Design of orthogonal filter banks using a multi-objective genetic algorithm for a speech coding scheme

Abdelkader Boukhobza, Nasreddine Taleb, Abdelmalik Taleb-Ahmed, Abdennacer Bounoua

Alexandria Engineering Journal (2022) 61, 7649–7657

HOSTED BY

ELSEVIER

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Design of orthogonal filter banks using a multi-objective genetic algorithm for a speech coding scheme

Abdelkader BOUKHOBZA [a,*], Nasreddine TALEB [b], Abdelmalik TALEB-AHMED [c], Abdennacer BOUNOUA [b]

[a] *Department of Electronics, Hassiba Benbouali University, Chlef, 02000, Algeria*
[b] *RCAM Laboratory, Djillali Liabes University, Sidi Bel Abbes, 22000, Algeria*
[c] *IEMN UMR CNRS 8520, UPHF, Valenciennes, France*

Received 23 August 2021; revised 12 December 2021; accepted 2 January 2022
Available online 14 January 2022

KEYWORDS

Discrete wavelet transform;
Orthogonal filter banks;
Speech coding;
Optimized wavelet;
NSGAIII algorithm

Abstract In this work, we propose an optimization scheme based on a multi-objective Genetic Algorithm (GA) for the design of orthogonal filter banks for speech compression. A parameterization is adopted to assure that the resulting filter banks satisfy perfect reconstruction and have at least two vanishing moments. We search for a parameter set that optimizes the coding gain and the frequency selectivity. As the objectives are conflicting, we investigate the solution that realizes the best compromise between the objectives criteria using the Non-dominated Sorting Genetic Algorithm (NSGAIII). Experimental results have shown that the optimized filter banks provide a significant gain in coding performances when comparing with the Daubechies orthogonal filter banks for test speech signals.

1. Introduction

Speech signals play a key role in the human communication system. Similar to other digital signals, speech signals need to be encoded and compressed. The fundamental purpose of speech compression is to characterize it with the least number of bits while maintaining its perceptual quality [1]. The speech compression is essential either for reducing memory storage requirements or accordingly for reducing transmission bandwidth requirements. Speech compression is required for long-distance communication, multimedia applications, video conferencing systems, digital cellular communications, and many others. To keep step with the rapid advances in these areas, different techniques [2] were developed to meet the growing demand for better speech compression algorithms.

During the last two decades, Discrete Wavelet Transform (DWT) has emerged as a powerful mathematical tool in many areas of science and technology, especially in the field of speech and image compression, which is implemented usually using multi-resolution filter banks for analyzing and extracting infor-

* Corresponding author.
E-mail address: aek_boukhobza@yahoo.ca (A. BOUKHOBZA).

Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2022.01.017
1110-0168

mation from non-stationary signals. Filter banks separate input signals into subband components representing the frequency-localized signal energy within each band. Subband processing units are then adapted to the characteristics of these subband components. The application of wavelet transform in speech coding is studied in [3] for speech analysis, coding, and synthesis. In this domain, the choice of filter banks used to implement DWT is a key problem which affects coding performance as well as the design of the coding system [4,5]. This paper deals with orthogonal filter banks, also known as conjugate quadrature filter (CQF) banks. This class of filter banks has some interesting properties, such as energy preservation, that is critical in the design of quantization and bit allocations algorithms [6].

The problem of orthogonal filter bank design has become the subject of many research activities [7–15]. Some of the proposed design methods are oriented to speech coding applications. For example, a new class of optimized wavelet filters for speech compression was introduced in [11]. In this case, the wavelet filter coefficients are obtained by simple linear optimization using various techniques, such as Kaiser and Blackman windows. The authors in [12] proposed a method for designing a two-channel quadrature mirror filter (QMF) bank based on Kaiser Windowing. In this technique, the cut-off frequency of the prototype filter and the shape of the Kaiser window are optimized using a genetic algorithm (GA). The resulting QMF filter bank is used as the mother wavelet in the Discrete Wavelet Transform (DWT) tree for speech compression. The study proposed in [13] presents a new speech compression technique combining a psychoacoustic model and uniform filter bank which is designed via optimization. The purpose of the psychoacoustic model is to decide which portions of the speech signal to discard without losing the quality of the human ear. In [14], an optimized filter bank is developed to improve speech perception by incorporating masking techniques in the algorithm to reduce the noise effect.

The first challenge of any design method is to construct a filter bank that satisfies perfect reconstruction (PR) in order to recover exactly the original signal from subbands. Several parameterization approaches have been proposed to express PR constraint in terms of free design parameters [7,9]. These parameters can be used by filter designers to achieve the aspirated characteristics, such as: symmetry, frequency selectivity, and so on. In addition, the number of vanishing moments (VM), which is related to the regularity of wavelets and scaling functions, is a crucial property of most signal processing applications, and should also be considered.

In this study, we will focus on the design of orthogonal wavelet filter banks in the speech compression domain. The optimization design approach presented in this work is a continuation of the improvement of earlier works in this subject. First, we adopt the factorization of the orthogonal filter bank presented in [7] to guarantee that the designed filter bank has a perfect reconstruction with at least two vanishing moments. Particularly in this work, and in order to achieve optimal filters for speech coding, we have considered criteria of practical significance in this field, namely: energy compaction capability or coding gain and frequency selectivity. Accordingly, the design problem consists in searching for the filter bank parameters that optimize simultaneously these two criteria. Really, in multi-objective problem, it is impossible to satisfy all criteria maximally.To solve this problem, a multi-objective genetic method called NSGAIII is used to find a set of compromised solutions [16] from which the final filter bank can be selected.

The structure of this article is as follows. Section 2 is a brief review of the wavelet transform. Section 3 describes the design criteria of the filter bank in detail. In Section 4, we present a short review of genetic algorithms and multi-objective optimization principles. Next, we define the optimization problem formula and the multi-objective genetic algorithm used to design the filter bank. In Section 5, we evaluate the compression performance of the optimized filter bank on a set of test speech signals. Finally, in Section 6, we summarize our work and present some perspective suggestions.

2. Review of wavelet transform

Wavelet transform is an analysis tool that can be used to mathematically describe the information content of a signal, from a coarser to a higher resolution [17]. The discrete wavelet transform is frequently constructed by using a perfect reconstructed finite impulse response filter bank. In this case, the signal is divided into different decomposition (resolution) levels. These decomposition levels include many subbands, which contain coefficients representing the time–frequency characteristics of the original signal.

As an example, a three level wavelet filter bank is shown in Fig. 1. Filters $H_0(z)$ and $H_1(z)$ are respectively the low-pass and high-pass filters, $G_0(z)$ and $G_1(z)$ are their corresponding synthesis filters. The outputs of the low-pass and high-pass branches are called approximations and details, respectively.

Recently, many wavelet-based methods have been devised in order to compress speech signals [4,18,19]. The most important characteristic of wavelet transforms, relative to data compression, is that it tends to concentrate the energy of the input signal into a relatively small number of wavelet coefficients. Compared with the complete signal, the coding of these coefficients requires fewer binary resources while maintaining satisfactory quality of the reconstructed signal. In wavelet speech coding scheme (see Fig. 2), the three most commonly used steps are:

DWT decomposition: In this step, the filter bank is designed, and then the DWT decomposition is performed on the speech signal. In addition to the filter bank coefficients, the speech coding performances are closely related to filter banks and the number of decomposition levels [3]. In our experiments, four wavelet decomposition levels are employed. It was recommended in [3] that the adequate number of decomposition levels for speech compression should be less or equal to five, without any further advantages in scales of more than five.

Thresholding: Is a simple non-linear technique in which each coefficient is compared against a threshold; if the coefficient is smaller than threshold, then set to zero; otherwise it is maintained or modified. In general, thresholding methods can be classified into two categories, namely, global and level thresholding methods. In global thresholding, the decomposed signal is classified based on single threshold value estimated from its wavelets coefficients. In level thresholding, this value is evaluated for each level of the wavelet decomposition tree. In this study, global thresholding is applied. The threshold value is given by:

$$\lambda = \sqrt{2\log(N)} \tag{1}$$

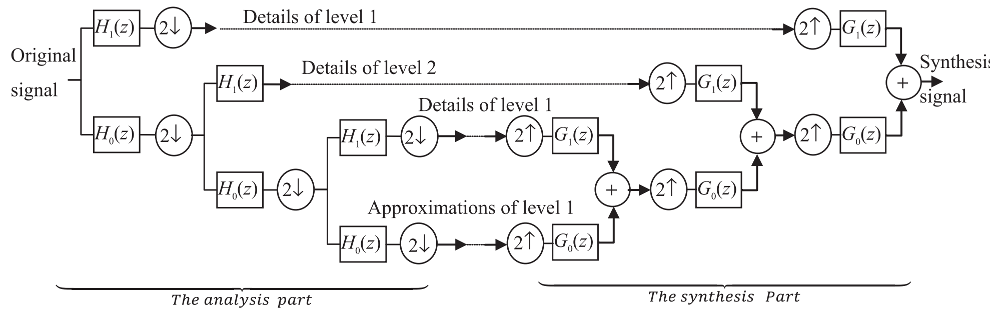


Fig. 1 a three level wavelet filter bank.

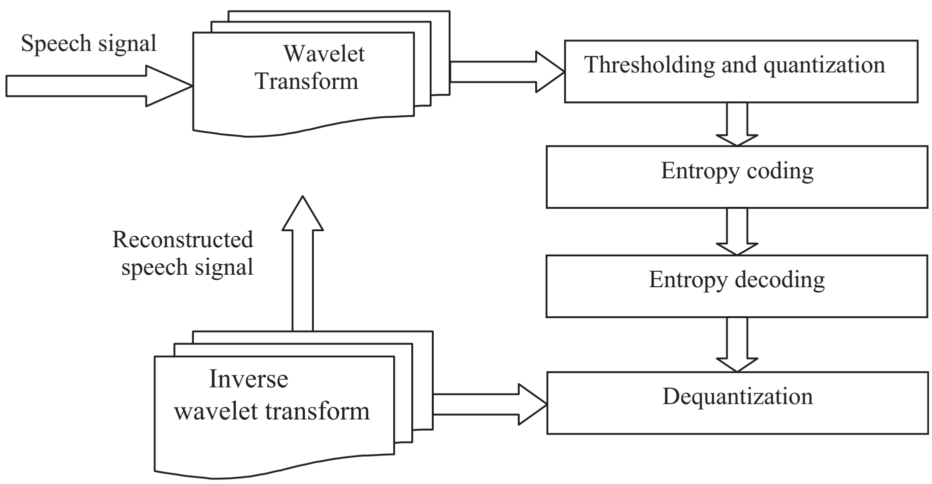


Fig. 2 A block diagram of wavelet based speech compression.

Where N is the number of samples in the signal.

Quantization: The purpose of this technique is to mask irrelevant information of the signal. This minimizes the number of bits needed to store the transformed coefficients by reducing the precision of these values. In our coding system, wavelet coefficients are quantized using the uniform step size computed by:

$$\Delta = \frac{x_{max} - x_{min}}{L} \tag{2}$$

Where, x_{max} and x_{min} are respectively the maximum and the minimum values in the signal and L is the number of quantization levels.

Entropy coding: This technique reduces the statistical redundancy in the quantized coefficients by using variable length coding techniques such as arithmetic coding, Huffman coding, etc.

3. Design criteria

The principle of speech coding is to use the least number of bits to represent the speech signal while maintaining an acceptable perceptual quality [2]. Our main goal is to design optimal orthogonal wavelet filter banks for speech compression. For this reason, we look for filter banks with some interesting characteristics, namely perfect reconstruction, high coding gain, good frequency selectivity and certain regularity (at least two vanishing moments). In the following subsections, we will describe these properties and their corresponding appropriate measurement functions, which are defined to quantify the effectiveness of filter banks in speech compression schemes.

3.1. Perfect reconstruction condition

Fig. 3 exposes an analysis-synthesis of a two-channel filter bank. In this system, the relationship between the input signal and the output one in the Z domain is given by:

$$\widehat{X}(z) = \frac{1}{2}\left[G_0(z)H_0(z) + G_1(z)H_1(z)\right]X(z) \tag{3}$$

$$+\frac{1}{2}\left[G_0(z)H_0(-z) + G_1(z)H_1(-z)\right]X(-z)$$

Setting: $G_0(z) = H_1(-z)$, $G_1(z) = -H_0(-z)$ the synthesis filter eliminates the aliasing term (i.e. term of $X(-z)$). We get:

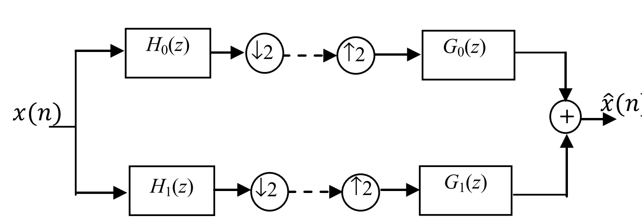


Fig. 3 Two-channel PR filter bank.

$$\begin{aligned}\widehat{X}(z) &= \tfrac{1}{2}[H_0(z)H_1(-z) - H_0(-z)H_1(z)]X(z) \\ &= T(z)X(z)\end{aligned} \tag{4}$$

Then, the PR condition can be achieved by imposing the transfer function $T(z)$ to be a pure delay of the form $T(z) = z^{-d}$, where d is the delay of the system.

$$T(z) = z^{-d} \tag{5}$$

In the two-channel FIR orthogonal filter bank, the relation between the analysis filters is defined by [20]:

$$H_1(z) = -z^{-(L-1)}H_0\left(-z^{-1}\right) \tag{6}$$

Accordingly, in time domain, the analysis and synthesis filter bank that constitutes the filter bank should satisfy the following equations:

$$\begin{aligned} h_1(n) &= (-1)^n h_0(L-n-1) \\ g_0(n) &= h_0(L-n-1) \qquad n = 0, \cdots, L-1 \\ g_1(n) &= h_1(L-n-1) \end{aligned} \tag{7}$$

L is the length of filters that must be even.

In this case, the PR condition of (eq.4) is fulfilled if:

$$H_0(z)H_0\left(z^{-1}\right) + H_0(-z)H_0\left(-z^{-1}\right) = 2 \tag{8}$$

This is equivalent to the following $L/2$ inequality constraints [20]:

$$\sum_{n=0}^{L-1-2k} h_0(n)h_0(n+2k) = \delta(k) \tag{9}$$

$$for \qquad k = 0, 1, \cdots, (L-2)/2$$

$\delta(k)$ is the Kronecker delta function.

Vaidyanathan [20] solved the design problem of orthogonal filter banks by using lattice factorization in polyphase domain. The lattice parameterizations proposed in [7,9] provide the possibility to design a PR orthogonal filter bank through unconstrained parameters with at least two vanishing moments.

3.2. Coding gain

Coding gain (CG) is a widely accepted coding performance metric used to estimate the energy concentration capability of a filter bank. By modelling natural speech as a one-dimensional Markov source with a correlation factor ρ and assuming that the quantization errors are uncorrelated, Katto and Yasuda [21] derived the filter-dependent coding gain:

$$CG(\rho) = 10\log 10\left(\prod_{k=0}^{M-1}(A_k B_k)^{-\frac{1}{\alpha_k}}\right) \tag{10}$$

where: $A_k = \sum_i \sum_j h_k^{'}(i)h_k^{'}(j)\rho^{|j-i|}$, $B_k = \sum_i g_k^{'}(i)^2$

For orthogonal filters, we have: $B_k = \sum_i g_k^{'}(i)^2 = 1$. Consequently, we obtain:

$$CG(\rho) = 10\log 10\left(\prod_{k=0}^{M-1}(A_k)^{-\frac{1}{\alpha_k}}\right) \tag{11}$$

where $h_k^{'}$ and $g_k^{'}$ are respectively the k^{th} analysis and synthesis filter of the M channel nonuniform filter bank equivalent to the N_d $(M = N_d + 1)$ level tree structured filter bank (e.g., Fig. 4), α_k is the corresponding subsampling ratio, and ρ is the correlation factor.

In addition, we have:

$$H_i^{'}(z) = \begin{cases} H_1(z) & if \quad i = 0 \\ H_1\left(z^{\alpha_i/2}\right)\prod_{k=0}^{i-1} H_0\left(z^{\alpha_k/2}\right) & if \quad 1 \le i \le M-2 \\ \prod_{k=0}^{i} H_0\left(z^{\alpha_k/2}\right) & if \quad i = M-1 \end{cases} \tag{12}$$

where:

$$\alpha_k = \begin{cases} 2^{i+1} & if \quad 0 \le i \le M-2 \\ 2^i & if \quad i = M-1 \end{cases} \tag{13}$$

In our work, we have used a correlation factor ρ = 0.95 and a six-level binary tree structure subband decomposition, because in experiments, this level number provides the best performance for various speech types, and is often used to evaluate wavelet speech coding algorithms.

3.3. Frequency selectivity

The advantage of frequency selectivity in speech coding is that the cost of coarse quantization in unimportant subbands is lower, because the errors will be restricted to the frequency bands where they appear. A popular criterion for evaluating frequency selectivity in subband coding theory is to make the two analysis low-pass and high-pass filters close to their ideal versions, respectively.

To measure the filter frequency selectivity, we use the filter bank Transition Band Energy (TBE) expressed by:

$$TBE = \int_0^{\pi} |H_0(\omega)H_1(\omega)|^2 \mathrm{d}\omega \tag{14}$$

where $H_i(e^{j\omega})$ is the frequency response of filters $H_i(z)$.

Using parseval's relation, we find:

$$TBE = \pi\sum_{n=0}^{L-1}|h_0(n) * h_1(n)|^2 \tag{15}$$

This function is a measure of the deviation from an ideal low-pass and high-pass filter pair [22]. If the overlap between the filters is zero, which is only possible for ideal filters, then TBE is zero.

4. Optimization problem formulation

Genetic algorithms form the main subset of evolutionary algorithms. In an optimization application, a GA starts with an initial random population of solutions. GA optimization use three operators to generate new solutions from the existing ones: evaluation, crossover and mutation. In the evaluation step, each solution is evaluated based on its objective functions values for fitness which represents a measure of the solution quality. Crossover simply combines a certain number of solutions to form offspring. The mutation is achieved by randomly altering solutions according to a given mutation rate. Each iteration of this process is called a *generation*. The procedures: evaluation, selection, crossover and mutation continue until a maximum number of generations is reached or another stopping criterion is satisfied.

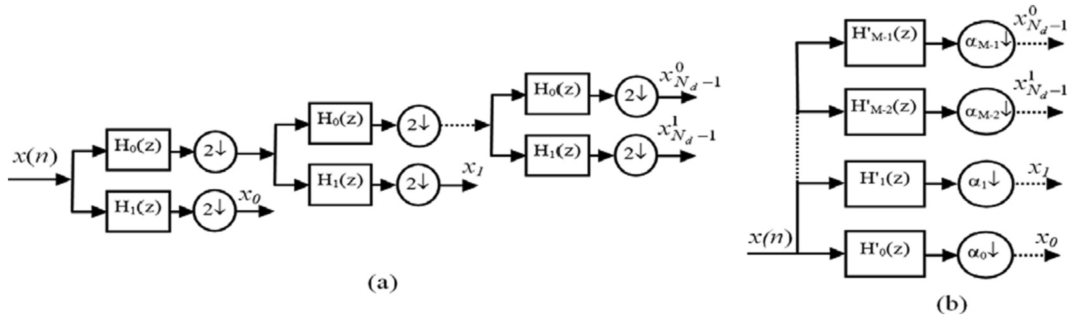


Fig. 4 M-band filter banks ($M = N_d + 1$) **(a)** N_d level tree structured filter bank **(b)** Equivalent M channel nonuniform filter bank.

However, a multi-objective optimization problem considers more than one objective function. These objective functions are usually conflicting with each other. In such cases, instead of one optimal solution, there is a set of non-dominated solutions where none is best for all objectives. These optimal solutions are largely known as the Pareto-optimal solutions. Because GAs use a population of solutions, they are able to capture several members of this set in a single optimization run. A detailed study of the approaches based on this principle can be found in [23–26]. Among the best of these methods, the non-dominated genetic algorithm NSGA III [16] has been employed in this work. The flowchart of NSGAIII is shown in Fig. 5.

Using the method presented in [7] and selecting $L/2-1$ and $L/2-2$ angles (L is the length of filter), we can construct an orthogonal perfect reconstruction filter bank with 1 and 2 vanishing moments, respectively. In order to design an effective speech compression filter bank, we search for the angles $\{\boldsymbol{\theta}_k, k = 1, \cdots, L/2 - N_{vm}\}$($N_{vm}$ is number of vanishing moments) that maximize the coding gain and minimize the energy of the transition zone. Our optimization problem is multi-objective, as described below:

$$min_{\theta_k}(Objf_1, Objf_2), and \begin{cases} Objf_1 = TBE \\ Objf_2 = \frac{1}{CG} \end{cases} \tag{16}$$

In our work, a set of angles is regarded as a chromosome and is optimized by the multi-objective genetic algorithm NSGA III to obtain a set of filter banks to minimize all pre-specified objective functions at a satisfactory level.

5. Results and discussion

Before evaluating our results, we give some important parameters of the genetic algorithm used in this work. The probabilities of the crossover and mutation operators are 0.5 and 0.02, respectively. The population size is set to 80, and the chromosomes of the initial population are obtained by randomly generating angles in the [0, π] interval. The maximum number of generations is set to 200.

First, our design method is applied to design orthogonal filter banks of length $L = 8$ with at least one and two vanishing moments. The design procedure of the optimal filter bank is achieved via the following steps:

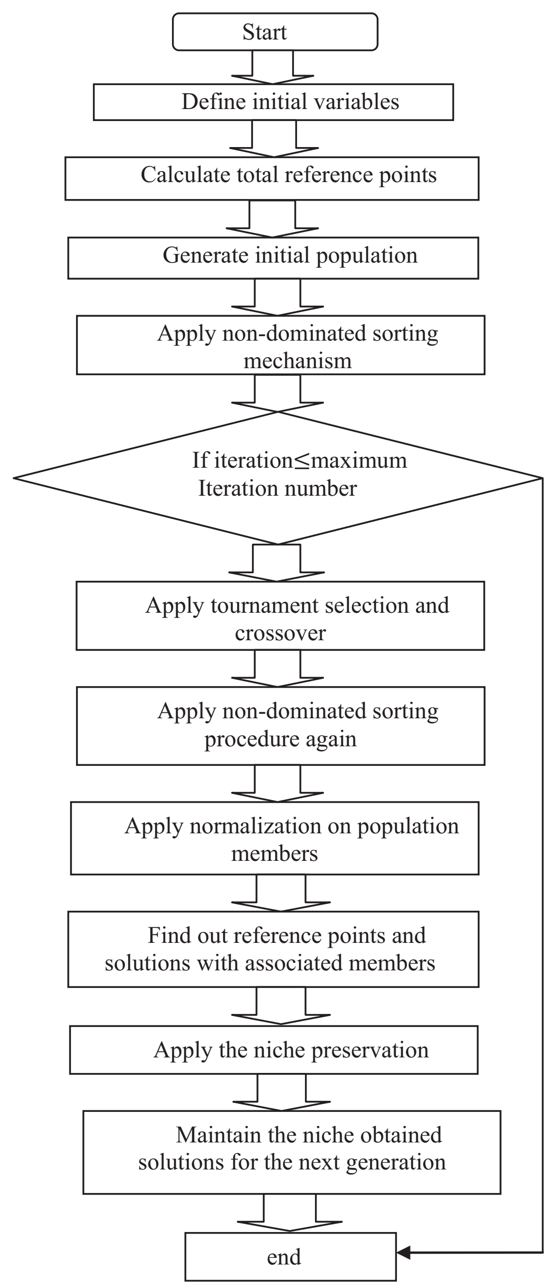


Fig. 5 Flowchart of the NSGAIII.

- The Multi-objective genetic method NSGAIII is first employed to find the Pareto optimal solutions (angles θ_k) that optimize the coding gain and TBE.
- The best compromised parameter θ^{opt} is selected from this set.
- A final filter bank (filters coefficients) is computed from these parameters using the parameterization proposed in [7].

Fig. 6 shows a scatter plot of a set of Pareto optimal solutions obtained for two optimized filter banks. In each case, a final solution is selected from a set of Pareto optimal solutions. The selected solutions labelled "Opt_1" and "Opt_2" (the index indicates the number of vanishing moments) correspond to the optimal angles θ_1^{opt} and θ_2^{opt}, respectively:

$$\theta_1^{opt} = \{1.63241293, 2.74983605, 1.93924546\}$$
$$\theta_2^{opt} = \{0.50451695, 1.86743322\}$$

The compression is expressed in terms of compression ratio as:

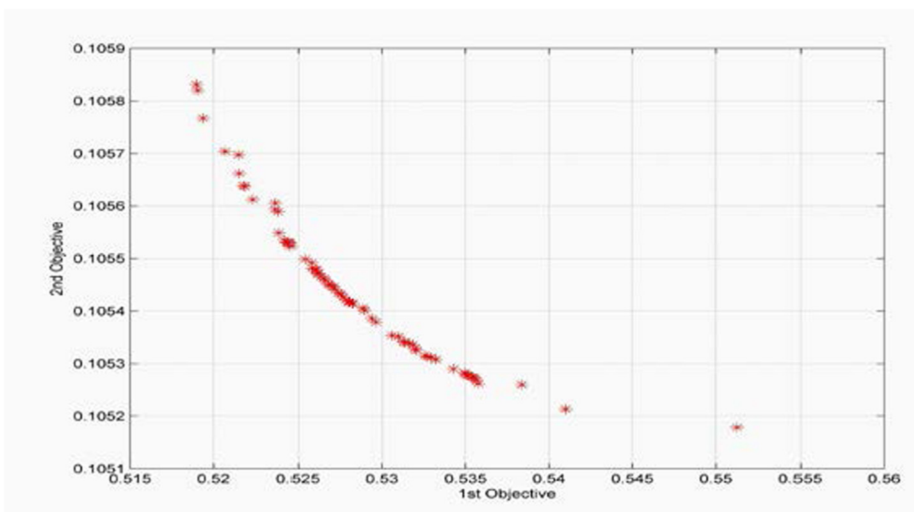
$$bitrate_{bit\ per\ sample} = \frac{number\ of\ bits\ of\ the\ compressed\ signal}{number\ of\ samples\ of\ the\ original\ signal} \tag{17}$$

Here, the entropy is used to estimate the bitrate.

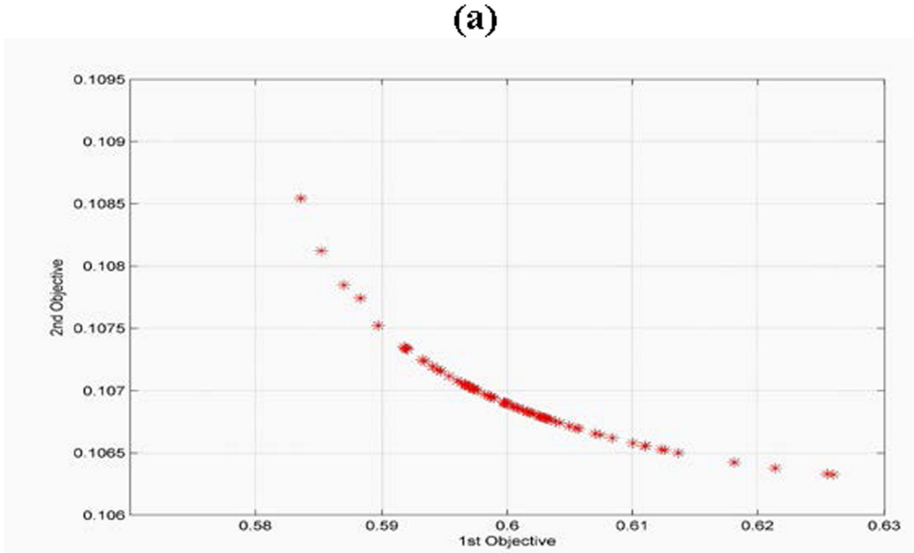
The distortion is measured in terms of Signal-to-Noise Ratio SNR. It is expressed as:

$$SNR_{dB} = 10log_{10}\frac{\sum_{i=1}^{N} x^2(n)}{\sum_{i=1}^{N}(x(n) - \widehat{x}(n))^2} \tag{18}$$

where, $x(n)$ and $\widehat{x}(n)$ are respectively the original and the reconstructed signals. N is the length of the signal. Here, three speech signals obtained from the TIMIT database are used for the evaluation of the filter banks coding performance.



(a)



(b)

Fig. 6 3-D scatter plot of the Pareto optimal solutions obtained for filter bank **Opt1 (a)** and **Opt2 (b)**.

Table 1 Speech coding performances of the proposed design method.

Speech Signals	Bitrate (bps)	Db4 Filter bank	Optimized filter bank "Opt1"	Optimized filter bank "Opt2"
	0.50	7.23	7.34	**7.43**
« SA1»	0.85	10.21	**10.40**	10.34
	1.00	11.30	**11.45**	11.37
	1.30	13.26	13.63	**13.63**
« SX37»	0.50	10.26	**10.74**	10.70
	0.85	14.19	14.24	**14.33**
	1.00	15.48	15.63	**15.63**
	1.30	17.98	**18.16**	18.15
« SI1027»	0.50	7.49	**7.60**	7.57
	0.85	10.55	**10.63**	10.61
	1.00	12.01	**12.15**	12.12
	1.30	14.30	14.43	**14.62**

Table 2 Comparison between characteristics of filter banks.

Filter bank	N_{vm}	TBE	CG_{dB}
Db4	4	0.80	9.66
Opt_1	1	0.55	9.78
Opt_2	2	0.65	9.74

In experimentation, the wavelet based system discussed in Sect. 2 has been used for speech compression. In order to verify the effectiveness of our design method, we compare the performance of the optimized filter bank with the performance of Daubechies orthogonal filter bank labelled "Db4" [27] in Table 1. We present the SNR results of all test signals at 0.5, 0.85, 1, 1.3bps (bit per sample), highlighting the best results in each case.

As it can be easily observed from Table 1, for all speech signals and in the majority of bitrates, our optimal filter banks outperform the Db4 filter bank. The best SNR gain is achieved by the filter bank "Opt1" in the case of « SX37» for bitrate 0.5bps, which has the value 0.48 db.

To justify the improvement of performances obtained with our filter banks, we have compared the values of objective functions to those of the Db4 filter bank in Table 2. Ideally, we wish to have filter banks that achieve high coding gain and low transition band energy. As desired, our design method leads to a significant diminution in the approximate error energy of the filter with respect to Db4. In addition, the filter bank that we designed provides the highest coding gain, which has a good correlation with coding performance.

In Fig. 7, we have compared the frequency responses of the low-pass filters of different filter banks mentioned above. This figure shows that our optimized filters Opt_1 and Opt_2 have a stepper transition band than that of Db4. It is clear from this figure that the Filter bank Opt_2 provides the best transition band which can be predicted from its TBE values presented in Table 2. On the other hand, the Db4 filter bank has the most flatness in magnitude responses at $\omega = \pi$, which is due to his higher number of vanishing moments (i.e. N_{vm}).

Fig. 8 illustrates the test speech signal "SX37" compressed at 0.5 bps using different filter banks. While for this bitrate, all reconstructed signals suffer from undesirable distorsions, it

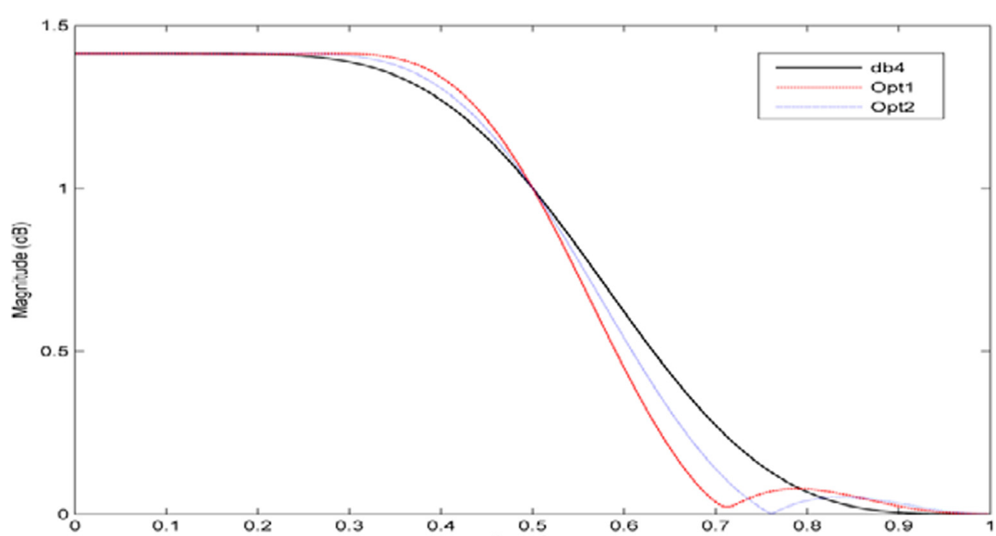


Fig. 7 Comparison between amplitudes of the frequency responses of the three analysis low-pass filters.

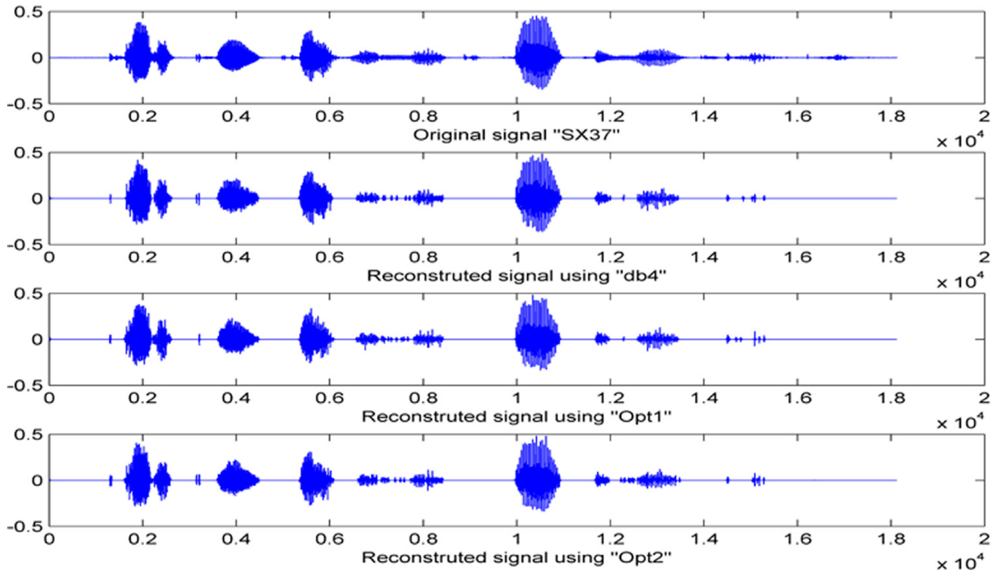


Fig. 8 (a) Original speech signal (SX37) and his reconstructed signals.

Table 3 Coding performances comparison of the optimized filter bank with the Db10 filter bank and the filter banks of [11].

Speech Signals	Bitrate (bps)	Db10 Filter bank	Kaiser10 Filter bank	Blackman10 Filter bank	Optimized filter bank*
	0.50	7.20	7.34	7.33	**7.40**
« SA1»	0.85	10.44	10.35	10.35	**10.45**
	1.00	11.55	11.51	11.52	**11.74**
	1.30	13.68	13.71	13.71	**13.82**
« SX37»	0.50	10.71	10.90	10.87	**11.04**
	0.85	**14.55**	14.54	14.51	14.52
	1.00	15.88	**16.06**	15.99	15.95
	1.30	18.38	18.43	18.39	**18.45**
	0.50	7.32	7.47	7.49	**7.60**
	0.85	10.76	10.84	10.77	**10.86**
« SI1027»	1.00	12.15	12.15	12.16	**12.19**
	1.30	14.51	14.53	14.50	**14.59**

$^{*}\boldsymbol{h}_0 = \{-0.0007754435\ \ -0.0039807964\ 0.0148295728\ 0.0157404073$
$-0.0458952614 \cdots -0.0258408218\ 0.0518062654 - 0.0334591674 - 0.1151005801$
$0.1627206002 \cdots 0.6337223808\ 0.6833961936\ 0.2001006647 - 0.1731227719 - 0.0590434941$
$\cdots 0.0987661624\ 0.0314434730\ \ -0.0178884682\ \ -0.0039807964\ 0.0007754435\}$

can be perceived that the optimized filter banks produce the best ones.

Table 3 shows a comparison of the proposed wavelet filter banks with the filter banks obtained by the optimization technique developed in [11] (see Table 1 in [11]) and the Daubechies filter of length 20 labelled "Db10" [27]. It is important to note that, in order to ensure a fair comparison, the filters should have the same length. As it can be seen, our optimized filter bank outperforms those of Db10 as well as the filter banks of [11] for the majority of bitrates.

6. Conclusion

In this work, we have proposed an optimization method for designing wavelet orthogonal filter banks in a speech coding scheme via a multi-objective genetic algorithm. A parameterization is used to implement a perfect orthogonal FIR filter bank having one and two vanishing moments. The optimal parameter set is the parameter set that maximizes the design criteria (i.e., coding gain and frequency selectivity). Since the optimization problem is multi-objective, we have exploited the non-dominated sorting genetic algorithm NSGAIII to solve it. This algorithm searches for a solution that achieves a compromise between diverse design criteria, which is called the Pareto optimal solutions.

From the obtained experimental results, it is shown that our optimized filter banks outperform the Daubechies filter bank as well as the optimized filter banks developed in [11].

We can state that by sacrificing the high degree of regularity of orthogonal wavelet filter banks, superior speech compression performance can be achieved with filters that demonstrate good energy compaction and low subband aliasing.

The optimality of filter banks is based on the stated design criteria. It could be interesting to search and exploit other criteria related to the objective quality of the speech signal as well as the subjective one.

Multi-objective GAs have shown that they are able to work well in the design of orthogonal wavelet filter banks. This design method could also be extended to other types of filter banks in order to be exploited in future works.

In addition, there exist other novel approaches developed for solving nonlinear optimization problems as those in [28] and [29]. These methods can be used to design effective filter banks and may be able to improve the coding performances.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research being reported in this publication was supported by the Algerian Directorate General for Scientific Research and Technological Development (DGRSDT).

References

[1] J. Liang, Filter bank design for multimedia coding and digital communications: Fast DCT, Lapped Transform And Beyond, LAP Lambert Academic Publishing (2010).
[2] J.D. Gibson, Speech coding methods, standards, and applications, IEEE Circuits Syst. Mag. 5 (4) (2005) 30–49.
[3] J.I. Agbinya, Discrete wavelet transform techniques in speech processing, Proceedings of Digital Processing Applications (TENCON'96), 1996, pp. 514–519.
[4] S.M. Joseph, P. Babu Anto, Speech compression using wavelet transform, in: Proceedings of International Conference on Recent Trends in Information Technology (ICRTIT2011), 2011, pp. 754–758.
[5] S. Joseph, P. Anto, The optimal wavelet for speech compression. Communications in Computer and Information Science, ACC 2011, Part III, CCIS 192, (2011), 406-414.
[6] J.R. Winkler, Orthogonal wavelets via filter banks theory and applications, The University of Sheffield Department of Computer Science, 2000.
[7] L.-K. Shark, C. Yu, Design of optimal shift-invariant orthonormal wavelet filter banks via genetic algorithm, Journal of signal processing, Elsevier 83 (12) (2003) 2579–2591.
[8] D.B.H. Tay, Zero-pinning the Bernstein polynomial: A simple design technique for orthogonal wavelets, IEEE Signal Process Lett. 12 (12) (2005) 835–838.
[9] H.M. Paiva, M.N. Martins, R.K.H. Galvao, J.P.L.M. Paiva, On the space of orthonormal wavelets: Additional constraints to ensure two vanishing moments, IEEE Signal Process Lett. 16 (2) (2009) 101–104.
[10] J. Yan, W.S. Lu, Towards global design of orthogonal filter banks and wavelets, Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim) ,Victoria, BC, Canada (2009) 187–192.
[11] A. Kumar, G.K. Singh, G. Rajesh, K. Ranjeet, The optimized wavelet filters for speech compression, Int. J. Speech Technol. 16 (2) (2013) 171–179.
[12] N. Aloui, B.N. Mohamed, A. Cherif, Genetic algorithm for designing QMF banks and its application in speech compression using wavelets, Int. J. Speech Graphics Signal Process. 5 (6) (2013) 1–8.
[13] T. Mourad, M. Bouhlel, A psychoacoustic model and a filter bank design using optimization for speech compression, Adv. Modell. Anal. B 61 (2) (2018) 80–87.
[14] V.V. Bolisetty, I.S. Prabha, A Novel Speech Compression Technique using optimized wavelet transform to improve the quality of auditory perception under low snr conditions, Int. J. Smart Security Technolog. (IJSST) 19 (6) (2019) 1473–8031.
[15] R. V. Ravi, K. Subramaniam, T. Roshini, S. P. B. Muthusamy, G. K. D. Prasanna Venkatesan, Optimization algorithms, an effective tool for the design of digital filters; a review. Journal of Ambient Intelligence and Humanized Computing, (2019).
[16] K. Deb, H. Jain, An Evolutionary Many-Objective Optimization Algorithm using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints, IEEE Trans. Evolutionary Comput. 18 (4) (2014) 577–601.
[17] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1998) 674–693.
[18] S. Kornsing, J. Srinonchat, Enhancement speech compression technique using modern wavelet transforms, International Symposium on Computer, Consumer and Control, 2012.
[19] S.M. Joseph, P. Babu Anto, Speech coding based on orthogonal and biorthogonal Wavelet, Procedia Technol. 6 (2012) 397–404.
[20] P.P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice-Hall, Englewood Cliffs, NJ, 1993.
[21] J. Katto, Y. Yasuda, Performance evaluation of subband coding and optimization of its filter coefficients, in: Procedings of SPIE, Visual Communications and Image Processing, 1991, pp. 95–106.
[22] M. Lightstone, E. Majani, S.K. Mitra, Low bit-rate design considerations for wavelet-based speech coding, Multidimension. Syst. Signal Process. 1997 (8) (1997) 111–128.
[23] C. A. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen, Evolutionary algorithms for solving multi-objective problems. Genetic and Evolutionary Computation Series, 1st edition, Springer, (2007).
[24] C. A. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen, Evolutionary algorithms for solving multi-objective problems. Genetic and Evolutionary Computation Series, 5, 2nd edition, (2014), Springer.
[25] X. Ma, J. Yang, N. Wu, Z. Ji, Z. Zhu, A comparative study on decomposition-based multi-objective evolutionary algorithms for many-objective optimization, IEEE Congress on Evolutionary Computation (CEC), 2016.
[26] K. Li, R. Wang, T. Zhang, H. Ishibuchi, Evolutionary Many-Objective Optimization: A Comparative Study of the State-of-the-Art, IEEE Access, 2018, pp. 26194–26214.
[27] I. Daubechies, Orthonormal bases of compactly supported wavelets, Commun. Pure Appl. Math. 41 (7) (1988) 909–996.
[28] C.H. He, An introduction to an ancient Chinese algorithm and its modification, Int. J. Numer. Meth. Heat Fluid Flow 26 (8) (2016) 2486–2491.
[29] S. Zulqurnain, A. Zahoor Raja, L. Dac-Nhuong, A. Aly Ayman, A neuro-swarming intelligent heuristic for second-order nonlinear Lane-Emden multi-pantograph delay differential system, Complex Intelligent Syst. (2021).