EmoTune - Changing Emotional Response to Images

Katharina Schwarz, Christian Fuchs, Manuel Finckh and Hendrik P. A. Lensch; University of Tübingen; Tübingen, Germany

Figure 1. *Tuning an arbitrary input image (center) to evoke a rather negative (left) or positive (right) emotional response. Based on a single input parameter the EmoTune filter alters the brightness, saturation and color temperature to steadily increase the valence of an observer.*

Abstract

Images and videos can be touching, triggering emotional responses or even changing the mood of the observer. The response is influenced both by the image content as well as the appearance of the image. In this paper, we investigate how solely simple image processing, i.e., modifying the brightness, the saturation, or the color temperature, actually changes the emotional perception. We collect empirical data on images associated with emotion labels and analyze the valence ratings of the different modifications and their strengths. We show that these relationships tend to be linear only in a limited range while sometimes stronger modifications lead to the opposite effect. Pushing the modifiers towards the boundaries we derive from those ranges and combining them successfully shifted the emotional affect on 92% on around 80 samples. From these findings we derive our EmoTune filter which allows for almost linear control by combining specific modes and demonstrate successful application to both images and videos.

Introduction

When viewing an image or watching a movie, a complex emotional response is evoked in the observer depending on multiple factors such as the displayed content, the observer's memories, or the appearance of the image. For sure, the evoked emotions are highly subjective. Thus, if artists provoke a certain mood they exaggerate their stylistic devices to carry out an impression. To complete the affecting atmosphere of a film, the movie industry performs color grading not only to correct exposure errors or balance color but also to create a certain style influencing the audience's perception. Professional colorists make shots appear warmer or cooler and are able to set a different mood of a scene [32].

Generally, an appealing look of an image influences if the viewer likes it, but the evoked emotional feeling can still differ. In this work, we explore the emotional affect, more precisely, we aim at *tuning* an image to predictably alter the emotional perception ("EmoTune" – Fig. 2). While the image content plays a major role in the viewer's emotional response, the semantics behind objects and their arrangements are particularly difficult to assess. Thus, we concentrate on content independent visual characteristics.

Overall, our work investigates how simple changes in basic global image modifications have a large impact on the emotional perception. We perform four types of experiments on images from two different sets where each image has been previously labeled

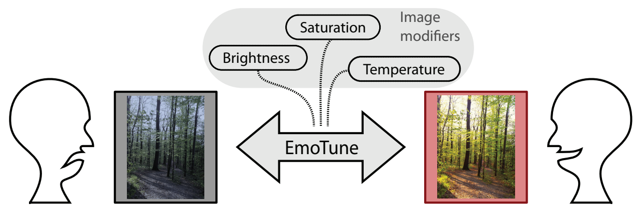

Figure 2. *EmoTune. Brightness, saturation and color temperature are altered to predictably change the evoked emotion of an observer.*

with the associated emotion. One of them is a new broad image set we assembled by hierarchically querying images from Flickr for specific emotional categories [21]. It allows to manipulate arbitrary images independent of their content. Further, in our empirical study, we investigate how the manipulation of basic image characteristics, namely hue, brightness, saturation, color temperature and local contrast affects the emotion an image evokes. The emotional state is expressed by subjective self-assessment of the so–called *valence* [3], rating positive or negative emotions on a five point Likert scale. Our observations show that depending on the mode and the strength of the modification one can boost, attenuate or even flip the emotional response in a rather controlled way. The study further revealed limits within which the modifiers do operate as expected. Beyond those particular thresholds they quickly negate the intended change. As the derived modifications are additive to each other, they can be applied jointly to achieve an even stronger tuning. Based on these insights we develop a novel heuristic filter that changes a given image such that the triggered emotion can be tuned relatively to the input using a single parameter (Fig. 1). We compare our tuned results to color gradings of an expert and establish the same visual trend. Thus, our developed *EmoTune* slider can support laymen as well as experts to influence the evoked emotion of films, games, advertisements, or website content to their liking.

Related Work

Emotion Categorization. In order to distinguish between different emotions, categorical approaches assign discrete labels to specific emotions [8, 11], whereas dimensional approaches map emotions into a space [34, 24, 27]. Due to the high growth rate in this field, Ekman [12] explored general agreements of scientists

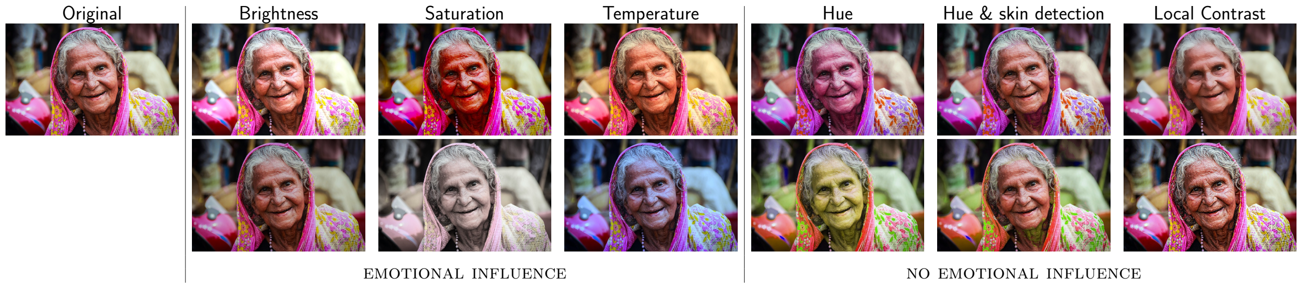

Figure 3. *All evaluated modifiers. First column original image, column 2-7 modified versions. Top row positive effect expected, second row negative effect expected. For the three rightmost columns no clear change in the emotional response was observed.*

about emotion revealing that both directions are similarly fundamental. To provide a standardized collection of visual emotional stimuli Lang et al. [16] used a dimensional method to develop the IAPS (International Affective Picture System) upon which Mikels et al. [21] defined emotional categories. We make use of both, the IAPS and the categories to explore emotional affect of visual data.

Color Manipulation. Manipulating image color to change the visual appearance has been mainly addressed in example-based color transfer to map the color concept or look from a reference image to another photo [25, 26, 1], a video sequence [6], or from a reference video to another video clip [2]. Further, multiple images are considered to alter a single image based on attributes like time [29] or season [15]. In contrast, we modify the color appearance of an image without any additional reference picture.

Color and Emotion. The effect of color on emotions has been reported as relationship between perceived valence and hue (wavelength) revealing that blue is the most pleasant color, yellow the least pleasant, and less bright as well as more saturated colors are more arousing [31, 30]. Further, emotional response to images has successfully been used to categorize images by machine learning approaches exploiting the IAPS [35, 17, 20]. Manipulating images in context of emotional affection is mainly explored in the field of non-photorealistic rendering to study the affect of renderings [10] or dampen the viewer's emotional response [36, 22]. Contrarily, we focus on provoking a stronger emotional response rather than dampen it. Other approaches enhance image harmony by shifting specific colors [7], study color themes to derive preferred ones [23], or exploit such themes to adjust the appearance based on an affective input word [33]. Further, learning semantic concepts between words and the image structure has been successfully explored to enhance images and provide re-renderings with spatially varying color enhancements depending on the concept [18, 19]. We now aim to modify the visual appearance of an image independent of its content or representation and without deploying prior learning steps. Following previous insights, we will investigate the response to color and intensity in our experiments.

Data Sets

IAPS. The mentioned IAPS [16] data set consists of images evoking a particular emotion with measured certainty. Additional labels assign the emotional categories defined by Mikels et al. [21]: Negative norms (anger, disgust, fear, sadness), positive norms (amusement, awe, contentment, excitement) and undifferentiated ones. Unfortunately, the IAPS mainly consists of images selected based on strong content (blood, insects, faces, ...) rather than displaying some mood or a certain atmosphere. Besides, most of the images in the IAPS are of rather low image quality, e.g., noisy, blurred, or low resolution.

Flickr. Inspired by the IAPS, but aiming at a diversity of images in adequate quality which are independent of content, we download images from Flickr based on the same emotional categories [21]: We define the set of negative norms as $\kappa_{neg} = \{anger, disgust, fear, sadness\}$ and the positive ones as $\kappa_{pos} = \{amusement, awe, contentment, excitement\}$. To query for precisely matching images, we combine their terms into the powersets $\mathscr{P}(\kappa_{neg})$ and $\mathscr{P}(\kappa_{pos})$ respectively. As stated by Schwarz et al. [28], very specific textual queries result in higher semantic precision with regard to the retrieved images but might return only a small amount. Thus, to assemble a large enough set with images of high precision we apply their hierarchical querying approach to both powersets separately. In detail, for each subset $P = \{p_1, ..., p_N\} \in \mathscr{P}(\kappa)$ we combine the elements $p_n \in P$ by conjunction to query for specific images. The next level is a disjunction of conjunctions of all subsets of the $N-1$ elements and so on. The most general level queries for the disjunction of all N elements p_n. This method leaves us with different levels of specialized images and enables further selection within a broad set. Overall, we obtain around 25K negative and 17K positive ones. For further processing, we manually select a small controllable subset of about 450 images, 214 negative and 237 positive ones, ensuring their mapping into the correct positive – negative class.

Videos. In order to apply our derived image filter to videos, we also crawl single short videos that are presented under creative commons license (CC0) from `https://vimeo.com`.

Image Modifiers

Our main objective is to influence the viewer's emotional response regardless of the image content. As indicated in previous work, color is a very powerful device to change the appearance of an image without performing a semantic analysis. Thus, we aim to identify a set of modifiers that is capable of changing the observer's evoked emotional response relatively to a given input image I towards more negative I^- or more positive I^+. Therefore, we evaluate simple but effective image modifications (Fig. 3):

Hue (H). The dimension hue in HSV color space defines how similar colors appear. It has been explored to have noticeable impact [31]. Our modifier I_H regulates this channel. Further, observers react highly sensitive if humans and, especially, faces are shown. As shifting the hue might result in a rather unnatural look, we apply face detection to handle them separately.

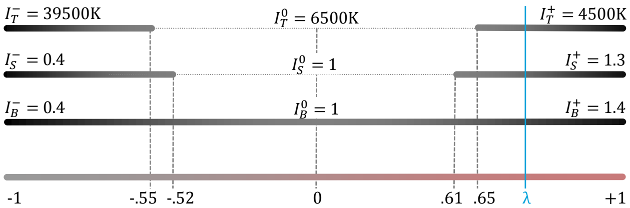

Figure 4. *Contribution of final modifiers. Estimated parameters, bounds and fadings are mapped into tuning space to derive the final filter. The emotional response is shifted towards negative or positve relatively to the input image.*

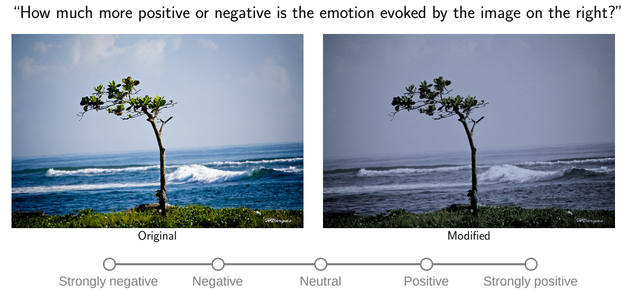

Figure 5. *Example as seen by AMT workers. The task is to answer the question (top row) on a 5pt Likert scale (bottom row).*

Saturation (*S*)**.** The saturation in HSV determines the colorfulness C of a color relative to its own brightness B.

$$S = \begin{cases} 0 & B = 0 \\ C/B & otherwise \end{cases}$$

This means, that it is inherently limited by the minimum and maximum possible brightness. Saturation already played an important role in research on color and emotion [31, 30] and, thus, we evaluate an adequate modifier I_S.

Brightness (*B*)**.** The third dimension in HSV represents the brightness. When viewed in the usual RGB space, changing the brightness results in a change of the mean RGB value of the image. In digital images, this change is limited by the dynamic range of the imaging pipeline. As light, or the intuitive imagination of day against night largely influences triggered feelings, we apply a brightness modifier I_B on the luminance channel of an image.

Color Temperature (*T*)**.** Color temperature denotes the color of light emitted by a black body heated to temperature T and is expressed in kelvin K. As certain color categories are closely related to the impression of an image giving it a rather cool or warm tone, we also incorporate a modifier I_T. We choose $6500K$ (color temperature of daylight) as the basis of I_T.

Local Contrast (*L*)**.** In [9, 5] it has been reported that there is only a small relationship between emotional impression and spatial frequencies, especially in the higher frequency bands. However, we think that boosting detail or smoothing an image might still influence the overall impression. In order to prevent destroying the visual content by modifying the global contrast, we implemented the local contrast modifier I_L based on edge-avoiding wavelets (EAW) [13, 14] to evaluate boosting and smoothing details on three scales of the EAW.

Combining Modifiers. Further, we aim at exploring the relation between the strength of the single modifiers and their resulting response, the limits within which the modifiers operate as expected, and suitable combinations to strengthen effects. Individual modifiers and combinations are evaluated in a user study.

EmoTune Model

Based on the previously mentioned modifiers, we develop a heuristic image filter that changes a given image such that the triggered emotion can be tuned relatively to the input using a single parameter. The final composition is visualized in Fig. 4. The later described experiments reveal that only the modifiers B, S and T control the valence in a reliable way. Thus, we measure how much each single modification I_m with $m \in \{B, S, T\}$ changes emotion relatively to the input state I^0 of an image as well as compared to the combination of several modifications. We further exploit boundaries for each modifier towards the negative I^- as well as the positive I^+ direction. Those boundaries as well as empirically established relative strength allow us to project them into a *tuning-space* which maps the possible range of emotional changes onto a simple scale. The final image filter is derived as follows: Let $\lambda \in [-1; 1]$ be the desired relative change in emotional response. Then, the final image O is computed from the original image I as

$$O = (\mathrm{T}_{\gamma(\lambda)} \circ \mathrm{S}_{\beta(\lambda)} \circ \mathrm{B}_{\alpha(\lambda)})(I). \quad (1)$$

The final parametrization of the model is estimated in an extensive user study which is presented in the following.

Human Evaluation

We want to derive a suitable mapping from the presented modifiers to our final EmoTuner such that the viewer's evoked emotion is predictably changed. Thus, as emotions are highly subjective, we performed experiments on Amazon Mechanical Turk (AMT) to obtain general ratings from a wide variety of people on a high demographically diverse level [4]. Overall, 86 Turkers participated during our testphase and completed a total of around $57K$ HITs (Human Intelligence Tasks).

In our setups, each task is rated on a five point Likert scale in the range $[-2; 2]$, noted as {*strongly negative, negative, neutral, positive, strongly positive*}, and is presented to 10 of the mentioned 86 Turkers. To explore the particular effects of the modifications, we designed experiments of four different types:

Generate Ground Truth. As the IAPS is a very well evaluated image database, we only verify ground truth for the subset of around 450 images we queried from Flickr. In this experiment, each image is categorized by Turkers based on the mentioned 5pt scale. For further processing, we only keep the images that were rated into the same category as they have been downloaded before and images of either category that have been consistently rated as neutral. Out of both, IAPS and Flickr, we randomly select 150 images: 50 negative, 50 neutral and 50 positive ones.

Evaluate Single Modifiers. This test identifies which modification influences emotion in terms of direction and amount. In a side by side setup, we display the original image and a modified version (Fig. 5) and ask Turkers to rate "how much more positive or negative" the modified image seems. This time the scale item is to be selected relative. We generate modified versions by emphasizing or dampening the image property by a fixed value (e.g., brightened to 140% or darkened to 60% of the original intensity) and produce the final test set by random permutation. This experiment revealed that only *brightness*, *saturation*, and *color temperature* tend to control the valence in a reliable way whereas it was hardly altered by shifting *hue* or change in *local contrast*. Thus, we focus on the successful modifiers B, S and T.

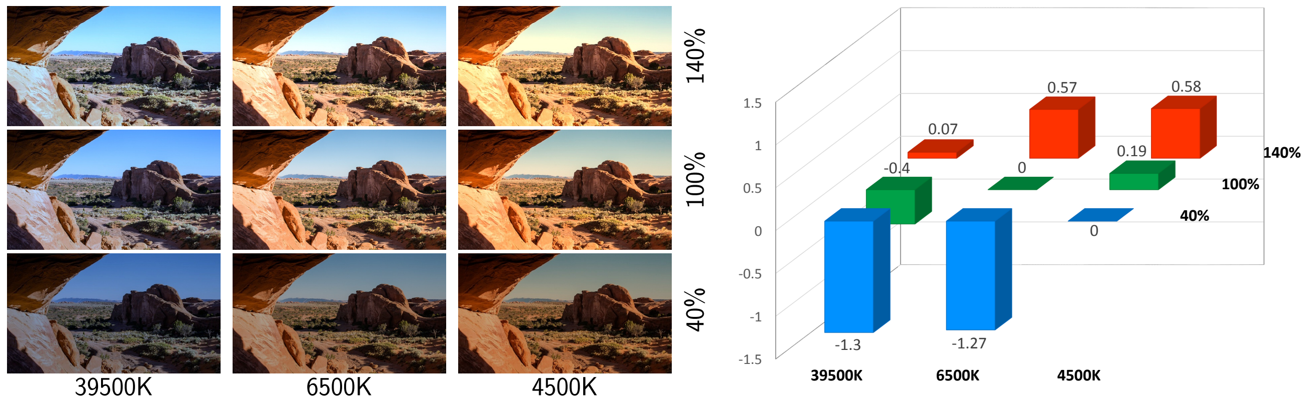

Figure 6. Left: Example combining color temperature T (cols) and brightness B (rows). Right: Average changes for combining T and B over all images arranged as on the left. The modifiers can amplify or dampen each other.

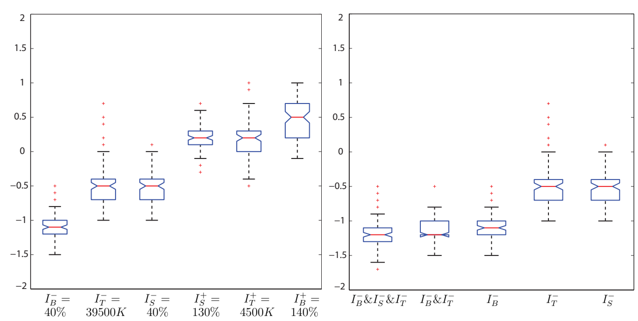

Figure 7. Boxplots. Left: Our empirical boundaries. Right: Single modifiers and combined ones leading to negative direction of the EmoTune filter.

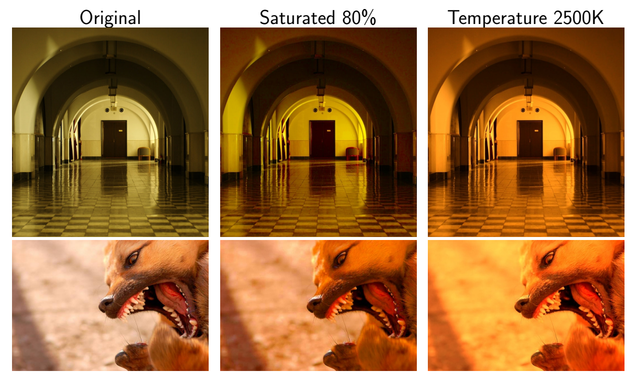

Figure 8. Example of extreme saturation and color temperature on neutral rated input (top row) and negative rated input (bottom row). Both values lead to more negative impression (scary or aggressive).

Evaluate modifier limits. With the same setup as in the previous experiment, we evaluate the strength and consistency of the emotional change with respect to the applied strength of the modifier. This means that, for example, besides a 40% modification in the previous experiment we further test on 50% and 60% in both directions to see how the modification operates in extreme settings. For each modifier we determine the limit up to which the assumed effect is observable.

Evaluate Combined Modifiers. The successful modifiers are further tested on their joint effect when applied simultaneously. We generate images with all possible combinations, tuples, and triples for the previously determined modifier strength limits and perform a similar experiment as the evaluation of the single modifiers. The results show that the various modifiers are additive but influence the emotional response to different degrees. From these findings we derive relative weights for intuitive EmoTuning.

Results of Experiments

For all previously described experiments we analyze the mean μ and the variance σ^2 of the emotional changes induced by the modification over all images in each ground truth category. Observed trends were the same on the IAPS and the Flickr set.

Results for Individual Modifiers. First, we investigate the effect of the three successful modifiers (Fig. 3, col 2-4). Their produced shifts in emotional response are summarized in Fig. 7 (left).

Brightness. Changing the brightness of an image has the most profound effect. Darkening an image pushes the reaction towards negative, whereas brightening the image makes it more positive to look at, even if rather negative content is shown. In our tests, increasing the brightness to about 140% achieved the maximum shift in response while the darkening is stopped at a minimum of 40%. In the entire range, the response is monotonically increasing as indicated by the third experiment.

Saturation. Desaturating the colors of an image makes the emotional response more negative, while saturating the colors typically yields more positive responses albeit not as strongly. For saturation there is an upper limit to how much it can be increased before the trend gets actually less predictable and tends to be more negative (Fig. 8). The predictable range is estimated to be between 40% and 130% of the original saturation.

Color Temperature. Changing the temperature towards cooler colors (i.e., more bluish) makes an image generally more negative. In the opposite direction, using warmer colors initially has the desired effect of provoking a more positive reaction. But, if the latter modification is applied too much, it can lead to a rather aggressive appearance (especially red tones) and push the emotional response towards the negative (Fig. 8, right). However, changing the temperature within the range of 4500K to 39500K only led to false labels in less than 5% of the input images.

Modifier	negative boundary	mid-point	positive boundary
Brightness (B)	0.4	1	1.4
Saturation (S)	0.4	1	1.3
Temperature (T)	39500K	6500K	4500K

Figure 9. *Estimated boundaries where modifiers still produce predictable monotonous changes. B, S are expressed with respect to 1 being the input image. For T the input image is assumed to be given at 6500K.*

Results for Combined Modifiers. In the last experiment we analyze the effect of combining the successful filters. Each modifier is used up to its estimated reliable bounds. Fig. 7 left shows plots for the final modifiers and the determined bounds. The boxes indicate that the positive trend is detected for the positive modifiers and, similarly, the negative trend for the negative modifiers. Overall, the negative trend is detected stronger than the positive one. Further, our experiment showed that the effect on the emotional response is actually boosted if multiple modifiers are combined beyond what is possible with just a single modifier (see Fig. 6). Again, the example indicates a stronger effect towards negative direction, e.g., negative boundary of temperature (39500K) combined with darkening (40%) leads to a negative change of -1.3 on the rating scale $[-2;2]$. Those examples also visualize that applying one modifier in the opposite direction of another modifier partially neutralizes the effect. More specifically, combining negative temperature and positive brightness (140%) results in a minimal positive average change of 0.07, and, in the other direction, positive temperature combined with negative brightness even in no predictable change. In all tested combinations, the change in brightness had the most dominating effect followed by saturation and color temperature (Fig. 7, right). Combining the modifiers strengthens the effect and also slightly increases the variance. Besides, pushing the modifiers to their evaluated boundaries and exploiting their derived combinations correctly shifted the emotional response in 92% on around 80 test cases.

Parameterization of EmoTuner. Based on the performed experiments, we derive a set of final modifiers with suitable parameters that are joined into the final EmoTuner (Fig. 4). The provided empirical evidence indicates that the amount of change in emotion can be steadily increased by parallel application of the final modifiers within boundaries in which they produce predictable monotonous change (Fig. 9). The parameters for the model (Eq. 1) are computed as:

$$\alpha(\lambda) = \begin{cases} 1.4\cdot\lambda & \text{if } \lambda \geq 0 \\ 0.4\cdot\lambda & \text{if } \lambda < 0 \end{cases}$$

$$\beta(\lambda) = \begin{cases} 1.3\cdot(\lambda-0.61)/0.39 & \text{if } \lambda \geq 0.61 \\ 0.4\cdot(\lambda+0.52)/0.48 & \text{if } \lambda < -0.52 \end{cases}$$

$$\gamma(\lambda) = \begin{cases} 6500-2000\cdot(\lambda-0.65)/0.35 & \text{if } \lambda \geq 0.65 \\ 6500+33000\cdot(\lambda+0.55)/0.45 & \text{if } \lambda < -0.55. \end{cases}$$

The modification of brightness is dominant over the entire range while its gain reduces towards extreme values. The modifications in saturation and color temperature are faded in towards the boundaries to support a steady growth in response. The particular thresholds when the other modifiers fade in are derived by the relative strength of the modifiers. For example, normalized to the maximum extend of the negative brightness modification, the negative saturation only achieves 48% in the evoked valence scale and therefore is applied only in 48% of the negative range of λ.

Discussion

Overall, the modifications tend to influence the valence more strongly towards negative than towards positive emotions (Fig. 10). However, the absolute scale needs to be analyzed in a more detailed experiment. One aspect that we could not control in the AMT study is the display setting which can affect the strength of the individual responses. Anyway, by presenting the reference image on the same monitor the relative trend should still be kept.

input label \ achieved change	- -	-	0	+	++
- -	–	✓	✓	✓	–
-	✓	✓	✓	✓	–
0	✓	✓	✓	✓	–
+	✓	✓	✓	✓	✓
++	–	✓	✓	✓	–

Figure 10. *Depending on the label associated with the input image the possible range of achievable valence tuning can be limited.*

One limitation of the proposed modifiers is that they cannot always overrule emotions triggered by semantic connotation of the image content. As shown in Fig. 10 original images that are labeled extremely positive or negative can only be emotionally tuned to a smaller extent than the average input image. Further, we demonstrated the importance to adhere to specific range limits for color saturation and temperature. Otherwise, one can for example easily produce the impression of being close to fire (Fig. 8).

Comparison to Expert Grading. Additionally, in order to find out how an expert modifies images to evoke a rather positive or negative emotional response and if our tuned results indicate a similar direction, we requested a professional colorist to perform such modifications. On a subset of 30 images from the previously generated ground truth, we asked the expert to globally modify each of the images such that the evoked emotion is as positive or negative as possible. Fig. 11 shows an example comparing our

Figure 11. *Example comparison of our results (top row) to modifications of an expert (bottom row) to evoke a rather negative emotion (left) or to shift an image towards a more positive emotion (right). Although final results might still vary, a similar trend can be observed.*

tuned results (top row) to the ones of the expert (bottom row). The original image (middle) is altered to evoke a rather negative emotion (left) or shifted towards a more positive emotion (right).

Overall, it seems that our proposed combination of modifications aims towards a similar trend. However, final results might still differ, e.g., compared to our negative version (top left), the negative suggestion of the expert (bottom left) seems extremely bluish. Nevertheless, an additional inking step could easily be applied on top of our version. Thus, although our filter can reduce workload for experts, they might need to add a finalizing step on top to fully cover their individual intentions.

Applications

Finally, our "EmoTuner" can be applied as a single slider to any input image and even works on videos.

EmoTune Slider Application. In a small application, we implemented the derived combination of modifiers as a single slider that allows the user to tune the affect of an arbitrary image towards a stronger positive or a stronger negative response on the range $[-1;1]$ relative to the initial image (Fig. 1). The example in Fig. 12 demonstrates the full range of the EmoTune slider. As

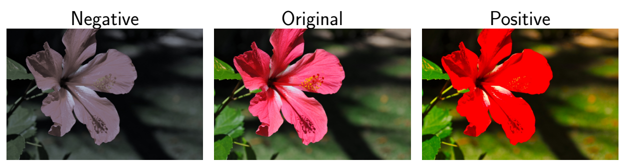

Figure 12. Demonstration of the full range of the EmoTune slider. The estimated negative boundary (left) is relatively dark whereas the positive extreme (right) indicates an almost artificial look.

any image can be selected as input, tuning towards the negative affect might already appear relatively dark whereas the positive extreme can result in an almost artificial look. However, crossing those derived boundaries usually does not provide more benefit but can even turn the direction back towards the opposite affect as indicated in Fig. 8.

Application to Videos. Further, we extended our emotional tuning application to work on videos. On individual scenes with little changes in lighting, the required emotional tuning state is simply applied to each frame. In videos with multiple scenes, we assume that the observer might request a different emotional appearance from scene to scene and support the user with an interactive tool (Fig. 13). It allows the user to select single frames and to tune each of those selected frames separately to a required emotional state. To obtain a smooth tuning along the video, linear interpolation is applied between the tuned states of the selected

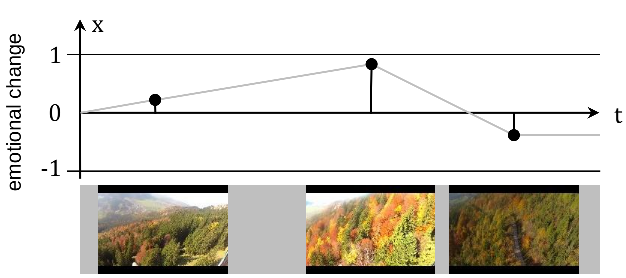

Figure 13. User-supported tuning of video stream. The user can pick several frames (bottom row) and slide them to a required emotional state. Linear interpolation allows for a smooth tuning along the stream.

frames. As long as the distance between the selected key frames is not too small, a temporally coherent video is generated.

Conclusion

We introduce a straightforward tool to modify the emotional response to images and videos. In a user study, we analyze the influence of simple image modifications on the evoked valence and reveal specific ranges where the tuning of brightness, saturation and color temperature work most efficiently or might produce unintended results. Based on these findings, we design our EmoTune filter to actively control the emotional response and combine the three successful manipulations to strengthen the effective change. Finally, we demonstrate applications to "tune" arbitrary images or smooth video sequences in real-time.

Of course, the filtered result cannot always overrule the emotion evoked by the depicted image content, e.g., a very negative image is rarely tuned towards positive. However, the filter works quite well in both directions for a broad range of input images. Besides, pushing the modifiers towards their bounds and combining them successfully shifted the emotional affect on 92% on around 80 samples. This demonstrates the already large impact of those simple manipulations. In future, even stronger emotional affect could be obtained by evaluating other dimensions than valence or even modifying the image content and context.

Acknowledgments

We thank Jan Fröhlich for helpful discussions on color and Andreas Minuth for supporting us with his expertise in color grading. Further, we thank all the Flickr users for supporting our research with their photographs. Original images from AntyDiluvian 1, Anwar Shamim 3, Hernan Vargas 5, mrhayata 8 (top), Laurentiu Nica 8 (bottom), Bruce Nagel 11. This work was supported by the DFG: SFB 1233, Robust Vision and the BCCN (BMBF/FKZ 01GQ1002).

References

[1] S. Bae, S. Paris, and F. Durand. Two-scale Tone Management for Photographic Look. *ACM Trans. Graph.*, 25(3):637–645, 2006.

[2] N. Bonneel, K. Sunkavalli, S. Paris, and H. Pfister. Example-Based Video Color Grading. *ACM Trans. Graph.*, 32(4):39:1–39:12, 2013.

[3] M. M. Bradley and P. J. Lang. Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. *J. Behav. Ther. Exp. Psychiatry*, 25(1):49–59, 1994.

[4] M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspect. Psychol. Sci.*, 6(1):3–5, 2011.

[5] L. Carretié, J. A. Hinojosa, S. López-Martín, and M. Tapia. An electrophysiological study on the interaction between emotional content and spatial frequency of visual stimuli. *Neuropsychologia*, 45(6):1187–1195, 2007.

[6] Y. Chang, S. Saito, and M. Nakajima. Example-based Color Transformation for Image and Video. In *Proc. of GRAPHITE*, pages 347–353, 2005.

[7] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y.-Q. Xu. Color Harmonization. *ACM Trans. Graph.*, 2006.

[8] C. Darwin. *The expression of the emotions in man and animals*. John Murray, London, 1872.

[9] S. Delplanque, K. N'diaye, K. Scherer, and D. Grandjean. Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *J. Neurosci. Methods*, 2007.

[10] D. Duke, P. Barnard, N. Halper, and M. Mellin. Rendering and Affect. *Comp. Graph. Forum*, 22(3):359–368, 2003.

[11] P. Ekman. *Basic Emotions*. John Wiley & Sons, Ltd., 1999.

[12] P. Ekman. What Scientists Who Study Emotion Agree About. *Perspect. Psychol. Sci.*, 11(1):31–34, 2016.

[13] R. Fattal, M. Agrawala, and S. Rusinkiewicz. Multiscale Shape and Detail Enhancement from Multi-light Image Collections. *ACM Trans. Graph.*, 26(3), 2007.

[14] J. Hanika, H. Dammertz, and H. P. A. Lensch. Edge-Optimized À-Trous Wavelets for Local Contrast Enhancement with Robust Denoising. *Comp. Graph. Forum*, 30(7):1879–1886, 2011.

[15] P.-Y. Laffont, Z. Ren, X. Tao, C. Qian, and J. Hays. Transient Attributes for High-level Understanding and Editing of Outdoor Scenes. *ACM Trans. Graph.*, 33(4), 2014.

[16] P. J. Lang, M. M. Bradley, and B. N. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8, The Center for Research in Psychophysiology, Univ. of Florida, 2008.

[17] S. Li, Y.-J. Zhang, and H. Tan. Discovering latent semantic factors for emotional picture categorization. In *ICIP*, pages 1065–1068. IEEE, 2010.

[18] A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk. Joint Statistical Analysis of Images and Keywords with Applications in Semantic Image Enhancement. In *Proc. of MM*, pages 489–498. ACM, 2012.

[19] A. Lindner and S. Ssstrunk. Semantic-Improved Color Imaging Applications: It Is All About Context. *IEEE Trans. Multimedia*, 17(5):700–710, 2015.

[20] J. Machajdik and A. Hanbury. Affective Image Classification Using Features Inspired by Psychology and Art Theory. In *Proc. of MM*, pages 83–92. ACM, 2010.

[21] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz. Emotional category data on images from the International Affective Picture System. *Behav. Res. Methods*, 37(4):626–630, 2005.

[22] D. Mould, R. L. Mandryk, and H. Li. Emotional Response and Visual Attention to Non-Photorealistic Images. *Computers & Graphics*, 36(6):658–672, 2012.

[23] P. O'Donovan, A. Agarwala, and A. Hertzmann. Color Compatibility from Large Datasets. *ACM Trans. Graph.*, 30(4):63:1–63:12, 2011.

[24] C. Osgood, G. Suci, and P. Tannenbaum. *The Measurement of Meaning*. Univ. of Illinois Press, Urbana, USA, 1957.

[25] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color Transfer between Images. *IEEE Computer Graphics and Applications*, 21(5):34–41, 2001.

[26] E. Reinhard and T. Pouli. Colour Spaces for Colour Transfer. In *Proc. of CCIW*, pages 1–15. Springer, 2011.

[27] J. A. Russell. A circumplex model of affect. *J. Pers. Soc. Psychol.*, 39(6):1161–1178, 1980.

[28] K. Schwarz, P. Rojtberg, J. Caspar, I. Gurevych, M. Goesele, and H. P. A. Lensch. Text-to-Video: Story Illustration from Online Photo Collections. In *Proc. of KES*, pages 402–409. Springer, 2010.

[29] Y. Shih, S. Paris, F. Durand, and W. T. Freeman. Data-driven Hallucination of Different Times of Day from a Single Outdoor Photo. *ACM Trans. Graph.*, 32(6):200:1–200:11, 2013.

[30] D. R. Simmons. The association of colours with emotions: A systematic approach. *J. Vis.*, 6(6), 2006.

[31] P. Valdez and A. Mehrabian. Effects of color on emotions. *J. Exp. Psychol. Gen.*, 123(4):394–409, 1994.

[32] A. van Hurkman. *Color Correction Handbook: Professional Techniques for Video and Cinema*. Peachpit Press, 2010.

[33] X. Wang, J. Jia, and L. Cai. Affective Image Adjustment with a Single Word. *Vis. Comput.*, 29:1121–1133, 2012.

[34] W. Wundt. *Grundriss der Psychologie*. Leipzig: Engelmann, 1896.

[35] V. Yanulevskaya, J. C. van Gemert, K. Roth, A. K. Herbold, N. Sebe, and J. M. Geusebroek. Emotional valence categorization using holistic image features. In *ICIP*, pages 101–104. IEEE, 2008.

[36] M. Zhao and S.-C. Zhu. Sisley the Abstract Painter. In *Proc. of NPAR*, pages 99–107. ACM, 2010.

Author Biography

Katharina Schwarz received her diploma in computer science from the University of Ulm and is currently a PhD student at the University of Tübingen. Her work focuses on the semantic relatedness between textual and visual sources as well as the exploration of the visual appearance of images.