Dynamic Guidance for Decluttering Photographic Compositions

Jane L. E
ejane@cs.stanford.edu
Stanford University
Stanford, CA, US

Kevin Y. Zhai*
kevin.y.zhai@gmail.com
Air, Inc
Brooklyn, NY, US

Jose Echevarria
echevarr@adobe.com
Adobe Research
San Francisco, CA, US

Ohad Fried
ofried@idc.ac.il
The Interdisciplinary Center
Herzliya, Israel

Pat Hanrahan
hanrahan@cs.stanford.edu
Stanford University
Stanford, CA, US

James A. Landay
landay@stanford.edu
Stanford University
Stanford, CA, US

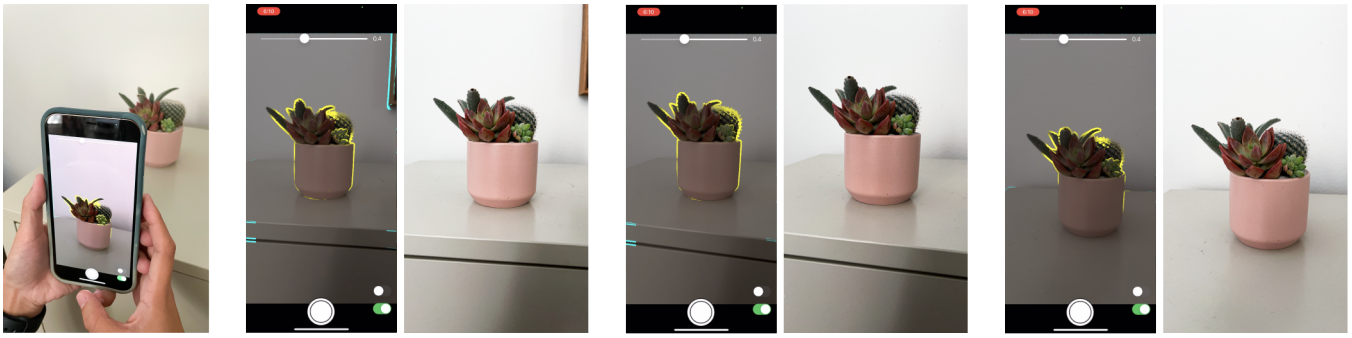

Figure 1: We present a photography guidance tool implemented as a capture-time app. Our tool allows the user to analyze the current clutter in the scene, and adjust until they achieve an image with less clutter. Here we show a sequence of overlays that the user sees as they make intermediate adjustments from their initial framing to achieve their final image. For the initial and final images, we show the overlay on the left and the corresponding photo on the right. The user notices the corner of the painting peeking into the photo in the initial frame, shifts the camera to the left to remove it from frame, but still notices the additional clutter caused by the strong line on the shelf in the foreground of the image in the intermediate frame. The user repositions the plant to the corner of the shelf to achieve the final image with a clean background.

ABSTRACT

Unwanted clutter in a photo can be incredibly distracting. However in the moment, photographers have so many things to simultaneously consider, it can be hard to catch every detail. Designers have long known the benefits of abstraction for seeing a more holistic view of their design. We wondered if, similarly, some form of image abstraction might be helpful for photographers as an alternative perspective or "lens" with which to see their image. Specifically, we wondered if such abstraction might draw the photographer's attention away from details in the subject to noticing objects in the background, such as unwanted clutter. We present our process for designing such a camera overlay, based on the idea of using abstraction to recognize clutter. Our final design uses object-based saliency and edge detection to highlight contrast along subject and image borders, outlining potential distractors in these regions. We describe the implementation and evaluation of a capture-time tool that interactively displays these overlays and find that the tool is helpful for making users more confident in their ability to take decluttered photos that clearly convey their intended story.

CCS CONCEPTS

• **Human-centered computing** → **Graphical user interfaces**; • **Computing methodologies** → **Graphics systems and interfaces**; *Computational photography*.

KEYWORDS

photography, camera interfaces, declutter, composition

ACM Reference Format:
Jane L. E, Kevin Y. Zhai, Jose Echevarria, Ohad Fried, Pat Hanrahan, and James A. Landay. 2021. Dynamic Guidance for Decluttering Photographic Compositions. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21), October 10–14, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3472749.3474755

*This work was done while Kevin Y. Zhai was at Stanford University.

UIST '21, October 10–14, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8635-7/21/10…$15.00
https://doi.org/10.1145/3472749.3474755

1 INTRODUCTION

Storytelling is an important aspect of photography. Photographers employ a wide range of visual techniques to communicate their intended story or narrative to a viewer [8, 9, 21, 27]. The definition of "story" in this sense of the word is often very open. It's more about capturing a moment or invoking an emotion, perhaps a sense of tension or inspiration—in general, it is just about capturing what caught the photographer's eye, e.g., a subject performing an action in their current environment. In fact, it is desirable for the photo to leave a gap in the narrative such that certain aspects are left up to the viewer's imagination. Two important factors to telling a clear story are picking a strong subject, and keeping visual distractions to a minimum.

Amateurs frequently make the mistake of taking too few photos in the moment, relying on editing to improve their photos. However, changes that can be made at the editing stages are limited, and often many mistakes cannot be fixed without returning to the photo location. Having unwanted background clutter is one such mistake that can be harder to fix without more significant changes, and can be particularly frustrating as they distract from the main focus of the image. While still at the scene, the photographer has many more options to reframe the image to clarify the subject and remove distractors, helping to improve these images and better direct viewers' eyes.

On the other hand, experienced photographers tend to capture many photos of a given scene: they know how to consider different options (e.g., composition, lighting, or pose), and recognize the challenges of not having the option to physically move the camera or elements in the image at edit-time. We see similarities in this photographic process and the design process—photographers are essentially iterating on their design (or image) in the camera as they consider these different aspects of photography and storytelling.

Designers have long known the benefits of quickly testing many ideas. Bill Buxton describes the benefits of ideating through sketching due to its flexibility "enabl[ing] ideas to be explored quickly and cheaply" [10]. Designers use sketching to externalize ideas that are still vague in their minds. In doing so, they can spot potential unexpected issues and refine or be spurred to explore new ideas based on these observations [23, 34]. Similarly while taking photos, it is cheap to iterate and generate more "prototypes" or photos in the moment. The photographer can test out ideas just by moving the camera around in space, or by taking photos and immediately reviewing them, and similarly notice any issues like unexpected clutter. Upon leaving the location, these photos become "high fidelity prototypes"—the photographer is more committed to these photos as potential changes are limited and more expensive.

One significant difference between sketching ideas and trying them out in the camera is this level of fidelity. At any point in time, a photo has complete detail. Depth of field can enable some blurring of the background, but each pixel of the image is pigmented, every object in the frame is captured. A sketch on the other hand is a selective representation, an abstracted view of the idea or concept being explored. Removing these low-level details allow the designer to more quickly explore a broad range of high-level concepts. Specifically in photography, it can be easy for a photographer to be too focused on the primary subject of a photo and miss objects immediately surrounding it in the background. Abstracting the image can help them view the image as a whole rather than focus on perfecting the subject's pose or expression.

We are interested in designing camera interfaces that can encourage users to incorporate these exploratory stages of the design process into their photographic process. In particular, we wondered how we might be able to bring some of these benefits of sketching to photography to promote this behavior of intentional exploration, and if that might help users notice unexpected mistakes, such as unwanted clutter in their photos. Photography instructors often give feedback through annotations on the photo [20]. We aimed to design feedback through such annotations directly in the camera. We describe our process towards designing an abstracted annotation of a photo and study how that influences how users address capturing decluttered photos.

Our final annotation design involves color-coded edge highlighting (very similar to the focus peaking feature that can be found on a number of commercial cameras [2, 32]) focused on regions around the subject(s) and image borders where clutter can be most distracting. Edges along the subject borders are colored in yellow, edges along the image borders are colored in cyan, other edges are hidden by default. In our in-camera app implementation, users also have the option to toggle on all edges (see Figure 7). We show edges within the subject(s) also in yellow and background edges not within the subject(s) and image borders in white (see Figure 1).

Specifically in this paper, we contribute:

- our **design process** for determining what an abstraction overlay for decluttering might look like,
- an **interactive in-camera app** that shows this edge highlighting overlay to users as decluttering guidance, and
- a **user evaluation** comparing this overlay to a grayscale overlay, a baseline method that many photographers currently employ for decluttering photos. The evaluation shows that the tool was helpful for making users more confident in their ability to take a clear and uncluttered photo.

In order to realize the proposed interaction, we additionally present a proposal of an **algorithm** for visually annotating potential clutter by highlighting relevant edges around salient objects and around the image border. In this work, our main focus is on the design process and the prototype app—our goal is not to propose an exact abstraction interface, but to understand how this style of in-camera interaction might influence users' photographic process, specifically with regards to decluttering.

2 RELATED WORK

We contextualize our work of designing an abstraction-based camera overlay within the most relevant work in image manipulation and camera guidance.

Image Abstraction and Simplification. Graphics researchers have taken a nonphotorealistic rendering lens on designing algorithms to create stylized image abstractions [15, 22, 36]. Many have taken the approach of trying to realize them as approximate image illustrations with strokes and colored regions, employing a

number of edge detection algorithms for extracting lines for the strokes [7, 11], and smoothing filters or superpixel algorithms for quantizing the color [1, 18, 35]. DeCarlo and Santella consider visual perception through tracking eye movement to determine a structural hierarchy of the image for understanding the importance of different lines in the simplification process [15].

The goal of this line of work tends to be to serve one of two purposes, generating an artistic result, or reducing data for visual communication [22]. Both are useful for our goal of designing visual overlays that are friendly for users to process interactively. Additionally, while these methods focus on generating output from existing media, Winnemöller et al.'s method performs at interactive speeds and produces temporal results, making it appropriate for interactive use on a live camera feed [36].

Image Decluttering. Image processing methods also enable removing clutter in post using a variety of techniques. A photographer can choose to drag a slider to incrementally remove automatically detected distractors [19], or select a region with clutter to be automatically filled based on the surrounding content [5]. A photographer can also choose to segment out the foreground of a portrait to place on a dramatic black or custom backdrop [31], a feature also available directly in-camera on some commercial phones (i.e., "Stage Light" in iPhone's Portrait mode). Nonetheless, these aren't intended to train the user to be more aware of background clutter while they are taking photos. Additionally, they all suffer from potentially jarring artifacts since they rely on pixel manipulation rather than adjusting the real physical objects in the image.

Capture-time Guidance. A range of contextual in-camera guidance exists both in research and in commercial cameras. Some existing work assists users in guiding users towards better lighting for their portraits [16, 24]. Others help the user navigate and position their cameras to achieve better compositions [6, 12, 26], or displaying preferred view proposals [25, 28]. Other work more generally guides the user by interactively displaying a measure of aesthetic quality [25, 29, 33]. While these aesthetics evaluation algorithms may consider clutter, they provide limited information to the user for interpreting these quality measures. We haven't seen work that focuses on helping users declutter their compositions in the camera.

Most similar to our work is the focus peaking feature in commercial cameras and the work of E et al. [17] for highlighting composition. Focus peaking is relevant to our work due to its methods of visualization—it similarly highlights subsets of edges in the images, but instead for the purpose of highlighting regions that are in focus [2, 32]. However, similar to a lot of other existing feedback in commercial cameras, the focus of this feature is more to assist users to more quickly arrive at the appropriate camera settings. By providing these forms of feedback (e.g., the light meter which helps measure overall exposure, or zebra stripes which help determine if regions of the image are overexposed), the user can more quickly evaluate their decisions. We are inspired by the iterative loop of testing and evaluating that these forms of feedback encourage. However, rather than focusing on accuracy and refinement of camera settings, we are interested in understanding how similar overlays can be used for promoting exploration of creative concepts and understanding how to apply higher level photography principles.

E et al. begins to tackle this for overall photo composition by interactively highlighting the perceived composition. We look to extend this idea by designing visual overlays that help the user identify clutter. We find decluttering particularly of interest because it starts to draw on the storytelling aspects of photography. Rather than being tied to a single concept, the process of decluttering an image can involve solutions that touch on several photographic concepts (e.g. composition, lighting, color).

3 ABSTRACTION DESIGN PROCESS

In this section, we describe the steps we took to design our abstraction interface for decluttering images. We hypothesized that abstracting an image could help evenly spread the photographer's attention across the image, effectively drawing the photographer's attention away from the details of the main subject to other areas of the image.

Through our design process, we aimed to answer two questions: (1) Will an abstracted visualization be effective in encouraging the user to see parts of the image outside of the main subject? If so, how does this make them change how they capture the scene? and (2) What is the right visualization to use for a photo "abstraction" to best invoke this type of awareness?

3.1 Wizard-of-Oz Prototype

To test out the concept of abstraction guidance, we started with a low-fidelity Wizard-of-Oz (WoZ) prototype [14] where experimenters manually drew abstraction overlays. These abstraction overlays took shape as rough outlines of the objects in the scene, mostly approximated by basic 2D geometric primitives (see Figure 2 for a few examples).

We informally tested our low-fidelity prototype with 19 participants (9 male, 10 female), 18 to 41 years old ($\mu = 24$). Participation was voluntary and no compensation was given. Two experimenters were involved in the prototype testing, and when possible, both were present. We ran the study using the iPhone's default Camera and Photos apps. Participants were asked to stage a scene of a person (for convenience, this was often one of the experimenters) interacting with an object of their choice, in order to have a more direct concept of "story" in the photos. We then handed them the phone to frame and take a photo of the scene. After capturing a first photo, one of the two experimenters drew the abstraction in dry erase marker(s) on a transparency and taped it to the phone (see Figure 3). The experimenter would return the phone and ask the participant to review the photo with the overlay. The participant would then be given an opportunity to take a new photo of the scene or stick with their current one. We would then apply a new transparency and draw the corresponding abstraction overlay for the new photo for the participant to review. While we didn't ask if they wanted to take another photo after this review step, occasionally (6) participants would ask if they could take a third photo and we would repeat the process above, presenting them with a third overlay.

3.1.1 Prototype Insights. From our WoZ prototype studies, we saw promising signs that the participants in fact noticed high level

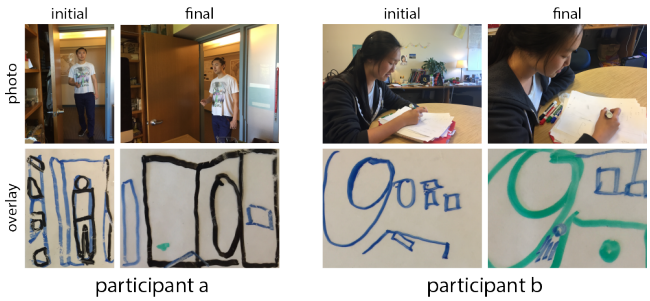

Figure 2: Two pairs of photos (top) from two participants from our WoZ prototype: the initial photo and the final photo upon seeing the photo with the transparency overlay (below). Left to right: (a) The participant notices the clutter on the bookshelves, and rotates the camera and changes perspective to focus attention more on the action of the subject entering the office. (b) The participant notices the clutter on the tables and wall in the background, and tilts the camera angle down to focus attention on the action of the subject writing on the paper.

changes to make to their photograph. In particular, it indeed showed promise as assistance for decluttering. In all but one case, participants noticed clutter in the background upon seeing their photo with the abstracted overlay. In the one case where the participant did not make adjustments to address clutter, a whiteboard mostly filled up the background of the image, and was involved in the action in the image and therefore did not clutter the image, but instead helped tell the intended story.

Participants tended to either adjust the camera position/angle (17) such that the clutter was no longer in frame, or move the clutter (3) out of the scene (or a mix of both). Figure 2 shows two such examples. In both cases, in removing clutter, the participant is also more intentional about framing the elements that are key to the story aspects of the image.

Moving the camera instead of the objects in the scene was often more practical because the objects could not be moved. Participants came up with a range of tactics for adjusting the camera to address clutter. In one situation the participant positioned the subject's head to block the unwanted clutter. Several (5) even discovered that "zooming in" by moving closer to the subject allows them to more easily remove/block clutter and focus on the subject, a technique often taught in photography resources. In doing so, participants additionally changed the overall compositions of the images. In a few cases, they chose to change orientation from portrait to landscape (6) to achieve a decluttered composition that they preferred. One participant noted that the first overlay made her realize that she hadn't achieved the composition she intended to because she was focused on other aspects of staging the scene. This initial photo featured the subject standing at the center of a portrait photo, soda in hand. In her second photo, she focused on capturing the intended composition, shifting the camera to the left to frame a less centered composition. Upon reviewing this second image and overlay, she

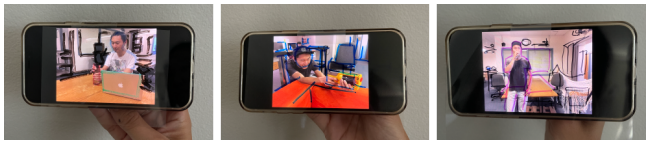

Figure 3: A range of transparencies overlaid on phone as shown to participants saw during the review step of the WoZ prototype study. Transparencies illustrate a basic abstraction of the image through rough outlining.

again reconsidered this intended composition and rotated the camera to landscape to further emphasize the off-centeredness of the composition.

Note that while these participants' behaviors matched what we hoped for, there were limitations to our study design such that they cannot be directly mapped to how a user might respond to seeing this style of overlay interactively in the camera. The user takes a photo before seeing the overlay, and the process of reviewing (with or without) the overlay can influence the photographic process. To test the interactive experience, we needed to find a way to automatically generate these overlays.

3.2 Visualization Designs

Given the observations from our low-fidelity prototype, we were motivated to continue with this concept and move onto the step of answering the second question: What is an "abstraction" of a photo? What should such a visualization look like, and how might we implement them? In particular, we saw that the abstraction overlay seemed most helpful for the purposes of noticing unwanted clutter in an image, so we decided that we would target designing an abstraction overlay that provides decluttering guidance.

3.2.1 Decluttering Principles. To answer our questions on how to design an abstraction overlay for decluttering, we looked to existing literature to better understand how photographers think about directing the viewer's attention for effective storytelling [9, 20, 30].

Photography books describe that contrast is key to directing attention [20]. Our eyes are drawn to regions of high contrast. To achieve the most contrast, the photographer should place light objects on a dark background, or dark objects on a light background. It is good practice to have this contrast around the subject as it will help make the subject distinct from the background. The contrast will clarify the story and declutter the overall image. In this paper, we will refer to this as **subject-background separation** (SBS) In art, this is more commonly referred to as the figure-ground relationship, a Gestalt Psychology principle [20]. We use subject-background separation in our work to make it easier to relate the concept to more familiar terms.

On the other hand, contrast in other regions especially the border of the image, will distract, causing the eye to be attracted away from the focal subject. In particular, contrast near the border of the image can draw the viewer's attention outwards rather than within the image—we call this **image border flicker** (IBF) [20]. Again, this is better known as edge flicker, but we chose to specify "image border" to differentiate it from the term "edge," as "edge" is

Figure 4: A few methods employed by photographers for highlighting contrast as applied on this painting by Emily Friant (left). Left to right: Original painting; A blurred and higher contrast version of the image, a representation of what it might look like to squint at the image; A grayscale version to focus on contrast without aspects of color; A recreation of Glover's outline annotation such that areas where the contrast is possibly too low between the subject and background are shown as gaps [20].

Painting by Emily Friant. Public Domain.

Figure 5: Two sets of example images addressing the two decluttering principles. For each set, we show the original photo that has poor SBS (top) or distracting IBF (bottom), the outlined original with the subject outlined where there is good contrast with the background, and an updated photo where the decluttering principle is addressed. On the left, the texture and color of the origami swan blends in with the blinds in the background, creating unclear SBS in the original photo. However, against the darker backdrop in the updated photo, the outline of the subject becomes much clearer. On the right, the bottle and dark corner of the chair against the light wall along the left border of the original photo cause distracting IBF. Adjusting the camera angle removes the distractors from frame in the updated photo, and even improves the SBS.

often used in computer science with regards to edge detection and identifying segments of sharp discontinuities [7, 11].

Given these principles, we wondered what annotation methods photographers currently used for highlighting clutter. Photographers recommend a number of methods to be able to more easily see the contrast in an image (see Figure 4). These range from squinting at the image to better focus on the contrast with low-level details blurred, to viewing the image in grayscale to better focus on contrast in the absence of color, to explicitly outlining boundaries along which there is clear contrast between the subject and background [9, 20, 30].

Figure 5 shows two examples of images that do not satisfy the decluttering principles, along with approximations of Glover's suggested outlining to emphasize contrast around the subject [20]. We note that while this overlay is helpful for identifying potential issues along the subject-background boundary, it does not help to draw any attention towards potential clutter to address along image borders.

3.2.2 Abstraction Visualization Options. Inspired by this idea of using outlines to highlight contrast and the lack of contrast, we hoped to recreate this outlining as an overlay directly in the camera (see Figure 5), while also extending it to contrast along the image borders. We looked at different approaches to executing this concept as one potential direction to pursue for our abstraction overlay. This concept leads us to two components: a method of line drawing to determine the potential outline, and a method for considering location context relative to the subject/image frame to determine which lines in the image are relevant to the decluttering principles (e.g., if they are along the subject-background boundary, along the image border, or neither).

We additionally were inspired by a line of research in non-photorealistic rendering to generate stylized image abstractions [15, 22, 36]. Specifically, since we knew our goal would be to display the overlay interactively in the camera, we focused on Winnemöller et al.'s real-time and temporally coherent implementation of image abstraction [36]. These papers break down the abstraction process into two components as well: a method of color flattening to smooth out detailed regions, and a method of line drawing.

Combining concepts from these two abstraction ideas, we have three components in total to consider: line drawing, location context, and color flattening. We additionally give ourselves line color (for the purposes of color-coding lines based on their location context), and image darkening (to provide more contrast against the line drawings) as parameters to consider in designing abstraction overlays.

Adjusting these parameters, we designed and implemented a range of potential overlay proposals. Since we did not yet need these to work directly on a camera, these proposals were prototyped in Python. Aiming to capture visual and conceptual diversity in the overlays, we narrowed down the set of overlay options to the 6 in Figure 6 to study further:

(a) color flattening + line drawing (in black)
(b) color flattening
(c) line drawing (in white)

Figure 6: Overlay options as presented to design survey participants for videos. Alongside the video with the set of overlays, we showed still frames of the beginning, middle, and end of the video with the same overlays. This video clip aims to show an example of someone initially noticing clutter (outlet) in the background (top) and moving the object to a cleaner background for the final photo (bottom).

(d) darkened image + line drawing (color-coded: yellow is subject, white is background, cyan is image border)
(e) d without lines within subject
(f) e with white lines removed

Note: overlay (a) features the full image abstraction from Winnemöller et al. and (b) is just the color flattening (or region smoothing) component.

3.2.3 Design Survey. We ran an informal design survey through Qualtrics with 29 participants (demographics information not collected) to try to understand if these overlay visualizations were interpretable by novice photographers, and if there were strong preferences between the overlay options. Participation was voluntary and no compensation was given.

The survey consisted of 12 pieces of visual media (8 photos, 4 videos) shown with the set of different abstractions overlaid. For a single photo, participants would see a row of images, first the original photo and then the 6 overlays applied. For a video, participants would see a concatenated video including the original video side-by-side with the set of overlays applied. The video would automatically loop, but participants also had controls to pause and play. Along with the videos would be 3 rows of still frames from the video to allow participants to compare the overlays across a few frames. Figure 6 shows a set of these overlays for a video as they were presented in the survey. We chose to also include a few videos to show the process of making adjustments while framing a photo, to mimic how the overlay might appear if being used in the camera.

overlay	SBS	IBF	total (weighted total)	% (weighted %)
a	83	29	112 (69.5)	10.4% (10.0%)
b	100	38	138 (102.5)	12.8% (14.7%)
c	112	59	171 (116.0)	15.9% (16.7%)
d	87	142	229 (137.2)	21.3% (19.7%)
e	64	155	219 (124.7)	20.4% (17.9%)
f	73	134	207 (147.2)	19.2% (21.1%)
overall	519	557	1076 (696)	100% (100%)

Table 1: Breakdown of design survey results: count for SBS, IBF, and total per overlay and overall. Since participants were allowed to pick up to 3 overlay choices per question, the weighted totals are computed by dividing a vote for an overlay by the total number of choices selected for that particular question.

Participants were presented with some training describing the two decluttering principles (see Section 3.2.1). They are also informed of our overall goal—to determine which overlay is best for evaluating images based on the two decluttering guidelines. For each of the 12 photos/videos, participants were asked:

- Which overlay helps you to best determine if the image has good **subject-background separation**? (If you don't have a preference between a few similar overlays, feel free to pick up to 3)
- Which overlay helps you to best determine if the image has good **image border flicker**? (If you don't have a preference between a few similar overlays, feel free to pick up to 3)
- (Optional) You will be asked to explain your preferences at the end of the survey, but feel free to explain any specific thoughts you have based on this image/video here.

Finally at the end of the survey, we asked participants to summarize their choices: "Please provide a brief explanation for your choices—why did you find these overlays most helpful? Are there specific characteristics of the overlays that you like (e.g., line drawing, color flattening, image darkening, or color)?" We hoped to gain some understanding of whether or not they had a general feeling of which overlays were helpful (maybe for different scenarios and considerations), and why. Materials for the study can be found in supplemental materials.

3.2.4 Design Survey Insights. Table 1 presents the results from our design survey. The results do not show a clear-cut "best" overlay design. However, overall (d)–(f) were more popular at 21.3%, 20.4%, and 19.2% of total overlay selections made, respectively. These each had the darkened image with selective color-coded line drawing based on their location context in the image. Counts show that the participants found (d)–(f) especially helpful for identifying IBF, whereas (a)–(c) were overall more helpful for identifying SBS.

Both of these conclusions were also supported by the qualitative feedback from participants' general impressions. Participants were quite thorough in their explanations of their interpretations of the components of these proposed overlays. One participant very clearly summarized the potential reasoning for this: *"if the most important thing for subject background separation is the contrast, then the overlays should try to mute lines as much as possible. If the subject is still well separated, then you know you have the contrast. For edge flicker, it almost seems like the opposite. If there's no contrast, then your eye doesn't notice objects on the border. So if an overlay draws outlines of those objects or lights them up, then they're easier to notice"* (P10).

Overall, almost all (23) participants expressed interest in some form of line drawing. Participants noted that the edges helped to define objects (18) and that they were helpful for noticing edges around the image border (5), supporting the observation shown by the overall counts on (d)–(f) being more useful for addressing IBF. Therefore we decided line drawing should be part of our final overlay design. Additionally, some are already somewhat familiar with selective outlining of edges from focus peaking features in commercial cameras, which selectively highlight edges in the image that are in focus [2, 32].

Many (10) participants also mentioned that the darkening of the image was particularly helpful for seeing the lines due to the contrast, but with the caveat that it made the original image harder to see. Participants (4) mentioned that color-coding was helpful, particularly accentuated in videos (P29). However, they also noted that outlining everything can be noisy (9), and especially did not want too many lines within the subject (6).

Further discussing the latter observation that (a)–(c) were more helpful for SBS, several (8) participants specifically noted that the color flattening was helpful for noticing SBS. They mentioned the smoothing of colors being helpful *"because it essentially simplifies the image and makes it easier to analyze the major color contrasts"* (P27) and noting that it helped in being able to *"dissociate the images from what I expect to see into what the colors actually are"* (P24). However, a few (3) noted that this flattening caused blurring that made the subject unclear. Both color flattening and image darkening made the image less clear, so we wanted our final overlay to only have one of these components. Since many participants specifically noted that the darkening was crucial for benefiting from the line drawing, we chose the image darkening component over color flattening for our abstraction overlay implementation.

Given the elements supported by the qualitative feedback, (f) appeared to be the best candidate for our overlay—it included the color-coded line drawings to help define the subject and possible distractors along the borders, while being less noisy than showing all outlines. However, since (d) was the most popular based on overall counts, we decided to go with a hybrid approach for our final abstraction overlay, enabling participants to switch back and forth between showing all lines, including those within the subject and background for additional context, and hiding these extra lines for minimal distraction.

4 IMPLEMENTATION

Results of both steps of our design prototyping process made us hopeful of the potential of an abstraction-based overlay. However, we had yet to try these overlays in an interactive manner. To do so, we needed an implementation that would run interactively on a phone. In this section, I describe our final interface along with the implementation and algorithm to enable it. Note that this final implementation is separate from those used for the prototype overlays for the design survey in Section 3.2.3.

Figure 7: Here we show a few states of the decluttering visualization on the phone. Left to right: The default visualization shows the edges around the detected subject border as well as around the image border over an opaque black background. The user can choose to reduce the opacity of the background (here the user adjusts the opacity from 1.0 to 0.4) to show some of the camera view. The user can also toggle on all edges to see edges within the subject and background as well. Finally, the user can either use the bottom toggle or touch anywhere on the screen to hide the visualization completely.

4.1 Interaction

Based on our learnings from the design survey (Section 3.2.3), we chose to go with a design inspired by a combination of overlay options (d) and (f). Thus, our final abstraction overlay is a context-aware line drawing.

To implement this, our final camera tool has 3 layers: the camera view, a black layer of varying degrees of opacity, and a color-coded outlines layer. As shown in Figure 7, by default the black layer is opaque (opacity is 1.0) and only the edges most relevant to our two decluttering principles (Section 3.2.1), subject-background separation and image border flicker, are visible.

As a reminder, the lines are color-coded such that lines within and immediately around the subject are yellow, lines along the image border are cyan, and remaining lines in the background are white. Lines are color-coded to help users more easily interpret the edges. A solid and defined yellow outline of the subject would mean that there is likely good SBS. Gaps along this edge might signal a lack of contrast—this could mean that the subject is blending into the background, or there are objects directly around the subject that are interfering with the clarity of the subject border. The presence of many cyan edges suggests there might be noise and clutter near the image border that could also take attention away from the subject. The additional white lines help to complement either yellow or cyan lines. These are especially helpful if the main subject is incorrectly identified by the saliency algorithm.

From the default state, the user can adjust the slider at the top to adjust the opacity of the black layer, bringing in more or less of the image color. They can use the toggle to bring in the remaining edges in the subject and background. They can also turn off the visualization by holding a finger down anywhere on the screen or using the toggle at the bottom right.

We choose the most abstracted form of the overlay (solid black, minimal edges) as the default as we imagine users starting in a more exploratory stage. As the user refines the image, they can reduce the opacity or bring in more edges to draw attention to lower level details.

4.2 Mobile Implementation

Our overlay tool is build on top of a basic iOS camera app. To allow participants to focus solely on creating decluttered compositions during the user study, this app has just a camera shutter button and no other camera functionality. On top of this, we implemented our camera overlays and added the necessary UI elements to adjust settings on our overlay. To generate the overlays, the edge detections (for line drawings) and saliency maps (for determining location context) of the current camera image are continuously computed in the background using existing Apple libraries [3, 4]. This app requires iOS 13.0 or higher. Video demonstration of the tool can be found in the accompanying video figure.

4.3 Context-Aware Line Drawing Algorithm

Figure 8 walks through our algorithm for generating our abstraction overlay. Given an image, our tool detects edges throughout the image for the line drawing [3]. However, we want to be able to determine the relevant context in order to focus on edges related to SBS and IBF. In particular, we need to identify a border around the subject. We realized we could estimate the image subjects using object-based saliency maps [4]. We used both algorithms off the shelf. For determining appropriate parameters for the edge detection [3], we iteratively tested parameters to achieve a balance between having enough definition in the edges and having too much noise.

Given this saliency map, we segment the image into regions describing the subject, subject border, image border, and remaining background. Since the saliency maps were object-based, they try to identify entire objects rather than general regions of visual attention. Thus, these maps directly served as a mask of the regions corresponding to salient subjects in the image. We then used simple morphological operations to obtain a subject border mask—specifically, we subtracted a slightly eroded version of this mask from a dilated version to capture the boundary between the subject(s) and the background immediately surrounding. The image border was a fixed pixel width border along the outer borders of the image, and the remaining background was anything between the subject border and this image border. These segmentations are used to classify and color code the edges: yellow designates edges within and around the subject, cyan for edges around the image border, and white for the remaining background edges. We merge these to form two edge-based overlays: one showing all edges color-coded, and the other only showing the relevant edges around the subject and image borders.

5 USER EVALUATION

We wanted to study how users would react to our abstraction guidance tool. We conducted a small formative pilot study to inform our summative user study design. In particular, in running our pilot, we also wanted to understand if a no guidance interface was a reasonable baseline to compare our tool against. Participants for

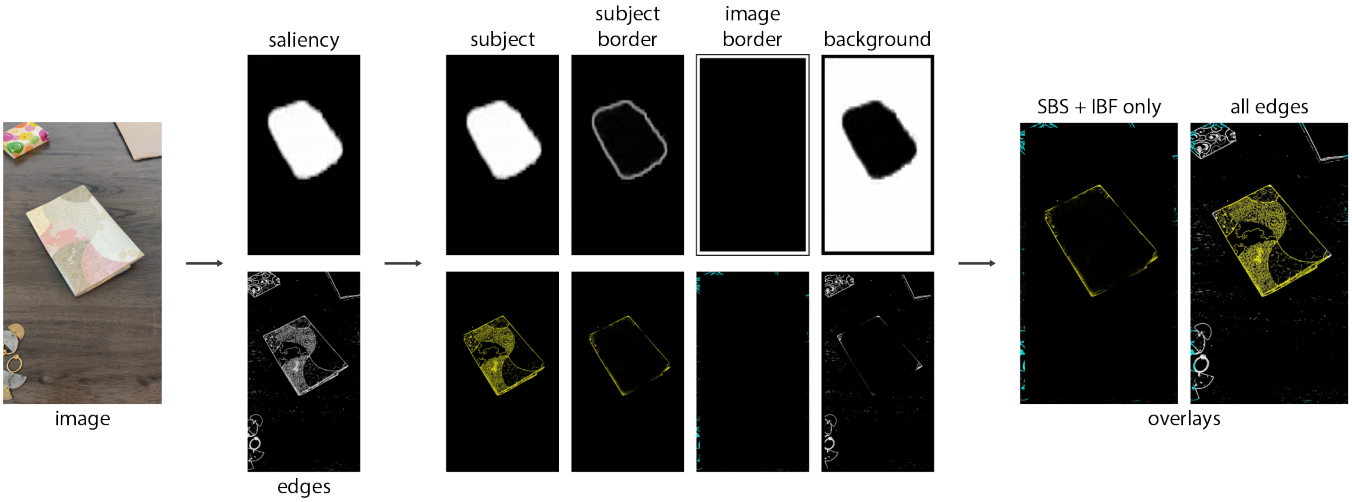

Figure 8: The breakdown of components of our abstraction overlay algorithm. On the left we have the original image. Next, we show results of computing the edges for the overall line drawing, and the object-based saliency map for location context. The next set of images are the intermediates for generating the overlay. On the top, we have the masks based on location context, and on the bottom, the edges masked and color-coded appropriately based on the context: yellow for subject and subject border, cyan for image border, and white for background. At the far right we have the two versions of the final abstraction overlay: showing the subset of edges relevant to SBS and IBF, and then showing the full color-coded line drawing.

both the pilot and larger summative study were compensated $15 for around an hour of their time.

5.1 Study Procedure

Due to the COVID-19 pandemic, we had to run our studies remotely. Overall, we approximately followed the final study design of E et al. for their capture-time composition guidance tool [17] and modify the design for work in remote settings. In their studies, they define two sets of 3 tasks of increasing scale. We liked how the difference in scale resulted in a range of complexity while composing. To modify this task design to work for remote studies, we decided to keep the structure of having 3 tasks per condition at different scales (small, medium, and large), but instead of specifying the subjects, we asked that participants choose their own subjects to photograph. Thus, participants completed 6 photo tasks: 3 at each of 2 locations of their choosing, using the baseline condition at one location and the abstraction guidance at the second. Figure 9 shows a few example images to demonstrate the scale of images chosen by the participants. Since the locations were all different, here we just counterbalanced condition to avoid biases from learning effects.

We ran these user studies over Zoom, asking the participant to adjust the webcam when possible to keep their photographing within view. Additionally to prep for the study, we confirmed beforehand that the participant had a phone running iOS 13 or higher (or found a way to drop off a device in a socially distant manner outdoors). We distributed the app using TestFlight. After getting the participant's consent, we started by walking through having them install TestFlight and subsequently our guidance tool. As additional preparation, we also stepped through using the iPhone screen recording functionality with the microphone to ensure sound was also captured as we asked participants to think aloud as they captured photos for each task.

Participants were provided with a document describing the two decluttering principles (see Section 3.2.1 as training for how to think about decluttering their photo compositions). They were told that following these guidelines would be helpful in telling a clear story in their photos, but also that these guidelines are just to provide some possible perspectives to consider and that participants are by no means required to follow them (e.g., if communicating their story involved intentionally having the subject blend into the background they should do so). After reading the document, we asked that the participants briefly describe the principles in their own words to confirm understanding. For each photo task, we asked participants to focus on the overall clarity of the story. We encouraged them to explore the process of framing the image, but to limit each task to around 1-2 minutes. We quickly walked all participants through the basic (no guidance) camera app. If they were using the tool condition first, they were additionally provided with a brief tutorial describing the overlay.

After each condition, participants were asked to complete surveys with a number of Likert questions (on a 7-point scale) about their experience using the tool along with the Creativity Support Index (CSI) questions (0 to 100) [13]. Following both conditions, they were asked to favorite a single photo per task (for a total of 6), and asked to rate these based on each of the two decluttering principles as well as whether or not they liked the photo in general. We ended the study with open-ended interviews asking about what they liked/disliked about the tool and how the interaction influenced their thought process as they took photos. Again these aspects of the study matched that of the E et al. composition

Figure 9: Participants were asked to choose their own subjects for each task. Here are a few example participant photos at each task scale: (1) small, (2) medium, and (3) large.

study [17], with questions adapted to highlight decluttering principles rather than composition. Materials for the study can be found in supplemental materials.

5.2 Pilot Study

We ran a pilot study (n = 5) to test this study design in a remote setting. Participants (3 male, 2 female) were 23 to 31 years old ($\mu = 29$). For these pilots, we compared our tool to a no guidance baseline condition.

5.2.1 Pilot Study Results. We found that overall, this study design worked reasonably well in the Zoom environment. However, even with just 5 participants we were seeing significant results suggesting that our tool was preferable to no guidance for these photo tasks. In terms of the tools influence on the process, participants felt more confident in their ability to capture a clear photo using the abstraction overlay (Mdn = 5, IQR = 5-6), versus no guidance (Mdn = 4, IQR = 4-4) [Wilcoxon signed-rank test $V = 0$, $p < .05$].

When evaluating their favorited photos per task, participants did believe that the photos captured using our tool had better subject-background separation (Mdn = 6, IQR = 4-5), than those captured using the no guidance baseline (Mdn = 4, IQR = 2-5) [$V = 0$, $p = .003$]. Participants found that the tool helped them achieve these more clear photos (no guidance: Mdn = 4, IQR = 2-4; tool: Mdn = 5, IQR = 5-5) [$V = 0$, $p = .001$]. They also liked their photos more when using the guidance tool (no guidance: Mdn = 5, IQR = 3-5; tool: Mdn = 6, IQR = 5-7) [$V = 0$, $p = .003$].

Thus, even though we did not find significant changes in CSI, we decided that we should compare against a baseline that provided a little more assistance. As we described earlier, photographers will sometimes use a grayscale display in their current practice, to help emphasize contrast in order to consider overall clarity and decluttering. We therefore use grayscale as our baseline condition. It is an active method employed by photographers and thus in some ways does encourage novices to see the image in the ways that expert photographers do.

5.3 Summative Evaluation

We ran remote studies over Zoom with 18 participants (6 male, 10 female), 24 to 32 years old ($\mu = 29$), to understand if the tool would help users declutter photos, and if users felt creative while using the tool. Users experienced two different conditions of the tool: a baseline grayscale overlay, and our tool highlighting edges along the subject and image borders.

5.3.1 Summative Study Results. Again we saw that the tool made the participants more confident in their ability to address the decluttering principles of subject-background separation and image border flicker (Mdn = 6, IQR = 5-7), versus no guidance (Mdn = 5, IQR = 4-6) [$V = 8$, $p = .03$]. For each favorited photo per task, we asked participants to self-assess them based on subject-background separation and image border flicker. Though we actually did not see a significant improvement in overall self-assessed quality in terms of these principles, participants did believe that the tool was helpful for the task of capturing clear and decluttered images (Mdn = 6, IQR = 5-6), versus no guidance (Mdn = 4, IQR = 4-6) [$V = 134$, $p = .003$].

Therefore we see that while participants are more confident in their ability to take clear photos, and found the tool helpful for achieving their favorite resulting photos, they didn't necessarily find that their photos were better with regard to their personal preferences or the decluttering principles.

5.3.2 Summative Study Discussion. We also did not find significant differences in CSI, but did find support for the increased confidence and descriptions of how the tool encouraged participants to explore more in the qualitative feedback. See accompanying video figures for more specific examples from the user studies (including those described here).

Confidence. Participants described feeling like they could take fewer photos, because they could be more confident in each photo they took: *"Usually when I take photos, I take a ton at once, but didn't do that here. I didn't need to because I was being so precise. I noticed myself reconsidering the composition more: e.g. should I have these things in the edges?"* (P15).

For example in Figure 10, this participant (P7) refines the camera angle until the edges in the overlay look the way she wants. She is watching the hook on the top left, making sure it doesn't end up in her shot. Thus upon capturing the final photo, she has intentionally refined some details that were brought to her attention by the overlay, and is therefore more confident that this photo achieves her goals.

Exploration. In addition to encouraging more confidence, we found that the tool further encouraged creativity through exploring the space in new ways: *"It was really helpful with how I take photos because normally it's more just snap and done. This one was more like, can I move things out of the background, can I move the subject to frame it to not have a distracting background? Another thing that I don't normally do is pivot the camera and usually just move within a flat plane"* (P12).

Another participant described that the external representation provided by the interface assisted in the process of exploring the scene and quickly evaluating different options: *"It caused me to experiment more… didn't see it as a rule that I needed to minimize lines, but the tool made it easy to move around and check by that metric, how 'good' it was"* (P9).

Figure 11 shows how an unexpected edge highlight encourages the participant to explore different backgrounds and compositions. This participant (P15) notices an unwanted edge at the bottom of the plant from the stove top, so she tries out different camera angles

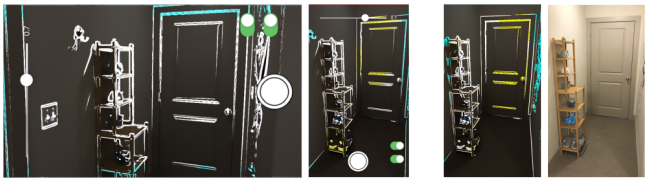

Figure 10: Frames from a participant using the tool—on the left are a set of intermediate steps, on the right is the final overlay that the participant saw as they took a photo and the resulting final photo. Left to right: (a) This participant frames a large scale landscape photo of this area of the room, but notices a lot of extra clutter, (b) upon switching to portrait to better focus on the region of interest, the hook becomes distracting clutter right along the top left image border. The participant makes small adjustments to refine the framing to remove this clutter for the final image.

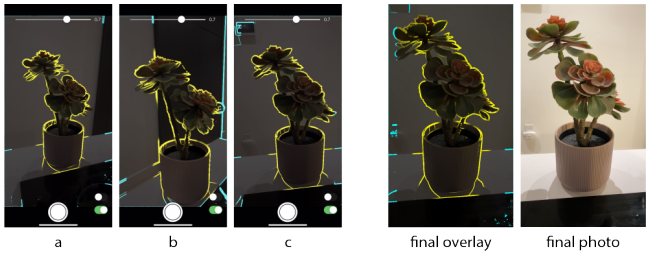

Figure 11: Frames from a participant using the tool—on the left are a set of intermediate steps, on the right is the final overlay that the participant saw as they took a photo and the resulting final photo. Left to right: (a) The participant frames a photo of this plant on the counter, but noticed that the strong edge at the bottom of the pot provides extra unwanted contrast and separation. (b) Attempting to move this edge out of frame, the participant tries to reframe the image with a different background, but doesn't like the dark wall in the background either. (c) The participant turns the camera towards the other direction to continue exploring the background, but notices the thermostat highlighted as it enters the frame at the top left. She keeps the initial edge, but frames a tighter shot of the plant to remove the high contrast thermostat screen from the background.

to remove the stove, but notices the dark panel in the background, which she doesn't want. She turns the camera back towards the white wall, but notices the high contrast object at the top left entering her shot. In this case, the participant actually ends up using the same background as she started with and still includes the edge that she identified as clutter. However, notice how through this process, this choice became much more intentional as she was aware of the alternative options she had in this space and decided that this best suited her preferences.

Subject Identification. An interaction that we didn't expect that emerged from the studies was that participants used the tool's ability to identify subjects to assess how someone might view their photos.

One participant (P1) observed that the tool jumps back and forth in its highlighting of a subject, She interprets this to mean that there is no clear subject in her photo, and confirms that this is consistent with her own perception. Therefore she tries to find a different way to capture this room with a more clear subject.

On the other hand, another participant (P18) had a slightly different interpretation of the lack of a consistent "subject" or in this case, no identified subject. He decided that since the large scale task shouldn't have a single focal subject, the lack of yellow aligned with the expectations of this goal.

For either of these interpretations, the tool has given the participant another perspective from which to consider what it means to clearly capture a subject. As a result, the participant has more clearly considered the concept of telling the story of their subject in this space.

6 LIMITATIONS AND FUTURE WORK

In studying our abstraction interface design, we discovered insights as well as limitations that suggest directions for future work in exploring the potential of abstraction for encouraging rapid iteration in the camera.

6.1 Other Abstraction Designs

Our design survey (Section 3.2.3) provided us with a lot of interesting insight on how participants might imagine an abstraction overlay to assist with decluttering. For example, we saw the potential benefits of color flattening especially for SBS. In fact, a few participants observed that the goals differed between finding clear separation between the subject and background and finding distracting constrast in the background/along the image border. In the first case, contrast is good—so it should be subdued to guarantee the contrast is still clear without details. In the second case, contrast is bad—so it should be emphasized to make sure the photographer realizes it's presence and has a chance to remove it from the frame. Thus, participants actually suggested using a mixed visualization where regions around the subject used color flattening and distractors near the edge are highlighted through line drawings.

6.2 Abstraction for Composition?

Additionally, we found in our low-fidelity prototype that the abstraction also encouraged participants to be more aware of composition. What might our abstraction overlay look like if our focus were instead exploring composition overall rather than decluttering? We would be interested in running a similar design survey asking participants to select overlays based on which best assisted in understanding the overall composition and seeing if reactions significantly differed from our survey focused on decluttering.

In fact, we would be interested in comparing how such an interface might compare to other grid-based composition guidance [17], as well as understanding how these different camera interfaces

might be able to play together. For instance, how they might consider the context from the user's process to determine when each might be most relevant to show.

6.3 Algorithmic Implementation

In designing our overlays, we were somewhat limited in the methods that we used in order to produce something that could be computed interactively. Our proposed technical approach was more a means of enabling the interaction of interest and not a core contribution. We additionally found a simpler algorithm was easier to explain to users when describing the tool for studies, and that users could interpret the visualizations in different contexts more easily.

While there isn't necessarily a clear definition of failure, the main "failure" case with our current algorithm is when the main subject is not properly captured in the saliency map. We considered this in our design and tried to design the visualization to be robust to this case by providing the option of showing all lines. However, even this "incorrect" information can be informative to the user. It communicates that it is hard to identify a subject, and in many cases influenced the user's interpretation of the photo (see Section 5). Due to the modular nature of our implementation, better object detection or object-centric saliency methods methods can be used in the future.

It additionally could be interesting to experiment more thoroughly with different types of edge detection to see which would best match what humans actually perceive as "noise"—e.g., in our user studies, we often found that textures like carpet ended up appearing as a lot of noise.

If we remove the requirement of our camera guidance running interactively, there are further approaches that can be considered both for identifying objects in the scene for location context and for generating line drawings. We can imagine instead of the interactive guidance, having an overlay that is shown upon reviewing previously captured photos (similar to our low-fidelity prototype interaction). These algorithms can run for a little longer without having too much of an impact on the overall interaction. The tool can also perform calculations in the background as the user is focused on taking a series of photos. For example, Fried et al. [19] take a non-edge based approach to finding distractors that better captures what humans annotate as being distractors than most edge detection algorithms. There also are a range of segmentation algorithms that could provide additional location context.

6.4 More Directed Guidance

In this paper, we have described a type of annotation-based guidance. The goal in pursuing this design is to give the user creative flexibility and avoiding having the system express its opinion on how the final image should look [17]. However through the design prototyping and user studies, we've noticed that decluttering seems difficult for people to address, perhaps because it is a new and unfamiliar concept. As a result, people still made mistakes even while being aware of specific clutter to consider and carefully refining their photos. Considering the two decluttering principles of SBS and IBF, they also are somewhat unique in that there is a more defined good and bad (unless the photographer is choosing to intentionally blend the subject into the background or have a very cluttered background). The creativity in decluttering comes at the stage of determining how to address the clutter. Therefore we wonder if it is ok to have a slightly more "opinionated" interface (at least for some training) as is suggested by one of our design survey participants: *"Making it very clear about the distinction between whether the overlay is used to indicate if the two properties are good, bad, or both would be very helpful"* (P18). In this case, we could imagine having the interface give specific guidance such as rotating the phone, removing an object from the scene, or moving the camera closer to the subject. Giving the phone even more authority, one could imagine integrating Fried et al.'s method for automatically remove clutter from the scene [19] directly into the phone.

7 ACKNOWLEDGEMENTS

We would like to thank Griffin Dietz for her assistance in transitioning to virtual user studies, Mitchell Gordon for his feedback and insightful discussions, and Matthew Cong and Charlene Seto Kung for their help with figure photos. We would also like to thank all of our participants for their time and feedback. Our research is supported by the Brown Institute for Media Innovation, Adobe Research, the Microsoft Research Dissertation Grant, and the Hasso Plattner Institute Design Thinking Research Program.

8 CONCLUSION

Much of the effort photographers put into designing a photo is to help more clearly communicate their intended story. Clutter in an image can greatly detract from the power of an image both in terms of effective communication and in terms of visual appeal. However, it can be easy for photographers to miss this clutter in the moment while paying attention to details in the subject. Inspired by the use of sketching in design to capture higher level structure, our goal was to bring some of the benefits of the abstraction in a sketch-like representation to photography. In this paper, we walked through the design process behind our abstraction overlay and guidance tool. We demonstrated this capture-time tool and how it encourages users to explore creative options to address the concept of decluttering.

REFERENCES

[1] Radhakrishna Achanta and Sabine Susstrunk. 2017. Superpixels and Polygons using Simple Non-Iterative Clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*
[2] Eric Adams. 2021. How to Use Focus Peaking to Take Better Photos. https://www.gearpatrol.com/tech/a330370/how-to-use-focus-peaking-photography/
[3] Apple. [n.d.]. CILineOverlay. https://developer.apple.com/documentation/coreimage/cilineoverlay
[4] Apple. 2019. Framework: Vision. https://developer.apple.com/documentation/vision
[5] Shai Avidan and Ariel Shamir. 2007. Seam carving for content-aware image resizing. In *ACM SIGGRAPH 2007 papers*. 10–es.
[6] Soonmin Bae, Aseem Agarwala, and Frédo Durand. 2010. Computational rephotography. *ACM Trans. Graph.* 29, 3 (2010), 24–1.
[7] Mitra Basu. 2002. Gaussian-based edge-detection methods-a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 32, 3 (2002), 252–260.
[8] Finn Beales and Alex Strohl. 2020. *The Photography Storytelling Workshop: A five-step guide to creating unforgettable photographs.* White Lion Publishing.
[9] Bruce Block. 2013. *The Visual Story: Creating the Visual Structure of Film, TV and Digital Media.* Routledge.
[10] Bill Buxton. 2010. *Sketching user experiences: getting the design right and the right design.* Morgan kaufmann.
[11] John Canny. 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* 6 (1986), 679–698.

[12] Scott Carter, John Adcock, John Doherty, and Stacy Branham. 2010. NudgeCam: Toward targeted, higher quality media capture. In *Proceedings of the 18th ACM international conference on Multimedia*. 615–618.

[13] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 21.

[14] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies—why and how. *Knowledge-based systems* 6, 4 (1993), 258–266.

[15] Doug DeCarlo and Anthony Santella. 2002. Stylization and abstraction of photographs. *ACM transactions on graphics (TOG)* 21, 3 (2002), 769–776.

[16] Jane L. E, Ohad Fried, and Maneesh Agrawala. 2019. Optimizing Portrait Lighting at Capture-Time Using a 360 Camera as a Light Probe.. In *UIST*. 221–232.

[17] Jane L. E, Ohad Fried, Jingwan Lu, Jianming Zhang, Radomír Měch, Jose Echevarria, Pat Hanrahan, and James A Landay. 2020. Adaptive Photographic Composition Guidance.. In *CHI*. 1–13.

[18] Adam Finkelstein and David H Salesin. 1994. Multiresolution curves. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. 261–268.

[19] Ohad Fried, Eli Shechtman, Dan B Goldman, and Adam Finkelstein. 2015. Finding distractors in images. In *Proceedings of the IEEE Conference on Computer Vision and pattern Recognition*. 1703–1712.

[20] Tavis Leaf Glover. 2014. *Canon of Design: Mastering Artistic Composition.* Tavis Leaf Glover.

[21] About JasenkaG. 2020. The 6 Fundamentals Of Storytelling Through Photography. https://www.lightstalking.com/storytelling-through-photography/

[22] Henry Kang, Seungyong Lee, and Charles K Chui. 2008. Flow-based image abstraction. *IEEE transactions on visualization and computer graphics* 15, 1 (2008), 62–76.

[23] James A Landay and Brad A Myers. 1995. Interactive sketching for the early stages of user interface design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 43–50.

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

[25] Shuai Ma, Zijun Wei, Feng Tian, Xiangmin Fan, Jianming Zhang, Xiaohui Shen, Zhe Lin, Jin Huang, Radomír Měch, Dimitris Samaras, and Hongan Wang. 2019. SmartEye: Assisting Instant Photo Taking via Integrating User Preference with Deep View Proposal Network. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 471, 12 pages. https://doi.org/10.1145/3290605.3300701

[26] 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.

[27] Pedro Quintela. 2019. How to Create Storytelling Photography That Will Captivate Your Audience. https://www.photoblog.com/learn/storytelling-photography-tips/

[28] Yogesh Singh Rawat and Mohan S. Kankanhalli. 2015. Context-Aware Photography Learning for Smart Mobile Devices. *ACM Trans. Multimedia Comput. Commun. Appl.* 12, 1s, Article 19 (Oct. 2015), 24 pages. https://doi.org/10.1145/2808199

[29] Katharina Reinecke and Krzysztof Z Gajos. 2014. Quantifying visual preferences around the world. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 11–20.

[30] Fred Ross. [n.d.]. *The Art of Composition: A Simple Approach for Creating Masterful Designs.* The Art of Composition.

[31] Xiaoyong Shen, Aaron Hertzmann, Jiaya Jia, Sylvain Paris, Brian Price, Eli Shechtman, and Ian Sachs. 2016. Automatic portrait segmentation for image stylization. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 93–102.

[32] Shawn C. Steiner. 2015. What is Focus Peaking? https://www.bhphotovideo.com/explora/photography/tips-and-solutions/what-focus-peaking

[33] 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*. 1213–1216.

[34] Masaki Suwa and Barbara Tversky. 2002. External representations contribute to the dynamic construction of ideas. In *International Conference on Theory and Application of Diagrams*. Springer, 341–343.

[35] Carlo Tomasi and Roberto Manduchi. 1998. Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*. IEEE, 839–846.

[36] Holger Winnemöller, Sven C Olsen, and Bruce Gooch. 2006. Real-time video abstraction. *ACM Transactions On Graphics (TOG)* 25, 3 (2006), 1221–1226.