I.J. Engineering and Manufacturing, **2017, 2, 39-49**
Published Online March 2017 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijem.2017.02.04

Available online at http://www.mecs-press.net/ijem

Voice Pathology Identification: A Survey on Voice Disorder

N.A. Sheela Selvakumari [a], Dr. V.Radha [b]

[a] *Research Scholar, Department of Computer Science, Avinashilingam Institute for Home Science and HigherEducation for Women, Coimbatore, Tamil Nadu*
[b] *HOD, Department of Computer Science, Avinashilingam Institute for Home Science and HigherEducation for Women, Coimbatore, Tamil Nadu*

Abstract

Nowadays, Identification and Classification of voice pathology plays a major role in the field of speech processing. This paper explores and compares various things like input database, parameters, features extraction techniques, methodology and classification techniques used by the researchers in the problem of identifying the voice pathology. In this paper, we compared seven research works done in the field of voice pathology identification and classification. By analyzing the data's mentioned in these research papers and by considering these research papers as a base study, we wish to do the further research on voice pathology identification.

Index Terms: Voice Pathology, Classification, Acoustic Analysis, Vocal Fold, MDVP Parameters.

1. Introduction

In recent years, a significant attention has been given to the domain of voice pathology identification and monitoring. In the voice pathology treatment, patients have to frequently visit the doctor for their voice therapy. But the patients are waiting for a long time to consult, they are spending a lot of money to find the pathology because the experts have to find the problems in the vocal folds using some endoscopic instruments only. Totally it is an expensive as well as a time-consuming process. Hence such things made the patients feel discomfort.

This situation paves the way for the research in finding an automated tool to identify the voice pathology. The basic purpose of this automated tool is to help the patients for identifying the pathological problems for their further progress.

Basically, the voice pathology may cause due to the faults in the speech organs, Autism, mental illness, Hearing Impairment, Paralysis or multiple disabilities. Clinically there is a wide range of guidelines and

* Corresponding author.
E-mail address: sheelaselvakumariphd@gmail.com, radharesearch@yahoo.com

methods to find the voice pathology and also these methods are subjective and invasive. But it is not the case for the automatic tool for the voice pathology.

Most of the researches were done based on the three methodologies namely, 1.Acoustics, 2.Parametric & Non-Parametric and 3.Statistical methods. The Acoustics features identify the pathology based on the functioning and condition of various speech organs. The major parameters used in the Acoustic analysis were Fundamental Frequency, Jitter, Shimmer, Harmonic to Noise Ratio and Intensity. In parametric & Non-parametric methodology [1,2], the parametric approaches were based on the analysis of speech signals and identifying the glottal signals and the Non-parametric approaches were based on the Time-Frequency, Magnitude Spectrum, and Amplitude Modulation.

Finally, the Statistical Methodology applies classification techniques to separate the normal and pathology signal using pattern recognition techniques. Thus Statistical methods implement the Artificial Intelligence concept to identify and predict the normal and pathology voice.

This paper indicates the literature review in section (2). A detailed study on the database, parameters, features, methodology and results for the related works done by the researchers in the field of Voice Pathology detection in section (3). Finally, the conclusion in noted in the section (4) and plan for future work is discussed in section (5).

2. Related Works

The Acoustic Parameter evaluation depends on the fundamental frequency, alone is somewhat difficult in finding the Voice pathology, but in the case of [3-6], the combination of Feature extraction techniques like Fundamental frequency, Jitter, Shimmer, Harmonic to Noise Ratio, Intensity, makes the evaluation quite easy to compare between normal and pathology voice. In [7] the Acoustic Features MFCC were used along with the GMM Classifier to train the HMM to classify the normal and pathology.

An automatic detection of speech pathology was developed based on HMM technique in [8], which emphasize the detection of pathologies that affect the speech using fricative, Prolongation, vowels and nasal. The language used for this research is "polish" language. In [9], Multilayer network methodology is used in finding the voice disorder classification. The wavelet energy coefficients were given as input for the Multilayer Neural Network. The database used for this research work was collected from the Tunisia National Hospital. This methodology gives 100% result in classifying the Normal voice and Pathology voice.

The classification methodology implemented for classifying voice pathology in [10] are GMM and SVM. The databases compressed and stored in MP3 format were used to detect the pathology that present in the human voice. A Cross-validation strategy was implemented in [11], which gives an improved performance while classifying the normal and pathological voice. The Database used in this work was Massachusetts Eye & Ear Infirmary Database. The performance of the system was evaluated by using the measurements DET and ROC Curves.

3. Related Works on Voice Pathology

In this section, some related works in the field of Voice pathology Identification and Classification were discussed based on their process, database, methodology, and results.

3.1. Multiband Approach

In this subsection, the research work "Detection of Voice Pathology using Fractal Dimension in a Multiresolution Analysis of Normal and Disordered Speech Signals" done by Zulfiqar Ali, Irraivan Elamvazuthi, Mansour Alsulaiman, Ghulam Muhammad were studied and analyzed.

Domain: Frequency

Dataset: MEEI Database

Pathology Disease: Adductor
Vocal Nodules
Keratosis
Vocal Fold Polyp
Paralysis

No. of Samples: 173 pathological & 53 Normal

Feature Extraction: KATZ (4) + MDVP (22)

Methodology: Multiband Approach

Classification: Support Vector Machine

Accuracy: 96.56%

Work Flow:

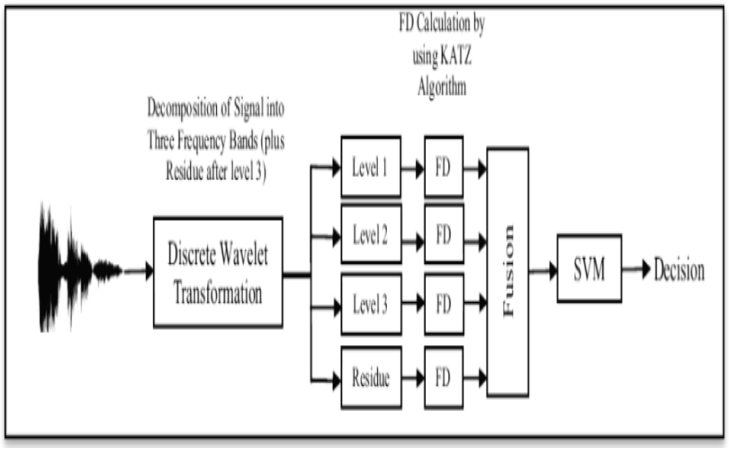

Fig.1. Research Plan for the Multiband Approach.

3.2. Multidimensional Approach

In this subsection, the research work "Automatic Assessment of Pathological Voice Quality Using Multidimensional Acoustic Analysis Based on the GRBAS Scale" done by Zhijian Wang, Ping Yu, Nan Yan, Lan Wang, Manwa L. Ng were studied and analyzed.

Domain: Frequency
Dataset: Voice samples collected from People's Liberation Army General Hospital

Pathology Disease: Vocal fold lesions,
Vocal cord paralysis,
Arytenoid granuloma,

Pre-cancerous vocal cord lesions,
Vocal cord carcinoma
Laryngectomies

No. of Samples: 320 pathological & 100 Normal (805 Samples)

Feature Extraction: LDA

Methodology: Multidimensional Acoustic Analysis

Classification: Extreme Learning Machines

Accuracy: 80.58%

Work Flow:

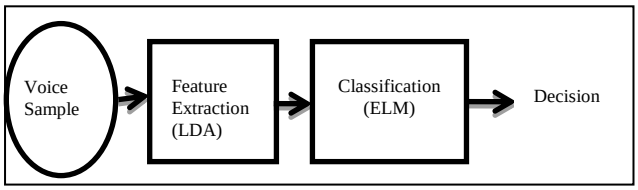

Fig.2. Research Plan for the Multidimensional Approach.

3.3. Multiple Feature Analysis

In this subsection, the research work "Voice Data Mining for Laryngeal Pathology Assessment" done by Daria Hemmerling, Andrzej Skalski, Janusz Gajda were studied and analyzed.

Domain: Voice samples of sustained vowels /a/, /i/ and /u/

Dataset: Saarbruecken Voice Database

Pathology Disease: Hyper functional dysphonia,
Vocal cord paresis
Laryngitis
Leukoplakia

No. of Samples: 705 pathological & 705 Normal

Feature Extraction: Principal Component Analysis

Methodology: Acoustic Analysis based on multiple features

Classification: Random Forest

Accuracy: 100%

Work Flow:

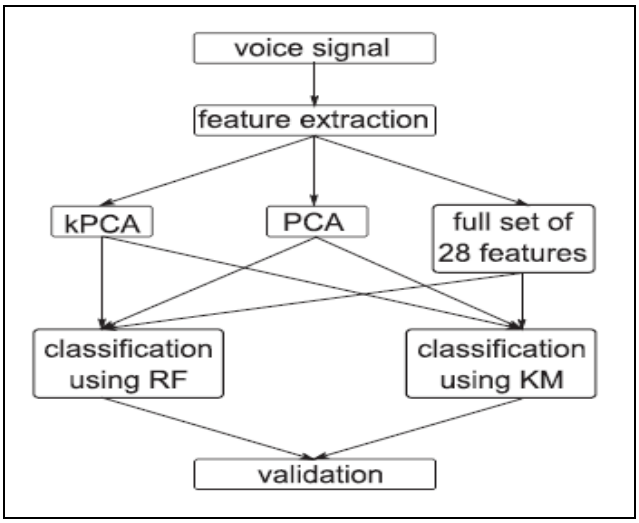

Fig.3. Research Plan for the Multiple Feature Analysis.

3.4. Band Wavelet Approach

In this subsection, the research work "Voice Disorder Signal Classification Using M-Band Wavelets and Support Vector Machine" done by Pouria Saidi, Farshad Almasganj were studied and analyzed.

Domain: Voice & Speech Signals

Dataset: 4337 database from Kay Elemetrics Corporation

Pathology Disease: Vocal fold paralysis
Vocal fold paresis
Nodules
Polyps
Edema

No. of Samples: 653 pathological & 57 Normal

Feature Extraction: M-band wavelet decomposition

Methodology: Five-band wavelet using Genetic Algorithm

Classification: Support Vector Machine

Accuracy: 99.3%

Work Flow:

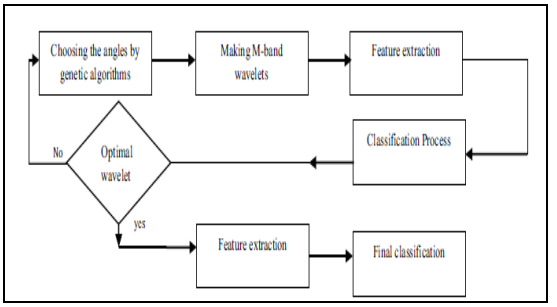

Fig.4. Research Plan for the Band Wavelet Approach.

3.5. Clustering Approach

In this subsection, the research work "Fusion of voice signal information for detection of mild laryngeal pathology" done by Adas Gelzinis, Antanas Verikas, Evaldas Vaiciukynas, Marija Bacauskiene were studied and analyzed.

Domain: Voice samples of vowel sound /a/

Dataset: Own Database

Pathology Disease: Laryngeal

No. of Samples: 689 pathological & 139 Normal

Feature Extraction: Ad-hoc decomposition of the 26 distinct features set

Methodology: Feature-oriented cluster-based partitioning

Classification: Random Forest

Accuracy: 98.2%

Work Flow:

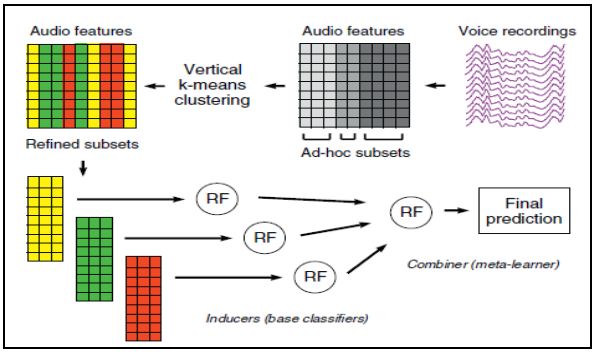

Fig.5. Research Plan for the Clustering Approach.

3.6. Acoustic Voice Analysis

In this subsection, the research work "Towards developing a Voice Pathologies Detection System" done by M. El Emary, M. Fezari, F. Amara were studied and analyzed.

Domain: Speech Signal

Dataset: Saarbruecken Voice Database

Pathology Disease: Spasmodic Dysphonia
Laryngeal

No. of Samples: 38 Pathological & 63 Normal

Feature Extraction: Mel-Frequency Cepstral Coefficients

Methodology: Acoustic Voice Analysis

Classification: Gaussian mixture model

Accuracy: 100%

Work Flow:

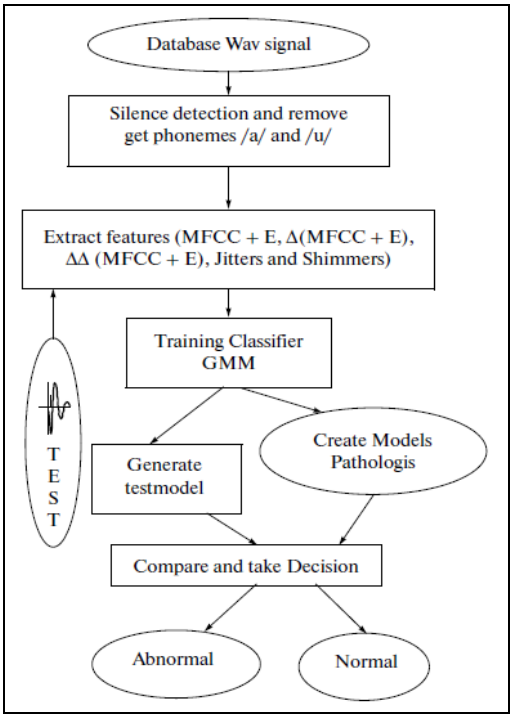

Fig.6. Research Plan for the Acoustic Voice Analysis.

3.7. Acoustic Feature Analysis

In this subsection, the research work "Classification of Normal and Pathological Voice Using SVM and RBFNN" done by V. Sellam, J. Jagadeesan were studied and analyzed.

Domain: Speech Signal

Dataset: Own Dataset

Pathology Disease: Vocal fold lesions,
Vocal cord paralysis,
Arytenoid granuloma,
Pre-cancerous vocal cord lesions,
Vocal cord carcinoma
Laryngectomies

No. of Samples: 10 Pathological & 10 Normal

Feature Extraction: Signal Energy
Pitch
Formant frequencies
Mean Square Residual signal
Reflection coefficients
Jitter and Shimmer

Methodology: Acoustic Feature Analysis

Classification: RBFNN

Accuracy: 91%

Work Flow:

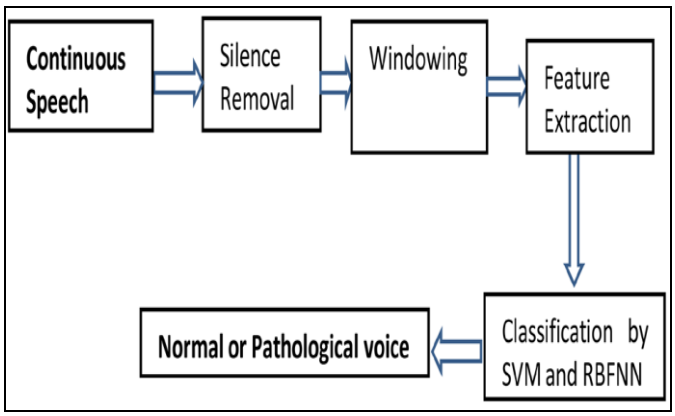

Fig.7. Research Plan for the Acoustic Feature Analysis.

4. Conclusion

A Comparative Study is made by analyzing the works done by the research community in the field of Voice Pathology Identification. The comparison is done based on their Domain, Database, No. of Samples taken for the experiment, Pathology Disease took for the research, Feature Extraction Techniques, Methodology adopted, Classification techniques adopted, the Workflow of the Research and the System Performance. By analyzing the above works, it shows that all the works are done based on the Acoustic Feature Analysis and Classification techniques. Most of the research papers mentioned that the Acoustic features play the main role in finding the pathology. Some research papers, mentioned above used the own database and with few samples the experiments were executed, if it is the case there may be fluctuations present in the results.

5. Future Work

Based on the review study of the research papers, we wish to do two things, one is to do a detailed study on the Acoustic Features and their combination to give a good result in our further research work, and the second one is to adopt a standard database for my further experiments.

References

[1] Raissa Tavares , Nathalia Monteir, "Optimizing laryngeal pathology detection by using combined cepstral features'' Proceedings of 20th International Congress on acoustics, ICA 2010 23- 27 August 2010, Sydney, Australia ICA 2010.

[2] Julian D. Arias-Londono,"Complexity Analysis of Pathological Voices by means Markov Entropy measurements'', 31st Annual International Conference of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6, 2009.

[3] Darcio G. Silva, Luıs C. Oliveira and Mario Andrea "Jitter Estimation Algorithms for Detection of Pathological Voices" Hindawi Publishing Corporation, EURASIP Journal on Advances in Signal Processing Volume, Article ID 567875, 2009.

[4] Miltiadis Vasilakis, Yannis Stylianou "Voice Pathology Detection Based on Short-Term Jitter Estimations in Running Speech'' Folia Phoniatr Logop, PP. 153–170, 2009.

[5] Sonu, R. K. Sharma "Disease Detection Using Analysis of Voice Parameters'' International Journal of Computing Science and Communication Technologies, VOL.4 NO. 2, January 2012.

[6] Jacques Koremana, "Correlates of Varying Vocal Fold Adduction Deficiencies in Perception and Production: Methodological and Practical Considerations" Folia Phoniatr Logop, PP: 305–320, 2004.

[7] Alireza A. Dibazar, Theodore W. Berger, and Shrikanth S. Narayanan "Pathological Voice Assessment" IEEE EMBS, NEW YORK, 2006.

[8] J. H. L. Hansen et al., "A Non Linear operator based speech feature analysis method with application to vocal fold pathology assessment," IEEE Trans. Biomed. Eng., 1995.

[9] L. Salhi, T. Mourad, and A. Cherif, "Voice Disorders Identification using Multilayer Neural Network," Int. Arab. J. Inf. Technol., Vol. 7, PP. 177–185, 2010.

[10] N. Saenz Lechon, et al., "Methodological issues in the development of automatic systems for voice pathology detection," Biomed. Signal Processing Control 1, PP. 120– 128, 2006.

[11] J. I. Godino Llorente, "Support Vector Machines Applied to the Detection of Voice Disorders", Springer Verlag, Berlin, PP. 219–230, 2005.

[12] Zulfiqar Ali, Irraivan Elamvazuthi, Mansour Alsulaiman, Ghulam Muhammad, "Detection of Voice Pathology using Fractal Dimension in a Multiresolution Analysis of Normal and Disordered Speech Signals", Journal of Medical Systems, Springer, 2015.

[13] Zulfiqar Ali, Irraivan Elamvazuthi, Mansour Alsulaiman, Ghulam Muhammad, "Automatic Voice

Pathology Detection with Running Speech by Using Estimation of Auditory Spectrum and Cepstral Coefficients Based on the All-Pole Model", Journal of Voice, Elsevier, 2015.
[14] Zulfiqar Ali, Mansour Alsulaiman, Irraivan Elamvazuthi, Ghulam Muhammad,"Voice Pathology Detection based on the Modified Voice Contour and SVM", Biologically Inspired Cognitive Architectures (BICA), Elsevier, 2015.
[15] Zhijian Wang, Ping Yu, Nan Yan, Lan Wang, Manwa L. Ng, "Automatic Assessment of Pathological Voice Quality Using Multidimensional Acoustic Analysis Based on the GRBAS Scale. Signal Processing Systems 82(2), PP: 241-251, 2016.
[16] Ping Yu, Zhijian Wang, Shanshan Liu, Nan Yan, Lan Wang, Manwa L. Ng, "Multidimensional acoustic analysis for voice quality assessment based on the GRBAS scale", ISCSLP, PP: 321-325, 2014.
[17] Pouria Saidi, Farshad Almasganj, "Voice Disorder Signal Classification Using M-Band Wavelets and Support Vector Machine", CSSP 34(8), PP: 2727-2738, 2015.
[18] Nafise Erfanian Saeedi, Farshad Almasganj, Farhad Torabinejad, "Support vector wavelet adaptation for pathological voice assessment", Comp. in Bio. and Med. 41(9), PP: 822-828, 2011.
[19] Nafise Erfanian Saeedi, Farshad Almasganj, "Wavelet adaptation for automatic voice disorders sorting". Comp. in Bio. and Med. 43(6), PP: 699-704, 2013.
[20] Evaldas Vaiciukynas, Marija Bacauskiene, Virgilijus Uloza, "Exploring similarity-based classification of larynx disorders from human voice", Speech Communication, 54(5), PP: 601-610, 2012.
[21] Evaldas Vaiciukynas, Antanas Verikas, Adas, "Fusing voice and query data for non-invasive detection of laryngeal disorders". Expert Syst. Appl. 42(22), PP:8445-8453, 2015.
[22] Adas Gelzinis, Evaldas Padervinskis, "Exploring sustained phonation recorded with acoustic and contact microphones to screen for laryngeal disorders". CICARE , PP: 125-132, 2014.
[23] Karmele Lopez de Ipina, Harkaitz Egiraun, "Feature Extraction Approach Based on Fractal Dimension for Spontaneous Speech Modelling Oriented to Alzheimer Disease Diagnosis". NOLISP , PP: 144-151, 2013.
[24] Karmele Lopez de Ipina, Blanca Beitia, "Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach", Computer Speech & Language 30(1), PP: 43-60, 2015.
[25] Karmele Lopez de Ipina, Harkaitz Eguiraun", On Automatic Diagnosis of Alzheimer's Disease Based on Spontaneous Speech Analysis and Emotional Temperature", Cognitive Computation 7(1) PP: 44-55, 2015.
[26] Fethi AMARA, Mohamed FEZARI, Hocine BOUROUBA, "An Improved GMM-SVM System based on Distance Metric for Voice Pathology Detection", Applied Mathematics & Information Sciences, May 2016.
[27] M. M. El Emary, M. Fezari , F. Amara, "Towards developing a voice pathologies detection system", Journal of Communications Technology and Electronics, Nov 2014.

Author's Profiles

Mrs. N.A. Sheela Selvakumari, has 4 years of teaching experience. Currently she is working in PSGR Krishnamal College for Women, Coimbatore. Her area of Specialization is Voice Recognition and Synthesis. She has 2 Publications at national and International Level.

Dr. V. Radha, Associate Professor, has more than 20 years of teaching experience. Her Area of Specialization is Image Processing, Optimization Techniques, Voice Recognition and Synthesis. She has 26 Publications at national and International level conferences.

How to cite this paper: N.A. Sheela Selvakumari, V.Radha,"Voice Pathology Identification: A Survey on Voice Disorder", International Journal of Engineering and Manufacturing(IJEM), Vol.7, No.2, pp.39-49, 2017.DOI: 10.5815/ijem.2017.02.04