

SPLIT MULTI-STAGE VECTOR QUANTIZATION BASED STEGANOGRAPHY FOR SECURE WIDEBAND SPEECH CODER

Merouane BOUZID and Bakkar LASKAR

Speech Communication and Signal Processing Laboratory,
University of Sciences and Technology Houari Boumediene (USTHB),
Electronics Faculty, P.O. Box 32, El-Alia, Bab-Ezzouar, Algiers, 16111, Algeria

ABSTRACT

Speech steganography is a technique of covert communication which conveys secret speech hidden in cover digital speech signal in such a way that the existence of the secret speech is concealed. In this paper, we develop a steganographic speech coding system based on embedding coded secret speech into host public speech coded by the AMR-WB (ITU-T G.722.2) speech coder. For the compression of the secret speech signal, we used the 2.4 kbits/s MELP speech coder. The embedding process of the secret bit stream is carried out into the split-multistage vector quantization (S-MSVQ) indices of G.722.2 immittance spectral frequencies (ISF) by modifying the mechanism of the S-MSVQ second stage.

KEYWORDS

Multi-stage vector quantization, steganography, data hiding, ISF parameters, secure speech, wideband speech coder, MELP, AMR-WB

1. INTRODUCTION

Steganography is the art of sending secret information in a cover media without arousing suspicion. Indeed, modern steganography techniques exploit the characteristics of digital media by using them as carriers (covers) to hold hidden information. Thus, the sender embeds secret information in a digital cover file to produce a stego-file, in such a way that the contents of hidden data and its existence cannot be detected by an observer during the transmission process [1]. The secret information can be extracted only at the authorized user's side.

In this work, we focalize particularly on speech steganography techniques which consist in hiding a secret speech signal into a cover (host) signal. A variety of speech/audio steganography methods have been proposed in the past, where most of them are based on the temporal domain, the transform domain and the compression domain. An extended review of the current state-of-art literature in digital audio/speech steganography techniques in each domain is given in [2]. In compression domain, speech steganography techniques based on vector quantization (VQ) have been getting more and more popular, since they enhance the conventional VQ coding by adding the possibility of data hiding.

In [3], Chang and Yu proposed a dither-like data hiding method to embed hidden data in the multistage vector quantizer (MSVQ) of the Mixed-Excitation Linear Predictive (MELP) and ITU-T G.729 speech coders. In [4], Geiser and Vary developed a steganographic method to embed digital data in the bitstream of an ACELP speech coder. In [5], Laskar and Bouzid proposed two

variants of VQ-based speech steganography binning schemes (SBS) for G.722.2 secure speech communication system. They showed that the two steganographic SBS methods carried out by balanced and unbalanced VQ codebook partitioning can generate stego-speech signals with similar quality to cover speech signals. In [6], an AMR-WB speech steganography system was proposed based on diameter-neighbor codebook partition method. It was shown that speech steganographic system can provide higher and flexible embedding capacity without noticeable decrease in speech quality and better performance against statistical steganalysis.

Since the Adaptive Multi-rate Wideband AMR-WB (Rec. G.722.2) [7], [8] speech coder still a good candidate for cover medium in speech steganography, we develop in this paper a steganographic AMR-WB coding system based on the dither-like data hiding idea. It's about modifying the mechanism of the second stage of the split-multistage vector quantizer (S-MSVQ) of G.722.2 immittance spectral frequencies (ISF) parameters to embed a secret speech coded by the 2.4 kbits/s MELP [9] speech coder.

An outline of this paper is as follows. In section 2, we first review briefly the basics of the conventional VQ, the split vector quantizer (SVQ) and the MSVQ. Then, we present the dither-like data hiding method applied on the MSVQ scheme. In section 3, we describe the design principle of a steganographic G.722.2 speech coding system developed according to the S-MSVQ based data hiding method. Experimental results are provided in section 4 to evaluate the performance of our speech steganographic system. Conclusions are given in section 5.

2. DITHER-LIKE DATA HIDING ON MSVQ QUANTIZER

Several data hiding methods, based on conventional VQ, have been proposed in literature [1], [10]. One of the most popular quantization-based data hiding method is probably the quantization index modulation (QIM) [11].

Before presenting the dither-like data hiding method applied on MSVQ scheme, let us first review briefly the basics of the conventional VQ, the split vector quantizer (SVQ) and the MSVQ.

2.1. Basic principle of VQ, SVQ and MSVQ schemes

A k -dimensional VQ of rate R bits/sample (bps) is a mapping of k -dimensional Euclidean space \mathcal{R}^k into a finite codebook $Y = \{y_0, \dots, y_{L-1}\}$ composed of $L = 2^{kR}$ codevectors [12]. The design principle of a VQ consists of partitioning the k -dimensional space of source vectors x into L non overlapping cells $\{R_0, \dots, R_{L-1}\}$ (partition) and associating with each cell R_i a unique codevector y_i such that the total average distortion D is minimized [12]. Various algorithms for the optimal design of VQ have been developed in the past. The most popular one is certainly the LBG algorithm [12]. This algorithm is an iterative application of the two optimality (nearest neighbor and centroid) conditions such as the partition and the codebook are iteratively updated.

In other hand, an N part k -dimensional SVQ (noted N -SVQ) is composed of N classical VQs of smaller sizes and dimensions [13]. Its basic principle consists of partitioning the set of the training base vectors x of dimension k in N subsets of sub-vectors of smaller dimension k_i (with $\sum_{i=1}^N k_i = k$). Then, for each part, the corresponding VQ codebook will be designed by using the LBG-VQ algorithm. Compared to a conventional unstructured k -dimensional VQ, of rate R bps and size $L = 2^{Rk}$, an N -SVQ is thus composed of N codebooks of smaller sizes $L_i = 2^{R_i k_i}$ (where $L = \prod_{i=1}^N L_i$ and R_i is the partial rate in bps). Figure 1 shows a bloc-diagram of an N -SVQ quantizer.

Concerning the conventional MSVQ, we can say that it is a kind of cascaded VQ where the output of one stage is given as an input to the next stage and the bit rates used for quantization are divided among all successive MSVQ stages [14], [12]. The first MSVQ stage performs a

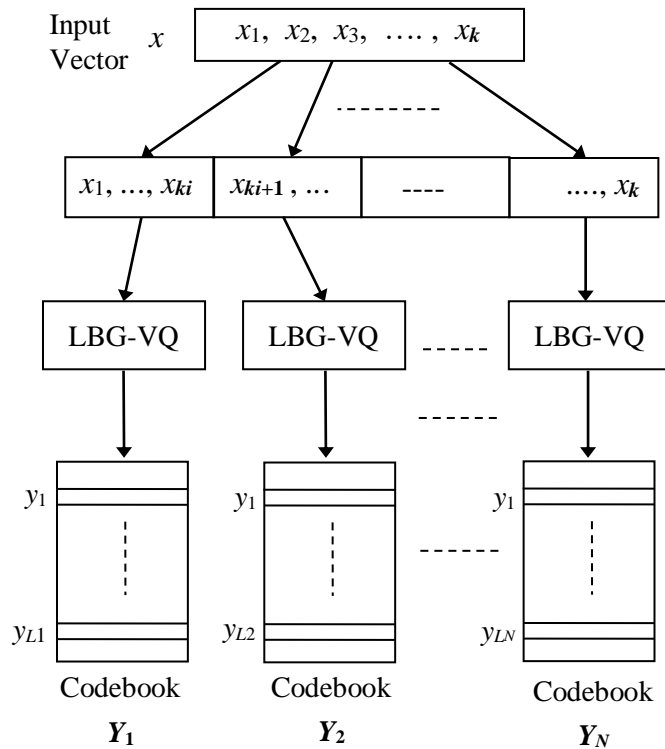


Figure 1. Bloc diagram of N -SVQ quantizer

VQ quantization of the input vector. Then, the second stage operates on the error vector between the original input vector and quantized first stage output. Practically, the quantized error vector (called residual) provides a second approximation to the original input vector leading to a more accurate representation of the input. A third stage may then be used to quantize the second stage error vector to provide further accuracy and so on. The final quantized version of the input vector is obtained by summing the output codevectors of all stages. The coding bit rate R of an M stages MSVQ is the sum of bit rates allocated to each MSVQ stage ($R = \sum_{i=1}^M R_i = \sum_{i=1}^M \log_2 L_i$). Figure 2 presents an example of a two stages MSVQ encoder/

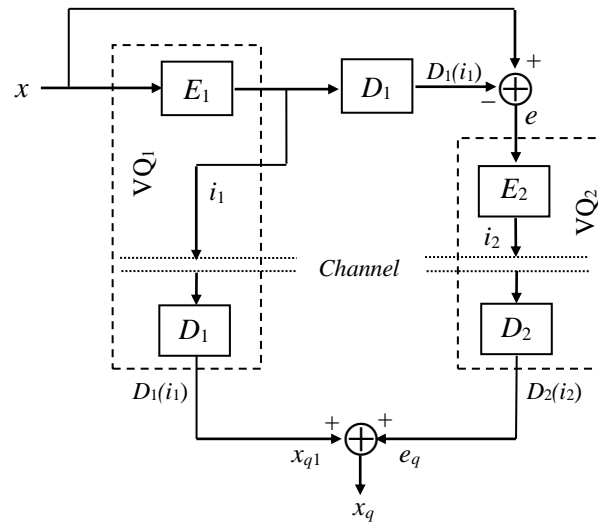


Figure 2. Two stages MSVQ encoder/decoder

decoder, where the VQ_1 of the MSVQ first stage includes the pair of the encoder E_1 and the decoder D_1 . The MSVQ quantized version of the input vector x is given by: $x_q = D_1(i_1) + D_2(i_2) = x_{q1} + e_q$.

2.2. Dither-like data hiding

Dithered quantization is a well-known technique used to reduce or to eliminate the statistical dependence between the original signal and quantization error. This is most often achieved by adding (pseudo-) random noise signal (called dither signal) to the original input signal prior quantization [11], [15].

In subtractive dithering, the dither signal is further subtracted from the quantizer's output. Thus, the total quantization error can be rendered statistically independent of the input signal as well as rendering error samples separated in time statistically independent of one another. This ensures that the power spectrum of the total error is independent of the system input, and that it is spectrally flat (white) even if the dither signal is not.

In a non-subtractive dithered system, this subtraction operation is omitted. In the subtractive dither-like data hiding (noted here SDDH) method proposed by Chang and Yu [3], the binary last stage codevector index of an MSVQ scheme are replaced with secret data bits. Thus, the last stage codevector, which is indexed now by the secret bits, is first subtracted from the original input vector to be quantized before running the MSVQ.

The SDDH idea [3] comes from the fact that in an MSVQ scheme the signals in last stages tend to be less correlated [12]. Consequently, if the codevector binary index of the last stage is replaced by the secret bits sequence to be hidden, the last stage can be viewed as a random noise that generates uncorrelated data with previous stages, which is the same as subtractive dithering [11]. By subtracting this noise data from the input of the MSVQ encoder, and adding it back at the MSVQ decoder, the degradation caused by hiding secret data can be reduced compared to the traditional non-subtractive dither (noted here NDDH) system. Let us note that in the case of conventional NDDH system, the secret data bits replace simply the MSVQ last stage binary index without vector subtraction at the first stage.

Examples of conventional NDDH and SDDH schemes are shown respectively in Figure 3-(a) and Figure 3-(b). The data hiding systems are applied to a two stages MSVQ with the second

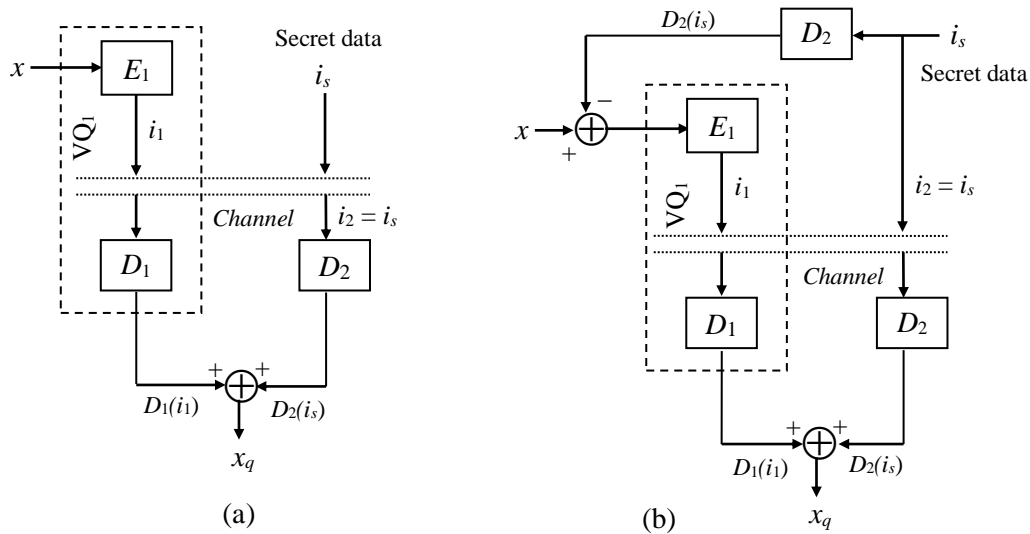


Figure 3. Two stages MSVQ dither-like data hiding: (a)- SDDH scheme, (b)- NDDH scheme

stage index used to hide secret data i_s . In the SDDH scheme, the MSVQ second stage codevector y_{is} (i.e., $D_2(i_s)$) indexed by secret sequence i_s is first extracted and subtracted from the input vector to be quantized x . Then, the first stage MSVQ is run as usual with $x - D_2(i_s)$ as input vector. The second stage MSVQ is never operated. The second codevector index i_2 , supposed to be delivered by this stage, is replaced by secret sequence i_s . Thus, the decoder receives the indices i_1 and $i_2 = i_s$ and performs exactly the same procedure as a MSVQ decoder to reconstruct the quantized version of x : $x_q = D_1(i_1) + D_2(i_s)$. At the same time, the secret sequence i_s is obtained as i_2 .

3. SPEECH STEGANOGRAPHIC SYSTEM: APPLICATION OF THE G722.2 S-MSVQ DATA HIDING SCHEME

In this section, we present a steganographic AMR-WB (G.722.2) speech coding system developed according to the S-MSVQ based data hiding method. Its basic principle consist in modifying the mechanism of the second stage of the S-MSVQ of G.722.2 ISF parameters to embed a secret speech coded by the 2.4 kbits/s MELP coder. Before presenting our speech steganographic system, let us first review briefly the S-MSVQ scheme principle.

3.1. Split MSVQ

The S-MSVQ is a hybrid scheme based on a MSVQ scheme combined with SVQ. Indeed, the S-MSVQ is a modified MSVQ with N -SVQ stages. It is structured in several successive stages, where each stage is represented by an N -SVQ instead of a simple VQ as in the conventional MSVQ. An example of a two stages S-MSVQ scheme is given in Figure 4.

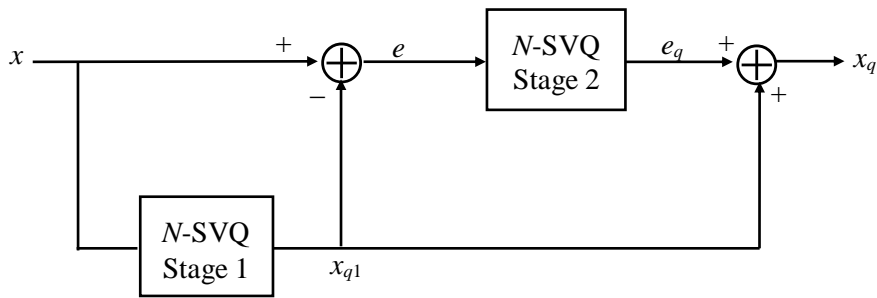


Figure 4. Two stages S-MSVQ scheme

The input vector x is first quantized by the N -SVQ of the first stage. Then, the quantization error (residual) e is used as an input to the second stage N -SVQ to obtain the quantized version e_q of the first stage residual error e . The final quantized version of x is simply the sum of the two output codevectors x_{q1} and e_q . Notice that the total number of bits allocated for quantization is divided among the S-MSVQ stages and the N split regions of each stage.

3.2. G.722.2 S-MSVQ-based data hiding scheme

Recall that the G.722.2 ISF parameters are quantized by a two stages S-MSVQ with 1st order MA predictor [7]. The standard G.722.2 S-MSVQ uses seven VQ codebooks, where two codebooks at the first stage (named here CB_{11} and CB_{12}) and five codebooks (named CB_{21} , CB_{22} , CB_{23} , CB_{24} , CB_{25}) at the second stage. Notice that the G.722.2 S-MSVQ works at 36 bits/frame for the lowest bit rate mode 0 (6.6 Kbits/s); and at 46 bits/frame for the other eight higher bit rate modes (8.85 to 23.85 Kbits/s).

For the standard 46 bits/frame S-MSVQ, 16-dimensional residual ISF vector $f_r = (f_r^1, \dots, f_r^{16})$ is split into two subvectors of dimension 9 ($f_{r1} = (f_r^1, \dots, f_r^9)$) and 7 ($f_{r2} = (f_r^{10}, \dots, f_r^{16})$), respectively. The 2 subvectors are then quantized in two stages. In the first stage, each subvector is quantized using 8 bits. In the second stage, the two quantization error subvectors $e_1 = f_{r1} - \hat{f}_{r1}$ and $e_2 = f_{r2} - \hat{f}_{r2}$ are split respectively into 3 and 2 subvectors according to the part divisions (3-3-3) and (3-4). The bit allocation for each subvector in the second stage are (6, 7, 7) bits and (5, 5) bits, respectively.

In our speech steganographic G.722.2 S-MSVQ-based data hiding system developed according to the SDDH principle, the codevectors binary indices of anyone one of the five second stage codebooks can be used to hide the secret bits sequences. The binary indices of the last stage are replaced simply by the secret bits sequence to be hidden. For a comparative evaluation, we developed also a speech steganographic G.722.2 system based on the NDDH approach.

It is important to note that in our steganographic G.722.2 systems, we can use a combination of more than one second stage codebook to perform the data hiding. Thus, the five codebooks CB_{21} , CB_{22} , CB_{23} , CB_{24} , CB_{25} can all be used in hiding process. Thus, some (or all) of the bit rate allocated to the second stage G.722.2 S-MSVQ can be used to embed secret bits. Figure 5 present an example of speech steganographic G.722.2 system where the second stage S-MSVQ CB_{25} is used for hiding secret bits sequence i_s . Notice that the bloc VQ_j in the S-MSVQ includes the pair of the encoder E_j and the decoder D_j .

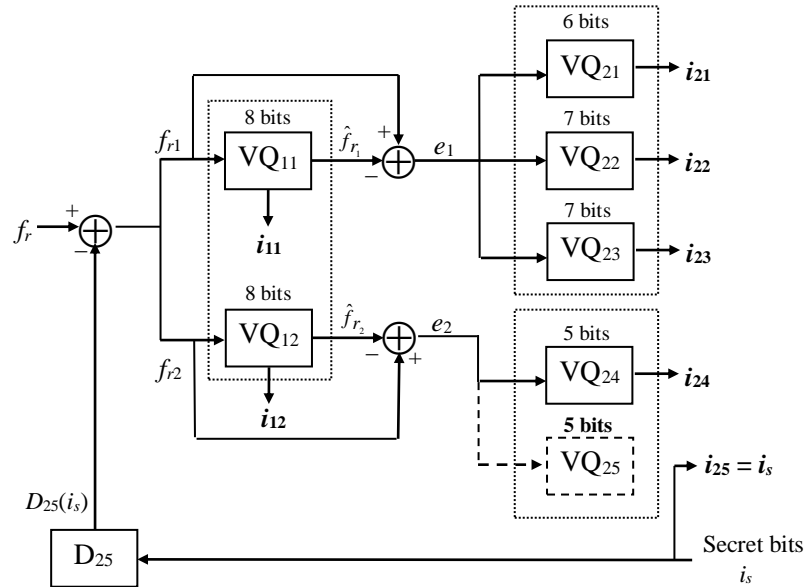


Figure 5. Example of steganographic G.722.2 S-MSVQ where the VQ CB₂₅ is used for hiding

4. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of our steganographic G.722.2 speech coding systems, designed based on modifying the mechanism of the second stage of the G.722.2 S-MSVQ quantization of ISF parameters (ISFs). The data hiding S-MSVQ modification concept was carried out according to the basic idea of SDDH and NDDH approaches. The steganographic systems were called respectively S-MSVQ-SDDH and S-MSVQ-NDDH.

In our applications, the main purpose is to hide a secret speech signal coded by the 2.4 kbps MELP into a host public speech coded by the G.722.2. Notice that in all simulations, we used the G.722.2 in mode 12.65 kbits/s where the ISFs are coded by an S-MSVQ of 46 bits/frame.

4.1. Performance evaluation criteria

Performance evaluation of the implemented speech steganographic systems will be done according to the hiding capacity represented by the embedding rate of the secret speech and to the transparency (imperceptibility) represented by the perceptual quality of the speech stego-signal synthesized by the G.722.2 with embedding procedure.

The total embedding rate is given by the ratio of the number of hidden secret bits and the length of the host speech coder frame (i.e., 20 ms in G.722.2). Let us note that in our steganographic systems, the embedding rate is variable according to the combination of codebooks used in the data hiding process. Table 1 gives the embedding rates (in bits/frame and in bits/s) when using each second stage S-MSVQ codebook individually.

Table 1. Embedding rates of steganographic S-MSVQ systems

Used Codebooks	Embedding rate (bits/frame)	Embedding rate (bits/s)
CB ₂₁	6 bits	300
CB ₂₂	7 bits	350
CB ₂₃	7 bits	350
CB ₂₄	5 bits	250
CB ₂₅	5 bits	250

It should be noted that the real total embedding rate is the sum of the individual embedding rates of the codebooks used in combination in the embedding process. If we use, for example, the binary indices of the VQ codebooks CB₂₁ and CB₂₂ to hide the secret bits sequence, the embedding rate is then equal to 650 bits/s. Thus, according to the possible codebook combinations, we can use several embedding bit rates ranging from 250 bits/s (minimum embedding rate) to 1500 bits/s (maximum embedding rate).

On the other hand, for imperceptibility, we use the ITU-T Rec. P.862.2 known under the abbreviation WB-PESQ (Wide Band extension of Perceptual Evaluation Speech Quality) [16] to evaluate the coded cover/stego speech signals quality. The hidden speech signal is imperceptible if a listener is unable to distinguish between the cover and the stego speech signals; which means that the WB-PESQ difference between the two cover/stego signals is negligible.

The performance of the steganographic S-MSVQ quantizer will be also evaluated by the well-know average spectral distortion (SD) measure. The spectral distortion of each frame i is given, in decibels, by [13], [17]:

$$SD_i = \sqrt{\frac{1}{n_1 - n_0} \sum_{n=n_0}^{n_1-1} \left[10 \log_{10} \frac{S(e^{j2\pi n/N})}{\hat{S}(e^{j2\pi n/N})} \right]^2}, \quad (1)$$

where $S(e^{j2\pi n/N})$ and $\hat{S}(e^{j2\pi n/N})$ are respectively the original and quantized power spectra of the LPC synthesis filter, associated with the i^{th} frame of speech signal.

Generally, we can get transparent quantization quality if we maintain the three following conditions [13]: 1)- The average spectral distortion (SD) is about 1 dB, 2)- No Outliers frames with SD greater than 4 dB, 3)- The percentage of Outlier frames having SD within the range of 2-4 dB must be less than 2%.

4.2. Performances of steganographic S-MSVQ coding systems

For each embedding rate, we performed an optimization procedure of our steganographic systems. It consists in finding the best choice of second stage S-MSVQ codebooks that can be combined in the hiding process to obtain the best possible performance.

The speech database used in the experiments consists of 60 minutes of speech taken from the international TIMIT database ($f_s = 16$ kHz) [18]. To construct the ISF database, we used the same LPC analysis function of the G.722.2, where a 16-order LPC analysis, based on the autocorrelation method, is performed every analysis frame of 20 ms. Thus, a database of 180000 ISF vectors was constructed.

For embedding rates varying between 5 and 30 bits/frame, the SD performances of speech steganographic G.722.2 S-MSVQ-SDDH and S-MSVQ-NDDH coding systems are shown in Table 2, where the secret bits sequences are generated randomly.

For a given embedding rate, the VQ codebooks noted in the table are only the second stage codebooks (CB₂₁, CB₂₂, CB₂₃, CB₂₄, CB₂₅) in which "1" means that the corresponding codebook is used in the embedding procedure. The bit rate of this VQ codebook is then reserved for the secret bits sequence to be hidden. For example, the notation "18 (1-0-1-0-1)" means that for an embedding rate of 18 bits/frame, the codebooks CB₂₁, CB₂₃ and CB₂₅ of the modified S-MSVQ second stage were selected as best choice to be used in hiding 18 bits per each frame.

Table 2. Performance of steganographic S-MSVQ-SDDH and S-MSVQ-NDDH systems

Embedding rate (Bits/frame)	S-MSVQ-NDDH systems			S-MSVQ-SDDH systems		
	Av. SD (dB)	Outliers (in %)		Av. SD (dB)	Outliers (in %)	
		2 - 4 dB	> 4 dB		2 - 4 dB	> 4 dB
5 (0-0-0-0-1)	1.65	21.63	0.08	1.48	14.82	0.06
6 (1-0-0-0-0)	2.34	60.79	3.20	2.03	46.36	1.64
7 (0-1-0-0-0)	1.92	39.39	1.11	1.79	31.98	0.64
10 (0-0-0-1-1)	2.20	58.38	0.78	1.92	39.05	0.40
11 (1-0-0-0-1)	2.72	75.34	6.40	2.35	62.14	2.70
12 (0-0-1-0-1)	2.39	65.97	2.53	2.13	51.38	1.37
14 (0-1-1-0-0)	2.61	71.19	5.66	2.40	64.21	3.31
16 (1-0-0-1-1)	3.09	80.70	12.65	2.67	73.41	5.93
17 (0-1-0-1-1)	2.81	81.93	6.35	2.48	69.77	3.20
18 (1-0-1-0-1)	3.27	76.25	18.82	2.84	75.79	9.05
20 (1-1-1-0-0)	3.41	71.99	24.02	3.10	76.91	14.50
23 (1-1-0-1-1)	3.57	69.77	28.64	3.12	77.40	14.95
24 (0-1-1-1-1)	3.29	79.05	17.93	2.95	80.70	9.85
25 (1-1-1-0-1)	3.67	65.67	33.00	3.29	76.03	19.30
30 (1-1-1-1-1)	3.98	53.12	46.57	3.54	69.37	28.53

These results show that the SD performance degradation due to embedding process is not proportional to embedding rate. For example, for an embedding rate of 10 bits/frame, the SD degradation caused by hiding in the last second stage codebooks CB₂₄ and CB₂₅ binary indices is less than that caused by hiding in the first codebook CB₂₁ binary indices of the 6 bits/frame case. Indeed, the degradation is rather related to the importance of the used codebook in frequency domain. Knowing that the human auditory system (HAS) is more sensitive in low frequencies bands, therefore the codebooks which represent the high frequencies are less important than those of the low frequencies.

On the other hand, these SD comparative results show that the steganographic S-MSVQ-SDDH system outperform the S-MSVQ-NDDH coding system.

4.3. Performance evaluation of speech steganographic G.722.2 with ISFs quantized by modified S-MSVQ

The cover public speech database used in the following evaluations is composed of 10 speech sequences of 32s extracted from the same TIMIT database. The secret bit stream was generated by the 2.4 kbps MELP from a speech sequence of $f_s = 8$ kHz extracted from a phonetically balanced Arabic speech database [19].

Table 3 presents WB-PESQ performance comparative evaluation of the global G.722.2 where its ISF parameters were quantized by the 46 bits/frame steganographic S-MSVQ in which the second stage structure is modified according to the basic concept of SDDH and NDDH, respectively. Notice that an embedding rate of 0 bits/frame means the original standard G.722.2 without steganography (i.e assessment of the cover speech signal). Note also that for each embedding rate, the best choices of the used steganographic second stage S-MSVQ codebooks are the same as those mentioned in Table 2.

Table 3. Performance of the global speech steganographic G.722.2 coding system

Embedding rate (Bits/frame)	G.722.2 with S-MSVQ-NDDH	G.722.2 with S-MSVQ-SDDH
	WB-PESQ	WB-PESQ
0	3.790	3.790
5	3.775	3.738
6	3.265	3.321
7	3.684	3.728
10	3.651	3.651
11	3.233	3.337
12	3.665	3.700
14	3.556	3.568
16	3.210	3.279
17	3.579	3.574
18	3.224	3.312
20	3.054	3.110
23	3.093	3.161
24	3.480	3.553
25	3.072	3.105
30	3.033	3.139

These simulation results show that for some embedding rates (5, 7, 10, 12, 14, 17 and even 24 bits/frame) the overall quality of stego-speech is almost identical to quality of cover public speech; which means that developed steganographic S-MSVQ-SDDH and S-MSVQ-NDDH techniques are practically imperceptible. Most WB-PESQ scores of the stego-signals are higher than 3.55. Hence, a good speech quality was obtained and no perceptual degradation was caused by the embedding process.

On the other hand, steganographic S-MSVQ-SDDH system yields slight improvement to the G.722.2 WB-PESQ performance compared to steganographic S-MSVQ-NDDH system.

5. CONCLUSIONS

In this paper, we developed a steganographic S-MSVQ quantizer for G.722.2 secure speech communication system. The embedding process of secret bits was carried out into the second stage S-MSVQ indices of G.722.2 ISFs according to the basic idea of subtractive (non-subtractive) dither-like data hiding. The global steganographic speech coding system was then based on embedding MELP coded secret speech into host public speech coded by the AMR-WB (ITU-T G.722.2) speech coder.

The simulation results showed that when the G.722.2 second stage S-MSVQ sub-codebooks of high frequencies bands are involved in the embedding process, our steganographic S-MSVQ-

SDDH and S-MSVQ-NDDH systems are practically imperceptible. Indeed, for some embedding rates (5, 7, 10, 12, 14, 17 and even 24 bits/frame), the G.722.2 (with S-MSVQ-SDDH) can generate stego-speech signals with similar quality to cover speech signals. Hence, the developed steganographic S-MSVQ-SDDH system can ensure a good transparency with a maximal embedding rate of 24 bits/frame (1200 bits/s). On the other hand, we can reach a maximum embedding capacity of 1500 bits/s but with a significant degradation in terms of SD and WB-PESQ. Robustness against intentional and non-intentional attacks has not been investigated in this work; it will be studied in future work.

REFERENCES

- [1] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, T. Kalker. *Digital Watermarking and Steganography*, Second Edition, Morgan Kaufmann Publishers, USA, 2008.
- [2] F. Djebbar, B. Ayad, K. A. Meraim, H. Hamam, "Comparative study of digital audio steganography techniques," *EURASIP Journal on Audio, Speech, and Music Processing*, Springer, vol. 25, pp. 1-16. 2012.
- [3] P. C. Chang, H. M. Yu, "Dither-like data hiding in multistage vector quantization of MELP and G.729 speech coding," *Thirty-Sixth Asilomar Conf. on Signals, Systems and Computers*, Monterey, CA, vol. 2, 2002, pp. 1199–1203.
- [4] B. Geiser, P. Vary, "High rate data hiding in ACELP speech codecs," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2008)*, Las Vegas, Nevada, USA, March 30-April 4, pp. 4005-4008.
- [5] B. Laskar, M., Bouzid, "Vector quantization based steganography for secure speech communication system," in *Proc. 14th International Conference on Security and Cryptography (SECURITY 2017)*, vol. 4, 24-26 July 2017, Madrid, Spain, pp. 407-412. Available: <https://www.scitepress.org/Papers/2017/63983/63983.pdf>
- [6] J. He, J. Chen, S. Xiao, X. Huang, and S. Tang, "A Novel AMR-WB Speech Steganography Based on Diameter-Neighbor Codebook Partition," *Security and Communication Networks*, vol. 2018. DOI:10.1155/2018/7080673, 2018.
- [7] B. Bessette, R. Salami, R. Lefebvre, M. Jelínek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, K. Järvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 8, pp. 620-636, 2002.
- [8] ITU-T Recommendation G.722.2. *Wideband coding of speech at around 16 kb/s using Adaptive Multi-rate Wideband (AMR-WB)*, 2003.
- [9] A. McCree, K. Truong, E. B. George, T. P. Barnwell, V. Viswanathan, "A 2.4 kbits/s MELP Coder Candidate for the New U.S. Federal Standard," in *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP'96)*, 1996, pp. 200-203.
- [10] P. Moulin, R. Koetter, "Data-Hiding Codes," in *Proceedings of The IEEE*, vol. 93, pp. 2083-2126, 2005.
- [11] B. Chen, G. W. Wornell, "Quantization index modulation methods: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.
- [12] A. Gersho, R. M. Gray, *Vector quantization and Signal compression*, Kluwer Acad. Publishers, USA, 1992.

- [13] K. K. Paliwal, B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 1, pp. 3-14, 1993.
- [14] B. H. Juang, A. H. Gray, "Multiple Stage Vector Quantization for Speech Coding," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'1982)*, Paris, France, 1982, pp. 597-600.
- [15] S. P. Lipshitz, R. A. Wannamaker, J. Vanderkooy, "Quantization and Dither: A Theoretical Survey," *J. Audio Eng. Soc.*, vol. 40, no.5, pp.355-375, May 1992.
- [16] ITU-T Recommendation P.862.2. Wideband Extension to Recommendation P.862 for the Assessment of Wideband Telephone Networks and Speech Codecs, Geneva, 2005.
- [17] S. Cheraitia, M. Bouzid, "Robust coding of wideband speech immittance spectral frequencies," *Speech Communication*, Elsevier, vol. 65, pp. 94-108, July 2014.
- [18] J. S. Garofolo et al., *DARPA TIMIT Acoustic-phonetic Continuous Speech Database*. National Institute of Standards and Technology (NIST), Gaithersburg, October 1988.
- [19] M. Boudraa, B. Boudraa, B. Guerin, "Mise en place de phrases arabes phonétiquement équilibrées," in *Proc. of XIX^{èmes} Journées d'Etude sur la Parole (JEP'92)*, Bruxelles, 1992.