Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/37/20230494

A comparison of application of Fourier Transform and Wavelet Transform on image compression

Yunfei Kan

College of Engineering, University of Wisconsin-Madison, Madison, WI-53715, USA

ykan5@wisc.edu

Abstract. This paper explores the application of Fourier Transform and Wavelet Transform in image compression, comparing their performance in terms of various parameters. As data compression continues to evolve to minimize storage and transmission costs, image compression techniques play a crucial role. Among the major compression methods, Fourier Transform and Wavelet Transform stand out. In this study, we introduce both transforms, elucidate their respective image compression methodologies, and undertake a comparative analysis of their compression qualities using multiple metrics. The mathematical underpinning of both transforms is discussed, and several quality metrics will serve as standards for comparison. In comparing the two techniques, Wavelet Transform consistently demonstrates better image compression quality, albeit with higher computational complexity. This study underscores the potential for further algorithmic enhancements to improve Fourier Transform-based compression and encourages an understanding of the trade-offs between compression quality and processing efficiency in the context of image compression techniques.

Keywords: Fourier transform, wavelet transform, image compression.

1. Introduction

Data compression has been developed and continuously improved since the invention of computer science to reduce the cost of data storage and transmission. Uncompressed data like images require explicitly large amount of capacity to store and bandwidth to transmit. For decades, researchers has presented various ways to achieve better image compression, and two of the major compression techniques are Fourier Transform and Wavelet Transform. In the rest of the paper, we will briefly introduce both transforms, describe their respective methods of image compression, and finally compare their compression qualities based on multiple parameters.

2. Mathematical transforms

2.1. Fourier transform

Fourier Transform (FT) describes a mathematical transforming process that takes a time-domain signal to a representation in the frequency domain. The fundamental idea of Fourier Transform is that any measured continuous signal, like sound waves or electrical signals, can be decomposed to linear combinations of weighted cosine signals of different frequencies. Breaking down a signal into its constituent cosine waves makes a waveform in time domain, which in normal cases is complicated to interpret, able to be represented by merely vertical lines in the frequency domain. This also helps analyzing and processing the signal, since it is easier to filter out certain frequencies than times.

2.2. Discrete fourier transform

Fourier Transform breaks down a time domain signal into easily understood sinusoidal waves that are easily understood, and the outputs are complex numerical so that the amplitudes and phases are maintained. When Fourier Transform is applied on a discrete signal, what can be observed from the frequency domain is an infinite sequence of values, and this process is called the Discrete Time Fourier Transform (DTFT) expressed as:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-jn\omega} \tag{1}$$

However, the output of DTFT, $X(e^{j\omega})$, is a continuous function of ω and cannot be directly analyzed by a digital computer. First replace ω with 2πf, and then sample the infinite sequence with a selected number N, we obtain the Discrete Fourier Transform (DFT) of a finite sequence

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{\frac{-j2\pi kn}{N}} \tag{2}$$

where k is a sequence of integers from zero to N minus one.

There is a variant form of DFT called 2-D DFT, and as its name indicates, 2-D DFT transforms a two dimensional signal defined over a finite 2-D dimension m times n. The 2-D DFT of a two dimensional signal, f(m,n), is represented as:

$$F(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)e^{\frac{-j2\pi lm}{M}} e^{\frac{j2\pi kn}{N}} \tag{3}$$

This 2-D DFT is widely applied in areas like image processing. However, directly calculation of 2-D DFT has a high time complexity. Hence, researchers created Fast Fourier Transform algorithm to solve such issue.

2.3. Fast fourier transform

The Fast Fourier Transform (FFT) is an optimized version for the implementation of DFT. Since the original DFT contains N components, the main idea of FFT is to first separate the sequence into even and odd indexed sub sequences so that the two sequences can be processed concurrently. After two N/2-point DFT are acquired, the FFT algorithm goes into the next level to obtain four N/4-point DFT, and this process continues until it reaches an one-point DFT. Then, with some unique methods implemented, the system outputs the transformed representation of each index of the original signal. By comparison, the regular DFT requires a complexity of N^2 operations, where N is the data size, while FFT allows the same calculation with $N(\log_2 N)$ operations. The ratio of DFT to FFT complexity is much more significant when N approaches infinity.

2.4. Wavelet transform

The simple definition of a wavelet is an oscillation existing in a limited amount of time, whose average value is equal to zero. The wavelet has two parameters, scale and location. Scale defines how 'compressed' the wavelet is, the lower the scale, the more compressed the wavelet. Location, as its name implies, sets the location of a wavelet in time.

One major property of Fourier Transform is that FT gives a global average over the entire signal when processing, and this may obscure some local information. Hence it is not that practical to use Fourier Transform in certain cases where short intervals of oscillation is a common scene. An alternative approach is to use the Wavelet Transform that decomposes a signal into wavelets.

DOI: 10.54254/2755-2721/37/20230494

In Wavelet Transform, the scale of a wavelet is selected and used to compare with the original signal, then that wavelet is convoluted across the signal, and multiplication takes place at each step. This forms an integral and the final value gives a coefficient for that wavelet at the chosen scale [1]. Then the scale is changed and the process repeated until the original signal is disintegrated into wavelets. The adjustable characteristic of Wavelet Transform allows a more accurate local description of signals. Through WT, the signal transformation becomes more robust, since Fourier Transform might cause component signals overlap in time [2].

Both Fourier and Wavelet Transform have various usages in fields of engineering. The following parts of the paper put more focus on the comparison of implementation of both methods on image compression. The description of each method application will be in the rest of this paper.

3. Quality metric of evaluation

3.1. Mean squared error (MSE)

Mean Squared Error is an error metric used to quantize and compare qualities of a compressed image. It measures the average square of difference between the two images, original and compressed. A lower MSE means less error occur during compression. MSE depends heavily on image intensity scaling, so it is a disadvantage that it might not perfectly fit the perception of ‘quality’ from a human perspective [3].

3.2. Peak signal-to-noise ratio (PSNR)

The signal-to-noise ratio compares the strength of signal of noise, and gives a ratio of two values, expressed in decibels (dB). The peak SNR indicates the scale of the maximum pixel intensity to the noise intensity in an image. A better image compression quality means it has higher PSNR statistic. The increase of PSNR value also indicates a lower MSE value. It is also noticeable that a PSNR value greater than 35dB indicates a perfect reconstructed image.

4. Fourier transform image compression

4.1. Method

When processing an image, it can be treated as a three-dimensional signal. The two horizontal axes are the spatial coordinates of the screen, and the vertical axis is the value of intensity or brightness of the pixel with coordinates (x,y). For image compression via Fourier Transform, for example, the Joint Photographic Experts Group (JPEG) method, divides an image into blocks of 8-by-8 or 16-by-16 pixels, and the Fast Fourier Transform is performed for each block [4].

When the FFT is taken, it results in an 8-by-8 spectrum for each block, and each value in the spectrum serves as an amplitude of a basis function. The FFT calculates the spectrum by correlating the pixel block with each of the basis functions. In the mentioned spectrum, lower frequencies resides in the top-left corner while higher frequencies are in the diagonally opposite corner [5]. In image processing, the section of the image with high frequency usually have a sharp change of color or brightness, which contains noise. It is noteworthy to mention that the high frequencies tend to have low FFT coefficient values, typically close to zero. The most commonly used compression after the FFT calculation is to discard some of the 64 spectrum values. As mentioned, most signals are contained in low frequency components, thus the higher frequency parts can be eliminated with a filter at the cost of a small fraction of quality.

4.2. Improved techniques

Although the Fast Fourier Transform significantly increased the calculation speed of Discrete Fourier Transform, it is still computationally intensive and usually involves complex components. Thus many researchers has purposed new algorithms, or to combine FFT with other mechanisms, in order to improve the compression techniques.

4.2.1. Novel algorithm for FFT. In the normal FFT algorithm exists irregularities such as complex addition, multiplication, and twiddle factors. As discussed earlier, the FFT of a signal is equivalent to the sum of cosine and sine components of the input. Anitha T G and S. Ramachandran purposed to use these components to overcome the irregularities of the twiddle factor. In their interpretation, the 2-D FFT is interpreted as an m by n matrix for various values of (u,x) and (v,y) as sine and cosine terms [6]. Similar to the JPEG algorithm, the input signal is divided into 8 by 8 sine and cosine matrices, then each is evaluated by equations 4 and 5.

$$C(u+1, x+1) = \cos((\pi/4)*(u*x)) \tag{4}$$

$$S(u+1, x+1) = \sin((\pi/4)*(u*x)) \tag{5}$$

With these two equations, the image is analyzed as 8 by 8 blocks successively, with each sinusoidal transform acquired. The purposed 2-D FFT, F(u,v), is then computed by summing the sinusoidal transform at each index shown in Equation 6:

$$F(u,v) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \begin{array}{l}\{ f(x,y)(\cos(2\pi(ux+vy)/N) \\ + j\sin(2\pi(ux+vy)/N)\}\end{array} \tag{6}$$

They also purposed an Inverse FFT algorithm by simply subtracting the sinusoidal transforms. After IFFT, the processed picture can be visually compared with the original and their PSNR calculated for a more objective comparison. The experiment results were great in quality with higher PSNR values than regular FFT results.

4.2.2. Implementing other technique. One of the steps when applying FFT to image compression is to set a threshold, or a filter, for the complex Fourier coefficients, and apply IFFT to restore the image. Such threshold has to be carefully chosen or the image could end up either not compressed properly or too lossy. Surinder K et al. concluded from experiments that to obtain higher qualities of compressed picture, each image should have its own thresholds [7]. However, finding a fit threshold individually is computationally extensive, and thresholds for images with colors are a set of three values rather than simply light intensity, because such images have different RGB values. For which, Surinder K et al. purposed the Intelligent Water Drop (IWD) algorithm to assist such complex task.

The IWD is an optimization algorithm that is normally used in problems in the form with parameters (N,E), where N and E are respectively the number of nodes and edges [8]. Each IWD begins calculating its solution by traversing the nodes along the edges of the graph until one iteration is complete. The iteration-best solution is then updated and the IWD continues with new parameters. Ideally, the IWD can be used to estimate the RGB triplet's values, but producing each value individually may not be the optimal way. Thus, the research team attempted to first achieve three parameter values for the triplet, which are then used to multiply the absolute maximum of the RGB channel yield [7].

The experimental results were not too significant, but quite an improvement compared with the standard methods. At similar data loss percentages, using the IWD algorithm was able to increase the SSIM values by 0.04 on average while maintaining the visual qualities.

5. Wavelet transform image compression

The basic logic of Wavelet Transform in image compression is similar to that of Fourier Transform. When applying WT, a signal is broken down to its component wavelets, and after this some of the coefficients can be ditched to remove some details and therefore reduce the amount of data used. As mentioned before, wavelets have a property called scale, thus lower-scaled and higher-scaled wavelets can locate fine and coarse details of a image, respectively. There are numbers of families of wavelets, Haar Wavelet, Daubechies Wavelet, etc. Each has its specialty in approximating signals, but the underlying ideas used in image compression are not too different.

Similar to the JPEG algorithm discussed above, the original image is partitioned into non-overlapping tiles, which are later compressed independently; each tile is considered an independent image and ready to be further decomposed. Within each level of 2D Wavelet Transform, the horizontal data is first transformed by low and high pass filters, then the produced horizontal approximation and details are transformed again on columns, producing two groups of vertical approximation and details [9]. This step separates the image into four parts: approximation of the original, and details for horizontal, vertical, and diagonal parts, each resides in a corner.

Still, in calculation, the image is transformed into coefficients, and after the Wavelet Transform is done, a threshold is set to discard unimportant coefficients by setting them to zero to save data. After the previous step, the major data are stored in the approximation image on the top-left, while the other three quadrants hold less crucial information, so zeroing out data in these quadrants becomes a better choice. Typically, people choose from global and level-dependent thresholding [10]. Global thresholding uses a single value for all decomposition levels, whereas level-dependent thresholding applies a different value for each decomposition level. In practice, the PSNR results of level-dependent thresholding are usually notably higher than global thresholding.

Also, with each addition to the level of transform, the top-left part is decomposed again into four parts. As the decomposition level increases, there is a larger proportion of coefficients received from detail quadrants. These groups usually have a shorter range of values than approximations, typically close to zero. Thus, it is easier to filter out values, or to compress data, with threshold zero when the decomposition level is higher.

6. Comparison

Before comparing the performances of Fourier Transform and Wavelet Transform on image compression, we can first take a look at their respective strengths on 1D signal compression. K A A Samsul et al. conducted an experiment, in which the two transforms are applied to compress different types of signals: a block signal, a heavy sine wave, and a mixture of both. The results show that the Wavelet Transform has better compression rate and lower MSE when compressing block wave signal and mixed signal, while Fourier Transform performs better on sine wave signal [11]. This could indicate that Wavelet Transform is more fit than Fourier Transform to process signals with sharp edges, so it is reasonable to predict that Wavelet Transform is stronger at compressing images with sudden changes in color or light intensity. Holding this prediction, we can take a further step to compare the application of two transforms on image compression.

Apart from the JPEG compression method introduced earlier, another algorithm called JPEG2000 was approved in 2003, which features a wavelet transform based compression method. As predicted, JPEG, using Fourier Transform, is less effective at processing details than JPEG2000 using Wavelet Transform. Comparing side by side, the pictures processed by JPEG had significant blurring and high level of blockiness. Details like the hair of the Lena picture, the sweat pores of a fingerprint picture, sharp transition of colors, and even straight lines of a construction blueprint, were either smeared or blocked, making them undetectable after applying JPEG compression. This is because the limitations of Fourier Transform makes JPEG algorithm simply throwing away the highest frequencies (thus the smallest details) in an image. On the other hand, JPEG2000 compression retained most fine details of a picture without much blurring under the same compression rate, but at the cost of a considerable increase in complexity and processing time [12]. Taking a look at mathematical parameters, under the same compression ratio, JPEG2000 always obtains a higher PSNR value than JPEG, meaning that it retains more information and details [13].

7. Conclusions

In the paper above, we discussed two transforms, Fourier and Wavelet, and their applications in image compression. The basic algorithm of both transforms use in image compression are similar, but their specific properties make wide differences in the results. By comparison, the compression quality of Wavelet Transform outperforms that of Fourier Transform in both visual and statistical qualities. However, better performance is achieved at the cost of longer processing time and computational complexity. The quality of Fourier Transform compression can be improved by inventing new algorithms to fortify the capabilities of FT, or combine it with mathematical models to optimize its performances.

References

[1] Burrus C S, Gopinath R and Guo H 2015 "Wavelets and Wavelet Transforms". Accessed: Aug. 25, 2023. [Online]. Available: http://cnx.org/content/col11454/1.6/

[2] Sifuzzaman M, Islam M R and Ali M Z 2009 "Application of Wavelet Transform and its Advantages Compared to Fourier Transform". *Journal of Applied Science,* **vol 13**, pp.121-134

[3] Mathworks 2023 Image Quality Metrics. Accessed: Aug 25, 2023. [Online]. Available: https://www.mathworks.com/help/images/image-quality-metrics.html

[4] Hassan M A, Al-Azawi R G and Gittaffa S A 2006 "Digital Image Compression Using Fourier Transform and Wavelet Technique". *Journal of Engineering and Development,* **vol 10, no. 4**, pp.153-165

[5] Smith S W 1997 *The Scientist and Engineer's Guide to Digital Signal Processing.* San Diego, California: California Technical Pub

[6] Anitha T G and Ramachandran S 2013 "Novel Algorithms for 2-D FFT and its Inverse for Image Compression". *2013 International Conference on Signal Processing, Image Processing and Pattern Recognition (ICSIPR)* pp.1-4

[7] Kaur S et al. 2022 "Optimizing Fast Fourier Transform (FFT) Image Compression using Intelligent Water Drop (IWD) Algorithm". *International Journal of Interactive Multimedia and Artificial Intelligence,* **vol 7**, **no. 7**, pp.48-48

[8] Shah-Hosseini H 2009 "The intelligent water drops algorithm: a nature-inspired swarm-based optimization algorithm". *Int. J. Bio-Inspired Computation,* **vol 1**, pp.71-79

[9] Lees K 2002 "Image Compression Using Wavelets" pp. 1-40

[10] Yadav R K, Gangwar S P and Singh H V 2012 "Study and analysis of wavelet based image compression techniques". *International Journal of Engineering, Science and Technology,* **vol 1, no.1**, pp.1-7

[11] Karim S A A, Karim B A, Hasan M K, Kamarudin M H and Sulaiman J 2011 "Wavelet Transform and Fast Fourier Transform for Signal Compression: A Comparative Study". *2011 International Conference on Electronic Devices, Systems and Applications (ICEDSA)* pp.280-285

[12] Samulski M 2005 "Compression of Images, Fourier and Wavelets". pp.1-29

[13] Ebrahimi F, Chamik M and Winkler S 2004 "JPEG vs. JPEG2000: An Objective Comparison of Image Encoding Quality". *The International Society for Optical Engineering (SPIE)*