J. Vis. Commun. Image R. 25 (2014) 874–878

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Visual-PSNR measure of image quality

Alexander Tanchenko

Synopsys Inc., st. Popova, app. 23.D, 197376 Saint-Petersburg, Russia

ARTICLE INFO

Article history:
Received 26 February 2013
Accepted 11 January 2014
Available online 24 January 2014

Keywords:
Image quality
Objective measure of image quality
Peak signal-to-noise ratio
Block-based compression algorithm
Subjective image quality
Image database
Mean opinion score
Human visual system

ABSTRACT

Objective assessment of image quality is important in numerous image and video processing applications. Many objective measures of image quality have been developed for this purpose, of which peak signal-to-noise ratio PSNR is one of the simplest and commonly used. However, it sometimes does not match well with objective mean opinion scores (MOS). This paper presents a novel objective full-reference measure of image quality (VPSNR), which is a modified PSNR measure. It will be shown that VPSNR takes into account some features of the human visual system (HVS). The performance of VPSNR is validated using a data set of four image databases, and in this article it is shown that for images compressed by block-based compression algorithms (like JPEG) the proposed measure in the pixel domain matches well with MOS.

1. Introduction

A wide variety of distortions that appear during compression, transmission, and post-processing may result in a decrease of visual quality of digital images. There are two kinds of quality assessments: *subjective assessments* — for methods where quality scores are evaluated by humans and *objective assessments* — for measures that can automatically estimate perceived image quality. These days several methods within the Video Quality Experts Group (VQEG) are developed [1] on how images are to be presented to experts and how their mean opinion scores (MOS) are collected. However, in practice subjective assessments are too inconvenient and expensive. On the other hand, a good subjective measure should match well with MOS.

An objective image-quality measure can be used in many important image and video processing applications such as rate-control systems and parameters optimization algorithms in image and video coders. For instance, quantization parameters can be chosen to have uniform values of measures over an image. At last, it can be used to benchmark post-processing systems such as deblocking and denoising filters.

Objective image-quality measures can be classified as follows: *full-reference measures* where the original image and the distorted image are available for comparison; *reduced-reference measures* where the original image is only partially available and *no-reference* or *blind measures* meaning that the original image is not available.

This paper deals with full-reference image-quality measures. The simplest and most widely used measure is the mean squared error (MSE) along with the related quantity of peak signal-to-noise ratio (PSNR). They are simple to calculate, have clear physical meanings but they do not match well with MOS [2]. In the last decades many objective quality measures that take into consideration known characteristics of the human visual system (HVS) have been developed. The most well-known of them are SSIM (Structural Similarity Index Metric) [3], MSSIM (Multi-scale Structural Similarity Index Metric) [4], IW-SSIM (Information Content Weighted Structural Similarity Index) [5] and VIF (Visual Information Fidelity) [6]. These metrics have a good correlation with MOS but they are more complex to calculate than MSE and PSNR.

Our goal was to develop a full-reference image quality measure that satisfies the following conditions:

(1) has a good linear correlation with MOS for images that are compressed by block-based image or video compression methods (like JPEG),
(2) is block-based and additive (it means that the value of the measure between two images is evaluated from calculations for all non-overlapped blocks on these images). This condition is necessary to measure can be calculated without additional memory allocation in an image block-based codec.
(3) has the computational complexity comparable with one for PSNR

E-mail address: atanchen@synopsys.com

http://dx.doi.org/10.1016/j.jvcir.2014.01.008

(4) has unit of measurement in decibel (like PSNR). The logarithmic unit of measurement (decibel) allows to measure a wide range of distortions and one simplifies the process of transition to the new measure.

In this paper we propose a novel visual-PSNR (VPSNR) measure that meets the above conditions. The VPSNR measure is block based, additive and has simple calculation formula, which allows it to be used in real-time optimization algorithms in image and video codecs. VPSNR measure models the contrast masking aspect of human visual system (HVS) [7]. VPSNR is based on a weighted MSE measure and in this sense VPSNR is similar to the idea of Noise Visibility Function [8] and implemented as wPSNR measure in [9]. But as we shall see later from experiments described below, VPSNR measure outperforms several existing objective measures including SSIM, MS-SSIM, IW-SSIM, VIF and wPSNR.

2. Definition of the VPSNR measure

The VPSNR measure is evolved from the PSNR measure. In this section we revise the calculation of the PSNR measure and show how to change the PSNR to obtain the VPSNR measure.

Suppose I_x and I_y are the original and distorted images that are being compared. The images are of the same size, but they have different visual properties. Let I_x be the image that has perfect visual quality, then VPSNR can serve as a quality measure of the image I_y. Let $B_k^x = \{x_i^k\}_{i=1}^n$, $B_k^y = \{y_i^k\}_{i=1}^n$, $k = 1, 2, \ldots, N$ be corresponding to non-overlapping luminance pixel blocks of two images I_x and I_y so that

$$I_x = \bigcup_{k=1}^{N} B_k^x, \quad I_y = \bigcup_{k=1}^{N} B_k^y \tag{1}$$

and $B_k^x \bigcap B_m^x = \emptyset$, $B_k^y \bigcap B_m^y = \emptyset$ for $k \neq m$.

The mean square error (MSE) values between the blocks B_k^x, B_k^y are calculated as

$$\text{mse}_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^k - y_i^k\right)^2, \quad k = 1, 2, \ldots, N. \tag{2}$$

The resulting PSNR measure between images I_x and I_y is defined as follows:

$$\text{psnr} = 10 \cdot \log_{10}\left[\frac{\left(2^d - 1\right)^2}{\overline{\text{mse}}}\right]. \tag{3}$$

In this equation, d is the number of bits used to represent samples. For example, in case of 8-bit samples $2^d - 1 = 255$. Further, $\overline{\text{mse}}$ is the mean of mse_k values for all blocks in the image

$$\overline{\text{mse}} = \frac{1}{N}\sum_{k=1}^{N}\text{mse}_k. \tag{4}$$

The VPSNR measure is evolved from PSNR measure with only difference in Eq. (2). To obtain the value of the VPSNR measure between images I_x and I_y the values of mse_k in Eq. (2) are replaced with visual MSE (VMSE) of the blocks

$$\text{vmse}_k = \frac{\text{mse}_k}{1 + 0.5 \cdot \sqrt{\sigma_x^k \sigma_y^k}}, \quad k = 1, 2, \ldots, N, \tag{5}$$

where the unbiased estimate of the standard deviation of the block is calculated as follows

$$\sigma_x^k = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i^k - \mu_x^k\right)^2} \tag{6}$$

and

$$\mu_x^k = \frac{1}{n}\sum_{i=1}^{n}x_i^k \tag{7}$$

is the mean intensity of block B_k^x (σ_y^k is calculated similarly).

It is obvious that VPSNR has the unit of measurement in decibel (like PSNR) and vmse_k is equal to mse_k if $\sigma_x^k = 0$ or $\sigma_y^k = 0$, i.e., there is no variance in pixel luminance. At last, VPSNR is a block-based measure. It means that in order to evaluate VPSNR for original and distorted images it is sufficient to evaluate the measure for all non-overlapped blocks on these images.

3. Application to images

In this section we describe some simple physical observations which led us to VPSNR measure.

HVS based metrics model psychophysical characteristics of the vision system to compute visual quality. HVS involves several different aspects, but we concentrate on contrast masking aspect. Contrast masking refers to the reduction in visibility of a distortion due to the presence of another signal component of similar frequency and orientation in a local spatial neighborhood. In Fig. 1 we have three different images with equal number of pixels N: the original image (a) and distorted images (b) and (c). The mse values between images (a), (b) and (a), (c) are about the same

$$\text{mse}(a, b) = 0.25 \cdot N, \quad \text{mse}(a, c) = 0.26 \cdot N. \tag{8}$$

The images (a) and (b) have equal values of the standard deviation

$$\sigma(a) = \sigma(b) = 0.35 \cdot \sqrt{N} \tag{9}$$

and the value of standard deviation of the image (c) is zero $\sigma(c) = 0$.

It can be observed that the images (a) and (b) are different but they have equal subjective image quality and this discrepancy will be increase with increasing the values of $\sigma(a)$. This observation can be modeled with the following formula of the visual MSE of the block

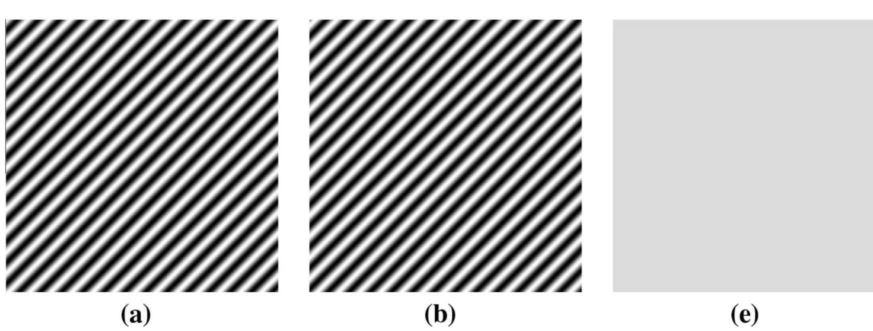

Fig. 1. $\text{mse}(a, b) = 0.25 \cdot N$, $\text{mse}(a, c) = 0.26 \cdot N$, $\sigma(a) = \sigma(b) = 0.35 \cdot \sqrt{N}$, $\sigma(c) = 0$.