

Article

Not peer-reviewed version

Lossless and Near-Lossless L-Infinite Compression of Depth Video Data

[Mohammad Ali Tahouri](#)^{*}, Alin Adrian Alecu, [Leon Denis](#), [Adrian Munteanu](#)^{*}

Posted Date: 21 January 2025

doi: 10.20944/preprints202501.1571.v1

Keywords: depth map; lossless and near-lossless compression; foreground-background segmentation; L-infinite; quantization; MAXAD



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Lossless and Near-Lossless L-Infinite Compression of Depth Video Data

Mohammad Ali Tahouri ^{1,*}, Alin Adrian Alecu ², Leon Denis ¹ and Adrian Munteanu ^{1,*}

¹ Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

² Universitatea Nationala de Stiinta si Tehnologie Politehnica Bucuresti, Splaiul Independentei 313, Bucharest, Romania

* Correspondence: mohammad.ali.tahouri@vub.be, Adrian.Munteanu@vub.be

Abstract: Depth information acquisition is critically important in medical applications, such as monitoring of elderly or human biometrics extraction. In such applications, compressing the stream of depth video data plays an important role due to bandwidth constraints on the transmission channels. This paper introduces a novel lightweight compression system that encodes the semantics of the input depth video and can operate in both lossless and L-infinite near-lossless compression modes. A quantization technique that targets the L-infinite norm for sparse distributions and a new L-infinite compression method that sets bounds on the quantization error are proposed. The proposed codec enables controlling the coding error on every pixel in the input video data, which is crucial in medical applications. Experimental results show an average improvement of 45% and 17% in lossless mode compared to standalone JPEG-LS and CALIC codecs, respectively. Furthermore, in near-lossless mode, the proposed codec achieves superior rate-distortion performance and reduced maximum error per frame compared to HEVC. Additionally, the proposed lightweight codec is designed to perform efficiently in real time when deployed on an embedded depth camera platform.

Keywords: depth map; lossless and near-lossless compression; foreground-background segmentation; L-infinite; quantization; MAXAD.

1. Introduction

Nowadays, the importance of image and video compression is undeniable in medical and computer vision applications, where the compression of depth video data is crucial for optimizing bandwidth usage while preserving the integrity of the original scene. Effective video compression is vital in enhancing healthcare delivery, which calls for further research in order to improve the technological capabilities of subsequent tasks making use of compressed video data. Many standards have been developed during the past decades for image and video compression. The standards are designed to encode signals using the available pixel information while optimizing for maximum compression efficiency and performance.

The basic concept of image compression is to reduce the size of the original image by representing the input with the fewest possible bits. Image compression techniques can be broadly categorized into two types based on the presence of error: lossless and lossy (or near-lossless) compression [1]. In lossless compression, the original image is perfectly reconstructed at the decoder, ensuring that there is no error between the original and reconstructed images. In lossy compression, some data is shrunk or discarded to achieve a significantly higher compression ratio, resulting in a potential loss of image quality.

Several studies provide theoretical reviews of image compression techniques with a particular focus on lossless compression in medical applications [2,3], while others have concentrated on the practical part [4–7]. In practical applications, these papers mainly utilize CT grayscale or MRI images. In [8], a spatial domain image codec is introduced. This technique involves dividing the image into multiple blocks and calculating the difference between minimum and maximum pixel values within each block. Then, the minimum pixel value of the block is subtracted from each pixel, and the resulting

values are stored and encoded for the new block. [9] used the same idea of spatial domain image compression, but the Lempel-Ziv-Welch (LZW) algorithm is applied to encode the subtraction value. Both methods are designed only for lossless compression of images. Lastly, a comprehensive review of classical and novel image compression methods based on the region-of-interest (ROI) [10] is presented in [11] that emphasizes the growing importance of image compression in various aspects of our digital life.

Depth images are data sources that encode a depth value in each pixel, representing the distance of that pixel from the depth camera. The depth map is a representation of a scene in two dimensions grayscale. These types of data sources have already found applicability in numerous domains including the medical one, where major hospitals and medical centers world-wide commonly use cameras for patient monitoring. In this respect, various algorithms have been proposed in literature that aim at compressing the depth map images using advanced methods. For instance, the work of Tabus et al [12] encodes both the contours and the depth value from the depth map. Their method, referred to as CERV (crack-edge-region-value), begins from an initial representation of the image, utilizing binary vertical and horizontal edges of the region contours, along with the depth values assigned to each region. This approach is then used in [13] to introduce the Anchor Points Coding (APC) technique, where the anchor points of contours are encoded using a context tree algorithm. Schiopu and Tabus also developed their work of Greedy rate-distortion Slope Optimization (GSOm) [14] by proposing a progressive coding technique for lossy-to-lossless compression methods [15]. Furthermore, [16] suggests the single depth intra mode to efficiently encode the smooth area within a depth image. The idea behind this is to simply reconstruct the current Coding Unit (CU) as a smooth area with a single depth sample value, thereby enhancing the encoding efficiency of that area within the depth map by incorporating a leaf merging of the pixel values. [17] proposed a bit depth compression by estimating the packet size based on the data sensor pattern, the algorithm performing only in lossless mode.

For video coding, many standards and algorithms have been developed. The foundational standards of video compression have evolved over the decades, beginning with the introduction of the Advanced Video Coding (AVC/H.264) standard in 2003 [18], followed by the release of the improved High Efficiency Video Coding (HEVC/H.265) standard in 2012 [19] culminating with the completion of the last standard Versatile Video Coding (VVC/H.266) in 2021 [20]. The standards employ both inter-coding and intra-coding. In intra-coding, each frame in the video is encoded independently without reference to other frames. In contrast, in inter-coding, the frames are encoded using prediction mechanisms based on the previously decoded reference frames. Inter-coding improves compression performance compared to intra-coding at the expense of increased computational complexity. Existing survey papers show that compression performance is steadily improved over the years by adopting increasingly complex prediction and compression mechanisms[21–23].

Typical lossy compression systems as cited above, employ the L2 metric to drive the compression process. However, L2 is a global metric which enables minimizing the overall coding error for a given bit rate. Essentially, the coding system minimizes the L2 distortion, corresponding to minimizing the mean square error between the original and reconstructed images for any given bit budget. However, such a global metric may be globally optimal but it may also be subject to large local errors. In depth image and video data, the pixel values correspond to actual depth from the scene to the camera plane. Controlling the coding error at the level of every pixel is thus of critical importance in order to prevent locally large depth errors that may affect subsequent processing tasks and medical decisions. What would be of interest in medical applications is to minimize the coding error at the level of every pixel for a given bit budget. This corresponds to the use of L-infinite metric to drive the coding process.

This paper proposes a novel lossless and near-lossless L-infinite codec for video processing in medical applications, with a focus on elderly monitoring and fall detection. The compression module is of critical importance in reducing the size of streamed video, but in this respect, it is essential that the encoding process introduces either no errors (lossless) or minimal, controlled errors (near-lossless). The work described herein addresses L-infinite compression, thus guaranteeing a bounded distortion

(error) in the reconstructed video. The proposed L-infinite codec is computationally lightweight, being implemented on an embedded platform (depth camera sensor), demonstrating the practical applicability of the proposed system on embedded devices.

The primary contributions of this work are as follows:

- A novel L-infinite codec for depth video data compression that preserves the semantic interpretation of the scene.
- An original quantizer that targets the L-infinite norm for sparse (discontinuous) residual distributions.
- A lightweight encoder optimized for real-time deployment on embedded platforms used in medical applications.

The remainder of this paper is organized as follows. Section 2 reviews the related works. Section 3 details the principal methodology in which we describe the architecture of our codec. Section 4 presents the experimental results following the deployment of the codec. Section 5 discusses the results. Finally, the conclusion is presented in Section 6.

2. Related Works

A lossless compression technique for high-precision depth maps based on pseudo-residuals is introduced in [24]. The entire process involves two sub-processes: pre-processing of depth maps followed by deep lossless compression of the processed maps. The proposed method is then used for the compression of depth images, rather than for streaming of depth video data. Another approach for depth compression is described in [25], where a two-dimensional gradient feature - Corner Points (CPs) is used to simultaneously accelerate the Coding Unit/Prediction Unit (CU/PU) size decision and intra-mode selection in 3D-HEVC. A corner point is defined as a point that has two main directions around its neighborhood and is used to detect a multi-directional edge. An experimental comparison of recent novel video compression standards is detailed in [26]. However, it is essential to note that all these algorithms do not bound the error on the decoder side, which means there is no possibility to control the localized distortion.

Depth medical image compression based on an optimized JPEG-Xt [27] algorithm is introduced in [28]. This algorithm is applied to CT and MRI images for lossless and lossy compression. The algorithm in [29] is proposed for near-lossless and 3D reconstruction using depth-of-field segmentation and uses a method to extract the foreground and encode using the JPEG2000 standard [30].

Among the versions of the JPEG family, the JPEG-LS [31,32] encoder is a standard known for low computational complexity and memory requirements. Consequently, we selected JPEG-LS because it has a more straightforward implementation and lower complexity on embedded devices such as depth camera sensors. The process begins with a predictive step where the value of each pixel is estimated based on neighboring pixels using a simple yet effective prediction model. The differences between the actual and predicted values, known as prediction residuals, are then computed. To adapt to varying image characteristics, JPEG-LS employs a context model that categorizes regions of the image based on local gradients. These residuals are encoded using Golomb-Rice codes, which are optimal for the statistical distributions of prediction residuals, ensuring efficient representation. Finally, it should be noted that JPEG-LS supports both lossless and lossy (near-lossless) encoding. However, despite the performance of these JPEG family standards for depth compression, according to our experiments, they do not achieve satisfying compression rates when they are used as a standalone solution. To improve the performance, we added a preprocessing block to our proposed codec, enhancing overall compression efficiency.

3. Methodology

In this section we describe the architecture of the proposed coding system for video depth data. The proposed codec performs intra-frame coding wherein the frames are encoded independently. This design choice enables low computational complexity, deployment on lightweight embedded

platforms and L-infinite compression of each depth frame when operating in lossy compression mode. Additionally, we use information from successive frames to create a reference background. For the encoding part, we employ two inputs for each video frame. Specifically, the depth frame itself which is a 16-bit depth map, respectively an 8-bit segmentation map associated to each frame. Details are given next.

3.1. Encoder Architecture

Figure 1 shows the main architecture of the encoding process. We split the foreground and background in the first stage, using the segmentation map. Foreground-background segmentation is carried out using a lightweight segmentation whereby humans are semantically segmented from the background. The extracted background is then subtracted from the reference background in order to obtain a residual image. The reference background is another crucial component of the architecture that significantly impacts the compression performance. The reference background is computed based on a specific number of static frames acquired at the beginning of the encoding process. We note that, whenever there is a change in the camera's position, the reference background is updated. To create the reference background, we select the pixels that are labeled as a static object across all the stored frames, then choose the most frequent value as a depth value for the reference background. This approach ensures that the reference background accurately represents the static elements in the scene, facilitating more efficient compression. The residual image then needs to be refined according to the reference background and foreground. In the refinement block, the residual pixels labeled foreground are set to zero.

The next important block is the quantization block. It should be noted that the proposed codec supports two types of compression. For lossless compression we bypass the quantization block, while for near-lossless compression (L-infinite compression) we apply quantization only for the residual image and losslessly encode the foreground image. The last block has the role of losslessly encoding the resulting the residue data and foreground, which we achieve using JPEG-LS encoding. We detail the encoder blocks in the following sections.

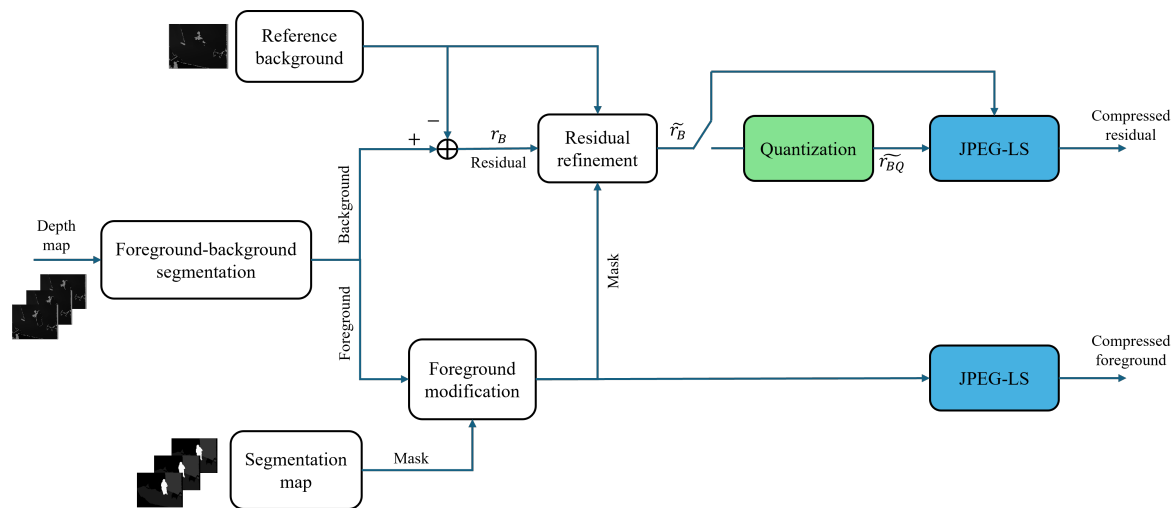


Figure 1. Block diagram architecture of the proposed encoder for depth video compression

3.1.1. Foreground-Background Segmentation

The first block in our architecture is foreground-background segmentation, where we separate each frame into the background and foreground components. In this sense we use a segmentation map generated using a machine-learning algorithm with multitasking objects. First, object segmentation is applied, and then a tracking ID is assigned to each object by multi-tracking across time. According to the labels in the segmentation map, we generate a mask which we subsequently use to extract precise

foreground and background information. The foreground contains only the moving objects of the scene.

3.1.2. Quantization

The quantization process involves mapping pixel values within a specific range (or interval) to a representative value, known as the assignment point. The quantization block helps us to lower the bit rate by merging similar values within an interval, thus reducing data redundancy. In our approach, we employ L-infinite-oriented quantization, which ensures that the error for every single pixel is bound to a specific value, known as the Maximum Absolute Difference (MAXAD). Thus, our coding scheme guarantees that the error for any pixel does not exceed a predefined MAXAD value. Moreover, the L-infinite quantization method helps preserving the semantic integrity of the scene, as the pixel value errors are controlled. As a result, the entire frame can be reconstructed with high local quality at the decoder.

For the proposed coding scheme we have designed piecewise uniform mid-tread quantizers equipped with a deadzone. The quantization method is applied to the residual images which contain both positive and negative values but exhibit a highly sparse (discontinuous) distribution. The piecewise uniform quantizers are adapted to the nature of the input distribution by dropping and merging redundant bins, and then the quantization indices are reassigned. This technique reduces the number of indices in the residual image, resulting in an increase in the compression performance. Furthermore, the foreground image, which contains critical information about the moving objects in the scene, is not quantized and remains losslessly compressed. By this technique, the essential information of the scene, particularly the moving objects, can be reconstructed without losing quality at the decoder.

In the experimental results section we also report compression results obtained by equipping our coding scheme with simple uniform quantizers. In this respect it is important to point out that the method employed to encode the quantized data is slightly different between the two quantization schemes. For uniform quantizers it is sufficient to transmit the bin size to the decoder, followed by encoding of the bin indices. In contrast, for piecewise uniform quantization, the actual assignment points are also transmitted to the decoder, so that the quantized values can be correctly mapped back to their original values in the residual image.

3.1.3. Lossless coding

Lossless coding represents the final stage in our architecture for completing the compression process. The algorithm chosen here is JPEG-LS, because of its highly efficient way of combining predictive compression with entropy coding. Furthermore, compared to other encoders, JPEG-LS has the most straightforward implementation with low complexity on the embedded camera sensor. In our approach, encoding is done on a frame-by-frame basis, applying the JPEG-LS algorithm to both the residual image and the foreground image independently. Additionally, in the near-lossless coding mode, we compress the list of quantization assignment indices using the Zlib [33] library. The compressed bytes are concatenated with a header, forming a complete packet for each frame. The header contains essential metadata, including the packet ID, packet size, MAXAD value, and the compression size for each separated image. The combination of the packet ID and packet size facilitates the identification of each frame's packet within the received stream during decoding. This packet structure facilitates efficient transmission and ensures that both the residual and foreground data can be accurately reconstructed in the decoder.

3.2. Decoder

The decoder block performs the inverse operation of the encoding procedure. It processes the received packets for each frame individually, extracting the residual and foreground images using the information provided in the header. For near-lossless mode, we also rearrange the quantized residual according to the assignment indices. Since the reference background is received as the first packet and

stored, the background of each frame can be reconstructed by adding the residual to the reference background. In addition, the refinements applied during encoding are reversed in this stage. Finally, the foreground and background are combined to reconstruct the frame. This process ensures that the compressed data is decoded efficiently and that the frame is reconstructed accurately. 1

4. Experimental Assessment

In this section, we introduce the dataset we used for our experiments, and explain the implementation of our codec in the embedded platform. Finally, we show the results of our codec compared with other algorithms capable of performing L-infinite compression.

4.1. Dataset

The proposed codec has been evaluated using video datasets captured by the depth camera sensor in various hospital rooms. These datasets include videos from elderly patients, nurses, beds, and other objects present in the rooms in typical monitoring setups for elderly care. Each video sequence comprises 16-bit depth maps for which the spatial resolution is downsampled to 120*160. The depth values are represented in millimeter units. The 8-bit segmentation maps of the datasets are produced using a machine learning algorithm executed before the encoding block, and are employed in order to perform foreground-background segmentation. In principle, the proposed codec can accommodate any foreground-background segmentation method. However, in our experiments, a proprietary semantic segmentation technique trained on video depth data for elderly monitoring applications has been employed. In the segmentation map, each pixel label corresponds to a specific object in the scene. Specifically, label 0 represents non-valid pixels (redundant pixel information of the scene), label 1 corresponds to the ground, label 2 represents the bed, if present in the scenes, and labels 10 and above denote humans, which are our moving objects in the scene. This labeling system improves the precision of the segmentation by clearly identifying static and dynamic elements within the frame. Figure 2 illustrates a sample depth frame of each dataset from 5 different hospital rooms, while Figure 3 shows the corresponding segmentation map of those sample depth frames.

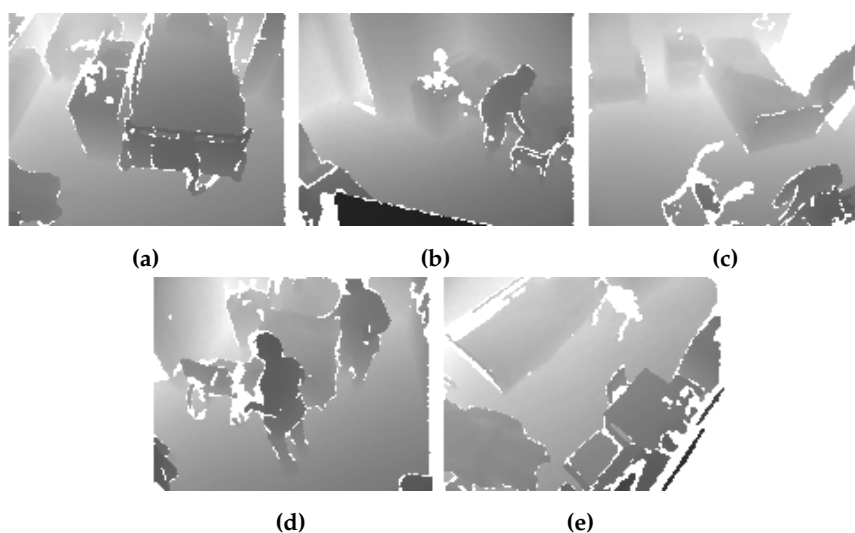
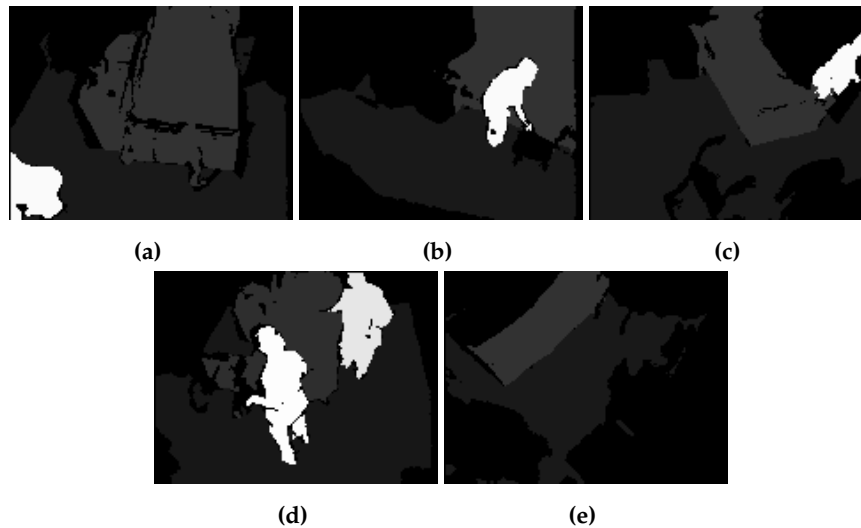


Figure 2. Sample depth map of 5 different hospital rooms: (a) Room 1, (b) Room 2, (c) Room 3, (d) Room 4, (e) Room 5

Table 1. Compression performance (rates in bpp) for lossless coding

Room No.	JPEG-LS	CALIC	Proposed Codec
Room 1	7.616	4.781	4.243
Room 2	7.186	4.915	3.909
Room 3	6.250	4.233	3.170
Room 4	7.177	4.566	4.185
Room 5	7.424	4.079	3.859

**Figure 3.** Sample segmentation map of 5 different hospital rooms: (a) Room 1, (b) Room 2, (c) Room 3, (d) Room 4, (e) Room 5.

4.2. Codec Implementation

Our proposed compression scheme features a lightweight encoder for real-world development on embedded platforms. Therefore, the encoder component capable of performing both lossless and L-infinite near-lossless compression is implemented directly in the depth camera sensor. This depth camera is a light-field sensor capable of capturing RGB, depth, and point cloud data from the scene. The encoder component is implemented using the C++ language in the depth camera, while the decoder part is implemented using Python and runs on the desktop platform. The implementation ensures efficiency in real-time operations and robust processing in the camera, which we will show the runtime results later.

4.3. Experimental Results

The first step of the experiment is conducted in lossless mode, wherein all video frames from the datasets are encoded, the resulting rates (in bpp) being reported in Table 1. We compare our work with two other well-established lossless coding schemes, namely CALIC [34] and JPEG-LS [31].

The results in Table 1 show that our encoder, which integrates a residual-based approach with segmentation and lossless JPEG-LS coding, outperforms the standalone CALIC and JPEG-LS in terms of bit rate. The reported results prove that the additional processing has a significant effect in lowering the bit rate, while losslessly preserving the input data, encoding the semantic segmentation on the input video depth and, as shown later, offering full real-time encoding capabilities on low complexity embedded devices.

In the near-lossless mode of compression, we achieve a higher compression rate by adding the quantization block. Moreover, the use of L-infinite quantizers allows us to bound the error for every single pixel to a specific value. Here, we chose to losslessly preserve the foreground, which contains moving objects or humans, as this type of compression is necessary for accurate fall detection. Tables 2–6 illustrate the results obtained using our proposed L-infinite coding scheme equipped with simple

uniform and piecewise uniform quantizers, respectively, while guaranteeing the depicted MAXAD values throughout the entire video sequence. Similar results are reported for the near-lossless versions of standalone CALIC and JPEG-LS that are operating at the same maximum distortion values.

Table 2. Near-lossless compression results for Room 1. Rates (in bpp) are shown, together with guaranteed maximum distortion (MAXAD) values.

MAXAD	JPEG-LS	CALIC	Proposed L-infinite Codec	
			uniform	piecewise uniform
1	6.621	4.432	3.62	3.029
5	5.056	3.168	2.932	2.678
10	4.304	2.64	2.559	2.456
15	3.822	2.345	2.117	2.091
20	3.465	2.177	1.669	1.676
25	3.166	2.067	1.431	1.433
30	2.922	1.972	1.264	1.277

Table 3. Near-lossless compression results for Room 2. Rates (in bpp) are shown, together with guaranteed maximum distortion (MAXAD) values.

MAXAD	JPEG-LS	CALIC	Proposed L-infinite Codec	
			uniform	piecewise uniform
1	6.302	4.611	3.389	2.975
5	4.768	3.569	2.662	2.421
10	3.994	3.089	2.464	2.297
15	3.51	2.762	2.351	2.215
20	3.215	2.529	2.228	2.12
25	2.964	2.371	2.071	1.988
30	2.774	2.24	1.913	1.849

Table 4. Near-lossless compression results for Room 3. Rates (in bpp) are shown, together with guaranteed maximum distortion (MAXAD) values.

MAXAD	JPEG-LS	CALIC	Proposed L-infinite Codec	
			uniform	piecewise uniform
1	5.54	3.919	2.709	2.356
5	4.335	3.009	2.15	1.959
10	3.653	2.564	1.92	1.815
15	3.151	2.279	1.757	1.692
20	2.8	2.091	1.64	1.587
25	2.514	1.969	1.467	1.431
30	2.301	1.855	1.254	1.248

Table 5. Near-lossless compression results for Room 4. Rates (in bpp) are shown, together with guaranteed maximum distortion (MAXAD) values.

MAXAD	JPEG-LS	CALIC	Proposed L-infinite Codec	
			uniform	piecewise uniform
1	6.32	4.192	3.587	3.07
5	4.73	3.263	2.838	2.551
10	3.978	2.791	2.556	2.381
15	3.576	2.471	2.41	2.286
20	3.3	2.27	2.238	2.14
25	3.063	2.14	1.915	1.852
30	2.876	2.026	1.618	1.588

Table 6. Near-lossless compression results for Room 5. Rates (in bpp) are shown, together with guaranteed maximum distortion (MAXAD) values.

MAXAD	JPEG-LS	CALIC	Proposed L-infinite Codec	
			uniform	piecewise uniform
1	6.722	4.265	3.297	2.86
5	5.386	3.348	2.564	2.317
10	4.655	2.872	2.29	2.16
15	4.199	2.564	2.094	2.02
20	3.892	2.364	1.851	1.811
25	3.604	2.227	1.63	1.592
30	3.36	2.111	1.495	1.483

Tables 2–6 show that for near-lossless L-infinite compression, the proposed codec achieves higher compression rates (lower bit rates) compared to standalone CALIC and JPEG-LS. The use of L-infinite quantizers allows us to successfully lower the bit rate compared to lossless compression, while maintaining a guaranteed bound on the maximum error for each pixel. Additionally, the codec design ensures that the critical part of the scene, which is the foreground, is preserved. Finally, the results demonstrate that piecewise uniform quantization outperforms uniform quantization in terms of compression performance.

In the following, we also compare our proposed L-infinite codec with HEVC. The performance rate of HEVC varies depending on parameters such as the Constant Rate Factor (CRF) and the preset. The preset parameter determines the encoding speed in HEVC, where a slower preset mode increases the compression rate at the cost of increasing encoding time. In the first instance, in Figure 4, we compare the Rate-Distortion (RD) curves obtained using our method, respectively HEVC operating under two encoding speed presets: medium and slow. Specifically, for the same rate we report the maximum distortion introduced across all video frames. The choice of HEVC presets is justified as medium being the typical default value, while slow being the value that offers the highest compression efficiency. Then, in Figure 5, we focus on a specific average MAXAD value of 25 to evaluate the maximum distortion across all frames over time for an equivalent bit rate. This comparative analysis highlights the efficiency of the proposed L-infinite codec in maintaining a lower bounded local distortion while achieving competitive bit rates. As can be seen in these figures, it is evident that HEVC lacks the ability to bound the distortion in the decoder, while our proposed L-infinite codec maintains the quality of the video for each frame. Moreover, at the same bit rate, on average, the proposed codec offers a lower and constant distortion bound.

The implementation of our encoder on the depth camera sensor allows us to show-case the real-time performance of the encoder, which is compiled and optimized for fast and real-time operation in both lossless and near-lossless (L-infinite) modes. Table 7 illustrates the average runtime per frame for the encoder in all different modes, for the five hospital rooms. This table confirms that our proposed

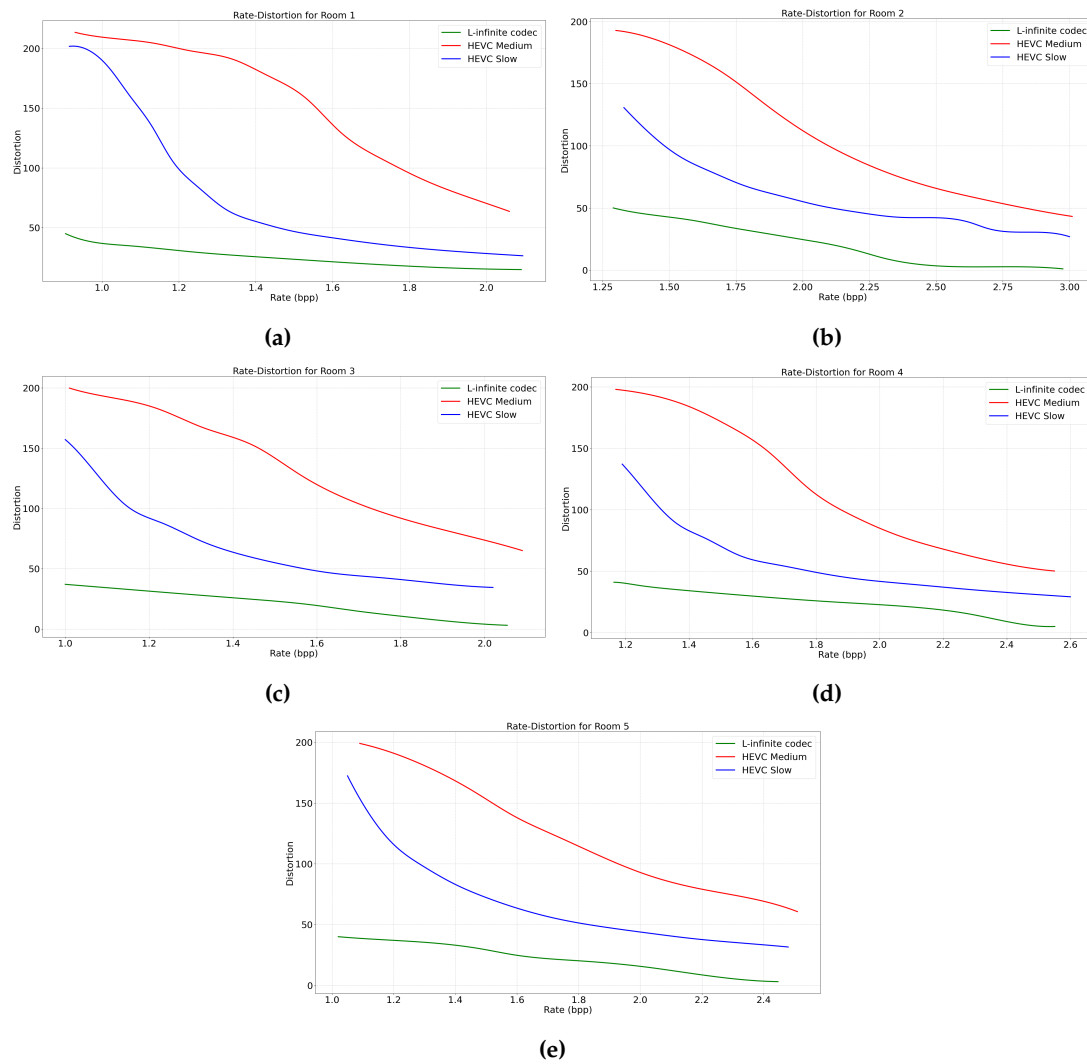


Figure 4. Rate-Distortion curve: (a) Room 1, (b) Room 2, (c) Room 3, (d) Room 4, (e) Room 5

codec operates with low computational complexity and is capable of compressing video data in real time. The low runtime per frame, even in near-lossless (L-infinite) mode, demonstrates the efficiency of the encoder. This efficient performance is crucial for applications like fall detection in medical settings, where both speed and accuracy are essential.

Table 7. Average runtime per frame of encoder, for different modes

Room No.	Lossless (ms)	Proposed L-infinite Codec (ms)	
		uniform	piecewise uniform
Room 1	6.3	8	9.5
Room 2	6.1	7.2	9
Room 3	5.2	6.5	8
Room 4	6	7.2	8.8
Room 5	5.7	7	8.4

5. Discussion

In the first set of experiments, for all datasets, we compared the performance of our proposed method against the standalone implementations of the JPEG-LS and CALIC algorithms. In lossless mode, we observe that the proposed codec clearly outperforms both JPEG-LS and CALIC, for which we report an average improvement of 45% in bit rate over JPEG-LS and 17% over CALIC respectively.

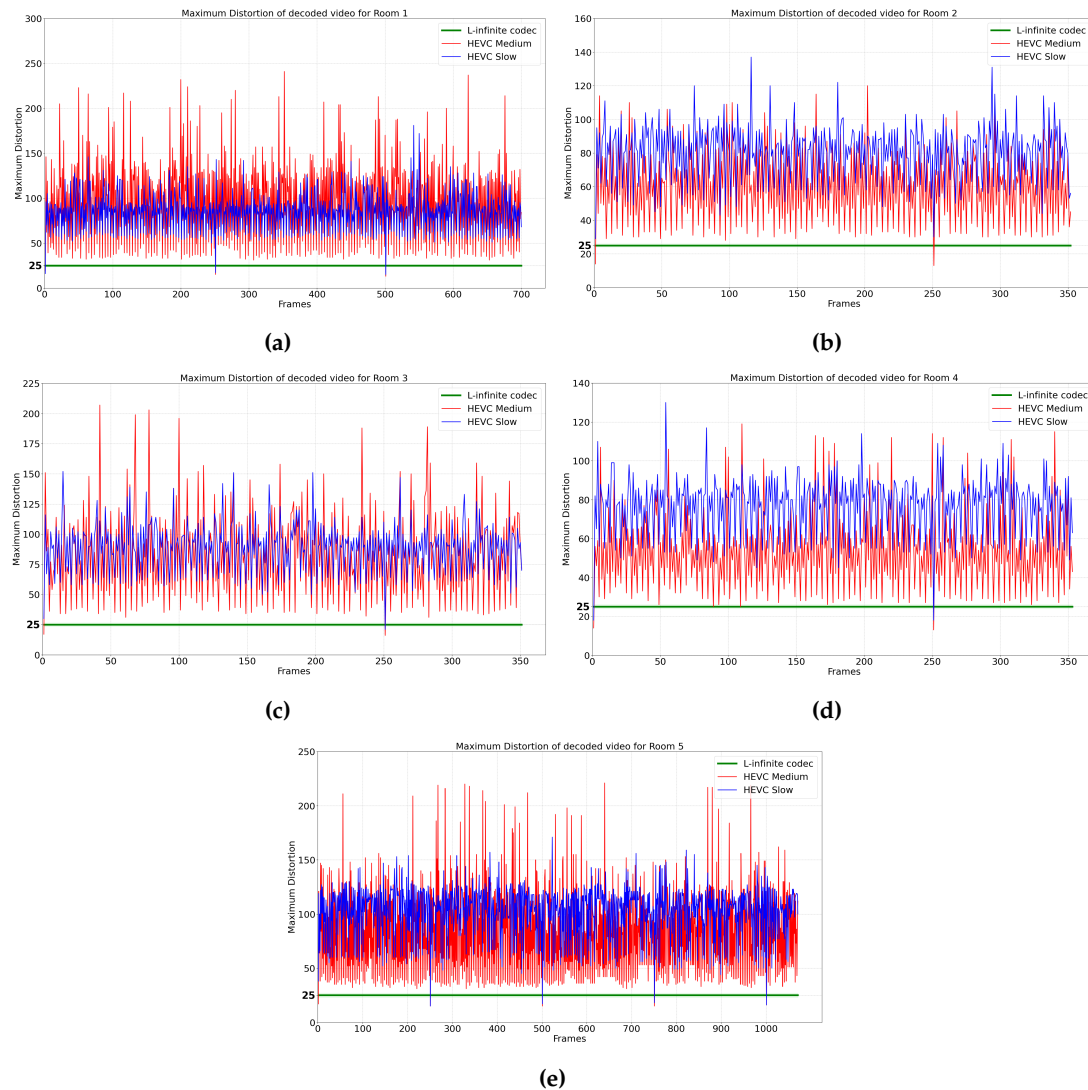


Figure 5. Maximum distortion of video across all frames: (a) Room1, (b) Room 2, (c) Room 3, (d) Room 4, (e) Room 5.

A more detailed examination of the L-infinite mode across different datasets also shows significant gains. In Room 1, our method obtained an average improvement of 50.4% over JPEG-LS and 22.1% over CALIC, for various MAXAD values. For Room 2, the improvements are 40.2% over JPEG-LS and 23.2% over CALIC; for Room 3, 48.7% over JPEG-LS and 30.5% over CALIC; for Room 4, 41.8% over JPEG-LS and 15.9% over CALIC; and for Room 5, 55% over JPEG-LS and 27.3% over CALIC. These results prove the significant impact of L-infinite quantization employed by the proposed encoder.

In the second set of experiments, it should be noted that HEVC can achieve varying compression rates depending on the encoder parameters. We compared the maximum distortion across all frames for each rate, and also we showed the rate-distortion curve for each dataset. The results show that HEVC cannot bound the distortion while our proposed L-infinite codec consistently achieves lower and guaranteed L-infinite distortion compared to HEVC.

Furthermore, the proposed lightweight codec is shown to be highly efficient and capable of real-time operation when implemented on a depth camera sensor, making it suitable for practical applications in medical environments. The runtime figures for the proposed codec are reported in Table 7 for all datasets in all different modes. We note that, while the slower mode of HEVC provides a lower bit-rate, it also costs higher encoding time, which makes it unsuitable for real-time applications on embedded devices.

6. Conclusion

In this paper, we have proposed a novel lightweight compression system for video depth data. Our codec is designed to operate in lossless or L-infinite near-lossless compression. In the latter case, we bound the distortion for each pixel by a specific value referred to as the Maximum Absolute Difference (MAXAD). This feature has the ability to control the coding error on each depth pixel and to preserve the semantic interpretation of the scene, which is crucial in medical applications. In both modes, we show that the proposed codec outperforms state-of-the-art coding schemes. Furthermore, our codec performs real-time compression when implemented in an embedded platform such as a depth camera sensor, making it practical for real-world use.

7. Acknowledgment

This work is funded by Innoviris within the research project MUSCLES funded by Innoviris and by the National Program for Research of the National Association of Technical Universities - GNAC ARUT 2023.

References

1. Dhawan, S. A review of image compression and comparison of its algorithms. *International Journal of electronics & Communication technology* **2011**, 2, 22–26.
2. Me, S.S.; Vijayakumar, V.; Anuja, R. A survey on various compression methods for medical images. *International journal of intelligent systems and applications* **2012**, 4, 13.
3. Ukrit, M.F.; Umamageswari, A.; Suresh, G. A survey on lossless compression for medical images. *International Journal of Computer Applications* **2011**, 31, 47–50.
4. Amri, H.; Khalfallah, A.; Gargouri, M.; Nebhani, N.; Lapayre, J.C.; Bouhlef, M.S. Medical image compression approach based on image resizing, digital watermarking and lossless compression. *Journal of Signal Processing Systems* **2017**, 87, 203–214.
5. Ansari, M.; Anand, R. Comparative analysis of medical image compression techniques and their performance evaluation for telemedicine. In Proceedings of the Proceedings of the International Conference on Cognition and Recognition. Citeseer, 2014, pp. 670–677.
6. Clunie, D.A. Lossless compression of grayscale medical images: effectiveness of traditional and state-of-the-art approaches. In Proceedings of the Medical Imaging 2000: PACS Design and Evaluation: Engineering and Clinical Issues. SPIE, 2000, Vol. 3980, pp. 74–84.
7. Alzahrani, M.; Albinali, M. Comparative Analysis of Lossless Image Compression Algorithms based on Different Types of Medical Images. In Proceedings of the 2021 International Conference of Women in Data Science at Taif University (WiDSTaif). IEEE, 2021, pp. 1–6.
8. Hassan, S.A.; Hussain, M. Spatial domain lossless image data compression method. In Proceedings of the 2011 International Conference on Information and Communication Technologies. IEEE, 2011, pp. 1–4.
9. SHARMA, K.; KUMAR, N. Spatial Domain Lossless Image Data Compression **2018**.
10. Jangbari, P.; Patel, D. Review on region of interest coding techniques for medical image compression. *International Journal of Computer Applications* **2016**, 134, 1–5.
11. Ungureanu, V.I.; Negirla, P.; Korodi, A. Image-Compression Techniques: Classical and “Region-of-Interest-Based” Approaches Presented in Recent Papers. *Sensors* **2024**, 24, 791.
12. Tabus, I.; Schioppa, I.; Astola, J. Context coding of depth map images under the piecewise-constant image model representation. *IEEE Transactions on Image Processing* **2013**, 22, 4195–4210.
13. Schioppa, I.; Tabus, I. Anchor points coding for depth map compression. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014, pp. 5626–5630.
14. Schioppa, I.; Tabus, I. Lossy depth image compression using greedy rate-distortion slope optimization. *IEEE Signal Processing Letters* **2013**, 20, 1066–1069.
15. Schioppa, I.; Tabus, I. Lossy-to-lossless progressive coding of depth-maps. In Proceedings of the 2015 International Symposium on Signals, Circuits and Systems (ISSCS). IEEE, 2015, pp. 1–4.
16. Lin, J.L.; Chen, Y.W.; Chang, Y.L.; An, J.; Zhang, K.; Huang, Y.W.; Lei, S. Advanced texture and depth coding in 3D-HEVC. *Journal of Visual Communication and Image Representation* **2018**, 50, 83–92.
17. Hwang, S.H.; Kim, K.M.; Kim, S.; Kwak, J.W. Lossless data compression for time-series sensor data based on dynamic bit packing. *Sensors* **2023**, 23, 8575.

18. Wiegand, T.; Sullivan, G.J.; Bjontegaard, G.; Luthra, A. Overview of the H. 264/AVC video coding standard. *IEEE Transactions on circuits and systems for video technology* **2003**, *13*, 560–576.
19. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on circuits and systems for video technology* **2012**, *22*, 1649–1668.
20. Bross, B.; Wang, Y.K.; Ye, Y.; Liu, S.; Chen, J.; Sullivan, G.J.; Ohm, J.R. Overview of the versatile video coding (VVC) standard and its applications. *IEEE Transactions on Circuits and Systems for Video Technology* **2021**, *31*, 3736–3764.
21. Zakariya, S.; Inamullah, M.; et al. Analysis of video compression algorithms on different video files. In Proceedings of the 2012 Fourth International Conference on Computational Intelligence and Communication Networks. IEEE, 2012, pp. 257–262.
22. Deshpande, R.G.; Ragha, L.L. Performance analysis of various video compression standards. In Proceedings of the 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC). IEEE, 2016, pp. 148–151.
23. Grois, D.; Marpe, D.; Mulyoff, A.; Itzhaky, B.; Hadar, O. Performance comparison of h. 265/mpeg-hevc, vp9, and h. 264/mpeg-avc encoders. In Proceedings of the 2013 Picture Coding Symposium (PCS). IEEE, 2013, pp. 394–397.
24. Wu, Y.; Gao, W. End-to-end lossless compression of high precision depth maps guided by pseudo-residual. *arXiv preprint arXiv:2201.03195* **2022**.
25. Fu, C.H.; Chan, Y.L.; Zhang, H.B.; Tsang, S.H.; Xu, M.T. Efficient depth intra frame coding in 3D-HEVC by corner points. *IEEE Transactions on Image Processing* **2020**, *30*, 1608–1622.
26. Petreski, D.; Kartalov, T. Next Generation Video Compression Standards–Performance Overview. In Proceedings of the 2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, 2023, pp. 1–5.
27. Srinivasan, S.; Tu, C.; Regunathan, S.L.; Sullivan, G.J. HD Photo: a new image coding technology for digital photography. In Proceedings of the Applications of Digital Image Processing XXX. SPIE, 2007, Vol. 6696, pp. 101–119.
28. Li, Z.; Ramos, A.; Li, Z.; Osborn, M.L.; Li, X.; Li, Y.; Yao, S.; Xu, J. An optimized JPEG-Xt-based algorithm for the lossy and lossless compression of 16-bit depth medical image. *Biomedical Signal Processing and Control* **2021**, *64*, 102306.
29. von Buelow, M.; Tausch, R.; Schurig, M.; Knauthe, V.; Wirth, T.; Guthe, S.; Santos, P.; Fellner, D.W. Depth-of-Field Segmentation for Near-lossless Image Compression and 3D Reconstruction. *Journal on Computing and Cultural Heritage (JOCCH)* **2022**, *15*, 1–16.
30. Christopoulos, C.; Skodras, A.; Ebrahimi, T. The JPEG2000 still image coding system: an overview. *IEEE transactions on consumer electronics* **2000**, *46*, 1103–1127.
31. Weinberger, M.J.; Seroussi, G.; Sapiro, G. The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS. *IEEE Transactions on Image processing* **2000**, *9*, 1309–1324.
32. Rane, S.D.; Sapiro, G. Evaluation of JPEG-LS, the new lossless and controlled-lossy still image compression standard, for compression of high-resolution elevation data. *IEEE Transactions on Geoscience and Remote sensing* **2001**, *39*, 2298–2306.
33. Deutsch, L.; Gailly, J. RFC 1950: ZLIB compressed data format specification version 3.3, May 1996. *Status: INFORMATIONAL*.
34. Wu, X.; Memon, N. CALIC-a context based adaptive lossless image codec. In Proceedings of the 1996 IEEE international conference on acoustics, speech, and signal processing conference proceedings. IEEE, 1996, Vol. 4, pp. 1890–1893.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.