On the Stimulation Frequency in SSVEP-based Image Quality Assessment

Sebastian Bosse*, Milena T. Bagdasarian*, Wojciech Samek*, Gabriel Curio†, and Thomas Wiegand‡
*Department for Video Coding and Analytics, Fraunhofer HHI, Berlin
†Department of Neurology and Clinical Neurophysiology, Chartié, Berlin
‡Image Communication Laboratory, Technical University Berlin, Berlin

***Abstract*—Steady-state visual evoked potentials (SSVEP) are brain responses elicited by periodic visual stimuli. Recently it was shown that the use of SSVEP in quality studies allows for accurate psychophysiological assessment of perceived visual quality, but the influence of the stimulation frequency is still unclear. This paper studies experimentally the relation between the SNR of the neural signal and the stimulation frequency in an psychophysiological quality assessment setup. For various source images tested at different distortion magnitudes over the range of 6 different stimulation frequencies, we show physiologically plausible results that provide insights into the temporal dynamics of neural distortion processing. Our findings inform a rational choice of stimulation frequency in SSVEP-based image quality assessment studies. This potentially improves the experimental setup of future image quality assessment studies exploiting the SSVEP paradigm.**

***Index Terms*—Image quality assessment, SSVEP, stimulation frequency, electroencephalography**

I. Introduction

Accurate estimation of perceived visual quality is crucial for operating, optimizing and evaluating most modern image and video communication systems, and has been an active research field since decades [1]. In recent years, a lot of progress has been made, yet a precise model for the perception of visual quality is not at hand. Thus, quantifying the perceptual properties of image processing and communication systems still relies on the collection of human responses when presented with a stimulus of interest, such as an image or video of a certain quality. Conventionally, these responses are obtained as overt, behavioral quality ratings during psychophysical tests [1]. Ratings are typically averaged over subjects and perceptual quality is reported as mean opinion score (MOS) [2]. These ratings are the results of a conscious process, typically reported in Likert-style questionaires [2], and as such, prone to subjective factors, such as decision strategies, expectations or fatigue [3]. Semantic annotations of rating scales given to participants during psychophysical tests may also fail at reflecting the participants' appraisal of the stimulus and thereby mislead given responses [3]. Another limitation of psychophysical approaches to multimedia quality assessment is its restriction to supra-threshold stimuli and its insensitivity to sub-threshold stimuli. Depending on explicit conscious responses these methods provide only limited insight into internal perceptual and cognitive processes underlying the decision making in quality assessment [4], [5]. These limitations of psychophysical methods led researchers to study psychophysiological approaches in the context of quality assessment. Psychophysiological approaches are complementary to classical psychophysical ones – in fact, Gustav Theodor Fechner postulated *inner* psychophysics already 1907 as a neural foundation of *outer* psychophysics [6]. Due to its relative modest cost in acquisition and operation, and its adequately high temporal resolution, electroencephalography (EEG) is one of the most popular methods used in psychophysiological quality assessment and other domains of BCI (Brain Computer Interfacing) [5].

Most studies investigate the relation between event related potentials (ERPs) to perceived quality for different types of media modalities. ERPs are brain responses that are elicited by a specific sensory or cognitive event, such as an unexpected change in quality. [7]–[10] show significant correlations between ERPs and MOS for distorted audio and speech signals. [11] studies the assessment of JPEG distortions exploiting ERPs. [12]–[15] show that also video quality can be assessed using ERPs. [16] shows a relation between vertical disparities in stereoscopic images and the neural signal measured as ERPs. Another line of research shows a change in the spectral power distribution related to perceptual quality for 3D videos subject to coding distortions [17], [18]. Also other psychophysiological signals such as electromyography (EMG) or eye movements are studied for quality assessment [19], [20].

In contrast to ERPs, being transient responses elicited by a single stimulus change, steady-state visual evoked potentials (SSVEPs) are neural responses elicited by a train of changes [21]. As a key feature, SSVEPs are confined to a set of narrow band frequency bins centered at the frequency of the stimulus change and its harmonics. Therefore, the signal-to-noise ratio (SNR) is typically high as only a small fraction of the broad band noise lies in the same frequency bins [21] the signal lives in. For image quality assessment SSVEPs have been studied at a stimulation frequency of $f_{stim} = 1.5\,\mathrm{Hz}$ [22]–[25]. While it is known that for different cognitive tasks the stimulation frequency has an impact on the properties of the neural signal, e.g., for face discrimination [26] or perception of written text [27], its influence on the neural detectability of changes in image quality is not clear.

This work provides insights into the impact of the stimulation frequency on image quality perception. 6 distorted images based on 3 source reference images at 2 distortion magnitudes are presented in an SSVEP alternation paradigm at 6 different

stimulation frequencies. Our results show a dependency of the SNR of the extracted SSVEP on the stimulation frequency. Our findings inform a rational choice of stimulation frequencies. In Sec. II the experimental setup is explained. Sec. III describes the data analysis. Results are presented and discussed in Sec. IV. Sec. V concludes the article.

II. Experimental Setup

The presented experiment consisted of two parts. In the first part, perceived quality was assessed in a conventional psychometric test, in the second part EEG data was recorded during stimulus presentation. Both experimental parts were conducted under identical viewing conditions: Stimuli were shown on a 27" screen (DELL U23711) with a native resolution of 2560×1440 pixels at a refresh rate of $f_r = 60\,\mathrm{Hz}$. The screen was calibrated according to the specifications in [28]. Stimuli were shown in native resolution without any scaling. The viewing distance was set with regard to the resolution of the HD source content (see Sec. II-B) leading to a viewing distance of $1.0\,\mathrm{m}$ [28]. The psychometric part took around $10\,\mathrm{min}$, the neurophysiological part lasted about $1\,\mathrm{h}$, excluding preparation time. 9 subjects (all male, in the age group 25-28) with normal or corrected-to-normal vision participated in the experiment and were monetarily compensated for participation.

A. Choice of Studied Stimulation Frequencies

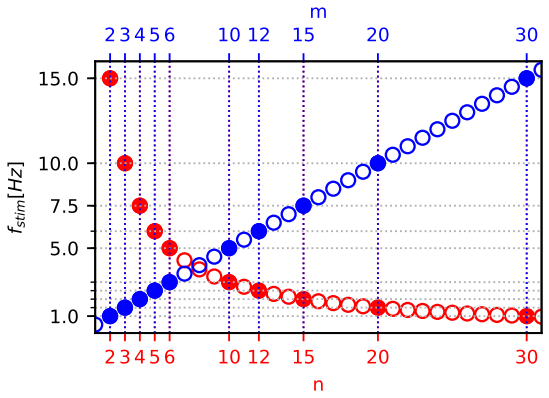


Fig. 1: Constraints on the set of possible stimulation frequencies for a display refresh rate $f_r = 60\,\mathrm{Hz}$ and a stimulation duration $T_{stim} = 2\,\mathrm{s}$. Blue circles: Constraint posed by stimulus duration. Red circles: Constraint posed by display refresh rate. Possible are only stimulation frequencies conforming to both constraints, indicated by filled circles.

As it is not possible to update the stimulus on screen without refreshing the display the set of possible stimulation frequencies for an SSVEP-base quality study is constraint by rate f_r of the display device used.
Since SSVEPs are effectively represented in the frequency domain, a second constrained is raised by the implicit assumption of a periodic signal by the Fourier transform. In order to avoid leakage, full cycles of the periodic stimulation have to fit into the temporal analysis window of the length T_{stim} [29]. Thus, a stimulation frequency f_{stim} has to satisfy the two conditions

$$f_{stim} \cdot T_{stim} = m \tag{1}$$

$$\frac{f_r}{2 \cdot f_{stim}} = n \quad \text{with} \quad m, n \in \mathbb{N}. \tag{2}$$

In order to achieve a frequency resolution of $0.5\,\mathrm{Hz}$, a stimulation duration of $T_{stim} = 2\,\mathrm{s}$ is used. The display frame rate is $f_r = 60\,\mathrm{Hz}$. Stimulation frequencies allowed by $T_{stim} = 2\,\mathrm{s}$ as given by Eq. 1 are indicated by blue circles, stimulation frequencies allowed by $f_r = 60\,\mathrm{Hz}$ (Eq. 2) are indicated by red circles. Filled circles indicate stimulation frequencies that satisfy both constraints. The resulting set of possible stimulation frequencies $[1,2,2.5,3,5,6,7.5,10,15,30]\,\mathrm{Hz}$ is reduced to $[2,3,5,6,7.5,10]\,\mathrm{Hz}$ in order to decrease the number of parameters in the experiment, as pilot experiments revealed a low SNR on the first two harmonics for stimulation frequencies higher than $10\,\mathrm{Hz}$ and lower than $2\,\mathrm{Hz}$.

B. Stimulus Material

Stimuli were based on three 768×768 pixel sized image patches cropped from HD video sequences. These patches, shown in Fig. 2, were selected in order to provide roughly texture-like yet diverse visual properties and served as source reference images. Visual quality of each of the source reference images was degraded to two different quality levels. Distortions were introduced by coding these images using the HM16.0 test model [30] of High Efficiency Video Coding standard (HEVC) [31] using *intra only* setting [32]. Quantization Parameters (QPs) used to generate the distorted patches were selected in order to meet quality levels of approximately similar perceptual qualities for the source reference image. For this, perceptual quality was estimated using HaarPSI [33], aiming at target values of 0.8 and 0.65. The resulting QP values were {36,41} for Woods, {35,42} for CrowdRun, and {39,47} for SunFlower.

C. Measurement of Behavioral Responses

Quality was assessed psychophysically following the degradation category rating (DCR) procedure with simultaneous presentation (SP) [28]. The undistorted source reference image was shown on the left hand side simultaneously with the distorted image on the right hand side for duration of $10\,\mathrm{s}$ or until subjects reported a quality judgement. During presentation, subjects were asked to report their quality judgement using a slider on screen. For rating, a five-grade degradation scale was used with the semantic annotation 5-*Imperceptible*, 4-*Perceptible, but not annoying*, 3-*Slightly annoying*, 2-*Annoying*, and 1-*Very annoying*.

D. Measurement of Neurophysiological Responses

In order to elicit SSVEP, during the neurophysiological part undistorted and distorted versions of the stimuli were presented in alternation at stimulus frequencies $f_{stim} \in [2, 3, 5, 6, 7.5, 10]\,\mathrm{Hz}$. Image sequences were presented for $10\,\mathrm{s}$

with the first and last second being excluded from further analysis. Each combination of source reference image, distortion level and stimulation frequency was presented in 7 trials in random order. Each trial started with a 1 s-presentation of a fixation point in the center of the active screen.

EEG was recorded at a sampling frequency of 1 kHz using BrainAmp amplifiers and an ActiCap active electrode system with 64 sensors (both by Brain Products, Munich, Germany). The electrodes were positioned at Fp1,2, AF3,4,7,8, Fz, F1-10, FCz, FC1-6, FT7,8 Cz, C1-6, T7, CPz, CP1-6, TP7,8, Pz, P1-10, POz, PO3,4,7,8, Oz, O1,2. The electrode that in the standard EEG montage is placed at T8 was placed under the right eye and used to measure eye movements. All electrodes were referenced to the left mastoid using a forehead ground. All impedances were kept below 10 kΩ.

III. Data Analysis

A. Behavioral Data

In psychophysical tests, some observers might give inconsistent responses that can distort the result of the test. Those observers can be identified by screening and should be rejected for further analysis as recommended in [34]. Mean opinion scores (MOS) are obtained by averaging condition-wise over the ratings reported by individual observers.

B. Neurophysiological Data

For offline analysis, EEG data is bandpass filtered using a 3rd order Butterworth filter with a −3 dB passband from 0.8 Hz to 40 Hz in order to attenuate line noise and to remove drifts and DC-offset. EEG data is downsampled to 120 Hz. The influence of horizontal eye movement is regressed out from the difference of the signals recorded at F7 and F8, vertical eye movement from the difference of the signals from VEOG and FP2. EEG data is re-referenced to common average reference (CAR). EEG data is epoched into 4 non-overlapping segments of the original 8 s trial records. This essentially increases the number of trials for the price of a reduction of stimulus duration, and, thus, to a spectral resolution of 0.5 Hz. Epochs with more than 20% of samples exceeding ±25 µV are excluded. Typically, these epochs are associated with strong eye movements, blinks or other body movement that could not be regressed out. The SSVEP was extracted from 2 s-epochs as discrete Fourier transform (DFT) coefficients. SNR per frequency bin is estimated as the ratio between the power in a frequency bin and the mean of the power in the two neighboring frequency bins. EEG data from O1, O2, Oz is coherently averaged over all trials, subjects, and channels. For data analysis the Wyrm toolbox [35] was used.

IV. Results

The MOS values of the stimuli as obtained in the behavioral part of the study are shown in Fig. 3. The results do not show a split of the stimuli into two groups of distortion level of similar perceptual quality. Thus, further analysis considers all stimuli jointly by averaging neural data for all conditions.

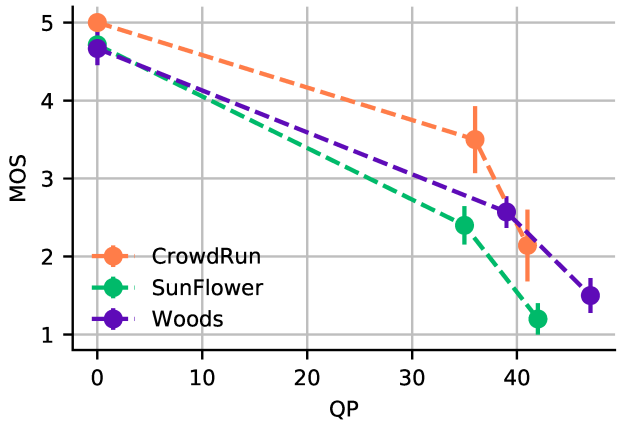
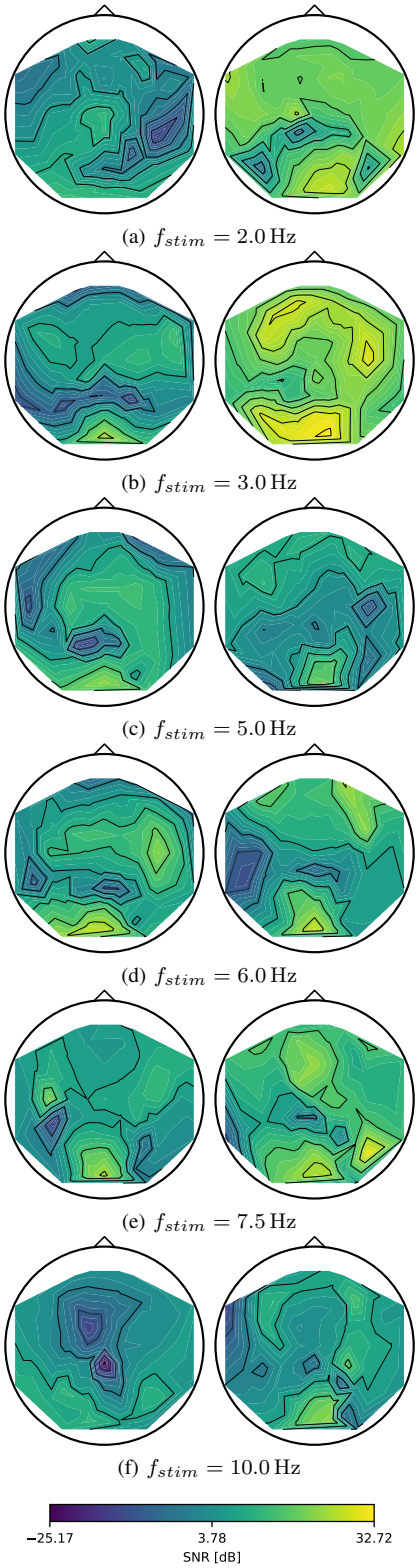


Fig. 3: MOS values of the stimuli used in the experiment

Fig. 4 exemplify the scalp topographies of the SNR for the first 2 harmonics $1f_{stim}$, $2f_{stim}$ of the stimulation frequency f_{stim} for $f_{stim} \in [2, 3, 5, 6, 7.5, 10]$ Hz for one subject. While for $f_{stim} \in [3, 5, 6, 7.5]$ Hz the SNR concentrates around occipital electrode positions, this concentration is partially reduced only for $f_{stim} \in [2, 10]$ Hz, in particular for the first harmonic.

Accordingly, the time course of the EEG signal averaged over the channels O1, Oz and O2 is shown in Fig. 6. For all considered stimulation frequencies $f_{stim} < 10$ Hz the EEG signal shows a clear modulation with the visual stimulation. For $f_{stim} = 10$ Hz this modulation decreases, also resulting in a lower amplitude of the signal. The according spectra of the SNR are shown in the 2nd and 4th row of Fig. 6, with red denoting the SNR on the harmonics of the stimulation frequency. For all stimulation frequencies the spectral power distribution of the signal is dominated by the first 2 harmonics. The dependency of the SNR of the stimulation frequency on the first 4 harmonics is summarized in Fig. 5. For $f_{stim} \in [5, 6]$ Hz the SNR at the 1st harmonic exceeds equals the SNR at the 2nd harmonic. SNR at 2nd harmonics exhibit peaks for $f_{stim} = 3$ Hz and $f_{stim} = 6$ Hz and a drop at $f_{stim} = 5$ Hz where the SNR of the 1st harmonic peaks. Higher harmonics (3rd and 4th) both exhibit a peak in SNR at $f_{stim} = 5$ Hz (and possibly below 2 Hz). SNR at the 4th harmonic displays a drop for $f_{stim} = 6$ Hz and a peak for $f_{stim} = 7.5$ Hz. SNR at the 3rd and 4th harmonic drops for $f_{stim} = 3$ Hz and for $f_{stim} \geq 7.5$ Hz.

The relation between SNR and stimulation frequency can be understood in parts from the power spectral density (PSD) of the EEG, as the theta band (approx. 4–7 Hz) has a lower level of activity as compared to the delta (approx. 1–3 Hz) and alpha (approx. 8–12 Hz) bands [36]. The PSD of the EEG recorded from unstimulated, open-eyed subjects is shown for comparison in Fig. 7. The 3rd and 4th harmonics of $f_{stim} = 3$ Hz are buried in the alpha activity, as well as the 2nd harmonic of $f_{stim} = 5$ Hz and the 1st harmonic of $f_{stim} = 10$ Hz, whereas the 1st and 2nd harmonics of $f_{stim} = 2$ Hz are affected by delta activity, leading to a



(a) $f_{stim} = 2.0\,\text{Hz}$

(b) $f_{stim} = 3.0\,\text{Hz}$

(c) $f_{stim} = 5.0\,\text{Hz}$

(d) $f_{stim} = 6.0\,\text{Hz}$

(e) $f_{stim} = 7.5\,\text{Hz}$

(f) $f_{stim} = 10.0\,\text{Hz}$

Fig. 4: Topographical distribution of the SNR on the first two harmonics for one subject (left column 1st harmonic, right column 2nd harmonic). Yellow indicates higher values, green indicates lower values (see colorbar).

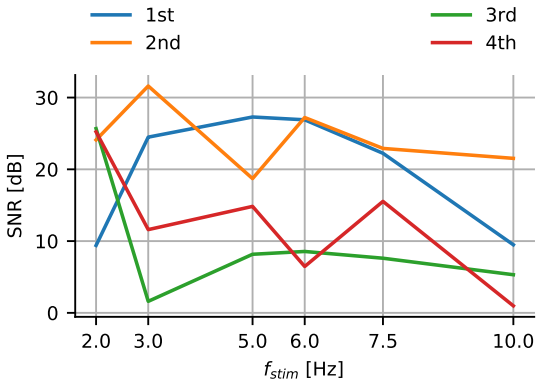


Fig. 5: SNR over different stimulation frequencies for first 4 harmonics averaged over channels O1, Oz, O2. (grand average over all subjects).

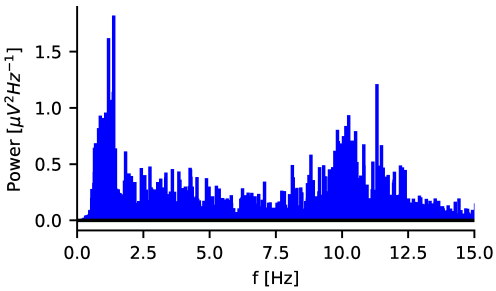


Fig. 7: Power spectrum of the EEG recorded from subjects having eyes open without controlled stimulation.

decrease of the related SNRs. 3rd and 4th harmonics of $f_{stim} = 2\,\text{Hz}$, and 1st harmonic of $f_{stim} = 6\,\text{Hz}$ on the other hand are less affected by the low background activity in the theta band.

In [24] a high correlation between the SSVEP at the amplitude at the 4th harmonic for a stimulation frequency $f_{stim} = 1.5\,\text{Hz}$ is reported but not explained. Extrapolating from Fig. 5 suggests that for $f_{stim} = 1.5\,\text{Hz}$ the SNR of the 4th harmonic is higher than e.g. the SNR of the 1st harmonic. This explains the superior suitability of the amplitude of the 4th harmonic at $f_{stim} = 1.5\,\text{Hz}$ as a neural marker in quality assessment.

Even harmonics of the SSVEP represent responses to onset and offset of the stimulus change and capture low-level properties of the stimulus such as contrast or luminance changes. Odd harmonics are responses to the onset only and often related to higher-level properties of the stimulus change [21]. The relation between even and odd, respectively, harmonics and perceptual quality is not clear. A balanced SNR between the first odd and the first even harmonic of the neural signal therefore seems to be a reasonable criterion. This balance is achieved at $f_{stim} = 6\,\text{Hz}$ (see. Fig. 5). Moreover, a stimulation frequency $f_{stim} = 6\,\text{Hz}$ shows the highest SNR at the first harmonic, which can be assumed to carry the signature of

neural distortion processing as it captures the stimulus onset.

V. Conclusion

We presented experimental results on the influence of the stimulation frequency in a SSVEP-based image quality assessment setup. Due to the curse of dimensionality imposed by the number of considered stimulation frequencies only 6 different stimuli could be used in the experiment. This renders the estimation of correlations between the neural signal and behavioral responses, e.g. MOS unreliable and thus the SNR was used as a proxy quantity. Our results show that a stimulation frequency of $f_{stim} = 6\,\mathrm{Hz}$ achieves a high SNR at the first harmonic and a balanced relation of SNR at the first two harmonics, indicating to be a favourable choice in quality assessment studies. On a reduced set of considered stimulation frequencies and a larger set of stimuli this should be validated in terms of correlation. Also the relation between odd and even harmonics, respectively, and perceived quality should be investigated in order to arrive at a final conclusion. Potential dependencies on the source reference images and its spatial statistics should be taken into account. Sophisticated channel decomposition techniques such as spatio-spectral decomposition (SSD) [37] were shown to be useful for SSVEP-based quality assessment [24] and could be used to analyze subject dependencies in future studies.

References

[1] Y. Chen, K. Wu, and Q. Zhang, "From QoS to QoE: A Tutorial on Video Quality Assessment," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 1126–1165, 2015.

[2] R. C. Streijl, S. Winkler, and D. S. Hands, "Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives," *Multimedia Systems*, vol. 22, no. 2, pp. 213–227, 2016.

[3] D. R. Riskey, "Use and Abuses of Category Scales in Sensory Measurement," *Journal of Sensory Studies*, vol. 1, no. 3-4, pp. 217–236, dec 1986.

[4] U. Engelke, D. P. Darcy, G. H. Mulliken, S. Bosse, M. G. Martini, S. Arndt, J.-N. N. Antons, K. Y. Chan, N. Ramzan, K. Brunnstrom, and K. Brunnström, "Psychophysiology-Based QoE Assessment: A Survey," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 1, pp. 6–21, feb 2017.

[5] S. Bosse, K.-R. Müller, T. Wiegand, and W. Samek, "Brain-Computer Interfacing for Multimedia Quality Assessment," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 002 834–002 839.

[6] G. T. Fechner, *Elemente der Psychophysik. 2*. Breitkopf & Härtel, 1907.

[7] J.-N. Antons, R. Schleicher, S. Arndt, S. Möller, A. K. Porbadnigk, and G. Curio, "Analyzing Speech Quality Perception Using Electroencephalography," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 6, pp. 721–731, oct 2012.

[8] A. K. Porbadnigk, J.-N. Antons, B. Blankertz, M. S. Treder, R. Schleicher, S. Möller, and G. Curio, "Using ERPs for assessing the (sub) conscious perception of noise," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, vol. 2010. IEEE, jan 2010, pp. 2690–2693.

[9] A. K. Porbadnigk, M. S. Treder, B. Blankertz, J. N. Antons, R. Schleicher, S. Möller, G. Curio, and K.-R. Müller, "Single-trial analysis of the neural correlates of speech quality perception." *Journal of Neural Engineering*, vol. 10, no. 5, p. 056003, oct 2013.

[10] S. Uhrig, S. Arndt, S. Möller, and J.-N. Voigt-Antons, "Perceptual references for independent dimensions of speech quality as measured by electroencephalography," *Quality and User Experience*, vol. 2, no. 1, p. 10, 2017.

[11] L. Lindemann and M. Magnor, "Assessing the quality of compressed images using EEG," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 3109–3112.

[12] S. Arndt, J.-N. Antons, R. Schleicher, S. Möller, and G. Curio, "Using electroencephalography to measure perceived video quality," *IEEE Journal on Selected Topics in Signal Processing*, vol. 8, no. 3, pp. 366–376, 2014.

[13] L. Lindemann, S. Wenger, and M. Magnor, "Evaluation of video artifact perception using event-related potentials," in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization - APGV '11*, 2011, pp. 53–58.

[14] M. Mustafa, S. Guthe, and M. Magnor, "Single-trial EEG classification of artifacts in videos," *ACM Transactions on Applied Perception*, vol. 9, no. 3, pp. 1–15, 2012.

[15] S. Scholler, S. Bosse, M. S. Treder, B. Blankertz, G. Curio, K.-R. Müller, and T. Wiegand, "Toward a direct measure of video quality perception using EEG." *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2619–29, 2012.

[16] F. S. Avarvand, S. Bosse, K.-R. Müller, R. Schäfer, G. Nolte, T. Wiegand, G. Curio, and W. Samek, "Objective quality assessment of stereoscopic images with vertical disparity using EEG," *Journal of Neural Engineering*, vol. 14, no. 4, p. 046009, 2017.

[17] E. Kroupi, P. Hanhart, J.-S. Lee, M. Rerabek, and T. Ebrahimi, "EEG correlates during video quality perception," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2014, pp. 1–4.

[18] S. Arndt, J.-N. Antons, R. Schleicher, and S. Möller, "Using electroencephalography to analyze sleepiness due to low-quality audiovisual stimuli," *Signal Processing: Image Communication*, vol. 42, pp. 120–129, 2016.

[19] P. Arnau-González, T. Althobaiti, S. Katsigiannis, and N. Ramzan, "Perceptual video quality evaluation by means of physiological signals," *Proceedings of the 9th International Conference on Quality of Multimedia Experience (QoMEX)*, 2017.

[20] J. Radun, M. Nuutinen, J.-N. Antons, and S. Arndt, "Did You Notice It? How Can We Predict the Subjective Detection of Video Quality Changes From Eye Movements?" *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 1, pp. 37–47, 2017.

[21] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *Journal of Vision*, vol. 15, no. 6, pp. 1–46, 2015.

[22] S. Bosse, L. Acqualagna, A. K. Porbadnigk, B. Blankertz, G. Curio, K.-R. Müller, and T. Wiegand, "Neurally informed assessment of perceived natural texture image quality," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*. IEEE, oct 2014, pp. 1987—-1991.

[23] S. Bosse, L. Acqualagna, A. K. Porbadnigk, G. Curio, K.-R. Müller, B. Blankertz, and T. Wiegand, "Neurophysiological assessment of perceived image quality using steady-state visual evoked potentials," in *SPIE Optical Engineering+ Applications*, vol. 9599. International Society for Optics and Photonics, 2015, pp. 959 914–959 914.

[24] S. Bosse, L. Acqualagna, W. Samek, A. K. Porbadnigk, G. Curio, B. Blankertz, K.-R. Müller, and T. Wiegand, "Assessing Perceived Image Quality Using Steady-State Visual Evoked Potentials and Spatio-Spectral Decomposition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8215, no. c, pp. 1–1, 2017.

[25] L. Acqualagna, S. Bosse, A. K. Porbadnigk, G. Curio, K.-R. Müller, T. Wiegand, and B. Blankertz, "EEG-based classification of video quality perception using steady state visual evoked potentials (SSVEPs)." *Journal of Neural Engineering*, vol. 12, no. 2, p. 026012, 2015.

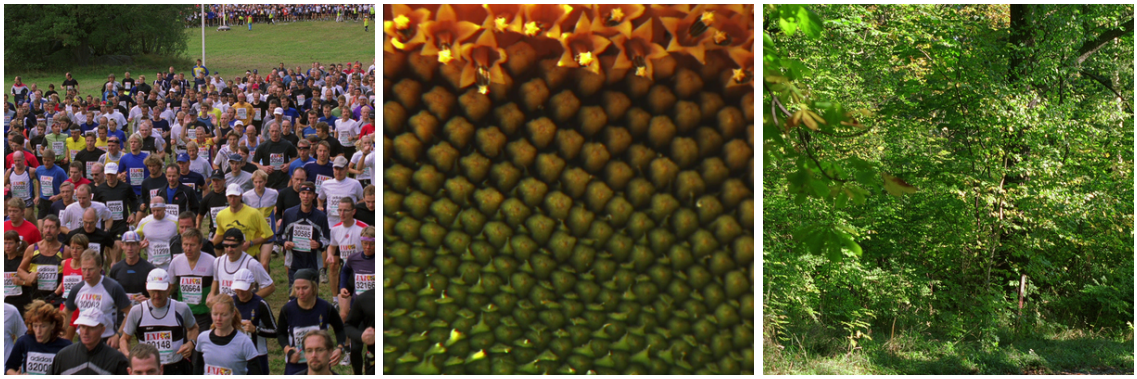
[26] E. Alonso-Prieto, G. Van Belle, J. Liu-Shuang, A. M. Norcia, and B. Rossion, "The 6 Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex." *Neuropsychologia*, vol. 51, no. 13, pp. 2863–75, nov 2013.

[27] J. D. Yeatman and A. M. Norcia, "Temporal tuning of word and face selective cortex," *Journal of Cognitive Neuroscience*, vol. 28, no. 11, pp. 1820–1827, 2016.

[28] ITU-T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications," International Telecommunication Union, Geneva, Switzerland, Tech. Rep., apr 2008.

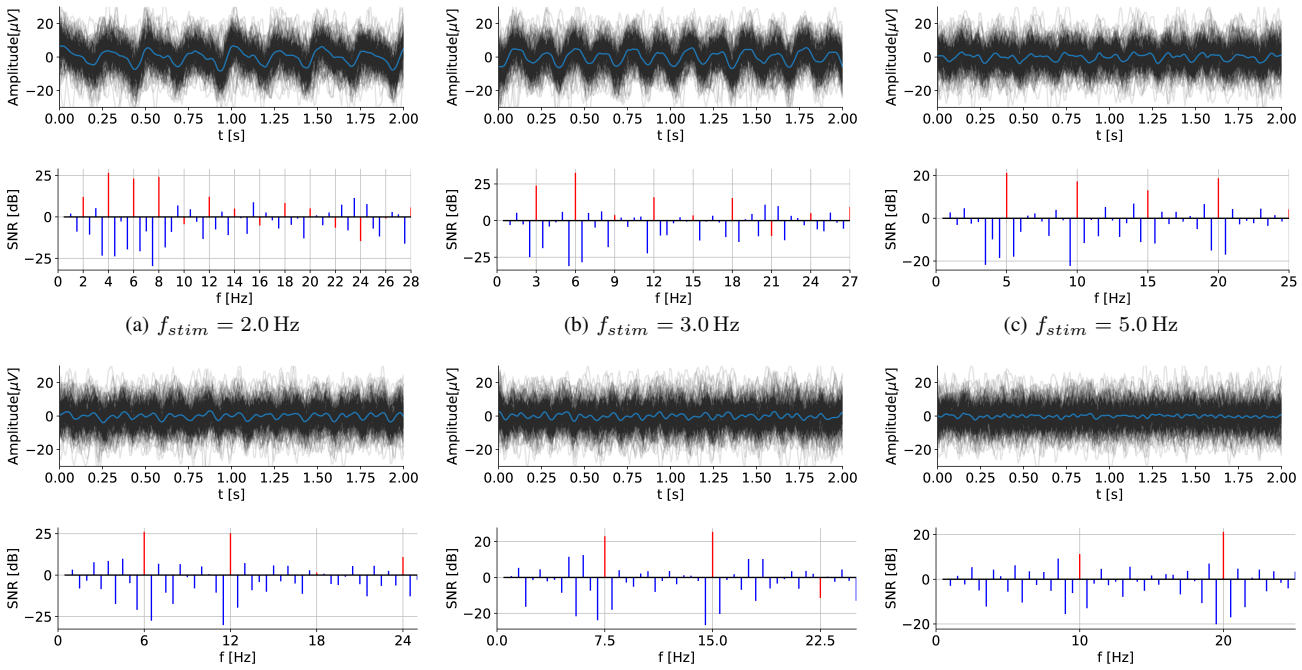
[29] M. Bach and T. Meigen, "Do's and dont's in Fourier analysis of steady-state potentials." *Documenta ophthalmologica. Advances in ophthalmology*, vol. 99, no. 1, pp. 69–82, jan 1999.

[30] JCT-VC, "Subversion Repository for the HEVC Test Model reference software," 2014.
[31] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
[32] F. Bossen, "Common test conditions and software reference configurations Output," document JCTVC-H1100 of JCT-VC, San José CA, USA, pp. 1–10, 2013.
[33] R. Reisenhofer, S. Bosse, G. Kutyniok, and T. Wiegand, "A Haar wavelet-based perceptual similarity index for image quality assessment," *Signal Processing: Image Communication*, vol. 61, 2018.
[34] ITU-R Recommendation BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," International Telecommunication Union, Geneva, Switzerland, Tech. Rep., 2012.
[35] B. Venthur, S. Dähne, J. Höhne, H. Heller, and B. Blankertz, "Wyrm: A brain-computer interface toolbox in python," *Neuroinformatics*, vol. 13, no. 4, pp. 471–486, 2015.
[36] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.
[37] V. V. Nikulin, G. Nolte, and G. Curio, "A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition," *NeuroImage*, vol. 55, no. 4, pp. 1528–1535, 2011.



(a) CrowdRun (b) SunFlower (c) Woods

Fig. 2: Source reference images used in experiment



(a) $f_{stim} = 2.0\,\mathrm{Hz}$ (b) $f_{stim} = 3.0\,\mathrm{Hz}$ (c) $f_{stim} = 5.0\,\mathrm{Hz}$

(d) $f_{stim} = 6.0\,\mathrm{Hz}$ (e) $f_{stim} = 7.5\,\mathrm{Hz}$ (f) $f_{stim} = 10.0\,\mathrm{Hz}$

Fig. 6: Time course (top and 3rd row) and SNR spectrum (2nd and bottom row) for different stimulation frequencies averaged over O1,Oz and O2 channels. Time courses: Grey indicates single trial time courses for and stimuli, average for one subject is denoted by blue. SNR spectra: SNR for harmonic frequencies is denoted by red, all other frequencies by blue.