https://doi.org/10.2352/ISSN.2470-1173.2018.12.IQSP-234

Perceptual Strengths of Video Impairments that Combine Blockiness, Blurriness, and Packet-Loss Artifacts

Alexandre Fieno da Silva; Federal Institute of Triângulo Mineiro; Paracatu, MG/Brazil. Mylène C.Q. Farias; University of Brasília; Brasília, DF/Brazil.

Abstract

In this paper, we report the results of a set of psychophysical experiments that measure the perceptual strengths of videos with different combinations of blockiness, blurriness, and packet-loss artifacts and the overall annoyance. Participants were instructed to search each video for impairments and rate the strength of their individual features (artifacts). A repeated measure Anova (RM-ANOVA) performed on the data showed that artifact physical strengths have a significant effect on annoyance judgments. We tested and reported a set of linear models on the experimental data and we found that all these models give a good description of the relation between individual artifact perceptual strengths and the overall annoyance. In other words, all models presented a very good correlation with the experimental data, showing that annoyance can be modeled as a multidimensional function of the individual artifact perceptual strengths. Additionally, results show that there are interactions among artifact signals.

Introduction

In the past decades, several quality assessment methods have been proposed with the goal of automatically measuring the video quality at the user side [1, 2]. Most available methods are full-reference video quality assessment (FR-VQA) methods, which require the use of the original (reference) video [3, 4]. The design of blind methods, i.e. no-reference video quality assessment (NR-VQA) methods, is still a challenge [5, 6]. Given that the overall video quality can be estimated by combining the individual artifact *perceptual* strengths [7–12], there is a considerable number of NR-VQA methods that use this 'multidimensional' approach [13–15]. These methods, known as distortion-specific (DS) VQA methods, estimate the strength of a set of artifacts and combine these estimates to obtain an overall quality score for the impaired video. Among the state-of-the-art DS-VQA methods, we can cite the papers of Hadizadeh & Bajic [16], Bahrami & Kot [17], Golestaneh & Chandler [18], and Li *et al.* [19–21].

The performance of DS-VQA methods depends on: the accuracy of the perceptual artifact strength models, which give an estimate of the perceived strength for each artifact, and the accuracy of the combination quality models, which compute a overall quality scores considering the perceptual strengths of each artifact and their interactions. So, the design good artifact metrics requires a good understanding of the perceptual characteristics of each artifact, as well as a knowledge of how the strength of each artifact contributes to the overall quality [22]. Up to our knowledge, besides the work by Farias *et al.*, little work has been performed to study and characterize the appearance and perception of combined artifacts [3]. As a consequence, currently there is no clear knowledge on how spatial and temporal video artifacts combine perceptually and how their impact depends on the physical properties of the video.

In this work, we study the characteristics of two spatial artifacts (blockiness and blurriness) and one temporal artifact (packet-loss), which are among the most commonly found artifacts in digital transmission scenarios. More specifically, we analyze the relationship between the *perceptual strengths* of these artifacts and the overall *annoyance*. We also analyze the relationship between physical and perceptual artifact strengths and study masking effects between artifacts. With this goal, we perform a set of six psychophysical experiments in which subjects estimated the strength and annoyance of blockiness, blurriness, and packet-loss artifacts, either in isolation or in combinations. With these goals, we performed an analysis of the subjective data obtained from these experiments and tested a set of combination models with the goal of predicting overall annoyance from the *perceptual strengths* and *physical strength parameters* of these three artifacts. This work is a follow-up on a previous work [23,24], in which we investigated the impact of *physical strength parameters* of blockiness, blurriness, and packet-loss on overall annoyance.

Experimental Methodology

Experiments were performed using a PC computer with test sequences displayed on a Samsung LCD monitor of 23 inches (Sync Master XL2370HD) with resolution 1920×1080 @60hz (FullHD 1080p). We used a constant illumination of approximately 70 lux and participants were kept at a fixed distance of 0.70 meters from the monitor using a chin-rest. The experimental methodology was the single-stimulus with hidden reference, with a 100-point continuous impairment scale [24, 25]. The participants, mostly graduate students from the author's institutions, were considered naive of most kinds of digital video defects and the associated terminology.

The experimental session started with a brief oral introduction. Then, participants performed a training, which consisted of watching highly impaired and pristine sequences to get acquainted with the typical artifact combinations and strengths. After the training, the scoring session started. Since initial judgments are generally erratic, we included 5 practice trials, which were not recorded [25]. Besides eliminating erratic answers, practice trials exposed subjects to a good range of impairments and gave them a chance to try the scoring interface. Experimental trials were performed with the complete set of test sequences presented in a random order. Videos were played once and subjects were not allowed to go back and watch them again. Experimental sessions lasted between 45 and 60 minutes. To avoid fatigue, experimental sessions were broken into sub-sessions.

As mentioned earlier, a total of six experiments were performed using the same set of test sequences. In three of the experiments (experiments 1, 3, and 5), we asked participants to estimate the annoyance of the test sequences. More specifically, participants were asked to estimate the annoyance caused by the artifacts, by giving a score between '0' and '100'. Artifacts as annoying as the worst artifacts in the training session should have be given a '100', artifacts half as annoying a '50', and so on. For each test sequence, these experiment provided one annoyance score per participant. We computed the Mean Annoyance Value (MAV) corresponding to each test sequence by taking the average of the annoyance scores provided by all participants.

In the other three experiments (experiments 2, 4, and 6), participants were asked to give a strength score to each individual type of artifact of the test sequence. Artifacts as strong as those seen in the training session should be given a '100', artifacts half as strong a '50', and so on. The number of artifacts in each test sequence varied. Data gathered from these three experiments provided up to three Mean Strength Values (MSV) for each test sequence: MSV_{bloc}, MSV_{blur}, and MSV_{pck}, which correspond to MSVs for blockiness, blurriness, and packet-loss, respectively. MSVs are computed by averaging the strength values over all subjects for each video and artifact type.

To study how the artifact strengths combine to predict the perceived annoyance of videos impaired by multiple and overlapping artifacts, we fit a set of linear models to the MSV subjective data and the MAV data collected (for the same test sequences) in the previous experiments [23, 24]. To estimate the performance of the models, we calculate the Pearson correlation coefficient (PCC) and the Spearman Rank Order Correlation Coefficient (SCC) between the subjective and predicted scores. To test the effect of the artifact parameters on annoyance, we perform a repeated-measure ANOVA (RM-ANOVA) with a significance level of 95% ($\alpha = 0.05$).

Genration of Test Sequences

We used seven high definition original videos, chosen with the goal of generating a diverse content, with spatial resolution of 1280×720, temporal resolution of 50 frames per second (fps), and duration of 10 seconds. Three types of artifacts were used: blockiness, blurriness, and packet-loss (more details about these artifacts can be found in ITU Recommendation P.930 [26]). To add artifacts to the originals, we used a system for generating artifacts [27, 28] that allowed a control of the artifact combination, visibility, and strength, what would be difficult using, for example, a compression codec.

To add blockiness to each video frame in our dataset, we calculated the average value of each 8×8 block of the frame and of the 24×24 surrounding block, then added the difference between these two averages to the block. To generate blurriness, we used a simple low-pass filter, as suggested by Recommendation P.930 [26]. Although we can vary the filter sizes and the cut-off frequencies to control the amount of blurriness, we used a simple 5×5 moving average filter. We generated test sequences with combinations of blockiness and blurriness by linearly combining the original video with blockiness and blurriness artifact signals in different proportions (i.e. 0.4, 0.6, and 0.8) [27, 28] .

To generate packet-loss artifacts, we first compressed the videos at high compression rates, what avoids inserting additional artifacts. Then, packets from the coded video bitstream were randomly deleted using different percentages of deleted packets (PDP), with higher percentages corresponding to higher levels of degradation [24]. To vary the time interval between consecutive artifacts, we changed the number of frames (M) between I-frames.

Table 1: Exps. 1-2: Combinations of the parameters PDP and M used for each of the 7 originals.

Comb	M	PDP	Comb	M	PDP	Comb	M	PDP
1	4	0.7	5	8	0.7	9	12	0.7
2	4	2.6	6	8	2.6	10	12	2.6
3	4	4.3	7	8	4.3	11	12	4.3
4	4	8.1	8	8	8.1	12	12	8.1

Table 2: Exps. 3-4: Set of combinations used for each of the 7 originals: 'bloc' and 'blur' correspond to the blockiness and blurriness intensities, respectively.

Comb	(bloc;blur)	Comb	(bloc;blur)	Comb	(bloc;blur)
1	(0.0;0.0)	5	(0.4;0.4)	9	(0.6;0.6)
2	(0.0;0.4)	6	(0.4;0.6)	10	(0.0;0.8)
3	(0.0;0.6)	7	(0.6;0.0)	11	(0.8;0.0)
4	(0.4;0.0)	8	(0.6;0.4)		

Experimental Setup

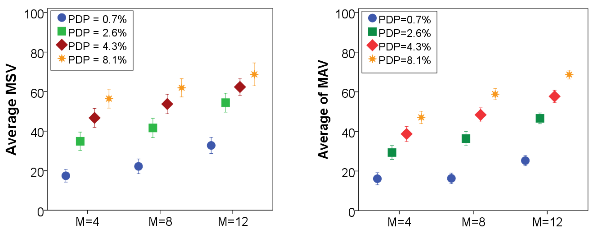
As mentioned earlier, this works analyzes the data collected from six different experiments. Sixteen participants rated the annoyance in Experiment 1, whilst fourteen participants rated the artifacts' strength in Experiment 2. In both experiments, test sequences contained only packet-loss artifacts. The set of PDP and M parameters used in this experiment are given in Table 1. A total of 7 originals and 12 parameter combinations were used, resulting in 91 test sequences.

Sixteen participants rated annoyance in Experiment 3, whilst fifteen participants rated the artifacts' strength in Experiment 4. In both experiments, test sequences contained different strengths of blockiness and blurriness artifacts, presented in isolation or in combination. These combinations are represented by a vector (bloc; blur), where 'bloc' is the blockiness signal intensity and 'blur' is the blurriness signal intensity. The experiment contained a set of videos with all possible combinations of the two artifact types (full factorial design: $3^2 = 9$), plus two additional combinations of strong pure blockiness and pure blurriness. Table 2 shows all combinations used in the experiment. A total of 7 originals and 10 combinations were used in this experiment, resulting in 77 test sequences.

Twenty-three participants rated annoyance in Experiment 5. whilst thirty-five participants rated the artifacts' strength in Experiment 6. In both experiments, the test sequences contained different strengths of blockiness, blurriness, and packet-loss artifacts, presented in combinations. The strength combinations are represented as a vector (PDP;bloc;blur), where 'PDP' is the intensity of packet-loss, 'bloc' is the intensity of blockiness, and 'blur' is the intensity of blurriness. To limit the number of artifact combinations, we selected a subset of all possible artifact combinations considering the results of the previous experiments (1-4). Table 3 shows all combinations used in this experiment, which include three intensities for each artifact type. Again, 7 originals and 19 combinations were used, resulting in 140 test sequences.

Table 3: Exps. 5-6: Combinations for each original: 'bloc' corresponds to the blockiness intensity, 'blur' to the blurriness intensity, and 'PDP' to the percentage of deleted packets.

Comb.	(PDP;Bloc;Blur)	Comb.	(PDP;Bloc;Blur)	Comb.	(PDP;Bloc;Blur)
1	(0.0;0.0;0.0)	8	(8.1;0.0;0.6)	15	(0.7;0.6;0.0)
2	(0.0;0.6;0.0)	9	(0.7;0.4;0.0)	16	(8.1;0.6;0.0)
3	(0.0;0.0;0.6)	10	(8.1;0.4;0.0)	17	(0.7;0.6;0.4)
4	(8.1;0.0;0.0)	11	(0.7;0.4;0.4)	18	(8.1;0.6;0.4)
5	(0.7;0.0;0.4)	12	(8.1;0.4;0.4)	19	(0.7;0.6;0.6)
6	(8.1;0.0;0.4)	13	(0.7;0.4;0.6)	20	(8.1;0.6;0.6)
7	(0.7;0.0;0.6)	14	(8.1;0.4;0.6)		



(a) (b)

Figure 1: Exps. 1-2: (a) MSV_{pck} plots, and (b) Average MAV plots for clustered error for M = 4, 8, and 12.

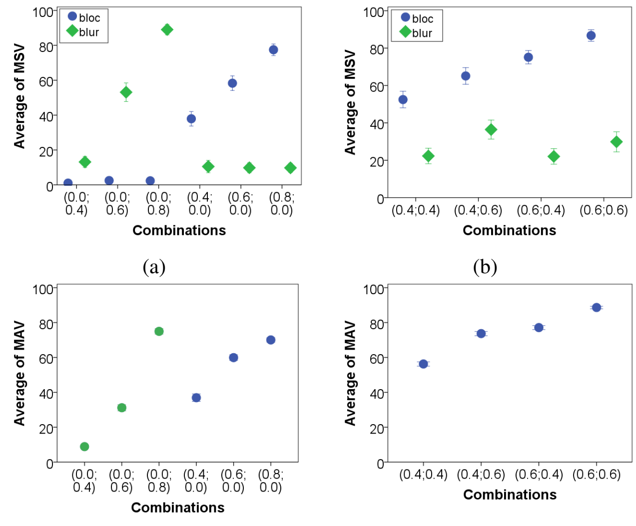
Experiments 1 and 2

Figure 1 show graphs of the average MSV_{pck} and the average MAV versus PDP, grouped according to the M value. Notice that, for M = 4, 8, and 12, both the highest MSV_{pck} and MAV always correspond to the strongest artifact (e.g. PDP = 8.1%). Although average values increase with both PDP and M, PDP seems to have a bigger effect than M. In fact, an RM-ANOVA test shows that the effect of PDP on MSV_{pck} and MAV, for any pair of M values, is statistically significant.

Experiments 3 and 4

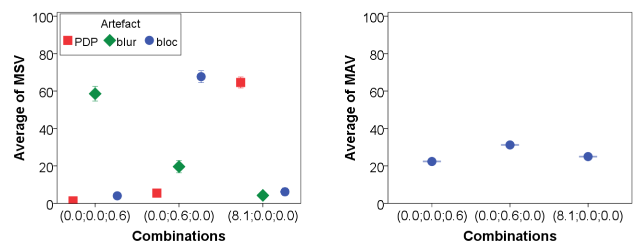
Figure 2 (a) and (c) show graphs of the average MSV_{blur} (green) and MSV_{bloc} (blue), and the average MAV, for test sequences containing combinations of only-blurriness and only-blockiness, respectively. Notice that, the highest MSVs are obtained for the combinations with higher artifact strengths. In addition, we can notice that the average MAVs increase with the artifact strength. An RM-ANOVA was performed to check if MSVs and MAVs differences for different blockiness and blurriness strengths are significant. The results show that there are significant statistical differences in MAVs and MSVs for all pairs of different strengths in only-blockiness and only-blurriness sequences. These results indicate that participants correctly perceived the different artifact strengths introduced in the videos.

Figure 2 (b) and (d) show graphs of the average MSV_{blur} and MSV_{bloc}, and the average MAV for all combinations of blockiness and blurriness (e.g. (0.4;0.4), (0.4;0.6), (0.6;0.4), and (0.6;0.6)), respectively. An RM-ANOVA test shows that differences between MSVs and MAVs obtained for any two combinations of blockiness and blurriness are statistically significant. The only exceptions are for the combination pairs (0.4;0.4) and (0.6;0.4) for which MSV_{blur} differences are not statistically significant, and for the combination pairs (0.4;0.6) and (0.6;0.4) for which MAV differences are not found significant. This means that a change in the artifact strength was perceived by the participants.



(a) (b) (c) (d)

Figure 2: Exps. 3-4: Plots for combinations (bloc;blur) with: (a) MSV for only-blockiness and only-blurriness, (b) MSV for blockiness and blurriness, (c) MAV for only-blockiness and only-blurriness and, (d) MAV for blockiness and blurriness.



(a) (b)

Figure 3: Exps. 5-6: (a) MSV and (b) MAV for combinations (0.0;0.0;0.6), (0.0;0.6;0.0), and (8.1;0.0;0.0).

Experiments 5 and 6

Figure 3 shows the MSV and MAV plots for combinations of pure blockiness, blurriness, and packet-loss. Notice that the highest MSVs correspond to the only artifact in the video. An RM-ANOVA shows that there are significant statistical differences in MSV, for any pair of combinations, with exception of the combination pair (8.1;0.0;0.0) and (0.0;0.6;0.0). The average MSV and MAV is higher for blockiness, followed by packet-loss, and blurriness. In a similar way, for stronger artifacts, MAV also increases.

Figure 4 shows the MSVs and MAVs for combinations with two types of artifacts ((PDP;bloc;0.0) or (PDP;0.0;blur)). The strongest artifact received the highest MSV. Nevertheless, an increase in the strength of a particular artifact signal does not always result in a proportional increase in this artifact perceived strength. For example, for (PDP;0.0;blur) combinations, an increase in the strength of blurriness causes a decrease in the perceived strength of the packet-loss (see Fig. 4 (a)). An RM-ANOVA test shows that there are significant statistical MSVs differences between all combinations of (PDP;0.0;blur). The only exceptions are the combination pairs ((0.7;0.0;0.4), (8.1;0.0;0.4)) and ((0.7;0.0;0.6), (8.1;0.0;0.6)), whose MSV_{blur} differences are not statistically significant. For these two combinations, only the

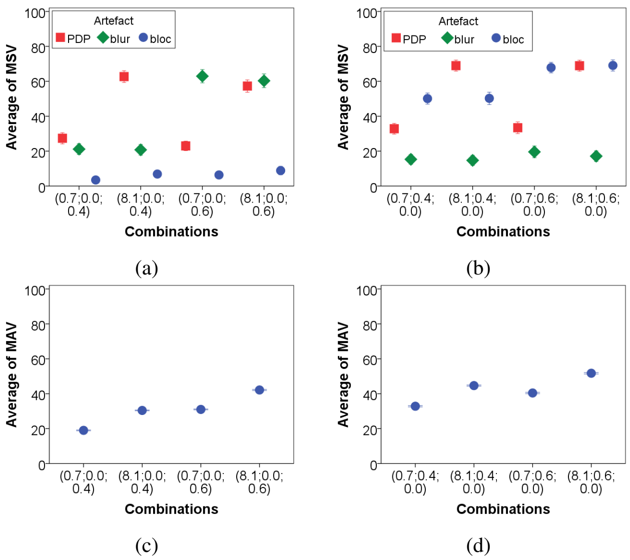


Figure 4: Exps. 5-6: (a) MSV and (c) MAV for combinations (PDP;0.0;blur), and (b) MSV and (d) MAV for combinations (PDP;bloc;0.0).

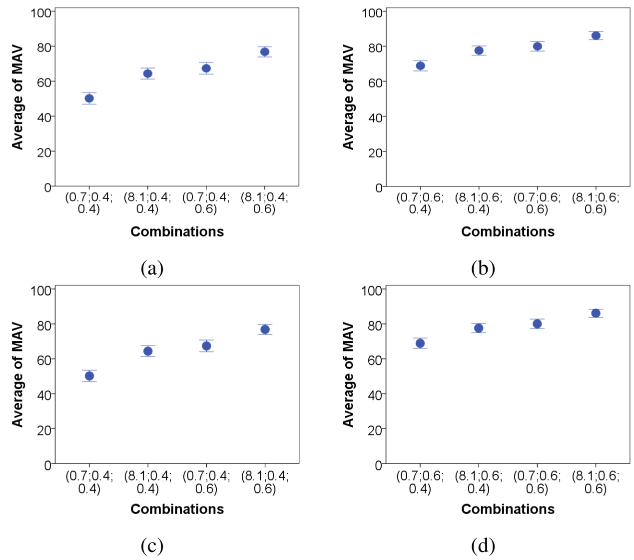


Figure 5: Exps. 5-6: (a) MSV and (c) MAV for combinations (PDP;0.4;blur), and (b) MSV and (d) MAV for combinations (PDP;0.6;blur).

packet-loss strength changes while the blurriness strength is kept constant. This result suggests that blurriness may be masking the perceived strength of packet-loss.

The presence of packet-loss in the (PDP;bloc;0.0) combinations changes the perceived strength of the blockiness artifact (see Fig. 4 (b)). This indicates that increasing the packet-loss strength in a (PDP;bloc;0.0) combination can intensify the perceived strength of blockiness. This may be caused by the similarity of blockiness and packet-loss artifacts, which are both characterized by the presence of rectangular areas distributed over the video frames. An RM-ANOVA test shows that there are significant statistical differences in MSV_{pck} for all combinations pairs (PDP;bloc;0.0). The only exceptions are the combination pairs (0.7;0.4;0.0), (0.7;0.6;0.0)) and ((8.1;0.4;0.0), (8.1;0.6;0.0). Another RM-ANOVA test also shows that there are significant statistical differences in MSV_{bloc} for the combination pairs ((0.7;0.4;0.0), (8.1;0.4;0.0)) and ((0.7;0.6;0.0), (8.1;0.6;0.0)). Also, an RM-ANOVA test shows that there are significant statistical MAVs differences between all combinations of (PDP;0.0;blur) and (PDP;bloc;0.0).

For combinations that correspond to videos with the three types of artifact signals, the average MSV_{bloc} is higher than the average MSV_{pck} and MSV_{blur}. Figure 5 shows plots of combinations with different values of packet-loss, blockiness, and blurriness strengths. An RM-ANOVA test shows that there are significant statistical differences between MSVs for most combinations of (PDP;bloc;blur), except for the combination pairs ((0.7;0.4;0.4), (0.7;0.4;0.6)) and ((8.1;0.4;0.4), (8.1;0.4;0.6)) in MSV_{pck}. Although only the strength of blurriness vary in both combination pairs, MSV_{bloc} also increases as MSV_{blur} increases. This result suggests that the blockiness is affected by increasing the blurriness. For the combination pairs ((0.7;0.4;0.4),(8.1;0.4;0.4)) and ((0.7;0.4;0.6), (8.1;0.4;0.6)), the MSV_{bloc} and MSV_{blur} differences are not statistically significant. For these combinations, the MSVs variations are higher for MSV_{bloc} than for MSV_{blur}. These results support the assumption that packet-loss artifacts intensify the perception of blockiness artifacts. An RM-ANOVA shows that there are significant statistical differences between MAVs for most combinations of (PDP;bloc;blur), except for the combination pairs ((8.1;0.4;0.4), (0.7;0.4;0.6)) and ((8.1;0.6;0.4), (0.7;0.6;0.6)).

Table 4: Exps. 5-6: Fitting parameters for linear model without intercept ($PA_{E3,L1}$) (* Significant at 0.05 level.)

Coefficient	Estimate	Std. Error	t-value	$Pr(>\|t\|)$	PCC	SCC
α	0.340	0.022	18.330	$<2e-16*$		
β	0.470	0.020	23.210	$<2e-16*$	0.937	0.936
γ	0.413	0.026	16.040	$<2e-16*$		

Table 5: Exps. 5-6: Fitting parameters for linear model with intercept ($PA_{E3,L2}$). (* Significant at 0.05 level.)

Coefficient	Estimate	Std. Error	t-value	$Pr(>\|t\|)$	PCC	SCC
δ	3.846	1.870	2.057	0.042*		
α	0.370	0.026	14.313	$<2e-16*$	0.937	0.937
β	0.456	0.021	21.448	$<2e-16*$		
γ	0.371	0.033	11.326	$<2e-16*$		

Annoyance Models

We tested a set of linear models fitting them on the MSV and MAV data of Experiments 5-6. We chose these experiments because they contained all three artifacts. The first tested linear model is a simple linear model, without any interaction term:

$$PA_{E3,L1} = \alpha \cdot MSV_{pck} + \beta \cdot MSV_{bloc} + \gamma \cdot MSV_{blur}. \quad (1)$$

Next, we adapt Eq. 1 to include an intercept coefficient (δ):

$$PA_{E3,L2} = \delta + \alpha \cdot MSV_{pck} + \beta \cdot MSV_{bloc} + \gamma \cdot MSV_{blur}. \quad (2)$$

Tables 4 and 5 show the fitting results for both models. Notice that all coefficients (i.e. δ, α, β, and γ) are statistically significant (see Columns 5 in Tables 4 and 5, respectively).

Table 6: Exps. 5-6: Fitting parameters for the linear metric with interactions $PA_{L3,E3}$ (* Significant at 0.05 level).

Coefficient	Estimate	Std. Error	t-value	Pr $(>\lvert t\rvert)$	PCC	SCC
α	5.476e-01	3.572e-02	15.327	$<2e-16*$		
β	5.470e-01	4.535e-02	12.062	$<2e-16*$		
γ	4.432e-01	3.530e-02	12.558	$<2e-16*$		
ρ_1	-2.918e-03	1.054e-03	-2.768	0.006*	0.956	0.947
ρ_2	-3.414e-03	1.321e-03	-2.585	0.011*		
ρ_3	-1.855e-04	1.277e-03	-0.145	0.885		
ρ_4	1.908e-05	2.834e-05	0.673	0.502		

Table 7: Exps. 5-6: Fitting parameters for the linear metric with interactions and an intercept term $PA_{L3,E4}$ (* Significant at 0.05 level).

Coefficient	Estimate	Std. Error	t-value	Pr $(>\lvert t\rvert)$	PCC	SCC
δ	-1.857e+01	2.768e+00	-6.710	5.22e-10*		
α	8.516e-01	5.488e-02	15.516	$<2e-16*$		
β	8.411e-01	5.888e-02	14.286	$<2e-16*$		
γ	7.670e-01	5.713e-02	13.424	$<2e-16*$	0.965	0.957
ρ_1	-7.729e-03	1.161e-03	-6.654	6.93e-10*		
ρ_2	-8.740e-03	1.393e-03	-6.274	4.66e-09*		
ρ_3	-5.488e-03	1.360e-03	-4.036	9.17e-05*		
ρ_4	1.062e-04	2.778e-05	3.821	0.000*		

Since we are also interested in understanding if the perceptual strengths interact with one another, we tested a linear model with interactions, as given by:

$$\begin{aligned} PA_{E3,L3} &= \alpha \cdot MSV_{pck} + \beta \cdot MSV_{bloc} + \gamma \cdot MSV_{blur} \\ &+ \rho_1 \cdot MSV_{pck} \cdot MSV_{bloc} + \rho_2 \cdot MSV_{pck} \cdot MSV_{blur} \\ &+ \rho_3 \cdot MSV_{bloc} \cdot MSV_{blur} + \rho_4 \cdot MSV_{pck} \cdot MSV_{bloc} \cdot MSV_{blur}. \end{aligned} \quad (3)$$

We also adapt Eq. 3 to include an intercept coefficient (δ):

$$\begin{aligned} PA_{E3,L4} &= \delta + \alpha \cdot MSV_{pck} + \beta \cdot MSV_{bloc} + \gamma \cdot MSV_{blur} \\ &+ \rho_1 \cdot MSV_{pck} \cdot MSV_{bloc} + \rho_2 \cdot MSV_{pck} \cdot MSV_{blur} \\ &+ \rho_3 \cdot MSV_{bloc} \cdot MSV_{blur} + \rho_4 \cdot MSV_{pck} \cdot MSV_{bloc} \cdot MSV_{blur}. \end{aligned} \quad (4)$$

Tables 6 and 7 show the fitting results for both models. Notice that most first, second, and third order coefficients are statistically significant (Columns 5 in Tables 6 and 7, respectively). The exceptions are ρ_3 and ρ_4 in $PA_{E3,L3}$ (see Table 6), which correspond to the interaction of (bloc;blur) and (PDP;bloc;blur). Notice also that most second order coefficients are negative, what may indicate masking effects, i.e. when two artifacts are present, one of them may attenuate the strength of the other artifact(s).

The interaction coefficient with highest magnitude corresponds to the interaction (PDP;blur), which suggest that packet-loss artifacts affect how blurriness artifacts are perceived.

Conclusions

We presented the methodology, statistical analysis, and conclusions of six psychophysical experiments. The goals of these experiments were to measure the artifact strengths and annoyance scores of videos with different combinations of blockiness, blurriness, and packet-loss artifacts. We also wanted to understand how these artifacts combine and interact to produce overall annoyance. The results showed that, when the artifact signals were presented alone at a high strength, subjects were able to identify them correctly. At low strengths, on the other hand, other artifacts were reported. Annoyance increased with both the number of artifacts and their strength.

Annoyance models were obtained by combining the artifact perceptual strengths using linear models, with and without intercepts and with and without interaction terms. Performing RM-ANOVA tests, we found that all types of artifact signal strengths had a significant effect in the overall annoyance. The tests also indicated that there were interactions among some of the artifact perceptual strengths. In summary, results show that annoyance can be modeled as a multidimensional function of the individual artifact strengths.

These results indicate that a blind image quality assessment method based on artifact measurements is indeed a valid approach. Nevertheless, although annoyance cannot be predicted using only one individual artifact signal measurement, it is not necessary to use all possible artifacts. It suffices to use the most (perceptually) significant artifacts. For example, blockiness seems to have the biggest effect on MAV. Finally, results show that there are interactions among artifact signals. Therefore, while designing quality models, it is important to take this into consideration to avoid underestimating or overestimating quality.

Acknowledgments

This work was supported in part by Conselho Nacional de Desenvolvimento Cientfico e Tecnológico (CNPq), in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and in part by the University of Brasília (UnB).

References

[1] M. C. Farias, No-reference and reduced reference video quality metrics: New contributions, Ph.D. thesis, University of California, Santa Barbara, California (2004).

[2] W. Lin, C.-C. J. Kuo, Perceptual visual quality metrics: A survey, Journal of Visual Communication and Image Representation 22 (4) (2011) 297–312.

[3] A. K. Moorthy, A. C. Bovik, Visual quality assessment algorithms : What does the future hold?, Intern. Journal of Multimedia Tools and Applic., Vol: 51 (2) (2011) 675–696.

[4] M. H. Pinson, S. Wolf, An objective method for combining multiple subjective data sets, in: Visual Communications and Image Processing 2003, International Society for Optics and Photonics, 2003, pp. 583–592.

[5] A. Ebrahimi Moghadam, P. Mohammadi, Subjective and objective quality assessment of image: A survey, Majlesi Journal of Electrcial Engineering 9.

[6] D. M. Chandler, Seven challenges in image quality assessment: past, present, and future research, ISRN Signal Processing 2013.

[7] A. F. Silva, M. C. Farias, J. A. Redi, Assessing the influence of combinations of blockiness, blurriness, and packet loss impairments on visual attention deployment, in: IS&T/SPIE Electronic Imaging, Vol. 9394, 2015.

[8] A. Ahumada, C. H. Null, Image quality: A multidimensional problem, Digital images and human vision (1993) 141–148.

[9] V. Kayargadde, J.-B. Martens, Perceptual characterization of images degraded by blur and noise: model, JOSA A 13 (6) (1996) 1178–1188.

[10] M. Nijenhuis, F. Blommaert, Perceptual error measure for sampled and interpolated images, Journal of Imaging Science and Technology 41 (3) (1997) 249–258.

[11] H. de Ridder, Minkowski-metrics as a combination rule for digital-image-coding impairments, in: SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology, International Society for Optics and Photonics, 1992, pp. 16–26.

[12] H. de Ridder, G. M. Majoor, Numerical category scaling: an efficient method for assessing digital image coding impairments, in: SC-DL tentative, International Society for Optics and Photonics, 1990, pp. 65–77.

[13] J. Caviedes, J. Jung, No-reference metric for a video quality control loop, Proc 13 (2001) 290–5.

[14] Z. Wang, H. R. Sheikh, A. C. Bovik, No-reference perceptual quality assessment of jpeg compressed images, in: Image Processing. 2002. Proceedings. 2002 International Conference on, Vol. 1, IEEE, 2002, pp. I–477.

[15] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, Perceptual blur and ringing metrics: application to jpeg2000, Signal processing: Image communication 19 (2) (2004) 163–172.

[16] H. Hadizadeh, I. V. Baji, No-reference image quality assessment using statistical wavelet-packet features, Pattern Recognition Letters 80 (2016) 144 – 149.

[17] K. Bahrami, A. C. Kot, A fast approach for no-reference image sharpness assessment based on maximum local variation, Signal Processing Letters, IEEE 21 (6) (2014) 751–755.

[18] S. A. Golestaneh, D. M. Chandler, No-reference quality assessment of jpeg images via a quality relevance map, Signal Processing Letters, IEEE 21 (2) (2014) 155–158.

[19] L. Li, W. Lin, H. Zhu, Learning structural regularity for evaluating blocking artifacts in jpeg images, Signal Processing Letters, IEEE 21 (8) (2014) 918–922.

[20] L. Li, Y. Zhou, W. Lin, J. Wu, X. Zhang, B. Chen, No-reference quality assessment of deblocked images, Neurocomputing 177 (2016) 572 – 584.

[21] L. Li, H. Zhu, G. Yang, J. Qian, Referenceless measure of blocking artifacts by tchebichef kernel analysis, Signal Processing Letters, IEEE 21 (1) (2014) 122–125.

[22] M. C. Farias, S. K. Mitra, No-reference video quality metric based on artifact measurements, in: Image Processing, 2005. ICIP 2005. IEEE International Conference on, Vol. 3, IEEE, 2005, pp. III–141.

[23] A. F. Silva, M. C. Farias, J. A. Redi, Annoyance models for videos with spatio-temporal artifacts, In: 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, 2016.

[24] A. F. Silva, M. C. Farias, J. A. Redi, Perceptual annoyance models for videos with combinations of spatial and temporal artifacts, IEEE Transactions on Multimedia.

[25] International Telecommunication Union, ITU-T Recommendation BT.500-8: Methodology for the subjective assessment of the quality of television pictures (1998).

[26] International Telecommunication Union, ITU-T Recommendation P.930: Principles of a reference impairment system for video (1996).

[27] M. C. Farias, S. K. Mitra, Perceptual contributions of blocky, blurry, noisy, and ringing synthetic artifacts to overall annoyance, Journal of Electronic Imaging 21 (4) (2012) 043013–043013.

[28] M. C. Farias, J. M. Foley, S. K. Mitra, Detectability and annoyance of synthetic blocky, blurry, noisy, and ringing artifacts, IEEE Trans. on Signal Processing, v 55 (2007) 2954–2964.

Author Biography

Alexandre F. Silva is a professor in the Federal Institute of Triângulo Mineiro. He received his M.Sc. degree in computer science from the University of Uberlândia (UFU), Brazil, in 2006. He received his Ph.D. in computer science from the University of Brasília (UnB), Brazil, in 2017. He also spent 6 months in TU Delft (in Netherlands) as a Ph.D. researcher. His current interests include video quality assessment and video quality metrics.

Mylene C.Q. Farias received her B.Sc. degree in electrical engineering from Federal University of Pernambuco (UFPE), Brazil, in 1995 and her M.Sc. degree in electrical engineering from the University of Campinas (UNICAMP), Brazil, in 1998. She received her Ph.D. in electrical and computer engineering from the University of California Santa Barbara, USA, in 2004 for work in no-reference video quality metrics. Dr. Farias has worked as a research engineer at CPqD (Brazil) in video quality assessment and validation of video quality metrics. She has also worked as for Philips Research Laboratories (The Netherlands) in video quality assessment of sharpness algorithms and for Intel Corporation (Phoenix, USA) developing no-reference video quality metrics. She is currently an assistant professor at the Department of Electrical Engineering of the University of Brasília (UnB), Brazil. Her current interests include video quality metrics, video processing, multimedia, watermarking, and compressive sensing. Dr. Farias is a member of IEEE and IEEE Signal Processing Society.