iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** VI **Month of publication:** June 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A Novel Approach for Waveform Compression

Dilpreet Singh[1], Parminder Singh[2]
[1]M.Tech. Student, [2]Associate Professor
CSE Department, Guru Nanak Dev Engineering College, Ludhiana

Abstract— Waveform Compression is a field of digital signal processing that focuses on reducing byte-rate and sample rate of waveform speech signals to increase transmission speed and storage capacity in multimedia devices. This paper explores resampling methodology for compressing the speech signal of wave file format. Resampling methodology is based on removing unessential samples from input waveform signal. In resampling method, channel bandwidth is utilised properly by reducing the byte rate and sample rate, which results in compression and thus, increasing the transmission speed and storage capacity. A major objective of waveform speech compression is to represent a speech signal with few bits that satisfies the level of wave file quality. The results illustrate the effectiveness of the proposed method in the field of data compression.

Keyword — Waveform, Digital Signal processing, Pulse Code modulation, Speech Samples.

I. INTRODUCTION

Storing large amount of data on devices like super computers is not a big issue, but the problem occurs for portable devices. The small scale platform devices as mobile phones have limited amount of storage memory to store files, so we have to find a compression technique to shrink the size of the data. Also, usage of the internet in the past few years is highly increased considerably and it provides us a platform to transfer files. Most of the high quality speech signals are used in online multimedia applications. These applications are widely used but they suffer from data loss because the amount of data being transferred is more than the transferring capacity of the bandwidth. Although, high bandwidth services are available today for transferring data across the internet, but still many people suffer from problems like uploading and downloading failures while using the internet. To overcome these problems, data compression is used. The method for data compression proposed in this paper is based on resampling. Speech is a medium which provides communication between individuals. People use language to exchange information and emotions while sitting in the same room or all over the world. Humans talk at a frequency of 4 kHz [1, 2]. A wave file has large number of samples. In various other audio file formats, compression is already done, for example, MP3 file format. The input signal is always digitised. The data compression, in general, removes the unnecessary information of the wave file. The proposed model is based on removing the discontinuity between neighbouring samples while preserving the quality of the actual data. The byte-rate of the wave signal is compressed in which the unnecessary bytes are removed. The proposed method compresses the wave file and splits it into different categories. In initial time of research, various speech compression techniques were established. The compression techniques are categorized into two types: dedicated techniques and general techniques [3]. In dedicated techniques, there is minimum distortion in the output speech. Whereas, the general techniques such as differential pulse code modulation, sub-band coding, and vector quantization have sound mathematical foundation.

II. RELATED WORK

Stylianou [4] performed the harmonic plus noise model (HNM) performs concatinative text-to-speech (TTS) synthesis. In HNM, waveform is represented as a time-varying harmonic component plus a modulated noise component. The decomposition of a waveform signal into these two components provides durable waveform modifications of the signal. The parametric representation of speech using HNM provides a simple method of smoothing discontinuities of acoustic units around concatenation points. The HNM provides high-quality speech synthesis while outperforming other models for synthesis (e.g., TD-PSOLA) in intelligibility, naturalness, and pleasantness. Nagy and Rozinaj, Implemented HNM (harmonic plus noise) model which was further extended by transient model for the compression of Solovak speech [5]. Furthermore, the method used for noise modelling in the proposed HNM system is not the same as the one used in classical HNM model. To process sounds like plosives, the transient model was added. The HNM method compresses the parameterized speech into a format which makes it easy to take prosodic modifications of the speech needed for speech synthesis. This approach of speech description and compression allows us to reduce the database size of waveform segments. Authors have also discussed the application of harmonic plus noise (HNM), which is further extended using the transient model. It is used for the construction of compressed waveform database in a format that is useful for prosodic modification of the synthesized speech. The HNM method substantially decreases the database of speech segments for a concatenative speech synthesis too [6].

Chompun et al. proposed a slightly modified flexible Multi-Pulse based Code Excited Linear Predictive (MP-CELP) coder to

©IJRASET 2015: All Rights are Reserved

evaluate the bit-rate for tonal speech language in the mobile applications. The coder comprises of a core coder and bit-rate scalable tools. The high pitch delay resolutions are applied to the versatile codebook of core coder to increase the waveform quality. The bit-rate scalable tool employs multi-stage excitation coding based on an embedded-coding approach. The output has the waveform quality of the desired coder and its speech quality is better than that of the former coder without waveform pitch-resolution adaptation [7].

Rajesh et al. described that the speech compression is a part of digital signal processing that works on reducing the bit-rate of waveform to enlarge processing speed of the wave signal and storage for fast developing multimedia. It based on a transform based methodology for compression of the waveform. In this methodology, different types like Discrete Wavelet Transform (DWT), Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) are exploited. A comparative study was done in terms of Signal-to-Noise ratio (*SNR*), Peak Signal-to-Noise ratio (*PSNR*) and Normalized Root-Mean Square Error (*NRMSE*) and it was found that DWT gives higher compression with respect to DCT [8]. Sunitha and Chitneedi, elaborated the Discrete wavelet transform and Adapive Kalman filter technique successfully compress and reconstruct words with perfect output by using both waveform coding. The heavy data is stored into small devices. The low bit-rate speech coder deliver compressed toll-quality speech. The output from Wavelet Coding was compared to Adaptive Kalman with Wavelet Coding. It was found out that the output of Wavelet Coding with Adaptive Kalman Filter was better than wavelet transforms [9].

Cai et al. discussed that the basic principle of the linear prediction is reviewed, and the common optimal linear prediction method is improved to get a new optimal linear prediction method that maps integers to integers [10]. An appropriate bit-recombination mark coding approach was explored according to the characteristics of prediction errors sequence. In the end, they proposed a new lossless compression method for acoustic waveform data based on linear prediction and bit-recombination mark coding. The compression performances of this method and several other lossless compression methods were compared and analysed. Test results validate the correctness of their method and demonstrate its advantages. The new method is potentially applicable to acoustic waveform data compression. Kaur discussed that the speech compression is the digital signal which is compressed by using various transmission techniques. Compression of the speech signal is done using transform methodology. Speech is compressed by DWT technique, afterward compressed signal is again compressed by DCT and then the compressed signal is decompressed by DWT. The quality of speech signal is measured on the basis of Peak Signal to Distorted Ratio (PSNR) and Mean Square Error (MSE) by using other filters of wavelet family [11].

III. TECHNIQUES

There are two major categories of compression: lossless and lossy compression. A lossless compression generates the same copy of the input original file after decompression method applies on it. The most common example of the lossless compression is the ZIP format. This compression method is useful on a range of files. The lossy compression does not generate the same copy of the original file after decompression method applies on it. The example of the lossy compression is JPEG format and the MP3 format used in coding audio data. The lossy compression is supported on psychoacoustics which takes into compassion the forecast behaviour of the human ear. Human can hear the frequency range of 20 to 20 kHz. Classification of compression method can be done by using three methods:

A. Direct Methods

The samples of the wave form signal are directly manipulated to deliver compression.

B. Transformation Methods

There are mainly 3 methods used for waveform compression. In Discrete Cosine Transform method, the energy of speech signal is concentrated in a few transform coefficients which yield good compression. In Fourier Transform, a waveform signal having periodic function of time is examined or synthesized as a number of harmonically related sine and cosine signals. Wavelet Transform delivers a compact representation of a waveform signal in terms of time and frequency.

C. Parameter Extraction Methods

Some features are extracted in advance using pre-processor. These features are later used to compress the wave form signal.

IV. METHODOLOGY

The 8-bit mono channel WAV sampled at 22,050 Hz (Hertz) would take 22,050 bytes per second. A 16-bit stereo WAV with a sampling rate of 44.1 kHz (kilohertz) takes 176,400 bytes per second (44,100/second * 2 bytes * 2 channels) [4]. In the proposed method the waveform compression, the input wave file is divided into several data blocks. The size of the wave file is based on its sample rate and byte rate. The wave file requires more storage space when the byte rate is high and less space as the

©IJRASET 2015: All Rights are Reserved

byte rate is low. The advantage of this approach is that the accurate compression rate can be selected adaptively to compress the input audio file according to the probability distribution characteristics of input audio file [8].

The resampling method takes a wave file as input and analysed its header part and data part. Header part specifies the attributes of audio file and data part contains the actual data. Bytes from 0 to 43 are the attributes of a file and remaining bytes represent the actual data. The header part is formed of three parts, first is RIFF(Resource Interchange File Format) chunk, second is format sub-chunk and last is data sub-chunk. The data part has actual bits which depend upon the size of the input file. The size of sample that is used is 16 bit. An array is created to store output data. A flag is formed and initialized to zero. The ratio is obtained by dividing the actual sample rate of input file by the desired sample rate. Each input sample from sample index is processed using ratio and flag. Then, process each sample by increasing flag one by one. Thus new samples obtained after processing the input samples are stored in an array. If sample index reaches to maximum bound then stop otherwise proceed to the next sample. When all the input samples are processed and stored in array file, then process the header part and combine both the files, this is called indexing. The output header part has different values than input header part. This method has achieves a good compression factor. Fig.1. shows the work flow of the proposed method.

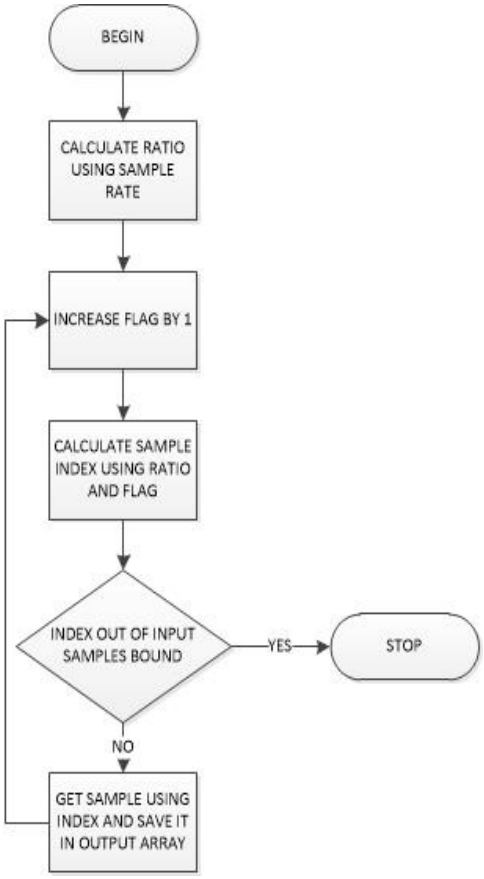


Fig. 1 Flow chart of Wave compression

A. Proposed Algorithm

Main steps of the algorithm are discussed below:

Flag which stores numeric values without decimal factor is set; this flag will be used to calculate other parameters in current algorithm.

Calculate the ratio of input sample rate and desired sample rate both will be numeric and then store the result in a variable including decimal factor of numeric value.

In the next step, an infinite loop is used, which will break, when a particular condition is satisfied, otherwise it will continue looping, from step D onwards.

Increment flag value, and calculate the index of current sample array from input file, using flag and ratio of input sample rate and desired sample rate, and discard the decimal portion of result by rounding off the numeric value.

©IJRASET 2015: All Rights are Reserved

Check that the index calculated in previous step, is inbound of current sample array range or not.
If the calculated index, inbound of current sample array range, just pick the sample from input sample range using that calculated index, and continue with the loop, started in Step C.
Otherwise break the loop, which is started in step C.

B. Bit Depth Conversion

In the bit depth conversion method, decrease the actual size of bytes occupied by each sample in the audio file, using different techniques like shifting most significant 8 bits to right and then storing these 8 bits in sample of 8 bits. As bit depth changed to 8 bits from 16 bits that is also known as Wave to pulse code modulation (PCM) conversion. But in Bit depth conversion, distortion takes place at the background, and voice of the output file is also distorted. Graphical representation of the wave signal, after bit depth conversion from 16 bit to 8 bit is shown in Fig. 2. In the diagram, the input signal wave form cycles in sequence shape and proper format. But the output signal of the waveform cycles is not in a proper format.

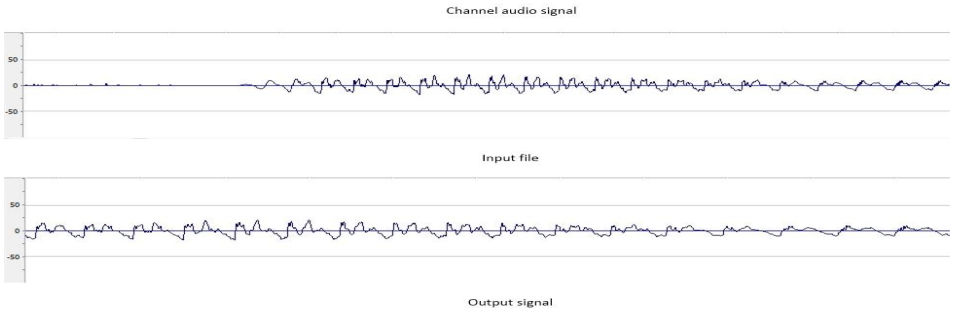


Fig. 2 Wave to PCM

C. Resampling

After Bit Depth Conversion from 16 bits to 8 bits, the resampling size of wave file is decreased, but it also leads to distortion in the background of voice. The signal is represented as a total number of sinusoids which compare frequencies, amplitudes and stages. This problem is removed with the help of Resampling method. Resampling method is used to decrease the size of wave file as it discards unwanted samples in the wave file and changes the sample rate and byte rate of audio file. The resampling method decreases the sample rate from 44100 samples per second to 16000 samples per second and byte rate from 88200 bytes per second to 32000 bytes per second. But during resampling, the bits per sample remain the same because it produces distortion as wave format is not supported while converting from 16 bit to 8 bit. The data compression is based on the transformation field. The graphical representation of wave signal compression, input wave file and compressed wave file are shown in Fig. 3.

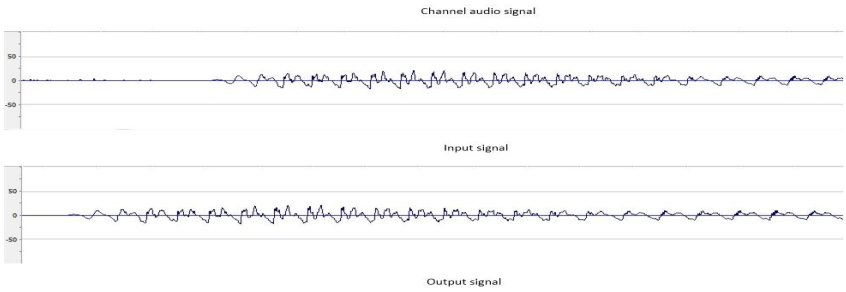


Fig. 3 Resampling

V. CONCLUSION

The proposed compression technique has reasonable size reduction in the input wave file. In the resampling method, the compression rate is 64-71% as the compression rate varies according to pitch of the wave file signal. The compressed waveform speech quality is useful for mobile phone speech synthesis. This method modifies the signal and increases the naturalness and quality of the compressed output signal of the speech. This approach digitizes each sample data using the minimum bit-rate, and the output can be further modified to achieve better results. The output signal describes the accuracy of the compression. Using this wave compression, we have achieved a reasonable data size without distortion that is easy to transfer and store. The resampling method is used at the time of reconstruction of the compressed wave file, and it achieves a high compression ratio.

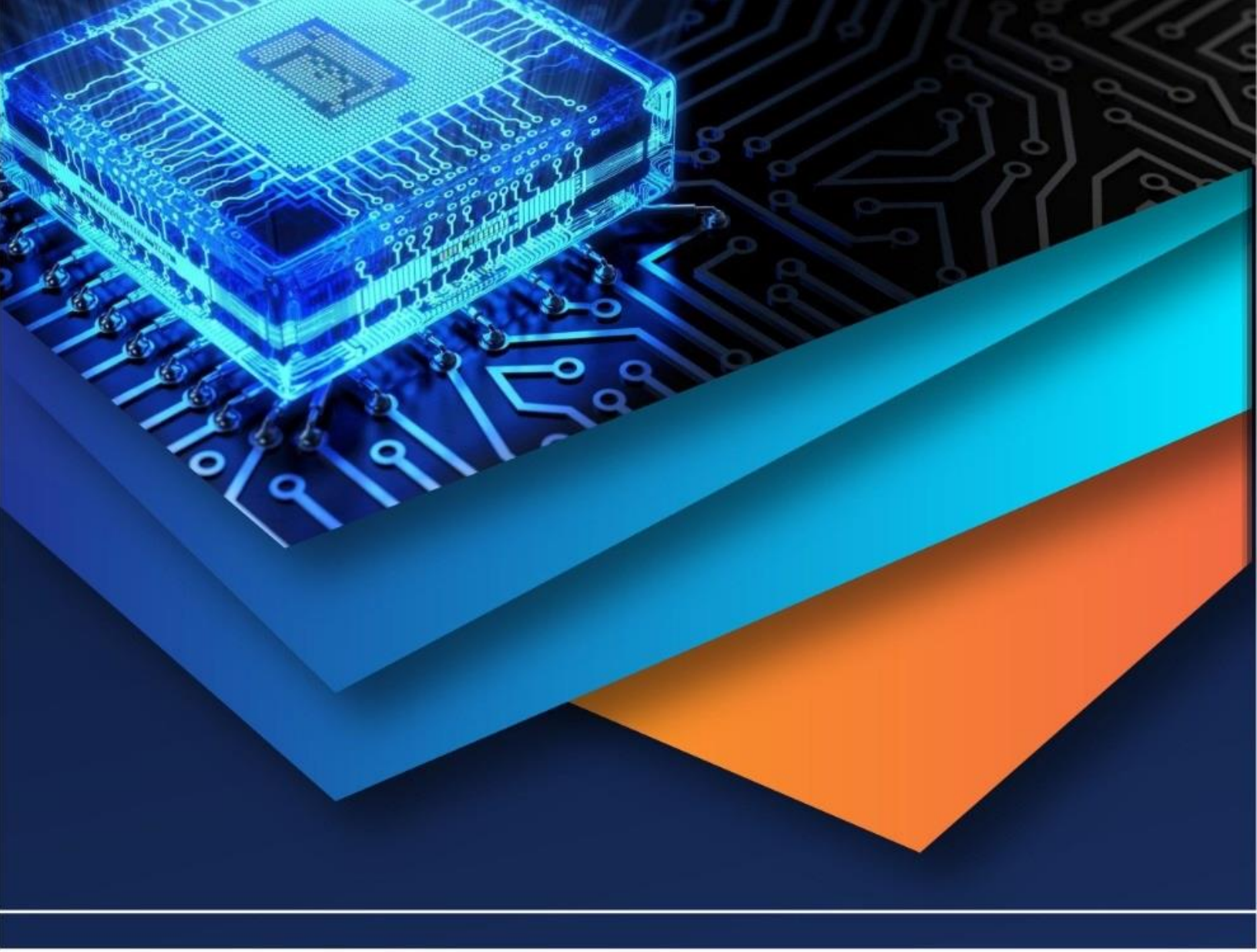
©IJRASET 2015: All Rights are Reserved

Better results are achieved by inputting an audio having high pitch and resolution delay.

REFERENCES

[1] H. Elaydi, M.I. Jaber and M.B. Tanboura, "Speech Compression Using Wavelet", International Journal for Applied Sciences, Vol. 2, pp.1-4, 2011.
[2] W. Chong and J. Kim, "Speech and Image Compression by DCT, Wavelet, and Wavelet Packet", in Proc. Information, Communication and Signal processing IEEE, Vol. 3, pp.1353-1357, 2002.
[3] R.S.H. Istepanian, A. Sungoor and J.C. Nebel, "Linear Predictive Coding and Wavelet Decomposition for Robust Microarray Data Clustering", IEEE, pp.4629-4632, 2007.
[4] Y. Stylianou, "Applying the Harmonic Plus Noise Model in Concatenative Speech Synthesis", IEEE Transaction on Speech and Audio Processing, Vol. 9, No. 1, pp.21-29, 2001.
[5] M.T. Nagy and G. Rozinaj, "Compression of a Slovak Speech Database Using Harmonic, Noise and Transient Model", in Proc. 52th International Symposium ELMAR-2010, Zadar, Croatia, pp.363-366, 2010.
[6] M.T. Nagy, G. Rozinaj and P. Hviš_, "Parametrization of a Slovak Speech Database for Mobile Platform Speech Synthesis", in Proc. 51th International Symposium ELMAR-2009, Zadar, Croatia, pp.225-228, 2009.
[7] S. Chompun, S. Jitapunkul and D. Tancharoen, "Novel Technique For Tonal Language Speech Compression Based on A Bit-rate Scalable MP-CELP Coder", in Proc. Information Technology: Coding and Computing, IEEE Computer Society ITCC 2001 April 2-4, Las Vegas, Nevada, USA, 2001.
[8] G. Rajesh, A. Kumar and K. Ranjeet, "Speech Compression using Different Transform Techniques", in Proc. International Conference on Computer & Communication Technology (ICCCT), pp.146-151, 2011.
[9] P. Sunitha and S.P. Chitneedi, "Speech Compression for Better Audibility Using Wavelet Transformation with Adaptive Kalman Filtering", International Journal of Engineering Research and General Science, Vol. 2, Issue 4, pp.379-384, 2014.
[10] M. Cai, W. Qiao, X. Ju and X. Che, "Lossless Compression Method for Acoustic Waveform Data Based on Linear Prediction and Bit-recombination Mark Coding", World Congress on Engineering and Computer Science (WCECS) San Francisco, USA, Vol. 1, 2013.
[11] H. Kaur and R. Kaur, "Speech compression and decompression using DWT and DCT", International Journal of Computer Technology & Applications (IJCTA), Vol. 3, Issue 4, pp.1501-1503, 2012.

©IJRASET 2015: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)