Exploring three views on image enhancement for Pixel Privacy

Simon Brugman
Radboud University, the Netherlands
simon.brugman@cs.ru.nl

Maciej Wysokiński
Universidad Complutense, Spain
maciwyso@ucm.es

Martha Larson
Radboud University, the Netherlands
m.larson@cs.ru.nl

ABSTRACT

The aim of the MediaEval 2018 Pixel Privacy task is to increase image appeal while blocking automatic inference of sensitive scene information. We investigate three different views from which we could consider enhancement: the view of the image aesthetics field, the view of automatic large-scale aesthetics inference models, and the view of social media users who reflect on their own photographic practices. Systematic image editing can do better than one-size-fits-all-filters with helping casual social media users find the desired photo look. Machine learning aesthetics assessment falls short when inferring individual preferences. A qualitative user study gives insight into the diversity and complexity of preferences.

1 INTRODUCTION

The MediaEval Pixel Privacy task aims at protecting users from large-scale inference of sensitive information while increasing image appeal. As we develop Pixel Privacy technologies, we want to understand how to apply and assess image enhancement. In this paper, we consider three views on image enhancement.

- The view from the **field of image aesthetics**: Here we explore what aspects of overall colour harmony we can systematise without full understanding of the content of the image.
- The view of the field of machine learning on **automatic inference of aesthetics**: We would like to better understand the potential of this technology for aesthetics evaluation of image enhancement in the Pixel Privacy task.
- The view of social media users: We survey a small group of participants who have the habit of consciously reflecting on their own photographic practices. This qualitative **user study** aims at discovering strong and weak points of our image enhancements.

In the following sections, we discuss each view in turn. Note that in this work we assume an interconnection between enhancement and appeal. Consistently with [11, 14] we consider that improving aesthetics also improves appeal.

2 SYSTEMATIC IMAGE EDITING

We consider the field of image aesthetics in order to discover aspects of photos that can be changed systematically, leading to an aesthetic improvement or an increase of appeal without full knowledge of what is being depicted in the photo. Such aspects would lend themselves well to automatization. Our interest in automatization is related to the observation that automatic filters are currently in widespread use and assume that transformations must be fast to match the speed of what is currently offered by apps.

The number of amateur photographers is growing as smartphone usage increases [16]. Popular camera mobile apps attract a large amount of activity, such as Instagram (more than 1B monthly active users [15]) and Flickr (the iPhone is the most used camera [7]). These apps allow users to edit images internally, for example, applying filters, supporting extremely fast sharing of edited images.

(a) Original image (b) Enhanced image

Figure 1: The original image (left) is classified by ResNet50 as `hotel/outdoor`, the enhanced image as `fire_escape`.

Currently, the state of the art in mobile apps for cameras is predefined filters, which can change in hue, saturation or lightness or add visual effects like blur or noise. Filtered photos, especially with increased colour temperature, exposure, and contrast, are more likely to be viewed (+21%) and commented on (+45%) than unfiltered photos [2]. These filters have the disadvantage of depriving users of editing control. Predefined filters are the same each time the filter is used and may limit the ability of users to achieve the desired photo look. Here, we aim to discover contributions from the field of image aesthetics that would allow us to improve the flexibility of photo filters to increase image aesthetics and add user appeal.

The users of image sharing networks can be divided into people with aesthetic knowledge and casual photographers. The former group tends towards smooth changes, supported by manual editing, the latter usually prefers to achieve more dramatic change [2]. Our goal is to discover dramatic changes consistent with image content, but not requiring full image understanding. Early explorations have directed our attention to colour grading and cropping for image enhancement. Figure 1 shows a colour transformation whose goal is to increase the appeal of the original image. The example was chosen because it is one of the promising cases where the classifier used in the Pixel Privacy task [9] is misdirected by a transformation.

What aspects of overall colour harmony can we systematise without full image understanding? As an initial attempt, we convert the input image to HSV colour space [19], obtaining pixel values expressed in terms of the three-dimensional nature of human colour perception [20]: (1) *hue*, which refers to pure colour, (2) *saturation* from white light to pure colour and (3) *value*, which refers to illumination values. Assigning the hue values to the specific ranges in the RGB colour wheel [13] (primary, secondary and tertiary colours) it is possible to identify dominant values in order to carry out an overall harmony shift sensitive to tones, tints, and shades.

In this experiment, we manipulate only hue values shifting them to different ranges in the RGB colour wheel according to the nearest detected harmony: monochromatic, analogous, complementary,

MediaEval'18, 29-31 October 2018, Sophia Antipolis, France

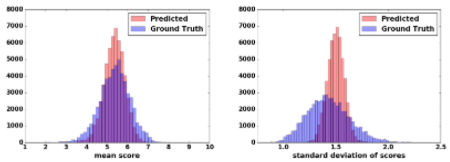

Figure 2: A histogram of the per image mean and standard deviation as calculated on ground truth and as predicted by NIMA, figure from [18].

double complementary, split complementary, triadic complementary [6] (pages 22-28). This methodology is similar to the geometrical formulation of classical colour harmony by Moon-Spencer [12]. We also applied a forced crop that considers the rule of thirds.

Note that visual perception involves both *a form* corresponding to structure and *colours* as a feature of reflected light [3] (page 20). Our experiment disregards form. The juxtaposition of the original and harmony-shifted images can intensify the sensation of artificial colours. However, colours are a response to light, and convincing looking colours are not absolute but can vary. The perceptual difference is reduced by the colour constancy phenomenon [1] (page 6-9). Note that the desired colour harmony can differ for each user and also may be more or less suitable for a given original image.

3 MACHINE LEARNING AESTHETICS

Technology for large-scale aesthetic inference is widely available and can be used by different multimedia applications for which user appeal is important, such as search engines (e.g., [11, 14]) and automated photo album management systems (e.g., [4]). We consider state-of-the-art advances in automatic image aesthetic assessment for evaluating appeal of image enhancement. In the task, our interest is focused on the user's personal perspective when sharing a picture on social media, as this is likely to lead to adoption of the privacy-preserving image enhancements.

A recent survey on image aesthetics assessment discusses visual features (hand-crafted and deep features), data set characteristics and evaluation metrics [5]. Neural-network-based machine learning models are able to assess image aesthetics more accurately than traditional approaches. They do not require explicit incorporation of expert knowledge of photography. There are efforts to improve on the state of the art. In [8, 11], the user ratings are extended with rater IDs, enabling user-specific models. NIMA [18] is a neural architecture for image assessment that predicts a distribution of ratings from one to ten. It improves handling of ground-truth ambiguity by optimizing the Earth Mover's Distance on ordered user score distributions. In addition to the mean user rating, the distribution of NIMA can capture agreement of user ratings. The loss used by NIMA can also be used for tuning image enhancement methods [17] and as a metric for perceptual distance [21].

From the nature of how NIMA is learned, it tends to avoid uncertain predictions when relevant information is missing. The image alone does not provide all information of the user state that could influence the rating (e.g., memories from the moment the user took the photo, current mood). Figure 2 illustrates that there is a mismatch between the distributions of per-image means and standard deviations when comparing the ground truth to the predictions of NIMA. NIMA predicts the overall reception of an image by users and does not attempt to predict reception of images by single users. The Pixel Privacy task could benefit from automatic assessment that treats all users equally in terms of the prediction error of their appeal judgements.

4 PERCEPTION OF IMAGE ENHANCEMENT

The user study is aimed at gathering qualitative insight into aspects of image enhancements important for user preference. The study compares three approaches: (1) systematically increasing overall colour harmony and improving composition (cf. Section 2), (2) enhancing the images intuitively, carried out by an artist, who restricted the enhancements to the same sort of manipulations that were applied systematically in (1), and (3) the style transfer approach described in [10]. Each approach is used to generate an enhanced image from ten original images from the manual test set of the 2018 Pixel Privacy task, resulting in 30 image pairs. The order of the pairs is randomised.

For each pair the original and enhanced image are randomly assigned to be Image A and Image B. Study participants look at both images and then answer the question "Which image would you prefer to share?" using a 5-point scale running between A and B. Additionally they give qualitative feedback by giving a short elaboration on their preference. The interface allows the user to toggle between image A and image B. Toggling makes the interface more closely resemble the user interfaces in existing applications (e.g., Instagram) and is also intended to eliminate unwanted direct comparisons of the two images. We had access to a group of people with conscious knowledge about images (e.g., photography or computer vision expertise), and, for this preliminary, we selected the study participants (ten in total) from this group. The rationale is that this group would be better able to identify which of their reactions is related to image transformations (as opposed to content) and express their reactions in words.

On average, study participants preferred the original image over the enhanced image. For systematic enhancement (1), enhanced images were preferred in 2 of the 10 cases, compared to 3/10 for intuitive enhancement (2). We identified several high-level categories capturing generalisations in the reasons given by study participants for their image preferences: colours (harmony, cold/warm), composition (ratio, perspective, focus, information, framing), no difference, authenticity (water is not purple). Notable was that for the systematic enhancement, composition change has an effect on the perceived authenticity and image quality (with respect to focus).

5 OUTLOOK

In this paper, we have explored three different views on image enhancement. Aspects from the field of image aesthetics can be systematised for specific image enhancement, making the change more dependent on the content of the image, while not requiring full understanding of the image. For the Pixel Privacy task, machine learning aesthetic assessment does not treat users equally in terms of the error of prediction of their appeal judgements, which is a potential limitation. A user study with experts gave valuable insight into the diversity of preferences for hue and composition.

ACKNOWLEDGMENTS

This work is part of the Open Mind research programme, financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] George A Agoston. 2013. *Color Theory and its Application in Art and Design.* Vol. 19. Berlin, Heidelberg.

[2] Saeideh Bakhshi, David A Shamma, Lyndon Kennedy, and Eric Gilbert. 2015. Why We Filter Our Photos and How It Impacts Engagement. In *Proceedings of the 9th International Conference on Web and Social Media (ICWSM).* AAAI, 12–21.

[3] José María Cuasante, Cuevas María, and Blanca Fernández Quesada. 2005. *Introducción al Color.* Akal, D.L., Tres Cantos (Madrid).

[4] Jingyu Cui, Fang Wen, Rong Xiao, Yuandong Tian, and Xiaoou Tang. 2007. EasyAlbum: An Interactive Photo Annotation System based on Face Clustering and Re-ranking. In *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 367–376.

[5] Yubin Deng, Chen Change Loy, and Xiaoou Tang. 2017. Image aesthetic assessment: An experimental survey. *IEEE Signal Processing Magazine* 34, 4 (2017), 80–106.

[6] Edith Anderson Feisner and Ronald Reed. 2013. *Color Studies.* Fairchild Books, New York.

[7] Flickr. 2018. Camera Finder. https://www.flickr.com/cameras/. (2018). Accessed: 2018-11-14.

[8] Shu Kong, Xiaohui Shen, Zhe Lin, Radomir Mech, and Charless Fowlkes. 2016. Photo Aesthetics Ranking Network with Attributes and Content Adaptation. In *Proceedings of the 14th European Conference on Computer Vision (ECCV).* Springer, 662–679.

[9] Martha Larson, Zhuoran Liu, Simon Brugman, and Zhengyu Zhao. 2018. Pixel Privacy: Increasing Image Appeal while Blocking Automatic Inference of Sensitive Scene Information. In *Working Notes Proceedings of the MediaEval 2018 Workshop.*

[10] Zhuoran Liu and Zhengyu Zhao. 2018. First Steps in Pixel Privacy: Exploring Deep Learning-based Image Enhancement against Large-scale Image Inference. In *Working Notes Proceedings of the MediaEval 2018 Workshop.*

[11] Ning Ma, Alexey Volkov, Aleksandr Livshits, Pawel Pietrusinski, Houdong Hu, and Mark Bolin. 2018. An Universal Image Attractiveness Ranking Framework. *arXiv preprint arXiv:1805.00309* (2018).

[12] Parry Moon and Domina Eberle Spencer. 1944. Geometric formulation of classical color harmony. *Journal of the Optical Society of America (JOSA)* 34, 1 (1944), 46–59.

[13] José María Parramón. 1998. *Teoría y Práctica del Color.* Parramón ediciones, Barcelona.

[14] Yale Song, Miriam Redi, Jordi Vallmitjana, and Alejandro Jaimes. 2016. To click or not to click: Automatic selection of beautiful thumbnails from videos. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (CIKM).* ACM, 659–668.

[15] Statista. 2018. Number of monthly active Instagram users from January 2013 to June 2018 (in millions). (2018). https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/ Accessed: 2018-11-14.

[16] Statista. 2018. Number of smartphone users worldwide from 2014 to 2020 (in billions). https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide. (2018). Accessed: 2018-11-14.

[17] Hossein Talebi and Peyman Milanfar. 2018. Learned perceptual image enhancement. In *2018 IEEE International Conference Computational Photography (ICCP).* IEEE, 1–13.

[18] Hossein Talebi and Peyman Milanfar. 2018. NIMA: Neural image assessment. *IEEE Transactions on Image Processing* 27, 8 (2018), 3998–4011.

[19] A Vadivel, Shamik Sural, and Arun K Majumdar. 2005. Human color perception in the HSV space and its application in histogram generation for image retrieval. In *Color Imaging X: Processing, Hardcopy, and Applications*, Vol. 5667. International Society for Optics and Photonics, 598–610.

[20] Stephen Westland, Kevin Laycock, Vien Cheung, Phil Henry, and Forough Mahyar. 2007. Colour Harmony. *Journal of the International Colour Association (JAIC)* 1 (2007), 1–15.

[21] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*