Experimental Comparison of PSNR and SSIM Metrics for Video Quality Estimation

Zoran Kotevski and Pece Mitrevski

Faculty of Technical Sciences, Ivo Lola Ribar bb, 7000 Bitola, Macedonia
{zoran.kotevski,pece.mitrevski}@uklo.edu.mk

Abstract. Since the development of digital video technology, due to the nature of digital video, the approach to video quality estimation has changed. Basically there are two types of metrics used to measure the objective quality of processed digital video: purely mathematically defined video quality metrics (DELTA, MSAD, MSE, SNR and PSNR) where the error is mathematically calculated as a difference between the original and processed pixel, and video quality metrics that have similar characteristics as the Human Visual System – HVS (SSIM, NQI, VQM) where the perceptual quality is considered in the overall video quality estimation. In this paper, an overview and experimental comparison of PSNR and SSIM metrics for video quality estimation is presented.

Keywords: Digital video, video compression, video quality metrics, HVS, SNR, PSNR, MSE, SSIM, NQI, VQM.

1 Introduction

Digital video quality estimation is concerned primarily with estimating the video quality of compressed video by means of mathematical calculations. The main goal of this compressed video quality estimation is to calculate the quality using mathematical calculations instead of estimating the quality "by hand" i.e. using larger number of human estimators [8]. If we try to define what video quality is we would end up with a conclusion that video quality is a state of perception by the Human Visual System [1]. So, this means that the best video quality estimator most definitely is the HVS. But, in real world situations, everyday availability of larger number of estimators is a huge problem and video quality metrics comes in handy. Basically there are two types of parameters for measuring the quality of processed digital video: mathematically defined metrics (DELTA, MSAD, MSE, SNR and PSNR) [5], [10], [13], [14] and metrics that have similar characteristics as the Human Visual System – HVS (SSIM, NQI, VQM) [2], [3], [4]. In order to evaluate and validate some of these video quality metrics, an experiment is conducted in which a larger number of differently processed video sequences are created and their PSNR and SSIM are measured. The results are basic charts that present these metrics dependence on the most common changes in processed video i.e. changes in brightness, contrast, hue,

saturation and noise. This paper is concerned with experimental comparison of the performance of the most widely used video quality metrics – PSNR and SSIM.

2 Introduction to PSNR (Peak Signal to Noise Ratio)

The PSNR parameter is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is usually expressed in terms of the logarithmic decibel scale.

The PSNR is most commonly used as a measure of quality of reconstruction of lossy video compression coders [9], [11]. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing coders it is used as an approximation to human perception of reconstruction quality. In some cases one reconstruction may appear to be closer to the original than other, even though it has a lower PSNR. Normally, higher PSNR indicates that the reconstruction is of higher quality. In ideal case the value of PSNR would be 100 dB, but in reality, in the field of image processing, typical values for PSNR are between 30 dB and 40 dB. PSNR is calculated using the mean squared error (MSE) [5], [10], [13] by the equation:

$$PSNR = 10 \cdot \log_{10}\left(\frac{255^2}{MSE}\right) \text{ [dB]} \tag{1}$$

and

$$MSE = \frac{1}{mn}\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}\left[f(x,y) - g(x,y)\right]^2, \tag{2}$$

where:

$f(x,y)$ – is the input variable (color value of the original pixel)
$g(x,y)$ – is the output variable (color value of the processed pixel)
m – is the number of pixels in horizontal direction
n – is the number of pixels in vertical direction

According to the mathematical equations for calculating the previously mentioned metrics (MSAD, MSE, SNR) [5], [10], [13], a conclusion can be drawn that they represent similar error values i.e. the calculated error is of the same degree. Because of this, PSNR can be considered as an unofficial representative of all the above mentioned video quality metrics. Considering its quite convenient characteristics, PSNR metric is still the most widely used metric for video quality estimation in many video processing systems, especially in video compression systems.

But, is it valid enough for us to rely on? Does it perform well enough to be taken for granted? This analysis is presented to compare PSNR to the newer SSIM metric, answer some of the questions and issues about PSNR video quality measurement and give directions to which situations it can or cannot be used.