Artistic Vision: Providing Contextual Guidance for Capture-Time Decisions

Jane L. E
Stanford University
Stanford, 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. Inspired by critique sessions common in in-person art practice classes, we propose design principles for creative learning. My dissertation research focuses on designing new interfaces and interactions 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 to make more informed and intentional decisions during capture. In this paper, we describe our design principles, and apply these principles in the design of two guided photography interfaces: one to explore lighting options for a portrait, and one to refine contents and composition of a photo.

Author Keywords

photography; guidance; context; lighting; composition

CCS Concepts

•**Human-centered computing** → User interface design; Graphical user interfaces; Mobile devices

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. In fact, the photographer is making many decisions, both technical and aesthetic, many of which are happening passively and without his awareness due to the sheer number of options to consider.

UIST'18 Adjunct, October 14–17, 2018, Berlin, Germany.

ACM ISBN 978-1-4503-5949-8/18/10.
DOI: http://dx.doi.org/10.1145/3266037.3266128

Until recently, cameras have solely used visual information from a single lens—helping users automatically set exposure, white balance, etc. Newer smartphone cameras now have two lenses to provide stereo information. I want to take this a step further in my research, proposing the use of other sensors and algorithms, such as 360° camera technology, depth sensors, or accelerometers, as additional sources for capturing the context of the scene to augment traditional capture. This additional data can be overwhelming to a human, but can provide helpful context for automated computation; I am interested in this intersection between automation and human decisions. In particular, I want to take advantage of the strengths of computation, to provide additional information in the form of in-camera guidance. This guidance can aid humans during the creative process of capturing a photo, enabling them to focus on the artistic decisions they are making based on their aesthetic preferences.

In my work, I realize this guidance through visualizations of automated results and suggestions, a new lens of sorts with which to see the camera view. Users are shown proposed orientations, framings, etc. and need to actively decide whether to follow the guidance of the algorithms, to adjust the suggestions based on their personal aesthetic preferences, as well as to completely ignore the guidance. Regardless of the decision, the user is made more aware of the relationship between the automated suggestions and their final photo, as well as the decisions she made to capture that image.

My goal is to leverage computation to provide users with content-based 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. By providing users with such interactive in-camera interfaces, we aim to help users take better photos by teaching them about visual elements of photography. In this paper, I approach this goal in relation to two capture phases: ***exploration*** and ***refinement***.

RELATED WORK

Automation in Creativity Support Tools

We focus on designing tools that allow automation to assist in the creative process, but ultimately allow the user to make the artistic decisions. Such creativity support tools exist across many domains.

Merrell et al. presented a furniture layout tool that suggests functional and visually pleasing layouts [8]. Users can manipulate the layouts based on their personal preferences, and the

system will iteratively update the suggestions to satisfy the constraints. Automation can assist in the more tedious aspects of an artist's work and thus allowing them to focus more on the creative aspects of the task. Cong et al. presented an artist-directed facial simulation system that adapts physically based simulations to respect the artist's animation intentions [3].

Image Editing and Analysis

Such systems also exist in the photography domain. SelPh is a system that is able to learn the user's photo editing preferences to automatically edit similar photos [5]. This system relies on a variety of image analysis algorithms to compute distance between images, feature vectors, as well as a user's personal preference model. Numerous algorithms exist to gain an understanding of both image quality [11] and image content [6]. As more of this work in the graphics and vision community becomes interactive, there are more opportunities to use and build on these methods to design further interactive guidance interfaces for photography.

Guided Photography

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 [9]. Li et al. 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 [7]. My research aims to further understand how such guidance impacts the quality of the photographs captured by users, as well as their ability to improve and learn the basics of photography.

GUIDED PHOTOGRAPHY DESIGN PRINCIPLES

My 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 [1]. We focus on two of these principles, practice and feedback, that most closely align with the common 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; it is important not to hinder the "artist's hand." My research aims to provide guidance in a manner that is not overly prescriptive and distracting, but instead provides subtle assistance. While we try to promote these principles, it is hard to guarantee their satisfaction during the design process. Therefore, quantitative and qualitative feedback from user studies is crucial for understanding the impact of my work on these aspects of practicing photography.

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? Does the interface make the user more comfortable sharing the 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?

PHOTO CAPTURE PHASES

The photo capture process can be broken down into two phases: exploration and refinement. These phases have different considerations and challenges, and therefore different user needs.

Exploration

The user starts in the exploration phase. She has something in mind that she wants to photograph—this can be an object, a person, a general location, etc. Regardless of what it is, the countless visual options can be intimidating and overwhelming. At this stage, the user is looking to explore the range of options in the scene, e.g. different lighting styles or compositions; she is looking to be inspired to pursue one.

Refinement

Once she settles on an idea of an image that inspires her, she enters the refinement phase. There still might be many things that could influence the quality of the final photo, such as distracting clutter or misalignment. At this stage, the user wants to know that she hasn't overlooked anything in the image and that the final photo accurately represents the image she has in her head; she is looking to precisely execute her vision. In interviews, I observed that challenges in using a camera can cause people to doubt their abilities as photographers, despite having strong aesthetic senses. This is especially true for those with visual and motor impairments [2, 10].

USER INTERFACES

Here we describe two projects that tackle the different challenges for each phase through specific visual elements: for the exploration phase, we guide users through the process of realizing varied ideas and selecting amongst them through exploring portrait lighting styles; for the refinement phase, we guide users to carefully and precisely execute a vision through decluttering and refining composition. Our preliminary user studies and prototypes support our hypothesis that these interfaces successfully promote creative learning through following the aforementioned design principles.

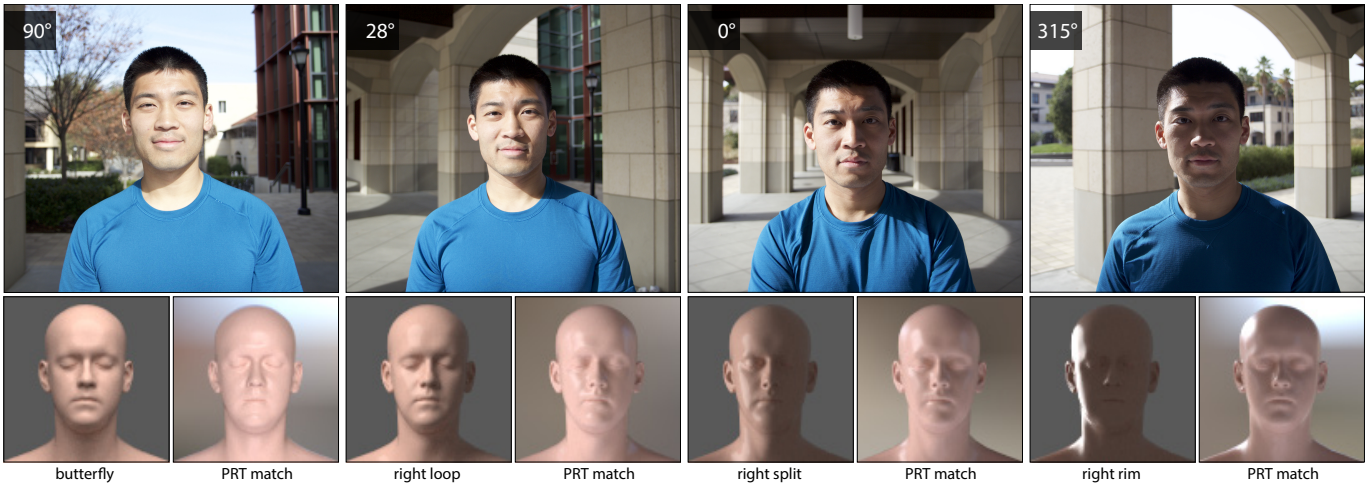

Figure 1. Lighting has a large impact on the quality of photos and videos. In a fixed lighting environment, photographers can produce many different lighting styles (e.g. butterfly, right loop, right split and right rim) just by rotating the subject in place without changing their location. Given an HDR environment map from a 360 camera at some initial orientation and a target lighting style (bottom left), the tool automatically identifies the optimal angle for reorienting the subject to match the desired lighting—e.g. 90° for butterfly lighting. A precomputed radiance transfer-based method on a generic head and skin model is used for efficiently optimizing lighting orientation and for visualizing the best orientation match (bottom right).

Exploration: Portrait Lighting

One of the most challenging and impactful considerations 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. With current cameras, capturing this information would require the user to aim the camera towards all light sources. Thus, we propose adding a low-resolution 360° camera to easily capture the full environment map.

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 an informal user study with 8 self-identified casual photographers (e.g. amateurs with an interest in photography), several participants expressed belief that using this interface would help them learn lighting (4). *"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. I can imagine that I will learn even more by using this interface more often even at a subconscious level when I just know myself how to turn relative to the light sources when I want to achieve a particular target"* (P4).

In addition to learning, the studies also provide some initial support that the interface satisfies the creativity principles. Participants expressed that it made them feel more confident and in control (6), and that the interface increased their awareness and intentionality (6) with regards to lighting. In particular, they appreciated the gallery of targets (4) as they *"never realized how many there were to choose from"* (P5), allowing them to *"feel like [they] can take much more dramatic and varied photos in a limited space"* (P8).

Refinement: Abstraction

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. I propose abstracting the camera image with the goal of allowing the photographer to be more aware of all elements of the image. My initial low-fidelity prototyping has shown promise. In the photos shown in Figure 2, 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.

In moving towards an automated, computationally-driven prototype, there are various potential abstraction algorithms, such as edge detection or image segmentation, each providing an opportunity to design a variation on the abstraction lens. Through user studies comparing these different conditions, I can understand differences in how the additional information provided by these lenses is perceived by users and how it influences their capture process.

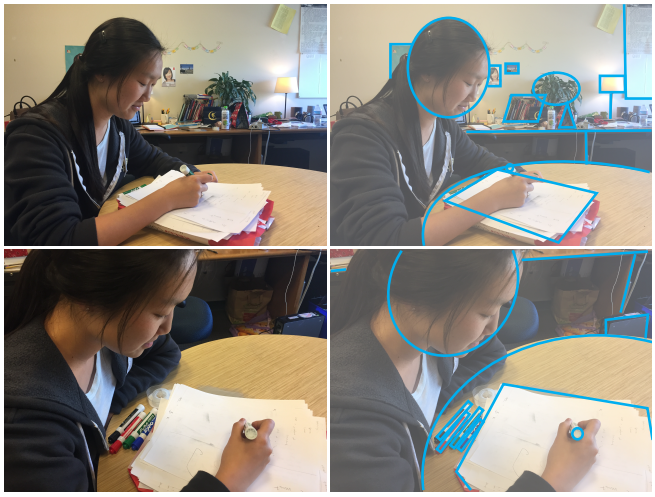

Figure 2. Pilot study results of my low-fidelity prototype of the abstraction interface. Users were asked to frame a shot of a person interacting with an object of their choice. After taking a first photo (top left), the experimenter draws an overlay that approximates the image as geometric shapes (top right). The image provided is a computer generated representation for legibility. During the study, these were drawn with whiteboard markers on transparencies and overlaid on the photo. 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).

CONCLUSION

In this paper, I presented my approach to designing guided photography interfaces for learning and described two concrete projects that take this approach, one that helps with the exploration phase of photography, and another that helps with the refinement phase of photography. These interfaces both showed promise in terms of both the learning and creativity principles defined. I am excited to further build on this approach and to hopefully touch on other learning principles, such as personalization, as well as other visual principles, such as balance and alignment. The goal of my work is to allow people to more actively develop an understanding of their aesthetic preferences, as well as to more reliably capture their artistic intentions. I hope to design camera interfaces and interaction paradigms that empower users of all skill and ability levels with photographic artistic expression.

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. Cynthia L. Bennett, Jane L. E, Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 76, 12 pages. DOI:http://dx.doi.org/10.1145/3173574.3173650
3. Matthew Cong, Kiran S. Bhat, and Ronald Fedkiw. 2016. Art-directed Muscle Simulation for High-end Facial Animation. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '16)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 119–127. http://dl.acm.org/citation.cfm?id=2982818.2982835
4. Fil Hunter, Steven Biver, and Paul Fuqua. 2015. *Light Science & Magic: An Introduction to Photographic Lighting*. Focal Press.
5. Yuki Koyama, Daisuke Sakamoto, and Takeo Igarashi. 2016. SelPh: Progressive Learning and Support of Manual Photo Color Enhancement. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2520–2532. DOI: http://dx.doi.org/10.1145/2858036.2858111
6. Li-Jia Li, Richard Socher, and Li Fei-Fei. 2009. Towards Total Scene Understanding: Classification, Annotation and Segmentation in an Automatic Framework. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2036–2043.
7. 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
8. Paul Merrell, Eric Schkufza, Zeyang Li, Maneesh Agrawala, and Vladlen Koltun. 2011. Interactive Furniture Layout Using Interior Design Guidelines. In *ACM SIGGRAPH 2011 Papers (SIGGRAPH '11)*. ACM, New York, NY, USA, Article 87, 10 pages. DOI: http://dx.doi.org/10.1145/1964921.1964982
9. 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.
10. Martez E. Mott, Jane L. E, Cynthia L. Bennett, Edward Cutrell, and Meredith Ringel Morris. 2018. Understanding the Accessibility of Smartphone Photography for People with Motor Impairments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 520, 12 pages. DOI: http://dx.doi.org/10.1145/3173574.3174094
11. Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, Winston H. Hsu, and Shao-Yi Chien. 2011. Scenic Photo Quality Assessment with Bag of Aesthetics-preserving Features. In *Proceedings of the 19th ACM International Conference on Multimedia (MM '11)*. ACM, New York, NY, USA, 1213–1216. DOI: http://dx.doi.org/10.1145/2072298.2071977