An Approach to Interactive Multimedia Systems Through Subjective Video Quality Assessment in H.264/AVC Standard

ZORAN MILIČEVIĆ[1], ZORAN BOJKOVIĆ[2], KAMISETTY R. RAO[3]
[1]Department of Telecommunications and IT GS of SAF
11000 Belgrade
REPUBLIC OF SERBIA
[2]University of Belgrade
11000 Belgrade
REPUBLIC OF SERBIA
[3]The University of Texas at Arlington, TX
UNITED STATES OF AMERICA
mmilicko@eunet.rs, z.bojkovic@yahoo.com, rao@uta.edu

Abstract: - This paper seeks to provide an approach to subjective video quality assessment in H.264/AVC standard. In general, methods for video content quality evaluation can be divided on objective and subjective one. While objective methodology uses mathematical methods for presenting human visual system (HVS), subjective method uses group of subjects (observers) to present decoded video sequences. Video content quality estimation enabled to evaluate in a subjective way the relevant parameter i.e., signal-to-noise ratio. For this purpose, special software program is developed for subjective quality estimation of all tested sequences using the recommendation ITU-T P.910 which is convenient for testing different multimedia applications. Testing for subjective quality evaluation is realized in two steps. The user's interface preparation is carried out in the first step, while in the second one, the presentation of the processed test sequences through user interface is realized. Using the subjective analysis in processed test sequences, it is confirmed that there exists a small difference in signal-to-noise ration, i.e., video quality between sequence in reference source format and sequence under the test condition. To prove this, the observer's ratings are mainly 4 or 5 and mean estimation values varying in the range from 4,00 to 4,95.

Key-Words: - H.264/AVC, Peak signal-to-noise ratio, Selective intra-prediction method, Optimized inter-prediction method, Subjective video quality assessment.

1 Introduction

Today video streaming over networks has become a popular application. For example, a wide range of interactive multimedia services such as video conferencing, interactive internet protocol television (IPTV), e-learning, surveillance and other real-time multimedia content distribution rely on multimedia streaming techniques. An important issue for success of such applications is to deal with dynamic end- user characteristics including decoding and display capabilities and, of course, network bandwidth limitations.

Subjective quality assessment is a crucial step to develop such a strategy and reach the goal of optimal content delivery, because it is the most accurate and reliable way to measure perceived multimedia quality and perform quantitative quality evaluation and comparison. When one conducts a subjective test, the methodology of the test needs to be determined according to the objectives of the test, the quality levels or artefacts in the stimuli, the number of stimuli, etc.

In order to specify, evaluate and compare video communication systems, it is necessary to determine the quality of the video images displayed to the viewer. Measuring visual quality is a difficult and often imprecise task because there are so many factors that can effect the results. Visual quality is inherently *subjective* and is influenced by many factors that make it difficult to obtain a completely accurate measure of quality. For example, a viewer's opinion of visual quality can depend on the task at hand, on actively participating in a videoconference, communicating using sign language, trying to identifying a person in a surveillance video scene. There are many factors influancing subjective quality. Our perception of a visual scene is formed by a complex interaction between the components of the Human Visual System (HVS), the eye and the brain. However, a viewer's opinion of 'quality' is also affected by other factors such as the viewing environment, the observer's state of mind as well as the extent to which the observer interacts with the visual scene. A user carrying out a specific task that requires concentration on part of a visual scene, will have a quite different requirement for „good" quality than a user who is passively watching a movie. Other important influences on perceived quality include visual attention. It means that an observer perceives a scene by fixating on a sequence of points in the image rather than by taking in everything simultaneously. Also, the so-called „recency effect" must be mentioned because our opinion of a visual scene is more heavily influenced by recently-viewed material than older video material [1, 2]. All of these factors make it very difficult to measure visual quality accurately and quantitatively [3].

Subjective tests for determining quality are very time consuming.

In order to provide better compression of video, compared to previous standards, H.264 MPEG-4 part 10 [4] video coding standard was developed by the Joint Video Team (JVT) [3,5,6,7,8]. H.264 fulfills significant coding efficiency, simple syntax specifications, and seamless integration of video coding into all current protocols and multiplex architectures. Thus, H.264 can support various applications (video broadcasting, video streaming, video conferencing over fixed and wireless networks as well as over different transport protocols), where subjective video quality becomes hot topic [9, 10].

Perceptual and subjective testing procedure should constitute an integral part of coder design and evaluation. Even when perceptually objective messages are known, some amount of subjective testing is needed.

The coding standard H.264/AVC (Advanced Video Coding) provides the improvement in coding performance which comes mainly from the prediction part [11]. Prediction must be always performed before texture coding for both inter and intra macroblocks [12].

The paper is organized as follows. After a brief description of related work, we describe the background of the subjective video quality assessment. Next, we emphasize the proposed framework concerning general conditions in the process of subjective video quality evaluation. Experimental results and discussion are presented in the second part of the paper, followed by the conclusion.

2 Related Work

The H.264/AVC standard achieves much higher coding efficiency than the H.263, MPEG-2 and MPEG-4 standard, due to its improved inter- and intra-prediction modes at the expense of higher computation complexity [13]. In order to reduce the computational complexity of the H.264/AVC encoder without significant rate-distortion performance degradation, a combined selective mode decision for fast intra-prediction mode selection and optimized inter-prediction method can be proposed [14].

In order to reduce the temporal and spatial redundancy more effectively, motion compensation uses variable block sizes, while directional intra-prediction investigates all available coding modes to decide the best one [15]. Rate–distortion optimization (RDO) is a very efficient tool for deciding a macroblocks coding mode, which has been adopted by H.264/AVC codec and brings higher coding efficiency.

However, due to precise calculation of current macroblocks distortion and coding-bits for each coding modes, RDO-based mode decision involves high computational complexity [16, 17].

The R–D optimization for video encoding using the Lagrange multiplier technique is addressed in [18]. Although different video standards support various coding options, the Lagrange multiplier optimization technique provides a systematic way to select the optimal coding mode [19]. In order to achieve good rate–distortion performance, the Rate–Distortion Optimized (RDO) mode selection process [4] evaluates the distortion and rate of each candidate mode prior to selecting the mode for the current MB. In the Joint Model (JM) reference encoder [20], this is carried out by coding the macroblock in each of the possible modes and choosing the mode that minimizes a rate–distortion cost function.

The intra-prediction reduces spatial redundancies by exploiting the spatial correlation between adjacent blocks in a given picture.

In proposed combine algorithm, selective intra-prediction mode decision methods for fast intra-mode decision stems from the fact that the dominating direction of a bigger block is similar to that of smaller block [21]. The best prediction mode of 4 × 4 luma block within 16 × 16 block has the same direction as that of 16 × 16 luma block [22]. The computation of the intra-prediction and the chroma prediction can be reduced on the base of the overall edge information from the 16 × 16 intra-prediction result.

H.264/AVC encodes the macroblock by iterating all the luma intra-decisions for each possible chroma intra-prediction mode for the best coding efficiency. Number of mode combinations for luma and chroma components in an macroblock is C8 × (L4 × 16 + L16), where C8, L4, and L16 represent the number of modes for chroma prediction, 4 × 4 luma prediction and 16×16 prediction, respectively. It means that, for an macroblock, it has to perform 4 × (9 × 16 + 4) = 592 different RDO calculations before a best RDO mode is determined.

The encoder with the selective intra-mode decision algorithm need to perform only 1× (4 × 16 + 4) = 68 for the best case. Thus selective intra-prediction algorithm has reduced the number of RDO modes calculation compared to the 592 modes that are used in the current RDO calculation in H.264/AVC video coding.

The second part of the proposed scheme for an H.264/AVC encoder is the optimized inter-prediction method for complexity reduction.

Inter-prediction creates a prediction model from one or more previously encoded video frames or fields using block-based motion compensation. Important differences from earlier standards include the support for a range of block sizes (from 16 × 16 down to 4 × 4) and fine subsample motion vectors (quarter-sample resolution in the luma component).

In order to optimize inter-prediction procedure we considered the impact of different partition size on compression performance. We proposed method to optimize large number of possible combinations (modes) for macroblock partitions and sub-macroblock within each macroblock. Also, we decided to avoid a small portion sizes (from 8 x 8 to 4 x 4), because they already contain large partition sizes. In that way, numbers of possible combinations (modes) for inter-prediction were decreasing.

Only next possible combinations (modes) for intra-I and inter-P slices prediction were considered: intra-4 x 4, intra-16 x 16, P slice skip, P slice 16 x 16, P slice 16 x 8, P slice 8 x 16. On the other hand, possible combinations for B slices are: B slice Direct, B slice 16 x 16, B slice 16 x 8, B slice 8 x 16, Bi-prediction 16 x 16, Bi-prediction 16 x 8 and Bi-prediction 8 x 16.

To evaluate the performance of the H.264/AVC codec, the proposed method was implemented into H.264/AVC reference software JM 16.0 and tested with various Quantization Parameters (QP=24, QP=28 and QP=32).

Different test sequences in Standard Definition and High Definitions are processed, when P and B pictures are analyzed. The selected sequences were of SD (720×576 pixels) and two HD resolutions (1280×720 and 1920×1088 pixels). The test sequences have been selected to emphasize different kind of motions and contents. For the experiments, the first 100 frames of the 4 different test sequences (Pedestrian, Blue sky, River bed and Rush hour) are used.

Comparisons with the case of exhaustive search were performed with respect to the change of average peak signal to noise ration-PSNR (ΔPSNR), the change of average data bits (ΔBit rate), and the change of average encoding time (ΔTime), respectively.

The main contributions were [14]:

- Selective intra prediction algorithm has reduced the number of RDO candidate modes for about 90% while inter prediction method gives the reduction of 40%.
- When proposed method is used and all analyzed test sequences are in SD resolution, encoding time saving is from about 15% to 17% depending on QP values. However, there is negligible loss in term Δ PSNR and bit rate values slightly decreased.
- When all test sequences are analyzed in two different HD resolutions, experimental results show that proposed method gives encoding time saving over 15% in average for all QP values. On the other hand, there is negligible loss in term Δ PSNR.

3 Background

Subjective video quality assessment (SVQA) represents methodologies for video quality evaluation which is watched by observer and presents observer opinion about video. Because the experiments with subjective evaluation are complicated with more psychological aspects and watching conditions, different methodologies for subjective quality evaluation are defined. They require sufficiently time and efforts for its application, realization as well as observers providing. Also, they are significantly more expensive comparing to the objective methods, but they offer more available and precise results, because there are addressed to the video information user.

Generally, the process of a subjective video quality evaluation includes the following steps:

- Codec's choice for testing and defining their parameters such as: pre-filtering, buffering, distance between key-frames;
- Test sequences choice (non-compressed);

- Coding conditions defining, like bit rate, resolution;
- Test sequences compression;
- Choice of the methodology for evaluation and establishing repetition of test conditions;
- The process of testing organization, subjects (observers) invitation, representation of the test sequences to subjects and finally, subjects rating;
- Collecting observers results, statistical analysis performance of results and non-coexistent subjects removal;
- Notification of test results.

For subjective video quality assessment, we used recommendation ITU-T P.910 (04/2008) which is intended to define non-interactive subjective assessment methods for evaluating the quality of digital video images coded at bit rates specified in classes for TV3, MM4, MM5 and MM6, for applications such as videotelephony, videoconferencing as well as storage and retrieval applications. The methods can be used for several different purposes including, but not limited to, selection of algorithms, ranking of video system performance and evaluation of the quality level during a video connection [23].

Measurement of the perceived quality of images requires the use of subjective scaling methods. The condition for such measurements to be meaningful is that there exists a relation between the physical characteristics of the "stimulus" (video sequence presented to the subjects in a test) and the magnitude as well as nature of the sensation caused by the stimulus.

A number of experimental methods have been validated for different purposes. The final choice for the methods for a particular application depends on several factors, such as the context, the purpose and where in the development process the test is to be performed.

An important issue in choosing a test method is the fundamental difference between methods that use explicit references like Degradation Category Rating (DCR) and methods that do not use any explicit reference, e.g., Absolute Category Rating (ACR), Absolute Category Rating with Hidden Reference (ACR-HR), and Pair Comparison Method (PC).

The DCR method should be used when testing the fidelity of transmission with respect to the source signal. This is frequently an important factor in the evaluation of high quality systems. For a long time, DCR has been a key method for the assessment of television pictures whose typical quality represents the extreme high levels of videotelephony and videoconferencing. The specific comments of the DCR scale (imperceptible/perceptible) are valuable when the viewer's detection of impairment is an important factor. Thus, when it is important to check the fidelity with respect to the source signal, the DCR method should be used.

The degradation category rating (DCR) implies that the test sequences are presented in pairs: the first stimulus presented in each pair is always the source reference, while the second stimulus is the same source presented through one of the systems under test. This method is also called the double stimulus impairment scale method. The time pattern for the stimulus presentation can be illustrated by Fig.1.

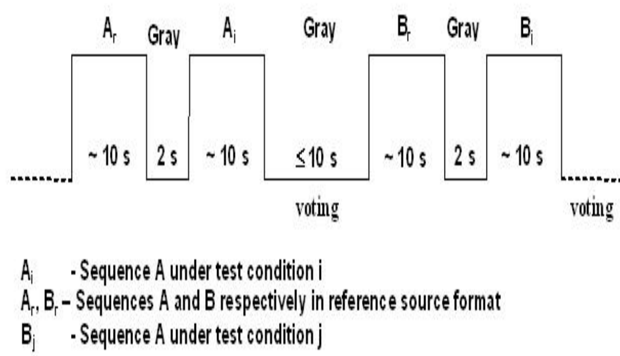


Fig. 1 Stimulus presentation in the DCR method.

If a constant voting time is used, then it should be less than or equal to 10 s. The presentation time may be reduced or increased according to the content of the test material.

The subjects are asked to rate the impairment of the second stimulus in relation to the reference one. The following five-level scale for rating the impairment should be used: 5 - Imperceptible, 4 - Perceptible but not annoying, 3 - Slightly annoying, 2 – Annoying, 1 - Very annoying. The necessary number of replications could be obtained for the DCR method by repeating the same test conditions at different points of time in the test.

4 Proposed Framework

Based on the related work, we focused on fact that proposed algorithm in different test scenarios gave negligible loss in term Δ PSNR. For example, when all analyzed test sequences were in SD resolution, there was negligible loss in term ΔPSNR for luma component of picture: it was only 0.49, 0.47 and 0.48 dB in average depending on QP values. On the other hand, when all analyzed test sequences were in HD resolution, there was negligible loss in term ΔPSNR for luma component of picture: it was only 0.40, 0.40 and 0.44 dB in average depending on QP values. Finally, when all analyzed test sequences were in full HD resolution, there was negligible loss in term ΔPSNR for luma component of picture: it was only 0.29, 0.32 and 0.36 dB in average depending on QP values [14].

Also, the PSNR values of luma (Y) component of pictures were used which were based on the equations below:

$$\overline{PSNR} = 10\log_{10}\left(\frac{255^{2}}{\overline{MSE}}\right) \quad (1)$$

where MSE is the mean square error. The average Δ PSNR was the PSNR difference expressed as a percentage between JM 16.0 encoder and the proposed method, respectively.

We would like to emphasize that we measured PSNR only for Y, because human visual system is more sensitive to luma then to chroma components of pictures.

In order to evaluate the objective PSNR results in [14] on the subjective way, we developed a software program. It was developed in accordance with recommendation ITU-T P.910, since it is suitable for the testing of multimedia applications [23].

In the process of subjective video quality assessment, the following general conditions are taken into account:

- The DCR method was used;
- The SD (Standard Definitions), HD (High Definitions) and full HD test sequences were used in YUV 4:2:0 video format;
- The sequence in reference source format and sequence under the test condition are shown simultaneously on the projection board;
- The sequences are displayed in the two window display put side by side within a 50% grey background;
- The sequences were perfectly synchronized which means that they both were started and stopped at the same frame.
- Each video sequence was repeated twice;
- The sequence in reference source format was placed always on the left side, while the sequence under the test condition on the right;
- 20 observers belonging to the population from 24 to 27 age participated in the experiment;
- Evaluation of image quality was not part of the work of observers, so they were not experienced estimators;
- Observers had normal visual acuity and normal colour recognition;
- In order to reduce the eye movement to switch the attention between the two windows, the viewing distance was 8H, where H indicated the picture height.
- Observers were used to assess the scale with five levels;
- Before starting the experiment scenario, applications that were found in the system under test are presented to observers;
- The descriptions of the evaluation type, as well as the assessment scale, together with simulation presentation are given in the written form.

5 Experimental Results and Discussion

User interface for the subjective assessment of quality of test sequences has been developed in accordance with the rules defined for the DCR test method (Fig. 2).

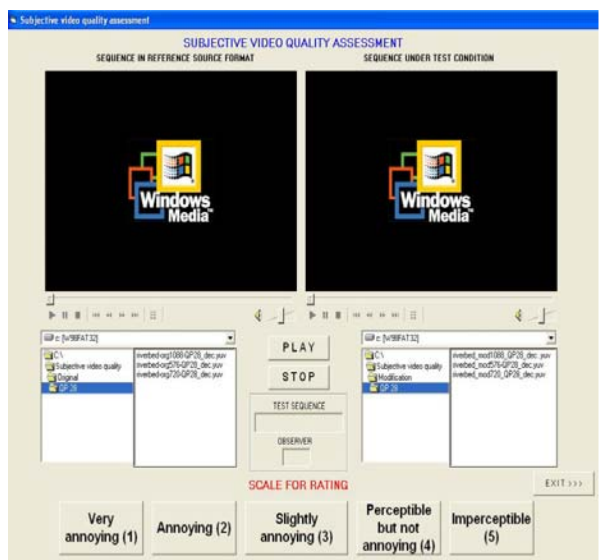


Fig. 2 User interface for subjective video quality assessment.

Before the testing procedure was carrying out, firstly it was necessary to fulfil all conditions relating to the observers, and then they have accessed the personal evaluation. The testing included two phases. The first phase entailed the entry of sequence name and serial number of the observers and then entering (loading) the sequence in reference source format as well as sequence under the test condition in an appropriate format shown in Fig. 3 for Riverbed test sequence as an example.

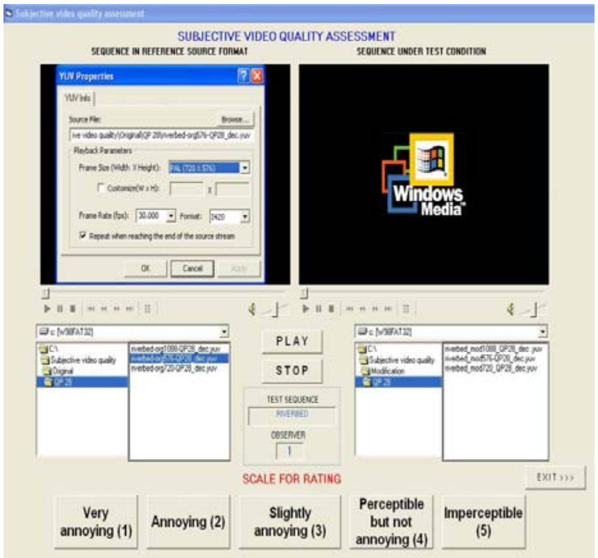


Fig. 3 An example: The first phase for Riverbed test sequence in an appropriate format.

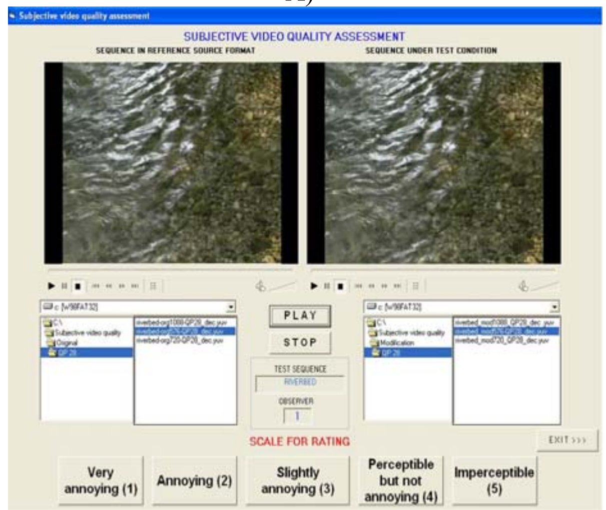
After that, the second phase - processing and rating in subjective video quality assessment for Riverbed test sequence in Standard, High and full High resolution for QP=28 are shown in Fig 4. All ratings given by observers were stored in a *Microsoft Access* database.

We choose the Riverbed test sequence as an example because this test sequence is much more complex in amount of the motion (motion vectors) in comparing to other processed video test sequences.

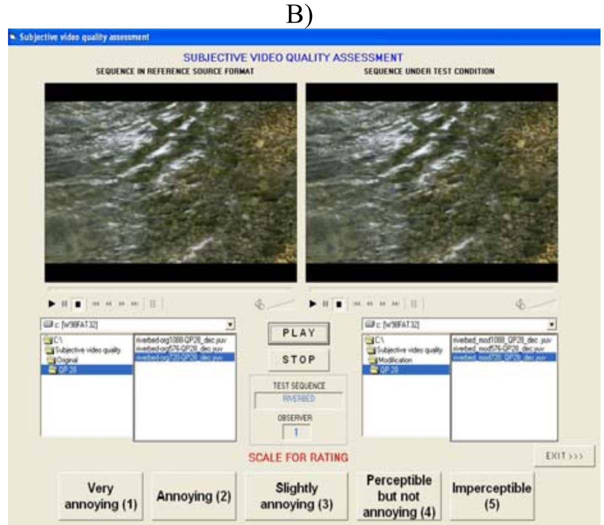
In the proposed algorithm for selective intra- and optimized inter-prediction, the obtained results for the subjective video quality evaluation between sequence in reference source format and sequence under the test condition, show the there exists small difference in the signal-to-noise ratio (SNR), i.e. the image quality. To prove this, there are mean estimation values given by the viewers, varying in the range from 4,00 to 4,95 for each video sequence in the different video resolution and with a different quantization parameter (QP).

Tables 1, 2, 3 contain overviews of the unique rating given by the observers for each of test video sequence (Blue sky, Pedestrian area, Riverbed and Rush-hour) in the different video resolutions (SD, HD and *full* HD) and for different QP values (QP=24, QP=28 and QP=32). As it can be seen, the observer's ratings are mainly 4 or 5.

A)



B)



C)

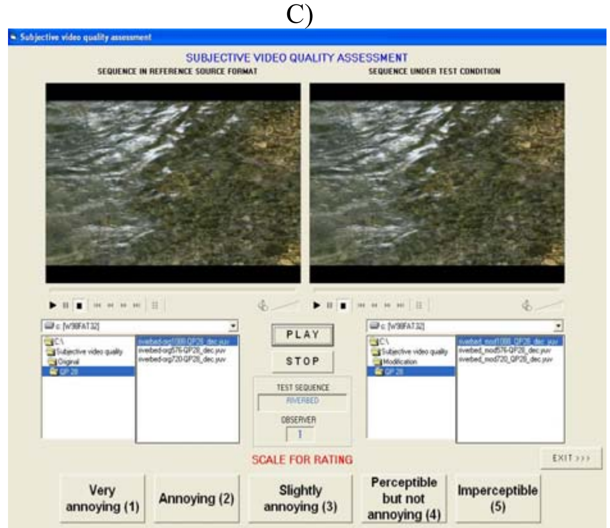


Fig. 4 An example: the second phase-processing and rating for Riverbed test sequence in different resolutions - A) Standard, B) High and C) full High resolution.

Table 1 Overview of the unique rating given by the observers for each of test video sequence in the different video resolutions (SD, HD and full HD) and QP=24.

Observers	QP=24					
	Blue sky (SD)	Blue sky (HD)	Blue sky (full HD)	Pedestrian area (SD)	Pedestrian area (HD)	Pedestrian area (full HD)
1	4	3	4	4	5	5
2	5	4	5	4	4	4
3	5	4	5	5	5	5
4	4	5	5	5	4	5
5	5	4	5	4	4	3
6	4	5	5	5	5	4
7	4	5	5	4	5	4
8	5	5	5	5	5	5
9	5	5	5	4	5	4
10	4	5	4	5	4	5
11	4	5	5	5	5	5
12	4	4	5	4	5	5
13	4	5	4	5	4	4
14	5	4	4	5	4	5
15	5	4	5	5	5	5
16	5	5	5	5	4	3
17	3	4	5	2	3	4
18	5	4	4	5	5	4
19	5	5	5	5	5	5
20	4	4	4	5	5	5
Average	4,45	4,45	4,70	4,55	4,55	4,45

Observers	QP=24					
	Riverbed (SD)	Riverbed (HD)	Riverbed (full HD)	Rush-hour (SD)	Rush-hour (HD)	Rush-hour (full HD)
1	3	3	3	5	4	5
2	5	4	3	5	4	4
3	4	4	5	4	4	5
4	4	4	5	3	4	4
5	5	4	3	5	4	5
6	4	3	4	5	5	4
7	4	5	4	5	4	4
8	5	5	5	5	5	5
9	4	4	5	5	5	4
10	3	3	3	5	5	5
11	4	4	4	5	5	5
12	5	4	4	4	4	5
13	5	5	5	5	5	5
14	4	4	5	5	5	5
15	4	4	4	5	5	5
16	5	4	4	5	5	5
17	4	4	4	4	5	5
18	3	3	4	5	5	5
19	5	5	5	5	5	5
20	5	5	5	5	5	5
Average	4,25	4,05	4,20	4,75	4,65	4,75

Table 2 Overview of the unique rating given by the observers for each of test video sequence in the different video resolutions (SD, HD and full HD) and QP=28.

Observers	QP=28					
	Blue sky (SD)	Blue sky (HD)	Blue sky (full HD)	Pedestrian area (SD)	Pedestrian area (HD)	Pedestrian area (full HD)
1	4	5	4	5	3	4
2	3	3	4	5	5	5
3	4	4	5	5	5	4
4	3	4	5	4	3	5
5	4	5	5	5	4	3
6	3	3	4	5	5	5
7	5	5	4	5	4	4
8	5	5	5	5	5	5
9	5	5	4	4	5	5
10	3	4	4	4	4	4
11	5	5	5	5	5	5
12	3	4	5	4	4	5
13	4	5	4	5	5	5
14	4	5	4	5	5	4
15	5	5	4	5	5	4
16	5	5	4	5	4	3
17	4	4	5	4	5	5
18	2	3	4	5	5	5
19	5	5	5	5	5	5
20	5	5	5	5	5	5
Average	4,05	4,45	4,45	4,75	4,55	4,50

Observers	QP=28					
	Riverbed (SD)	Riverbed (HD)	Riverbed (full HD)	Rush-hour (SD)	Rush-hour (HD)	Rush-hour (full HD)
1	5	4	4	5	5	5
2	4	3	4	5	5	5
3	4	4	4	5	4	4
4	4	5	5	5	4	5
5	5	4	5	5	5	5
6	4	4	4	5	5	5
7	5	5	4	5	5	4
8	5	5	5	5	5	5
9	5	5	5	5	5	4
10	4	4	3	5	4	4
11	4	4	4	5	5	4
12	4	3	5	5	5	5
13	5	5	5	5	5	5
14	3	3	4	5	5	3
15	4	3	4	5	5	4
16	5	5	3	5	5	5
17	5	4	5	4	4	5
18	5	5	3	5	5	5
19	5	5	5	5	5	5
20	5	5	5	5	5	5
Average	4,50	4,25	4,30	4,95	4,80	4,60

Table 3 Overview of the unique rating given by the observers for each of test video sequence in the different video resolutions (SD, HD and full HD) and QP=32.

Observers	QP=32					
	Blue sky (SD)	Blue sky (HD)	Blue sky (full HD)	Pedestrian area (SD)	Pedestrian area (HD)	Pedestrian area (full HD)
1	4	5	4	5	4	3
2	3	4	3	5	3	3
3	4	4	5	4	4	4
4	4	4	4	3	3	5
5	3	4	4	5	4	3
6	4	4	5	5	4	5
7	5	5	5	5	5	4
8	5	5	5	5	5	5
9	5	5	4	5	5	4
10	3	4	4	5	4	5
11	3	4	5	5	4	3
12	2	3	4	4	4	4
13	4	5	4	5	4	4
14	4	5	5	4	5	5
15	4	4	5	5	5	4
16	5	5	4	5	4	2
17	4	4	5	5	5	5
18	4	3	4	5	4	4
19	5	5	5	5	5	5
20	5	4	4	5	5	5
Average	4,00	4,30	4,40	4,75	4,30	4,10

Observers	QP=32					
	Riverbed (SD)	Riverbed (HD)	Riverbed (full HD)	Rush-hour (SD)	Rush-hour (HD)	Rush-hour (full HD)
1	4	4	4	5	5	5
2	5	4	3	5	4	5
3	4	4	5	4	4	5
4	5	5	3	5	5	5
5	5	4	3	5	4	5
6	5	5	3	5	5	5
7	5	5	4	5	5	4
8	5	5	5	5	5	5
9	5	5	5	4	5	5
10	3	4	4	4	5	5
11	3	4	4	5	5	5
12	3	4	4	4	4	5
13	5	5	5	5	5	5
14	3	3	4	5	5	4
15	4	4	4	5	5	5
16	4	4	3	5	5	4
17	5	5	5	4	5	5
18	5	5	5	4	5	5
19	5	5	5	5	5	5
20	5	5	5	5	5	5
Average	4,40	4,45	4,15	4,70	4,80	4,85

Some the observers ratings are less than 4, because they focused on some detail in video sequence and place behind global view on video sequence structure. Also, for each of the video test sequence in different resolutions (SD, HD and *full* HD) an overview of the ratings mean values is presented. The obtained results of the subjective quality evaluation show that there exists a small difference between sequence in reference source format and sequence under the test condition.

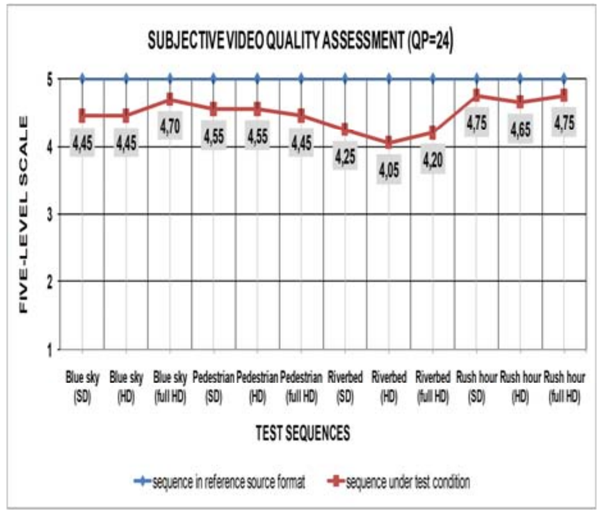


Fig. 5 Subjective video quality assessment results for different resolutions (SD, HD and *full* HD) and QP=24.

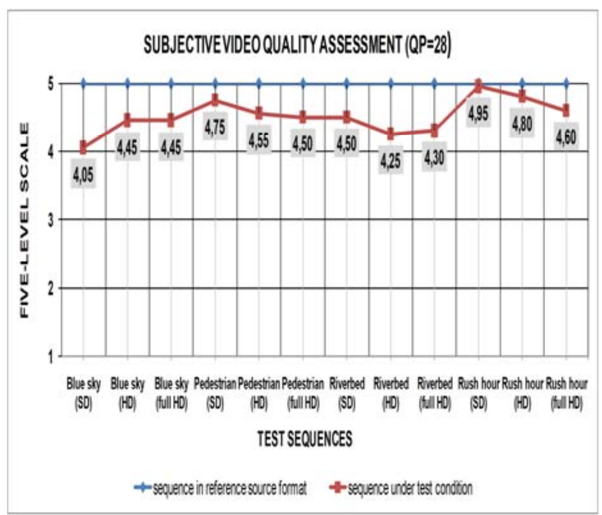


Fig. 6 Subjective video quality assessment results for different resolutions (SD, HD and *full* HD) and QP=28.

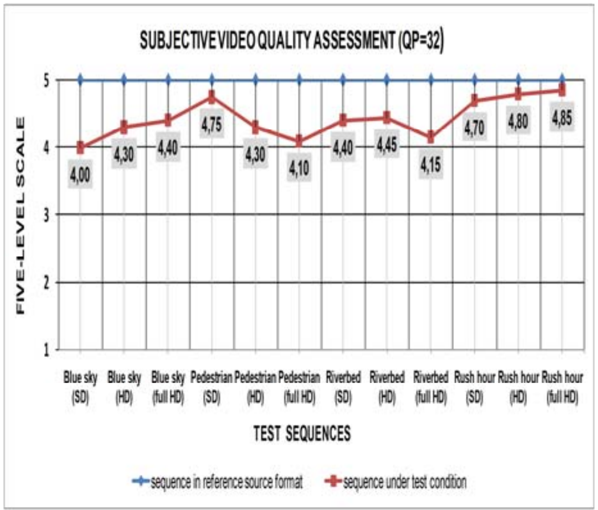


Fig. 7 Subjective video quality assessment results for different resolutions (SD, HD and *full* HD) and QP=32.

Figures 5, 6, 7 contain the results for the subjective evaluation of the tested video sequences quality in SD, HD and *full* HD resolution for different QP values. Besides the results concerning estimation the mean values for video test sequences in SD, HD and full HD resolution, respectively, the results for each of the resolution are presented, too.

6 Conclusion

In attempt to present subjective video quality estimation in H.264/AVC standard, special software program is developed using the recommendation ITU-T P.910, which is convenient for testing different multimedia applications.

The user interface for subjective test sequences quality evaluation is developed according to the rules defined for the Degradation Category Rating (DCR) testing method. Concerning observers all tests conditions are taken into account. After that, the subjective quality evaluation over the test sequences is carried out. Testing is realized in two steps. The user's interface preparation is carried out in the first step, while in the second one, the presentation of the processed test sequences through user interface is realized. The subjective rating about the quality, i.e. the observed difference between sequence in reference source format and sequence under the test condition, was given by the observers.

The results for subjective evaluation of the video quality using the proposed algorithm for selective intra- and optimized inter-prediction, confirmed that there exit a small difference in the signal-to-noise ratio, i.e., in video quality. To prove this, there are mean estimation values given by the viewers, varying in the range from 4,00 to 4,95 for each video sequence in the different video resolution and with a different quantization parameter (QP).

Using the subjective analysis in processed test sequences, it is confirmed that there exits a small difference in signal-to-noise ration, i.e., video quality between sequence in reference source format and sequence under the test condition.

References:

[1] Wade N, Swanston M., *Visual Perception: An Introduction*, 2nd edition, Psychology Press, London. 2001.

[2] Aldridge R., Davidoff J., Hands D., Ghanbari M., and Pearson D.E., Recency effect in the subjective assessment of digitally coded television pictures, in: Heriot-Watt University, *Proc. Fifth International Conference on Image Processing and its Applications* (July 4-6, 1995 Edinburgh, UK), 1995, 4, pp. 336-339.

[3] Richardson I.E.G., *H.264 and MPEG-4 video compression video coding for next-generation multimedia*, The Atrium, Southern Gate, Chichester, West Sussex; John Viley & Sons Ltd., 2003.

[4] Weigand T., Sullivan G., Bjontegaard G., Luthra A., Overview of the H.264/AVC video coding standard, *IEEE Transactions on Circuits and System for Video Technology*, vol. 13, no. 7, 2003, pp.560-76.

[5] Puri A., Chen H., Luthra A., Video coding using the H.264/MPEG-4 AVC compression standard, *Signal Processing: Image Communication*, vol.19, is. 9, 2004, pp. 793-846.

[6] H.264. International telecommunication union, recommendation ITU-T H.264: advanced video coding for generic audiovisual services, ITU-T, 2003.

[7] Kwon S., Tamhankar A., Rao K.R., Overview of the H.264/MPEG-4 part 10, *Journal of Visual Communication and Image Representation*, vol. 17, is.9, 2006, pp. 186-215.

[8] Ostermann J., Bormans J., List P., Marple D., Narroschke M., Pereira F., Stockhammer T., Wedi T., Video coding with H.264/AVC: tools, performances and complexity, *IEEE Circuits and System Magazine*, 2004, pp. 7-28.

[9] Pang W. L., Nadia N., Chieng D., Enhanced layer 3 service differentiation for WLAN, *WSEAS Transactions on Systems*, vol. 8, is. 5, 2009, pp. 649-658.

[10] Saraireh M. S., Al-Dmour N. A., Novel mechanisms for quality of service improvements in wireless Ad Hoc networks, *WSEAS Transactions on Systems*, vol. 10, is. 7, 2011, pp. 209-225.

[11] Zhang L., Xi L., Zhou B., Image retrieval method based on entropy and fractal coding, *WSEAS Transactions on Systems*, vol. 7, is. 4, 2008, pp. 332-341.

[12] Huang Y.W., Hsieh B.Y., Chen T.Ch., Clan L.G., Analysis fast algorithm and VLSI architecture design for H.264/AVC intra frame coding, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.15, is.3, 2005, pp. 378-401.

[13] Milicevic Z. M., Bojkovic Z. S., Intra/inter algorithm for B frame processing in H.264/AVC encoder (extended version), North Atlantic University Union - *International Journal of Communications*, vol. 1, is. 1, 2007, pp. 10-15.

[14] Milicevic Z., Bojkovic Z., H.264/AVC standard: A proposal for selective intra- and optimized inter- prediction, *Journal of Network and Computer Applications (JNCA)*, vol.34, is.2, 2011, pp. 686-691.

[15] Moorthy T., Chen, P. P., Ye A., A scalable architecture for H.264/AVC variable block size motion wstimation on FPGAs, *WSEAS Transactions on Signal Processing*, vol. 7, is. 1, 2011, pp. 23-33.

[16] Choi I., Lee J., Jeon B., Fast coding mode selection with rate-distortion optimization for MPEG-4 part-10 AVC/H.264, *IEEE Transactions on Circuits and System for Video Technology*, vol.16, no.12, 2006, pp. 1557–1561.

[17] Milicevic Z., Bojkovic Z., Computational time reduction using low complexity skip prediction for H.264/AVC standard, *WSEAS Transactions on Information Science and Applications*, vol. 14, is. 1, 2007, pp. 59-63.

[18] Wiegand T., Lightstone M., Mukherjee D., Campbell T., Mitra S.K., Rate distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard, *IEEE Transactions on Circuits and System for Video Technology*, 1996, vol.6, no.2, pp. 182–190.

[19] Sullivan G., Wiegand T., Rate-distortion optimization for video compression, *IEEE Signal Processing Magazine*, vol.15, no.11, 1996, pp. 74–90.

[20] JVT-K049, Joint model reference encoding methods and decoding concealment methods, ISO/IEC MPEG and ITU-T VCEG Joint Video Team, 2004.

[21] Park J.S., Song H.J., Selective intra prediction mode decision for H.264/AVC encoders, *Transactions on Engineering, Computing and Technology*, vol.13, 2006, pp. 51–55.

[22] Miličević Z., Bojković Z., An approach to selective intra coding and early inter skip prediction in H.264/AVC, *FACTA UNIVERSITATIS (NIŠ)*, Ser: Elec. Energ., vol. 21, no. 1, 2008, pp. 107-119.

[23] P.910, International telecommunication union, recommendation ITU-T P.910: subjective video quality assessment methods for multimedia applications, ITU-T, 2008.