Voice Parameter Analysis for the disease detection

Dixit[1], Vikas Mittal[2], Yuvraj Sharma[3]

[1]*(M.Tech. Student, Department of ECE, Maharishi Markandeshwar University, Mullana, Haryana, India.)*
[2]*(Asst. Prof, Department of ECE, Maharishi Markandeshwar University, Mullana, Haryana, India.)*
[3]*(Asst. Prof, Department of ECE, Maharishi Markandeshwar University, Mullana, Haryana, India.)*

Abstract: *The analysis of the human voice has arisen as an important area of study for its various applications in medical as well as engineering sciences. Voice analysis basically deals with extraction of some parameters from voice signal for processing of voice in desirable applicability by using suitable techniques. This paper states the certain common medical conditions which affect voice patterns of patients in evidence to leading research studies that had verified the voice alterations as diagnosis symptom in respective medical conditions. As well as, a comparative study of voice analysis techniques will be presented and special emphasis is given to certain prominent biomedical tools that are commercially available and are fundamentally based on voice analysis technology. Recent advancements in the field of voice analysis systems are also given.*
Keywords: *Voice analysis systems, pathological voice conditions, voice analysis techniques, voice feature extraction, time domain, frequency domain, and dysphonia.*

I. Introduction To Voice Analysis Systems

Voice medium is the most primitive and natural form of communication known to human civilization. Vocal communication is an essential skill required in man's professional as well as in personal life for expressing feelings, earning livelihood and in other day to day social interactions. People suffering from pathologic voices have to face many difficulties in their daily lives. The voice pathologic disorders are associated with respiratory, nasal, neural and larynx diseases. Thus, analysis and diagnosis of vocal disorders have become an important medical procedure. This has inspired a great deal of research in voice analysis measure for developing models to evaluate human verbal communication capabilities. The earliest attempt of this type was an idea of mechanical mimic of the human vocal apparatus by Wolfgang von Kempelen in 1791 [1] that was realised 40 years later by Charles Wheatstone [2]. The researchers of present era have exploited vast aspects of voice analysis due to the advancements in digital technology which had equipped them with highly versatile, fast and low cost digital processors. Moreover, voice being a time-varying signal has all the signal information present in its time variable characteristics.Voice analysis involves the transformation of voice signal into a set of signals or a set of parameters with an objective to simplify the voice signal to extract features directly pertinent for different applications and to suppress redundant aspects of the signal. These techniques assist speech pathologists and medical professionals to study and diagnose the voice signals with ease. Voice analysis can be done either in the time domain or in the frequency domain. The time domain waveform of voice signal carries all the auditory information like phonemes, syllables, prosody levels, etc [3] while the frequency domain represents the spectral characteristics of the voice signal such as frequency, pitch, etc. Table 1 illustrates information that can be extracted from voice parameters by time-frequency-based analysis. Speech Recognition technology is renowned medical application of voice analysis.

Table 1: Time-frequency-based extracted features.

Parameters	Features	Information extracted
Frequency	Jitter (local, absolute, ppq5, ddp).	Variation in the pitch of voice.
Pulse	Standard deviation of period, period of pulses.	Speech rate of speaker.
Amplitude	Shimmer (local, db, dda, apq11, apq5, apq3).	Variation in loudness of voice.
Voice	Unvoiced frames, voiced frames.	Structure of voice.
Harmonic	Noise- to- harmonic, Harmonic-to-noise.	Relative highness or lowness of voice.
Pitch	Maximum/Minimum pitch, Mean/Median pitch.	Peaks of the sound spectrum of voice.

II. Medical Conditions Marked By The Irregular Voice Patterns

Linguist and speech pathologists have diagnosed characteristic irregularities in voice patterns of patients with certain medical conditions like Asthma, Depression, Autism Spectrum Disorders, Parkinson's disease (PD), Alzheimer's disease, Schizophrenia, Larynx cancer, etc. Scientists as well as medical researchers are exploring the correlation to quantify these variations in voice corresponding to various medical conditions for designing the voice analysis systems for diagnosis and treatment of respective health disorders in a controlled clinical environment [19]. These voice analysis methods facilitate clinicians in screening patients and

scaling the progress of ongoing treatment. Numerous research studies have been conducted for classifying these acoustic [6], prosodic [7], emotional [14] or lexical voice features for extracting health information of the subjects.

2.1 Voice Dysphonia associated with Asthma condition:

Asthma is very common chronic inflammatory disorder of the airways characterized by the increased airway hyper-responsiveness, involuntary bronchospasm, reversible airflow obstruction, breathlessness and the associated symptoms are wheezing, allergies, coughing, sore throat and inflammation around the nasal track. Inflammation along the passageways of the nose goes down to the vocal cords of the larynx. Swollen, inflamed cords do not vibrate efficiently that makes the voice sound hoarse and thus voice quality is impaired [4]. Any qualitative and/or quantitative alteration of the voice is defined as dysphonia. Dysphonia can be due to both organic factors (organic dysphonia) and other factors (dysfunctional dysphonia) [8]. Spectral "noise" is strictly linked to air flow turbulences in the vocal tract, mainly due to irregular vocal folds vibration and/or closure, causing dysphonia. Such symptom requires a set of endoscopic analysis (using Video laryngoscope, VLS) for accurate analysis [5]. However, early detection of dysphonia can assist patients in timely recovery of asthma as health conditions deteriorate in later stages.

2.2 Aphasia as evaluation symptom of Alzheimer:

Dementia of Alzheimer Type (DAT) is the most common form of dementia. There is no cure for the disease and state of the patient worsens as it progresses which eventually leads to death. The common symptoms are: memory loss, confusion, irritability, aggression, trouble with language, and mood swings [26]. The most significant effect on functional communication skills that accompanies Alzheimer's disease (AD) is aphasia [27], a loss of oral communicative ability like breakdowns in semantic processing, shallow vocabularies and word-finding difficulties leading to the deterioration of spontaneous speech [28]. The Mini-Mental State Exam (MMSE) is a cognitive grading scale used in the assessment of patients first described by Folstein et al. [29] in 1975 and Verbal picture descriptions later became the other precise measures for assessing spontaneous speech in AD. [28]. Current methods of assessing DAT involve diagnosis of patients by Bucks et al. [25] procedure of lexical analysis, by taking statistics data of spontaneous speech. The authors concluded that the stylometric attributes had sufficient discriminating power in distinguishing between the language models of AD sufferers and control subjects.

2.3 Vocal impairment in People With Parkinson's (PWP):

Parkinson's disease is the second most pandemic neurodegenerative disorder after Alzheimer. Parkinson's disease is defined as a disorder of the central nervous system that results from the loss of the neurons, mainly from the substantia nigra region of the brain. The substantia nigra cells produce dopamine, a chemical messenger responsible for coordination of body movements. The loss of dopamine causes neurons to fire without normal control, leaving patients with the defects in motor as well as non-motor body functions. The initial PD symptoms are non-motor neuropsychiatric problems like cognition, voice disorder, behavioral alterations, sensory and sleep difficulties [31] followed by motor symptoms that include tremor, rigidity, slower movements and postural instability [31]. Parkinson's disease can be categorized by significant variations in voice (Dysphonia) [30]. The characteristic symptoms of Dysphonia in Parkinson's Disease are monotonous, reduced pitch, inability to vary pitch, variable rate, short rushes of speech, imprecise consonants, inability to sustain prolonged vowel phonation and a breathy and harsh voice [33] [34]. Parkinson's patients are unable to produce loudness, pitch, voice modulations and rhythm patterns required for expressing certain emotions [32]. Studies have found that people with Parkinson (PWP) produce fewer words, take frequent pauses, and have abnormal speech rate [36]. These variations in voice patterns become worse as the disease progresses [35]. Research has shown that speech may be a useful signal for discriminating PWP from healthy individuals [37]. On the basis of clinical evidences which suggests that the vast majority of PWP typically exhibit some form of vocal disorder [38] and these vocal impairments are among the earlier PD symptoms preceding the actual Parkinson's disease., even five years prior to clinical diagnosis [39], we can reinforce the fact that voice may reflect disease status.

2.4 Correlating voice acoustics' in Depression patients :

Depression is a psychiatric disorder defines by the state of low mood, thoughts, behaviour, senses, ailments and feelings. Depressed people become anxious, fatigue, irritable, worried and hopeless; face problems in sleeping, memorising, deciding, loses concentration or interest in activities that once were pleasurable, one can attempt or even commit suicide. Research studies have correlated the acoustic features of the depressed and suicidal subjects by analysing their voices. Research at Georgia Tech and the Medical College of Georgia [18] evaluated the patients suffering from clinical depression on the basis of voice features extracted from the glottal

waveform. Alpert et al. [6] stated that clinical impressions were substantially related to acoustic parameters like fluency and prosody of depression patients. Pestian et al. [21] designed a screening tool that predicts the likelihood of an individual being in depression or in suicidal state. The authors implied the results of linguistic analysis from an act of writing a suicide note from individuals who had committed suicide and normal individuals to extract features: word count, pronouns, anagrams, emotional words, etc. The classifier reached 79% accuracy in discriminating between subjects. Mundt et al. [19] presented a study of voice acoustic patterns of depressed patients to map a track of disease from the early stage to the treatment stage on the scale of extracted features. These findings suggest that acoustic measures of the patient's voice provide objective procedures to evaluate the depression.

2.5 Surface-language model features in Schizophrenia:
Schizophrenia is a psychological disorder of genetic origin that defect the patient's mental and emotional reflexes. It causes the delusions of hearing, vision, poor thinking, and discoursed language that affect the patient's social as well as professional life.. Elvevag et al. [12] presented the Latent Semantic Analysis (LSA) by evaluating patients based on their free-form verbalizations. Elvevag et al. [13] in another study analyzed the speech transcripts of schizophrenic patients, their relatives, and healthy unrelated individuals for three types of measures: statistical language models, semantic measures and surface features. The results discriminated the patients and non-patients on the basis of surface features and language model features while semantic features are required to discriminate the patients from their healthy relatives. Bitouki et al. [9] had used the automatic emotion recognition approach to diagnose the schizophrenia. They used Mel-Frequency Spectral Coefficients to capture information about expressivity and emotions, computed from the phonemes: stressed vowels, unstressed vowels, and consonants contain in voice samples. These experimental results withstood the significance of voice analysis in diagnosing Schizophrenia.

2.6 Autism Spectrum Disorders, a linguistic disorder:
Autism spectrum disorder is a range of neurodevelopment disorders that is characterized by a 'triad of impairments' in social interaction, behavioural and communication skills. Kanner [22] was among the first to describe the behavioural disorders in autistic spectrum related to voice and language. Kanner [22] suggested that in lack of simple confirmatory test, ASD can be diagnosed by a set of psychological and physical assessments. Many studies have related ASD to the language disorders, incoherent discourse, prosody and communication problems. ASD is characterized with certain particular prosody. Prosody-Voice Screening Profile (PVSP) is a standardized screening method for studying prosodic deficits in individuals with High-Functioning Autism (HFA) and Asperger Syndrome (AS) [23]. Diehl et al. [11] use Praat [10] to extract fundamental frequency (f0) and showed that HFA individuals had a higher average standard deviation in f0 than controls. Le Normand et al. [17] study prominence and prosodic contours in different types of speech acts for visualizing the sound in the Praat editor [10]. Van Santen et al. [24] correlated the acoustic measures: lexical stress, focus and phrasing with diagnostic characteristics. Hoque [16] proved the variations in voice parameters of ASD patients.

2.7 Emotional classification of Cancer patients:
Medical sciences termed Cancer as malignant neoplasm, is a wide group of diseases which involve unregulated cell growth forming tumors and later conquering nearby regions of the body. Many researchers studied the presence and extent of emotional expressions in the cancer patients' voice. Oxman et al. [20] were earlier among those who attempted to diagnose cancer patients using automatic textual analysis. The authors found that the lexical lookup method provides the best predictions of diagnostic classification. Automatic Speech Recognition (ASR) has been verified as an effective tool for evaluating intelligibility in patients suffering from cancer of the head and neck. In 2010, Maier et al. found that patient groups showed significantly lower word recognition rates than the control group. ASR significantly correlated word recognition rates with experts' evaluation of intelligibility. Thus it can be concluded that ASR serves as the best means for quantifying the aspects of pathologic speech. Recently, Graves et al [15] also discovered vocal differences in breast cancer survivors in terms of emotional expressions.

III.   Voice Analysis Techniques
The complex nature of voice signal that it varies in time makes it difficult to analyze. It also poses a big challenge for scientists in creating an efficient voice analysis system. That's why most of the voice analysis techniques designed so far aim at extracting the time-varying features of voice signal to simplify the steps of evaluation, decomposition or modification of the signal [40]. Voice signal contains voiced segments that correspond to periodicity and signal's spectrum energy and unvoiced segments of the signal contain non periodic components [44]. An ideal representation of voice signal would be a transformed signal or a set of signals or a set of features in reference to the original signal such that all useful information can be presented in

more predictable and comprehensive form. The information present in voice signal can be extracted in the form of features computed by using algorithms or techniques comprising of basic mathematical functions. The voice signal can fundamentally be analysed in the Frequency Domain and Time Domain.

3.1 Short-time window analysis of voice signal:

Short-time window analysis of voice signal relies on capturing the time varying properties of voice signal. In this technique, voice is analysed for a small interval of time window for which signal properties remain constant. The extracted parameters are computed multiple times from distinct time windows and results are then averaged to achieve high accuracy and for studying the effect of time on the selected features. Window function is represented by $W(n)·S(n)$, where $S(n)$ is voice signal, $W(n)$ corresponds to the region of interest in signal spectra [42]. The shape, duration and size of the window depend on the features to be analyzed in the desired application. The window size is preferably taken as small as possible to avoid random noise in the signal [43]. Table 2 represents the basic window functions used commonly [41]. The most commonly used window is rectangular window (1) that has a high frequency resolution, equal weighing functions, but high band leakage causing noise at the output. Better temporal resolution is obtained by using short window and high frequency resolution using a longer window [42]. Hence a suitable size of the window must be chosen to represent the harmonic structure accurately. The rectangular window can be represented as:

$$w_R(n) = \begin{cases} 1 & \text{if } n = 0,\ldots,R-1, \\ 0 & \text{elsewhere,} \end{cases} \tag{1}$$

Table 2: Basic Window functions.

Window name	w(n) in (R-1⁄2)≤n≤(R-1⁄2)	Main-lobe width, π/R	Side-lobe Level dB
Rectangular	1	4	-13.3
Hann	$1/2(1+\cos(2\pi n/R))$	8	-31.5
Hamming	$.54+.46\cos2\pi n/R$	8	-42.7
Blackman	$.42+.5\cos2\pi n/R +.08\cos 4\pi n/R$	12	-58.1

3.2 Voice analysis via time- domain parameters:

Time domain analysis of voice signal is defined as the transform of voice signal into a set of parameters that show a negligible transition in time to be analysed easily [45]. Fig.1 (a.) represents the time domain analysis of voice as amplitude vs time realized by using Fourier analysis technique [47]. The time domain methodology is classified include Zero-crossing Rate technique and Short-Time Auto correlation technique. Zero-crossing Rate technique extracts spectral information. Zero crossing can occur when the waveform crosses the time axis or changes the algebraic sign [46]. Short-Time Auto correlation of the time signal is the inverse Fourier transform of the energy spectrum that contains the information of periodicity, harmonics and amplitude [46].Time domain representation of voice signal (2) in terms of weighted function is:

$$P(n) = \sum_{m=-\infty}^{\infty} T[s(m)]\, w(f-m), \tag{2}$$

where $S(n)$ is the voice signal, $W(f)$ is weighted function, $P(n)$ represents smaller phonetic time units called phonemes and amplitude variations across two phonemes correspond to pauses in voice.
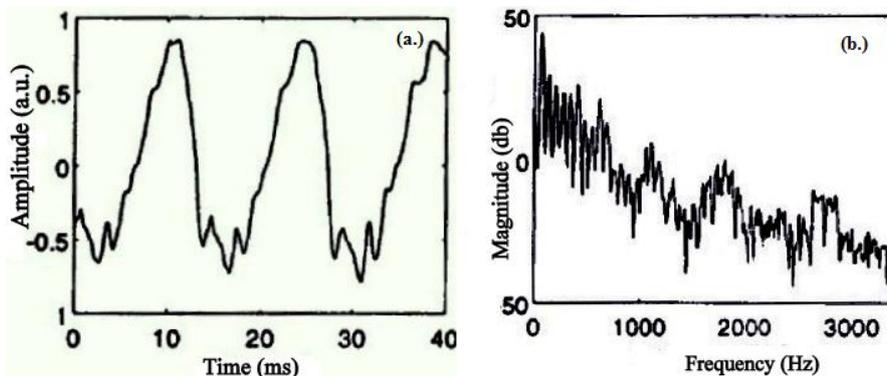

Fig.1: (a.) The time domain analysis of voice signal, (b.) the frequency domain analysis of voice signal.

3.3 Voice analysis via frequency-domain parameters:

Frequency domain represents energy components of signal spectrum, so frequency domain aspects of voice are more significant in terms of information features than its phase or timing aspects(3). Fig. 1(b.) shows frequency domain represented by energy vs frequency [47]. The frequency domain parameters can be extracted

from Filter bank Analysis and Short-Time Fourier Transform Analysis. Filter bank Analysis employs a set of band pass filters for representing the spectral distribution of energy in the spectral envelope [48]. Short-Time Fourier Transform Analysis represents the voice in terms of amplitude and phase as a function of frequency. The Fourier transform of a voice signal S(n) is defined as: $S_n(e^{jw}) = \sum_{m=-\infty}^{\infty} S(m)e^{-jwn}W(n-m)$, (3)

where W(n) represents a low pass filter, $S_n(e^{jw})$ denotes amplitude and phase of the input voice [10].

3.4   Linear predictive coding [LPC] technique:

Linear predictive coding is an efficient and widely accepted means to achieve voice analysis [50]. LPC establishes the short-time correlation in the voice samples by using short order filter [42]. The LPC is used to estimate the voice harmonics, vocal tract function, frequency and bandwidth of the signal. Estimation of voice features from a sample is based on a linear combination of its previous samples (4). The LPC implements the design of Vocoder (an analysis scheme where the spectrum of a source signal is weighted by the spectral components) [49]. Spectral envelope extraction achieved by LPC analysis is very accurate. Thus, LPC is useful in the stationary spectrum representation. The steady state function [42] of a system can be represented as:

$$H(z)= \hat{s}(n) / U(z) \text{ ,and } \hat{s}(n)=\sum_{k=1}^{p}a_k\hat{s}(n-k)+G\sum_{l=0}^{q}b_l u(n-l), \text{ therefore } H(z)=G((1+\sum_{l=0}^{q}b_l z^{-1}) / (1-\sum_{k=1}^{p}b_l z^{-1})), \quad (4)$$

where ŝ(n) is synthesised speech signal, G is gain factor; p, q denotes poles and zeros.

The LPC can be applied to signal in two ways: least-square autocorrelation and least-square covariance. The least-square autocorrelation technique minimizes the mean energy in the error signal in a sample frame by analyzing the voice signal in the time limited window. In the least-square covariance method the error signal e(n) is windowed instead of the input speech signal S(n). Pitch prediction and detection is an advance stage of LPC analysis achieved by introducing some considerable variations in the spectrum. The residual signal will have long-term correlations with the voiced region of the speech. Hence, the residual signal is spectrally flatten in second stage of prediction. The window size is large enough to accommodate 16 samples to 160 samples. Pitch predictor can be represented as: $P(z)=(1/1-\sum b_j z^{-(j+T)})$, T is the pitch period, b_j is the pitch gain [42]. Pitch symbolises the fundamental frequency of a signal during the analysis of the voice signal. The pitch can be calculated either from the periodicity in time domain or from the regularly spaced harmonics in the frequency domain.

IV.   Commercial Biomedical Voice Analysis Tools

The scenario of research in voice related applications today has become much vast owing to advances in mathematical tools (algorithms) and computers. The voice processing technology holds limitless potential in modern biomedical and communication systems. That's why the voice processing has been viewed as the main driving force behind many new innovations and researches in medical and computing engineering like digital signal processors (DSPs).

4.1   Spectral Analysis Systems :

Spectral Analysis Systems are used for determining the acoustic structure of the human voice. Traditional acoustic analysis tools are based on electromechanical hardware platform, but nowadays most of the tools employed in speech pathology clinics are software based.

4.1.1   Frequency Analyzer System invented by Harry R. Foster, East Orange, and Elmo E. Crump [51], at Ohmega Laboratories, is a sound spectrograph device, protected under U. S. Patent 2615078A. The Sona-Graph is the commercial form of Frequency Analyzer System manufactured by KayPENTAX , a benchmark acoustic analysis tool used in the diverse fields of medicine, speech pathology, linguistics, bioacoustics, forensics, and military intelligence.

4.1.2   The Visi-Pitch is a widely used clinical tool designed by KayPENTAX for speech-language pathology. Visi-PitchIV provides the high-fidelity hardware, robust data acquisition, speech therapy tools, efficient user interface and latest features for extracting pitch parameters and spectral characteristics. It has Spectrogram analysis tool, Multi-Dimensional Voice Program, Waveform Editor and Auditory Feedback Tools. Sona-SpeechII is the software prototype of Visi-PitchIV, developed by KayPENTAX.

4.1.3   The Spectrum Analyser (Freq51) is acoustic software tool for biomedical signal analysis. It states real time power spectrum of signal as a function of frequency with multiple settings for sampling, Fourier transforms and time averaging.

4.1.4   Spectrogram (v5.1.7) is a spectral analysis Windows program that can produce spectrograms from real time voice, recorded biological sounds and instrumental music.

4.1.5   Cool Edit 2000 developed by David Johnston is useful software tool in acoustic and music applications as it provides a wide range of features for recording, editing, and filtering and spectrogram presentation of voice data.

4.1.6   Praat is an efficient computer software package for the analysis of speech in phonetics. It is developed by Paul Boersma and David Weenink on 1995 [10]. It can run on a wide range of operating systems including, Linux, UNIX, Mac and Microsoft Windows. Its latest version is Praat 5.3.39.

4.1.7   The Matlab also offers many efficient tools for analyzing voice data as: VOICEBOX (written by Mike Brookes is a voice analysis and processing toolbox), GLOAT (designed by Thomas Drugman, provides glottal analysis of voice), COLEA (records voice directly into Matlab).

4.2   Speech synthesizing voice systems:

Speech synthesizing voice systems are used for producing human speech artificially on a computer platform called a speech synthesizer. The modern Speech Synthesizing system can convert normal text/symbolic linguistics into speech, Such systems aid the people with dyslexia, severe speech impairment, reading difficulties and pre-literate children.

4.2.1   The earliest known speech synthesizer systems are VOCODER and VODER invented by Homer W. Dudley at Bell Labs. Dudley and Bell Labs were granted U.S.Patent 2121142 on June,1938 for the system under the name "System for the artificial production of vocal or other sounds" [52].

4.2.2   Speech-generating devices (SGDs), a kind of Speech synthesizing systems are used to supplement or replace speech for individuals with severe speech impairments, by enabling the verbal communication. Auditory Sciences, Dynavox Mayer-Johnson, LC Technologies, Tobii Technology, Prentke Romich Company, Words, Text Speak Design and Saltillo Corporation are main companies that produce SGD.

4.2.3   Text-to-speech technology (TTS) offers synthesis of speech from text format and many companies offering TTS APIs to their customers include: iSpeech.org, Nuance, AT&T, Ivona, Apple, Microsoft, and Google Android.

4.3   Voice Stress Analysis systems:

Voice Stress Analysis (VSA) systems are designed to record psycho physiological stress responses that are present in a human voice. Voice Stress Analysis technology is based on the fact that micro-muscle tremors (MMT) occur in the muscles of the vocal tract of a person under stress. These muscular tremors are transmitted through the voice and could be analyzed easily.

4.3.1   Psychological Stress Evaluator (PSE) is the earliest known voice stress analyser tool invented by Allan D. Bell *et al.*, is defined as a truth verification device or lie detector. PSE is manufactured by Dektor Counterintelligence and Security Inc. protected under U.S. Patent 3971034 A [53].

4.3.2   Computer Voice Stress Analyzer II (CVSA) is manufactured by National Institute for Truth Verification Federal Services (NITV FS). CVSA is equipped with Voice Imaging Technology, Report Auto-Write and Final Analysis Confirmation Tool (FACT), protected by the U.S. Patent 7,321,855 and 7,571,101 used by local and state federal agencies, as well as by the US Military and Intelligence units.

4.4   Voice recognition systems :

Voice recognition systems find the identity of speaker by analyzing certain features in voice. Such systems are employed by medical professionals in creating medical reports, correspondence, documentation, secure and effective exchange of patient data such as: initial evaluations, daily notes, lab results, patient reports, discharge information, referral notes, visit summaries and e-health record systems.

4.4.1   Dragon Medical 360 voice recognition system by Nuance Healthcare protected under U.S. Patent No. 7,274,775 is best known voice recognition solution in healthcare today.

4.4.2   M*Modal Fluency Direct of M*Modal IP LLC, (formerly known as Medquist Holdings, Inc.) -enables EHRs, clinical documentation systems, radiology information systems (RIS), and picture archive and communication systems (PACS) by voice command.

V.   Recent Advancements

The recent developments in biomedical voice analysis systems include Portable speech therapy device, invented by Fred Evangelist in 2003 [54] (U.S.Patent 6623273). It enables self monitoring of voice exercises for cancer patients. In 2004 J.Masteron, *et al*. invented the Speech analysis therapy system (U.S.Patent 6732076) [55], that presents a methodology for the assessment of the pathological voice. In 2005, M. Rothenberg [56] invented the System for measuring velar function during speech (U.S. Patent 6850882).The device deals with the acoustic analysis of voice by considering different spectral features. In 2007, J. Meyerhoff and J.H.L. Hansen [57] presented the Methods and systems for detecting, measuring and monitoring stress in speech under

the U.S.Patent 7283962. Later, getting financial assistance from the US government in 2008, M. Rosalyn *et al.* [58] published the U.S.Patent 7457753, for the telephone pathology assessment, enabling the remote evaluation of patients with vocal impairments. In 2011 V.E. Mikhajlovich *et al.* [59] proposed a Method for detection in speech pathology under patent RU2433488. The method divides the spectrum in sets of pair of low-frequency harmonics, and high frequency harmonics for calculating ratio of the total energy about high overtones to the total energy about low-frequency harmonic, then these coefficients of harmonics for healthy and diseased are compared. Recently, in 2013, Tao Zhi *et al.* [60] published Patent CN103258545A by the name of the Pathological voice subdivision method, in which values of pathological voices are set precisely to enhance pre-diagnosis and recovery measuring tools implemented in vocal diseases.

VI.   Conclusion

The purpose of this paper was to state the significance of voice analysis systems in the prediagnosis of certain medical conditions which later may transform into fatal or incurable diseases. This paper had presented the various voice analysis techniques that are implemented for tracking characteristic variations in voice patterns of patients. The research papers reviewed, verify the success of the voice analysis techniques in the practical environments. But each of these approaches is not suitable for mapping all types of dysphonia patterns. So there is a need of a standard voice analysis tool and methods for clinical assessment, and develop a code of clinical practice. Moreover, a new methodology that ensures high accuracy, patients ease, economical and less time consuming in predicting symptoms at a very early stage is the need of the hour. Although, this elementary area of biomedical engineering, which deals with a variety of procedures for disease-related voice diagnosis, is now attracting the attention of many researchers and health care giants. Thus we can conclude that the beginning of the third millennium opens new horizons for the advancement of voice diagnosis systems. As it seems now, the process of evolution of biomedical technology has already begun and the way it is nurturing the humankind, looks very promising for the future generations.

References

[1]   Kempelen W.V., "Le mechanisme de la parole, Suivi de la description d'une machine parlante", J.V.Degen, 1791.
[2]   Wheatstone C., "The scientific papers of Sir Charles Wheatstone Taylor and Francis", London, 1879.
[3]   Sakar B.E., Isenkul M.E., Sakar C.O., Sertbas A., Gurgen F., Delil S., Apaydin H., Kursun O., "Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings", IEEE Journal of Biomedical and Health informatics, 17(4), 828-34, 2013.
[4]   Lee L., Chamberlain L.G., Loudon R.G., Stemple J.C., " Speech segment durations produce by healthy and asthmatic subjects" , J Speech Hear Disorder, 53(2), 186-193, May 1988.
[5]   Lee L., Loudon R.G., Jacobson B.H., Stuebing R., "Speech breathing in patients with lung disease", Pub Med, 147(5), 1199-1206, May 1993.
[6]   Alpert M. et al, "Reflections of depression in acoustic measures of the patient's speech", Journal of Affective Disorders, 66, 59–69, 2001.
[7]   Ang J. et al, "Prosody-based automatic detection of annoyance and frustration in human-computer dialog", ICSLP, 2002.
[8]   Boyanov B., Hadjitodorov S., "Acoustic analysis of pathological voices: a voice analysis system for screening of laryngeal diseases", Proc. IEEE Engineering in Medical and Biology, 16 (4), 74-82, 1997.
[9]   Bitouk D. et al. (2009), "Improving Emotion Recognition using Class-Level Spectral Features", Interspeech, 2009.
[10]   Boersma P., Weenink D., "PRAAT: Doing phonetics by computer", Version 4.3.14, Computer program, 2005.
[11]   Diehl J. et al, "An acoustic analysis of prosody in high-functioning autism", Applied Psycholinguistics, 30(3), 2009.
[12]   Elvevaag B., Foltz P., Weinberger D., Goldberg T., "Quantifying Incoherence in Speech: an Automated Methodology and Novel Application to Schizophrenia", Schizophrenia Research, 93, 304–316, 2007.
[13]   Elvevaag B., Foltz P., Rosenstein M., DeLisi L., "An automated method to analyze language use in patients with schizophrenia and their first degree-relatives", Journal of Neurolinguistics, 2009.
[14]   Lee C.M., Narayanan S., "Towards detecting emotions in spoken dialogs", IEEE Transaction on Speech and Audio Processing, 2004.
[15]   Graves K. et al., "Emotional expression and emotional recognition in breast cancer survivors: A controlled comparison", Psychology and Health, 20, 579–595, 2005.
[16]   M. E. Hoque et al. (2009), "Exploring Speech Therapy Games with Children on the Autism Spectrum", Interspeech, 2009.
[17]   Normand M.L. et al, "Prosodic disturbances in autistic children speaking French, Speech Prosody," Campinas, 2008.
[18]   Moore E. II, Clements M., Peifer J., Weisser L., "Investigating the Role of Glottal Features in Classifying Clinical Depression", IEEE EMBS, Cancun, 2003.
[19]   Mundt, J. et al "Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology," Journal of Neurolinguistics,20(1), 50–64, 2007.
[20]   Oxman T., Rosenberg S., Schurr P., Tucker G., "Diagnostic Classification Through Content Analysis of Patient Speech," American Journal of Psychiatry, 145, 464–468, 1988.
[21]   Pestian J., Matykiewicz P., Grupp-Phelan J., Lavanier S. A., Combs J., Kowatch R., "Using Natural Language Processing to Classify Suicide Notes", ACL BioNLP Workshop, 96–97, 2008.
[22]   Kanner L., "Irrelevant and metaphorical language in early infantile autism", American Journal of Psychiatry, 103, 242–246, 1946.
[23]   Shriberg L. et al, "Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome," Journal of Speech, Language, and Hearing Research, 44(5), 2001.
[24]   Santen J.V. et al, "Automated assessment of prosody production", Speech Communication, 5, 1082–1097, 2009.
[25]   Bucks R., Singh S., Cuerden J.M., Wilcock G., "Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analyzing lexical performance", Aphasiology, 14(1), 71-91, 2000.
[26]   A. P. Association, "Diagnostic and Statistical Manual of Mental Disorders", 4th ed., 1994.

[27] Cummings J., Benson F., Hill M., Read S., "Aphasia in dementia of the Alzheimer type", Neurology, 35, 394-397, 1985.
[28] Forbes K., Venneri A., Shanks M., "Distinct patterns of spontaneous speech deterioration: an early predictor of Alzheimer's disease", Brain and Cognition, 48(2-3), 356-361, 2002.
[29] Folstein M., Folstein S., McHugh P., "Mini-mental state. a practical method for grading the cognitive state of patients for the clinician", Journal of Psychiatric Research, 12, 189-198, 1975.
[30] R. Arefi Shirvan, E. Tahami. "Voice Analysis for Detecting Parkinson's Disease Using Genetic Algorithm and KNN", IEEE Conference on BioMedical Engineering, 978-1-4673-1005-5/11, December 2011.
[31] Jankovic J. "Parkinson's disease: clinical features and diagnosis" ,J. Neurol. Neurosurg. Psychiatr. 79 (4): 368–76., April 2008.
[32] Canter, G.J., "Speech characteristics of patients with Parkinson's disease:I. Intensity, pitch, and duration", Journal of Speech and Hearing Disorders, 28(3), 221–229, 1963.
[33] Canter, G.J., "Speech characteristics of patients with Parkinson's disease: II. Physiological support for speech", Journal of Speech and Hearing Disorders, 30(1), 44–49, 1965.
[34] Canter, G.J., "Speech characteristics of patients with Parkinson's disease: III. Articulation, diadochokinesis and over-all speech adequacy", Journal of Speech and Hearing Disorders, 30(3), 217–224, 1965.
[35] Skodda, Sabine; Rinsche, Heiko; Schlegel, Uwe, "Progression of dysprosody in Parkinson's disease over time-A longitudinal study", Movement Disorders, 24 (5), 716–22, April 2009.
[36] Hammen, V.L., & Yorkston, K.M., "Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria", Journal of Communication Disorders, 29, 429–444, 1996.
[37] A. Ho, R. Iansek, C. Marigliani, J. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," Behav. Neurol., 11, 131–137, 1998.
[38] B. Harel, M. Cannizzaro, and P. J. Snyder, "Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease:A longitudinal case study," Brain Cognition, 56, 24–29, 2004.
[39] LeWitt P. A., "Parkinson's disease: Etiologic considerations", Parkinson's disease and movement disorders: Diagnosis and treatment guidelines for the practicing physician in C. H. Adler & J. E. Ahlskog (Eds.), 91-101, Totowa, NJ: Humana Press, 2000.
[40] Yegnanarayana B., Veldhuis R.N.J., "Extraction of vocal-tract system characteristics from speech signals", IEEE Transactions on Speech and Audio processing, Vol. 6, 313-327, July 4, 1998.
[41] Mitra S. K., " Digital Signal Processing: A computer-Based Approach" ,McGraw-Hill, New York, 1998.
[42] Kondaz A.M., "Digital Speech coding for low bit rate Communication Systems", John Wiley & sons, 1999.
[43] Shaughnessy D.O., "Speech Communication", 2nd edition, University press Ltd, 2001.
[44] Gold B., Morgan N., "Speech and Audio Signal Processing", John Wiley & Sons, 2006.
[45] Paulraj M.P., Yaacob S., Hariharan M., "Diagnosis of vocal fold pathology using time-domain features and systole activated neural network", IEEE 5th International Colloquium on Signal Processing & Its Applications, 29 – 32, March 2009.
[46] Yang X., Tan B., Ding J., Zhang J., Gong J.,"Comparative Study on Voice Activity Detection Algorithm" , IEEE International Conference on Electrical and Control Engineering (ICECE), 599 – 602, June 2010.
[47] Moulines E., Verhelst W., "Time-Domain and Frequency-Domain Techniques for Prosodic Modification of Speech", Elsevier Science B.V., Speech Coding and Synthesis, 1995.
[48] Allen J. B., Rabiner L. R., "A unified approach to short-time Fourier analysis and synthesis", IEEE Proceedings, Vol.:65 , 1558 – 1564, Nov. 1977.
[49] Arfib D., Keiler F., Zolzer U., "Source-filter processing", Digital Audio Effects, 299–372, John Wiley and Sons, Ltd., 2002.
[50] Schroeder M. R., "Computer Speech: Recognition, Compression, and Synthesis", Springer Verlag, 1999.
[51] Foster H.R., Orange E., Crump E.E., "Frequency Analyzer System", U. S. Patent 2615078A, October 21, 1952.
[52] Dudley H.W., Riesz R., "System for the artificial production of vocal or other sounds", U.S. Patent 2121142, June, 1938.
[53] Bell A.D., Ford W.H., McQuiston C.R., "Psychological Stress Evaluator (PSE)", U.S. Patent 3971034 A.
[54] Evangelist F., "Portable speech therapy", U.S. Patent 6623273, September 23, 2003.
[55] Masteron J., Bernhardt B., Albert V.S., "Speech analysis and therapy system and method", U.S. Patent 6732076, May 4, 2004.
[56] Rothenberg M., "System for measuring velar function during speech", U.S. Patent 6850882, February 2, 2005.
[57] Meyerhoff J., Hansen J.H.L., "Methods and systems for detecting, measuring and monitoring stress in speech", U.S. Patent 7283962, August 13, 2007.
[58] Rosalyn M., Reilly R., Chazal P.D., Mullane B. O., Lacy P., "Telephone pathology assessment", U.S. Patent 7457753, November 25, 2008.
[59] Mikhajlovich V.E., Stanislavocich D.S., Viktorovich L.D., Karpovich M.K., Vladimirovich M.A., F.O. Gennadevna, "Method of detecting pathology of voice leading speech", RU Patent 2010104610, August 20, 2011.
[60] Zhi T., Qiang Z., Xiaojun Z., Di W., Zhe X. Z., Jingjing J., "Pathological voice subdivision method", CN Patent 103258545A, August 21, 2013.