

Cepstral Analysis of Speech using Discrete Hartley Transform

Madhukar. B. N.
Senior Assistant Professor, ECE Dept,
New Horizon College of Engineering,
Bangalore, INDIA.

Dr. P. S. Satyanarayana
Senior Professor, ECE Dept,
Don Bosco Institute of Technology,
Bangalore, INDIA.

Abstract —This paper presents Cepstral analysis of speech signal using Discrete Hartley Transform. Hitherto, the cepstral analysis of speech was carried out using in temporal domain and in the frequency domain using Discrete Fourier Transform (DFT) based approach. A new approach of finding the cepstral coefficients in the frequency domain using DHT, rather than using the DFT is proposed. DFT being a complex transform takes more computation time for finding the cepstral coefficients of the speech signal, but DHT being a real transform takes less computation time to do the same with less memory requirement. The relationship between DFT and DHT is made use of for finding the cepstral coefficients, rather than using the DFT directly. The usage of DHT method is found to be optimal than using the direct DFT approach thereby saving implementation cost substantially in the cepstral analysis of speech signals. The analysis is done by comparing the usage of DFT directly on the speech signal and then using DHT, thereby seeing the performance of these mathematical transforms based on computation time which is used as a performance metric for validating the veracity of the two discrete transforms used. This work is implemented using MATLAB R2014a software.

Keywords—DHT, DFT, Cepstrum.

I. INTRODUCTION

The Discrete Hartley Transform (DHT) is a variant of the Discrete Fourier Transform (DFT) which is renowned and is one of the largely used transforms in the Communication Engineering and Signal Processing. Discrete Hartley Transform (DHT) was developed by Ronald N. Bracewell, a famous Australian physicist, engineer, and mathematician, in the 1980s. DHT is a discretized version of the Continuous Hartley Transform (CHT) invented by the U.S. electronics researcher Ralph Vinton Lyon Hartley (1888 A.D. – 1970 A.D.) in 1942 [1]. Though it existed in the technical literature, it remained in the oblivion for a number of years until Bracewell invented and published a discretized version of it in 1983. Hartley also invented the Hartley Oscillator and also contributed towards the development of Information Theory during its stage of infancy. In this paper, Hartley Transform and its discrete variant are reviewed followed by their application to the cepstral analysis of speech signal [2].

II. THE CONTINUOUS HARTLEY TRANSFORM

The Continuous Hartley Transform (CHT) is an orthogonal integral transform. Originally, Hartley defined the CHT of a continuous – time function $f(t)$ as

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \cdot \text{cas}(\omega t) dt \quad (1)$$

Here, ω is the angular frequency in rad/sec. Note that $\text{cas}(\alpha) = \cos \alpha + \sin \alpha = \sqrt{2} \sin\left(\alpha + \frac{\pi}{4}\right) = \sqrt{2} \cos\left(\alpha - \frac{\pi}{4}\right)$ is the cosine-and-sine or Hartley kernel. In Signal Processing terms, this transform takes a signal (function) from the time-domain to the Hartley spectral domain (frequency domain). Eq. (1) is called the Analysis Equation of the CHT [3]. The Inverse Continuous Hartley Transform (ICHT) was again defined by Hartley originally as

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(\omega) \cdot \text{cas}(\omega t) d\omega \quad (2)$$

Eq. (2) is called the Synthesis Equation of the CHT [4]. The Hartley transform has the convenient property of being its own inverse. Hence, it is an Involution Integral or Transform [5].

The properties of the $\text{cas}[\cdot]$ function follow directly from Trigonometry and its definition as phase-shifted trigonometric functions as given below.

$$\text{cas}\theta_2 = \sqrt{2} \sin\left(\theta_2 + \frac{\pi}{4}\right) = \sqrt{2} \cos\left(\theta_2 - \frac{\pi}{4}\right).$$

It has angle-addition identity as given below.

$$\text{cas}(\theta_1 + \theta_2) = \{\cos \theta_1 \text{cas} \theta_2 + \sin \theta_1 \text{cas}(-\theta_2)\}$$

Also,

$$\text{cas}(\theta_1 + \theta_2) = \{\cos \theta_2 \text{cas} \theta_1 + \sin \theta_2 \text{cas}(-\theta_1)\}$$

The first-order derivative of $\text{cas}\theta_1$ w.r.t. θ_1 is given by the following expression [6],

$$\text{cas}'(\theta_1) = \frac{d}{d\theta_1} [\text{cas}\theta_1] = \cos \theta_1 - \sin \theta_1 = \text{cas}(-\theta_1).$$

III. RELATIONSHIP BETWEEN THE CONTINUOUS HARTLEY TRANSFORM AND THE COMPLEX FOURIER TRANSFORM

The CHT is closely related to the Continuous Time Fourier Transform or Infinite Fourier Transform or the Complex Fourier Transform. CHT differs from the classic Fourier transform in the choice of the kernel. It is well-known that the Continuous Time Fourier Transform is a complex integral transform due to the use of the complex exponential, $e^{-j\omega t}$ or $e^{-j2\pi ft}$ as its kernel. But the Continuous Hartley Transform (CHT), unlike the Continuous Time Fourier Transform, is a real transform because its kernel $\text{cas}(\theta) = \cos \theta + \sin \theta$ is real. Even its backward or inverse transform is also real [7].

The Continuous – Time Fourier Transform (CTFT) of a continuous – time function $f(t)$ is defined as,

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (3)$$

The Inverse Fourier Transform of $F(\omega)$ is defined by the following equation,

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (4)$$

The CTFT can be directly obtained by CHT by the following expression,

$$F(\omega) = \frac{1}{2} [H(\omega) + H(-\omega)] - j \frac{1}{2} [H(\omega) - H(-\omega)] \quad (5)$$

That is, the real and imaginary parts of the Fourier transform are simply given by the even and odd parts of the Hartley transform, respectively. Conversely, for a real-valued function $f(t)$, the Hartley transform is given from the Continuous – Time Fourier transform's real and imaginary parts [8] as

$$H(\omega) = F_R(\omega) - F_I(\omega) = \text{Re}[F\{f(t(1+j))\}] \quad (6)$$

Here, F is the Fourier Transform operator, and $\text{Re}[\cdot]$ depicts the real part of the entity. It is easy to extend the definition of Continuous Hartley Transform (CWT) and its inverse to N – dimensions.

IV. THE DISCRETE HARTLEY TRANSFORM (DHT)

The Discrete Hartley Transform (DHT) of a discrete – time signal, $x(n)$, is defined as [9]

$$X_H(k) = \text{DHT}[x(n)] = \sum_{n=0}^{N-1} x(n) \text{cas}\left(\frac{2\pi kn}{N}\right) \quad (7)$$

$$\forall 0 \leq n, k \leq N - 1.$$

$$\text{where, } \text{cas}\left(\frac{2\pi kn}{N}\right) = \cos\left(\frac{2\pi kn}{N}\right) + \sin\left(\frac{2\pi kn}{N}\right).$$

$$\text{Also, note that } \text{cas}\left(-\frac{2\pi kn}{N}\right) = \cos\left(\frac{2\pi kn}{N}\right) - \sin\left(\frac{2\pi kn}{N}\right) \quad (8)$$

The Inverse Discrete Hartley Transform (IDHT) of $X_H(k)$ is defined as,

$$x(n) = \text{IDHT}[X_H(k)] = \sum_{k=0}^{N-1} X_H(k) \text{cas}\left(\frac{2\pi kn}{N}\right) \quad (9)$$

$$\forall 0 \leq n, k \leq N - 1.$$

Eqs. (7) and (9) are respectively called the Analysis and Synthesis Equations of DHT. Note that $\text{cas}\left(\frac{2\pi kn}{N}\right)$ is the kernel of DHT [10]. It is a real term due to which the DHT is a real transform. The same kernel is made use of in the computation of the IDHT. Thus, unlike in the DFT, where there is a change of sign in the kernel in the synthesis equation, in the case of the DHT, there is no such thing which means that a single algorithm can be used to compute both the forward and backward transforms of DHT, which is a major advantage over the conventional DFT. This is the equivalent to the case encountered in Continuous Hartley Transform. Hence, the DHT also is its own inverse. Thus, the DHT involves only real operations and hence, the computational load and memory requirement are considerably reduced by 50%, which stands out as a good merit over the conventional DFT. Efficient algorithms called the Fast Hartley Transforms (FHT) have been developed for computing the DHT and have been in use in many DSP applications [11].

V. RELATIONSHIPS BETWEEN DHT AND DFT

A. Relationship between the Analysis Equations of DFT and DHT

The DFT of a discrete – time sequence $x(n)$ is defined as

$$X(k) = \text{DFT}[x(n)] = \sum_{n=0}^{N-1} x(n) e^{-\frac{2\pi kn}{N}} \quad (10)$$

$$\forall 0 \leq n, k \leq N - 1.$$

Using Euler's Theorem, $e^{\pm j\gamma} = \text{cis}(\pm\gamma) = \cos(\gamma) \pm j\sin(\gamma)$, we get,

$$X(k) = \sum_{n=0}^{N-1} x(n) [\cos\left(\frac{2\pi kn}{N}\right) - j \sin\left(\frac{2\pi kn}{N}\right)] \quad (11)$$

Due to the absence of the j term in $X_H(k)$, DHT is purely real. Next, consider the following expression,

$$e^{j\alpha} = \text{cis}(\alpha) = \cos(\alpha) + j\sin(\alpha).$$

$$\text{We know that } \cos \alpha = \frac{e^{j\alpha} + e^{-j\alpha}}{2} \text{ and } \sin \alpha = \frac{e^{j\alpha} - e^{-j\alpha}}{2j}$$

$$\Rightarrow \text{cas}\alpha = \cos \alpha + \sin \alpha = \frac{(1-j)e^{j\alpha}}{2} + \frac{(1+j)e^{-j\alpha}}{2} \quad (12)$$

Consider the following assumptions

$$\chi = e^{j\alpha}, k = \text{cas}\alpha, \xi = \frac{(1-j)}{2}, \delta = \frac{1+j}{2} \quad (13)$$

Then, we have the quadratic equation,

$$\xi\chi^2 - k\chi + \delta = 0 \quad (14)$$

Solving for χ , we get,

$$\Rightarrow \chi = \frac{\rho}{2\xi} \pm \frac{1}{2\xi} j (\cos \alpha - \sin \alpha) \quad (15)$$

Considering only the negative sign and simplifying, we get,

$$\chi = e^{j\alpha} = \frac{1+j}{2} \text{cas}\alpha + \frac{1-j}{2} \text{cas}(-\alpha) \quad (16)$$

$$\text{Also, } \chi^{-1} = \frac{1}{\chi} = e^{-j\alpha} = \frac{1+j}{2} \text{cas}(-\alpha) + \frac{1-j}{2} \text{cas}(\alpha) \quad (17)$$

Now, the DFT of $x(n)$ is given by the well known equation

$$X(k) = \text{DFT}[x(n)] = \sum_{n=0}^{N-1} x(n) e^{-\frac{2\pi kn}{N}} \quad (18)$$

Making use of Eq. (17) in the above equation, we get,

$$X(k) = \frac{1+j}{2} X_H(-k) + \frac{1-j}{2} X_H(k)$$

$$\therefore X(k) = \frac{1}{2} [X_H(k) + X_H(-k)] + j \frac{1}{2} [X_H(-k) - X_H(k)]$$

Since $X(k)$ is complex, we can write the above equation as,

$$X(k) = X_R(k) + jX_I(k) \quad (19)$$

Thus, the DFT coefficients in terms of DHT coefficients are given by the following two equations,

$$X_R(k) = \frac{1}{2} [X_H(k) + X_H(-k)] \quad (20)$$

$$X_I(k) = \frac{1}{2} [X_H(-k) - X_H(k)] \quad (21)$$

Subtracting Eq. (21) from Eq. (20), we get the DHT coefficients which are expressed in terms of DFT [12].

$$X_H(k) = X_R(k) - X_I(k) \quad (22)$$

B. Relationship between the Synthesis Equations (Inverses) Of DFT and DHT

If $x_H(n) = \text{IDHT}[X_H(k)]$, then it is possible to express the IDHT of $X_H(k)$ in terms of the IDFT of $X(k)$.

$$x_H(n) = x_R(n) - x_I(n) \quad (23)$$

Here, $x_R(n)$ and $x_I(n)$ are the real and imaginary parts of $x(n)$ respectively. Note that $x_R(n) = \text{IDFT}[X_R(k)]$ and $x_I(n) = \text{IDFT}[X_I(k)]$ respectively. DHT finds applications in a variety of domains such as Speech Processing, Image Processing, Biomedical Signal Processing, Data Compression, etc [13].

VI. SPEECH ANALYSIS USING CEPSTRUM

Cepstrum was invented jointly by B. P. Bogert, M. J. R. Healy, and J. W. Tukey in 1963. The name "cepstrum" was derived by reversing the first four letters of "spectrum." Operations on cepstra are labelled *quefrency analysis*, or *cepstral analysis*. A cepstrum is the result of taking the Inverse Fourier Transform (IFT) of the logarithm of the estimated spectrum of a signal. There is a complex cepstrum, a real cepstrum, a power cepstrum, and phase cepstrum. The power cepstrum in particular finds applications in the analysis of human speech. This was the one that was originally invented [12].

Speech is considered to be the output of a system, the vocal tract $V(z)$ to an input $p(n)$ which is either a periodic impulse train due to the vibration of vocal chords or the white noise due to the air flow. Based on the type of the input, the speech signal is broadly classified as voiced for the periodic impulse train input and as unvoiced speech for the white noise input [13].

Speech is composed of excitation source and vocal tract system components. The main theme of cepstral analysis of speech signal is to segregate its excitation and vocal tract components without any advanced knowledge about source and/or system. Voiced sounds are generated by exciting the time varying system characteristics with periodic impulse signals. The unvoiced sounds are generated by exciting the time varying system with a stochastic noise sequence. Hence, the resultant speech signal is the linear convolution of the corresponding excitation sequence and vocal tract filter characteristics. Let $a(n)$ and $h(n)$ be the excitation sequence and vocal tract filter sequences. Then, the speech signal is given by the one-dimensional linear convolution of $a(n)$ with $h(n)$.

$$s(n) = a(n) \otimes h(n) \quad (24)$$

where, the symbol ' \otimes ' denotes discrete – time linear convolution. Taking Discrete – Time Fourier Transform (DTFT) on both sides of Eq. (24) yields the frequency spectrum,

$$S(e^{j\omega}) = A(e^{j\omega}) \cdot H(e^{j\omega}) \quad (25)$$

Eq. (25) can also be written as

$$S(\omega) = A(\omega) \cdot H(\omega) \quad (26)$$

The speech signal is then deconvolved into the excitation and vocal tract components in the temporal domain. Taking magnitudes on both sides of Eq. (26), we get,

$$\Rightarrow |S(\omega)| = |A(\omega)| \cdot |H(\omega)| \quad (27)$$

Taking Napierian or Natural logarithms on both sides of Eq. (27), we get,

$$\log_e |S(\omega)| = \log_e \{|A(\omega)| \cdot |H(\omega)|\} \quad (28)$$

$$\text{Or } \ln |S(\omega)| = \ln |A(\omega)| + \ln |H(\omega)| \quad (29)$$

The logarithmic operation is used for transforming the magnitude speech spectrum where the excitation component and vocal tract component are multiplied, into a linear combination of these components [14]. The segregation is done by taking the Inverse Discrete Fourier Transform (IDFT) of $\ln |S(\omega)|$, which yields the cepstral coefficients in the temporal domain. In the cepstral domain, the vocal tract

components are represented by the slowly varying components concentrated near the lower cepstral region and excitation components are represented by the fast varying components at the higher cepstral region [15].

$$\begin{aligned} c(n) &= IDFT[\ln|S(\omega)|] \\ \Rightarrow c(n) &= IDFT[\ln|A(\omega)| + \ln|H(\omega)|] \end{aligned} \quad (30)$$



Figure 1 Cepstral Computation using DHT Approach.

Sometimes, variations occur in the lower cepstral region due to the vocal tract characteristics and the rapid varying strata of the cepstrum towards the upper cepstral region that is represented by the excitation characteristics of the short time speech segment. A method that is used for extracting vocal tract and excitation characteristics is liftering operation which is done in the temporal domain. Liftering operation is similar to filtering operation in the frequency domain where a desired cepstral region for analysis is selected by multiplying the whole cepstrum by a Boxcar apodization function at the desired position [12]. Liftering is a useful and meaningful process with the real cepstrum for obtaining an estimate of the log spectrum of either of the separated components. That is, we can apply a useful linear operation to the real cepstrum. The output of this process in the quefrency domain is a real cepstrum. However, if the objective is to return to the original time domain with an estimate of the separated signal, the real cepstrum will fail, because its "linearizing" operation is not invertible. To complete this task, we would need a phase – preserving linearizing operation [14].

VII. APPLICATION OF THE DISCRETE HARTLEY TRANSFORM (DHT) IN DIGITAL SPEECH PROCESSING

DHT, since it is a real transform, is very much useful in the cepstral analysis of speech. If $x(n)$ is the discrete speech signal, firstly, its N -point DHT, $X_H(k)$ is computed. Then, the DFT of $x(n)$ is calculated by exploiting the relationship between DFT and DHT, by using the following relations.

$$X_R(k) = \frac{X_H(k) + X_H(-k)}{2} \quad (31)$$

$$X_I(k) = \frac{X_H(-k) - X_H(k)}{2} \quad (32)$$

$$X(k) = X_R(k) + jX_I(k) \quad (33)$$

We next express $X(k)$ in Steinmetz form or polar form as, $X(k) = |X(k)|e^{j\theta(k)}$ (34)

But from Digital Speech Processing theory, it is known that $\hat{x}(n) = IDFT[\log_e\{X(k)\}] = IDFT[\ln\{X(k)\}]$

$$\Rightarrow \hat{x}(n) = IDFT[\ln\{|X(k)|e^{j\theta(k)}\}]$$

$$\hat{x}(n) = IDFT[\ln|X(k)| + j\theta(k)] = IDFT[\hat{X}(k)] \quad (35)$$

DHT is then computed by using the relation, $\hat{X}_H(k) = \hat{X}_R(k) - \hat{X}_I(k)$, where, $\hat{X}_R(k) = \ln|X(k)|$ and $\hat{X}_I(k) = \theta(k)$.
 $\Rightarrow \hat{X}_H(k) = \ln|X(k)| - \theta(k)$ (36)

Finally, the cepstral coefficients $\hat{x}(n)$, are computed by taking the IDHT of [15] $\hat{X}_H(k)$, i.e.,
 $\hat{x}(n) = IDHT[\hat{X}_H(k)]$ (37)

VIII. RESULT ANALYSIS AND VALIDATION

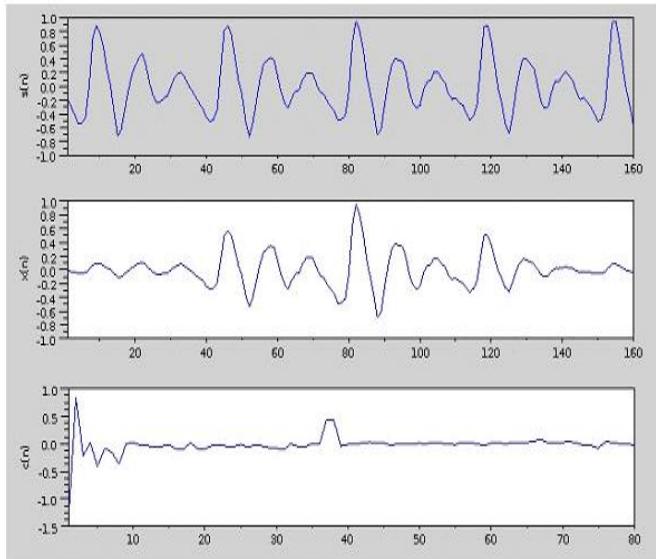


Figure 2 Cepstrum of voiced speech segment using DHT.

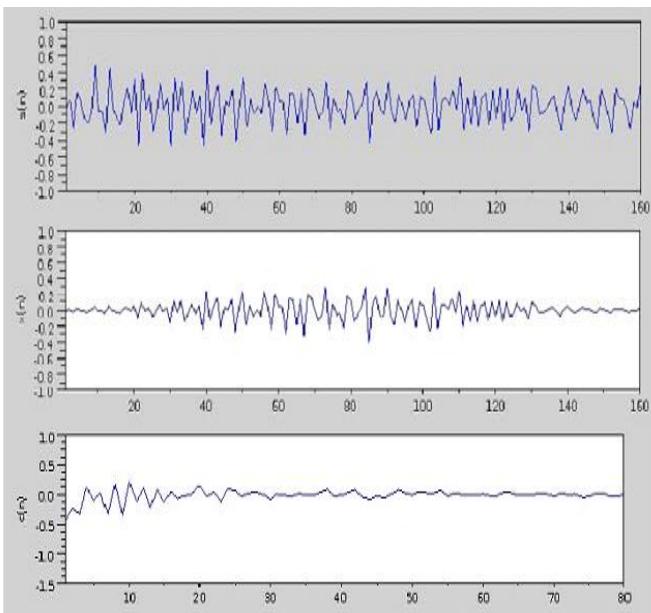


Figure 3 Cepstrum of unvoiced speech segment using DHT.

The cepstra of voiced and unvoiced speech segments using DHT approach are shown in Figures 2 and 3 respectively. The main theme of cepstral analysis of speech signal is to segregate its excitation and vocal tract components. The input speech signal is converted into short-term segments of duration 15 – 20 msec. The frame size is maintained to 20 msec and then each frame is multiplied by Hamming Window. Then, the cepstral representation of short-

term speech is computed by finding the IDFT of the log magnitude spectrum. In this work, instead of using the IDFT directly, we have used IDHT to compute IDFT. This is done by using IDHT rather than using IDFT directly. Figure 2 shows a 20 msec voiced frame and its cepstrum in the temporal domain. It can be clearly seen that the vocal tract components are concentrated in the lower cepstral region and the excitation components are concentrated in the higher cepstral region. Figure 2 also shows $s(n)$ which is the voiced frame considered and $x(n)$ which is the windowed frame. Also shown is the cepstral coefficients $c(n)$. Note that $c(n)$ is symmetrical in the cepstral domain.

CONCLUSIONS

The DHT is two to three times faster than DFT because of no complex arithmetic being involved. All the four quadrants of the Hartley domain data must be used. In the computation of DHT and its inverse, only one quadrant of sines and cosines need to be calculated due to symmetry. Also, the multiplication in DHT is real but complex in the case of DFT. The DHT butterfly loop requires less memory space than the DFT because all of the data are stored in arrays of real numbers. The DFT butterfly loop on the other hand uses complex arrays, which require twice the memory space of a real array. In the DFT computation, there is one multiplication and two additions of complex numbers, which adds up to four multiplications and six additions of floating point numbers for each iteration. Also, the DHT has four multiplications and six additions of floating point numbers for each iteration. The butterfly loop of the DHT loops from 2 to $\frac{N}{2}$, while the butterfly loop of the DFT loops from 1 to N . Since the DHT loops half the number of times as the DFT, the DHT algorithm has two multiplications and three additions for every four multiplications and six additions of the DFT. Also, the DHT has no multiplications for the zero and Nyquist frequencies, which is a major advantage of DHT over DFT. DHT needs less memory to store numbers than the DFT because DHT does not use complex numbers. The results of the DHT can be stored in the same memory space as the original data set, thus eliminating the need to allocate more disk space. Also converting from Hartley domain to Fourier domain and vice versa is a direct and simple procedure.

In this work carried out by the authors, cepstral analysis of voiced speech segment and unvoiced speech segment using both DFT and DHT approaches were considered. Usage of direct DFT approach required 1171.5085 complex additions and 585.75425 complex multiplications. Usage of DHT approach required 292.9 multiplications and 585.823 additions, which indicate that DHT has saved the computation load by 50%. Thus, by using the DHT approach, the computation, storage time and the memory requirement are saved by 50% in the cepstral analysis of speech. This is a good improvement and the DHT based method is good when compared to the traditional DFT approach.

FUTURE SCOPE AND STUDY

The present work is based on the concept of the basic Discrete Hartley Transform. Other variants of Hartley Transform and its amalgamation with Wavelet Transform need to be worked within the nature of quasiperiodicity and the nonstationarity of the speech signal. The behaviour of the speech signal spectrum to such robust algorithms needs to be analysed.

ACKNOWLEDGEMENT

M.B.N likes to thank Dr. Mohan Manghnani, Chairman, New Horizon Education Institution, for providing the necessary infrastructure and encouraging us to carry out this research work. M.B.N is grateful to Dr. Manjunatha, Principal, New Horizon College of Engineering, and Dr. Sanjay Jain, Professor & HOD, ECE Department, New Horizon College of Engineering for their constant support to achieve his endeavour. Dr. P.S.S likes to thank the management of Don Bosco Institute of Technology for their wholehearted support in carrying out this work.

REFERENCES

- [1] Alan V. Oppenheim and Ronald W. Schafer, Discrete – Time Signal Processing,” 3rd Edition, Pearson India Inc., New Delhi, India, 2014.
- [2] John G. Proakis and Dimitris G. Manolakis, “Digital Signal Processing: Principles, Algorithms and Applications,” 4th Edition, Prentice-Hall of India Private Limited, New Delhi, India, 2007.
- [3] Agostino Abbate, Casimer M. DeCusatis and Pankaj K. Das, Wavelets and Subbands: Fundamentals and Applications, Birkhauser, Paris, France, 2003.
- [4] Martin Vetterli and Jelena Kovacevic, Wavelets and Subband Coding, 1st Edition, Prentice-Hall, New Jersey, U.S.A., 1995.
- [5] Sanjit K. Mitra, Digital Signal Processing – A Computer Based Approach, 4th Edition, Tata McGraw – Hill India Private Limited, New Delhi, India, 2014.
- [6] A. Jensen and A. la CourHarbo, Ripples in Mathematics – The Discrete Wavelet Transform, 1st Edition, Springer – Verlag India Private Limited, New Delhi, India, 2003.
- [7] John R. Deller, Jr., John H.L. Hansen, and John G. Proakis, Discrete-Time Processing of Speech Signals, 1st Edition, IEEE Press, Wiley India Private Limited, New Delhi, India, 2014.
- [8] Lawrence R. Rabiner, and Ronald W. Schafer, Digital Processing of Speech Signals, 1st Edition, Pearson India Inc., New Delhi, India, 2005.
- [9] Douglas O'Shaughnessy, Speech Communications: Human and Machine, 2nd Edition, Cambridge University Press India Private Limited, New Delhi, India, 2007.
- [10] Lawrence R. Rabiner, and Biing-Hwang Juang, Fundamentals of Speech Recognition, 1st Edition, Pearson India Inc., New Delhi, 2005.
- [11] Thomas F. Quatieri, Discrete-Time Speech Signal Processing, 1st Edition, Pearson India Inc., New Delhi, 2006.
- [12] A.M. Kondoz, Digital Speech, 2nd Edition, Wiley India Private Limited, New Delhi, India, 2009.
- [13] Ben Gold and Nelson Morgan, Speech and Audio Signal Processing: Processing and Perception of Speech and Music, 1st Edition, Wiley India Private Limited, New Delhi, India, 2009.
- [14] Shaila D. Apte, Speech and Audio Processing, 1st Edition, Wiley India Private Limited, New Delhi, India, 2014.
- [15] Chris Rowden (Editor), Speech Processing, 1st Edition, The Essex Series in Telecommunications & Information Systems, McGraw-Hill Publishing Company, Berkshire, England, 1992.