# OBJECTIVE AND SUBJECTIVE MEASURUMENTS FOR VIDEO COMPRESSION SYSTEM

Sonja Grgić, Mislav Grgić, Branka Zovko-Cihlar
Faculty of Electrical Engineering and Computing
Unska 3, HR-10000 Zagreb, Croatia

**Abstract**: *In this paper the extent to which the quality of video signal is degraded by process of compression and decompression is considered. Traditional objective video signal quality measurements based on static test signals are not able to characterise the picture degradation due to compression and decompression. These measurements are indirect and resulting distortions of test signals determine video-processing characteristics. Signal compression algorithms reduce bit rates by removing redundant or irrelevant information. Redundancy is a function of picture content. Therefore objective measurements have limited effectiveness in predicting the quality of compressed images as seen by the observes. There is no general method known for this type of evaluation except subjective evaluation. Subjective measurement is geared toward properties of human visual system and permits an integrated evaluation of picture quality. But, subjective measurement takes a large amount of time and resources and results are not always repeatable. In this paper we present an overview of the methodologies for measurement of picture quality using objective and subjective evaluation procedure. The results of objective testing using a set of test signals in compressed video system are given. The paper outlines a need for new objective picture quality measurement method providing good correlation to subjective measurements.*

***KEYWORDS: Video Signal Compression, Test Signals, Picture Quality Measurement***

## INTRODUCTION

At one time, it was considered that digital video signal transmissions should always produce noiseless and distortion-free pictures. This ideal standard means small or no difference in quality between the source pictures produced using the studio production equipment and the reconstructed pictures as observed by the viewer.

The ideal standard condition can only be achieved if the available bandwidth is large enough to cope with worst case pictures, which contain many details and high level of movement, even if they only occur rarely.

Today, this high ideal is relaxed and some degree of distortion is acceptable on certain rare scenes but on the majority of scenes high resolution should be achieved. The optimum trade-off between picture quality and bandwidth (bit rate) requirements demands detailed statistical analysis and subjective testing of the viewers tolerance to selective distortion which is determined by the level of the bit rate reduction. For example, digital television signal requires bit rate of 216 Mb/s for component 4:2:2 system as defined in ITU-R Recommendation BT-601, [1]. This level is far too high for a practical TV delivering system and multimedia application. Consequently, a certain amount of video signal compression is necessary to reduce the bit rate to a more practical level [2-6].

The main goal of digital video compression is to achieve the minimum possible distortion for a given coding rate or equivalently, to achieve a given acceptable level of distortion with the least possible coding rate, [9]. Compression systems not only remove redundant information but also modify the picture. The resulting picture may contain perceptible impairments. The level of impairments can be specified either by an objective measure such as signal to noise ratio or by a subjective measure such as mean opinion score (MOS), [7, 8].

## 1. OBJECTIVE MEASURES OF PICTURE QUALITY

Television evolved as an analog system and test procedure was optimised for analog parameters. The objective assessment of the performances of video system has conventionally been assessed by a set of test signals [11]. These signals have been developed and refined over a number of years to test linear and nonlinear system parameters initially for monochrome signal later for composite coded color signal (PAL, NTSC, SECAM). Traditional video signal quality measurements using static test waveforms do not characterise the picture degradation due to compression and decompression leading to the need for new objective measures of picture quality. The number of compression systems incorporates temporal prediction technique. In these systems objective assessment using static test signals give invalid results because static test signals do not explore any temporal characteristics of the system.

A standard objective measure of picture quality in digital video system is reconstruction error. Suppose that one has a system in which an input picture element block {x(n)}, n=0,1,...,N-1, is reproduced as {y(n)}, n=0,1,...,N-1. The reconstruction error r(n) is defined as the difference between x(n) and y(n), [9]

$$r(n) = x(n) - y(n) \tag{1}$$

The variances of x(n), y(n) and r(n) are $\sigma_x^2$, $\sigma_y^2$ and $\sigma_r^2$. In the special case of zero-means signals, variances are simply equal to respective mean square values measured over appropriate sequence length M:

$$\sigma_z^2 = \frac{1}{M}\sum_{n=1}^{M} z^2(n), \qquad z = x, y \text{ or } r \tag{2}$$

A standard objective measure of coded picture quality is the ratio of signal variance to reconstruction error variance (signal to noise ratio - S/N) usually expressed in decibels (dB)

$$\frac{S}{N}(dB) = 10\log_{10}\left(\frac{\sigma_x^2}{\sigma_r^2}\right) \tag{3}$$

When the input signal is R-bit discrete variable, the variance or energy can be replaced by the maximum input symbol energy $(2^R-1)^2$. For the common case of 8 bits per pel of input images, the peak signal to noise ratio can be defined as

$$\left(\frac{S}{N}\right)_P (dB) = 10\,log_{10}\left(\frac{255^2}{\sigma_r^2}\right) \tag{4}$$

Signal to noise ratio is not adequate as a perceptually meaningful measure of digitized picture quality because the reconstruction errors in general do not have the character of signal-independent additive noise and seriousness of the impairments cannot be measured by a simple power measurement. Small impairment of a picture can lead to a very large value of $\sigma_r^2$ and consequently a very small value of S/N, in spite of the fact that the perceived image quality can be very acceptable. For example, a slight spatial shift of an image causes a large numerical distortion but no visual distortion and oppositely, a small average distortion can result in a damaging visual artifact if all the errors are concentrated in a small but important region. In fact, in picture coding systems the truly definitive measure of coded signal quality is perceptual quality as measured by careful subjective experimentation.

## 2. CURRENT SUBJECTIVE QUALITY ASSESSMENT METHODS

Both ITU-R and ITU-T have been developing recommendations on subjective quality assessment methodologies. Originally, CCIR (subsequently names ITU-R) was basically addressing methodologies for evaluating audio and video quality for broadcasting and entertainment services, while CCITT (subsequently named ITU-T) was merely addressing methodologies for evaluating speech quality in telephony. However, the conventional ITU-R methodologies have not been sufficient to assess short extract of digitally coded video material because quality fluctuates widely depending on the scene content and impairments that may be short lived. Therefore, the conventional ITU-R methodologies have been changed and a new methodology suitable for picture quality evaluation in digital video systems was added.

The ITU-R, besides the Rec.BT.500, that provides the fundamental description of the subjective video quality assessment methods, has produced recommendations describing the complete procedures, based on the methods illustrated in recommendation BT.500-7 [14], but adapted to specific systems as High Definition Television - HDTV [15], and enhanced PAL and SECAM, [16].

The ITU-T, besides the recommendation P.80 [17], that describes the subjective speech quality assessment methods, has produced recommendations particularly suited for audio-visual communication services (e.g. videophone, videoconference, co-operative working etc.) as P.910 [18], P.920 [19] and P.930 [20]. P.910 presents method very close to those presented in ITU-R

Recommendation BT.500-7, but test procedures have been adapted to the case of low and medium bit-rates (up to 2 Mb/s). P.920 is specific for audio-visual communications and it is particularly suitable to evaluate the transmission delay. P.930 describes an adjustable video reference system that can be used to generate the reference conditions necessary to characterise the subjective picture quality of video produced by compressed digital video systems. So only this last recommendation is not specific for a class of application. Thus the general trend is to define test methods tailored on specific classes of services. This is because quality requirements can strongly change from service to service and the traditional subjective test methods are often not suited to provide representative evaluations.

## 3. OBJECTIVE ASSESSMENT USING TEST SIGNALS

In analog and full-bandwidth digital video system signal quality measurements using test waveforms can give good characterisation of picture quality. It can be true for the compressed video system, which reduce bit rates by removing intraframe spatial redundancy and do not use interframe temporal prediction. To illustrate the objective assessment using static test waveforms in compressed video system, we used the measurement system that is shown in Figure 1.

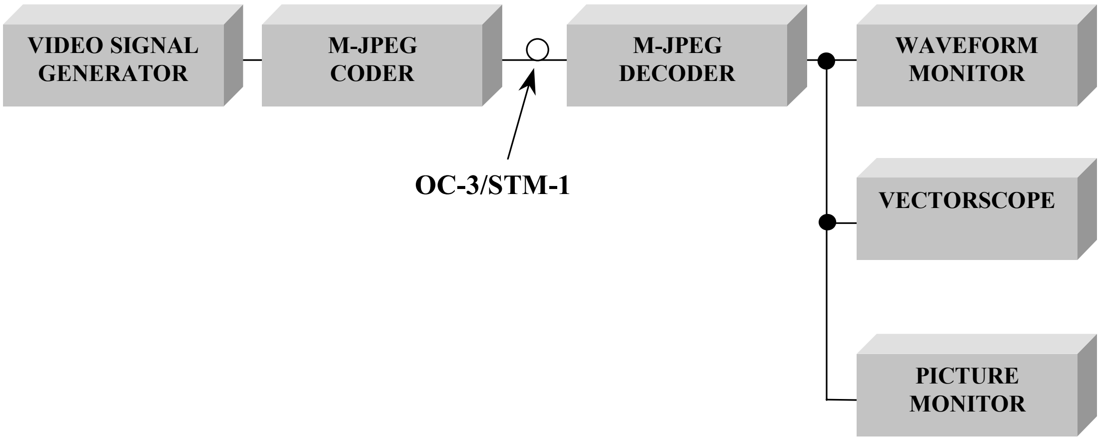
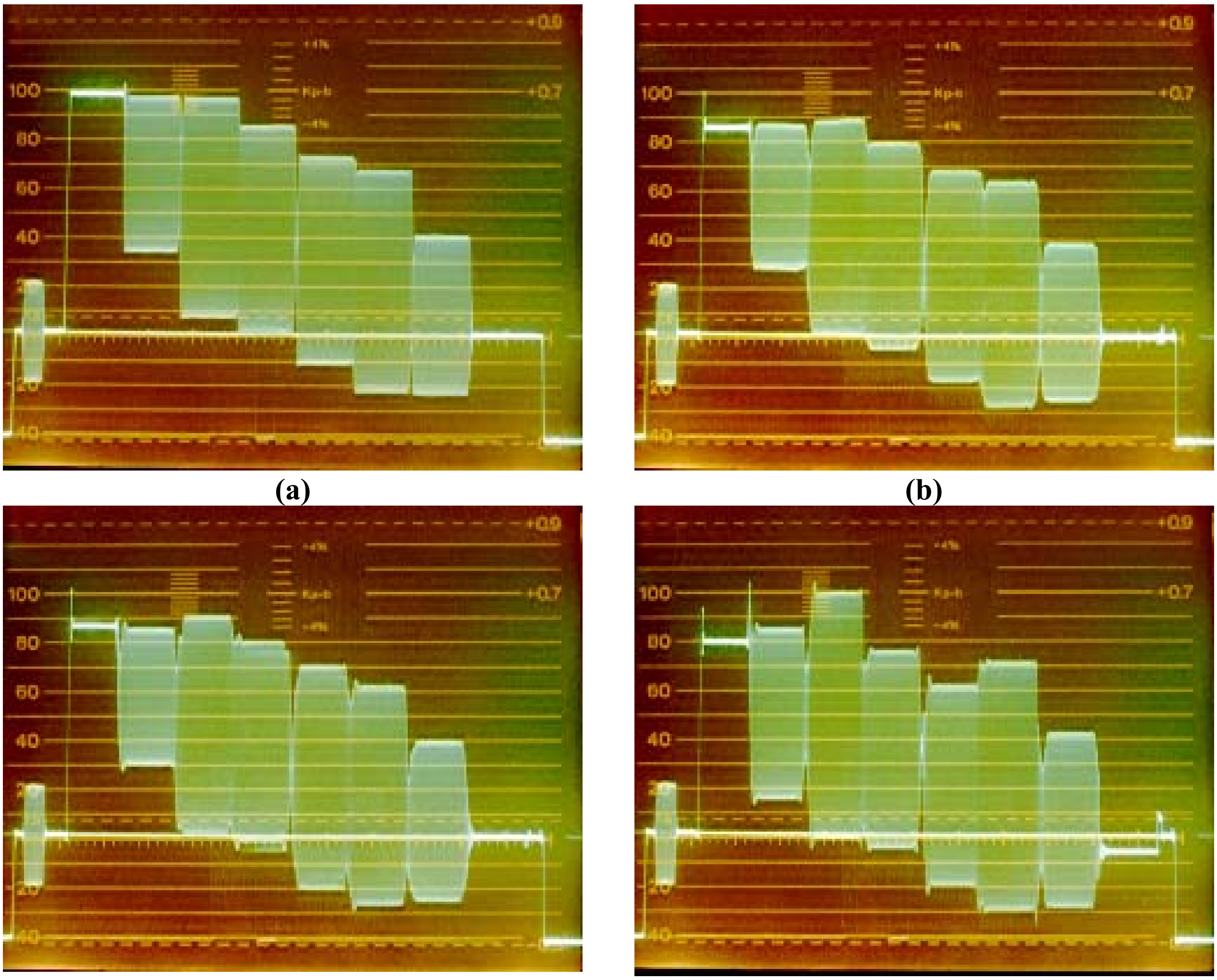


*Figure 1-The measurement system configuration*

The M-JPEG coder digitizes a full frame rate interlaced video from television signal generator for transmission over Asynchronous Transfer Mode (ATM) Network, [12]. It uses Motion-JPEG compression [2]. Transmit coder contains ATM interface to 155 Mb/s OC-3/STM-1 multimode fibre. Coder provides adaptive quantization factor (Q-factor) compression. Larger value of Q-factor means lower picture quality and higher level of compression. When the data rate approaches the channel capacity, the system tries to reduce the data rate by dropping the frames, causing the picture to appear jerky and broken. In this system, in anticipation of high load, the Q-factor (and compression factor) will be temporary increased reducing the quality of the frames so that they can be transported within the limit. This has a two-fold advantage:

- the video can be transmitted at lower data rates using less of the available network channel capacity which can be redeployed elsewhere
- the video, which contains extremely complex types of image, can be delivered without dropping frames but with lower picture quality.



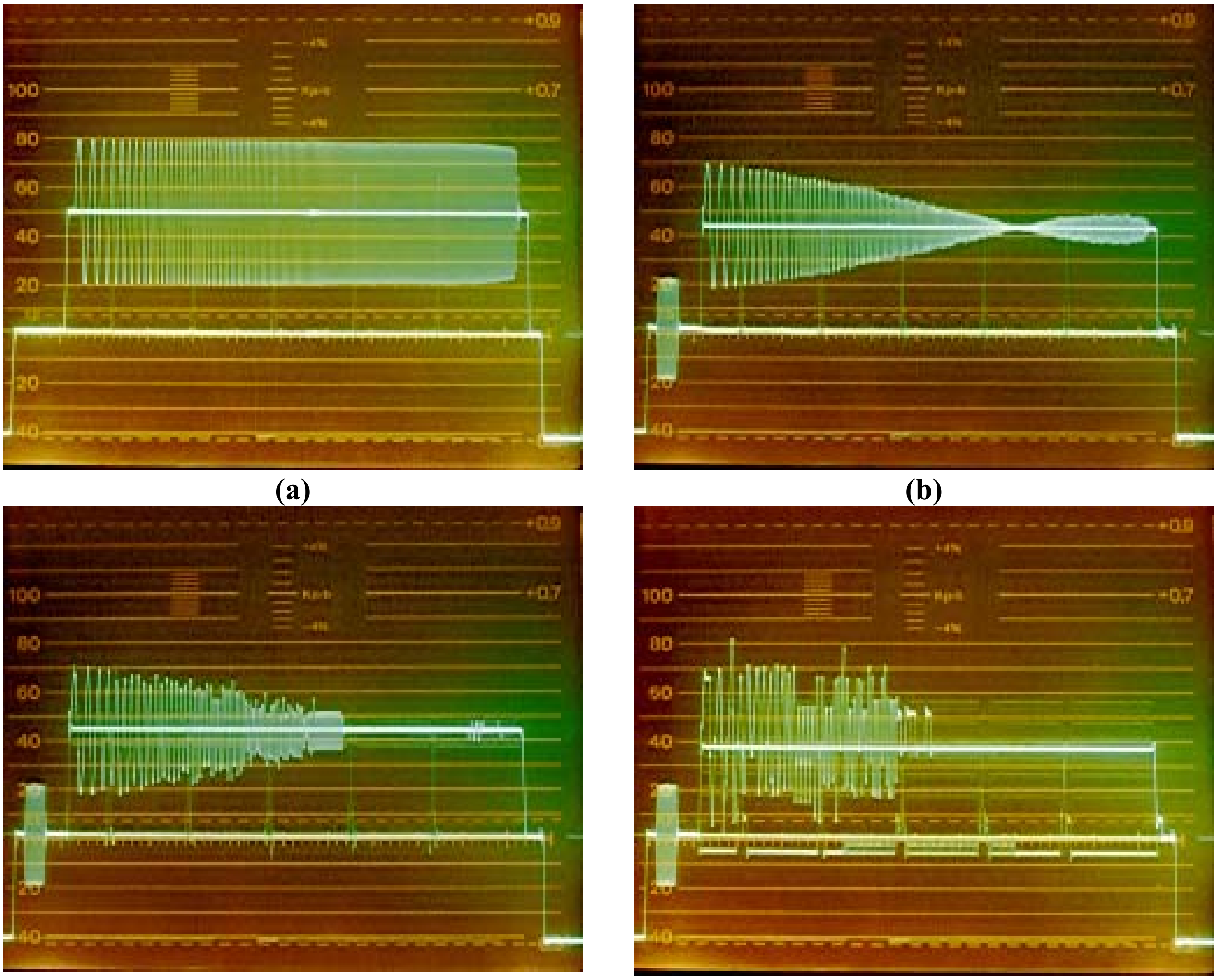
**(a)** **(b)**

**(c)** **(d)**

*Figure 2 - Waveforms for original color bar signal (a), and color bar signal after compression with different compression factors (b) Q-20, (c) Q-200, (d) Q-1024*

The video signal generator's output (see Fig.1) is connected to the input of the transmit coder. The OC-3/STM-1 output of the coder is connected to the input of transmission channel. The video output of the M-JPEG decoder is connected to the waveform monitor, vectorscope and video monitor. The video signal generator produces the test signals that are transported by transmission system and decoded video is observed on the measurement equipment. The test signals used in testing the system are: color bar, line sweep, multiburst and modulated ramp [11]. The different compression factors were used: Q-20 (high quality and low compression), Q-200 (average quality and moderate compression) and Q-1024 (low quality and high compression). The Q-factor determines quantizer scale factor and consequently, the number of quantized coefficients that are used in picture reconstruction. The camera is used to record the input waveform responses and output waveform responses. The difference between input and corresponding output waveforms is measure of picture quality. This difference increases by increasing the Q-factor and level of compression.

Color bar test signal was used for chrominance phase and amplitude measurements using vectorscope and waveform monitor (see Fig.2). For the compression factor Q-1024 the presence of large amount of distortion is indicated as phase shifts and amplitude deviation. Chrominance phase shifts result in a hue error while amplitude deviation result in color saturation error.

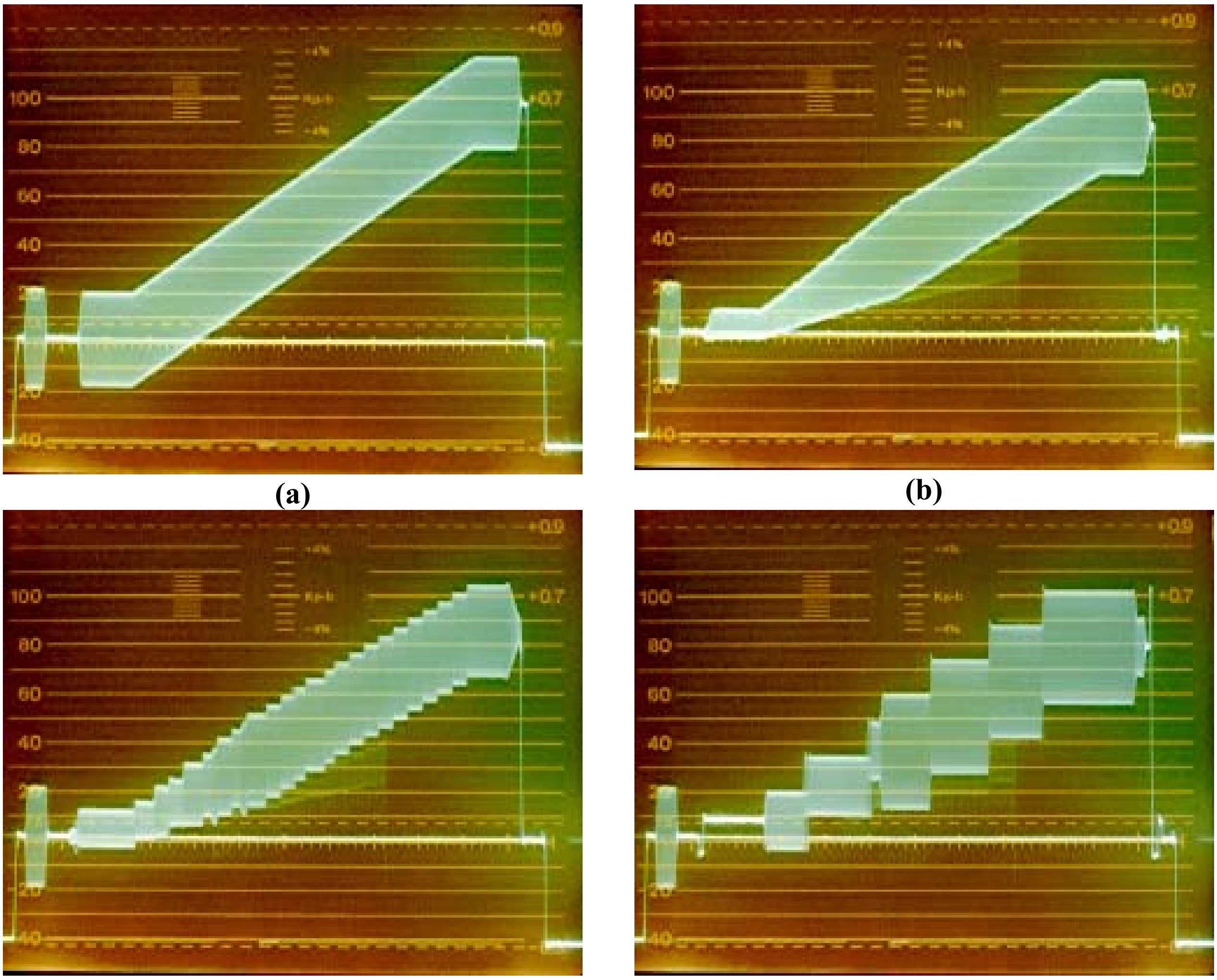


(a) (b)

(c) (d)

*Figure 3 - Waveforms for original line sweep signal (a), and line sweep signal after compression with different compression factors (b) Q-20, (c) Q-200, (d) Q-1024*

The sweep and multiburst test signals are used for frequency response measurements that evaluate the system's ability to transfer signal components of different frequencies without affecting their amplitudes. In a sweep signal the frequency of the signal sine wave is continuously increased (0.5-6 MHz) over the interval of a line (Fig.3(a)). Multiburst signal includes six packets of discrete frequencies that fall within the TV passband (0.5 to 5.8. MHz). These test signals evaluate system's amplitude response over the entire video spectrum or system's ability to transform signal components of different frequencies without affecting their amplitudes. Figure 3 shows input (a) and output (b, c, and d) waveforms for line sweep test signal. The difference between input and output signals can be noted and this difference increases by increasing the compression factor. High frequencies are compressed and lost. The distortions in the frequency response are represented on the picture monitor as loss of spatial resolution due to the coarse quantization.



**(a)** **(b)**

**(c)** **(d)**

*Figure 4 - Waveforms for original modulated ramp signal (a), and this signal after compression with different compression factors (b) Q-20, (c) Q-200, (d) Q-1024*

Modulated ramp test signal was used for differential gain measurements. It is test signal that contains uniform amplitude chrominance superimposed on different luminance levels. Differential phase distortion is present if a signal's chrominance phase is affected by luminance level. It is result of a system's inability to uniformly process the high frequency chrominance information at all luminance levels. Differential gain is present if chrominance gain is dependent on luminance levels. This type of error is shown in Figure 4. When differential gain is present, color saturation is not correctly reproduced. In M-JPEG coder high compression is achieved by neglecting high frequency DCT coefficients and luminance and chrominance are changed after the process of coding and decoding. Typical artifacts due to reducing the data rate following the DCT are blockiness. High frequency DCT coefficients, which are neglected, influence the picture quality and luminance and chrominance components are changed after the process of coding and decoding.

## 4. SUBJECTIVE EVALUATION PROCEDURE

To analyse the correlation between objective and subjective testing results, subjective evaluation of picture quality was performed. The experiment was based on the single-stimulus method in accordance with the ITU-R Rec. BT.500, [14]. The protocol of single stimulus (SS) method is based on the use of long sequences (Fig.5). These longer segments comprise a more representative sample of scene-dependent video within a single assessment period. In this method, no comparison with an unimpaired reference condition is invited during the presentation and only one test sequence is shown each time.

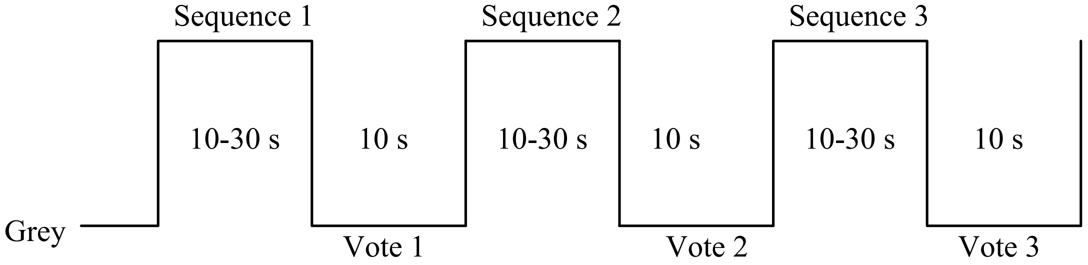


*Figure 5 - Structure of presentations for SS method*

To perform subjective assessment of picture quality, we used very similar measurement system configuration to that showed in Figure 1 (signal generator is replaced by video tape recorder). M-JPEG compression was applied to three different types of video sequences with different statistical characteristics to obtain a balance of critical and moderately critical material:

A. One person is facing the camera directly and talking (a video with little motion),
B. Two persons talking (a video in which the motion of the speakers is large),
C. The groups of people (a video with many details and large motion ).

The viewers value the quality of sequences using five-grade quality scale (5-excellent, 4-good, 3-fair, 2-poor, 1-bad). The results are evaluated using mean opinion score (MOS) as the measure of picture quality. In total 15 viewers participated in the test. They were non-experts with normal visual acuity. The size of test sequences was 750 frames (30s). The picture quality of the test sequence is changed using different compression factors: Q-20, Q-200 and Q-1024. Figure 6 shows the MOS for the test sequences coded with Q-20, Q-200 and Q-1024. For the highest compression factor sequence C. (many details and subjects movement) has the lowest MOS. This sequence can be reconstructed only with many approximations that result in very low picture quality. For the lowest compression factor (Q-20) all sequences have MOS larger then 4. It means that coder works by removing redundancy of data without approximations. For all quality levels the sequence A. with the most redundant data has the largest MOS. It confirms that codec works better for sequence with more redundant data.

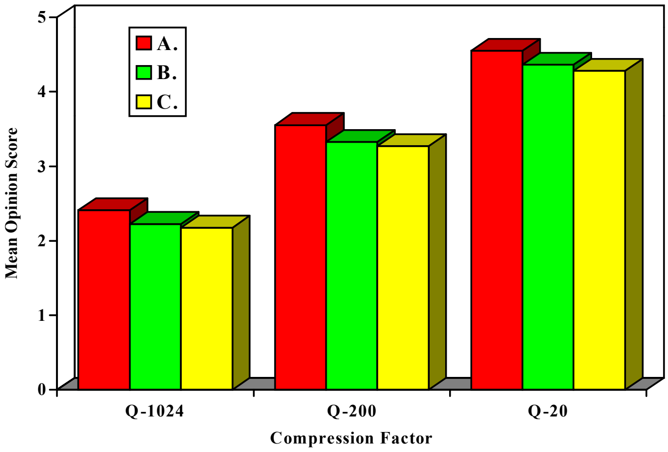


*Figure 6 - The MOS for the test sequences coded with Q-20, Q-200 and Q-1024*

The differences in MOS values for the sequences coded with the same compression factor are very small. It means that the picture quality is determined primarily by compression factor and secondarily by picture content. The results of objective measurements show good correlation with results of subjective testing. It indicates that objective measurements can be used for picture quality evaluation in compression system, which removes only spatial redundancy within a frame.

## 5. CONCLUSION

The evaluation of picture quality in digital systems is a complex affair, and it is mistake to believe that few simple quality measures characterise a system. The objective measures such as distortion of static test waveforms, Signal to Noise Ratio and Mean Squared Error do not correlate well with subjective quality measures in compression system which removes temporal redundancy between frames. In compressed video system, which removes temporal redundancy, distortion is a function of picture content. In this system quality measurements using test waveforms do not characterise the picture degradation due to compression and temporal characteristics of the system are not explored. But objective measures have good correlation with subjective results in compressed video system, which reduce only spatial redundancy within single frame. Objective measurements are repeatable and do not depend on viewing conditions or the mood of the viewers.

In real video sequences many picture defects are hidden by the picture transition or content with which they are associated. Therefore subjective measures are needed. But, subjective measures of picture quality depend on viewing conditions. The measurement takes a large amount of time and resources and the results are not always repeatable. Ideally, a combination of subjective and objective test methods is the most effective way to test video compression system.

## REFERENCES

[1] ITU-R Rec.BT.601:"Encoding Parameters of Digital Television for Studios", ITU, Geneva, 1993.

[2] G.K.Wallace:"The JPEG Still Picture Compression Standard", Comm.ACM, vol.34, No.4, Apr.1991, p.p.31-44.

[3] ITU-T Recommendation H.261:"Video Codec for Audio-Visual Services at Px64 kbit/s", Mar.1993.

[4] ISO/IEC IS 11172-2:"Information Technology-Coding of Moving Picture and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s: Video", Aug.1993.

[5] ISO/IEC IS 13818-2:"Information Technology-Generic Coding of Moving Pictures and Associated Audio Information: Video", Draft, Mar.1994.

[6] M.Drury:"Picture Quality Issues in Digital Video Compression", IBC'95, Conf. Pub. No. 413, Amsterdam, 1995, p.p.13-18.

[7] B.Zovko-Cihlar, S.Bauer, D.Modrić:"Coding Techniques in Multimedia Communications", 2nd International Workshop on Image and Signal Processing, Budapest, 1995, p.p.24-32.

[8] S.Bauer, B.Zovko-Cihlar, M.Grgić:" The Influence of Impairments from Digital Compression of Video Signal on Perceived Picture Quality", 3rd International Workshop on Image and Signal Processing, Manchester, Nov.1996.

[9] N.Jayant, P. Noll: "Digital Coding of Waveforms: Principles and Applications to Speech and Video", Prentice Hall, Washington, 1984.

[10] B.Zovko-Cihlar: "Noise in Radiocommunications", Školska knjiga, Zagreb, 1989.

[11] L.E. Weaver:"Television Video Transmission Measurements", Marconi Instruments, 1973.

[12] R.O. Onvural:" Asynchronous Transfer Mode Networks: Performance Issues", Artech House, 1994.

[13] M.Ardito, M.Visca:"Correlation between Objective and Subjective Measurements for Video Compressed Systems", IBC'95, Conference Publication No. 413, Amsterdam,1995, p.p.7-12.

[14] ITU-R Rec. BT. 500:"Methods for the Subjective Assessment of the Quality of Television Pictures"

[15] ITU-R Rec.BT.710:"Subjective Assessment Methods for Image Quality in HDTV"

[16] ITU-R Rec. BT.811-1:"Subjective Assessment of Enhanced PAL and SECAM Systems"

[17] ITU-T Rec. P.80:"Methods for Subjective Determination of Transmission Quality"

[18] ITU-T Rec.P.910:"Subjective Video Quality Assessment Methods for Multimedia Application"

[19] ITU-T Rec. P.920:" Interactive Test Methods for Audio-Visual Communications"

[20] ITU-T Rec. P.930:"Principles of Reference Impairments System for Video"