Artistic Vision: Providing Contextual Guidance for Capture-Time Decisions

Jane L. E
Stanford University
Stanford, CA 94305, USA
ejane@stanford.edu

Abstract
With the increased popularity of cameras, more and more people are interested in learning photography. People are willing to invest in expensive cameras as a medium for their artistic expression, but few have access to in-person classes to help improve upon their artistic skills. Inspired by critique sessions common in in-person art practice classes, we propose design principles for creative learning. We focus on applying these principles to design new interfaces that provide contextual in-camera feedback to aid users in learning visual elements of photography. We interactively visualize results of image processing algorithms as additional information for the user at capture-time. In this paper, we describe our design principles, and apply these principles in the design of three guided photography interfaces: one to explore lighting options for a portrait, one to highlight overall composition, and one to aid in de-cluttering.

Author Keywords
photography; guidance; context; lighting; composition

CCS Concepts
•**Human-centered computing** $\rightarrow$ **User interface design;** *Graphical user interfaces; Mobile devices;*

C&C '19, June 23–26, 2019, San Diego, CA, USA
ACM 978-1-4503-5917-7/19/06.
https://doi.org/10.1145/3325480.3326555

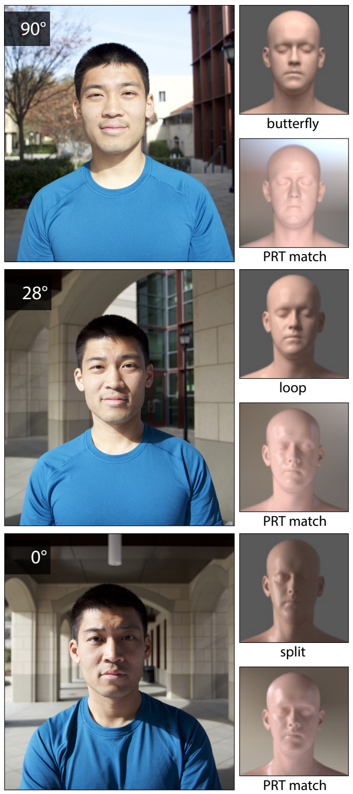

Figure 1: Butterfly, loop, and split lighting styles generated at a fixed location. Given a 360° HDR environment map and a target lighting style (top right), the tool automatically identifies the optimal angle for reorienting the subject to match the desired lighting (top left) and visualizes the scene at this orientation (bottom right).

Introduction

As cameras become smarter and more pervasive, more people want to learn to be better content creators. However, currently cameras provide limited aid in improving the aesthetic quality of the user's photographs. For an amateur who is interested in photography, but has limited training and equipment, the prospect of trying to take a "good" photo can be somewhat daunting. There are many variables to adjust: camera settings, location, and when a subject is present, their pose and expression, etc. In the moment, while framing a photo, these numerous considerations can be distracting and challenging to navigate.

Our goal is to leverage computation to provide users with context-aware feedback that allows them to capture in a more informed and intentional way, without impeding their ability to make their own aesthetic and creative decisions. In particular, we want to take advantage of the strengths of computation, to provide additional information in the form of in-camera feedback. Realizing this feedback through visualizations of computational results provides users with new lenses of sorts with which to see the camera viewport. The goal of these lenses is to allow users to become more aware of their artistic decisions and preferences, and ultimately help users take better photos by implicitly teaching them about visual elements of photography.

Related Work

While much work exists for automatically improving images, in-camera guidance interfaces to aid the user are limited. Mitarai et al. discovered that professionals tend to use multiple compositions in a single image. They presented a system that detects photographic elements (lines and saliency) in a captured photo to determine the closest composition. It displays visual guidance to propose an additional composition [6]. Li and Vogel presented a system specifically for capturing selfies. They learn aesthetic models for face position, face size, and lighting direction, and display guidance to help users achieve the idea camera distance and orientation [5]. My research aims to further understand how to design in-camera guidance that allows users to better consider artistic choices based on photography concepts.

Guided Photography Design Principles

Our approach is to design interfaces that provide guidance to help users learn as they practice photography. This requires addressing both learning and creativity principles.

Learning Principles

Ambrose et al. describe seven principles around how students learn based on the Science of Learning [1, 3]. We focus on two of these principles, practice and feedback, that most closely align with the practice of having critique sessions in in-person art practice classes.

Practice. *"To develop mastery, students must acquire component skills, practice integrating them, and know when to apply what they have learned"* [1]. Photography guidance should help users apply visual principles directly in the context of the image they are currently trying to capture.

Feedback. *"Goal-directed practice coupled with targeted feedback enhances the quality of students' learning"* [1]. Photography guidance should suggest possible goals and provide feedback on progress towards these goals.

Creativity Principles

When designing for creative education, it is also important to consider the tools' impact on the creative process and artistic expression.

Confidence. Creative works tend to be particularly abstract and subjective to assess, making it especially challenging for beginners to develop confidence. Does using the guidance interface help the user feel more confident in the quality of the final photo?

Ownership. Does the user feel ownership over the resulting photo? Did the user feel like they were able to express their personal aesthetic style? Across users, is there variation in the photos or do they begin to converge in style? It is important not to hinder the "artist's hand." Thus, we aim to provide guidance subtly, in a manner that is not overly prescriptive.

User Interfaces

Here we describe three projects that tackle challenges associated with important visual elements of photography: lighting, composition, and de-cluttering. Our preliminary user studies and prototypes support our hypothesis that these interfaces successfully promote creative learning by following the aforementioned design principles.

Figure 2: A set of (image, composition template) pairs computed by our algorithm.

Portrait Lighting

One of the most challenging and impactful considerations in photography is lighting. In a portrait studio, it is common to have a main light, fill light, and background light, as well as rim lights, hair lights, kickers, etc., positioned in a way to achieve a specific lighting style [4]. Non-experts generally don't have access to such equipment or have knowledge of how to arrange them. However, even relying on available light, the lighting on a face can vary drastically by just rotating the subject (Figure 1).

We leveraged this observation to design and implement an interface that shows the photographer a gallery of possible lighting styles achievable in the current environment, and helps the photographer orient their subject to capture their selected look. Determining this orientation requires knowledge of the environment—specifically, the position of the subject relative to lights in the scene. We do this by capturing the environment with a 360° camera.

With regards to the learning principles, our interface encourages users to practice making portrait lighting decisions in the context of the current subject and location. The reorientation guidance provides feedback on how close the user is to the target orientation/lighting. In a user study with 28 participants (20 amateur, 8 expert), several participants expressed belief that using this interface would help them learn: *"I feel like I learned a lot about portrait lighting by just using the interface once and I will definitely think more intentionally about the lighting style that I want to achieve when I take photos."* (P4).

In addition to learning, the studies also provide some initial support that the interface satisfies the creativity principles. When using the tool, participants were significantly more confident in their performance (NASA TLX) of the task of capturing a well-lit portrait (Wilcoxon-Mann-Whitney Test [7]: Z = 2.8, p = .01).

Composition

One way of interpreting composition is by looking at alignment of visually important elements in an image with the lines/intersection points of a composition grid. Currently in articles, books, etc., photographers will describe photographic composition by manually highlighting individual lines from such a grid to emphasize alignment choices of the image. We want to automate this process to enable an interactive guidance interface that provides these annotations as feedback to help the user discover such alignments directly in the context of the current camera image.

In order to better understand how to select lines to highlight as well as understand the potential of such an interface, we had 10 experts each annotate around 10 photos (5 provided by us, 5 of their own) with these composition lines. Given these images with an overlaid composition grid, experts immediately highlighted certain lines that they would've seen even without the grid. However, many noted additional alignments or near-alignments that they noticed due to the overlaid grid, and suggested potential adjustments they may have tried in camera with this new perspective. This suggests that even experts who already have training in intuitively visualizing these grids in their minds may be influenced by such an interface.

Regarding implementation, many saliency algorithms exist that identify objects of visual prominence [2]. Thus we generate composition templates to capture the set of lines that best align with a saliency map. Figure 2 shows a few composition templates generate by our heuristic algorithm.

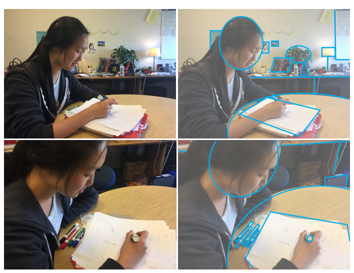

Figure 3: After the user takes a first photo (top left), the experimenter draws an overlay that approximates the image as geometric shapes (top right). The user is shown the photo with the overlay and is given the option to take another photo. Here, the user chooses to make some adjustments for the second iteration (bottom).

De-Clutter

Once the photographer has decided on an approximate orientation and framing, a range of adjustments in framing can still significantly impact the quality of the final shot. With the photographer mainly focusing on the subject and the action, it can be easy for some unwanted objects in the background to go unnoticed. We propose abstracting the camera image in order to allow the photographer to be more aware of all elements of the image. The initial low-fidelity prototyping has shown promise. In the photos shown in Figure 3, the user sees in the overlay, the clutter on the desk in the background. For her final photo, she decided to take the photo from a higher perspective angled downwards towards the main subject of the photo (the person writing on a stack of paper). This greatly reduced the prominence of the office clutter and keeps the image more focused. Most participants exhibited similar behavior; upon seeing a photograph with an abstracted overlay, users noticed unwanted clutter and moved objects and/or reframed the image to keep them out of the frame.

Evaluation Challenges

As described, we aim to design interfaces that encourage creative learning. However, the goals of promoting creativity and learning are somewhat at odds. To avoid removing any creative ownership, our algorithm avoids making explicit suggestions; this also means all "teaching" happens implicitly–through users' making more informed decisions due to feedback the in-camera guidance.

Expert and self-assessment of the quality/diversity of produced work can provide some measurement of both creativity and learning. Additionally, we have seen some qualitative evidence that users believe they are learning. However, it is difficult to concretely evaluate learning. A possible direction would be to try to measure levels of active, engaged, meaningful, and socially interactive learning (the four "pillars of learning") [3]. We are still exploring how to best evaluate based on these pillars, as well as considering other metrics for evaluating creative learning.

Conclusion

In this paper, we presented our approach to designing guided photography interfaces for learning and described three concrete projects that take this approach. While these projects are at varying stages of completion in terms of algorithm and design, they have shown promise in terms of both the learning and creativity principles defined. We are excited to further explore this approach by building upon these prototypes, and studying them with photographers of varying skill levels.

References

[1] Susan A. Ambrose, Michael W. Bridges, Michele DiPietro, Marsha C. Lovett, and Marie K. Norman. 2010. *How Learning Works: Seven Research-Based Principles for Smart Teaching*. John Wiley & Sons.

[2] Zoya Bylinskii, Tilke Judd, Ali Borji, Laurent Itti, Frédo Durand, Aude Oliva, and Antonio Torralba. 2015. Mit saliency benchmark. (2015).

[3] Kathy Hirsh-Pasek, Jennifer M. Zosh, Roberta Michnick Golinkoff, James H. Gray, Michael B. Robb, and Jordy Kaufman. 2015. Putting Education in "Educational" Apps: Lessons From the Science of Learning. *Psychological Science in the Public Interest* 16, 1 (2015), 3–34. DOI: http://dx.doi.org/10.1177/1529100615569721 PMID: 25985468.

[4] Fil Hunter, Steven Biver, and Paul Fuqua. 2015. *Light Science & Magic: An Introduction to Photographic Lighting*. Focal Press.

[5] Qifan Li and Daniel Vogel. 2017. Guided Selfies Using Models of Portrait Aesthetics. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 179–190. DOI: http://dx.doi.org/10.1145/3064663.3064700

[6] Hiroko Mitarai, Yoshihiro Itamiya, and Atsuo Yoshitaka. 2013. Interactive photographic shooting assistance based on composition and saliency. In *International Conference on Computational Science and Its Applications*. Springer, 348–363.

[7] Markus Neuhäuser. 2011. Wilcoxon–mann–whitney test. In *International encyclopedia of statistical science*. Springer, 1656–1658.