International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

New Methodology of Scaling for Predefined Metrics based on Feature Extraction Using Fundamental Concept of Convolutional Neural Networks

Amitesh Kumar Singam

Blekinge Institute of Technology, Sweden
Halmstad University, Sweden

Abstract: ***In the field of Video Processing, cause of I frame effect is mainly happened within static region of frame due to coarse quantization during frame-by-frame prediction which is most annoying when intra coding is deployed instead of periodically inserted intra frames and skipping of I frame effect was possible only because of default rate distortion optimization in JM reference coder which is based on H.264 standards during encoding process as mentioned in [1], [2]. Moreover, we encountered I frame effect effect during encoding processing because of above reason not because of multiple frames. In our case, we considered principles of Convolutional Neural Network within machine learning fundamentals for Feature Extraction and Classification i. e, it is based on User Experience and also deployed technique manually in the field of Information theory to skip I frame effect, i. e, as author stated in [4], [5].***

Keywords: I frame effect, Multiple Frames, Computer Vision, Feature Extraction, Data Compression, User Experience.

Classification CS. CV (Computer Vision), CS. IT (Information Theory)

1. Introduction

In field of Computer Vision, feature extraction and its classification or annotation towards extracted features is quite essential and moreover quality of bit stream information contains information at nal, slice and macroblock level as author stated et al. [5]. Generally the assessment of video quality is classified into two methods and those are subjective and objective Analysis. so basically Subjective analysis is conducted under predetermined test condition and recommendations while analysis is based on human perception for every single subject, since it is concerned with streaming data which is perceived by a viewer or subject and expresses subjects opinion on a particular video sequence in comparison with its original video sequence without awareness of subject viewing original video.

Subjective Video Quality Assessment

The subjective tests are performed under certain test environment conditions, in our case, its based on ITU - Recommendations. Since Human perception is considered as the true judgment, so precise measurement of perceptual quality became quite expensive and tedious in terms of time such as preparation, running and human resources but in our case, while considering principles of convolutional neural networks which is based on user experience and it is following present trend. More over user experience is referred as scaling of quality metric, so for quantifying, we need to consider that our metric which falls under scaling between consistency and inconsistency. In some cases Objective quality assessment is also essential because it is designed based on HVS (Human Visualizing System) characteristics and more over with some aspects of Human Visualisation system, characteristics of features such as color perception contrast, spatial and temporal masking effects, orientation sensitivity and frequency selectivity are incorporated in within quality metrics.

Test Condition based on ITU - R Recommendations

Results of subjective assessment largely dependents on the factors like selection of test video sequences and welly defined evaluation procedure. In our research work, we carefully employed the specifications recommended by ITU - R BT 500 - 10 as mentioned and VQEG Hybrid test plan in which is explained briefly in following sections. In our research, we considered 6 different set of video sequences, i. e, resolution of CIF and QCIF were selected in raw progressive format based on six different motion content and including various levels of coding complexity recommended by ITU - R P.910. The measurement of spatial - temporal information is not essential due to flickering effect and more over quality of transmitted video sequence is highly dependent on the whereabouts of assessment. Each of 120 video sequences with duration of 10 seconds long with fast, medium and slow motion content were selected for subjective analysis. In our research, the generation or encoding process of Test videos sequences were processed using JM reference software 16.1 based on H.264 standards.

Misconception of I frame effect

In the field of signal processing and communications, for a video transmitted in channel, a coded frame contains high quality information but gradually reduced for following sequence of P - frames and the occurrence of I frame effect will have huge influence on subjective test but we can skip I frame effect for default rate distortion algorithm as stated by author et. al [5]. But in our case its not channel because in the field of telecommunications, networking becomes quite essential and therefore flickering issue has been raised. The feature extraction process for H.264 coded bitstream data was performed in two main steps. First the encoded video were decoded using a preferred settings in configuration file of JM reference software 16.1 in order to generate an XML file of coding parameters for each video sequence. These XML files contained video information at macroblock level such as quantization parameters, absolute and difference motion vectors, and the type of macro block.

Features or Variables at frame level

For instance, its not per video, we evaluated our proposed model at frame level, Moreover, these features were extracted out of H.264 bit stream data and also These features are expected to have high correlation with corresponding perceptual quality score of the selected video, mainly feature extraction of 120 videos has been processed using JM Reference software Version 16.0. The properties of encoded videos are acquired from bit stream data of H.264/AVC which has been generated as a trace file while encoding process. Rather than using completely decoded frame, our interest lies in reversing the entropy encoding of bit stream. By analyzing three successive Nal, Slice and Macro Block Layer and the following features were extracted as stated by author et al. [1], [2].

- Avg QP - Average Quantization Parameter.
- Inter [%] - Percentage of Inter Coded Macro Blocks
- Intra [%] - Percentage of Intra Coded Macro Blocks
- Skip [%] - Percentage of Skip Coded Macro Blocks
- P16x16 [%] - Percentage of Inter Coded Macro Blocks with 16x16 subdivision
- P8x8 [%] - Percentage of Inter Coded Macro Blocks with 8x8 subdivision
- P4x4 [%] - Percentage of Inter Coded Macro Blocks with 4x4 subdivision
- MV_X - Average of Horizontal absolute Motion Displacement
- MV_Y - Average of Vertical absolute Motion Displacement
- MVD_X - Average of the Motion Displacement difference in horizontal direction

$$MVD_X = |MV_x(i_l.j) - MV_x(i_r,j)|$$

Where (i_l,j) and (i_r,j) positions at left and right edge image or frame

- MVD_Y - Average of the Motion Displacement difference in vertical direction

$$MVD_Y = \left|MV_y(i_l.j) - MV_y(i_r,j)\right|$$

Where (i_l,j) and (i_r,j) positions at left and right edge image or frame

- Zero MVs [%] - Percentage of Zero absolute Motion vectors
- Zero MVDs [%] - Percentage of Zero Motion vector Difference
- Motion Intensity

$$MI1 = \sum_{i=0}^{N} \sqrt{MV_{X^2{}_i} - MV_{Y^2{}_i}}$$

where N is the total number of macro blocks in each frame. MV_{X_i} and MV_{Y_i} are the absolute motion vector of the ith macro block in Horizontal (X) and Vertical (Y) directions respectively.

- Motion Intensity II

$$MI2 = \sqrt{MV_{X^2} - MV_{Y^2}}$$

where MV_X and MV_Y are the average of absolute motion vectors in each frame in X and Y directions respectively.

- IinPframes [%] - Percentage of Intra coded macro blocks in P frames.

Scaling of Predefined Quality Metrics

In our research, a non linear fitting formulation was based on Y (typical independent variables) not on input features (X). In the point of research aspects, even though its an regression analysis, it is still an hypothetical approach which resulted in improvised version of development in scaling of predefined quality metrics i. e, range of scaling between consistency and inconsistency to skip I frame effect as stated in [10].

- **Error Concealment**, it quantifies concealed error within neural network applications as author stated in [4], [5], [6].
- **Motion Dynamics**, it quantifies consistency with Motion Intensity I and II
- **Rate Distortion Optimization**, as author stated et al. [7]
- **Correlation Coefficient**, deployed for linear fitting application towards regression analysis as author stated in [3]
- **Covariance Coefficient**, deployed for nonlinear fitting application towards regression analysis stated in [10]
- **Data Annotation**, Feature Extraction, Classification and Regression analysis as author stated et al. [1], [2].
- **Human in loops**, Deep learning for Knowledge Transfer as stated in [8].
- **Human Perceptions** it translates human understanding at perception state as mentioned in [9]

2. Conclusion

our extensive exploration into the realm of subjective video quality assessment underscores the intricate balance between human perception and technical analysis. By harnessing the capabilities of convolutional neural networks alongside the structured methodologies recommended by ITU - R, we have endeavored to bridge the gap between subjective experience and objective metrics. Our research meticulously evaluates video quality through a multi - faceted lens, incorporating both the nuanced understanding of human visual systems and the precision of algorithmic analysis. Through the detailed examination of frame - level features extracted from H.264 bitstream data, we highlight the significant correlation these features hold with perceived video quality. Moreover, our work sheds light on the complexities of encoding processes and their impact on video perception, emphasizing the necessity of considering both subjective opinions and objective assessments in the quest for improved video quality metrics. As we move forward, it is clear, that the integration of human insights with advanced computational methods will continue to be pivotal in advancing our understanding of video quality assessment. Our findings not only contribute to the ongoing discourse in the field but also pave the way for future research aimed at enhancing user experience in the ever - evolving landscape of video technology. Finally, we concluded that scaling of existing quality metric between

Paper ID: SR24303213849 DOI: https://dx.doi.org/10.21275/SR24303213849

consistency and inconsistency can be only used to skip the I frame effect but cannot be resolved.

References

[1] Singam, Amitesh Kumar and Shahid, Muhammad, Full Reference Video Quality Assessment Based on Statistical Based Transform Coefficients (January 26, 2023). Available at SSRN: http: //dx. doi. org/10.2139/ssrn.4386339.

[2] Singam, Amitesh Kumar and Wlode, Kulesza, Revised One, a Full Reference Video Quality Assessment Based on Statistical Based Transform Coefficient (July 15, 2023). Available at SSRN: http: //dx. doi. org/10.2139/ssrn.4511335

[3] Amitesh Kumar Singam, Venkat Raj Reddy Pashike, "Progressive Studies Over Sparseness of Least Squares Support Vector Machines based on Hypothetical Assumption", International Journal of Science and Research (IJSR), Volume 11 Issue 11, November 2022, pp.1059 - 1060, https: //www.ijsr.net/getabstract. php?paperid=SR221118143746.

[4] Amitesh Kumar Singam, Muhammad Shahid. A Just Resembling Translations of Feed Forward and Back Propagation Neural networks. *ESS Open Archive*. October 14, 2023. DOI: 10.22541/essoar.169724934.43005575/v1

[5] Singam, Amitesh Kumar and Pashike, Venkat Raj Reddy and Shahid, Muhammad, An Error Concealment based on Translations of Generalized Neural Networks (August 19, 2023). Available at SSRN: http: //dx. doi. org/10.2139/ssrn.4545855

[6] Singam, Amitesh. (2023). Rationale Neural Networks, Deployment of Activation Function and Its Outcome.

[7] Singam, Amitesh & Pashike, Venkat. (2023). Optimization of Rate Distortion Control Based on Fundamentals of Convolutional Neural Networks. .10.21203/rs.3. rs - 3644320/v1.

[8] Singam, Amitesh. (2023). Humans in Loops for Deep Learning Transfer based on Principles of Recursive Neural Networks.10.31224/3344.

[9] Pashike, V. R. R., Amitesh Kumar Singam and Muhammad Shahid.2023 "Human Perceptions Based on Translations of Recurrent Neural Networks Principles for Low Latency Applications" Preprints. https: //doi. org/10.20944/preprints202308.1931. v2

[10] Singam, Amitesh. (2023). Computational Complexity of Human Visualization Characteristics based on Well Known Metric.10.36227/techrxiv.22242523. v1.