MARYAM MUSTAFA
mustafa@cg.cs.tu-bs.de
Computer Graphics Lab, TU Braunschweig

STEFAN GUTHE
stefan.guthe@tu-clausthal.de
TU Clausthal

Prof. Dr. Ing. MARCUS MAGNOR
magnor@cg.tu-bs.de
Computer Graphics Lab, TU Braunschweig

The Human in the Loop: EEG-driven Photo Optimization

Technical Report 2013-11-20

November 14, 2013

Computer Graphics Lab, TU Braunschweig

Contents

Abstract

This paper investigates the brain's response to appealing and unappealing versions of images. We present results from several ElectroEncephaloGraph (EEG) experiments using images with varying levels of 'pleasingness' as stimuli, which shed light on the preference and perception of pleasing and displeasing image versions. An analysis of the EEG data shows a distinct and reliable difference in the neural response to image versions with the same content but different parameter values for saturation, contrast and brightness. We use this EEG data to create a neural-feedback loop to automatically optimize these parameters to render more pleasing image variations.

1 Introduction

This paper looks into the question of the 'pleasingness' of images and the human neural response to aesthetically pleasing or displeasing image variations. How are good and bad images perceived? Is there a way to measure the idea of the 'pleasingness' of a photograph or image and to model that measure to create more appealing versions of the original image?

A great deal of time, money and effort is spent on editing photographs, be they from a vacation, a wedding, a birth or a graduation. This is especially true of late with the ubiquitous use of smart phones and applications like Instagram, Aviary and Photoshop Express that allow you to edit, change and personalize photographs. Given this proliferation of photo editing applications it becomes increasingly important to understand the perception of and the preference for an image with a different 'look','feel' or 'personalization'. Is it possible to determine the individual preference for an image and to quantify that preference to create individualized image optimizations?

The problem of defining visual art and its perception is not an easy one and has an entire field of scientific study dedicated to it. Neuroaesthetics examines the neural basis of the observation and experience of works of art [Zek00]. Many cognitive and psychophysical experiments have been conducted to understand the aesthetic perception of art and the emotions it evokes [CVCM09, WKK$^+$07]. These experiments have shown that pleasant and unpleasant images evoke a reliable and measurable neural response [PBL$^+$08, SJWH04]. However, most of this work focuses on understanding the aesthetics of art and does not attempt to use the results to create more pleasing images.

There are several advantages to developing techniques that merge neuroaesthetics with image editing and optimization. An analysis of the emotional and visual information processing that occurs in the brain and that enables us to make aesthetic judgements about visual media allows boosting traditional image editing software. It encourages the development of methodologies that can predict and automatically personalize an image for an individual user. Also, unlike contemporary computer algorithms the human perceptual system perceives an image or photograph in its entirety and context, instead of just relying on low-level pictorial description or mid-level content information [WFC$^+$09]. To create aesthetically pleasing images it is important to understand global, high-level image perception that incorporates personal tastes, experiences and preferences.

In this study, we hypothesize that images with the same content but different image parameter values for saturation, contrast and brightness, that create a different 'look' or 'feel', evoke distinct, reliable and measurable neural responses. We also hypothesize that based on this neural response, it is possible to optimize the image parameters to create a more appealing, personalized version of the input image.

This paper addresses two main questions: Firstly, is it possible to measure the preference for different image versions using EEG? To answer this question

we analyze the brain's responses to professionally retouched,deliberately degraded and neutral versions of the same image (Experiment 1). This data is then used to train a support vector machine (SVM) to learn what the brain's response to a pleasing (professional) or a displeasing image (degraded) version looks like.

The second main question is if it is possible to optimize images based on EEG signals? We address this question by presenting a neural-feedback loop that varies image display parameters based on single-trial EEG data (Experiment 2). The subject's neural response to the displayed image is measured and a 'visual appeal' score is determined in real-time by the previously trained SVM. This score is then used to drive a numerical optimization routine that varies the image display parameter values, changing in turn the image being observed by the subject. The neural-feedback loop drives the displayed output towards an aesthetic optimum for that individual user (Experiment 2).

2 Related Work

Neuroaesthetics and Perception of Art Much has been written and studied about art, its aesthetic experience and perception. There have been numerous studies that have looked at the perception of art ranging from neurophysiology [Zek00, CVCM09] to psychophysical and perceptual experiments about the categorization of art [WKK+07, WFC+09]. However, our review will be limited to the most relevant works within neuroaesthetics.

Wallraven et al. [WFC+09] conduct a study to see if it is possible for an algorithm to cluster paintings the way a human can. They implement and test several computational measures sensitive to color, texture and spatial composition and found that none of the computational measures correlated with the human data. They hypothesize that this was due to the higher-order processing of the human brain when viewing works of art.

Zeki [Zek00] argues that no discussion about aesthetics of art is complete without an understanding of its neural basis. Similarly over recent years many experiments have been conducted to understand the neuroanatomical correlates of aesthetic preference [VG04, SJWH04]. Most relevant to our work are the studies which show a distinct and reliably different Event Related Potential (ERP) when viewing pleasant, unpleasant and neutral images [PBL+08, CSB+00]. Cuthbert et al. [CSB+00] conduct experiments to assess the brain's reactivity to emotional pictures by recording event-related potentials. Their results show an extended centroparietal slow wave that is significantly larger for affective than neutral pictures and was maintained over a 6-second period.

Most of this work focuses on understanding the aesthetics of art without aiming to use the results to create pleasing images. The case we consider here is that the content is neutral and has little inherent valence but the appeal of the image is governed by stylistic means. A classic example are the photographs by Ansel Adams whose aesthetic appeal is governed predominantly by contrast and brightness [ABdC80].

EEG and emotion There is considerable research that has shown an EEG to be a reliable means for determining the emotional state of a user during the performance of various tasks [Bos06, HDR08, CKGP06, MRN+08]. One hypothesis for the perception of emotion in the brain is the valence model which postulates that left frontal inactivation is an indicator of a withdrawal response, linked to negative emotion, while right frontal inactivation is a sign of a positive reaction [SW86]. For a complete overview of the models of emotional processing please refer to [DEYH05]. Based on these earlier studies we use the EEG to measure the emotional and visual response to images.

Parameter optimization Different approaches have been proposed to determine optimal parameter values for various photo editing algorithms [CCV03]. Techniques based on user feedback include Interactive Evolution [Sim91], Inverse Design [SDS+93, KPC93], and Design Galleries [MAB+97]. A recent example for automatic lighting parameter optimization was proposed by Shacked et al. [SL01].

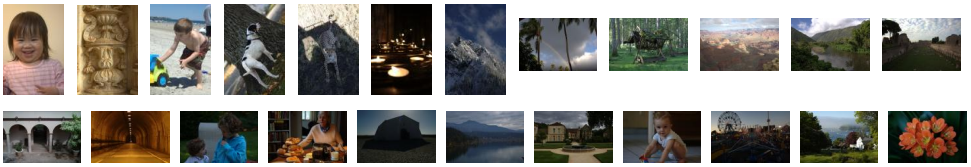


Figure 1: Original image versions used for SVM training (from MIT-Adobe FiveK Dataset)

A considerable amount of work has also been done in the area of automatic image enhancement [KLW12, HKK12]. Commercial applications like Microsoft's Office Picture Manager, Adobe's Lightroom and Adobe Photoshop also feature tools for the enhancement of photographs.

Related to our optimization approach is the work of Bychkovsky et al. [BPCD11]. In their work, a database of 5 sets of 5000 photos was adjusted by professional photographers. The result was used for supervised machine learning, focusing on global tonal adjustments. Once trained, the system is able to learn an individual's preference for future adjustments from a small set of images retouched by the user. We make use of the same dataset, and our approach also makes use of supervised learning, albeit of user EEG recordings. In contrast to Bychkovsky et al. [BPCD11], however, our method does not require each new user to go through the effort of retouching a number of photos first. The motivation behind our approach is to use EEG to quantify the notion of visual appeal and that emotional sense of "liking" a particular variation of a picture, a feeling that often-times we are unable to translate into objective image properties.

3 Experiment 1

Our studies were conducted in two steps. The first set of experiments was to determine if it is even possible to reliably detect a measurable neural response to different image versions with varying levels of pleasingness where image content stays the same.

2 male and 8 female, healthy participants of an average age of 25 years and with normal or corrected-to-normal vision took part in collecting the EEG data. All participants had average experience with digital footage and no involvement in professional image rendering or editing. The participants were instructed orally and received a training session with a small image set but the same experimental set-up to prepare them for the procedure. It is important to note that these 10 participants who provided the data for this experiment were different from the ones that took part in the neural-feed back loop testing in Experiment 2.

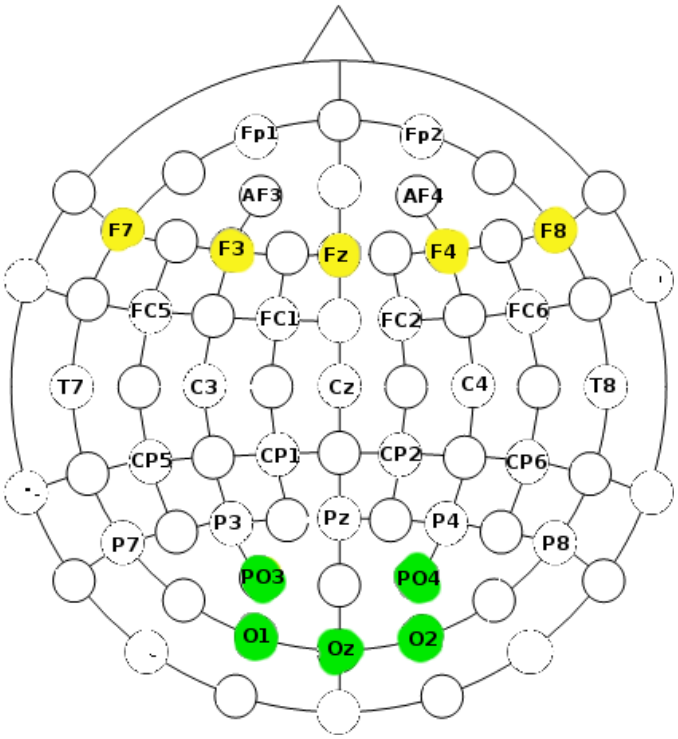


Figure 2: EEG 32 electrodes layout according to the international 10-20 system

3.1 Stimuli

The basic stimuli for this experiment consisted of 23 randomly selected images from the MIT-Adobe FiveK Dataset [BPCD11], Fig. 1. The database has a total of six stylistically different versions of each image, the original photo plus five different, professional photographer-modified variations. This dataset was chosen because we wanted to use a publicly available dataset that had not only the original image, but also professionally enhanced versions to teach the classifier what 'good' enhancements are. We also deliberately chose a dataset with neutral images, as opposed, i.e, to IAPS [LBC99], because we wanted to measure the emotional response to the 'look' of an image rather than to its content. Since all images in the database are aesthetically pleasing to a certain degree, we created two additional image versions by either over-saturating or under-exposing them. This was necessary to be able analyze the differences between the brain's response to a visually appealing image (expertly retouched) and one that is not (over-saturated or under-exposed). Our assumption that deliberately degraded image versions are perceived as less appealing than the original version or the professionally retouched versions was evidenced by user responses.

3.2 Procedure

The participants, upon arrival at the lab were asked to read and sign consent forms. They were told they would see a sequence of images and were instructed to mentally judge their preference or liking for each image. They were made to feel comfortable while the Biosemi 32-electrode cap and electrodes were attached [bio13](Fig. 2). An EEG was recorded with 32 electrodes according to the international 10-20 system. Additionally, a 4 channel EOG and mastoids were recorded which were used as a reference to remove data with accidental eye movements. The recorded data was referenced to the mastoids and filtered with a high-pass filter with a cut-off frequency of 0.1 Hz to remove DC-offset and drifts.

To gather our data, we presented all participants with the original photo, two of the expert-retouched versions, as well as our over-saturated and under-exposed versions, totalling 5 versions per photo and 115 different images overall. Each participant was shown the same 115 images. The order in which the different photos were shown was randomized, with all versions of the same photo appearing after each other but in random order. Each image was shown for a total of 2 seconds, followed by a fixation screen (gray background with a white dot in the center) shown for 700ms.

At the end of each trial (all versions of one photo), the participants were shown all versions of the photo as thumbnails and asked to select the one they liked least. This was timed so as to not give them any opportunity to compare and analyse the images, but to quickly select one based on their initial preference. This task was given to focus and keep their attention on the images and to validate our assumption that the bad versions were disliked the most.

3.3 Results

The user responses obtained during the experiment show that the two 'bad' versions were selected 79% The participants response to image preference correlates with the EEG data where the bad version 2 was disliked the most 48% of the time whereas bad version 1 was disliked the most 31% of the time. This data verifies our assumption that the bad versions really were disliked the most and the photographer versions were more appealing.

To analyze the data we averaged all the trials from all participants over electrodes P04,PO3,Oz,O1 and O2 for each image category (bad1,bad2,p4,p5 and original). We then ran two tailed t-tests on the data to determine if the response to each image category was in-fact distinct. The results are as follows:

1. Original vs. Bad Version 1 : p = 5.2×10^{-16}
2. Original vs. Bad Version 2: p = 5.5×10^{-5}
3. Original vs. Photographer Version 4: p = 2.9×10^{-13}
4. Original vs. Photographer Version 5: p = 0.91

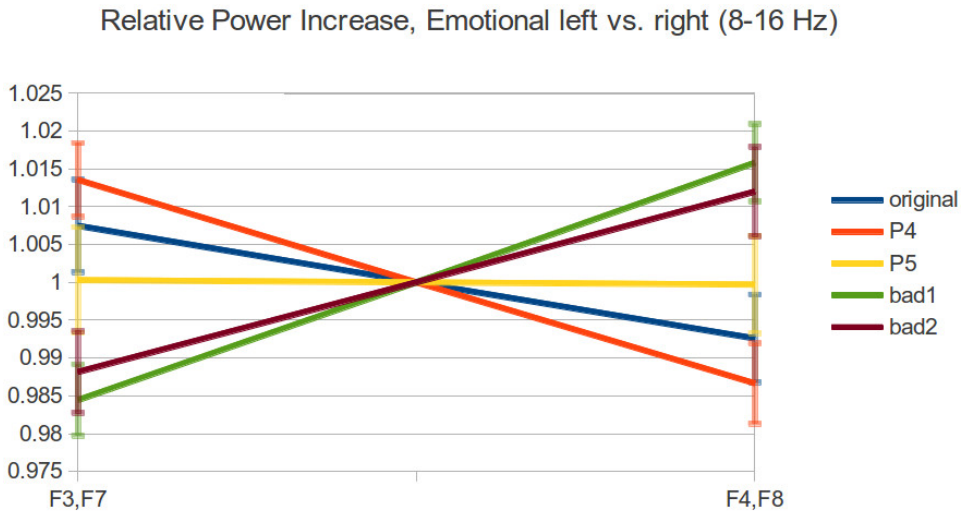


Figure 3: Left Hemisphere (F3/F7) versus Right Hemisphere (F4/F8) response to different image versions.

5. Bad Version 1 vs. Photographer Version 4: p = 2.8×10^{-7}
6. Bad Version 1 vs. Photographer Version 5: p = 0.021
7. Bad Version 2 vs. Photographer Version 4: p = 5.1×10^{-20}
8. Bad Version 2 vs. Photographer Version 5: p = 1.3×10^{-8}

Given the rejection of the null hypothesis ($p < 0.05$) in the above cases there is sufficient evidence for the statistical significance of the results. The only case where the null hypothesis is not rejected is between the original image and the photographer version 5. This means that the original and the photographer version 5 evoked a similar visual response in the brain. The response to the bad and the photographer enhanced versions was distinct not only from each other but also the original.

We conducted a second analysis of the EEG data from the frontal electrodes which also revealed a distinct response to the presented image versions. Previous data from EEG studies and emotion has provided evidence of lateralization of emotion in the frontal cortex [SW86] which predicts right hemisphere dominance for negative emotions and left for positive. Fig. 3 shows the response from electrodes F3/F7 (averaged) and F4/F8 (averaged) which are located in the front of the head corresponding to the frontal cortex. F3/F7 are in the left frontal hemisphere and F4/F8 in the right (Fig. 2). The data shows a higher relative power increase for the bad images over electrodes F4 and F8 (right hemisphere) while electrodes F3 and F7 (left hemisphere) show a higher power increase for images P4 and P5 (photographer enhanced versions). Given the complexity of the human brain and the many hypothesis concerning the experience and perception of emotion [DEYH05], this

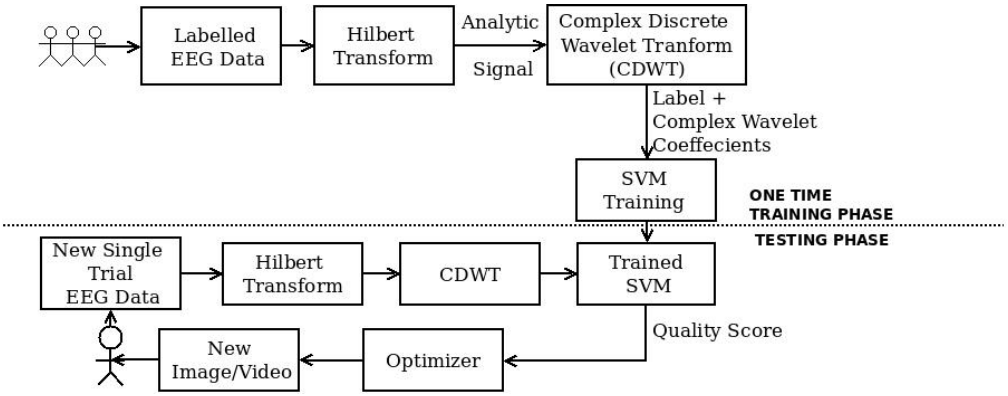


Figure 4: The framework for the optimization loop requires a training phase, which needs to be conducted only once (Experiment 1). The loop optimizes the image until an optimal version is created.

is but one explanation of the difference in potential observed between the right and left hemisphere.

4 Experiment 2

For the second experiment we were interested in investigating if it is possible to train a classifier to detect a positive or negative response to an image version from a single trial. This classifier was incorporated into a neural feedback loop to determine a score of 'pleasing' or 'displeasing' for each image version shown. An optimization algorithm then varied the image parameters towards an appealing optimum.

Fig. 4 shows the main components of our EEG-driven optimization loop. There is a one-time training phase required to teach the classifier the difference between the neural response to 'good' versus 'bad' visual stimuli (Experiment 1). The user sees an image, and the EEG is measured and sent to the Support Vector Machine (SVM)-based classifier. The classifier calculates a score, in real-time from the EEG data reflecting the user's 'liking' for the image. The optimizer, in turn, varies three image parameters, contrast, brightness and saturation, in response to the score. The algorithm then re-renders the image corresponding to the new parameter values, which is again displayed to the user.

The optimization framework is applicable to any rendering algorithm whose output varies depending on some set of parameter values. The modular set-up allows for easy replacement of the rendering component to optimize different parameters.

4.1 Single Trial EEG Data Analysis

A well-known problem in analyzing single trial EEG data is the poor signal-to-noise ratio. Traditionally, this has been tackled using the Discrete Windowed Fourier Transform (DWFT), producing a full frequency spectrum for a short time-span (usually one second). This approach either has poor temporal resolution, e.g. one frequency spectrum for each second of data, or produces large amounts of redundant information. On the other hand, the Discrete Wavelet Transformation (DWT) only produces a very rough frequency spectrum consisting of octaves only. However, it also adapts the temporal resolution to the frequency resulting in no redundant information being generated. Given a discrete input signal $f(t)$ and the wavelet filter pair consisting of a low-pass filter $g(t)$ and a high-pass filter $h(t)$, the general discrete wavelet transform is defined as follows.

$$s_0(t) = f(t) \tag{1}$$

$$s_{n+1}(t) = \sum_{k=-\infty}^{\infty} s_n(k)\, g(2t-k) \tag{2}$$

$$d_{n+1}(t) = \sum_{k=-\infty}^{\infty} s_n(k)\, h(2t-k) \tag{3}$$

This transformation can also be written as a series of lifting steps via factorization [SS00]. For the cubic B-Spline wavelets, this leads to the following definition.

$$\begin{aligned} d_{n+1}(t) &= s_n(2t+1) \\ &\quad -\frac{9}{16}s_n(2t+2) + \frac{1}{16}s_n(2t+4) \\ &\quad -\frac{9}{16}s_n(2t) + \frac{1}{16}s_n(2t-2) \end{aligned} \tag{4}$$

$$\begin{aligned} s_{n+1}(t) &= s_n(2t) \\ &\quad +\frac{9}{32}d_{n+1}(t) - \frac{1}{32}d_{n+1}(t+1) \\ &\quad +\frac{9}{32}d_{n+1}(t-1) - \frac{1}{32}d_{n+1}(t-2) \end{aligned} \tag{5}$$

The advantage of this factorization is the easy invertability of the above equations. In fact the cubic B-Spline wavelet can be defined by writing inverse wavelet transform in lifting steps to produce cubic B-Spline functions for a given single wavelet coefficient rather than creating the original functions $g(t)$ and $h(t)$ in frequency space. When comparing to Fourier Transforms, the DWT somewhat resembles the Discrete Cosine Transform (DCT) as the kernel functions are symmetric and it does not produce any phase information. While this is an advantage as far as redundancy goes, it causes issues with phase-shift of input signals as frequencies get dropped if they are off-phase. To produce the same frequency response, even if an input signal is shifted in time, we use the Complex Discrete Wavelet Transform

(CDWT) as implemented by Olkkonen et al. [OPOZ06] using a separate Hilbert Transform followed by two wavelet transformations. To analyze the EEG data, we cut out a two second chunk starting half a second before the image is presented. We then remove the baseline drift from this signal to avoid introducing erroneous high frequencies in the Hilbert Transformation. After the Wavelet Transformation, we only use the coefficients $d_n(t)$ corresponding to the one second starting right when the image is being presented. The additional data was required for padding in the Hilbert Transform. Similar to Mustafa et al. [MGM12a], we only keep the frequency bands corresponding to the 5Hz - 20Hz range. The complex wavelet coefficients $s_n(t)$ are then multiplied with their complex conjugate in order to get the power in each frequency band.

4.2 Classifier

We use a standard support vector machine [CL11] for all classification tasks. For the training, we performed a standard 5-fold cross-correlation test to search for the best set of training parameters. The data is split randomly into 5 groups of equal number trials. Using a C-SVM with a Radial Basis Function $e^{-g|x_i-x_j|^2}$ (RBF) classifier and a set of fixed parameters, the support vector machine is trained with data from 4 groups and tested against the trials in the remaining group. This process is repeated until all trials have been classified. As proposed by Chang and Lin [CL11], the process is repeated until the best set of parameters has been found. The final SVM is then trained with all of the training data and stored for running the optimization loop as it takes several hours to search the best parameters with the above approach. To minimize over-fitting and to reduce the amount of data being processed, we average the wavelet coefficients like Mustafa et al. [MGM12a] over time down to 8 complex coefficients per channel for one second of EEG data. Once the SVM has been trained, we use the probabilistic version, i.e. the one that produces not only a classification but also a confidence value of how accurate the classification is, for generating a score value for any new EEG input. The score value is simply defined as the confidence of the input belonging to the class of good images. Note that a confidence below 50% means the image is more likely to be to belong to the class of bad images.

4.3 Optimizer - Nelder Mead

In order to optimize our given set of parameters to produce a pleasing image, we not only need a score function that tells us how good any given image is but also an optimization algorithm that changes the parameters. Given the amount of noise contained in our score function, we cannot use derivative information and thus have to resort to direct evaluation methods. Furthermore, we assume that our score function has a single maximum, i.e. it is unimodal, and we are interested in finding a good value close to this maximum rather than its exact location. For the actual optimization, i.e. the task of finding new parameters to test for, we chose the

very simple Nelder-Mead heuristic [NM65]. Even though it is know to sometimes not converge beyond a certain point, it is still sufficient to produce an improved image [KLT03] given the amount of noise we encounter. The main advantage of the Nelder-Mead heuristic is its fast convergence compared to other direct evaluation methods. The initial step in every Nelder-Mead driven optimization is to set up the initial simplex $x_1 ... x_{n+1}$, consisting of $N+1$ locations, i.e. parameter settings, for N parameter, and evaluate the score function at these locations. In our case this means starting with some initial parameter settings, presenting the corresponding image, capturing the EEG data, processing the data and running it through the SVM to produce the score value. Once the initial simplex has been set up (Fig. 5(a)), new parameter settings are generated based on a specific set of rules and a cost function (Figure 5(b-f)), i.e. one minus the score value in our case. The basic idea is to "walk´´ the cost function $f(x)$ downhill and eventually reduce the size of the simplex. After each step, the new simplex contains the parameter settings with the lowest $N+1$ corresponding cost values, i.e. the ones with the images rated highest. In detail, the Nelder-Mead algorithm works as follows after setting up the initial simplex (letter correspond to Figure 5, $\alpha = 1$, $\gamma = 2$, $\rho = -\frac{1}{2}$,$\sigma = \frac{1}{2}$)

1. Order vertices so that $f(x_1) \leq f(x_2) \leq ... \leq f(x_{n+1})$. If we are happy with the results so far go to (f).
2. Reflect x_{n+1} at the center of gravity of all other vertices $x_O = \frac{1}{n}(x_1 + ... + x_n)$ of the simplex to produce $x_r = x_O + \alpha(x_o - x_{n+1})$ and evaluate the cost function. If $f(x_r) < f(x_1)$ go to step (e), otherwise if $f(x_r) < f(x_n)$, replace x_{n+1} with x_r and go back to (a).
3. Contract the simplex by generation $x_c = x_O + \rho(X_O - x_{n+a})$ so that it lies between x_r and x_O. If $f(x_c) < f(x_{n+1})$, replace x_{n+1} with x_c and go back to (a).
4. Reduce the simplex by replacing all x_i expect x_1 with $x_i = x_1 + \sigma(x_i - x_1)$, evaluate the cost function for all new vertices and go back to (a).
5. Expand the simplex by calculating $x_e = x_o + \gamma(x_o - x_{n+1})$. If $f(x_e) < f(x_r)$, replace x_{n+1} with x_e and go back to (a). Otherwise, replace x_{n+1} with x_r and go back to (a).
6. Output x_n as the best vertex, i.e. set of parameters.

Once the simplex is sufficiently small enough or after a given number of iterations, we stop the optimization process and simply output the parameters corresponding to the highest encountered score value.

4.4 Evaluation

To test our loop, we varied the basic image parameters saturation, brightness and contrast to see if we could obtain aesthetically pleasing versions of the original image.

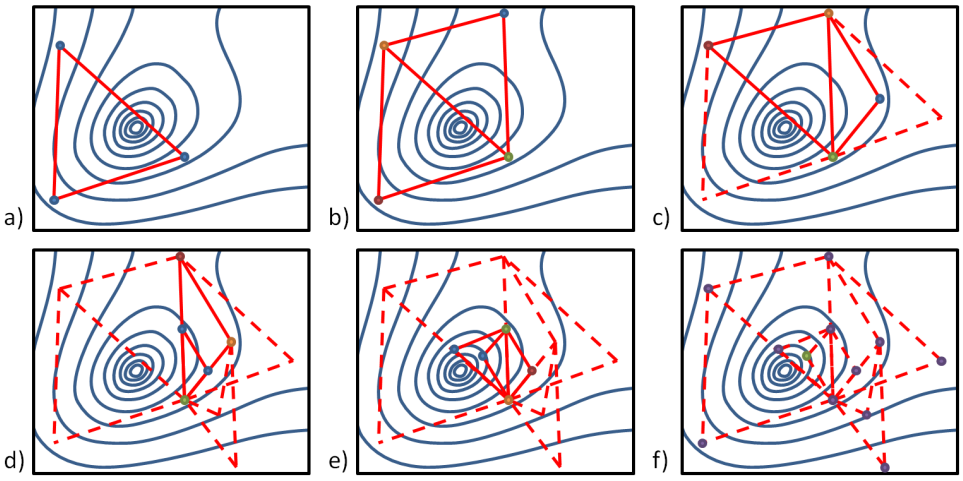


Figure 5: Iteration steps in Nelder-Mead heuristic: (a) setting up initial simplex, (b) reflection, (c) contraction, (d) reduction, (e) expansion, (f) end of optimization

4.5 Procedure

We evaluated our method with 15 users who did not participate in the SVM training phase (Experiment 1). Their average age was 25 years, and they had no professional experience in image editing. The evaluation experiment was done on 12 randomly selected photos from the MIT-Adobe FiveK Dataset [BPCD11] which had not been used for SVM training, Fig. 8. For each photo, the optimization loop was initialized with the original image version and each image version was shown for 2 seconds. The optimization loop was varied the three parameters saturation, brightness and contrast. The possible range for each parameter was set to allow significant variations in image appearance but without becoming unnatural. This was done to limit the otherwise large solution space and shorten optimization time. For each photo, the user was shown subsequently optimized versions based on the classifier score while watching the images. The users were asked to look at the rendered pictures and think about how they liked it. They were then shown a rating screen (1 - 5) and asked to rate the image. Optimization terminated either when it converged or after a maximum of 20 iterations 8(1-5 minutes total). This was because for most people the images converged within this number of iterations after which the same image versions started to repeat. To be able to evaluate the SVM classifier-generated visual appeal scores, the classifier scores generated for each image version during the experiment were displayed to the experimenter only (not the participant) during testing.

Results

The optimization loop created distinct versions of the original images (Fig.7,8). These versions were unique to each individual and distinct from the photographer

Figure 6: EEG data snippets for all participants and all images viewed, over the electrodes PO4, PO3,Oz, O1 and O2 (Fig.2), split by classifier-assigned score to 'good' (green) and 'bad' (red) image versions during optimization.

enhanced-versions. After each image optimization we had detailed discussions with the participants regarding the optimized images. Most of the participants preferred their own optimized version over the original image. To validate these results, we compared the user rating obtained during optimization for the original image with the final, optimized version: 86% of the time the optimized version was preferred to the original. Interestingly, often times it was not the accuracy of the image in terms of color or details that the users were interested in but more often than not, their preference for the enhanced image was based on how it made them feel. This is explained to some extent by the data gathered from the previous experiment where the emotional response to different images as seen in Fig. 3 was distinct based on image aesthetics. Our discussions with the users also revealed that they preferred working with images that offered the possibility for creative enhancements. For example, all users enjoyed looking at and enhancing the City photo, Fig. 7a, while the house photo was considered rather uninteresting, Fig. 8o.

How valid is the classier assigned score in terms of how it correlates with the actual neural response? To investigate how the classifier works and to see what the actual EEG data that it is classifying looks like, we plotted the power change over time, for all the users over all optimization runs (Fig.6). The data was split into two groups based on the classifier assigned score: low scored EEG data(< 0.5) and highly scored EEG data (> 0.5). Fig. 6 shows the distinct difference in the power increase for EEG data that the classifier gave a low or high score to. The p-value from the t-test between the data for high and low scores was < 0.05, indicating a statistically significant difference between the two groups. The image

(a) Original input image

(b) Optimized version 1

(c) Optimized version 2

(d) Optimized version 3

(e) Optimized Version 4

Figure 7: EEG-Optimized results are unique for each indivual and competitive in terms of visual appeal with the photographer enhanced versions.

versions which evoked the greatest visual response at $\approx 200ms$ after stimulus onset, were given a low score by the classifier since they were least liked [MGM12b]. We observe that the power over time goes down for low rated EEG data whereas for EEG data with a high score the power stays up. This analysis shows that the classifier is indeed able to predict a score that corresponds to the subjects neural reaction to an image version. Although it is possible to generate a score which to an extent corresponds to the actual neural reaction, it is an estimate. This is because the classifier scores are derived from a binary decision regarding the image. The classifier is trained to detect good and bad images from the EEG feedback. This binary decision is then converted into a weighted average score based on the classifier's confidence and so the scores contain a certain amount of noise.

We also randomly split the data from each group (< 0.5 and > 0.5) into two separate groups to see how closely they were related. Each sub-group (a and b) is very close to the average of the two, i.e., that the individual EEG data within the larger group corresponds well to the average. To determine if there is statistically a difference between each subgroup and the average we ran two-tailed t-tests with the resulting p-values indicating that each subgroup belong to the same population as the main group:

1. (< 0.5) vs.(> 0.5): < 0.05
2. (< 0.5) vs. (< 0.5a): 0.85
3. (< 0.5) vs. (< 0.5b): 0.85
4. (< 0.5a) vs.(< 0.5b): 0.75
5. (> 0.5) vs. (> 0.5a): 0.9
6. (> 0.5) vs. (> 0.5b): 0.9
7. (> 0.5a) vs. (0.5b): 0.85

These p-values indicate that within the two groups of EEG data with scores assigned either < 0.5 or > 0.5, individual EEG trials belong to the same population.

5 Conclusions and Outlook

Our study has shown that it is possible to detect a difference in the neural response to image versions with a different 'look' or 'feel'. An analysis of the ERP and the power increase over different electrodes has shown a statistically significant difference in the perception of images which have been retouched differently. We have also shown that it is possible to use these neural responses to train a support vector machine to 'learn' the difference between the neural response to pleasant and unpleasant image versions. This is validated using EEG data from the testing of the optimization loop and conducting t-tests that showed a statistical difference between EEG data given a high and a low classifier score.

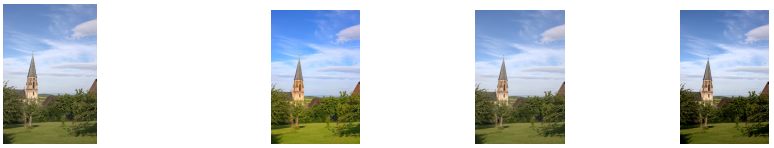
Our method can be applied in conjunction with any computational algorithm where varying certain parameters will produce a visually different output. Here, we have tested it on images and with a limited number of parameters. Given the NM optimizer's polynomial complexity, increasing the number of variable parameters from 3 to 4 is likely to increase the time to convergence by about 50%.

Our experiments indicate that it requires at least one second of presentation time and recorded EEG data to reliably estimate the "visual appeal" of an image (in our experiments we showed each image for 2 seconds). In our current implementation, one iteration step takes 30ms for data processing. Our framework works well with applications that elicit large changes in visual stimuli; it does not work well very well with small changes in the image. This is because an EEG does not pick up small neural changes in response to minute changes in images.

This initial study opens up exciting areas for further work. Would it be possible to optimize the content of an image as opposed to just its display parameters? How well will this approach work with more complicated image manipulations, like blurring, editing the background or foreground of the image? Is it possible to detect a neural response to changes in texture/edges information within an image? Of particular interest to us for further exploration is the possibility of gathering enough EEG to be able to create a predictive model for this optimization loop that mimics what an average person's response to an image might be.

Another intriguing application would be, for example, to use EEG-driven photo personalization in conjunction with Facebook's Instagram application [Lon11]. At the moment, however, a few obstacles to its use outside the lab remain to be solved. One challenge is, of course, the bulky EEG system itself, including head cap, 32 electrodes and wires. We would like to test our approach with the Emotiv EEG neuroheadset [Emo12] which is wireless and gel-less and can be used anywhere. The Emotiv headset contains most of the same electrodes that we used in this study for our data analysis (Fig.2) and so conceivably the shift would be possible. This would allow our method to be accessible to everyone, requiring no skill or knowledge of EEG measurements or image editing.

(a) Original Swan Image (b) EEG-optimized versions

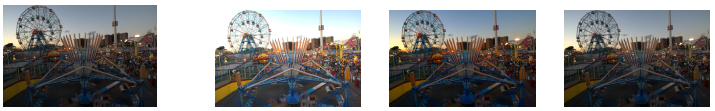


(c) Original Church Image (d) EEG-optimized versions

(e) Original Flower Image (f) EEG-optimized versions

(g) Original Gas Station Image (h) EEG-optimized versions

(i) Original House Image (j) EEG-optimized versions



(k) Original Carnival Image (l) EEG-optimized versions

(m) Original People Image (n) EEG-optimized versions

(o) Original Red House Image (p) EEG-optimized versions

Figure 8: Original images used for testing (left) along with some of the EEG optimized versions from different participants

References

[ABdC80] Ansel Adams, Robert Baker, and Alexandre Roberto de Carvalho. *The camera*. Little, Brown Boston, 1980.

[bio13] Biosemi activetwo system, 2013. `http://www.biosemi.com/products.htm`.

[Bos06] Danny Oude Bos. Eeg-based emotion recognition. *The Influence of Visual and Auditory Stimuli*, 2006.

[BPCD11] V. Bychkovsky, S. Paris, E. Chan, and F. Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 97–104. IEEE, 2011.

[CCV03] Paulo Cezar, Pinto Carvalho, and Luiz Velho. Mathematical optimization in graphics and vision. In *Course Notes Siggraph*, 2003.

[CKGP06] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun. Emotion assessment: Arousal evaluation using eegs and peripheral physiological signals. *Multimedia content representation, classification and security*, pages 530–537, 2006.

[CL11] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[CSB+00] Bruce N Cuthbert, Harald T Schupp, Margaret M Bradley, Niels Birbaumer, and Peter J Lang. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological psychology*, 52(2):95–111, 2000.

[CVCM09] Gerald C Cupchik, Oshin Vartanian, Adrian Crawley, and David J Mikulis. Viewing artworks: contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain and cognition*, 70(1):84–91, 2009.

[DEYH05] Heath A Demaree, D Erik Everhart, Eric A Youngstrom, and David W Harrison. Brain lateralization of emotional processing: historical roots and a future incorporating "dominance". *Behavioral and cognitive neuroscience reviews*, 4(1):3–20, 2005.

[Emo12] Emotiv. Epoc neuroheadset. `http://www.emotiv.com/store/hardware/epoc-bci/epoc-neuroheadset/`, 2012.

[HDR08] R. Horlings, D. Datcu, and L.J.M. Rothkrantz. Emotion recognition using brain activity. In *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, page 6. ACM, 2008.

[HKK12] Sung Ju Hwang, Ashish Kapoor, and Sing Bing Kang. Context-based automatic local image enhancement, 2012.

[KLT03] Tamara G. Kolda, Robert Michael Lewis, and Virginia Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, 45:385–482, 2003.

[KLW12] L. Kaufman, D. Lischinski, and M. Werman. Content-aware automatic photo enhancement. In *Computer Graphics Forum*. Wiley Online Library, 2012.

[KPC93] John K. Kawai, James S. Painter, and Michael F. Cohen. Radioptimization: goal based rendering. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '93, pages 147–154, New York, NY, USA, 1993. ACM.

[LBC99] Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. International affective picture system (iaps): Technical manual and affective ratings, 1999.

[Lon11] Ben Long. Why we love instagram, hipstamatic, and the lo-fi photo trend. `http://www.macworld.com/article/1161514/whywelovelofi.html`, 2011.

[MAB+97] J. Marks, B. Andalman, P.A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, et al. Design galleries: A general approach to setting parameters for computer graphics and animation. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 389–400. ACM Press/Addison-Wesley Publishing Co., 1997.

[MGM12a] Maryam Mustafa, Stefan Guthe, and Marcus Magnor. Single-trial eeg classification of artifacts in videos. *ACM Transactions on Applied Perception (TAP)*, 9(3):12, 2012.

[MGM12b] Maryam Mustafa, Stefan Guthe, and Marcus Magnor. Single trial EEG classification of artifacts in videos. *ACM Transactions on Applied Perception (TAP)*, 9(3):12:1–12:15, July 2012.

[MRN+08] M Murugappan, M Rizon, R Nagarajan, S Yaacob, D Hazry, and I Zunaidi. Time-frequency analysis of eeg signals for human emotion detection. In *4th Kuala Lumpur International Conference on Biomedical Engineering 2008*, pages 262–265. Springer, 2008.

[NM65] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[OPOZ06] Hannu Olkkonen, Peitsa Pesola, Juuso Olkkonen, and Hui Zhou. Hilbert transform assisted complex wavelet transform for neuroelectric signal analysis. *Journal of Neuroscience Methods*, 151(2):106 – 113, 2006.

[PBL+08] M Carmen Pastor, Margaret M Bradley, Andreas Löw, Francesco Versace, Javier Moltó, and Peter J Lang. Affective picture perception: Emotion, context, and the late positive potential. *Brain research*, 1189:145–151, 2008.

[SDS+93] C. Schoeneman, J. Dorsey, B. Smits, J. Arvo, and D. Greenberg. Painting with light. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 143–146. ACM, 1993.

[Sim91] Karl Sims. Artificial evolution for computer graphics. *SIGGRAPH Comput. Graph.*, 25(4):319–328, July 1991.

[SJWH04] Harald T Schupp, Markus Junghöfer, Almut I Weike, and Alfons O Hamm. The selective processing of briefly presented affective pictures: An erp analysis. *Psychophysiology*, 41(3):441–449, 2004.

[SL01] R. Shacked and D. Lischinski. Automatic lighting design using a perceptual quality metric. In *Computer graphics forum*, volume 20, pages 215–227. Wiley Online Library, 2001.

[SS00] Wim Sweldens and Peter Schroder. Building your own wavelets at home. *Computer*, (1995:5):15–87, 2000.

[SW86] E.K. Silberman and H. Weingartner. Hemispheric lateralisation of functions related to emotion. *Brain and Cognition*, 5:322–353, 1986.

[VG04] Oshin Vartanian and Vinod Goel. Neuroanatomical correlates of aesthetic preference for paintings. *Neuroreport*, 15(5):893–897, 2004.

[WFC+09] Christian Wallraven, Roland Fleming, Douglas Cunningham, Jaume Rigau, Miquel Feixas, and Mateu Sbert. Categorizing art: Comparing humans and computers. *Computers & Graphics*, 33(4):484–495, 2009.

[WKK+07] Christian Wallraven, Kathrin Kaulard, Cora Kürner, Robert Pepperell, and Heinrich H Bülthoff. Psychophysics for perception of (in) determinate art. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 115–122. ACM, 2007.

[Zek00] Semir Zeki. Inner vision: An exploration of art and the brain. 2000.