Volume 6, Issue 3, March 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Study on Application of Digital Signal Processing in Voice Disorder Diagnosis

[1]N. A. Sheela Selvakumari*, [2]Dr. V. Radha
[1] Asst. Prof., Department of Computer Science, PSGR Krishnammal College for Women, Coimbatore, Tamilnadu, India
[2] Professor, Department of Computer Science, Avinashiligam Institute for Home Science and Higher Education for Women, Coimbatore, Tamilnadu, India

Abstract* — The Investigation of the human voice has become an important area of study for its numerous applications in medical as well as industrial disciplines. 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. Voice analysis mainly deals with extraction of some parameters from voice signals, which is used to process the voice in appropriate for the particular application by using suitable techniques. The use of these techniques combined with classification methods provides the development of expert aided systems for the detection of voice pathologies. This Study paper states the certain common medical conditions which affect voice patterns of patients and the tests which are used as diagnosis for voice disorder.

***Keywords*— Digital Signal Processing, Voice Disorder, Pattern Classification, Pathological Voices**

I. INTRODUCTION

Digital signal processing (**DSP**) is the numerical manipulation of signals, usually with the intention to Measure, Filter, Produce or Compress continuous Analog Signals. It is characterized by the use of digital signals to represent these signals as Discrete Time, Discrete Frequency, or other Discrete Domain Signals in the form of a sequence of numbers or symbols to permit the digital processing of these signals.

Theoretical analyses and derivations are typically performed on discrete-time signal models, created by the abstract process of sampling. Numerical methods require a digital signal, such as those produced by an Analog-to-Digital Converter (ADC). The processed result might be a frequency spectrum or a set of statistics. But often it is another digital signal that is converted back to analog form by a Digital-to-Analog Converter (DAC). Even if that whole sequence is more complex than analog processing and has a discrete value range, the application of computational power to signal processing allows for many advantages over analog processing in many applications, such as error detection and correction in transmission as well as data compression.

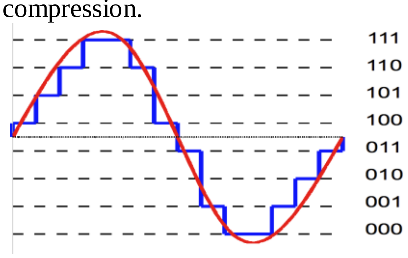

Fig. 1- Analog to Digital Conversion

As mentioned in [1], Digital Signal Processing and Analog Signal Processing are subfields of Signal Processing. DSP applications include Audio and Speech Signal Processing, Sonar and Radar Signal Processing, Sensor Array Processing, Spectral Estimation, Statistical Signal Processing, Digital Image Processing, Signal Processing for Communications, Control of Systems, Biomedical Signal Processing, and Seismic Data Processing. DSP algorithms have long been run on standard computers, as well as on specialized processors called Digital Signal Processors, and on purpose-built hardware such as Application-Specific Integrated Circuit (ASICs).

Currently, there are additional technologies used for Digital Signal Processing, including more powerful general purpose Microprocessors, Field-Programmable Gate Arrays (FPGAs), Digital Signal and Stream Processors. Digital signal processing can involve linear or nonlinear operations. Nonlinear signal processing is closely related to nonlinear system identification and can be implemented in the Time, Frequency, and Spatio-temporal domains.

This paper established in cue with the Signal Sampling techniques in section (II), and Domains used by the Digital Signal Processing Techniques were discussed in the section (III). In section (IV) the Digital Signal Processing Applications were discussed. The implementation of Digital Signal Processing was discussed in section (V) and in section (VI) the Test and Diagnosis for Voice Disorder was discussed. Finally, the Conditions which cause Voice Disorders were discussed in section (VII) and Conclusion is noted in the section (VIII).

II. SIGNAL SAMPLING

The increasing use of computers has resulted in the increased use of, and need for, Digital Signal Processing. To digitally analyze and manipulate an Analog Signal, it must be digitized with an Analog-to-Digital converter. According to [3], Sampling is usually carried out in two stages, Discretization and Quantization. In the Discretization stage, the space of signals is partitioned into equivalence classes and quantization is carried out by replacing the signal with representative signal of the corresponding equivalence class. In the Quantization stage, the representative signal values are approximated by values from a finite set.

The Nyquist–Shannon sampling theorem states that a signal can be exactly reconstructed from its samples if the sampling frequency is greater than twice the highest frequency of the signal, but this requires an infinite number of samples. In practice, the sampling frequency is often significantly higher than twice that required by the signal's limited bandwidth.

Some continuous-time periodic signals become non-periodic after sampling, and some non-periodic signals become periodic after sampling. In general, for a periodic signal with period T to be periodic with N after sampling with sampling interval T_s, the following must be satisfied:

$$T_s N = Tk \qquad \text{….Equ(1)}$$

where *k is an integer.*

III. DOMAINS USING DIGITAL SIGNAL PROCESSING

In DSP, engineers usually study digital signals in one of the following domains: time domain, spatial domain, frequency domain, and wavelet domains. They choose the domain in which to process a signal by making an informed assumption as to which domain best represents the essential characteristics of the signal. A sequence of samples from a measuring device produces a temporal or spatial domain representation, whereas a discrete Fourier transform produces the frequency domain information, that is, the frequency spectrum. Autocorrelation is defined as the cross-correlation of the signal with itself over varying intervals of time or space.

A. Time and Space Domains

The most common processing approach in the time or space domain is enhancement of the input signal through a method called filtering. Digital filtering generally consists of some linear transformation of a number of surrounding samples around the current sample of the input or output signal. There are various ways to characterize filters, Such as:

- A "Linear" filter is a linear transformation of input samples; other filters are "non-linear". Linear filters satisfy the superposition condition, i.e. if an input is a weighted linear combination of different signals, the output is a similarly weighted linear combination of the corresponding output signals.
- A "Causal" filter uses only previous samples of the input or output signals; while a "non-causal" filter uses future input samples. A non-causal filter can usually be changed into a causal filter by adding a delay to it.
- A "Time-Invariant" filter has constant properties over time; other filters such as adaptive filters change in time.
- A "Stable" filter produces an output that converges to a constant value with time, or remains bounded within a finite interval. An "unstable" filter can produce an output that grows without bounds, with bounded or even zero input.
- A "Finite Impulse Response" (FIR) filter uses only the input signals, while an "infinite impulse response" filter (IIR) uses both the input signal and previous samples of the output signal. FIR filters are always stable, while IIR filters may be unstable.

A filter can be represented by a block diagram, which can then be used to derive a sample processing algorithm to implement the filter with hardware instructions. The output of a linear digital filter to any given input may be calculated by convolving the input signal with the impulse response.

B. Frequency Domains

Signals are converted from time or space domain to the frequency domain usually through the Fourier transform. The Fourier transform converts the signal information to a magnitude and phase component of each frequency. Often the Fourier transform is converted to the power spectrum, which is the magnitude of each frequency component squared. The most common purpose for analysis of signals in the frequency domain is analysis of signal properties. The engineer can study the spectrum to determine which frequencies are present in the input signal and which are missing.

In addition to frequency information, phase information is often needed. This can be obtained from the Fourier transform. With some applications, how the phase varies with frequency can be a significant consideration. Filtering, particularly in non-real time work can also be achieved by converting to the frequency domain, applying the filter and then converting back to the time domain. This is a fast, O(n log n) operation, and can give essentially any filter shape including excellent approximations to brick wall filters.

There are some commonly used frequency domain transformations. For example, the Cepstrum converts a signal to the frequency domain through Fourier transform, takes the logarithm, then applies another Fourier transform. This emphasizes the harmonic structure of the original spectrum.

Frequency domain analysis is also called Spectrum- or Spectral Analysis.

C. Z-Plane Analysis

Whereas analog filters are usually analyzed in terms of transfer functions in the s plane using Laplace transforms, digital filters are analyzed in the z plane in terms of Z-transforms. A digital filter may be described in the z plane by its characteristic collection of zeroes and poles. The z plane provides a means for mapping digital frequency to real and imaginary z components,

where $\mathbf{Z = re^{jw}}$ for continuous periodic signals and

$\mathbf{w=2\pi F}$ (**F** is the digital frequency).

This is useful for providing a visualization of the frequency response of a digital system or signal.

D. Discrete Wavelet Transformation

In numerical analysis and functional analysis, a Discrete Wavelet Transform (DWT) is any Wavelet Transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information.

IV. APPLICATIONS

The main applications of DSP are audio signal processing, audio compression, digital image processing, compression, speech, speech recognition, digital communications, radar, sonar, financial signal processing, seismology and biomedicine. Specific examples are speech compression and transmission in digital mobile phones, room correction of sound in hi-fi and sound reinforcement applications, weather forecasting, economic forecasting, seismic data processing, analysis and control of industrial processes, medical imaging such as CAT scans and MRI, MP3 compression, graphics image, Hi-Fi loud speaker crossovers and equalization, and audio effects for use with electric guitar amplifiers.

A. Application of Digital Signal Processing in Voice Disorder Diagnosis

The generation of vocal sounds is generally referred to as "phonation," while the action of generating word sounds is referred to as "speech" or "articulation." The organs involved in phonation and/or speech are the oral cavity, nasal cavity, pharynx, larynx, trachea, bronchus, lungs, thorax, and diaphragm; these multiple organs work in a coordinated manner to perform complex integrated movements.

According to [5] Speech Signal mainly deals with extraction of features from voice signals and combined with the pattern classification methods provides the expert systems for the detection of voice pathologies shown in Fig.2.

In the phonation mechanism, in response to a command to vocalize from the cerebral cortex, the respiratory muscles contract and expiratory flow from the lungs is moved upwards towards the trachea and larynx as a power source, while at the same time both vocal cords are adducted through the recurrent laryngeal nerve, closing the glottis.

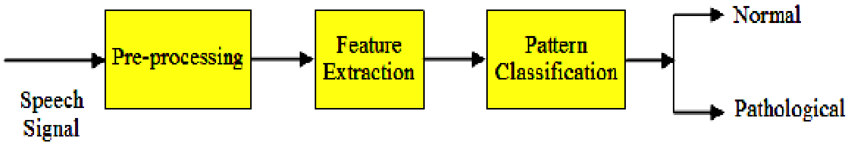

Fig. 2 - Finding the Voice Disorder level based on the Speech Signal

The expiratory flow raises the sub glottal pressure, causing the vocal cords to vibrate and generate sound, which passes through the vocal tract (which acts as a resonance chamber) to produce vocal sound. Consonants and vowels are articulated, becoming speech, and are generated continuously to produce spoken words. If any part of these phonation control or motion mechanisms becomes impaired, a voice disorder occurs.1) The clinical state of voice disorders can be classified from phonation mechanisms as follows: 1) glottal closure disorder; 2) affected vocal cord stiffness; 3) vocal cord asymmetry; 4) respiration/ resonance chamber disorder; and 5) psychological factors. Of these, 1) to 3) are abnormalities in the shape and motility of the larynx and are the main causes of voice disorders. When examining patients who are complaining of voice disorders, first of all they are asked about their main complaints, medical history, degree and quality of hoarseness, past history, occupation, and daily lifestyle habits and social background related to phonation, and possible causes of the voice disorder are estimated.

A physician who is a voice specialist can estimate the patient's condition just by listening to their voice. Furthermore, performing indirect laryngoscope or laryngeal endoscopy enables the diagnosis of many laryngeal disorders. Initial responses to patients by general practitioners differ according to whether the disorder they have is benign or malignant, acute or chronic.

Although it is thought that patients with benign disorders such as acute Corditis Vocalis with a common cold are in many cases initially treated at internal medicine clinics, in cases where there is a high degree of hoarseness and the hoarseness has not improved in two weeks or more, the patient is referred to a physician specializing in ear, nose, and throat disorders. Loosely examining a patient without looking at the larynx can result in malignancies such as laryngeal cancer and thyroid cancer being overlooked. Of course, patients with acute epiglottitis or other airway stenotic disorders should be referred to a specialist physician urgently.

B. Diagnosis for Voice Disorders

With regard to the patient's chief complaints, they are asked about how phonation has been impaired. It is important to obtain a present illness including the time since onset of the symptoms and treatment at other hospitals. With regard to contributory factors, the patient is asked about voice misuse, past operations on the larynx, and past operations for which the patient was under general anaesthesia. If the patient has experienced symptoms of heartburn, acid reflux, or reflux esophagitis in the past, the possibility of laryngeal granuloma can be considered. If the patient complains of respiratory difficulties, there is the possibility of airway stenosis and the patient is referred to a specialist. With regard to occupation, misuse the voice and vocal cord nodules are easily formed in cases with vocal abuse. Dysphonia plicae ventricularis can be observed amongst Buddhist monks. For restaurant and service industry employees, smoking, drinking, and karaoke singing also exert an effect, with vocal cord polyps and polypoid vocal cords occurring more commonly.

C. Inspection and Palpation

With regard to inspection of the oral cavity and oropharynx, the clinical features of acute inflammation, such as mucosal reddening or adhesion of purulent mucus, are observed. With regard to palpation of the neck, for acute disorders the presence/absence and location of tenderness is checked, and for cases in which a malignant disorder is suspected, special attention is paid to cervical lymphadenopathy and thyroid tumours.

D. Maximum Phonation Time

For Maximum Phonation Time (MPT), the maximum length of time a patient can vocalize after taking a deep breath is measured. In general, 10 seconds or less is abnormal, and 5 seconds or less interferes with daily living. Disorders which shorten MPT include recurrent nerve paralysis and vocal cord atrophy.

E. Auditory-Perceptual Evaluation of Hoarseness

An auditory-perceptual evaluation method for hoarseness is the GRBAS scale of the Japan Society of Logopaedics and Phoniatrics, which gives scores of 0, 1, 2, or 3 for the Grade of hoarseness; Roughness, Breathiness, Asthenia, and Strain, where 0 is normal, 1 is a slight degree, 2 is a medium degree, and 3 is a high degree. Rough hoarseness is a rasping, rattling sounding voice that can be heard mainly in disorders such as vocal cord polyps, polypoid vocal cords, and laryngeal cancer. Breathy hoarseness is a whispery voice that can be heard in such disorders as recurrent nerve paralysis, vocal cord nodules, laryngeal cancer, acute Corditis Vocalis, and vocal cord atrophy. Asthenic hoarseness is a small, weak voice which is heard in such disorders Aspsychosomaticaphonia and myasthenia gravis. Strained hoarseness is produced with the throat constricted; this condition occurs in such disorders as spasmodic voice disorders and laryngeal cancer. In addition, disorders in which the voice becomes muffled include Peritonsillar abscess and acute epiglottitis; these are potentially lethal disorders which cause airway stenosis and must not be overlooked.

V. IMPLEMENTATION

Depending on the requirements of the application, Digital Signal Processing tasks can be implemented on general purpose computers. Often when the processing requirement is not real-time, processing is economically done with an existing general-purpose computer and the signal data (either input or output) exists in data files.

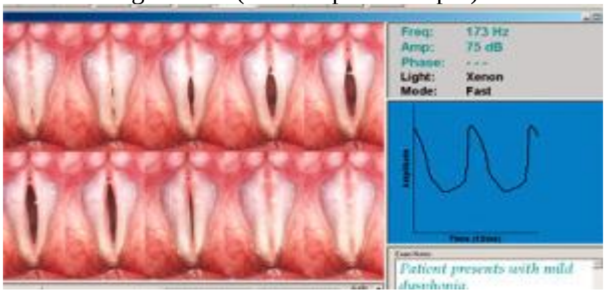

Fig. 3 - Quick Snapshot of the Patient's Vocal Fold Structure and Vibratory Pattern

Such process is same as any other data processing technique, except when Digital Signal Processing used mathematical techniques, such as the FFT, and the sampled data is usually assumed to be uniformly sampled in time or space. For example, processing digital photographs with software such as Photoshop.

Based on [6], when the application requirement is real-time, DSP is often implemented using specialized microprocessors such as the DSP56000, the TMS320, or the SHARC. These often process data using fixed-point arithmetic, though some more powerful versions use floating point. For faster applications FPGAS might be used. Beginning in 2007, multicore implementations of DSP's have started to emerge from companies including Freescale and Stream Processors Inc. For faster applications with vast usage, ASICS might be designed specifically. For slow applications, a traditional slower processor such as a microcontroller may be adequate. Also a growing number of DSP applications are now being implemented on embedded systems using powerful pcs with multi-core processors.

VI. TESTS AND DIAGNOSIS FOR VOICE DISORDER

Based on [2], the tests and the diagnosis methodology used for voice disorder were listed below:

A. Laryngeal Endoscopy and Indirect Laryngoscopy

Larlyngeal endoscopy and indirect laryngoscopy are the most useful tests, and looking at the larynx makes the diagnosis of many laryngeal disorders possible. Stroboscopy is an examination in which vocal cord vibrations are observed and is useful in detecting minute pathological lesions in vocal cords.

B. Tests related to Voice Pitch and Strength

Voice pitch measures Speaking Fundamental Frequency (SFF) and Vocal Range. Voice strength measures the sound pressure level at comfortable phonation as well as at the maximum and minimum strengths of phonation. Voice profile tests and other examinations are available.

C. Aerodynamic Tests

Aerodynamic tests include measurement of MPT, average airflow rate during Phonation, and Larynx Efficiency.

D. Acoustic Analysis Tests

These are tests which enable quantitative evaluation and include parameters such as pitch period perturbation's Amplitude Perturbation, and Laryngeal Noise components as well as Spectrograms.

E. Auditory-Perceptual Evaluation

The GRBAS scale is the main tool used in auditory-perceptual evaluation, which physicians and speech pathologists use to subjectively assess the degree and quality of hoarseness. There is also a voice handicap index which patients use to subjectively evaluate social and lifestyle limitations Such as, functional aspects, voice and larynx condition Such as, physical aspects, and what the patient is feeling Such as, emotional aspects.

VII. CONDITIONS WHICH CAUSE VOICE DISORDERS

The Patient should be referred to a specialist physician, if the hoarseness does not improve in disorders (1) or (2) below, urgently in the case of disorder (3) and swiftly in the case of disorders (4) or (5) based on [4].

A. Vocal Cord Polyps

These are the most common organic disorders of the larynx which cause voice disorders. A common site for polyps is from the front third to the center of the membranous portion of the vocal cord, and in many cases the sides are asymmetrical, regardless if the polyp occurs on one side or both sides of the vocal cords. Polyps are thought to be caused by submucosal bleeding of the vocal cords, and contributing factors include voice misuse and smoking. With regard to hoarseness, in many cases the patient is found to have rough hoarseness or breathy hoarseness. At our institution, the abovementioned specialized tests are carried out and vocal sound before and after treatment is evaluated.

B. Acute Corditis Vocalis

The Larynx becomes inflamed due to a common cold, etc., causing the voice to become whispery and hoarse, in some cases the patient's voice becomes aphonic. There is diffuse reddening and swelling of both vocal cords, and histologically this is regarded as being caused by inflammatory cell invasion of the superficial Lamina Propria, Edema, and Vasodilatation. Due to the rapid swelling of the vocal cords, the mucosa is extended excessively, reducing its mobility; mucosal waves are diminished, with asymmetrical vocal cord vibrations.

C. Acute Epiglottitis

Even when symptoms of acute inflammation such as pharyngeal pain or fever causing a muffled voice are observed, examination of the oral cavity may overlook the inflammation finding, making it extremely important that the larynx also be examined. This is a potentially lethal disorder that causes airway stenosis and must not be overlooked. This disorder is frequently observed in men in the prime of their lives; it can easily cause medical disputes as the disease progresses quickly with the patient dyspnoea, resulting in hypoxic encephalopathy or death.

D. Recurrent Nerve Paralysis

After leaving the ambiguous nucleus of the medulla oblongata as the vagus nerve, the recurrent nerve, which controls the movement of the vocal cords, travels a long distance before reaching the larynx, and may become paralysed due to damage sustained along its various parts. In many cases this is caused by neck or chest disorders, and it may also be caused by tumour's such as thyroid cancer, esophageal cancer, and lung cancer; lymph node metastasis, aortic aneurysm, and complications of surgeries performed to cure these; infectious diseases such as viruses; and tracheal intubation. It may also be caused by idiopathic and intracranial/skull base diseases. Symptoms for unilateral recurrent nerve paralysis include hoarseness and mis-swallowing, while bilateral paralysis causes airway stenosis with respiratory difficulties. The position in which the vocal cords are fixed determines the status of glottal closure and the degree of hoarseness also changes depending on vocal cord position. Although strictly speaking these differ, terms such as laryngeal paralysis, vocal cord paralysis, and vocal cord fixation are also used.

E. Laryngeal Cancer

Laryngeal cancer is divided into Supraglottic, Glottic, and Subglottic types. The Glottic type is the most common, and when the cancer invades the Mucosal Lamina Propria and muscle layer of the vocal cords, it impairs vocal cord vibration, causing hoarseness. In the Supraglottic and Subglottic types, patients are slow to become aware of the hoarseness and detection of the cancer is slower than that for the Glottic type. When the cancer invades the Cricoarytenoid Articulation, the vocal cord becomes fixed, developing incomplete Glottal Closure, causing air leak and breathy hoarseness during vocalization.

VIII. CONCLUSIONS

In the past few decades, medicine has received noteworthy contributions, which have allowed for great advancement in medical activity within various contexts, for example, improvement of surgery techniques, description of the human genome, or even assistance in medical diagnosis. In particular, regarding medical diagnosis, digital signal processing techniques have been employed recently as an efficient, non-invasive, and low-cost tool to analyse vocal signals, with the aim of detecting and classifying alterations in the production of sounds that may be associated with larynx pathologies. The human voice is an important communication tool and any inadequate behaviour may have deep implications in an individual’s social and professional life. The acoustic analysis is a complementary technique to methods based on the direct inspection of vocal folds, which may reduce the frequency of invasive exams, since such measurements are able to reveal important physiological characteristics of the vocal tract. In general, most of the techniques involving the analysis of dysfunctions available in the literature employ several sorts of pre-processing stages in order to extract useful characteristics for the classification of diverted patterns and to obtain performance improvement of classifiers.

REFERENCES

[1] Karplus, Kevin, and Alex Strong. “Digital Synthesis of Plucked-String and Drum Timbres.” Computer Music Journal, V: 7, I: 2, PP: 43-55, 1983.

[2] Jacobson BH, Johnson A, Grywalski C, “The Voice Handicap Index (VHI): Development and Validation”, American Journal of Speech-Language Pathology, V: 6, PP: 66–70, 1997.

[3] Mitra, Sanjit K., Gordana Jovanovic-Dolecek, and Mikhail K. Tchobanou. "On the Design of One-Dimensional Sparse Arrays with Apodized End Elements." IEEE Conference on Signal Processing, PP: 2239-2242, 2004.

[4] Verdolini K., Rosen C., & Branski R., “Classification Manual for Voice Disorders-I”, Psychology Press, Vol. I, 2006.

[5] Patricia Henríquez, “Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics”, IEEE Transactions On Audio, Speech And Language Processing, V: 17, I: 6, 2009.

[6] Lotfi Salhi, Talbi Mourad, and Adnene Cherif, “Voice Disorders Identification Using Multilayer Neural Network”, The International Arab Journal of Information Technology, V: 7, I: 2, 2010.

