Image Loss

A Qualitative and Quantitative Approach

Analysis conducted by **Gauntlet Consulting** for **Skeyeon, Inc**.

Table of Contents

Introduction

In this white paper, we address the definition and evaluation of loss in image compression using both a qualitative and quantitative approach, specifically focusing on the JPEG XS compression algorithm developed by IntoPIX. We describe a mathematical method in which the loss associated with lossy compression is quantified. There will also be a discussion regarding the subjective opinion of when an image is visually lossless to the human eye.

Loss vs. Lossless

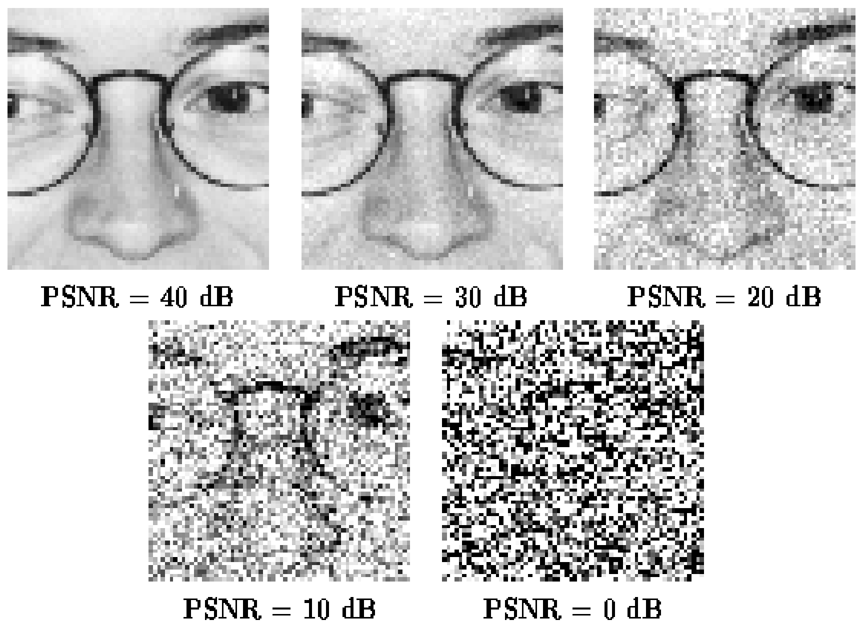
Lossy compression is a common form of compression in which image data is permanently removed to save space. However, the method in which the information is removed prioritizes subjective human visualization. This results in techniques such as chroma subsampling which involves averaging color values while retaining brightness values as brightness is more discernible to the human eye. The loss in lossy compression will differ depending on the image; images with similar neighboring pixels will benefit more as averaging the color values will have a more negligible effect.

On the other hand, lossless compression involves condensing image data based on exact neighboring color values, resulting in a reversible compressed image. However, the image size remains much higher than with lossy compression.

Objective Evaluations

When it comes to analyzing the loss of a certain compression algorithm, metrics based on mathematical formulas can quantify the change in image quality. One of the leading metrics is the Peak Signal to Noise Ratio, which is the ratio between the maximum signal power and the power of the corrupting noise introduced by the compression algorithm. A greater PSNR represents a greater image quality after compression. This metric is used to quantify the image quality loss after compression, and can be represented as the following mathematical formula where the logarithm on the right side is the logarithm of the Mean Square Error between the original frame C and the compressed frame C'. In a scenario with a fixed set of content images and different compression algorithms, PSNR can compare the relative quality more accurately than alternative metrics such as Structural Similarity, Universal Image Quality Index, Noise Quality Measure, and others. However, this does not mean that the PSNR can serve as a complete representation of the quality loss introduced by a compression algorithm. Objective evaluations must be used in tandem with subjective assessment on the visual loss perceived by the human eye.

$$PSNR = 10 \cdot \log I_{max}^2 - \frac{10}{3N} \cdot \log \sum_c \left(\sum_{(i,j)} \left(C(i,j,c) - C'(i,j,c) \right)^2 \right).$$



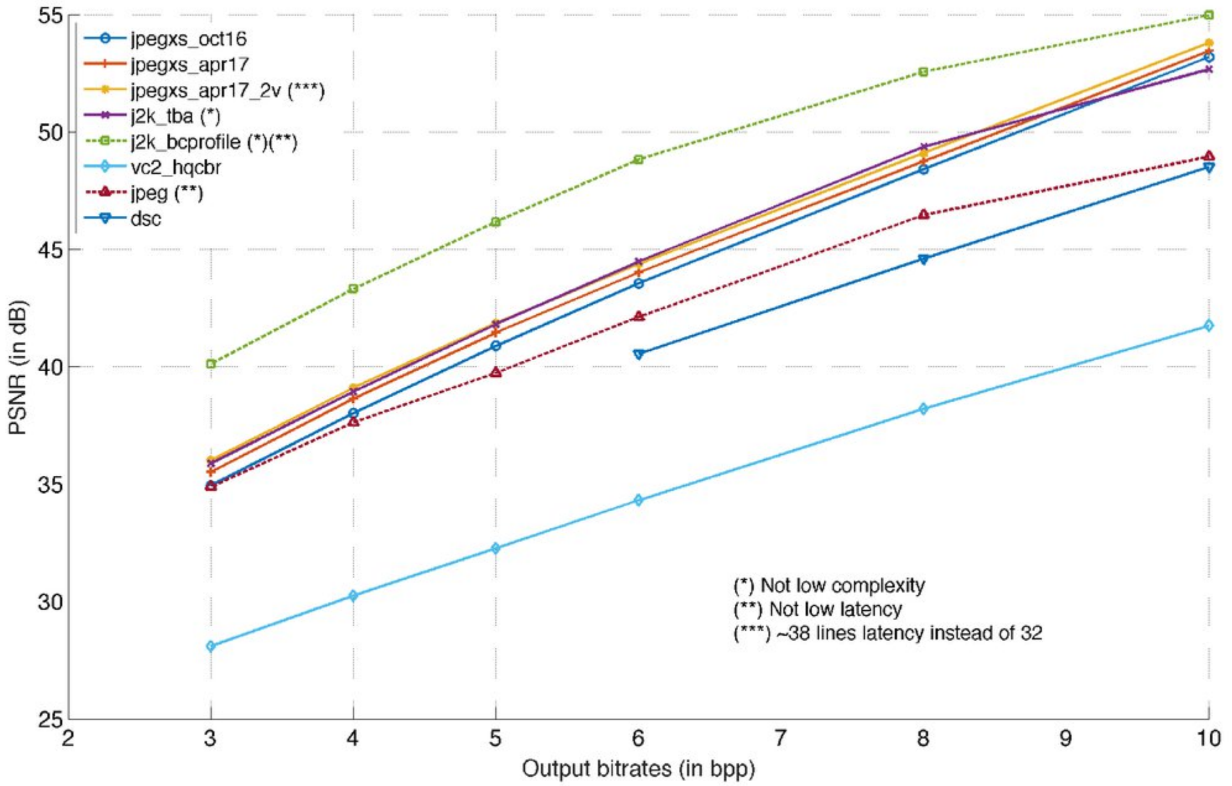
As stated earlier, the PSNR is specifically applicable when comparing different compression algorithms across a fixed sequence of images. Antonin Descampe, a compression technologist at intoPIX, conducted a study on a set of 8 still RGB 8 bit images where he tested three different versions of JPEG XS:

1. The baseline version from October 2016
2. The latest available version at the time from April 2017
3. The latest version but with two vertical wavelet decompositions as opposed to 1, which increases complexity and latency, titled jpegxs_apr17_2v

Then, they selected five anchor points to serve as reference points:

1. A JPEG 2000 configured with tile-based allocation, which is not low complexity, titled J2k_tba
2. A standard JPEG 2000 Broadcast profile, which is not low latency or low complexity, titled j2k_bcprofile
3. Display Stream Compression, a competing visually lossless algorithm.
4. VC-2, another competing visually lossless algorithm.
5. Standard JPEG compression method, which is not low latency.

Descampe compressed each image with each algorithm and benchmarked their average PNSR ratios:



Note: Descampe's paper utilized RGB 444 8-bit images, but the distortion curves reach 10 output bits per pixel. It is important to note that the paper is unclear on how a PSNR was calculated for an output bitrate of 10 bpp when the ground truth of the images is only 8 bpp.

By looking at the distortion curves, it is evident that as the output bitrate increases, the resulting PSNR metric increases. Moreover, it is clear that VC-2 and DSC, the competing visually lossless algorithms, do not reach the same level of quality as the various versions of JPEG XS when it comes to the specific image set utilized. Moreover, the Broadcast profile of JPEG 2000 performs better than all others but has significantly higher complexity and latency. We can also observe significant improvements in the first baseline JPEG XS version from Oct

2016 to the most recent version at the time of the study, from April 2017. JPEG XS is nearing JPEG 2000 Tile Based allocation, which has a higher complexity. JPEG XS with two vertical wavelet decompositions even beats JPEG2000 TBA at low and high output bitrates. It is essential to keep in mind that these PSNR values only represent the algorithm's performance regarding the images selected as part of the evaluation. It would be very valuable to compare the most recent versions of these algorithms with a set of satellite images to more accurately represent Skeyeon's use case.

Subjective Evaluations

In support of objective metrics such as PNSR, a subjective evaluation is essential. One of the most appropriate ways to assess whether a particular compression algorithm is visually lossless is to follow the flicker test method outlined in the document ISO/IEC 29170-2 created by the JPEG Advanced Image Coding and Evaluations group. In this evaluation, viewers are shown "flicker sequences" where the screen is split into two parts. On one side, a crop of the original still image is displayed, and on the other side, a temporal interleaving of the original and compressed frame where they each appear four times per second is shown. Viewers are then asked to discern which side of the screen displays the interleaved content or cast a "no decision" vote. Then, a score is calculated based on the number of correct answers and no decision votes.

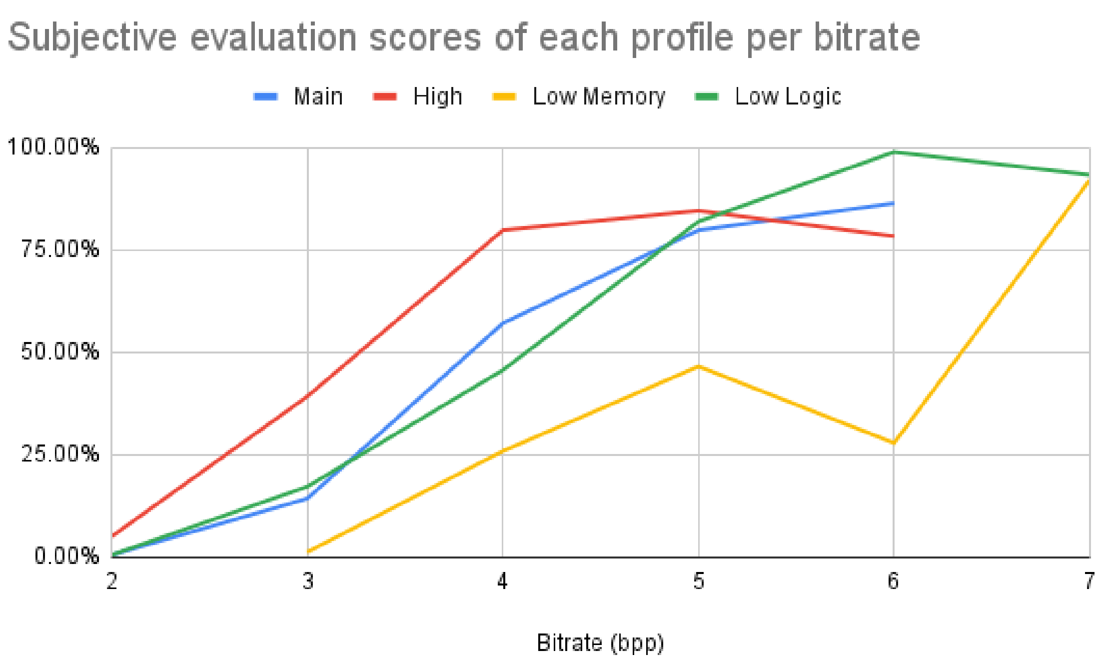
Alexandre Willeme, a member of the JPEG group, conducted this subjective evaluation on four different configurations of JPEG XS. They tested a main, high, low memory, and low logic profile.

1. The Main profile is an intermediate trade-off between complexity, buffering needs, and coding efficiency. It targets use cases such as broadcast production, frame buffer compression in digital TV and compression over display links.
2. The High profile--which is somewhat similar to JPEGXS_apr17_2v from the previous study because it also allows two vertical wavelet decompositions--aims to provide the best coding efficiency (and visual quality preservation) at the cost of increased complexity and more buffering. It targets display links in high-end digital TV, cinema remote production and broadcast contribution.
3. The Low Memory profile is intended at reducing the buffering needs significantly at the cost of a coding efficiency drop. The Low memory profile uses no vertical wavelet decompositions. Its targeted use cases include in-camera compression and hardware-cost sensitive situations.
4. Finally, they tested a Low Logic profile which models applications that can only afford the lowest complexity implementations at the cost of a slight coding efficiency drop. It is intended to be used in broadcast production and industrial cameras.

Features	Main	High	Low Memory	Low Logic
Benefits	Intermediate trade-off	Best coding efficiency and visual quality preservation	Reducing significantly the buffering needs	Lowest complexity implementations
Costs	Intermediate trade-off	Increased complexity and more buffering	Coding efficiency drop	Slight coding efficiency drop

These four profiles were tested with the subjective evaluation methodology for a “flicker test” presented in the ISO/IEC 29170-2 document.

Before analyzing the results, it is vital to keep in mind that this study was conducted with a small sample image size; only six images were selected. Hence, there is a somewhat significant random error introduced in the results--this is evident when looking at the "Low Memory" profile's curve as it dips up and down irregularly. This can be easily solved by recreating the experiment with a much larger sample size of particular images representing the research study's objectives. In other words, this same subjective methodology could be conducted with around 25 satellite images that represent Skeyeon's use case. Keeping that in mind, in this particular case,



the Main profile needs a minimum bitrate of 4bpp to pass the test. The High profile needs 4bpp as well. On the other hand, the Low Memory profile needs 7bpp, and the Low Logic needs 5bpp. These results are not perfectly representative. They only represent the performance of these profiles in terms of the 6 images that were chosen to be part of the experiment. Nonetheless, these results are valuable because they can give an idea of the maximum compression while remaining visually lossless. They serve as an excellent way to quantify the effects of compression on the visual quality of the images that were included in the data set.

Conclusion

There are many different ways to assess the loss introduced into an image by a compression algorithm. They fall into two categories, objective evaluations, and subjective evaluations. Objective methods serve to quantify loss by measuring different metrics directly. Metrics measure specific differences between the original and compressed images. Because they are based on mathematical formulas, the metrics can detect slight changes that are not necessarily observable to humans but are very significant in terms of the raw data. For objective methods, the most common and generic metric is the Peak Signal to Noise Ratio (PSNR). It can serve to create a relative comparison between the performance of different algorithms in an image set. The flicker test is the most suitable method for subjective methods to quantify the visual loss introduced into an image. However, subjective methods only measure the loss that is observable by humans. Nonetheless, it can be helpful to determine the maximum compression ratios that still maintain visually lossless.

Evaluations work best when objective and subjective methods are used in tandem. Each technique can highlight different types of loss; therefore, they are not self-sufficient and must be used in support of each other.

JPEG XS fills a fundamental niche in the image compression subspace. This algorithm attempts to juggle all the needs such as complexity, latency, and buffering to output a visually lossless image.

The research papers presented in this presentation are from 2017. There has been little public documentation on JPEG XS’ performance in objective and subjective evaluations since. It would not be too far off to consider that the performance of compression algorithms has improved since. Therefore, a new analysis of both objective and subjective methods must be conducted with Skeyeon’s particular use case of satellite image compression in mind.

Works Cited

Descampe, Antonin, et al. “JPEG XS, a New Standard for Visually Lossless Low-Latency Lightweight Image Compression.” ‘*Applications of Digital Image Processing XL*, 2017, https://doi.org/10.1117/12.2273625.

Mahmoudpour, Saeed, et al. “Overview of the JPEG XS Core Coding System Subjective Evaluations.” *Applications of Digital Image Processing XLI*, 2018, https://doi.org/10.1117/12.2323435.

Veldhuizen, Todd. “Measures of Image Quality.” *CVonline*, https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/VELDHUIZEN/node18.html.

Willème, Alexandre, et al. “Overview of the JPEG XS Objective Evaluation Procedures.” *Applications of Digital Image Processing XL*, 2017, https://doi.org/10.1117/12.2273168.