

# Performance Analysis of a Hybrid K-Means -Huffman Image Compression Scheme in Rayleigh Fading Wireless Environments

Priyanka Hazowary<sup>1</sup>, Sushanta Kabir Dutta<sup>2</sup> and Rupaban Subadar<sup>3</sup>

Department of Electronics and Communication Engineering  
North- Eastern Hill University, Shillong, Meghalaya, 793022, INDIA  
Corresponding author: [priyankahazowary16@gmail.com](mailto:priyankahazowary16@gmail.com):

## ABSTRACT

The demand for transmitting digital images over wireless networks necessitates efficient compression techniques that balance compression ratio, image quality, and robustness to channel errors. This paper presents a performance analysis of a hybrid image compression scheme that combines K-Means clustering for lossy vector quantization and Huffman coding for lossless entropy encoding. We evaluate the performance of this scheme in a simulated Rayleigh fading wireless environment, which is characteristics of non-line-of-sight mobile communication channels. Key performance metrics, including Compression Ratio (CR), Peak Signal-to-Noise Ratio (PSNR), and Bit Error Rate (BER) on reconstructed image quality, are analyzed. The results demonstrate that while the hybrid scheme achieves a high compression ratio, its performance is highly sensitive to channel errors.

**Keywords:** K-Means clustering, Huffman coding, Rayleigh fading, image compression.

## I. INTRODUCTION

Image compression is a process to reduce image size in order to transfer over a digital network or save in a smaller memory size while maintaining considerable similarity with original image data. Image size reduction brings us remarkable advantages in many fields like teleconferencing, multimedia, medical applications and many more. Image compression can be classified into two ways, lossy compression and lossless compression. The lossless compression permits the image to be compressed without loss of any data. In short, lossy compression may have higher compression ratio but it may lose some data during this process. Lossy compression is advantageous in some circumstances where certain amount of error is tolerable such as streaming media, videoconferencing, facsimile transmission. By exploiting human visual properties, lossy compression succeeds high compression ratio [1]. Wireless image transmission faces significant

challenges due to bandwidth constraints and multipath fading, which degrade signal quality and reliability. Multipath fading occurs when signals travel multiple paths to the receiver, causing destructive interference that attenuates signal strength and increases bit error rates (BER). Wireless communications have drawn a lot of attention lately. However, wireless channel multipath effects result in the received signal being a total of attenuated, delayed, and phase-shifted replicas of the transmitted signal [2]. With the growing demand for wireless multimedia applications, efficient image compression and reliable transmission over noisy channels have become critical challenges. Image compression techniques are essential to reduce the bandwidth and storage requirements of high-resolution images. However, when these compressed images are transmitted over wireless networks, particularly in fading environments, they are highly susceptible to channel impairments such as noise, interference, and signal fading. Hybrid image compression schemes have attracted considerable attention for their ability to balance compression efficiency with computational simplicity. In particular, the combination of K-Means clustering, a lossy quantization technique, with Huffman coding, a lossless entropy encoding method, has been shown to provide promising results in reducing image data while preserving perceptual quality. K-Means reduces spatial redundancy by grouping similar pixel values, while Huffman coding further reduces statistical redundancy by assigning variable-length codes to the quantized output [3] [4]. On the other hand, Rayleigh fading is a common wireless channel model that represents multipath propagation in non-line-of-sight (NLOS) conditions, such as urban or indoor environments. In such environments, signal fluctuations can severely degrade the quality of received images, especially when compression algorithms are not designed with channel impairments in mind [5] [6] [7]. Although various studies have focused on image compression techniques and their performance in ideal or Additive White Gaussian Noise (AWGN) channels, limited research has analyzed their robustness in realistic wireless fading channels. Furthermore, most existing

works treat compression and channel modeling as separate domains and there is limited evidence that studies higher-order QAM (8/16/32) choices to end image quality after decompression for clustering based compressed sources under Rayleigh fading. This paper aims to fill this gap by integrating hybrid compression schemes with Rayleigh fading channel modeling to evaluate image quality and transmission reliability [8].

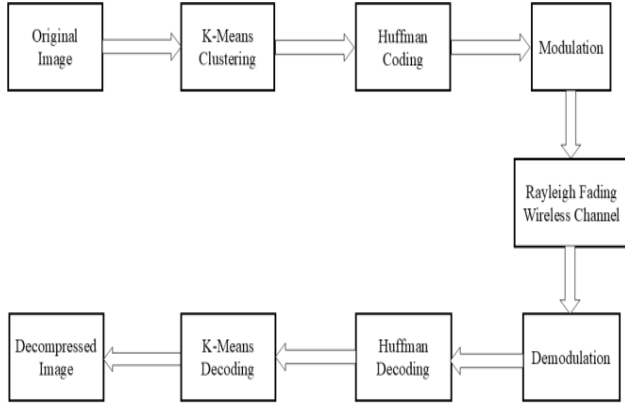


Fig. 1. A schematic diagram showing comprehensive performance analysis of a hybrid image compression scheme.

The structure of this paper is organized as follows: Section II discusses the Overview of the Proposed hybrid compression models framework. Section III presents about the Wireless channel model. Compression performance evaluation is provided in Section IV. Section V include the simulation results and discussions, followed by the conclusion and future work in Section VI.

## II. OVERVIEW OF THE PROPOSED HYBRID COMPRESSION MODELS FRAMEWORK

### A. *K-means clustering*

Is an unsupervised learning approach for clustering data. It produces a particular number of independent and non-hierarchical groups depending on attributes [9]. It is used to split data when it contains unlabelled data. In k-means clustering, the total given data is partitioned into k number of clusters. Basically, this algorithm has two stages where in the first stage it calculates the k means which called centroids and in the second stage assigns each point to the cluster which has the lowest distance between the cluster's centroid and the respective data point. However, when the partition process is done, we need to recalculate the k centroids (means of each cluster). The distances between each centroid and data point must be recalculated, and the minimal distance is used for re-clustering. Each cluster within the partition is identified by its member objects, as well as its centroid.

The sum of the distance from all the objects over the cluster is minimized at the centroid [21].

The main target of K-means algorithm is to partition 'm' data points in 'k' number of clusters. The constraints of the K-means clustering algorithm include that the user must set the 'k' value, and it is sensitive to data points far from the clusters [16]. The process of K-means algorithm goes through 4 steps which are briefly summarized below:

- First, the value for 'k' is set and according to that k initial centroid values are picked randomly.
- Second, it is ensured that each observation is assigned to the nearest centroid by calculating the distance between them. Any distance metric can be used for this calculation.
- Third, initial centroids are moved to the new mean positions calculated for each cluster.
- Second and third steps are continued until all the centroids do not move further and it is ensured that the clusters do not overlap [17] [18].

### B. *Huffman coding*

Huffman coding, introduced by Dr. David A. in 1952, is a lossless data compression technique that constructs minimum redundancy codes based on the frequency of symbols. It uses a greedy algorithm to assign shorter binary codes to more frequent symbols and longer codes to less frequent ones, optimizing data representation [10]. Huffman coding is highly effective for images with high redundancy, achieving significant compression without quality loss, though it may be less efficient for images with complex or irregular pixel distributions [11]. The first stage in the Huffman algorithm is to generate a sequence of source reductions by ranking the probability of each symbol and combining the two least likely symbols into a single symbol. This will be utilised in the subsequent source reduction stage to generate codes based on the frequency of input symbols.

Nodes and leaves are the main parts of a Huffman tree. We calculate which two leaves have the lowest probability at each stage and combine them to create a node. With N being the number of symbols, the tree is constructed in N-1 steps using a bottom-up approach. Every route that leads to the right is given a 1; every path that leads to the left is given a 0. To develop the code for a certain symbol, move down the tree in a top-down approach and build up the code. Huffman codes assign the shortest codes to the characters with the highest frequency [15].

### III. WIRELESS CHANNEL MODEL

The wireless channel is random in nature and has various reflectors in the environment, resulting in fading. Multiple reflections generate multiple paths for a signal to travel from transmitter to receiver, which is known as multipath propagation [20]. An efficient bandwidth channel with additive white Gaussian noise, a greater bitrate, a shorter propagation time, and reduced power consumption are all required for smooth picture transmission over a wireless channel [12]. The corruption of a wireless transmission channel is caused by errors that are bursty in nature. Physical channel modelling is a challenging task that also relies on the transmitter, receiver, and other objects moving along the signal path [13].

#### A. Mathematical analysis of Rayleigh fading Model

Rayleigh fading is a statistical model that describes the fast variations of a signal as it travels through a wireless communication channel and there is no dominant propagation along a line of sight between the transmitter and receiver. In this case, the probability density function (PDF) of channel fading amplitude can be given as [14]

$$P_{\alpha}(\alpha) = \frac{2\alpha}{\Omega} \exp\left(-\frac{\alpha^2}{\Omega}\right), \quad \alpha \geq 0 \quad (1)$$

where  $\Omega = E[\alpha^2]$  and  $E[\cdot]$  is the expectation operator. A Rayleigh RV  $\alpha$  can be modeled as  $\alpha = |Z| = |X + jY|$ , where  $X$  and  $Y$  are independent Gaussian RVs with zero mean and equal variances.

The SNR pdf of Rayleigh fading channel can be derived from (1) by performing the random variable transformation according to the following formula

$$\gamma = \frac{E_b}{N_0} \alpha^2, \quad (2)$$

where  $E_b$  is the energy per bit and  $N_0$  is the power spectral density of Gaussian noise.

The final result available in [14]

$$f_{\gamma}(\gamma) = \frac{1}{\bar{\gamma}} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right) \quad (3)$$

By taking the derivative of the product CDF of each input branch, the output SNR pdf for the SC receiver may be obtained. The CDF of single branch can be calculated from (5) by performing the integration, which is given below [14]

$$\begin{aligned} \Pr[\gamma \leq \gamma_{th}] &= \int_0^{\gamma_{th}} \frac{1}{\bar{\gamma}} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right) d\gamma \\ &= 1 - \exp\left(-\frac{\gamma_{th}}{\bar{\gamma}}\right) \end{aligned} \quad (4)$$

If we consider  $L$  number of branches, in the SC receiver the CDF of the output SNR can be obtained by taking product of  $L$  terms of equation (6).

$$F_{\gamma_{sc}}(\gamma_{sc}) = \left[1 - \exp\left(-\frac{\gamma_{sc}}{\bar{\gamma}}\right)\right]^L \quad (5)$$

By taking the derivative of (7) we obtained the output SNR PDF of SC receiver as [15]

$$f_{\gamma_{sc}}(\gamma_{sc}) = \frac{L}{\bar{\gamma}} \left[1 - \exp\left(-\frac{\gamma_{sc}}{\bar{\gamma}}\right)\right]^{L-1} \exp\left(-\frac{\gamma_{sc}}{\bar{\gamma}}\right) \quad (6)$$

The above expression can be used to derive the various performance measures of the communication system.

The error probability for various modulation schemes over Rayleigh fading channel considering SC receiver can be obtained by averaging the conditional probability of error

( $p_e(\in \gamma)$ ) over the PDF given in (8).

$$p_e = \int_0^{\infty} p_e(\in \gamma_{sc}) f_{\gamma_{sc}}(\gamma_{sc}) d\gamma_{sc} \quad (7)$$

For binary coherent modulation

$p_e(\in \gamma_{sc}) = Q(\sqrt{2a\gamma_{sc}})$ , where  $a=0.5$  for binary FSK and  $a=1$  for BPSK. Now putting the values in (9) we obtained

$$\begin{aligned} p_e &= \int_0^{\infty} Q(\sqrt{2a\gamma_{sc}}) \frac{L}{\bar{\gamma}} \left[1 - \exp\left(-\frac{\gamma_{sc}}{\bar{\gamma}}\right)\right]^{L-1} \\ &\quad \times \exp\left(-\frac{\gamma_{sc}}{\bar{\gamma}}\right) d\gamma_{sc} \end{aligned} \quad (8)$$

The above integral is solved by using the binomial theorem and using the equations [14(A-8a and A-6)]

$$\begin{aligned} p_e &= \frac{L\Gamma(1.5)}{2} \sqrt{\frac{a\bar{\gamma}}{\pi}} \sum_{k=0}^{L-1} \binom{L-1}{k} \\ &\quad \times (-1)^k \frac{{}_2F_1\left(1, 1.5, 2, \frac{k+1}{k+1+a\bar{\gamma}}\right)}{(k+1+a\bar{\gamma})^{1.5}} \end{aligned} \quad (9)$$

In a similar way the BER of SC receiver can be evaluated for various modulation schemes.

#### IV. COMPRESSION EVALUATION PERFORMANCE

There are several metrics to determine the image quality, but the most commonly used ones are the PSNR, compression ratio and BER. This section discusses different parameters.

##### A. PSNR

PSNR stands for peak signal to noise ratio. The PSNR is used for the determination of the standard quality a digital image and is measured by

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) (dB) \quad (10)$$

##### B. Compression Ratio (CR)

It is a calculated based on number of bits to represent the original image and the compressed image. Compression ratio shows that how much times the image has been compressed.

$$CR = \frac{N_1}{N_2} \quad (11)$$

Where  $N_1$ ,  $N_2$  represent the number of bits required the original and compressed image respectively.

##### C. Bit Error Rate (BER)

Bit Error Rate is defined as the ratio of the number of bits received in error to the total number of bits broadcast via a communication channel within a certain period.

$$BER = \frac{N_e}{N_t} \quad (12)$$

Where  $N_e$  = Number of bits received and  $N_t$  = Total number of transmitted bits.

#### V. RESULTS AND DISCUSSIONS

This study employs the K-means algorithm and Huffman coding for image compression, utilizes 8-QAM, 16-QAM, and 32-QAM for modulation, and simulates transmission over Rayleigh fading channels, which serve as a realistic model for urban and mobile wireless environments with AWGN, followed by equalization, demodulation, and image reconstruction. The pixel quantity of an image reduces by a considerable amount after quantization of colour shades using K-means algorithm. Clustering is faster on images with lesser colour variation. Therefore, a way to reduce execution time for high colour variation images is processing a lesser number of pixels. The algorithms are carried out by MATLAB. The simulated results are tabulated to compare the PSNR of different algorithms, the compression ratio, Eb/No (SNR per bits) values and the BER, where the PSNR is most frequently

used as a metric to assess the quality of the reconstructed image for image compression schemes.

First, both of the images are compressed by clustering all the pixel data of the images by using some specific 'k' values. The evaluation results are tabulated in Table I. Fig.2 shows PSNR vs Eb/No and Fig.3 shows BER vs Eb/No for 8-QAM, 16-QAM and 32-QAM. However, Fig.4 to Fig.6 displays the reconstructed images of Cameraman and Peppers of size  $256 \times 256$  using Bit Error Rate (BER) performance of 8-QAM, 16-QAM and 32-QAM modulation. are also shown below:

**Table I: Comparison of PSNR and BER vs. Eb/No for Cameraman images**

8-QAM		
Eb/No	PSNR	BER
5dB	15.5703	0.095
10dB	19.5894	0.038
15dB	23.8830	0.013
20dB	29.1665	0.0042
25dB	32.7352	0.0013
16-QAM		
Eb/No	PSNR	BER
5dB	17.4307	0.098
1dB	21.4037	0.04
15dB	25.8778	0.014
20dB	30.5966	0.0046
25dB	35.5109	0.0015
32-QAM		
Eb/No	PSNR	BER
5dB	14.1327	0.13
10dB	17.5288	0.059
15dB	21.8447	0.022
20dB	26.5204	0.0073
25dB	31.6062	0.0024

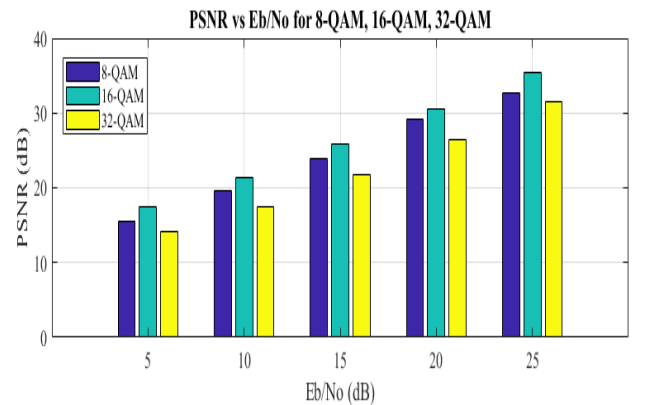


Fig.2. PSNR vs Eb/No for 8-QAM,16-QAM, 32-QAM

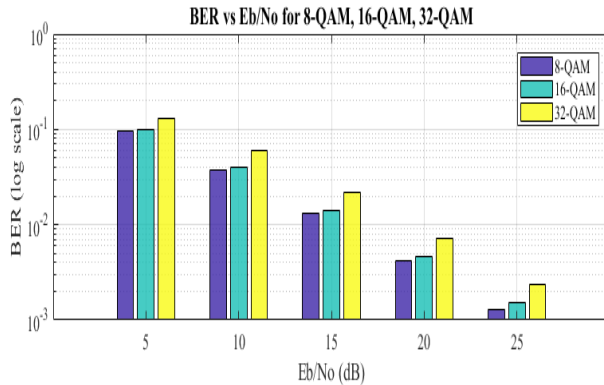


Fig.3. BER vs Eb/No for 8-QAM, 16-QAM, 32-QAM

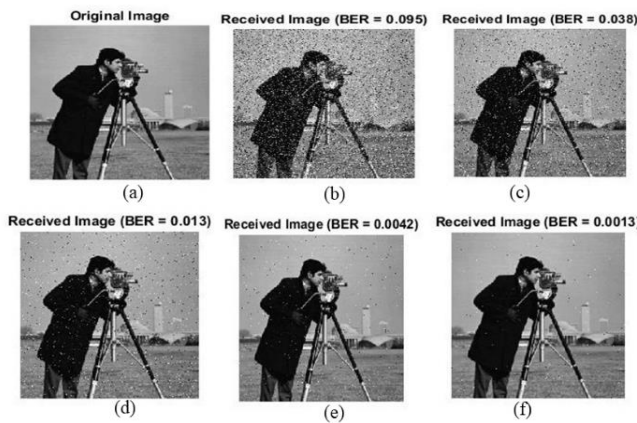


Fig.4. Bit Error Rate (BER) performance of 8-QAM modulation for Cameraman image

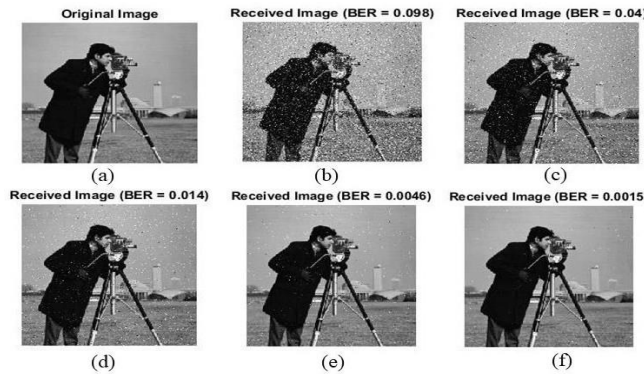


Fig.5. Bit Error Rate (BER) performance of 16-QAM modulation for Cameraman image

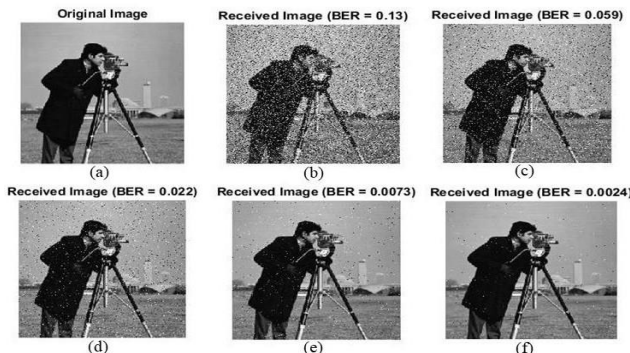


Fig.6. Bit Error Rate (BER) performance of 32-QAM modulation for Cameraman image

Through simulation-based evaluation under varying Eb/No (5 dB to 25 dB) and with different modulation schemes ( $M = 8$ ,  $M = 16$  and  $M = 32$ ), several critical insights were drawn. The hybrid scheme significantly reduced data size by applying K-Means to quantize the colour space, grouping similar pixel intensities based on visual perception. Using Huffman coding to encode the quantized data efficiently without further information loss. The hybrid K-Means–Huffman scheme is highly dependent on Eb/No and modulation type.  $M = 8$  is most robust to channel noise, with lowest BER, making it suitable for low SNR environments, though at the cost of lower image quality and data rate.  $M = 16$  performs best overall, delivering high PSNR and low BER under moderate SNR.  $M = 32$  should be avoided in poor channel conditions, as it requires higher Eb/No to maintain image integrity. Rayleigh fading introduces significant distortion at lower SNR levels. The BER was notably high at low Eb/No (up to 13% for  $M = 32$  at 5 dB), which severely degraded image quality. However, as Eb/No increased, BER dropped exponentially, improving image reconstruction accuracy. This confirmed that the hybrid scheme's error resilience improves with channel quality. K-means has been widely used for vector quantization in image compression by clustering pixels and representing them using cluster centroids. However, K-means introduces quantization error. To further improve compression, entropy coding typically Huffman coding is applied to the quantized indices, leveraging symbol probability distributions for lossless coding. Hybrid methods combining multiple compression tools including clustering, transform coding, and entropy encoding have been shown to achieve better compression ratios while preserving image quality. In Rayleigh fading, the BER increases non-linearly as the compression ratio rises and the Signal-to-Noise Ratio (SNR) diminishes still, our approach demonstrates strong performance, especially when integrated with channel diversity or equalization techniques. In comparison to traditional transform-based or independent entropy coding techniques, the K Means–Huffman approach offers a beneficial compromise reduced encoder complexity and comparable reconstruction quality, even in fading conditions.

## VI. CONCLUSION AND FUTURE WORK

The hybrid K-Means–Huffman coding is effective for wireless image transmission over Rayleigh fading channels, achieving significant compression while maintaining acceptable visual quality, provided channel conditions are suitable. The study provides a high compression ratio by effectively combining the lossy nature of K-Means-based vector quantization with the lossless efficiency of Huffman coding. The hybrid method effectively diminishes image data size via lossy quantization succeeded by lossless entropy coding, achieving a compromise between compression efficacy and computational simplicity. As  $E_b/N_0$  rises, the PSNR markedly enhances, signifying superior reconstructed image quality. The BER diminishes as  $E_b/N_0$  increases, indicating improved resilience to channel noise. The suggested technique demonstrates strong resilience to multipath fading, rendering it appropriate for wireless multimedia transmission. However, in future work we can combine K-Means and Huffman coding with transform coding like DCT, DWT to exploit both spatial and spectral redundancies.

## REFERENCES

- [1] Moulick, H.N. and Ghosh, M., 2013, June. Image compression using k-means clustering and nuclear medicine image processing. In *International Journal of Innovative Research in Computer and Communication Engineering* (Vol. 1, No. 4).
- [2] Lineswala, V. P., & Patel, J. N. (2009, December). JPEG image compression and transmission over wireless channel. In *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies* (pp. 643-645). IEEE.
- [3] Kumar, S., & Singh, R. (2017). Lossless Huffman coding for image compression and decompression based on block and code book size using K-Means algorithm in spatial and frequency domain. *Multimedia Tools and Applications*, 76(1), 1-16.
- [4] Kaur, P., & Singh, S. (2016). Hybrid colour image compression based on FMM and Huffman coding. *Procedia Computer Science*, 85, 464-471.
- [5] Rappaport, T. S. (2002). *Wireless Communications: Principles and Practice* (2nd Edition). Prentice Hall.
- [6] Zhou, Y., & Wang, X. (2008). JPEG Image Transmission over Rayleigh Fading Channel Using Error Correction Codes. *2008 International Conference on Computer Science and Software Engineering*.
- [7] Kumari, A., & Kumar, S. (2024). Compressed Sensing Image Communication in Rayleigh Fading Channel Using Polar Code. *Microwave Review*, 30(2), 3–14.
- [8] Jain, A. K., & Gupta, P. (2021). A review on image compression and transmission techniques over wireless channels. *Wireless Personal Communications*, 117(2), 1365–1383.
- [9] M. R. Mohammed, “Improved K-mean algorithm for background subtraction in video surveillance”, MSc, thesis, University of Technology, Iraq, (2018).
- [10] Shaikh, A. A., & Gaddekar, P. P. (2015). Huffman coding technique for image compression. *Compusoft*, 4(4), 1585.
- [11] Sandeep, P., Reddy, K. N., Teja, N. R., Reddy, G. K., & Kavitha, S. (2023, July). Advancements in Image Compression Techniques: A Comprehensive Review. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)* (pp. 821-826). IEEE.
- [12] Chandra, M., Agarwal, D., & Bansal, A. (2016, July). Image transmission through wireless channel: A review. In *2016 IEEE 1st international conference on power electronics, intelligent control and energy systems (ICPEICES)* (pp. 1-4). IEEE.
- [13] L. N. Kanal and A. R. K. Sastry, “Models for channels with memory and their applications to error control,” *Proc. IEEE*, vol. 66, pp. 724–744, July 1978.
- [14] T. S. Rappaport, “Wireless communication principles and practice,” New Jersey: Prentice Hall, 1996.
- [15] Bhatti, A. T., & Kim, J. (2015). Implementation of lossless Huffman coding: image compression using K-means algorithm and comparison vs. random numbers and message source. *International Research Journal of Engineering and Technology*, 2(5), 497-505.
- [16] Paul, R. K., Jena, S., Chandran, S., Bandyopadhyay, A., & Swain, S. (2024). Image Compression Scheme based on Optimized K-means Clustering and Higher-Level Decomposed DWT. *Procedia Computer Science*, 235, 642-655.
- [17] Arora, K., Shukla, M. and Gobindgarh, I., Performance Comparison of K-means and Rectangle Segmentation Algorithms in Compression of Colour Images.
- [18] Wicaksono, Y.A., Rizaldy, A., Fahriah, S. and Soeleman, M.A., 2017, October. Improve image segmentation based on closed form matting using K-means clustering. In *Application for Technology of Information and Communication (iSemantic)*, 2017 International Seminar on (pp. 26-30). IEEE.
- [19] Usha, S. M., & Nataraj, K. R. (2016). Bit error rate analysis using QAM modulation for satellite communication link. *Procedia Technology*, 25, 456-463.
- [20] Hazowary, P., Talukdar, N., Dutta, S. K., & Subadar, R. (2024, February). A Comparative Study of Jpeg and Similar Compression Standards for Wireless Communication. In *NIELIT's International Conference on Communication*,

*Electronics and Digital Technologies* (pp. 117-129). Singapore: Springer Nature Singapore.

[21] Dhanachandra, N., Manglem, K. and Chanu, Y.J., 2015. Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54, pp.764-771.

---

*Cite this article as:*

*Priyanka , Sushanta and Rupaban , " Performance Analysis of a Hybrid K-Means -Huffman Image Compression Scheme in Rayleigh Fading Wireless Environments" , ACT2024: Special Proceedings of Applied Computer Technology, Published in December 2025,*  
*Link:* <http://actsoft.org/science/act2024-pro/75-ccsn2025.pdf>,  
*AOI:* 10.11234.2024.00036

*@Copyright to 'Applied Computer Technology', Kolkata, WB, India. Website: [actsoft.org](http://actsoft.org), Email: [info@actsoft.org](mailto:info@actsoft.org), published on: December 2025. ISBN: [978-81-985770-3-0](https://www.isbn-international.org/product/978-81-985770-3-0)*