

Tariq A. Hassan*, Rageed Hussein AL-Hashemy and Rehab I. Ajel

Speech Signal Compression Algorithm Based on the JPEG Technique

https://doi.org/10.1515/jisys-2018-0127 Received January 17, 2018; previously published online April 28, 2018.

Abstract: The main objective of this paper is to explore parameters usually adopted by the JPEG method and use them in speech signal compression. Speech compression is the technique of encoding the speech signal in some way that allows the same speech parameters to represent the whole signal. In other words, it is to eliminate redundant features of speech and keep only the important ones for the next stage of speech reproduction. In this paper, the proposed method is to adopt the JPEG scheme, which is usually used in digital image compression and digital speech signal compression. This will open the door to use some methods that are already designed to fit two-dimensional (2D) signals and use them in a 1D signal world. The method includes many priori steps for preparing the speech signal to make it more comparable with the JPEG technique. In order to examine the quality of the compressed data using the JPEG method, different quantization matrices with different compression levels have been employed within the proposed method. Comparison results between the original signal and the reproduced (decompressed) signal show a huge matching. However, different quantization matrices could reduce the quality of the reproduced signal, but in general, it is still within the range of acceptance.

Keywords: Data compression, discrete cosine transforms, speech signal coding, speech reconstruction, DCT matrix, discrete cosine transforms.

1 Introduction

Since the beginning of the computer era, data growth continued until it reached unprecedented levels. Therefore, there is a need to think about efficient ways that will help store a huge amount of data using a limited amount of space.

The general idea of data compression encompasses two fundamental parts: lossy compression and loss-less compression [4, 8]. Lossless compression is the task when all original data can be recovered when the file is uncompressed; it is assumed that no part of the file could change during the compress/uncompress processing. Thus, the system must be assured that all of the information is completely restored, and this is done by selecting the parameter values that represent the compressed copy of the original file. Lossy compression, on the other hand, refers to the case when some information will be permanently eliminated [4]. In other words, when the file is uncompressed, only part of the original file will be there. It is generally assumed that almost all redundant information will be deleted.

One of the most important applications of the signal compression is video signal compression that relies totally on dealing with each image presented in the video stream [6, 22].

Speech signal compression, on the other hand, has many practical applications ranging from cellular technologies to digital voice storage. Compressing the signal could allow messages with a longer period to be

^{*}Corresponding author: Tariq A. Hassan, Computer Science Department, College of Education, University of Almustansirya, Baghdad, Iraq, e-mail: tariqmono2016@uomustansiriyah.edu.iq

Rageed Hussein AL-Hashemy: Computer Science Department, College of Education, University of Almustansirya, Baghdad, Iraq Rehab I. Ajel: Computer Science Department, College of Science, University of Almustansirya, Baghdad, Iraq

stored on a limited memory size. Typically, speech signals are sampled at rate of 8 K (or 16 K) samples/second. With lossy compressing techniques, the 8 K rate can be reduced to a level that is barely noticeably in speech.

The key point of this work is to use one compression method that, to some extent, reduces the amount of lost information presented in the speech signal. Therefore, the design of the proposed method will focus on the lossless compression and also taking into account that the compressed data are as small as possible to get good compression rate.

Finally, the compression rate is totally relayed on the quantization matrix. In this paper, different quantization matrices are adopted on many speech samples to examine the ability of the different quantization matrices with different compression rates to compress and retrieve the speech signal.

The objective of this paper is to test the ability of the JPEG method in speech compression and examine to what extent such method can compress the speech without distortion or losing some important information presented in the speech signal.

As the JPEG method is mainly designed to deal with two-dimensional (2D) signals, such as images, the proposed method will include some modifications on the speech data in two aspects. First, is to break-down the speech signal into 1D fixed length frames. Then, these frames are translated into 2D plane of speech data. Second, the speech data, itself, need to be converted into an acceptable form suitable for the JPEG compression method. While the JPEG algorithm only accepts positive input values, the speech data can be both positive and negative values. So, the speech signal needs to be converted into only positive values to match the JPEG method requirements. The absolute value cannot be used as it will distort the speech signal. Instead, adding a specific bias value in the range of 0-1 when the speech signal is normalized will lift up the speech signal to the positive level. This will make the speech data suitable for the JPEG algorithm and, at the same time, keep the speech data intact.

The main contribution in this work is to apply techniques that aim to:

- Translate the 1D speech signal into 2D image like signal that comparable with the JPEG algorithm shape.
- Modify speech data values in order to make them acceptable by the JPEG algorithm, especially in term of the negative values of the speech data.

The rest of the paper is organized as follows: the literature review is presented in Section 2. Section 3 explains in some detail the compression model used in this paper and presents the discrete cosine transform (DCT) equations for the 2D signals. Section 4 presents the main steps of the JEPG algorithm used for speech signal compression and explains the major modifications that should be applied on the speech signal before the compression processing step. Section 5 deals with some theoretical background such as quantization matrices of the different quantization levels, equations used for compression, and reconstruction of the signal. Section 6 presents the obtained results using the JPEG method in speech signal compression. Section 7 is the conclusion of this work.

2 Literature Review

Many methods with different amounts of performance have been used in the last few years. One of the most common signal compression techniques is the linear predictive coding (LPC) [10]. The LPC method, however, is commonly used in speech coding techniques, the act of converting the speech signal into a more compact form. The system presented in Ref. [10] suggests that using the LPC method, the speech compression would be highly lossy in a way that the output signal does not necessarily be exactly the input signal. The work in Ref. [18] suggested that using the LPC method with the quadrature phase shift keying (QPSK) modulation could overcome the LPC system in terms of speech signal compression.

A method in Ref. [13] presents a hybrid lossless compression technique that adopts the LPC, DWT, and Huffman used to compress the medical images. The combination of many techniques in the system has improved the system performance and maximized the compression ratio.

Another method that is usually used for speech compression is the wavelet technique. Wavelets are used to decompose the speech signal into different frequency bands based on different scales and positions of the wavelet function. These frequency bands represent the function coefficients that are used as parameters for signal compression. One big challenge for the wavelet compression method is the selection of an appropriate mother wavelet function. Choosing a suitable mother wavelet will play an important role in minimizing the difference between the output and input signal [20].

The technique presented in Ref. [17] suggested using the LPC and DCT as a combination of more than one technique in order to improve the speech compression ratio. The proposed technique outperforms the LPC or DCT alone.

The neural network is also used in image compression. A technique used in Ref. [7] combines the neural network, as a predictive system, with the discrete wavelet coding. The system suggests that the errors between the original and the predicted values can be eliminated using the neural network as a predictor system.

The work presented in Ref. [1] relies on the fast Hartley transform that used to encode the speech signal into the number of coefficients. These coefficients are quantized and coded into a binary stream, which represents the compressed form of the input signal. The results showed that the fast Hartley transform used in compressing the speech signal can exceed the performance of the wavelet transform [1].

The pulse code modulation (PCM) method has been used to compress the speech signal or, more precisely, to encode the analog signal into the digital stream [19]. In this method, the signal is sampled into the regular intervals, then, each sample is quantized to the closest number within the limit of the digital values. The PCM, however, can suffer from sampling impairments and quantization errors compared to the other methods [19].

The low bit rate vocoder-based system is adopted by Ref. [14] to compress the speech signal to the limit that allows to transmit the signal within low-bandwidth transmission lines (military line for example).

3 The Compression Model

General speech compression techniques are normally based on spectral analysis, using linear predictive coding or wavelet techniques [5, 16]. These techniques are usually used to estimate the best parameters that represent the original data when they are uncompressed. JPEG method, on the other hand, is a lossy compression scheme adopted for image compression techniques. The JPEG algorithm benefits from the fact that humans could not realize colors at high frequencies [11]. These high frequencies can be regarded as redundant data of the image that requires a large amount of storage space. Therefore, such frequencies can be eliminated during the compression task.

In speech signal compression, the system takes advantage of the fact that the human auditory system pays little attention to the high frequencies of the speech signal [12]. So, in order to compress a speech file using the lossy compression method (such as the JPEG scheme), the system will regard the information presented in the high frequencies as redundant data and can be eliminated without a huge effect on the speech signal. In other words, only the more important frequencies remain and are used to retrieve the source image in the decompression process [3].

Images in the DCT are divided into 2D square blocks (8*8 blocks) or larger. The DCT is applied for each block in order to calculate the DCT parameters (code parameterization), which represent the compressed values of the source data. Quantization and coded steps are utilized to compress the data to the desired compression level. On the other hand, the receiver should be able to decode the compressed data and retrieve the original signal. In this case, the inverse DCT process should be used and recollect blocks back together into a single unit of data (image or signal). In this process, however, many of the DCT parameters could have zero values. These can be ignored without crucial impact on the reconstructed image quality. The 2D DCT equation can be defined as follows:

$$\mathbf{D}(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \mathbf{sig}(n, m) \cos\left[\frac{(2n+1)i\pi}{2N}\right] \cos\left[\frac{(2m+1)j\pi}{2N}\right]$$

$$C(x) = \begin{bmatrix} \frac{1}{\sqrt{2}} & \text{if } x = 0\\ 1 & \text{if } x > 0 \end{bmatrix}$$
(1)

where, \mathbf{D} is the DCT output matrix of the speech signal elements \mathbf{sig} , at the pixel at the matrix location (\mathbf{n} , \mathbf{m}), \mathbf{N} corresponding to the dimensional length of the matrix (N=8 in the case of an 8*8 matrix).

The DCT presented in Eq. (1) produces a new set of elements that represent the DCT parameters of the input frame. Actually, in the proposed system, the DCT transform matrix has been used to calculate the DCT parameters because it is easy to deal with and much clear than the DCT formula [2, 23]. This matrix is derived from Eq. (1) and represented as follows:

$$\mathbf{Mx}(i,j) = \begin{cases} \frac{1}{\sqrt{N}} & \text{if } i = 0\\ \sqrt{\frac{2}{N}} \cos\left[\frac{(2j+1)i\pi}{2N}\right] & \text{if } i > 0 \end{cases}$$
 (2)

The matrix *Mx* is the DCT coefficient matrix with almost fixed values depending on the *N* variable. This matrix is an important part for the compression and decompression processes.

Using the DCT matrix, the system is ready to compute the DCT parameters of the source signal frame. In the experiment, and as the speech data values are normally ranged between –1 and 1, we noted that the best DCT parameters can be obtained by enlarging the speech data values (multiply them by a scale factor) to make them big enough for the DCT matrix.

The DCT parameters are computed using Eq. (3):

$$\mathbf{D}_{\mathsf{Coef}} = \mathbf{M} \mathbf{x} \cdot \mathbf{BLK} \cdot \mathbf{M} \mathbf{x}^{\mathsf{T}} \tag{3}$$

In this equation, D_{coef} is the DCT parameters of the input frame and represents (8*8) the DCT coefficients of one speech signal frame, and **BLK** is the 2D form of the input speech signal frame (signal values).

Basically, the upper left corner of the DCT parameter matrix (see Figure 3) aggregates with the low frequencies of the signal. Moving toward the lower right corner of the matrix, the system is heading to the high frequencies of the signal. For the speech signal, the prominent idea is that the low part of the signal spectrum (low frequency) can carry much important information than those in the high frequency [15]. Therefore, the trend is to eliminate the values within the high-frequency area and retain only those in the low-frequency part of the matrix.

4 The Speech Signal JPEG Compression Processing

The JPEG compression technique employs the discrete cosine transform in its process. The original image is usually divided into 8*8 blocks, and the DCT is applied on each of the partitioned images. Compression is then achieved by performing quantization processing on the output of the DCT process. When the JPEG scheme is used to compress a 1D signal (speech), many priori steps are needed to be considered in order to make the speech signal more compatible with the compression method. One of the very important steps is to convert the 1D signal into the 2D matrix. The problem is that the speech signal is variable depending on the duration of the word being said. Hence, there is no guarantee that it is a fixed-size image such as a 256*256 pixel. One solution is to divide the speech signal into fixed-length frames, and each frame is then converted independently into the 2D array of speech samples.

- In general, the main steps of the compression model can be summarized as follows:
- Step 1: The speech signal is divided into the fixed-length frames (64 samples in the conducted experiment, which facilitates converting it into the 8*8 matrix. The samples in the 8*8 data array represent the time-domain signal values.
- Step 2: For each frame of the signal, the DCT is applied in order to transform the speech signal into the frequency domain and establish a compressed data by concentrating most of the speech signal in the lower spectral frequencies.
- Step 3: The DCT parameters for each frame (matrix) are then uniformly quantized using the 8*8 element quantization table (QT). The same QT table will later be used (at uncompressing step) to recover the original speech data.
- Step 4: The original speech data are recovered from the quantized parameters (the compressed data) using the inverse discrete cosine transform (IDCT).

Figure 1 shows the main steps of the proposed JPEG-based speech compression model along with its decompression steps.

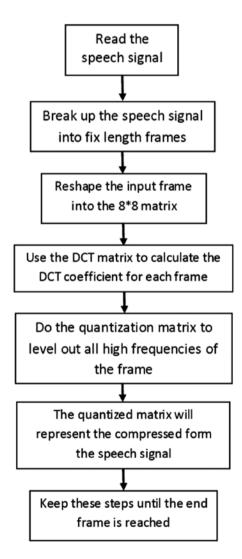


Figure 1: The Main Steps of the Out-Proposed Compression System.

5 Quantization and Coding Process

Once the DCT parameters are constructed, the speech data will be ready to be compressed by the quantization processing. Quantization is simply achieved by multiplying one of quantization matrices by the DCT parameters matrix.

One of the best characteristics of the JPEG techniques is that there are many options for the quantization matrices depending on the quality of the compression level and the amount of available space. The JPEG technique provides quantization quality levels ranging from 1 to 100 in which the highest compression level is in level 1, while the poorest compression level is in level 100. This works oppositely with the quality of the retrieving data, where the poorest quality is at level 1, and the highest is at level 100 [9]. Regardless of the level that will be selected, the required quantization matrix should maintain the signal from being distorted and, at the same time, not to exceed the available memory space. In this paper, two quantization levels have been examined to see the effects of the different compression levels on the speech signal quality. The JPEG standard quantization matrix has level qualities Q_{50} and Q_{90} . The quality matrix Q_{50} can provide a fine quality and good compression level [21]. The Q_{90} quality matrices provide an excellent compression ratio with some level of distortion in signal quality. The JPEG standard quantization matrix with a quality levels Q_{50} and Q_{90} are shown in Figure 2.

The quantization parameters are obtained using the following equation:

$$\mathbf{Comp} = \mathbf{round} \frac{\mathbf{D}_{\mathbf{coef}}}{\mathbf{Q}}$$

where, D_{coef} is the DCT parameter matrix, Q is the quality matrix, and Comp represents the compressed form of the data input matrix (speech frame). This equation divides each element in the DCT matrix by its counterpart element in the quality matrix. Then the result is rounded to the nearest integer. So, the output will contain only integer values grouped in the upper left corner of the output matrix. An example of an output (compressed frame block) is depicted in Figure 3.

The final step of the compression process is the coding stage. In this stage, the output matrix of the quantization step is converted into a binary data stream. The JPEG technique encodes the quantized elements by arranging them into a zigzag sequence. Arranging the quantized elements will facilitate the encoding by putting the non-zero values first before the zero values. Encoding will not be implemented in this work as the objective is to match the original speech signals with the retrieved ones after doing the compression/ decompression process.

Once the signal is compressed and coded, it will be ready for transmitting or storing. However, when receiving such data (or in the case of retrieving), the system needs to reconstruct the source signal. This is done by the decompression process. Decompression is the activity of restoring the source data from the

$$Q_{50} = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix} \qquad Q_{90} = \begin{pmatrix} 3 & 2 & 2 & 3 & 5 & 8 & 10 & 12 \\ 2 & 2 & 3 & 4 & 5 & 12 & 12 & 11 \\ 3 & 3 & 3 & 5 & 8 & 11 & 14 & 11 \\ 3 & 3 & 3 & 4 & 6 & 10 & 17 & 16 & 12 \\ 4 & 4 & 7 & 11 & 14 & 22 & 21 & 15 \\ 5 & 7 & 11 & 13 & 16 & 12 & 23 & 18 \\ 10 & 13 & 16 & 17 & 21 & 24 & 24 & 21 \\ 14 & 18 & 19 & 20 & 22 & 20 & 20 & 20 \end{pmatrix}$$

Figure 2: JPEG Standard Quantization Matrices of Quality Levels Q_{50} and Q_{90} .

							_
-7		0	-1	0	-1	0	0
4	3	1	0	0	0	0	0
13	1	-1	1	0	0	0	0
-1	-1	0	0	0	0	0	0
	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	ر ہ

Figure 3: An Example of the Compressed Form of One Speech Signal Frame.

compressed counterpart. This process could change some values of the original data depending on the compression rate used in the system. As mentioned earlier, the higher the compression rate, the higher is the loss in the image/speech signal quality resulting in many values of the signal undergoing a dramatic change. Regardless of the compression rate that was used, the following equation is used to reconstruct of the spectrum of the original speech signal:

$$\mathbf{R} = \mathbf{Q} * \mathbf{Comp} \tag{5}$$

where, **R** is the spectrum of the reconstructed signal out of the quantization matrix Q, and **Comp** is the compressed signal. Inverse DCT (or inverse FFT) will be applied to Eq. (5) in order to obtain the time domain of the reconstructed signal. Equation (6) is used to generate the speech signal (recovery) from the reconstructed signal:

$$\hat{\mathbf{sig}} = \operatorname{round}(\mathbf{M}\mathbf{x}^{\mathsf{T}} * \mathbf{R} * \mathbf{M}\mathbf{x}) \tag{6}$$

In this equation, **ŝig** is the recovery speech signal (the decompressed signal), and **Mx** is the DCT matrix coefficients. According to Eq. (6), the major parameter of the signal quality in the recovery stage is the R matrix, which, in some way, depends on the Q or the quantization matrix. In other words, if there is proper selection of the quantization matrix, fine recovering values will be obtained.

6 Experimental Results and Discussion

Two strategies were experimented to test the proposed method. The first one is to modify the JPEG parameters. This includes the parameters of Eq. (1) and converting the quantization matrix into 1D vector. Modifying the parameters in Eq. (1) is relatively easy, but the hardest part in this case is to find a way to convert the 2D quantization matrix into the 1D quantization vector. Moreover, choosing the right quantization parameter values is very hard as there is no quantization vector generated that depends on the nature of the speech signal.

The second strategy, which is adopted in this work, is to convert the speech signal from a 1D vector into an image-like 2D matrix. This includes the essential preparations that are required to make the speech data values more appropriate to work with the JPEG algorithm, especially in terms of negative values. In order to examine the proposed method, the speech signal is divided into a fixed length of 5-ms duration that represents 64 samples. The 64-sample vector is converted into the 8*8 data matrix. This process is applied for the whole speech signal. Some of these frames have no (or little) information about the speech, like silence or just noise. So, one important role in the signal preparation is to remove the silence frames (with low energy) from the speech, as they have little effect on the speech signal.

The selected frames are collected in one matrix to generate one image-like speech matrix. The compression process will apply to the matrix of the signal so that a set of parameters will be generated. These will represent the key parameters in the decompression process.

Figures 4-7 show examples of the original and decompressed signal frames with different compression rates.

The behavior of the compressed and decompressed signal reveals two crucial views about the speech signal compression techniques.

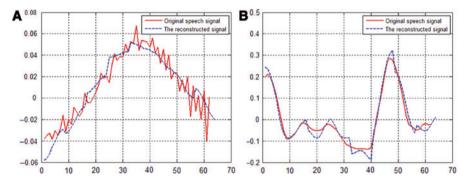


Figure 4: The Compression/Decompression Process of Quality Matrix Q_{10} for Two Different Energy Level Frames. (A) Frame energy = 0.05, (B) frame energy = 0.85.

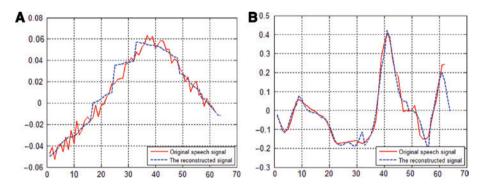


Figure 5: The Compression/Decompression Process of Quality Matrix Q_{30} for Two Speech Frame with Energies. (A) Frame energy = 0.08, (B) frame energy = 1.32.

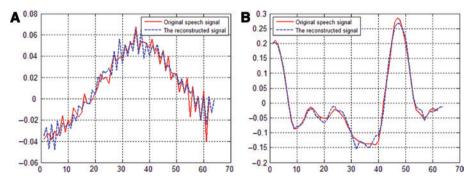


Figure 6: The Compression/Decompression Process of Quality Matrix Q_{50} for Two Different Energy Level Frames. (A) Frame energy = 0.09, (B) Frame energy = 1.34.

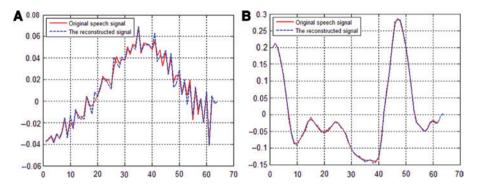


Figure 7: The Compression/Decompression Process of Quality Matrix Q_{90} for Two Different Energy Level Frames. (A) Frame energy = 0.09, (B) Frame energy = 1.34.

First, the reconstructed signal is highly affected by the energy of the compressing frame. This, to some extent, can be justified by the fact that a tiny change in low-energy signals could cause a noticeable change in the signal. Figure 4 shows two different signal frames (of the same speech signal) with different energies, both compressed using the Q_{10} quality matrix. The differences between the original frame signal and the reconstructed one are clear to notice. Figure 4A shows that the reconstructed signal is highly distorted. This is because the energy of the frame signal is low (quite silent speech frame). Figure 4B shows that the reconstructed signal is quite identical to the original one. Similar cases are depicted in Figures 5–7. This will prove that the signal energy has a huge effect on the compressing quality regardless of the method or parameters that are used for signal compression.

Second, the proper selection of the quantization matrices can minimize the differences between the two signals. This is clearly the case in Figures 4–7. The reconstructed signal is the best using Q_{90} (Figure 7), and it is less with other quality matrices (Figures 4 and 5).

Therefore, in order to get the best match between the reconstructed signal and the original one, the quality matrix as well as the signal energy should be focused on. Both play an important role in reconstructing the compressed signal in a way that some information will not be lost during the compression/decompression processing.

The error (distortion) of the low-energy frames could happen because these frames can hold little information about the word being said; in other words, noise-like signals can suffer more distortion than the real informative speech signals.

The type of quality matrix should be selected carefully for compression/decompression processing. The high-quality matrix has a good (less distorted) reconstructed signal but, at the same time, less compression rate in terms of file size. The low-quality matrix could cause more distortion (or even wipe off) in the reconstructed signal. So, the chosen quality matrix should compromise between the quality and size.

Comparing the methods presented in the literature review, some important points about the proposed model can be seen. First, the model suggests that the perceptual quality of the images can be used in speech signal processing (compression in this model). This proves that the quantization matrices suggested for the image compression technique can be adopted in the speech signal processing in terms of compression or encoding. Second, in terms of the accuracy, the proposed method gives a low accurate similarity between the compressed and reconstructed signals, especially with low-level quantization matrices and low-energy signal (noise-like). This is the case with the LPC technique [10] and the PCM technique [19]. In order to improve the result quality, some systems suggest using a combination of many techniques on one model [7, 13, 17]. This, however, will add some complexity to the system and can increase the required time for the system. The proposed method can overcome this problem by selecting a good quantization matrix quality and/or increase the signal energy. No more parameters are needed to improve the system performance.

Third, in the case of the signal being buried under an outside noise, the compression process needs to be preceded by a filtering step (de-noising). This subject is out the scope in this paper. However, many filter types can be adopted for this purpose but a filter that is highly accurate in denoising will be preferred.

7 Conclusion

This paper has introduced a new compression strategy that explores the potential characteristics of the JPEG method to compress the speech signal. The comparison results have demonstrated the system robustness in reconstructing the speech signal with little change especially in the case of the low-energy parts of the signal. Although the system is highly accurate when using the quality level matrix Q_{00} , the problem with the lowenergy frame makes it a bit far from ideal reconstructing. So, a new set of quality matrices or new strategy is needed to solve the low-energy part of the speech signal.

The main contribution of this research is modifying a 1D signal (speech) in a way that makes it appropriate with a 2D compression algorithm like JPEG. The modification includes two stages: first, the speech signal is brroken down into fixed-length frames and arranging the accepted ones (depending on their energies) in 2D form. Second, the speech signal data usually involve both the positive and negative values. This is absolutely not accepted by the JPEG method, so the system does some steps to overcome this problem by increasing the base value of the speech signal data that guarantees that all speech data are converted into the positive values. The increment parameter will vary depending on the speech sample on hand.

In general, the proper compression rate will highly rely on two major factors: the first, is the energy of the speech signal; the higher the signal energy is, the best results will be achieved. Second, the higher the quality of the matrix applied on the signal is, the less is the produced distorted signal.

Bibliography

- [1] N. Aloui, S. Bousselmi and A. Cherif, DSP real-time implementation of an audio compression algorithm by using the fast Hartley transform, Int. J. Adv. Comput. Sci. Appl. 8 (2017), 472-477.
- [2] S. Bouguezel, M. O. Ahmad and M. N. S. Swamy, A fast 8*8 transform for image compression, in: 2009 International Conference on Microelectronics - ICM, Columbus, OH, USA, pp. 74-77, 2009.
- [3] K. Cabeen and P. Gent, Image compression and the discrete cosine transform, Math 45, College of the Redwoods, 1998.
- [4] J. S. Chitode, Information coding techniques, Technical Publications, Maharashtra, India, 2007.
- [5] E. B. Fgee, W. J. Phillips and W. Robertson, Comparing audio compression using wavelets with other audio compression schemes, in: Engineering Solutions for the Next Millennium. 1999 IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No.99TH8411), vol. 2, Edmonton, Alberta, Canada, pp. 698-701, 1999.
- [6] R. F. Haines and S. L. Chuang, The effects of video compression on acceptability of images for monitoring life sciences experiments, NASA technical paper. National Aeronautics and Space Administration, Office of Management, Scientific and Technical Information Program, 1992.
- [7] A. J. Hussain, D. Al-Jumeily, N. Radi and P. Lisboa, Hybrid neural network predictive-wavelet image compression system, Neurocomputing 151 (2015), 975-984.
- [8] P. B. Khobragade and S. S. Thakare, Image compression techniques a review, IJCSIT 5 (2014), 272–275.
- [9] J. D. Kornblum, Using jpeg quantization tables to identify imagery processed by software, Digit. Investig. 5 (2008),
- [10] A. R. Madane, Z. Shah, R. Shah and S. Thakur, Speech compression using linear predictive coding, in: Proceeding of the International Workshop on Machine Intelligence Research, MIR Day GHRCE- Nagpur, pp. 119-121, 2009.
- [11] M. Marcuss, Jpeg Image compression, 2014. http://www.cs.darmouth.edu/mgm/finalreport.pdf.
- [12] J. C. Middlebrooks and D. M. Green, Sound localization by human listeners, Annu. Rev. Psychol. 42 (1991), 135-159.
- [13] A. Mofreh, T. M. Barakat and A. M. Refaat, A new lossless medical image compression technique using hybrid prediction model. SPIJ 10.3 (2016), 20.
- [14] P. Rajeswari and L. Krishna, Implementation of low bit rate vocoder for speech compression, Int. Res. J. Eng. Technol. 4 (2017), 965-969.
- [15] H. Reetz and A. Jongman, Phonetics: transcription, production, acoustics, and perception, Blackwell Textbooks in Linguistics, John Wiley & Sons, 2011.
- [16] T. Sen and K. J. Pancholi, Speech compression using voice excited linear predictive coding, Int. J. Emerg. Technol. Adv. Eng. 2 (2012), 306-209.
- [17] J. A. Sheikh and S. Akhtar, Realization and performance evaluation of new hybrid speech compression technique, Int. J. IJMECE 4 (2016) 2321-2152.
- [18] J. A. Sheikh, S. Akhtar, S. A. P. Post and G. M. Bhat, On the design and performance evaluation of compressed speech transmission over wireless channel, in: 2015 Annual IEEE India Conference (INDICON), New Delhi, India, pp. 1-5, 2015.

- [19] J. Singh, Speech compression techniques: an overview, Int. J. Recent Innov. Trends Comput. Commun. 5 (2017), 637-640.
- [20] T. Siva Nagu, K. Jyothi and V. Sailaja, Speech compression for better audibility using wavelet transformation with adaptive Kalman filtering, Int. J. Comput. Appl. 53 (2012), 1-5.
- [21] M. Tuba and N. Bacanin, Jpeg quantization tables selection by the firefly algorithm, in: 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, Morocco, pp. 153–158, 2014.
- [22] G. K. Wallace, The jpeg still picture compression standard, IEEE Trans. Consum. Electron. 38 (1992), xviii-xxxiv.
- [23] A. B. Watson, Perceptual optimization of DCT color quantization matrices, in: Proceedings of 1st International Conference on Image Processing, vol. 1, Austin, TX, USA, pp. 100-104, 1994.