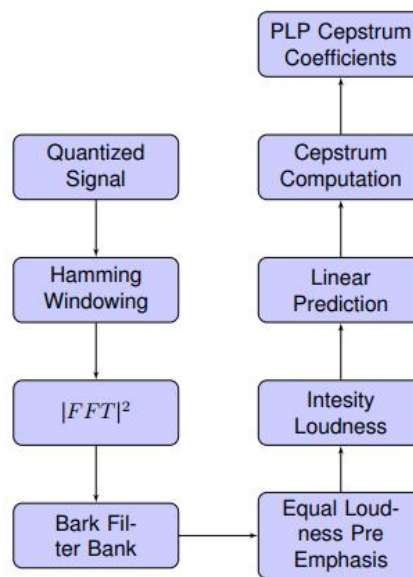


Figure 4: Perceptual Linear Prediction (PLP) Feature Extraction Process [10]

First, the quantized sign is windowed. This is to limit the sign discontinuities. At that point, Hamming Window is utilized and the power range of the windowed sign is determined to utilize FFT. The three stages of recurrence distorting, smoothing and examining are incorporated into a solitary channel bank called Bark Filter Bank [11]. To invigorate the affectability of human hearing an equivalent uproar pre-accentuation is utilized to loads the channel bank yields [12]. The yield that is the sound-related twisted line range is then prepared by the Linear Prediction to organize [13]. The last advance is the calculation of the Cepstral Coefficients.

3.4 Perceptual Linear Prediction (PLP) Relative Spectral Perceptual Linear Prediction (RASTA-PLP)

To remove short-term noise variations a special band-pass filter was added to each frequency sub-band in traditional PLP algorithm [14]. This is called RASTA-PLP and it is the filtering method used for removing the conventional disturbances. In traditional PLP, it has limited capability in

dealing with distortions. This is overcome by the introduction of RASTA-PLP. The input speech signal will go through the process of spectral analysis by using MFCC or PLP. This is then modified by the compressing the static non-linearity and will be filtered by band pass filter. Then another filter banks is used to expand the non-linearity and coefficients are produced. This is shown in Figure 5.

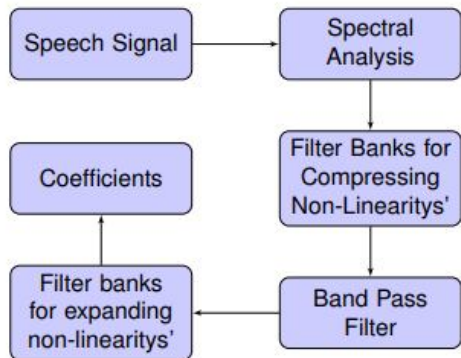


Figure 5: RASTA-PLP Feature Extraction Process [15]

3.5 Wavelet Transform (WT)

The wavelet transform is another method that has a similarity with how human ear processes sound. Therefore, it is suitable for speech processing. Discrete Wavelet Transform (DWT) and Wavelet Packet Decomposition (WPD) are explained and discussed in the next section.

3.5.1 Discrete Wavelet Transform (DWT)

DWT can extract information of non-stationary signals and it is very suitable for speech data. It is better in performance and computationally effective and efficient for feature extraction in speech. It has a varying window sizes therefore, it is efficient in all frequency ranges. Signal are passes through two filters that are low-pass filter and a high-pass filter and it produces two signals [17]. The output of a low pass filter is called as approximation coefficients and the output of highpass filter is called as detail coefficients [17].

3.5.2 Wavelet Packet Decomposition (WPD)

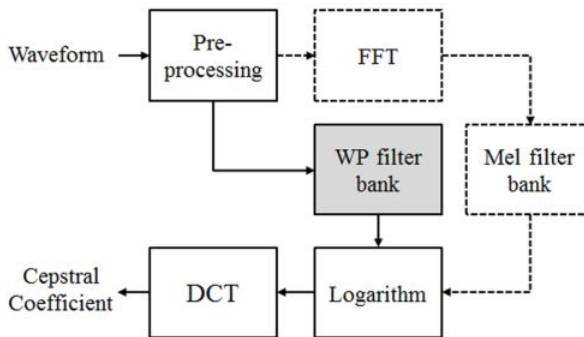
A generalization of DWT is actually WPD. Therefore, WPD is more flexible. Similar to DWT, WPD is decomposed into low frequency components and high frequency components. The difference is that in WPD it applies the transform step to the low pass and high pass results whereas in DWT it only apply to low pass results [18].

4. COMPARISON OF FEATURE EXTRACTION TECHNIQUES

Feature selection and extract are very crucial to speech recognition system. In most cases, it is domain or applications oriented. Table 1 presented the strengths and the weaknesses of the most commonly used feature extractions methods. Applications related to each method are also highlighted.

5. FEATURE EXTRACTION TECHNIQUES

Earlier a few features extraction techniques that are single techniques with their strengths and weaknesses were



presented. Studies show that better performance can be obtained by combining a few methods together to extract relevant features. These hybrid methods can be further investigated. A few significant hybrid feature extraction techniques and their comparison will be discussed in the next following sections.

5.1 Discrete Wavelet Packet Decomposition (DWPDP)

For discourse improvement and to conquer the impediments of DWT and WPD, new cross breed strategies were presented. This new half breed technique is called Discrete Wavelet Packet Decomposition (DWPDP) and it joins the highlights of both DWT and WPD. It comprises of three stages process where from the outset the discourse sign is part into two groups that are High and Low-recurrence band signal. At that point, WPD is connected to the high-recurrence segments and DWT is connected to the low-recurrence segments. In conclusion, the highlights delivered from the two techniques are joined and a component vector set is shaped [24].

The half and half calculation DWPDP has a couple of focal points, for example, the high-recurrence band are disintegrated into more parcels. This will expand the presentation and computationally increasingly successful and produce a higher acknowledgment rates [25], [26].

5.2 Phase Autocorrelation Bark Wavelet Transform (PACWT)

Phase Autocorrelation Bark Wavelet Transform (PACWT) combines the benefits of phase autocorrelation (PAC) with bark wavelet transform. It is a hybrid method and improve the robustness based on alternative measure of autocorrelation. The process of PACWT is shown in Figure 6.

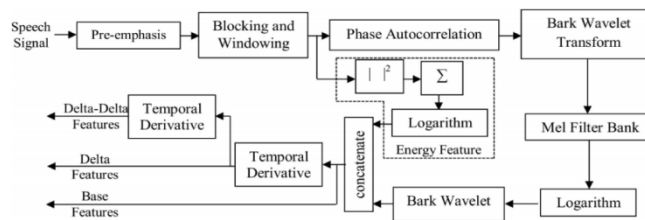


Figure 6: Block diagram of the PACWT Feature Extraction [28]

Table 1: Selected Feature Extraction Techniques

Methods	Applications	Strengths	Weaknesses
Linear Predictive Coding (LPC)	Tonal analysis, Musical instrument	<ul style="list-style-type: none"> • LPC method is easy to implement and the mathematics are very precise and simple [19]. • Low dimension feature vectors are represented for the spectral envelope [19], [21]. 	<ul style="list-style-type: none"> • Feature components are highly correlated [19] • The representation of speech production or perception based on the linear scales are not adequate [20]. • A priori information on the speech signal under test cannot be included [19].
Mel Frequency Cepstral Coefficients (MFCC)	Voice recognition system for security purpose	<ul style="list-style-type: none"> • It's not based on linear characteristics; hence, similar to the human auditory perception system [19], [20] • Low correlation between coefficients [19] • Provides good discrimination 	<ul style="list-style-type: none"> • Limited representation of speech signals since only the power spectrum is considered [19] • Low robustness to noise [19],[20]
Perceptual Linear Predictive Analysis (PLP)	Speech analysis	<ul style="list-style-type: none"> • Low dimensional for the resultant feature vector [19] • Voiced and unvoiced speech has reduction in the discrepancy [19] 	<ul style="list-style-type: none"> • Spectral balance is easily altered by the communication channel, noise, and the equipment used [19] • Dependent on the whole spectral balance [19].
Relative Spectral Perceptual Linear Prediction (RASTA-PLP)	Spectrum factor analysis	<ul style="list-style-type: none"> • Spectral components that change slower or quicker than the rate of change of the speech signal are suppressed [19] • These features are best used when there is a mismatch in the Analog input channel between the development and fielded systems [20] 	<ul style="list-style-type: none"> • Poor performance in clean speech environments [22]
Wavelet Transform (WT)	Multiresolution analysis, Time frequency localization, and Multirate filtering	<ul style="list-style-type: none"> • Capable of compressing a signal without major degradation [19] • Able to perform efficient time and frequency localizations [19],[23] 	<ul style="list-style-type: none"> • Not flexible as same basic wavelets have to be used for all speech signals [19]

First, the speech signal is pre-emphasized where Hamming window is used for a given frame on the pre-emphasized signal. Then computing correlation coefficients produce autocorrelation sequence during the Phase Autocorrelation. This is followed by simply applying the bark wavelet transform to the signal that passes through the Mel-filter bank. Finally, PACWT feature coefficients are produced. Then the first and second derivatives of the time sequence of each base feature are also calculated. Final PACWT feature coefficients set were produced by the concatenation of the derivatives to the base feature set.

5.3 Wavelet Based Mel-Frequency Cepstral Coefficients (WPCC)

Wavelet Based Mel-Frequency Cepstral Coefficients (WPCC) is a hybrid of the wavelet transform method and the MFCC. Firstly, the wavelet transform is applied to the speech signal into two different frequency channels to decompose them. High frequency channel components have all the details and the low frequency channel are only the approximations. Then the MFCC of the approximations and details channels are calculated. This is for capturing the characteristics of individual speakers [29]. This will ease the calculation of the coefficients. The process of WPCC is shown in Figure 7.

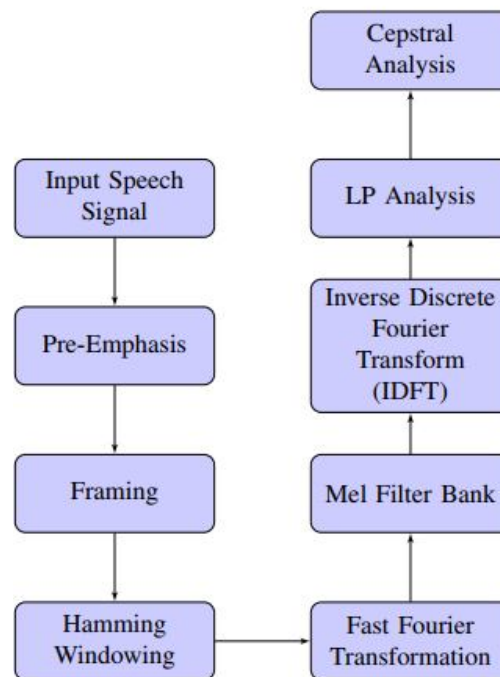


Figure 8: Feature Extraction using RPLP [32]

5.5 Bark frequency cepstral coefficients (BFCC)

BFCC is a hybrid of PLP and Bark filter bank. BFCC is very similar to MFCC except that it uses the bark filter bank in comparison to Mel filter bank [34]. As mentioned earlier bark filter are sensitive to human hearing. The signal is compressed and finally DCT is used to de-correlate the features.

Analysis shows that wavelet based DWPD are much more efficient, the performance is higher and the computational complexity are reduced. The dimensionality reduction is efficient with wavelet based DWPD and it produces better vector size. It increases the accuracy and suitable for non-stationary signals. Comparison between Phase Autocorrelation Bark Wavelet Transform (PACWT) and MFCC shows that PACWT are better for male voice data compared to female voice data. This is because it is better in low-SNR conditions. Revised Perceptual Linear Prediction Coefficients (RPLP), are mostly used in spoken language identification. It has the advantage of the pre-emphasis filter, Mel scale filter bank, LP and cepstral analysis. MFCC and BFCC shows good performance, however in noise environments, MFCC shows better performance. However, Wavelet Packets (WPs) shows better performance in comparison to MFCC due to its rich coverage of time-frequency properties. Table 2 highlights the strengths and the weaknesses of the presented hybrid methods.

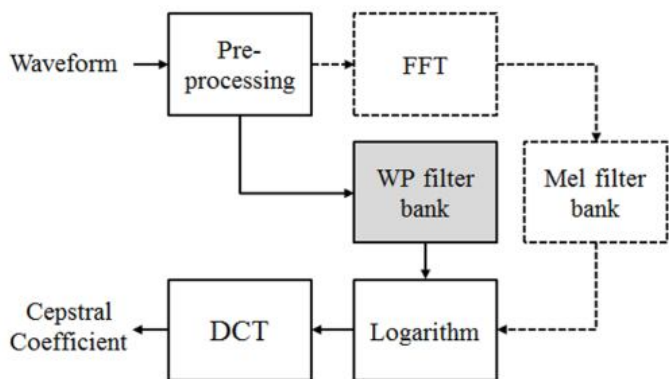


Figure 7: Mel-Frequency Cepstral Coefficient (MFCC) using Mel filter bank and Wavelet Packet Cepstral Coefficient (WPCC) using wavelet packet (WP) filter bank [29]

5.4 Revised Perceptual Linear Prediction (RPLP)

RPLP is a hybrid feature extraction based on PLP and MFCC. It uses Mel Filter bank instead of bark filter bank. First, the input signal is pre-emphasized then the segmentation and FFT spectrum is processed by applying Mel scale filter bank. The output is converted to the cepstral coefficients using LP analysis. The first six steps are similar to MFCC steps. Then it is followed by PLP steps. This steps can be seen in Figure 8. After all these steps IDFT, LP analysis and Cepstral analysis were applied in the same way as in PLP features.

Table 2: Hybrid Feature Extraction Techniques

Methods	Applications	Strengths	Weaknesses
Discrete Wavelet Packet Decomposition (DWPD)	Speaker independent digits recognition	<ul style="list-style-type: none"> • Computational complexity is reduced because it can decompose high frequency band into more partitions [24], [26]. 	<ul style="list-style-type: none"> • Performance reduce for stationary signals [25].
Phase Autocorrelation Bark Wavelet Transform (PACWT)	Robust Speech Recognition and Speaker Identification	<ul style="list-style-type: none"> • The PACWT feature extraction method is generally noise-robust compared to MFCC, particularly in high-noise (low-SNR) environments [28]. • Recognition performance was significantly better for male data than for female data [28]. 	<ul style="list-style-type: none"> • In clean speech MFCC has a higher recognition rate than the PACWT [28].
Wavelet-Based Mel-Frequency Cepstral Coefficients (WPCC)	Speaker Identification System	<ul style="list-style-type: none"> • For clean speech, it provides better performance compared to MFCC features [30], [31]. • It reduces the problem of noise and improves efficiently the recognition rate [31]. 	<ul style="list-style-type: none"> • WPCC does not show the robust performance in ASR [30], [29].
Revised Perceptual Linear Prediction (RPLP)	Spoken Language Identification	<ul style="list-style-type: none"> • RPLP features increase the accuracy of the recognition relatively better than the standard MFCC [50]. • Improve of recognition accuracy against PLP under noisy conditions [32]. 	<ul style="list-style-type: none"> • Identification accuracy vary; depends on different classifier
Bark Frequency Cepstral Coefficients (BFCC)	Speech Recognition in noisy environments	<ul style="list-style-type: none"> • Higher identification accuracy is produced for infinite distance in comparison with other feature extraction methods [33]. 	<ul style="list-style-type: none"> • MFCC perform better than the conventional BFCC method and sometimes performance degrade under noisy environment [29], [33].

6. CONCLUSION

Speech processing involves with big amount of speech signal data. Therefore, data reduction is very important in reducing the computational complexity and increase the performance. However, data reduction can results in losing important speech signals. Selecting feature extraction technique is very important in preserving important speech signals. Careful selection of methods can be decided and the applications should also be considered. This paper presented a number of commonly used feature selection methods and a few hybrid methods. The strengths and the weaknesses of the methods were presented and discussed. FFT, LPC and MFCC has higher computational complexity. Basically, they are much better for stationary signals compared non-stationary signal. Wavelet based methods provide less computational complexity and give higher performance. The accuracy are also higher in comparison with non-wavelet based method. From the literature, studies show that wavelet based method is a recommended method for speech signals. It shows that different applications require different feature extraction methods. However, in most cases wavelet based methods gives better accuracy with higher performance. Hybrid methods provide better results in comparison with single methods. The wavelet based Mel-Frequency Cepstral Coefficients (WPCC) shows higher accuracy for speech processing applications and provide standard coefficient for classifications. Further improvement can be achieved by incorporating optimization algorithms. This can further provide higher accuracy with reduced computational complexities especially under noisy conditions.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Universiti Sains Islam Malaysia (USIM) for the supports and facilities provided. This research study is sponsored by Universiti Sains Islam Malaysia (USIM) under USIM Competitive Grant [PPP/UTG-0114/FST/30/11414].

REFERENCES

[1] Pooja V. Janse, Ratnadeep R. Deshmukh, 2014, **Design and Development of Database and Automatic Speech Recognition System for Travel Purpose in Marathi**, OSR Journal of Computer Engineering (IOSR-JCE), Volume 16, Issue 5, Ver. IV, PP 97-104
<https://doi.org/10.9790/0661-165497104>

[2] Urmila Shrawankar, **Techniques for Feature Extraction in Speech Recognition System: A Comparative Study**, Available from: <https://arxiv.org/ftp/arxiv/papers/1305/1305.1145.pdf>

[3] Ms. Yogita A. More, Mrs. S. S. Munot(Bhabad), 2016, **Effect Of Combination Of Different Features On Speech Recognition For Abnormal Speech**, International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 5 Issues 8,Page No. 17590-17592

[4] Monica Mundada, Bharti Gawali ,Sangramsing Kayte,2014, **Recognition and Classification Of Speech**

And Its Related Fluency Disorders, IJCSIT, Vol. 5 (5) , 6764-6767

[5] Sayf A. Majeed, Hafizah Husain, Salina A. Samad, 2015, **Phase Autocorrelation Bark Wavelet Transform (PACWT) Features for Robust Speech Recognition**, PAN – IPPT, Archives Of Acoustics, Vol. 40, No. 1, pp. 25–31
<https://doi.org/10.1515/aoa-2015-0004>

[6] Ms. Pratibha Saroj, Mrs.Shilpa Verma, 2015, **Speech Recognition Of Deaf And Hard Of Hearing People By Using Neural Network**, International Journal of Emerging Technology and Innovative Engineering Volume 1, Issue 8, ISSN: 2394 6598.

[7] Pratik K. Kurzekar, Ratnadeep R. Deshmukh, Vishal B. Waghmare, Pukhraj P. Shrishrimal, 2014 , **Issues and Challenges of Voice Recognition in Pervasive Environment**, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Issue 12.

[8] Hariharan Muthusamy, Kemal Polat, Sazali Yaacob, 2015, **Improved Emotion Recognition Using Gaussian Mixture Model and Extreme Learning Machine in Speech and Glottal Signals**, Mathematical Problems in Engineering, Volume 20, Article ID 394083
<https://doi.org/10.1155/2015/394083>

[9] Pratik K. Kurzekar, Ratnadeep R. Deshmukh, Vishal B. Waghmare, Pukhraj P. Shrishrimal,2014, **A Comparative Study of Feature Extraction Techniques for Speech Recognition System**, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Issue 12.
<https://doi.org/10.15680/IJRSET.2014.0312034>

[10] E.Chandra, K.Manikandan, M. Sivasankar, 2014, **A Proportional Study on Feature Extraction Method in Automatic Speech Recognition System**, International Journal Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering, Vol. 2, Issue 1.

[11] Sascha Disch, Harald Popp, 2012, **Apparatus and Method for Determining a Plurality of Local Center Of Gravity Frequencies of a Spectrum of an Audio Signal**, United States Patent Application Publication, US 2012/0008799 A1.

[12] A.Nagesh, 2016, **A Comparison of Feather Extraction Methods for Language Identification using GMM**, International Journal of Engineering Trends and Technology (IJETT), Volume 31, Number 4.
<https://doi.org/10.14445/22315381/IJETT-V31P239>

[13] Namrata Dave, 2013, **Feature Extraction Methods LPC, PLP and MFCC in Speech Recognition**, International Journal for Advance Research in Engineering and Technology, Volume 1, Issue VI.

[14] Inshirah Idris, Md Sah Salam, 2016, **Improved Speech Emotion Classification from Spectral Coefficient Optimization**, Advances in Machine Learning and Signal Processing, pp 247-257.
https://doi.org/10.1007/978-3-319-32213-1_22

[15] Sanjivani S. Bhabad, Kamaraj Naidu, 2014, **RASTA-PLP for Speech Recognition of Articulatory Handicapped People**, International Journal off

- Scientific Research and Education, Volume 2, Issue 11, Pages-2313-2321.
- [16] Rish Loughran, Alexandros Agapitos, Ahmed Kattan, Anthony Brabazon, 2017, **Feature selection for speaker verification using genetic programming**, Evolutionary Intelligence, Volume 10, Issue 12, pp 121
<https://doi.org/10.1007/s12065-016-0150-5>
- [17] Xuechuan Wang, Kuldip K. Paliwal, 2002, **A Modified Minimum Classification Error (MCE) Training Algorithm for Dimensionality Reduction**, Journal of VLSI Signal Processing 32.
- [18] Marko V. Jankovic, Masashi Sugiyama, **Probabilistic Principal Component Analysis Based on JoyStick Probability Selector**, Available from: <https://www.researchgate.net/publication/221533826>
- [19] Wenzhi Liao, Aleksandra Piurica, Paul Scheunders, Wilfried Philips, Youguo Pi, 2013, **Semisupervised Local Discriminant Analysis for Feature Extraction in Hyperspectral Images**, IEEE Transactions On Geo Science And Remote Sensing , Vol. 51, No. 1.
<https://doi.org/10.1109/TGRS.2012.2200106>
- [20] Lahiru Dinalankara, 2017, **Face Detection and Face Recognition Using Open Computer Vision Classifies**. Available from: <https://www.researchgate.net/publication/318900718>
- [21] Anusuya, M., Katti, S., 2011, **Front end analysis of speech recognition: a review**, Int. J. Speech Technol., 14, (2), pp. 99145
<https://doi.org/10.1007/s10772-010-9088-7>
- [22] Sonia Sunny, David Peter S., K Poullose Jacob, **A Comparative Study of Wavelet Based Feature Extraction Techniques in Recognizing Isolated Spoken Words**, Available from: <http://www.ijsp.com/uploadfile/2013/0710/20130710105020955.pdf>
- [23] S. Kadambe , P. Srinivasan, 1994, **Application of adaptive wavelets for speech coding**, Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis.
- [24] Korba, M.C.A., Messadeg, D., Djemili, R.H.B., 2004, **Robust speech recognition using perceptual wavelet denoising and mel-frequency product spectrum cepstral coefficient feature**, Informatica, 32, pp. 283288.
- [25] Zhou, P., Tang, L.Z., Xu, D.F., 2009, **Speech recognition algorithm of parallel subband HMM based on wavelet analysis and neural network**, Inf. Technol. J., 8, pp. 796800
<https://doi.org/10.3923/itj.2009.796.800>
- [26] Veisi, H., Sameti, H., 2011, **The integration of principal component analysis and cepstral mean subtraction in parallel model combination for robust speech recognition**, Digit. Signal Process, 21, (1), pp. 3653
<https://doi.org/10.1016/j.dsp.2010.07.004>
- [27] Lee, J.Y., Hung, J. , 2011, **Exploiting principal component analysis in modulation spectrum enhancement for robust speech recognition**, Eighth Int. Conf. on Fuzzy Systems and Knowledge Discovery (FSKD), Shanghai, pp. 19471951
<https://doi.org/10.1109/FSKD.2011.6019893>
- [28] Garau, G., Renals, S., 2008, **Combining spectral representations for large vocabulary continuous speech recognition**, IEEE Trans. Audio Speech Language Process., 16, (3), pp. 508518
<https://doi.org/10.1109/TASL.2008.916519>
- [29] Fontaine, V., Ris, C., Leich, H., 1996, **Nonlinear discriminant analysis with neural networks for speech recognition**, Proc. EUSIPCO 96, EURASIP, pp. 15831586
- [30] Venkateswarlu, R.L.K., Kumari, R.V., Jayasri, G.V., 2011, **Speech recognition using radial basis function neural network**, Third Int. Conf. on Electronics Computer Technology (ICECT), 2011, Kanyakumari, pp. 441445
<https://doi.org/10.1109/ICECTECH.2011.5941788>
- [31] Dengfeng, K., Shuang, X., Bo, X., 2008, **Optimization of tone recognition via applying linear discriminant analysis in feature extraction**, Third Int. Conf. on Innovative Computing Information and Control (ICICIC), Dalian, Liaoning China, pp. 528531
- [32] Sonia Sunny, David Peter S, K Poullose Jacob, 2013, **Design of a Novel Hybrid Algorithm for Improved Speech Recognition with Support vector Machines Classifier**, International Journal of Emerging Technology and Advanced Engineering, vol.3, pp.249-254.
- [33] P. Kumar, A. Biswas, A .N. Mishra and M. Chandra, 2010, **Spoken Language identification using hybrid feature extraction Methods**, Journal of telecommunication, vol. 1, pp. 11-5.
- [34] Shaurya Agarwala, Pushkin Kachroob, Emma Regentovab, 2016, **A hybrid model using logistic regression and wavelet transformation to detect traffic incidents**, IATSS Research, Volume 40, Issue 1, Pages 56-63.
<https://doi.org/10.1016/j.iatssr.2016.06.001>