Volume - 11 | Issue - 12 | December - 2022 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

CLASSIFICATION REVIEW OF RAW SUBJECTIVE SCORES TOWARDS STATISTICAL ANALYSIS

Statistics

Amitesh Kumar Singam	Postgraduate, IEEE Signal Processing Society.
Bharat Reddy Ramancha	Master of Science.
Venkat Raj Reddy Pashike	Master of Science.
Muhammad Shahid*	Ph.D. *Corresponding Author

ABSTRACT

We proposed a no reference model for quality estimation based on compression artifacts which are principal factor for determining the QoE compared to varying depth range within compression. Moreover, in our past research work, contents with strong compression have been rejected by the Subjects who graded quality of distorted videos and judged unacceptable. The results of our proposed approach confirm that our subjective scores of 120 videos are related to control of visible spatial artifacts of reconstructed videos not distorted and with consideration of User Experience (UX), It was traced out by refining raw subjective scores, it happened due to coding of spatial and temporal perceptual information according to H.264 standards mentioned by ITU only for video quality assessment, i.e., instead of coding spatial of chrominance and temporal information of luminance plane, selected test sequences were coded within luminance plane for both spatial and temporal information

KEYWORDS

NR-VQM, MOS, H.264, SSCQE

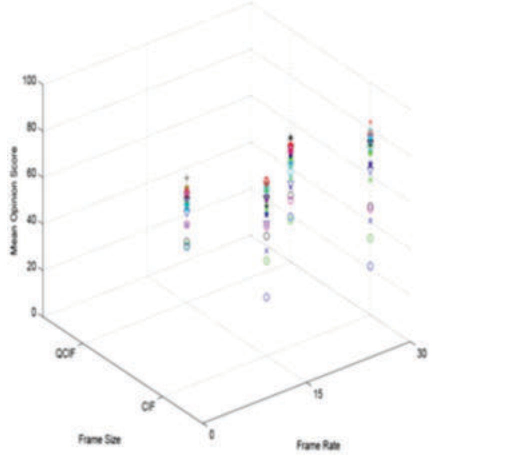
INTRODUCTION

The true judges of video quality are humans as end users of the video services. The scientific process of evaluation of video quality by humans is called subjective quality assessment. However, subjective evaluation is often too inconvenient, time-consuming, expensive and it must be done by following special recommendations in order to produce reproducible and standard results. These reasons give rise to the need of some intelligent ways of automatically predicting the perceived quality that can be performed swiftly and economically

GENERATION OF VIDEO SEQUENCES

Mainly, the proposed method involves extraction of visual quality relevant bitstream parameters and building of a machine learning based model for quality prediction. These parameters were selected carefully to keep the complexity in control and to get reasonable coding information which can represent the coding distortions. Following is the description of the extracted parameters and the rationale of making a parameter a part of the proposed model.

Structural Information: In H.264, the 16x16 macro block (MB) can be sub partitioned into blocks of 4x4, 4x8, 8x4 or 8x8, 8x16, 16x8 depending on the coding mode being chosen for the sake of minimal error in prediction. These values provide an estimate of the structural information of a video. The occurrence of each of these alternatives was taken as a percentage of the total blocks.



Spatial and Temporal complexity: In H.264/AVC, a choice between 4x4 or 16x16 intra coding mode (and 8x8 for high profiles) is made based on texture complexity present in the video frame. The ratio of 4x4 blocks out of total blocks in an intra frame would give an estimate of spatial complexity. And the ratio of intra blocks out of the total blocks in an inter (B or P) frame would give an estimate of temporal complexity present in a video sequence

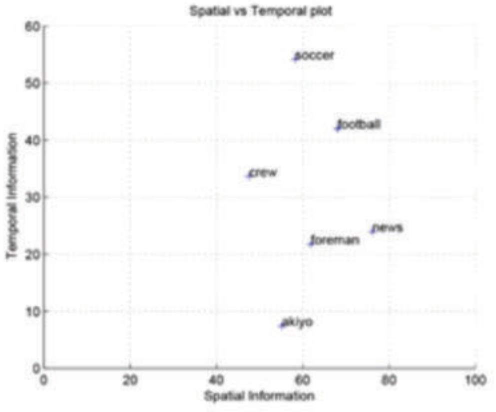


Figure 1 *Spatial versus Temporal Information of CIF*

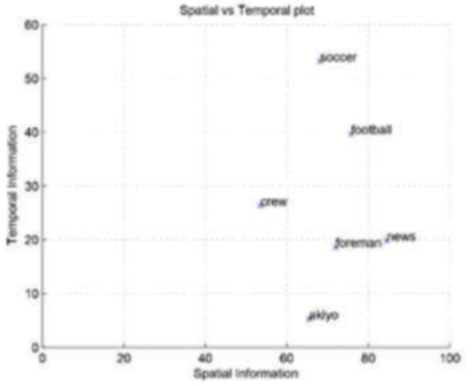


Figure 2 *Spatial Versus Temporal Information Of QCIF*

 PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr

Motion Contents: A set of motion vector (mv) based statistics were calculated to characterize the motion contents of a video. Ratio of zero mvs out of total mvs in a frame and average mv value in a frame were calculated. Moreover, two indicators of motion intensity were calculated. As in H.264/AVC difference in motion vector is sent and not the absolute value of mv; we also included this information as an estimate of change in motion.

Coding Distortion: Quantization is the main source of distortion introduced during the video encoding and QP value is the parameter which is used to steer the scale of quantization. Hence, average QP value used for encoding a video was included in the prediction model.

SUBJECTIVE DATA SCREENING

In our past research work, we have considered the recommendations given by ITU-R BT 500-12[1] lab setup of our experiments. Particularly, the method followed was Single stimulus continuous quality evaluation (SSCQE) where a test video sequence is shown once without presence of any explicit reference, corresponds to the reality where users see only the processed version of videos The tests were conducted in a lab set-up designed in accordance with standards. A flat LCD screen with non-glare surface treatment was used for displaying the video sequences. The used monitor had resolution 1440x900 with 5ms response time and its color temperature was set at 6500K in sRGB mode. Other hardware includes a desktop HP system having 3 GHz AMD processor and 4 GB RAM. A comfortable seating arrangement was made for the subjects at three to four times the high of the display screen.

A software tool developed at the department was used to automate the process of presenting the videos in the center of the screen. Videos were played in a random order for each subject with insertion of the standard intervals (10 sec.) in between for grading. Viewers were not given the privilege to repeat any video and software front end had no controls available for the subjects to alter the intended processes in anyway. The software automatically stored the results in an excel sheet. The grading scale used was 0-100 and the scores were mapped to the 1-5 scale afterwards for further use.

Non-expert subjects who were invited to participate in the tests were mainly international students in different master programme offered at the university and some staff members also took part in the grading campaign. The viewers were introduced to the tests by dictating a common text saying that they are supposed to grade a set of videos on visual quality basis. To ensure no viewer fatigue, the test sessions were kept around half an hour length. In order to obtain reliable results out of raw subjective scores, a twostep filtering method was employed to refine the results. The first step was to detect and discard the observers who exhibited large change of votes compared to the average scores. The second step involved the screening of inconsistent observers without any thought of systematic change. The algorithmic details of these steps are reported in Annex 2 of After performing the refining process, the outliers were removed, and we were left with 18 subjects. Mean opinion score (MOS) was calculated from the scores of these subjects for each test condition.

NEW REFINING ALGORITHM BASED ON H.264 STANDARDS

To Confirm that obtained scores for each time window of test configuration is normal distribution or not, a test was conducted and for achieving first step of refining raw scores, mean, standard deviation and the coefficient for each of all the time windows of each test configuration was computed. This process helps in rejecting observers based on scores significantly far-off from average scores. this step detects and discards the observers based on consistency of votes given and similarly, the distribution of scores is normal or not is confirmed by the means of test. To achieve final step of refining raw scores, mean standard deviation and the coefficient for each of the time windows of each test configuration are calculated. Mean opinion score (MOS) was calculated from the refined scores of subjects for test configuration k as following, where i represents score given by one subject and N is the number of subjects after outlier removal.

$$MOS = \sum_{i=1}^{N} x_k(i)$$

VALIDATION OF PROPOSED MODEL

To obtain the 95% confidence interval between subjective scores and total number of subjects we need to find upper and lower bounds based on sampling distribution i.e., and now it is possible with normal distribution because our proposed algorithm discard inconsistency with raw scores. So, we considered distribution of test configuration for two scenarios i.e., before and after Refining the scores and our model is validated by comparing 95% confidence for two scenarios as illustrated and following mathematical expression for sampling distribution.

$$95\%\ CI = \tan h(\arctan h(r) \pm z_{\frac{\alpha}{2}} * \frac{1}{\sqrt{n-3}})$$

where $z_{\frac{\alpha}{2}}$=1.96
standard deviations, n=number video sequences.

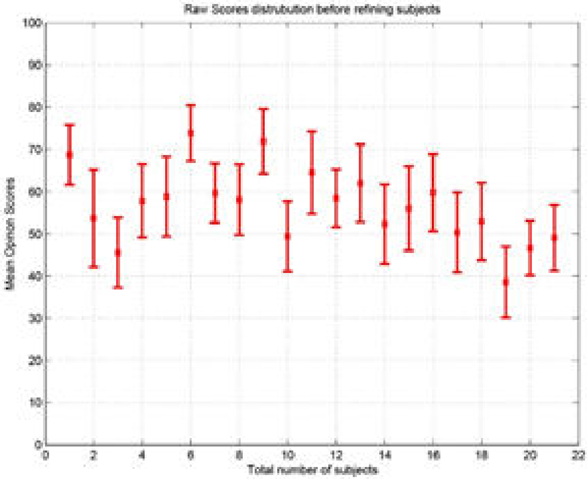


Figure 2 *95% of Confidence interval before refining Scores*

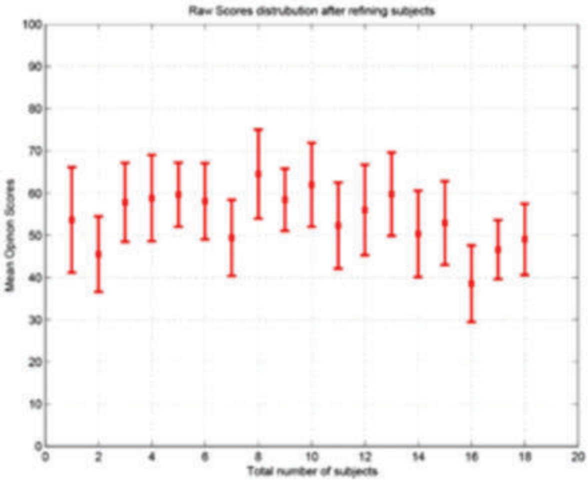


Figure 3 *Figure 2 95% Of Confidence Interval After Refining Scores*

CONCLUSIONS

We concluded that even after achieving consistency within subjective scores, hypothetically we must assume that our test configuration as sampling distribution not normal because after our investigation we concluded that human virtualization characteristics are considered as independent variables.

Acknowledgments

I was confident enough to support and contribute based on my experience and work together with my research group towards achieving good results in the end of my research.

REFERENCES:

[1] Subjective video quality assessment methods for multimedia applications, ITU-T, Recommendation ITU-R P910, September 1999.