Hypothesis

Not peer-reviewed version

Human Perceptions Based on Translations of Recurrent Neural Networks Principles for Low Latency Applications

Venkat Reddy Raj Pashike , AMITESH KUMAR SINGAM * , Muhammad Shahid

Posted Date: 31 August 2023

doi: 10.20944/preprints202308.1931.v2

Keywords: Neural Networks; Data Screening; Human Perception

Hypothesis

Human Perceptions Based on Translations of Recurrent Neural Networks Principles for Low Latency Applications

Venkat Raj Reddy Pashike [1], Amitesh Kumar Singam [1,*] and Muhammad Shahid [3]

[1] Dublin City University
[2] Halmstad University
[3] Blekinge Institute of Technology
* Correspondence: amisin23@student.hh.se

Abstract: The Necessity of Human resources beyond perception of human understanding towards the Evaluation of video quality or Data screening methodology is conducted based on human perception level since it is concerned with how visual content is perceived by a observer based on observations with his/ her perception on a particular video sequence. Therefore, we considered that the subject has to grade the encoded video sequences under certain test environment conditions based on ITU-Recommendations. Since Human perception is considered as the true judgment and precise measurement of visual content, data screening has became quite essential and quite comfortable to general public due to introduction of User Experience(UX) concept by User Experience community. The translations of A recurrent neural network is based on certain principles, for instance we considered natural language processing which is certainly adaptable towards understanding sequential data and use patterns to predict the consistency within observers. In our research, we adapted principles based on Recurrent Neural Networks while assuming consistency within observers for predicting video quality within data screening environment towards subjective experiments. Moreover,this research work explores the trade offs between Human perception on visual content and consistency of observations within individual observer.

Keywords: neural networks; data screening; human perception

1. Introduction

Even though it is computationally very complex to incorporate spatial and temporal information in real time issues which are useful for a wide range if it correlates well with human perception. The impairments visibility is subjected to spatial and temporal properties of visual content, more over spatial and temporal complexity is quite expensive and time conservative method while considering within Human Visual characteristics.

2. Reference works of ITU Recommendations

Since our methodology of data screening is limited to visual content excluding audio because we considered out test video sequences data at frame by frame level or macro block level where complexity of motion content is essential and audio visual data is negligible. 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 [1] as mentioned and VQEG Hybrid test plan in which is explained briefly in following sections.

a) Test Video Sequences: Shahid et.al [2] considered six different video sequences of CIF and QCIF spatial resolutions were selected in raw progressive format based on different motion content and including various levels of spatial-temporal complexity recommended by ITU-R P.910.

b) Specifications of Data Screening Methodology: The involvement of human observers within laboratory viewing environment specified by ITU-R BT.500-12 Standards are mentioned for Single

 doi:10.20944/preprints202308.1931.v2

Stimulus Continuous Quality Evaluation(SSCQ) process out of Single Stimulus which includes Stimulus Comparison Quality Evaluation.

3. Estimation of Motion Dynamics while considering User Experience(UX)

Investigation of unidentified error at decoder side has been considered into account because of missing motion vectors in reconstructed frames which resulted in increase computational complexity within motion vectors and its not due to poor coding or compression nor because of delay, More over based on assumptions if b frame size is less than predetermined threshold the we must not consider motion intensity feature which was traced out by User Experience, so decision making tree decides either to consider the motion intensity features out of all or not based on Hypothesis.

4. Translations of Recurrent Neural Networks

The measurement of spatial-temporal information is essential due to quality of transmitted video sequence is highly dependent on this whereabouts. The formulations for quantifying spatial and temporal perceptual information of test sequences.

a) Spatial Information: SI is calculated for each frame of video sequence within luminance plane, which leads in obtaining spatial information for n frames.

b) Temporal Information: TI is calculated for amount of temporal changes of a video sequence on luminance plane and measurement of temporal perceptual information

4.1. Principles based on Translations

- Structural Information and Motion Content: Out of all existing features within structural information of bit stream data, motion vector plays quite essential role for quantifying dedicated features such as motion intensity, more over Motion vector complexity is quite high at macro block layer as mentioned in [3].
- Coding Distortion: The effectiveness of changes for identifying the errors within data transmitted due to interruption in signal within a channel is completely based on coding theory and amitesh et al. [4] explained in detail information about rate distortion control and information theory.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2023 doi:10.20944/preprints202308.1931.v2

Figure 1. Spatial and Temporal information computed for luminance component of selected CIF and QCIF videos.

5. Confidence Interval of observations and consistency based on Decision Making Tree

This above box plot illustrates inconsistency between outcome of two possibilities based on decision making tree but one of two decision captured 99 % Confidence Interval between observations and consistency which concludes that there is huge difference in prediction between comparison of two decisions and secondly consistency is based on assumptions.

6. Conclusion

We concluded that our approach is based on principles of Recurrent Neural Networks and our assumptions based on RNN priciples are within reach of Human Perception level because subjective scores are considered as true or standard scores of video quality. Since all the 120 video sequences which we generated for our research work are encoded in JM Reference 16.1 based on H.264/AVC software uses Rate distortion optimization algorithm for improving video quality while video compression and distortion measure in JM encoder is Mean Square Error by default [5].

 doi:10.20944/preprints202308.1931.v2

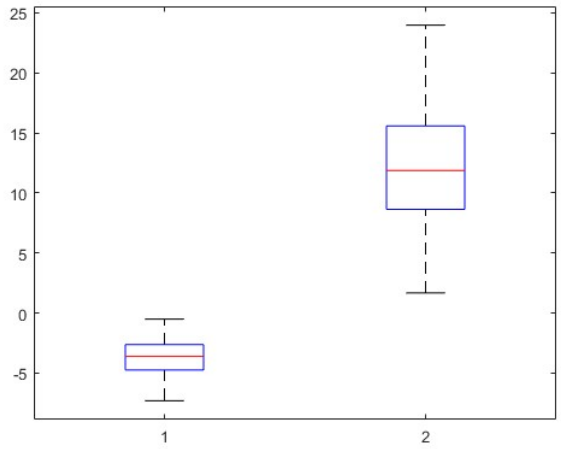

Figure 2. 99% Confidence Interval between observations and consistency for Two Decisions.

References

1. ITU-R Radio communication Sector of ITU, Recommendation ITU-R BT.500-12, 2009. http://www.itu.int/.
2. Shahid, M.; Singam, A.K.; Rossholm, A.; Lovstrom, B. Subjective quality assessment of H.264/AVC encoded low resolution videos. *2012 5th International Congress on Image and Signal Processing* **2012**, pp. 63–67.
3. Singam, A.K.; Wlode, K. Revised One, a Full Reference Video Quality Assessment Based on Statistical Based Transform Coefficient. *SSRN Electronic Journal* **2023**.
4. Singam, A.K. Coding Estimation based on Rate Distortion Control of H.264 Encoded Videos for Low Latency Applications. *arXiv e-prints* **2023**, p. arXiv:2306.16366, [arXiv:cs.IT/2306.16366]. doi:10.48550/arXiv.2306.16366.
5. Singam, A.; Wlodek, K.; Lövström, B. Classification Review of Raw Subjective Scores towards Statistical Analysis. *SSRN* **2023**.