UTILIZATION OF MATLAB FOR PICTURE QUALITY EVALUATION

Tomáš Kratochvíl, Pavel Šimíček

Institute of Radio Electronics, Brno University of Technology

Abstract

The contribution deals with the measurement and picture quality evaluation in area of DTV (Digital Television). Subjective and objective picture quality evaluation methods and measures overview are presented in the paper and test pictures selection criteria are outlined. Then the simplified PQES (Picture Quality Evaluation System) features applied in Matlab are presented including the laboratory workplace utilization and example of quality evaluation. Finally the comparison of subjective and objective quality evaluation approach is concluded.

1 Introduction

There are several dimensions of PQE (Picture Quality Evaluation) generally splitted into subjective and objective measurements. *Subjective measurements* are the result of human observers providing their opinion of the video quality and *objective measurements* are performed with the aid of instrumentation, calibrated scales and mathematical algorithms. The test pictures and scenes are used for both direct measurements (subjective and objective picture quality). The objective quality is well established and can detect the picture quality distortions that are too small for the human to see. The subjective quality based on group of observers is very time consuming and exhausting. Very interesting possibility of testing appears from the Human Visual System (HVS), but there is still no universal solution [1], [2].

2 Picture Quality Evaluation

Subjective measurement - according to ITU-R BT.500 recommendation has been used with a set of standard methods. A number of observers are selected, tested for their visual capabilities, shown a series of test scenes and asked to score the quality of the scenes. Perception based on subjective evaluation is quantified by *MOS* (Mean Opinion Score). There is a list of common subjective measurement method, the first and second is commonly used in PQE analysis of DTV systems [3]:

- *DSCQS* (Double Stimulus Continuous Quality Scale) – multiple reference scene pairs with the reference and degraded scenes randomly, scoring is on a continuous quality scale from excellent to bad where each scene of the pair is separately rated but in reference to the other scene,
- *SSCQE* (Single Stimulus Continuous Quality Evaluation) – continuous program evaluated over long period, scoring is a distribution of the amount of time, this method relates well to the time variant qualities of compressed television system,
- *DSIS* (Double Stimulus Impairment Scale) – multiple reference scene, degraded scene pair, reference is always first, scoring is commonly knows as a five point scale from imperceptible to very annoying,
- *SS* (Single Stimulus) – no repetition of test scenes, different scoring methods (adjectival – 5 half-grade scale, numerical – 11-grade scale, non-categorical – continuous scale),
- *SSMR* (Single Stimulus Multiple Repetition) – additional multiple times repetition,
- *SC* (Stimulus Comparison) – differences between scene pairs are scored with adjectival 7-grades from -3 to +3 scale or non-categorical.

Objective measurement - of picture quality is of high importance to compare algorithms designed for multimedia signal coding or signal enhancement. Perception based on objective evaluation presents *PQS* (Picture Quality Scale) [4]. In many optimization approaches, the aim is the minimization of the squared error, which is related to the shortest Euclidian distance between sample magnitudes. Any objective measurement system must have good correlation with subjective results for the same DTV system and test scenes. The methods are based on comparison of reference with degraded picture and may be divided into two categories - feature extraction and picture differencing.

The picture differencing measures [5] contain evaluation of

- *MSE* (Mean Square Error) $$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j) - f'(i,j)\right]^2 , \quad (1)$$

- *PSNR* (Peak Signal to Noise Ratio) $$PSNR = 10\log\frac{\left(2^n - 1\right)^2}{MSE} , \quad (2)$$

- *AD* (Average Difference) $$AD = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j) - f'(i,j)\right], \quad (3)$$

- *MD* (Maximum Difference) $$MD = Max\left(\left|f(i,j) - f'(i,j)\right|\right), \quad (4)$$

- *NK* (Normalized Cross-Correlation) $$NK = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j)\cdot f'(i,j)\right]}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j)\right]^2} , \quad (5)$$

- *MAE* (Mean Absolute Error) $$MAE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left|f(i,j) - f'(i,j)\right| , \quad (6)$$

- *NAE* (Normalized Absolute Error) $$NAE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j) - f'(i,j)\right]}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j)\right]^2} , \quad (7)$$

- *SC* (Structural Content) $$SC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f(i,j)\right]^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[f'(i,j)\right]^2} \quad (8)$$

etc., where original picture $f(i, j)$ and degraded picture $f'(i, j)$ have resolution M x N pixels.

3 Test Pictures Selection Criteria

Depending on the original picture, the test value and evaluation are not always correlated with the impression of quality of a subjective observation. There is spatial measure, the *SFM* (Spatial Frequency Measure) indicates the overall activity level in a picture, defined by *R* (Row Frequency) and *C* (Column Frequency). In spectral domain the *SAM* (Spectral Activity Measure) is defined as a measure of picture predictability. The evaluation deals with the DFT coefficients of picture and *SAM* has the dynamic range of [1, infinity). Higher values of *SAM* imply higher predictability and active images are difficult to code [5]. Mentioned measures are determined by

- *SFM* (Spatial Frequency Measure) $$SFM = \sqrt{R^2 + C^2} , \quad (9)$$

  where $$R = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=2}^{N}\left[f(i,j) - f(i,j-1)\right]^2} , \quad (10)$$

  and $$C = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=2}^{N}\left[f(i,j) - f(i-1,j)\right]^2} . \quad (11)$$

- *SAM* (Spectral Activity Measure) $$SAM = \frac{\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left|F(i,j)\right|^2}{\left[\prod_{i=0}^{M-1}\prod_{j=0}^{N-1}\left|F(i,j)\right|^2\right]^2} , \quad (12)$$

  where $F(i,j)$ is DFT spectrum and its coefficients.

4 Evaluation System Applied in Matlab

Simplified evaluation system called PQES (Picture Quality Evaluation System) was implemented in Matlab (see Fig. 1) [6]. It contains GUI (Grafical User Interface) and it uses dialogs for reference and degraded picture selection and evaluation of its *SAM* and *SFM* values. The objective PQE contains the feature extraction and picture differencing approach and directly enumerates measures that were presented in the paper. The subjective PQE deals with applied algorithms of ITU-R BT.500 recommendation and respects an order of test pictures. Used methods of subjective PQE were mentioned in the paper and the result of the evaluation is enumeration of *MOS* and *PQS* measures. The objective PQE is simplified in a way that only one person evaluates the quality. The laboratory workplace for PQES is aimed into fast experimental evaluation and education of students.

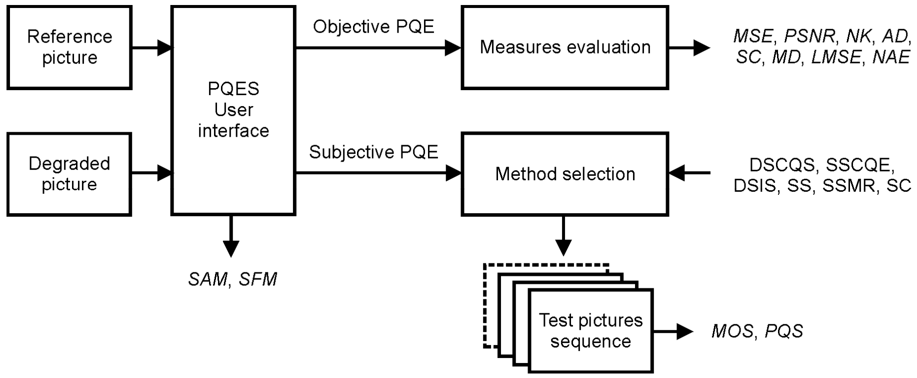

Figure 1: PQES block diagram with methods and evaluated measures.

Presented example deals with the non-compressed static image (see Fig. 2a) in analog TV resolution (720 x 576 pixels / 8 bps) that was I-frame MPEG-2 compressed in according to DVB standard (see Fig. 2b). The Table 1 presents some experimental results of objective PQES evaluation that are available immediately after the test pictures selection and Fig. 3 presents the PQES GUI.

Figure 2: Test image for objective PQE a) reference picture (non-compressed static picture), b) degraded picture (DCT based compressed I-frame MPEG-2).

Table 1: Experimental results of PQES (only luminance component) applied on pictures in Fig.2.

Objective measure	*Result*	*Objective measure*	*Result*
Mean Square Error	169.32	Normalized Cross-Correlation	0.9910
Mean Absolute Error	9.2759	Average Difference	0.2092
Normalized MSE	0.0598	Maximum Difference	88.247
Normalized Absolute Error	0.0763	Structural Content	1.0082
Signal to Noise Ratio	12.227	Spatial Frequency Measure	29.564
Peak Signal to Noise Ratio	25.843	Spectral Activity Measure	195.10

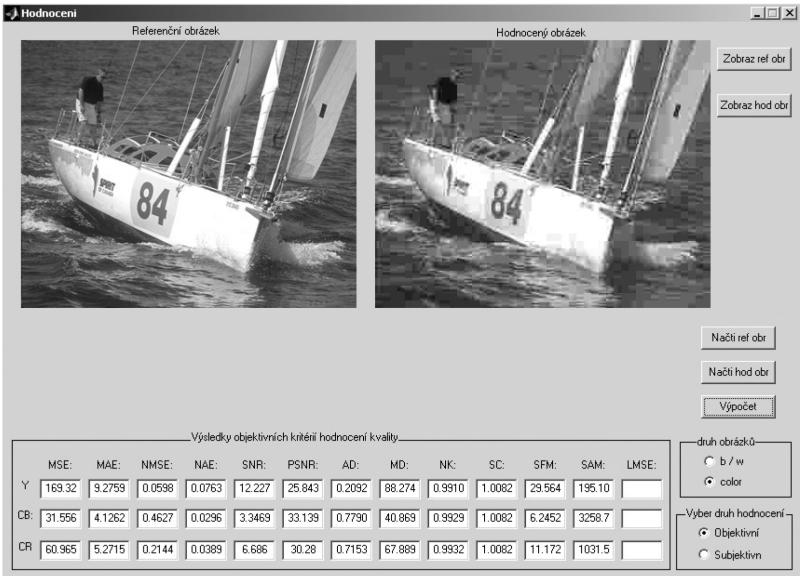

Figure 3: PQES Graphical User Interface – objective measures results.

5 Conclusion

Advantages of subjective measurement is obtaining of valid results for conventional and compressed DTV systems and evaluation of scalar *MOS* (Mean Opinion Score) 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 and test must be considered, many observers must be selected and it is very time consuming. The objective measurements with picture differencing correlates best with subjective results. Combination of these methods gives the best results and correlation between subjective and objective scores, but it is still not technology independent [7].

Acknowledgement

This contribution was supported by the grant project of Research programme of Brno University of Technology MSM0021630513 "*Electronic Communication Systems and New Generation Technology (ELKOM)*" and Czech Science Foundation GACR no. 102/05/2054 "*Qualitative Aspects of Audiovisual Information Processing in Multimedia Systems*".

References

[1] U. Riemers. *DVB, The Family of International Standards for Digital Video Broadcasting, 2nd edition*. Springer-Verlag, 2004.
[2] W. Fisher. *Digital Television, A Practical Guide for Engineers*. Springer-Verlag, 2004.
[3] Rec. ITU-R BT.500-11. *Methodology for the subjective assessment of the quality of television pictures*. ITU-R, 1974-2002.
[4] *A guide to Picture Quality Measurements for Modern TV Systems*. Tektronix Inc.,1997.
[5] M. Mrak, S. Grgić, M. Grgić. *Picture Quality Measures in Image Compression Systems*. In Proceedings of the Eurocon 2003 conference, p. 233-237, Ljubljana, Slovenia, 2003.
[6] P. Šimíček. *Image quality evaluation in television and video technique area*. Bachelor thesis, BUT Brno, 2005 (in Czech).
[7] *Comparing Objective and Subjective Picture Quality Measurements*. Tektronix, 1998.

Tomáš Kratochvíl, M.Sc., Pavel Šimíček, Bc.
Institute of Radio Electronics Brno University of Technology, Purkyňova 118, 612 00 Brno,
Tel: +420 541 149 113, Fax: +420 541 149 244, E-mail: kratot@feec.vutbr.cz