17.2: Detectability and Annoyance of Synthetic Blocky and Blurry Video Artifacts

Mylène C.Q. Farias†, Michael S. Moore†, John M. Foley, and Sanjit K. Mitra†*

†Department of Electrical and Computer Engineering,* Department of Psychology

University of California Santa Barbara, Santa Barbara, CA 93106 USA

Abstract

In order to study video artifacts and their effect on perceived quality, it is desirable to create synthetic artifacts that are less complex than compression artifacts, yet have the same appearance. These should be produced by a standard method, be adjustable and combinable, and produce a wide range of quality. This paper describes a method for generating two synthetic artifacts, shows that they have these properties and that their thresholds and annoyance values are closely related to those of MPEG-2 artifacts.

1. Introduction

Current video compression systems such as MPEG-2 are known to introduce various distortions or impairments in the decoded video sequences such as image blocking, mosquito noise, and blurring [1]. We do not yet have a good understanding of how the perceived quality of compressed video images depends on their physical properties. One approach is to work with synthetic artifacts that look like compression artifacts, yet are simpler, purer, and easier to describe and control. Such artifacts offer advantages for experimental research on video quality and are necessary components of the kind of reference impairment system recommended by the ITU-T for the measurement of image quality [2]. Our focus is on compression artifacts produced by MPEG-2.

There are several properties that are desirable in synthetic artifacts if they are to be useful for these purposes. The synthetic artifacts should:

- be generated by a precisely defined and easily replicated algorithm,
- be relatively pure and easily adjusted and combined to match the appearance of the full range of compression impairments, and
- produce psychometric functions and annoyance functions that are similar to those for compression artifacts.

In this paper we describe algorithms for the creation of blocky and blurry synthetic artifacts. These are simpler than those described in the ITU-T recommendation, yet they can be combined to look like artifacts produced by MPEG-2 compression. We combined them to create blocky-blurry artifacts and we then did a psychophysical experiment to determine the psychometric functions for detection and the annoyance functions for these combined artifacts. We compare these functions with those from an earlier experiment in our laboratory in which MPEG-2 artifacts were inserted in the same video sequences [3].

2. Method

The usual approach to subjective quality testing is to degrade a video by a variable amount and ask the test subjects for a quality rating [4]. Since both the type and the strength of artifacts vary from frame to frame and region to region, this method cannot be used to measure the visibility and annoyance produced by specific artifacts at specific strengths. To do this we use an experimental paradigm in which impairments are restricted to an isolated region (defect zone) of the video clip for a short time interval. The rest of the video is left in its original state [3].

The test subjects were instructed to search each video for defective regions and to indicate the annoyance value caused by any defects seen. The regions where the defects or artifacts appear are varied to prevent the test subjects from learning the locations where they appear. The same artifact at the same strength will vary in both visibility and annoyance depending on where it appears in the video sequence [3].

Our test subjects were drawn from a pool of students in the introductory psychology class at UCSB. The students are thought to be relatively naive concerning video artifacts and the associated terminology. They were asked to wear any vision correcting devices (glasses or contacts) that they normally wear to watch television.

To generate the test video sequences, we start with a set of five original video sequences of assumed high quality. These were the MPEG standard Bus, Cheerleader, Flower-garden, Football and Hockey sequences. The video clips are all 5 seconds long and contain scenes that we think are typical of normal television.

2.1 Generation of synthetic artifacts

Blocking is a distortion of the image characterized by the visibility of an underlying block encoding structure [1] and is often caused by coarse quantization of the spatial frequency components during the encoding process. We produced blocking artifacts by using the difference between the average of each block and the average of the surrounding area to make each block stand out.

The algorithm for producing the blocking artifacts is as follows: The first step is to calculate the average of each 8×8 block of the frame and of the 24×24-surrounding block, which has the current 8×8 block as a center. The next step is to calculate the difference, $D(i,j)$, between these two averages for each block. This difference is added to each block of the original frame $X(k,l)$:

$$A(k,l) = X(k,l) + D(i,j) \,, \qquad (1)$$

where $i = round(k/8)$, with $1 \leq i \leq Rows/8$, and $j = round(l/8)$, with $1 \leq i \leq Cols/8$. $\boldsymbol{X}$ is the original frame, $\boldsymbol{A}$ is the frame with artifacts, *Rows* is the total number of rows of the frame, and *Cols* is the number of columns of the frame. The same algorithm was applied to luminance and the two color components of the video. Three other minor modifications of the artifacts were made: block average was matched to the original, borders of defect zones were faded, and saturation was controlled.

Blurring is the reduction in sharpness of edges and spatial detail

[1]. In compressed images blurring is often caused by trading off bits to code resolution and motion. Blurring artifacts were generated by applying a symmetric, two-dimensional FIR (Finite duration Impulse Response) low-pass filter to the digital image array. We used a 5×5 mean filter in this experiment [5]. Different filters with varying cut-off frequencies could be used to control the amount of blurriness introduced.

The artifacts that we used were a mix of blurry and blocky artifacts, which were visually similar to the artifacts present in highly compressed MPEG-2 reconstructed videos. The artifact mixtures, ***C***, (blocky-blurry artifacts) were obtained using a fixed combination of a pure blurry and blocky artifacts given by the following proportion:

$$\boldsymbol{C} = 0.5 \cdot \boldsymbol{A}_{\boldsymbol{blocky}} + 0.5 \cdot \boldsymbol{A}_{\boldsymbol{blurry}} , \qquad (2)$$

where $\boldsymbol{A}_{blocky}$ is the sequence with block artifacts, $\boldsymbol{A}_{blurry}$ is the sequence with blurring artifacts. This combination rule was chosen based mostly on the basis of visual comparisons. We searched for a proportion that looked as similar as possible to a MPEG-2 sequence compressed at 7.5 Mbps, then reconstructed. We also measured the averages and variances of the luminance of the created video sequences to assure a rough match with the MPEG-2 reconstructed ones. This combination rule produced artifacts that were the most perceptually similar to MPEG-2 artifacts compressed to 7.5 Mbps. They also roughly matched the MPEG-2 artifacts in the mean and variance of their luminance distributions.

2.2 Generation of test sequences

Since one of the goals of the experiment was to compare the visibility and annoyance value of synthetic artifacts with the real MPEG-2 artifacts, the same original sequences and defect zones used in the previous experiment were used. To generate the test sequences, we combined the original video with the combined artifact in different proportions. By varying their relative weights, we could weaken the artifact (allowing the original to dominate), strengthen the artifact (allowing the artifact to dominate), or even exaggerate the artifact (boost the difference between the artifact and original). The basic formula is:

$$\boldsymbol{M} = \boldsymbol{X} \cdot (1 - r) + r \cdot \boldsymbol{C} , \qquad (3)$$

where $\boldsymbol{M}$ is the result, $\boldsymbol{X}$ is the original, $\boldsymbol{C}$ is the sequence with the combined artifacts, and r is the weighting factor ($0 \le r \le 1$). It is important to emphasize at this point that the artifacts were added only to the defect zones. All other areas were kept untouched.

The total squared error *(TSE*) is defined as:

$$TSE = \sum (\boldsymbol{M} - \boldsymbol{X})^2 = \sum (r(\boldsymbol{X} - \boldsymbol{C}))^2 , \qquad (4)$$

The *TSE* was used as our objective error measure. The test sequences with synthetic artifacts had *TSE* values that were between the *TSE* values of the MPEG-2 sequences coded at 1 Mbps and at 7.5 Mbps.

2.3 Experimental Design

There were five original sequences. Artifacts were inserted into three defect zones in each at six different strengths. Each of these 95 sequences was presented once to each subject. The experimental trials were preceded by showing some extreme examples to set the scale for the annoyance judgments and by five practice trials.

2.4 Data Analysis

One of our goals was to measure the error energy (TSE) detection threshold for each of our artifacts. To do this we need the probability of detection for each artifact. The threshold is defined as the log error energy such that the artifact was seen by 50% of our subjects. We estimated the probability of detecting each artifact by counting the number of people who detected the artifact and dividing by the number of observations. The probability as a function of the $\log_{10}$ *TSE* (Eq. (4)) was fitted using the Weibull function, which has an S-shape similar to our data and is defined as

$$P(x) = 1 - 2^{-(S \cdot x)^k} , \qquad (5)$$

where $P(x)$ is the probability of detection, x is $log_{10}TSE$, $1/S$ is the 50% detection threshold in log total squared error, and k is a constant that indicates the steepness of the function.

Figures 1 and 2 depict the psychometric functions for both the synthetic artifacts and the MPEG-2 artifacts coded at 7.5 Mbps. The strength of the MPEG-2 artifacts was varied in the same way as the synthetic artifacts (Eq. (3)). The video is Cheerleaders with the defect zone filling the top one third of the frame. Figures 3 and 4 depict the same function for the video Flower with the defect zone over the houses. As expected, the curves have the same form.

Table 1 Parameters of psychometric and annoyance functions for the synthetic and MPEG-2 artifacts inserted in the same regions of the sequences. Bold numbers indicate values that are statistically significant at $p = 0.05$.

	S		k		$\bar{x}$		β	
Sequences	Synthe.	MPEG	Synthe.	MPEG	Synthe.	MPEG	Synthe.	MPEG
BusBottom	0.26	0.27	19.15	9.60	4.39	4.08	0.27	0.45
BusMiddle	0.27	0.29	14.96	26.28	4.18	3.93	0.22	0.25
BusTop	0.28	0.32	10.18	12.71				
CheerBottom	0.28	0.27	19.71	13.03	4.18	4.11	0.28	0.33
CheerMiddle	0.30	0.28	13.12	18.30	3.92	4.02	0.27	0.45
CheerTop	0.28	0.27	10.57	19.27	4.42	4.39	0.34	0.46
FlowerGarden	0.24	0.25	17.09	14.98	4.62	4.59	0.29	0.37
FlowerHouses	0.29	0.28	12.18	15.71	3.99	4.03	0.26	0.29
FlowerSky	0.28	0.28	16.44	11.65				
FootballLeft					3.87	3.79	0.38	0.28
FootballMiddle	0.32	0.32	15.77	22.02	3.68	3.42	0.25	0.25
FootballRight					3.94	3.53	0.40	0.36
HockeyLeft	0.32	0.33	11.11	18.01	3.58	3.46	0.30	0.39
HockeyMiddle	0.33	0.36	19.79	17.78	3.55	3.37	0.25	0.30
HockeyRight					3.55	3.45	0.28	0.27
Mean	0.288	0.293	15.006	16.612	3.99	3.859	0.292	0.342
T-test, p =	0.281		0.414		**0.007**		**0.043**	
Correlation	**0.831**		-0.252		**0.930**		0.262	

The parameters *S* and *k*, for all test sequences, are shown in Table 1 for both synthetic and MPEG artifacts. The empty spaces in Table 1 correspond to cases where a fit was not possible. The MPEG-2 data are from Reference [3]. The sensitivity parameters, *S*, for both types of artifacts are highly correlated. The small differences in the parameter values for *S* and *k* are not statistically significant (T-test, two-tailed, $p > 0.05$).

We used the standard methods [5] for analyzing the annoyance judgments provided by the test subjects and computed the mean annoyance values. The mean annoyance values for each test sequence were fitted with the standard logistic function [5]:

$$y = y_{\min} + \frac{(y_{\max} - y_{\min})}{\left(1 + \exp\left(-\frac{(x - \bar{x})}{|\beta|}\right)\right)}, \qquad (6)$$

where *y* is the predicted annoyance and *x* is the $log_{10}(TSE)$. The parameters $y_{\min}$ and $y_{\max}$ establish the limits of the annoyance value range. The parameter $\bar{x}$ translates the curve in the x-direction and the parameter β is inversely related to the steepness of the curve.

Figures 5 and 6 depict the mean annoyance values versus the log total squared error for the synthetic artifacts and the MPEG-2 artifacts coded at 7.5 Mbps, corresponding to the video Cheerleaders with the defect zone filling the top part of the frame. Figures 7 and 8 depict the same function for the video Flower with the defect zone over the houses. Again, the curves have similar form. Most of the annoyance functions for the synthetic artifacts are shifted slightly to the right implying that the same error produces slightly less annoyance. Also, most of these functions are steeper for the synthetic artifacts, implying that annoyance grows faster with TSE for these artifacts.

The parameters of the annoyance functions ($\bar{x}$ and β) are shown in Table 1. Here the small differences in the parameters for the corresponding MPEG-2 and synthetic artifacts are statistically significant. The $\bar{x}$ values, for the two types of artifacts are highly correlated.

3. Conclusions

In this paper, a new method for generating synthetic artifacts for digital video sequences was presented. The proposed method satisfies all the conditions specified in [2], except for the fact that at present time it only generates two types of artifacts – blockiness and blurriness. The synthetic artifacts look very realistic and have objective error measures within the same range as the MPEG-2 artifacts. A comparison of the results of this experiment with the results from a previous experiment carried out in our Laboratory at UCSB, using MPEG-2 artifacts [3] showed that the form of the functions for MPEG-2 and synthetic artifacts is the same and the *S* and $\bar{x}$ parameters are highly correlated across the two artifact types. Thresholds are essentially the same, but synthetic artifacts are slightly less annoying than MPEG-2 artifacts at the same TSE.

6. Acknowledgements

This work was supported in part by a National Science Foundation Grant CCR-0105404 and in part by CAPES-Brazil.

7. References

[1] Michael Yuen, and H.R.Wu, "A Survey of Hybrid MC/DPCM/DCT Video Coding Distortions," Signal Processing 70 (1998), pp. 247-278.

[2] Recommendation ITU-R BT.500-930, "Principals of a reference impairment system for video," ITU-T 1996.

[3] Michael S. Moore, John M. Foley, and Sanjit K. Mitra, "Detectability and Annoyance Value of MPEG-2 Artifacts Inserted into Uncompressed Video Sequences," Proceedings of the SPIE, Human Vision and Electronic Imaging V, San Jose, CA, vol. 3959, pp. 99-110, January 2000.

[4] ITU Recommendation BT.500-8, "Methodology for Subjective Assessment of the Quality of Television Pictures," 1998.

[5] Rafael C. Gonzalez, and Richard E. Woods, "Digital Image Processing," Addison Wesley.

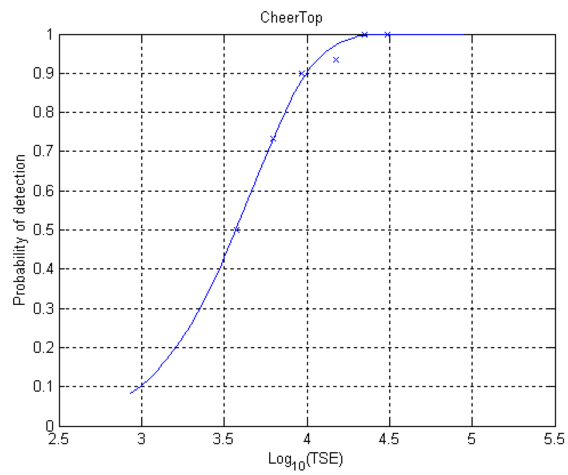

Figure 1. Psychometric functions for the synthetic blocky-blurry artifact inserted in the video Cheerleaders-Top.

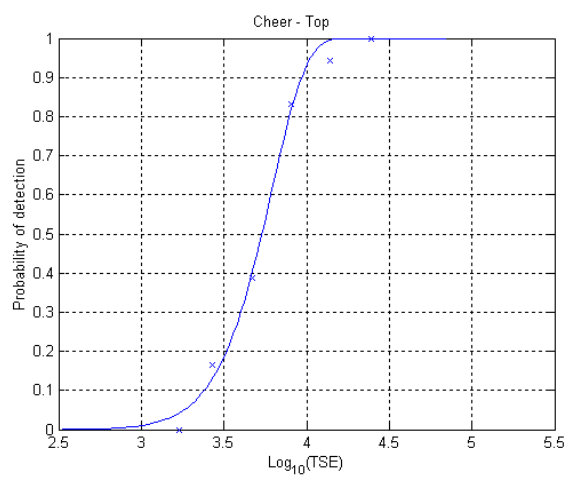

Figure 2. Psychometric function for MPEG-2 artifact inserted in the same region of the video Cheerleaders-Top.

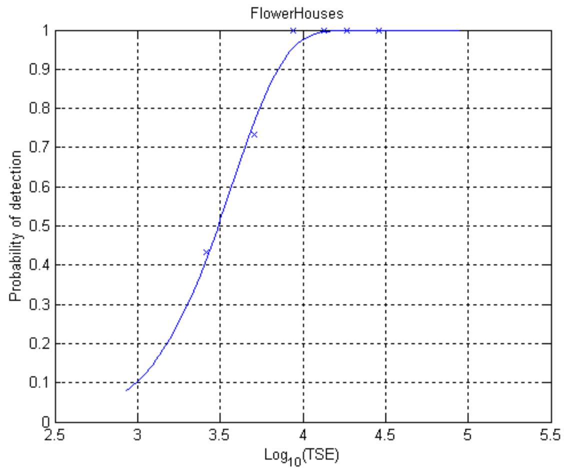

Figure 3. Psychometric functions for synthetic blocky-blurry artifact inserted in the video Flower-Houses.

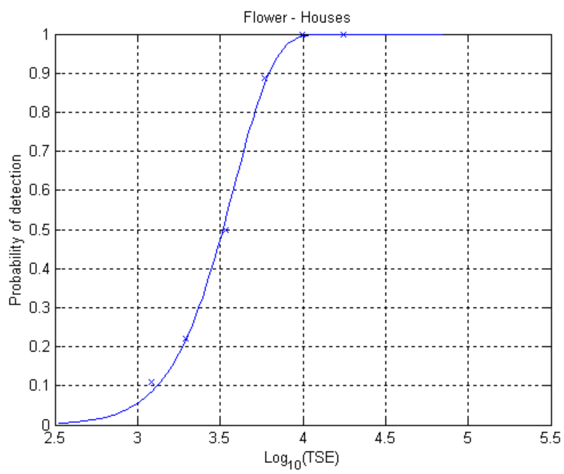

Figure 4. Psychometric function- for MPEG-2 artifact inserted in the same region of the video Flower-Houses.

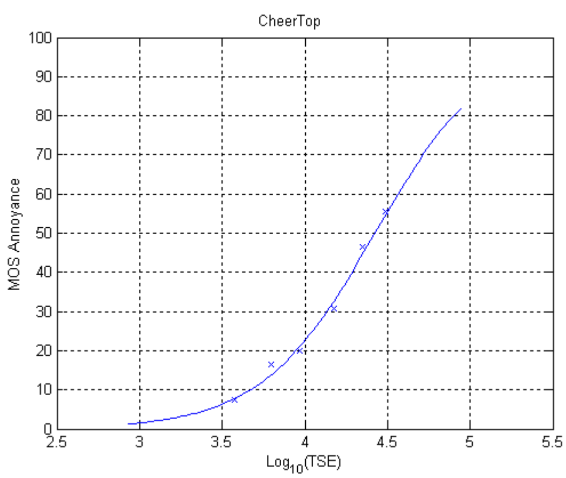

Figure 5. Mean annoyance curves for synthetic artifacts, corresponding to the video Cheerleaders-Top.

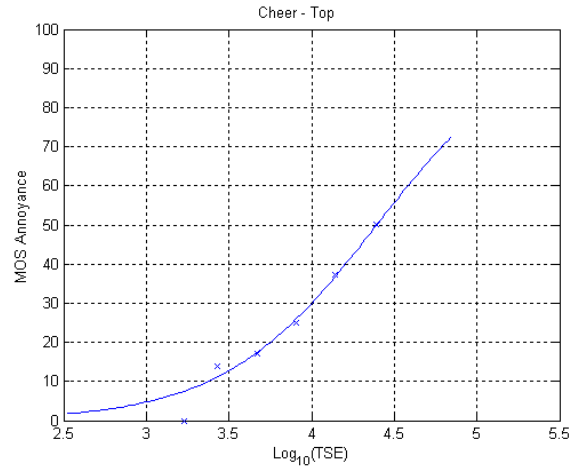

Figure 6. Mean annoyance curves for MPEG-2 artifacts, corresponding to the video Cheerleaders-Top.

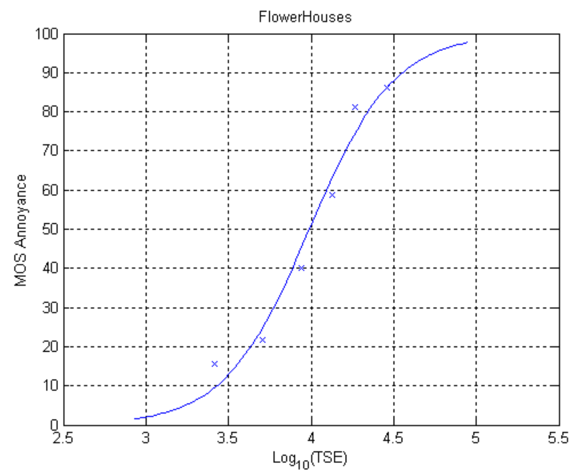

Figure 7. Mean annoyance curves for synthetic artifacts, corresponding to the video Flower-Houses.

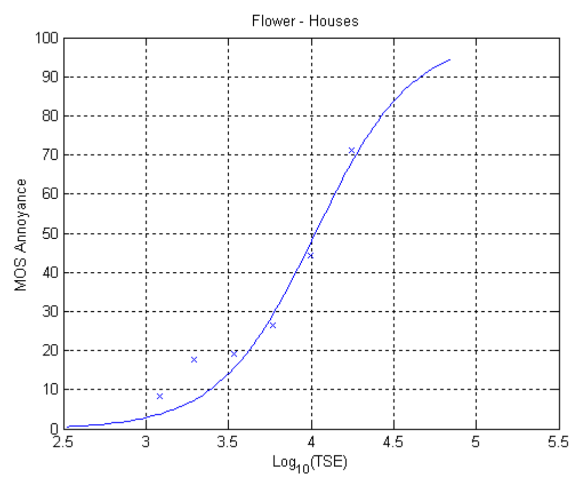

Figure 8. Mean annoyance curves for MPEG-2 artifacts, corresponding to the video Flower-Houses.