Implementation of variable bitrate data hiding techniques on standard and proposed GSM 06.10 full rate coder and its overall comparative evaluation of performance

Ninad Bhatt · Yogeshwar Kosta

Abstract Paper deals with implementation of variable bit rate steganographic data transmission over ETSI GSM 06.10 FR coder at five different bitrates. Then, few modifications are suggested in Regular Pulse Excitation section of ETSI GSM FR coder which ultimately claims to produce state of the art proposed GSM FR coder. In contrast with ETSI GSM FR coder, proposed coder also exhibits same bit rate steganographic data transmission. Here, in order to facilitate the same, few RPE pulses are identified and being utilized for embedding and hiding the information bits into them. Key element of this research is to allow for joint speech coding and data hiding and that is accomplished with two different approaches like Fixed and Joint Approach. These both approaches are implemented on both Standard and Proposed coders for their overall analytical evaluation of performance using Subjective (Mean opinion Score and Degraded MOS) and Objective (Perceptual Evaluation of Speech Quality) analysis. Small data information is represented as stego signal which can be embedded over different encoded wave files (chosen from NOIZEUS corpus) that serve as carrier signal. Simulation results for both coders reveal the trade off between data embedding rate and recovered speech quality (for both approaches). It is quite evident from both Subjective and Objective analysis that proposed coder offers comparable performance at the same time with lesser simulation delay because of its inherent constructional difference. It remains the fact that for both the coders, Joint approach performs better but at the cost of more simulation delay.

N. Bhatt (✉)
VNSGU, Surat, Gujarat, India 395007
e-mail: bhattninad@gmail.com

Y. Kosta
Marwadi Education Foundation, Rajkot, Gujarat, India 365001
e-mail: ypkosta@gmail.com

1 Introduction

As far as various wireless communication networks are concerned, till date numerous steganographic data transmission techniques have been evolved in order to embed and transmit the secret data by establishing virtual communication link within the encoded and transmitted host (career) signal. Such information data could be small text, audio, image or any other means of multimedia signals. Conventional data hiding techniques offered direct embedding of information bits on digital encoded speech signals or equivalently transform domain techniques have also been investigated for data embedding which aim to reduce audibility of embedded watermark (Geiser and Vary 2008). A novel approach has been addressed in recent past to jointly embed and encode the speech signal which is popularly known as Joint Source Coding and Data Hiding. When issues related to joint embedding and data hiding are highlighted, occupancy of embedded watermark data and the speech quality offered by modified host signals are surely considered to be important factors. In general, potential applications of watermark data hiding are authentication and digital rights management but in contrast to that, this research focuses on steganographic transmission of data over wireless link and hence performance of stego signals for robustness against deliberate attacks could be less relevant in comparison with higher variable steganographic data transmission rate, constant (minimum) data rate and robustness against transmission errors (Shahbazi et al. 2010). The approach of embedding and data

hiding which is cited above is popularly known as compressed domain watermarking compared to other classical data hiding approaches and the reason is while embedding steganographic data, speech signal is already compressed and encoded.

In this work, modification in the grid position selection strategy for RPE pulses have been proposed and then requantization could be taken place to provide room for embedding steganographic data into the bit steam of host signal (Bhatt and Kosta 2011). GSM voice channel generally uses one of the Full Rate (FR), Half Rate (HR), Enhanced Full Rate (EFR), and Adaptive Multi Rate (AMR) standard speech Codecs.

This research aims at implementation and overall performance evaluation of variable bitrate steganographic data embedding and transmission on wireless link over encoded bitstream of Standard and Proposed GSM FR coder. Partial programming and tweeting using Joint and Fixed approaches (for both the coders) target to mitigate the dependence of variation in recovered speech quality with respect to variation in embedding hidden data bits. The goal of this research is to analyze the overall behavior and performance comparison of both Standard and Proposed coders under variable embedding bit rate conditions. Set of Subjective and Objective analysis parameters are then used to ensure that proposed coder out performs in comparison with its counterpart Standard coder.

2 ETSI GSM 06.10 FR coder and proposed modifications

GSM Full Rate 06.10 Speech Coder is classified in Hybrid Coder which explores Analysis by Synthesis principle to provide attractive trade off between Waveform coder and Vocoders. It exhibits superior speech quality at moderate transmission bit rates but at the cost of comparatively higher implementation complexity (Malkovic 2003).

GSM FR 06.10 speech coder has been standardized by ETSI (ETSI 1998). Full Rate Coder Consists of three major blocks which are Linear Predictive Coding Section, Long Term Predictive Section and Regular Pulse Excitation Section. The proposed modifications are suggested in RPE Section in selection strategy of grid positions as per Bhatt and Kosta (2011). A new proposed grid selection strategy is shown in Fig. 1 and is mathematically expressed as follows.

$$X_m(k) = X(m + 4k)$$

$$m = 0, 1, 2, 3; \ k = 0, 1, \ldots, 9 \tag{1}$$

where m = no. of grids per sub-frame and k = no. of samples per grid.

As can be witnessed from Fig. 1, in the case of proposed coder, number of samples per grid reduces by three and
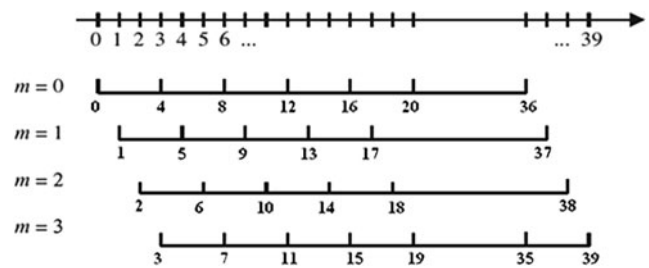

Fig. 1 Sampling grids used in position selection for proposed GSM FR Coder (Bhatt and Kosta 2011)

Table 1 Bit allocation for proposed GSM full rate speech coder (Bhatt and Kosta 2011)

Parameter	No. per frame	Resolution	Total bits/ frame
LPC	8	6, 6, 5, 5, 4, 4, 3, 3	36
Pitch period	4	7	28
Long term gain	4	3	12
Grid position	4	2	8
Peak magnitude	4	6	24
Sample amplitude	$4 * 10$	3	120
Total			224

eventually it results into reduction of total 36 bits per frame which can then be utilized for steganographic data transmission over wireless link (Bhatt and Kosta 2011). In comparison with ETSI GSM FR coder, implementation of Proposed coder offers computationally efficient performance and lesser simulation delay time because of reduced grid size for each subframe. The proposed modification in GSM FR offers a new bit allocation as shown in Table 1.

Each produced time frame of proposed coder consists of encoded bitstream of 224 bits and 36 bits spared for steganographic data transmission. In stark contrast with ETSI GSM FR coder having 260 bits frame size, proposed coder embeds and hides 36 spared bits in Class I_b as per Channel Coding standards GSM 05.03 (ETSI 1999).

The reason of selection of Class I_b is quite obvious as exchange of bits in that class has error protection using Convolution encoder at the same time overwriting of bits in that class offer marginal degradation of recovered speech quality at receiver. Class I_a offers highest error protection but a single bit error introduced because of embedding and overwriting may lead to significant degradation in speech quality. Class II has inherent advantage of embedding and hiding data bits into that class because chances of degradation of speech quality is negligible but as there is no error protection in the said class, hence, it is not viable to pad data bits into it as embedded data itself may be lost because of

burst error. As per (Bhatt and Kosta 2011) (Table 4) proposed modifications in GSM 05.03 (Table 2) (ETSI 1999) have been suggested where data bits d146–d181 (total of 36 bits) have been included into class I_b on order to provide room for steganographic data embedding and transmission.

3 Joint source coding and data hiding

Aggressive research has been carried out in recent past about steganographic information transmission over encoded speech bitstream. Few popular methods are Least Significant Bit (LSB) insertion, Spread Spectrum, Echo and Phase Coding, auditory masking and Quantization Index Modulation (QIM) etc. In contrast with these methods, current research has been focused on Joint Source Coding and Data Hiding techniques. In this section joint source coding and data hiding techniques implemented on Standard and Proposed GSM FR coders are discussed.

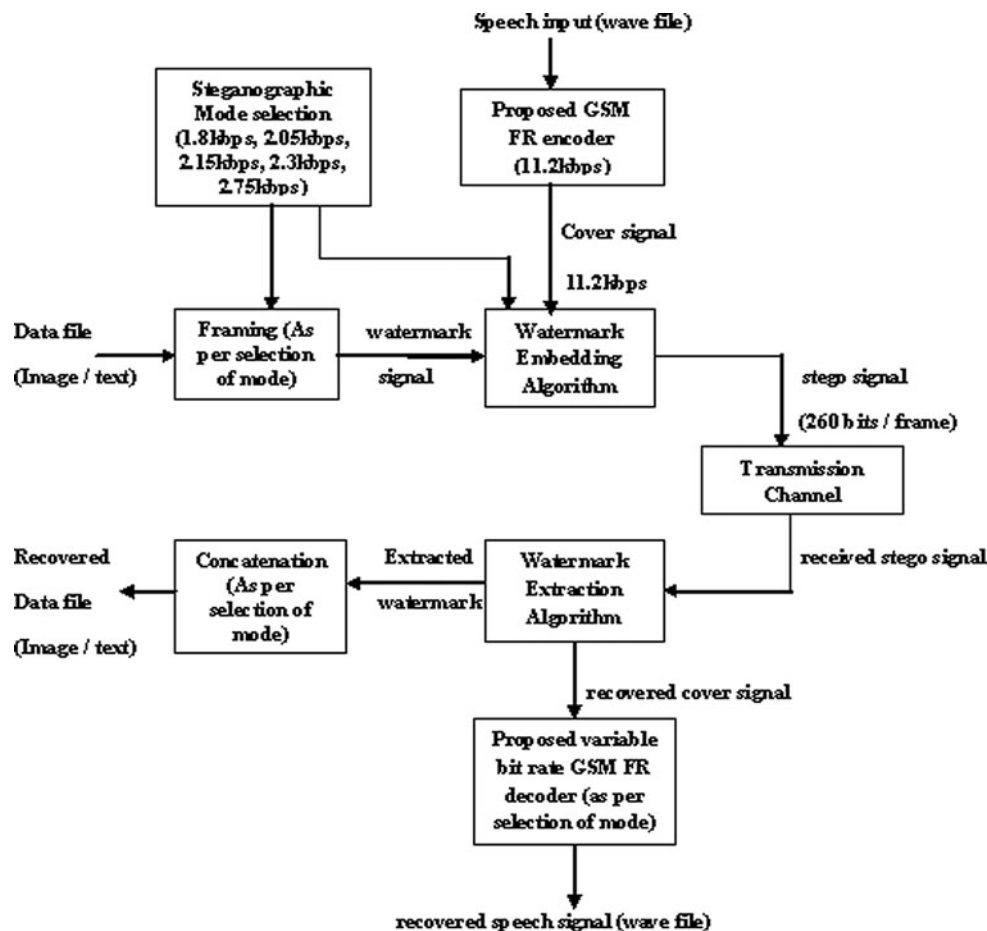### 3.1 Variable bitrate data hiding on proposed GSM FR coder

In this work, five different bitrate steganographic modes have been developed on Proposed coder. Among them the first steganographic mode is 1.8 kbps which is produced by sparing 36 bits/frame as discussed previously. The other four steganographic modes are 2.05 kbps (41 bits/frame), 2.15 kbps (43 bits/frame), 2.3 kbps (46 bits/frame) and 2.75 kbps (55 bits/frame) where each time frame consists of 20 ms as per ETSI GSM FR (ETSI 1998). As can be described in Bhatt and Kosta (2011) (Table 4 Class I_b), few RPE pulse no. 13–22 (bit no. d127–d136) and 27–35 (bit no. d137–d145) with bit index one having been chosen for data embedding and overwriting as per Shahbazi et al. (2010). The major reason for identifying few above mentioned RPE pulses for steganographic data embedding is because overwriting of those specified bits, offer only marginal degradation in terms of recovered speech quality as per Shahbazi et al. (2010). Thus selection of all the steganographic modes is on the basis of identification of RPE pulses as per Shahbazi et al. (2010) over which embedding and masking of information data bits may result into marginal degradation of speech quality at receiving terminal.

Figure 2 demonstrates the proposed joint coding and data hiding techniques carried out in this research work. Initially data information (like text/image/audio contents) has been converted into frames. Here it is to be noted that size of frames should be made variable, depending upon selection of steganographic mode, between 36 bits to 55 bits. Cover (host) signal can be generated by performing encoding operation on developed proposed GSM FR coder. Frame size

Table 2 Embedding positions of steganographic data on proposed GSM FR coder at different mode of bitrates

Embedding rates (Kbps)	1 (RPE 13)	2 (RPE 14)	3 (RPE 15)	4 (RPE 16)	5 (RPE 17)	6 (RPE 18)	7 (RPE 19)	8 (RPE 20)	9 (RPE 21)	10 (RPE 22)	11 (RPE 27)	12 (RPE 28)	13 (RPE 29)	14 (RPE 30)	15 (RPE 31)	16 (RPE 32)	17 (RPE 33)	18 (RPE 34)	19 (RPE 35)
1.8 kbps	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2.05 kbps	●				●				●				●				●		
2.15 kbps	◀			◀			◀			◀			◀			◀			◀
2.3 kbps	▶		▶		▶		▶		▶		▶		▶		▶		▶		▶
2.75 kbps	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆

of cover signal for proposed coder is 224 bits and as discussed earlier for steganographic data embedding in the case of last four modes, RPE pulses are chosen and bits are embedded by overwriting. Role of the watermark embedding algorithm is to combine host signal with steganographic bitstream and eventually it should produce stego signal where each frame contains 260 bits per 20 ms time frame at the original 13 kbps bitrate of standard GSM FR coder. In this work, transmission channel and its analysis has not been touched upon. As per mode selection, watermark extraction algorithm extracts and separates recovered cover signal and steganographic data bitstream. Recovered cover signal is fed to decoding section of proposed GSM FR coder for reproduction of speech signal and simultaneously frame wise received steganographic bitstream are finally concatenated in order recover original data file (text/image/audio etc.).

3.1.1 Fixed approach

In this approach, positions of RPE pulses (having bit index equal to one from Class I_b) for embedding and hiding information data are made fixed as per Table 2. If the embedded bit (from given information file) to be overwritten on given

RPE pulse (as per Table 2) is different from that RPE bit then fixed approach produces quantization error of decimal value two and in turn it results into marginal degradation of recovered speech quality. If both embedding and RPE bits are same then error is zero. As can be observed from Table 2 except 2.75 kbps mode, in all other modes few RPE pulses are not at all utilized for embedding hidden data.

As discussed in Sect. 2, each RPE pulses are encoded by three bits using Adaptive Pulse Code Modulation according to ETSI GSM FR standards. The coded RPE pulses are represented by x_c. Let us assume that x is the magnitude of RPE pulses and y is the magnitude of decoded RPE pulses. x_c is three bit encoded value of RPE pulses which are denoted as $x_1 x_2 x_3$ and $x_{c'}$ is new generated bitstream after embedding information bits. Information bit to be embedded is denoted as x_i (ETSI 1999).

As each RPE pulses are encoded by three bits, embedding of hidden data bit into it at location of bit index one, produces quantization error. Out of eight possible combinations of RPE pulses x_c, for both the cases $x_i = 0$ and $x_i = 1$, in four combinations the quantization error is decimal values two and zero in the remaining four combinations.

3.1.2 Joint approach

In this approach rather than embedding steganographic data in bit index one of given RPE pulses, RPE pulses having bit index one and zero both are jointly utilized for embedding steganographic data. In order to minimize quantization error, x_2 and x_3 are modified jointly for embedding in the following way. Assuming if information bit (to be embedded) is $x_i = 1$ and $x_c = 010$ then $x_{c'} = 000$ for fixed approach but $x_{c'} = 011$ as per following algorithm of joint approach (ETSI 1999).

if $(x_i > x_2)$
　　$x_2 = x_i$
　　$x_3 = 0$
　　else
　　$x_2 = x_i$
　　$x_3 = 1$
end

In the case of Joint Approach, out of eight possible combinations of RPE pulses x_c, for both the cases of $x_i = 0$ and $x_i = 1$, only in two combinations quantization error has decimal value two where as in four combinations quantization error is observed to be having decimal value one and remaining two combinations reflect quantization error with decimal value zero.

3.2 Variable bitrate data hiding on standard ETSI GSM FR coder

Here in the case of steganographic data hiding over standard GSM FR coder, for implementing both fixed and joint approaches, few GSM encoded parameters like RPE pulses, block amplitudes and Log Area Ratios are identified that belongs to Class I_b as per channel coding standards 05.03 (Table 2). With reference to Fig. 2, in this variable data hiding approach for standard GSM FR coder, input speech wave file is applied to Standard GSM FR encoder of 13 kbps (in place of proposed GSM FR Encoder) and the role of watermark Embedding algorithm is to sort out few encoded GSM FR parameters for embedding and overwriting data to generate final stego signal. For implementing different steganographic bitrate modes, FR encoded parameters from class I_b as per GSM 05.03 (Table 2) have been chosen and overwritten with bits of steganographic data bitstream. For first mode of 1.8 kbps RPE pulses number (having bit index one from class I_b) 20–25, 30–42, 47–59, 64–67 with total 36 bits/frame have been chosen and overwritten with steganographic data bits. For 2.05 kbps mode, in addition to above 36 bits other RPE pulses 15–19 (having bit index one from class I_b as per GSM 05.03) are added that results into 41 bits/frame which have been chosen and overwritten. For 2.15 kbps mode in addition to above 41 bits RPE pulses 13–14 have been added to sum up to 43 bits/frame. In the case

of 2.3 kbps mode above 43 bits per frame are added to block amplitude parameter number 29, 46, 63 (each having bit index one from class 1B) resulting into total of 46 bits per frame for steganographic embedding and overwriting. Finally for 2.75 kbps mode above calculated 46 bits are added to block amplitude parameter number 12 (bit index one) and 63 (bit index two), Log Area Ratio number 1, 5, 7 (bit index one), Log Area Ratio number 2, 3, 8, 4 (bit index two) that results into total 55 bits per frame. The above mentioned GSM FR encoded parameters are selected with reference to Shahbazi et al. (2010), Hu and Wang (2006) considering the fact that embedding and overwriting of these parameters affects the degradation of recovered speech quality the least.

The selected bits as per above strategy (for all steganographic bitrate modes) are embedded by overwriting and transmitted using both fixed and joint approaches (as discussed in previously) as a stego signal. At decoder side, this stego signal is then extracted by watermark extraction algorithm to recover both steganographic data file and speech signal from standard GSM FR decoder.

4 Overall performance comparison between variable bitrate steganographic GSM FR and proposed GSM FR coders

This work is splitted into two sections where in first phase Proposed GSM FR coder is implemented for five different steganographic bitrate modes using fix and joint approaches and in next phase ETSI Standard GSM FR coder with the same provisions. To judge and compare the overall performance of both the Standard and Proposed Steganographic coders, here, six speech wave files have been chosen from NOIZEUS corpus (NIOZEUS 2009). Also small text and image files have been selected for steganographic data transmission. Each narrow band speech corpus are sampled by 8 KHz and encoded by 16 bits mono. The length and size of steganographic information embedding is to be made dependent upon size and length of carrier signal i.e. no. of samples in speech wave files and it also depends upon the selection of steganographic mode. In order to compare the overall performance of both the above mentioned steganographic coders, Subjective (Mean Opinion Score and Degraded MOS) and Objective (Perceptual Evaluation of Speech Quality) analysis have been conducted.

4.1 Results obtained for subjective analysis

In this work, two different types of subjective analysis have been carried out. As far as the categories of Subjective analysis are concerned, Mean Opinion Score (MOS) belongs to Absolute Category Ratings (ACR) and Degraded MOS belongs to Degraded Category Ratings (DCR).

Table 3 MOS comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of standard GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps	
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint
Sp01	3.07	3.05	3.12	3.13	3.11	3.14	3.15	3.20	3.22	3.28
Sp02	2.71	2.70	2.73	2.81	2.79	2.85	2.88	2.89	2.88	2.92
Sp06	3.07	3.15	3.08	3.13	3.11	3.12	3.17	3.20	3.25	3.24
Sp21	3.17	3.20	3.22	3.25	3.30	3.30	3.31	3.38	3.40	3.39
Sp25	3.22	3.25	3.23	3.25	3.30	3.28	3.32	3.33	3.38	3.40
Sp30	3.40	3.40	3.31	3.43	3.33	3.42	3.45	3.47	3.48	3.52

Table 4 MOS comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of proposed GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	
Sp01	3.07	3.09	3.09	3.14	3.12	3.14	3.16	3.21	3.22
Sp02	2.80	2.80	2.85	2.88	2.84	2.86	2.87	2.85	2.88
Sp06	3.10	3.11	3.13	3.14	3.16	3.15	3.21	3.22	3.24
Sp21	3.26	3.29	3.31	3.30	3.35	3.38	3.36	3.40	3.40
Sp25	3.28	3.30	3.32	3.35	3.35	3.33	3.34	3.36	3.36
Sp30	3.37	3.39	3.39	3.40	3.44	3.46	3.44	3.45	3.48

4.1.1 Results of mean opinion score ratings

In this analysis, thirty untrained listeners have been chosen to participate into the analysis. Out of thirty, fifteen male and fifteen female listeners have been provided with high quality headphones and subjected to quiet sound proof environment. Each listener has been assigned with decoded wave files of all cases (all six wave files, for all five steganographic bitrate modes and for both joint and fixed approaches) for both the Standard and Proposed implemented GSM FR coders. The scores registered by individual listener for each specific case have then been averaged to obtain the final MOS score.

As can be witnessed from Tables 3 and 4, almost for all cases of bitrate modes and for all decoded wave files for both standard and proposed coders, MOS score value keeps reducing with increase in the steganographic bitrate mode from 1.8 kbps to 2.75 kbps. The reason behind selection of upper bound of bitrate mode of 2.75 kbps is because of the fact that with increase in the steganographic bitrate mode, there should be comparable speech quality at receiving end. It can also be highlighted while comparing both fixed and joint approaches for all cases and for both coders, almost in all cases joint approach results into better values with respect to its counterpart fixed approach. It should be brought to notice that fixed and joint approaches are not possible in 1.8 kbps mode case of proposed coder as this mode is a parent mode developed because of offering proposed modifications on standard GSM FR however in standard GSM FR coder both approaches are implemented and analyzed. Obtained and tabulated results for both standard and proposed GSM FR coders are quite comparable.

4.1.2 Results of degraded mean opinion score ratings

For DMOS analysis, as discussed previously, same procedure has been referred and performed. Initially all six original clean speech files are offered to all listeners (for all cases of bitrate modes, for both fixed and joint approaches and for both standard and proposed coders) before offering decoded speech files and then the ratings of individual listeners are noted down and then scores are averaged to achieve final DMOS scores for individual wave files.

Tables 5 and 6 advocate the overall performance of standard and proposed coders for DMOS ratings. As expected marginal decrement in the values of DMOS are quite evident from the results obtained in the case of both the coders with respect to increase in the steganographic mode from 1.8 kbps case to 2.75 kbps for both the joint and fixed approaches of implementation. Still for majority of the cases, it remains the fact that joint approach offers marginally better obtained results in comparison with its counterpart. In stark contrast it is also visible from Tables 5 and 6 that DMOS scores for both the coders are quite comparable and satisfactory.

4.2 Results obtained for objective analysis

In the objective category of analysis, performance of both the coders for both of the implemented approaches have been studied, evaluated and compared using Perceptual Evaluation of Speech Quality scores as per ITU-T (2001), Hu and Loizou (2008). The measurements of PESQ scores

Table 5 DMOS comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of standard GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps	
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint
Sp01	3.04	3.19	3.06	3.21	3.08	3.20	3.12	3.19	3.17	3.22
Sp02	2.70	2.80	2.75	2.79	2.80	2.84	2.85	2.83	2.89	2.90
Sp06	3.06	3.20	3.14	3.19	3.25	3.28	3.23	3.26	3.30	3.29
Sp21	3.15	3.18	3.21	3.23	3.19	3.29	3.22	3.30	3.33	3.31
Sp25	3.20	3.26	3.24	3.33	3.23	3.29	3.32	3.37	3.39	3.45
Sp30	3.38	3.37	3.36	3.46	3.35	3.44	3.39	3.49	3.41	3.52

Table 6 DMOS comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of proposed GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	
Sp01	3.12	3.12	3.17	3.19	3.15	3.21	3.20	3.23	3.23
Sp02	2.77	2.80	2.80	2.85	2.84	2.83	2.87	2.90	2.92
Sp06	3.12	3.11	3.15	3.20	3.19	3.24	3.22	3.22	3.28
Sp21	3.13	3.12	3.19	3.22	3.17	3.24	3.20	3.22	3.26
Sp25	3.25	3.28	3.30	3.30	3.27	3.29	3.30	3.34	3.34
Sp30	3.32	3.33	3.35	3.37	3.38	3.42	3.39	3.42	3.45

Table 7 PESQ score comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of standard GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps	
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint
Sp01	2.48	2.53	2.49	2.54	2.52	2.51	2.51	2.53	2.55	2.58
Sp02	2.21	2.29	2.24	2.28	2.28	2.31	2.25	2.29	2.27	2.32
Sp06	2.43	2.45	2.44	2.45	2.40	2.43	2.48	2.46	2.51	2.55
Sp21	2.43	2.50	2.54	2.56	2.55	2.58	2.51	2.55	2.54	2.58
Sp25	2.69	2.73	2.68	2.75	2.69	2.75	2.74	2.73	2.77	2.81
Sp30	2.91	2.96	2.92	2.98	3.02	3.03	3.00	3.02	3.01	3.04

Table 8 PESQ score comparison for various wave files for different steganographic bitrate modes (between fixed and joint approaches) of proposed GSM FR coder

Wave file	2.75 kbps		2.30 kbps		2.15 kbps		2.05 kbps		1.8 kbps
	Fix	Joint	Fix	Joint	Fix	Joint	Fix	Joint	
Sp01	2.43	2.45	2.48	2.48	2.47	2.49	2.49	2.51	2.52
Sp02	2.23	2.26	2.25	2.28	2.28	2.27	2.27	2.30	2.30
Sp06	2.45	2.50	2.48	2.51	2.47	2.52	2.46	2.52	2.53
Sp21	2.45	2.48	2.52	2.51	2.52	2.53	2.50	2.52	2.54
Sp25	2.66	2.69	2.69	2.71	2.70	2.70	2.70	2.68	2.73
Sp30	2.83	2.88	2.86	2.89	2.88	2.87	2.89	2.90	2.92

for both standard and proposed coders are cited in Tables 7 and 8.

Tables 7 and 8 depict the comparative analysis and performance of both coders for obtained PESQ scores. As identical to subjective analysis, here in PESQ analysis also marginal but gradual reduction in the score is observed with steganographic bitrate mode increment for both the fixed and joint approaches.

Further to conduct the complete analysis, Standard GSM Full Rate coder (13 kbps) has been implemented and its PESQ scores have been computed for selected set of utterances. The obtained values have been denoted as $PESQ_{max}$.

$PESQ_{min}$ values have been measured for the case of highest mode (i.e. 2.75 kbps mode) for both coders and for both approaches.

Table 9 Overall maximum percentage reduction comparisons between PESQ scores for fixed approach

Wave file	13 kbps standard GSM FR (PESQ$_{max}$)	Proposed GSM FR 2.75 kbps mode with fixed approach (PESQ$_{min}$)	Standard GSM FR 2.75 kbps mode with fixed approach (PESQ$_{min}$)	Percentage reduction in PESQ score (%) for proposed GSM FR 2.75 kbps mode	Percentage reduction in PESQ score (%) for standard GSM FR 2.75 kbps mode
Sp01	2.68	2.43	2.48	9.32 %	7.46 %
Sp02	2.45	2.23	2.21	8.97 %	9.79 %
Sp06	2.61	2.45	2.43	6.13 %	6.89 %
Sp21	2.69	2.45	2.43	8.92 %	9.66 %
Sp25	2.92	2.66	2.69	8.90 %	7.87 %
Sp30	3.03	2.83	2.91	6.60 %	3.96 %

Table 10 Overall maximum percentage reduction comparisons between PESQ scores for joint approach

Wave file	13 kbps standard GSM FR (PESQ$_{max}$)	Proposed GSM FR 2.75 kbps mode with joint approach (PESQ$_{min}$)	Standard GSM FR 2.75 kbps mode with joint approach (PESQ$_{min}$)	Percentage reduction in PESQ score (%) for proposed GSM FR 2.75 kbps mode	Percentage reduction in PESQ score (%) for standard GSM FR 2.75 kbps mode
Sp01	2.68	2.45	2.53	8.58 %	5.59 %
Sp02	2.45	2.26	2.29	7.75 %	6.53 %
Sp06	2.61	2.50	2.45	4.21 %	6.13 %
Sp21	2.69	2.48	2.50	7.80 %	7.06 %
Sp25	2.92	2.69	2.73	7.87 %	6.50 %
Sp30	3.03	2.88	2.96	4.95 %	2.31 %

As can be demonstrated in Tables 9 and 10, both standard and proposed coders are quite comparable (for both fixed and joint approaches) with respect to maximum percentage reduction of PESQ score in the case of highest steganographic bitrate mode of 2.75 kbps. Overall percentage reduction ranges between 2 to 10 % for all cases. For majority of cases joint approach offers less percentage reduction of PESQ in contrast with its counterpart fixed approach. It is truly fact that overall percentage reduction in PESQ above 10 % for the case of both ETSI and Proposed GSM FR coders are not advisable for its recovered speech quality performance, hence it imposes limit on upper bound of selection of steganographic mode not beyond 2.75 kbps.

Practically, there exists a trade off between obtaining and maintaining comparable recovered speech quality by compromising upper bound of steganographic selection of bitrate mode (in this case 2.75 kbps) for information embedding and hiding or vice versa.

4.3 Computational comparison analysis of simulation delay

Tic and Toc commands in MATLAB are explored to calculate the total simulation time taken by simulation algorithm for embedding steganographic data frame-wise into narrow band bitstream, decoding and data extraction at receiver. In aggregate, joint approach takes more simulation time for both the coders. Simulation time for both the coders have been examined and despite the fact that Standard coder offers comparative results for both subjective and objective tests, it takes 1.36 times more simulation time (an average of all bitrate modes and for all wave files) compared to Proposed coder for the same simulations. Proposed grid selection strategy plays a major role for the reduction of simulation time (for all cases) in proposed coder. At this juncture of time a prerequisite to be considered is that throughout the above discussed analysis, data file and its length to be embedded for steganographic transmission, has to be made constant and fix. Further, because of inherent implementation complications, joint approach reflects into more simulation time and increment in delay time is a proportional element with respect to increase in bitrate modes. In case of real time implementation (which is not implemented in this research) on any digital signal processor, for the analysis of any given bitrate mode, proposed coder may offer less execution and algorithmic delay time and hence less complexity (in MIPS) in comparison with its counterpart.

5 Discussions and concluding remarks

This research focuses on two parallel implementation phases and their performance cross-comparisons. This work uti-

lizes few modifications suggested in grid selection strategy to produce Proposed GSM FR coder (which in turn offers the parent steganographic bitrate mode of 1.8 kbps) for steganographic bitstream transmission. Further, research investigates few GSM encoder parameters (selected RPE pulses from class I_b as per GSM 05.03 standards) for embedding and hiding variable bitrate steganographic information bitstream (depending upon selected bitrate mode between 2.05 kbps and 2.75 kbps) in each transmitted frame of Proposed GSM FR coder having effective bitstream of 260 bits in 20 ms time frame.

Then, in order to implement and execute the same all five steganographic bitrate modes over ETSI standard GSM FR coder, once again some encoder parameters are chosen from class I_b. Selection of such parameters solely dependent upon their subjective importance so that embedding and hiding over the bits of those encoder parameters affect the received speech quality the least. In this research, embedding and extraction of small text and image files have successfully been conducted for all steganographic bitrate modes over six different wave files as a cover signals for both Standard and Proposed GSM FR coders.

This study, implementation and analysis reveal the tradeoff between speech quality and embedding capacity that in fact impose an upper bound on selection of highest steganographic bitrate mode (here 2.75 kbps) along with acceptable recovered speech quality. As can be witnessed from both PESQ (objective) and MOS as well as DMOS (subjective) analysis that almost for all cases of wave files and for all bitrate modes gradual reduction in speech quality is quite evident with reference to proportional increment in embedding bitrate modes from 1.8 kbps to 2.75 kbps.

As depicted from the analysis carried out for both the approaches, as a whole joint approach (for both coders) performs slightly better in terms of recovered speech quality but at the expense of higher simulation delay. As far as comparison between Standard and Proposed coder are concerned, Objective and Subjective analysis results obtained for Proposed Coder (for all bitrate modes and for all wave files) are quite comparable with Standard coder. While computing

maximum percentage reduction in PESQ scores, the range of percentage reduction is found between 2 % and 10 % for all cases. For both fixed and joint approaches, maximum percentage reductions in PESQ scores were quite comparable between Standard and Proposed coders. Moreover because of the inherent structural benefit of proposed coder, simulation time taken by proposed coder is significantly lesser in stark contrast with Standard coder for all bitrate modes. Though not touched upon in this research, if both coders are implemented in real time on any digital signal processor, the overall algorithmic delay and computational complexity in the case of Proposed coder can be found less compared to Standard coder.

References

Bhatt, N., & Kosta, Y. (2011). Proposed modifications in ETSI GSM 06.10 full rate speech codec and its overall evaluation of performance using MATLAB. *International Journal of Speech Technology*, *14*(3), 157–165.

ETSI (1998). *Digital cellular telecommunications system (phase 2+), full rate speech, transcoding* (GSM 06.10 version 7.0.0 Release 1998), pp. 10–59.

ETSI (1999). *Channel coding* (GSM 05.03 version 8.9.0 (2005-01), release 1999); 12–19 & 98.

Geiser, B., & Vary, P. (2008). High rate data hiding in ACELP speech codecs. In *Proc. of ICASSP-2008*.

Hu, Y., & Loizou, P. C. (2008). Evaluation of objective quality measures for speech enhancement. *IEEE Transactions on Audio, Speech, and Language Processing*, *16*(1), 229–238.

Hu, L., & Wang, S. (2006). Information hiding based on GSM full rate speech coding. In *Proc. of WiCOM IEEE conference*, Wuhan, Sept. 2006.

ITU-T Recommendation P.862 (2001). *Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, February 2001; 1–18.

Malkovic, D. (2003). Speech coding methods in mobile radio communication systems. In *17th international conference on applied electromagnetics and communications*, Oct. 2003, Croatia.

Shahbazi, A., Soltanmohammadi, E., Rezaie, A. H., Sayadiyan, A., & Mosayyebpour, S. (2010). Content dependent data hiding on GSM full rate encoded speech. In *International conference on signal acquisition and processing*.

The NIOZEUS database (2009). Available on http://www.utdallas.edu/~loizou/speech.