# Methods of Measure and Analyse of Video Quality of the Image

Iulian UDROIU(1, Ioan TACHE(2, Nicoleta ANGELESCU(1, Ion CACIULA(1
1 – "VALAHIA" University of Targoviste, Romania
2 – POLITEHNICA University of Bucharest, Romania
iudroiu@valahia.ro, ioan.tache@gmail.com, nicoletaangelescu@yahoo.com, caciula@valalahia.ro

*Abstract:* - This paper present an analysis of the method for the evaluation of quality of the images from the video signal. The results are based on the simulation of the human perception. The device used for the evaluation is a TEKTRONIX PQA500 Picture Quality Analyzer.

*Key-Words:* -image quality, video quality, PQR, DMOS, PSNR.

## 1 Introduction

In [1], is presented a state of art of the methods for the assessment of video quality. The study analyses the quality of the MPEG - 2 streams.
Many organizations use in order to evaluate the video content quality, some subjective evaluating methods based on the observations of a person or a group who have the skill to detect and to interpret some defects of the images due to the digitizing processes or a communication transmission chain.
In order to uniform these methods, the researchers have developed recommendations (e.g. ITU-R BT.500, which present the selection methods and configuration of the video sequent which will be analyzed. These evaluating methods of the image quality are very expensive and time consuming. However, many organizations can not agree such as methods for trying to optimize the design of the subassemblies and the optimization of their characteristics [2].
In [3], the author presents the meaning of the redundancy versus human perception in the video and audio domain.
In [4], the authors propose a DVC(Distributed Video Coding) scheme and a pixel-domain Wyner-Ziv (PDWZ) video coding scheme in wireless multimedia sensor network for multimedia broadcasting. The results show an improvement of perceptual quality.
In [5], the authors propose a method to reduce the bandwidth in a mobile network using the H.264/AVC video compression standard and to enhance the video quality at same bit-rate.
On the other hand, the actual quality measurements depend on the comparison of a video reference sequence with a test sequence where the processing and transmission chain can bring some errors.
Evaluation of the quality decrease of the reference sequence in regard to the test sequence can be made in virtue of the pixel by pixel comparison. This method can lead at the quality errors. Therefore, if a man should evenly perceive the video content, the quality evaluation of this content would be made and pixel by pixel comparison of some reference image with the test image can be made by computing the mean squared error (MSE) which would represents in this case the noise introduced by system or the signal processing chain.
Though, the people can evaluate the video content mechanically and are not influenced by many factors which can affect the skill in order to perceive the differences between the video sequences.

## 2 The human perception

Clearly, the human perception is not equivalent with simple noise detection and the measurement of the image quality is made only by measurement of the noise difference between the reference video sequences and the test sequence (PSNR) is not enough clear. In order to adopt the subjective evaluations with the objective measurements of quality is necessary to take into consideration the characteristics of the human perception. This thing is shown bellow in Fig. 1.

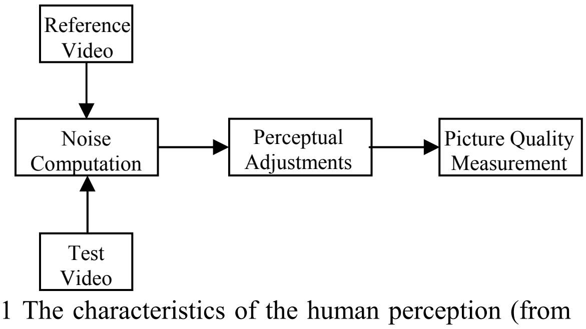


Fig. 1 The characteristics of the human perception (from [7] )

## 2.1 Details of experiment

Even if the measurement of PSNR is necessary in diagnosing the defects in the video processing and gives a general indication concerning the quality changes, nevertheless this measurement is not adapted good enough at the subjective human criteria.

Another adjusting alternative of PSNR measurement at these criteria is presented bellow in Fig. 2, in which adaptation at the human perception criteria is made by evaluation of the contrast difference of image.

This kind of evaluation includes the relationship between the contrast information perception and luminance and the different masking behaviors of the human perception. Practically, the measurement consists in computing the contrast difference perceived between the test image and the reference image, first than evaluation of PSNR difference between other two images (Fig. 2).

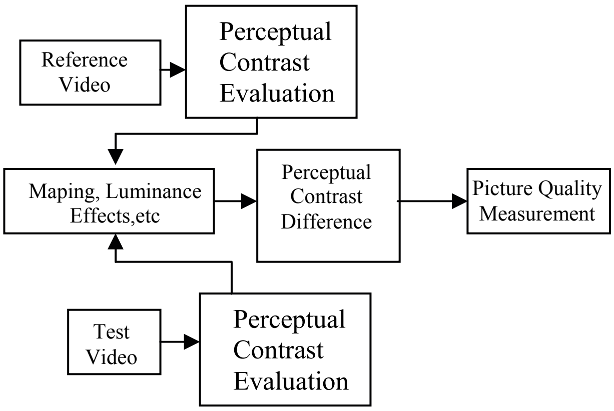


Fig. Picture qualitz measurement using Perceptual Contrast Difference (from [7])

Tacking into consideration these problems above mentioned, if it wants the replacement of the subjective tests, based on the human perception with a measurement fulfilled with a measuring equipment and can lead to the repeatable results, the used measurement method has to take into consideration and to include the corrections which has to take into account of the different behavior of the human factor in the video image perception.

For this, some parameters have to be defined, taking into account of the human perception.

## 2.2 Defining of PQR parameter

On basis of the observations above mentioned, it was introduced the PQR(Picture Quality Ratio) concept which converts the perception of the contrast difference between the reference image and the test image in a value which represents the skill of onlooker in order to perceive this difference.

Within the experiments of the subjective perception, this capacity of the onlooker in order to observe the differences between two images are evaluated by means of a parameter called JNDs ( Just Noticeable Differences ). The PQR value of 1 is equal with value of 1 JND. For a good assessment of the human perception conditions, the measuring method which has to take into consideration of the luminance information which there is within the test image and the reference image and also, the characteristic of the display which is used [10].

Taking into consideration the perception sensitivity experiments and measurements which are achieved considering the limit of this parameter, for PQR interpretation was realized the following scale:

- 0 : reference image and test image are identical;
- < 1 : perceptual contrast is less than 0.1% (or JND less than 1) – person who assesses can distinguish yet, the differences between other two images: test and reference;
- 1: perceptual contrast between the test image and the reference image is about 0,1% or JND is 1. Onlooker can assess the contrast differences due to the video processing or display technology or video conditions;
- 2-4: evaluater can distinguish the differences between the test and reference images. Between these limits, there are PQR values for MPEG encoder of high quality and wide frequency band. Video quality is between excellent and good;
- 5-9: evaluater can distinguish very easy the differences between the test image and reference image. PQR values between these limits correspond MPEG encleders of narrow band for current applications.
- >10: evaluater can observe big differences between the test image and reference image. The quality of the test images is worst.

Therefore, the PQR measurements are mainly useful to the video applications of high quality (for values between 2-4).

Because this assessment assumes the subjective perception of the images on a screen of a display, for PQR measurement is necessary a configuration of the measurement system which takes into account the human perception and its characteristics.

Therefore, it is necessary to be the possibility for selecting the visualization technology (CRT, LCD) and also, the characteristics of the equipments.

In order to implement the measuring technology based on the human perception, it was developed a model of the human perception based on the answer

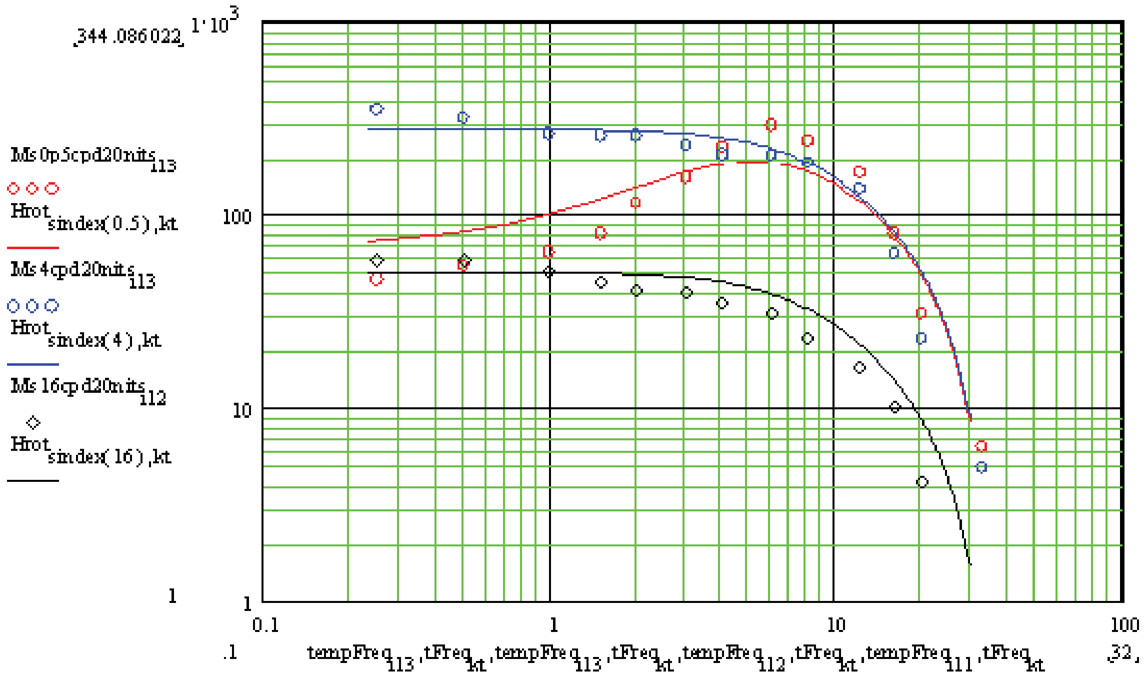


Fig.3 .Modulation Sensitivity versus Temporal Frequency (from [7])

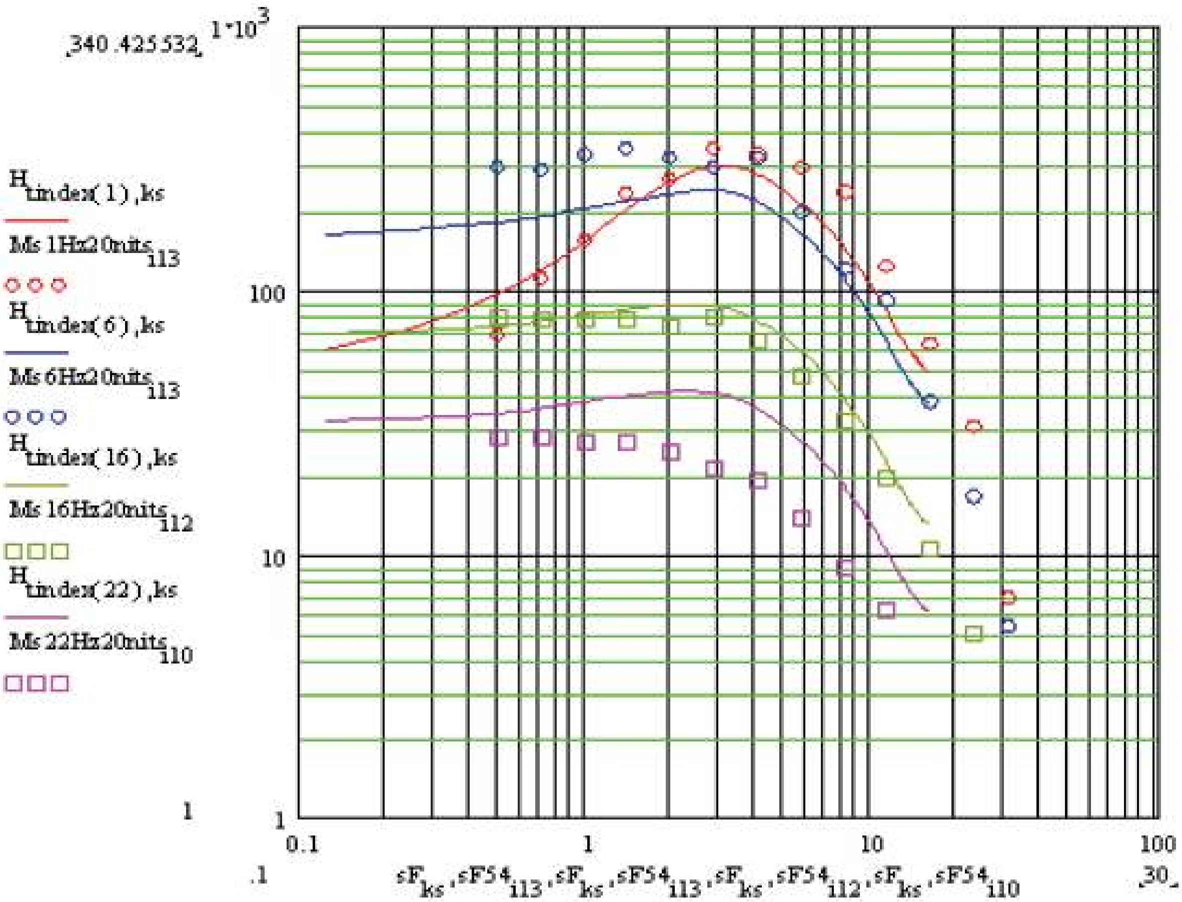


Fig. 4 Modulation sensitivity versus Spatial Frequency (from [7])

of the light stimuli according to the following parameters: contrast, medium luminance, spatial frequency, temporal frequency, temporal and angular extension, eccentricity, orientation, adaptation effects. The charts from Fig. 3 and Fig.4 represent the scientific data due to at the calibration of the system type of human visualization in metering device.

### 2.3 Defining of DMOS parameter

In the ITU-R BT.500-11 recommendation are presented some methods for the subjective perception of the television perception and for their assessment is used a scale of quality of four equal levels.

Each position on this scale is converted in a numerical value representing the score which is given the tested video image. Excellent area is placed between the values from 0 to 20, while the area having the worst quality is placed between the values of the score of 80 – 100[10].

These values are interposed depending on the number of the participants at this test and create the medium score of the opinions (MOS- Mean Opinion Score) for each particular sequence which are analyzed.

This parameter has a relative character. Its value depends on the values range in the worst case and the best case used preliminarily in the training sessions.

The relative character of DMOS reflects the relative quality scale for the different types of applications. Therefore, the quality scale for mobile video applications is different of the scale which refers to the cinema digital applications.

DMOS represents the perceptual contrast difference between the best case and the worst case for a particular measurement of video sequence.

Considerations concerning the influence of the image visualization on evaluater are available like in the case of PQR measurement. DMOS measurement takes into account the behavior of compression type for the human assessment which are shown bellow (Fig. 5).

Due to this behavior, the subjective evaluations are quantitative similar in S zone of the curve from this figure, having a major impact on the DMOS values which correspond the worst case because the images which are considered as having a low quality which correspond at DMOS values ranged at the maximum level of the values (about 100) can be performed as having values of about 65.

This thing is made as within the data analysis coming from a subjective evaluation allowing to the form of the curve and the compression phenomenon. Concerning the values scale, DMOS in range 0-20,

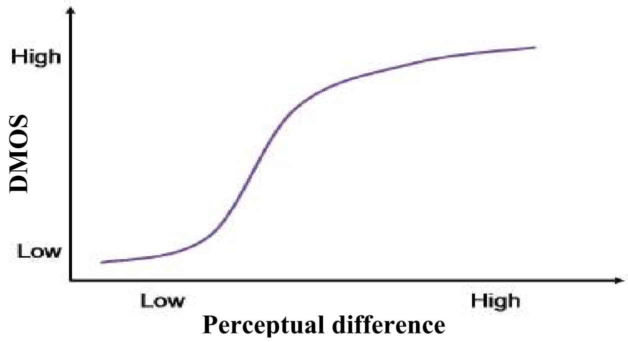


Fig. 5 DMOS vs Perceptual Difference(from [10])

there are the excellent or good images, while in the assessing range 21-40, there are the acceptable images or not so good quality.

Values of DMOS over 40 give the poor quality images which refer to the reference images. Within a measurement based on the subjective criteria, the assessment of DMOS parameter is made by an ample calibration process with the visual perception system.

### 2.4 Measurement of PSNR parameter

In the set of measurements based on the perceptual assessments, PSNR measurement does not have any adjustment depending on these[10].

PSNR value is computed as being the square root between the reference image and test image reported to the top value, which are computed for each frame from the test image and for the whole test sequence according the following formula[10]:

$$PSNR(f_n) = 20 * \log_{10}\left[\frac{235}{Diff}\right] \quad (1), \text{ where :}$$

$$Diff = \sqrt{\frac{1}{N_v N_h}\sum_{j=0}^{N_v-1}\sum_{i=0}^{N_n-1}\left[Y_{ref}(i,j,f_n) - Y_{last}(i,j,f_n)\right]^2}$$

$$PSNR_{seq} = 20 * \log_{10}\left[\frac{235}{Diff}\right] \quad (2), \text{ where}$$

$$Diff = \sqrt{\frac{1}{MN_v N_h}\sum_{n-0}^{M-1}\sum_{j=0}^{N_v-1}\sum_{i=0}^{N_n-1}\left[Y_{ref}(i,j,f_n) - Y_{last}(i,j,f_n)\right]^2}$$

According this formula, the result of the computations is not finally a ratio, this referring to the differences on the frame and on the whole sequence.
If the test image and the reference test are identical, PSNR has the values of 80dB.
If the reference image is an image having high quality, PSNR has the values of 40dB, if the test image is also of high quality.PSNR values under 30dB give the test video images having a low quality.
Combination of PSNR measurements with measurements based on the perceptual conception can create a full image concerning the impact of the difference between the test image and the reference image.
Such as the measurement is shown bellow (Fig. 6) where are drawn the curves which correspond to PSNR and DMOS.

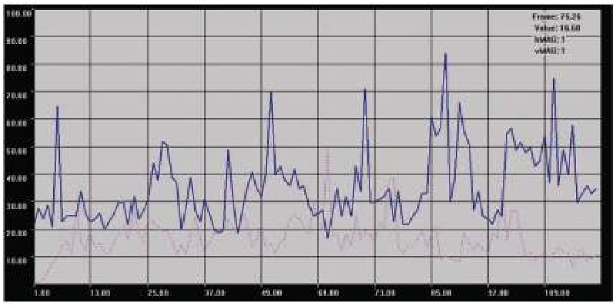

Fig .6 PSNR and DMOS curves

Comparing the map of the differences created between PSNR measurement and PQR or DMOS measurement, can be relieved all zones where there are such as problems, helping the engineers to convert these problems in hard or soft solutions.

## 3. Measurements

Measurements were performed by means of an equipment of TEKTRONIX PQA500 type, which is the first equipment based on the measuring technologies above mentioned.
As reference, it was used a video file in YUV format uncompress, which was especially chosen in order to relieve all measurement possibilities of the equipment used.
In order to obtain the test file, the reference file was passed by a coding-decoding chain MPEG 4.
The used measurement scheme is shown in Fig. 7.
In this case the inlet file was generated by a computer in the format mentioned and has a clip which contains 486 frames, which was passed by the measurement system.

### 3.1 Measurement of PSNR

A faster method in order to identify the possible errors which are by transmission of the video file in

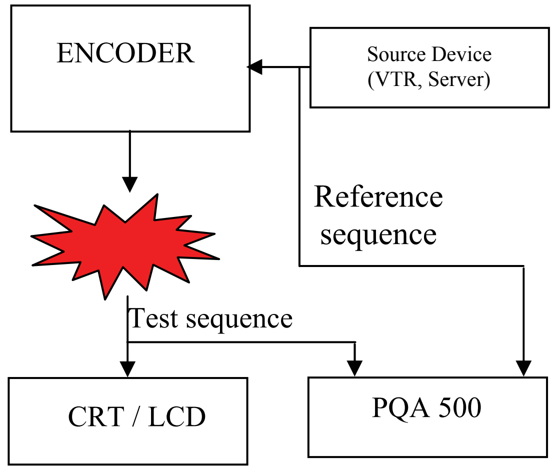


Fig. 7 The measurement scheme

the testing chain is to measure PSNR, an objective measurement and which in the case PQA500 gives the difference between the reference sequence and other test without filtration. This method is very useful in order to determine the differences between the reference sequence and other test pixel by pixel, this measurement not being a prediction on which a human observer can perceive as differences between images [6].
The measurement equipment PQA500 has the possibility to display on the screen the reference sequence, test sequence and PSNR values as chart or the difference image, as in figure bellow shown (Fig. 8).

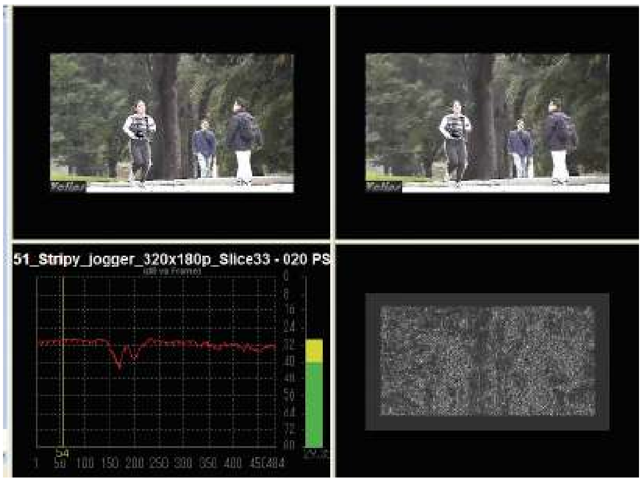


Fig. 8 The measurement equipment PQA500

PSNR is seen like the noise power obtained by the extraction of the reference sequence from other test.

The brightness zones correspond of the differences between other two sequences, there are not the between the reference sequence and other test. If differences, PSNR image would be black. In this case, PSNR has a good enough value and does not create big problems.
It remarks the fact, that in the PSNR computing operation, this equipment carried a spatial alignment of the sequences which allows the sequence comparison with the different resolutions.
The measurement gives an objective image on the differences between the reference sequences and other test introduced by the video processing chain.

### 3.2 Measurement based on the human perception

In order to carry the measurements based on the human perception, it has to be defined a measuring sequence of the equipment [2], [7].
This sequence is shown in Fig. 9:

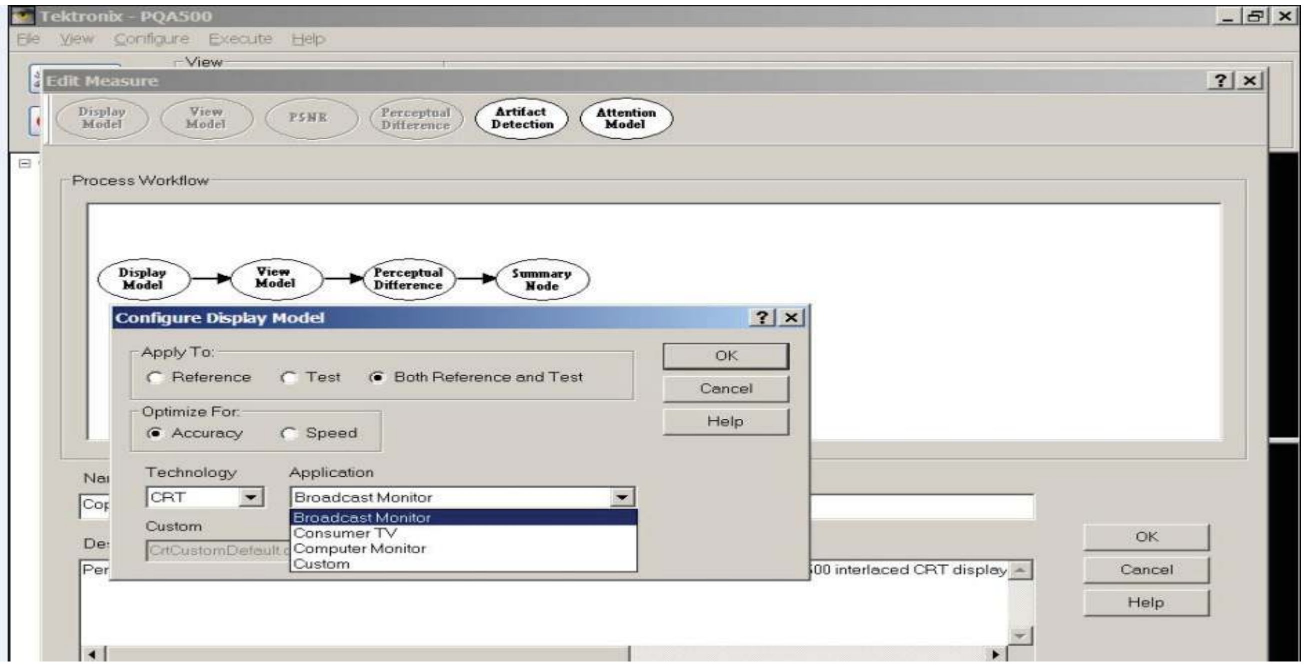

Fig.9 The configuration of display model

The first phase is to define the display model, for which the equipment has the configuration screens.
Data used for measurement are presented in Fig. 10, Fig. 11 and Fig.12. In the second phase, it is defined the visualization model, which supposes the setting of the visualization and perception conditions. Configuration screens are presented bellow.
By means of these measurement, this equipment creates the perceptual difference maps, which indicates the differences between the test

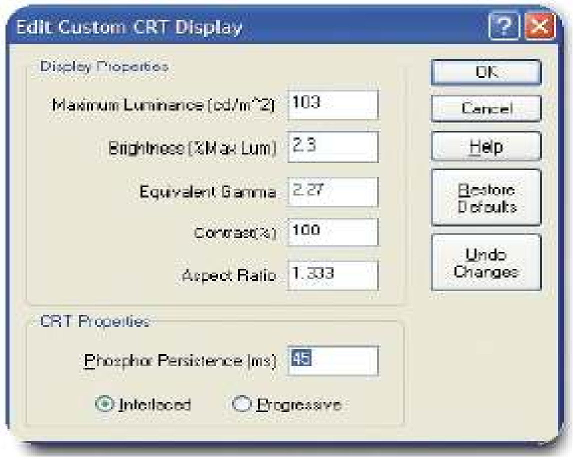

Fig. 10 Data used for measurement

sequences and reference sequence measured in some predefined conditions, these being relieved on image by brighter zones. After the difference zones are found between the reference sequences and test sequences, investigation can be carried on in order to improve the video processing stages by detecting the artifact type which produced the damage of image.
In this sense, the equipment PQA500 can measure all artifact types which there are in the compression process like Blurring, ringing/msquito noise, edge Blockiness si DC Blockiness. Such as measurement is presented bellow (Fig. 13). The equipment indicates on screen the reference sequence, test sequence, Perceptual Difference Map and appropriate chart.

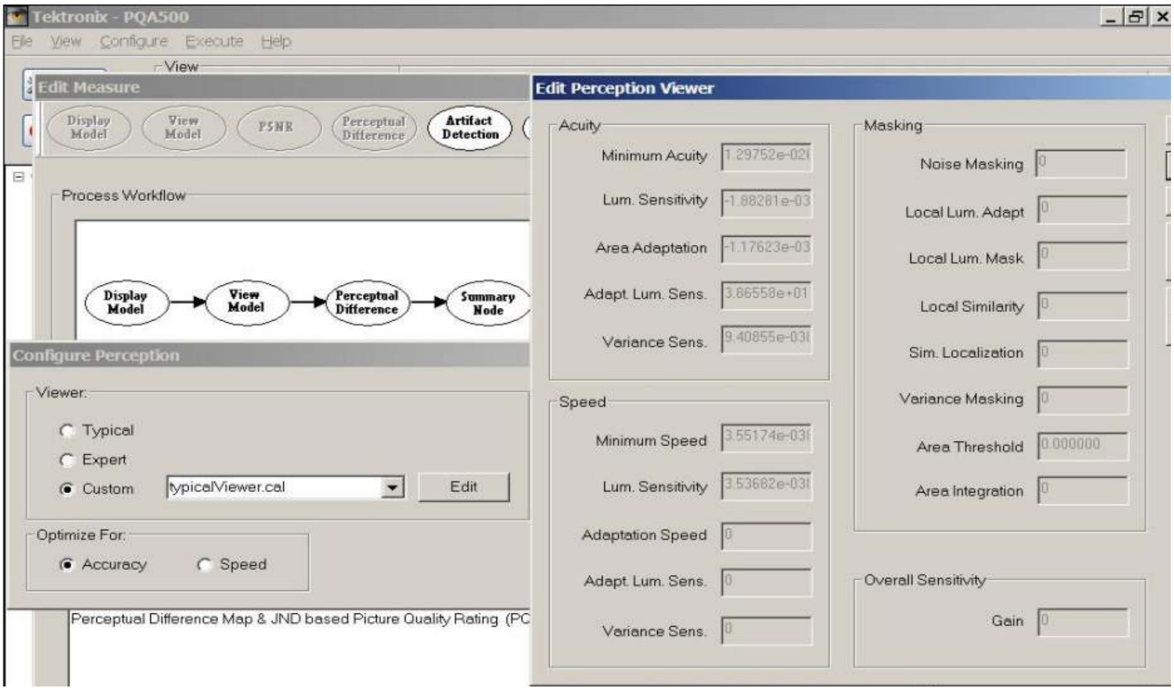

Fig. 11 Configure viewing conditions

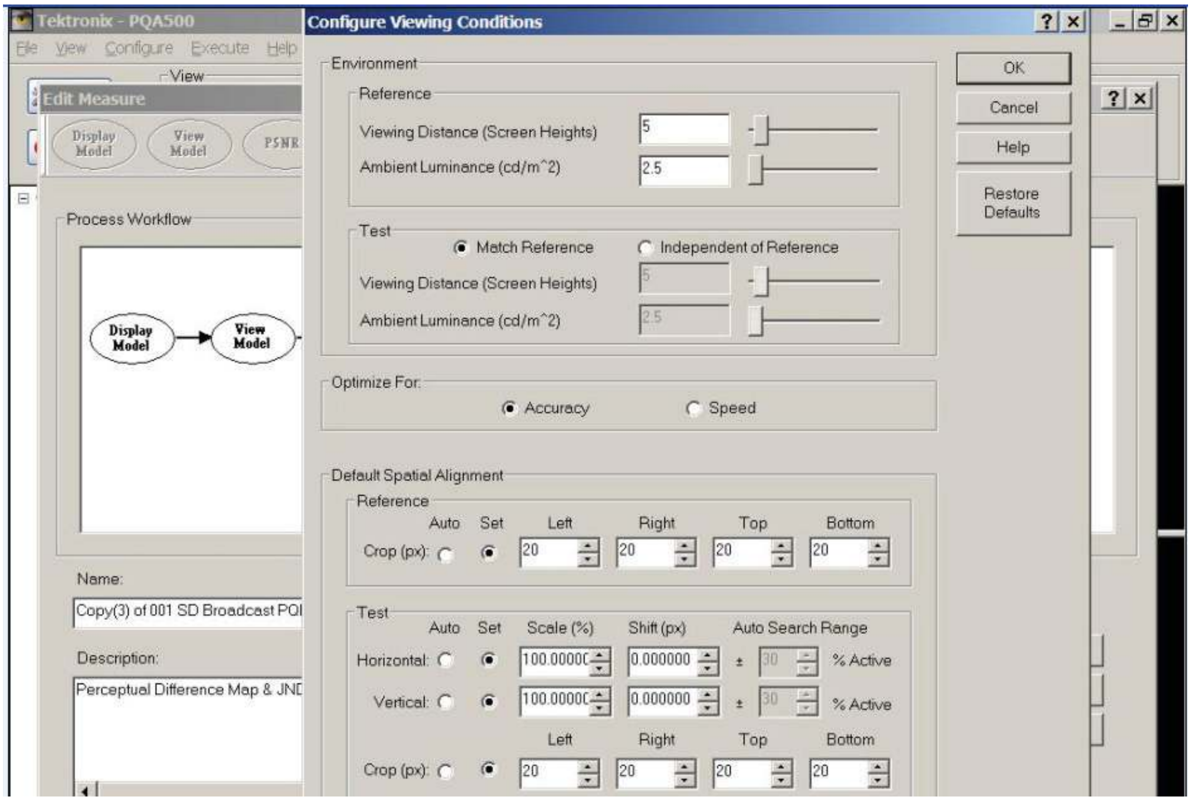

Fig. 12 Setting the perception viewer

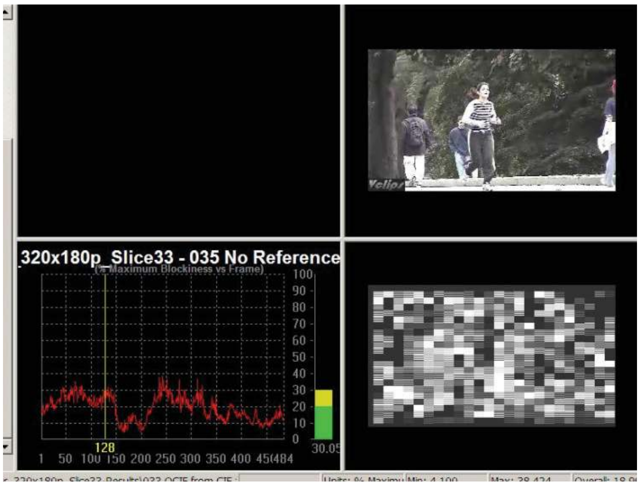


Fig. 13 The reference sequence, test sequence, Perceptual Difference Map and appropriate chart- the setup test

Measurements are made in presetted conditions (for broadcast SD,HD, cinema, etc) .Table 1, Table 2 and Table 3 contains the values measured for the first five frames using PQA500.

Table1 Measured parameters using PQA500 for the first five frames: Min,Max, Mean Abs. and StdDev

| Frame <Ref Test> | Min | Max | Mean Abs | StdDev |
|---|---|---|---|---|
| #1.0<1.0 1.0> | 0 | 56.274 | 7.594 | 118.691 |
| #2.0<2.0 2.0> | 0 | 59.400 | 19.000 | 129.842 |
| #3.0<3.0 3.0> | 0 | 61.434 | 19.929 | 135.769 |
| #4.0<4.0 4.0> | 0 | 56.181 | 19.197 | 141.082 |
| #5.0<5.0 5.0> | 0 | 34.761 | 14.661 | 64.0700 |

Table 2 Measured parameters using PQA500 for the first five frames:Minkowski,DMOS,PQR and PSNR

| Frame <Ref Test> | Minkowski | DMOS | PQR | PSNR |
|---|---|---|---|---|
| #1.0<1.0 1.0> | 21.846 | 99.999 | 68.321 | 20.491 |
| #2.0<2.0 2.0> | 23.387 | 99.999 | 69.073 | 19.898 |
| #3.0<3.0 3.0> | 24.317 | 99.999 | 69.521 | 19.541 |
| #4.0<4.0 4.0> | 23.876 | 99.999 | 69.309 | 19.736 |
| #5.0<5.0 5.0> | 17.396 | 99.999 | 66.069 | 22.352 |

The video sequence used in tests contains 486 reference frame and is encoded at video resolution 480x180. The graphic of PSNR versus reference frames and DMOS versus reference frames are shown in the Fig, 14 and Fig, 15.

According to Table 1 and Table 2, the artifact percent is reduced and does not degrade significant the image, this thing observing by the test image on the display.

Possibility of artifact detecting is very important because creates the possibility of DMOS improvement by taking the adequate measurements at the circuit level of video processing.

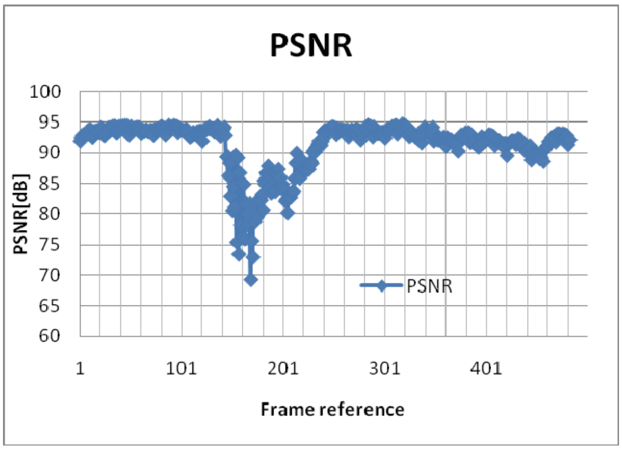


Fig 14 PSNR versur Frame reference graph

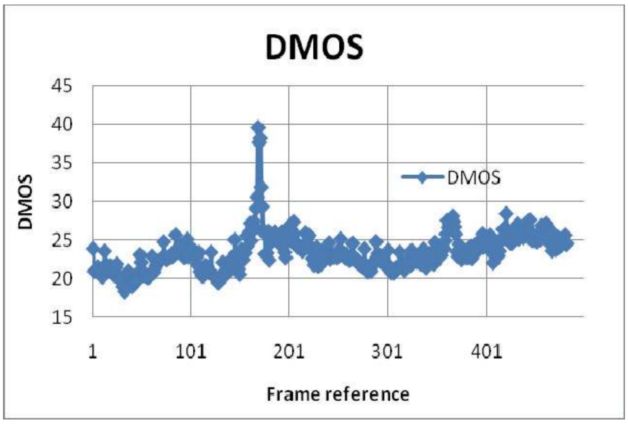


Fig 15 DMOS versur Frame reference graph

Thus, if edge-blocking artifact influences DMOS, much more than blurring artifacts, then it has to take into account a blocking filtration.

It can be determined as well, the influence of the artifact on PSNR, depending on the type or artifact detection can be measured together with any measurement combination of the image quality in order to observe what artifact can be visible on the image.

In the same time, the possibility of optimization of video processing stages, PQA500 lets the definition of an Attetion Model.

This thing reffers to the fact that the quality analysis Is focused on another content type. Thus, if we are talking about football meci, our attention is focused on player and the ball position in the playing field and less on the greensward or in tribunes.

If we refer to the news programmes as well, the attention is focused on speaker and not on the background.

Defining of an Attention Model allows an optimization of the image processing based on content, so that to obtain the high values of DMOS and PQR. In order to measure in this manner, PQA500 has a configuration meniu of this model and which is presented in Fig. 16:

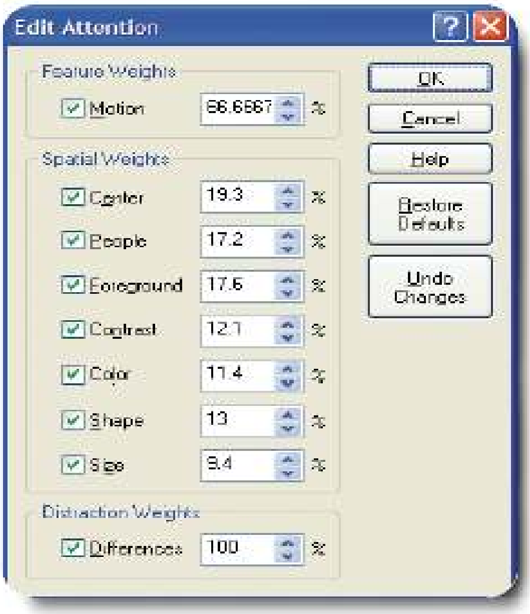


Fig 16 Defining an attention model

After setting of the atention conditions, the measurement can be made on the screen being displayed the test and reference sequences, as well as Attention Map and appropiate chart (Fig.17).

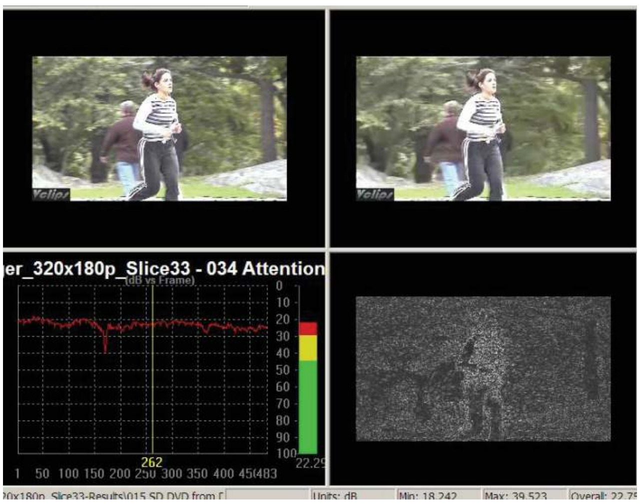


Fig. 17 The reference sequence, test sequence, Perceptual Difference Map and appropriate chart

This measurement is important from the optimization process point of view because the map indicates us as bigger is the perceptual difference in the image interest zones, in sense that indicates where bit resources have to be allocated in order to improve the image quality.

As in the case of Perceptual Difference Map, the image is followed by table filled with values of DMOS and PQR which give us the detailed informations concerning the image quality and it does not the significant damage situations.

## 4. Conclusions

Paper presented new methods of the quality measurement of images based on the simulation of the human perception, which come to complete the classical methods of quality assessment by introduction of the subjective factor.

Advantage of these methods consists in the fact that allows the repeatable measurements and are not time consumings.

On the other hand, thise measurements relieve some aspects concerning the signal damage on transmission chain which are not mentioned in the assessment methods of the causes which lead to the quality damage (DMOS) as well as the optimization of video processing stages concerning the improvement of the image quality based on an Attention Model.

By connection of the measurement system in the different sides of the transmission system, can be localized the points where the damage of the quality image is significant, with mention that PQA equipment knows to make only measurements on images uncompressed.

In order to make full measurements on a digital transmission system of images, such as measurements should be completed as well, with the objective measurements obtained at the stream-urilor MPEG level carried e.g. with MPEG Analyzer Tektronix of MTM400A type or MTS400EA ( software version), which can relieve some problems at the level of compressed signals which are not different at thefirst measurement.

*References:*


[1] S. Olsson, M. Stroppiana, J. Bafna, Objective methods for assessment of video quality : state of the art, , *IEEE Transactions on Broadcasting*, Vol. 43, Issue 4, Dec 1997,pp. 487 – 495

[2] K.Ferguson –Tektronix Inc.Portland OR97077- An adaptable Human Vision Model for Subjective Video Quality Rating Prediction among CIF, SD, HD and E-cinema.

[3] Blank M, Redundancy versus video and audio human perception, *WSEAS Transactions on Communications*, Vol.7, Issue 7, July 2008, pp.749-764

[4] Xue Z, Loo K.K, Cosmas J, Yip P.Y, Distributed video coding in wireless multimedia sensor network for multimedia broadcasting, *WSEAS Transaction on Communications*,Vol. 7, Issue 5, May 2008, pp.418-427

[5] Liu Y.-P., Zhang S.-Y, Xu S.-C.,Zhang Y., Study of significant technology of H.264 in Mobile Network, *WSEAS Transactions on Communications,* Vol. 8, Issue 1, January 2009, pp. 51-60

[6] Tektronix AN – Picture Quality Measurements for Video Application.

[7] Tektronix – PQA 500 Data Sheets.

[8] Tektronix Technical Brief – Comparing Objective and Subjective Picture Quality Measurements.

[9] Tektronix AN – A Guide to Picture Quality Measurements for Modern Television Systems

[10] Tektronix AN – Understanding PQR, DMOS and PSNR Measurement