VISUAL DATA COMPRESSION FOR MACHINE VISION APPLICATIONS

Omar M. Salih[1], József Vásárhelyi[2]
[1]PhD Student, [2]Associate Professor
[1,2]Institute of Communication and Info-Communications,
Faculty of Mechanical Engineering and Informatics
[1]omar.salih@uni-miskolc.hu, [2]jozsef.vasarhelyi@uni-miskolc.hu

ABSTRACT

This paper addresses the challenges of computer vision and highlights the importance of visual data compression due to the limitations of real-time processing and large data volumes. This research proposes a lossy-based compression method to seamlessly integrate it into image processing applications, optimize memory and improve real-time performance.

1. INTRODUCTION

Machine vision has become a transformative force in recent years, revolutionizing how computers perceive and interpret visual information. As an intersection of computer science, artificial intelligence, and image processing, machine vision empowers machines to replicate and boost human visual capabilities. This capability enables computers to understand, analyze, and respond to visual data, laying the foundation for groundbreaking applications across various industries [1]. While machine vision has recorded significant success, it still has problems, such as real-time processing requirements and massive data size [2]. Addressing these challenges presents opportunities for innovation and further advancements in the field. Visual data compression within machine vision applications is crucial due to the difficulties in efficiently handling extensive data. Many machine vision systems utilize high-resolution images or video recordings to recognize sophisticated details and to make accurate decisions. However, the data generated by these sources is so large that it poses a significant computational and storage burden [3]. In real-time processing, autonomous systems, and remote sensing, the massive stream of raw visual data can consume computational resources, leading to a loss of response and a lack of real-time performance [4].

Visual data compression is a critical component of solving these problems by decreasing the size of images or videos while leaving important information unchanged. This procedure involves reducing the size of images while retaining relevant information, optimizing the storage and transmission of information, and improving real-time processing [5]. Two primary approaches of image compression are lossless and lossy. Lossless compression retains the original data during the compression-decompression cycle, ensuring a perfect reconstruction of the original data. In contrast, lossy compression involves sacrifices of details to achieve higher compression levels. Less important information is discarded to reduce the size of the files [6]. The contribution of this paper is to present a lossy-based image compression method to be successfully integrated within machine vision applications without adversely affecting the application performance.

2. A CASE STUDY: THREE-DIMENSIONAL RECONSTRUCTION

Three-dimensional reconstruction is a fundamental process in machine vision, transforming 2D images into 3D models that accurately represent the spatial arrangement of objects. The basic idea behind 3D reconstruction is to create a volumetric representation from a succession of 2D images or frames. The complexity of 3D reconstruction involves massive computational resources and efficient processing of large-scale visual data [7]. Image compression has proven to be a key component in solving storage and transmission bandwidth challenges. However, the unique properties of 3D data, such as depth maps and point clouds, bring new challenges and opportunities for customized compression strategies. Adaptive compression strategies are crucial in dynamic environments where scenes change over time, balancing the trade-off between compression ratio and temporal fidelity. Real-time applications such as robotics and augmented reality can significantly benefit from compression techniques that adapt to the dynamic nature of visual data [8].

3. METHODOLOGY

As the interplay between image compression and 3D reconstruction continues to be revealed, exploring novel algorithms, adaptive techniques, and real-time implementations provides affluent prospects for further research. Addressing challenges related to varying scene complexity, compression artefacts, and computational efficiency is crucial to improving the applicability of image compression in 3D reconstruction scenarios. This paper presents a developed JPEG method to address the limitations of a JPEG compression scheme. JPEG's inherently lossy compression can produce artefacts that affect the accuracy of 3D reconstructions. In multi-view 2D images, the loss of fine details can lead to inaccurate 3D reconstruction, especially in geometrically complex areas. As illustrated in

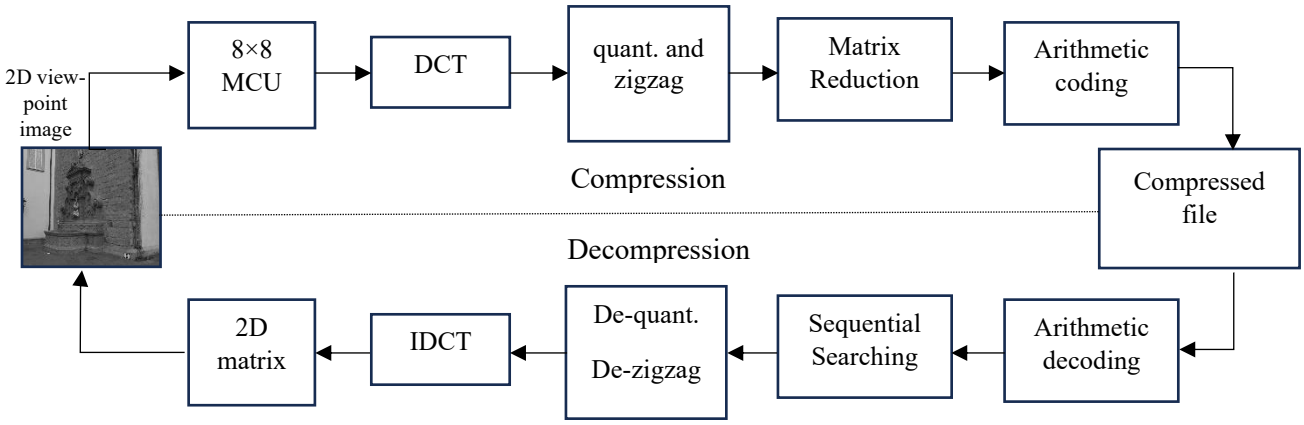


Fig. 1.
The proposed compression and decompression workflow

Fig 1, the proposed method starts by applying the Discrete Cosine Transformation (DCT) for every 8×8 Minimum Coded Unit (MCU), followed by a uniform quantization process.

Afterwards, the quantized values from all MCUs are zigzag scanned and saved in a 1D array. The matrix reduction [9] is applied to encode the data using three randomly generated keys and compresses every three quantized values into a single encoded value. The arithmetic coding is then applied, and the data are converted to a compressed binary stream. At the decoding phase, this process is performed in reverse order. The compressed data is firstly transformed using Inverse Discrete Transform (IDCT), de-quantized, and applied to sequential search to retrieve the decoded data using the duplicate generated keys in the coding stage. The retrieved data are then arithmetically decoded to obtain the reconstructed image pixels. Fig 2 shows the block diagrams of the matrix reduction and sequential search performed in the compression and decompression stages. The symbols k1, k2, and k3 refer to the generated three keys. While CD_i and ED_i represent the i^{th} coefficient data and encoded data, respectively.

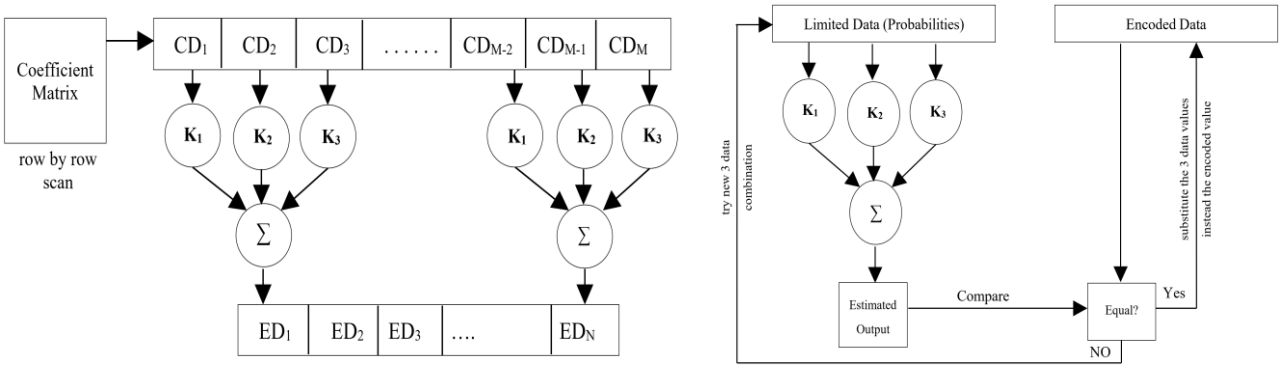


Fig. 2.
Matrix reduction block diagram (left) and sequential search block diagram (right)

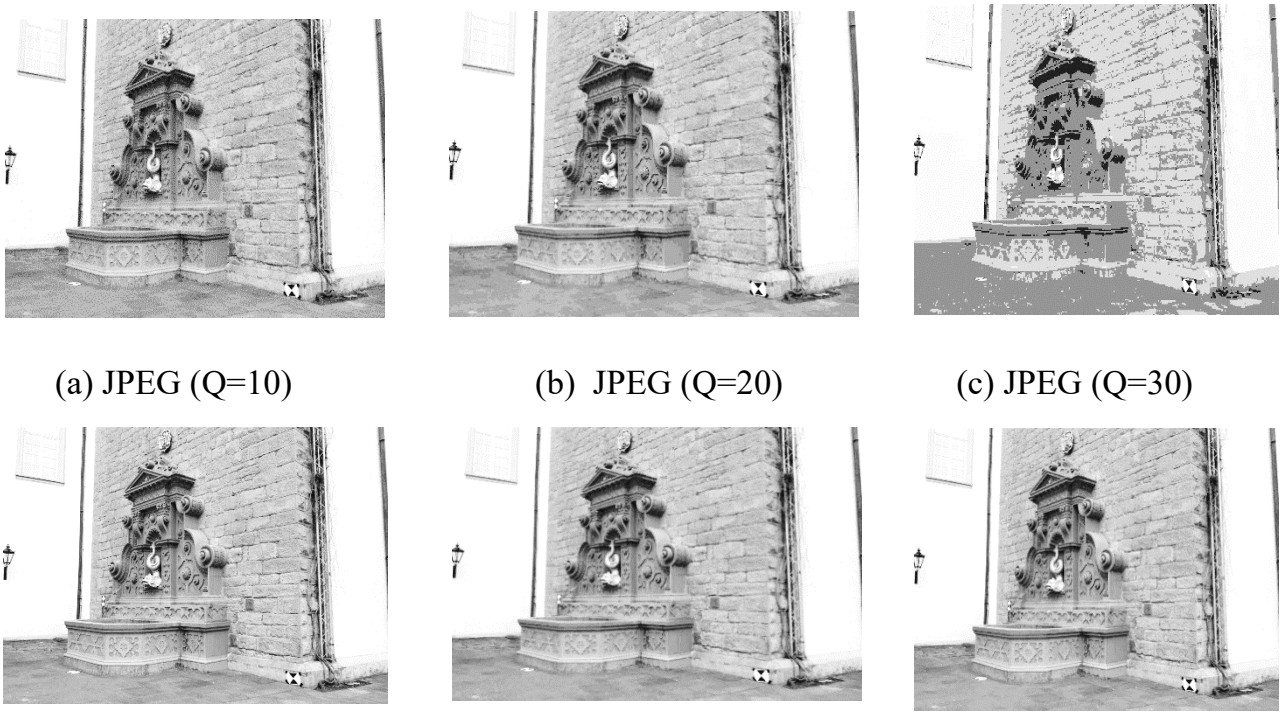
4. EXPERIMENTAL RESULTS

The demonstrated method has been applied to a series of 11 multiple viewpoints grayscale 2D images to reconstruct the 3D representation. The technique has been developed in MATLAB 2021b, and the computer specification is CPU Intel core i7-12700H speed 2.3 GHz and 32 GB of RAM. The results are compared to the JPEG compression through different evaluation metrics. Table 1 compares the performance of the two methods for the first three viewpoint images. To offer a comprehensive evaluation framework, two widely used metrics for this purpose are considered: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR is a quantitative measure that assesses the fidelity of the compressed image concerning the original image. The higher the PSNR, the better preservation of image quality. SSIM evaluates the structural information and perceived quality of close images by comparing local patterns of pixel intensities.

Table 1.

Compression results as a comparison of the proposed method to JPEG compression

Original Size MB	Dim.	Multi-view images	Quant. value	JPEG				Proposed method			
				Comp. size KB	CP %	SSIM	PSNR	Comp. size KB	CP %	SSIM	PSNR
6	3072 × 2048	View 1	20	209.3	96.5	0.4711	42	211	96.5	0.6230	43
			30	166.2	97.3	0.3447	40	162.7	97.4	0.5206	42.2
			40	142.5	97.6	0.1956	37.9	138.4	97.7	0.4542	41.6
		View 2	20	231.5	96.2	0.5354	41.9	228.3	96.3	0.6633	43.1
			30	179.5	97	0.3868	40.2	176.3	97.1	0.5674	42.3
			40	152.3	97.5	0.2192	38	149.7	97.6	0.5021	41.6
		View 3	20	242.1	96	0.5715	42	235.5	96.2	0.6821	43.5
			30	179.5	79	0.3961	40.5	181.5	97	0.5850	42.6
			40	159.1	97.4	0.2219	38.3	154.4	97.5	0.5174	42

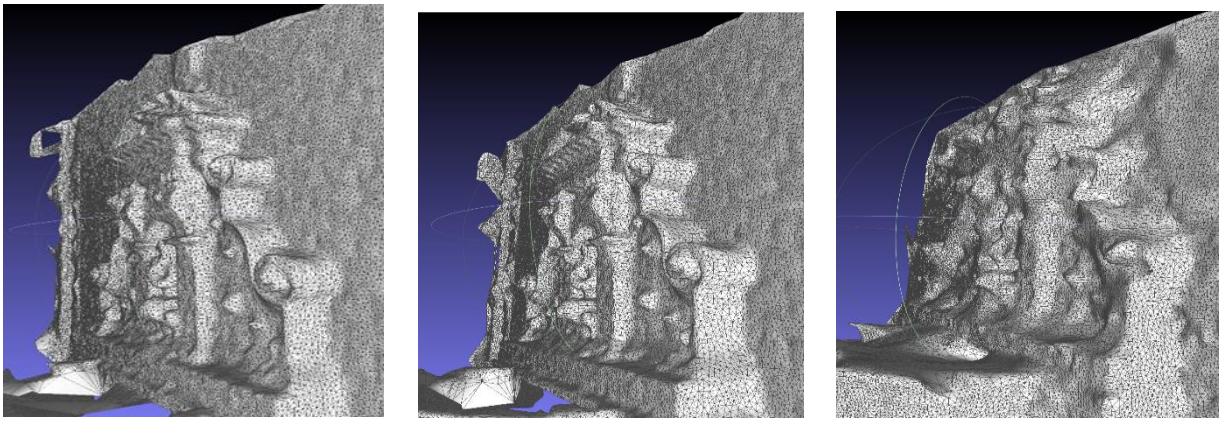


(a) JPEG (Q=10) (b) JPEG (Q=20) (c) JPEG (Q=30)

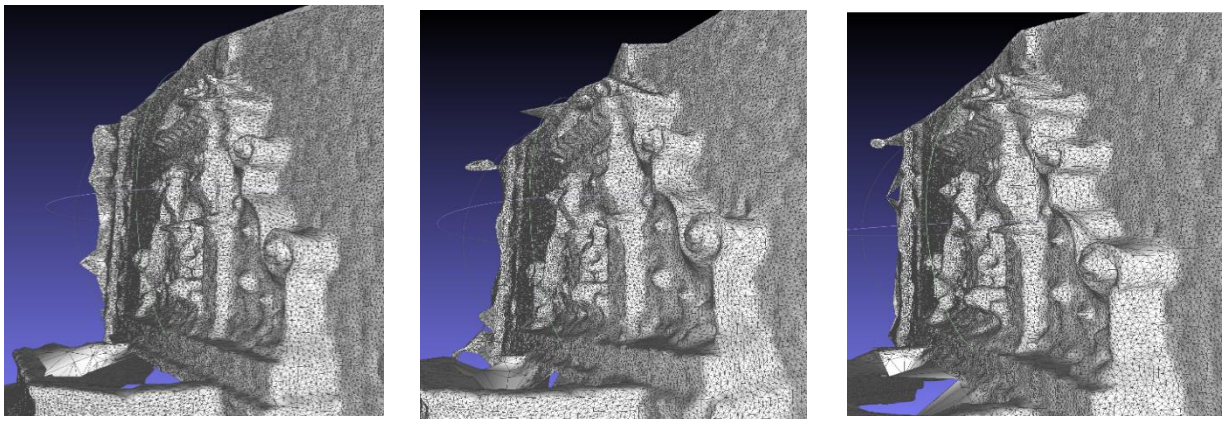
(f) Proposed (Q=30) (d) Proposed (Q=10) (e) Proposed (Q=20)

Fig. 3.

The first viewpoint image is compressed using JPEG (a-c) and the proposed method (d-f) SSIM ranges between 0 to 1. The closer the value is to 1, the better the perceptual image quality. The table also compares the image compressed size and compression percentage (CP). For each tested image, three quantization values are used to show the compression performance in high compression ratios. The results show that the proposed method overcomes JPEG compression through the described metrics and can perceive the image quality while increasing quantization. JPEG showed a significant decrease in decompressed image quality when increasing quantization. Fig 3 illustrates the decompressed images of the first viewpoint image using the two methods. The visualization results can be shown in Fig 4 of the 3D reconstructed object using the decompressed viewpoint images.



(a) JPEG (Q=10) (b) JPEG (Q=20) (c) JPEG (Q=30)



(d) Proposed (Q=10) (e) Proposed (Q=20) (f) Proposed (Q=30)

Fig. 4.
The visualization views of the reconstructed 3D object using JPEG (a-c) and the proposed method (d-f)

As can be seen, the 3D reconstructed objects obtained from the proposed method offer perceiving details' accuracy during different compression rates. In contrast, the 3D reconstructed object by JPEG compression shows the degrading of the object features and missing information as the compression increases.

5. CONCLUSION AND FUTURE WORK

This paper has presented and demonstrated a 2D multi-viewpoint image compression method and evaluated it in the 3D reconstruction application. The results are compared to the well-known JPEG compression standard to show its performance and efficiency. The findings collectively affirm the efficacy of the proposed compression method for 3D reconstruction, emphasizing its ability to preserve image quality and 3D object details, particularly under high compression ratios. The demonstrated advantages underscore the

potential practical applications of the proposed method in scenarios where efficient compression and accurate 3D reconstruction are paramount. The proposed process's efficient implementation and speeding up are required for future work since it adds additional computational load and is more time-consuming.

REFERENCES

[1] SMITH, M. L., SMITH, L. N. and HANSEN, M. F.: **The quiet revolution in machine vision- a state-of-the-art survey paper, including historical review, perspectives, and future directions.** Computers in Industry. 2021 Sep 1; 130:103472.

[2] WANG, S, WANG, Z., WANG, S. and YE Y.: **End-to-end compression towards machine vision: Network architecture design and optimization.** IEEE Open Journal of Circuits and Systems. 2021 Nov 23; 2:675-85.

[3] CAO, J., Yao, X., ZHANG, H., JIN, J., ZHANG, Y. and LING, B. W.: **Slimmable Multi-Task Image Compression for Human and Machine Vision.** IEEE Access. 2023 Mar 24; 11:29946-58.

[4] BOUDERBAL, I., AMAMRA, A. and BENATIA, M. A.: **How would image down-sampling and compression impact object detection in the context of self-driving vehicles?.** In International Conference on Computing Systems and Applications 2020 Dec 14 (pp. 25-37). Cham: Springer International Publishing.

[5] ZHANG, Q., WANG, S., ZHANG, X., MA, S. and GAO, W.: **Just recognizable distortion for machine vision oriented image and video coding.** International Journal of Computer Vision. 2021 Oct;129(10):2889-906.

[6] ZHANG, Y.: **A Rate-Distortion-Classification approach for lossy image compression.** Digital Signal Processing. 2023 Sep 1; 141:104163.

[7] SALIH, O. M., RASHEED, M. H., SIDDEQ, M. M. and RODRIGUES, M. A.: **Image compression for quality 3D reconstruction.** Journal of King Saud University-Computer and Information Sciences. 2022 May 1;34(5):2271-87.

[8] GOLTS, A. and SCHECHNER, Y. Y.: **Image compression optimized for 3D reconstruction by utilizing deep neural networks.** Journal of Visual Communication and Image Representation. 2021 Aug 1; 79:103208.

[9] SALIH, O. M., ROSTUM, H. M., VÁSÁRHELYI, J. and SIDDEQ, M. M.: **Fast Joint Image Compression-Encryption Algorithm used for 3D Reconstruction.** In2023 24th International Carpathian Control Conference (ICCC) 2023 Jun 12 (pp. 400-405), IEEE.