



**HAL**  
open science

## How Do People Edit Light Fields?

Adrian Jarabo, Belen Masia, Adrien Bousseau, Fabio Pellacini, Diego Gutierrez

► **To cite this version:**

Adrian Jarabo, Belen Masia, Adrien Bousseau, Fabio Pellacini, Diego Gutierrez. How Do People Edit Light Fields?. ACM Transactions on Graphics, 2014, 33 (4), 10.1145/2601097.2601125 . hal-01060860

**HAL Id: hal-01060860**

**<https://inria.hal.science/hal-01060860>**

Submitted on 4 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# How Do People Edit Light Fields?

Adrian Jarabo<sup>1\*</sup> Belen Masia<sup>1\*</sup> Adrien Bousseau<sup>2</sup> Fabio Pellacini<sup>3</sup> Diego Gutierrez<sup>1</sup>  
<sup>1</sup>Universidad de Zaragoza <sup>2</sup>Inria <sup>3</sup>Sapienza University of Rome



**Figure 1:** Example results of light fields edited by different users. Top: A synthetic light field (vase), with ground truth depth information. Bottom: Two real light fields (toys and motorbike) captured with the Lytro camera. In this work we evaluate the benefits of different light field interaction paradigms and tools, and draw conclusions to help guide future interface designs for light field editing.

## Abstract

We present a thorough study to evaluate different light field editing interfaces, tools and workflows from a user perspective. This is of special relevance given the multidimensional nature of light fields, which may make common image editing tasks become complex in light field space. We additionally investigate the potential benefits of using depth information when editing, and the limitations imposed by imperfect depth reconstruction using current techniques. We perform two different experiments, collecting both objective and subjective data from a varied number of editing tasks of increasing complexity based on local point-and-click tools. In the first experiment, we rely on perfect depth from synthetic light fields, and focus on simple edits. This allows us to gain basic insight on light field editing, and to design a more advanced editing interface. This is then used in the second experiment, employing real light fields with imperfect reconstructed depth, and covering more advanced editing tasks. Our study shows that users can edit light fields with our tested interface and tools, even in the presence of imperfect depth. They follow different workflows depending on the task at hand, mostly relying on a combination of different depth cues. Last, we confirm our findings by asking a set of artists to freely edit both real and synthetic light fields.

**CR Categories:** I.3.4 [Computer Graphics]: Graphics Utilities—Paint systems; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

**Keywords:** light field, edit, interfaces

\*Joint first authors

## 1 Introduction

Light fields [Levoy and Hanrahan 1996; Gortler et al. 1996] are rapidly gaining popularity as an alternative to digital photographs. Consumer light field cameras already exist (such as *Raytrix*<sup>TM</sup> or *Lytro*<sup>TM</sup>), even compact enough to be included in mobile devices [Venkataraman et al. 2013]. As the number of captured and shared light fields increases, the need for editing tools arises as well. However, current light field editing techniques are limited to changing perspective or focus, or to applying some pre-defined filters [Lytro Inc. 2013; Vertical Horizon 2013]. As opposed to the well-established editing of 2D images, user interfaces to edit light fields remain largely unexplored.

Editing light fields is a challenging task for several reasons. First, a light field is a four-dimensional data structure while the majority of displays and input devices today are designed for two-dimensional content. Second, light fields are redundant which implies that any local edit on a light field needs to be propagated coherently to preserve this redundancy. Finally, while light fields provide a vivid sense of depth, this depth information is not encoded explicitly. Light field user interfaces must take these properties into account to present the visual information in a legible way and to minimize redundant work for the user.

Current literature proposes two general paradigms to interact with light fields on two-dimensional devices. *Multiview* [Zhang et al. 2002; Shum et al. 2004; Wang et al. 2005] relies on parallax to convey disparity. Interaction is constrained to the angle of view: Users can change the viewpoint, equivalent to limited displacement of a virtual camera. *Focus* interfaces [Davis et al. 2012], on the other hand, rely on depth-of-field blur. Interaction is constrained to the focus plane, which the users can change. However, neither of these have been thoroughly analyzed or validated.



In this paper, our goal is to explore light field editing interfaces from a user perspective, and to provide a quantitative and qualitative evaluation of the effectiveness of different approaches. To do this, we build interfaces based on the two main existing light field editing paradigms, conduct a wide range of experiments with novice users, and analyze their performance and preferences. The experiments grow in complexity, and are divided in two studies. In the first one, we ask participants to perform tasks with both interfaces, based on simple edits and using synthetic light fields to be able to use accurate depth information. From the gained insights, we design a third interface, which additionally incorporates a series of new tools. This new interface and tools are used in our second study, which covers more advanced tasks (e.g. changing an object's appearance, or colorization). Moreover, this second study uses real light fields as input, which means that the recovered depth is not perfect. In designing the experiments, we seek to answer questions such as: Can users edit light fields, similar to common image editing? How do people interact with the 4D structure of a light field? What is the best method to locate a position in such 4D space? What is the preferred workflow for editing? Is the currently imperfect depth in real light fields good enough for editing? Does it affect performance?

**Conclusions** The analysis of the collected objective and subjective data allows us to draw the following conclusions:

- novice users can edit captured light fields of different nature, although with varying workflows;
- users rely on depth information frequently, finding it essential in many tasks;
- inaccuracies in depth information do not hinder users' performance when editing;
- a *multiview* approach is essential for navigating and viewing the light field edits;
- on the other hand, a *focus* approach is attractive to users because of the feeling of control it conveys; alternative tools, however, can also provide that feeling of control without focus;
- we also present tools with which users exploit the extra dimensionality of light fields and which, in combination with the editing paradigms, support artistic exploration.

**Limitations and scope** The findings of our work can only be conclusive to the interfaces, tools and depth reconstruction methods used in the experiments. While the interfaces are the two most common interfaces found in previous literature on light field manipulation, other types of interfaces could be proposed and studied. We limit our study to local point-and-click tools (as opposed to global filters or propagation-based editing tools), but they represent common edits in most image editing software; additionally, point-and-click interaction represents the building block for more complex editing tools, such as selection or local filtering. Finally, our analysis on the influence of errors in depth is limited to the real depth maps reconstructed with three state-of-the-art depth reconstruction methods. However, despite these issues, we believe our work provides a solid building block for the design of light field editing interfaces, a reference for future researchers and interface designers. Thus, to further validate our methods and allow others to build on our work, we make our interfaces, raw data and analysis available online <sup>1</sup>.

## 2 Related Work

Different interfaces and interaction paradigms have been explored in the fields of computer graphics and human-computer interaction, for instance for lighting [Kerr and Pellacini 2009; Ou et al. 2012], material editing [Kerr and Pellacini 2010; Colbert et al. 2006], video editing [Santosa et al. 2013; Goldman et al. 2008] or painting [Haerberli 1990]. In this paper, we perform the first study focusing on the particular topic of light field editing.

**Light field editing** Most light field editing systems perform consistent operations over multiple views. Seitz and Kutulakos [2002] estimate a voxel-based representation of a light field to propagate local edits, such as painting and scissoring, between multiple views of a scene. Jarabo et al. [2011] propagate sparse edits in a light field based on pixel affinity. Related methods estimate depth in a stereo pair to perform consistent painting and copy/pasting [Stavakis and Gelautz 2004; Lo et al. 2010; Price and Cohen 2011; Kim et al. 2013b]. While depth estimation assumes static scenes, other approaches rely on feature matching to propagate edits over image collections containing deformable objects [Hasinoff et al. 2010; Ha-Cohen et al. 2011; Yücer et al. 2012].

In contrast, other systems require user intervention to indicate correspondences between two or more views. Zhang et al. [2002] and subsequent work by Wang et al. [2005] morph between two light fields by first requiring users to position polygons in several views, constrained by epipolar geometry. Users then indicate corresponding polygons in a second light field to guide the morph. In Pop-Up Light Field [Shum et al. 2004], users segment the light field into multiple depth layers by adjusting a polygon around the silhouette of each object in multiple views. Chen et al. [2005] segment the light field into multiple volumes defined by the user, which can be deformed, while Horn and Chen [2007] present a shader-like language to edit and compose light fields.

There are, additionally, a few prototype displays that allow light field editing using gesture tracking [Marton et al. 2012] or 3D light pen [Tompkin et al. 2012]. However, we choose to focus on interfaces that can be available to a wide audience without the need for what today still is specialized hardware. We further reject alternatives such as editing the light field in its epipolar volume form [Kim et al. 2013a], or in a stereo interface, based on a pilot study.

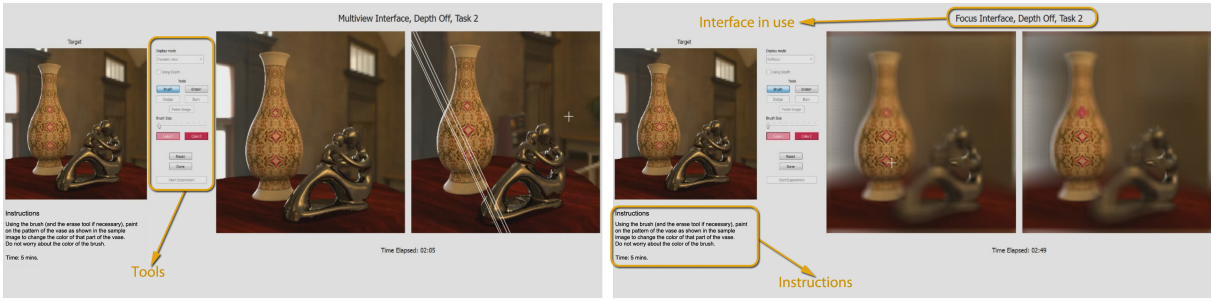
The goal of these works is to develop specific editing tools, or to propagate edits consistently across views, but none of them provide an analysis of the interface itself. In contrast, we focus on analyzing how the user interacts with the four-dimensional structure of a light field in order to specify common editing operations, with the final goal of learning insights to design a light field editing interface.

## 3 Overview

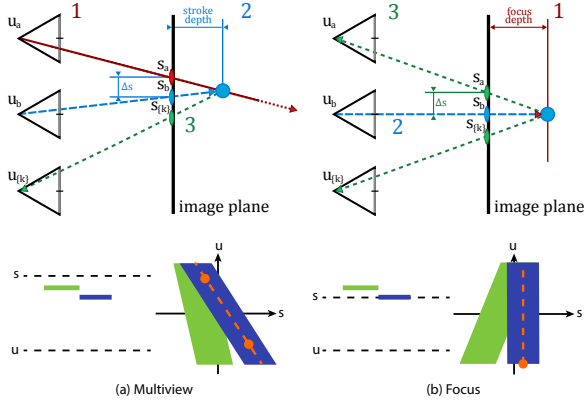
**Interfaces** In this paper, we first analyze two basic interfaces, based on interaction paradigms found in the light field editing literature. In one interface, which we call *multiview*, users navigate between the views of the light field to specify correspondences that locate their edits in space. This class of interfaces is the most common in existing work on non-automatic light field editing. We also investigate an alternative navigation interface, which we call *focus*, where the light field is shown with a synthetic shallow depth-of-field. This is based on Davis et al. [2012], which relies on defocus, computed following Isaksen et al. [2000], to guide the capture of unstructured light fields.

These interfaces rely on different *depth cues* to allow users to visualize the disparity of edits between the views of the light field.

<sup>1</sup><http://giga.cps.unizar.es/ajarabo/pubs/lfeiSIG14/>



**Figure 2:** User interfaces used in our tests. Left: multiview paradigm. In it, windows are inter-changeable, and show two different views of the light field; epipolar lines (white) mark the trajectory a currently active stroke will follow when moved in depth. Right: focus paradigm, in which the in-focus plane marks the depth at which strokes will be placed; the window on the right shows the “un-blended” strokes (please see main text for details).



**Figure 3:** Workflow when drawing a stroke in each paradigm. **Top row, left:** multiview. (1) The user first draws a stroke  $s_a$  in one view  $u_a$ . (2) Depth is then adjusted on a different view  $u_b$  by displacing the stroke along the epipolar line. (3) The stroke is then projected onto the other views  $u_{\{k\}}$  of the light field, yielding  $s_{\{k\}}$ . **Top row, right:** focus. (1) The user first specifies depth by placing the focal plane. (2) A stroke  $s_b$  is drawn on the central view  $u_b$  at the specified depth. (3) The stroke is projected onto the other views of the light field, yielding  $s_a$ , and  $s_{\{k\}}$ . **Bottom row:** A light field interpretation of the two paradigms. In the multiview paradigm, the user specifies two correspondences (orange dots), which provide the disparity of the 3D point (slant in the light field). In the focus paradigm, the user first places the point of interest in focus by shearing the light field. As a result, all images of the point are aligned and one scribble is enough to edit all the views.

While multiview uses *parallax* to convey disparity, *focus* relies on *depth-of-field blur* to visualize depth. Both cues are reminiscent of the way people experience 3D in the real world, and it is unclear if one cue is preferable to another, even though most existing light field editing tools have chosen so far to rely on the *multiview* approach. Furthermore, they do not require specialized viewing hardware [Wetzstein et al. 2012; Masia et al. 2013] and do not *intrinsically* necessitate additional data such as scene depth. We also test whether the use of depth information is helpful during editing.

**Implementation** The interfaces tested share the same screen layout, shown in Figure 2. On the left, there is a description and example image of the current task. Next to it, a control panel and two working windows, named Window 1 (W1) and Window 2 (W2). Interface manipulation is performed with a mouse or tablet. In the *multiview* interface (Figure 2, left), the user is presented with two

views of the light field, whose viewpoints are independently manipulated by panning and tilting. This allows the artist to view the light field from a different viewpoint than the one used for manipulation, which is a common workflow in 3D software packages. In this interface, the windows are equivalent. The workflow for placing a stroke in *multiview* is described in Figure 3, top left. In the *focus* interface (Figure 2, right), the scene is rendered with a wide synthetic aperture that blends all views of the light field [Isaksen et al. 2000]. Points that are in focus appear sharp because their images are aligned, while points that are out of focus appear blurry because of the disparity between their images. By construction, this alignment gives us the position of any in-focus point in all views of the light field. In this interface, the user cannot alter the viewpoint, but can adjust the depth of the focal plane of the scene, i.e. the relative disparity of the views [Vaish et al. 2004]. Additionally, with depth enabled, the views focus directly on the visible surface below the mouse. Here, W2 displays the edit without blending it with the views of the light field, to show its area of influence in other views and help determine occlusions. The workflow for placing a stroke in *focus* is described in Figure 3, top right.

Many editing tasks require users to pick locations in 3D space to, e.g., draw strokes. With depth information disabled, we place the user strokes on a plane *parallel* to the camera, so all points in the stroke share the same depth. Otherwise, we snap the strokes at the depth under the pointer. We found no effect due to depth quantization when using depth to edit. In both cases the strokes have zero-width in the depth domain. We refer the reader to the supplementary video for a practical demonstration of the workflow with each of them.

**Experiments** We perform two separate user studies. In our first study, we analyze the performance of the two paradigms described above (multiview and focus) when used exclusively, and considering their use with and without depth information. This yields a total of *four* interfaces. We evaluate those interfaces by making subjects perform several simple tasks. This first study allows us to gain an initial but formal understanding of the light field editing process by precisely characterizing the strengths and weaknesses of each paradigm in simple, controlled tasks.

In our second study, we analyze users’ workflow on more complex, but natural, editing tasks, informed by the analysis of the first study. In this case, we let users switch paradigms and activate or not the use of depth; we also include additional tools based on previous feedback. Furthermore, we use captured light fields exhibiting imprecise reconstructed depth.

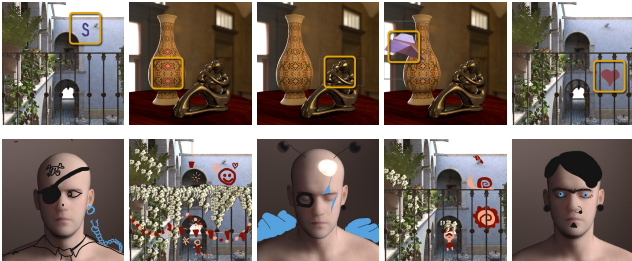
## 4 Experiment 1: Synthetic Scenarios

**Goal** We want to compare the two interfaces with respect to their effectiveness, efficiency and subjective preference. With *effectiveness* we refer to how well the intended task is accomplished, *efficiency* is related to the effort of obtaining a particular output, and *subjective preference* is based on qualitative data, i.e. user opinions on ease of use, learning curve, among others.

**Light fields** We use three different synthetic light fields, depicting different types of scenes (see Figure 4 and the supplementary material): a complex architectural scene (*San Miguel*), a still-life scene (*vase*), and a human head (*head*). These scenes have different depth, geometry and reflectance complexities. We use synthetic scenes to have precise depth information and be able to compute the error of the performed edits. We render the scenes with a light field camera implementation in the physically-based renderer PBRT [Pharr and Humphreys 2010]. We use  $17 \times 17$  views with a resolution of  $400 \times 400$ , in order to achieve real-time interactions at roughly 30 frames-per-second. We up-sample the rendered images to  $600 \times 600$  during display to facilitate more accurate placement of the strokes.

**Editing tools** The set of tools for this study consists of: *brush painting*, *erasing*, *dodging & burning* and *pasting* of pre-loaded billboards parallel to the camera plane. All these edits are directly controlled by strokes locations. We choose these operations since they are common in most image editing software, they are well-known to users, requiring little training, and they represent simple operations from which more complex edits can be performed.

**Tasks** We asked users to perform two types of tasks: *directed tasks*, where the user has specific instructions on what to edit; and *open tasks*, where the user is only given general guidelines. We refer the reader to the supplementary material for the specific instructions and example target images.



**Figure 4:** Top row: Target images given to users in the directed tasks of the first study (S1 to S5). The small highlighted areas have been added to this figure for visualization purposes and future reference. Bottom row: Example results of user edits for the open tasks, S6 and S7.

*Directed tasks* are performed for all four interface configurations. We use the central view of an edited light field as the target image. Users are not required to match the target image precisely, but rather to match the depth at which the strokes have to be positioned using the target image as a visual reference. Only one tool is available for each task, plus the eraser which is always available; the color brush is limited to one color, to avoid unnecessary distractions. Time to completion is limited to 5 minutes. These tasks, and the specific light field for each, are the following (see Figure 4):

**S1** Draw a simple object on a surface of constant depth (*San Miguel*)

**S2** Paint a simple pattern on a non-planar surface (*vase*)

**S3** Increase the brightness of the specular reflections on a curved surface (*vase*)

**S4** Place an object billboard within a certain depth range in free space (*vase*)

**S5** Draw on a partially occluded surface (*San Miguel*)

The five tasks have been chosen to cover a wide range of use cases: Tasks S1 and S2 are devised to test general editing of surfaces, while Task S3 deals with the particular case of specular highlights, which do not lie on the surface of the object. Task S4 investigates how to work in free space, while Task S5 tests how to best deal with occlusions.

After performing the directed tasks, subjects complete two *open tasks*, where real-world photos are given as a source of inspiration, and participants are free to use all the tools at will, plus two different colors for the brush. Time to completion is limited to 12 minutes. The tasks vary based on interface selection by the user:

**S6** The user is allowed to select whether to use depth information or not during editing. The task is done twice per subject, once with the *multiview* paradigm (toggling freely between using or not depth), and once with *focus* (also with or without depth). The task is carried out on the *head* light field.

**S7** The user is allowed to freely change between the four interfaces. The task is done on the *San Miguel* light field.

**Experimental procedure** Twenty paid subjects participated in the experiment (6 female, 14 male). All of them had previous knowledge on image editing, 3D modeling or 3D sculpting software, with either an artistic or technical background. Most of them (90%) had no previous knowledge of light fields, which were briefly introduced to them in the beginning.

Each subject used each interface, with and without depth, first for all directed tasks and then the open ones. We randomize the order of each interface configuration to mitigate learning effects and record all users actions and screen. Subjective preferences were collected with questionnaires filled after each interface and at the end of the experiment. The full experiment took around four hours per subject, including training and breaks.

### 4.1 Analysis

Throughout the experiment we collected both quantitative data on task errors and timings, qualitative data on performance and difficulty of both tasks and interfaces, and free-form comments on interface effectiveness. We report here the analysis and main findings and include additional data in the supplementary material. For brevity, we will refer to our interface configurations in the rest of the analysis as *M* (multiview without depth), *MD* (multiview with depth), *F* (focus without depth) and *FD* (focus with depth).

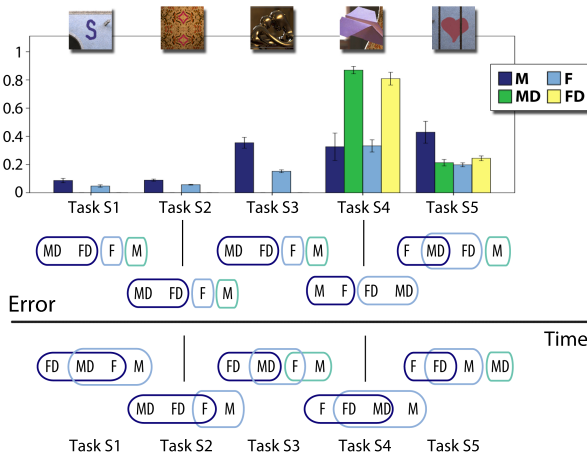
**Procedure** We use repeated measures ANOVA for the analysis of error, timings and ratings, to test whether the levels of a factor (e.g. interface is a factor; the types of interface are its levels) have influence on the observed data. It is a repeated measures scheme because we measure the same independent variables (e.g. error) using the same participants under the different conditions. Additionally, Kruskal-Wallis (a nonparametric extension of ANOVA) is used to analyze rankings, since they are an ordinal variable and typically cannot be assumed to follow a normal distribution. In all tests, we use a *p*-value of 0.05 to indicate significance. When sphericity is violated, according to Mauchly's test, we report significance values adjusted with the Greenhouse-Geisser correction [Cunningham and Wallraven 2011]. In all figures, error bars represent the standard

error of the mean. We perform outlier rejection on the measured error data, based on the interquartile difference, with a factor of 2.2 [Hoaglin and Iglewicz 1987]. This led to dropping one user in Tasks S2, S3 and S5, two users in Task S1 and three users in Task S4.

**Error in depth** To evaluate how well a user can specify locations in the light field, we measure the *error in depth* of the stroke. We choose this over measuring image-based differences since our tasks are not pure matching tasks. Specifically, for each view of the light field, we first compute the  $L_1$  distance between the depth of the stroke and the target depth, for each pixel of the stroke, and divide it by the number of pixels covered by such stroke. We then average across all views of the light field. Our experiments showed that  $L_1$  averaged across views approximates a normal distribution better than other metrics, which facilitates the subsequent analysis. Note that in Task S4 (positioning in free space) there is not a single fixed target depth, but a valid range between the vase and the sculpture. We thus compute error in depth with respect to the limits of such range, assigning a value of zero within it.

Figure 5 (top) shows the per interface mean error for each of the directed tasks (S1–S5). The error is highly dependent on the task, which accounts for 73% of the variance. When taking into account interfaces, the ANOVA yielded significant differences between interfaces for all tasks, as summarized in Table 1. The figure additionally illustrates significant differences between interfaces according to the pairwise comparisons tests.

This analysis suggests that while interfaces with depth information work well when manipulating surfaces without occlusions, not using depth information is actually more effective when occlusions are present, or when the editing task requires positioning in free space. This is due to the fact that the edits will snap to the underlying depth of the light field, which is not desirable in those particular cases. The supplementary material contains a video showing Task S5 (handling occlusions) being performed with all four interfaces, as well as Tasks S1 to S4 performed with different interfaces.



**Figure 5:** Top: Mean error per interface for each task and pairwise comparisons for the error in each task. Items in the same set are statistically indistinguishable. Bottom: Pairwise comparisons for the time to completion in each task.

**Time to completion** We illustrate statistically significant differences between interfaces in time to completion for each directed task (S1–S5) in Figure 5 (bottom). In general, users were able to complete the tasks in the allocated time with all interfaces. We

	$H$	$(df_1, df_2)$	$p$	$\eta^2$ (%)
<b>S1</b>	25.230	(1.517, 25.792)	0.000	59.7
<b>S2</b>	138.745	(1.305, 23.491)	0.000	88.5
<b>S3</b>	70.390	(1.612, 20.108)	0.000	79.6
<b>S4</b>	24.951	(1.861, 29.779)	0.000	60.9
<b>S5</b>	6.275	(1.264, 22.760)	0.015	25.8

**Table 1:** Results of the repeated measures ANOVA for the interface factor for the error in depth in each of the tasks.  $H$  is the test statistic,  $df_1$  and  $df_2$  the between-group and within-group degrees of freedom, respectively,  $p$  the associated significance and  $\eta^2$  is indicative of the proportion of variance of the data that the interface factor explains.

observe that having depth information leads to faster editing when painting on surfaces, as long as no occlusions are present;  $M$  tends to take longer than the rest, apparently being less intuitive for users, while  $F$  performs well in most situations, specially when dealing with occlusions or positioning in free space, which are the two most challenging scenarios in our tests.

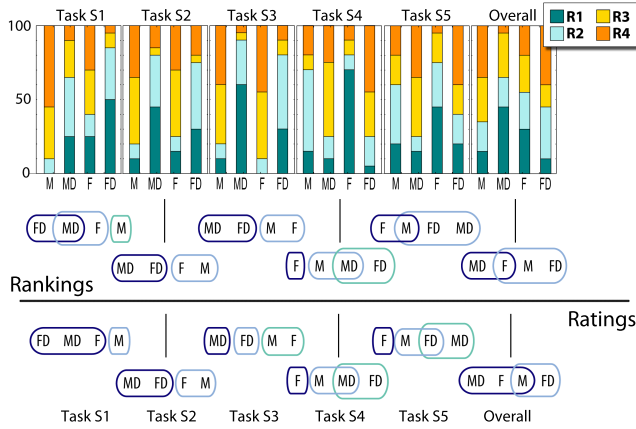
**Rankings and ratings** The final questionnaire contained eleven questions in which the users had to rank and rate the four interfaces. Five questions referred to the preference of interface for each of the five directed tasks, and one to the overall preference. The remaining five questions investigate preference in more general aspects, namely positioning in depth and on a plane ( $x$ - $y$ ), erasing, perceived accuracy of the interface and difficulty of use. For each ranking, we also compute the rank product per interface  $\Psi(\vartheta) = (\prod_i r_{\vartheta,i})^{1/m}$ , where  $r_{\vartheta,i}$  is the ranking received by interface  $\vartheta$  in a specific question and  $m$  the number of subjects [Rubinstein et al. 2010]. We use rank products to sort the interfaces when grouping them in statistically different groups (Figure 6, top).

Rankings for the different tasks (Figure 6, top) exhibit again a large between-task variability, in accordance with the error and time to completion. In Tasks S2 and S3  $MD$  and  $FD$  are ranked significantly higher ( $p \leq 0.035$ ) than no-depth interfaces ( $M$  and  $F$ ). The difference between the interfaces with depth ( $MD$  and  $FD$ ) is not significant ( $p = 0.160$ ). In Tasks S4 and S5 the trend is again reversed: there is a clear preference for interfaces without depth, and in particular for  $F$  ( $p \leq 0.035$  and  $p \leq 0.008$  in Tasks S4 and S5, respectively).

When asked about the overall ranking,  $MD$  ranks first, significantly higher than  $FD$  and  $M$  ( $p \leq 0.011$ ), but there is no significant difference between the rest. This is probably due to the large dependency on the task, shown by previous analyses. Despite the similarity of both interfaces in the rest of the tests, the users reported that the *multiview* paradigm allowed them a better visualization of the light field and the edits. Results of pairwise comparisons between interfaces for mean ratings for preference per task and overall preference are shown in Figure 6 (bottom). These ratings strongly correlate with rankings (Spearman’s  $\rho = 0.80$ ,  $p \simeq 0.000$ ). This is meaningful, indicating that users have clear opinions regarding the interfaces for the different tasks.

**Preferences in open tasks** During open tasks, when users can freely toggle the use of depth and interface, we record the time spent on each interface, and what actions are performed in each of them. Specifically, we track the time they spend drawing, erasing, changing the view point, and adjusting depth. When given complete freedom, the preferred workflow is to perform edits with the *focus* paradigm, use *multiview* to inspect the changes from different





**Figure 6:** Top: Rankings from final questionnaire for questions on preference for each task and overall preference ( $R_i$ : rank  $i$ ), and pairwise comparisons between interfaces for the rankings. Items in the same set are statistically indistinguishable. Bottom: Pairwise comparisons between interfaces for the ratings. There is a clear agreement between both ratings and rankings

viewpoints, and then go back to editing with *focus* again. This is supported by Table 9 in the supplementary material, which shows the total number of times users went from one interface to another during editing, and by the users’ feedback in the final interview, where they agree in their preference to edit with *focus* due to the feeling of control it provides. Finally, based on their answers to the after-task questionnaires, 85% of the users found the possibility of switching between interfaces during Task S7 helpful or very helpful (as indicated by a rating of 4 and 5 on a scale from 1 to 5). Together with the additional data, a video showing sample editing sessions by subjects for these tasks can be found in the supplementary material.

**Main findings and interface improvements** In this experiment we make the following observations, which lead to improvements in our editing system for the next experiment:

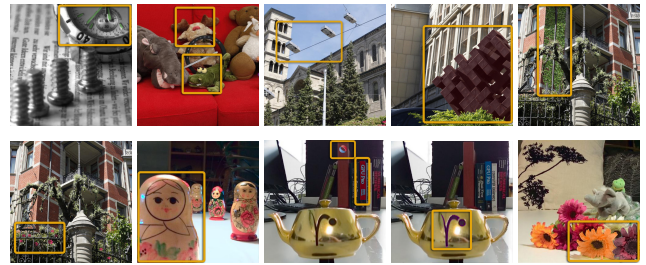
- (i) There is a high between-task variability with respect to the choice of interface in user preferences and error. Furthermore, the vast majority (85%) of users report that switching between paradigms and toggling depth on/off is very helpful.
  - We merge both interfaces into one in which users can switch between focus and multiview, and turn depth on or off.
- (ii) Handling occlusions is challenging for users, to the point that they prefer not using depth information if they are present.
  - We introduce a *Depth Selection* tool that creates a selection mask by clicking on a location to set the depth and specifying a depth threshold; this avoids unwillingly editing occluding parts when trying to edit occluded ones.
- (iii) Focus gives users better understanding of where they are positioned in the light field.
  - We introduce a *Visual Aid* mode that helps visualizing the active depth range (either the one selected or the one potentially selectable, if a threshold is set) by placing a semi-transparent checkerboard on the parts of the light field which are not within the active depth range.
- (iv) In interviews, users demanded selection tools based on color to be combined with depth-based ones.
  - We introduce a *Color Selection* tool that creates a selection mask based on picking a color and setting a threshold around it in color space. This threshold is independent of the depth

threshold, and thus each one yields a different selected area. This tool can be combined with the *Depth Selection* tool, in which case the selected area is the intersection of both.

## 5 Experiment 2: Real Scenarios

**Goal** In the second experiment, we investigate common edits on real, captured light fields with *reconstructed* depth. Our goal is to evaluate whether users can perform such edits in the presence of imprecise depth, and what interaction paradigms and tools they prefer. We incorporate the changes mentioned above and also include a *Hue Brush* for more versatility. Quantitative evaluation of the effect of imprecise depth could in principle be performed by introducing errors in depth procedurally; while this would allow a controlled error analysis, we found that the different existing depth-reconstruction algorithms produce very different types of artifacts, which cannot be modeled with simple functions. Thus, such quantitative analysis would hardly generalize to real scenarios. We choose instead to test whether such reconstructed depth maps from a varied set of real scenes are good enough to allow users to satisfactorily perform convincing edits.

**Light fields** We use eight captured light fields, depicting different types of scenes, shown in Figure 7. The first scene, *watch*, is captured with a Raytrix camera and its depth obtained with the algorithm by Wanner and Goldlücke [2012]. The next four—*couch*, *church*, *statue* and *mansion*—are captured with a camera gantry, and depth information is obtained with the algorithm by Kim and colleagues [2013a] (we use light fields and depth maps from their own database). The final three scenes—*matrioska*, *lab*, and *frog*—are captured with a Lytro light field camera, and the depth maps obtained with the Lytro SDK. These scenes have different depth, geometry and reflectance complexities, feature different baselines, from small (Lytro and Raytrix) to large parallax (Kim et al.) between views, and cover a wide range of capture setups and depth reconstruction methods found today. Light field angular resolutions are  $9 \times 9$  in all cases except for the horizontal-parallax-only ones [Kim et al. 2013a], whose resolution is  $17 \times 1$ . Spatial resolution is the same as in the first study.



**Figure 7:** Target images given to users in the tasks of our second study. The small highlighted areas have been added to this figure for visualization purposes, marking the area to edit. In reading order, Tasks R1 to R10. Different depth reconstruction methods are used: R1 is based on Wanner and Goldlücke [2012], R2 to R6 are based on Kim et al.[2013a], and R7 to R10 are from Lytro.

**Tasks** We asked users to perform a variety of *directed tasks* (see Figure 7) where textual instructions and an example target image (corresponding to the central view of the light field) are given to the user, who is asked to perform the given edit as accurately as possible. We chose tasks that span a wide range of use cases, in terms of editing operations, tools used, complexity of the geometry

of the light field, quality of the depth reconstruction, and type of edition (occlusion, free space, curved surface, etc):

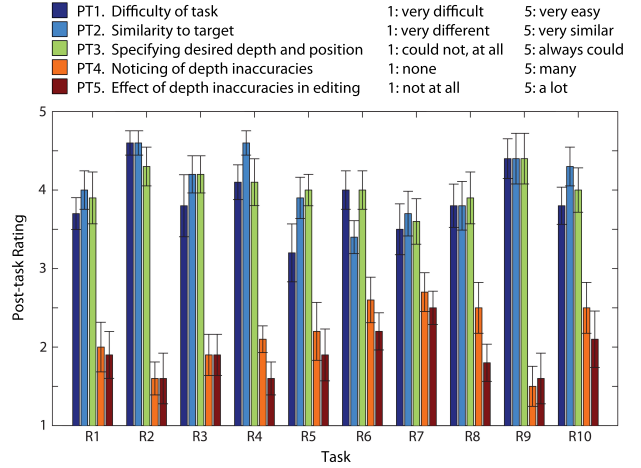
- R1** Colorization: Colorize the arrows and the marks on the inner circumference of the watch, so that users draw on a slanted surface (*watch*)
- R2** Correcting small parts: Change the color of the reindeer’s nose and of the eyes of the crocodile, so that users draw on curved surfaces (*couch*)
- R3** Cloning an object: Place another light, so that users work in free space (*church*)
- R4** Altering a material or color: Change the hue of the statue, so that users deal with selection and complex geometries (*statue*)
- R5** Texture brush: Add ivy to the wall, so that users deal with occlusions and slanted surfaces (*mansion*)
- R6** Importing an object (billboard-like): Add flowers to the bush, so that users deal with occlusions and work on areas of complex reconstruction (*mansion*)
- R7** Change luminances in the scene: Use dodge to brighten-up the matrioska in the foreground, so that users edit curved surfaces with a coarse depth reconstruction (*matrioska*)
- R8** Altering the color of an object: Change one of the books’ color from blue to red, to deal with color selection. Users are also asked to paste a SIGGRAPH logo on a book, another import-and-paste task (*lab*)
- R9** Improving small details: Change the RenderMan logo in the teapot to a purplish color to make it more salient (*lab*)
- R10** Harmonizing the colors in the scene: Change the colors of the two foremost flowers in the scene to match those of the rest of the flowers, so that users deal with selection of complex geometries (*frog*)

**Experimental procedure** Ten subjects participated in the experiment (four female, six male). All of them had previous knowledge on image editing, 3D modeling or 3D sculpting software, with either an artistic or technical background. Half of them had previous knowledge of light fields. The full experiment took around two hours per subject, including training and short breaks. The order of the tasks is randomized to compensate for possible learning effects. We recorded the screen during all the experiments. Each task had a time to completion limited to ten minutes. After finishing each task, users had to fill in a five-question questionnaire and at the end of the ten tasks the users completed a final questionnaire. All questionnaires can be found in the supplementary material. Each participant completed the study with a debriefing.

**Data** In this case, since we have no ground truth depth, we cannot measure error. Our objective measurements correspond to time of use of the different interaction paradigms (focus/multiview, with/without depth), the different tools, and the time to completion. Subjective ratings were collected in the questionnaires. For each task, we asked to rate, on a scale [1..5], the difficulty of the task, the similarity of the obtained result to the target, the ability to correctly position editions in space and depth, the perceived inaccuracies in the depth information, and whether those inaccuracies affected the editing process. Finally, we asked them to rate and rank each task according to their difficulty. In addition, we record all their actions.

## 5.1 Analysis

**Overall** In general, subjects were able to perform the tasks satisfactorily in the allotted time, as indicated by results for post-task questions (see Figure 8 for the results and questions): All subjects were satisfied with the result obtained (mean rating across subjects and tasks for similarity to the target is  $\mu_{PT2} = 4.09$ , 5 being “very similar” and 1 “very different”), and they did not find the editing



**Figure 8:** Ratings in post-task questionnaires in our second study, averaged across subjects, for each of the five questions. Bars represent the standard error of the mean. Note that in questions one to three higher is better, while in the two last questions lower is better.

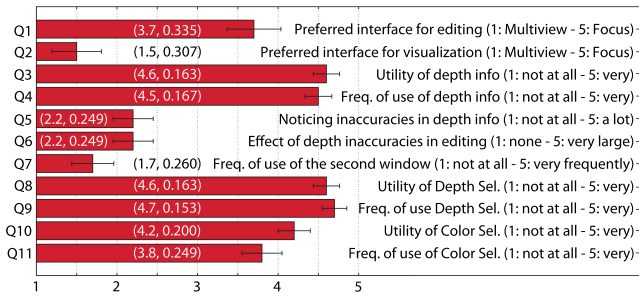
very difficult (mean rating across tasks and subjects for difficulty is  $\mu_{PT1} = 3.89$ , 5 being “very easy” and 1 being “very difficult”). Further, they state that they were “almost always”,  $\mu_{PT3} = 4.04$ , capable of placing the edits in the (3D) position they intended.

**Interaction paradigms** Figure 10 (left) shows the mean times, across subjects and per task, during which the focus and multiview paradigms were used, in percentage with respect to time to completion of the task. Differences are not significant for any task except for task R3 (*church*), where multiview is preferred, and task R7 (*matrioska*), in which focus is preferred. The first one is to be expected, since blur (or de-focus) does not offer a clear depth cue due to the nature and configuration of the scene. In R7, however, focus is more frequently used, the difference with respect to multiview being significant; we suspect this is due to the scene having a small number of very clearly separated depth planes.

In the final questionnaire, subjects rated and ranked multiview and focus, both in terms of visualization and of editing per se. For visualization, multiview is required by almost all subjects (see mean rating in Figure 9, Q2), in a manner consistent with results in our first study; also consistently, rank product for multiview for visualization is 1.07, meaning most of the time it is ranked first of the two. However, when it comes to editing, subjects are almost divided between both paradigms, meaning there is no preference for one over the other (rank product for multiview for editing is 1.62). When asked in the debriefing, many subjects would state that, given the tools at hand and the Visual Aid, there was not a significant difference between one paradigm and the other for editing.

**Depth information** With respect to the use of depth information we consider two aspects: First, whether inaccurate depth is good enough for editing light fields; and second, whether the use of depth information is preferred. When looking at post-task questions, perhaps surprisingly, for most of the tasks users noticed almost no inaccuracies in the depth information (mean rating is  $\mu_{PT4} = 2.16$ , with 5 being “many inaccuracies” and 1 being “none”). More importantly, the inaccuracies noticed did not significantly affect their editing (mean rating for the corresponding question is  $\mu_{PT5} = 1.91$ , 5 being “they affected editing a lot” and 1 being “they did not affect editing at all”). These averages are slightly

lower for the more sophisticated method of Kim et al. [2013a], which also makes use of higher resolution light fields, than for the Lytro light fields ( $\mu_{PT4, Kim} = 2.08$  vs.  $\mu_{PT4, Lytro} = 2.30$ , and  $\mu_{PT5, Kim} = 1.84$  vs.  $\mu_{PT5, Lytro} = 2.00$ ), although the difference is not significant. These conclusions on depth inaccuracies are confirmed by subjects’ responses in the final questionnaire (see Figure 9, Q5 and Q6). In terms of usefulness of depth information, users’ opinion has improved in the second study. When looking at the final questionnaire responses (Figure 9, Q3 and Q4), we see that they find it “useful” or “very useful” and use it “frequently” or “very frequently”. Timings confirm this, as shown in Figure 10 (right), which depicts the mean percentage over total time, across subjects and per task, during which depth information is activated or de-activated. This confirms the need for taking into account depth information but coupled with tools that exploit it adequately. It should be noted that while the task performed and the reconstruction quality might influence users preferences, it is not our goal to perform an accurate evaluation of depth reconstruction methods, but to derive usable guidelines for light field editing taking into account state-of-the-art capture and reconstruction methods.

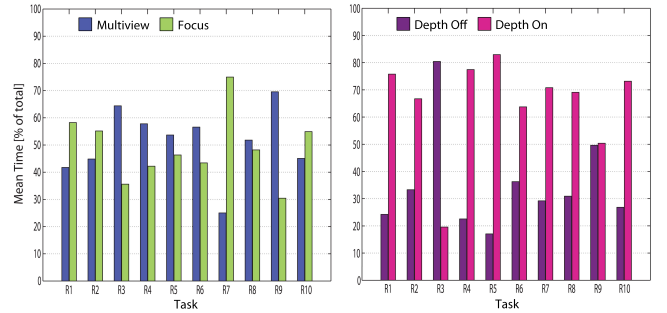


**Figure 9:** Mean ratings for questions in the final questionnaire of our second experiment, averaged across subjects. Bars represent the standard error of the mean (SEM). For each question, the exact mean  $\mu$  and SEM are shown in the form  $(\mu, SEM)$ . Note that not always higher is better.

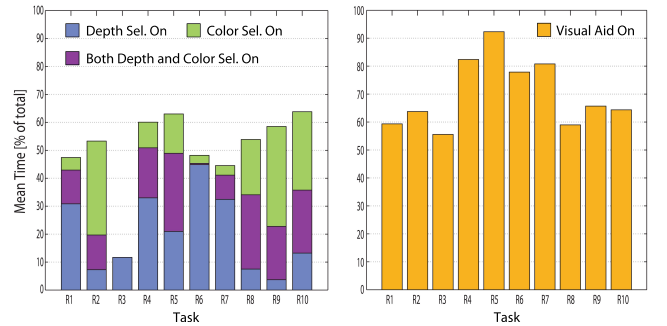
**Newly incorporated tools** When asked about the Depth Selection tool in the final questionnaire, responses by subjects were almost unanimous: the vast majority find it “useful” or “very useful”, and use it “frequently” or “very frequently” (see Figure 9, Q8 and Q9). Similar to the Depth Selection tool, the Color Selection tool is well received by subjects (Figure 9, Q10 and Q11; mean rating for utility of Color Selection is 4.20, 5 being “very useful” and 1 being “not useful at all”), although they state that the use of the Color Selection tool is less frequent than the Depth Selection tool ( $\mu_{FQ, DepthUse} = 4.70$  vs.  $\mu_{FQ, ColorUse} = 3.80$ ). The tasks are designed in such a way that, four of them can be done only with Color Selection (R2, R8, R9, R10), four of them require Depth Selection (R1, R5, R6, R7), and one both (R3 does not require either of them). Yet, the frequency of use of the Depth Selection tool is significantly higher, as stated above, meaning that subjects tend to favor the Depth Selection, or a combination of both even if only Color Selection could suffice.

We also look at the temporal data to evaluate the use of the newly incorporated features: Depth Selection, Color Selection, and Visual Aid functionalities. The times of use show that in most of the tasks, at least 50% of the time either the Depth Selection or the Color Selection tools (or both) were activated, as shown in Figure 11 (left). The frequency of use of the Visual Aid feature can be seen in Figure 11 (right); on average, the time Visual Aid is activated exceeds 50% of the time depth information is used in all tasks, confirming its utility.

**Conclusions of Experiment 2** We have seen that subjects are able to satisfactorily perform the variety of tasks with the interface and tools provided in the allotted time. Surprisingly, they very seldom notice inaccuracies in the depth information, and in any case these do not significantly affect their editing process. Further, for editing, there is not a significant difference between focus or multiview, whereas as expected, for visualization an approach that lets users see the different views individually is favored. We can conclude that for editing, provided the adequate tools and depth management functionalities (e.g. Visual Aid and Depth Selection), the differences between the two interaction paradigms are blurred out. Another relevant finding is the fact that users do leverage the extra dimensionality of a light field: Even in cases where selection based on color would suffice, they complement it with selection based on depth. Consequently, depth information, together with depth management functionalities, are favored and extensively used, easing previously complicated tasks such as occlusion handling.



**Figure 10:** Mean percentage across subjects, in realistic scenarios, of total time spent with multiview or focus (left) and with depth activated or deactivated (right), with respect to the total time to completion.



**Figure 11:** Left: Mean across subjects of time of activation of only Depth Selection, only Color Selection, and both tools simultaneously, in percentage over total time to completion of the task. Right: Mean times of use of the Visual Aid, in percentage over total time using depth information.

## 6 Discussion and Conclusions

Our findings support the potential for light field editing, currently underdeveloped and not on-par with image editing software. Overall, the most important general findings are: First, that users can perform editing tasks on light fields with our interface and tools; and second, that they do exploit the extra dimensionality of the light field. In fact, the real light fields in the second experiment incorporate reconstructed depth information, which subjects used constantly. Surprisingly, even though these reconstructed depth maps





**Figure 12:** Top row: Central view of the original light field. Bottom row: Sample edits performed by advanced users. From left to right: desaturating the background in motorbike2 (Lytro); modifying the couch, the hippo, and the crocodile’s body in couch (Kim et al. [2013a]); and changing his hair color and lighting her up in couple (Lytro).

have different degrees of inaccuracy (as shown, e.g., in Figure S.5 in the supplementary material), users reported that they barely noticed these inaccuracies, or that they did not significantly affect the editing process. We believe this is due to the nature of the inaccuracies and to the set of tools provided. For instance, a combination of Depth Selection and Color Selection typically overcomes the influence of noise. This is true even in low-parallax light fields, which do not allow for geometry recovery beyond a series of coarse discrete depth planes.

Moreover, occlusions, which are a very specific and relevant aspect in light field editing, were handled gracefully using the Depth Selection tool in the second experiment. However, having depth information does not mean using it all the time. In fact, users reported that toggling depth off offered them a *high degree of control* over the edits, although at the cost of longer editing times. This kind of versatility of the interface in the second study favored artistic exploration, as reported by subjects and shown in Figures 1 and 12 and in the supplementary video.

Another interesting finding is the fact that all users edited specular highlights as if they were a feature on the surface of the object (Task S3). This is physically inaccurate since they are actually detached from the reflecting surface [Templin et al. 2012], but nothing was reported during the interviews. This seems to confirm previous findings on the inability of the human visual system to correctly assess the physical accuracy of reflections and highlights [Fleming et al. 2003]. Finally, in terms of workflow and layout, the newly incorporated tools facilitated editing tasks to the point that the right window of the interface was rarely used (see Figure 9, Q7). Thus, for the advanced edits shown in the supplementary video and in Figure 1, we opted to add a button to open the right window only when needed.

**Future Work** Interfaces for light field editing remain largely unexplored. As such, there are many more opportunities for future studies. We have presented a number of diverse scenes and common tools which we hope help in future research. Additionally, our findings may be applied to editing RGB+D data, or even stereoscopic editing, which can be regarded as particular cases of light field editing.

As with any user study, our conclusions are only strictly valid for

the tested scenarios. Exploring other complex interaction procedures, such as selecting a volume in a light field, using edit propagation tools or applying perspective corrections when pasting objects, are interesting avenues of future work; these would require extending the local point-and-click edits used in our experiments to other global or non-local manipulations. Further, we test here a limited set of spatial and angular resolutions; it is worth looking into whether larger resolutions would affect the editing workflow and user performance. Additionally, while our results suggest that users’ editing is not significantly affected by depth errors, a quantitative analysis of the effect of imprecise depth remains an open question. Finally, we have explored the two most common paradigms found in literature (*multiview* and *focus*); devising and exploring new paradigms might lead to new interaction workflows for light field editing. It is also important to remark that we have used structured light fields [Levoy and Hanrahan 1996], such as those acquired with existing commercial plenoptic cameras (Lytro™, Raytrix™); unconstrained light fields [Gortler et al. 1996] are much less common, and out of the scope of this work.

All the material covered in this paper, including code for the interfaces and to record and analyze data; compiled, functioning versions of all the interfaces; and the raw data itself, are publicly available in the web. We hope that other researchers will be able to build on our work by extending the functionality of the interfaces, sharing some new test light fields or devising new tasks. We believe that our framework and principled evaluation methodology can help gain a better understanding of editing in the multidimensional space of light fields.

## Acknowledgements

We want to thank the reviewers for their insightful comments, the participants of the experiments (in particular Sara Lopez, Patrick Moosebrugger, Carlos Aliaga and Jose I. Echevarria), Chia-Kai Liang for his help with the Lytro SDK, the models in the captured light fields (Paz Hernando, Julio Marco, Laura Serra and Raul Buisan), and the members of the Graphics and Imaging Lab and the REVES Group for valuable discussions and feedback. We also want to thank the authors of [Kim et al. 2013a] and [Wanner and Goldluecke 2012], especially Sven Wanner, for sharing their light fields; Guillermo Leal Llaguno for the *San Miguel* scene; Emmanuelle Chapoulie and the AIM@SHAPE project for *vase*; and Infinite Realities for the *head* data. This research has been partially funded by the European Commission, 7th Framework Programme, through projects GOLEM, VERVE and TROPIC, the Spanish Ministry of Science and Technology, project TAMA, Intel Corporation and Adobe. Belen Masia has additionally been supported by an Nvidia Graduate Fellowship.

## References

- CHEN, B., OFEK, E., SHUM, H.-Y., AND LEVOY, M. 2005. Interactive deformation of light fields. In *Proc. of 13D’05*, 139–146.
- COLBERT, M., PATTANAIK, S., AND KŘIVÁNEK, J. 2006. BRDF-shop: Creating physically correct bidirectional reflectance distribution functions. *IEEE Comput. Graph. Appl.* 26, 1, 30–36.
- CUNNINGHAM, D., AND WALLRAVEN, C. 2011. *Experimental Design: From User Studies to Psychophysics*. AK Peters.
- DAVIS, A., LEVOY, M., AND DURAND, F. 2012. Unstructured light fields. *Comp. Graph. Forum* 31, 2pt1, 305–314.
- FLEMING, R. W., DROR, R. O., AND ADELSON, E. H. 2003. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision* 3, 5, 347–368.



- GOLDMAN, D. B., GONTERMAN, C., CURLESS, B., SALESIN, D., AND SEITZ, S. M. 2008. Video object annotation, navigation, and composition. In *Proc. of UIST'08*, 3–12.
- GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The Lumigraph. In *Proc. of SIGGRAPH'96*, 43–54.
- HACOHEN, Y., SHECHTMAN, E., GOLDMAN, D. B., AND LISCHINSKI, D. 2011. Non-rigid dense correspondence with applications for image enhancement. *ACM Trans. Graph.* 30, 4, 70:1–70:10.
- HAEBERLI, P. 1990. Paint by numbers: Abstract image representations. *SIGGRAPH Comput. Graph.* 24, 4, 207–214.
- HASINOFF, S., JOZWIAK, M., DURAND, F., AND FREEMAN, W. 2010. Search-and-replace editing for personal photo collections. In *Proc of ICCP'10*, 1–8.
- HOAGLIN, D. C., AND IGLEWICZ, B. 1987. Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association* 82, 400, 1147–1149.
- HORN, D. R., AND CHEN, B. 2007. Lightshop: interactive light field manipulation and rendering. In *Proc. of I3D'07*, 121–128.
- ISAKSEN, A., MCMILLAN, L., AND GORTLER, S. J. 2000. Dynamically reparameterized light fields. In *Proc. of SIGGRAPH'00*, 297–306.
- JARABO, A., MASIA, B., AND GUTIERREZ, D. 2011. Efficient propagation of light field edits. In *Proc. of SIACG'11*, 75–80.
- KERR, W. B., AND PELLACINI, F. 2009. Toward evaluating lighting design interface paradigms for novice users. *ACM Trans. Graph.* 28, 3, 26:1–26:9.
- KERR, W. B., AND PELLACINI, F. 2010. Toward evaluating material design interface paradigms for novice users. *ACM Trans. Graph.* 29, 4, 35:1–35:10.
- KIM, C., ZIMMER, H., PRITCH, Y., SORKINE-HORNUNG, A., AND GROSS, M. 2013. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph.* 32, 4, 73:1–73:12.
- KIM, Y., WINNEMÖLLER, H., AND LEE, S. 2013. WYSIWYG stereo painting. In *Proc. of I3D'13*, 169–176.
- LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proc. of SIGGRAPH'96*, 31–42.
- LO, W.-Y., VAN BAAR, J., KNAUS, C., ZWICKER, M., AND GROSS, M. H. 2010. Stereoscopic 3D copy & paste. *ACM Trans. Graph.* 29, 6, 147:1–147:10.
- LYTRO INC., 2013. The Lytro camera. <http://www.lytro.com>.
- MARTON, F., AGUS, M., GOBETTI, E., PINTORE, G., AND RODRIGUEZ, M. B. 2012. Natural exploration of 3D massive models on large-scale light field displays using the fox proximal navigation technique. *Computers & Graphics* 36, 8, 893–903.
- MASIA, B., WETZSTEIN, G., DIDYK, P., AND GUTIERREZ, D. 2013. A survey on computational displays: Pushing the boundaries of optics, computation and perception. *Computers & Graphics* 37, 1012–1038.
- OU, J., KARLÍK, O., KŘIVÁNEK, J., AND PELLACINI, F. 2012. Toward evaluating progressive rendering methods in appearance design tasks. *IEEE Comput. Graph. Appl.* 33.
- PHARR, M., AND HUMPHREYS, G. 2010. *Physically based rendering: From theory to implementation*. Morgan Kaufmann.
- PRICE, B. L., AND COHEN, S. 2011. Stereocut: Consistent interactive object selection in stereo image pairs. In *Proc. of ICCV'11*, 1148–1155.
- RUBINSTEIN, M., GUTIERREZ, D., SORKINE, O., AND SHAMIR, A. 2010. A comparative study of image retargeting. *ACM Trans. Graph.* 29, 5, 160:1–160:10.
- SANTOSA, S., CHEVALIER, F., BALAKRISHNAN, R., AND SINGH, K. 2013. Direct space-time trajectory control for visual media editing. In *Proc. of CHI'13*, 1149–1158.
- SEITZ, S. M., AND KUTULAKOS, K. N. 2002. Plenoptic image editing. *Int. Journal of Computer Vision* 48, 2, 115–129.
- SHUM, H.-Y., SUN, J., YAMAZAKI, S., LI, Y., AND TANG, C.-K. 2004. Pop-up light field: An interactive image-based modeling and rendering system. *ACM Trans. Graph.* 23, 2, 143–162.
- STAVRAKIS, E., AND GELAUTZ, M. 2004. Image-based stereoscopic painterly rendering. In *Proc. of EGSR'04*, 53–60.
- TEMPLIN, K., DIDYK, P., RITSCHEL, T., MYRSKOWSKI, K., AND SEIDEL, H.-P. 2012. Highlight microdisparity for improved gloss depiction. *ACM Trans. Graph.* 31, 4, 92:1–92:5.
- TOMPKIN, J., MUFF, S., JAKUSCHEVSKIJ, S., MCCANN, J., KAUTZ, J., ALEXA, M., AND MATUSIK, W. 2012. Interactive light field painting. In *ACM SIGGRAPH 2012 Emerging Technologies*, 12:1–12:1.
- VAISH, V., WILBURN, B., JOSHI, N., AND LEVOY, M. 2004. Using plane + parallax for calibrating dense camera arrays. In *Proc. of CVPR'04*, 2–9.
- VENKATARAMAN, K., LELESCU, D., DUPARRÉ, J., MCMAHON, A., MOLINA, G., CHATTERJEE, P., MULLIS, R., AND NAYAR, S. 2013. PiCam: An ultra-thin high performance monolithic camera array. *ACM Trans. Graph.* 32, 6, 166:1–166:13.
- VERTICAL HORIZON, 2013. Lightfield Iris. <http://www.verticalhorizon-software.com/LFI/index.html>.
- WANG, L., LIN, S., LEE, S., GUO, B., AND SHUM, H.-Y. 2005. Light field morphing using 2D features. *IEEE Trans. Vis. Comput. Graph.* 11, 1, 25–34.
- WANNER, S., AND GOLDLUECKE, B. 2012. Globally consistent depth labeling of 4d light fields. In *Proc. of CVPR'12*, 41–48.
- WETZSTEIN, G., LANMAN, D., GUTIERREZ, D., AND HIRSCH, M. 2012. Computational displays: combining optical fabrication, computational processing, and perceptual tricks to build the displays of the future. In *ACM SIGGRAPH 2012 Courses*.
- YÜCER, K., JACOBSON, A., HORNUNG, A., AND SORKINE, O. 2012. Transfuse image manipulation. *ACM Trans. Graph.* 31, 6, 176:1–176:9.
- ZHANG, Z., WANG, L., GUO, B., AND SHUM, H.-Y. 2002. Feature-based light field morphing. *ACM Trans. Graph.* 21, 3, 457–464.