Multimed Tools Appl
DOI 10.1007/s11042-009-0441-2

Comparison of HDTV formats using objective video quality measures

Emil Dumic · Sonja Grgic · Mislav Grgic

Abstract In this paper we compare some of the objective quality measures with subjective, in several HDTV formats, to be able to grade the quality of the objective measures. Also, comparison of objective and subjective measures between progressive and interlaced video signal will be presented to determine which scanning emission format is better, even if it has different resolution format. Several objective quality measures will be tested, to examine the correlation with the subjective test, using various performance measures.

Keywords Video quality · PSNR · VQM · SSIM · TSCES · HDTV · H.264/AVC · RMSE · Correlation

1 Introduction

High-Definition Television (HDTV) acceptance in home environments directly depends on two key factors: the availability of adequate HDTV broadcasts to the consumer's home and the availability of HDTV display devices at mass market costs [6]. Although the United States, Japan and Australia have been broadcasting HDTV for some years, real interest of the general public appeared recently with the severe reduction of HDTV home equipment price. Nowadays many broadcasters in Europe have started to offer HDTV broadcasts as part of Pay-TV bouquets (like BSkyB, Sky Italia, Premiere, Canal Digital ...). Other major public broadcasters in Europe have plans for offering HDTV channels in the near future. The announcement of Blu-Ray Disc and Game consoles with HDTV resolutions has also increased consumer demand for HDTV broadcasting.

The availability of different HDTV image formats such as 720p/50, 1080i/25 and 1080p/50 places the question for many users which HDTV format should be used with which compression algorithm, together with the corresponding bit rate. 720p/50 format is a format

E. Dumic (✉) · S. Grgic · M. Grgic
University of Zagreb, Faculty of Electrical Engineering and Computing,
Department of Wireless Communications, Unska 3/XII, HR-10000 Zagreb, Croatia
e-mail: emil.dumic@fer.hr
URL: http://www.vcl.fer.hr/

Published online: 29 January 2010

with 1280 horizontal and 720 vertical pixel resolution and progressive scanning at 50 frames/sec as specified in SMPTE 296M-2001 [18]. 1080i/25 format is a format with 1920 horizontal and 1080 vertical pixel resolution and interlaced scanning at 25 frames/sec or 50 fields/sec as specified in SMPTE 274M-2005 [17] and ITU-R BT. 709-5 [11]. 1080p/50 format is a format with 1920 horizontal and 1080 vertical pixel resolution and progressive scanning at 50 frames/sec as specified in SMPTE 274M-2005 [17] and ITU-R BT. 709-5 [11]. A debate in Europe about the best HDTV format for emission led to the Technical Recommendation R112 - 2004 [3] of the European Broadcasting Union (EBU) which recommends that emission standards for HDTV should be based on progressive scanning: 720p/50 is currently the optimum solution, but 1080p/50 is an attractive option for the longer term. Objections were raised by those who had decided in the past to adopt the 1080i/25 HDTV format.

Virtually all future HDTV displays sold in Europe will be matrix displays that require de-interlacing of the interlaced video signal (1080i/25). The video quality achieved after the de-interlacing process would partly depend on the complexity of the de-interlacing algorithm that would affect the price of the consumer display. Progressive scanning at 50 fps (frames per second) provides better motion portrayal than interlaced scanning with 25 fps or 50 fields per second. This is particularly important for critical HDTV genres such as sport.

The choice of raster and scanning algorithm is only one element that should be considered when deciding about the policy for HDTV broadcasting. Other important elements are the economic basis for the service, the content arrangements, the delivery platform, the format and compression system for audio and the compression systems for video. Any HDTV services launched later than the middle of 2005 use one of the ‘advanced coding’ schemes (MPEG-4 Part 10, i.e. AVC/H.264 [9] or possibly VC1 [19]), that allow better picture quality at lower bit rates. The choice of bit rate for HDTV broadcasting needs to take into account the economic savings associated with having more HDTV programmes against the picture quality benefit. Any HDTV broadcast will need to stand picture quality comparison (in the home) with downloaded and packaged HDTV media, which have the capability to use relatively high bit rates. These requirements put challenges on the video compression format applied in HDTV systems, particularly on the trade-off between the bit rate and video quality.

In this paper we examine the efficiency of MPEG-4 Part 10 Advanced Video Coding (AVC) compression system in compression system that uses progressive and interlaced HDTV formats. It is often claimed that progressive video is compressed more efficiently than interlaced video. Therefore, it provides better image quality at a given bit rate. The tests will be done using three resolution formats: 720p/50, 1080i/25 and 1080p/50, and using different compression ratios achieved by H.264/AVC compression algorithm.

This paper is organized as follows. In section 2 objective video quality measures are presented. Section 3 explains Triple Stimulus Continuous Evaluation Scale (TSCES) as subjective measure of image quality. Section 4 presents encoder settings. Performance measures are given in Section 5. Section 6 compares different objective video quality measures with the results of subjective assessment, using different performance measures. Finally, section 7 draws the conclusion.

2 Objective video quality measures

To be able to compare original and compressed sequences, we used the following three objective quality measures:

- PSNR (Peak Signal to Noise Ratio) [4, 24];

- SSIM (Structural Similarity Index) [22];
- VQM (Video Quality Measure) [26].

PSNR is the ratio between the maximum possible signal power and the noise power:

$$PSNR = 10\log_{10}\frac{255^2}{MSE}$$
$$MSE = \frac{\sum_i \sum_j \left(a_{i,j} - b_{i,j}\right)^2}{x \cdot y} \tag{1}$$

It is usually expressed in terms of the logarithmic decibels. In (1) $a_{i,j}$ and $b_{i,j}$ are pixels from original and compressed image, x and y describe height and width of an image and MSE stands for Mean Square Error. When PSNR is calculated for a color video sequence, there are more possibilities to calculate the final PSNR. Human Visual System (HVS) has considerably less ability to sense detail in color information than in lightness [8]. Therefore, information for color difference (Cr and Cb planes) in images or video sequences can be ignored if it is of no importance. In this paper we will also ignore color information when calculating PSNR. There are two possibilities to calculate PSNR of a video sequence from PSNR of each of the images: average or global (or overall) PSNR. Average PSNR calculates the final PSNR as the mean value of all PSNR values for each frame (or field). Overall PSNR is done after calculating the mean MSE for all frames in the video sequence, thus solving the problem of having the perfect frame in the sequence. Average PSNR would then calculate infinite PSNR, while overall PSNR will have finite value. In this paper we will use average PSNR (for luminance component) because there is no perfect frame and moreover, tested encoder has already implemented average PSNR for luminance component.

The Structural Similarity (SSIM) is a novel method for measuring the similarity between two images [22]. It is computed using three image measurement comparisons: luminance, contrast and structure. Each of these measures is calculated over the 8×8 local square window moved pixel-by-pixel over the entire image. At each step, the local statistics and SSIM index are calculated within the local window. Because resulting SSIM index map often exhibits undesirable "blocking" artifacts, each window is filtered with a Gaussian weighting function (11×11 pixels). In practice, one usually requires a single overall quality measure of the entire image, so Mean SSIM (MSSIM) index is computed to evaluate the overall image quality. The SSIM represents a quality measure of one of the images being compared, while the other image is regarded as being perfect. SSIM gives results between 0 and 1, where 1 means excellent quality and 0 means poor quality. In this paper we will use average SSIM across all frames (or fields), only for luminance component.

Video Quality Measure (VQM) presents another approach in image quality measuring that correlates more with the HVS. Figure 1 shows an overview of the flowchart of VQM [26].

The first step is color transform to the YUV color space. After that original and compressed image are transformed using DCT transform. This step separates incoming images into different spatial frequency components. Third step is converting each DCT coefficient to local contrast (LC). After this step, most values lie inside [−1, 1]. Fourth step converts LC to just-noticeable differences. The DCT coefficients are converted to just-noticeable differences by multiplying each DCT coefficient with its corresponding entry in the SCSF (Spatial Contrast Sensitivity Function) matrix. MPEG default quantization matrix is used for the static SCSF matrix. For the dynamic matrix each entry in the static SCSF

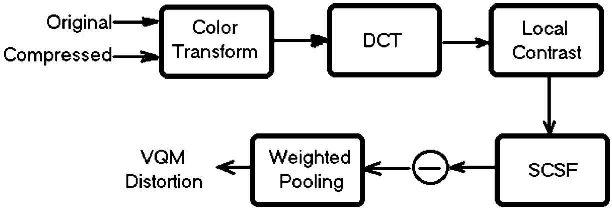


Fig. 1 VQM measuring algorithm

matrix is raised to the power decided by the frame rate of video sequences. The final step is weighted pooling of the mean and the maximum distortion. Here the two sequences are subtracted. At this step, VQM also incorporates contrast masking into a simple maximum operation and then weights it with the pooling mean distortion. This reflects the facts that a large distortion in one region will suppress sensitivity to other small distortion, because weighted maximum distortion into pooled distortion is much better than the pooled distortion alone [26]. VQM measure can obtain results between few grades and zero, where values which are near zero mean almost identical tested and original video sequence. In this paper we will use VQM, only for luminance component of a video sequence.

3 Subjective video quality measure

Subjective quality measure TSCES (Triple Stimulus Continuous Evaluation Scale) for sequences Crowdrun and Parkjoy was taken from [7], for distances "3H" or "4H" (where H stands for display height). Basically, in this subjective measure, a non-expert and an expert viewers, male and female of average age, were selected as observers after screening for normal vision. Training sequences and an explanation were given before the viewings, and short relaxation breaks between the testing series were offered to the observers. Assessors were presented with three equal HDTV monitors (Grade-1 type). Vertical angles of the three displays are adjusted so that a viewer at an eye height of 1.2 meter maintains a constant viewing distance (of 3 times picture height) from all three displays. Following settings were used for each of the displays [7]:

- ITU-R BT.500-11 viewing environment [10];
- all three displays have to show the same scene at the same time;
- all three displays have to be of the same Grade-1 type;
- top display shows video sequence that is regarded as of the perfect quality (in this case 1080p);
- middle display shows the video sequence under the test;
- bottom display serves as display showing video of the worst quality (in this case 576i format downsampled from 1080i format and then compressed using H.264/AVC algorithm at the bit rate of 3 Mbit/s, so that the degradation would be of the same type as on the middle display).

Assessors are given clear instructions on how to vote their results prior the test. They were asked to mark on the line (Fig. 2) where the quality of the sequence on the central monitor falls between the quality of the sequence shown on the top and bottom monitor. The results can be afterwards mapped on the 5 impairment scale or onto a 100 point continuous quality scale. It is also suggested that upper and lower reference video

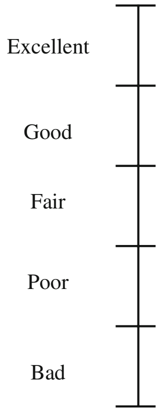


Fig. 2 Scale used for voting

sequences should be included in the test measurement. If assessors do not recognize them within a 20% threshold their grades should be removed, because this indicates that assessors were either not concentrated or did not understand the method.

4 Encoder settings

H.264/AVC compression standard makes delivery of HDTV broadcasts more quality/cost effective. H.264/AVC is a general purpose video coding standard which allows different bit rates and quality settings. The general idea of this standard is to provide HD quality image at relatively low bit rates [23]. Figure 3 shows typical video coding/decoding chain and scope of video coding standardization, decoding.

To be able to compare three of the above mentioned objective video quality measures, we have used subjective quality results from [7]. Test sequences used in this comparison are Crowdrun and Parkjoy, which source material can be downloaded from [15]. Sequence in three different resolutions (1080p, 1080i and 720p) was first converted from .sgi to .yuv format and 4:2:0 chroma subsampling format [2] using sgi2yuv program [20]. Afterwards, these uncompressed .yuv sequences were converted to .264 raw bytestream using H.264/AVC compression (High Profile) and freeware x264 Encoder [25]. We could not use the recommended reference H.264/AVC JM Encoder [21], that was used in [7], because its software implementation is not optimized for daily-used computers (hardware configuration used: AMD Athlon64 X2 4200 MHz, 4 GB RAM, Windows Vista 64). For the first few frames of Crowdrun 720p uncompressed sequence the average time for calculating one compressed frame using H.264/AVC JM was between 5 and 10 min (about 0.002–0.003 fps). On the same hardware configuration the average time in fps (frames per second) for x264 Encoder is shown in Table 1.

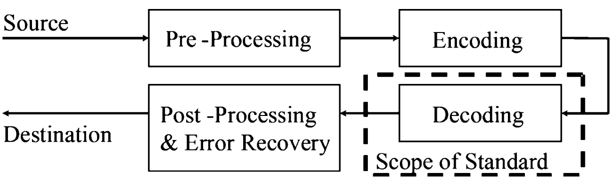


Fig. 3 Video coding steps

Basic settings for H.264/AVC x264 Encoder, for all resolutions, are the following:

- 1 slice per frame;
- search shape: 8×8;
- CABAC entropy coding;
- default quantization matrix;
- uneven multi-hexagonal search as pixel motion estimation method (for higher values of motion vector search range);
- variable bit rate, average bit rate values: 6, 8, 10, 13, 16 and 18 Mbit/s.

Specific settings for 720p resolution format are:

- HP@level4.0;
- direct temporal mode used as motion vector prediction mode;
- maximum motion vector search range 96 pixels;
- 24 picture GOP.

Specific settings for 1080i resolution format are:

- HP@level4.0;
- spatial temporal mode used as motion vector prediction mode (because direct mode does not work with option 'interlaced' yet);
- 'interlaced' mode used (using MbAFF—Macroblock Adaptive Frame/Field coding);
- maximum motion vector search range 128 pixels;
- 6 picture GOP.

Specific settings for 1080p resolution format are:

- HP@level5.0;
- direct temporal mode used as motion vector prediction mode;
- maximum motion vector search range 128 pixels;
- 24 picture GOP.

We did not use hierarchical B-frame coding (option 'b-pyramid') as suggested in [7], because in used encoder this option is not supported with 'mb-tree' (macroblock tree ratecontrol) option and because 'mb-tree' option gave slightly better results than option 'b-pyramid'. In the end, the original sequence (that has the same resolution format as tested and 4:2:0 chroma subsampling format) and the compressed sequences were compared using PSNR, SSIM and VQM quality measures.

5 Performance measures

To be able to compare different objective video quality measures with TSCES, we have used several different performance measures:

- Pearson's product-moment correlation coefficient;

Table 1 Average time in fps required for x264 Encoder to compress different sequences, using settings described below

	720p	1080i	1080p
Crowdrun	3.028	0.628	1.08
Parkjoy	3.277	0.728	1.098

- RMSE (Root Mean Square Error);
- Spearman's rank-order correlation coefficient.

Pearson's product-moment correlation coefficient is calculated as:

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \overline{x}))(y_i - \overline{y})}{(n-1) \cdot s_x \cdot s_y}, \mathrm{i} = 1, ..., n \tag{2}$$

where x_i and y_i are sample values (x are results for different objective measures and y are results for TSCES), and x and y are sample mean:

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i, \overline{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i \tag{3}$$

$$s_x = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - x)^2} \tag{4}$$

$$s_y = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{5}$$

s_x and s_y are standard deviations (calculated using n-1 in the denominator).

Pearson's correlation reflects the degree of linear relationship between two variables, from −1 to 1, where 0 means that there is no relationship and ±1 means perfect fit.

We calculated Pearson's correlation coefficient before and after nonlinear regression. The nonlinearity chosen for regression for each of the methods tested was a 5-parameter logistic function (a logistic function with an added linear term), as it was proposed in [16]:

$$Q(x) = b_1 \cdot \left(\frac{1}{2} - \frac{1}{1 + e^{b_2 \cdot (x - b_3)}}\right) + b_4 \cdot x + b_5 \tag{6}$$

However, this method has some drawbacks. First, logistic function and its coefficients (b_1, ..., b_5) will have direct influence on correlation (e.g. if someone chooses another function or even the same function with other parameters, results can be quite different). Another drawback is that function parameters are calculated after the objective measures calculation, which means that resulting parameters will be defined by the used video sequences. Different sequences can again produce different parameters. Coefficient parameters are given in Tables 2, 3, and 4. We have used three different methods to find the best fitting coefficients:

- Trust-Region method [14];
- Levenberg–Marquardt method [12], [13];
- Gauss–Newton method [1].

The final method for finding fitting coefficients for nonlinear regression was the one which computed better results for performance measures (lower RMSE and higher

Table 2 Coefficient parameters for logistic function, for Crowdrun sequence

Measure	b_1 (95% confidence bounds)	b_2 (95% confidence bounds)	b_3 (95% confidence bounds)	b_4 (95% confidence bounds)	b_5 (95% confidence bounds)
PSNR	−105.1	−0.5288	28.25	−4.987	195
	(−1804, 1593)	(−4.877, 3.82)	(24.01, 32.49)	(−123.6, 113.6)	(−3136, 3526)
SSIM	−78.94	−18.45	0.7885	−94.35	124.7
	(−1741, 1584)	(−236.2, 199.3)	(0.5374, 1.04)	(−3860, 3672)	(−2805, 3054)
VQM	12.73	66.03	2.396	−20.38	120
	(−4.591, 30.04)	(−1997, 2129)	(0.4366, 4.354)	(−29.33, −11.43)	(95.42, 144.5)

Pearson's and Spearman's correlation). RMSE (Root Mean Square Error) is calculated as:

$$RMSE = \sqrt{\frac{1}{n-k} \cdot (x-y)^2} \tag{7}$$

In (7) n is the number of tested video sequences, modified by a correction for degrees of freedom (k=5 in our case, because we have 5 parameters in the fitted function, Equation (6)), x is TSCES measure and y is fitted objective measure after nonlinear regression.

Spearman's correlation coefficient is a measure of a monotone association that is used when the distribution of the data makes Pearson's correlation coefficient undesirable or misleading. Spearman's coefficient is not a measure of the linear relationship between two variables. It assesses how well an arbitrary monotonic function can describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables [5].

6 Results

6.1 Video quality measures

Figures 4 and 5 show results for different video quality measures relative to bit rate. For each resolution objective quality measures were calculated on slightly different bit rates,

Table 3 Coefficient parameters for logistic function, for Parkjoy sequence

Measure	b_1 (95% confidence bounds)	b_2 (95% confidence bounds)	b_3 (95% confidence bounds)	b_4 (95% confidence bounds)	b_5 (95% confidence bounds)
PSNR	−60.73	−0.761	26.49	−0.1102	58.12
	(−661.8, 540.3)	(−5.958, 4.436)	(19.72, 33.25)	(−42.66, 42.44)	(−1023, 1139)
SSIM	−688.2	−4.427	0.5183	−252	69.7
	(−3.354e+005, 3.34e+005)	(−1095, 1086)	(−80.31, 81.34)	(−1.13e+005, 1.124e+005)	(−1.219e+004, 1.233e+004)
VQM	119.3	−1.474	3.725	18.14	−5.033
	(−2968, 3207)	(−18.33, 15.38)	(2.824, 4.626)	(−640.5, 676.8)	(−2452, 2442)

Table 4 Coefficient parameters for logistic function, for both video sequences

Measure	b_1 (95% confidence bounds)	b_2 (95% confidence bounds)	b_3 (95% confidence bounds)	b_4 (95% confidence bounds)	b_5 (95% confidence bounds)
PSNR	−80.77	−0.5549	27.15	−3.052	138.8
	(−806.9, 645.4)	(−3.542, 2.433)	(21.6, 32.7)	(−48.81, 42.71)	(−1075, 1353)
SSIM	−397.3	−7.714	0.7497	−472.9	396.6
	(−2.912e+004, 2.833e+004)	(−249.3, 233.8)	(−0.3289, 1.828)	(−3.162e+004, 3.068e+004)	(−2.223e+004, 2.302e+004)
VQM	181.5	−1.197	3.632	31.96	−56.83
	(−3225, 3588)	(−10.79, 8.401)	(2.837, 4.427)	(−567.8, 631.8)	(−2238, 2125)

because x264 Encoder could not compress sequences at exactly targeted bit rate (final and targeted bit rates could be different up to 12% for resolution 1080i and about 5% for 720p and 1080p). Only TSCES is shown at exactly targeted bit rates (according to [7]).

Figure 6 shows comparison between objective quality measures (PSNR, SSIM, VQM) and subjective quality measure (TSCES) before and after nonlinear regression, for both video sequences together (36 sequences).

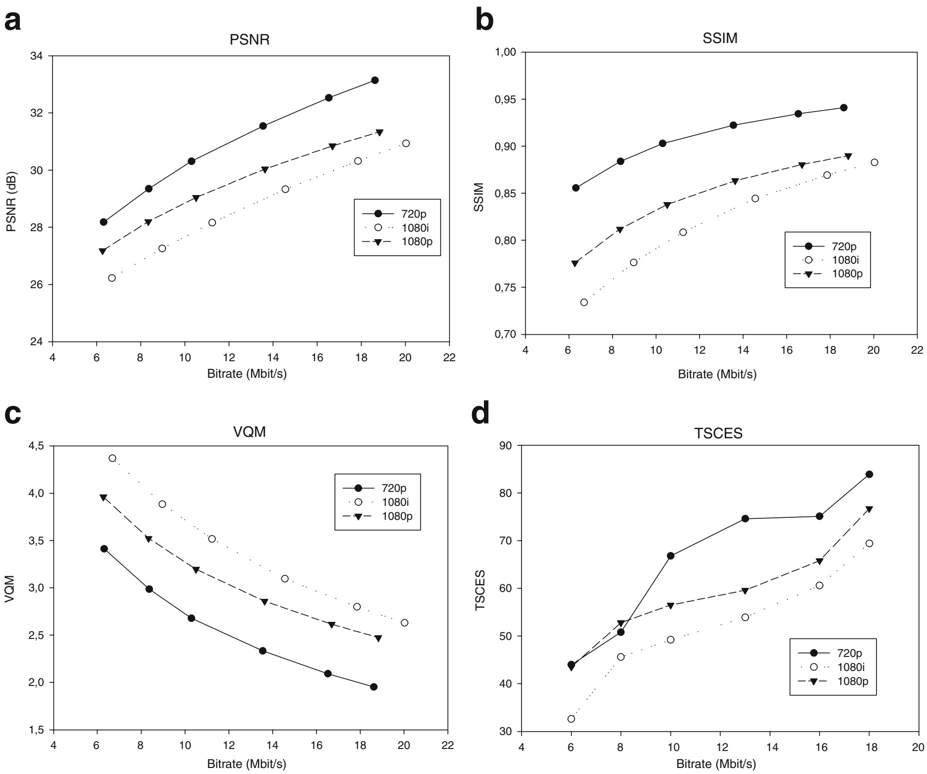


Fig. 4 Different video quality measures relative to bit rate, Crowdrun sequence: (**a**) PSNR, (**b**) SSIM, (**c**) VQM, (**d**) TSCES

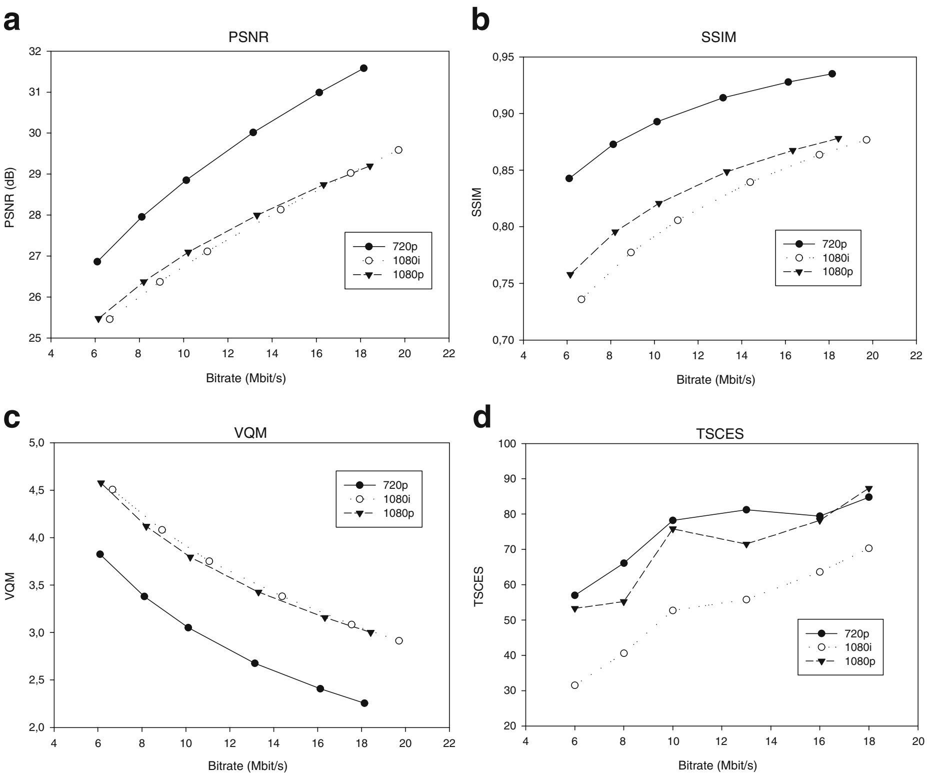


Fig. 5 Different video quality measures relative to bit rate, Parkjoy sequence: (**a**) PSNR, (**b**) SSIM, (**c**) VQM, (**d**) TSCES

TSCES had to be corrected because it was calculated for exactly round bit rate (according to [7]) and our encoder only coded sequences at rates near given bit rates. It was assumed that TSCES changes linearly with rate in the near region (higher bit rate produces linearly higher TSCES).

6.2 Results of the RMSE, Spearman's and Pearson's correlation

In this section we examine how well each objective measure fits each video sequence separately, as well as overall results for both sequences, before and after nonlinear regression used in previous section. Results for coefficient parameters for logistic function are presented in Tables 2, 3, and 4. RMSE, Spearman's and Pearson's correlation parameters for each video sequence are given in Fig. 7 and for both sequences together in Fig. 8. When calculating correlation coefficients, those which are calculated before nonlinear regression are denoted on figures with black bars, and after nonlinear regression with gray bars.

6.3 Discussion of the results

In section 6.2 we presented RMSE, Pearson's and Spearman's correlation for Crowdrun, Parkjoy and both video sequences together. For Crowdrun sequence, VQM gives best results for both correlations and lowest RMSE (after nonlinear

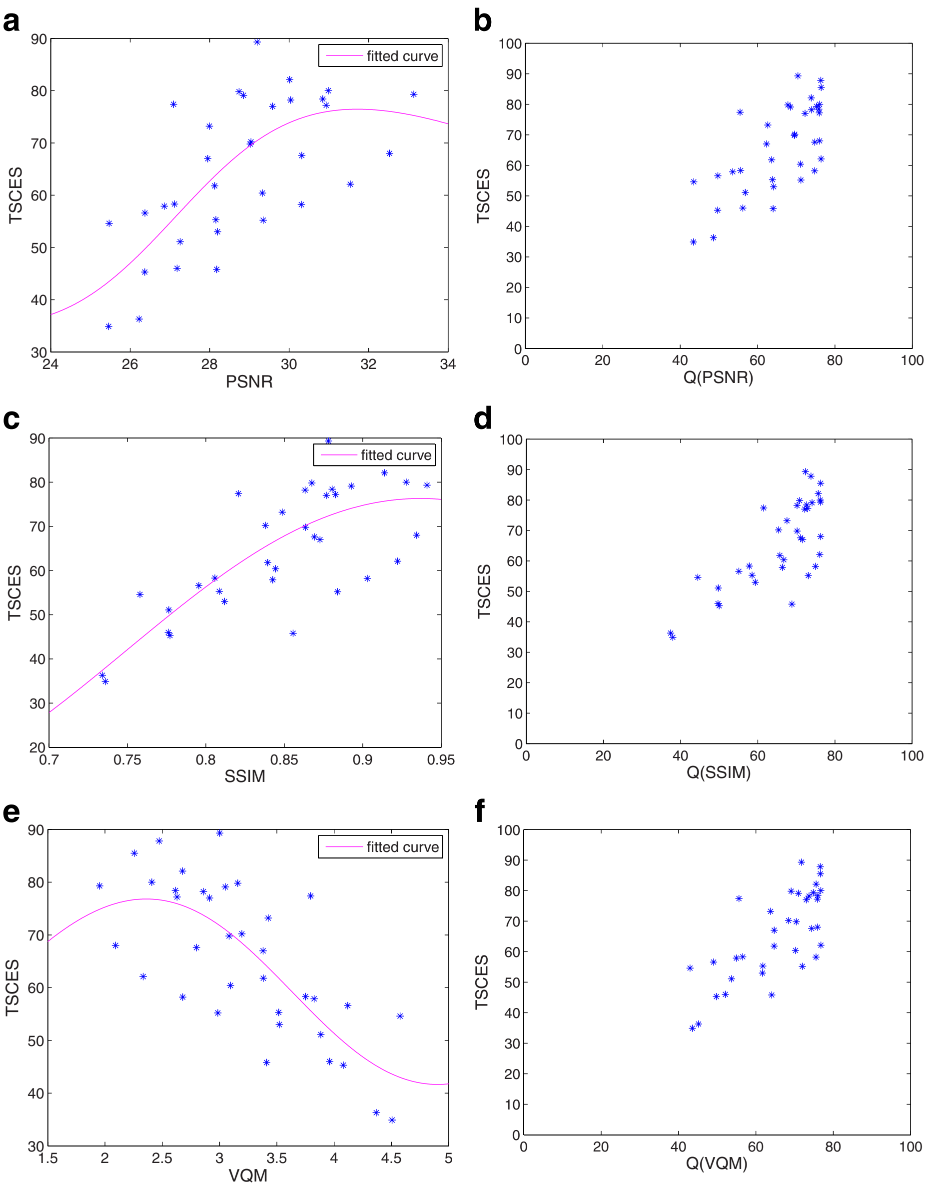


Fig. 6 Comparison of all 36 degraded video sequences and objective quality measures with TSCES, before and after nonlinear regression: (**a**) PSNR (before) - TSCES, (**b**) PSNR (after) - TSCES, (**c**) SSIM (before) - TSCES, (**d**) SSIM (after) - TSCES, (**e**) VQM (before) - TSCES, (**f**) VQM (after) - TSCES

regression), PSNR is somewhat worse and SSIM gives worst results for all three performance measures. For Parkjoy video sequence, SSIM gives best results in all performance measures, slightly worse is PSNR and worst is VQM. However, in this video sequence all objective measures give very similar results, only with minor differences.

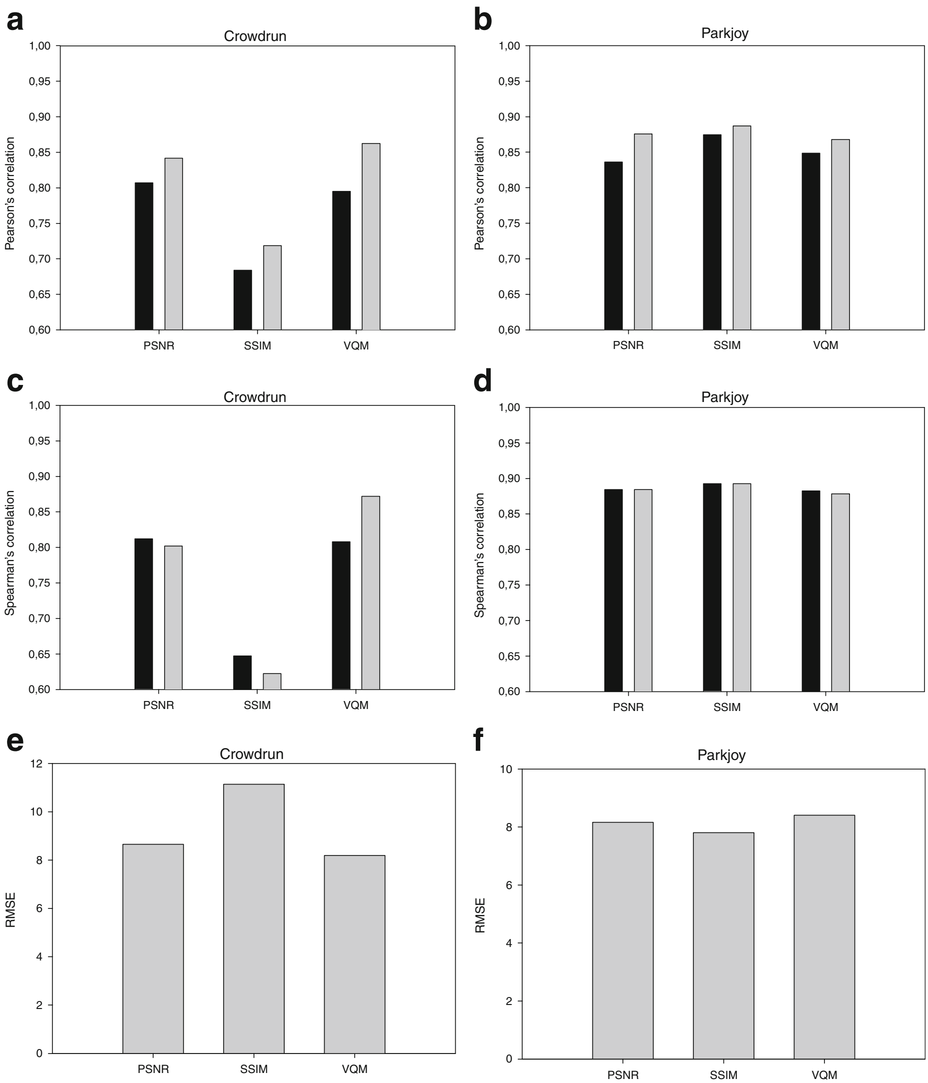
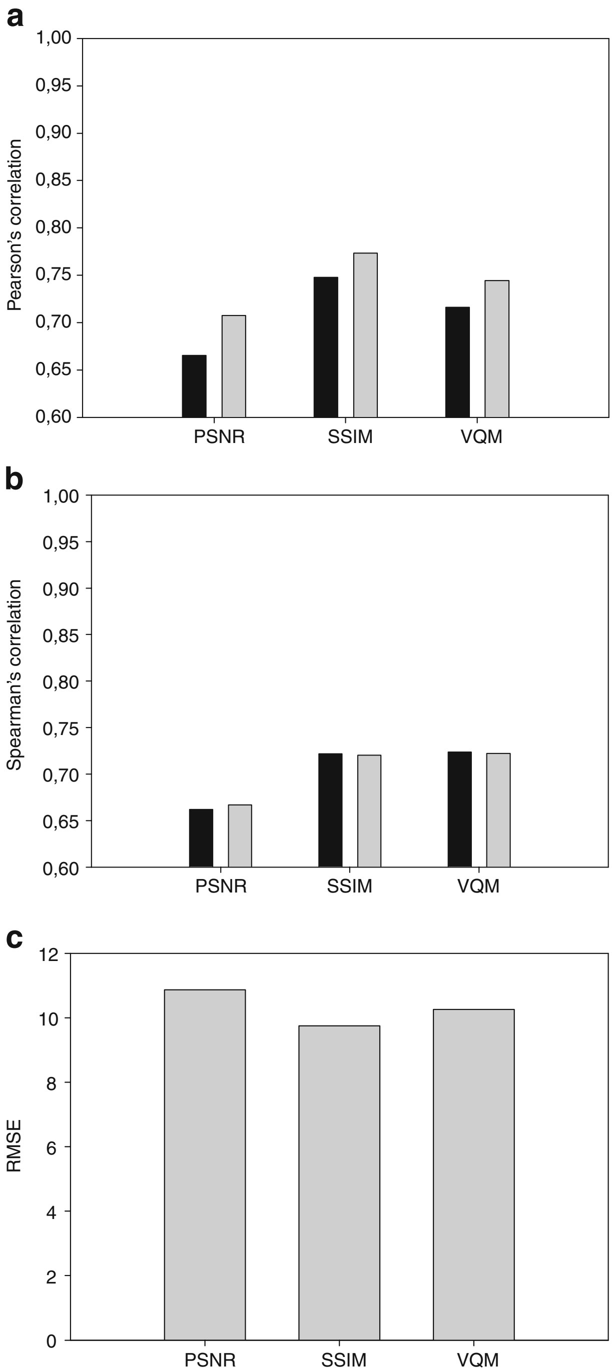


Fig. 7 Comparison of RMSE, Spearman's and Pearson's correlation coefficient, for Crowdrun [(**a**), (**c**) and (**e**)] and Parkjoy [(**b**), (**d**) and (**f**)] video sequences. Black bars denote coefficients that are calculated before nonlinear regression; gray bars denote coefficients that are calculated after nonlinear regression

When comparing objective results of both test sequences (Fig. 8), SSIM gives best results for Pearson's correlation and RMSE, while VQM gives best Spearman's correlation. PSNR gives worst results for all three performance measures.

It can be concluded that for the same video sequence, even in different resolution formats, PSNR and VQM give reasonable good results (SSIM was good measure for only 1 sequence), while for more sequences combined together it is better to use SSIM or VQM measure.

Fig. 8 Comparison of RMSE, Spearman's and Pearson's correlation coefficient, for both video sequences. When calculating correlation coefficients, black bars denote coefficients that are calculated before nonlinear regression; gray bars denote coefficients that are calculated after nonlinear regression



7 Conclusion

In this paper we compared several video quality measures with subjective measure. Firstly, subjective results of different HDTV formats show that progressive scanning should be considered rather than interlaced for all future HDTV emissions. For the same type of video sequence, PSNR and VQM measures yield reasonable good performance results in comparison with subjective testing. If we want to compare more different video sequences, it will be better to use SSIM or VQM measure (although all three measures did not give nearly good results to be compared with subjective measurement). In the future, we will make our own subjective tests under controlled conditions, in order to have exact subjective grades of each compressed video sequence. Also, by using more video sequences it will be possible to conclude more precisely which objective measure has higher correlation with subjective testings.

Acknowledgement The work described in this paper was conducted under the research projects: "Picture Quality Management in Digital Video Broadcasting" (036-0361630-1635), and "Intelligent Image Features Extraction in Knowledge Discovery Systems" (036-0982560-1643), supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

References

1. Dennis JE Jr (1977) Nonlinear Least-Squares. In: Jacobs D (eds) State of the Art in Numerical Analysis. Academic Press, pp 269–312
2. Dumic E, Mustra M, Grgic S, Gvozden G (2009) Image Quality of 4:2:2 and 4:2:0 Chroma Subsampling Formats. Proceedings of the 51st International Symposium ELMAR-2009, Zadar, Croatia 28–30:19–24
3. EBU Technical Recommendation R112 - 2004 (2004) EBU Statement on HDTV Standards, EBU Technical Committee, October
4. Grgic S, Grgic M, Mrak M (2004) Reliability of Objective Picture Quality Measures. J Electr Eng 55(1–2): 3–10
5. Hauke J, Kossowski T (2007) Comparison of Values of Pearson's and Spearman's Correlation Coefficient on the Same Sets of Data. Proceedings of the MAT TRIAD 2007 Conference, Bedlewo, Poland, 22–24 March
6. Hoffmann H, Itagaki T, Wood D, Bock A (2006) Studies on the Bit Rate Requirements for a HDTV Format with 1920x1080 Pixel Resolution, Progressive Scanning at 50 Hz Frame Rate Targeting Large Flat Panel Displays. IEEE Trans Broadcast 52(4):420–434
7. Hoffmann H, Itagaki T, Wood D, Hinz T, Wiegand T (2008) A Novel Method for Subjective Picture Quality Assessment and Further Studies of HDTV Formats. IEEE Trans Broadcast 54(1):1–13
8. Hubel DH (1988) Eye, Brain and Vision. Scientific American Library, New York
9. ISO/IEC 14496-10 (2003) Information Technology-Advanced Video Coding: A Codec for Video Signals (also called AVC and is technically identical to the ITU-T H.264), Standard ISO/IEC, Geneva
10. ITU-R BT.500-11 (2003) Methodology for the Subjective Assessment of the Quality of Television Pictures", International Telecommunication Union, Geneva, Technical Report ITU-R BT. 500-11
11. ITU-R BT.709-5 (2002) Parameter Values for the HDTV Standards for Production and International Programme Exchange, International Telecommunication Union, Geneva
12. Levenberg K (1944) A Method for the Solution of Certain Problems in Least Squares. Q Appl Math 2:164–168
13. Marquardt DW (1963) An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J Soc Ind Appl Math 11(2):431–441
14. Moré JJ, Sorensen DC (1983) Computing a Trust Region Step. SIAM J Sci Statist Comput 4(3):553–572
15. Uncompressed video sequences - source material available at:ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat/
16. Sheikh HR, Sabir MF, Bovik AC (2006) A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms. IEEE Trans Image Process 15(11):3440–3451
17. SMPTE 274M-2005 (2005) Television - 1920x1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates, Society of Motion Pictures and Television Engineers, New York

18. SMPTE 296M-2001 (2001) Progressive Image Sample Structure—Analog and Digital Representation and Analog Interface, Society of Motion Pictures and Television Engineers, New York
19. SMPTE 421M-2006 (2006) VC-1 Compressed Video Bitstream Format and Decoding Process, Society of Motion Pictures and Television Engineers, New York
20. sgi2yuv program available at: http://www.ldv.ei.tum.de/Members/tobias/videotools/sgi2yuv.zip
21. Suehring K. H.264/AVC Software Coordination, JM Software Download available at: http://iphome.hhi.de/suehring/tml/
22. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans Image Process 13(4):600–612
23. Wiegand T, Sullivan GJ, Bjontegaard G, Luthra A (2003) Overview of the H.264/AVC Video Coding Standard. IEEE Trans Circuits Syst Video Technol 13(7):560–576
24. Winkler S, Mohandas P (2008) The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics. IEEE Trans Broadcast 54(3):660–668
25. x264 Encoder available at: http://x264.nl/
26. Xiao F (2000) DCT-based Video Quality Evaluation. Final Project for EE392J, winter

Emil Dumic received the B.Sc. in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 2007. He is currently a Ph.D. student at the Department of Wireless Communications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. His research interests include image interpolation, wavelet transforms, and digital satellite television.

Sonja Grgic received the B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 1989, 1992 and 1996, respectively. She is currently full Professor at the Department of Wireless Communications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. Her research interests include television signal transmission and distribution, picture quality assessment and wavelet image compression. She has participated in 10 domestic and international scientific projects. Currently, she is a project leader of the research project “Picture Quality Management in Digital Video Broadcasting” financed by the Ministry of Science, Education and Sports of the Republic of Croatia. She is author or co-author of 16 papers published in scientific journals, more than 120 papers published in conference proceedings of international scientific conferences as well as of 15 reviewed studies and expert works. She was editor of 6 international conference proceedings. She is a member of IEEE and SMPTE.

Mislav Grgic received B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 1997, 1998 and 2000, respectively. He is currently an Associate Professor at the Department of Wireless Communications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. He participated in 5 scientific projects financed by the Ministry of Science, Education and Sports of the Republic of Croatia and 3 international COST projects of the European Commission. Currently, he is a project leader of the research project: “Intelligent Image Features Extraction in Knowledge Discovery Systems” financed by the Ministry of Science, Education and Sports of the Republic of Croatia and a member of the Management Committee of the European project COST Action IC0604. He published more than 110 papers in books, journals and conference proceedings in the area of image and video compression, content-based image retrieval, face recognition and digital mammography (computer-aided detection and diagnosis of breast cancer). Prof. Grgic is a senior member of IEEE.