Image processing using DCT and wavelet transform

SITI AISYAH*, J. S. MANDEEP
Department of Electrical, Electronic & Systems Engineering ,Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia,43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

The DCT works by separating images into parts of differing frequencies. During a step called quantization, where part of compression actually occurs, the less important frequencies are discarded. Only the most important frequencies that remain are used to retrieve the image in the decompression process. It is similar to the discrete Fourier transform where it transforms a signal or image from spatial domain to frequency domain. Wavelet transform on the other hand is a multi-resolution transform that allows a form of time-frequency analysis. It provides a progressive encoding of the image at various scales, which is more flexible. The wavelets comprise a normalized set of orthogonal functions on which the image is projected. The aim of this project is to compare the performance of the DCT and the wavelet transform in image processing. Most images contain some amount of redundancy that can sometimes be removed when the image is stored and replaced when it is reconstructed, but eliminating this redundancy does not lead to high compression. Fortunately, the human eye is not very sensitive to a wide variety of information loss. An image can be changed in many ways that are either not detectable by the human eye or do not contribute to degradation of the image. Compression of an image allows the number of bits to be reduced to represent the coded image which contains a number that is smaller than the original format. This number is variable and it depends on how the image is compressed. Different types of wavelets and different stages of DCT compression ratio have been used to perform the transform of a test image. The results were analyzed with the amount of errors introduced during the compression process.

(Received December 15, 2011; accepted February 20, 2012)

Keywords: DCT, JPEG2000, Wavelet, Image processing

1. Introduction

With the advent of the multimedia era and the growth of digital packet networks, the total amount of image data accessed and exchange by users daily has reached the huge of several petabytes. Therefore, the compression of continuous still image, either greyscale or colour, has grown tremendously in importance. Image processing is used for two different purposes which are to improve the visual appearance of image to human viewer and to prepare the images for measurement of the features and structures present [1]. The main concerned in image compression is to minimize the number of bits required to represent an image with no significant loss of information. Moreover, the rapid growth of various images applications consequently making the storage requirements of digital imagery growing explosively. It is very crucial when it comes to represent an image that contain large amounts of information that requires much storage space, large transmission bandwidths and long transmission times[2]. Therefore, it would be an advantage if the compression process stored only the essential information needed to reconstruct the image. An image can be represented as a matrix of pixel values. In order to compress the image, redundancies must be exploited. Imagine areas where there is little or no change between pixel values. It will have large redundancies [3]. However, if the images have frequent large changes in colour, it will have less redundant and harder to compress.

For still image compression, the Joint Photographic Experts Group (JPEG) standard has been established by International Standards Organization (ISO) and International Electro-Technical Commission (IEC). Generally, the performance of these coders will degrade at low bit-rates mainly because of the underlying block-based Discrete Cosine Transform (DCT) scheme. Meanwhile, the wavelet transform has emerged as a cutting edge technology, within the field of image compression. It provides substantial improvements in picture quality at higher compression ratios. JPEG use DCT as a technique to compress the image. The DCT can perform faster and therefore it is widely used. It can easily compress the data that is highly correlated. DCT also have a fixed basis images that can gives good compromise between information packing ability and computational complexity [4]. The Inverse Discrete Cosine Transform (IDCT) can be used to retrieve the image from its transform representation. In contrast, wavelet transform can be used to reduce the size of the image without losing much of the resolutions computed and values less than a pre-specified threshold are discarded. Thus it reduces the amount of memory required to represent the given image. Because of their inherent multi-resolution nature, wavelet-coding schemes are especially suitable for applications where scalability and tolerable degradation are important. Recently the JPEG committee has released its new image coding standard, JPEG-2000, which has been based upon wavelet transform.

Although DCT can perform much faster than wavelet transform, wavelet transform has recently emerged as the tool of choice for image compression because of their inherent multi-resolution nature. Wavelet-coding schemes are especially suitable for applications where scalability and tolerable degradation are important. For this project, evaluation must be made between these two compression methods and the result is then compared [5-7].

2. Wavelet description

There are many wavelets being used now a day for decomposition of signals and images. According to Akthar [2], the main types are Daubechies, Biorthogonal, Coiflets, Symlets and Dmey. The major characteristics of different types of wavelets families are briefly described here.

2.1 Daubechies

The Daubechies wavelets are compactly supported and have highest number of vanishing moments. The types can be defined as db1, db2 until db45. It can perform with Discrete Wavelet Transform and continuous wavelet transform. The Daubechies are not symmetrical. The length of the filter is 2N while the number of vanishing moments is N. Each step of the wavelet transform applies the scaling function to the data input. If the original data set has N values, the scaling function will be applied in the wavelet transform step to calculate $N/2$ smoothed values. In the ordered wavelet transform the smoothed values are stored in the lower half of the N element input vector. The Daubechies filter enjoys wide popularity and is used by a number of researchers. It has maximum regularity and vanishing moments for its size, but is relatively long and irrational [8].

2.2 Biorthogonal

The Biorthogonal wavelets are compactly supported wavelets for which symmetry and exact reconstruction is possible with Finite Impulse Response (FIR) filters. The types are bior1.1, bior1.3, bior1.5, bior2.2, bior2 and many more. Discrete and continuous wavelet transforms are possible with these. Designing Biorthogonal wavelets allows more degrees of freedom. One additional degree of freedom is the possibility to construct symmetric wavelet functions. Thus, Biorthogonal wavelets have been the de facto standard for image compression applications. The viability of symmetric extension with Biorthogonal wavelets is the primary reason cited for their superior performance. Furthermore, the Biorthogonal wavelets indicate a slight performance advantage for low frequency images [9]

2.3 Symlets

The Symlets wavelets are compactly supported wavelets with highest number of vanishing moments. The types are sym1, sym2, sym3, sym4, and sym5. Discrete and continuous wavelet transform is possible with these types of wavelets. Filter length is 2N. Associated scaling filters are near linear-phase filters. The Symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the db family. Therefore, the properties of the two wavelet families are similar.

2.4 Coiflets

Coiflets wavelets are compactly supported wavelets with highest number of vanishing moments for both psi and phi for a given support width. These wavelets must be generated numerically from their filter coefficients since there is no closed formula for them. Coiflets appear more symmetric than Daubechies orthonormal scaling and wavelets have compact support, but are not orthonormal. The wavelet is near symmetric, their wavelet functions have $N/3$ vanishing moments and scaling funcstions $N/(3-1)$, and have been used in many applications. Discrete and continuous wavelet transform is possible with these.

2.5 Discrete Meyer

Discrete Meyer wavelet is FIR based approx of the Meyer wavelet. Meyer wavelet is infinitely regular orthogonal wavelet. The Meyer wavelet is useful because it is analytic, symmetric, and its Fourier transform is completely supported. It can also be used to observe the correspondence between the scaling behaviours. Although wavelet does not decay exponentially fast, it decays faster than any polynomial, which is sufficient for many purposes.

The Discrete Cosine Transform (DCT) is the heart of JPEG compression. The input to the DCT is a set of numeric values and the output is a set of the same size. The DCT is an invertible transform, which means that its output coefficients can be used to recreate the original input values. The reverse of the DCT is called the Inverse Discrete Cosine Transform (IDCT). The DCT is often referred to as the Forward DCT (FDCT). The DCT transforms the set of input values into a set of coefficients to cosine functions with increasing frequencies. Each of the original values is transformed into the sum of cosines. In this the DCT is closely related to the Fourier transform. The DCT is commonly used to process data organized in either one or two dimensions. The number of input values is usually a power of two. In JPEG, the DCT and IDCT are always in two dimensions on data arranged in 8×8 blocks.

JPEG is primarily a lossy method of compression.JPEG was designed specifically to discard information that the human eye cannot easily see. DCT separates images into parts of different frequencies where less important frequencies are discarded through quantization and important frequencies are used to retrieve the image during decompression. Some of DCT advantages compared to other input dependent transform is stated below.

It has been implemented in single integrated circuit.

It has the ability to pack most information in fewest coefficients.

It minimizes the block like appearance called blocking artifact that results when boundaries between sub-images become visible.

The JPEG standard was formed in the 1980's and has been an effective first solution to the standardization of image compression. JPEG is always known to provide a very useful strategy for DCT quantization and compression but it was only developed for low compressions. The 8 × 8 DCT block size was chosen for speed and not for performance.

3. Basic process

JPEG stands for Joint Photographic Experts Group which is a commonly used method of compression for photographic images. The degree of compression can be adjusted, allowing a selectable tradeoff between storage size and image quality. The following is a general overview of the JPEG process [10].

a. Original image is divided into blocks of 8 × 8.

b. Pixel values of a black and white image range from 0 to 255 but DCT is designed to work on pixel values ranging from -128 to 127. Therefore each block is modified to work in the range.

c. DCT is applied to each block by multiplying the modified block with DCT matrix on the left and transpose of DCT matrix on its right.

d. Each block is then compressed through quantization.

e. Quantized matrix is then entropy encoded.

f. Compressed image is reconstructed through reverse process.

g. Inverse DCT is used for decompression.

A block diagram of JPEG sequential coding is shown in Fig. 1. The image is divided in 8×8 blocks. For each block, the DCT is computed, its coefficients are quantized, and then entropy-coded. The quantization and entropy coding tables are written in the compressed file along with the entropy coded data. JPEG follows the common transform coding paradigm, according to which the image undergoes a linear invertible transform and then the transform coefficients are quantized and entropy-coded. The first operation performed on each component is level shifting. The unsigned integer sample values are subtracted $2^{(N-1)}$, where N is the number of bits per sample on which the samples are represented. Then, they are input to the DCT. In JPEG, each block of 8×8 samples is independently transformed using the two-dimensional DCT. This is a compromise between contrasting requirements which is larger blocks would provide higher coding efficiency, whereas smaller blocks limit complexity.

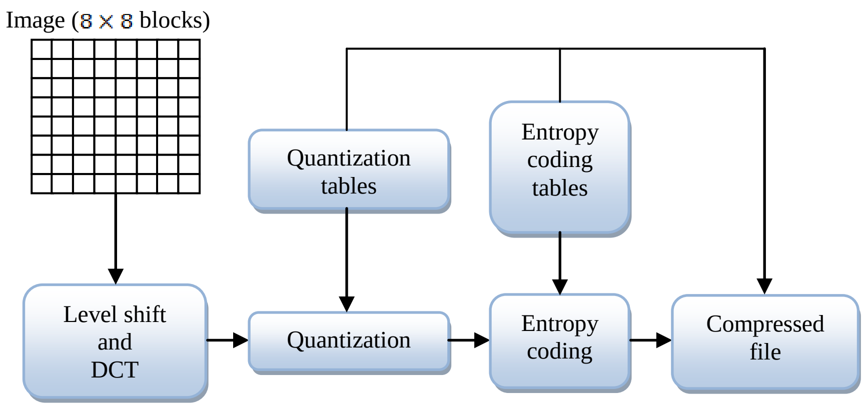


Fig. 1. Block diagram of JPEG sequential coding.

In wavelet decomposition, the image is decomposed into a set of subimages which also represent details at different scales. The subimages also represent details with different spatial orientations such as edges with horizontal, vertical, and diagonal orientations. Assume that the images to be analyzed are rectangular with N×M pixels. The number of pixels in wavelet decomposition is only NM.

Fig. 2 depicts an analysis filter bank, with one input $x(n)$ and two outputs $x_0(n)$ and $x_1(n)$. The input signal $x(n)$ is processed through two paths. In the upper path, $x(n)$ is passed through a lowpass filter $H_0(e^{jw})$ and decimated by a factor of two. In the lower path, $x(n)$ is passed through a highpass filter $H_1(e^{jw})$ and also decimated by a factor of two. For convenience, the following assumption is made.

a. The number N of available samples of $x(n)$ is even.

b. The filters perform a circular convolution which is equivalent to assuming that $x(n)$ is a periodic signal.

Under these assumptions, the output of each path is periodic with period equal to $N/2$ samples. Hence the analysis filter bank can be thought of as a transform that maps the original set $\{x(n)\}$ of N samples into a new set $\{x_0(n), x_1(n)\}$ of N samples.

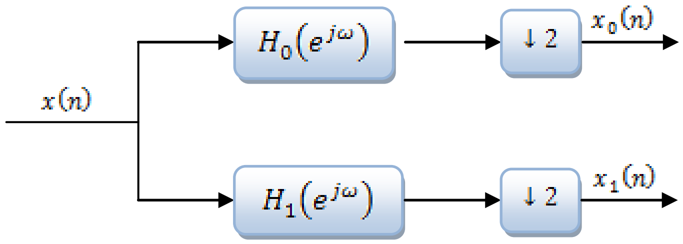


Fig. 2. Analysis filter bank with lowpass filter $\boldsymbol{H_0}(e^{j\omega})$ *and highpass filter* $\boldsymbol{H_1}(e^{j\omega})$.

Fig. 3 shows a synthesis filter bank. Here there are two inputs $y_0(n)$ and $y_1(n)$, and one single output $y(n)$. The input signal $y_0(n)$ is upsampled by a factor of two and filtered using a lowpass filter $G_0(e^{j\omega})$. The output $y(n)$ is obtained by summing the two filtered signals. Assume that the input signals $y_0(n)$ and $y_1(n)$ are periodic with period $N/2$. This implies that the output $y(n)$ is periodic with period equal to N. So the synthesis filter bank can also be thought of as a transform that maps the original set of N samples into a new set $\{y_0(n), y_1(n)\}$ of N samples $\{y(n)\}$.

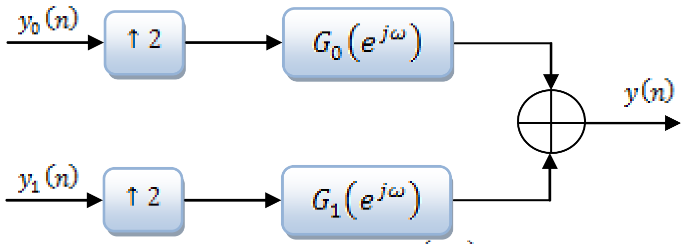


Fig. 3. Analysis filter bank with lowpass filter $\boldsymbol{G_0}(e^{j\omega})$ *and highpass filter* $\boldsymbol{G_1}(e^{j\omega})$.

When the output $x_0(n)$ and $x_1(n)$ of an analysis filter bank is applied to the input of a synthesis filter bank, under some specific conditions on the four filters $H_0(e^{j\omega})$, $H_1(e^{j\omega})$, $G_0(e^{j\omega})$ and $G_1(e^{j\omega})$, the output $y(n)$ of the resulting analysis and synthesis system is identical to its input $x(n)$. This condition is known as perfect reconstruction.

Fig. 4 shows the sequence of operations that generate the compressed image. After the optional multicomponent transformation, each component is tiled, wavelet transformed, quantized, and entropy-coded. The rate allocator decides which compressed data portions will be inserted in the codestream. Each image component is considered separately and is first divided into rectangular tiles. If the samples are unsigned integers, level shifting is performed as in JPEG, by subtracting 2^{N-1} from each image sample, where N is the bit depth.

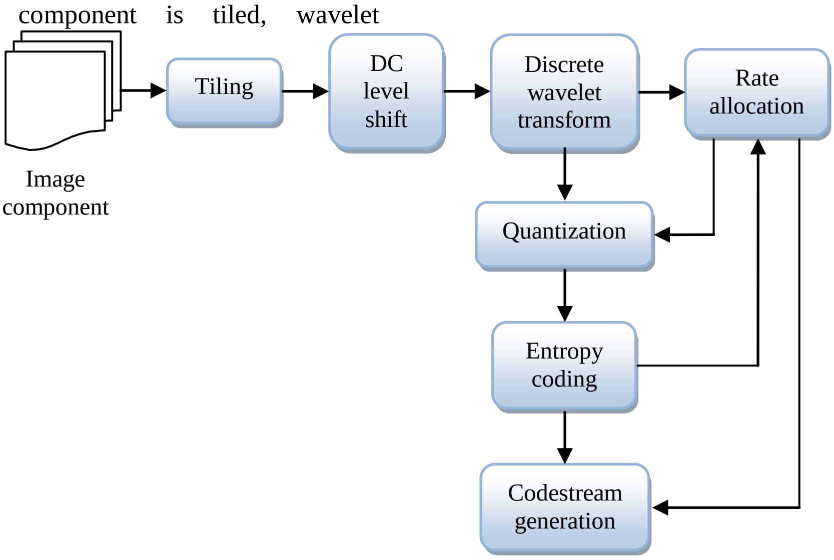
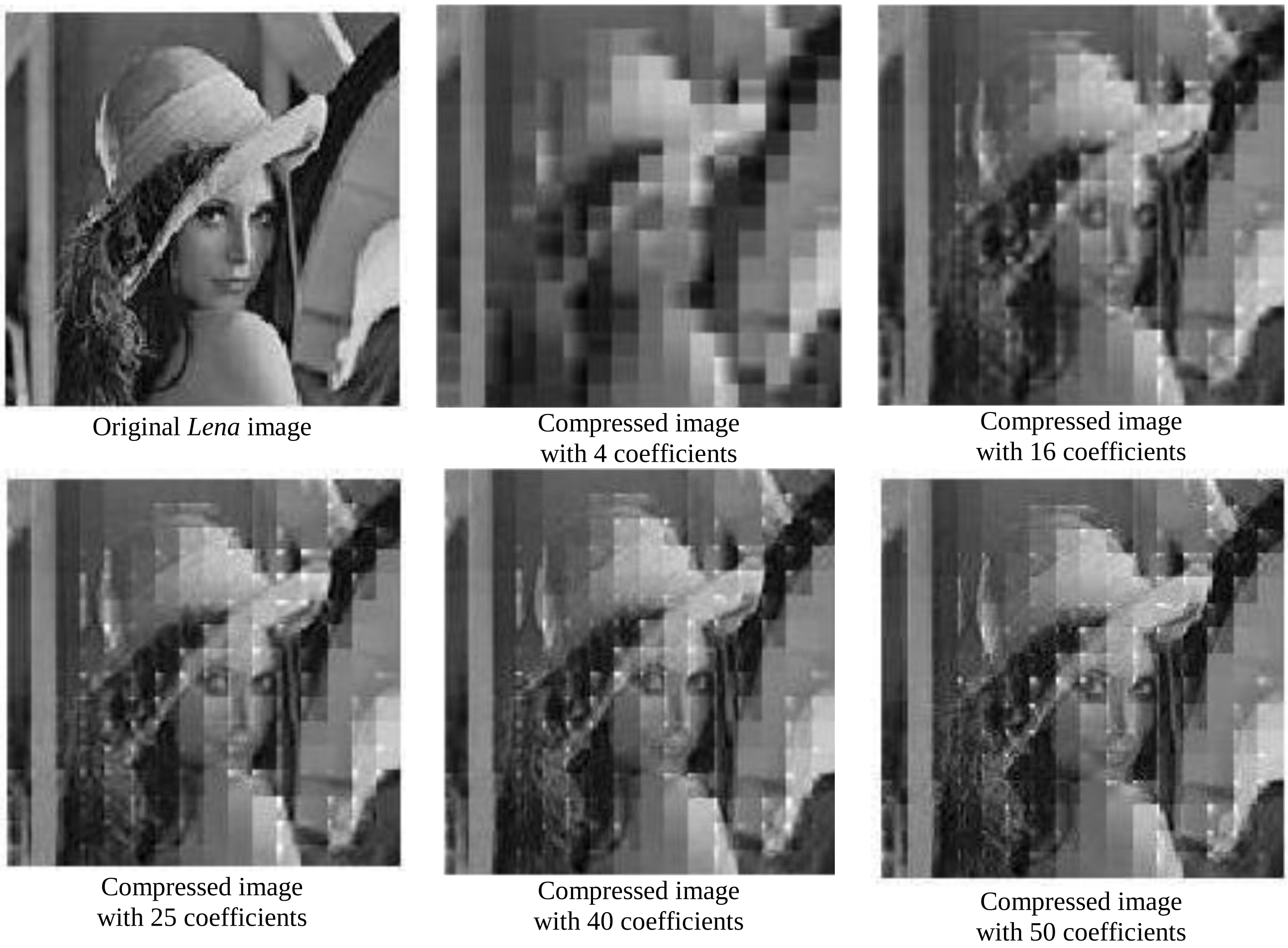


Fig. 4. Basic block diagram of JPEG 2000.

4. Results and discussion

Fig. 5 represented the compressed image for the original Lena image after taking various numbers of coefficients for quantization. As the number of coefficients increases, the quality of the image decreases whereas compression ratio continues to increase. The array of compressed blocks that constitute the image is stored in a drastically reduces amount of space. These imaged is then reconstructed through decompression as in Fig. 6 by using IDCT.

The function of image compression is to minimize the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space. It also reduces the time required for images to be sent over the Internet or downloaded from web pages. Higher compression factor will produce a small image file size but this will affect the picture quality. Generally, the higher the compression ratio, the poorer the quality of the resulting image. Basically, it is well known that JPEG performs rather poorly at high compression ratios. It is important to consider the relationship between compression ratio and picture quality when compressing an image.



Original *Lena* image

Compressed image
with 4 coefficients

Compressed image
with 16 coefficients

Compressed image
with 25 coefficients

Compressed image
with 40 coefficients

Compressed image
with 50 coefficients

Fig. 5. Lena image compressed using various numbers of coefficients.

Fig. 6. Lena image restored using various numbers of coefficients.

Seven wavelets family and JPEG 2000 standard have been used to obtain the results. Lena image of size 256×256 is used as a subject for this compression. Second level decomposition approximation is chosen for simple compression. The Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) are the two error metrics used to compare image compression quality. The MSE represents the cumulative squared error between the compressed and the original image, whereas PSNR represents a measure of the peak error. A lower value for MSE means lesser error. The relationship between MSE and PSNR is inversely proportional and this translates to a high value of PSNR if there is a lesser error in that image. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher. The term signal here is the original image and the noise is the error in reconstruction. The PSNR block computes the peak signal-to-noise ratio, in decibels, between two images. This ratio is often used as a quality measurement between the original and a compressed image. The higher the PSNR, the better the quality of the compressed or reconstructed image.

From Table 1, it can be seen that Discrete Biorthogonal CDF 9/7 from JPEG 2000 standard has the higher PSNR meaning that this type of compression is better than other type of wavelet transform. The compression ratio which is the size of the compressed file compared to the uncompressed file is also higher than others. A nice feature of JPEG 2000 is at each layer, it can store parts of the code block. To get a better image quality, it must send more layers.

Table 1. Summarize compression for wavelet transform.

Type	File size (Kb)		Compression Ratio	MSE	PSNR (Db)	Max different
	Before	After				
Haar		63.49	71.5408	55.5328	30.6853	39
Daubechies		64.09	70.3936	22.1222	17.7114	219
Coiflets		64.24	67.8482	37.9597	15.3664	216
Symlets		63.71	70.0711	42.6998	14.8553	212
Discrete Meyer	66.4	64.77	48.6822	650.6752	19.9972	207
Biorthogonal		64.00	70.9814	71.6523	29.5785	48
Reverse Biorthogonal		63.01	70.0769	122.1660	27.2613	111
CDF 9/7	35.73	29.84	83.5120	30.5003	33.2878	46
CDF *5/3*		28.89	80.8531	34.2364	32.7859	50

6. Conclusion

DCT and wavelet transform is a type of powerful tools and extremely useful when it comes to image processing. Both method has its own specialty but it depend on the application and usually involves adjustment between several factors including hardware or software, the allowable coding delay and the required compression level when it comes to choose which method suit the best.

References

[1] Acharya, T. & Ray, A.K. 2005. Image Processing: Principles and Applications. First Edition. Canada: John Wiley and Sons, Inc.
[2] J. Akhtar, M. Y. Javed, IEEE-ICET International Conference on Emerging Technologies **11**, 133 (2006).
[3] M. Antonini, M. Barlaud, P. Mathieu, IEEE Transactions on Image Processing **1**(2), 285 (1997).
[4] S. Ayub, S. M. Beg, T. T. Fook, JPEG 2000 and wavelet compression. IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering **10**(1), 590 (2002).
[5] M. Barni, Document and Image Compression (Signal Processing and Communications). First Edition. US: CRC Press 2004.
[6] A. Bovik, The Essential Guide to Image Processing. First Edition. San Diego: Academic Press 2009.
[7] K. Cabeen, P. Gent, Image compression and the discrete cosine transform. College of the Redwoods. UK 2009.
[8] J. M. Chen, Biorthogonal Wavelets. Slide. German: University of Trier 2001.
[9] Chen, J.M. 2001. Daubechies Wavelets. Slide. German: University of Trier.
[10] Christensen, M. 2009. JPEG image compression implemented in MATLAB. Technical Report, Department of Electrical and Computer Engineering, College of Engineering.

*Corresponding author: sitiaisyah05@gmail.com