SoftwareX 28 (2024) 101961

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Software Update

Version [7.1] – [IV-PSNR: Software for immersive video objective quality evaluation]

Jakub Stankowski, Adrian Dziembowski *

Institute of Multimedia Telecommunications, Poznan University of Technology, 60-965 Poznań, Poland

A R T I C L E I N F O

Keywords:
Image quality assessment
Video quality assessment
Immersive video
Structural similarity

A B S T R A C T

This paper describes a new version of the IV-PSNR software, developed for the effective objective quality assessment of immersive video. Version 7.1 includes the calculation of structural similarity between compared sequences using the IV-SSIM metric, designed to properly handle the unique characteristics of immersive video, as well as the classic SSIM and MS-SSIM metrics. Moreover, by introducing new modes, IV-PSNR 7.1 is adapted to assess the quality of novel approaches to multiview video processing, based on radiance fields and implicit neural visual representations. Currently, this version of the software is used by the ISO/IEC MPEG VC standardization group for the evaluation of the second edition of the MIV coding standard, and in works aimed at the development of a future standard for radiance field representation and compression.

Metadata

Nr	Code metadata description	*Please fill in this column*
C1	Current code version	v7.1
C2	Permanent link to code/ repository used for this code version	https://github.com/jstankowski/ivpsnr
C3	Permanent link to reproducible capsule	https://github.com/jstankowski/ivpsnr/tree/main/example
C4	Legal code license	BSD-3-Clause
C5	Code versioning system used	git
C6	Software code languages, tools and services used	C++
C7	Compilation requirements, operating environments and dependencies	CMake 3.15 or newer, C++17 conformant compiler (e.g., GCC 8.0 or newer, clang 5.0 or newer, MSVC 19.15 or newer), external libraries: fmtlib, libspng, and miniz – automatically downloaded during CMake build process
C8	If available, link to developer documentation/ manual	https://github.com/jstankowski/ivpsnr#readme
C9	Support email for questions	jakub.stankowski@put.poznan.pl

1. Description of the software-update

The main application of the IV-PSNR software is objective quality assessment for immersive video systems, where a user can freely navigate within a three-dimensional scene [1,2] watching video created by a view rendering process [3]. Such video is characterized by specific artifacts, such as slight shifts of reprojected pixels and global changes in scene illumination [4]. The IV-PSNR software is designed to efficiently handle those artifacts providing reliable immersive video objective quality, highly correlated with the subjective perception of the rendered views.

This paper describes IV-PSNR v7.1, which introduced four major changes, including an extended list of quality metrics (Section 1.1), support for reading PNG files (Section 1.2), color space conversion mode (Section 1.3), and several performance improvements (Section 1.4). These changes significantly increased the usability of the software, adapting it for various applications requiring pixel reprojection (Section 1.5).

1.1. Extended list of quality metrics

Early versions of the IV-PSNR software (up to v5.0 published in [5])

DOI of original article: https://doi.org/10.1016/j.softx.2023.101592.

Refers to Jakub Stankowski, Adrian Dziembowski, IV-PSNR: Software for immersive video objective quality evaluation, SoftwareX, vol. 24, 101,592, December 2023, DOI:10.1016/j.softx.2023.101592.

* Corresponding author.
E-mail addresses: jakub.stankowski@put.poznan.pl (J. Stankowski), adrian.dziembowski@put.poznan.pl (A. Dziembowski).

https://doi.org/10.1016/j.softx.2024.101961
Received 22 August 2024; Received in revised form 26 September 2024; Accepted 29 October 2024
Available online 13 November 2024

included the calculation of the IV-PSNR metric (Peak Signal-to-Noise Ratio for Immersive Video) [3], as well as two other state-of-the-art metrics commonly used for immersive video: PSNR and WS-PSNR (Weighted-to-Spherically-uniform PSNR) [6]. While the IV-PSNR metric can be considered highly correlated with subjective quality for immersive video [1] applications, it is based on PSNR, which performs much worse than metrics based on structural similarity estimation [7].

Therefore, the Authors developed a new metric, combining the advantages of structural similarity and IV-PSNR, resulting in the IV-SSIM metric (Structural Similarity Index Measure for Immersive Video) [8], which clearly outperforms state-of-the-art metrics in immersive video applications (Fig. 1A) and performs well in other scenarios (Fig. 1B). Given the high correlation between IV-SSIM and subjective video quality, the new version of the IV-PSNR software (v7.1) includes the calculation of IV-SSIM. As a side effect, since the implementation was already completed, the software outputs also the value of the state-of-the-art SSIM [7] and MS-SSIM (Multi-Scale SSIM) [9] metrics.

As presented in Fig. 1, IV-SSIM [8] not only outperforms the IV-PSNR metric in immersive video assessment but also is more robust in typical, non-immersive applications making it a more versatile metric. Therefore, the authors believe that the new version of the IV-PSNR software is more useful regardless of the use case.

1.2. PNG input mode

In addition to extending the metric list for better correlation with subjective quality, the new version of the IV-PSNR software includes two features that enhance its usefulness in different multi-view video applications: working with PNG files and the RGB color space.

By default, the IV-PSNR software compares two raw video files in the YUV format. Since v7.1, however, the software also allows quality assessment for lists of PNG files. This change was motivated by the need to adapt the IV-PSNR software for the purposes of implicit neural visual representations (INVR), where output files are commonly stored as still PNG images [16,17].

1.3. Color space mode

Originally, the IV-PSNR software was designed for accurate quality assessment of immersive video coding, thus it assumed that the input videos use the YCbCr color space, where the color of a pixel is defined by luma Y and two chromas: Cb and Cr. This color space is commonly used in video transmission and video processing applications [27]. In other applications, however, such an assumption may not be correct, and the files may use the RGB color space (e.g., if the PNG files are used).

Therefore, in order to avoid the need to use external conversion tools, which may introduce additional errors and generate unreproducible results, IV-PSNR v7.1 is able to operate both in YCbCr and RGB color space. By default, the software assumes the input files are in a generic YCbCr color space and calculates the metrics in the same color space. When the color space mode is enabled, the IV-PSNR software may accept RGB input files and calculate the metrics in the RGB color space, or it can perform a color space conversion (YCbCr input files assessed in the RGB color space or RGB files assessed in the YCbCr color space). The software incorporates several conversion algorithms, compliant with commonly used YCbCr color space definitions: BT.601 [18], BT.709 [19], and BT.2020 [20].

1.4. Performance improvements

Besides algorithmic improvements, the Authors have focused on further increasing the performance of the IV-PSNR software, i.e., on the reduction of the time required for quality assessment.

Version 6.0 introduced a fast SIMD (Single Instruction, Multiple Data) implementation of the Kahan-Babuška-Neumaier accumulation, allowing for the use of 256-bit SIMD vectors (containing four double-precision numbers). Such an implementation is much faster than the scalar one due to the parallel calculation of four values and two paths, from which the result is selected. In the scalar implementation, such behavior is achieved by using a computationally demanding branching mechanism.

Additionally, the “reasonable auto” mode was introduced for determining the number of threads used for quality metrics calculation. Previously, the software used all available threads by default. This behavior was not optimal for high core-count systems, as merging the results from many threads required a significant amount of time. The reasonable auto mode limits the number of used threads for PSNR-based metrics to eight, and for structural similarity metrics – which are better scalable – to sixty-four.

The third improvement – introduced in version 7.0 – reduced the time measurement complexity by replacing the std::chrono::high_resolution_clock with the RDTSCP (ReaD Time-Stamp Counter and Processor ID) instruction [21], which is more precise and has less overhead. The RDTSCP instruction is available on all modern x86–64 platforms, while on others the software falls back to the high_resolution_clock method.

Finally, all added structural similarity metrics, and the new color space conversion mode are implemented with the use of efficient SIMD processing, including SSE (Streaming SIMD Extensions), AVX2 (Advanced Vector Extensions), and AVX512. The results output by the

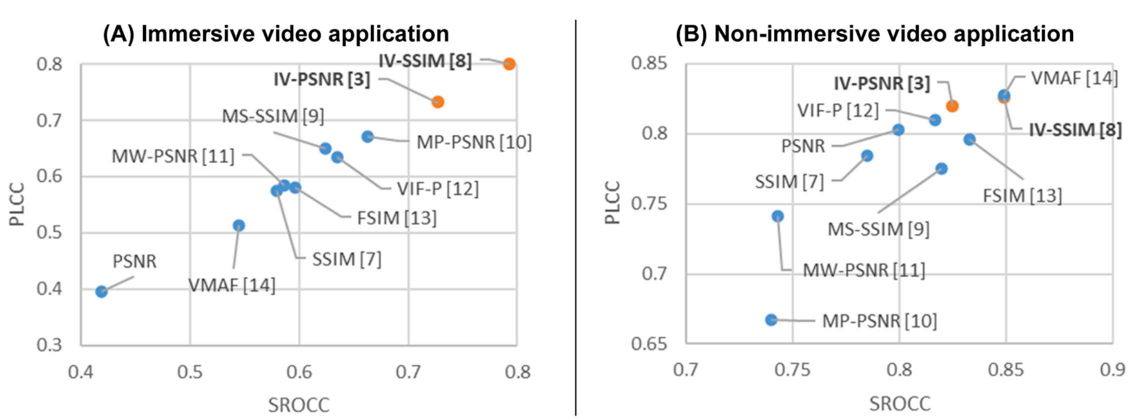

Fig. 1. Correlation between objective and subjective quality – PLCC (Pearson Linear Correlation Coefficient) and SROCC (Spearman Rank Order Correlation Coefficient) [15] values for considered metrics in two applications: A: immersive video, B: non-immersive video. Results from [8]. State-of-the-art metrics used in comparison: PSNR, IV-PSNR [3], SSIM [7], IV-SSIM [8], MS-SSIM [9], MP-PSNR (Morphological Pyramid PSNR) [10], MW-PSNR (Morphological Wavelet PSNR) [11], VIF-P (Pixel-based Visual Information Fidelity) [12], FSIM (Feature Similarity Index Measure) [13], and VMAF (Video Multimethod Assessment Fusion) [14].

IV-PSNR software are exactly the same, regardless of the instruction set used.

1.5. Usability

The IV-PSNR software was initially created for the purpose of developing the ISO/IEC 23090–12 MPEG immersive video (MIV) coding standard [22]. It was included in the MIV common test conditions (MIV CTC) [23] and widely used for immersive video coding research. The modifications presented in this paper have expanded its use beyond immersive video coding, making it suitable for other multiview video applications such as radiance field representation and compression. Currently, it is used (under the name QMIV – Quality Metrics for Immersive Video [24]) for the evaluation of the second edition of MIV [25] and supports ongoing ISO/IEC MPEG's works aimed at the development of a novel radiance fields coding standard [26].

CRediT authorship contribution statement

Jakub Stankowski: Writing – review & editing, Software, Conceptualization. **Adrian Dziembowski:** Writing – original draft, Validation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jakub Stankowski reports financial support was provided by Ministry of Education and Science of the Republic of Poland. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by the Ministry of Science and Higher Education of Republic of Poland.

Data availability

No data was used for the research described in the article.

References

[1] Wien M, Boyce J, Stockhammer T, Peng WH. Standardization status of immersive video coding. IEEE J Emerging Selected Topics Circuits Systems 2019;9(9):5–17.
[2] Boyce J, et al. MPEG immersive video coding standard. Proceede IEEE 2021;119 (9):1521–36.
[3] Dziembowski A, Mieloch D, Stankiewicz O, Domański M. Virtual view synthesis for 3DoF+ video. In: 2019 Picture Coding Symposium (PCS); 2019. p. 1–5.
[4] Dziembowski A, Mieloch D, Stankowski J, Grzelka A. IV-PSNR – the objective quality metric for immersive video applications. IEEE Transact Circuit Syst Video Technol 2022;32(11):7575–91.
[5] Stankowski J, Dziembowski A. IV-PSNR: software for immersive video objective quality evaluation. SoftwareX 2023;24(1-6):101592.
[6] Sun Y, Lu A, Yu L. Weighted-to-spherically-uniform quality evaluation for omnidirectional video. IEEE Signal Process Lett 2017;24(9):1408–12.
[7] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error measurement to structural similarity. IEEE Transact Image Process 2004;13(4): 600–12.
[8] Dziembowski A, Nowak W, Stankowski J. IV-SSIM – the structural similarity metric for immersive video. Applied Sciences 2024;14(16):1–14.
[9] Wang Z, Simoncelli EP, Bovik AC. Multiscale structural similarity for image quality assessment. In: 37th Asilomar Conference on Signals, Systems & Computers; 2003. p. 1398–402.
[10] Sandić-Stanković D, Kukolj D, Callet PLe. DIBR synthesized image quality assessment based on morphological pyramids. In: 3DTV-CON; 2015. p. 1–4.
[11] Sandić-Stanković D, Kukolj D, Callet PLe. DIBR-synthesized image quality assessment based on morphological multi-scale approach. EURASIP J Image Video Process 2016;2017(1):1–23.
[12] Sheikh HR, Bovik AC. Image information and visual quality. IEEE Transact Image Process 2006;15(2):430–44.
[13] Zhang L, Zhang L, Mou X, Zhang D. FSIM: a feature similarity index for image quality assessment. IEEE Transact Image Process 2011;20(8):2378–86.
[14] The Netflix Tech Blog, 2016. Accessed: Sep. 2023. [Online]. Available: https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652.
[15] Chikkerur S, Sundaram V, Reisslein M, Karam L. Objective video quality assessment methods: a classification, review, and performance comparison. IEEE Transactions on Broadcasting 2011;57(2):165–82.
[16] Singh R, Shukla A, Turaga P. Polynomial implicit neural representations for large diverse datasets. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023. p. 2041–51.
[17] Bang G, Do J, Kang JW, Bea SJ, Lee H, Lee J, Kim S. Implicit neural visual representation compression of 3D scenes. In: International Workshop on Advanced Imaging Technology (IWAIT); 2023.
[18] ITU-R. Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios. Recommendat ITU-R BT 2011;601.
[19] ITU-R. Parameter values for the HDTV standards for production and international programme exchange. Recommendation ITU-R BT 2015;709.
[20] ITU-R. Parameter values for ultra-high definition television systems for production and international programme exchange. Recommendation ITU-R BT 2015;2020.
[21] Intel, "Intel 64 and IA-32 Architectures Software Developer's Manual," 2024.
[22] Vadakital VKM, et al. The MPEG Immersive Video Standard – Current Status and Future Outlook. IEEE MultiMedia 2022;29(3):101–11.
[23] ISO/IEC. Common test conditions for MPEG immersive video. In: Document ISO/IEC JTC1/SC29/WG4 MPEG VC N0539; 2024.
[24] ISO/IEC. Software manual of QMIV. In: Document ISO/IEC JTC1/SC29/WG4 MPEG VC N0535; 2024.
[25] ISO/IEC. *Text of ISO/IEC DIS 23090-12 MPEG immersive video* (2nd edition). In: Document ISO/IEC JTC1/SC29/WG4 MPEG VC N0533; 2024.
[26] ISO/IEC. Common test conditions on radiance field representation and compression. In: Document ISO/IEC JTC1/SC29/WG4 MPEG VC N0561; 2024.
[27] Poynton C. Digital video and HDTV: algorithms and interfaces. San Francisco, USA: Morgan Kaufmann Publishers; 2003. p. 94–123.