UDC 004.02
https://doi.org/10.32362/2500-316X-2022-10-4-7-17

REVIEW ARTICLE

Comparative analysis of compression algorithms for four-dimensional light fields

RETRACTED

Roman G. Bolbakov @,
Vladimir A. Mordvinov,
Artem D. Makarevich

MIREA – Russian Technological University, Moscow, 119454 Russia
@ Corresponding author, e-mail: bolbakov@mirea.ru

Abstract
Objectives. The widespread use of systems for capturing light fields is due to the high quality of the reproduced image. This type of capture, although qualitatively superior to traditional methods to capturing volumetric images, generates a huge amount of data needed to reconstruct the original captured 4D light field. The purpose of the work is to consider traditional and extended four-dimensional image compression algorithms, to perform a comparative analysis and determine the most suitable.
Methods. Mathematical methods of signal processing and methods of statistical analysis are used.
Results. Algorithms are compared and analyzed in relation to the compression of four-dimensional light fields using the PSNR metric. The selected evaluation criterion is affected not only by the dimension of the compression algorithm, but also by the distance of the baseline of the capture setting, since the difference between images increases with the distance between the optical centers of each camera matrix. Thus, for installations consisting of an array of machine vision cameras located on racks and placed in a room, the obvious choice would be to use conventional image compression methods. Furthermore, based on the assessment of the arbitrariness of video compression methods, it should be noted that the XVC algorithm remains undervalued, although its results are higher, algorithm AV1 can be considered the next in order of importance. It has been established that the latest compression algorithms show higher performance if compared to their predecessors. It has also been shown that with a small distance between the optical centers of the captured images, the use of video compression algorithms is preferable to the use of image compression algorithms, since they show better results in both three-dimensional and four-dimensional versions.
Conclusions. A comparison of the results obtained shows the need to use algorithms from the video compression family (XVC, AV1) on installations with a long baseline (mounted on camera stands). When working with integrated light field cameras (Lytro) and setting the capture with a short baseline, it is recommended to use image compression algorithms (JPEG). In general, video compression algorithms are recommended, in particular XVC, since on average it shows an acceptable level of PSNR in both the case of a short and long installation baseline.

Keywords: 3D visualization, 4D light field, light field compression

• **Submitted:** 10.01.2022 • **Revised:** 04.05.2022 • **Accepted:** 20.06.2022

For citation: Bolbakov R.G., Mordvinov V.A., Makarevich A.D. Comparative analysis of compression algorithms for four-dimensional light fields. *Russ. Technol. J.* 2022;10(4):7–17. https://doi.org/10.32362/2500-316X-2022-10-4-7-17

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ОБЗОР

Сравнительный анализ алгоритмов сжатия четырехмерных световых полей

RETRACTED

Р.Г. Болбаков @,
В.А. Мордвинов,
А.Д. Макаревич

МИРЭА – Российский технологический университет, Москва, 119454 Россия
@ Автор для переписки, e-mail: bolbakov@mirea.ru

Резюме
Цели. Широкое распространение систем захвата световых полей обусловлено высоким качеством воспроизводимого изображения. Этот вид захвата, хоть и качественно превосходит традиционные подходы к захвату объемных изображений, генерирует огромное количество данных, необходимых для восстановления исходного заснятого четырехмерного светового поля. Цель работы – рассмотреть традиционные и расширенные до четырехмерной размерности алгоритмы сжатия изображений, провести их сравнительный анализ и определить наиболее подходящие из них.
Методы. Использованы математические методы обработки сигналов и методы статистического анализа.
Результаты. Проведены сравнение и анализ алгоритмов применительно к сжатию четырехмерных световых полей с использованием метрики PSNR. Установлено, что на выбранный критерий оценивания влияет не только размерность алгоритма сжатия, но также и расстояние базовой линии установки захвата, так как разница между изображениями увеличивается в зависимости от расстояния между оптическими центрами каждой матрицы камеры. Так для установок, состоящих из массива камер машинного зрения, находящихся на стойках, расставленных в помещении, очевидным выбором будет применение обычных методов сжатия изображений. Также, исходя из оценки произвольностей методов сжатия видео, замечено, что алгоритм XVC остается недооцененным, хотя его результаты оказываются выше остальных. Следующим по значимости можно считать алгоритм AV1. Установлено, что новейшие алгоритмы сжатия показывают более высокую производительность по отношению к своим предшественникам. Продемонстрировано, что при небольшом расстоянии между оптическими центрами запечатленных изображений применение алгоритмов сжатия видео более предпочтительно, чем применение алгоритмов сжатия изображений, так как они показывают более высокие результаты как в трехмерном, так и в четырехмерном варианте.
Выводы. Сравнение полученных результатов показывает необходимость применения на установках с длинной базовой линией (установленных на стойках камеры) алгоритмов из семейства сжатия видеозаписей (XVC, AV1). При работе с интегрированными камерами светового поля (Lytro) и установкой захвата с короткой базовой линией рекомендуется использовать алгоритмы сжатия изображений (JPEG). В общем случае рекомендуется использовать алгоритмы сжатия видео, в частности XVC, поскольку в среднем он показывает приемлемый уровень PSNR как в случае с короткой, так и с длинной базовой линией установки.

Ключевые слова: 3D-визуализации, 4D-световое поле, сжатие световых полей

• **Поступила:** 10.01.2022 • **Доработана:** 04.05.2022 • **Принята к опубликованию:** 20.06.2022

Для цитирования: Болбаков Р.Г., Мордвинов В.А., Макаревич А.Д. Сравнительный анализ алгоритмов сжатия четырехмерных световых полей. *Russ. Technol. J.* 2022;10(4):7–17. https://doi.org/10.32362/2500-316X-2022-10-4-7-17

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RETRACTED

INTRODUCTION

By its nature, light field information describes the parameters of light emitted from a point in space captured, for example, by a large number of cameras [1] or a light-field imaging camera [2] based on the light field principle [3]. Requirements for the storage and transmission of such data often come down to increasing the volume of their storage and improving the performance of transmission channels. Therefore, the development and research of compression methods for light fields has become increasingly important in recent years. Although there are many effective compression formats for still and moving images, little research has been reported in the literature on the influence of these methods on the properties of light field images [4]. In this work, we assessed the impact of modern methods of image and video compression on the quality of images obtained based on light field data. These methods include the latest video compression standards, especially AV1[1] [5] and XVC[2] [6]. In order to take full advantage of the potential of common image compression algorithms on 4D light field images, we extend these techniques to 3D and 4D measurements. The paper demonstrates that 4D light field data can be compressed much higher than independent still images while maintaining the same visual quality of the perceived image.

In order to describe a three-dimensional scene from any possible position of the observer, we define the function $P(x, y, z, \varphi, \psi)$ [2], where (φ, ψ) is the viewing angle of the camera (in spherical coordinates); (x, y, z) is the absolute position of the viewpoint (Fig. 1). The result of the execution of the P function is a color that describes the visual representation of the observed point in space. The parameter t (time) can be added to the definition of the function P in order to describe a dynamic scene.

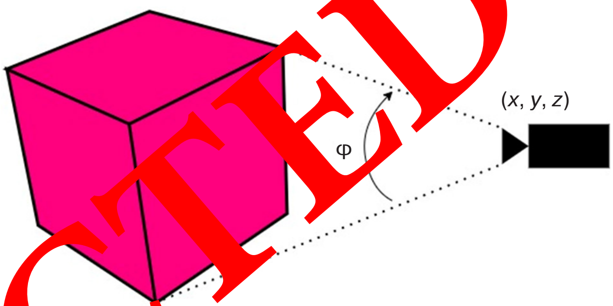

Fig. 1. Capturing a scene from one viewing position. For simplicity, the viewing angle is specified for one spherical coordinate

Our objective is to describe the scene by shooting either with multiple cameras (camera array) or with a single compact array with microlenses in front of it, such as in Lytro. In this case, the aperture can be depicted as a grid of views (cameras) located on a two-dimensional plane. This case is illustrated in Fig. 2, where the distance on the line passing through the base points between the individual views is described by the parameter d. This representation is often referred to as a 4D light field (LF) since we are dealing with a light field function L sampled in four dimensions (k, l, m, n), where (m, n) are pixel coordinates; (k, l) are subaperture image indexes.

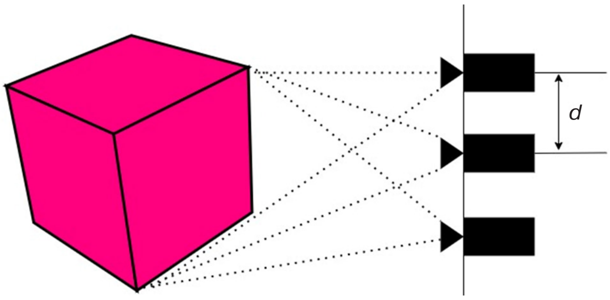

Fig. 2. Capturing a 4D light field with a camera array

The light fields received by one compact single device have limitations on the viewing angle. Light fields based on multiple cameras provide large viewing

[1] AOMedia Video 1 (AV1) is an open video compression standard for encoding video transmitted over the Internet. It replaces the VP9 video encoding format developed by Google. According to [5], AV1 outperforms H.265/HEVC by 17% and VP9 by 13% over a wide range of bitrates/resolutions. Developed by the Alliance for Open Media (AOMedia), created in 2015 and comprised of electronics, video-on-demand, and web browser companies (AMD, Apple, Arm, Broadcom, Intel, Nvidia, Amazon, Facebook, Google, Hulu, Netflix, Mozilla, Microsoft). Timothy B. Terriberry. Progress in the Alliance for Open Media (slides). URL: https://people.xiph.org/~tterribe/pubs/lca2017/aom.pdf (18 January 2017). Accessed June 22, 2017.

[2] XVC is a video encoding format with a strong focus on low bitrate streaming applications. The official website (URL: https://xvc.io/. Accessed December 04, 2021) claims that the codec is superior to AV1, H.265/HEVC, and VP9.

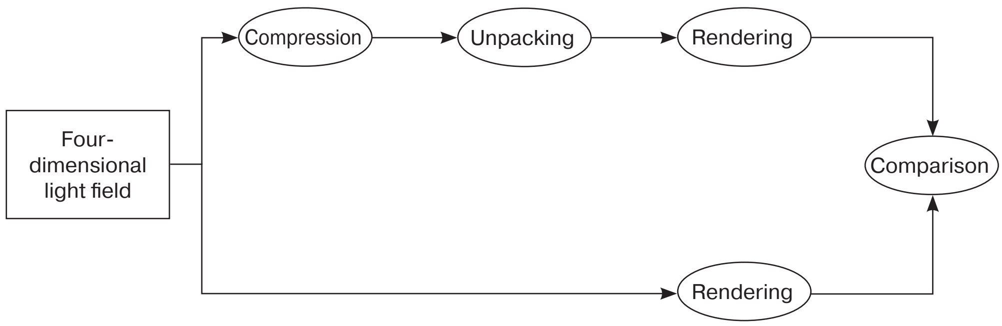

Fig. 3. Data flow diagram of the technique for comparing compression algorithms

angles by distributing camera arrays in the space around the object. In practice, the number of views located on a two-dimensional plane varies from a couple units. Given the high resolution of the sensors, it is not surprising that the volume of light field data is enormous. As an example, consider the "Treasure Chest" light field[3] (Fig. 3) taken from the Stanford Light Field Archive. This field is captured using a 17 × 17 grid of cameras with an image resolution of 1536 × 1280 pixels. The size of uncompressed data exceeds 1 Gb. When using photo sequences to organize footage, storage and transmission requirements will increase proportionately.

METHODS FOR COMPRESSING 4D LIGHT FIELDS

Recently, several methods for compressing 4D light fields have been proposed in the literature [4, 5, 7–19]. Using some of these methods, researchers are trying to directly compress data received from sensors that are preceded by microlenses (lens images). Others compress the resulting 4D light field instead of processing the original "raw" data. In this article, we will focus only on the latter.

Let us compare various present-day compression methods applicable to 4D light field data. These methods include the latest video compression standards, especially AV1 (approved in June, 2018) and XVC (version released in July, 2018). In order to make a comparison, we refocus the original and decompressed light field. The evaluation is then carried out using the PSNR (peak signal-to-noise ratio) metric as a full-scale quality benchmark.

Separate displays from the original light field are usually not rendered. Therefore, it makes no sense to directly compare the original and decompressed light fields, although such methodology is usually used to evaluate the performance of one kind of compression. For this reason, the compression performance estimation methodology for multi-focus rendering from [4] will be used. This methodology consists basically of evaluating the rendering quality of views for multiple breakpoints. Rendered displays are obtained by combining pixels from different 4D light field views for different focal planes. The average distortion is calculated as the average of the PSNR for several rendered views in the focal plane. The comparison technique is illustrated in Fig. 3. Note that PSNR is calculated from the mean square of the error over all three color components.

A 4D light field contains a 2D grid of 2D views captured from cameras. The baseline length between individual views varies from a few millimeters (microlenses) to a few centimeters (camera array). Therefore, it is natural to expect a high level of similarity between views adjacent in either of the two grid directions. This similarity opens the way to understanding 4D light field data as a video sequence moving between viewpoints. Alternatively, we can consider a 4D light field as 3D or directly as a 4D body. The approaches described above can also be found in light field compression using an image, video, 3D or 4D image coding system (although other approaches are possible, for example, using 3D video).

In recent years, the compression performance of various approaches to light field imaging have been compared and evaluated.

In [4], the authors evaluated the performance of the main image coding standards with independent views and H.265/HEVC[4] with independent views. The label "with independent views" indicates that the individual views have been compressed independently of each other. Approaches to video encoding were not evaluated in the work. As expected, the H.265/HEVC internal profile proved to be the most efficient compression method.

In [7], the authors compared the compression performance of three strategies using the H.265/HEVC codec. The first strategy performs compression directly

[3] Standford light-field dataset. URL: http://lightfield.stanford.edu/?gclid=Cj0KCQiA47GNBhDrARIsAKfZ2rD2CB3lMtzHJXPr0uXM_KJm_tElIZIZLviFERCFsasV9JygG55uBIaAtRTEALwwcB. Accessed December 04, 2021.

[4] H.265/HEVC is a high efficiency video coding; MPEG-H is a video compression standard developed as a successor to the widely used H.264/AVC (MPEG-4).

RETRACTED

on the original light-field image. The following strategy organizes the views of the 4D light field into a pseudo-temporal sequence in a spiral order and subsequently compresses them. In the latter strategy, compression occurs on a subset of images extracted from the light-filed image by transforming into a four-dimensional light field. The results of the work carried out by the authors show that the encoding of a four-dimensional light field leads to better performance when compared to direct encoding of images received from cameras.

The authors of [8] compared the performance of JPEG[5], JPEG 2000, and SPIHT[6] directly on images obtained from camera matrices. The comparison was carried out using the same methodology as in this article. As you might expect, JPEG 2000 shows the best compression performance.

In [9], the authors proposed to rearrange the 4D light field as a tiled representation of a large rectangular image. This image was then compressed using a JPEG 2000 encoder. The proposed scheme was compared with standard image encoding algorithms, namely JPEG 2000 and JPEG XR. However, it is not clear how accurately these standard encoding algorithms were applied to 4D light field data.

In [10], the author reconstructs a four-dimensional light field into a three-dimensional body. The 3D volume is then encoded using a 3D DCT scheme of 8 × 8 × 8 blocks similar to the JPEG coding system.

Besides conventional coding methods, there is also an alternative approach [11] which uses deep learning to estimate a 2D representation from sparse sets of 4D representations. Another approach [12] proposes its own sparse coding scheme for the entire four-dimensional light field, based on several optimized key representations.

The method described in [13] is based on the construction of superbeams which limit the superpixels that form a given superbeam. This constraint is necessary in order that superbeams can be used to support angular dimensionality reduction based on low-rank matrix approximation. Then, an approximation of the low-rank matrix for the superbeams is calculated with inconsistency compensation using singular value decomposition (SVD). The base vectors are then encoded using HEVC or JPEG-Pleno VM 1.1 for each individual representation.

In [14, 15], the authors propose a hierarchical coding structure for four-dimensional light fields. The 4D light field is broken down into several views and then organized into an encoding structure according to spatial coordinates. All representations are coded hierarchically. The scheme is implemented in the H.265/HEVC reference software.

In [16], the authors propose an encoding scheme which divides a four-dimensional light field into several central views and other adjacent views. The adjacent views are subtracted from the center views, and then both groups are encoded using the H.265/HEVC codec. The authors of [17, 18] transfer the four-dimensional light field to the H.265/HEVC codec using the inter-prediction mode for individual LF views. Finally, great attention has been paid to compression approaches based on convolutional neural networks [19, 20]. It can be seen from the above that JPEG 2000 and especially H.265/HEVC coding schemes are quite popular when compressing 4D light fields.

In this article, we compare the performance of the main presently methods for compression with losses. These methods can be divided into four groups depending on how the 4D LF data is processed. The first group covers the following image encoding methods: JPEG and JPEG 2000. In [21] they are called methods based on self-similarity. The second group includes video encoding methods: H.265/HEVC, AV1, VP9, and XVC. They are generally referred to as pseudo-sequence-based methods. The third group extends methods for encoding images in three dimensions. This group consists of JPEG 3D and JPEG 2000 3D. Note that JPEG 3D refers to a 3D image and not to a pair of stereoscopic images. The fourth group extends image encoding methods in four dimensions. However, there is only one method in this group—JPEG 4D.

The following codecs are used to assess the above methods: OpenJPEG, x265, libaom (AV1 codec library), libvpx (VP8/VP9 codec SDK), and XVC codec.

INITIAL DATA

This section presents the data set, the multifocal imaging method, the experiments performed with this data set, and the results obtained.

The data set consists of four 4D light fields based on two types of capture devices. Two light fields were taken with a Lytro Illum B01 light-field camera (manufactured by Lytro, USA), and the other two were taken with conventional DSLR cameras.

The first conventional camera light field was captured with a multi-camera array, and the second with a simple motorized setup equipped with a Canon Rebel XTi digital camera (manufactured by Canon, Japan). The corresponding resolutions and corresponding image divergence ranges are listed in Table. The values in the last column—image divergence—describe the difference in pixels in the location of the same 3D object projected onto images taken by the camera or calculated from the light-field camera image, in the case of Lytro.

[5] JPEG—Joint Photographic Experts Group, titled after the name of the developer.

[6] SPIHT is a set partitioning in hierarchical trees.

Fig. 4. Data set used for comparison.
Left to right: Danger de mort, Chessboard, Treasure Chest, Palais du Luxemburg

RETRACTED

As can be seen, the range of discrepancies will be narrow (from −1 to +1 pixels) for a light field with a dense sample (short baseline) in the case of a Lytro camera and wide (from 40 to 90 pixels) for images obtained by an array of cameras. These values obviously correlate with the focal length of the camera and the distance between the centers of the cameras or light-field lenses in the case of Lytro. For convenience, the central view for each light field is shown in Fig. 4.

The first and [illegible] light fields shown in Fig. 4 are taken with [illegible]; Chessboard is captured using an array of cameras; Treasure Chest is captured using a [illegible] with a camera affixed to it.

Digital refocusing of images in the virtual focal plane is achieved using the shift-sum algorithm [22]. This algorithm shifts subaperture images (views) according to the optical center of the camera relative to the camera baseline relative to the reference frame and accumulates the corresponding pixel values. The refocused image will be the average of the converted images. The calculation of the pixel value at point (m, n) of the refocused image E_d is given by the equation:

$$E_d(m,n) = \frac{1}{N}\sum_{k,l} L(k,l,m+\alpha k,n+\alpha l), \qquad (1)$$

where N is the number of summed images; α is the distance of the synthetic plane from the main lens; k and l are indices of the subaperture image of the light field representation; αk and αl are the shift parameters with respect to the reference system. Linear interpolation was also performed in the last two 4D measurements, in order to convert the sampled light field function to a continuous one.

EXPERIMENT 0

It is worth clarifying in advance whether it is really necessary to evaluate the quality of images displayed for several focal points, rather than on the original data (that is, directly compare the original and decompressed

Table. Dataset used for comparison

Description	Source	Resolution	Discrepancy in pixels
Danger de mort	EPLF dataset	15 × 15 × 625 × 434	From −1 to 1
Chessboard	Saarland University	8 × 8 × 1920 × 1080	From 40 to 90
Treasure Chest	Stanford Computer Graphics Lab	17 × 17 × 1536 × 1280	From −1 to 7
Palais du Luxemburg	EPLF dataset	15 × 15 × 625 × 434	From −1 to 1

RETRACTED

set of images). A quick experiment shows that there is a big difference between these approaches (Fig. 5). It is about 10 decibels in PSNR depending on the bit rate and compression method. This can be explained by the fact that any pixel in the displayed form is the sum of the pixels from the 4D light field, so this sum together reduces compression artifacts. In other words, we can afford to compress 4D light fields much more than independent images while maintaining the same visual quality of the displayed image.

Figure 5 shows the difference in quality evaluation using a 4D light field directly compared to using images rendered at virtual focal planes. The illustration is shown on the Danger de mort light field.

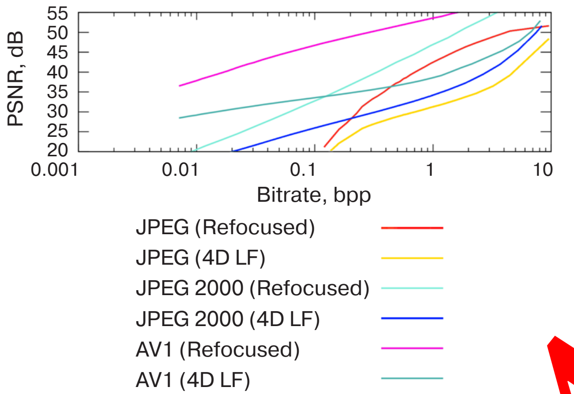

Fig. 5. Experiment 0: difference in quality evaluation

EXPERIMENT 1

As can be seen from the literature review, most present-day approaches to light field compression process either 2D data or its sequence (video compression). The compression of 4D light field images is still a relatively unexplored area. Since the 4D light field is sequences of 2D images (viewpoints), 2D compression techniques can be used to independently encode the viewpoints. However, such methods do not allow the use of pixel correlations in all four dimensions. A similar reasoning can be applied to 3D methods. In our Experiment 0, we were interested in studying the effects of contraction of light fields in three and four dimensions. In order to fairly evaluate compression performance, the same compression method must be used for the 2D, 3D, and 4D cases. Thus, in the work we use a custom implementation of the JPEG compression method with the ability to process 2D, 3D or 4D data. In addition, there exists the JPEG 2000 standard with the ability to compress 2D and 3D data in the same way. Unfortunately, JPEG 2000 does not work with 4D images. Since the similarity of neighbouring pixels in 3D and 4D is highly dependent on the camera baseline, different results can be expected. The result of this experiment is shown in Fig. 6. In each graph, the horizontal axis shows the bitrate (bits per pixel) and the vertical axis shows the average PSNR for multiple rendered focal plane viewpoints.

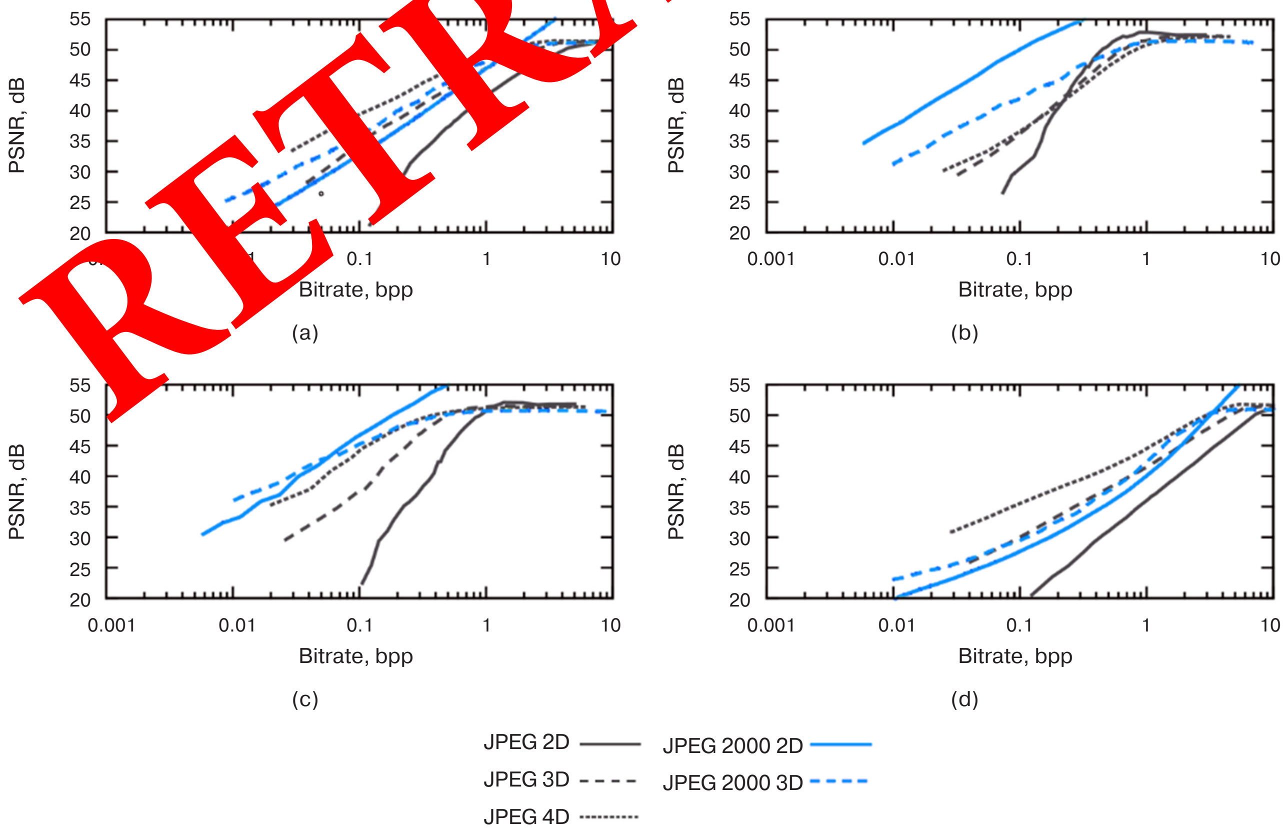

Fig. 6. Experiment 1: performance comparison of image compression methods: (a) Danger de mort, (b) Chessboard, (c) Treasure Chest, (d) Palais du Luxemburg

For light fields with a small baseline (Danger de mort and Palais du Luxemburg), both 3D compression methods clearly outperform their 2D counterparts over the entire bitrate range. Likewise, the 4D JPEG method is clearly superior to its 3D counterpart.

This is not surprising since pixels in the same spatial position in neighboring views are highly correlated. However, the situation changes as the baseline increases. In doing so (Treasure Chest and Chessboard), adjacent views become less and less similar, resulting in higher amplitudes of base transform coefficients. Consequently, the situation is changing in favor of compression methods with smaller sizes.

Considering the JPEG method, Treasure Chest is a special case since it contains a large number of black pixels. It turns out that it is more efficient to compress these solid areas at once with a single 4D block rather than with multiple 3D blocks. Likewise, it is more efficient to use one 3D block than multiple 2D blocks.

EXPERIMENT 2

The second thing to note in the previous section is the use of video compression standards. The question arises whether it is better to compress 4D light fields as a sequence of 2D frames or as a multidimensional body? Therefore, we measured the performance of all the above video compression standards. The results can be seen in Fig. 7. This time the results for only two light fields are shown for brevity.

Interestingly, the XAVC codec actually showed better compression performance than HEVC and AV1.

In order to answer the question, "What is the best compression method for these light fields?", an additional comparison of the results with the most efficient methods from Experiment 1 was carried out. The overall comparison is shown in Fig. 8. Interestingly, video compression methods perform better than all image compression methods, even better than their 3D and 4D extensions.

CONCLUSIONS

The objective of this work was to evaluate existing methods suitable for compressing 4D light fields with losses. Since the light field in the original version is a set of images captured by an array of cameras, image compression methods will be the first thing to choose from in compression problems. The experiment shows that methods which process 4D light fields directly in

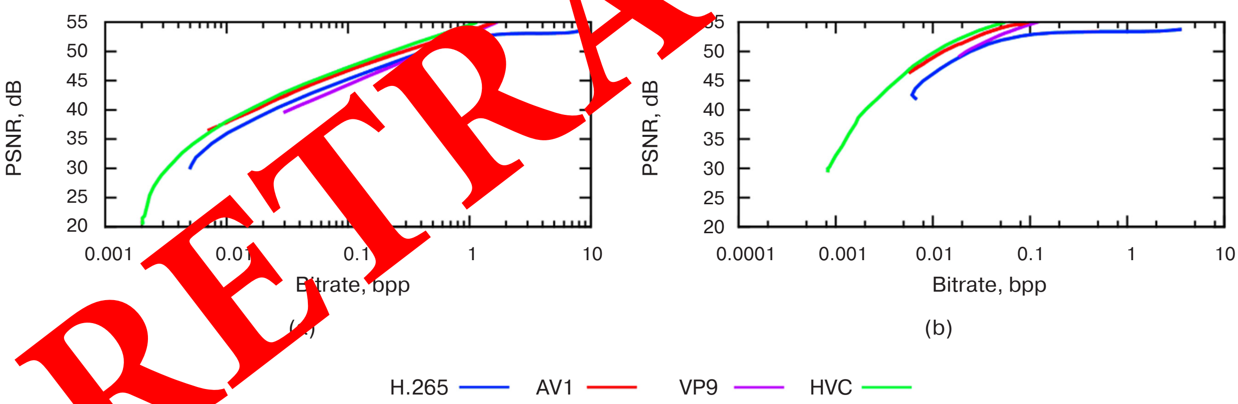

Fig. 7. Experiment 2: performance comparison of video compression methods. Fields used: (a) Danger de mort, (b) Chessboard

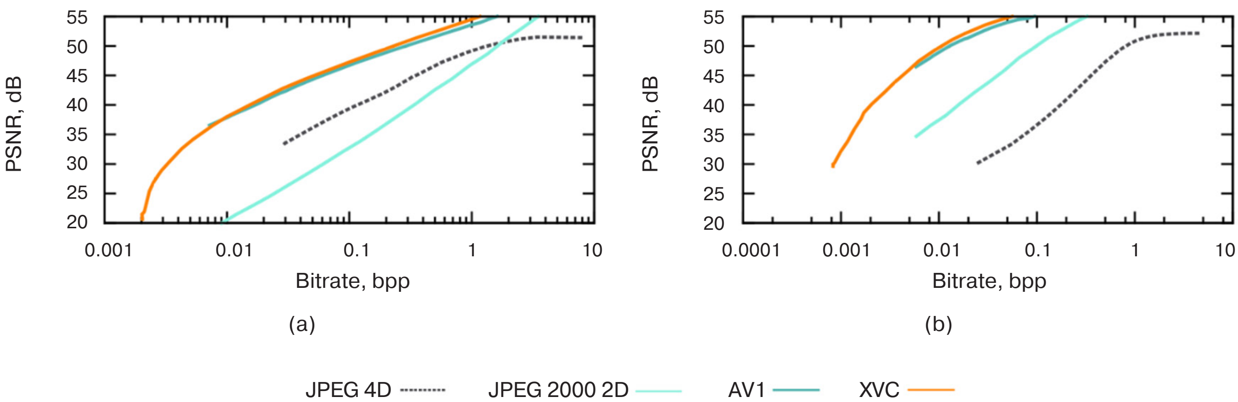

Fig. 8. Performance comparison of video compression methods in relation to image compression algorithms. Fields used: (a) Danger de mort, (b) Chessboard

RETRACTED

four or three dimensions can achieve better compression results than classical implementations of 2D image compression.

However, it should be noted that the chosen evaluation criterion, namely PSNR, is affected not only by the dimension of the compression algorithm, but also by the baseline distance, since the difference between images increases with the distance between the optical centers of each camera matrix. Thus, for installations consisting of an array of machine vision cameras located on racks and placed in a room, the obvious choice would be to use conventional image compression methods.

In addition, based on the evaluation of the arbitrariness of video compression methods, we can see that the XVC algorithm remains underestimated, although its results are higher. Algorithm AV1 can be considered the next in order of importance. This confirms the fact that the latest compression algorithms show a higher performance in relation to their predecessors.

Also, when the distance between the optical centers of the captured images is small, the use of video compression algorithms is preferable to the use of image compression algorithms, since they show better results in both three-dimensional and four-dimensional versions.

Authors' contributions

R.G. Bolbakov—idea, development of research design, consultation on the problems of carrying out all the stages of the study.

V.A. Mordvinov—analysis of scientific work, revision with the contribution of valuable intellectual content.

A.D. Makarevich—literature analysis, article writing, research planning; carrying out all stages of the study, formalization of the list of references.

REFERENCES

1. Broxton M., Flynn J., Overbeck R., Erickson D., Hedman P., Matthew Duvall M., Dourgarian J., Busch J., Whalen M., Debevec P. Immersive light field video with a layered mesh representation. *ACM Trans. Graph*. 2020;39(4):86. https://doi.org/10.1145/3386569.3392485
2. Georgiev T., et al. Lytro camera technology: theory, algorithms, performance analysis. In: *Proc. SPIE. Multimedia Content and Mobile Devices*. 2013;8667:86671J. https://doi.org/10.1117/12.2013581
3. Adelson E.H., Bergen J.R. The plenoptic function and the elements of early vision. In: Landy M., Movshon J.A. (Eds.). *Computational Models of Visual Processing*. London: The MIT Press Cambridge; 1991. P. 3–20. https://doi.org/10.7551/mitpress/2002.003.0004
4. Alves G., Pereira F., da Silva E.A.B. Light field imaging coding: Performance assessment methodology and standards benchmarking. In: *2016 IEEE International Conference on Multimedia & Expo Workshops* (*ICMEW*). IEEE; 2016. С. 1–6. https://doi.org/10.1109/ICMEW.2016.7574774
5. Zabrovskiy A., Feldmann C., Timmerer C. Multi-codec DASH dataset. In: *Proceedings of the 9th ACM Multimedia Systems Conference*. 2018. P. 438–443. https://doi.org/10.1145/3204949.3208140
6. Samuelsson J. The XVC video code" a revolutionary software-defined video compression format. *SMPTE Motion Imaging Journal*. 2019;128(10):1–8. https://doi.org/10.5594/JMI.2019.2937737
7. Viola I., Řeřábek M., Ebrahimi T. Comparison and evaluation of light field image coding approaches. *IEEE Journal of Selected Topics in Signal Processing*. 2017;11(7):1092–1106. https://doi.org/10.1109/JSTSP.2017.2740167
8. Dlabaja D. 4D-DCT based light field image compression. *Image*. 2019;28(1):128. Available from URL: http://excel.fit.vutbr.cz/submissions/2019/[illegible]/21.[illegible]
9. Brites C., Ascenso J., Pereira F. Lenslet light field image coding: Classifying, reviewing and evaluating. *IEEE Transactions on Circuits and Systems for Video Technology*. 2020;31(1):[illegible]–354. https://doi.org/10.1109/TCSVT.2020.[illegible]76784
10. [illegible] A. A [illegible]D DCT compression algorithm for omnidirectional integral images. In: *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. IEEE; 2006:2:9142643. https://doi.org/10.1109/ICASSP.2006.1660393
11. [illegible] D., et al. View synthesis-based light field image compression using a generative adversarial network. *Inf. Sci*. 2021;545:118–131. https://doi.org/10.1016/j.ins.2020.07.073
12. Pratapa S., Manocha D. RLFC: random access light field compression using key views and bounded integer sequence encoding. In: *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. 2019. P. 1–10. https://doi.org/10.1145/3306131.3317018
13. Dib E., et al. Super-ray based low rank approximation for light field compression. In: *2019 Data Compression Conference* (*DCC*). IEEE; 2019. P. 369–378. https://doi.org/10.1109/DCC.2019.00045
14. Joshitha R., Sharma M. A hierarchical coding scheme for glasses-free 3d displays based on scalable hybrid layered representation of real-world light fields. In: *2021 IEEE International Conference on Systems, Man, and Cybernetics* (*SMC*). IEEE; 2021. P. 3491–3498. https://doi.org/10.1109/SMC52423.2021.9658584
15. Ravishankar J., Sharma M. A novel hierarchical light field coding scheme based on hybrid stacked multiplicative layers and Fourier disparity layers for glasses-free 3D displays. arXiv preprint arXiv:2108.12399. 2021. https://doi.org/10.48550/arXiv.2108.12399
16. Tu W., et al. Efficient content adaptive plenoptic video coding. *IEEE Access*. 2020;8:5797–5804. https://doi.org/10.1109/ACCESS.2020.2964056
17. Chao Y.H., et al. Pre-demosaic graph-based light field image compression. arXiv preprint arXiv:2102.07883. 2021. Available from URL: https://arxiv.org/pdf/2102.07883.pdf
18. Mathew A.G., Mathew A.A. Intra coding of plenoptic images in HEVC. In: *Emerging Technologies for Sustainability*. CRC Press; 2020. P. 459–465. Available from URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9780429353628-58/intra-coding-plenoptic-images-hevc-ashlin-george-mathew-anu-abraham-mathew

19. Bakir N., et al. RDO-based light field image coding using convolutional neural networks and linear approximation. In: *2019 Data Compression Conference* (*DCC*). IEEE; 2019. P. 554. https://doi.org/10.1109/DCC.2019.00066
20. Zhong T., et al. Light field image compression using depth-based CNN in intra prediction. In: *2019 IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP*). IEEE; 2019. P. 8564–8567. https://doi.org/10.1109/ICASSP.2019.8682820
21. Yuan X., Haimi-Cohen R. Image compression based on compressive sensing: End-to-end comparison with JPEG. *IEEE Transactions on Multimedia*. 2020;22(11): 2889–2904. https://doi.org/10.1109/TMM.2020.2967646
22. Li Y., et al. Coding of focused plenoptic contents by displacement intra prediction. *IEEE Transactions on Circuits and Systems for Video Technology*. 2015;26(7):1308–1319. https://doi.org/10.1109/TCSVT.2015.2450333

About the authors

Roman G. Bolbakov, Cand. Sci. (Eng.), Associate Professor, Head of the Department of Instrumental and Applied Software, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: bolbakov@mirea.ru. Scopus Author ID 57202836952, RSCI SPIN-code 4210-2560, http://orcid.org/0000-0002-4922-7260

Vladimir A. Mordvinov, Cand. Sci. (Eng.), Professor, Department of Instrumental and Applied Software, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: mordvinov@mirea.ru. RSCI SPIN-code 9390-1540, http://orcid.org/0000-0003-3622-8448

Artem D. Makarevich, Postgraduate Student, Department of Instrumental and Applied Software, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: artemmakarevich1997@gmail.com. http://orcid.org/0000-0002-2211-1241

Об авторах

Болбаков Роман Геннадьевич, к.т.н., доцент, заведующий кафедрой инструментального и прикладного программного обеспечения Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: bolbakov@mirea.ru. Scopus Author ID 57202836952, SPIN-код РИНЦ 4210-2560, http://orcid.org/0000-0002-4922-7260

Мордвинов Владимир Александрович, к.т.н., профессор кафедры инструментального и прикладного программного обеспечения Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: mordvinov@mirea.ru. SPIN-код РИНЦ 9390-1540, http://orcid.org/0000-0003-3622-8448

Макаревич Артём Денисович, аспирант кафедры инструментального и прикладного программного обеспечения Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: artemmakarevich199[illegible]gmail.com. http://orcid.org/0000-0002-2211-1241

Translated from Russian into English by E. Shklovskii
Edited for English language and spelling by Dr. David Mossop