Developing a Simulator Applied to Audio Coding Process MPEG-4 AAC

Mauricio Harff, Marcio Garcia Martins, Arthur Tórgo Gómez
Postgraduate Interdisciplinary Program in Applied Computing
University of Vale do Rio dos Sinos
São Leopoldo, Brazil
e-mail: mharff@gmail.com, marciog@unisinos.br, breno@unisinos.br

Abstract—**This paper proposes the development of a simulator to find the internal parameters of the MPEG-4 Advanced Audio Coding (AAC), so as to optimize the perceptual audio quality for a given bit rate. The implementation of the simulator was developed in ANSI C programming language, using the Tabu Search and Genetic Algorithm in a hybrid structure. Through the minimization of the Average Noise-to-Mask Ratio (ANMR) metric, the simulator identifies the best configuration of internal parameters of the MPEG-4 AAC and improves the perceptual audio quality.**

Keywords-audio compression; MPEG-4 Advanced Audio Coding; Metaheuristics.

I. INTRODUCTION

During the encoding process, the MPEG-4 AAC encoder must dynamically choose the parameters that determine the best perceptual audio quality. This process of determining the best parameters was defined as the AAC Encoding Problem [1]. In this process, the reference encoder uses an iterative method known as Two Loop Search (TLS) [2] to define the parameters for a particular audio frame. However, this search method of parameters used, with respect to encoder MPEG-4 AAC, does not solve optimally the AAC Encoding Problem [3]. This paper proposes a simulator applied to the AAC encoder, based on a hybrid algorithm of metaheuristics Tabu Search (TS) [4] and Genetic Algorithm (GA) [5], that through experiments may define the parameters through the obtaining of solutions of good quality. Thus, the simulator proposed realizes the function of the dynamic definition of internal parameters of encoder configuration.

This paper is organized as follows. Section 2 presents related work. Section 3 presents and discusses the MPEG-4 encoder. Section 4 presents the mathematical formulation proposed. Section 5 shows the computational architecture of the simulator. Section 6 presents the results obtained. And finally, Section 7 presents the conclusions.

II. RELATED WORK

In the literature, one can find some works that address the AAC Encoding Problem, using different solution techniques and some variations in the structure of the solution model. These approaches may simplifying the model, proposing simpler methods to solve it, or make the model more faithful to reality, by the addition of decision variables and generate results with higher perceptual audio quality. The problem with the addition of more decision variables is that the problem becomes more complex to solve, and so mathematical methods of solution require a higher computational time for its resolution. In the work of Aggarwall [2], an algorithm called Search Trellis, based on a trellis arrangement is introduced to improve the efficiency of the AAC encoder. This paper proposes a quantization step optimization in the encoder, so that get produce at lower bit rates when compared to the reference encoder. In [2], were used only two decision variables, the Scale Factors (SF) and Huffman Code Books (HCB). The model takes into account in the Objective Function (OF), the amount of bits needed to represent the quantized spectral coefficients and also the side information while preserving the perceptual quality of the audio, i.e., minimizing the ANMR metric. The results obtained in the Aggarwall's work show an improvement in perceptual audio quality, compared to the reference encoder to a same bit rate. The improvement achieved resulted in an increase of two times in the perceptual audio quality, which represents a decrease twice in the ANMR metric.

Continuing the work of Aggarwall [2], an improvement was proposed by Bauer [3]. Even using only two decision variables, the SF and HCB, the author proposed a different technique to solve the AAC Encoding Problem. The method, called Fast Trellis Search, provides a perceptual quality very close to Trellis Search algorithm, but solves the AAC Encoding Problem approximately 25 times faster. Due to the computational time needed to solve the problem, the old technique was unable to make an implementation of AAC encoder in real time.

In [2] and [6], the authors, also discuss the AAC Encoding Problem. However, these works use not only two decision variables in the model, and so the problem becomes more true to life and at the same time more complex. Using Mixed Integer Linear Programming (MILP), the problem is modeled by making use of four decision variables: blocks structure, grouping of blocks, SF and HCB. The mathematical model that describes the problem to four decision variables is complex and extensive. In this approach, the authors are able to obtain significant improvements in perceptual audio quality when compared with the structures of the reference encoding models and the others discussed before. However, due to the complexity of the problem, the computational processing time becomes very high. This makes it impossible an encoder with such structure of parameter optimization to be implemented in real time.

 ISBN: 978-1-61208-398-8

III. MPEG-4 AAC ENCODER

The dynamics of the MPEG-4 AAC encoder starts with data processing, the conversion of the audio signal in time domain to the frequency domain using the Modified Discrete Cosine Transform (MDCT). After this procedure, the coefficients of a signal frame are divided into blocks, may be formed one long block or 8 short blocks [7]. If the frame coefficients are divided into 8 short blocks, they can still be combined into sections to reduce the total bit rate. Within a long block or sections, spectral coefficients are divided again, however in the frequency domain, in Scale Factor Band (SFB). For each SFB, the encoder sets an amount of bits to represent the frequency coefficients. This procedure occurs dynamically, in real time, through the iteration of bit rate control loop and the distortion control loop [2]. These loops define the quantization interval and the HCB that better configure the encoder to a SFB of a frame. Furthermore, it is necessary that the encoder transmits in its output data flow, the value of SF and HCB associated to each SFB, as side information [8]. The encoder also uses some bits, to tell the decoder which the grouping used in the sections of the blocks. In this way, the overall bit rate needed to encode, a given audio signal, is the sum of the bits needed to the coding of the sections and of the blocks, with the sum of the bits needed to encode side information. Therefore, the four parameters to be managed by the MPEG-4 AAC encoder can be defined [3]:

- Blocks Structure (short block, long block);
- Blocks Grouping (only for short blocks);
- Quantization Interval (Scale Factors);
- Entropy Encoder (Huffman Code Books).

To ensure a low bit rate, the encoder must choose a good combination of the four parameters, so that the predefined bit rate not is exceeded and that the perceptual audio signal quality is guaranteed. To measure the perceptual encoded signal quality, the ANMR metric is used. This is a metric that incorporates the human ear model in its conception [9]. The noise or quantization errors for specific frequency bands (critical bands) [9] are computed, so that the level of importance of noise at a given frequency range is obtained through the response curve of the human ear model [10]. In this way, it is possible to represent the perceptual audio quality through an exact parameter [11].

In [1][3] and [6], the authors define the ANMR optimization process as the AAC Encoding Problem. This problem resembles a classical combinatorial optimization problem, the Part Selection Problem. The problem of selecting parts belongs to the group of NP-Complete complexity problems [10].

Figure 1 shows the architecture of the MPEG-4 AAC encoder, with the simulator structure based on metaheuristics inserted in the diagram.

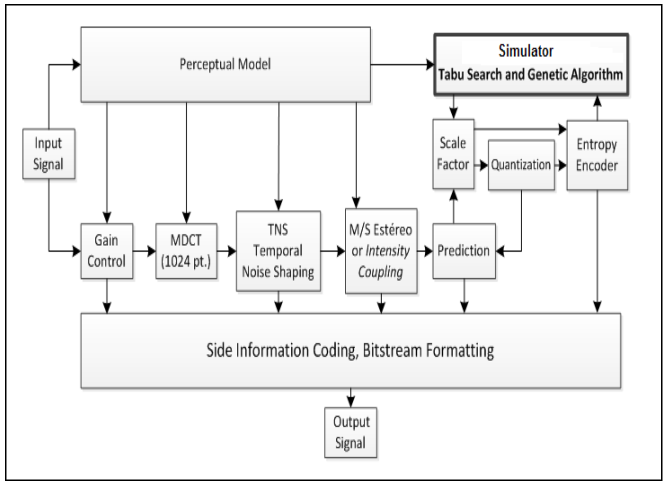

Figure 1. The architecture of the MPEG-4 AAC encoder with the proposed simulator.

Find the optimal solution to the problem of AAC encoding is a complex optimization problem, due to the interdependencies of the four parameters: blocks structure, blocks grouping, and quantization interval and entropy encoder [3]. The approaches found in the literature about the resolution of the AAC Encoding Problem, do not allow real time implementations. Thus, it is necessary simplify the problem. The AAC Encoding Problem is significantly simplified if the relationship between some parameters is neglected, as for example, the interdependence of adjacent Scale Factors, which are encoded differentially.

The method proposed by the standards [7][8] to solve the AAC Encoding Problem, is the iterative method TLS. However, some important disadvantages of the TLS procedure are that it does not necessarily converge, ignores the inter-band correlations of Scale Factors and Huffman Code Books, and is not intended to minimize the total distortion due to consider their analysis separately by bands [11] [12].

IV. MATHEMATICAL FORMULATION

The mathematical model, used in this paper, is based on the approach of Claus Bauer [3]. However, are considered three parameters of the problem: the quantization intervals, represented by SF, the HCB for each SFB, and the grouping of blocks for each frame, if it is configured as short block. The choosing process of the type of block, between short block and long block, is responsibility of the MPEG-4 AAC, because it uses the Perceptual Entropy metric [10], which identifies the minimum number of bits required to represent the frame. Thus, the objective of the simulator is to find the best combination of the three parameters, for a given bit rate pre-defined, so that the perceptual audio quality is maintained or improved.

Due to the mathematical formulation of the AAC Encoding Problem be very extensive, with several expressions to represent the constraints, in this paper, the mathematical model will be represented in a summary form. In [3], the full model is presented.

$Minimize$

$$ANMR(Z) = \frac{1}{N}\sum_{g=0}^{G+1}\sum_{i=1}^{8R}\sum_{a=1}^{M_1}\sum_{b=1}^{M_2} e_{g,i,a}Z_{g,i,a,b} \quad (1)$$

$Subject\ to$:

$$R(Z,A,B,g,U,X) \leq R_t \,, \quad (2)$$

where

$$R(Z,A,B,g,U,X) = \quad (3)$$

$$= \sum_{g=0}^{G+1}\sum_{i=1}^{120}\sum_{a=1}^{M_1}\sum_{b=1}^{M_2} Z_{g,i,a,b}Q_{g,i}(a,b) \quad +$$

$$+ \sum_{i=2}^{T}\sum_{a=0}^{2M_1} A_{L,i,a}F^*(a) + \sum_{i=2}^{120}\sum_{a=0}^{2M_1} A_{S,i,a}F^*(a) +$$

$$+ \sum_{i=2}^{T}\sum_{b=0}^{2M_2} B_{L,i,b}G^*(b) + \sum_{i=2}^{120}\sum_{b=0}^{2M_2} B_{S,i,b}G^*(b) \quad +$$

$$+ 4x_1 + 3\sum_{j=0}^{7}(x_{2,j} + x_{3,j} + x_{4,j}) + c\sum_{j=1}^{7} u_j \,,$$

where:

- N = Total number of Scale Factor Bands within a frame;
- R = Total bit rate;
- R_t = Total maximum bit rate specified;
- i = The i-th Scale Factor Band;
- j = The j-th part of a frame encoded as short block;
- g, G = Type of Block and Block Grouping;
- a, A = Scale Factor (SF);
- b, B = Huffman Code Books (HCB);
- $Z_{g,i,a,b}$ = Specific parameters of quantization error;
- $e_{g,i,a}$ = Distortion weight for a specific configuration;
- $Q_{g,i}(a,b)$ = Number of bits required to encode spectral coefficients for a particular configuration of SF and HCB;
- M_1 = Highest value admissible for the SF;
- M_2 = Highest value admissible for the HCB;
- L = Parameters associated with a long block;
- S = Parameters associated with a short block;
- $F^*() =$ Function that represents the number of bits to differentially encode the SF: $a_i - a_{i-1}$ (side information);
- $G^*() =$ Function that represents the number of bits for encoding the HCB: $b_i - b_{i-1}$ (side information);
- x, X = Auxiliary variables for calculating the quantity of bits (side information);
- U, u = Number of bits to represent the grouping of blocks (side information).

Equation (1) is the Objective Function (OF) of the AAC Encoding Problem that must be minimized. This function is a weighted sum of quantization errors (distortion) of audio signal in each Scale Factor Band (SFB). The quantization error is defined by $Z_{g,i,a,b}$ for the i-th SFB, to a configuration g of blocks, where the i-th SF is chosen equal to a and the i-th HCB equal to b. The inverse of the masking threshold, which defines the behavior of the human ear, is defined by the weight $e_{g,i,a}$ of the i-th SFB in the g-th configuration. In this way, the lower the value of the sum of ANMR, the greater the perceptual audio signal quality.

The transmission rate is the main constraint on the model of AAC encoder, because once determined by the user, this bit rate should not be exceeded. The maximum bit rate defined by the user is represented by R_t in (2). Thus, the bit rate produced by the encoder should always be less than the rate R_t.

The constraint (3) represents every portion of the bit rate, which together, comprise the total bit rate stream output of MPEG-4 AAC encoder. The first term of summations in (3) represents the amount of bits used to encode the spectral coefficients. The following two summations in (3) represent the bit rate to encode the Scale Factors of long blocks and short blocks respectively. Analogously, the following two summations represent the amount of bits to encode the Huffman Code Books, used in long and short blocks respectively. The last two summations represent the number of bits used to encode the arrangement structure of blocks, as side information. This side information is fundamental since it allows the decoder interpret how the data are organized within the data stream.

V. SIMULATOR ARCHITECTURE

The simulator architecture is composed by three modules: Initial Solution, Genetic Algorithm and Tabu Search. Thus, at the end of the implementation of metaheuristics, the best solution obtained to the configuration of the internal parameters is delivered to the MPEG-4 AAC encoder. It should be noted that this structure is performed individually for each audio frame.

Figure 2 shows the hybrid architecture developed to solve the AAC Encoding Problem, which is explained below.

Step 1: The initial solution used by the Genetic Algorithm is generated during the TLS iteration existing within the MPEG-4 AAC reference encoder. So, after run an iteration of external loop and an iteration of internal loop of the TLS algorithm [8], a solution is obtained. The solution still is randomly modified, with probability 0.05, before inserted into the initial population.

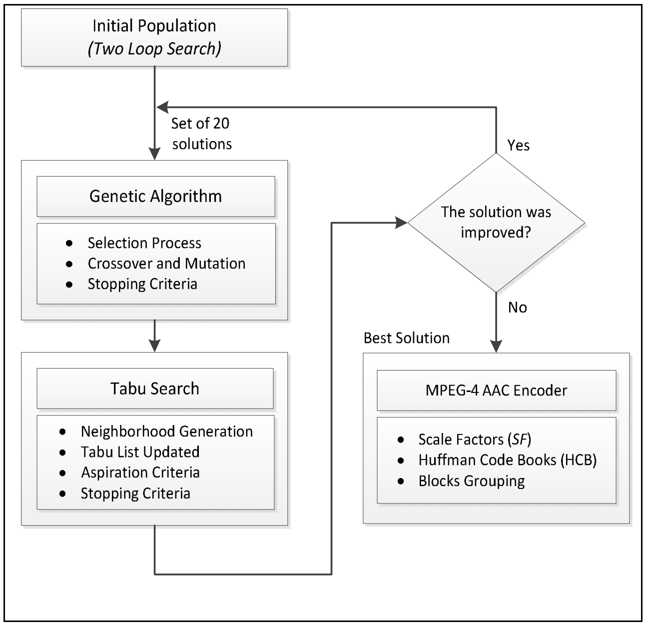

Figure 2. Hybrid architecture of the simulator.

Step 2: Utilizing the initial population, the GA generates new individuals, through crossover operator, and diversifies the population, through mutation operator. The process of selection by tournament, with a probability of 0.8 for the fittest, was used to identify individuals that will suffer crossover process. After the stop criterion is satisfied, the best solution obtained in the GA process is assumed as the initial solution of the TS. The adopted stopping criterion for TS process was a predefined number of iterations without improvement in OF value.

Step 3: From the initial solution, the TS generates neighbor solutions in an iterative process. The best solutions obtained in this process are stored in a new population, which is used as the input of the GA. TS is executed until the stop criterion is reached.

When the algorithm is finalized, the best solution obtained is passed to the MPEG-4 AAC encoder. From the data generated by the simulator, the frame encoding can be performed. The data of the ANMR are archived, for every frame, in order to allow the evaluation of the perceptual quality of the coded audio signal.

VI. COMPUTATIONAL EXPERIMENTS

The audio encoder and the simulator were programmed using the ANSI C language. Experiments and validation tests were performed through audio files. In this way, audio files in WAV format were used as input of encoder, which generates compressed audio files in AAC extension in its output. For the coding experiments a set of audio test was necessary, covering different types of music and sounds. Thus, based in [3][13][14][15], an audio files library was prepared to be used in the experiments of this work. The list of sounds composing the audio library test is shown in Table I. All WAV files listed have one audio channel with a sampling frequency of 44100Hz, and each sample is represented by 16 bits. To validate the developed computational model, experiments were performed in order to identify the influence of the three decision variables addressed in this study on perceptual quality metrics ANMR and bit rate. Through these experiments, the individual behavior of each one of the decision variables was studied. Thus, the value of only one of the variables was changed, while the values of other variables were held constant, totaling three validation experiments. As said before, the implementation was developed in ANSI C language, and it was inserted in the code of the Ffmpeg reference encoder, and was run on an Intel® Core™ i3 computer.

TABLE I. TEST AUDIO FILES

Index	Sound Type	Description	Duration
1	Drums	DG Samples – Rock beat Drums 02	8,15s
2	Blues	All Blues – Kora Jazz Band	15.03s
3	Bass	The Clairvoyant – Iron Maiden	9.50s
4	Choral	Symphony N^0 9 – Beethoven	10.03s
5	Crash China	DG Samples – Crash China	3.16s
6	Harpsichord	Oeuvres Pour Clavecin – François Couperin	10.02s
7	Narration	DG Samples – Vocals Shout 120	10.85s
8	Orchestra	Symphony N^0 9 – Beethoven	15.40s
9	Rock	It´s my life – Bom Jovi	15.07s
10	Sax	DG Samples – Sax Riff 128 C	8.24s

In the first experiment, to identify the individual behavior of SF, all frames are intentionally encoded as long blocks, and the fifth HCB was used, so that both variables would not influence on the measure of perceptual quality and bit rate. Following, the SF values were changed within a specific range [7] and [8], and its relation to the bit rate and perceptual quality is presented in Figure 3.

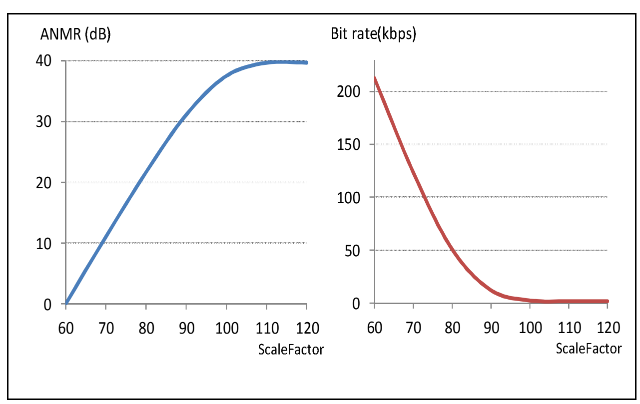

Figure 3. Scale Factor *vs* ANMR and bit rate.

The graphical analysis of the Figure 3 identifies that the smaller the value of SF, lower distortion is present in audio signal, i.e., greater the perceptual quality, and greater the number of bits required.

Similarly, the second validation experiment, all frames were also configured as long blocks, and the SF values were kept constant with a value 90. The values of HCB were changed in twelve possible values, and the results can be seen in Figure 4.

Through a graphical analysis of the Figure 4, the ranges of Largest Absolute Value (LAV) [7] and [8], from HCB are clearly identified, due the distortion of the audio signal does not vary for HCB of even values of LAV.

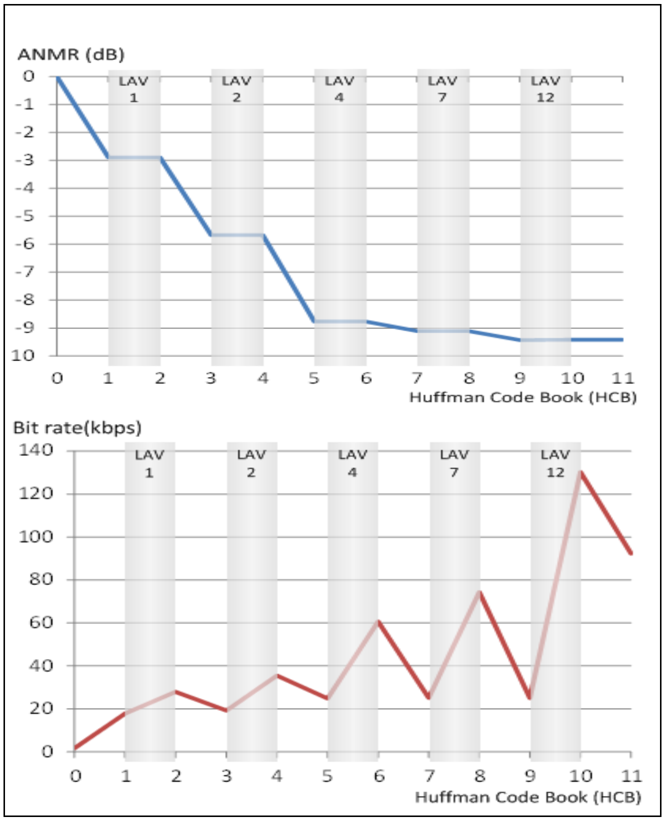

Figure 4. Huffman code book *vs* ANMR and bit rate

However, the bit rate is changed for HCB of even values of LAV, and increases as the value of LAV also increases [7]. The third experiment was performed to identify the influence of the grouping of blocks into sections, configured as a frame short block, on the metric ANMR and bit rate. Thus, all frames have been intentionally configured as short block, and the SF and HCB values were fixed at 90 and 5, respectively. In this experiment, combinations of seven sections were used, with the number of sections incremented from two to eight. It is observed that with the increased number of sections within one frame, the distortion value was reduced, and bit rate value was increased. This behavior follows the standard documentation [7] and [8] and is shown in Figure 5.

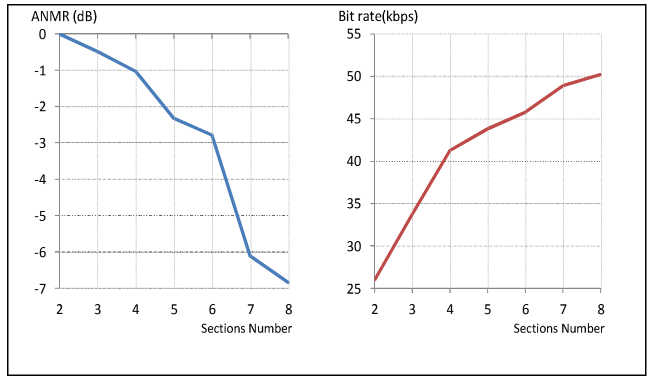

Figure 5. Huffman code book *vs* ANMR and bit rate

The dynamics observed in the three validation experiments, using the developed simulator, represent the behavior of the MPEG-4 AAC encoder [7] and [8] faithfully. Thus, the developed simulator is validated.

To set the calibration parameters of GA and TS, extensive experiments were conducted to identify which parameters combination that produced better values of OF. In GA the following values were set: mutation probability 0.9, crossover probability 0.8, and maximum number of generations 50. In the TS algorithm the value 200 was used for define the number of interactions without improvement in the FO value, utilized in the stop criterion, and the size of Tabu List. To evaluate the performance of the simulator developed for the MPEG-4 AAC encoder, the ten test audio files were encoded in four predetermined bit rates: 48kbps, 116kbps, 184kbps and 250kbps. Thus, it was possible to verify if the simulator model has produced better results regarding perceptual audio quality when compared to the reference encoder, using the TLS technique and Hybrid Algorithm (HA). The experimental results can be analyzed from Table II. Due to the stochastic behavior of metaheuristics, 50 encoding experiments for every audio file were executed, so that the results can be expressed by its mean and standard deviation values. For all audio files test, the HA had provided lower results for ANMR, i.e., a superior perceptual quality if compared to those obtained with TLS technique from reference encoder. The average execution time, per frame, for the rate of 48kbps was 50ms, 116kbps was 150ms, 184kbps was 300ms, and for 250kbps was 400ms. Table II presents the average of the attenuation of distortion and the respective standard deviation for the ten audio files tested, obtained by hybrid algorithm in relation to the process TLS, from reference MPEG-4 AAC encoder. The average values and the respective standard deviations were obtained from 50 runs of the simulator for all values shown in Table II.

TABLE II. RESULTS OF THE EXPERIMENTS.

File Index	Distortion (ANMR)							
	48kbps		116kbps		184kbps		250kbps	
	TLS	HA	TLS	HA	TLS	HA	TLS	HA
1	3.467	2.485 σ = 93	685	404 σ = 49	116	94 σ = 7	21	14 σ = 1
2	9.541	3.778 σ = 86	1.391	422 σ = 10	171	82 σ = 3	48	16 σ = 0,497
3	275	143 σ = 37	53	2 σ = 0,102	3	0,416 σ = 0,041	2	0,106 σ = 0,002
4	1.519	1.356 σ = 69	115	65 σ = 10	24	10 σ = 0,233	6	2 σ = 0,59
5	9.306	6.197 σ = 172	681	563 σ = 19	191	92 σ = 8	39	15 σ = 1
6	5.131	2.708 σ = 83	682	266 σ = 24	81	52 σ = 2	18	10 σ = 0,26
7	410	117 σ = 17	89	13 σ = 1	4	3 σ = 0,487	3	0,59 σ = 0,027
8	934	740 σ = 72	114	44 σ = 7	23	8 σ = 0,226	4	2 σ = 0,034
9	55.476	27.327 σ = 177	11.501	3.252 σ = 55	1.084	551 σ = 7	354	102 σ = 2
10	1.401	1.130 σ = 76	312	107 σ = 13	19	11 σ = 1	4	2 σ = 0,189

In Table III, it can be observed the gain (dB) of perceptual quality of the solutions found by the simulator,

TABLE III. HYBRID ALGORITHM *VS* TWO LOOP SEARCH (TLS).

ANMR (dB)	Bit rate			
	48 kbps	116 kbps	184 kbps	250 kbps
μ	-2,38	-5,17	-3,27	-4,92
σ	1,55	3,79	2,18	3,16

when compared to the solutions of reference MPEG-4 AAC encoder. The negative values represent a decrease of values obtained by the ANMR metric.

VII. CONCLUSIONS

This work has proposed a simulator that uses the GA in combination with the TS, which was developed to simulate the process of solve the AAC Encoding Problem. Currently, this problem is solved using the technique of the reference encoder MPEG-4 AAC, known as TLS.

The simulator was validated through extensive experiments, showing the dynamic behavior as expressed in the MPEG-4 AAC standard [9]. The experimental results showed a significant improvement in perceptual audio quality achieved by the HA proposed, when compared to the TLS technique from reference encoder. In all test runs, for different files and different bit rates, the simulator has produced the best results. For higher bit rates, processing time can be significant.

REFERENCES

[1] C. Bauer and M. Vinton,"Joint Optimization of Scale Factors and Huffman Code Books for MPEG-4 AAC," IEEE Transactions on Signal Processing, vol. 54, no. 1, Jan. 2006, pp. 177-189, doi: 10.1109/TSP.2005.861090.

[2] A. Aggarwal, S. L. Regunathan and K. Rose, "Near - Optimal Selection of Encoding Parametersf for Audio Coding," Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May, 2001, vol. 5, pp. 3269-3272, doi: 10.1109/ICASSP.2001.940356.

[3] C. Bauer, "The Optimal Choice of Encoding Parameters for MPEG4 AAC Streamed over Wireless Networks," Proc. 1st ACM Workshop on Wireless Multimedia Networking and Performance Modeling (WMuNep'05), Oct. 2005, pp. 93-100, doi: 10.1145/1089737.1089753.

[4] F. Glover, " Future Paths for Integer Programming And Links to Artificial Intelligence," Computers & Operations Research, May 1986, vol. 13, no. 5, pp. 533-549, doi:10.1016/0305-0548(86)90048-1.

[5] J. H. Holland, Adaptation in Natural and Artificial Systems: an introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press, Michigan, 1st edition, 1975.

[6] C. Bauer, M. Fellersand and G. Davidson, "Multidimensional Optimization of MPEG-4 AAC Encoding." Proc. International Conference on Acoustics, Speech, and signal Processing (ICASSP), May 2006, vol. 5, pp. 69-72, doi: 10.1109/ICASSP.2006.1661214.

[7] ISO/IEC,14496-3: Information Technology - Coding of Audio Visual Objects - Part 3: Audio, Geneva, 2005.

[8] ISO/IEC, 13818-7: Information technology - Generic Coding of Moving Pictures and Associated Audio Information : Part 7 - Advanced Audio Coding (AAC), Geneva, 2004.

[9] H. Fastl and E. Zwicker, Psychoacoustics: Facts and Models, Springer, New York, 3rd ed., 2007.

[10] J. D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal on Selected Areas in Communications, vol. 6, no. 2, Mar. 1988, pp. 314-323, doi: 10.1109/49.608.

[11] M. T. Ali and M. Saleem, "Improved Audio Quality at 48 Kbits/s for MPEG-4 AAC," Proc. Second International Conference on Electrical Engineering, Mar. 2008, pp. 1-8, doi: 10.1109/ICEE.2008.4553921.

[12] B. Wang, J. Zhang, J.,Y. Yao and L. Xie, "A New Bit-allocation Algorithm for AAC Encoder Based on Linear Prediction," Proc. 11ª International Conference on Communication Technology Proceedings, Nov. 2008, pp. 726-729, doi: 10.1109/ICCT.2008.4716222 .

[13] S. Wu and X. Qiu, "An Bit Allocation Method Based Rate-Distortion Control Algorithm for MPEG-4 Advanced Audio Coding," Proc. International Conference on Audio, Language and Image Processing (ICALIP2008), Jul. 2008, pp. 237-241, doi: 10.1109/ICALIP.2008.4590241.

[14] O. Derrien, P. Duhamel, M. Charbit and G. Richard," A New Quantization Optimization Algorithm for the MPEG Advanced Audio Coder Using a Statistical Subband Model of the Quantization Noise," IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 4, Jun. 2006, pp. 1328-1339, doi: 10.1109/TSA.2005.858041.

[15] N. C. Singh, "Measuring the Complexity of Sound," Pramana - Journal of Physics, vol. 77, no. 5, Nov. 2011, pp. 811-816, doi: 101007/s12043-011-0188-y.