Advances in Modelling and Analysis B
Vol. 61, No. 2, June, 2018, pp. 80-87
Journal homepage: http://iieta.org/Journals/AMA/AMA_B

A psychoacoustic model and a Filter Bank Design using optimization for speech compression

Talbi Mourad[1*], Med S. Bouhlel[2]

[1] Laboratoire des Semi-Conducteurs, Nanostructures et Technologie Avancée, Center of Researches and Technologies of Energy of Borj Cedria, Tunis 952050, Tunisia
[2] Sciences Electroniques, Technologie de l'Information et Télécommunications (SETIT), Sfax BP 11693029, Tunisia

Corresponding Author Email: talbi1969@yahoo.fr

https://doi.org/10.18280/ama_b.610205

Received: 7 April 2018
Accepted: 8 June 2018

Keywords:
speech compression, psychoacoustic model, Filter Bank, optimization, bits before/bits after compression

ABSTRACT

In this paper we propose a new speech compression technique employing psychoacoustic model and a general approach for Filter Bank Design using optimization. This technique is inspired from an audio compression technique using psychoacoustic model and a Modified Discrete Cosine Transform (MDCT) filter banks of 32 filters. In fact, in this proposed approach, we have used Uniform/Non-Uniform Filter Bank (which is designed using optimization) instead of a MDCT filter banks of 32 filters. The two techniques are evaluated and compared with each other by computing bits before and after compression. They are tested and applied to different speech signals. The simulation results obtained from the computation of the compressed files size and the Compression Ratios (CR), show that the proposed technique outperforms the second one. In term of perceptual speech quality, the outputs speech signals of the proposed compression system are with good quality. This is justified by the computation of SNR (Signal to Noise Ratio), PSNR (Peak Signal to Noise Ratio), NRMSE (Normalized Root Mean Square Error) and PESQ (Perceptual evaluation of speech quality). We have also compared the proposed technique to one previous research work which is a speech compression technique based on Discrete Wavelet Transform (DWT) and integrating a Voice Activity Detection (VAD) Module. This comparison is also based on the computation of SNR, PSNR, NRMSE, PESQ and CR and the obtained results show that the proposed technique outperforms this third technique based on DWT and VAD.

1. INTRODUCTION

The essential purpose in speech compression consists in representing with minimum number of bits, the digital speech waveform while preserving its perceptual quality [1-2]. The speech compression [3] is essential either for reducing memory storage requirements or for reducing transmission bandwidth requirements, without harming the speech quality. For example, digital cellular phones use some compression techniques for compressing, in real-time, the speech signal over general switched telephone networks. Speech compression is also needed for reducing the storage requirements for storing the voice messages or for mail forwarding of voice messages. All these applications depend on the efficiency of the speech compression technique. Consequently, in the past, different techniques [4] were developed to meet the rising demand for better algorithms of speech compression. Alike to other digital data compression techniques, speech compression ones can also be classified into two categories which are lossy compression and lossless compression. Lossless compression is frequently performed by waveform coding techniques. In these techniques [5] actual shape of the signal produced by the microphone and its associated circuits is conserved. A most popular waveform coding technique is pulse code modulation (PCM). Other lossless techniques such as differential quantization and adaptive PCM, make speech signals compression by localizing redundancy and optimizing or suppressing it by the quantization process. All such techniques require simple signal processing and lead to minimum distortion with small compression [6-7]. A detailed study on these techniques is presented in [7-8] and the references therein. Concerning lossy compression, the compressed data is a close approximation of the original data and not the same as it. Although, it leads to a much higher compression ratio than that of lossless compression. The literature review reveals that a considerable progress has been made on lossy compression techniques such as sub-band coding [8], linear predictive coding (LPC) [4] and turning point [5]. In subband decomposition, spectral information is divided into a set of signals that can then be encoded by using a diversity of techniques. Based on subband decomposition, different techniques have been devised for speech compression [9].

In this paper, we propose a new speech compression technique employing psychoacoustic model and a general approach for Filter Bank Design using optimization [13]. In fact this technique is inspired from an audio compression one proposed by Alex et al. [10-12] and this by replacing a Modified Discrete Cosine Transform (MDCT) filter banks of 32 filters used in [10-12] by a Uniform/Non-Uniform Filter Bank which is designed using optimization [13]. In our technique evaluation we have also used another research work which a third compression technique based on Discrete Wavelet Transform (DWT) and integrating a Voice Activity Detection System. This latter technique was proposed in [14]. In fact the DWT is a powerful tool used in many domains of signal and image processing such as in [15-16].

The rest of this paper is organized as follow, in section 2, we will detail the proposed compression technique. In section 3, we will present results and discussion and we will conclude in section 4.

2. THE PROPOSED TECHNIQUE

Before presenting in details the proposed speech compression technique, we first deal with Background on psychoacoustic Model. Then we will be interested in a general Approach for Filter Bank Design using Optimization.

2.1. Background on psychoacoustic model

The psychoacoustic model is based on many researches made on human perception. These researches have demonstrated that the average human hearing of all frequencies is not the same. Effects due to the human sensory system limitations and different sounds in the environment lead to facts that can be employed in order to remove unnecessary data contained in an audio signal [10-12]. The two principal human auditory system properties that make up the psychoacoustic model are the auditory masking and the hearing absolute threshold. Each of them provides a manner of determining which signal portions are indiscernible and inaudible to the average human, and can therefore be eliminated from a signal [10-12].

2.2. A general approach for Filter Bank design using optimization

Filter banks, as shown in Figure 1, have different applications in speech processing [13]. One of the principal necessities in filter bank design is Perfect Reconstruction (PR) which intuitively means that the filter bank cannot introduce some degradation on the signal. In general, filter banks can be classified into two main groups that are uniform filter bank in which all sampling rates, $\{n_1, n_2, \ldots, n_K\}$ are all similar while the non uniform filter bank in which at least one sampling rate does not equal to the others [13]. A general approach for Filter Bank Design Using Optimization, was detailed in [13] and references therein.

2.2.1 The perfect reconstruction condition [13]

Consider a filter bank (Fig.1.) having $\{n_1, n_2, \ldots, n_K\}$ as the integer sampling rates. The filter bank output in the Z-domain, can be expressed as follow:

$$\hat{X}(z) = \sum_{k=1}^{K} \frac{1}{n_k} F_k(z) \sum_{l=0}^{n_k-1} X\left(ze^{-j\frac{2\pi l}{n_k}}\right) H_k\left(ze^{-j\frac{2\pi l}{n_k}}\right) = \sum_{k=1}^{K} \frac{1}{n_k} F_k(z) X(z) H_k(z) + \sum_{k=1}^{K} \frac{1}{n_k} F_k(z) \sum_{l=1}^{n_k-1} X\left(ze^{-j\frac{2\pi l}{n_k}}\right) H_k\left(ze^{-j\frac{2\pi l}{n_k}}\right) = X(z)T_0(z) + \sum_{l=1}^{n_k-1} X\left(ze^{-j\frac{2\pi l}{n_k}}\right) T_l(z) \quad (1)$$

With:

$$T_0(z) = \sum_{k=1}^{K} \frac{1}{n_k} F_k(z) H_k(z) \quad (2)$$

$$T_l(z) = \sum_{k=1}^{K} \frac{1}{n_k} F_k(z) H_k\left(ze^{-j\frac{2\pi l}{n_k}}\right) \quad for \quad l = 1, \ldots, K \quad (3)$$

$T_0(z)$ and $T_l(z)$, $l \neq 0$ represent respectively the overall distortion transfer function and aliasing transfer function corresponding to $X(z \cdot \exp(-2j\pi l/n_k))$. When $T_l(z)$ is zero and $T_0(z)$ is a pure delay, then PR is obtained [13].

2.2.2. Filter bank design algorithm

The use of filter banks is for splitting a signal into a number of frequency sub-bands using analysis filters and then each sub-band is separately processed. Therefore the first requirement is to have analysis filters Hk(z) which satisfy some prescribed frequency specifications. The second condition is the requirement that the analysis and synthesis filters satisfy the PR condition [13]. In [13], the aim was to design a filter bank that achieves PR [13] and MRAF satisfy some prescribed requirements. In [13], just the MRAF (the magnitude response of the analysis filters) [13] was considered and the phase response of each individual analysis filter was ignored to increase the degrees of freedom for a given length N. Each designed filter may not have a linear phase but the PR conditions guarantee the linear phase of the whole filter bank. The design procedure needs finding the analysis and synthesis filters coefficients $\{h_k, f_k\}$ so that the PR conditions detailed in [13] (see Lemma 1 in [13]) and the prescribed specifications for MRAF are satisfied. The design approach is based on minimizing the following performance index, combining PR error (e_{PR}) and magnitude response error (e_F), with respect to analysis filters coefficients:

$$J = e_{PR} + e_F = W_{PR}\|Ax - b\|^2 + W_F \sum_{k=1}^{K} e_{F,k} \quad (4)$$

where $\|.\|$ is l_2 -norm and W_F and W_{PR} are optional weights. The expressions A, x, b and $e_{F,k}$ are detailed in [13]. The optimization parameters are the analysis filters coefficients. The synthesis filters coefficients are obtained as the least square solution of the following equation:

$$Ax = b \quad (5)$$

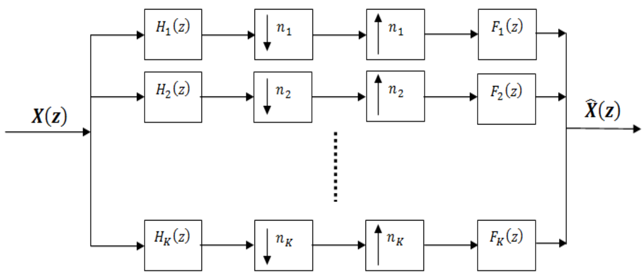


Figure 1. Filter bank [13]

2.3 Details of the proposed compression technique

Alex et al. [10-12] implemented a compression scheme that uses psychoacoustic modeling to determine which portions of the audio signal, they remove without loss of sound quality to the human ear. In their compression system, the original signal is run through cosine modulated perfect reconstruction filter banks having 32 filters. This MDCT filter banks of 32 filters are defined as follow [10-12]:

$$h(k,n) = w(n)\sqrt{2/M} \cdot cos\left((2 \cdot (n-1) + M + 1) \cdot \frac{(2 \cdot (k-1)+1) \cdot \pi}{4M}\right) \quad (6)$$

$$g(k, L-n+1) = h(k,n) \quad (7)$$

where we have:

$1 \le k \le M, 1 \le n \le L, L = 2M, M = 32$ and $w(n) = sin\left(\frac{\pi}{2M} \cdot (n-0.5)\right)$

The signal is divided by the filter banks into distinct frequency components and then it is quantized with a variable number of bits, which is based on the psychoacoustic model results. They have made analysis on this compressed version of the signal and by using different quantization schemes they can get 30 to 75% compression of the original signal. This difference is due to the overhead needed for decoding the quantized signal in each scheme. Here is a simplified block diagram of their scheme (Fig. 2.) [10-12]:

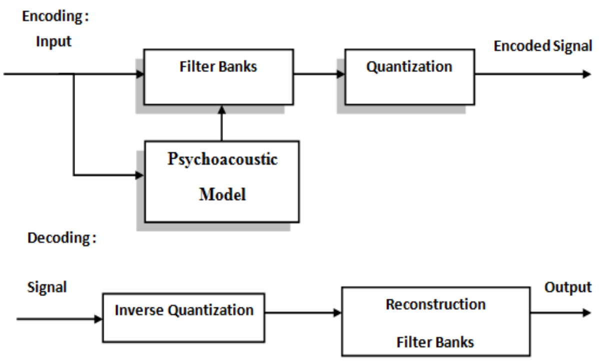


Figure 2. Encoding/Decoding systems [10-12]

In this paper, we have modified the compression system of Alex et al. [10-12] (Fig. 2) by replacing the MDCT (Modified Discrete Cosine Transform) filter banks of 32 filters by a Uniform/Non-Uniform Filter bank which is designed using optimization [13]. The goal is to design M analysis and synthesis FIR filters so that the analysis filters satisfy some frequency specifications and the filter bank (almost) meets the perfect reconstruction (PR) conditions. Both goals are achieved by minimizing the following performance index [13]:

$$J = w_1 \cdot (PR\ error) + w_2 \cdot (Frequency\ Specification\ error\,) \quad (8)$$

where w_1 and w_2 are optional weights.

The algorithm can design both uniform (critically/over sampled) and non-uniform filter banks [13]. Figure 3 illustrates the used uniform filter bank.

Where $H_0(z)$, $H_1(z)$, $G_0(z)$ and $G_1(z)$ are respectively the z-transforms of the impulse responses of the analysis and synthesis filters, $h_0(k)$, $h_1(k)$, $g_0(k)$ and $g_1(k)$. These impulses are obtained from optimization by minimizing the performance index given by (8). Therefore we have replaced the impulses responses h and g associated to MDCT filters banks of 32 filters given by equations (6) and (7), by h_0, h_1, g_0 and g_1. In Table 1 are listed the impulses coefficients obtained from optimization where h_0 and h_1 are designed for analysis and g_0 and g_1 are designed for synthesis.

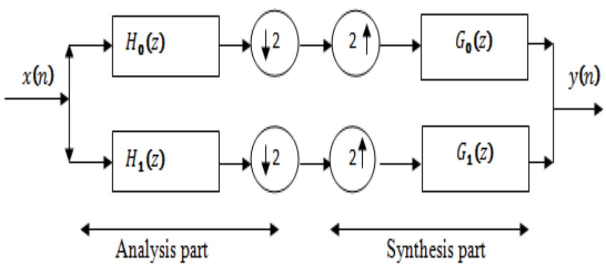


Figure 3. Analysis-Synthesis optimized filterbank

Table 1. Coefficients of Analysis-Synthesis responses impulses of the designed filter bank using optimization

h_0	h_1	g_0	g_1
-0.0052	0.0005	0.0004	0.0023
-0.0023	0.0002	0.0002	-0.0011
0.0479	-0.0196	0.0040	0.0882
-0.0047	-0.0060	0.0004	-0.0498
-0.0863	0.1729	-0.2292	-0.0377
0.1611	-0.0170	0.1119	-0.4325
0.5151	-0.3864	0.8658	1.0458
0.4322	0.5188	1.0158	-0.7691
0.0513	-0.2195	0.3254	-0.0450
-0.1103	-0.0171	-0.1663	0.3356
0.0001	-0.0221	-0.0152	-0.0101
0.0002	0.0455	0.0982	-0.0348
0.0000	-0.0003	-0.0040	0.0002
0.0003	0.0014	-0.0104	0.0006

In this work, the use of a uniform/non-uniform Filter Bank which is designed using optimization [13] is justified by two things. The first is that we want to reduce the computation complexity in our speech compression system by replacing the MDCT (Modified Discrete Cosine Transform) filter banks of 32 filters [10-12] by this uniform/non-uniform Filter Bank. The second thing is that this Filter Bank is with perfect reconstruction.

3. RESULTS AND DISCUSSION

Before presenting the results obtained from Matlab Simulation, we will be interested in the subsection 3.1 in performance evaluation.

3.1. Performance Evaluation

In this paper, we present the objective criteria used for evaluation and comparison between the proposed technique and that of Alex et al. [10-12]. These criteria are bits before and after compression, SNR (Signal to Noise Ratio), PSNR (Peak Signal to Noise Ratio), NRMSE (Normalized Root Mean Square Error) which is detailed in [14, 17] and PESQ (Perceptual Evaluation Speech Quality) which are detailed in [18-19].

3.1.1 File format and comparison

To determine compression ratios for our compression schemes we first have to determine the number of bytes that each file takes. We have used the same computation rules of files size (Original files size, 16-bit Compression, 8-bit Compression, Full Range Compression, Narrow Range Compression) as used in [10-12].

3.2. A comparative study

In Table 2. are listed the results concerning Bits before and after compression using the proposed technique and that of Alex et al. [10-12]. These two techniques are applied to two different speech signals which are ‘‘1.wav’’ and ‘‘3.wav’’. Those obtained results show clearly that the proposed technique outperforms the technique of Alex et al. [10-12] and this in term of size of output files. To solve the problem of speech degradation when using narrow range (Figure 4) in the proposed technique and the technique of Alex et al. [10-12], we have multiplied the psychoacoustic model threshold by an adjustment factor α. We have selected 30 as the value of α and this based on simulation results.

Figures 5, 6 and 7 show three different zoomed regions of the original and the compressed speech signals. According to these figures, we remark that the silent regions are zeroed and the others are thresholded and this due to the quantization procedure.

In Table 2 are listed the results concerning Bits before and after compression using the proposed technique and that of Alex et al. [10-12]. The two techniques are applied to two different speech signals.

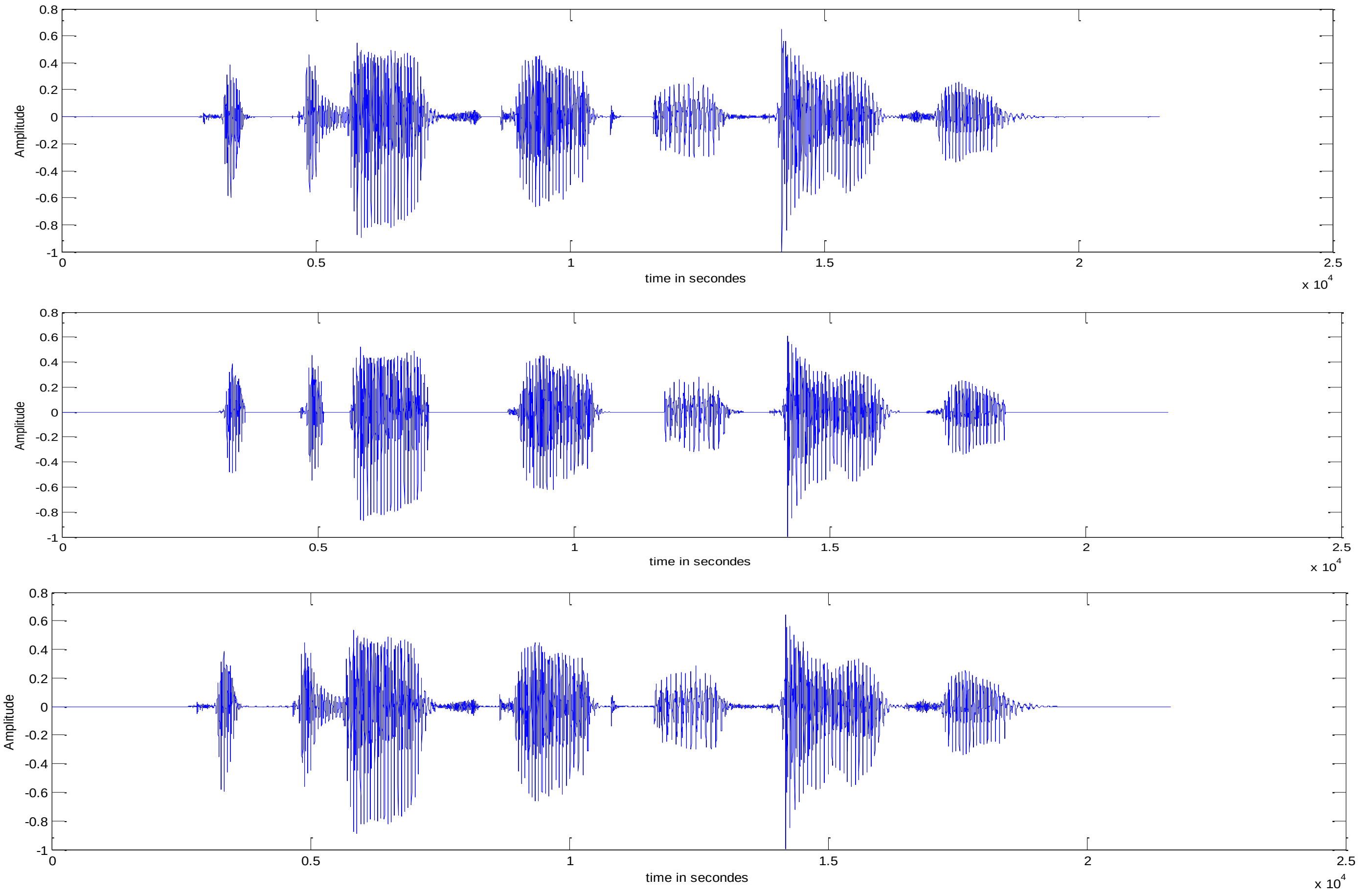
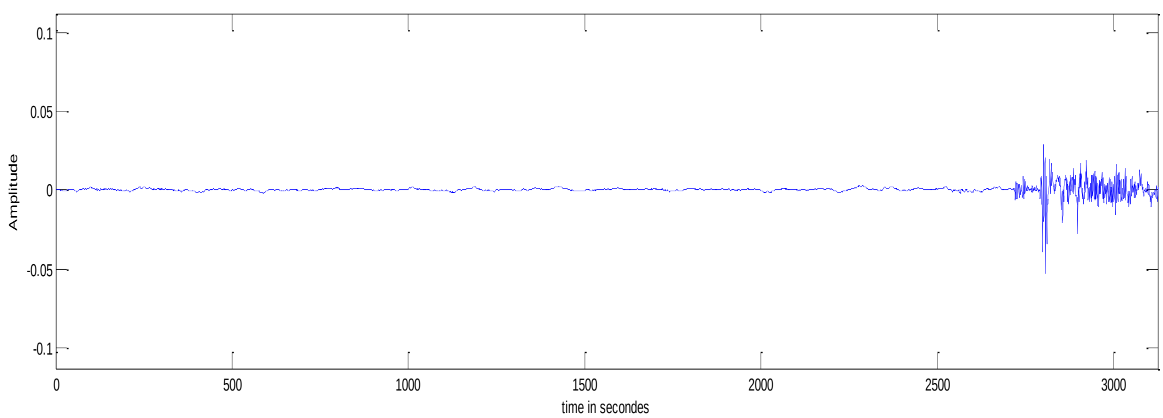


Figure 4. Original Speech signal at the top, degraded output speech signal (in the middle) of the proposed compression system without multiplying the threshold by α, output speech signal (at the bottom) of the proposed compression system with multiplying the threshold by α



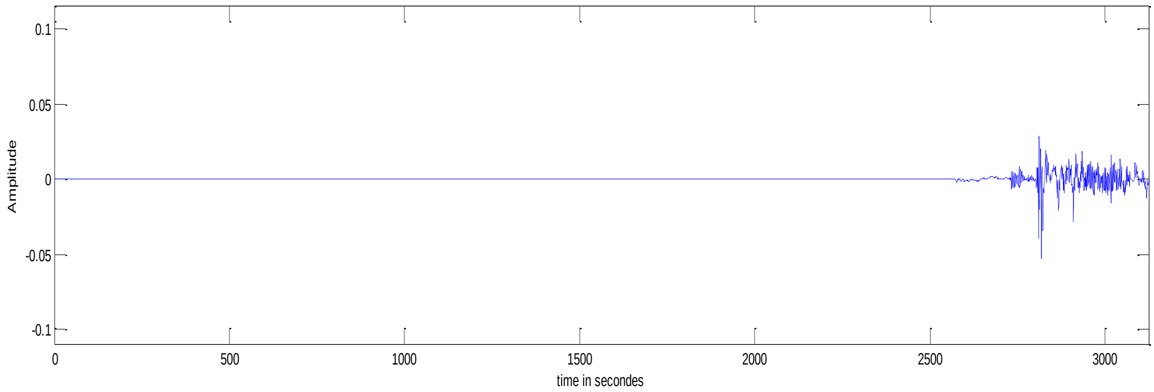


Figure 5. The beginning (at the top and zoomed) of the original speech signal represented in Fig.4, the beginning (at the bottom and zoomed) of the corresponding compressed speech signal represented in Figure 4 (at the bottom)

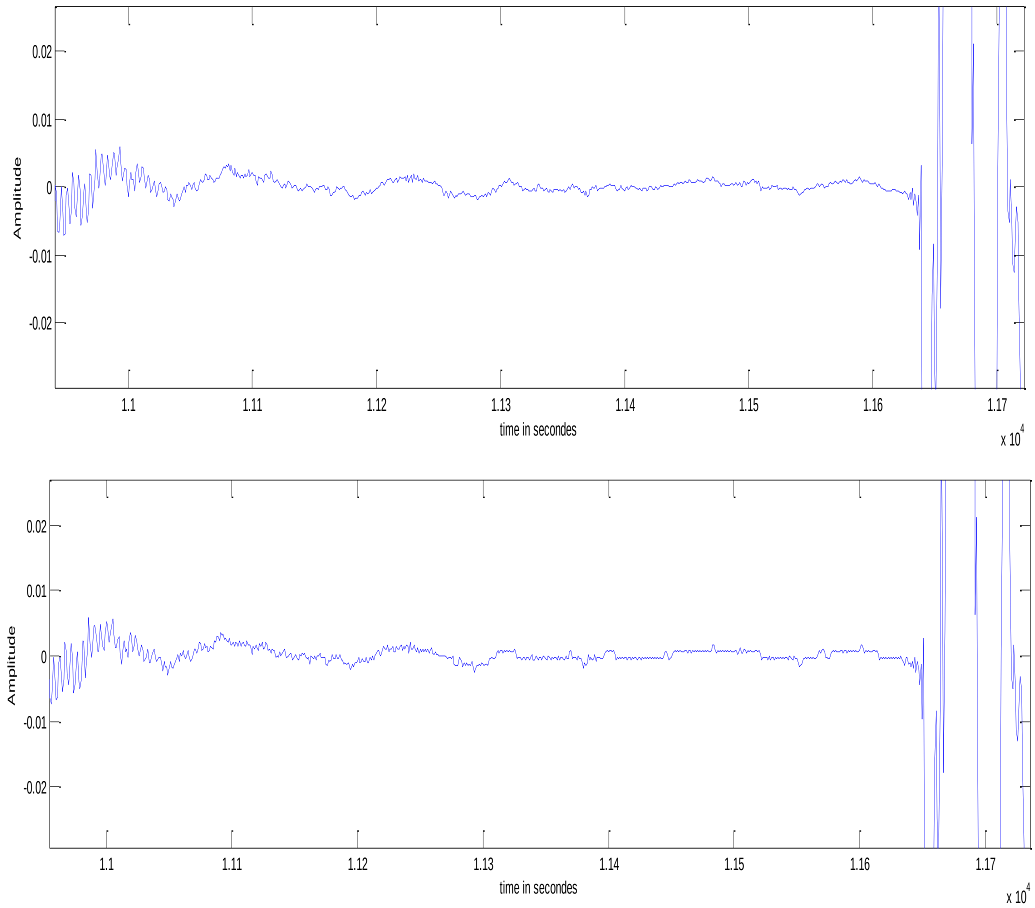
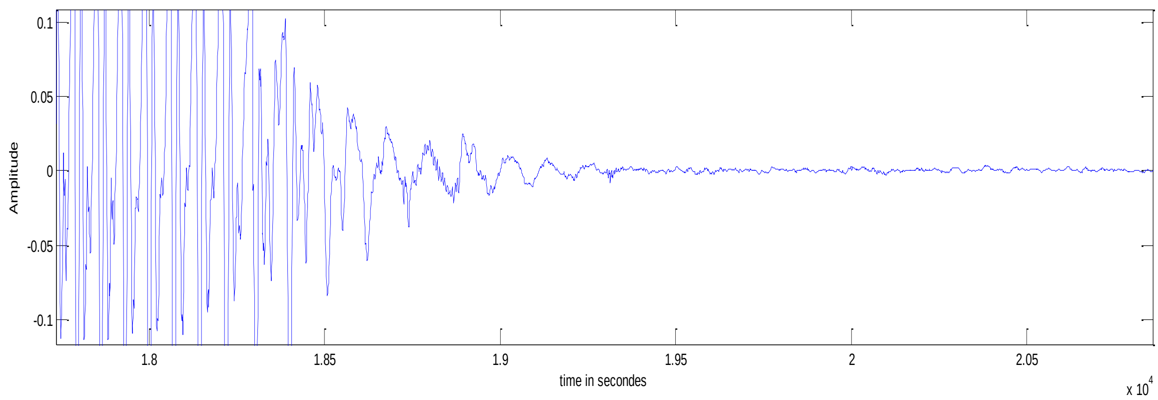


Figure 6. The middle (at the top and zoomed) of the original speech signal represented in Figure 4 (at the top), the middle (at the bottom and zoomed) of the corresponding compressed speech signal represented in Figure 4 (at the bottom)



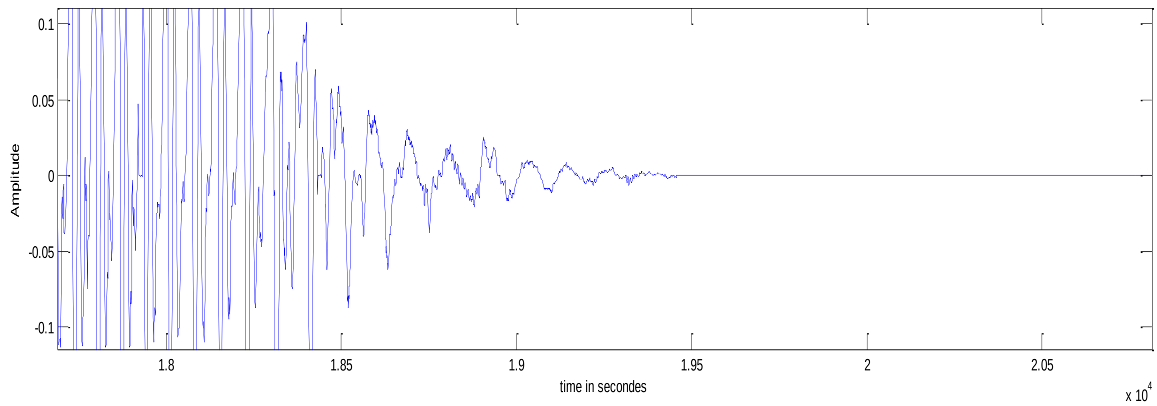


Figure 7. The end (at the top and zoomed) of the original speech signal represented in Fig.4 (at the top), The end (at the bottom zoomed) of the corresponding compressed speech signal represented in Figure 4 (at the bottom)

Table 2. Bits before and after compression

Compression Technique	Signal	Full Range Quantization	Narrow Range Quantization	8 bits Quantiz-ation	16 bits Quantiza-tion	Original 16*number of samples	Original .wav
Technique of Alex et al. [10, 11, 12]	Sig1 : ‘1.wav’	325632	365568	202752	399360	373152	372736
	Sig3 : ‘3.wav’	339200	380800	211200	416000	345680	345700
The proposed technique	Sig1 : ‘1.waw’	265320	279556	184680	369000	373152	372736
	Sig3 : ‘3.wav’	247632	249816	172368	344400	345680	345700

Table 3. Compression ratio and bits before and after compression for Narrow Range quantization in the two cases: multiplying and without multiplying by factor α

The compression technique	Speech Signal	Bits before Compression	Narrow Range Quantization	
			Multiplying by a factor α=30	Without multiplying by any factor
Technique of Alex et al. [10,11, 12]	Sig1: ‘‘1.wav’’	372736	242688 (CR = 1.5359)	365568 (CR = 1.0196)
	Sig2: ‘‘3.wav’’	385024	222464 (CR = 1.7307)	335104 (CR = 1.1490)
	Sig3: ‘‘5.wav’’	345700	252800 (CR = 1.5233)	380800 (CR = 0.9078)
The proposed technique	Sig1: ‘‘1.wav’’	372736	152460 (CR = 2.4448)	222464 (CR = 1.6755)
	Sig2: ‘‘3.wav’’	385024	142296 (CR=2.7058)	249816 (1.5412)
	Sig3: ‘‘5.wav’’	385088	195332 (CR = 1.9715)	279556 (CR = 1.2366)
		Compression Ratio (CR)		
Technique of Aloui et al. [17]	Sig1: ‘‘1.wav’’	CR = 2.0418		
	Sig2: ‘‘3.wav’’	CR = 2.2853		
	Sig3: ‘‘5.wav’’	CR = 2.0256		

Those obtained results show clearly that the proposed technique outperforms the technique of Alex et al. [10-12] and this in term of size of output files: the sizes of output files from the proposed compression system are smaller than the those of the output files from the system of Alex et al. [10-12].

In Tables 3 and 4 are listed the results obtained from the application of the proposed technique, the technique of Alex et al. [10-12] and that of Aloui et al. [17] to three different speech signals (‘‘1.wav’’, ‘3.wav’’ and ‘‘5.wav’’) and this in case of narrow range quantization (with/without multiplying the psychoacoustic model threshold by a tuning factor $\alpha = 30$). Those results are in term of bits before and after compression, Compression Ratio (CR) and PESQ. The speech compression system of Aloui et al. [14] is based on Discrete

Wavelet Transform (DWT) [20-21] and integrating a Voice Activity Detection (VAD) Module [14]. According to Aloui et al. [14], this VAD module avoids the computation of discrete wavelet coefficients during the inactive voice signal.

Tables 3 and 4 and Figure 4 (c) show clearly that by multiplying the psychoacoustic model threshold by a tuning factor $\alpha = 30$, we obtain output speech signals with acceptable perceptual qualities for the proposed compression system and with good quality for the compression system of Alex et al. [10-12]. Compared to the speech compression technique proposed by Aloui et al. [14], the proposed technique permits to obtain better values of Compression Ratio (CR) and also better values of PESQ.

The output speech signals of the proposed speech compression system are with a little constant delay compared to the original speech signals and this in case of narrow range quantization. To solve this problem we have suppressed this delay and we have obtained the following results listed in Table 5.

Table 4. PESQ values of the reconstructed speech signal in case of Narrow range

The compression Technique	Speech signal	Narrow range quantization: Multiplying by a factor α=30	Narrow range quantization: Without multiplying by any factor
Technique of Alex et al. [10, 11, 12]	Signal1 : ‘‘original1.wav’’	PESQ = 4.28	PESQ = 2.07
	Signal2 : ‘‘nhwy_jrmlr.wav’’	PESQ = 4.20	PESQ = 2.22
	Signal3: ‘‘1.wav’’	PESQ = 3.52	PESQ = 1.58
	Signal4: ‘‘3.wav’’	PESQ = 3.85	PESQ = 2.07
	Signal5: ‘‘5.wav	PESQ = 3.70	PESQ = 1.53
The proposed Technique	Signal1 : ‘‘original1.wav’’	PESQ = 3.28	PESQ = 1.59
	Signal2 : ‘‘nhwy_jrmlr.wav’’	PESQ = 3.43	PESQ = 1.66
	Signal3: ‘‘1.wav’’	PESQ = 2.43	PESQ = 0.55
	Signal4: ‘‘3.wav’’	PESQ = 2.81	PESQ = 0.84
	Signal5: ‘‘5.wav’’	PESQ = 2.48	PESQ = 0.59
Technique of Aloui et al. [17]	Signal1 : ‘‘original1.wav’’	PESQ = 2.7951	
	Signal2 : ‘‘nhwy_jrmlr.wav’’	PESQ = 2.5410	
	Signal3: ‘‘1.wav’’	PESQ = 2.3249	
	Signal4: ‘‘3.wav’’	PESQ = 2.6399	
	Signal5: ‘‘5.wav’’	PESQ = 2.3190	

Table 5. SNR, PSNR and NRMSE of Alex et al method and the proposed speech compression for “Narrow range quantization”

Compression Technique	Signal	SNR (Signal to Noise Ratio)	PSNR (Peak Signal to Noise Ratio)	NRMSE (Normalized Root Mean Square Error)	Compression Ratio (CR)
Technique of Alex et al. [10, 11, 12]	Signal1 :‘‘original1.wav’’	-0.89	17.25	1.10	1.5563
	Signal2 :‘nhwy_jrmlr.wav’	-1.68	15.61	1.21	1.5277
	Signal3: ‘‘1.wav	-3.68	11.90	1.52	1.5359
	Signal4: ‘‘3.wav	-3.76	14.92	1.54	1.3675
	Signal5: ‘‘5.wav	-4.36	12.54	1.65	1.5233
The proposed Technique	Signal1 : ‘‘original1.wav’’	16.83	34.93	0.14	1.9985
	Signal2 :‘nhwy_jrmlr.wav’	15.73	33.00	0.16	1.9719
	Signal3: ‘‘1.wav	15.73	31.29	0.16	2.4448
	Signal4: ‘‘3.wav	17.43	36.10	0.13	2.1710
	Signal5: ‘‘5.wav	14.17	31.08	0.19	1.9715
Technique of Aloui et al. [17]	Signal1 : ‘‘original1.wav’’	11.2420	29.3349	0.2741	1.8653
	Signal2 :‘nhwy_jrmlr.wav’	10.5927	27.8888	0.2954	1.5021
	Signal3: ‘‘1.wav	15.3391	30.9440	0.1710	2.0418
	Signal4: ‘‘3.wav	16.9001	35.5655	0.1429	2.2853
	Signal5: ‘‘5.wav	13.2419	30.1519	0.2177	1.6785

The obtained results from computation of SNR, PSNR, NRMSE and Compression Ratio show that the proposed compression technique outperforms the two others techniques of Alex et al. [10-12] and Aloui et al. [17]. In fact, in term of perceptual quality and according to SNR, PSNR and NRMSE, the reconstructed speech signals obtained after compression/de-compression using the proposed technique, have a better perceptual qualities compared to the two techniques [10-12, 17]. Moreover, in term of Compression Ratio, the proposed technique permits to obtain the best values of Compression Ratio.

4. CONCLUSION

In this paper, we have proposed a new speech compression technique integrating a psychoacoustic model and uniform Filter Bank which is designed using optimization. The role of the psychoacoustic model is to determine which portions of the speech signal to remove without loss of sound quality to the human ear. The proposed technique was evaluated and compared to a second technique based on psychoacoustic model and MDCT (Modified Discrete Cosine Transform) filter banks having 32 filters. This evaluation is performed by computing the bits before and after compressing, SNR, PSNR, NRMSE and PESQ. The obtained results from bits before and after compressing, SNR, PSNR, NRMSE and PESQ, show that the proposed technique outperforms the other compression technique (based on the psychoacoustic model and the MDCT filter banks of 32 filters). In fact, in term of perceptual quality and according to SNR, PSNR and NRMSE and PESQ, the reconstructed speech signals obtained after compression/de-compression using the proposed technique, have a better perceptual qualities compared to those obtained from the application of the second compression technique. Moreover, in term of sizes of the reconstructed speech signals and Compression Ratios, the proposed technique permits to obtain better Compression Ratios compared to those obtained by the second technique. We have also evaluating the proposed technique by comparing it to a third speech compression technique based on Discrete Wavelet Transform (DWT) and VAD (Voice Activity Detection). This comparision is also in term of SNR, PSNR, NRMSE and CR and the obtained results show that the proposed technique outperforms the third technique based on DWT and VAD.

ACKNOWLEDGMENT

We would like to thank all the people who contributed in some way to this work which was supported by CRTEn (Center of Researches and Technologies of Energy of Borj Cedria, Tunis and Ministry of Higher Education and Scientific Research.

REFERENCES

[1] Xie N, Dong G, Zhang T. (2011). Using lossless data compression in data storage systems: not for saving space. IEEE Transactions on Computers 60(3): 335–345.

[2] Gibson JD. (2005). Speech coding methods, standards, and applications. IEEE Circuits and Systems Magazine 5(4): 30–49.

[3] Riadh A, Salim S, Said G, Ali C, Taleb Ahmed A. (2014). Novel detection algorithm of speech activity and the impact of speech codecs on remote speaker recognition system. WSEAS Transactions on Signal Processing. E-ISSN: 2224-3488.

[4] Junejo N, Ahmed N, Unar MA, Rajput AQK. (2005). Speech and image compression using discrete wavelet transform. In IEEE Symposium on Advances in Wired and Wireless Communication, pp. 45–48.

[5] Arif M, Anand RS. (2012). Turning point algorithm for speech signal compression. International Journal of Speech Technology. 10.1007/s10772-012-9151-7, 2012.

[6] Gersho A. (1992). Speech coding. In A. N. Ince (Ed.), Digital Speech Processing, pp. 73–100. Boston: Kluwer Academic.

[7] Gersho A. (1994.). Advance in speech and audio compression. Proceedings of the IEEE 82(6): 900–918.

[8] Agbinya JI. (1996). Discrte wavelet transform techniques in speech processing. In IEEE Tencon Digital Signal Processing Applications Proceedings, New York: IEEE, pp. 514–519.

[9] Gershikov E, Porat M. (2007). On color transforms and bit allocation for optimal subband image compression. Signal Processing. Image Communication 22: 1–18.

[10] W.A.V.S. Compression. http://www.aamusings.com/project documentation/wavs/index.html

[11] http://www.360doc.com/content/05/0810/17/641_5382.shtml

[12] Giron-Sierra JM. (2017). ‘Digital signal processing with matlab examples. http://www.springer.com/us/book/9789811025334

[13] Moazzen I, Agathoklis P. (2014). A general approach for filter bank design using optimization. Technical Report. http://www.ece.uvic.ca/~imanmoaz/publications.htm

[14] Aloui N, Boussalmi S, Cherif A. (2015). Optimized speech compression algorithm based on wavelets techniques and its real time implementation on DSP. I.J. Information Technology and Computer Science 3: 33-41.

[15] Suryavanshi HE, et al. (2013). Digital image watermarking in wavelet domain. International Journal of Electrical and Computer Engineering (IJECE) 3(1): 1–6.

[16] Narmadha D, et al. (2014). An optimal HSI image compression using DWT and CP. International Journal of Electrical and Computer Engineering (IJECE) 4(3): 411– 421.

[17] Boussalmi S, Aloui N, Cherif A. (2016). Adaptive speech compression based on discrete wave atoms transform. International Journal of Electrical and Computer Engineering (IJECE) 6(5): 2150-2157.

[18] Talbi M, Barnoussa C, Cherif A. (2013). Speech compression based on psychoacoustic model and a general approach for filter bank design using optimization. ACIT’2013.

[19] Talbi M. (2017). Speech enhancement based on stationary bionic wavelet transform and maximum a posterior estimator of magnitude-squared spectrum. International Journal of Speech Technology 20(1): 75–88.

[20] Chattopadhyaya A, Chattopadhyay S, Bera JN, Sengupta S. (2016). Wavelet decomposition based skewness and kurtosis analysis for assessment of stator current harmonics in a PWM – fed induction motor drive during single phasing condition. Advances in Modelling and Analysis B 59(1): 1-14.

[21] Sun Q, Zhao X. (2016). A new subspace based speech enhancement algorithm with low complexity. Advances in Modelling and Analysis B 59(1): 164-176.