ISO/IEC JTC 1/SC 29/WG 04

MPEG Video Coding

Convenorship: CN

Document type:	Output Document
Title:	Software manual of IV-PSNR for Immersive Video
Status:	Approved
Date of document:	2020-10-23
Source:	ISO/IEC JTC 1/SC 29/WG 04
Expected action:	None
Action due date:	None
No. of pages:	5 (without cover page)
Email of Convenor:	yul@zju.edu.cn
Committee URL:	https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg4

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 04 MPEG VIDEO CODING

Title	**Software manual of IV-PSNR for Immersive Video**
Source	**WG 04, MPEG Video Coding**
Status	**Approved**
Editor	**Adrian Dziembowski**
Serial Number	**19743**

1. Introduction

IV-PSNR [M48093] is a PSNR-based objective quality metric adapted for Immersive Video applications. Compared to PSNR, two major modifications were added: Corresponding Pixel Shift and Global Color Difference. Corresponding Pixel Shift eliminates the influence of a slight shift of objects' edges caused by reprojection errors. Global Color Difference reduces the influence of different color characteristics of different input views.

IV-PSNR for YUV file is calculated as:

$$\text{IVPSNR}_{\text{YUV}} = \frac{\sum_{c=0}^{2} \text{IVPSNR}(c) \cdot \text{CCW}(c)}{\sum_{c=0}^{2} \text{CCW}(c)} \, ,$$

where $\text{CCW}(c)$ is the Color Component Weight for each color component c and $\text{IVPSNR}(c)$ is the IV-PSNR for that component:

$$\text{IVPSNR}(c) = 10 \cdot \log\left(\frac{\text{MAX}^2}{\text{IVMSE}(c)}\right),$$

where MAX is the maximum value of the color component (e.g. 1023 for 10-bit video) and:

$$\text{IVMSE}(c) = \frac{1}{W \cdot H} \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \min_{\substack{x_R \in [x-\text{CPS},x+\text{CPS}] \\ y_R \in [y-\text{CPS},y+\text{CPS}]}} (c_T(x,y,c) - c_R(x_R,y_R,c) + \text{GCD(c)})^2 \, ,$$

where W and H are width and height of the image, $c_T(x,y,c)$ and $c_R(x,y,c)$ are values of color component c in the position (x,y) in the test image and the reference image, respectively, CPS is the maximum Corresponding Pixel Shift between reference and test image, and GCD is the Global Color Difference for component c:

$$\text{GCD}(c) = \max\left(\frac{1}{W \cdot H} \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \left(c_R(x,y,c) - c_T(x,y,c)\right), \text{MUD}(c)\right),$$

where $\text{MUD}(c)$ is the Maximum Unnoticeable Difference for color component c.

In order to provide better quality assessment for omnidirectional video, WS-PSNR technique [Sun17] was applied (however, in the current version of the IV-PSNR software only the equirectangular projection is supported):

$$\text{WS-IVMSE}(c) = \frac{\displaystyle\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \min_{\substack{x_R \in [x-\text{CPS}, x+\text{CPS}] \\ y_R \in [y-\text{CPS}, y+\text{CPS}]}} \left(c_T(x,y,c) - c_R(x_R, y_R, c) + \text{GCD}(c)\right)^2 \cdot w_{x,y}}{\displaystyle\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} w_{x,y}},$$

where weight $w_{x,y}$ is calculated as:

$$w_{x,y} = \cos\frac{\left(y + 0.5 - \frac{H}{2}\right) \cdot \pi}{H},$$

where x, y is a position of the pixel in ERP image and H is the height of this image.

$\text{CCW}(c)$, $\text{MUD}(c)$ and CPS values are predefined:
- $\text{CCW}(c)$:
  - $\text{CCW}(0) = 1$ (luma component),
  - $\text{CCW}(1) = 0.25$ (1st chroma component),
  - $\text{CCW}(2) = 0.25$ (2nd chroma component),
- $\text{MUD}(c) = 1\%$ for all the color components,
- $\text{CPS} = 2$.

IV-PSNR is calculated separately for each frame of the sequence. In the end, the mean IV-PSNR value is returned.

The IV-PSNR quality metric is based on PSNR, therefore, the higher the number, the better is the quality.

2. Software manual

IV-PNSR v2.1 accepts parameters listed in table below:

General parameters

Cmd	ParamName	Description
-i0	InputFile0	YUV file path – reference
-i1	InputFile1	YUV file path – tested
-w	PictureWidth	Width of sequence
-h	PictureHeight	Height of sequence
-bd	BitDepth	Bit depth (optional, default: 8, up to 14)
-cf	ChromaFormat	Chroma format (optional, default: 420) [420, 444]
-s0	StartFrame0	Start frame (optional, default: 0)
-s1	StartFrame1	Start frame (optional, default: 0)
-l	NumberOfFrames	Number of frames to be processed (optional, default: -1 = all)
-o	OutputFile	Output file path (optional)

Equirectangular parameters

Cmd	ParamName	Description
-erp	Equirectangular	Equirectangular sequence (flag, default disabled)
-lor	LonRangeDeg	Longitudinal range of ERP sequence [°] (optional, default: 360)
-lar	LatRangeDeg	Lateral range of ERP sequence [°] (optional, default: 180)

Application parameters

Cmd	ParamName	Description
-t	NumberOfThreads	Number of worker threads if compiled with OpenMP (optional, default: -1 = all, suggested 4-8)
-v	VerboseLevel	Verbose level (optional, default: 2)

VerboseLevel description

Value	Printed data
0	final PSNR, WS-PSNR, IV-PSNR values only
1	0 + configuration + detected number of frames
2	1 + argc/argv + frame level PSNR, WS-PSNR, IV-PSNR
3	2 + computing time (LOAD, PSNR, WS-PSNR, IV-PSNR) (uses high resolution clock, could slightly slow down computations)
4	3 + IV-PSNR specific debug data (GlobalColorShift, R2T+T2R)

- The commandline parameters are position-intependent.
- When no parameters are used, syntax help is outputted.

3. Examples

1. IV-PSNR of SA_ref.yuv and SA_test.yuv. Sequence resolution is 4096×2048, YUV420, 10 bits per sample. Sequence format is ERP. Mean IV-PSNR calculated for the first 20 frames will be written into IV-PSNR.txt:

```
IV-PSNR -i0 SA_ref.yuv -i1 SA_test.yuv -w 4096 -h 2048 -bd 10 -erp -l 20 -o IV-PSNR.txt
```

2. IV-PSNR of SD_ref.yuv and SD_test.yuv. Sequence resolution is 2048×1088, YUV420, 8 bits per sample. Sequence format is perspective. Mean IV-PSNR calculated for all frames will be written into results.txt:

```
IV-PSNR -i0 SD_ref.yuv -i1 SD_test.yuv -o results.txt -w 2048 -h 1088
```

3. IV-PSNR of SC_ref.yuv and SC_test.yuv. Sequence resolution is 4096×2048, YUV420, 10 bits per sample. Sequence format is ERP, with lateral range equal to 90°. Mean IV-PSNR calculated for 5 frames (frames 0-4 of reference video and 10-14 of test video) will be written into o.txt:

```
IV-PSNR -i0 SC_ref.yuv -i1 SC_test.yuv -w 4096 -h 2048 -erp -lar 90 -l 5 -s1 10 -o o.txt
```

4. Software

MPEG Git Repository: http://mpegx.int-evry.fr/software/MPEG/MIV/RS/IVPSNR
Public read-only access: https://gitlab.com/mpeg-i-visual/ivpsnr
Software coordinator: Adrian Dziembowski, adrian.dziembowski@put.poznan.pl

5. References

[M48093]   A. Dziembowski, M. Domański,
            "[MPEG-I Visual] Objective quality metric for immersive video",
            ISO/IEC JTC1/SC29/WG11 MPEG/M48093, July 2019, Göteborg, Sweden.
[M54279]   J. Stankowski, A. Dziembowski,
            "[MPEG-I Visual] Fast implementation of IV-PSNR software",
            ISO/IEC JTC1/SC29/WG11 MPEG/M54279, July 2020, Online.
[M54896]   J. Stankowski, A. Dziembowski,
            "Even faster implementation of IV-PSNR software",
            ISO/IEC JTC1/SC29/WG04 MPEG/M54896, October 2020, Online.
[Sun17]   Y. Sun, A. Lu, L. Yu,
            "Weighted-to-Spherically-Uniform Quality Evaluation for Omnidirectional
            Video", IEEE Signal Processing Letters 24.9(2017):1408-1412.

6. Changelog

v2.1 [M54896]:
- support for parallel processing (using OpenMP),
- addition of PSNR and WS-PSNR [Sun17] values outputting,
- fixed WS-weight calculation for ERP sequences with non-180 lateral range,
- changed commandline arguments formatting,
- addition of detection of corrupted YUV files,
- change in compile-time parameters:
  - VERBOSE_LEVEL is now a commandline parameter,
  - WSPSNR_PEAK_VALUE_8BIT flag added (default: enabled), when enabled, the signal peak value for WS-PSNR computation is set to 255 << (BitDepth − 8). Otherwise, it is equal to 2^BitDepth − 1.

v2.0 [M54279]:
- addition of (rOff) and (tOff) commandline parameters,
- removal of redundant GCD calculations,
- usage of `uint16_t` data type and 4:4:4 chroma format for internal picture storage,
- new implementation of pixel-level processing steps,
- reduction of filesystem burden by coalescing read,
- detection of read errors – causes application to exit returning `EXIT_FAILURE`,
- implementation of Kahanand-Babuska-Neumaier accumulation,
- improved conversion of 8bps input sequences,
- improved interpolation for input sequences with 4:2:0 chroma format,
- addition of 3 compile-time parameters:
  - VERBOSE_LEVEL – controls number of per-frame printing; default = 0,
  - USE_KBNS – enables the Kahanand-Babuska-Neumaier accumulation; default: enabled,
  - USE_FIXED_WEIGHTS – enables faster 5×5 block search with fixed component weight (equal to 4:1:1); default = enabled,
- fixed possibility of reading from unallocated memory region during 5×5 block search,
- fixed GCD values rounding and clipping.

v1.0 [M45093].