Image & Video Quality Assessment and Human Visual Perception

Ravi Kumar Saini [#1], Mrs. Mamta Yadav [#2]

M.Tech Scholar, Assistant Professor

YCET, Narnaul (India)

Abstract— *Images and videos have become an essential part of day to day life, we observe the images and videos and draw the conclusion that a particular image or video is of good quality or bad quality as lots of time we say that the particular video is a high definition video so some questions arise how we assess the quality of an image and video? What are the factors and parameters that make the image of good or bad quality? How the images and videos are perceived by the human eyes? So this paper is concentrated around all the issues related to the quality assessment of the images and the videos. This letter describes the state of art related to the image and video quality by utilizing the some pre-existing quality metrics like PSNR, SSIM and VQM etc. And it also give emphasize on the human visual perception that describes the sensitivities of human eyes towards the images and videos.*

Keywords — *Error sensitivity, human visual system (HVS), image coding, image quality assessment, JPEG, JPEG2000, perceptual quality, structural information, structural similarity(SSIM), Video Quality Assessment (VQA), Peak Signal Noise Ratio(PSNR).*

I. Introduction-

Evaluating the image perceptual quality is a fundamental problem in image and video processing, and various methods have been proposed for image quality assessment(IQA).This letter presents IQA metrics such as Conventional IQA indices (mean squared error (MSE), signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR)), state-of-the-art IQA metrics(structural similarity based image quality assessment (SSIM),multi-scale-SSIM, non shift edge based ratio (NSER) and their limitations . In the non shift edge based ratio (NSER) method the procedures involved include computing the response of classical receptive fields, zero-crossing detection, and non-shift edge based ratio (NSER) calculation. This IQA metric is very simple but very effective and performs much better than most state-of-the-art IQA metric. During acquisition, processing, compression, storage, transmission and reproduction, digital images are subject to a wide variety of distortions any of which may result in a degradation of visual quality. For applications in which images are ultimately to be viewed by human beings, the only "correct" method of quantifying visual image quality is through subjective evaluation. In practice, however, subjective evaluation is usually too time-consuming, expensive and inconvenient. To develop quantitative measures that can automatically predict perceived image quality is the goal of research in objective image quality assessment.

Image quality assessment (IQA) has been becoming an important issue in numerous applications such as image acquisition, transmission, compression, restoration and enhancement, etc with the rapid proliferation of digital imaging and communication technologies. For many scenarios, e.g. real-time and automated systems the subjective IQA methods cannot be readily and routinely used, it is necessary to develop objective IQA metrics to automatically and robustly measure the image quality. It is anticipated that the evaluation results should be statistically consistent with those of the human observers. In the past decades the scientific community has developed various IQA methods. Objective IQA metrics can be classified as full reference (FR), no-reference (NR) and reduced-reference (RR) methods according to the availability of a reference image. Objective image quality metrics can be classified according to the availability of an original (distortion-free) image, with which the distorted image is to be compared. Most of the existing approaches are known as full-reference, meaning that a complete reference image is assumed to be known. In many practical applications, however, a no-reference or "blind" quality assessment approach is desirable and the reference image is not available. In a third type of method, the reference image is only partially available, in the form of a set of extracted features made available as side information to help evaluate the quality of the distorted image, this is referred to as reduced-reference quality assessment. This paper focuses on full-reference image quality assessment.

II. CONVENTIONAL IQA INDICES

The conventional metrics such as the mean squared error (MSE) and the peak signal-to-noise ratio (PSNR) operate directly on the intensity of the image. The mean squared error (MSE) is the simplest and most widely used full-reference quality metric, computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of peak signal-to-noise ratio (PSNR). These are appealing because they are simple to calculate; have clear physical meanings and they are mathematically convenient in the context of optimization. In the last three decades, the development of quality assessment methods that take advantage of known characteristics of the human visual system (HVS).The majority of the proposed perceptual quality assessment models have followed a strategy of modifying the MSE measure so that the errors are penalized in accordance with their visibility. An image signal quality can be evaluated as a sum of an undistorted reference signal and an error signal. The loss of perceptual quality is directly related to the visibility of the error signal is a widely adopted assumption. The simplest implementation of this concept is the MSE, which objectively quantifies the strength of the error signal. But two distorted images with the same MSE may have very different types of errors; some are much more visible than others. Most perceptual image quality assessment approaches proposed in the literature attempt to weight different aspects of the error signal according to their visibility, as determined physiological measurements in animals or by psychophysical measurements in humans. This approach was pioneered by Manos and Sakrison [2], and has been extended by many other researchers over the years

Limitations

1) They do not correlate well with the subjective fidelity ratings.
2) They are not very well matched to perceived visual quality
3) These do not remove the dependencies in input signal
4) It is not clear that error visibility should be equated with loss of quality as some distortions may be clearly visible but not objectionable.
5) Near-threshold models cannot be generalized to characterized perceptual distortions larger than threshold To overcome these limitations Structural Similarity Based Image Quality Assessment (SSIM) is proposed.

III. STRUCTURAL SIMILARITY BASED IMAGE QUALITY ASSESEMENT

Natural image signals are highly structured. Their pixels exhibit strong dependencies, and these dependencies carry important information about the structure of the objects in the visual scene, especially when they are spatially proximate. The Murkowski error metric is based on point wise signal differences, which are independent of the underlying signal structure. Although most quality measures based on error sensitivity decompose image signals using linear transformations. The motivation of this approach is to find a more direct way to compare the structures of the reference and the distorted signals. It is based on the assumption that the human visual system is highly adapted to extract structural information from the viewing field. It follows that a measure of structural information change can provide a good approximation to perceived image distortion. This new philosophy can be best understood through comparison with the error sensitivity philosophy. The problems of natural image complexity and decor relation are also avoided to some extent because this metric does not attempt to predict image quality by accumulating the errors associated with psychophysically understood simple patterns. Instead, this metric proposes to evaluate the structural changes between two complex-structured signals directly.

Limitations

1) SSIM index is a single-scale approach.
2) It achieves the best performance when applied at an appropriate scale, this is drawback of the method because the right scale depends on viewing conditions (e.g., display resolution and viewing distance), but a single scale approach lacks the flexibility to adapt to these conditions. To overcome this drawback multi-scale SSIM is proposed that weight the relative importance between different scales. Differences of Error sensitivity approach and Structural Similarity based IQA

Error sensitivity approach	**Structural Similarity based IQA**
1.Estimates errors to quantify the image degradation 2.It is difficult to explain why contrast-stretched image has very high quality 3.It is bottom up approach 4.It has the supera-threshold problem	1.Considers image degradation as change in structural information 2.It is easy to explain why contrast-stretched image has very high quality 3.It is top down approach 4.It overcomes supera-threshold problem as it does not rely on threshold values

IV. MULTI-SCALE STRUCTURAL SIMILARITY

The perceivability of image details depends the sampling density of the image signal, the perceptual capability of the observer's visual system and the distance from the image plane to the observer. In practice, the subjective evaluation of a given image varies when these factors vary. A single-scale method as described in the previous section may be appropriate only for specific settings. To incorporate image details at different resolutions Multi-scale method is a convenient way.

Limitation: This approach is still rather crude and ad-hoc it does not work under much more broader application.

V. NON SHIFT EDGE BASED RATIO: A NEW IMAGE QUALITY METRIC

NSER overcomes the drawback of MS-SSIM .This metric works robustly across different IQA databases. It achieves better performance than performance to state-of-the-art IQA metrics, such as MS-SSIM. NESR use the earliest vision features, more specifically, zero-crossing edges only, to measure the difference between reference and distortion images. The zero crossings is defined as the information carried by the edges is represented by their spatial locations in the image. When an image is distorted from the original one, the positions of edge points will change accordingly. If there is more distortion, then there will be higher the change in the degree of the edge positions. Therefore, it is a straightforward idea to compare the edge maps of the reference and distorted images to measure their difference. In this metric the reference and distorted images are well registered, which is commonly assumed in IQA research. It is difficult to pair the edge points in the reference and distorted images and comparing the locations of the same edge point in the reference and distorted images. Considering the fact that when the image is distorted the significant edges in an image won't easily change their spatial locations, the edges are investigated that stay in their original locations after the image is deteriorated, and those edges are Non-Shift Edges (NES) their map is defined as below:

$$\mathrm{NSE}(A,B) = F_A \cap F_B$$

Here, A and B denote the reference and distorted images, respectively F_A and F_B are the edge maps of them. An edge map is a binary image, where "1" denotes an edge point and "0" denote a non-edge point. Obviously, the NSE map can be calculated by the "AND" operation of the two binary edge maps, denoted by $F_A \cap F_B$. The variation of the number of edge points in NSE can be used to measure the image quality. Clearly, the more serious the distortion is, the fewer points the NSE map will have. By considering the different contents in different images, the number of edge points in NSE should be normalized by that in the reference image. The proposed algorithm is compared with state-of the-art IQA metrics of different classes: IFC and VIF which are based on the information theory framework, SSIM, UQI and MS-SSIM, and which are based on the structural distortion, NQM and VSNR which are based on the HVS model, as well as the L2 distance based PSNR. All of them work on the luminance component only.

Differences of MS-SSIM IQA metric and NSER IQA metric

MS-SSIM IQA metric	NSER IQA metric
1) MS-SSIM mimics functionally the IQA of HVS to build the metric. 2) MS-SSIM includes three distortion components: luminance, contrast and the structural-similarity, among which the structural-similarity is the core factor.	1) NSER uses only the early vision features (i.e., edges) in the IQA metric design. 2) NSER uses only the binary edge maps to measure the image quality in the form of NSE that can be considered as the "structural-similarity" in some sense

NSER still achieves comparable performance to MS-SSIM by using only the primitive zero-crossings. This shows that zero-crossings can be efficient for IQA and very effective. The NSE detection eliminates much information redundancy in the image and actually selects the most significant features in the reference and distorted images. The information lost in the process of binary edge detection is not so important for IQA. The pixels belonging to a structure are related to each other with a specific intensity distribution, and the information the structure carries is hidden behind this distribution. When an image is deteriorated, the structure and the distribution vary. This is why the information fidelity criteria [14] and the structural similarity indexes [15], work well for IQA.The image structure features used by the above IQA metrics are constructed from the basic primitive signals generated by ganglion and LNG neurons, and by Marr's theory [16], the information existed in the basic primitive signals can be represented by the zero-crossings and their spatial distribution. The structural variation caused by the image distortion will lead to the change of spatial distribution of zero-crossings. This change can be expressed and measured by using the NSE map and NSER metric.

VI. CONCLUTION

This letter presents the SSIM works according to the human visual perception but on the other hand PSNR doesn't so theoretically SSIM should be more consistent with the MOS and this we have proved experimentally also as we can observe that SSIM's coefficient of correlation is 0.990759 which is more near to 1 than the PSNR's coefficient of correlation that is 0.962583.

Hence we can draw some conclusions from this experiment

- SSIM produces more accurate quality scores than the PSNR
- As the noise in the image increases the quality scores will decrease.

REFERENCES

[1] T. D. Kite, B. L. Evans, A. C. Bovik, and T. L. Sculley, “Digital halftoning as 2-D delta-sigma modulation,” in Proc. IEEE Int. Conf. Image Proc., vol. 1, Oct. 1997.

[2] Q. Lin, “Halftone image quality analysis based on a human vision model,” in Proc. SPIE, vol. 1913, Feb. 1993.

[3] T. Mitsa, K. L.Varkur, and J. R. Alford, “Frequency channel based visual models as quantitative quality measures in halftoning,” in Proc. SPIE, vol. 1913, Feb. 1993.

[4] T. Mitsa and K. L. Varkur, “Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms,” in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, vol. 3, Mar. 1992.

[5] S. Daly, “The visible differences predictor: An algorithm for the assessment for image fidelity,” in Proc. SPIE Conf. on Human Vision, Visual Processing, Digital Display, vol. 1666, San Jose, CA, Feb. 1992.

[6] J. Lubin, “A visual discrimination model for imaging system design and evaluation,” in Vision Models for Target Detection and Recognition, Singapore: World Scientific, 1995.

[7] P. Teo and D. Heeger, “A model of perceptual image fidelity,” in Proc. IEEE Conf. Image Processing, vol. 2, Oct. 1995.

[8] “Perceptual image distortion,” in Proc. IEEE Conf. Image Processing, vol. 2, Nov. 1994.

[9] E. Peli, “Contrast in complex images,” J. Opt. Soc. Amer. Oct. 1990.

[10] T. Mitsa, “Image quality metrics for halftone images,” in Proc. SPIE, vol. 1778, Mar. 1992.

[11] R. Ulichney, Digital Halftoning. Cambridge, MA: MIT Press, 1987.

[12] B. A.Wandell, Foundations of Vision. Sunderland, MA: Sinauer, 1995.

[13] P. Barten, “Evaluation of subjective image quality with the square-rootintegral method,” J. Opt. Soc. Amer. A, vol. 7 Oct. 1990.

[14] SSIM BASED RANGE IMAGE QUALITY ASSESSMENT by W.S. Malpica and A.C. Bovik Vision Res, June 1986.

[15] Overview of scalable H.264 / MPEG4-AVC extension by Thomas Wiegand

[16] Evaluation of MPEG4-SVC for QoE protection in context of transmission errors by Yohann Pitery, Marcus Barkowsky, Patrick Le Callet, Romuald Pépion IRRCyN Lab , Image & Video Communication Group Ecole Polytechnique Nantes, France