23

Digital Video Image Quality

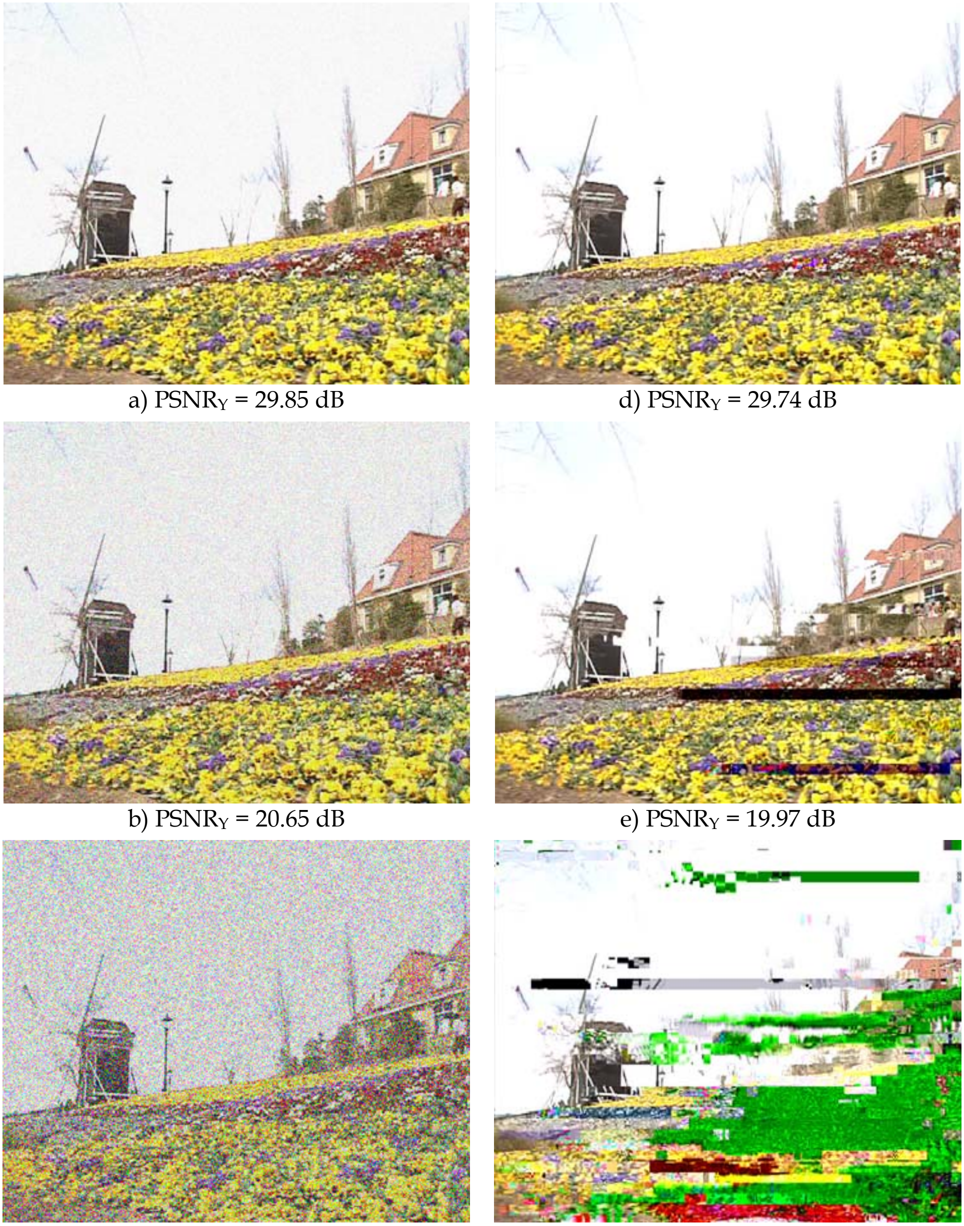
Tomáš Kratochvíl and Martin Slanina
Department of Radio Electronics, Brno University of Technology
Czech Republic

1. Introduction

The video image quality in digital television systems is subject to quite different effects and influences than that in analog television systems. There are mainly two sources which can disturb the digital video image quality and which can cause visible degradation of video image quality. These are the source coding and the related compression and transmission link from the modulator to the receiver. The perturbation, noise, transmission channel influence or transmission distortions can cause an increase of channel bit error rate and due to the error protection, e.g. FEC (Forward Error Correction) in DVB (Digital Video Broadcasting) (Fisher, 2008), included in the signal, most of the bit errors can be repaired up. It leads to QEF (Quasi Error Free) transmission conditions, and the errors are not noticeable in the video image. If the transmission channel is too noisy, the transmission totally breaks down. This situation is well known as the "fall of the cliff", or simply "cliff off". The linear or nonlinear distortion has no direct effect on the video image, but in an extreme case it can also lead to a breakdown. No matter if the picture quality is good, bad or indifferent, it needs to be evaluated differently and detected by different means in DTV (Digital Television) and DVB systems than in ATV (Analog Television). An example of the video image quality in ATV and DTV system is shown in Fig. 1.

There are several dimensions of digital video image quality evaluation, generally splitted into the subjective and objective methods. The subjective evaluation is a result of human observers providing their opinion on the video image quality. The objective evaluation is performed with the aid of instrumentation, calibrated scales and mathematical algorithms. Direct measurements are performed with the video images (picture quality measurement) and indirect measurements are made processing specially designed test signals in the same manner as the pictures (signal quality measurement) (Tektronix, 1997). The test video image sequences are used for both direct measurements, subjective and objective, but in a compressed digital video image system, they can not be used for the compression encoder/decoder part of the system because a comparison of the codec influence on the common test scenes and natural scenes is not possible. To specify, evaluate and compare digital video systems with video image artifacts caused by compression or transmission, the quality of the digital video and image presented to the observer has to be determined. Video image quality is inherently subjective and is affected by many subjective factors. It could be difficult to obtain accurate measures and results. Measuring video image quality using objective criteria results is an accurate and repeatable evaluation, but there is still no general objective evaluation. It should naturally cover the subjective experience of a human observer and performance of a video display and viewing conditions (Richardson, 2002).

Source: Digital Video, Book edited by: Floriano De Rango,
ISBN 978-953-7619-70-1, pp. 500, February 2010, INTECH, Croatia, downloaded from SCIYO.COM



a) $PSNR_Y$ = 29.85 dB

d) $PSNR_Y$ = 29.74 dB

b) $PSNR_Y$ = 20.65 dB

e) $PSNR_Y$ = 19.97 dB

c) $PSNR_Y$ = 13.58 dB

f) $PSNR_Y$ = 12.85 dB

Fig. 1. Examples of typical distortion aftifacts in the video transmission over noisy channels. Uncompressed video sequence a) low level of noise, b) average level of noise, c) high level of noise. MPEG-2 algorithm compressed video sequence d) low bit-error rate and QEF transmission, e) average bit error rate and blockiness, f) high bit-error rate and cliff off effect.

2. Subjective test procedures

The test procedures for subjective test are defined especially in ITU-R recommendation BT.500-11 (ITU, 2001). The most popular is evaluation by the DSCQS (Double Stimulus Continuous Quality Scale) method. An assessor evaluates a pair of digital video image short sequences, called A and B, one after another. Then he is asked to give a score to A and B sequences on a continuous scale. The scale is divided into five intervals of the subjective quality scores reaching from excellent through good, fair, poor to bad quality. The impairment scale related to the mentioned five intervals is in Tab. 1.

Score	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Table 1. Score and related subjective quality evaluation criteria

Fig. 2. Subjective methods and continuous quality rating forms for DSCQS and SSCQE

In a typical test session, the assessor is presented a series of sequence pairs and is asked to evaluate a grade of each individual pair, as shown in Fig. 2. In each pair of digital video image sequences, one is unimpaired (so-called reference sequence) and the other is the same sequence modified by a compression algorithm or process under test, e.g. video image compression or transmission. A typical example is a video coding system as shown in Fig. 3. In this case, the original sequence is compared to the same video image sequence which was subject to encoding and decoding. The order of the evaluated sequences is randomized during the evaluation, so the assessor does not know which sequence (original or impaired), he is currently evaluating. This prevents the assessor from predicting and prejudging the results. At the end of the test, the scores are normalized and the result is a score that indicates the relative video image quality of the impaired and reference sequences. The resulting score is denoted to as MOS (Mean Opinion Score).

The DSCQS test can be used as a realistic measure of subjective digital video image quality. In its application it must be considered that it suffers from several practical problems. The evaluation can vary significantly and depends on selection of the assessors and also on the characteristics of the video image sequence under test. This variation can be reduced by

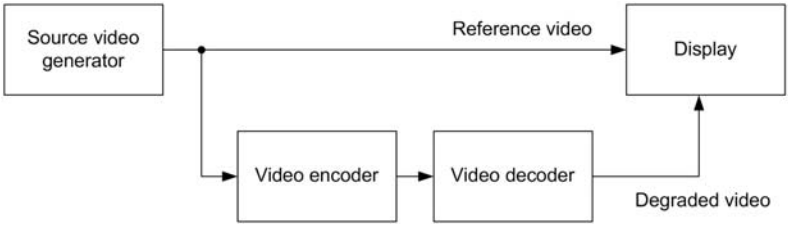


Fig. 3. General arrangement for the DSCQS method evaluation

repeating the test with several sequences and several assessors. An expert assessor, who is familiar with the digital video image artifacts caused by compression, can give a biased score. That is why the non-expert assessors are preferred. Additionally, non-expert assessors can quickly learn to recognize compression artifacts in the digital video image sequences. Subjective tests are also influenced by the viewing conditions. A test carried out in a comfortable environment will be evaluated with a higher score than the same test carried out in a less comfortable setting. It has been proved that the "recency effect" (Richardson, 2002) means that the assessor's opinion is significantly biased by the last few seconds of a video image sequence. The quality of the last section will strongly influence the score of a whole longer sequence. That is why the subjective test is really suitable only for short video image sequences.

The second popular subjective test defined by ITU-R BT.500 recommendation is the SSCQE (Single Stimulus Continuous Quality Evaluation) method (ITU, 2001). In this subjective test the assessor evaluates video image quality without the reference sequence. Since the SSCQE deliberately dispenses the reference sequence, this method can be used more widely in practice. In this method a group of test persons assesses the processed video image sequence only and evaluates the score again from excellent to bad, which also provides a video image quality profile over time (see Fig. 2).

The advantages of subjective testing are in obtaining valid results for conventional and compressed television systems and evaluation of scalar MOS that works over a wide range of still and motion picture applications. The disadvantages of subjective testing are in a wide variety of possible methods which must be considered for the test. Many observers must be selected and it is very time consuming in case the procedure respects all the requirements.

3. Objective tests

Objective test methods are based on automated, computational approach. Depending on the original video image sequence, the objective test results are not always correlated with the impression of quality in a subjective observation. The degree of correlation to subjective results can be considered a benchmark of subjective tests.

The first choice when selecting a metric for full-reference quality evaluation is usually the peak signal-to-noise ratio (PSNR). For video sequences, it can be easily computed as (Winkler, 2005)

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{m^2}{\mathrm{MSE}}, \ [\mathrm{dB}] \tag{1}$$

where m is the number of values by which a pixel can be represented (e.g. m = 255 for 8-bit luma samples) and MSE is the mean squared error, computed as

$$\mathrm{MSE} = \frac{1}{\mathrm{T}\cdot\mathrm{M}\cdot\mathrm{N}}\sum_{k=1}^{\mathrm{T}}\sum_{i=1}^{\mathrm{M}}\sum_{j=1}^{\mathrm{N}}\left[f(k,i,j)-\tilde{f}(k,i,j)\right]^2 \cdot \tag{2}$$

The constants M, N, T are the horizontal and vertical dimensions in pixels and the number of frames (fields), respectively, f and $\tilde{f}$ are the sample values (luma or chroma) of the degraded and the reference video sequence, respectively. The peak signal-to-noise ratio is very simple and easy to implement, but its disadvantage is in poor correlation with subjective tests.

A great effort has been devoted to developing objective video quality metrics in recent years. An ideal objective quality metric should closely simulate the results of subjective tests. Many approaches have been proposed to achieve this, with different success. To select the metrics suitable for real applications, two phases of tests were performed by the Video Quality Experts Group (VQEG) in 2000 and 2003, respectively (VQEG, 2000; VQEG, 2003).

In the first testing phase ten proposed quality evaluation algorithms were considered, and the correlation of their results with subjective scores obtained for video sequences with different characteristics (different scene contents) and subject to different quality degradations (compressed with H.263, MPEG-2 encoders using different settings, Betacam with drop-out) was examined. All the tested video sequences were in standard definition, considering both 625- and 525-line television systems. The first phase of testing was completed only with a limited success - the performance of the proposed quality evaluation algorithms was very close to the performance of PSNR. As a result, none of the tested algorithms was proposed by the VQEG to be included in an ITU Recommendation.

Another testing (Phase II) was realized by the VQEG a couple of years later, considering a set of six proposed quality evaluation algorithms. The testing procedures were very close to those performed in Phase I. However, out of the six proposed algorithms, four were selected to be included in an ITU Recommendation, published in 2004 as ITU-R Recommendation BT.1683. In the following subsections, the principles of the four standardized algorithms will be briefly described (ITU, 2004).

3.1 The BTFR algorithm

This metric was designed by the British Telecom, United Kingdom, and is denoted to as the BTFR algorithm (British Telecom full-reference automatic video quality assessment tool).

The algorithm computes several measures comparing the degraded video and the reference video, to finally combine the measures together to get a quality prediction. A simple diagram of the algorithm operation is shown in Fig. 4. The preprocessing block in the diagram consists of several steps. It includes format conversion, cropping, offset and matching operations. These are performed on the luma (Y) as well as chroma (U, V) components of both the degraded and the reference video sequences. Matching operations are also included in the preprocessing block. They consist in finding the best match for blocks within each degraded field from a buffer of neighboring reference fields. A matched video sequence is then used instead of the reference sequence in some of the following analyses:

- *PSNR analysis* – a PSNR calculation is performed using the degraded and matched reference sequences.

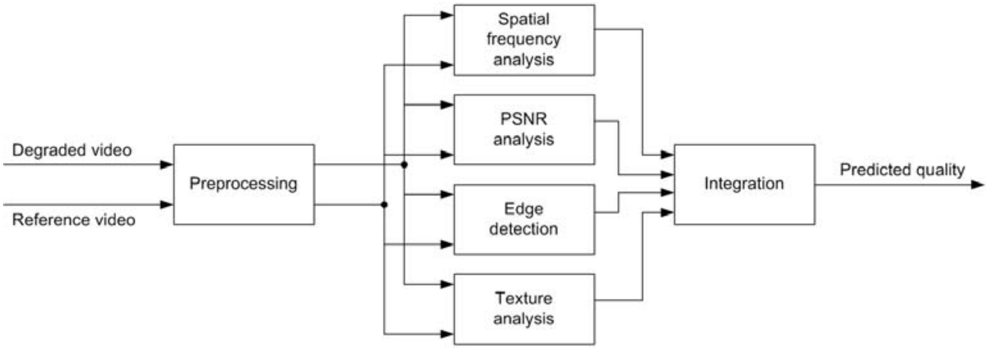


Fig. 4. Operation of the BTFR algorithm

- *Spatial frequency analysis* – a pyramid transformation of the degraded and matched reference sequences is performed. The differences between pyramid arrays are then calculated using SNR.
- *Edge analysis* – an edge detector is used to create edge maps of the matched reference and degraded video sequences. The total numbers of edge-marked pixels are then calculated for both edge maps.
- *Texture analysis* – the texture properties are measured by recording the number of turning-points in the luma signal along horizontal picture lines.

Finally, all the parameters gained from the analyses are put together in the Integration block to form the final quality measure. The integration is nothing but computing a linear combination of the parameters, with specified weights and offset.

3.2 The EPSNR algorithm

The second metric described in Rec. BT.1683 was designed in cooperation of the Yonsei University, Radio Research Laboratory and SK Telecom, Republic of Korea. It is based on the fact that human observers are very sensitive to degradations around edges – when the edges are blurred, the subjective scores are likely to be worse. Additionally, many compression algorithms tend to produce artifacts around the edges. The metric computes a value called Edge PSNR (EPSNR) and uses it as a quality measure after post-processing. A block diagram of the metric is shown in Fig. 5.

Using an arbitrary edge detection algorithm, edge areas are located in the first step using the reference video sequence. For each field (or frame when processing progressive video), an edge mask image is created – the algorithm operates on a field-by-field basis. Then differences between the reference and the degraded video fields are computed, based on simple mean squared error evaluation, limited to the edge areas. Finally, PSNR of the edge areas (EPSNR) is computed from the mean squared error.

In the final phase, post-processing is applied to the EPSR value of the actual field, taking into account the following:

- For high PSNR values, the EPSNR overestimates perceptual quality. The solution is in piecewise linear scaling (reduction) of EPSNR values over 35.
- If the degraded video is severely blurred (the number of edges detected in the degraded video is significantly lower than the number of edges in the reference video), the EPSNR is reduced.
- Scaling is performed at the end to reach the range of outputs between 0 and 1.

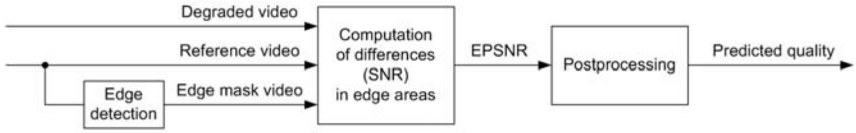


Fig. 5. Metric based on Edge PSNR

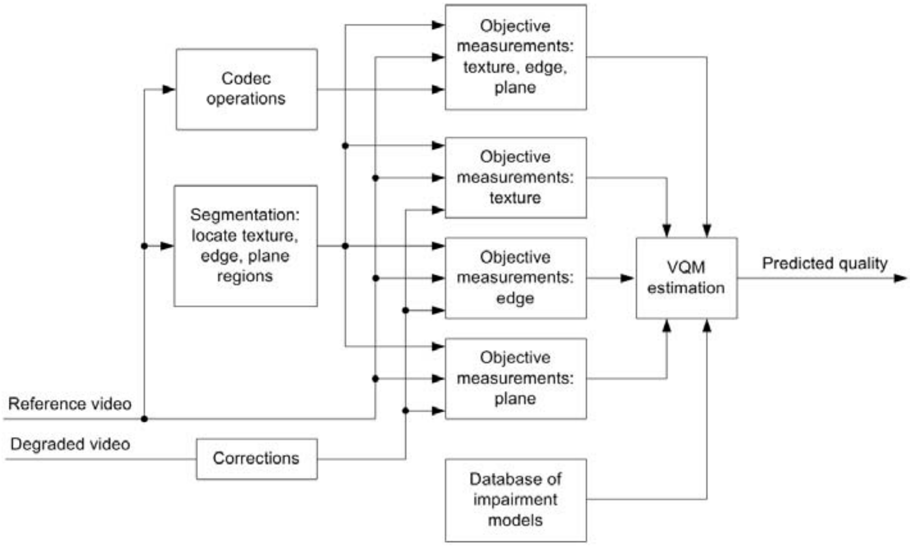


Fig. 6. Simplified block diagram of the IES metric

3.3 The IES algorithm

This metric is designed by the Center of Telecommunications Research and Development, Brazil (CPqD). It is denoted to as the IES system (image evaluation based on segmentation). Its simplified block diagram is shown in Fig. 6.

Based on analysis of the reference video sequence, each field (frame) is segmented into texture, edge, and plane regions. Each of these regions is then processed separately: objective quality measures are computed for the reference and the degraded video fields after correction (offset, gain). As each of the objective measures is performed for the luma as well as both chroma components, it results in nine parameters.

An interesting approach is in including a block called Codec operations. This is nothing but encoding and decoding the reference video sequence using two different codecs in parallel: the MPEG-2 4:2:0 and the MPEG-1 CIF algorithms (with fixed encoder settings). The resulting video sequences together with the reference video are then subject to the same objective analysis as the reference – corrected degraded video sequence pair from the input. Thus, for two codecs, three region types, and three components, we get the total of eighteen additional objective parameters for the quality estimation.

The IES system uses a database of impairment models for scenes different from the reference video scenes in order to estimate the subjective quality of the degraded video sequence. The database consists of information about sequences with different degrees of motion and

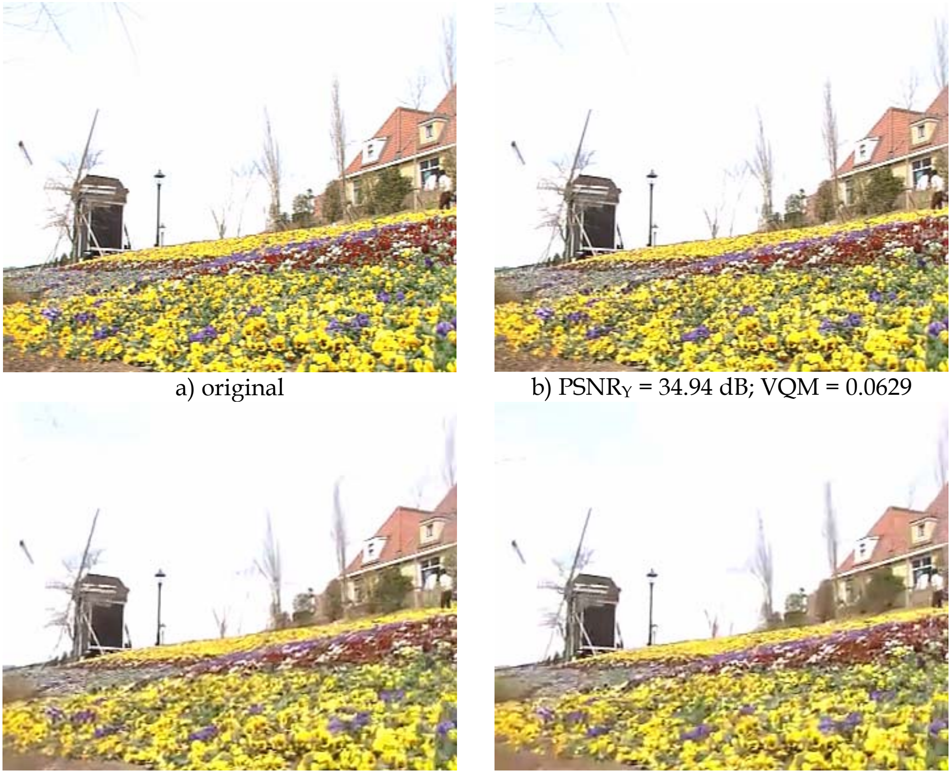
detail and different context. Together with spatial and temporal attributes extracted from the reference video sequence, it forms another input for the quality estimation algorithm.
Finally, subjective impairment levels are estimated from the objective parameters and the resulting predicted quality measure is achieved through a linear combination of the estimated impairment levels.

3.4 The VQM algorithm

The fourth and the last metric described in Rec. 1683 was designed by the National Telecommunication and Information Administration / Institute for Telecommunication Science, US. The acronym VQM used by the authors stands for Video Quality Metric. The metric is quite complex, involves preprocessing (matching) operations as well as a thorough evaluation of video sequence properties. An important feature is that this metric natively operates not only on separate fields (frames), but breaks the video sequences into S-T (spatial-temporal) sub-regions, including a block of pixels in several consequent fields.
The metric can also be called a reduced-reference metric, as it extracts a certain information (quality feature) from the reference and the degraded video sequences, and then forms a quality measure based on the extracted information only – just several values. The quality features include information on spatial gradients of the video scenes, chrominance information, contrast information, temporal information, etc. Their proper combination gives the metric output value. Typically, one output value is calculated for one video clip of 5 – 15 seconds in length.
The Fig. 7 shows frame of a video sequence compressed with the H.264/AVC algorithm using different settings. An original uncompressed sequence is shown in Fig. 7a). The sequence in Fig. 7b) is compressed with high bit rate, fine quantization, and the resulting PSNR is almost 35 dB. The output value of the VQM metric is almost zero, which means there will be probably no noticeable difference from the original. Indeed, both pictures look identical. Now look at the pictures in Fig. 7c) and Fig. 7d). Even though the PSNR computed for the whole sequence (100 frames) has almost the same values, the VQM differs considerably. By taking a close look at the pictures, especially on the bush in front of the house, much more blur is visible on the picture in Fig. 7d). This proves that the different degradation was not captured by the peak signal-to-noise ratio, but the VQM metric exhibits quite different values showing that the quality in the bottom right picture in Fig. 7d) is worse. The PSNR is computed only in the luminance channel, which has the highest impact on the perceived quality. The range of the VQM values is from zero to one, and the best quality with no degradation is represented by a zero.

4. Objective tests with no reference

The no-reference video quality assessment metrics cannot rely on any information about the original material. What information is then available at the receiver side and can be used for measurement? Usually, no-reference metrics use some a-priori information about the processing system. For example, a usual DVB-T broadcasting system using MPEG-2 source coding is known to have the block artifacts as the most annoying impairment (Fischer, 2008). Tracking these artifacts down in the video image may provide enough information to judge the overall quality. In the following text, metrics for still image quality evaluation will be considered as well as those for video sequences only. In fact, most of the still image metrics can be used for video sequences when applied on each video frame.



a) original

b) $PSNR_Y$ = 34.94 dB; VQM = 0.0629

c) $PSNR_Y$ = 23.59; VQM = 0.3953

d) $PSNR_Y$ = 23.23; VQM = 0.4713

Fig. 7. Frame of a compressed video sequence using the H.264 algorithm. PSNR and according VQM video image quality evaluation results are shown.

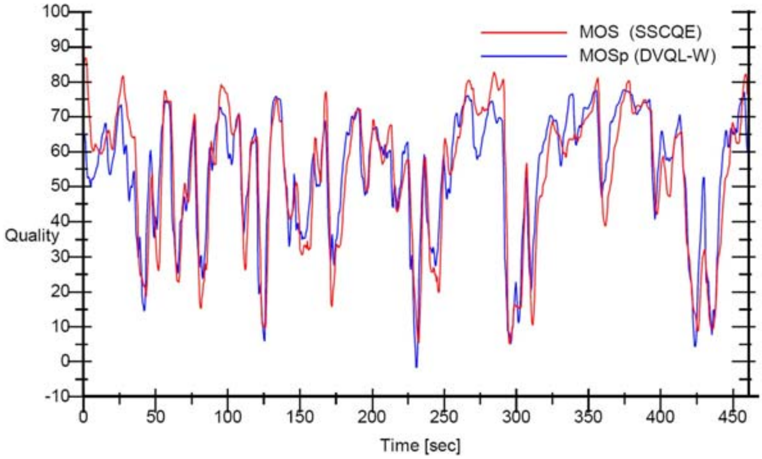
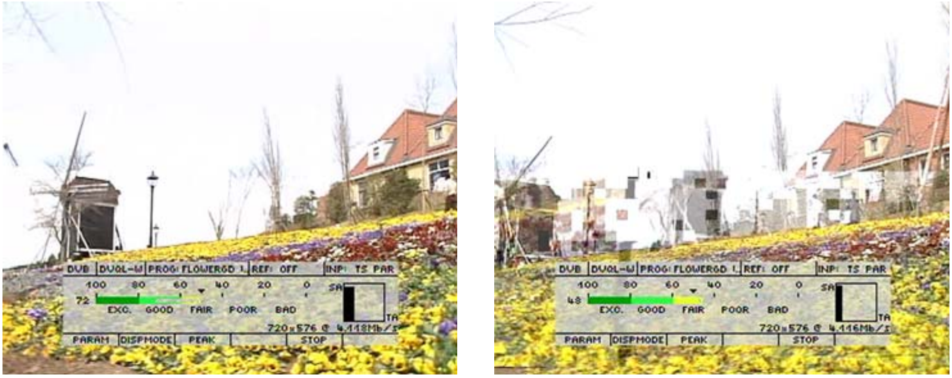


Fig. 8. Comparison of subjective and objective picture quality (Lauterjung, 1998)



a) b)

Fig. 9. Example of a real-time objective picture quality analysis using DVQ analyzer. The metric DVQL-W is used to demonstrate the achieved video image quality in one frame of the a) reference video, b) degraded video with blockiness and cliff-off effect present.

4.1 No-reference analysis using pixel values

The block artifact detection approach is used in DVQ analyzer - video quality measurement equipment supplied by Rohde & Schwarz and is briefly described in (Fischer, 2008). The principle of the method is in assumption that block artifacts create a regular grid with constant distances. Neighboring pixel differences are computed for the whole image and averaged in such manner that only 16 values remain (since MPEG-2 is supposed to create 16 x 16 blocks). If the average pixel value difference is significantly larger on block boundaries, a statement can be made that block artifacts are present in the image.

To bring the objective test results closer to the subjectively perceived quality, other quantities in the moving picture are also taken into consideration. These are spatial and temporal activities (Fischer, 2008). The spatial activity is a measure of the existence of fine structures in the video image and temporal activity is a measure of change and movement in successive frames. Both activities can render the blocking structure invisible or mask them. Such artifacts in the video image are then simply not seen by the human eye.

If masking is incorporated, the DVQL-W (Digital Video Quality Level – Weighted) metric applied in DVQ analyzer delivers a prediction of the MOS. With the masking included, the algorithm shows an excellent correlation with subjective assessment results as it is shown in the Fig. 8. The results of the subjective evaluation were obtained by the SSCQE method. The compiled test sequence consisted of 11 well-known test sequences such as “Flowergarden”, etc. The data rates for the sequences varied between 1 MBit/s and 9 MBit/s. From the subjective assessment about 1000 measurement values were obtained. Their scaling factor was re-based and a fixed delay of 1 second was introduced. With this optimization, an overall correlation of more than 94 % was achieved (Lauterjung, 1998).

An example of a real-time measurement using DVQ analyzer is shown in Fig. 9 and numerical results are in the Fig 10. The DVQL-W metric evaluates blocking structure in the video image of a selected DTV program in an MPEG-2 TS. It is obvious from Fig. 9 that the quality decreases with the blockiness in the video. The temporal and spatial activity and evaluation in the luminance and chrominance video channels are considered (Fisher, 2008).

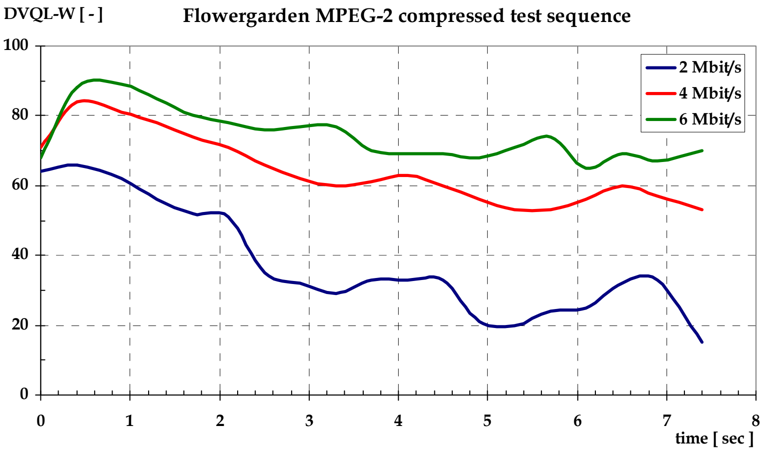


Fig. 10. Video image quality analysis of the „Flowergarden" test sequence. The DVQL-W metric and according quality varies with the MPEG-2 compressed test sequence bitrate.

In (Pan et al., 2004), a no-reference algorithm was presented capable of detecting block artifacts in a block-by-block manner and, as an extension, detects flatness of the image. A no-reference block and blur detection approach was introduced in (Horita et al., 2004), designed to measure quality of JPEG and JPEG2000 images. Another no-reference algorithm for block artifact detection was described in (Wang et al., 2000). Another common distortion, blur, can also be used for quality evaluation. Of course, depending on the characteristics of the processing system (whether or not the system is likely to introduce blur). An interesting metric was presented in (Ong et al., 2003). A very similar approach is also used in (Marziliano et al, 2002). In principle, these metrics analyze how steep the changes in pixel values within the line are. The main difference is that (Ong et al., 2003) analyzes not only the horizontal direction, but measures blur in four directions instead.

An interesting no-reference approach was used in (Tong et al., 2005), using a learning algorithm to assess the overall quality of an image. The metric uses pixel values of a decoded picture, which was subject to JPEG or JPEG2000 compression.

4.2 No-reference analysis using transform coefficients and encoded stream values

In (Sheikh et al., 2005), a metric was presented for JPEG2000 compressed static images. The JPEG2000 standard uses wavelet transform. The authors analyze the wavelet coefficients to gain a quality measure. An observation was made that in natural images, these coefficients have some characteristic properties. If the wavelet coefficients do not behave in a desired manner, quality degradation can be expected. However, this metric is only applicable for wavelet transform compressed images, and thus not applicable for any of the wide-spread present-day video compression standards. Anyway, coefficient analysis for video sequences is also possible.

In (Gastaldo et al., 2002), such analysis was performed for MPEG-2 compressed video sequences. First of all, a statistical distribution analysis was performed to say which of the features available in the MPEG-2 transport stream may be used for evaluation. Over twenty features were then used to feed an artificial neural network for learning and consequently for quality evaluation. For this metric, a correlation as high as 0.85 was achieved by the authors. A different approach was published in (Eden, 2007), where the author computes PSNR values of a H.264/AVC using transform coefficient and quantization parameter values, which means computation can be done on the encoded bit stream only.

4.3 No-reference metric developed at the Brno University of Technology

A no-reference quality metric was recently developed at the Brno University of Technology and published in (Slanina & Říčný, 2008). The metric operates on a compressed bit stream conforming to the H.264 / AVC standard. The idea is based on the fact that the encoder can adaptively select the sizes of blocks to be coded, and the coarseness of quantization of residual transform coefficients. A very simple artificial neural network is then used to process the input parameters, represented as ratios of the block sizes used by the encoder, the quantization parameter, and information on the quality of the reference frames for the inter predicted (using motion compensation) frames. The metric is not supposed to output values simulating subjective tests – the artificial neural network is trained to simulate PSNR values for a given compressed video sequence without reference.
The attainable correlation of the metric with real PSNR values is above 0.95. This value is somewhat lower than the correlation achieved in (Eden, 2007). Anyway, the algorithm is designed in such manner that it can be easily changed to predict different values than just the PSNR. The authors are currently working on predicting output values of the standardized full-reference algorithms. So far, it turns out that to achieve satisfactory results, the number of parameters extracted from the bit stream needs to be increased (the bit stream carries other low cost information, such as the bit rate, gop format, etc.).

5. Conclusion

Measuring video image quality is difficult and very often not precise. There are many factors that can affect the results and their interpretation. The advantages of subjective testing are in obtaining valid results for conventional and compressed television systems and the possibility of evaluating scalar MOS over a wide range of still and motion video image applications. Their disadvantages are in a wide variety of possible methods and tests to be considered, the high number of observers required and in high time demands. There are many objective testing approaches. It can be stated that the algorithms with video image feature analysis correlate better with subjective results than just simple pixel-based methods. A combination of different measurements and features gives the best results and correlation between subjective and objective scores but it is hardly technology independent.

6. Acknowledgement

This work was supported by the Research program of Brno University of Technology no. MSM0021630513, "Electronic Communication Systems and New Generation Technology (ELKOM)" and the research project of the Czech Science Foundation no. 102/08/P295, "Analysis and Simulation of the Transmission Distortions of the Digital Television DVB-T/H". The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement no. 230126.

7. References

Fisher, W. (2008). *Digital Video and Audio Broadcasting Technology. A Practical Engineering Guide*. 2nd edition. Springer, ISBN978-3-540-76357-4, Berlin

Eden, A. (2007). A no-reference estimation of the coding PSNR for H.264-coded sequences, *IEEE Transactions on Consumer Electronics*, Vol. 53, No. 2, May 2007, pp. 667-674, ISSN 0098-3063

Gastaldo, P.; Zunino, R. & Rovetta, S. (2002). Objective assessment of MPEG-2 video quality, *Journal of Electronic Imaging,* Vol. 11, No. 3, July 2002, pp. 365-374, ISSN 1017-9909

Horita, Y.; Arata, S. & Murai, T. (2004). No-reference image quality assessment for JPEG/JPEG2000 coding, *Proceedings of XII. European Signal Processing Conference EUSIPCO-2004,* Vol. 2, pp. 1301-1304, ISBN 3-200-00148-8, Vienna, September 2004, Vienna University of Technology

ITU (2001). *Methodology for the subjective assessment of the quality of television pictures.* Recommendation ITU-R BT.500-11. 2001.

ITU (2004). *Objective Perceptual Video Quality Measurement Techniques for Standard Definition Digital Broadcast Television in the Presence of a Full Reference*. Recommendation ITU-R BT.1683. 2004.

Lauterjung, J. (1998). *Picture Quality Measurement*. Rohde & Schwarz GmbH & Co KG. Reprint of paper done at IBC Sept. 1998, Amsterdam, 1998. [online] http://www2.rohde-schwarz.com/file_3609/PQM.pdf

Marziliano, P.; Dufaux, F.; Winkler, S. & Ebrahimi, T. (2002). A no-reference perceptual blur metric, *Proceedings of the International Conference on Image Processing,* Vol. 3, pp. 57-60, ISBN 0-7803-7622-6, Rochester, NY, September 2002

Ong. E. P. et al. (2003). A No-reference quality metric for measuring image blur, *Proceedings of the Seventh International Symposium on Signal Processing and Applications*, Vol. 1, pp. 469-472, ISBN 0-7803-7946-2, Paris, July 2003, IEEE

Pan, F et al. (2004). A Locally-Adaptive Algorithm for Measuring Blocking Artifacts in Images and Videos, *Proceedings of the 2004 International Symposium on Circuits and Systems ISCAS '04,* Vol. 3, pp. 925-928, ISBN 0-7803-8251-X, Vancouver, May 2004, IEEE Circuits and Systems Society, Piscataway

Richardson, I.E.G. (2002). *Video Codec Design. Developing Image and Video Compression Systems*, Willey, ISBN 0-471-48553-5, Chirchester

Sheikh, H. R.; Bovik, A. C. & Cormack, L. K. (2005). No-reference quality assessment using natural scene statistics: JPEG2000, *IEEE Transactions on Image Processing*, Vol. 14, No. 11, November 2005, pp. 1918-1927, ISSN 1057-7149

Slanina, M. & Říčný, V. (2008). Estimating PSNR in high definition H.264/AVC video sequences using artificial neural network, *Radioengineering,* Vol. 17, No. 3, September 2008, ISSN 1210-2512 [online] http://www.radioeng.cz/fulltexts/2008/08_03_103_108.pdf

Tektronix, Inc. (1997). *A guide to Picture Quality Measurements for Modern TV Systems.* Tektronix Inc., 1997. [online] http://www.tek.com/Measurement/App_Notes/PicQuality/25W_11419_0.pdf

Tong, H. et al. (2005). Learning no-reference quality metric by examples, *Proceedings of the 11th International Multimedia Modelling Conference MMM '05,* pp. 247-254, ISSN 1550-5502, Melbourne, 2005, IEEE Computer Society

VQEG (2000). *Final Report from the Video Quality Experts Group on thr Validation of Objective Models of Video Quality Assessment*. ITU, 2000. [online] http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI/COM-80E_final_report.pdf

VQEG (2003). *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. ITU, 2003. [online]. http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final _Report.pdf

Wang, Z.; Bovik, A. C. & Evans, B. L. (2000). Blind measurement of blocking artifacts in images, *Proceedings of 2000 International Conference on Image Processing*, Vol. 3, pp. 981-984, ISBN 0-7803-6297-7, Vancouver, September 2000

Winkler, S. (2005). *DigitalVideo Quality: Vision Models and Metrics.* Wiley, ISBN 0-470-02404-6, Chichester