



# Compression of High-Resolution Space Video Using Phase Grating

Anirban Patra<sup>1,3</sup>  · Arijit Saha<sup>2</sup> · Kallol Bhattacharya<sup>3</sup>

Received: 3 February 2023 / Accepted: 4 August 2023 / Published online: 28 August 2023  
© Indian Society of Remote Sensing 2023

## Abstract

Compression of high-resolution space video is one of the important concerns in many remote sensing application areas. As high-resolution video carries a lot of valuable information (which is required for detailed analysis), therefore compression of these videos is a challenging task. Many existing video compression algorithms are not applicable for high-resolution remote sensing and medical video due to too much information loss. Here, a fast, optical processing (phase grating)-based video compression technique has been discussed. The selected video is divided into frames initially. Each video frame is further divided into three planes—red, green, and blue which are to be modulated using phase grating. As a result of the modulation, three spectrums are generated for each frame. Only one spectrum is taken into consideration and sequentially placed on the canvas. At the receiving end, by applying inverse Fourier transform at selected spots of the canvas, planes are extracted and frames are reconstructed using the extracted planes. For quality checking, Peak Signal to Noise Ratio (PSNR) and correlation coefficient (to measure the closeness with original images) methods are used. The entire process is completed in 0.838 s (for 60 fps space video); hence, it can be expressed as a fast process. So, compared with a few existing video compression algorithms, this system has advantages like low noise affection, fast processing, etc.

**Keywords** Video processing · Phase grating · Inverse Fourier transform · PSNR · Correlation coefficient

## Introduction

The use of high-resolution space video in the research area is increasing day by day for analytical purposes in remote sensing work. As this type of video consumes a lot of different types of vital information, hence special care must be taken to compress the video. For a long time, researchers have been working on the compression of images and videos for different purposes. Existing algorithms are divided into two classes: lossless and lossy. Lossless compression is preferred for analytical purposes but few lossy compression algorithms provide better results in high-resolution image processing also.

The reversible color transforms with a double-embedded context model for lossless image compression is one of the

oldest and simplest algorithms (Gormish et al., 1997). Here, researchers claim that this system is suitable for lossless as well as nearly lossless compression, but the output images are highly noise affected. A modified version of the reversible transform is applied to eliminate the redundancy in both spatial as well as spectrum domains (Zhang & Wu, 2006). Compared to the previous system, this system takes less time. The integer form of the Tchebichef Transform (Xiao et al., 2015) is suitable for a different type of low-resolution image (including color images) for lossless compression. For real-time applications using hardware implementation to compress remote sensing images, an onboard FPGA (Field Programmable Gate Array) system with a CCD line (Rusyn et al., 2016) performs well. A different form of high-efficiency video coding (HEVC) is utilized to compress the videos by maintaining their quality. In general, HEVC is suitable for lossy video compression but a modified version of it can be utilized for lossless compression also (Choi & Ho, 2015). A new verification method of HEVC shows its effectiveness of it compared to other existing methods (Tan, 2016). Joint Exploration Model (JEM) is applied nowadays to enhance the capacity of HEVC in video compression (Chen et al., 2020). Earlier high-volume images were used vocationally; hence, lossless

✉ Anirban Patra  
anitublu@gmail.com

<sup>1</sup> Department of ECE, JIS College of Engineering, Kolkata, West Bengal, India

<sup>2</sup> Department of ECE, BP Poddar Institute of Management and Technology, Kolkata, West Bengal, India

<sup>3</sup> Department of Applied Optics and Photonics, University of Calcutta, Kolkata, West Bengal, India

image compression was applied to them. With the increase in high-volume image and video applications, a lot of lossy compression algorithms are proposed by researchers. Out of them, the prediction method is widely used for image as well as video compression due to the advantages of the simple encoding and decoding process (Ayoobkhan, Chikkannan, & Ramakrishnan, 2017). Huffman coding, a popularly known lossless image compression process, is used to compress medical videos after modification of their histogram (Rahman et al., 2018). The recent incorporation of neural networks in the modified prediction system makes it more appropriate for image compression (Bowen et al., 2021). Apart from this system, the neural network is also used in other video compression algorithms too (Ma et al., 2020). In a deep neural network (DNN)-based system, texture-based video coding is used to process raw video with the help of data-driven learning (Ding et al., 2021). In neural network-based video coding systems, performance can be improved either by improving the video codecs or by using a holistic method in end-to-end video compression processes (Birman & Segal, 2020). Spatial and temporal redundancy in video compression can be reduced using convolution neural networks in post-processing and pre-processing systems (Gorodilov et al., 2018).

In the optical domain, the sinusoidal grating is applied to compress medical as well as remote sensing images by using high spatial frequency. Due to the modulation, multiple spectrums are generated. Each spectrum contains image information; hence, only one spectrum is sufficient to extract the image. As only one spectrum is used for extraction, therefore, size of the extracted image is less than the original image (Patra et al., 2019, 2021). The efficiency of amplitude grating is low. Hence, to improve it, the authors have applied phase grating for modulation (Patra et al., 2020, 2022).

Recently, a modified version of the genetic algorithm (GA) of deep learning is used to compress a large data set which can optimize the utilization of storage space (Skandha et al., 2022). One of the main challenging tasks of long-distance image or video communication is to maintain security. Recently, different versions of hashing algorithms are used to protect the reconstructed image information so that it can't be tampered with easily (Shaik et al., 2022a, 2022b). Blockchain is also a powerful tool for the protection of valuable information (Bathula et al., 2022, 2023).

The main objective of this paper is to utilize optical processing (phase grating) in remote sensing video to compress it by maintaining proper quality so that it can be used for accurate analysis. Here, a few drawbacks of existing algorithms like noise affection, time-consuming, etc., are overcome.

Amplitude grating is not selected due to its poor diffraction efficiency (6.25% of the total optical power only) (Harvey & Pfisterer, 2019). The entire process is divided into several steps. Initially, the selected video is divided into thirty frames. Each frame is divided again into three planes—red, green, and blue, respectively. In the next process, each plane is modulated using a high grating frequency. Multiple spatial domain spectrum (only three properly visible) is generated due to modulation, out of which only one spectrum will be taken into consideration for further processing. The selected spatial domain spectrum is placed sequentially into a canvas. During the retrieval process, Inverse Fourier Transform is applied at selected areas to extract the frames. Extracted frames are joined together to reconstruct the video. To measure the quality of the frames, Peak Signal to Noise Ratio (PSNR) and correlation coefficient (to measure the closeness with original images) methods are used. A comparison study with previous methods has also been discussed.

## Methodology

In this section, the entire process is categorized into three parts: Modulation, Retrieval, and Quality Measurement.

**Modulation:** During the modulation process, using high spatial frequency (1000 lines/cm), video frames are modulated at 0 deg. orientation angle by phase grating. As a result of modulation, three visible spectrums are generated spatial domain.

Let us assume that  $f_1(x, y)$  represents 1st video frame which is selected to be modulated by phase grating of the form  $\exp\left[i\left(\frac{m}{2}\right)\sin 2\pi u_1 x\right]$ , where  $u_1$  and  $m$  represent the grating frequency and phase contrast, respectively.

Modulated object  $s_1(x, y)$  is given by

$$s_1(x, y) = f_1(x, y) \exp\left[i\left(\frac{m}{2}\right)\sin 2\pi u_1 x\right] \quad (1)$$

Therefore, after convolution, FT of the modulated object  $S_1(x, y)$  may be expressed as

$$\begin{aligned} S_1(u, v) &= [F_1(u, v) \otimes \left[ \sum_{q=-\alpha}^{q=+\alpha} J_q\left(\frac{m}{2}\right) \delta(u - qu_1, v) \right]] \\ &= \left[ \sum_{q=-\alpha}^{q=+\alpha} J_q\left(\frac{m}{2}\right) F_1(u - qu_1, v) \right] \end{aligned} \quad (2)$$

where  $J_q$  and ' $q$ ' represents the  $q$ th order Bessel Function and diffraction order, respectively.

Equation 2 indicates the diffraction pattern which consists of multiple spots which the video frame contains. Therefore, Eq. (2) may be rewritten by considering the zero-order and the first two orders of the spectrum as

$$S_1(u, v) = \left[ J_0\left(\frac{m}{2}\right) F_1(u, v) + J_{-1}\left(\frac{m}{2}\right) F_1(u + u_1, v) + J_{+1}\left(\frac{m}{2}\right) F_1(u - u_1, v) \right] \quad (3)$$

Figure 1 displays the spectrum of modulated Red plane of frame  $f_1(x, y)$ .

To modulate other frames, the same process is applied.

From the modulated spectrum plane which is shown in Fig. 1, only one spectrum is selected. In Fig. 2, the selected spectrum is marked by the Arrow sign.

In this process, a 60 FPS (Frame per Second) video is used for the research work. The selected video is divided into 60 frames initially. Each frame is divided into three planes—red, green, and blue, respectively. Hence, 180 spectrums are collected for transmission. All the spectrums are automatically placed in a sequential manner on the canvas.

The placed spectrum is displayed in Fig. 3

## Retrieval Process

To recover the original video frames, the inverse Fourier transform is used in selected zonal spots. The pixel resolution of each selected spot is  $1000 \times 700$ .

Using high-frequency grating, each video frame's red, green, and blue planes are modulated. During this modulation, three spectra (band or spot) are generated for each plane. The central band is the Fourier transform of the particular plane and the others are two sidebands (lower sideband and upper sideband). Each sideband contains the same amount of information; hence, only one sideband is considered for further processing.

During the initial stage of the modulation, one plane is selected for the video frame from the Red, Green, or Blue. The plane is modulated, and as a result, three bands are generated. The upper sideband is selected (the lower sideband also provides the same result) and it is placed in the predefined position of the final canvas. The same process is repeated for other planes of the same video frame.

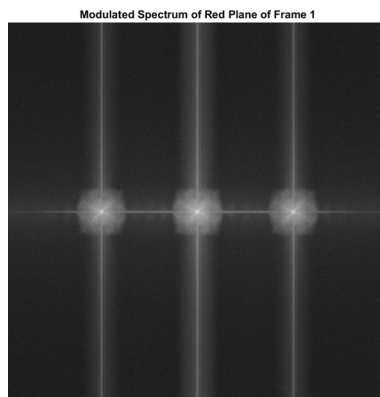


Fig. 1 Modulated spectrum of red plane of frame 1

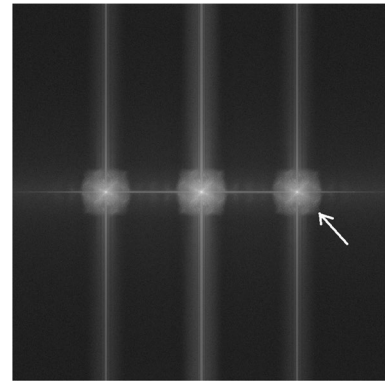


Fig. 2 The selected spectrum marked by arrow sign

**Automatic Selection:** As a fixed grating frequency is maintained for the modulation of all planes, the sideband's position in the entire spectrum is fixed in all cases. A pixel intensity graph is used to identify the location of the spectral component corresponding to the image. The highest intensity of the sidebands is 80% of the central order intensity. A low pass filter with a cutoff (radius) of 500 pixels (which is found to be adequate for high-resolution image

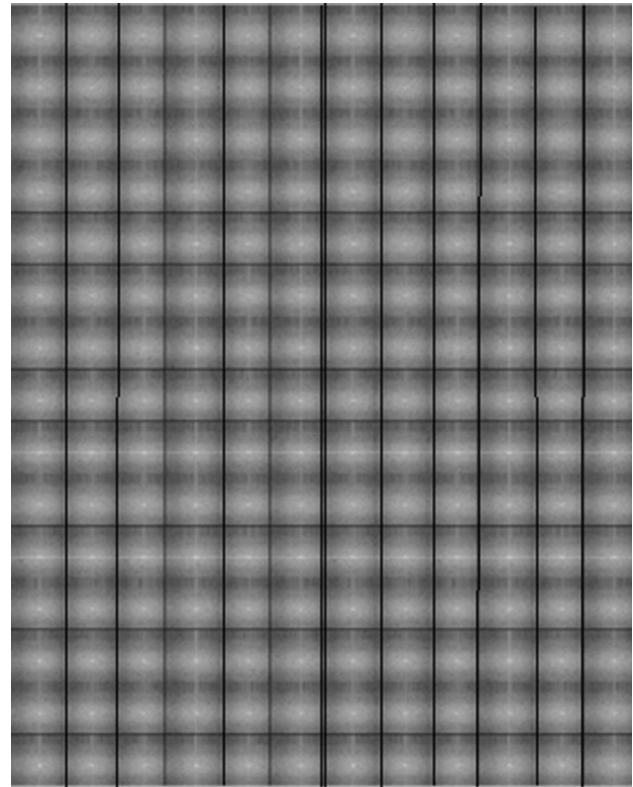


Fig. 3 Placed spectrum

reconstruction and to avoid aliasing with the neighboring order) is applied along with the inverse Fourier transform for the extraction of image information. The main reason for the selection of this cutoff frequency is that beyond this value, the performance is not satisfactory. As the position of the sideband is fixed, an algorithm is programmed in such a way that automatically the particular zone is selected for processing. The following diagram is used to explain the process (Figs. 4, 5).

To avoid the aliasing effect, a high grating frequency (1000 lines/cm) is selected. If the grating frequency is low, the bands overlap with each other (aliasing) and hence it would be difficult to extract information.

A vertical and horizontal gap is maintained between each selected spot in the main canvas to avoid the aliasing effect.

### Quality Checking of the Output Image

Peak signal to noise ratio (PSNR) and correlation coefficient methods are used to measure the quality of the extracted video frames.

### Results and Discussion

For our research work, a small video of the nebula is used. The entire video is divided into 60 frames (pixel size 1280 × 720). The whole simulation work is performed by MATLAB 2015 software, i5 Processor, and 32 GB RAM. No optical hardware system is used in this research work.

Figure 6 represents 60 frames that are used in this work.

Fig. 4 Pixel intensity graph along the dotted line

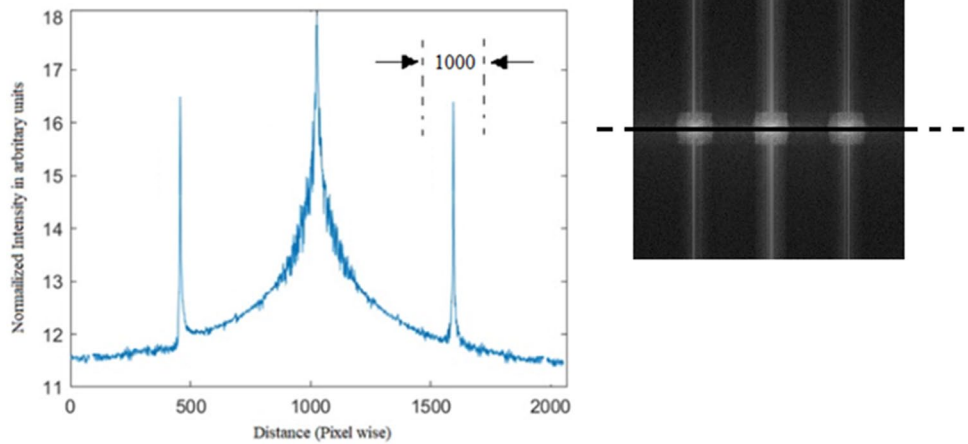
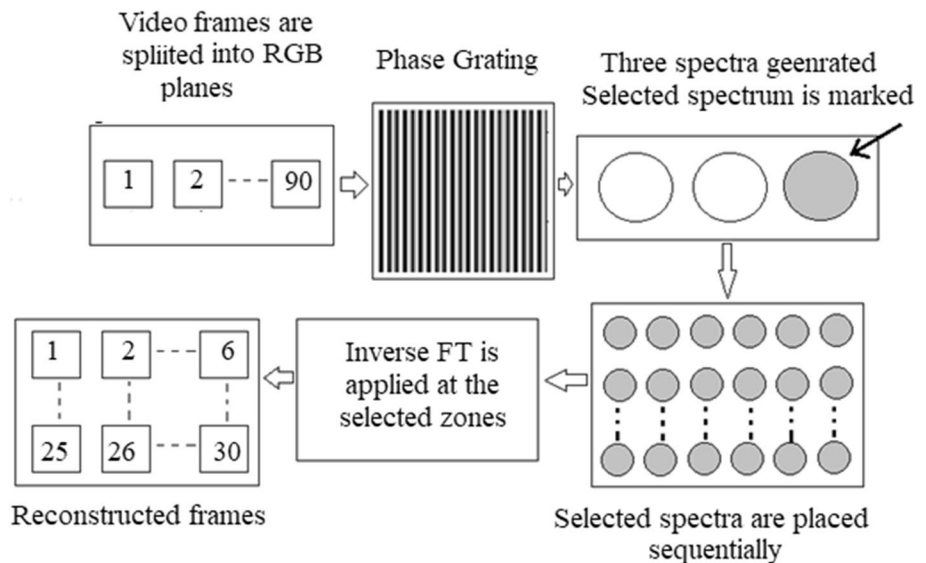
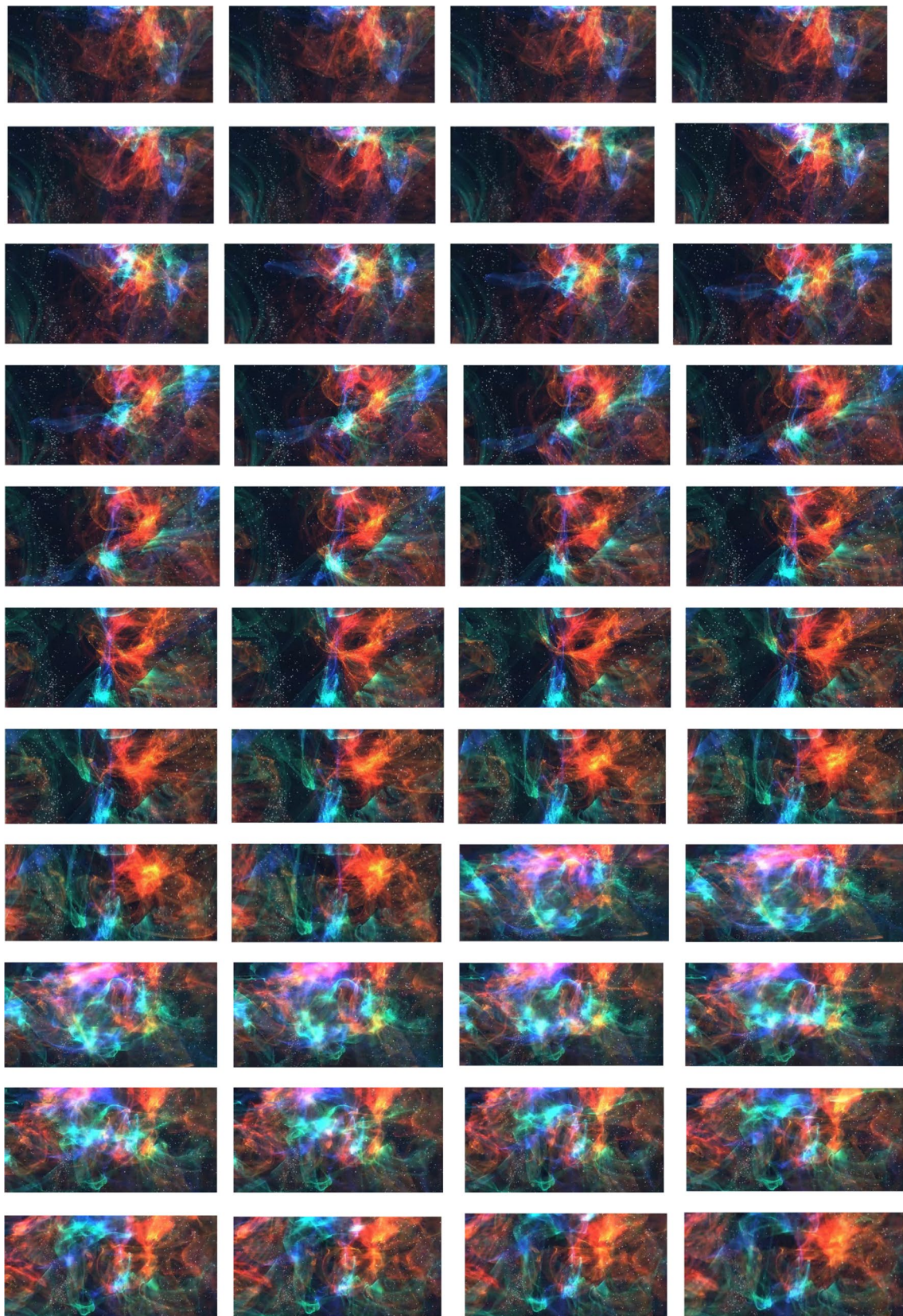


Fig. 5 Basic diagram of the entire process





**Fig. 6** 60 High-resolution frames of space video

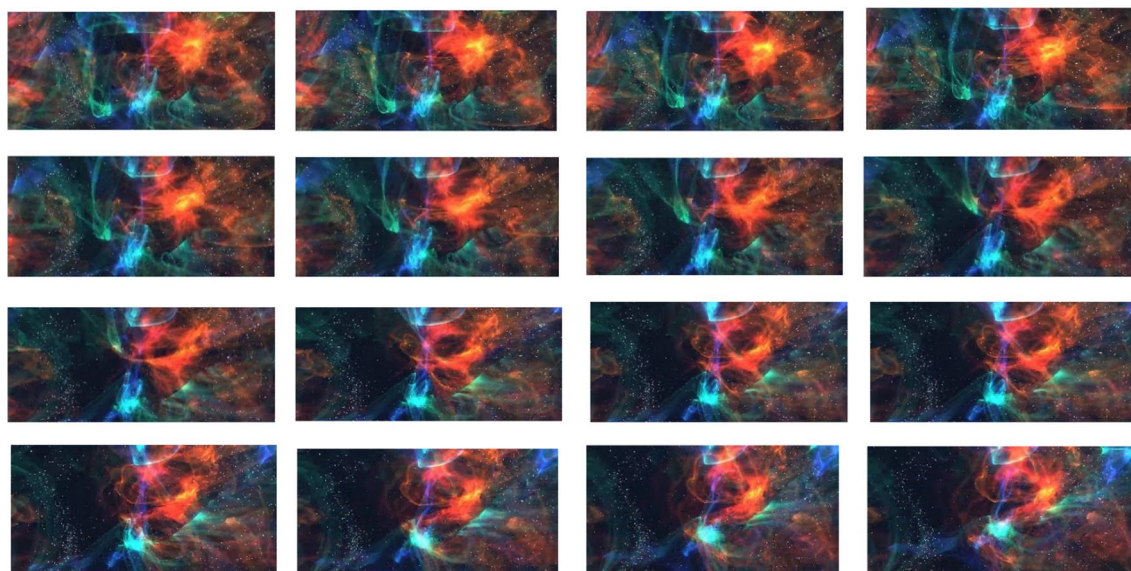


Fig. 6 (continued)

To retrieve the frames, from the encrypted canvas, inverse FT is applied in selected spots after decryption (dividing by random phase matrix). This process consists of two parts: automatic selection of the maximum intensity area followed by application of inverse FT in the area. Extracted frames are shown in Fig. 7.

To measure the quality of the extracted images, conventional PSNR and correlation coefficient methods are used. PSNR measures the effect of noise attack, whereas the correlation coefficient compares the closeness of the output and input frames. The result is illustrated in Table 1.

The high value of the correlation coefficient indicates the good quality of the extracted images. In our system, more than 99% of the image (frame) information is confined to the spectrum generated due to modulation. So, more than 99% of the image (frame) can be recovered using this method. As a result, the value of the correlation coefficient is very high and almost close to 1.

For extraction, only one sideband is used whose pixel resolution size ( $1000 \times 700$ ) is less than the original frame size. Therefore, the resolution of the extracted frame is less than the original frame. For the isolation of the frames, a small pixel gap is maintained between each frame. The pixel resolution size of the canvas is less than the total size of all original frames which is mentioned in Table 2

### Why a Low Compression Ratio is Selected

For a detailed analysis of the output images, a low compression ratio is maintained. If the compression ratio is high, due to too much loss of information, it may not be suitable for proper analysis.

The compression ratio should be low for space video to avoid too much loss of information. If high pixel resolution ( $> 1280 \times 780$ ) is considered, more storage space is required to store the original as well as the reconstructed video file. Moreover, it consumes more time for transmission and processing.

In the case of low pixel resolution ( $< 1280 \times 780$ ), less storage space is required for storing the original video, but the quality of the output image is degraded compared to the original value. The main reason is due to less information present in the low pixel resolution ( $< 1280 \times 780$ ) video.

Pixel resolution  $1000 \times 700$  is selected as it has been observed that at this particular value, the quality of the reconstructed frames, PSNR, and the correlation coefficient is satisfactory. Beyond this value ( $> 1000 \times 700$ ), the quality of the reconstructed frames is better, but it consumes more storage space.

Time required to process with 24 fps and 60 fps.

Time required to process 24 fps video—0.573 s.

Time required to process 60 fps video—0.838 s.

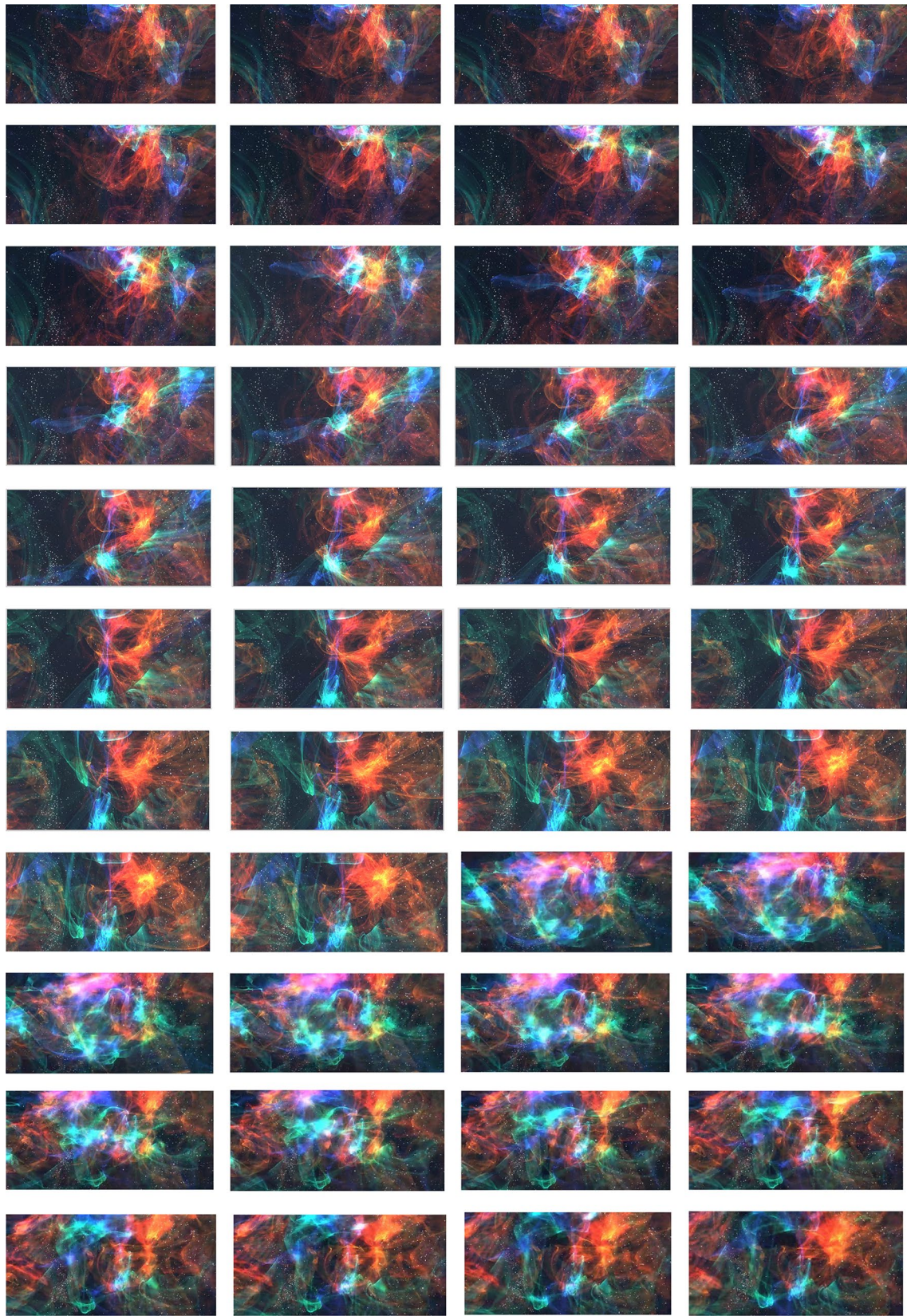


Fig. 7 Extracted frames of space video

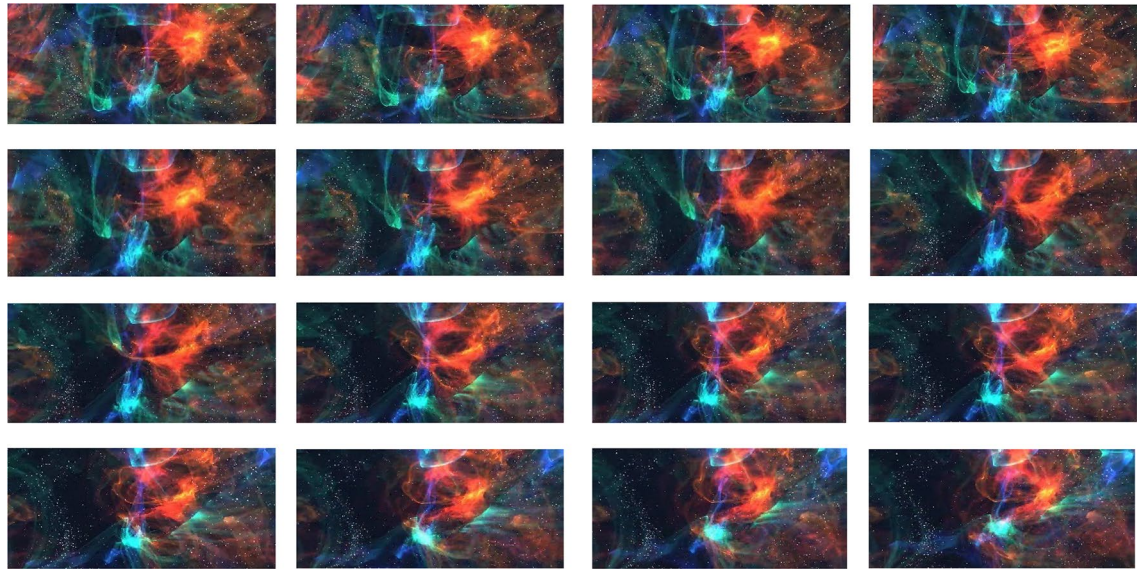


Fig. 7 (continued)

Table 1 Comparison table

S. no	Frame	PSNR	MSE
1	1–3	38.84	63.45
2	4–8	38.82	63.46
3	9, 10	39.28	63.42
4	11–14	39.11	63.42
5	15	38.97	63.43
6	16–19	39.17	63.42
7	20, 21	39.26	63.42
8	22–26	38.94	63.43
9	27–30	39.14	63.42
10	31–33	39.23	63.42
11	34	38.95	63.43
12	35–39	39.18	63.42
13	40, 41	39.11	63.42
14	42–45	38.87	63.45
15	46, 47	39.22	63.42
16	48	39.2	63.42
17	49–51	38.83	63.45
18	52	38.95	63.43
19	53–55	39.3	63.41
20	56, 57	39.17	63.42
21	58	39.22	63.42
22	59, 60	39.16	63.42

Table 2 Storage space calculation

Video FPS	The original size of the frames (MB)	Size of reconstructed frames (MB)
30	20.1	14.2
60	40.2	28.4

Compression Ratio (Uncompressed Size/ Compressed size)= 1.42:1

## Comparison Study with Previous Works

Method	Drawback	Our method
Reversible lossless Spectrum spatial transforms (Zhang & Wu, 2006)	High Compression Ratio (> 11); not suitable for Space Images	Low compression ratio (< 1.5); suitable for space research
HEVC lossless video Compression (Choi & Ho, 2015)	Time-consuming due to complex coding (> 1 s)	Fast process (takes 0.627 s)
Amplitude grating (Patra et al., 2019)	Poor efficiency (6.25%)	High efficiency as phase grating is used (33.8%)



## Possible Difficulties of the System

This proposed research work is tested on computational simulation only hence for real-time hardware-oriented systems (especially for encrypted videos with a high value of fps) the time requirement may be more due to the increase in the number of planes and complex encryption algorithms. Moreover, in applications where a high compression ratio is required, systems may not provide excellent output.

## Conclusion

Here, an effective high-resolution video compression process with the help of phase grating has been presented. HD video frames carry a lot of information; hence, to avoid too much information loss from the video frames (which is required for detail analysis), a low compression ratio is maintained. No post-processing system is used here as the quality of the extracted frames is reasonably good. (PSNR is greater than 38 and the correlation coefficient value is close to 1 in all cases.) A comparative study also showed the effectiveness of the method compared with a few existing algorithms.

**Funding** We further declare that we have No Financial or Non-Financial Competing interests regarding this manuscript.

## Declarations

**Conflict of interest** We declare that this manuscript is original and has received no funding from any source. We confirm that all of us are aware of the Conflicts of Interest policy of the journal.

## References

- Bathula, A., Gupta, S. k., Merugu, S., & Skandha, S.S. (2022). Academic projects on certification Mmanagement using blockchain—A review. In *2022 International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC)*, pp. 1–6.
- Bathula, A., Muhuri, S., & Gupta, S.k. et al. (2023). Secure certificate sharing based on Blockchain framework for online education. *Multimedia Tools and Applications*, 82, 6479–16500.
- Birman R., Segal Y. & Hada, O. (2020). Overview of research in the field of video compression using deep neural networks. *Multimedia Tools and Applications*, 79, 11699–11722.
- Bowen, L., Yu, C., Shiyu, L., & Hun-Seok, K (2021). Deep learning in latent space for video prediction and compression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 701–710; (2021).
- Chen, J., Karczewicz, M., Huang, Y., Choi, K., Ohm, J., & Sullivan, G. J. (2020). The Joint Exploration Model (JEM) for video compression with capability beyond HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(5), 1208–1225.
- Choi, J.-A. & Ho, Y.-S. (2015). Efficient residual data coding in CABAC for HEVC lossless video compression. *Signal, Image, and Video Processing*, 9.
- Ding, D., Ma, Z., Chen, D., Chen, Q., Liu, Z., & Zhu, F. (2021). Advances in video compression system using deep neural network: A review and case studies. *Proceedings of the IEEE*, 109(9), 1494–1520.
- Gormish Michael, J., Schwartz Edward, L., Keith Alexander, F., Boliek Martin, P., & Zandi, A. (1997). Lossless and nearly lossless compression for high-quality images. In *Proceedings of SPIE 3025, Very High Resolution and Quality Imaging II*.
- Gorodilov, A., Gavrilov, D., & Schelkunov, D. (2018). Neural networks for image and video compression. In *2018 International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI)*, pp. 37–41
- Harvey, J. E., Pfisterer, R. N. (2019). Understanding diffraction grating behavior: including conical diffraction and Rayleigh anomalies from transmission gratings. *Optical Engineering*, 58(8).
- Ayobkhan, M.U.A., Chikkannan, E. & Ramakrishnan, K. (2017). Lossy image compression based on prediction error and vector quantization. *Journal on Image and Video Processing*, 35.
- Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., & Wang, S. (2020). Image and video compression with neural networks: A review. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(6), 1683–1698.
- Patra, A., Chakraborty, D., Saha, A., & Bhattacharya, K. (2019). Compression of satellite images using sinusoidal amplitude grating. *International Journal of Electronics Engineering*, 11(1), 664–667.
- Patra, A., Saha, A., & Bhattacharya, K. (2020). Compression and multiplexing of medical images using optical image processing. *Computational Intelligence and Its Applications in Healthcare*, pp 63–71.
- Patra, A., Saha, A., Chakraborty, D., & Bhattacharya K. (2021). Compression of high-resolution satellite images using optical image processing; ED - Satellite Systems - Design, Modeling, Simulation, and Analysis.
- Patra, A., Saha, A., & Bhattacharya, K. (2022). Efficient storage and encryption of 32-Slice CT Scan images using phase grating. *Arabian Journal of Science and Engineering*, 47 (6).
- Rahman, M. A., Fazle Rabbi, M. M., Islam, M. M. & Islam M. R. (2018). Histogram modification based lossy image compression scheme using Huffman coding,”. In *4th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*
- Rusyn, B., Lutsyk, O., Lysak, Y., Lukenyuk, A., & Pohreliuk, L. (2016). Lossless image compression in the remote sensing applications. In *IEEE First International Conference on Data Stream Mining & Processing (DSMP)*
- Shaik, A. S., Karsh, R. K., Islam, M., & Singh, S. P. (2022a). A secure and robust autoencoder-based perceptual image hashing for image authentication. *Wireless Communications and Mobile Computing*
- Shaik, A.S., Karsh, R.K., Suresh, M., & Gunjan, V.K. (2022b). LWT-DCT Based Image Hashing for Tampering Localization via Blind Geometric Correction. *Lecture Notes in Electrical Engineering*, vol 783
- Skandha, S.S., Agarwal, M., & tkarsh, K. et al. (2022). A novel genetic algorithm-based approach for compression and acceleration of deep learning convolution neural network: an application in computer tomography lung cancer data. *Neural Computing & Applications*, 34, 20915–20937.
- Tan, T. K., Rajitha, W., Mrak, M., Ramzan, N., Baroncini, V., Ohm, J. R., & Sullivan, G. (2016). Video quality evaluation methodology and verification testing of HEVC compression performance. *In*

*IEEE Transactions on Circuits and Systems for Video Technology*, 26(1), 76–90.

Xiao, B., Lu, G., Zhang, Y., Li, W., & Wang, G. (2015). Lossless image compression based on integer discrete tchebichef transform. *Neurocomputing*, 214(C).

Zhang, N., & Wu, X. (2006). Lossless compression of color mosaic images. *IEEE Transactions on Image Processing*, 15, (6)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.