Full Reference Video Quality Assessment Metric on Base Human Visual System Consistent with PSNR

Anastasia Mozhaeva, Lee Streeter
The University of Waikato
Hamilton, New Zealand
am476@students.waikato.ac.nz, lee.streeter@waikato.ac.nz

Igor Vlasuyk, Aleksei Potashnikov
Moscow Technical University of Communications and Informatics
Moscow, Russian Federation
i.v.vlasiuk@mtuci.ru, a.m.potashnikov@mtuci.ru

***Abstract*—Quality assessment of video compression with a focus upon the perceptibly of distortions introduced by compression from the point of view of human perception is an important area of video quality research. Although the subjective assessment of video quality is more accurate regarding human perception of video than objective, it necessitates the very large design space of the subjective experiment. As the result in recent years, there has been a growing interest in the development of improved models of objective video quality. Based on the researches of fundamental limit to human vision, here we offer experiments of visual perception and measure the characteristics of the human visual systems on the "Research of the characteristics of the visual systems" software and test equipment we created. Also, we represent a new quality measuring method we developed, which takes into account the features of the Human Visual System and will be represented comparison estimates this method with current models used by quality assessment reference, and with subjective scores.**

I. INTRODUCTION

Numerous quality assessment (QA) methods have been proposed over the past years. At the present stage of development, all QA methods are classified into broad categories:

Subjective criteria for assessing quality. Such an assessment, in this case, is carried out by a person. As shown by Mohammadu, Ebrahimi-Moghadam and Shirani [1], since a person is a receiver in most digital image or videos processing applications, subjective quality (SQ) is the most accurate method for evaluating quality. Subjective evaluations are expensive and time-consuming, which makes them impractical. Moreover, subjective experiments are further complicated by many factors including viewing distance, a display device, lighting condition, subjects' vision ability, and subjects' mood.

Objective criteria for assessing quality (OQA). Here the evaluation is performed algorithmically. The goal of objective quality metric is to design mathematical models that can predict the quality of an image or video. OQA methods can be classified into three categories. The first category is full-reference quality assessment (FR-QA) where the undistorted, perfect quality reference is fully available. The second category is reduced-reference quality assessment (RR-QA) where the reference is not fully available. Instead, some features of the reference are extracted and employed as side information to evaluate the quality. The third category is no-reference quality assessment (NR-QA) where one does not have access to the reference. For objective measures to be useful to quality assessment, they must correlate with subjective impressions of quality and ideally should be both standardized and independent of the systems or processes involved [1]. According to Ying et al. [2], no-reference perceptual quality prediction is a difficult, unsolved problem of great consequence to the social and streaming media industries that impact billions of viewers daily. Unfortunately, popular NR prediction models perform poorly on real-world distorted pictures or videos.

Criteria for assessing visibility image quality, predicts the likelihood that a human observer will be able to detect differences between a pair of images. The visibility metric is the metric measuring whether introduced changes in images or video are visible or not. According to Ye [3], visibility metrics can provide localized information in the form of visibility maps in which each value represents a probability of detection. D. Chandler, Phan and Alam [4] and K. Wolski et al. [5] note that the key problem that hinders the development of better-quality metrics is limited training data. Also, experiments Ye and Wolski et al. showed that typical quality indicators are not accurate enough to predict the Visual Loss Threshold, and existing visibility metrics are based on a simplified version of models of the human visual system.

According to Chandler, Phan and Alam [4] without question, today's QA algorithms predict quality for a variety of images or video and distortion types remarkably well. However, the focus of QA research was a shift over the last decade from the previously broad objective of gaining a better understanding of how humans judge quality, to the currently limited objective of achieving a better fit to the available subjective data. Also need pay attention mention in [4], that the primary goal of the vast majority of research in visual psychophysics is to gain knowledge of how the human visual system (HVS) operates; any relations to image or video quality are usually secondary and are usually not extensively discussed in such studies. Consequently, it is often up to the designer of a QA algorithm to decide how the psychophysical findings relate to quality without a full view. For changing situation, it needs new knowledge and tests on relations how HVS interacts with QA. In this work, we introduce a VQA method, the results of which can be easily compared with PSNR. However, it will consider the knowledges about the peculiarities of human perception of temporal and spatial characteristics at different brightness levels.

We hope that our research efforts will help motivate further research on new fundamental knowledge about to study the properties of visual perception and measure the characteristics of the human visual systems in relation to image or video quality. The rest of this paper is organized as follows. Section II gives an overview of previous studies of quality metrics dependent on HVS. Section III discusses the RCVS software and test equipment, and the subjective testing procedure is discussed, and objective analysis of the result of the experiment is presented. In Section IV, the quality measuring method we developed is presented. Section V and VI we present result and discuss the findings.

II. Related Work

Research to emulate the performance of the biological visual system has received great attention in the past. Almost 60 years ago, Luizov [6] and de Lange [7] described a fundamental limit to human vision, representing experimental and theoretical studies of parameters of visual perception. Their work describes how sensitivity depends on the light level, the development of visual perception in time, establishing the relationship between the brightness acting on the eye and the effective brightness. Their results provided an enduring foundation for all subsequent studies of video quality with a focus on HVS. Only a relatively small number of existing video quality assessment (VQA) algorithms detect motion explicitly and interpret HVS.

Current models used by quality assessment reference include the FR Peak signal-to-noise ratio (PSNR), Structural similarity image metric (SSIM) [8], Video Multimethod Assessment Fusion (VMAF) [9]. In real-time applications, reference models are problematic. These models serve as a base for the NR VQA research. Another area of use of FR metrics is compression based on neural networks, where PSNR is used as the main quality metric in this time. However, according to Wang and Bovik [10], PSNR poorly correlates with visual quality assessment. PSNR does not have a spatial and temporal psycho-visual model, therefore, for example, even an insignificant shift of the reference and the estimated image in space or the desynchronization of a video sequence in time can give significantly underestimated metric results. The PSNR metric is the most common - almost all studies in the field of image quality, where it is possible to apply metrics with a full reference, use PSNR along with others, because PSNR, based on the standard deviation, never gives overestimated results. Pambrun and Noumeir [11] showed analytically and experimentally that for images with fragments of large or small values of medium brightness, local SSIM estimates are not stable. SSIM does not take account of different absolute luminance levels or viewing distance and do not correlate well with human perceptions of image quality. According to [12], VMAF has a few disadvantages, however, the metric is being improved. It is also necessary to pay attention to such metrics as Visible differences predictor (PDM) [13], Video Information Fidelity Criterion (V-IFC) [14], Video Quality Assessment Using a Statistical Model of Human Visual Speed Perception [15]. The analysis of the different components of the PDM revealed that visual quality metrics that are essentially equivalent at the threshold level can exhibit differences in prediction performance for complex sequences, depending on the implementation choices made for the color space and the pooling algorithm. V-IFC exploits image content but overlooked the impact of distortion types. In [15] the ways to combine the information content, the perceptual uncertainty measures, the computation of local image contrast and the estimation of motion vectors may be improved.

We give here in only a brief overview to elucidate important aspects of previous studies. According to Wang and Li, VQA is a highly challenging problem because the ultimate purpose is to emulate the performance of the biological visual system, which is extremely complicated. Also, in the development of VQA algorithms, it is desirable to maintain a good balance between accuracy and complexity. Therefore, a great deal of effort should be made to simplify the models without significantly losing their accuracy.

III. Research of the Characteristics of the Visual Systems, Subjective Test Setting, Procedure, Assessment Methodology, and Subjective Analysis

Based on the above information and [16], [17], we created a software and test equipment "Research of the characteristics of the visual systems" (RCVS). Our goal is to study new aspects of the dependence of the perceived contrast on the spatial and temporal frequency. The software and test equipment are designed to study the properties of visual perception and measure the characteristics of the visual systems. The interface allows the user to adjust the monitor parameters (brightness, amplitude, and period values for both temporal and spatial image parameters) and select the image type. RCVS has two parts. The first part to study the dependence of the perceived contrast on temporal frequency. The second part to study the dependence of the perceived contrast on the spatial parameters.

The temporal characteristics of vision are described by two main indicators summation time, and critical flicker frequency [16]. The HVS has a certain inertia: after the stimulus is turned on, it takes time for the appearance of a visual reaction. The visual impression does not disappear immediately, but only sometime after the cessation of the effect of light or image on the eye since the retina also takes time to restore the visual pigment. There is an equivalence between the intensity and duration of light exposure to the eye. The shorter the visual stimulus, the more intensity it must have to elicit a visual sensation. Fig. 1 demonstrated the framework of the RCVS.

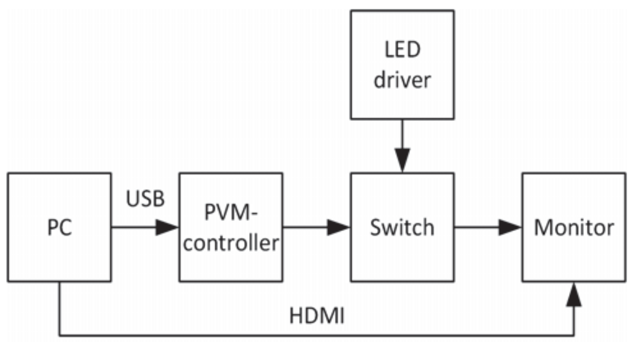

Fig. 1. The framework of the RCVS

We created RCVS so to get the opportunity to study of the time characteristics of the human visual system, based on the frequency dependence of the scintillation amplitude for different values brightness. To expand understanding how the dependence of the perceived contrast on the spatial parameters, we added to RCVS block of spatial frequency versus amplitude for black/white images.

The assessors are seated in a controlled environment. The view distance is strictly kept in 2 m (3.2 times of the picture height), from the center of the monitor to the seat. The total number of assessors is 25 consisting of 9 females and 16 males. Their age is distributed from 18 years to 40 years. Four of the assessors are students in the telecommunications field. Most assessors have no prior experience in the telecommunications area. Observers had a normal or corrected vision and normal color vision. Before the testing session, all assessors received instructions for performing experiments. The trial training session lasting 10 minutes was carried out individuality for each assessor. The duration of a test session ranges 25 min and designed for two people, one of whom conducts the experiment and fixed the values, and the second is the subject. The subjective test is conducted based on the two modules.

In first module, we study the temporal characteristics of the human visual system. The researcher sets the values: brightness, amplitude, period (in this test we use 3 levels of brightness: 64,128,239). Then, researcher selects: Flicker amplitude. The researcher opens the image in full screen mode by clicking the Show Image button. The researcher changes the amplitude with arrows to the right (increase) or left (decrease), while the subject selects the minimum amplitude at which flicker will be noticeable. The researcher records the obtained values of the period and amplitude in the report. In the same way, were selected the minimum amplitude of each level of brightness for the values of the period: 15,17,19,21,25,27,29,30,35,40,50,75,100,250,500.

The second module, the study frequency versus amplitude for b/w images. The researcher sets the values: brightness, amplitude, period. Then, researcher selects: Spatial amplitude. The researcher opens the image in full screen mode by clicking the Show Image button. Researcher changing the right or left arrows to change the amplitude in space until it becomes distinguishable for assessor. Researcher fixed the obtained values of the period and normalized distance, in the same way, were selected the normalized distance for the values of the period: 2, 4, 8, 16, 32, 64, 128, 256.

We collected 25 subjective scores of the minimum amplitudes for each option in two models RCVS. Overall, we gathered $25 \cdot 53 = 1325$ scores. Then we calculated arithmetic mean over all individual values for each option of all modules. Fig. 2 show the arithmetic mean of the contrast sensitivity and the amplitude of the temporal and spatial frequency obtained in the experiments. “Full reference video quality assessment metric on the base human visual system consistent with PSNR” this method is based on the results of adaptation to HVS.

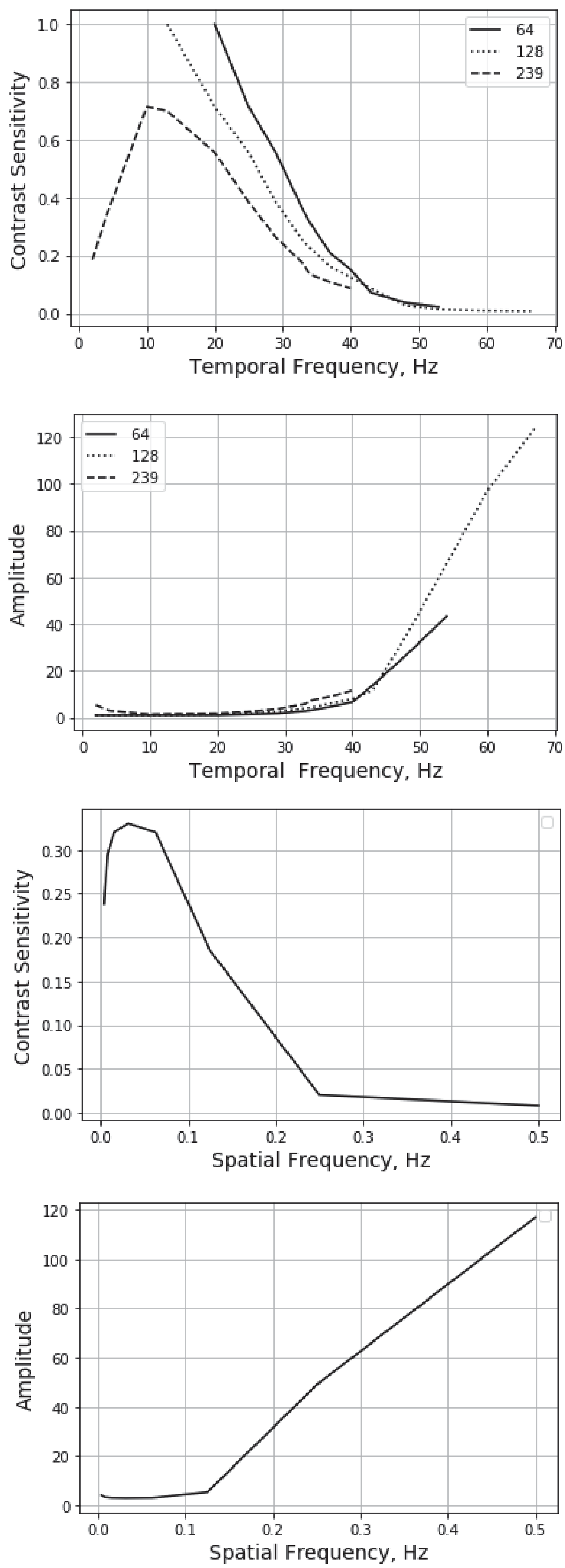

Fig. 2. The arithmetic means of the contrast sensitivity and the amplitude of the temporal frequency obtained in the experiments

IV. FR VQA METRIC ON-BASE HVS CONSISTENT WITH PSNR

In this section, we introduce a video quality method, the results of which can be easily compared with PSNR. However, it considers the knowledge of the peculiarities of human perception of video. The proposed method is (1):

$$\begin{aligned}&\mathrm{PSNR}'(I, I_R, t) = \\ &= \mathrm{PSNR}\big(IK(x, y, t), I_R K(x, y, t), t\big)\end{aligned} \tag{1}$$

where I is estimated frame, I_R is a reference, K - a overall weight coefficient. The quality of the distorted video is measured in the proposed method of two levels: the spatial component and the temporal component:

$$K = K_s K_t \tag{2}$$

Where in (2) K_s and K_t are spatial and temporal weight coefficients respectively. The framework of the proposed method is shown in Fig.3. The framework of the methodology for weight estimate (HVS model) is shown in Fig. 4.

Below considered obtaining the spatial component and the temporal components and an extended formula for PSNR.

A. Spatial component

In block Spatial weight function of HVS model, the impulse response is selected in (3):

$$h(x, y) \xrightarrow{F^{-1}} \big(H\big(f_x, f_y\big)\big) \tag{3}$$

where $f_{x,} f_y$ are spatial frequencies.

According to the study [18] H is isotropic in all directions, therefore:

$$H\big(f_x, f_y\big) = H_0\Big(\sqrt{f_x^2 + f_y^2}\Big) \tag{4}$$

Where in (4) H_0 is the space-time characteristic of the human visual system, obtained in subjective experiments outlined above.

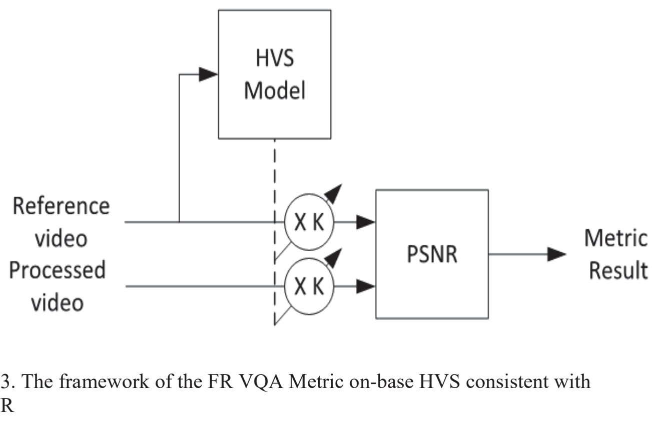

Fig. 3. The framework of the FR VQA Metric on-base HVS consistent with PSNR

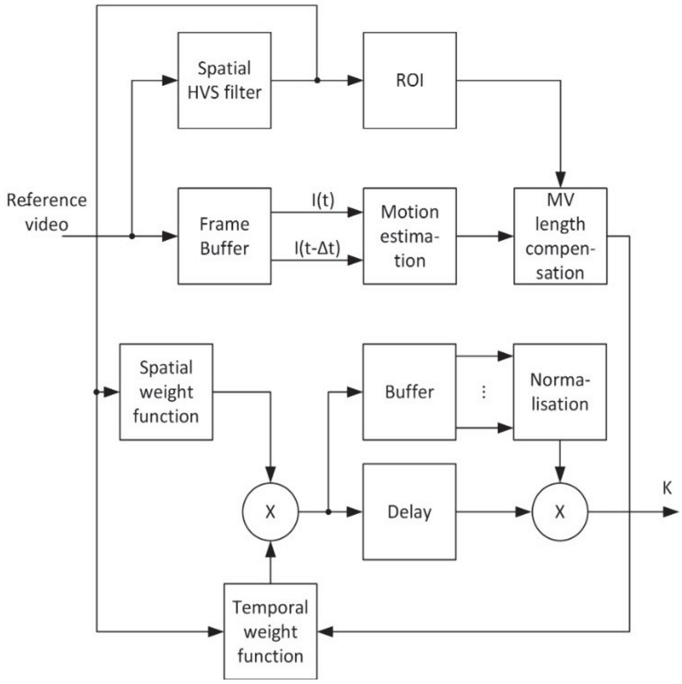

Fig. 4. The framework of the methodology for weight estimate (HVS model)

After that, filtering is done (5):

$$I'_R\big(x, y\big) = I_R\big(x, y\big) * h\big(x, y\big) \tag{5}$$

Then these results have combined into the final formula of weights of the pixels (6):

$$K_s(x, y) = \frac{I'_R\big(x, y\big)}{I_R\big(x, y\big)} \tag{6}$$

B. Temporal component

According to the data presented above, there is a need to consider the temporal aspects when assessing the video quality, following this it is necessary to find the weight coefficients. In block of “Motion estimation” of HVS model, the chosen candidate region becomes the predictor for the current 16 × 16 block. Search area 32 pixels in each direction. We calculate the intermediate vectors by interpolation, this is due to the limited computing power [19].

Block-based motion compensation is relatively straightforward and computationally tractable. In this work we use [20] method.

In block Temporal weight function of HVS model, the compensated motion vectors are found:

$$v_{com}(x, y, t) = v(x, y, t) - v_{ROI} \tag{7}$$

where in (7) v is the motion vector, $v_{ROI}(t)$ is found by the formulas:

$$v_{ROI}(t) = \text{avg}(v(x, y, t)) \tag{8}$$

for those pairs (x, y) in (8) of values that belong to the region of interest (ROI).

The task of ROI is to identify objects that are more significant for the HVS. In this work, the ROI distinguishes individual objects using the watershed method. No more than 5 objects are selected closer to the center, while the number of objects can be less if their total area is more than the initially specified one [20].

The temporary sampling rate based on the motion vectors is found:

$$f(x, y, t) = f_s f_V v_{com}(x, y, t) \tag{9}$$

where in (9) f_s is maximum frequency of contrast sensitivity function, f_V is the number of frames per second.. Finally, in (10) the weight coefficients are determined by considering the HVS characteristics.

$$K_t(x, y, t) = H_t(f(x, y, t), L(x, y, t)) \tag{10}$$

where $L(x, y, t)$ is the brightness of the video sequence, H_t is the time-frequency response of the human visual model shown in Fig. 5.

C. The proposed method (PSNR-M)

The methodology for weight estimate developed by us and presented above is introduced into the PSNR metric by weighing this function and the original video sequences in (11):

$$\text{PSNR}'(I, I_R, t) = \\ = 20\log_{10} \frac{255\sqrt{\sum_{t_n=-\frac{n}{2}}^{t_n=\frac{n}{2}} \sum_{x,y} K_s^2(x, y, t_n) K_t^2(x, y, t_n)}}{\sqrt{\sum_{x,y} (I - I_R)^2 K_s^2(x, y) K_t^2(x, y, t)(n+1)}} \tag{11}$$

where $n = \Delta t \cdot f_V$. It should be noted that an adequate metric estimation is possible for all frames of the sequence, excluding the first passing for Δt and the last frame in the video sequence.

The peculiarity of the expression is in the normalization of the obtained values of the metric. Normalization is carried out on the basis of knowledge that all human attention is directed to the controlled image during the period of adaptation of the HVS. Thus, attention can be redistributed within the specified period to the space-time domains of the monitored video sequence. As a result, the weighted coefficients on average on the adaptation interval and over the entire field of visibility are normalized to 1.

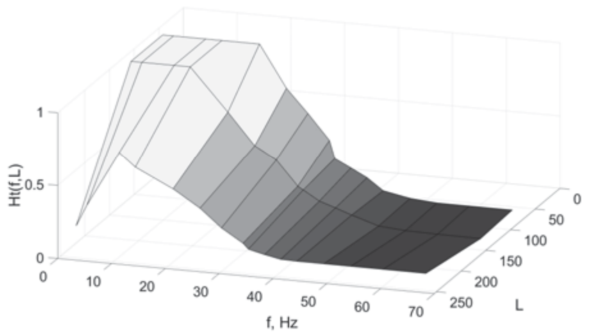

Fig. 5. The graph of time-frequency response of the HVS

To design the source code, we relied a Matlab software. To facilitate video quality research, we make our metric code available upon request to A.Mozhaeva, I.Vlasuyk.

V. RESULTS AND DISCUSSION

The LIVE-NFLX-II video quality [21] and LIVE-NFLX video quality database [22, 23] is used to evaluate the performance of the proposed VQA method. The LIVE-NFLX video quality database provides 14 source video clips, 122 distorted video clips, and the LIVE-NFLX-II database provides 15 source video clips, 420 distorted video clips, and their associated continuous-time scores capture the instantaneous. Quality of Experience, and retrospective scores which reflect the overall viewing experience. The original videos were used in the investigation, that was selected from databases. These databases are selected because they represented highly realistic contains with Quality of Experience responses to various design dimensions. Two measures are employed as performance indexes to evaluate the performance of the proposed VQA method: PSNR, SQ. In Fig. 6, we report the comparison among SQ, PSNR, and the proposed method for three videos.

The one video of LIVE-NFLX, AsianFusion and AirShow videos of LIVE-NFLX-II. When comparing the dependence of the studied metric on the frame number for the three test sequences used, the following can be noted. In the AirShow video, you can see significant fluctuations in PSNR metric estimates associated with a significant deterioration in image quality when the background/ROI moves rapidly. The estimates of the developed metric are more stable due to the redistribution of the estimate weights to the object of interest. However, in some cases this can lead to excessive sensitivity, for example, to an extremely low estimate in the case of artifacts in the region of interest (for example, distortion of glare on an airplane wing, AirShow about 400 frames), which was obtained as a result of normalizing the values of the developed metric to PSNR values. For a similar reason, you can see an extremely high score on the AsianFusion test sequence of about 600 frames (the video shows the transition through a white field and the indicated frames are identical to the reference sequence). In this case, were obtained the maximum values of metric. However, based on the theory HVS [24] the experts evaluating the video (SQ) simply did not respond to quality improvements.

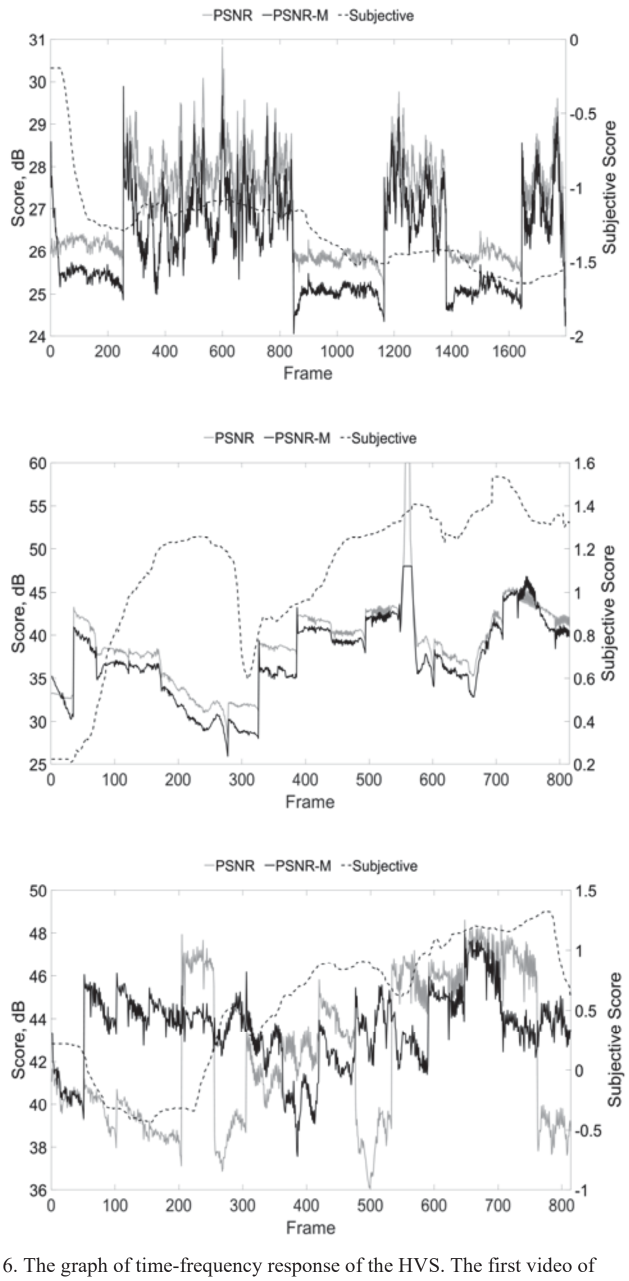

Fig. 6. The graph of time-frequency response of the HVS. The first video of LIVE-NFLX, second "AsianFusion" and third "AirShow" videos of LIVE-NFLX-II. Achieved method (PSNR-M), PSNR, SQ.

It follows from the above that, despite the objectively better correlation of the developed method and subjective assessments, there are problems associated with normalizing the values of the developed metric to PSNR. As a result of which outliers of its values can be observed during extreme short-term changes in the quality of the content.

Our method achieves the best overall correlation values for video on LIVE-NFLX and LIVE-NFLX-II, Table I. However, the correlation values of our method of one video, AirShow of LIVE-NFLX-II in Fig.6, is less than PSNR. This may be explained by the fact that this video consists the high temporal aspects, and in the block of "Motion estimation" of the HVS model need to change motion compensation method.

TABLE I. THE VALUES OF THE CORRELATION BETWEEN SQ AND PSNR/PSNR-M

FR METRIC	AIRSHOW	ASIANFUSION	LIVE-NFLX
PSNR	0.437	0.309	0.115
PSNR-M	0.017	0.326	0.134

It should be noted that the presented correlation coefficients are only estimates, since the test base used is too small to obtain reliable statistical data. For this reason, we do not provide an estimate of the uncertainty in these values. Obtaining quality results and related statistics on various text video sequences is one of the external sources of work. Based on this data, some parameters of the presented method can also be refined, for example, the size of objects attributable to ROI, the time interval for normalizing the obtained metric values for comparison with PSNR and the window used, the search area for motion vectors.

VI. CONCLUSION

In this work, we propose experiments of visual perception and measure the characteristics of the human visual systems on the "Research of the characteristics of the visual systems MTUCI" software and test equipment we created. Also, we propose a novel FR-VQA method which can be easily compared with PSNR and takes into account the features of the Human Visual System. Our proposed method is compared with two state-of-the-art methods on two publicly available VQA databases and achieves the better values then the PSNR without adaptation to HVS. This indicates that there is ample room for developing an objective model which correlates well with human perception. Because our metric is based on PSNR, then there is exist criticality in time and shifts of pixel values in space, brightness, and color. In a further study, we will consider embedding the automatic corrections. However, then the results are not compatible with PSNR, which was the purpose of our work. For objective measures to be maximally useful for quality assessment, they must be NR quality assessment. In future work, for NR quality assessment, we need more comprehensive tests with a focus on the dependence of the spatial and temporal frequency, the logarithmic illumination of the retina, and on the regions of interest for users in streaming video sequences.

REFERENCES

[1] P. Mohammadu, A. Ebrahimi-Moghadam, S. Shirani, "Subjective and Objective Quality Assessment of Image: A Survey", *Majlesi Journal of Electrical Engineering*, vol.9(1), Mar 2015, pp.55-83.

[2] Z. Ying, H. Niu, P.Gupta1, D. Mahajan, D. Ghadiyaram, A. Bovik, "From Patches to Pictures (PaQ-2-PiQ): Mapping the Perceptual Space of Picture Quality", *University of Texas at Austin, Facebook AI*, Dec 2019.

[3] N. Ye, "Visibility Metric for Visually Lossless Image Compression", Ph.D. dissertation, Department of Computer Science and Technology, University of Cambridge, Cambridge, United Kingdom, 2020.

[4] D. Chandler, T. Phan, M. Alam, "Seven challenges for image quality research", *SPIE(Newsroom)*, Examining the limitations of current research into image quality assessment opens doors for further studies., January 15.

[5] K. Wolski, D.Giunchi, N.Ye, P. Didyk, K. Myszkowski, R. Mantiuk, H. Seidel, A. Steed, R. K. Mantiuk, "Dataset and metrics for

predicting local visible diferences", *ACM Transactions on Graphics*, vol.36(4), July 2017, Article 1.

[6] A.V. Luizov, *Inertia of vision*., Moscow: Oborongiz, 1961.

[7] H. De Lange, "Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. I. Attenuation characteristics with white and colored light," *Journal of the Optical Society of America*, vol. 48, 1958, pp. 777-784.

[8] Z. Wang, A. Bovik, H.R. Sheikh, E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity.", *IEEE Transactions on Image Processing*, vol.3, April 2004, pp.600 - 612.

[9] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara. (Jun 2016) Toward a practical perceptual video quality metric. [Online]. Available: https://medium:com/netflix-techblog/ toward-a-practical-perceptual-video-quality-metric-653f208b9652.

[10] Z. Wang, A. Bovik, "Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures", *IEEE Signal Processing Magazine*, vol.26(1), Jan. 2009, pp.98-117.

[11] J.F. Pambrun, R. Noumeir. "Limitations of the SSIM quality metric in the context of diagnostic imaging", *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 2960–2963

[12] A. Zvezdakova, D. Kulikov, S. Zvezdakov, D. Vatolin, "BSQ-rate: a New Approach for Video-codec Performance Comparison and Drawbacks of Current Solutions", *Programming and computer software*, vol.46, May 2020, pp.183-194.

[13] S.Winkler, *Digital Video Quality: Vision Models and Metrics*, NJ: Wiley, Mar. 2005.

[14] Seshadrinathan, A.C. Bovik, "An information-theoretic video quality metric based on motion models", *Proc. Int. Workshop Video Processing & Quality Metrics for Consumer Electronic*s, Scottsdale, AZ, Jan. 2007.

[15] Z. Wang, Q. Li, "Video quality assessment using a statistical model of human visual speed perception," *J. Opt. Soc. Amer. A*, vol.24, Dec. 2007, pp. B61–B69.

[16] A. Potashnikov, I. Vlasuyk, S. Romanov, "Method of measuring temporary and spatial characteristics of the human visual systems", *X International industrial scientific and technical conference: collection of works. 2016,* pp. 102-103

[17] I. Vlasuyk, "Development of a model of the human visual system for the method of objective image quality control in digital television systems, *Telecommunications and Transportation*, S1. C., 2009, pp. 189-192

[18] A. Potashnikov, I. Vlasuyk,"Method of formation of pseudo random arrays of light filters in light-signal converters", *XII International Branch Scientific and Technical Conference 2018*, pp. 149-151

[19] A. Potashnikov, I. Vlasuyk, I. Augstkaln, "Analysis of methods for detecting moving objects of different types on video image", *Fundamental problems of radio electronic instrumentation,* vol.17, 2017, pp. 1201-1204.

[20] I. Vlasuyk, "Enhanced area coding in the JPEG 2000 image compression standard", *Telecommunications and Transportation*, S1. C., 2011.

[21] C.Bampis, Z. Li, I. Katsavounidis, T. Huang, C. Ekanadham, A. C. Bovik, "Towards Perceptually Optimized End-to-end Adaptive Video Streaming ", *arXiv preprint,* Aug 2018.

[22] C. Bampis, Z. Li, A. Moorthy, I. Katsavounidis, A. Aaron, A. Bovik, "Study of Temporal Effects on Subjective Video Quality of Experience", *IEEE Trans. Image Process.,* vol. 26, no. 11, 2017, pp. 5217–5231.

[23] C. Bampis, Z. Li, A. Moorthy, I. Katsavounidis, A. Aaron, A. Bovik, "LIVE Netflix Video Quality of Experience Database", *Online: http://live.ece.utexas.edu/research/LIVE_NFLXStudy/index.html, 2016.*

[24] D.H. Hubel, *Eye, brain, vision*, Moscow: Mir, 1990.