ज्ञान-विज्ञान विमुक्तये

UNIVERSITY GRANTS COMMISSION

BAHADUR SHAH ZAFAR MARG

NEW DELHI – 110 002.

PROFORMA FOR SUBMISSION OF INFORMATION AT THE TIME OF SENDING THE FINAL REPORT OF THE WORK DONE ON THE PROJECT

1.	**NAME AND ADDRESS OF THE PRINCIPAL INVESTIGATOR**	Dr. Subrata Bhattacharya Dept. Of Electronics Engineering, Indian Institute of Technology (Indian School of Mines) Dhanbad Pin - 826004
2.	**NAME AND ADDRESS OF THE INSTITUTION**	Indian Institute of Technology (Indian School of Mines) Dhanbad Pin - 826004
3.	**UGC APPROVAL NO. & DATE**	F. 42-120/2013 (SR) DATED 12TH MARCH 2013
4.	**DATE OF IMPLEMENTATION**	Commenced on 01/04/2013
5.	**TENURE OF THE PROJECT**	3 YEARS FROM 01/04/2013 TO 31/03/2016

Subrata Bhattacharya

6. **TOTAL GRANT ALLOCATED** Rs. 7,58,800/-

7. **TOTAL GRANT RECEIVED** 1st Instalment : Rs. 4,69,800 /-

2nd Instalament : NIL

8. **FINAL EXPENDITURE** Rs. 4,69,800

9. **TITLE OF THE PROJECT** REAL-TIME IMPLEMENTATION OF SPEECH CODERS

10. **OBJECTIVES OF THE PROJECT** The main objective of the proposed work is to implement a set of Speech coding algorithms on Digital Signal Processors and comparison of their real-time performances. The DSP platform to be used is TMS 320C6X (the particular DSK to be used in our case is TMS320C6416 DSK) from Texas Instruments and the parameters to be studied during real-time implementation are *synthesized speech quality, time for computation* and *memory used.*

11. **WHETHER OBJECTIVES WERE ACHIEVED** The objective as mentioned above was achieved for a set of speech coders.

12. **ACHIEVEMENTS FROM THE PROJECT**

1) Real-time performance of a set of speech coders that were selected after literature survey was investigated into through implementation using TMS320C6416 DSK
2) Research publications based on the findings have been made and are being made ready for communication in future.
3) Man-power development (manifested through expertise developed by the project fellow which is also beneficial for his academic development i.e. a higher degree like Ph. D.)

13. **SUMMARY OF THE FINDINGS (IN 500 WORDS)**

Subrata Bhattacharya

Based on the literature survey, a set of speech coders were selected for performance comparison: G.718 and G.729.1, Adaptive Multi Rate (AMR) coders – Narrowband & Wideband. Apart from TMS 3206416 DSK, the software MATLAB® 2013a and Code Composer Studio version 5.5 (CCSv5.5®) were used for real-time implementation. We have taken the speech files from the ITU-T website. The sentences we chose were in Hindi and English, spoken by Hindi and British English speaking people – both male and female.

It was observed that variation in some of the parameters e.g. codebook size, split vector quantization (SVQ) in AMR coders resulted in change in execution time and also in memory consumption. If we use less number of SVQ blocks (with larger codebook size), we require more memory as compared to the case when more number of SVQ blocks with smaller codebook size are used. Effect of pre- and post-filtering blocks on memory consumption and execution time were also made. It was found that use of pre- and post-processing blocks resulted in the program memory requirement getting increased.

Quality of synthesized speech in different cases was noted. Objective quality measurement through ITU-T P.862 perceptual evaluation of speech quality (PESQ) recommendation was carried out for different coders and with variation in parameters for a particular coder. For example, there was no variation in PESQ score with change in element per split vector in the case of AMR coder.

Some additional tasks like studying the effect of interpolation of the LSFs were also undertaken to make the investigation more exhaustive. As expected intuitively, the memory requirement for real-time implementation was more with LSF interpolation than that when interpolation was not done.

14.	**CONTRIBUTION TO THE SOCIETY (GIVE DETAILS)**	1. Speech coding being an integral part of present-day communication, with low bit rate and acceptable speech quality as the chief objective, the research undertaken is hoped to contribute in this endeavour, especially it being related to the actual implementation of the coders. 2. Man-power development in different ways: (i) Expertise developed by the project fellow (ii) Percolating the knowledge and expertise developed to the

Subrata Bhattacharya

UG and PG level through laboratory experiments / projects developed for them

(iii) Plans are being made to disseminate the knowledge to the common people to let them know the intricacies of the technology (in our case, speech processing) involved in development of the devices like mobile phones that they are so much familiar with nowadays. This will motivate them to update themselves and seek further guidance thus leading to inclusive growth of the nation.

15.	**WHETHER ANY Ph. D ENROLLED / PRODUCED OUT OF THE PROJECT**	One student (the project fellow) had enrolled for Ph.D (thesis not submitted).
16.	**No. of the Publications out of the project work (please attach reprints)**	1. Kumar, Sandeep, Bhattacharya, S., and Patel, Premanand. (2014) **"A new pitch detection scheme based on ACF and AMDF",** Proceedings *of IEEE International Conference on Recent Advanced Communication Control and Computing Technology (ICACCCT),* Syed Ammal Engg. College, Ramanathapuram, Tamilnadu (India), pp. 1235-1240.

Subrata Bhattacharya
(PRINCIPAL INVESTIGATOR)

1/2/19
(REGISTRAR / PRINCIPAL)

Dean (Research & Development)
Indian Institute of Technology
(Indian School of Mines)
Dhanbad - 826004 (INDIA)

For Director

Brief objective of the Project:

The main objective of the proposed work is to implement a set of Speech coding algorithms on Digital Signal Processors and comparison of their real-time performances. The Speech coding algorithms are to be selected after a detailed survey of the most recent algorithms. The DSP platform to be used is TMS 320C6X (the particular DSK to be used in our case is TMS320C6416T DSK) from Texas Instruments because of their widespread use in real-time implementation of DSP systems. The parameter to be studied during real-time implementation are *synthesized speech quality, time for computation* and *memory used.*

Methodology:

The steps involved in execution of this project are :

1. Survey of literature related to Speech coding techniques especially those related to Perceptual Speech coder.

2. Once the set of speech coders is selected, the next step is implementation of them on Digital Signal Processor (DSP). We are using a processor **TMS 320C6416T** (fixed point Digital Signal Processor) from Texas Instrument for implementation of them.

   The implementation stage includes creating a model using **MATLAB®** and **SIMULINK®**, corresponding to a particular speech coding technique. After the model is created in SIMULINK environment, the **Code Composer Studio® (CCS)** which is an Integrated development environment from Texas Instruments is to be used to compile and to run the real-time model after downloading the executable code to the target board. We are using **MATLAB 2013a** and **Code Composer Studio version 5.5 (CCSv5.5)** as the software component in real-time implementation. Though standard SIMULINK library contains blocks for a variety of signal processing operations and a few blocks for specific operations are available from the DSP block set, some blocks are not available in SIMULINK library for which we need to create the corresponding blocks using *Matlab Embedded function.*

3. Estimation of the parameters - *synthesized speech quality, time for computation* and *memory used* (for all these coders) and performance comparison of the coders in terms of these parameters.

Brief description of the work done

1. Extensive survey of literature related to speech coding techniques, especially those related to perceptual speech coders has been done. It also includes review of literature on real time implementation of different speech coders. Based on the literature survey, a set of speech coders were selected for performance comparison. The speech coders selected are : G.718 and G.729.1 - coders used for wireline services, Adaptive Multi Rate (AMR) coders - for wireline and wireless services. Apart from these coders, internet low bit-rate codec (iLBC) which is used in Voice-over-Internet was also thought of implementation.

2. Purchase of **TMS320C6416 DSK** (a DSP Starter kit based on the *fixed point Digital Signal Processor* TMS 320C6416) has been done. All real-time implementations were carried out using this DSK.

Subrata Bhattacharya

3. Familiarization with TMS320C6416 DSK (earlier we worked on TMS320C6713 DSK which is built using the floating point DSP TMS320C6713), implementation of some basic programs on it and measurement of some important quantities during real-time implementation (execution time and memory used).

4. Development of a new pitch detection technique based on *autocorrelation function (ACF)* and *average magnitude difference function (AMDF)* has been made. One research paper, based on this work, has been presented in an International conference.

5. (a) Development of a SIMULINK model to implement the Adaptive Multi Rate (AMR) coder and study of the effect of variation of some important parameters e.g. effect of using Split vector quantization (SVQ) instead of Vector quantization (VQ), effect of increasing codebook size, effect of pre- and post-processing filters were carried out. Observations were made of synthesized waveform, time taken during simulation and perceptual evaluation of quality (PESQ) of speech.

(b) Development of real-time model of the above coder (using TMS 320C6416 DSK as the target board) and effect of the variation of the parameters mentioned above. Alongwith execution time, memory used during real-time implementation was also measured.

6. Development of the simulation model for G.729.1 and observations made of speech waveforms, PESQ score and time taken during simulation.

7. Development of simulation as well as real-time model of AMR-WB and G.718.

A number of blocks used in the above models were not available in SIMULINK library and we had to build them using *Embedded MATLAB function*.

Subrata Bhattacharya