Case Study
UDC 004.021:621.396]:004.761AIoT
DOI 10.7251/IJEEC2601054J

Application of DCT Compression on Images in IoT Systems

Nevena Jeftenić, Tamara Rašić

Information Technology Function, Mtel a.d. – Telekom Srpske, Banja Luka, Republic of Srpska, Bosnia and Herzegovina

E-mail address: nevena.jakovljevic@mtel.ba, tamara.rasic@mtel.ba

Abstract— **Through the integration of sensors, software, and wireless connectivity into everyday activities, the Internet of Things (IoT) creates intelligent ecosystems whose applications are inevitably permeating all segments of life, including smart cities, agriculture, healthcare, and manufacturing. Such systems frequently utilize image and video data as information sources for machine learning and analysis, which are essential for anomaly detection, security surveillance, facial recognition, traffic optimization, and—in the case of this study—monitoring bee activity within a beehive. Given that digital images contain vast amounts of data transmitted through digital processing systems, it is vital to efficiently address the challenges of data transmission within resource-constrained IoT networks. The primary objective of the research presented in this paper is to demonstrate that the application of DCT (Discrete Cosine Transform) compression on images can significantly reduce data volume while simultaneously preserving the image quality required for further application. To evaluate the compression, DCT was applied to JPEG images obtained from smart beekeeping system cameras, utilizing the Python programming language with OpenCV and NumPy libraries. The images were sourced from publicly available datasets. The analysis involved applying DCT compression across six defined quality levels: 1, 5, 20, 50, 80, and 100, covering a range that allows for the assessment of the trade-off between compression ratio and image quality. As expected, the lowest quality factor yields the smallest file sizes but results in the lowest image quality. However, it is significant that even with a moderate factor of 20, substantial optimization is achieved — a fivefold reduction in data while maintaining satisfactory image quality for further processing. The optimal result is achieved at a factor of 50, which provides a twofold reduction in data and an output image where quality differences, compared to the original, are indiscernible to the naked eye.**

Keywords- DCT, quality factor, IoT, Python, image compression.

I. Introduction

It is well-established that bees, as the most significant pollinators in the ecosystem, are critical for sustaining life as we know it. However, the last few decades have seen an escalating decline in honeybee colonies worldwide. The phenomenon of bee disappearance is complex and is primarily referred to in the scientific literature as Colony Collapse Disorder (CCD). Scientific studies indicate that the loss of bee colonies is influenced by several factors, including exposure to pesticides, biotic factors, nutritional deficiencies, electromagnetic fields, and climate change [1].

In an effort to understand bee behavior and prevent their further decline, various hive monitoring systems have been developed. These systems are typically based on microcomputers, sensors, and video cameras [2].

Since apiaries are most frequently located beyond the reach of stable power supplies and robust internet connectivity, hive monitoring systems can be classified as resource-constrained IoT ecosystems. These systems typically have limited data storage capacity on internal memory or SD cards. Frequent image generation by the camera quickly exhausts the available storage space. Furthermore, the generated data is usually transmitted wirelessly to a server or the cloud—a process that is often unreliable and slow. Finally, such systems are predominantly battery-powered, and data transmission is one of the most energy-intensive processes [3].

When a camera generates an image, it produces a file of significant size (up to several MB), which directly conflicts with the aforementioned system constraints regarding storage, processing speed, and transmission. To resolve this limitation, the implementation of image compression is essential. By reducing the file size, all three issues are addressed simultaneously: transmission duration is shortened, thereby conserving battery power; data transfer rates are increased; and finally, longer data retention on the device itself is enabled.

Therefore, the compression of images obtained through hive monitoring does not merely represent an option, but a critical trade-off that guarantees the reliable and cost-effective functionality of the hive tracking system.

This paper is a revised and expanded version of the paper presented at the XXV International Symposium INFOTEH-JAHORINA 2026

A. Discrete cosine transform (DCT)

Transform coding represents a key component of modern image and video processing systems. It is based on exploiting the high correlation between adjacent pixels in an image, as well as between pixels in successive video frames, to map the original spatial data into the transform domain, producing decorrelated coefficients. In this manner, entropy is reduced, and more efficient compression is achieved within the source encoder. The Discrete Cosine Transform (DCT) is frequently utilized in contemporary systems for the compression of one-dimensional signals, such as bioimpedance, as well as for image and video data [2].

The DCT removes the correlation between pixels, allowing the resulting coefficients to be compressed independently and more efficiently. The definition of the one-dimensional (1-D) DCT for a sequence of length N is given as [4,5]:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) cos\left[\frac{\pi(2x+1)u}{2N}\right], \qquad (1)$$

$u = 0,1,2, \dots, N-1$, the inverse transform is given as:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) cos\left[\frac{\pi(2x+1)u}{2N}\right], \qquad (2)$$

$$x = 0,1,2, \dots, N-1.$$

In equations (1) and (2), α is defined as:

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases} \qquad (3)$$

The first coefficient of the DCT, for $u = 0$:

$$C(u=0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x) \qquad (4)$$

represents the average value of the sample sequence and is referred to as the DC coefficient, while all other coefficients are called AC coefficients.

The cosine functions used in the DCT form an orthogonal set of basis functions, meaning they are mutually independent: the multiplication of different functions followed by summation yields zero, whereas the multiplication of a function by itself yields a constant value. These functions enable the representation of a signal as a combination of independent components at various frequencies.

The objective of this paper is to examine the efficiency of DCT application to images; therefore, the analysis must be applied to a two-dimensional (2-D) space. The two-dimensional DCT represents a direct extension of the one-dimensional case and is given by:

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v) cos\left[\frac{\pi(2x+1)u}{2N}\right] cos\left[\frac{\pi(2y+1)v}{2N}\right], \qquad (5)$$

$$u, v = 0,1,2, \dots, N-1$$

The inverse transformation is defined as:

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v) cos\left[\frac{\pi(2x+1)u}{2N}\right] cos\left[\frac{\pi(2y+1)v}{2N}\right], \qquad (6)$$

$$x, y = 0,1,2, \dots, N-1$$

It is important to emphasize that the two-dimensional basis functions can be generated by multiplying horizontally oriented one-dimensional basis functions by a vertically oriented set of the same functions, which significantly facilitates their implementation.

B. Fundamental Properties of DCT Relevant to Image Processing Applications

The following section provides an overview of certain DCT properties that are relevant to image processing applications:

- Decorrelation: The primary advantage of image transformation is the removal of redundancy between adjacent pixels, resulting in uncorrelated transform coefficients that can be encoded independently. The low autocorrelation value obtained following the application of the DCT indicates that this transform possesses excellent decorrelation properties.
- Energy Compaction: The DCT effectively concentrates the image energy into a small number of coefficients, allowing for the discarding of low-value coefficients without a perceptible loss in the quality of the compressed image. This property is particularly pronounced in highly correlated images, where the majority of the energy is contained within the low-frequency region.
- Separability: The property of separability allows the 2-D DCT to be computed by applying successive 1-D DCT operations along the rows and columns of the image. This significantly reduces computational complexity and increases efficiency, with the same principle applying to the inverse DCT as well.
- Symmetry: The DCT possesses the characteristic of symmetry because the operations along the rows and columns are functionally identical. This property allows the transformation matrix to be precomputed and utilized in image processing, thereby significantly enhancing computational efficiency.
- Orthogonality: The orthogonality of the DCT allows the inverse transformation to be obtained via matrix transposition. Coupled with its strong decorrelation properties, this further reduces complexity and enhances computational efficiency.

C. Algorithm Implementation

The algorithm for DCT compression was modeled using the Python programming language (version 3.11) [6], along with the

OpenCV [7] and NumPy libraries [8]. OpenCV is a comprehensive library for image processing, while NumPy enables the application of mathematical functions and operations on arrays, specifically the matrices representing the images. Within the OpenCV implementation, a standard JPEG quantization matrix was utilized. This matrix remained unmodified but was automatically scaled according to the quality factor, thereby ensuring compatibility with IoT devices. The analysis was conducted on images extracted from publicly available dataset hosted at https://data.mendeley.com/datasets/8gb9r2yhfc/6 [9]. For the purposes of this research, samples recorded under stable daylight conditions were selected. The focus was placed on the daily activity cycle of the bees, during which the traffic density at the hive entrance peaks and the demand for visual information transmission is most frequent.

The algorithm was applied to the hive images with a resolution of 1920×1080, considering six different quality levels: 1, 5, 20, 50, 80, and 100. The quality level serves as a control variable that influences the quantization step within the DCT compression process. A high quality level reduces the quantization steps, ensuring that the DCT coefficients carrying fine-detail information are preserved. Conversely, a low value for this parameter increases the quantization steps, thereby reducing most high-frequency coefficients to zero. The result is an image that retains only the dominant, low-frequency components forming the basic outlines, while occupying minimal storage space as the resulting sequences of zeros are efficiently compressed. The reconstructed image was then compared to the original; the primary criterion for assessing compression acceptability was the image's utility for further processing — specifically, the ability to accurately detect bees. Images where quantization did not cause object merging or a loss of edge sharpness were considered acceptable.

II. Results and Discussion

The analysis was conducted on four beehive images captured in a natural environment. The main difference between the images relates to the variation in bee density; additionally, differences in the coloration of the hive itself were observed. All images have a resolution of 1920×1080 pixels, with original file sizes of 907 KB, 1 020 KB, 1 085 KB and 1 093 KB, respectively.

In the first phase of the research, an analysis of the impact of quality levels on file size was conducted. Observing the results presented in Table I, it can be concluded that the quality level and the compression ratio are inversely proportional. The lowest quality level enables the highest compression ratio — resulting in the smallest output file size — and vice versa.

TABLE I. JPEG compressed file sizes as a function of quality level (Q = 1–100) for test image Hive A. The original input image size is 907 KB.

Quality level	Input image size[KB]	Output image size [KB]	Compression percentage[%]
1	907	30.85	96.60
5		60.20	93.36
20		184.76	79.63
50		276.21	69.55
80		433.41	52.21
100		1199.44	—

The same conclusion can be drawn based on the results obtained for the remaining three images included in the analysis, which are presented in Tables II, III, and IV.

TABLE II. JPEG compressed file sizes as a function of quality level (Q = 1–100) for test image Hive B. The original input image size is 1020 KB.

Quality level	Input image size[KB]	Output image size [KB]	Compression percentage[%]
1	1 020	38.20	96.25
5		73.14	92.83
20		212.20	79.20
50		313.88	69.23
80		488.62	52.09
100		1353.04	—

TABLE III. JPEG compressed file sizes as a function of quality level (Q = 1–100) for test image Hive C. The original input image size is 1085 KB.

Quality level	Input image size[KB]	Output image size [KB]	Compression percentage[%]
1	1 085	45.2	95.83
5		81.7	92.47
20		230	78.80
50		341.56	68.52
80		525.32	51.58
100		1476.28	—

TABLE IV. JPEG compressed file sizes as a function of quality level (Q = 1–100) for test image Hive D. The original input image size is 1093 KB.

Quality level	Input image size[KB]	Output image size [KB]	Compression percentage[%]
1	1 093	45.57	95.83
5		82.34	92.47
20		232.36	78.74
50		344.39	68.49
80		529.97	51.51
100		1487.50	—

The results obtained for the Hive A image were used to analyze the significance of applying compression to improve the efficiency of digital image storage. The same analysis was applied to the results of the other images as well. The following graph illustrates the number of images that can be stored in 1 MB of storage space depending on the compression level:

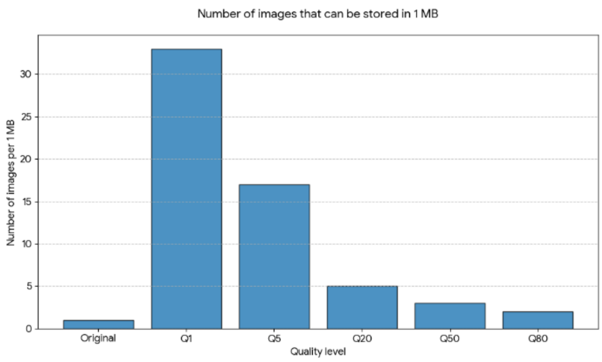

Figure 1. Dependence of the number of images that can be stored in 1 MB of memory on the quality level (Q = 1–80) when applying DCT compression to the test image of Hive A.

While the original image allows for the storage of approximately one image per 1 MB of memory, moderate quality levels (20–50) enable the storage of 3–5 times more images with an acceptable loss of quality. Extreme quality levels do not provide memory efficiency and are not practical for application. Therefore, compression is essential in systems for processing and archiving large amounts of image data.

In the second phase, it was necessary to determine the extent to which the achieved memory savings affect the degree of degradation in the compressed images.

Fig. 2 and 3 show the original image of Hive A and the output images obtained by applying different quality levels. Based on a subjective visual assessment, it can already be concluded that if image quality is to be preserved in terms of detail clarity, the quality level must be set high. If the quality level is low, the reconstructed image will also be of lower quality.

Figure 2. Original image of Hive A used as a reference for evaluating the impact of DCT compression on image size.

Figure 3. Comparison of compressed image versions at different quality levels (Q); lower Q values cause blocking artifacts and loss of sharpness, while higher values preserve better visual fidelity with reduced image size.

Setting the quality parameter to a low value (1 or 5) inevitably leads to the appearance of pronounced artifacts in the output image. A moderate quality level Q=20 produces a reconstructed image of satisfactory visual quality, while reducing its size by approximately five times compared to the original. A factor of 80 results in approximately half the original data size and an output image in which quality differences, relative to the original, are imperceptible to the naked eye.

Fig. 4 to 9 present a comparative overview of the original and compressed images of hives B, C, and D. They further confirm the uniformity of the presented results, as well as the consistency of the previously drawn conclusions regarding the effectiveness of the applied transformation.

Figure 4. Original image of Hive B used as a reference for evaluating the impact of DCT compression on image size.

Figure 5. Comparison of compressed image versions at different quality levels (Q); lower Q values cause blocking artifacts and loss of sharpness, while higher values preserve better visual fidelity with reduced image size.

Figure 6. Original image of Hive C used as a reference for evaluating the impact of DCT compression on image size.

Figure 7. Comparison of compressed image versions at different quality levels (Q); lower Q values cause blocking artifacts and loss of sharpness, while higher values preserve better visual fidelity with reduced image size.

Figure 8. Original image of Hive D used as a reference for evaluating the impact of DCT compression on image size.

Figure 9. Comparison of compressed image versions at different quality levels (Q); lower Q values cause blocking artifacts and loss of sharpness, while higher values preserve better visual fidelity with reduced image size.

Furthermore, in order to eliminate observer subjectivity and obtain precise data on information preservation within the image, the final phase involved an objective analysis using mathematical metrics. While subjective assessment verifies the practical usability of compressed images for further processing, objective metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) mathematically evaluate the extent to which the applied compression degrades key image details.

A. PSNR

The PSNR parameter measures the ratio between the maximum signal power and the corrupting noise power caused by quantization. It is calculated using the following equation:

$$PSNR = 10\ \log \frac{(L-1)^2}{\frac{1}{M}\frac{1}{N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[f(x,y)-\hat{f}(x,y)\right]^2} \qquad (7)$$

where L-1 represents the maximum possible pixel intensity. The product $\frac{1}{M}\frac{1}{N}$ denotes the total number of pixels in the image, while f(x,y) and $\hat{f}$(x,y) represent the pixel values of the original and compressed images at the same positions [10].

PSNR is expressed in decibels (dB), and typically ranges from 20 to 50 dB, where 50 dB indicates an almost perfect image. Values below 25 dB indicate visible degradation. While PSNR provides a reliable measure of absolute pixel-level error, it often fails to reflect how degradation is perceived by the human eye.

B. SSIM

The SSIM metric simulates human visual perception by relying on the fact that certain visual information holds lower importance than other information during visual processing in the human brain. The essence of this metric is based on the premise that digital images are characterized by a high degree of spatial redundancy, meaning that the values of individual pixels are strongly correlated with their surroundings and can be accurately predicted based on neighboring elements. Due to this interdependence, each individual pixel carries a small amount of information, which enables efficient compression without significantly degrading the perception of the overall image structure. Therefore, SSIM serves as a complementary metric to PSNR, specifically focusing on structural changes rather than just pixel-to-pixel differences [11].

The process of calculating the SSIM index is performed by comparing three key components: luminance, contrast, and image structure. Instead of analyzing the entire image at once, the procedure is applied to 8×8 pixel blocks, which enables precise detection of degradation occurring during image reconstruction. The final result is obtained by combining these parameters into a single index value, calculated as follows:

$$SSIM\ (x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \qquad (8)$$

where μ_x and u_y are the mean luminance values, σ_x and σ_y,are the standard deviations (i.e., contrast), and σ_{xy} is the covariance between images x and y.

The resulting index always ranges from 0 to 1 and represents a quality score that aligns with subjective human perception. The closer the value is to 1, the more visually similar the compressed image is to the original.

The results of the calculated metrics for the tested set of beehive images are presented in Tables V–VIII.

TABLE V. SSIM AND PSNR PARAMETER VALUES FOR DIFFERENT QUALITY LEVELS (Q) OF COMPRESSED IMAGES – HIVE A.

	Quality level	SSIM	PSNR
Hive A	1	0.5968	19.69 dB
	5	0.7715	22.48 dB
	20	0.9191	28.04 dB
	50	0.9714	33.01 dB
	80	0.9866	36.34 dB
	100	0.9992	49.04 dB

TABLE VI. SSIM AND PSNR PARAMETER VALUES FOR DIFFERENT QUALITY LEVELS (Q) OF COMPRESSED IMAGES – HIVE B.

	Quality level	SSIM	PSNR
Hive B	1	0.5655	20.66 dB
	5	0.7596	23.28 dB
	20	0.9173	28.82 dB
	50	0.9708	33.89 dB
	80	0.9861	37.20 dB
	100	0.9991	49.95 dB

TABLE VII. SSIM AND PSNR PARAMETER VALUES FOR DIFFERENT QUALITY LEVELS (Q) OF COMPRESSED IMAGES – HIVE C.

	Quality level	SSIM	PSNR
Hive C	1	0.6145	19.15 dB
	5	0.7857	21.94 dB
	20	0.9251	27.59 dB
	50	0.9743	32.37 dB
	80	0.9880	35.68 dB
	100	0.9994	47.21 dB

TABLE VIII. SSIM AND PSNR PARAMETER VALUES FOR DIFFERENT QUALITY LEVELS (Q) OF COMPRESSED IMAGES – HIVE D.

	Quality level	SSIM	PSNR
Hive D	1	0.6129	19.17 dB
	5	0.7847	21.83 dB
	20	0.9250	27.56 dB
	50	0.9741	32.31 dB
	80	0.9881	35.64 dB
	100	0.9994	47.10 dB

The obtained PSNR and SSIM values are strongly correlated with the subjective visual assessment, confirming that a decrease in these metrics is accompanied by a noticeable degradation of image details across all compression levels.

III. CONCLUSION

In this paper, a data compression method based on the two-dimensional DCT was presented and applied to images obtained from cameras in smart beekeeping systems.

The experimental results confirm that the quality level and compression ratio are inversely proportional, whereby lower quality levels enable significant memory savings, but with noticeable degradation of visual details. The objective metrics PSNR and SSIM show a high correlation with the subjective visual assessment, thereby confirming their reliability in evaluating the quality of compressed images. A particular contribution of this study is the identification of an "optimal point" at quality level Q = 50, where an average space saving of approximately 69% is achieved, while maintaining high visual integrity (SSIM > 0.97 and PSNR ~ 33 dB). The obtained results indicate that the application of compression is essential in systems designed for processing and archiving large amounts of image data, requiring a careful selection of quality parameters in accordance with the requirements of a specific application.

The consistency of these results is documented through comparative analyses of the four beehives (A, B, C, and D), demonstrating that the applied method maintains uniformity and stability across the entire test image set.

The proposed DCT compression method represents a promising solution for data management and transmission in systems that primarily rely on the visual monitoring of beehives.

However, the limitations of the proposed approach are most evident at low quality levels (Q < 20), where pronounced artifacts occur. Although such aggressive compression provides space savings of more than 90%, it leads to the loss of fine details, which may negatively impact algorithms designed for automatic bee counting or varroa detection. Furthermore, this study is limited to stable daylight conditions, while the system's performance under extremely low-light conditions or when using infrared cameras remains a subject for future work.

In addition, future research will focus on the development of adaptive quantization matrices that dynamically adjust to the image content, thereby processing the hive background differently from the bees themselves.

ACKNOWLEDGMENT

The authors express their gratitude to Prof. Mitar Simić of the Faculty of Electrical Engineering, University of Banja Luka, for the valuable discussions during the experimental work, as well as throughout the preparation of the paper.

REFERENCES

[1] G. Singh and A. Rana, "Honeybees and colony collapse disorder: understanding key drivers and economic implications," Indian National Science Academy, 2025. [Online]. Available: https://www.researchgate.net/publication/389138949_Honeybees_and_colony_collapse_disorder_understanding_key_drivers_and_economic_implications.

[2] M. Simić and T. J. Freeborn, "Discrete Cosine Transform-based Data Compression of Bioimpedance with Embedded Hardware," 24th International Symposium INFOTEH-JAHORINA, March 2025.

[3] V. Starčević, M. Simić, V. Risojević and Z. Babić, "Integrated video-based bee counting and multi-sensors platform for remote bee yard monitoring," 21st International Symposium INFOTEH-JAHORINA, March 2022.

[4] O. Anwar, "Smart beehive monitoring for remote regions," Ph.D. dissertation, Dept. Electron. Eng., 2022, pp. 31–33. [Online]. Available: https://www.researchgate.net/publication/365265478_Smart_beehive_monitoring_for_remote_regions

[5] S. A. Khayam, "The Discrete Cosine Transform (DCT): Theory and Application," March 10th 2003.

[6] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*, Scotts Valley, CA: CreateSpace, 2009.

[7] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, vol. 25, no. 11, pp. 120-123, 2000.

[8] C. R. Harris et al., "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020. doi: 10.1038/s41586-020-2649-2

[9] Sledevic, Tomyslav (2024), “Labeled dataset for bee detection and direction estimation on beehive landing boards”, Mendeley Data, V6, doi: 10.17632/8gb9r2yhfc.6

[10] R. C. Gonzalez and R. E. Woods, “Digital image processing,” 2nd Edition,2022.

[11] Wang et al., 2004, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.

Nevena Jeftenić, holds an MSc in Electrical Engineering and currently serves as a Lead Engineer for Mediation and Anti-Fraud Support at M:tel. Her research focuses on the Internet of Things (IoT), an area in which she has authored several scientific papers.

Tamara Rašić, MSc in Electrical Engineering, is currently employed at M:tel as Head of the Database and Applications Department. Alongside her career in the telecommunications industry, she is engaged in scientific research and has authored several scientific papers.