

A Retina-inspired Sampling Method for Coding and Compression for Enhancement Image Based on Discrete Double Density Wavelet Transformer

Ali K. Nahar¹

Department of Electrical Engineering, University of Technology, Baghdad, Iraq
ali.k.nahar@uotechnology.edu.iq

Abstract— The retina, a layer of light-sensitive tissue on the inner surface of the eye, is what allows for vision. That propose to design and analytically examine an image quantizer coding that is inspired by the way the retina manages and compresses visual features based on past studies. Retina-inspired coding is a new cipher algorithm which is very promising. Comparing the Double-Density Discrete Wavelet Transform (DDDWT) to earlier traditional approaches, it is more difficult but also produces excellent results. This efficient coding is credited to a streamlined account that, by the way, deals with 2-D and 3-D pictures, transformation matrices, and the conversion of the number DDWT as if via matrix multiplication. Naturally, this code makes a lot of interesting points of view available. In this proposal, we'll try to show how various theoretical model extensions can be applied to applications in video surveillance, information technology, and neuroscience. Therefore, that believe that using a multidisciplinary approach will help system achieve objectives. This strategy would apply signal processing methods with data from neurophysiology. There are hoping that our efforts will result in some innovative new coding algorithms. The proposed codes will be implemented in MATLAB for ease of experimentation. To implement this code more successfully, a low-level programming language is required.

Index Terms— DDDWT, Retina-inspired, Coding.

I. INTRODUCTION

In recent years, advancement in video compression algorithms has become extremely difficult, as improving existing standards appears to be both complex and time-consuming [1-3]. The most recent typical, High Efficiency Video Coding (HEVC), was published in 2015, 10 years after its predecessor, H.264 [4]. In terms of bit-rates, HEVC exceeds its predecessor by over 50%, which is a significant achievement [5]. The compression needs, which have increased over the past ten years, are thought to make this development insufficient [6]. One or two of the most pressing issues that call for faster development of the video compression algorithms are the widespread use of video surveillance cameras in security systems, the widespread use of 4K cameras, Facebook Surround 360 cameras, Ultra-high-definition television, smart phones, and the widespread use of the internet and social media [7-8].

Technology has improved, making people's lives significantly better and easier than they were previously [9]. High-resolution images and videos are among the features of this advancement that require adopted by the majority of scientific devices [10]. The following are some intriguing examples: Analog television was replaced by digital television, which broadcasts through cables, satellites, and the internet, and offers a wide range of television and channels shows [11]. Videotape, that once used to store television shows and cinemas, have given way to CDs, DVDs, Blu-Ray Discs (BDs), and other forms of digital video storage. Cellphones are utilized in addition to making calls and sending SMS, as

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

mobile phones, cameras, web browsers, social networking gadgets, and navigation systems. [12]. People's lives have improved as a result of technological advancements.

The integrate-and-fire (IF) model is one of the neuron models that are most frequently employed in neuroscience [14]. In the IF model, the neuron is represented as a leaky capacitor. The membrane is charged when the stimulus current is conducted. A spike is fired when the membrane voltage surpasses the threshold. The membrane potential is then returned to its resting state. The integrate and-fire model is simplistic in comparison to a genuine spiking neuron. The classic, on the other hand, has been widely employed in neuromorphic engineering to create sensory systems. Another popular neuron model is the Spike Response Model (SRM), which is an expansion of the leaky integrate-and-fire model [15]. According to the summary of the research, it will not serve the purpose of encryption after pressing the acceleration occurring now. Recent research suggests that employing DDWT code as the hermetic framework equivalent of Daubechies orthonormal wavelet transforms could alleviate two or more of these encoding issues [9]; The wavelet filters are minimal in length and adequate. Certain polynomial characteristics are significant when redundant sampling is used. Due to the fact that DDWT achieves poorer shift sensitivity than DWT [10], it contains twice as many wavelengths at each scale as DWT. DWT filters suffer from two main problems, sensitivity, poor directionality, and lack of phase information. The shift contrast characteristic due to the up sampling process. The reason why down sampling leads to shift anisotropy. The second drawback is the lack of directional selectivity DWT, Which causes major problems in images after compression. The methods used achieve high reliability and are a good method at present, but with the acceleration of the science of transmission and communication, that need an algorithm that addresses the problem statement that may arise due to the accuracy of images and the difficulty of transmitting at Encryption with multiplexing.

The purpose of this work is to visualize new image effects affected by the retina. However, for the encoding issue to remain important to save power and bandwidth there must be a change in the underlying paradigm behind creating new image encoders/decoders. Therefore, that proposal aims to lay the foundations for DDDWT to generate original image compression schemes based on biological visual system models. In this paper the focus is on the clear retina, the main encoder in the visual system and the device responsible for visual stimuli. It may be unknown in deep science regarding the visual system, hence the driving force and new concepts, especially if we can use DDDWT techniques to understand the neural structure of the retina.

II. CODING PRINCIPLE BASED ON DDDWT

The coding principle is a common architecture that all loss compression algorithms have adopted, plus that one specifies the process of encoding and decoding a signal. This input signal f , which could be audio, video, or a combination of these, will first be transformed into a more compressed format. In general, wavelets decompose an image at different scales using a pyramidal algorithm architecture. The decomposition is along the vertical and horizontal directions and maintains constant the number of pixels required to describe the image. Discrete Cosine Transform (DCT), Discrete Sine Transform (DST), and Fourier are a few of the transformations utilized in compression techniques. The output of this transform is identified by a quantizer, which determines whether any duplicate information should be removed. The only thing that can lead to distortion is quantization. As a result, lossless compression does not need quantization. Lossless entropy coding is a function that converts a signal's quantized intensity into code-words with lengths that change inversely with the frequency of occurrence. As was done in previous works, The Discrete Wavelet Transform (DWT) filters were used in conversion processing and achieved good results, but they suffer, as we said previously, from failure of directionality in some cases. By using DDWT, which contains three filters instead of first level and it

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

is divided into nine filters in the second dimension, it is treated. Therefore, the sampling method for encryption and image compression for DWT The performance improves even better with DDWT, as demonstrated by the results. The goal is to use filters of DDWT without complexity or adding greater cost to the encryption system with Retina-inspired sampling method. The proposed solution that is used of spike train decoding algorithms to reconstruct the texture using Huffman coding, allowing each point to be replayed with the use of DDDWT is able to reconstruct image sequences that are easy to view and more flexible than DWT, and is adaptable in addition to providing superior texture with fixed settings, depending on the data. Experimental.

Fig. 1(a) shows the oversampled analysis and synthesis filter bank is used to construct the DDDWT on discrete-time signals (1) [18]. One low pass filter ($h_0(n)$) and two distinct high pass filters ($h_1(n)$ and $h_2(n)$) make up the analysis filter bank's three analysis filters (n). As the input signal $X(N)$ moves through the system, the analysis filter bank breaks it down into three sub-bands, each of which is then down sampled by two. According to the process depicted in Fig. 1(b), $X_L(N/2)$, $X_{H1}(N/2)$ and $X_{H2}(N/2)$ are the low frequency (or coarse) sub-band and the two high frequency (or detail) sub-bands, respectively. The synthesis low pass $h_0(n)$ and two high pass $h_1(n)$ and $h_2(n)$ filters are then used to mix the up sampled signals in order to recover the original signal. It is worth noting that the filters used in the synthesis step aren't always the same as those used in the analysis stage. The time reversals of h_i for an orthogonal filter bank are called $h_i(n)$ (n). The number of wavelets in wavelet frames with the aforementioned shape is double what is necessary. Instead of 2, the filter bank is oversampled by 3/2. If the filter bank is iterated just once on its low-pass branch (h_0) or L , the overall oversampling rate will be 7/4. A three-stage filter bank has an oversampling rate of 15/8. The DDDWT coefficients of the previous step are stated as follows, as specified by the Mallat pyramid algorithm shown:

$$W(k, n) = \sum_i X_{LHj}(i, k - 1) * h(i - 2n) \tag{1}$$

Where, W_{ij} represents the the scaling and wavelet functions' respective dilation coefficients.

$$Y_{IJ} = W' * X * W \tag{2}$$

Where, Y_{IJ} is the image or signal transformation, (I) represents low and (J) high are respectively.

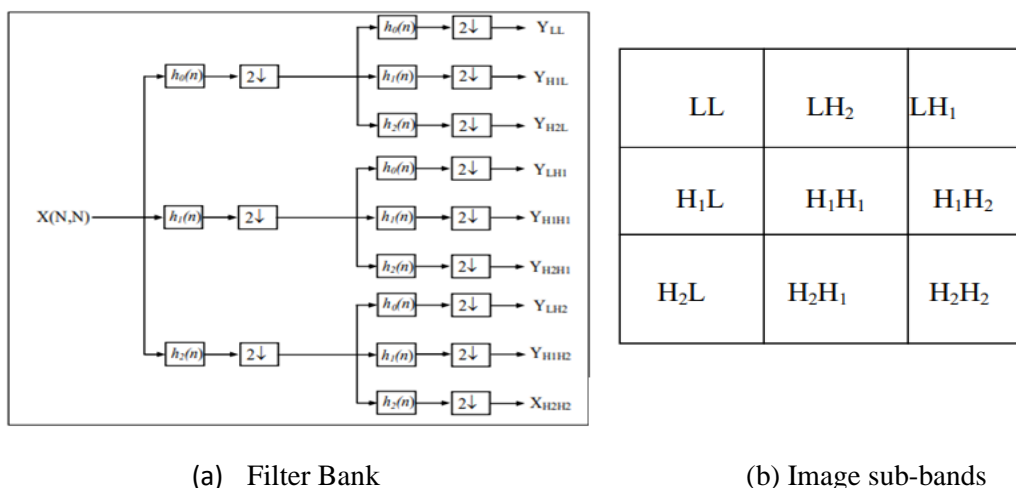


FIG. 1. CONSTRUCT 2-D LEVEL DECOMPOSITION FOR DDDWT [22].

III. ENTROPY CODING (SHANNON, HUFFMAN)

Suggested a systematic strategy are achieving Shannon's ideal compression ratio. In recognition of his achievement, the coding technique is known as "Huffman code." Huffman codes are used to compress and transmit digital data in various devices, including fax machines, modems, computer networks, and high-definition television (HDTV). Huffman coding has the benefit of making use of the frequency disparities. Less storage is needed for frequently occurring characters at the expense of greater storage being needed for each of the less common characters. A practice collection of information indicative of the information to be lossless compressed Image can be used to estimate the frequency of occurrence. Then, Shannon entropy is as followed:

$$H(s) = \sum_i p_i \log_2 \left(\frac{1}{p_i} \right) \tag{3}$$

In addition, that can describe unquestionable entropy:

$$s_i \forall \leq p_i \rightarrow E(s) = \sum_i \text{Min}(s_i^2, p_i^2) \tag{4}$$

Where s_i are the DDDWT coefficients, s is the signal, and p_i is a positive threshold Entropy is a widely used concept, particularly in the field of signal processing. Information-related characteristics for a precise representation of a particular signal are described by the traditional entropy-based criterion. Information-related characteristics for a precise representation of a particular signal are described by the traditional entropy-based criterion. Entropy, which has knowledge about the image's concentration, is frequently utilized in image processing. Condition an alphabet (ABC) has n the probabilities of occurrence are p_1, p_2, \dots, p_n and different symbols s_1, s_2, \dots, s_n , Rearranging the symbols results in $p_1 > p_2 > \dots > p_n$. Connect the two symbols with the lowest probabilities to a contraction process. Let's say there are two symbols, s_{n-1} and s_n . That substitutes a fictitious symbol for these two real symbols. Nearly, -1 , with a $p_{n-1} + p_n$ probability of occurrence. As a result, the new set of symbols has the following members: s_1, s_2, \dots, s_{n-2} , and H_{n-1} . Repeat the preceding procedure until there is only one member left in the final set. The binary tree is then traversed from the root to the leaf node corresponding to probability to get the code-word for each symbol s_i . An intriguing example five proprieties (A, B, C, D & E) of Huffman coding is given in Fig. 2.

S_i (P_i)	L evel 1	Level 2	Level 3	L evel 4
A(0.4)	A (0.4)	A(0.4)	B+E+D +C(0.6)	{ B,C,D,E} {A}(1)
B(0.25)	B (0.25)	D+E+ C(0.35)	A(0.4)	
C(0.2)	C (0.2)	B(0.2 5)		
D(0.1)	D +E (0.15)			
E(0.05)				

s	C_i
A	0
B	10
C	100
D	110
E	111

FIG. 2. HUFFMAN CODING.

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

IV. RETINA-INSPIRED FILTERING OF IMAGE

In this study, there are developing the dynamic retina filtering models to process images dynamically interested in changing the conventional coding principle shown *Fig. 3*. The visual system is a powerful tool that effectively processes input visual stimuli on the fly. Numerous neuroscientific models have been proposed in the literature in an effort to fit the neuroscientific measures of various organisms, such as salamanders, cats, primates, etc. That the complexity, computational expense, and power consumption of the video compression methods will be improved and benefited by the dynamic models.

These bio-inspired methods focus more on the modeling of peripheral idea. However, the sharp images are produced by the fovea, which is a structure in the center of the primate retina. [21]. Fremelet theory contains the possibility of spatiotemporal separation that adapts to the time-pulse response due to the three filters, as the tight-frame version of Daubechies' orthonormal wavelet transform; in an oversampled setting, the wavelet filters satisfy a number of crucial polynomial criteria and are of minimal length [22]:

$$Y(x_i, t) = H_i(t) * W(x) \quad (5)$$

Where $H_0(t)$ is typically band-pass for "transient" units and low-pass for "sustained" units (i equal 1 and 2). Building a multi scale bank of the DDDWT filters with each one's resolution determined by $H_i(t)$, which is generated by Wavelet stationary, was the aim of the theory's temporal function. Fremelet Transformer made another attempt to enhance the static the DDDWT filter [23]. The input signal is processed, and then random spike trains are generated using the firing rate inferred from the modified signal. As a Huffman process, each spike train exists.

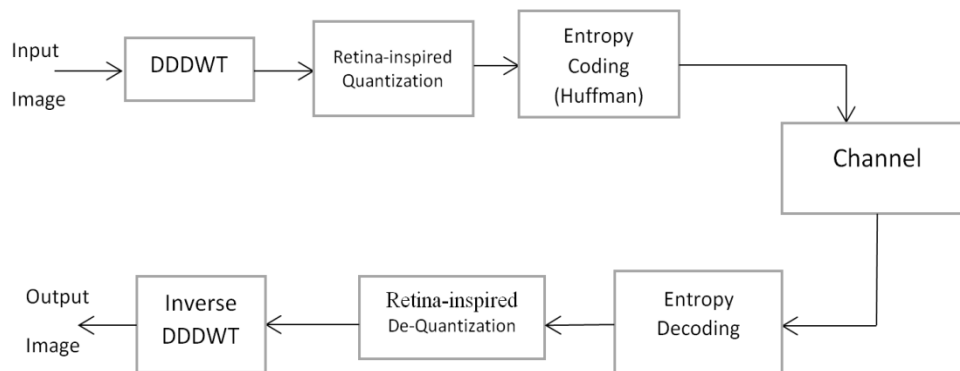


FIG. 3. DDDWT CODING PRINCIPLE.

For pixels, if they accumulate Intensities up to the transmission threshold T , as well, soft thresholding is a method of converting all factors into zero by a specific quantity determined through the value of threshold (T). A specific signal's Cucumber Threshold value is crucial for reducing noise. Transactions (on each level) made from zero and below the threshold (T) are transactions, and transactions made above the threshold are not changes. The firing of a spike indicates that the brightness is sufficient. The related complex is reset at the same time by draining all fees.

$$T \leq R = \int_0^t I dt \quad (6)$$

Where, I mention the brightness intensity, R stands for retina-inspired, and t for integration time.

Fig. 4(a) being motivated by the visual system model. In a mechanism camera, photoreceptors transform the light's intensity into a voltage. The composite at each pixel adds the intensity once the

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

analog-to-digital converter has finished signal conversion and output file digital luminance intensity. Various light intensities correspond to various Accumulation (A) rates. To address the issue of visual reconstruction, we provide a sampling model that is inspired by the retina and an optical reconstruction model. Commonly, DVS sensors and their conversion to a brand-new camera known as the Spike Camera cause problems. In Fig. 4(a), the sampling process is displayed. This particular graphic illustrates and evaluates the various mathematical models that have been put forth to explain how the retina is controlled to dynamically filter visual inputs. Therefore, starting with static spatial models that can be transformed into dynamic models with the aid of transformation and that fit more precisely with the neuroscience measurements that were done in Fig. 4(b); the models propose to approximate this retinal filter.

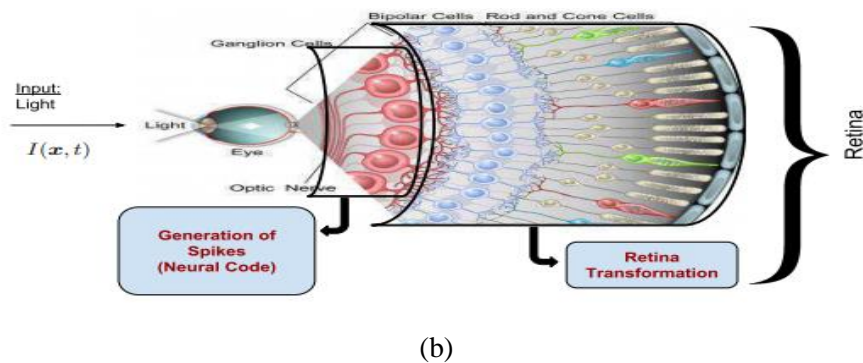
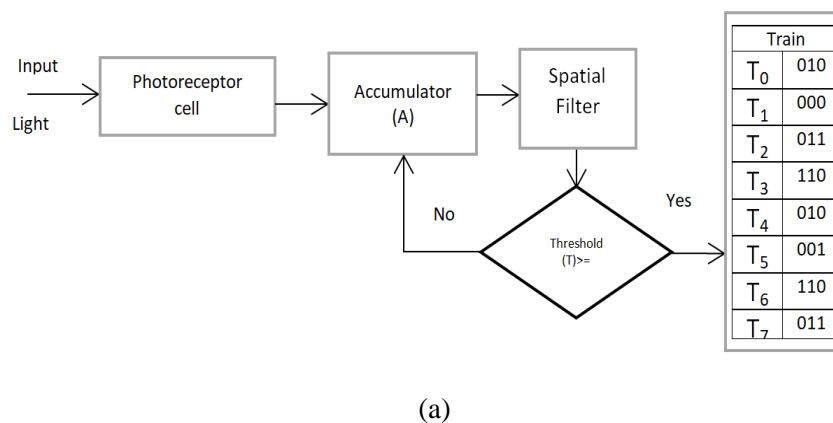


FIG. 4. PROCESS CAPTURED BY SPIKE'S CAMERA (A) A STREAMLINED PROCESS (B) EYE VISUAL PLANNING.

V. NUMERICAL RESULTS

The purpose of this section is to assess the decomposition-inspired state of the retina, which was set in accordance with the reasonable biometric parameters supplied in Section II and displayed in Fig. 5. These parameters can be changed, and DDDWT filters can be created, making them more suitable to other research areas. The purpose of my effort does not include changing these settings, but that test and analyze certain physiologically illogical criteria results. Additionally, Fig 6 and 7 show the cameraman and the woman for the two different scenarios in which my photo was taken. To create a convenient, smooth transformation, the second group is split into three levels. For the first group, the retina-inspired filter will reach its full potential in the first level.

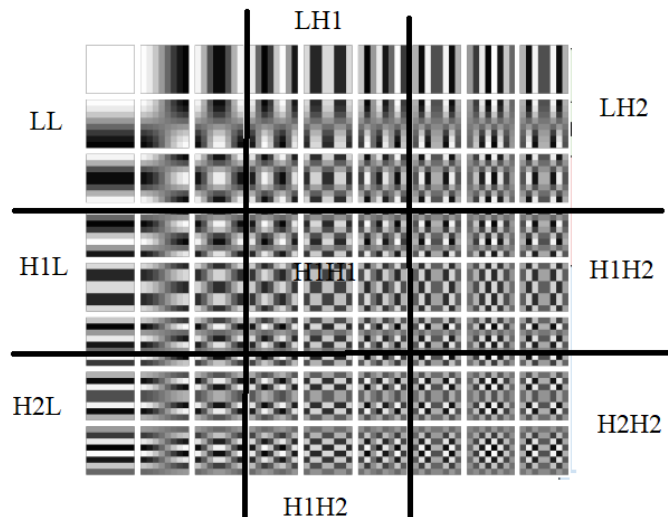
DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

FIG. 5. DECOMPOSITION FOR THE DDDWT IMAGE SUB-BANDS TWO-LEVEL.

Inverse of the first set is the second set, and the third group is equivalent to zero since it has zero affinity. *Fig. 5* makes it obvious that the first two groups have the exact same spectrum, at least for biologically realistic parameters. As a result, the first group—which is close to the ideal case is the subject of our analysis. When a possible bio-parameter image with the retina-inspired filter has $H_0(t) = LL$ in *Fig. 6(i)*, we show the filtering results *Fig. 6 (c)*. The DDDWT First and Level database houses the photos in the left and center columns. The image in the right column was obtained using the high bass (band pass filters H11 and H22) from *Fig. 7(d) & (e)*.

The images are $n = 256 \times 256$ pixels in size, and for each experiment, we decided to only display the separation of two layers. The bandwidth of *Fig. 5* is used to build the retina-inspired filter. We experimented with various PSNR and T fixed weight parameters in *Fig. 6*. *Fig. 7's* first column corresponds to the parameters that are biologically conceivable when $T = 1.5$. The parameters in the middle and left columns are biologically undesirable but are extremely similar to those in the left column. It is intriguing that while PSNR is decreasing, LL and $H_0(t)$ instances are being classified more strongly.

When comparing images, the mean squared error (MSE)—while simple to implement—is not highly indicative of perceived similarity. Structural similarity aims to address this shortcoming by taking texture into account. The structural similarity index measure (SSIM) is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. SSIM is used for measuring the similarity between two images as shown in Table I.

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

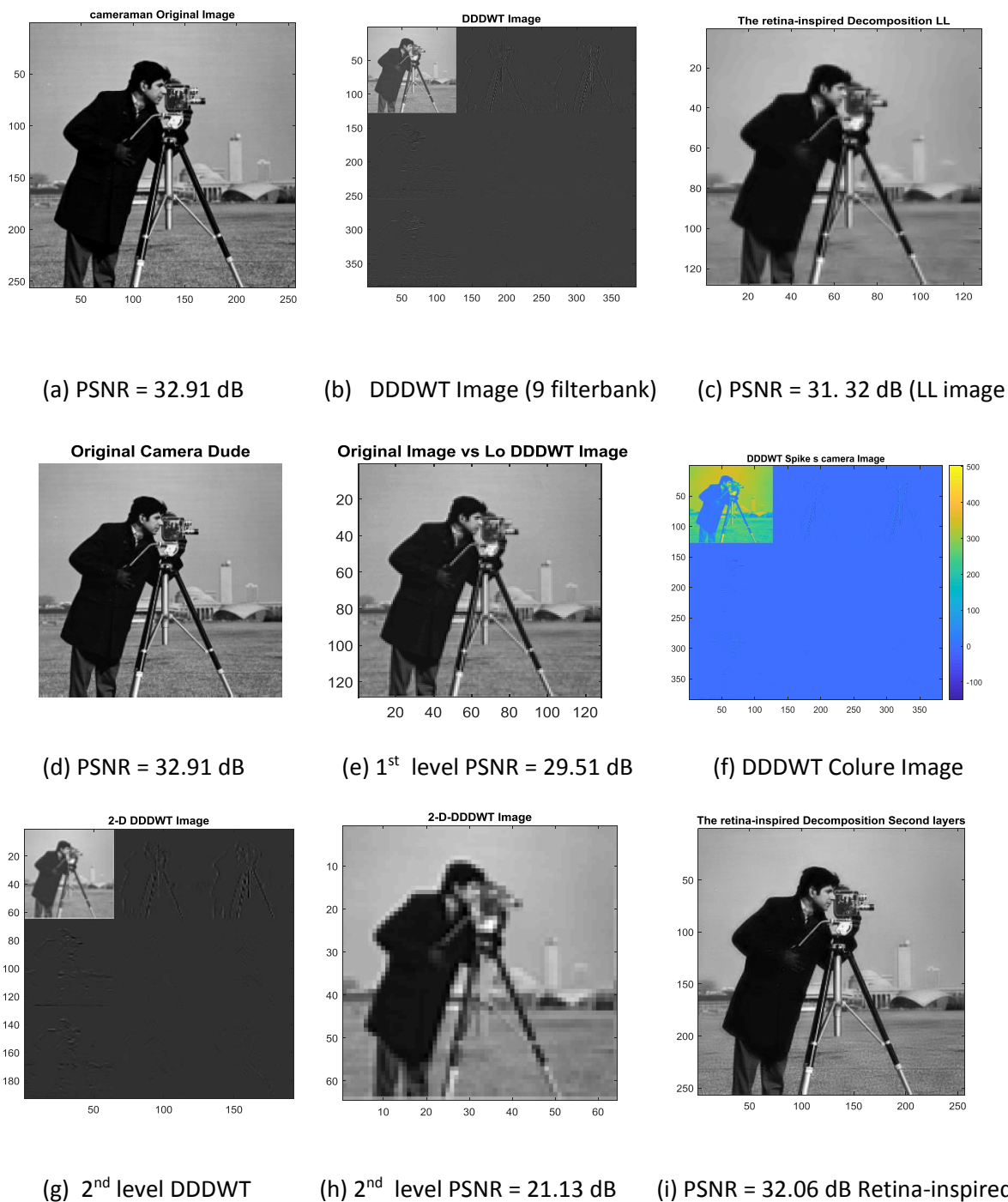


FIG. 6. IMAGE COMPRESSION BASED ON DDDWT ON VISUAL COMPARISON OF THE CODEC INSPIRED BY THE RETINA WITH THE "CAMERAMAN.TIF" IMAGE.

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

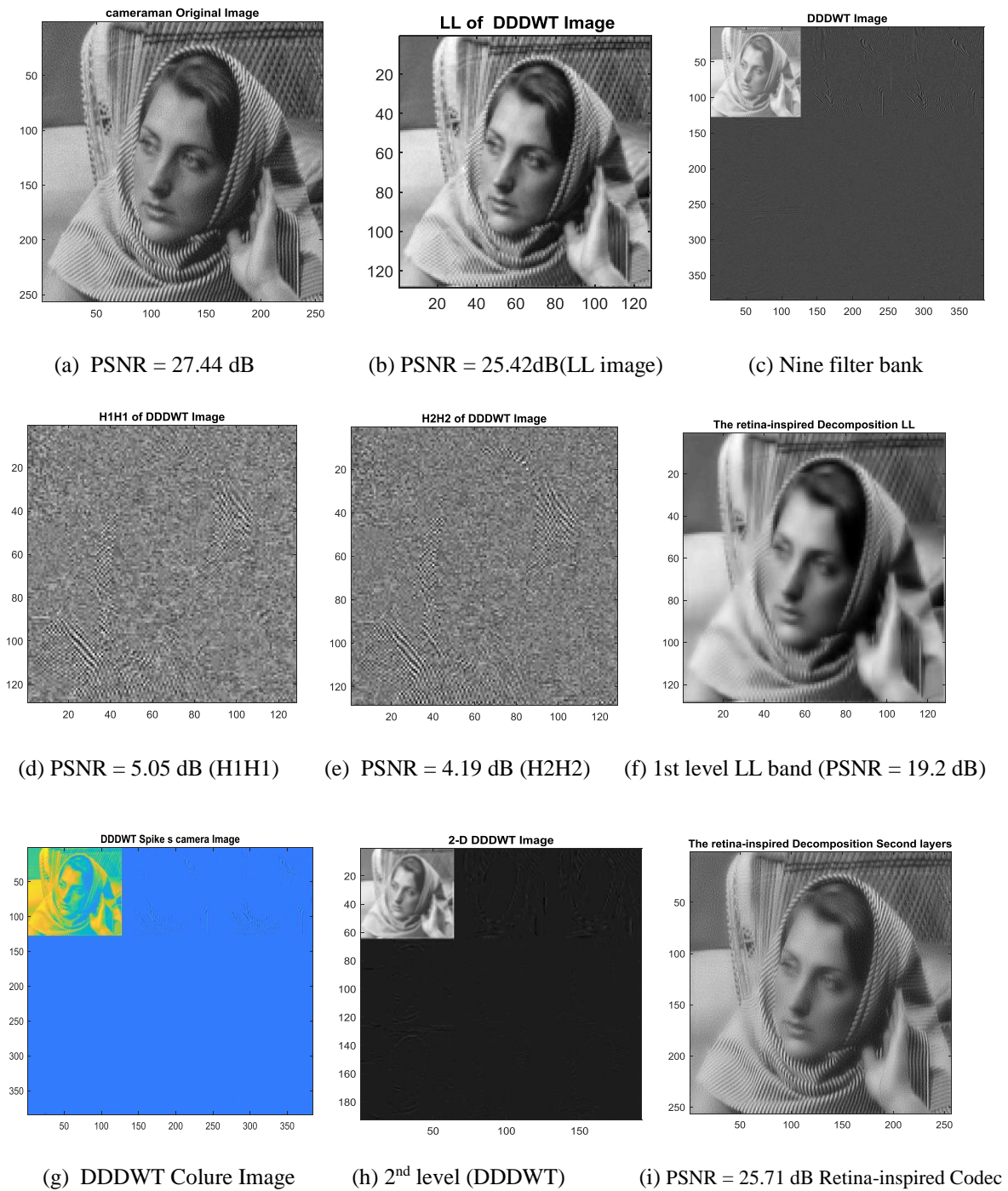


FIG. 7. DDWT-BASED IMAGE COMPRESSION COMPARISON OF THE "WOMAN" IMAGE AND THE RETINA-INSPIRED CODEC.

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

TABLE I. IMAGE COMPARISON (MSE VS SSIM)

Name of image	Original Image			Compression Image (LL)			Retina-inspired Codec		
	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR
cameraman	1	0	32.91	0.65	0.232	31.32	0.92	0.049	32.06
Woman	1	0	27.44	0.80	0.291	25.42	0.83	0.120	25.71

VI. CONCLUSIONS

This paper presents various models related to the decoding of the eye, using the DDDWT using a camera inspired by biological life imaging the retina. Algorithms have been proposed to decode the thorny train to reconstruct the texture with Huffman coding, allowing each point to be replayed. For real-time applications like object identification or action recognition tasks, Huffman encoding is preferable. The DDDWT is able to recreate display-friendly image sequences, and DDDWT is adaptive plus provides superior texture with static settings, according to experimental data. More efficient tissue reconstruction algorithms must be investigated in the future, especially in complex scenarios. This paper will be helpful to any scholars who wish to learn more about the visual system's capacity for perception. By comprehending its structure and operation, researchers can more effectively develop coding algorithms.

REFERENCES

- [1] L. Zhu, S. Dong, T. Huang and Y. Tian, "A Retina-Inspired Sampling Method for Visual Texture Reconstruction," 2019 IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, **2019**, pp. 1432-1437, doi: 10.1109/ICME.2019.00248.
- [2] E. Doutsis, L. Fillatre, M. Antonini and J. Gaulmin, "Retina-inspired video codec," 2016 Picture Coding Symposium (PCS), Nuremberg, **2016**, pp. 1-5.
- [3] M. S.Mohamed, S. Ange A.M. A.A.Malek, A.T. Bosakowski "Physics-based and Retina-inspired Technique for Image Enhancement" Conference: 9th International Conference on Bio-inspired Systems and Signal Processing January **2016**, DOI: 10.5220/0005709902390247.
- [4] E. DoutsisEffrosyni, D.Lionel F. Lionel F. Marc A. AntoniniJulien G.J. Gaulmin "Retinal-inspired filtering for dynamic image coding" Conference: ICIPAt: Quebec, Canada, September **2015**; DOI: 10.1109/ICIP.2015.7351456.
- [5] L. Zhu, S. Dong, T. Huang, Y. Tian "A Retina-Inspired Sampling Method For Visual Texture Reconstruction" Xiv:1907.08769v1 [eess.IV] 20 Jul **2019**.
- [6] H. Yaichi, A. Mohamed "A novel bio-inspired coding scheme for wireless sensor networks" MasmoudiComputer Science2016 11th International Design & Test Symposium (IDT) **2016**.
- [7] L. Zhu, S. Dong, J. Li, T. Huang and Y. Tian, "Retina-Like Visual Image Reconstruction via Spiking Neural Model," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, **2020**, pp. 1435-1443, doi: 10.1109/CVPR42600.2020.00151.
- [8] Y. Ye, S. Ci, A. Katsagelos, Y. Liu, and A. Y. Qian, "Wireless video surveillance: A survey". IEEE Access, 2013, pp:646-660.
- [9] R. R. S. Tomar, and K. Jain, "Lossless image compression using differential pulse code modulation and its application. International Journal of Signal Processing, Image Processing and Pattern Recognition" vol. 9, No.1, pp:197-202,2016.
- [10] G. Sullivan, J.-R. Ohm, W.-J Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard. IEEE Transaction on Circuits and Systems for Video Technology", vol. 22, No.12, pp:1649-1669, 2012.
- [11] L. U. Perrinet, "Biologically inspired computer vision, chapter Sparse models for Computer Vision" (Number 14) Wiley-VCH Verlag GmbH & Co. KGaA. 2015.
- [12] B.Pesquet-Popescu, M., Cagnazzo, and F. Dufaux, "Motion-estimation - A Video Coding Viewpoint. Academic Press Library in Signal Processing" Image and Video Compression and Multimedia, vol. 5 No.27, 2014.
- [13] M. C. Potter, B.Wyble, C. E. Hagmann, and E. S. McCourt, "Detecting meaning in RSVP at 13ms per picture" Attention, Perception & Psychophysics, vol76, No.2, pp:270-279, 2014.

DOI: <https://doi.org/10.33103/uot.ijccce.24.2.5>

- [14] A. N. Burkitt, "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input," *Biol. Cybern.*, vol. 95, pp. 1–19, 2006.
- [15] M. Avermann et al. S. Mensi, R. Naud, "Parameter extraction and classification of three neuron types reveals two different adaptation mechanisms," *J. Neurophys.*, vol. 107, pp. 1756–1775, 2007.
- [16] E. Doutsis, L. Fillatre, M. Antonini and J. Gaulmin, "Bio-inspired sparse representation of images", *Proc. Groupe d'Etudes du Traitement du Signal et des Images (Gretsi)*, pp. 1-4, 2017.
- [17] E. Doutsis, L. Fillatre, M. Antonini and J. Gaulmin, "Retina-Inspired Filter," in *IEEE Transactions on Image Processing*, vol. 27, no. 7, pp. 3484-3499, July 2018, doi: 10.1109/TIP.2018.2812079.
- [18] A K Nahar, M J Mobarek " Enhancement Medical Images Based on the Optimize Double Density Wavelet Transform Technique" *International Journal of Bioinformatics and Computational Biology*, vol.3, No.1, pp: 6-16, 2018.
- [19] E.Mishra, and K. K. ().Singh, "Comparison of various lossless image compression techniques". *International Journal of Engineering Research and Applications ISSN*, vol.5, No.6, pp: 36–39, 2015.
- [20] D.Mukherjee, J. Bankoski, A. Grange, J. Han, J. Koleszar, P. Wilkins, Y. Xu, and R. Blakeltje, "The latest open-source video codec vp9 - an overview and preliminary results". *Picture Coding Symposium (PCS)*, 2013, pp 390 – 393.
- [21] I. Zitouni R.Sarikaya "Arabic diacritic restoration approach based on maximum entropy models", *Computer Speech and Language*, 2009 pp: 23 257.
- [22] A.K.Nahar, A.S.Jaddar, "A compression original image based on the DDWT technique and enhancement SNR" *Int. J. Eng. Technol. Sci.* vol.5, No.3, pp: 73–89, 2018.
- [23] F. Arrebola et al. R. Marfil, E. Antunez, "Merging attention and segmentation: active foveal image representation," *BrainInspired Computing*, Springer International Publishing, 2013.