HDR Video Super-Resolution for Future Video Coding

Seiya Umeda, Niai Yano, Hiroshi Watanabe
Waseda University
3-14-9, Okubo, Shinjuku-ku, Tokyo, Japan
sana39_fs_mkoy@toki.waseda.jp

Tomohiro Ikai, Takeshi Chujoh, Norio Ito
Sharp Corporation
1-9-2, Nakase, Mihama-ku, Chiba, Japan
ikai.tomohiro@sharp.co.jp

***Abstract*— New coding tools are emerging in Future Video Coding (FVC). A novel video coding system, which combines FVC-era coding tools with super-resolution is proposed, aiming to demonstrate the state of the art coding results in ultra-high definition video coding. Experimental results show the proposed system outperforms convention system in both SDR and HDR images.**

Keywords—Future Video Coding, Super-resolution, HDR

I. INTRODUCTION

After the approval of High Efficiency Video Coding (HEVC) standard in 2013, Ultra high definition (UHD) video, such as 4K and 8K, has been rapidly deployed in video streaming, cable television, Blue-ray and broadcasting etc. In the name of Future Video Coding (FVC) [1], the next generation video coding project was launched with Joint Video Exploration Team (JVET) of ITU-T/VCEG and ISO-IEC/JTC1/SC29/WG11 (MPEG) [2] and much more video consumption with relatively low bitrate would be anticipated in the future. In the meantime, novel super resolution techniques [3] are proved to provide superior video quality. In parallel, deep learning based super resolution (DLSR) [4][5] has been discussed and its concerned high complexity would be resolved by wide range of GPU deployment and newly introduced DL specific chips. The deep learning based loop filter has been also tested [6].

These digital video experience is further strengthened by wide color and high dynamic range as guaranteed with ITU-R BT. 2100 standard which defines Perceptual Quantization (PQ) and Hybrid Log Gamma (HLG).

In this paper, we propose an ultra-high video coding system, which combines so-called FVC-era technologies and super-resolution technique to provide 4K HDR video in broader range of applications, including mobile video streaming etc.

II. SUPERRESOLUTION IN UNCOMPRESSED VIDEO

Before discussing compressed video, uncompressed video super-resolution was tested. Only the first frame of a video is picked up for our experiment. In this test, bicubic interpolation and a super-resolution A+ [3] was compared where input image was generated using 2 : 1 subsampling (both in the horizontal and vertical direction) with Lanczos filtering and interpolated images was measured by PSNR with the original image. As shown in Fig. 1, super-resolution A+ shows clearly better objective quality compared to Bicubic in both SDR and HDR. For 2K and 4K-size images, about 1.5 and 0.75 dB improvement can be obtained respectively. Fig. 2 shows the subjective effect of super-resolution applied to the uncompressed image. As shown, white dots in the tie and vertical lines in the background are more explicit in A+.

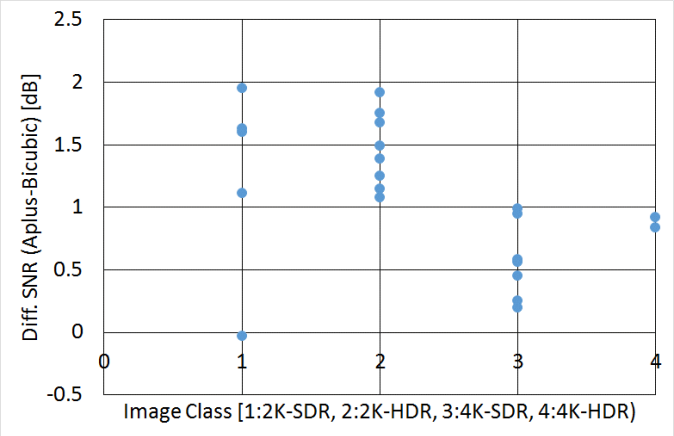
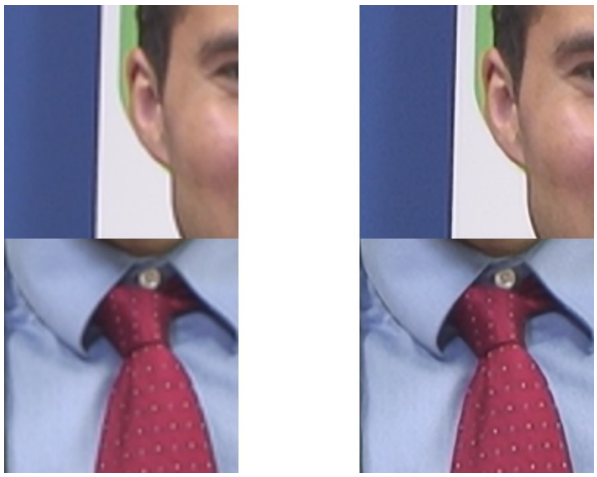


Fig. 1. PSNR results in uncompressed video [dB]



Interpolation by Bicubic Super-Resolution by A+

Fig. 2. Subjective results in uncompressed video (Johnny)

978-1-5386-2615-3/18/$31.00 ©2018 IEEE

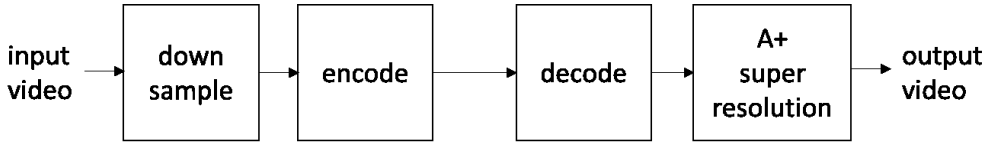


Fig. 3. Proposed video coding system

III. Proposed Vidoe Coding System

As shown in Fig. 3, the proposed system down-samples original input videos with Lanczos filter and encodes the down-sampled SDR and HDR video by the latest JEM 7.0 (Joint Exploration Model) encoder [7]. The encoded bitstream is to be decoded by the corresponding JEM decoder and output video is up-sampled with a super resolution technique A+.

IV. Experiment I

We employ A+ for super-resolution technique. A conventional interpolation with Bicubic was also tested as an anchor (anchor1). The sampling ratio of 2 : 1 was tested. full resolution encoding of 1 : 1 was also tested as a conventional coding (anchor2).

Firstly, we tested 8 and 10-bit SDR video images. The first frame was used. The file size (bitrate) of bitstreams were roughly matched between full resolution bitstream (anchor2) and half resolution bitstream (proposed and anchor1). Specifically, QP value 41-45 was used for anchor 2 (1 : 1) while QP value 31-39 was used for proposed and anchor 1 (2 : 1).

Experimental results are shown in TABLE II. The file size differences are between -7.9 % and + 0.8 % as shown in Table III. In TABLE II, it is shown the proposed video coding system outperforms the conventional full-size approach (anchor2) in PSNR in most cases. Fig. 4 shows subjective quality. From Fig. 4, A+ provides sharper edges in characters and fine details in a face. We also tested different QP cases, namely, Low-QP and High-QP. In general, improvement can be seen when QP has middle or high value.

Since bitrate cannot be exactly set to the same value, rate vs quality characteristics should be considered. Thus, PSNR difference vs. Rate difference characteristics for Low, Mid, and High QP values to SDR images is shown in Fig. 5. From Fig. 5, in SDR images, it is observed that **High-QP of A+ approach provides 4% bitrate reduction at the same quality** to the anchor2. However, there is no significant improvement in Low-QP. Especially, it is difficult to obtain gains in video sequences names Cactus, BasketBallDrive and BQTerrace at Low-QP values.

TABLE II PSNR results in SDR [dB] (Mid-QP)

	Proposed	Anchor1	Anchor2
Resolution	2 : 1	2 : 1	1 : 1
Method	A+	Bi-cubic	-
Tango2	**37.54**	37.54	37.08
Drum100	32.94	**32.95**	32.86
Campfire	**35.40**	35.18	34.68
ToddlerFountain2	**29.72**	29.73	29.68
CatRobot	**35.20**	35.14	34.68
TrafficFlow2	**33.66**	33.65	33.53
DaylightRoad2	**32.46**	32.45	32.38
RollerCoaster2	38.62	**38.64**	37.88
Kimono	33.43	**33.44**	33.23
ParkScene	**29.54**	29.52	29.53
Cactus	31.14	30.74	**31.28**
BasketballDrive	33.23	32.56	**33.76**
BQTerrace	27.03	26.36	**27.67**

TABLE. III File size information in SDR (Mid-QP)

	File size of full res. [bytes]	File size of half res. [bytes]	Diff [%]
Tango2	18439	18476	0.2%
Drum100	33481	31426	-6.1%
Campfire	37490	36564	-2.5%
ToddlerFountain2	40032	38924	-2.8%
CatRobot	39478	39775	0.7%
TrafficFlow2	17013	15974	-6.1%
DaylightRoad2	21822	20483	-6.1%
RollerCoaster2	22633	22824	0.8%
Kimono	6899	6671	-3.3%
ParkScene	9049	8695	-3.9%
Cactus	19050	18878	-0.9%
BasketballDrive	6451	6442	-0.1%
BQTerrace	19099	18925	-0.9%

2 : 1 Bicubic 2 : 1 A+ 1 : 1

Fig. 4. SDR subjective results (Mid-QP, Campfire [9])

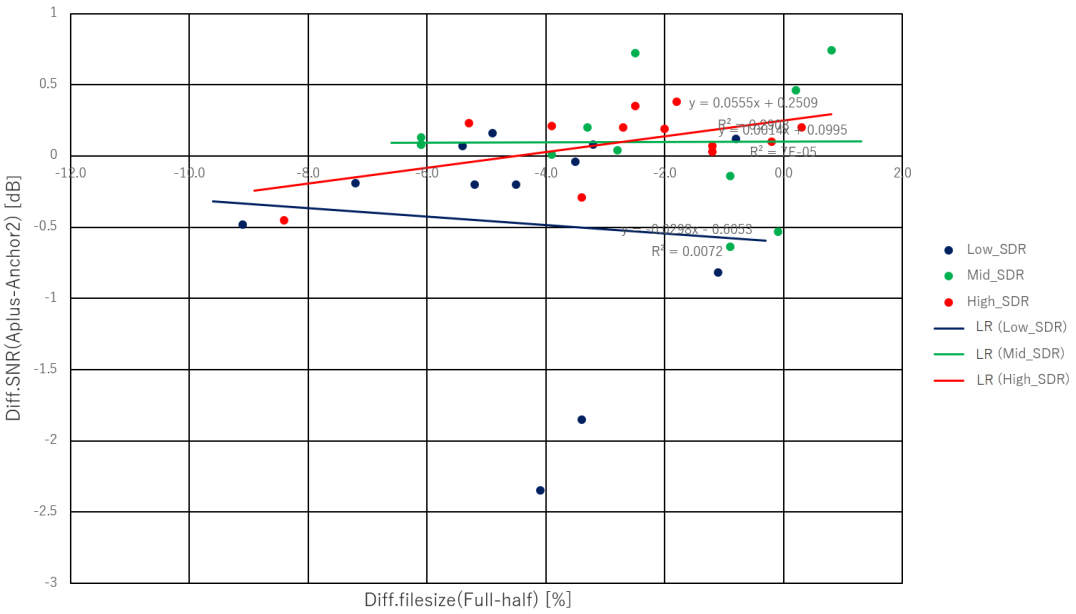


Fig. 5. PSNR difference vs Rate when QP changes for SDR

V. Experiment II

Secondly, we tested 10-bit HDR video images. Test sequences includes FireEater2, Market3, SunRise, ShowGirls2, BallonFestival, Hurdles, and Starting in 1080p PQ, FlyingBirds and SunsetBeach in 4K HLG. The up-sampling (super resolution) is computed in the coded representation domain rather than in linear light domain. The first frame was used. PSNR comparison is shown in TABLE IV. The file size differences are -9.1 % and + 0.4 % as shown in Table V.

Again, PSNR difference vs. Rate difference characteristics for Low, Mid, High-QP values to HDR images is shown in Fig. 6. From Fig. 6, it is observed that Mid and High QP of A+ approach provides **more than 10% bitrate reduction at the same quality** to the anchor2. However, there is no significant improvement in Low-QP. This trend is the same as in SDR.

TABLE IV PSNR results in HDR (Mid-QP)

	Proposed	Anchor1	Anchor2
Coded Resolution	2 : 1	2 : 1	1 : 1
Method	A+	Bi-cubic	-
FireEater2	**41.67**	41.67	40.81
Market3	30.76	30.53	**30.98**
SunRise	**36.24**	36.15	35.83
ShowGirls2	**32.95**	32.93	32.87
BallonFestival	**34.54**	34.39	33.89
Hurdles	**34.61**	34.51	34.17
Starting	32.76	32.65	**32.92**
Cosmos	**29.41**	29.41	29.33
FlyingBirds	**34.33**	34.33	34.17
SunsetBeach	33.57	33.45	**33.75**

TABLE V File size information in HDR (Mid-QP)

	File size of full res. [bytes]	File size of half res. [bytes]	Diff [%]
FireEater2	1252	1184	-5.4%
Market3	15710	14181	-9.7%
SunRise	3114	3054	-1.9%
ShowGirls2	7202	6548	-9.1%
BallonFestival	17758	17823	0.4%
Hurdles	9903	9832	-0.7%
Starting	14324	13053	-8.9%
Cosmos	5143	4698	-8.7%
FlyingBirds	19589	19570	-0.1%
SunsetBeach	24412	24179	-1.0%

HDR specific metrics of deltaE100 and PSNR-L100 were calculated using HDRtools [8]. PSNR-L100 that represents the distortion in the lightness domain of the CIELab colour space is shown in TABLE VI. In terms of the rate distortion comparison, the similar trend to Fig. 6. is confirmed. Fig. 7 shows subjective image quality in the HDR experiment. Straight edges of the proposed method are clearer than others.

VI. Conclusion

We propose and experimented the video coding system which uses super-resolution for both SDR and HDR. From experiments, the proposed super resolution system provides superior visual quality in the range of Mid-QP and High-QP. Moreover, the experiment shows compared to the normal coding without sub-sampling, 4 to 10 % bitrate saving at the same image quality can be achieved. The objective and subjective quality improvement is observed in HDR as well as SDR.

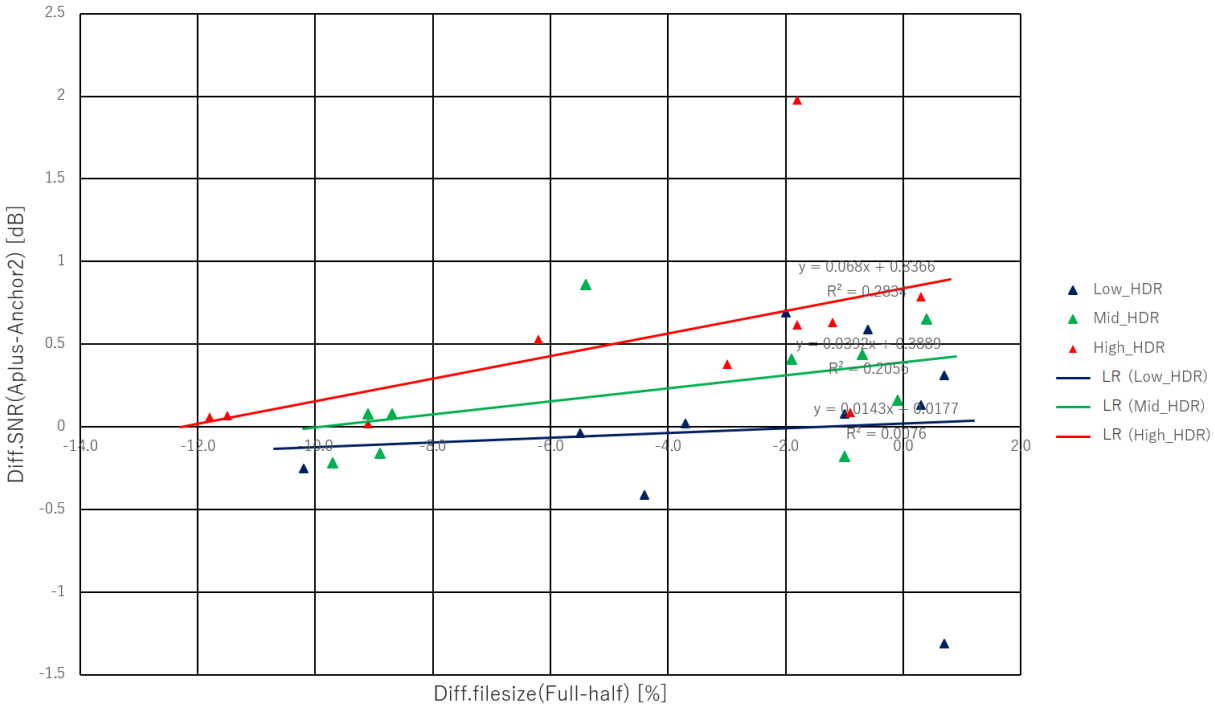
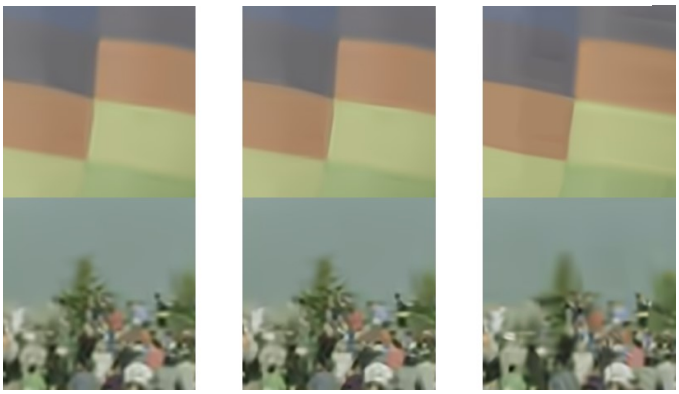


Fig. 6. PSNR difference vs Rate when QP changes for HDR

TABLE VI PSNR-L100 results in HDR (Mid-QP)

	Proposed	Anchor1	Anchor2
Coded Resolution	2 : 1	2 : 1	1 : 1
Method	A+	Bi-cubic	-
FireEater2	**42.244**	42.240	41.678
Market3	31.270	31.144	**31.386**
SunRise	**34.645**	34.618	34.396
ShowGirls2	**34.918**	34.904	34.857
BallonFestival	**34.222**	34.155	33.770
Hurdles	**31.685**	31.628	31.548
Starting	32.670	32.607	**32.831**
Cosmos	**34.026**	34.022	33.955
FlyingBirds	30.537	**30.538**	30.457
SunsetBeach	**31.386**	31.367	31.274



2 : 1 Bicubic 2 : 1 A+ 1 : 1

Fig. 7. HDR subjective results (BalloonFestival [10], Mid-QP)

REFERENCES

[1] http://mpeg.chiariglione.org/tags/future-video-coding.

[2] http://phenix.int-evry.fr/jvet/

[3] R. Timofte, V. D. Smet, and L. V. Gool, "A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution," In Asian Conference on Computer Vision (ACCV 2014), pp.111-126, Nov. 2014.

[4] C. Dong, C. C. Loy, K. He, X. Tang, "Image Super-Resolution Using Deep Convolutional Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.38, No.2, pp.295-307, June 2015.

[5] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, ", https://arxiv.org/abs/1609.04802, 2017.

[6] C. Dong, C. C. Loy, K. He, X. Tang, "CNN-based in-loop filtering for coding efficiency improvement", IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop, 11-12 July 2016.

[7] https://jvet.hhi.fraunhofer.de/svn/svn_HMJEMSoftware/

[8] https://gitlab.com/standards/HDRTools

[9] L. Song, X. Tang, W. Zhang, X. Yang, P. Xia, The SJTU 4K Video Sequence Dataset, the Fifth International Workshop on Quality of Multimedia Experience (QoMEX2013), Klagenfurt, Austria, July 3rd-5th, 2013.

[10] Copyright Cable Labs under CC-BY-ND 4.0.