EasyChair Preprint

№ 9452

Classification Review of Raw Subjective Scores Towards Statistical Analysis

Amitesh Singam

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

December 11, 2022

Classification Review of Raw Subjective Scores towards Statistical Analysis

Amitesh Kumar Singam, *IEEE Computational Intelligence Society Membership*

***Abstract*—In case of Supervised Learning methods, even hypothetically it's Impossible to combine Speech with Visual characteristics together. Based on ITU Recommendations, we initially assumed and considered the typical two-dimensional plane. In technical terms it should be Chrominance and Luminance plane. In order to quantify impairments of spatial and temporal domain, firstly it should be based on technical assumptions, i.e, we should do mathematical operations based on spatial information within chrominance plane and temporal information within luminance plane. Secondly colour domain exists between two planes and moreover, Scope of subjective quality assessment is essential towards subjective scores as independent variables. But even independent variables are limited to few concepts, out of limited issue, after investigating Human Visual Characteristics, selectively subjective scores are considered as true values judged by humans and 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 visualization characteristics are considered as independent variables**

***Index Terms*—SSCQ, ITU H.264, QoE**

I. 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 has to 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

II. OVERVIEW OF VIDEO COMPRESSION TECHNIQUES

In wireless networks, an uncompressed video needs huge amount of bandwidth and storage. End user cost is proportional to availability of bandwidth and data transmission capacity in network or channel. Therefore, data transmitted in the network is compressed with very effective and lossy compression algorithms. For live video streaming,the most common compression standards are H.263 standardized by International Telecommunication Union (ITU), MPEG-4 part 2 standardized by International Organization for Standardization, H.264 which is also known as Advanced video coding and MPEG-4 part 10 standardized by International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) and ITU.

a) : The initial phase in video generation is sampling in spatial, temporal and color domain. Spatial sampling refers to number of pixels in each frame, Temporal domain sampling refers to resolution in number of pictures per second and color sampling domain provides color space like Gray Scale and RGB.

b) : At present, Video coding algorithms are intended to support a combination of temporal and spatial prediction along with transform coding. Each frame is split into macro blocks, these macro blocks are paradigm in frames. Paradigm represents subset of macro blocks to decode independently. In video Compression we have three classes of frames, B-frames, I-frames, P-frames. They together are called as group of pictures. Since frames are segmented into macro blocks, I-frame is an intra-coded frame which contains intra macro blocks, P-frame is a predicted frame, B-frame is bi-predicted frame which contains intra and predicted macro blocks. A sequence of video which contains I-frames, P-frame and B-frames.

c) : The main reasons for video compression are limited network bandwidth for real time video transmission and limitations in storage capacity. The factors should be considered during compression are quality, compression rate, complexity and delay.Video compression usually utilizes two basic compression techniques, Inter and Intra frame compression. Inter frame compression is compression between the frames and it is designed to minimize temporal redundancy. Intra frame compression is compression within an individual frame, it is designed to minimize spatial redundancy. We employed H.264 standards for video generation, scaling and decoding.

VC/BR	1000kbps	800kbps	600kbps	400kbps	200kbps
Akiyo	+	*	▽	×	○
Crew	+	*	▽	×	○
Football	+	*	▽	×	○
Foreman	+	*	▽	×	○
News	+	*	▽	×	○
Soccer	+	*	▽	×	○

TABLE I
POINTER DELTAILS OF ENCODED VIDEOS

III. DATA SCREENING

In our past research work, we have considered the recommendations given by ITU-R BT 500-12 [1] specifications within lab setup of our experiments. Particularly, the method we followed was Single stimulus continuous quality evaluation(SSCQ), 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 This subjective experiment was conducted in a lab set-up designed in accordance with ITU 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.

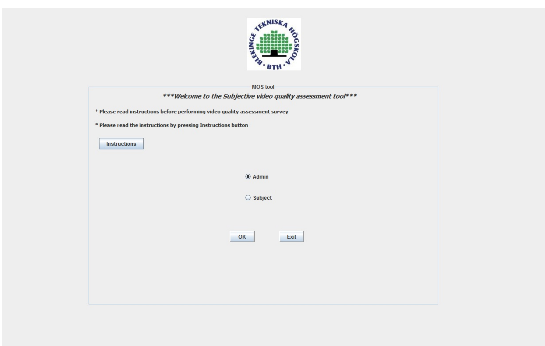

Fig. 1.

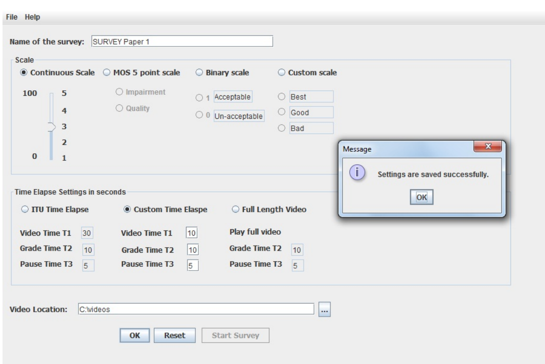

Fig. 2.

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

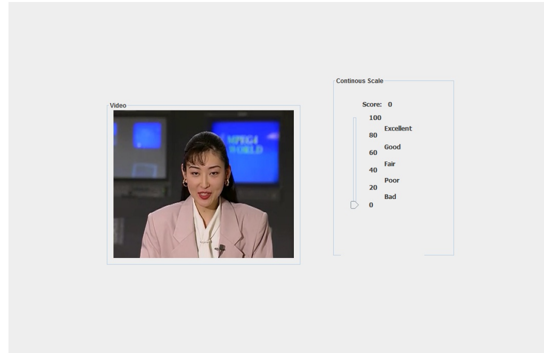

Fig. 3.

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 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

IV. New Refining Alogirthmn based on H.264 Standards

We implemnted this alogirthmn to overcome inconsistency within RAW Subjective scores for 27 subjects and even though subjective scores are true values jugded by human subjects after further investigation we understood that mean opnion Scores are considered as independent not dependant variables. We developed a refining algorithm based on H.264 standards while considering ITU recommendations to overcome inconsistency within predicted scores by validating. 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.

V. 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 discared 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 validted by comparing 95% confidence for two scenarios as illustrated and following mathematical expression for sampling distribution.

$$95\%CI = tanh(arctanh(r) \pm Z_{\frac{\alpha}{2}} * \frac{1}{\sqrt{n-3}}) \quad (1)$$

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

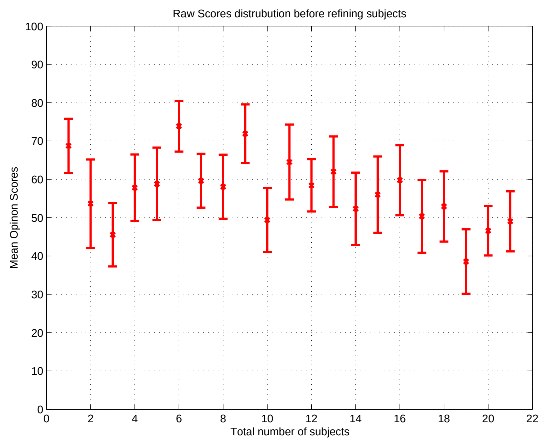

Fig. 4. 95 percent confidence interval of scores before refining

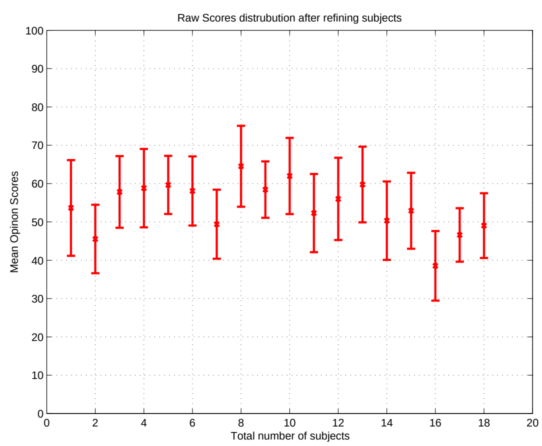

Fig. 5. 95 percent confidence interval of scores after refining

VI. Conclusion

We concluded that even after achieving consistency within subjective scores, hypothetically we must assume that our test configuration as sampling distrubition not normal because after our investigation we concluded that human vistulization characterstics 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.

VII. Biography Section

A mitesh Kumar Singam received the M.Sc. degrees from Blekinge institute of Technology, Karlskrona, Sweden in the year 2012. From the year 2013-2022, he worked in govt based food corporation Pvt Ltd company and his expertise skills are accounting and Billing. His current research interests lies within investigating and analyzing the issue based on decisions making without losing data integrity within any information and he is a member of IEEE since 2022