Objective Evaluation of WebP Image Compression Efficiency

Maurizio Pintus, Giaime Ginesu, Luigi Atzori, and Daniele D. Giusto

Department of Electronic Engineering, University of Cagliari, Italy
`m.pintus@gmail.com, {g.ginesu,l.atzori}@diee.unica.it, ddgiusto@unica.it`

Abstract. Performances of multimedia coding techniques are still improving in terms of compression ratio, coding features, and robustness against errors even if at a slower pace with respect to what we were used to up a decade ago. One of the latest codec which is expected to improve on the state of the art is the WebP algorithm released by Google. With the intent to evaluate the extent of this improvement, in this paper we provide an objective evaluation of the compression efficiency of WebP, by comparing it with alternative algorithms. From the results it appears that the performance of the proposed codec is in line with that of the alternative methods, without achieving any major improvement and lacking several features.

Keywords: Image compression, lossy coding, codec assessment.

1 Introduction

The compression of still images has undergone a significant improvement in the past decades. In the nineties, compression ratios experienced an increase from a 2:1 – 3:1 factor, with lossless entropy coders, to 20:1 and more thanks to the lossy JPEG standard. A decade later, a further 20% increase has been achieved though JPEG 2000. Improvements also involved the development and support for advanced features, such as progressive and lossless to lossy coding, multi-channel and HDR support or region of interest coding. Nowadays, the research and development community is mainly focused on moving pictures, as an extension of still image coding, whereas compression efficiency is improving at a slower pace. The success of a new compression technology then depends on both its performances and features, and is deeply influenced by other commercial factors, such as the presence of patents or licensing royalties and the support in major software packets. Nonetheless, each time such new technology is submitted to the attention of the community, there is the need to evaluate it. The evaluation is performed through comparative studies with existing technologies to test the compression efficiency achieved by the proposed coding algorithm, its computational complexity, and any additional functionalities. An example of such activity can be found in [1].

L. Atzori, J. Delgado, and D. Giusto (Eds.): MOBIMEDIA 2011, LNICST 79, pp. 252–265, 2012.

Compression efficiency expresses the ability of the coding algorithm to maximize the visual quality of a compressed image versus the number of bits used to represent it. Subjective evaluation consists in collecting quality statistics from a sample of users feedbacks and are expensive and time consuming. On the other hand, objective quality assessment makes use of computer algorithms in order to automatically estimate the perceived visual quality.

In this paper, we focus on the compression efficiency evaluation of the new image format from Google, WebP [2]. Released in late September 2010, WebP is a lossy compression algorithm to be used on photographic images, which features predictive coding and exploits variable block sizes. It is reported to offer 39.8% more byte-size efficiency than JPEG for the same quality. Performance evaluation is accomplished by comparing the results of WebP with three state of the art image compression formats (JPEG, JPEG 2000, JPEG XR) in terms of-two objective quality metrics (PSNR and SSIM). The paper is organized as follows: in Section 2 an overview of the competing coding algorithms and the objective quality metrics is provided, Section 3 provides the results, while in Section 4 conclusions are drawn.

2 Background

This section briefly illustrates the coding algorithms under analysis and the quality metrics used in the assessment.

2.1 Coding Algorithms

JPEG dates back to 1990, when the International Standard Organization created the Joint Photographic Experts Group with the task of developing an international compression standard for still pictures. The resulting standard was published in 1993 under the reference ISO/IEC 10918. JPEG compression can be described in six main steps: 8×8 pixels block decomposition, discrete cosine transform, thresholding and quantization, zig-zag scan, run length coding and variable length coding. JPEG compression can be either lossy or lossless.

JPEG 2000 has gained the status of international standard in 2000 as ISO/IEC 15444. The discrete cosine transform was replaced by a newly designed wavelet-based method. New features include: multiple resolution representation, random code-stream access and processing, also called Region of Interest and side channel spatial information. A more accurate description of the JPEG 2000 characteristics can be found in [3].

JPEG XR is based on a technology originally developed and patented by Microsoft. The codec was first announced as Windows Media Photo in 2006 and then renamed to HD Photo in the same year. Thanks to the collaboration with the Joint Photographic Experts Group, HD Photo gained the status of international standard ISO/IEC 29199 under the name JPEG XR. Differences between JPEG XR and JPEG include: 2-level hierarchical transformation within 16x16 macroblock regions, lossless integer transform employing a lifting scheme, optional overlap prefiltering step before each of its 4x4 transform stages, prediction of coefficient values across transform blocks applied to the DC and AC, adaptive reordering and Huffman coding for the coefficients, variable coefficients quantization step sizes inside the same color plane of the image.

WebP is a new format for lossy image compression developed by Google in 2010. It is based on the VP8 video codec with a Resource Interchange File Format (RIFF) container. VP8 innovations include an alternate or constructed reference frame, consisting of image data that is encoded into the bitstream but never displayed. It serves to improve the encoding of subsequent frames by providing an additional predictor than previously transmitted frames [4]. Intra prediction, actually used in the case of image compression, is mostly taken from H.264. Loop filtering is used to remove blocking artifacts introduced by quantization of DCT coefficients from block transforms. WebP uses VP8 intra predictive coding. Images are divided into blocks of pixels of variable sizes, whose values are predicted using the values in neighboring block, so that only the difference (residual) is encoded. Residuals are DCT and Hadamard transformed, quantized and entropy-coded through a non-adaptive arithmetic coder [5]. WebP does not currently provide for alpha channel or HDR support, nor lossless or lossy-to-lossless compression.

2.2 Quality Metrics

PSNR is considered as the most recognized and least complex quality metric. However, its output does not correlate well with the image quality degradation since as perceived by the Human Visual System (HVS).

Structure SIMilarity (SSIM) [7] defines the quality degradation as the product of luminance, contrast, and structural errors affecting the image structure. The structural error is defined as the residual error in the image after its normalization with respect to luminance and contrast. The general form of the SSIM between signal x and y is defined as:

$$\mathrm{SSIM}\,(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma}, \tag{1}$$

where α, β and γ are parameters that define the relative importance of the three components. If $\alpha = \beta = \gamma = 1$, the resulting SSIM index is given by:

$$\mathrm{SSIM} = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x + \sigma_y + C_2\right)}. \tag{2}$$

Although its sensitivity to relative translations, scaling and rotations of images, the SSIM index is quite simple and it performs quite well across a wide variety of image and distortion types. It is able to improve on the traditional PSNR by providing results with are more correlated with the image quality as perceived by the Human Visual System.

3 Assessment

3.1 System Description

All tests were done according to the process described in Fig. 1. All images from each dataset were compressed using the chosen codecs (JPEG, JPEG 2000, JPEG XR and WebP) and the original images were compared with the co-decoded images, on the basis of two quality metrics (PSNR and SSIM). Since not all the codecs provide a tool to directly set the compression ratio, but rather allow for controlling the resulting quality, a first set of coding trials has been performed. Then, the trials from the competing algorithms with the closest compression ratio have been matched. The presented results are the quality index averaged over all the images for each considered dataset. When reporting average results, the lowest bitrate displayed for each codec is the highest bitrate produced through the compression of all images at its lowest quality level. Similarly, the highest bitrate displayed for each codec is the lowest bitrate produced through the compression of all images at its highest quality level. For example, given the chosen output bitrate range from 0 to 8 bpp, with a step of 0.1 bpp, if there were 3 images in the dataset and the bitrates of the images compressed with the lowest quality were 0.15, 0.18 and 0.25, the lowest bitrate displayed in the plot would be 0.3. If the bitrates of the images compressed with the highest quality were 3.05, 3.5 and 4, the highest bitrate displayed in the plot would be 3.

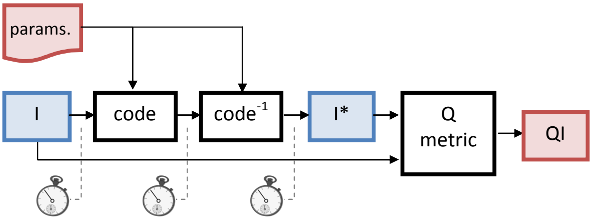

Fig. 1. Performance assessment process

For the experiments, both 24 bpp RGB and 8 bpp luminance images were used from 3 standard datasets. The Canon dataset [8] consists in 18 images with two different resolutions: 512×480 pixels and 512×512 pixels. The Kodak dataset [9] is made of 23 images with 768×512 pixels resolution.

Finally, the “The new test images” dataset contains 15 high-resolution images, ranging from 3008×2000 to 7216×5412 pixels [10]. Because of the limitations of the WebP implementation, images from this dataset have been subsampled by a 2 factor. Table 1 summarizes the software that was used for the compression and conversion between different image formats and the quality measurement.

Table 1. Software tools used for the assessment

Task		Tool
Conversion between lossless formats		GIMP [11]
JPEG codec		Convert Image [12]
JPEG 2000 codec		ImageMagick [13]
JPEG XR	coding	jpg2wdp [14]
	decoding	XnView [15]
WebP	coding	solution that can be found in [16]
	decoding	WebPConvert [17]
SSIM		The SSIM Index for Image Quality Assessment [18]

3.2 Results

Results are first shown as PSNR and SSIM for each greylevel dataset; average values are reported for each dataset, while bpp values are shown in logarithmic scale.

As expected, JPEG achieves the worst performance; the old standard is unable to achieve very high compression ratios and results in lower quality at low bitrates when compared to the other algorithms.

Considering the Canon (Fig. 2) and the Kodak (Fig. 3) datasets, JPEG 2000, JPEG XR and WebP show a similar behavior with slight differences. According to both PSNR and SSIM, WebP performs slightly better than the competing techniques for bitrates from 0.1 bpp to about 0.6 bpp for the Canon dataset and around 1 bpp for the Kodak dataset. For bitrates higher than 1 bpp, JPEG XR outperforms WebP, which provides even lower quality with respect to JPEG 2000. It has to be noted that the WebP codec is unable to run in the entire bitrate range where the other codecs operate. This is a drawback attributable to the used codec and not to the coding algorithm. However, the average values reported in the graph are only computed for bitrates available for all images in the dataset.

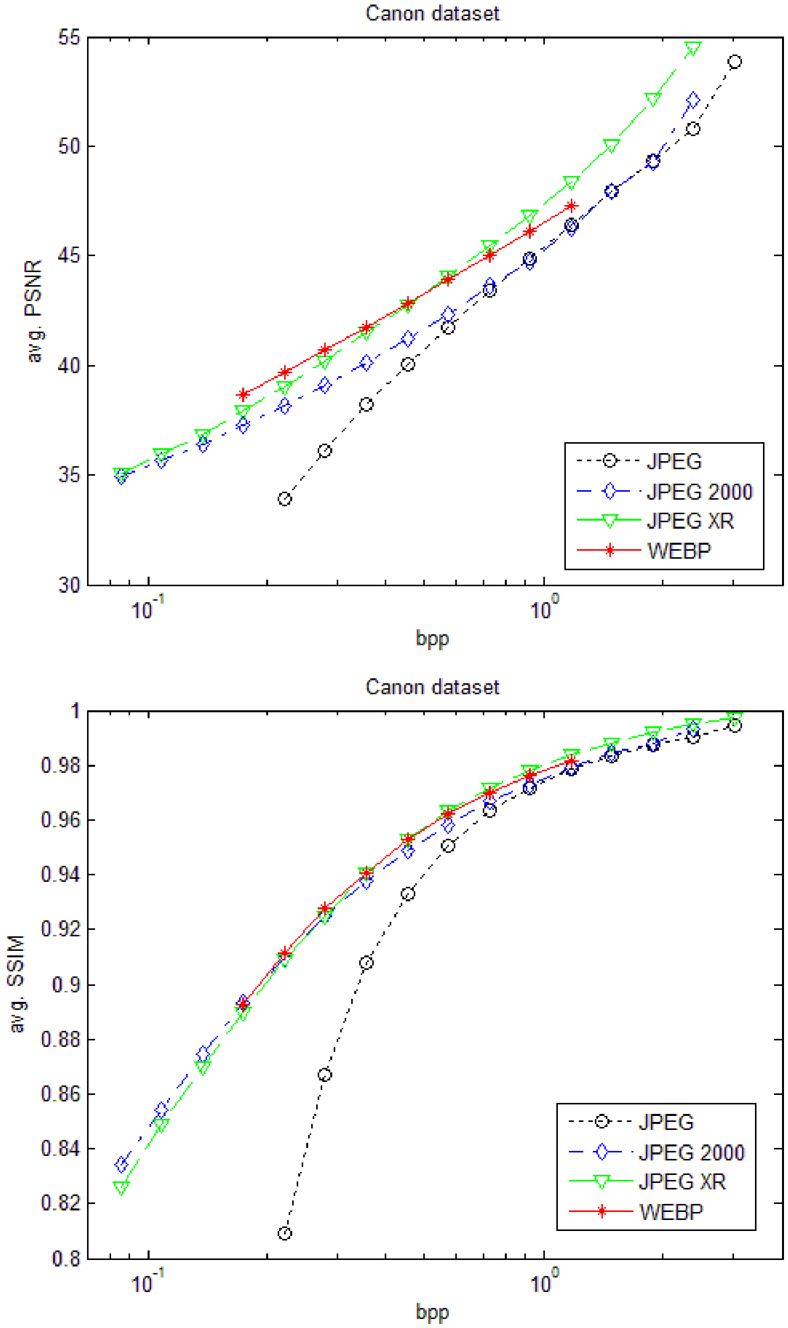

Fig. 2. PSNR and SSIM average results for the Canon dataset

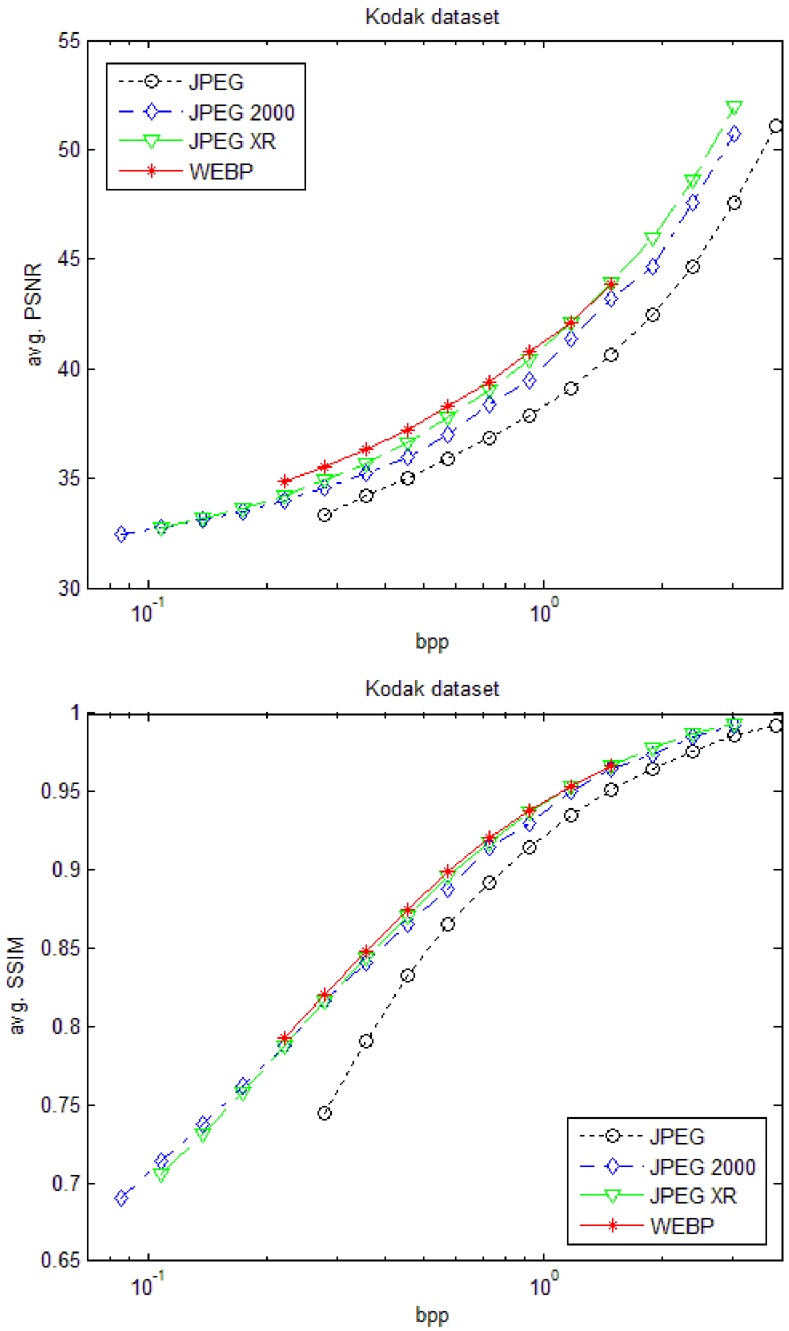

Fig. 3. PSNR and SSIM average results for the Kodak dataset

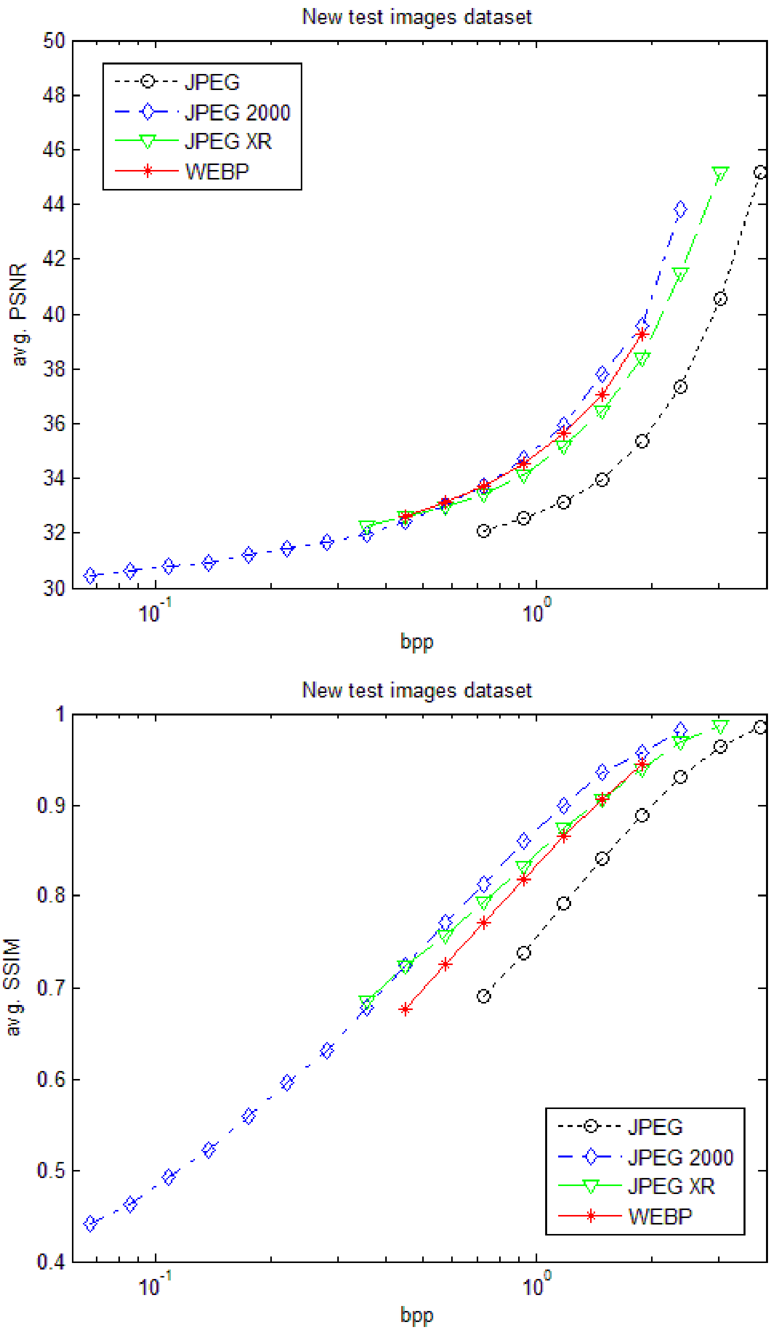

Fig. 4. PSNR and SSIM average results for the "new test images" dataset

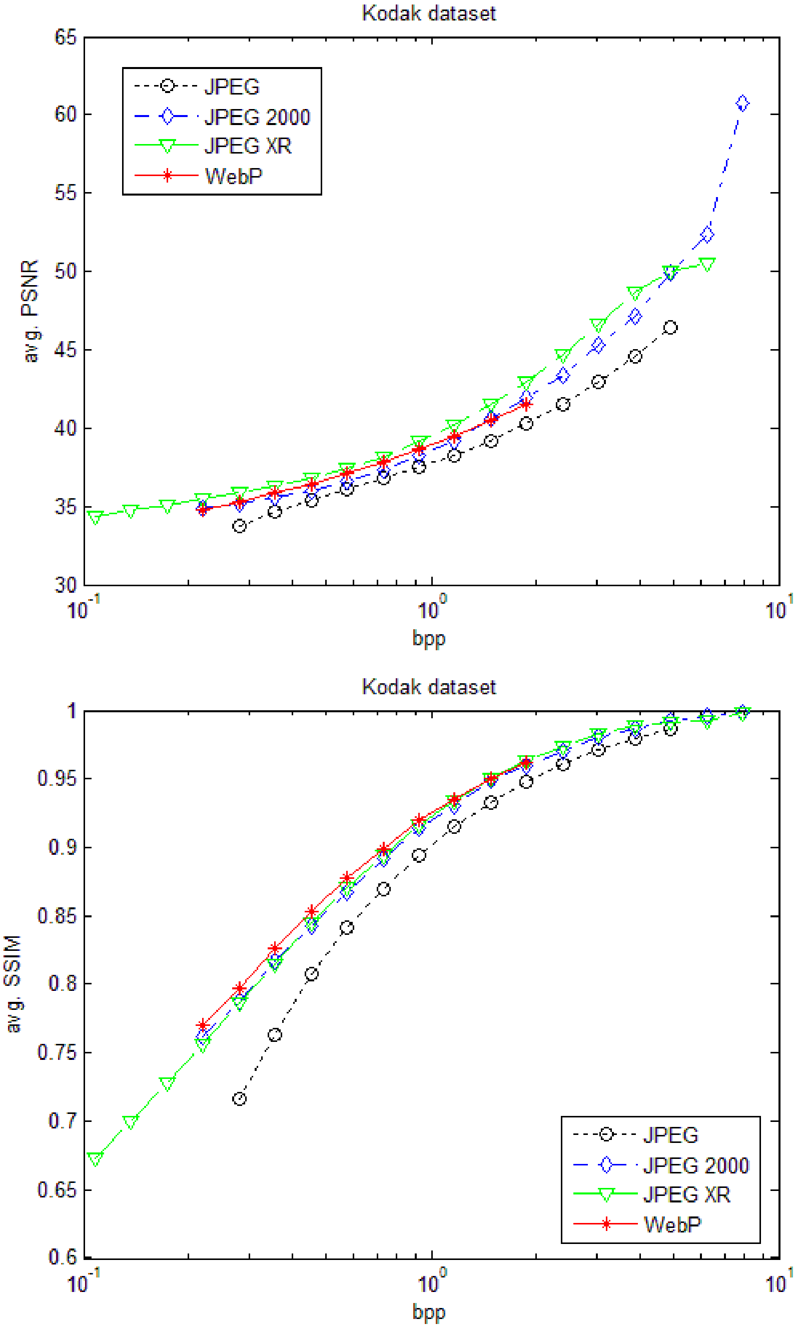

Fig. 5. Average RGB PSNR and SSIM for the Kodak dataset

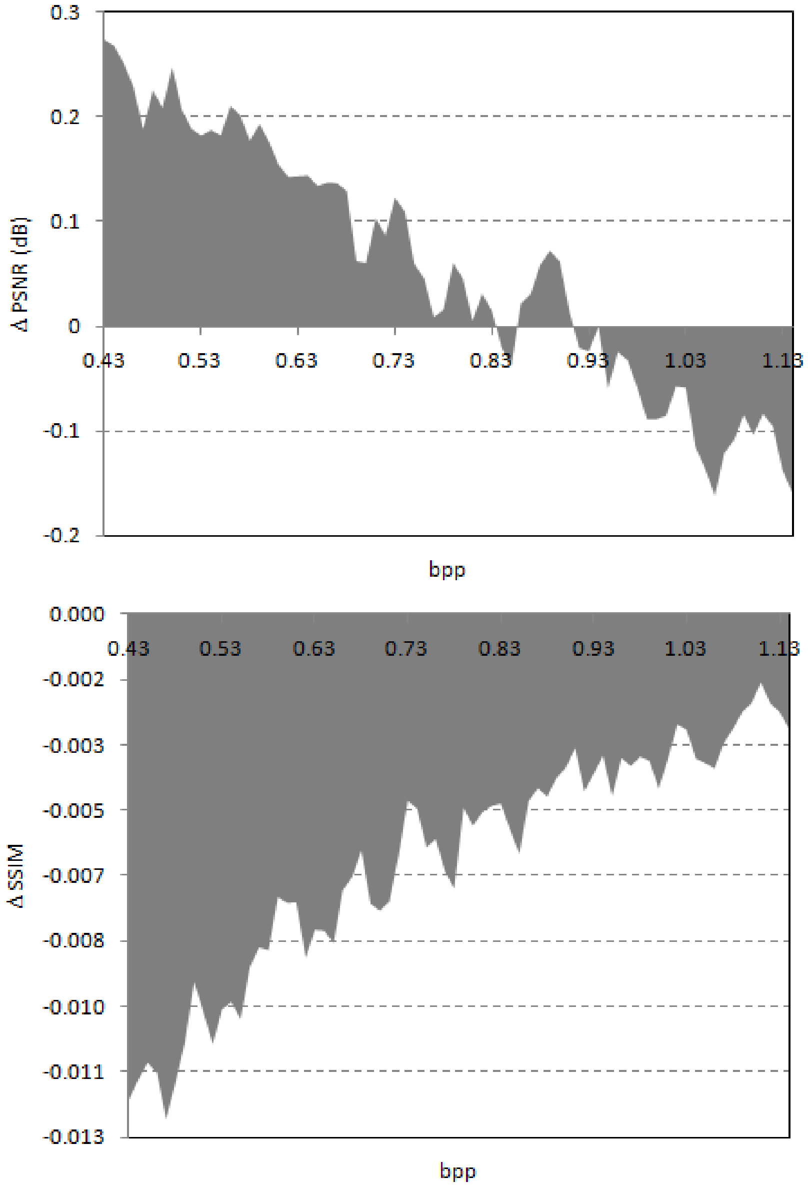

Fig. 6. Comparison between WebP and JPEG XR in terms of the average PSNR and SSIM difference

WebP performance seems to get worse with high-resolution and synthetic images. The average results from the "The new test images" dataset (Fig. 4) show that JPEG 2000 is the best among the competing algorithms in terms of both PSNR and SSIM, whereas WebP prevails over JPEG XR in terms of PSNR but not in terms of SSIM.

The dataset under consideration is made of high-resolution and highly detailed images, two of whom are non-natural scenes (computer graphic and 2D math plot respectively). As in the previous cases, the JPEG 2000 codec implementation is able to reach much higher compression ratios than WebP.

RGB results are not dissimilar from those presented in the previous figures. Fig. 5 presents the average PSNR and SSIM values for the Kodak dataset. Except for the JPEG XR superiority in the PSNR results, which is not in accordance with the SSIM quality, the greylevel results are mostly confirmed. The JPEG XR achievements could be explained with the superior color management offered by such standard. Such figures are common to the three datasets used for this study.

In Fig. 6 an overall comparison between WebP and JPEG XR is shown. PSNR and SSIM results have been averaged over all images from all datasets and only the difference between WebP and JPEG XR is shown. The limited bitrate range has been chosen in accordance with the available measurements. Regarding PSNR, WebP seems to perform better than JPEG XR at low bitrates. Such a superiority is as much as 0.25 dB at 0.43 bpp and linearly decreases to fade at 0.9 bpp, with the inversion of such trend at higher bitrates. On the other hand, SSIM results are in favor of JPEG XR, starting from a 0.011 gain at 0.44 bpp and decreasing as the compression ratio decreases.

Figs. 7-9 show a visual comparison between the selected algorithms for bitrates lower than 1 bpp. As expected, JPEG images present evident blocking artifacts and are perceptibly worse than those from competing algorithms. JPEG 2000, XR and WebP produce comparable outputs, although JPEG 2000 and XR seem to better preserve high-frequency components, while WebP sacrifices some details in favor of visual appearance. Such behavior can be observed by looking at either the petals in Fig. 7 (JPEG XR vs WebP) or the concrete wall in Fig. 8 (JPEG 2000 vs WebP). Some details are lost in the WebP image, even though the average appearance is clear, resembling the effect of bilateral filtering. This effect is not equally acceptable in the case of complex or synthetic images. Fig. 9 shows that JPEG 2000 better approximates the original “zone_plate” signal, followed by JPEG XR. In this case, WebP anisotropic filtering effect results into unnatural edges.

Fig. 7. Portion of image 7 (from the Kodak dataset) compressed at 0.26 bpp

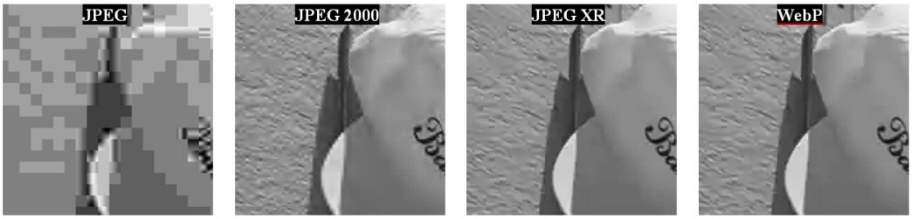

Fig. 8. Portion of image 3 (from the Kodak dataset) compressed at 0.15 bpp

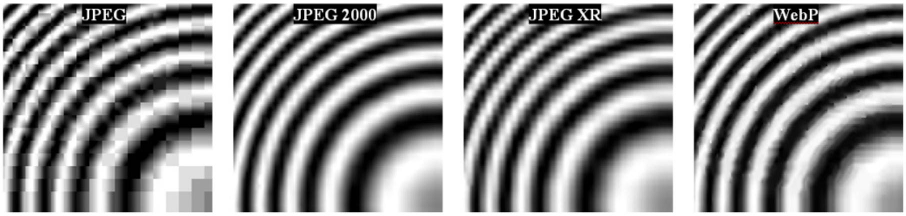

Fig. 9. Portion of the image “zone_plate” (from “The new test images” dataset) compressed at 0.95 bpp

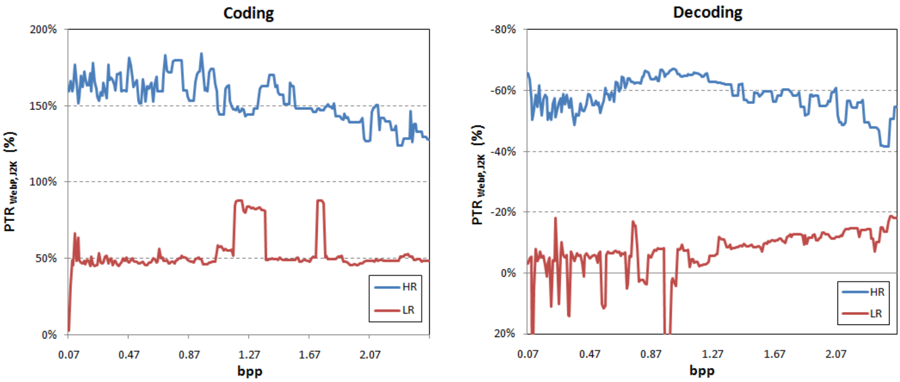

Fig. 10. Average processing time rate between WebP and JPEG 2000 as described in eq. 3 for the coding and decoding phase

Finally, Fig. 10 reports on the average processing time rate between WebP and JPEG 2000 for the coding and decoding of two groups of 10 images (LR: 512×480 and HR: 1500×1500) as:

$$\mathrm{PTR}_{\mathrm{WebP,J2K}} = \left(\mathrm{t}_{\mathrm{WebP}} - \mathrm{t}_{\mathrm{J2K}}\right)/\mathrm{t}_{\mathrm{J2K}}\ \%. \tag{3}$$

A consumer PC with Intel Core2 Duo T5800 CPU @ 2.00 GHz and 4 GB RAM has been used for the experiments. It can be noted that the WebP implementation is significantly slower than JPEG 2000 in the coding phase, while it outperforms the recent compression standard in the decoding. It must also be observed that such difference is relatively small in the case of low-resolution images, which are frequently used in web applications.

4 Conclusions

An objective assessment of the WebP image codec has been provided. The new compression algorithm from Google has been evaluated in terms of compression efficiency by comparing its experimental results with other state of the art codecs. The WebP algorithm shows good efficiency with natural images for bitrates up to about 1 bpp. In this scenario, WebP is often slightly superior than the competing techniques, both in terms of PSNR and HVS-based evaluation. On the other hand, its performance drops with high resolution/highly detailed or synthetic images and for bitrates higher than 1 bpp. Visual comparison reveals a blurry effect probably due to loop filtering, which can be pleasing or unnatural depending on the scene. Processing time evaluation shows that the current implementation of WebP is fairly inefficient in the coding phase and reasonably fast in the decoding of low-resolution images.

From such considerations, and with some caution due to experimenting through a single implementation, WebP classifies among the current state of the art algorithms for image compression, without providing any important innovations or performance boost. Moreover, WebP does not support several features that are common nowadays, such as transparency or lossy to lossless compression. Given such limitations, WebP should really provide further technical improvements, given the aim of its authors to replace the old JPEG, a task already failed by other standard competitors.

References

[1] De Simone, F., et al.: Subjective evaluation of JPEG XR image compression. In: SPIE Optics and Photonics, Applications of Digital Image Processing XXXII, San Diego, CA, vol. 7443 (2009)

[2] WebP homepage, `http://code.google.com/speed/webp/`

[3] JPEG 2000 Part I Final Committee Draft v. 1.0 (March 2000)

[4] Inside WebM Technology: The VP8 Alternate Reference Frame, `http://blog.webmproject.org/2010/05/inside-webm-technology-vp8-alternate.html`

[5] WebP code, `http://code.google.com/speed/webp/index.html`

[6] Chalmers, A., et al.: Image Quality Metrics. In: ACM SIGGRAPH, New Orleans, USA, July 23-28 (2000)

[7] Wang, Z., et al.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. on Image Processing 13(4), 600–612 (2004)

[8] Canon dataset, `http://www.cipr.rpi.edu/resource/stills/canon.html`

[9] Kodak dataset, `http://www.cipr.rpi.edu/resource/stills/kodak.html`

[10] The new test images dataset, `http://www.imagecompression.info/test_images/`

[11] GIMP, `http://www.gimp.org/`

[12] Convert Image, `http://www.softinterface.com/Convert-Image%5CConvert-JPG.htm`

[13] ImageMagick, `http://www.imagemagick.org/script/index.php`

[14] Herila, B.: jpg2wdp, `http://www.bherila.net/2010/01/easily-batch-convert-jpeg-to-hd-photo-with-jpg2wdp/`

[15] XnView, `http://www.xnview.com/`

[16] WebP Discussion, `https://groups.google.com/a/webmproject.org/group/webp-discuss/browse_thread/thread/b8b20e010565dc09/174256299ece6ec2`

[17] WebP for.NET, `http://webp.codeplex.com/`

[18] Wang, Z., et al.: The SSIM Index implementation, `http://www.ece.uwaterloo.ca/~z70wang/research/ssim/`