Adaptive Speech Coding Method Based on Singular Value Decomposition and Grey Wolf Optimization for Arabic Language

Original Scientific Paper

Hassan Kassim Albahadily
University of Mustansiriyah, College of Science, Department of Computer Science
Baghdad, Iraq
hassan@uomustansiriyah.edu.iq

Alaa A. Jabbar Altaay
University of Mustansiriyah, College of Science, Department of Computer Science
Baghdad, Iraq
alaaaltaay@uomustansiriyah.edu.iq

Jamal N. Hasoon
University of Mustansiriyah, College of Science, Department of Computer Science
Baghdad, Iraq
Jamal.hasoon@uomustansiriyah.edu.iq

Abstract *– Speech coding plays a crucial role in maintaining speech quality while optimizing network resources and expediting transmission, as well as facilitating the storage of speech data. In this paper, an adaptive method for speech coding using singular value decomposition (SVD), grey wolf optimization (GWO), and run-length encoding (RLE) was proposed. The proposed method streamlines the speech matrix through preprocessing, converting it into short intervals. Subsequently, each interval undergoes decomposition using SVD, followed by optimization of compression quality using GWO. Finally, RLE is employed as the last step to increase space-saving. The developed method was conducted on two datasets: Quran and LibriSpeech using PSNR, PSEQ, and MOS tests. The results demonstrate promising outcomes, achieving space-saving up to 89.80, 84.04, 74.76, 67.24, and 59.52, respectively, for different values of quality (10, 20, 30, 40, and 50). GWO was used to optimize the quality factor which varies in each block, further increasing the space-saving up to 85.77. The average value of PSNR was equal to 21.3, MOS = 4.71, and PSEQ was equal to 3.95. Lastly, the RLE method effectively reduced the size of speech matrices, yielding a highly satisfactory space saving of up to 90.77, while maintaining excellent speech quality.*

Keywords*:* *Adaptive speech compression, Singular value decomposition (SVD), Grey wolf optimization (GWO)*

Received: November 12, 2023; Received in revised form: January 12, 2024; Accepted: March 16, 2024

1. INTRODUCTION

Speech is a form of audio data with specific requirements that must be met for the compressed data to be understandable [1]. Speech compression plays a vital role in modern digital life, as it minimizes storage requirements and facilitates transmission over networks. In both scenarios, the main objective is to reduce costs and save time. Speech compression is particularly important in applications such as teleconferencing, where transmitting large amounts of data is not cost-effective. Therefore, any method capable of reducing transferred data is considered cost-effective.

Adaptive speech coding methods have a crucial role in speech compression. These methods have enabled efficient speech transmission across various communication systems, including mobile networks, voice-over IP (VoIP), and streaming services [2].

Various methods are used for speech coding, including transformations along with optimization techniques such as discrete cosine transform (DCT), discrete wavelet transform (DWT), and singular value decomposition (SVD) [3].

Transformation is commonly used in compression, occurring after the preprocessing stage, to convert the data distribution from the time domain into the frequency domain to identify the potential for quantizing new data distribution in a way that effectively reduces its size [4].

Certain compression methods use swarm intelligence in the quantization process to achieve the desired quality of compressed speech [5].

SVD is a powerful mathematical tool commonly used for data compression. In the realm of speech compression, SVD is useful for reducing the dimensionality of speech, thereby aiding in the reduction of storage and transmission signals associated with speech data [6].

The singular vectors are a set of orthonormal vectors that span the same space as the original matrix, while the singular values are scalars that represent the relative importance of each singular vector [7]. Widely used in data science and engineering, SVD is a powerful tool for analyzing and manipulating matrices. SVD finds applications in various fields, including image and signal processing, data compression, machine learning, and many other fields [8].

Optimization is required to enhance the compression process. Grey wolf optimization (GWO) is a metaheuristic optimization algorithm introduced as a novel technique for resolving complex issues [9]. It has proven to be an effective optimization method [10], performing well in both unimodal and multimodal problem scenarios. GWO has successfully tackled optimization problems such as feature selection, image compression, and power system optimizations [11]. One notable advantage of GWO is simplicity and ease of implementation, with only a few parameters requiring tuning [12].

The structure of this paper comprises an introduction, a literature review, sections on SVD, GWO, RLE, the proposed method, and a final segment dedicated to results and conclusions.

The contribution of this work lies in highlighting the significance of utilizing the GWO optimization algorithm in conjunction with the SVD method to enhance speech compression. This involves selecting an optimal quality key to maximize compression efficiency. The proposed method is specifically applied to a distinct type of speech (the Quranic intonation of the Arabic language).

2. LITERATURE REVIEW

Recently, several studies have investigated adaptive speech coding methods, and a selection of these efforts is outlined below.

Hosny et al. [13] introduced two voice compression methods based on wavelet transforms, zero wavelet transform and average zero wavelet transform. These methods decompose the speech signal into multiple-resolution components, eliminating low-energy components to improve compression. This approach achieved a space-saving of 16.56 with a PSNR = 27.

Vig and Chauhan [14] proposed a hybrid wavelet method for voice reduction, breaking down the speech signal into multi-resolution components and eliminating low-energy signals using Walsh and DCT. By adjusting the threshold, this method achieved a space-saving of 72.10 with a PSNR of 49.71.

Alsalam et al. [15] employed contourlet and wavelet transforms for voice compression. The one-dimensional wavelet-transformed voice is converted into a two-dimensional array for contourlet transformation, followed by applying the contourlet transform on the high wavelet coefficients. This method achieved a space-saving of 53 with SNR = 33.

Vatsa and Sahu [16] proposed a speech compression method using discrete wavelet transform (DWT) and DCT with RLE and Huffman encoding to remove redundancies. The developed method was evaluated through compression factor, PSNR, MOS, and normalized root mean square error, achieving a space-saving of up to 29.11 with a PSNR of 16.39.

Bousselmi [17] developed an adaptive speech compression method based on the discrete wave atoms transform. This approach involves truncating signals based on the SNR and then using RLE and Huffman coding. The researchers found that the wave atom transform outperforms other wave transforms, achieving a notable space-saving of up to 10.78 with a PSNR of 36.74.

Abduljaleel [18] proposed a method for compressing and encrypting speech signals based on Sudoku, fuzzy C-means, and the Threefish cipher. The initial step involves removing low frequencies, followed by the fuzzy C-means method. This method successfully achieved a space-saving of 50.20 with a PSNR of 41.40.

Elaydi [19] introduced a lossy compression scheme using the DWT, resulting in a space-saving of 3.33 with a PSNR of 44.85.

The mentioned studies developed new techniques of speech compression using SVD, DCT, and DWT with some modifications or enhancements. All developed methods were tested on different datasets using objective tests like SNR, PSNR, and Compression factor and they achieved a good ratio of compression.

3. SINGULAR VALUE DECOMPOSITION

SVD is a mathematical tool providing substantial theoretical and technical insights into linear transformations with algebraic features [20]. It is decomposing a matrix into three matrixes returning the original matrix if they are combined.

The SVD decomposition of a matrix A can be represented by the following equation.

$$A = U\,S\,V^{T} \tag{1}$$

Where:

$U = m \times m$ matrix of orthonormal eigenvectors of AAT

$S = n \times n$ matrix of diagonal elements.

VT = the transpose of an $n \times n$ matrix containing the orthonormal eigenvectors of ATA.

Fig. 1 illustrates the SVD for matrix A [20].

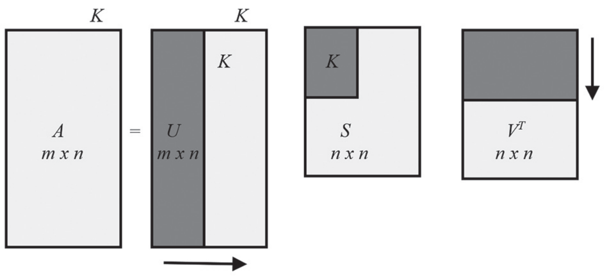

Fig. 1. SVD of matrix A

K represents the number of columns in the first matrix U, the number of elements from the diagonal of the second matrix S, and the number of rows in the third matrix V. The K parameter controls the quality of compressed speech and how many columns will be used in the decomposition process.

4. GREY WOLF OPTIMIZATION (GWO)

GWO is one of the swarm intelligent algorithms (metaheuristic algorithm) invented by Mirjalili *et al.* [21], which is modelling the hunting behavior of grey wolves [22]. The algorithm is designed to mimic the hierarchical structure and hunting strategy observed in grey wolves in the wild [23]. The GWO algorithm consists of four main components: social hierarchy, tracking/hunting, surrounding, and attacking prey [24, 25].

Grey wolves are categorized into four types based on their social hierarchy: alpha (α), beta (β), delta (δ), and omega (ω). The leadership hierarchy of wolves is illustrated in Fig. 2.

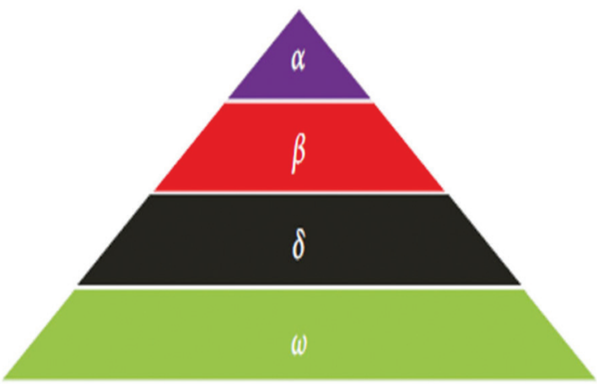

Fig. 2. The wolf leadership hierarchy

The primary decision-maker and leader is the alpha wolf, with beta and delta wolves supporting the alpha wolf in this role. The three leadership wolves (α, β, and δ), possessing the highest fitness levels, take charge of the hunting and optimization process, while the omega (ω) wolves follow their lead.

The following equations can be utilized to quantitatively represent the surrounding prey process:

$$X(t+1)=X(t)-A.D \tag{2}$$

where X_p represents the location of the prey, A defines the coefficient vector, and D is defined as:

$$parentD=|C.X_p\,(t)-X(t)| \tag{3}$$

where C stands for the coefficient vector, X defines the location of the grey wolf, and t represents the current iteration. The coefficient vectors A and C are determined by the following equations:

$$A=2a.r_1-a \tag{4}$$

where elements of a are linearly reduced from 2 to 0 over the sequence of iterations, and $r1$ and $r2$ define the random vectors in the range [0, 1].

Hunting: In terms of hunting, the first three prominent solutions (α, β, and δ) attained are stored and induce other search agents (including ω) to adjust their positions, considering the position of the best search agent. The positions of the grey wolves can be updated according to the following equations.

$$X(t+1)=(X\alpha+X\beta+ X\delta\,)/3 \tag{5}$$

5. RUN-LENGTH ENCODING (RLE)

RLE is a lossless approach that provides reasonable space-saving for specific data types by replacing consecutive data values in a file with a count number (run) and its value. The implementation of RLE sometimes depends on the data type being compressed. The operation of this method is illustrated in Fig. 3.

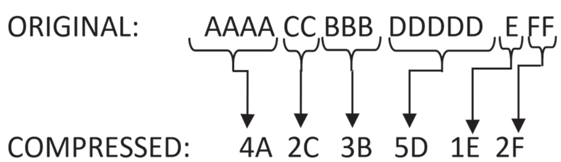

Fig. 3. The work of RLE algorithm

6. THE PROPOSED METHOD

The proposed adaptive method is introduced for speech compression, starting with the necessary preprocessing steps. This involves framing, removing silent intervals, and then reshaping a specific number of samples into a two-dimensional matrix for the decomposition process.

The framework of the proposed method is shown in Fig. 4.

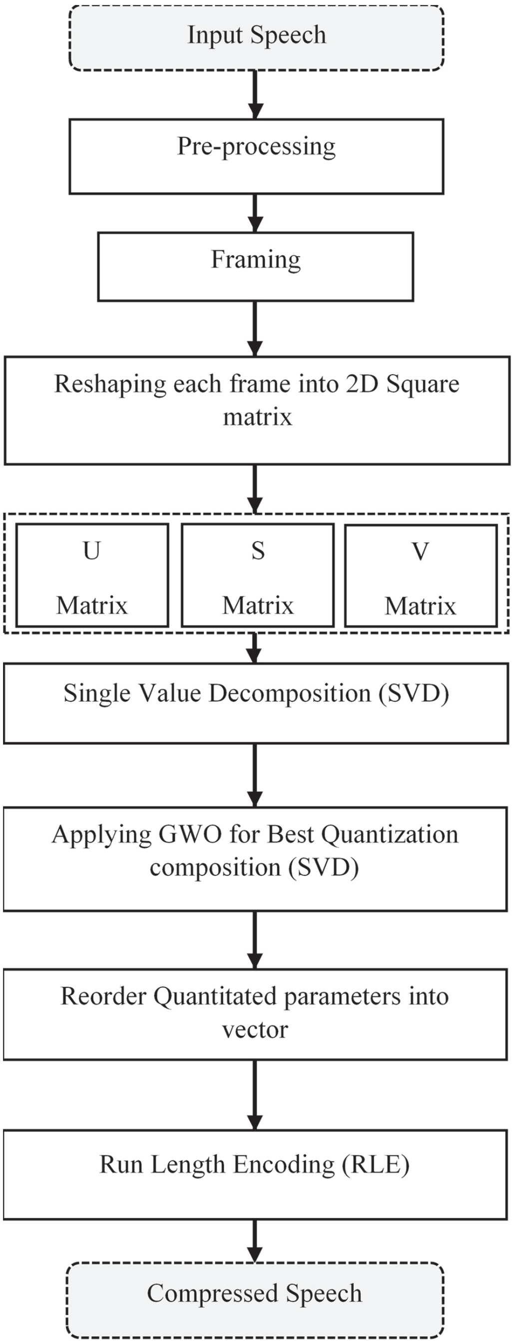

Fig. 4. The framework of the proposed method

6.1. PREPROCESSING OF SPEECH

The preprocessing of input speech involves two main steps: filtering the speech with a noise-removing filter called the Wiener filter and removing silence intervals.

The Wiener filter is designed to reduce the impact of additive noise in a signal while preserving the desired signal components. In the context of speech processing, the Wiener filter estimates the power spectral density of both the noise and the signal. Subsequently, it applies a frequency-dependent gain to the noisy signal, aiming to minimize the mean square error between the estimated clean signal and the noisy signal. This process enhances the SNR of the speech, making downstream speech processing tasks more effective.

Silence intervals, which are non-speech segments that contain no useful information, are then removed to reduce the computational data dimensionality, focusing solely on the speech data. This method involves establishing a threshold based on the amplitude or energy of the speech signal. Segments that fall below this threshold are identified as silence and removed. This step is beneficial in decreasing the amount of unnecessary data that needs processing, particularly in applications where only the speech content is relevant.

6.2. FRAMING OF SPEECH

Framing is the process of dividing a continuous stream of data representing speech signals into smaller segments called frames. This process is defined by specifying the size of the square block, such as 256×256, which is equivalent to 65,536 samples. After removing silent intervals, the remaining samples are partitioned into frames, each equal to this specified value. If the size is less than the selected value, the final frame is padded.

Before entering the decomposition process, the frames are reshaped into two-dimensional matrices. These matrices can have their dimensions (number of rows and columns) adjusted according to specific requirements by modifying the arguments passed to the reshape method.

6.3. SINGULAR VALUE DECOMPOSITION OF BLOCK RESHAPING

The SVD process is used to decompose a two-dimensional matrix into a one-dimensional matrix, allowing for efficient data manipulation. Each block will select k rows from the three matrices and ignore the remaining rows. The process of SVD decomposition is shown in Fig. 5.

6.4. APPLYING GWO FOR BEST QUANTIZATION COMPOSITION

The selection of k values for all blocks is enhanced through the application of GWO. GWO randomly selects numbers and then optimizes them to strike a balance between the output size and the quality of speech compression. This process involves using GWO to fine-tune the "k" values for data blocks, ensuring the desired balance between reducing data size and preserving information integrity is achieved.

6.5. REORDERING QUANTITATED PARAMETERS

In this step, the chosen parameters are rearranged into a single vector, which represents the compressed speech. The process is applied to each block, from start to end, and all other values are quantized to a specific decimal digit. This step is important for the following stage of lossless compression.

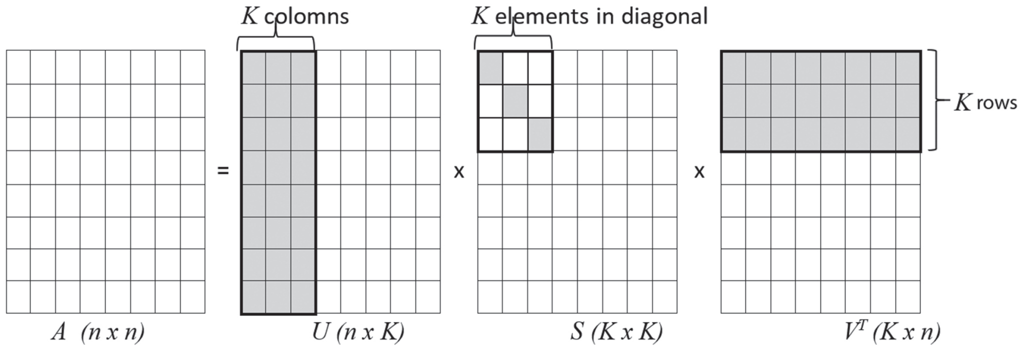

Fig. 5. Reshaping the matrix using SVD

6.6. RUN-LENGTH ENCODING OF QUANTIZED VECTOR

RLE is a highly efficient compression algorithm, especially effective when there is a long sequence of identical values in the data. It excels at reducing redundancy and significantly decreases the overall size of the data. This technique is applied to the truncated parameters, resulting in compressed data represented in pairs indicating the run and its length.

7. RESULTS

To assess the efficiency of the proposed method and validate the results, the method has been implemented on two datasets. The first dataset comprises twenty speech files of Quranic intonation S1-S5 with varying durations (3, 6, 9, and 12 seconds), frequency of 8000 KHz, and mono channel. The second dataset is LibriSpeech L1-L5 with durations (3, 6, 9, and 12 seconds), frequency of 8000 KHz, mono channel, yielding 300 seconds in total.

The experiments were conducted on the datasets using MATLAB 2020b on a computer with 2.4 GHz CPU frequency and 16 GB of memory under the Windows 10 operating system.

The space-saving factor was used as the reduction in size relative to the uncompressed size as shown in the following equation.

$$\text{Space Saving} = 1 - \frac{compresed\ Size}{uncompressed\ Size} \quad (6)$$

The obtained results, after applying SVD on the datasets, are presented in Table 1 and Table 2.

Table 1. The space-saving ratio for Quran intonation

	Duration (Second)			
***K* values**	**3**	**6**	**9**	**12**
10	89.91	88.29	89.77	91.61
20	84.16	84.10	83.69	84.23
30	75.34	76.10	74.65	72.94
40	66.29	67.35	68.16	67.16
50	59.08	59.23	60.20	59.54

The values in Table 1 represent the space-saving factor for test speech files of duration 3, 6, 9, and 12 seconds with five different values of the quality parameter (K) =10, 20, 30, 40, and 50. The results indicate that the average space-saving value is 89.90 when the quality factor K is set to 10. Subsequently, depending on the chosen compression quality K, the ratio decreases to 84.04, 74.76, 67.24, and 59.52 when K is set to 20, 30, 40, and 50, respectively.

Similar results were calculated for the dataset LibriSpeech as shown in Table 2.

Table 2. The space-saving ratio for LibriSpeech

	Duration (Second)			
***K* values**	**3**	**6**	**9**	**12**
10	91.57	90.29	90.15	90.22
20	83.29	82.22	82.06	84.21
30	75.04	75.55	75.52	75.43
40	68.33	66.82	67.39	68.04
50	60.31	60.82	59.11	59.87

The space saving is increased when K is set to a low value, implying good quality and the ratio decreases when K is set to a high value, indicating low speech quality. The relationship between the average space-saving and the quality factor K of values 10–50 is illustrated in Fig. 6.

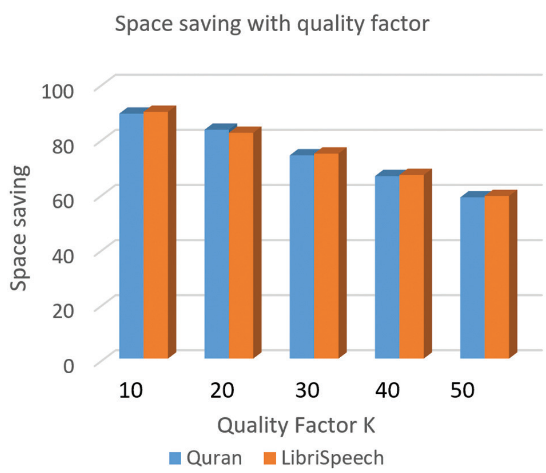

Fig. 6. The average space-saving of speech

To assess the quality of compressed speech and validate the results of the proposed method, several tests will be conducted on the recovered speech files. One commonly used test is PSNR, which measures the ratio between the maximum possible value of a signal and the power of distorting noise that impacts its quality.

The PSNR values are shown in Table 3 and Table 4.

Table 3. PSNR (dB) for Quran intonation

	Duration (Second)			
K values	3	6	9	12
10	6.93	4.76	4.70	4.42
20	9.14	10.96	10.38	8.26
30	17.11	16.17	14.34	17.54
40	17.31	16.40	15.04	21.05
50	28.86	24.77	23.09	27.17

Table 3 represents the PSNR values of test speech files of duration 3, 6, 9, and 12 seconds with five different values of quality parameter (K) =10, 20, 30, 40, and 50. The results indicate that the average PSNR value is 5.20 when the quality factor K is set to 10. Subsequently, depending on the chosen compression quality K, PSNR increases to 9.68, 16.29, 17.45, and 25.97 when K is set to 20, 30, 40, and 50, respectively.

Similar results were found for the dataset LibriSpeech as shown in Table 4.

Table 4. PSNR (dB) for LibriSpeech

	Duration (Second)			
K values	3	6	9	12
10	6.31	6.26	4.47	5.01
20	8.71	8.28	7.51	10.70
30	12.37	16.21	10.73	14.31
40	17.19	14.27	20.57	16.88
50	29.75	23.01	23.78	26.75

PSNR is increased when K is set to a high value, showing high noise and low speech quality. The relationship between the average PSNR and the quality factor K of values 10–50 is illustrated in Fig. 7.

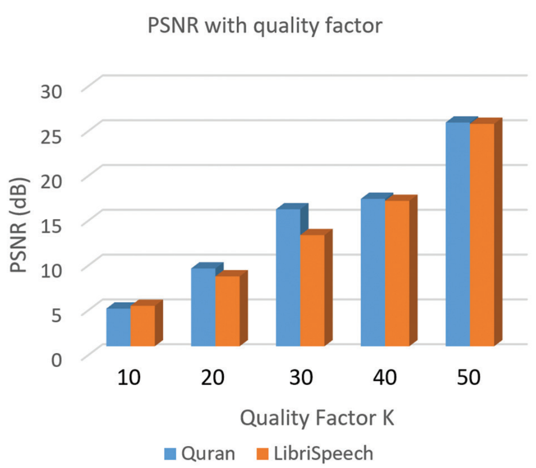

Fig. 7. The mean PSNR of the compressed speech

The speech frames contain varying data, and it is important to select a varying value of quality factor K according to the contents of the frame. The GWO is employed to make the selection of K varying, depending on the frame significance. The space-saving increases when using GWO optimization, albeit with a trade-off of relatively subdued speech quality. The results after applying GWO optimization are presented in Table 5 and Table 6.

Table 5. Space-saving with GWO for Quran intonation

	Duration (Second)			
Speech Files	3	6	9	12
S1	86.63	84.29	86.13	86.04
S2	85.21	84.91	85.64	85.88
S3	86.94	86.16	85.68	85.47
S4	84.97	85.28	86.12	85.79
S5	86.33	85.35	85.94	86.58
Average	86.02	85.20	85.90	85.95

Table 5 represents the space-saving factor after using GWO for test speech files of duration 3, 6, 9, and 12 seconds. The average values are listed in the last row. The results indicate that the average space-saving value is almost the same for the test files with an average = 85.77.

Similar results were calculated for the dataset LibriSpeech as shown in Table 6.

Table 6. Space-saving with GWO for LibriSpeech

	Duration (Second)			
Speech Files	3	6	9	12
L1	82.50	79.15	85.87	78.04
L2	84.51	80.15	82.11	80.53
L3	80.93	84.99	77.36	82.38
L4	80.84	84.87	80.17	77.33
L5	85.36	77.35	85.60	85.97
Average	82.83	81.30	82.22	80.85

The space-saving is increased for all files because the value of quality factor K is selected as the best value for each frame. The relationship between the average space-saving after using GWO is illustrated in Fig. 8.

Fig. 8. Space-saving of speech after GWO

The subjective test will be applied to check the quality of speech. The first test is Mean Opinion Score (MOS) which is used to evaluate results and measure the quality of compressed speech. MOS consists of five values (5 = excellent, 4 = very good, 3 = fair, 2 = poor, 1 = bad) to express the quality of speech as perceived by listeners [26]. The MOS test involves presenting both the original and compressed speech to ten native speakers individuals (five males and five females), who then assign values from 1 to 5 based on the quality they perceive. Forty speech files were used in the test, each file contains one sentence and several words depending on the duration of the file. The results of the MOS test are shown in Table 7 and Table 8.

Table 7. MOS test for Quran intonation

	Duration (Second)			
Speech Files	**3**	**6**	**9**	**12**
S1	3.76	4.36	3.54	3.93
S2	3.66	4.44	4.04	4.57
S3	4.02	4.20	4.55	4.18
S4	3.93	4.78	4.15	4.24
S5	4.64	3.70	3.58	4.95

Table 7 represents the average values of MOS for test speech files of duration 3, 6, 9, and 12 seconds. The results show that the average MOS values are 4, 4.30, 3.97, and 4.37 which indicate very good quality with an average value =4.71.

Similar results were calculated for the dataset LibriSpeech as shown in Table 8.

Table 8. MOS test for LibriSpeech

	Duration (Second)			
Speech Files	**3**	**6**	**9**	**12**
L1	4.84	4.60	4.70	3.76
L2	4.16	4.21	4.44	4.69
L3	3.74	3.85	4.06	3.96
L4	4.23	4.59	4.10	4.15
L5	4.48	4.11	4.55	4.43

The results obtained from Table 8 indicate that the quality was acceptable, and all compressed speech files were nearly identical to the original files.

Another subjective test is the Perceptual Evaluation of Speech Quality PSEQ, which analyzes speech signals and considers a good result if the score is above 3.5. The results of PESQ are shown in Table 9.

Table 9. PESQ test for Quran and LibriSpeech

Quran intonation		LibriSpeech	
Speech File	**PESQ**	**Speech File**	**PESQ**
S1	3.983	**L1**	3.822
S2	3.992	**L2**	3.914
S3	3.906	**L3**	3.899
S4	3.921	**L4**	3.904
S5	3.932	**L5**	3.903

Table 9 represents the average values of PESQ for test speech files of duration 3, 6, 9, and 12 seconds for both datasets Quran (S1-S5) and LibriSpeech (L1-L5). The results show that the average PESQ value is 3.95 for Quran files and 3.92 for the LibriSpeech dataset, which indicates very good quality for the compressed speech.

The final step involves applying the lossless compression method RLE, which relies on identifying repeated similar values present in the speech. The results from the RLE method increased the space-saving by 5–7%, resulting in a final space-saving of 90–92% for the collected dataset and 85-90% for the LibriSpeech dataset.

The suggested method has demonstrated promising results when compared with other referenced efforts. The comparison between the suggested method and other methods is presented in Table 10.

Table 10. Performance comparison with related work

Reference	Compression Method	Space Saving	MOS	PSNR
[12]	ZWT,AZWT	85.44	3.8	27.0
[13]	DCT	72.10	-	49.71
[22]	DWT, DCT	70.89	-	16.39
[23]	DWAT	89.22	-	36.74
[24]	C-Means	50.20	-	41.40
[25]	DWT	66.7	-	44.85
Suggested Method	SVD	85.36	4.16	21.3

According to the results from Table 10, and comparing the suggested method with the related studies, the suggested method provides good results and can be used in different cases and applications.

To validate the results and ensure the stability of the suggested method, the experiment was tested on the datasets by the authors as a trial presentation, then repeated two times with a total time equal to 150 minutes, in addition to the subjective test which takes 30 minutes which confirming the reliability of the tests to validate the judgments of the obtained results and the suggested method.

8. CONCLUSION

An adaptive method that combines SVD, GWO, and RLE has been proposed for compressing speech signals. The method has shown promising results, achieving up to a 92% reduction in the size of speech files compared to their original size. The quality of the compressed speech was evaluated using PSNR, which yielded a value of 21.3. The validation test was supported by the subjective tests MOS which was 4.71 and PESQ which yielded 3.95, indicating excellent speech quality.

For future work, it is essential to explore the application of the proposed method in speech storage and over VoIP protocols for transferring audio files through Internet-of-Things applications after encrypting the files. Additionally, it is recommended to utilize optimization algorithms based on AI and DNA, with a focus on saving the most frequently repeated words and generating their compressed equivalents.

ACKNOWLEDGMENT

The authors would like to thank the University of Mustansiriyah, Baghdad, Iraq, for the support and provision of open resources that facilitated the completion of this work.

9. REFERENCES

[1] R. Cox, C. Kamm, L. Rabiner, J. Schroeter J. Wilpon, "Speech and language processing for next-millennium communications services", Proceedings of the IEEE, Vol. 88, No. 8, 2000, pp. 1314-1337.

[2] H. Xie, Z. Qin, G. Li, B. Juang, "Deep Learning Enabled Semantic Communication Systems", IEEE Transactions on Signal Processing, Vol. 69, 2021, pp. 2663-2675.

[3] J. James, V. Jyothi, "A Comparative Study of Speech Compression using Different Transform Techniques", International Journal of Computer Applications, Vol. 97, No. 2, 2014, pp. 16-20.

[4] V. Maider, J. Amigo, "Pre-processing of hyperspectral images Essential steps before image analysis", Chemometrics and Intelligent Laboratory Systems, Vol. 117, 2012, pp. 138-148.

[5] M. Tuba, N. Bacanin, "JPEG quantization tables selection by the firefly algorithm", Proceedings of the International Conference on Multimedia Computing and Systems, Marrakech, Morocco, 14-16 April 2014, pp. 153-158.

[6] M. Wall, A. Rechtsteiner, L. Rocha, "Singular Value Decomposition and Principal Component Analysis", A Practical Approach to Microarray Data Analysis, Springer, 2003, pp. 91-109.

[7] S. Berkant, L. Eldén, "Handwritten digit classification using higher order singular value decomposition", Pattern Recognition, Vol. 40, No. 3, 2007, pp. 993-1003.

[8] A. Cichocki, "Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis", IEEE Signal Processing Magazine, Vol. 32, No. 2, 2015, pp. 145-163.

[9] M. Hemis, B. Boudraa, T. Merazi-Meksen, "Intelligent audio watermarking algorithm using Multi-objective Particle Swarm Optimization", Proceedings of the 4th International Conference on Electrical Engineering, Boumerdes, Algeria, 13-15 December 2015, pp. 1-5.

[10] A. Kaur, S. Sharma, A. Mishra, "An Efficient Opposition Based Grey Wolf Optimizer for Weight Adaptation in Cooperative Spectrum Sensing", Wireless Personal Communications, Vol. 118, 2021, pp. 2345-2364.

[11] L. Abualigah, A. Khader, E. Hanandeh, "A new feature selection method to improve the document clustering using particle swarm optimization algorithm", Journal of Computational Science, Vol. 25, 2018, pp. 456-466.

[12] A. Bilal, G. Sun, S. Mazhar, A. Imran, "Improved Grey Wolf Optimization-Based Feature Selection and Classification Using CNN for Diabetic Retinopathy Detection", Evolutionary Computing and Mobile Sustainable Networks, Vol. 116, Springer, 2021.

[13] N. Hosny, S. El-Ramly M. El-Said, "Novel techniques for speech compression using wavelet transform", Proceedings of the Eleventh International Conference on Microelectronics, Kuwait, 22-24 November 1999, pp. 225-229.

[14] R. Vig, S. Chauhan, "Speech Compression using Multi-Resolution Hybrid Wavelet using DCT and Walsh Transforms", Procedia Computer Science, Vol. 132, 2018, pp. 1404-1411.

[15] A. Alsalam, S. Razoqi, E. Ahmed, "Effects of Using Static Methods with Contourlet Transformation on Speech Compression", Iraqi Journal of Science, Vol. 62, No. 8, 2021, pp. 2784-2795.

[16] S. Vatsa, O. Sahu, "Speech Compression Using Discrete Wavelet Transform and Discrete Cosine Transform", International Journal of Engineering Research and Technology, Vol. 1, No. 5, 2012.

[17] S. Bousselmi, N. Aloui, A. Cherif, "Adaptive Speech Compression Based on Discrete Wave Atoms Transform", International Journal of Electrical and Computer Engineering, Vol. 6, No. 5, 2016, pp. 2150-2157.

[18] I. Abduljaleel, A. Khaleel, "Developed a Speech Signal Compression and Encryption Based on Sudoku, Fuzzy C-means and ThreefishF Cipher",

International Journal of Electrical and Computer Engineering, Vol. 11, No. 6, 2021, pp. 5049-5059.

[19] H. Elaydi, M. Jaber, M. Tanboura, "Speech Compression Using Wavelets", Proceedings of the International Arab Conference on Information Technology, Ajman, UAE, 6-8 December 2003, pp. 1-8.

[20] N. Albatayneh, K. Ghauth, F. Chua, "A Semantic Content-Based Forum Recommender System Architecture Based on Content-Based Filtering and Latent Semantic Analysis", Advances in Soft Computing, Vol. 287, 2014, pp. 369-378.

[21] K. Srikanth, "Meta-heuristic Framework: Quantum Inspired Binary Grey Wolf Optimizer for Unit Commitment Problem", Computers and Electrical Engineering, Vol. 70, 2018, pp. 243-260.

[22] J. Neeraj, "Applications of Grey Wolf Optimization in Control System", Acadimic Publishing, 2018.

[23] S. Mirjalili, A. Lewis, "Grey Wolf Optimizer", Advances in Engineering Software, Vol. 69, 2014, pp. 46-61.

[24] L. Wong, "Grey Wolf Optimizer for Solving Economic Dispatch Problems", Proceedings of the IEEE International Conference on Power and Energy, Kuching, Malaysia, 1-3 December 2014.

[25] B. Yang, "Grouped grey wolf optimizer for maximum power point tracking of doubly-fed induction generator-based wind turbine", Energy Conversion and Management, Vol. 133, 2017, pp. 427-443.

[26] P. Loizou, "Speech quality assessment", Multimedia Analysis, Processing and Communications, Springer, 2011, pp. 623-654.