Supplemental Materials for Exploratory Stage Lighting Design using Visual Objectives

1 Comparison with Design Galleries

In comparison to Design Galleries, our method allows users to very quickly focus on the part of the design space of interest to them. The same space can be explored with Design Galleries, however the full design space would have to be sampled and organized into a hierarchy. We show in Figure 1 the challenges with using random sampling to try to find a specific part of the design space in a limited number of samples.

2 Workflow Observations

Users demonstrated a variety of different methods of working with the objectives and the interface. We describe some of the common workflows to help interpret the results and discuss how the concepts could be integrated into the current lighting design process.

Primary Concepts. One of the first operations performed by almost all users of our interface is to pick two objectives, a color objective and intensity objective, and apply those two objectives to the entire stage. These objectives serve as the primary objectives used in the scene. If the results do not quite line up with what the user is looking for, they adjust the parameters of the primary objectives and redo the search. Once the stage looks roughly in line with their expectations, they remove the primary objectives and start refining their design. Users typically then pick a different objective to apply to a smaller region of the stage to further refine their design. Under this workflow, users typically avoid using pins and instead focus in on the parts of the stage that require their attention.

Targeting and Selection. Many users chose to limit the region of the stage the objectives apply to instead of excluding regions from their objectives. This is likely due to the small size of the light lab, where every area can be cleanly isolated, and pins were simply not necessary. Some users also had some difficulty intuiting what lights would be selected when a concept was applied. The difficulty comes from the fact that physical lights tend to spill outside of a particular selected region onto the floor or other background elements, and the selected lights are those that *influence* the selected region, not *contained within* the selected region. We provided users with a toggleable view that shows the selected lights to help visualize how selection works.

Fine Tuning. One of the most common complaints about our interface was the removal of the individual light slider controls found in the traditional interface from our interface. Users were often able to quickly get to the point with our interface that they wanted to start doing detailed adjustments to individual lights. While we did prove the views as a replacement for the sliders, users mostly ignored the views, and expressed a preference to just show all the lights in the list from the standard interface. The speed at which users reached this point is indicative of the effectiveness of our system, as it allows expert designers to quickly start doing tasks that require their expertise instead of spending those first few minutes just setting up the scene, and it allows intermediate users to create better designs more quickly.

As a result of the speed of the interface, users actually spent *less* time using the objectives by themselves than they did tweaking the results generated from the objectives. This is a positive indication that the interface we built accomplishes the goals of effectively exploring a reasonable space of designs as defined by the user's objectives by getting them to a point where they can perform fine-tuning operations quickly.

Layering Concepts. Many users chose to work with one objective at a time rather than using multiple objectives at the same time, and some were confused by the need to manipulate intensity and color concepts separately. The confusion may stem from some UI problems found in the version of the tool used in the study, but also suggests that people can only hold a few concepts in their mind when performing a complex design task. Keeping the list of active concepts displayed in the interface helps with keeping track of references while designing.

3 Prior User Studies

This section presents user studies run for an earlier draft of the paper. We ran two experiments, and present the results here.

The first study is modeled after a common theatrical scenario: a designer is given a prompt and asked to quickly prepare *multiple candidate lighting designs* to present to the director of a stage production. In this scenario it is desirable to create a diverse range of results for director to review, as the hope is that at least one of these designs aligns well with the director's vision. In this study (detailed in Section 3.1), we found that participants were always able to produce a high-scoring lighting design within ten minutes using the visual objectives interface. Participants did not always succeed in generating a high-scoring design using a traditional slider-per-light interface.

The second study (Section 3.2) was patterned after the previous lighting design interface study by Kerr and Pellacini [2009] and is designed to directly compare the performance of a lighting designer using the visual objectives interface against that of a baseline slider-per-light interface. In this study, participants used each interface to create a *single lighting design for each prompt*, and we assessed their experience using a post-study questionnaire. We also had expert lighting designers evaluate the resulting images. Participants found the visual objectives useful but were unable to produce scenes that scored well under all evaluators using either interface, motivating the need for exploratory design to generate several candidates.

3.1 Exploratory Design Study

In this study, we asked participants to act as the lighting designer for the director of a hypothetical theater production. To simulate a situation where a designer needs to provide quick feedback to the director, designers were only allowed to spend ten minutes lighting each scene. The goal is for the designer to produce a diverse set of designs for the director with the hope that the director highly approves of one of the designs. It is common for professional theater productions to feature over twenty scenes, so a limited amount of design time per scene is a realistic operating scenario in the early stages of development.

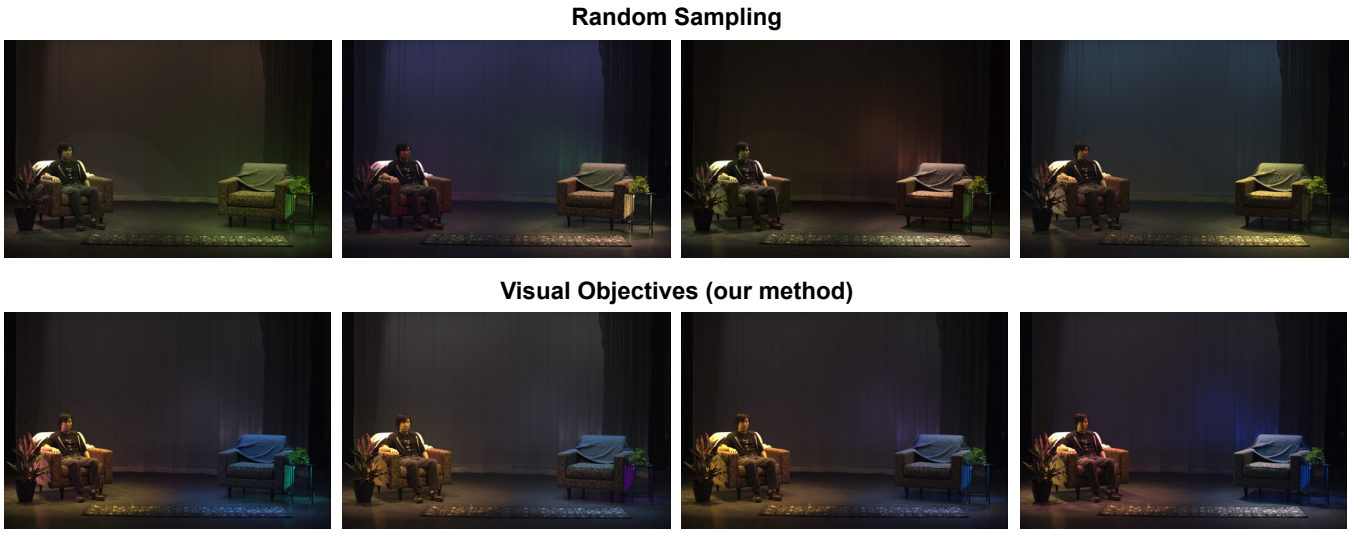


Figure 1: *Comparison between random samples and designs generated by our approach. The target design the user is trying to achieve is the same as Figure 2 in the main paper.*

3.1.1 Experimental Setup

Tasks. Each participant was asked to complete two lighting design tasks taken from in-class exercises in an intermediate undergraduate lighting design class. For each task, the user is given a general prompt describing the scene along with environmental annotations. The two prompts used for this experiment were: *"A man sits by a fire in an open field. The fire is comforting but the overall scene feels a little creepy,"* and *"A man walks across the path at the back of the stage at dawn, feeling relaxed."* By design, these tasks are open-ended, and different color and intensity configurations can successfully satisfy the prompt. All tasks are performed on the same lighting stage featuring 190 lights and 22 light groups.

Interfaces. We compare two interfaces: the *visual objective* interface and a baseline *sliders* interface. In the sliders interface, participants select individual lights or light groups and directly specify color using a standard color-selection interface. The sliders interface is comparable to the current lighting consoles from the theater industry and familiar to the expert participants. To focus the evaluation, we did not allow individual slider-based control of the lights in the visual objective interface, even though in practice we expect a complete tool would offer both interface paradigms.

Participants. The participants included three experienced theatrical and cinematic lighting designers and two novices. Each participant had normal color vision and was asked to perform two tasks. Participants started with a 20-minute tutorial session where they familiarized themselves with both interfaces. Then participants were given ten minutes with each interface to complete one of the tasks. The order of the interfaces used by each participant was randomized. Participants could produce as many designs as they desired in the allotted time.

Professional assessment. We recruited an expert lighting designer (a professor of theatrical lighting design) to serve the function of the "director" and judge the quality of the designs produced by participants. The judge was presented designs (with the corresponding task prompt) in random order and asked to assign a grade from 1 to 5 (with 5 being the best) based on how well the the design achieved the specified goal. We refer to this grade as the *task score*. The diversity of the designs created by each participant was also rated on a 1-to-5 scale; this was used to detect cases where the artist generated a large number of similar designs.

	Visual Objectives	Sliders
Num Designs Created (avg.)	3.6	2.4
Max Task Score (avg.)	5.0	4.2
Avg Task Score (avg.)	4.3	4.0
Diversity (avg.)	5	3.4

Table 1: ***Results of exploratory design study.*** *All five participants were able to create a top-scoring design using the visual objectives interface. Only three of five managed to do so using the sliders interface. (Average per-participant task scores, number of designs, and diversity are averaged across all participants for each interface.)*

3.1.2 Study Results

The results of this experiment are summarized in Table 1. Recall that the goal of this scenario is for the director to find at least one design that they like. We found that all participants were able to generate at least one scene with the maximum rating of 5 using visual objectives interface. Only three of the five participants were able to gain a score of 5 when using the sliders interface. Table 1 also suggest that the director determined that designs created using the visual objectives interface were more diverse than those created using the sliders interface. This greater diversity may have been useful in allowing the director to always find a design that they like.

Designs created by the participants in this study, as well as the individual scores for each design, can be found in the supplemental material.

3.2 Interface Comparison Study

Our second user study focuses on assessing the performance of designers using the visual objectives interface as well as on the experience of using the system. We designed this study to closely follow prior experiments on lighting design interfaces by Kerr and Pellacini [2009]. Participants use either the visual objectives interface and the baseline sliders interface to create a single lighting design

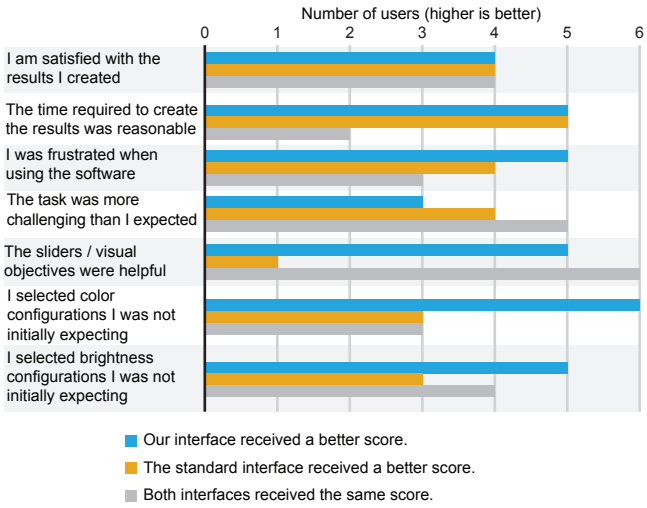


Figure 2: *Analysis of the results of our first user study questionnaire (Section 3.2). We compare our interface and the sliders by counting how many users gave a better score to ours. Depending on the question, a high score or a low score is better, and we define a "good score" accordingly (i.e., it is a high score if the question is about a positive property, and a low scores if it is about a negative property) so that a larger number of users in the chart above is always a better assessment. Our interface receives similar or better assessment for all the questions except the one about how challenging the task was.*

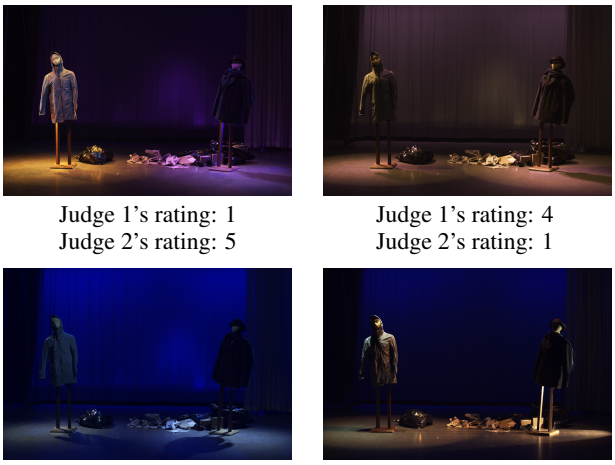
in response to a prompt. The results are judged for task relevance and overall quality. Designers were also asked about their experience with the interfaces to assess how well our interface assists the participants' design process.

3.2.1 Experimental Setup

This study uses the same setup as in Section 3.1.1 with the following changes. Instead of generating multiple designs, participants were asked to spend all ten minutes allotted to each task creating a single design. The study involved 12 participants, who each completed four design tasks: two with the visual objectives interface and two with the baseline sliders interface. After completing the tasks for each interface, the participants completed a survey.

The prompts for three of the tasks used in this study, along with example output from the participants is shown in Figure 6. The fourth task was a transfer task (referred to as an *open trial* by Kerr and Pellacini [2009]). In this task, the participant is presented with a photograph from a theatrical scene and asked to transfer the lighting configuration in the photograph to the target stage. This task is more constrained since the artist's choice of colors and relative intensities is intended to follow the input photograph as closely as possible. All tasks are performed on the same lighting stage, but different stage props are used in the three scenarios to contextualize the task (Figure 6). The stage features 44 lights and 13 light groups.

To judge the quality of resulting designs, we recruited an additional professional lighting designer (to augment the judge used in the first study) and asked both judges to evaluate each design's *task score* (relevance to task) and *quality score* (the overall quality of the lighting without regard to the task).



Judge 1's rating: 1
Judge 2's rating: 5

Judge 1's rating: 4
Judge 2's rating: 1

Judge 1's rating: 4
Judge 2's rating: 1

Judge 1's rating: 5
Judge 2's rating: 2

Figure 3: *Four results from our interface comparison study for task (A) in Figure 6. Due to individual preferences for particular lighting designs, expert judges showed strong disagreement in assessing the quality of the scenes.*

	Visual Objectives	Sliders
Avg Task Score	3.1	3.3
Avg Quality Score	3.0	3.2

Table 2: ***Results of interface comparison study.*** *Average scores for the two interfaces in our interface comparison study. All results were evaluated on a 1-to-5 Likert scale (higher is better).*

3.2.2 Study Results

Figure 6 provides a sampling of designs created by participants during the user study, and shows the visual objectives used to achieve these designs. The task prompts, all artist-made scenes, and the full results of the questionnaire are provided as supplemental materials.

Observations and Participant Experiences. The results of the user study questionnaire are displayed in Figure 5. Participants found visual objectives useful when performing the design tasks (avg. 4.25) and all participants had positive experiences with the UI for targeting visual objectives on stage (avg. 4.42). Responses suggest that participants found the design candidates generated by the system to be a useful part of their design process, even though sometimes these candidates did not line up with participant expectations (avg. 3.16). This could be viewed as a failure of sampling to capture visual objectives in some instances, but also could be viewed as a useful property that enables participants to encounter design candidates that would not have otherwise been considered.

Participants indicated that they were less frustrated using the visual objectives interface when performing tasks (avg. 2.83) compared to the baseline (avg. 3.08), but also reported that accomplishing tasks was slightly more difficult than expected using visual objectives (avg. 3.75 vs avg. 3.25 for the baseline). This may be due to unfamiliarity with the visual objectives interface, or that participants simply expected the interface to be easier to use after the tutorial session.

When working with the interface, we observed that participants typically began applying a color and intensity objective to the entire stage in order to quickly get to a point close to their intended design. If the objectives did not result in the desired effect, participants often chose a different set of visual objectives and repeated the

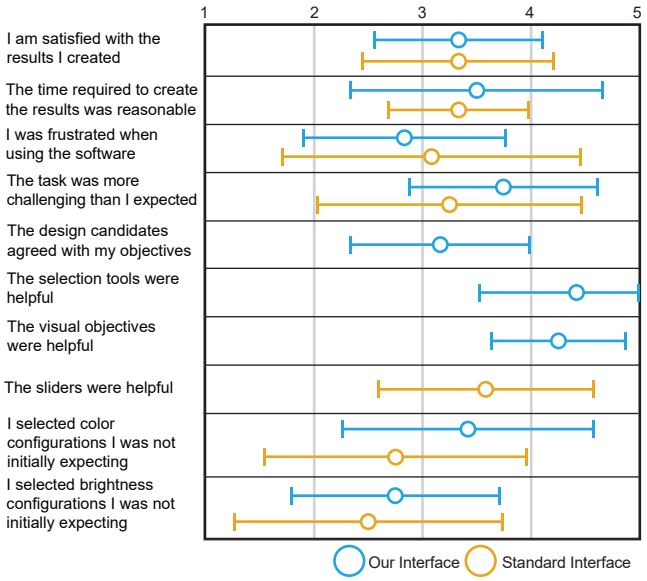


Figure 4: *Average response to each question on the interface comparison study questionnaire. Responses are on a 1-to-5 scale, with 1 indicating disagreement, and 5 agreement. Error bars represent one standard deviation.*

process. Participants were able to reach a design that they liked quickly but expressed a desire to use traditional sliders to perform small modifications on the scene to perfect it; as expected this desire was especially prevalent among experienced participants familiar with direct light manipulation interfaces. More detailed observations of the participants can be found in the supplemental materials.

Professional Assessment. Two expert lighting designers judged the quality of the participant's designs (results summarized in Table 2). In contrast to the first user study, overall task score for both interfaces, *when results are average across both judges*, was low—at most 3.3 (similar results hold for the quality score as well). Further examination revealed this lower average is due to high disagreement between the two expert judges. (The variance in task score was 0.9 on our 1-to-5 scale.) Although both expert judges assigned some scenes high ratings, they were often inconsistent in the designs they preferred. Given a condition like *"summer night in a city"*, the judges did not agree over an acceptable range of colors and intensities, and were critical of designs that were inconsistent with their expectations. Figure 3 provides examples where the judges disagreed by a spread of 4 or more.

While these results make it difficult to draw quantitatively conclusions about the performance of participants in the study, we believe they (along with the results of the first user study) suggest that there is notable value in the proposed exploratory design interface. The ability to rapidly explore different designs is valuable aid for both individual creative thought and also communication between creative professionals.

3.3 Novice Designer Study

Our second user study examined the extent to which the proposed interface could assist novice lighting designers in their design process. We recruited users with little to no lighting design experience and asked them to create designs for two different scenarios. Users were given as much time as they wanted to create the design.

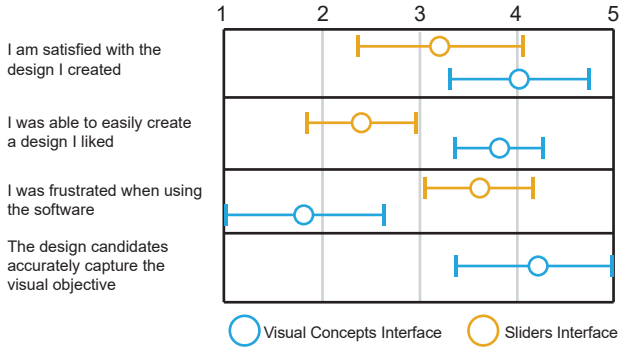


Figure 5: *Average response to each question on the novice study questionnaire in Section 3.3. Responses are on a 1-to-5 scale, with 1 indicating disagreement, and 5 agreement. Error bars represent one standard deviation.*

3.3.1 Experimental Setup

Each participant was asked to complete two lighting design tasks taken from in-class exercises in an intermediate undergraduate lighting design class. For each task, the user is given a general prompt describing the scene along with environmental annotations. The two prompts used for this experiment were: *"It is a late spring evening. A couple sits at a table in their garden watching the sunset.,"* and *"It is a cold winter morning. People have not yet begun the day. A man sitting in a armchair before a fire, looks out the window and enjoys the silence."* By design, these tasks are open-ended, and different color and intensity configurations can successfully satisfy the prompt. All tasks are performed on the same lighting stage featuring 44 lights and 13 light groups (shown in Figure 7).

Participants were trained to use two different lighting design interfaces: a traditional slider-based interface where each light parameters could be adjusted using slider controls, and our visual objectives interface. In the sliders interface, participants select individual lights or light groups and directly specify color using a standard color-selection interface. The sliders interface is representative of current lighting consoles from the theater industry. The chosen interface for each task was randomized for each participant.

We recruited five participants with little to no lighting design experience for this study. Participants were given as much time as needed to complete the lighting design tasks. Participants were also given a pre-selected library of reference images, but were allowed to find images on the internet if they desired (only one participant searched for their own images). After completing both tasks, participants completed a short survey where they rated the self-assessed quality of their design, along with their experience using each of the interfaces.

3.3.2 Results

While participants ended up with scenes that they were satisfied with in both interfaces, they reported a much higher level of frustration with the baseline sliders interface (Figure 5). For novice users, much of the frustration with the sliders interface comes from being unable to predict how changing one set of lights will affect the entire scene, leading to a tedious process of trial-and-error as they attempt to adjust light parameters to match their intended design. With the visual objectives interface, this process was made much easier and faster, since changing a design idea was as simple as selecting a different image from the image library.

(a) It is a summer night in the city. A suspicious man dressed in dark clothes steps out of an alley ready to follow the woman who is passing by.

(b) It is a cold winter morning. People have not yet begun the day. A man sitting in an armchair before a fire, looks out the window and enjoys the silence.

(c) It is a late spring evening. A couple sits at a table in their garden watching the sunset.

Figure 6: *Designs created by participants of the Interface Comparison study using the visual objectives interface. The reference images used by the participants to create their designs are shown in the top left of each design.*

(a) It is a late spring afternoon/evening. A person walks through a garden as the sun sets.

(b) It is a cold winter morning. One person goes about their morning tasks in their home.

Figure 7: *Designs created by participants of the novice study using the visual objectives interface. The reference images used by the participants to create their designs are shown in the top left of each design.*

Novice participants agreed that the system produced good design candidates from their selected images. Most users left model parameters extracted from reference images alone, however one user adjusted the color objective to contain colors that were not present in the original image, suggesting that none of the available research images exactly matched their ideal color palette. Rather than finding another reference image, this participant preferred to manually set their own palette. Some of the scenes created with our interface are shown in Figure 7.

References

KERR, W. B., AND PELLACINI, F. 2009. Toward evaluating lighting design interface paradigms for novice users. *ACM Trans. Graph. 28*, 3 (July), 26:1–26:9.