ISO/IEC JTC 1/SC 29/WG 11
Coding of moving pictures and audio
Convenorship: Japan (JISC)

Document type:	**Approved WG 11 document**
Title:	**Software manual of IV-PSNR for Immersive Video**
Status:	**Output document**
Date of document:	**2020-07-11**
Source:	**Convenor, ISO/IEC JTC 1/SC 29/WG 11**
No. of pages:	**5**
Email of convenor:	**ostermann@tnt.uni-hannover.de**
Committee URL:	**http://isotc.iso.org/livelink/livelink/open/jtc1sc29**

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 11 NError! Reference source not found.

Online – July 2020

Source:	Video
Title:	Software manual of IV-PSNR for Immersive Video
Editor:	Adrian Dziembowski

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:

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

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

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

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

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

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:

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

where $\mathrm{MUD}(c)$ 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):

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

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.

$\mathrm{CCW}(c)$, $\mathrm{MUD}(c)$ and CPS values are predefined:

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

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-PSNR v2.0 executable accepts 9 – 11 parameters:

```
IV-PSNR  ref.yuv  test.yuv  W  H  NOF  BPS  CS  ERP?  out.txt  (rOff)  (tOff)
```

Required:

ref.yuv:	path to reference .yuv file
test.yuv:	path to test .yuv file
W:	video width
H:	video height
NOF:	number of frames
BPS:	bits per sample
CS:	chroma subsampling format (420 and 444 formats are supported)
ERP?:	0 if perspective, 1 if ERP
out.txt:	path to output .txt file

Optional:

(rOff):	reference video frame offset (default: 0)
(tOff):	test video frame offset (default: 0)

When no parameters are used, syntax help is outputted.

3 Examples

1. IV-PSNR of SA_v4_ref.yuv and SA_v4_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 SA_v4_ref.yuv SA_v4_test.yuv 4096 2048 20 10 420 1 IV-PSNR.txt
```

2. IV-PSNR of SD_v8_ref.yuv and SD_v8_test.yuv. Sequence resolution is 2048×1088, YUV420, 8 bits per sample. Sequence format is perspective. Mean IV-PSNR calculated for first 100 frames will be written into IV-PSNR.txt:

```
IV-PSNR SD_v8_ref.yuv SD_v8_test.yuv 2048 1088 100 8 420 0 IV-PSNR.txt
```

3. IV-PSNR of SE_v3_ref.yuv and SE_v3_test.yuv. Sequence resolution is 1920×1080, YUV420, 10 bits per sample. Sequence format is perspective. Mean IV-PSNR calculated for 5 frames (frames 0-4 of reference video and 10-14 of test video) will be written into IV-PSNR.txt:

```
IV-PSNR SE_v3_ref.yuv SD_v3_test.yuv 1920 1080 5 8 420 0 IV-PSNR.txt 0 10
```

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 Changelog

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].

6 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.

[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.