INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/AG 5

ISO/IEC JTC 1/SC 29/AG 5 m57119
July 2021, Online

Title: **On the importance of chroma in immersive video quality assessment**

Source: **Adrian Dziembowski, Jakub Stankowski, Dawid Mieloch**
Poznań University of Technology

Abstract

This informative contribution presents the results of the experiments on assessing the influence of color component weighting on the performance of PSNR and IV-PSNR in immersive video applications. The results presented in this document were obtained for the purposes of the journal paper about the IV-PSNR quality metric (currently under review).

1 Experiments

1.1 Color component weighting in quality assessment

Many objective quality metrics assess the quality of the image or video separately for each color component (e.g. Y, Cb, and Cr). Then, the quality of each component is being combined, e.g. using a weighted average. A common example is the PSNR, where the combined quality for three components is usually calculated as [611]:

$$PSNR_{YUV} = \frac{6 \cdot PSNR_Y + PSNR_U + PSNR_V}{8}.$$

In IV-PSNR, the default weights are set differently [WG04 N0013]:

$$IV\text{-}PSNR_{YUV} = \frac{4 \cdot IV\text{-}PSNR_Y + IV\text{-}PSNR_U + IV\text{-}PSNR_V}{6},$$

but the principle is the same. In this document, we present what happens, if different weights for chroma components are used. For both metrics, we have tested 13 weighting schemes (w_Y:w_{Cb}:w_{Cr}): 16:1:1, 8:1:1, 6:1:1, 4:1:1, 2:1:1, 1:1:1 (equal weights for 3 components), 1:2:2, 1:4:4, 1:6:6, 1:8:8, 1:16:16, 1:0:0 (only luma), 0:1:1 (only chromas).

1.2 Methodology

For each quality metric, the correlation between its results and the subjective quality was estimated. We have used two rank-based coefficients: KROCC and SROCC (Kendall and Spearman rank correlation coefficients). Correlation between MOS and metrics was evaluated in 3 experiments, described in m57118.

1.3 Immersive video coding: CfP on 3DoF+ Visual [WG11 N18145]

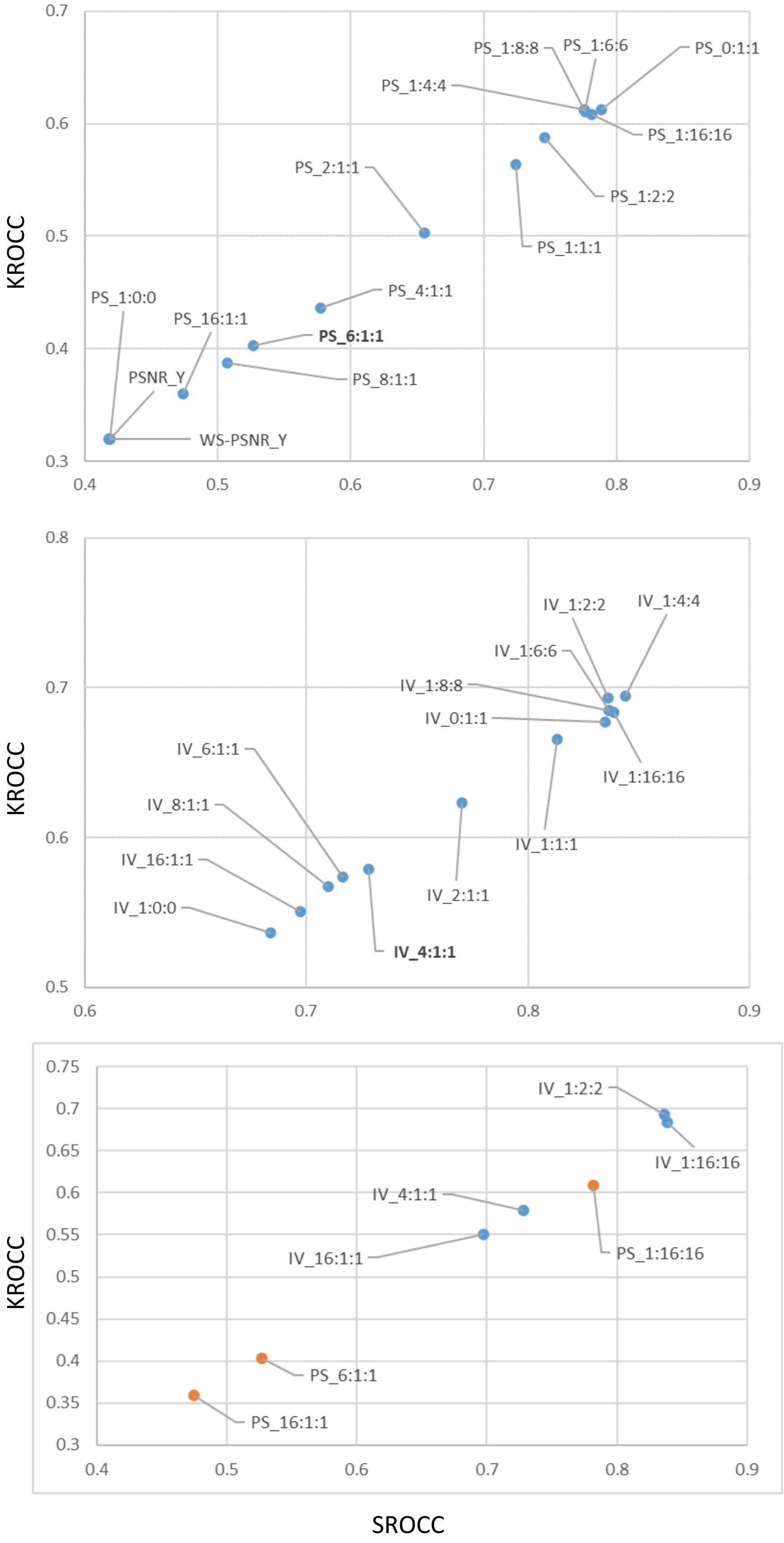

Fig. 1. KROCC and SROCC for all weighting schemes for PSNR, IV-PSNR, and chosen schemes for both metrics.

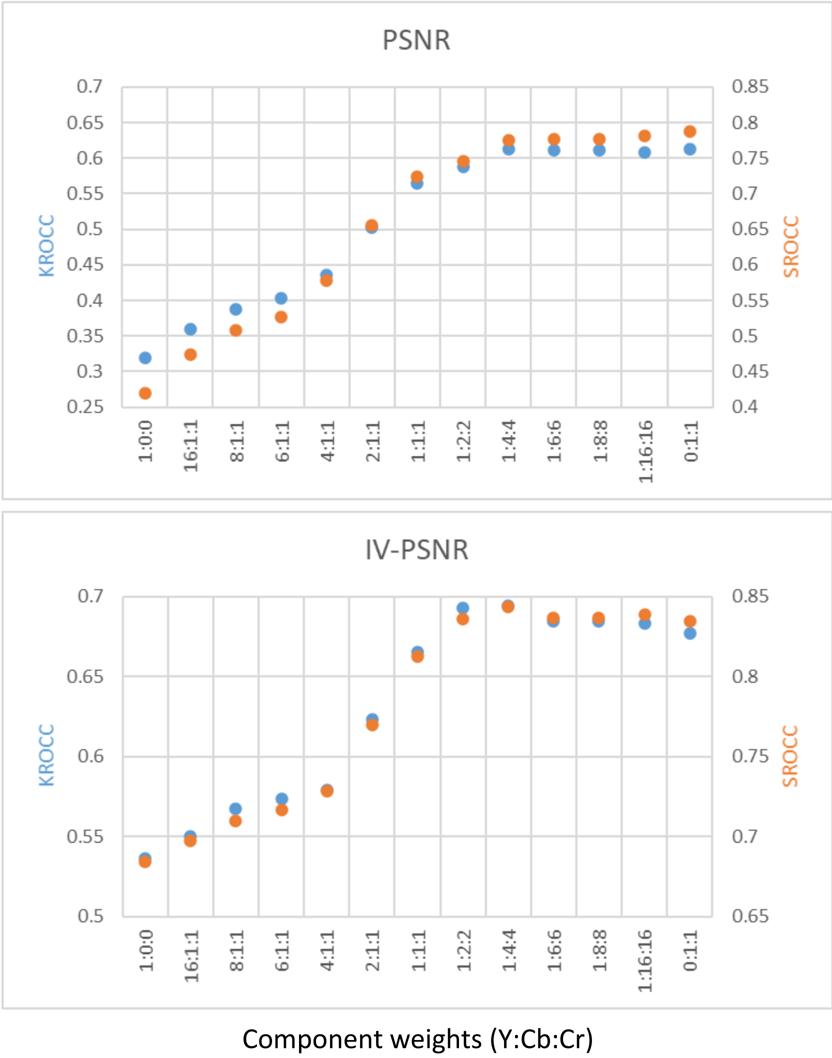

Fig. 2. The dependency between KROCC/SROCC and component weighting scheme.

The results for the immersive video coding are highly unexpected:

- chroma-focused weighting schemes correlate with MOS much better, than schemes with higher luma weight,
- 0:1:1 (which omits the quality of luma) is the best weighting scheme for PSNR, for IV-PSNR, it is only slightly worse than optimal schemes,
- commonly-used weighting schemes: 4:1:1 and 6:1:1 have a much worse correlation with MOS,
- PSNR with optimal weighting outperforms default IV-PSNR (6:1:1).

1.4 Immersive video processing: synthesis, color correction, filtration

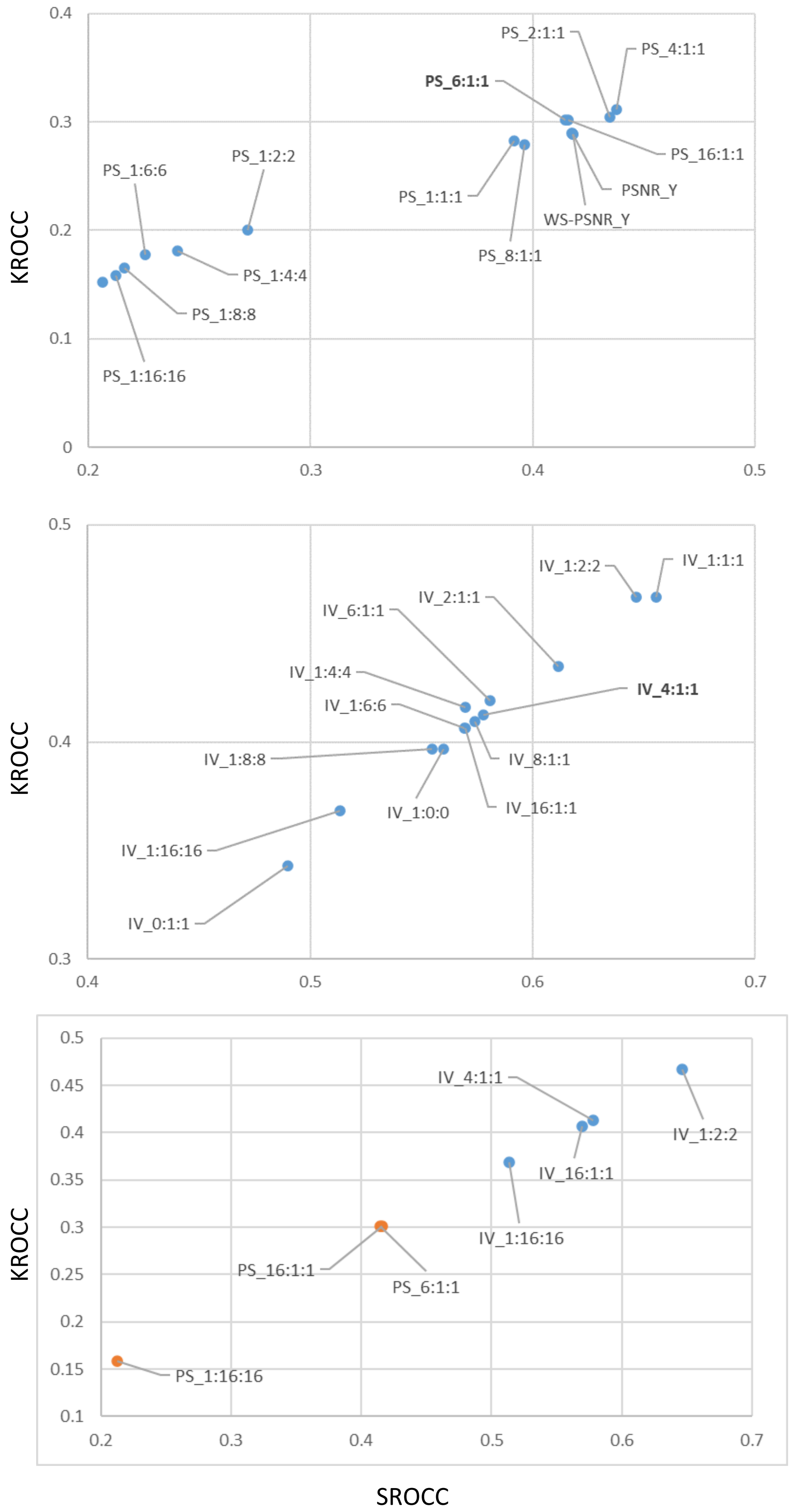

Fig. 3. KROCC and SROCC for all weighting schemes for PSNR, IV-PSNR, and chosen schemes for both metrics.

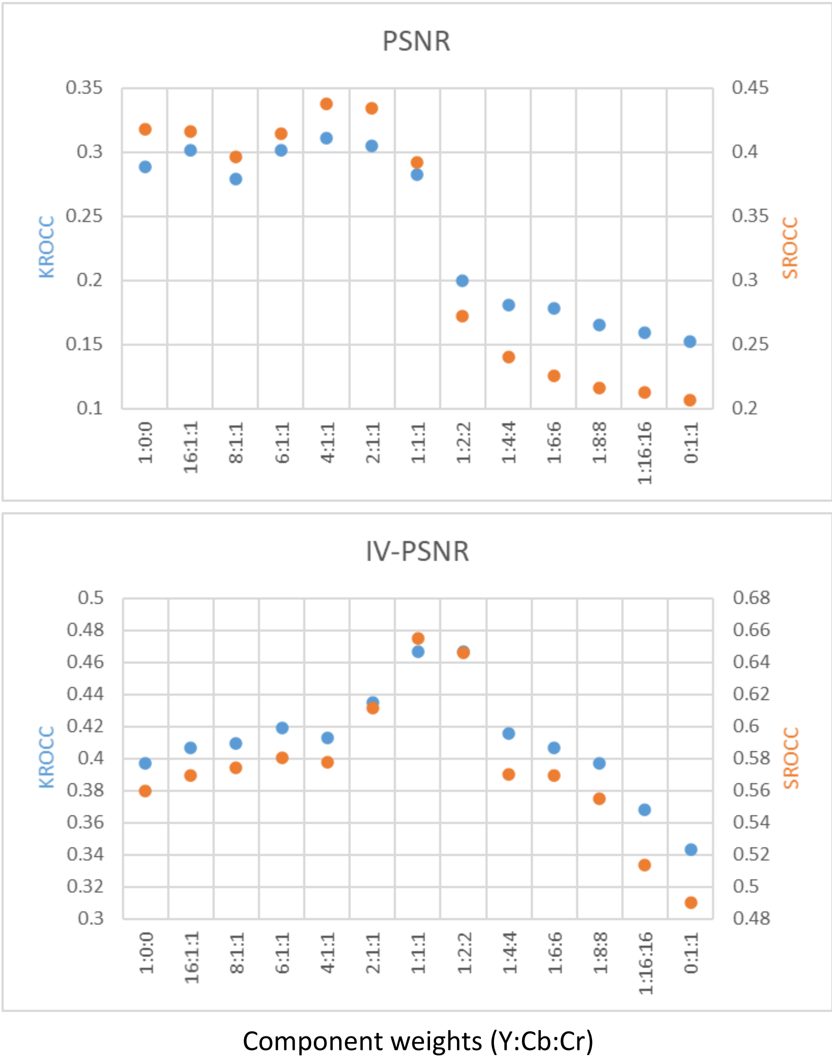

Fig. 4. The dependency between KROCC/SROCC and component weighting scheme.

The results for the immersive video processing are different, but several issues can be spotted:

- the general tendency is different for PSNR and IV-PSNR,
- opposing to the previous experiment, the 0:1:1 scheme is the worst one, both for PSNR and IV-PSNR,
- for both metrics, all the luma-focused schemes have a similar correlation with MOS,
- for IV-PSNR, there is a broad peak for schemes 2:1:1, 1:1:1, and 1:2:2, while such a phenomenon does not exist for PSNR,
- the best PSNR scheme is still worse than the worst IV-PSNR scheme,
- typical 6:1:1 weighting for PSNR seem to be a good choice,
- default 4:1:1 weighting for IV-PSNR is clearly outperformed by 1:1:1 and 1:2:2 schemes.

1.5 Non-immersive video applications: TID2013 database [TID2013]

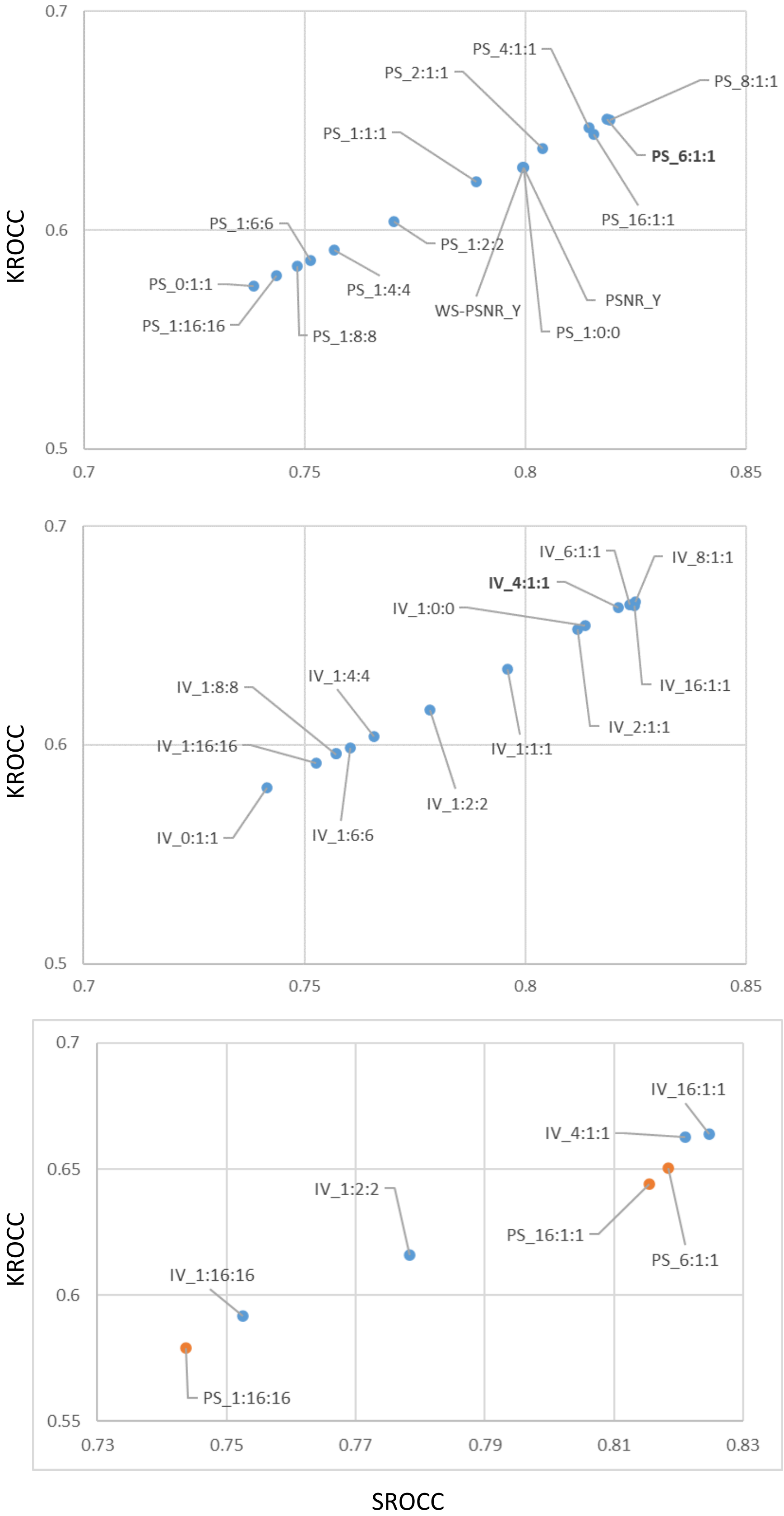

Fig. 5. KROCC and SROCC for all weighting schemes for PSNR, IV-PSNR, and chosen schemes for both metrics.

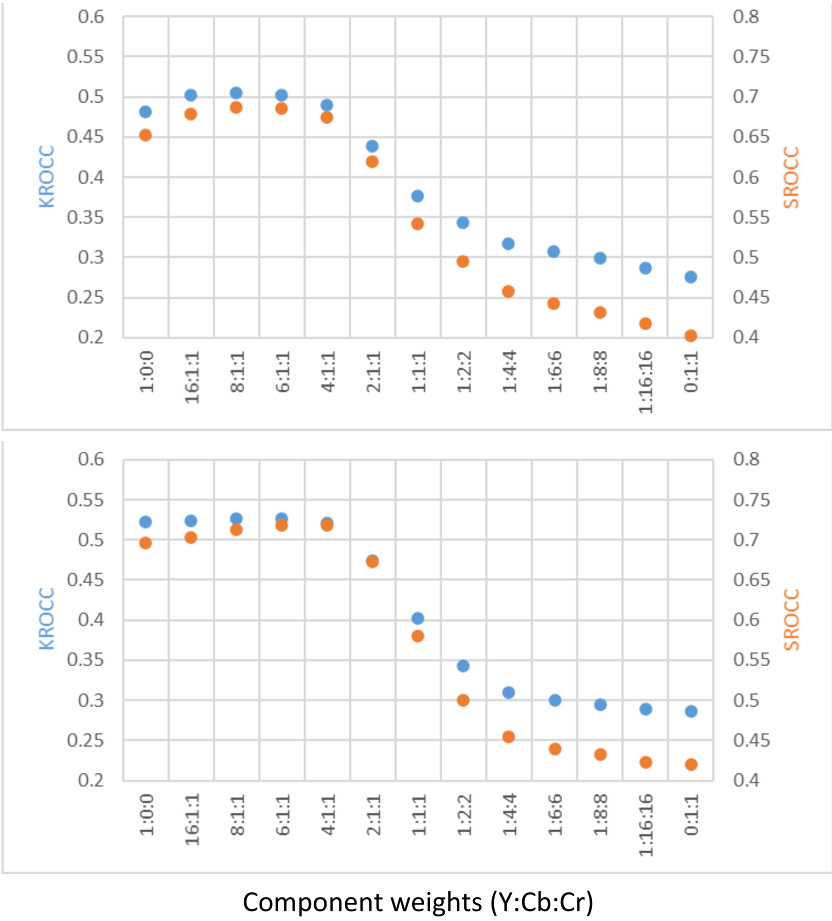

Fig. 6. The dependency between KROCC/SROCC and component weighting scheme.

Table 1. Ranks for all considered schemes of PSNR (the best metric for each distortion type is highlighted).

RANK	#	PS_1:0:0	PS_16:1:1	PS_8:1:1	PS_6:1:1	PS_4:1:1	PS_2:1:1	PS_1:1:1	PS_1:2:2	PS_1:4:4	PS_1:6:6	PS_1:8:8	PS_1:16:16	PS_0:1:1
AVG	0	6	3	**1**	2	4	5	7	8	9	10	11	12	13
Additive Gaussian noise	1	13	12	11	10	9	8	**1**	6	7	5	4	2	3
Noise in color comp.	2	2	3	**1**	4	5	6	7	8	9	12	10	11	13
Spatially correl. noise	3	13	11	12	10	9	8	7	6	5	4	3	2	**1**
Masked noise	4	13	12	11	10	9	8	7	6	5	3	4	2	**1**
High freq. noise	5	9	11	13	12	10	7	8	2	4	5	6	3	**1**
Impulse noise	6	13	12	11	10	9	2	**1**	3	5	4	6	7	8
Quantization noise	7	**1**	2	3	4	5	7	6	8	13	9	12	11	10
Gaussian blur	8	2	4	**1**	3	5	6	7	8	9	10	11	12	13
Image denoising	9	**1**	2	3	4	5	6	7	8	9	10	11	12	13
JPEG compression	10	8	6	4	3	2	**1**	5	7	9	10	11	12	13
JPEG2000 compression	11	8	5	2	**1**	3	4	6	7	9	10	11	12	13
JPEG transm. errors	12	3	**1**	2	4	5	6	7	8	9	10	11	12	13
JPEG2000 transm. errors	13	2	**1**	3	4	5	6	7	8	9	10	11	12	13
Non ecc. patt. noise	14	13	12	11	10	9	3	**1**	2	4	5	6	7	8
Local block-wise dist.	15	2	**1**	3	4	5	6	7	9	8	11	13	12	10
Mean shift	16	3	**1**	2	4	5	6	7	8	9	10	11	12	13
Contrast change	17	2	**1**	3	4	5	6	7	8	9	10	11	12	13
Change of color saturation	18	13	12	11	10	9	8	7	6	5	2	**1**	3	4
Multipl. Gaussian noise	19	9	8	7	10	11	13	12	6	5	4	3	2	**1**
Comfort noise	20	13	12	10	6	8	9	3	**1**	2	4	5	7	11
Lossy compr. of noisy images	21	12	13	11	10	9	2	**1**	3	4	5	6	8	7
Image color quant. w. dither	22	**1**	2	3	4	5	6	7	8	9	10	11	12	13
Chromatic aberrations	23	**1**	2	3	4	5	6	7	8	9	10	11	12	13
Sparse sampl. and reconstr.	24	8	5	2	**1**	3	4	6	7	9	10	11	12	13

Table 2. Ranks for all considered schemes of IV-PSNR (the best metric for each distortion type is highlighted).

RANK	#	IV_1:0:0	IV_16:1:1	IV_8:1:1	IV_6:1:1	IV_4:1:1	IV_2:1:1	IV_1:1:1	IV_1:2:2	IV_1:4:4	IV_1:6:6	IV_1:8:8	IV_1:16:16	IV_0:1:1
AVG	0	5	2	1	3	4	6	7	8	9	10	11	12	13
Additive Gaussian noise	1	7	5	4	3	1	2	6	8	9	10	11	12	13
Noise in color comp.	2	5	4	2	1	3	6	7	8	9	10	11	12	13
Spatially correl. noise	3	6	4	3	2	1	5	7	8	9	11	12	10	13
Masked noise	4	13	12	11	10	9	8	2	4	7	5	6	3	1
High freq. noise	5	7	6	4	2	1	3	5	8	9	10	11	12	13
Impulse noise	6	13	12	11	10	9	8	7	5	1	2	3	4	6
Quantization noise	7	13	12	10	11	9	8	7	6	5	3	2	4	1
Gaussian blur	8	1	2	3	4	5	6	7	8	9	10	11	12	13
Image denoising	9	2	1	3	4	5	6	7	9	10	11	12	13	8
JPEG compression	10	6	3	4	2	1	5	7	8	10	11	13	12	9
JPEG2000 compression	11	1	2	3	4	5	6	7	8	10	11	12	13	9
JPEG transm. errors	12	1	2	3	4	5	6	7	8	9	10	11	12	13
JPEG2000 transm. errors	13	1	2	3	4	5	6	7	8	9	10	11	12	13
Non ecc. patt. noise	14	13	11	3	1	4	2	5	6	7	8	9	10	12
Local block-wise dist.	15	1	2	3	4	5	6	7	9	11	13	12	10	8
Mean shift	16	9	7	6	5	3	2	1	4	8	10	11	12	13
Contrast change	17	1	2	3	4	5	6	7	8	9	10	11	13	12
Change of color saturation	18	13	12	11	10	9	8	4	6	5	1	2	3	7
Multipl. Gaussian noise	19	7	6	4	3	1	2	5	8	9	10	11	12	13
Comfort noise	20	13	12	11	10	9	8	7	6	5	4	3	2	1
Lossy compr. of noisy images	21	13	12	11	10	8	5	1	2	3	4	6	7	9
Image color quant. w. dither	22	3	1	2	4	5	6	7	8	9	10	11	12	13
Chromatic aberrations	23	1	2	3	4	5	6	7	8	11	12	13	10	9
Sparse sampl. and reconstr.	24	1	2	3	4	5	6	7	9	10	11	13	12	8

The results obtained for the non-immersive video applications are as expected:

- the luma-focused weighting schemes correlate with MOS better than the chroma-focused ones,
- however, it is valuable to consider the quality of chroma components, as the 1:0:0 scheme is worse than typical 6:1:1 and 4:1:1 schemes

1.6 General remarks and recommendations

Fig. 7 presents the results of the combination of two experiments on immersive video. Each dot presents the average SROCC and KROCC for both experiments.

As presented, typical/default weighting schemes 6:1:1 and 4:1:1 are not optimal for immersive video applications.

In order to achieve the highest correlation with MOS, the 1:1:1 for PSNR and 1:2:2 for IV-PSNR should be used (Table 3).

Table 3. Comparison of SROCC/KROCC between default and optimal weighting scheme.

Experiment	PSNR						IV-PSNR					
	SROCC			KROCC			SROCC			KROCC		
	6:1:1	1:1:1	Gain	6:1:1	1:1:1	Gain	4:1:1	1:2:2	Gain	4:1:1	1:2:2	Gain
Coding	0.53	0.72	**0.20**	0.4	0.56	**0.16**	0.73	0.84	**0.11**	0.58	0.69	**0.11**
Processing	0.41	0.39	**-0.02**	0.3	0.28	**-0.02**	0.58	0.65	**0.07**	0.41	0.47	**0.05**
Average	0.47	0.56	**0.09**	0.35	0.42	**0.07**	0.65	0.74	**0.09**	0.50	0.58	**0.08**

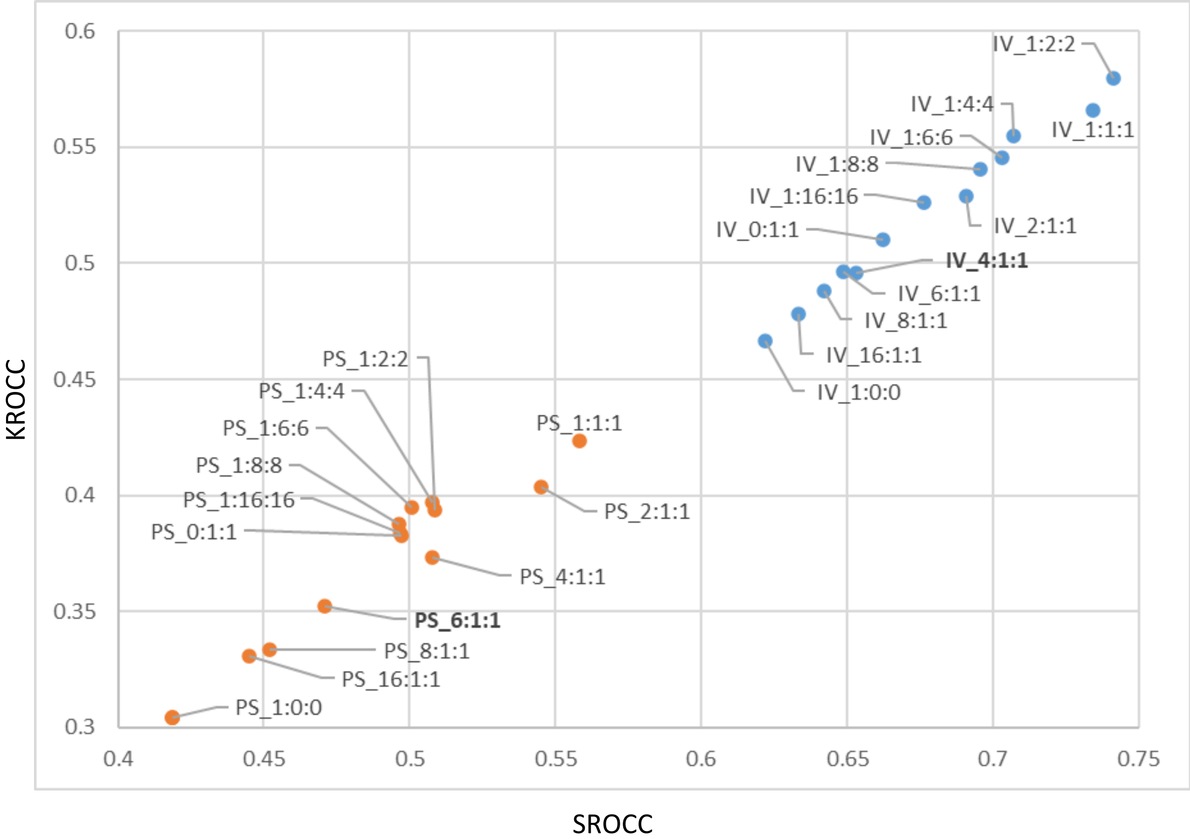

Fig. 7. Combination of experiments 1.3 (immersive video coding) and 1.4 (immersive video processing): average KROCC and SROCC values for all the considered weighting schemes and both metrics; PS: PSNR, IV: IV-PSNR.

Acknowledgement

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

The Authors would like to thank Vittorio Baroncini from EVATech and Giacomo Baroncini from GBTech for providing the results of the subjective quality evaluation of MPEG Call for Proposals. The Authors would like to thank also the proponents of MPEG CfP: Bart Kroon and Bart Sonneveldt (Philips Research Eindhoven), Vinod Kumar Malamal Vadakital (Nokia), Lu Yu and Bin Wang (Zhejiang University), Gérard Briand, Julien Fleureau, and Renaud Doré (Technicolor/ InterDigital) for providing executables and/or videos generated using their proposals for the evaluation of the IV-PSNR metric.

References

[611] Y. Huang, H. Qi, B. Li, and J. Xu, “Adaptive weighted distortion optimization for video coding in RGB color space,” *IEEE International Conference on Image Processing, ICIP 2014*, Paris, France, Oct. 2014.

[MVS] A. Dziembowski, A. Grzelka, D. Mieloch, O. Stankiewicz, K. Wegner, and M. Domański, “Multiview Synthesis – improved view synthesis for virtual navigation,” 32nd Picture Coding Symposium, PCS 2016, Nürnberg, Germany, Dec. 2016.

[TID2013] N. Ponomarenko et al., “Image database TID2013: Peculiarities, results and perspectives,” Signal Proc.: Image Comm., vol. 30, pp. 57-77, 2015.