IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Volume-3, Issue-3 E-ISSN: 2347-2693

A Survey on Image/Video Quality Assessment- some Challenges and Limitations

Renuka.H[1*] and M.Azath[2]

[1,2] *Department of computer Science, Mets School of engineering, Kerala, India.*

www.ijcseonline.org

Received: Feb/24/2015 Revised: Mar/07/2015 Accepted: Mar/20/2015 Published: Mar/31/2015

***Abstract*—** Image quality assessment has involved the comparison of a corrupted image with an original or perfect version of that given image. Many real time cases, this perfect image is not easily available. This research introduces a new metric, that measures visual quality of a single given image and also quality of video images is considered. Operating in this no-reference framework, one new method is suited for real-world applications, such as television monitoring and digital camera quality sensing. Most of the theoretical basis of this work centers on the notion of level-of-detail. Knowing whether an image is very smooth or highly detailed is important in both the detection and assessment of errors. At this time, there are three types of errors that commonly arise in practice are considered, that are namely blur, noise, and compression. Every given image is assigned a score reflecting its perceived quality. Human test cases may validate the new techniques. In this paper, we discuss several open challenges in an image and video quality research. These challenges coming from lack of complete perceptual models for: supra threshold distortions, natural images, interactions between images and distortions, images containing nontraditional and multiple distortions, and images containing enhancements. Here we also discuss the challenges related to computational efficiency.

.

***Keywords*—** Image Quality Assessment (IQA), NR Method, objective QA

I. INTRODUCTION

Image quality assessment (IQA) plays an important role in the field of image processing. Image quality estimation can be a complicated and hard work since each human have different opinion in physical and psychological parameters [4]. Out of many methods proposed for measuring the image quality, none of them should be treated a perfect one. Research on image quality assessment can be traced back to the early research on quality evaluation of optical systems and analog television broadcast/display systems. Image quality metrics are divided into two kinds subjective and objective. Human visual system (HVS) is an example of subjective IQM. Most IQM are related to the difference between two images (the original and distorted image) and this type is called reference IQM, other IQM are not related to the difference between the two images like reduce reference IQM and no reference IQM. Subjective experiments involve a panel of participants which are usually non-experts, also referred to as test subjects, assess the perceptual quality of given test material such as a sequence of images or videos. Subjective experiments are typically conducted in a controlled laboratory environment. Careful planning and several factors including assessment method, selection of test material, viewing conditions, grading scale, and timing of presentation have to be considered prior to a subjective experiment. Due to the time-consuming nature of executing subjective experiments, large efforts have been made to develop objective quality metrics, alternatively called as objective quality methods. Depending on the degree of information that is available from the original video as a reference in the quality assessment, the objective methods are further divided into full reference (FR), reduced reference (RR), and no-reference (NR).

II. OBJECTIVE QUALITY ASSESSMENT OF IMAGES

• FR methods: According to this method, the entire original image/video is available as the reference. Accordingly, FR methods are based on comparing distorted image/video with the original image/video [4]. In full reference IQA the reference image is want to be known and predict the visual quality by comparing the distorted signal against the reference image Mean Square error (MSE) and peak signal to noise ratio (PSNR) are mostly used. Measurement methods consider the human visual System characteristics to incorporate with perceptual Quality. The HVS measure use the psychophysical measurements to compute the visual quality and the image is decomposed to obtain the gain control model in the sub band decomposed domain.MSE[8] is used to evaluate the quality and is defined as:

$$MSE={}^{1}/_{M\times N}\sum^{N}_{i=1}\sum^{N}_{j=1}(x_{ij}-y_{ij})^{2}$$

Where x is the original image and y is the distorted image M, N are the width and height of the image. When MSE value increases as the compression ratio increases. If the MSE value decreases to zero then pixel by pixel matching of images become perfect. MSE is a very simpler one.

• RR methods: In this case, it is not required to give access to the original image/video but only to provide representative features about texture or other suitable characteristics of the original image/video. The comparison of the reduced information from the original image/video with the corresponding information from the distorted image/video provides the input for RR methods.

RR metric requires a feature vector from the reference image to evaluate the quality and these feature vectors are derived from 'm' the parameters of statistical models. The different approaches used in RR QA are first based on the modeling image distortions, second is based on HVS and third is based on NSS. Based on the HVS method, the features are extracted to provide a reduced description of the image and they are not directly related to any specific distortion system. Training for different types of distortion is needed. Unnaturalness will occur due to the distortions and is measured based on the natural image statistics and here training is not needed for it and is more relevant to the visual perception of image quality.

• NR methods: This class of objective quality methods does not require access to the original image/video but searches for artifacts with respect to the pixel domain of an image/video, utilizes information embedded in the bit stream of the related image/video format, or performs quality assessment as a hybrid of pixel-based and bit stream-based approaches. All these types of visual quality metrics are now considered for standardization by various groups, including the Video Quality Experts Group (VQEG) for video and JPEG Advanced Image Coding (AIC) for images.

III. SUBJECTIVE QUALITY ASSESSMENT OF IMAGES

In subjective methods the human subjects are utilized to perform the task of assessing visual quality. The advantages of this method are, it is most reliable methodology and it provides useful information for the subsequent modeling phase. It gives better understanding of mechanisms underlying the quality perception. The different methods used in are:

A. Single Stimulus (SS) Method

This method is used for evaluating the IQA algorithms i.e. here a set of stimuli is taken one at a time and include a reference image in that set and it is not informed to the observer. Observer evaluates the quality and score is expressed in a numerical category rating. Single judgment is required per assessment and then the average score has been calculated. The quality range will be spanned by the stimuli. But this method induces inconsistency so we move to the other method called quality ruler method.

B. Quality Ruler (QR) Method

This method is composed of a series of reference images and whose scale is already known and they are closely spaced in quality, but span a wide range of quality together. It detect the quality difference between them and the observer find the reference image closest in the quality to the test stimulus by visual matching and quality score is noted[5]. Compared to SS method it is more consistent and QR scores are highly correlated to objective measure of distortions than the SS scores.

C. Mean Opinion Score

Mean opinion score produce the accurate results with small number of scores. It is generated by averaging the results of a set of standard, subjective test and act as an indicator for the perceived image quality.

The other method used are *force-choice method* but it does not tell the difference between the quality of images where in pair wise similarity judgment method the quality difference between the two images are noted[6]. Disadvantages of subjective assessment are time consuming and are difficult to design and cannot be performed in real time.

IV. SUBJECTIVE QUALITY ASSESSMENT FOR VIDEO IMAGES

We take four experimental methods for quality assessment of video images. *Single* and *double stimulus*- methods represent categorical rating, in which observers judge the quality images on a fixed 5-point scale. These methods are dominant in video quality assessment. *Forced-choice*- pair wise comparison is an ordering method, which observers decide which of the two displayed images has higher quality [6]. The method is popular in computer Graphics. But it is very tedious work if large number of conditions needs to be compared.

In the pair wise *similarity judgment*- method observers not only choose which image has higher quality, but also estimate the difference in quality on a continuous scale. Such method is used in the functional measurement approach, which relies on relative judgments.

1) Single stimulus categorical rating

This method involves displaying an image for a short and fixed duration of time and then asks an observer to rate it using one of the five categories: excellent, good, fair, poor or bad such adjectives are commonly used in quality assessment as they give intuitive meaning to the numbers on an abstract quality scale. The five-point scale is a widely used for this approach. But there must be some methods that favor continuous rather than categorical scales to avoid quantization artifacts .This experimental method is also known as *Absolute category rating with hidden reference*. Though 5–10 s presentation time is recommended for video, we found in a study that3s presentation is sufficient to assess image quality, yet it does not slow-down the experiment too much time. Fixing presentation time ensures that a comparable amount of attention is devoted to each image. However, presentation time is a variable for all samples that also affects the overall length of the experiment and also the efficiency of the experimental method. All images are shown in random order and include reference images. There is no time limit in the voting stage but no image is shown during that time. The method is efficient as it requires only n+1 trials to assess n conditions (one additional trial for the reference image).

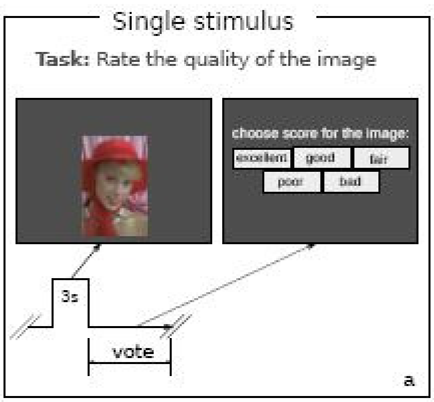

Fig 1: single stimulus rating

2) Double stimulus categorical rating

This is analogous to the single-stimulus method, but a reference image and a test image are presented in random order one after another for 3seconds each following that, a voting screen is displayed on which both images are assessed separately using the same scale as for the single Stimulus method. The method requires n trials to access n conditions [9].

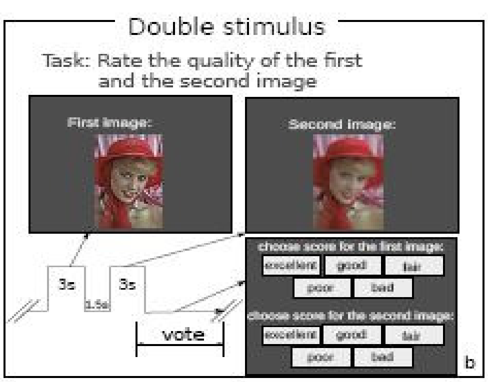

Fig 2: double stimulus rating

3) Forced-choice pair wise comparison

The observers are shown a pair of images (of the same scene) corresponding to different conditions and asked to indicate an image of higher quality .Observers are always forced to choose one image, even if they see no difference between them (thus a forced-choice design). There is no time limit or minimum time to make the choice. The Method is straightforward and thus expected to be more accurate than rating methods. But it also requires more trials to compare each possible pair of conditions: $0.5(n\cdot (n-1))$ for n conditions. The number of trials can be limited using balanced incomplete block designs in which all possible paired comparisons are indirectly inferred. But even more effective reduction of trials can be achieved if a sorting algorithm is used to choose pairs to compare. Efficient sorting algorithms, such as *quick sort*, can reduce the number of comparisons necessary to order a set of conditions to approximately n log n, which could be significantly less than the full comparison, especially if the number of conditions n is large.

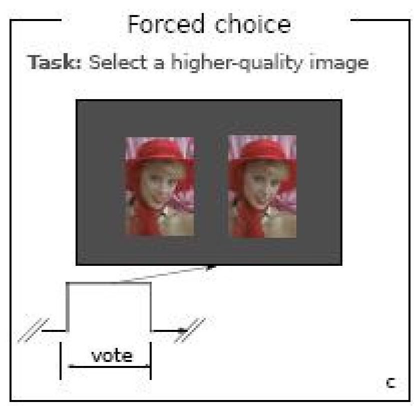

Fig
3: Forced choice method

4) Similarity judgment

While the forced-choice method orders images according to quality, it does not tell us how different the images are [10]. In pair wise similarity judgments observers are not only asked to mark their preference, but also to indicate on a continuous scale how large the difference in quality is between the two images. Observers can choose to leave the marker in the '0' position if they see no difference between the pair. The sorting algorithm used for the pair wise comparisons can also be used for the similarity judgments. The position of the marker (on the left or right side of '0') decides on the ranking of the image pair. If '0' is selected, the images are ranked randomly.

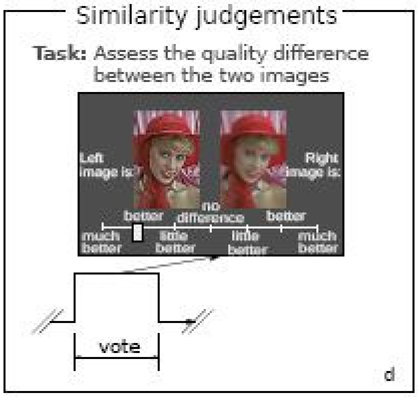

Fig 4: similarity judgment method

Table1: different methods for quality assessment

Methods	description	Result	Quality level
Single stimulus (SS)	A set of stimuli is taken one at a time and include a reference image in that set and it is not informed to the observer.	Score is expressed in a numerical category rating.	inconsistent
Quality ruler (QS)	A series of reference images and whose scale image is already to known	Detect the quality difference between their image and the observers image	Consistent compared to SS
Mean opinion square	It is generated by averaging the results of a set of standard, subjective test and act as an indicator for the perceived image quality.	Mean opinion score produce the accurate results with small number of scores.	consistent
Forced choice	User is given two images and forced to select between them	It does not tell the difference between the quality of images	Consistency is low

V. CONCLUSION

In this paper, we discuss some possible subjective and objective image quality assessment mechanisms. Quality assessment algorithms are needed to monitor the quality for real time applications. Subjective methods are impossible to implement in real time systems, so objective methods are more attracted in recent years. But accurate and efficient IQA measures help to enhance their applicability in real time applications.

REFERNCES

[1] Z. Wang, A. C. Bovik and L. Lu, "Why is image quality assessment so difficult?" IEEE International Conference on Acoustics, Speech, & Signal Processing, (2002).

[2] Damon M. Chandler, "Seven Challenges for Image Quality Research", *SPIE-IS&T/Vol. 9014 901402-2*

[3] Fan Zhang, Songnan Li, **"**Limitation and challenges of Image Quality Measurement".

[4] Deepa maria Thomas,"A survey on efficient No-reference Blur estimation methods", *IJRET, dec-2013 vol: 2,Issue:12.*

[5] Anna Geomi George,"A Survey On Different Approaches Used In Image Quality Assessment*", IJETAE,feb-2013,vol-3,issue-2.*

[6] Rafal K. Mantiuk, "Comparison of four subjective method for image quality assessment", computer graphics forum *Volume 0* (*1981*), *Number 0 pp. 1–13.*

[7] Zhou Wang, "Reduced-Reference Image Quality Assessment Using A Wavelet-Domain Natural Image Statistic Model", Human Vision and Electronic Imaging X, *Proc. SPIE, vol. 5666,SPIE.*

[8] Marius Pedersen and Jon Yngve Hardeberg, "Survey of full-reference image quality metrics", ISSN: 1890-520X ISBN: 978-82-91313-20-7.

[9] Song Nan Li, Lawrence Chun-Man Mak, and King Ngai Nag," Visual Quality Evaluation for Images and Videos", W. Lin et al. (Eds.): Multimedia Analysis, Processing & Communications, SCI 346, pp. 497–544.

[10] Springerlink.com©Springer-VerlagBerlin Heidelberg 2011.

[11] http://en.wikipedia.org/wiki/Recognition_memory.

AUTHORS PROFILE

Renuka.H has completed B Tech in CSE from SreeNarayana Mangalam College Of Engineering and Technology, Moothakunnam, Kerala, in 2010. Presently she is pursuing her M Tech in CSE from Met's School of Engineering, Thrissur, Kerala. Her area of interests include Database design,Image Processing and Artificial Intelligence.

Dr. M. Azath is Head of Department of Computer Science and Engineering, Met's School Of Engineering, Mala. He has received Ph.D. in Computer Science and Engineering from Anna University in 2011. He is a member in Editorial board of various international and national journals and also a member of the Computer society of India, Salem. His research interests include Networking, Wireless networks, Mobile Computing and Network Security.

