|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  14th Meeting: Vienna, AT, 25 July – 2 Aug. 2013 | Document: JCTVC-N0148 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AhG8: Guided Image Filtering for Screen Content Coding** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information, Draft | | |
| *Author(s) or Contact(s):* | Thijs Vermeir  Jürgen Slowack  Jan De Cock  Glenn Van Wallendael  Sebastiaan Van Leuven | Tel: Email: | +32 56 26 2379 [thijs.vermeir@barco.com](mailto:thijs.vermeir@barco.com)  [jurgen.slowack@barco.com](mailto:jurgen.slowack@barco.com)  [jan.decock@ugent.be](mailto:jan.decock@ugent.be)  [glenn.vanwallendael@ugent.be](mailto:glenn.vanwallendael@ugent.be)  [sebastiaan.vanleuven@ugent.be](mailto:sebastiaan.vanleuven@ugent.be) |
| *Source:* | Barco N.V., President Kennedypark 35, 8500 Kortrijk, Belgium  Ghent University – iMinds | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

Downsampling the chroma components, as required for coding in HEVC Main profile (4:2:0), generates visual artifacts for screen content. This contribution provides the results of experiments with a guided image post-processing filter to reduce these artifacts after decoding. This filtering step restores edges in the chroma component of screen-content, which have been affected due to the original 4:2:0 sub-sampling. The luma component is used as a guide to enhance the chroma components. By analyzing a local histogram per pixel, the radius of the guided image filter is reconfigured. Tests conducted on several sequences show that the PSNR of the Cb and Cr components for lossless compressed streams is increased with 0.63 dB and 0.99 dB on average, respectively. The presented results are based on on-going experiments, therefore also future work is proposed.

# Introduction

The Human Visual System (HVS) is more sensitive to variations in brightness than color (chroma). Therefore it is common in video compression to subsample the chroma components. This is also the case in HEVC, where the main supported color space is YUV 4:2:0. This reduction in (chroma) resolution does not create visual artifacts for natural content. However, screen content has different characteristics as it is often temporally static and contains sharp edges. The viewer is also looking at the screen in another way, since typically the viewer is closer to the screen and can look only to a specific area of the screen[1]. This combination can result in visual artifacts when screen content coding is performed using HEVC Main profile (due to YUV 4:2:0 subsampling)[2]. A non-subsampled color format is required to provide visually lossless quality for screen content. Such color spaces could include YUV 4:4:4 or RGB.

A trend in the professional visualization market is that (mobile) consumer electronics devices (like smartphones and tablets) should be able to integrate with the professional visualization system. Consequently, the system should encode the video stream in a format supported by these devices. Typically these devices have hardware accelerated decoding (and/or encoding), but the support is mostly limited to the most popular profiles (HEVC Main profile). This is in conflict with the previous requirement for a 4:4:4 color format, since HEVC Main profile only supports 4:2:0. Simulcasting the stream, with a limited amount of streams, could be a valid solution for this problem but this can make the system more expensive with the extra equipment that is needed. Also the extra complexity (device configuration, network bandwidth management, predictability) to manage the system should not be underestimated. In larger systems this is not a maintainable solution at all. Manufacturers of chip or SoC devices that support hardware acceleration for codecs also concentrate on the Main profile. High profile codecs are, compared to Main Profile codecs, not widely available or are rather expensive due to low sales volumes.

This contribution proposes a solution for this problem by allowing upgrading the visual quality of the HEVC Main profile stream on a professional encoder while still supporting the stream on consumer electronics. To enhance the quality in the professional equipment the output of the decoder is additionally filtered with a guided image filter to improve the visual quality for screen content. In this contribution the filter is only working after the decoder. There are no changes needed to the encoder. This could change in future work (see 5 Future work).

# Guided Image Filtering

The guided image filter [3] is an edge-preserving smoothing filter such as the bilateral filter but with better behavior near edges. It uses a linear model to generate the filtered output by using a guidance image, which can be either the image itself or another image. The filter transfers structural visual information from the guidance image to the filtered output. It can be implemented efficiently in linear time, regardless of the filter’s kernel size and intensity range. Currently this is one of the fastest edge-preserving filters.

The filter can be configured with two parameters: the radius and the regularization. The radius defines the size of the square region around the current pixel being processed. This region is used to generate the smoothing parameter for the filter and for edge detection. The ratio between the covariance (of the guide and image) and variance (of the guide) defines if the texture of the guide (like edges) should be transferred to the filtered output. With the regularization parameter this ratio can be fine-tuned.

In this contribution, we propose a guided image filter to filter the chroma using the luma as the guide. The filter is applied to each chroma component (Cb and Cr) independently. Typically screen content has sharp edges where the edge slope only spans a few pixels. When downsampling edges (with less steep transitions) no down- or up-sample artifacts are generated. Therefore the regularization parameter is not used dynamically and configured as a fixed property. The radius property is used to re-configure the filter for every pixel to adapt the smoothing behavior of the filter.

# Histogram-based analysis

Screen Content such as text and window border typically has sharp edges, with varying characteristics (e.g. due to font and operating system themes). This requires a dynamic configuration of the filter.

The guided image filter in [3] is extended with a histogram-based analysis tool to define the radius. A local histogram is created in the learning phase for each pixel in the luma component. This histogram, a square window around the pixel with a fixed radius (11 pixels), is analyzed by taking three parameters into account (i.e., the number of unique luma values, how many times these luma values are used and the difference in luminance between 2 luma values). The filter is applied for all radius values (from a radius value of 1 up to the fixed radius of the histogram). When all luma pixels are analyzed, the results are collected in a lookup table. This table allows to select the best radius for a set of pixel characteristics, and on average to generate the best output of the filter (PSNR based). In the following tests the same content is used in the training phase as in the evaluation phase, creating the best table for this limited set of sequences.

# Results

In a first phase the tests concentrate on evaluating this filter method on losslessly coded 4:2:0 streams. In a later phase this research will be extended to lossy coding. Screen content sequences from the HEVC test set are used (i.e., sc\_cad\_waveform, sc\_programming, sc\_video\_conferencing\_doc\_sharing and sc\_wordEditing) as well as two other sequences (i.e., Office\_Scada\_1 and Video\_Maps\_1) as depicted in Figure 1. Video\_Maps\_1 contains mixed content; it also has natural video from security cameras.

Figure 1: Additional test sequences used in this experiment.  
Left image is Office\_Scada\_1, right image is Video\_Maps\_1.

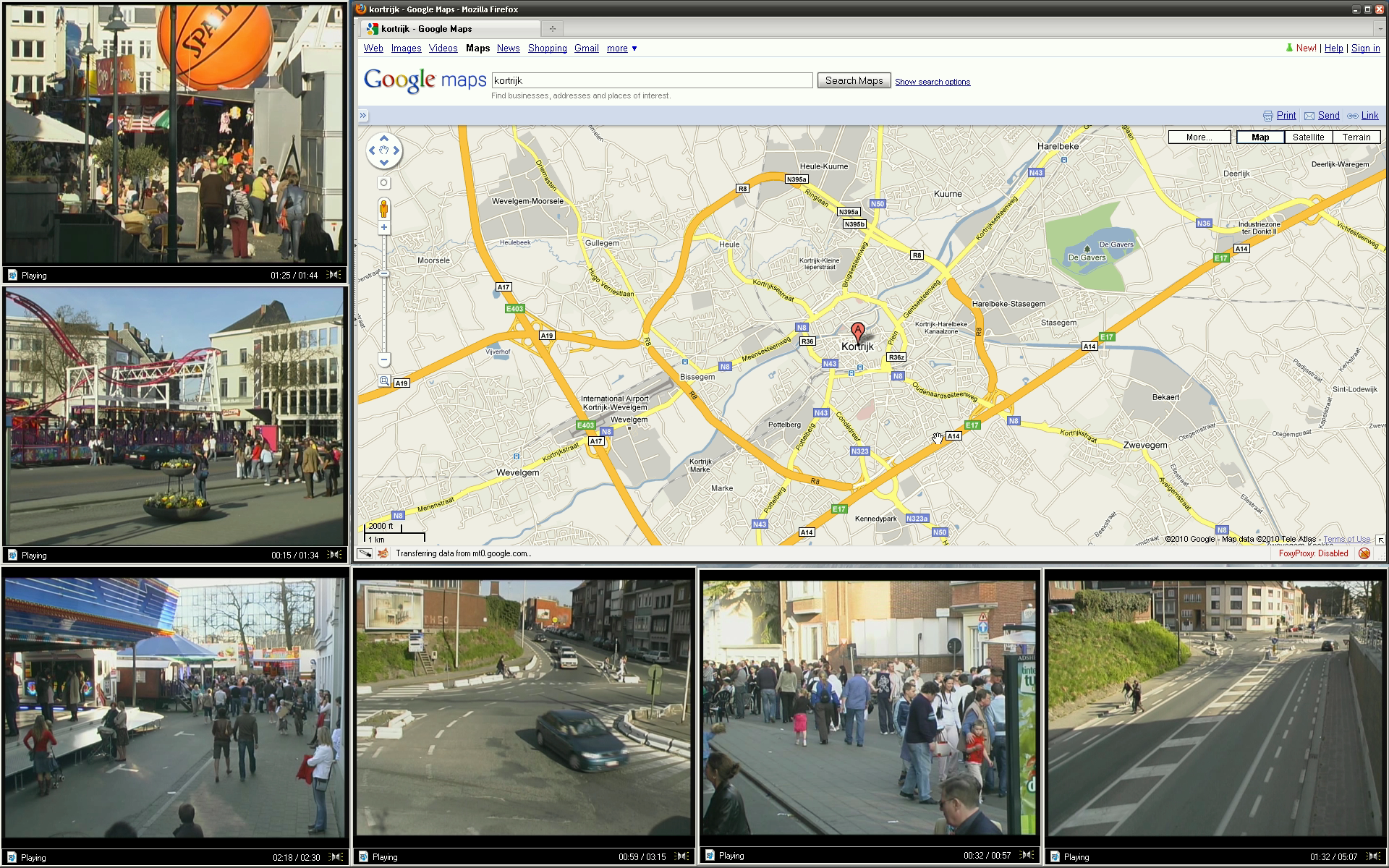
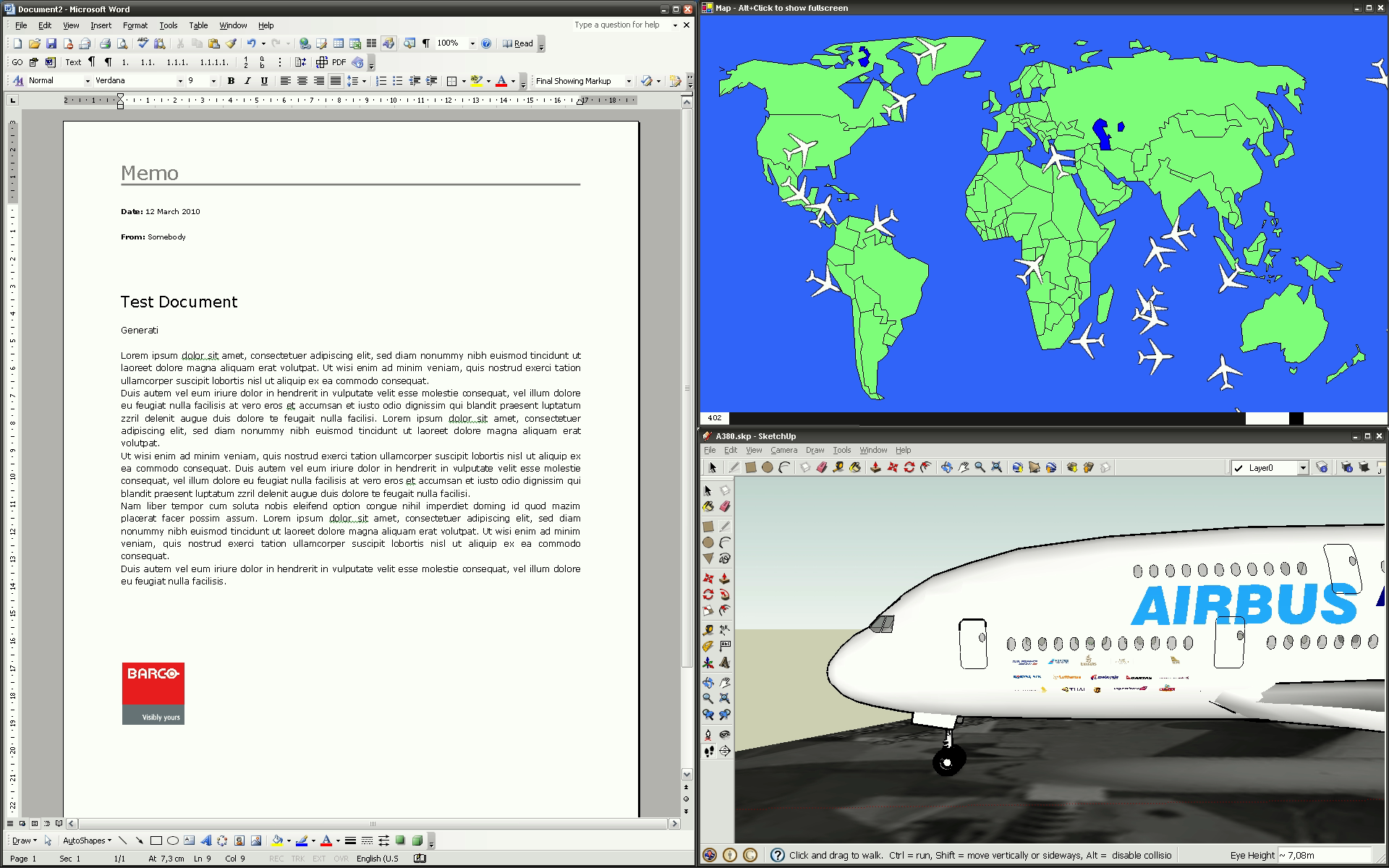


Table 1: Results of guided image filtering with fixed radius for all pixels for lossless compression, compared to the original 4:4:4 sequence

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Unfiltered** | | **Guided Image filtering with fixed radius** | | | | | | | | | | | |
|  |  |  | **1** | | **3** | | **5** | | **7** | | **9** | | **11** | |
| **Sequence** | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** |
| Office\_Scada\_1 | 36.67 | 38.58 | 36.67 | 38.58 | 37.28 | 40.33 | 35.07 | 40.35 | 33.63 | 40.08 | 32.68 | 39.74 | 32.01 | 39.42 |
| sc\_cad\_waveform | 24.97 | 23.06 | 24.97 | 23.06 | 26.05 | 24.35 | 26.44 | 24.45 | 26.45 | 24.27 | 26.36 | 24.09 | 26.26 | 23.91 |
| Video\_Maps\_1 | 37.95 | 42.95 | 37.95 | 42.95 | 38.98 | 43.97 | 38.00 | 42.96 | 36.95 | 41.77 | 36.02 | 40.75 | 35.22 | 39.89 |
| sc\_programming | 36.14 | 35.96 | 36.14 | 35.96 | 36.99 | 36.96 | 36.00 | 36.10 | 34.9 | 35.03 | 34.30 | 34.43 | 33.84 | 33.95 |
| sc\_video\_conferencing\_doc\_sharing | 32.70 | 31.12 | 32.70 | 31.12 | 32.91 | 31.26 | 32.64 | 31.01 | 32.4 | 30.81 | 32.24 | 30.64 | 32.09 | 30.48 |
| sc\_wordEditing | 37.95 | 36.28 | 37.95 | 36.28 | 37.24 | 36.48 | 36.03 | 35.81 | 35.51 | 35.50 | 35.13 | 35.30 | 34.74 | 35.14 |

Table 1 shows the results of applying the guided image filter to the chroma components using with the luma as the guide. In the first two columns the guided image filter is not applied, only the bilineair upsample filter is used for the chroma components. In the next columns the guided filter is used with a fixed radius on the whole frame. If a fixed radius of 3 pixels is used the best PSNR is obtained. However in the sc\_cad\_waveform sequence radius 5 is the best option.

Table 2: Results of guided image filtering when best radius is selected for each pixel, compared to the original 4:4:4 sequence

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequence** | **Unfiltered** | | **Best Radius GIF** | | **Gain** | |
|  | **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** |
| Office\_Scada\_1 | 36.67 | 38.58 | 39.99 | 42.04 | 3.32 | 3.46 |
| sc\_cad\_waveform | 24.97 | 23.06 | 28.11 | 26.1 | 3.14 | 3.04 |
| Video\_Maps\_1 | 37.95 | 42.95 | 41.55 | 46.26 | 3.60 | 3.31 |
| sc\_programming | 36.14 | 35.96 | 39.62 | 39.05 | 3.48 | 3.09 |
| sc\_video\_conferencing\_doc\_sharing | 32.7 | 31.12 | 34.57 | 32.74 | 1.87 | 1.62 |
| sc\_wordEditing | 37.95 | 36.28 | 40.03 | 39.02 | 2.08 | 2.74 |
|  |  |  |  | **Average** | **2.92** | **2.88** |

Table 2 shows the results when testing all radius values and selecting the radius with the highest PSNR. If the best radius is selected, the PSNR gain could go up to +2.92 dB and +2.88 dB on these sequences.

Table 3: Results of guided image filtering when radius is defined by histogram analysis, compared to the original 4:4:4 sequence

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequence** | **Unfiltered** | | **Histogram based guided image filter** | | **Gain** | |
| **Cb** | **Cr** | **Cb** | **Cr** | **Cb** | **Cr** |
| Office\_Scada\_1 | 36.67 | 38.58 | 37.29 | 40.52 | 0.62 | 1.94 |
| sc\_cad\_waveform | 24.97 | 23.06 | 26.75 | 25.01 | 1.78 | 1.95 |
| Video\_Maps\_1 | 37.95 | 42.95 | 38.91 | 43.86 | 0.96 | 0.91 |
| sc\_programming | 36.14 | 35.96 | 36.57 | 36.60 | 0.43 | 0.64 |
| sc\_video\_conferencing\_doc\_sharing | 32.70 | 31.12 | 33.11 | 31.25 | 0.41 | 0.13 |
| sc\_wordEditing | 37.95 | 36.28 | 37.54 | 36.63 | -0.41 | 0.35 |
|  |  |  |  | **Average** | **+ 0.63** | **+ 0.99** |

Table 3 shows the results when the histogram analysis is used to define the radius. The lookup table in this test is generated by analyzing the local histograms of all the sequences. On average there is a PSNR gain on the chroma planes of +0.63 dB and +0.99 dB. This method works best on the sc\_cad\_waveform sequence (see fragment in Figure 2) where a gain of +0.31dB and +0.56 dB is achieved on top of the best-fixed radius filter for this sequence (radius=5). The sc\_wordEditing sequence shows a PSNR loss for the Cb component. This is because this sequence has lots of (blue) transparency, which is not detected as screen content and consequently the chroma component is more smoothed than required. In spite of the lower PSNR there is still an improvement inn subjective screen quality, as illustrated in Figure 3.

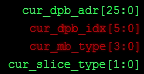
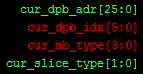


Figure 2: Fragment from the first frame of sc\_cad\_waveform. Left is the original image with the artifacts. Right is the improved image with guided image filtering.



Figure 3: Fragment from sc\_wordEditing: (top) unfiltered with PSNR=37.95 (bottom) filtered with PSNR = 37.54

# Future work



This contribution is a work in progress and not all options have been explored to define the best radius (see Table 2). How the radius can be defined better from the luma component needs more research. To reduce the computational complexity the information from the encoding process (CTB size, prediction information...) can be reused to define the radius.

The lookup table is dependent on the content, so a useful path for future work could also be to do a learning phase at the encoder side and transmit the table in a SEI message to the decoder. In this way the decoder can use the best table for all sequences and should not be pre-configured with the table.

# Conclusion

This contribution demonstrates the use of guided filtering for improving screen content coding for cost efficient video compression. The proposed algorithm provides an objective improvement of visual quality (PSNR) on the chroma planes of +0.63 dB and +0.99 dB. It also demonstrates subjective improvements in the provided test sequences. However this is still a work in progress. More research has to be done to integrate into the encoding chain for improved efficiency.

# References

[1] T. Vermeir, “Use cases and requirements for lossless and screen content coding”, JCTVC-M0172, Incheon, Korea, Apr. 2013.

[2] H. Yu, X. Wang, Y. Jing, “AHG8: More investigation on screen content”, JCTVC-M0320, Incheon, Korea, Apr. 2013.

[3] He, Kaiming, Jian Sun, & Xiaoou Tang. (2013). Guided Image Filtering. *IEEE Transactions on Pattern Analysis and Machina Intelligence* *, 35* (6), 1397-1409.

# Patent rights declaration(s)

**Barco N.V. may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

**Ghent University - iMinds may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**