EasyChair Preprint

№ 9481

Computational Complexity of Human Visual Characteristics Based on Well Known Metric

Amitesh Kumar Singam and Venkat Raj Reddy Pashike

December 15, 2022

Computational Complexity of Human Visualization Characteristics based on Well Known Metric

*

1st Amitesh Kumar Singam
Dept of Electrical Engineering
IEEE, Computational Intelligence Society
Hyderabad, India
0000-0002-2532-0989

2rd Venkat Reddy Raj Pashike
Dept of Electrical Engineering
Blekinge tekniska hogskola
Hyderabad, India
venkatreddy.pashike@gmail.com

***Abstract*—We proposed a well known model for quality estimation based on compression artifacts which are principal factor for determining the Quality of Experience(QoE). Moreover, contents with strong compression have been rejected by viewers and one who graded quality of distorted videos have judged unacceptable in my past research work. It happened due to User Experience because most of viewers were highly experienced one and are also related to same field in the same university. 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 which means it was coded within basic compression and moreover our research interest lies within validation of Subjective and objective quality assessment.We conclude that Subjective scores are considered as independent variables and input features of H.264 bit stream data as dependent variable and moreover input features are validated by correlation coefficients.**

***Index Terms*—VQM, MOS, NR-VQM, H.264, QoE.**

I. Introduction

A. Video Quality Assessment

With extremely huge growth between quality of experience(QoE) and User Experience(UX), the neediness of application developers to provide better quality than each other towards advanced technique to assess the video quality is in great demand now. these days, the competition within developers towards developing multimedia Applications in mobile or hand-held devices boosted the interest in no-reference objective video quality assessment where the availability of reference video is partially available. Due to the huge demand in this area, it is necessary to provide a required level of customer satisfaction given by the perceived video streaming quality. So, No-reference video quality estimation takes major role for these conditions.

a) : In General, the quality Assessment of video Streaming within network or transmission of video in a channel is classified into two methods and they are subjective and objective video quality assessment. Subjective quality assessment is conducted based on human perception since it is concerned with how video is perceived by a viewers and grades the scores for respective videos based on his/her perception. The subject has to vote for the quality of video under certain test configuration which are based on ITU-Recommendations. Since Human perception is considered as the true judgment or values of human towards grading video quality based on precise measurement of perceptual quality but it is quite expensive and tedious in terms of time such as preparation, running and human resources but not presently because of awareness of user experience(UX).

b) : In our survey, according to user Experience community, Objective quality assessment is not an essential because, its an blind perception unlike subjective scores. Even after deploying Human Visualizing Characteristics, in some aspects artifacts such as contrast, orientation sensitivity, spatial and temporal masking effects, frequency selectivity and color perception are completely not understandable by Machines and also it is computationally very expensive and complex to design a quality metric with above aspects. It its an hypothetical approach with human perception.

c) : The impairments visibility which is related to video processing system is subjected to spatial and temporal properties of video content and should be default considered since subjective analysis is much effective in research point of view than objective metric. In our thesis work we have done experiment on both subjective and objective analysis.

II. Related Work

A. Lossy and Lossless Compression

In a communication field, data compression plays a vital role towards designing algorithms in order to reduce size of data, overall compression is performed based on smaller strings of bits i.e., 0s and 1s towards eliminating redundant characters, Data compression is performed whenever there is a need to reduce the size of data. In order to understand this technique, basically we need to understand the difference between information and data. In general, raw data holds disorganized cluster of values, mean numbers, text, symbols, and etc. On the other hand data required to save the information will not only reduce size but perhaps the quality too, but the information will remain intact. Only after considerable loss in information, we can lose the data. This type of compression eliminates redundant of data instead of reducing size of any data through encoding or using any kind of formula and is not feasible even though it is essential, and data can be restored to its original state without loss of any information, but it is less effective for larger data. Theory of rate distortion control plays major role in information theory which provides the practical foundations for Lossy data compression and this type of compression addresses the problem of determining the minimal number of bits per symbol, which is measured by the rate R, that should be communicated over a channel, so that the data transmitted at source can be approximately reconstructed at the receiver without exceeding an expected distortion D.

B. An Overview of Video Compression Techniques

In wireless networks, an uncompressed video needs huge amount of bandwidth and storage and also end user cost is proportional to availability of bandwidth and data transmission in network or channel. Therefore data transmitted in network are compressed with very effective and Lossy compression algorithms. For video streaming in mobile, compression standards like h.263 standardized by (ITU), MPEG-4 part 2 standardized by International Organization for Standardization (ISO) , H.264 which is also known as Advanced video coding and Mpeg-4 part 10 are standardized by JVT(Joint Video Team) of experts both ISO/IEC(International Electro technical commission) and ITU are recommended.

The initial phase in video generation is sampling in spatial, temporal and color domain. Spatial sampling refers to number of pixels in each of the picture based on picture resolution, Temporal domain sampling refers to number of pictures per second based on frame rate, and Color sampling domain provides color space like Gray Scale and RGB. Video Compression Fundamentals

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. This macro blocks are paradigm in frames, which represents subset of macro blocks in order to decode independently.

In Video Compression we have three classes of frames. They are B-frames, I-frames, P-frames. 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 which contains either intra or predicted macro blocks, B-frame is bi predicted frame which contains intra and predicted macro blocks. A sequence of video which contains two key or I-frames ,unidirectional-predicted or P-frame and bi-directional predicted or B-frames

III. Objective Video Quality Assessment

Although subjective quality assessment of video is the ultimate ground-truth for quality measure but it is time consuming, slow and expensive. One has to rely on objective or computerized techniques for faster and efficient quality estimation. PSNR and MSE have been the traditional pixel by pixel comparison objective metrics which, however, don't correlate well with perceptual assessment of humans. Owing to the fact that its desirable to have perceptually relevant objective metrics, Structural Similarity Index and Perceptual Evaluation of Video Quality (PEVQ) are commonly used metrics due to there adherence to the human way of quality assessment. Peak signal to noise ratio (PSNR) is expressed as

$$PSNR = 10\log\frac{{MAX_p}^2}{MSE(n)}$$

MAX_t is maximum pixel value and MSE is average of square of difference between luminance values of corresponding pixels between two frames.

$$MSE = \frac{1}{XY}\sum_{u=1}^{U}\sum_{v=1}^{V}[I_R(x,y) - I_D(x,y)]^2 \tag{1}$$

$I_R(x,y)$ is intensity value of reference frame at pixel location (x, y) and $I_D(x,y)$ is intensity value of degraded frame at pixel location (x, y). X and Y are number of rows and columns in a video frame. PSNR calculated for an entire sequence of video of length N is expressed as

$$PSNR = \frac{1}{N}\sum_{n=1}^{N} PSNR(n) \tag{2}$$

SSIM which was defined for images is also used as an alternative for evaluation of video quality. SSIM considers quality degradations in the frames as perceived changes in the variation of structural information between frames of distorted and original video sequences.

$$SSIM(n) = \frac{[2\mu_{I_R}(n)\mu_{I_D}(n)+C_1][2\sigma_{I_R I_D}(n)+C_2]}{[{\mu_{I_R}}^2(n)+{\mu_{I_D}}^2(n)+C_1][{\sigma_{I_R}}^2(n)+{\sigma_{I_D}}^2+C_2]}$$

$\mu_{(I_R)}(n)$ and $\mu_{(I_D)}(n)$ are mean intensity of n^{th} frame of reference video (I_R) and degraded i video (I_D) respectively, $\sigma_{(I_R)}(n)$ and $\sigma_{(I_D)}(n)$ are contrast of n^{th} frame of reference video (I_R) and degraded video (I_D). C_1, C_2 are constants used in order to evade any instabilities in the structural similarity comparison. SSIM is calculated for entire sequence of video of length N

$$SSIM = \frac{1}{N}\sum_{n=1}^{N} SSIM(n) \tag{3}$$

Its a multi-scale structural similarity(MSSIM) approach which provides further flexibility than previous methods in integrating the variations of conditions like display resolution and viewing distance. MSSIM actually calibrate the factors that states the relative importance of different scales.

$$MSSIM(x,y,n) = [l_M(x,y)]^{\alpha_M}\prod_{j=1}^{M}[c_j(x,y)]^{\beta_j}[s_j(x,y)]^{\gamma_j} \tag{4}$$

$c_j(x,y)$ and $s_j(x,y)$ denotes calculation of contrast and structure comparison at j^{th} scale. $l_m(x,y)$ denotes computation of luminance comparison only at scale M. MSSIM is calculated for entire sequence of video of length N

$$MSSIM = \frac{1}{N}\sum_{n=1}^{N} MSSIM(x, y, n) \quad (5)$$

PEVQ measures quality of video based on mean square of two frames for luminance component. PEVQ is a standardized end-to-end measurement algorithm which estimates mean opinion scores of the video quality by modeling the behavior of the human visual system and it has become a part of ITU-T Recommendation J.247.

a) : Our interest lies in solving issues from above research papers which motivated us to take as challenge to solve issues such as over fitting problem of neural networks, improve the generalization performance and accuracy of our proposed model over the previous methods.

IV. Research Questions

In general terms, the trade-off between computational complexity and its respective reverse approach.

V. New and Innovative idea towards Reverse approach

Generally, in supervised learning methods larger data is projected into high dimensional induced space to perform regression or classification, but in my research point of view, even though we conducted subjective experiments for low resolution videos,our research data is in the format of multidimensional array so, instead of choosing supervised we considered following procedure i.e, transforming array into matrices through Bi-variate covariance method towards statistical analysis.

Normally, we transform arrays into matrices for performing mathematical operations, specifically In my research point of view, features extracted out of Bitstream data based on h.264 standards are correlated with objective scores based on bi-variate correlation method to perform statistical analysis, but in this research its reverse, we used bi-variate covariance method on subjective scores not input features towards performing analysis based on Multi Linear regression

A. Multi Linear Regression

In this method a linear fitting formulation is based on input features(X) not Y (typical independent variables). In terms of research aspects, even though its an regression analysis, it is still an hypothetical approach.

VI. Well Known Quality Metrics

A. Benchmark Measurements

According to VQEG phase, performance of an objective quality prediction model can be evaluated by three parameters which describe prediction accuracy, Monotonicity and consistency. These parameters are evaluated by the following methods.

Accuracy: Pearson linear correlation coefficient describes prediction accuracy of proposed prediction model. Mathematical model of Pearson linear correlation coefficient is given by:

$$r_p = \frac{\sum_{i=1}^{N}(\hat{Y}_i - \hat{Y}) * (Y_i - Y)}{\sum_{i=1}^{N}\sqrt{(\hat{Y}_i - \hat{Y})^2 * (Y_i - Y)^2}}. \quad (6)$$

where $\hat{Y}_i, Y_i$ represent estimated and target values respectively; $Y, \hat{Y}$ represents mean of target and estimated values respectively and N is the total number of each such values.

Monotonicity: Spearman rank order correlation coefficient is related to prediction monotonicity of proposed prediction model between estimated and true values. Mathematical model of Spearman rank order correlation coefficient metrics is:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \quad (7)$$

Where d denotes the difference of ranks between estimated and target value.

Consistency:The consistency feature of an objective metric is assessed by the outlier ratio and it is expressed as ratio of number of outlier points and total data points.

$$OutlierRatio = \frac{Number\ of\ Outliers}{Total\ number\ of\ data\ points} \quad (8)$$

VII. Conclusions and Future Work

We conclude that Within our Contemporary Approach, Subjective scores are considered as independent variables and excluding objective scores, solely, subjective scores are considered as co variance coefficients. Future work is based on developing new algorithm based on bi-variate covariance method.

References

[1] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.