

# SOFT COLOUR SEGMENTATION ON LIGHT FIELDS

*Pierre Matysiak, Mairéad Grogan, Weston Aenbacher, Aljosa Smolic*

V-SENSE Project, School of Computer Science and Statistics, Trinity College Dublin

## ABSTRACT

**In this work we explore methods for allowing advanced colour editing on light field images to be performed. This investigation is twofold. First we look at soft colour algorithms to decompose images into colour layers and the various ways it could be applied to light field data in order to ensure spatially consistent results. Then, with the purpose of colour editing in mind, we present an object-based layer separation method so that editing a layer does not wrongly affect specific objects. We further discuss the advantages and drawbacks that light field data present over regular single-view images for this purpose. Finally we present some editing results to show that our methods allow us to obtain visually appealing images that remain consistent across all light field views and minimise the colour artefacts inherent to layer decomposition methods.**

*Index Terms*— Light Fields, Soft Colour Segmentation, Object Segmentation, Colour Editing

## 1. INTRODUCTION

Light fields are a way of representing the light information contained in a finite volume of space [1]. They are an extension of traditional single view images and are typically represented by the 4D plenoptic function with two spatial and two angular dimensions. The most common methods of capturing real light field images are camera arrays [2], single cameras on a moving gantry [3], or consumer-grade plenoptic cameras [4].

While most of the research has been focused on applications such as rendering, depth estimation, or super-resolution, there is less work on image editing to take advantage of the higher dimensionality of light field data. In this work, we investigate the possibility of applying colour decomposition algorithms on light fields and detail the advantages and drawbacks of such methods. As far as we are aware, this is the first work looking at using these methods on light field images. Additionally we propose taking advantage of light fields to counteract one of the drawbacks of decomposition methods. Since the output of such algorithms is generally a number of colour layers based on a pre-computed palette, these layers can contain object and background information that have no semantic relation to each other. As a result any editing done on a single layer would affect all these objects, perhaps at the risk of causing unwanted artefacts, e.g. applying unnatural colours to human skin. In this work we present an automatic depth-based object-aware layer separation method to allow for easier colour editing.

---

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776.

## 2. RELATED WORK

### 2.1. Soft colour segmentation

Soft colour segmentation is a method of image decomposition which consists of separating the image into several semi-transparent layers containing pixel information close to a colour from a pre-computed palette. Initial works by Aksoy et al. used colour unmixing to satisfy minimisation functions, in which the colour palettes were computed by probability distributions obtained through pixel voting [5]. Tan et al. obtain a colour palette by simplifying a RGB convex hull of all observable colours [6]. The simplification can be adjusted to obtain a different number of colours in the palette.

Aksoy et al. further improved upon their previous techniques by implementing a more efficient voting scheme [7]. In order to obtain more consistent colour layers, Tan et al. use spatial coherence by extending their palette extraction method through a RGBXY convex hull [8]. A new technique by Koyama et al. decomposes images based on editing software blending modes [9]. The resulting layers may contain colour that do not appear in the original image and only using the proper blending modes for reconstruction to ensure a stable result.

One of the drawbacks of Tan et al.'s method is that by simplifying a RGB convex hull, it is quite possible to obtain palette colours that do not appear in the image. Feeling that obtaining a palette that would not be representative enough of the image can make editing work less intuitive for the user, Wang et al. use a similar method where a polyhedron is placed around the image colours in 3D space [10]. However they do not necessarily compress it to the convex hull and the palette they extract through an optimisation problem ends up being more accurate. Jeong et al. first sample pure colours, then build a hierarchical model by splitting each layer, and all the possible colours within it, into two layers where the colour variance is much smaller and the dominant colours are as distinct from each other as possible. [11]

### 2.2. Object Segmentation for Light Fields

Object segmentation on light field images has had a variety of techniques proposed. Mihara et al. propose a graph-cut method that works on sub-aperture views to find object edges and segments objects by enforcing a global consistency [12]. Hog et al. developed a method to exploit the redundancy of light fields to reduce the graph size of Markov random fields by using a ray bundle structure [13]. Their method is interactive and ensures stability and consistency across all light field views. Zhu et al. propose a super-pixel segmentation method which uses ray-tracing in the light field volume and accounts for the disparity between each super-pixel to provide a refocus-invariant segmentation [14]. A more recent method proposed by Khan et al. improves upon the

previous one by using a clustering step to enforce better consistency across the light field instead of simply propagating the results from the central view [15].

### 2.3. Light Field Editing

Jarabo et al. provided a comprehensive overview of different techniques used for light field editing, which included colour editing, inpainting, adding objects at various depth layers or drawing on partially occluded surfaces [16]. Le Pendu et al. propose a novel method for inpainting using low rank matrix completion which takes advantage of the redundancy of light field views [17]. Frigo et al. developed a method using epipolar planes to propagate edits, colour or inpainting, to the entire light field in a consistent manner. Zhang et al. created a method allowing object manipulation such as resizing or moving through depth planes [18]. They first decompose the central image in different depth layers, allowing the user to edit any of them. The patch-based method then reconstructs the image by transforming all possibly affected layers. Finally these edits are propagated from the centre view to the rest of the light field views. In this work we wish to provide another level of editing for artists and other light field users.

## 3. SOFT COLOUR SEGMENTATION

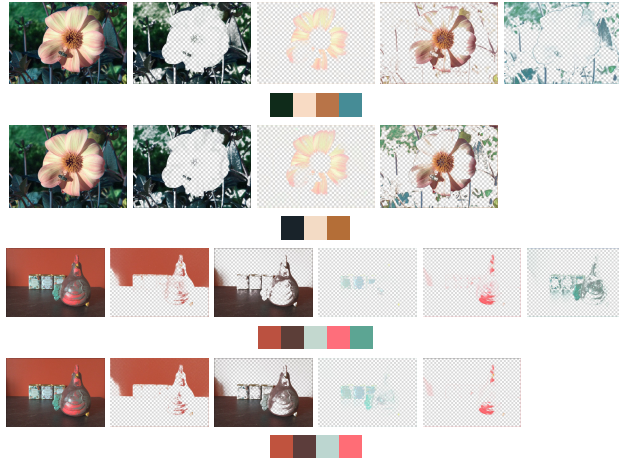
We base our investigations on our implementation<sup>1</sup> of the more recent method by Aksoy et al. [7]. This decision was motivated by the output of the method, providing layers of colours present in the image, which allows for more intuitive editing. As the method was developed for single images, we require a new strategy to apply it on light field data in the sub-aperture image representation, in order to enforce consistency in both the colour palette and the segmentation between all the views. In this section we describe some of the methods we used in achieving this goal.

### 3.1. Naive approach

For this initial approach, we apply the entire soft colour segmentation method to each individual sub-aperture view. That means a colour palette is computed for each image and the layer separation is done based on this palette. Because of slight variations in colour distribution caused by the disparity, the computed palette is not consistent across all views. This is visible in Fig. 1, where we show the layer decomposition of two consecutive views in a light field row. In many of the examples we worked on, the palette size varies by one or even two colours, leading to layers inconsistent in number and representation. Any kind of editing done with layers like these would result in erroneous results with flickering between views and artefacts clearly visible no matter which method is used to represent or view the light field.

### 3.2. Global approach

After seeing the output of our first method, we investigated ways of ensuring global consistency, especially regarding the computation of the colour palette. To this end, we first construct a mosaic image containing all the views from the light field, and perform an initial computation of a single global colour model,



**Fig. 1:** Results of our naive approach to perform soft colour segmentation on two consecutive views of images *bee\_2* and *chicken*, with associated colour palettes. The original view is on the left. In both cases, as with most images we studied, the number of layers can be different across views. In the top row of image *bee\_2* the last two layers end up mostly in a single layer in the decomposition of the bottom row, and in the top row of *chicken* the additional layer is composed mostly of pixels that end up in layers 2 and 3 in the decomposition on the bottom row. This results in major inconsistencies in these layers.

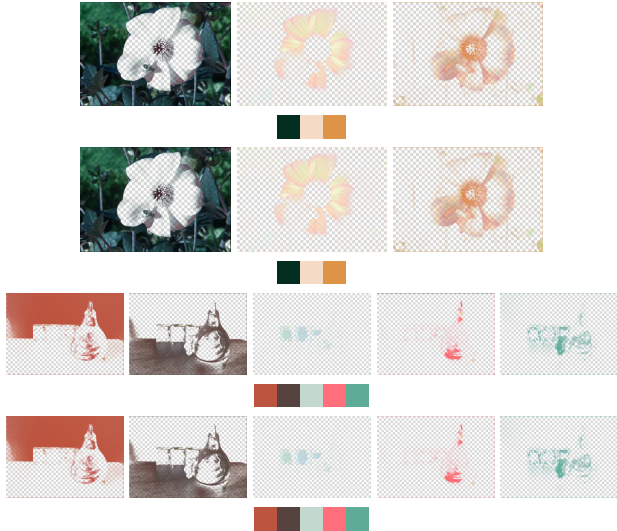
before applying the soft colour segmentation to each individual view using that single colour palette. The reasoning behind this was to ensure all colours from the light field would be represented in the colour palette, including some that might appear only in specific views due to occlusions. This method, as we can see in Fig. 2, produces more spatially consistent results with less variation between views, although some minor flickering can still be detected upon closer inspection (zoom in). The colour distribution in each layer is also generally more consistent and useful in comparison with the results of the naive method where we obtain a different number of layers. When using this global method we enforce consistency in both colour palette size and within the composition of each layer.

Using a global colour model additionally ensures all the views have the same number of layers and allows for situations where some objects could suffer from occlusion in some views but not others. In the case where the object is occluded, the layer would simply appear nearly empty. This is preferable to having the occluded objects appear in unrelated colour layers because of their low representation ratio.

### 3.3. Epipolar plane images

We additionally attempted to perform soft colour segmentation on epipolar plane images (EPI) instead of the sub-aperture views. We compute the colour palette using information from the whole light field rather than from a select view. Similar to the global method, the results are much more consistent globally, since the EPIs contain exactly the same colour information as the related views, even though the distribution is different. As this method offers no advantage while adding an extra computational step to generate the EPIs, we decide to use the method described in section 3.2 as our base for the rest of this paper.

<sup>1</sup>[https://github.com/V-Sense/soft\\_segmentation](https://github.com/V-Sense/soft_segmentation)



**Fig. 2:** Results of our global approach to perform soft colour segmentation on the same consecutive views as in Fig. 1. Compared to the naive approach here the number of layers is equal, and the colour distribution within layers is more consistent.

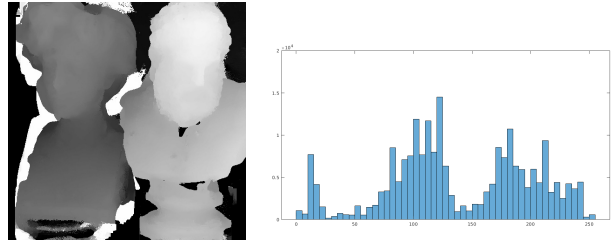
#### 4. OBJECT-BASED LAYER SEPARATION

In this section we present a method to separate objects in layers based on their depth. When analysing the results of traditional soft colour segmentation, it appears that many layers contain information from different objects which do not necessarily have a semantic relation. Editing the entire layer will affect all these objects and may cause some undesirable artefacts. Our intuition is to separate some of these layers into semantically relevant ones, which should make some editing tasks easier to perform.

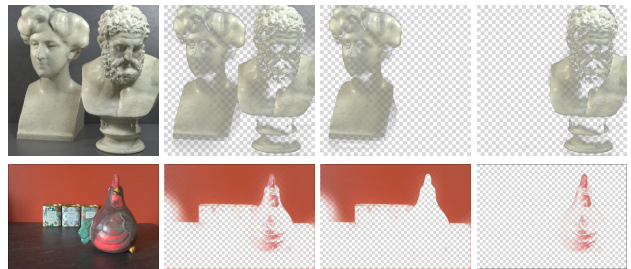
As light fields give us additional depth information, we chose to use it to separate objects that may appear in the same colour layers. We use the Spinning Parallelogram Operator method of Zhang et al. [19] to obtain a depth map of the light field. Using this information, we compute a histogram of the depth values of the image, as in Fig. 3. Once we obtain the histogram, we assume each peak represents a specific object, or at the least a depth layer containing mostly information from a single object. Values in and around the minima typically mark the separation. However to get the most precise split we use a gradient-based method described in the next paragraph. For instance in Fig. 3, it would appear there are three potential separate objects based on the histogram. Visual inspection, however, only shows two objects, while the values near zero belong to the background.

To rectify this as well as cases where there exists some overlap in depth between separate objects, we use a gradient-based metric which looks at continuity between the approximate centroids of the objects and the rest of the pixels. Starting from these centroids, we go outward and measure the difference in depth values until we detect sudden changes larger than a user-defined threshold typically chosen based on the range of depth values. Experimentation shows that setting this value as half the spread of depth values for the current object gives satisfying results. If that threshold is exceeded we determine the pixels belong to another object. Some visual results of layer splitting based on this can be observed in Fig. 4.

Even though the method splits the layer into separate, seman-



**Fig. 3:** Depth map of synthetic image *greek* (see Fig. 4). On the right, the histogram of depth values, cleaned to ignore the white values around the left statue and the black values belonging to the background.



**Fig. 4:** Examples of layer splitting using depth information on synthetic image *greek* and real image *chicken*. The original view is on the left.

tically coherent layers, some artefacts may occur. For instance in the image *greek* the layer containing the right statue still contains some background information. This is due to the depth estimation method incorrectly handling this boundary. Similarly in the image *chicken* some information from the metal boxes end up in the same layer as the background wall, instead of being in a separate layer, since the depth estimation puts them both on the same level.

#### 5. EXPERIMENTAL RESULTS

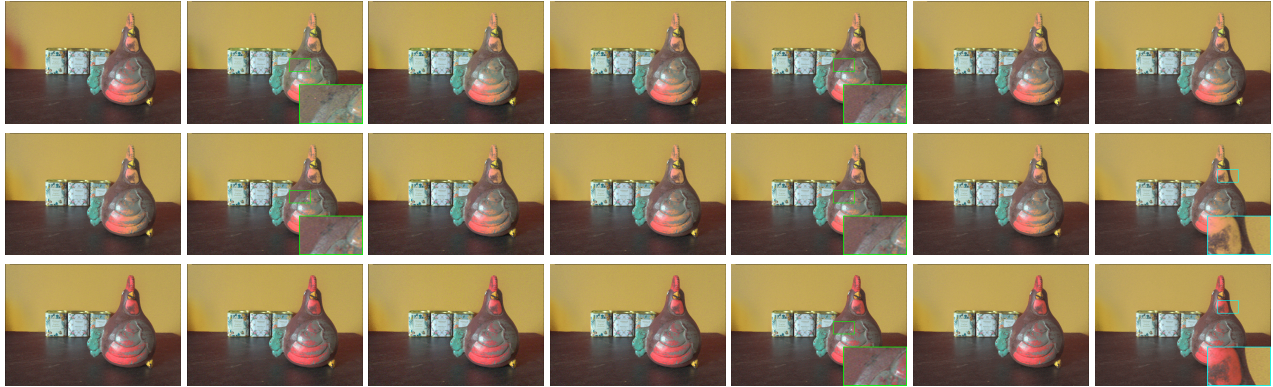
In this section we briefly detail the run time of the different parts of our work. We additionally present edited results taking advantage of our contributions. The images in this paper have been taken from the INRIA dataset [17], the HCI dataset [20] and some are our own Lytro images [21]. All of the Lytro data has been decoded, colour corrected and denoised using the methods of Matysiak et al. [21]. More information and additional results are available online<sup>2</sup>.

##### 5.1. Computation time

The computation time for a single light field view using the naive approach, i.e. computing the colour palette and doing the segmentation, takes on average 12 minutes for our C++ implementation of the method, which has not been optimised. This needs to be multiplied by the number of usable views in a light field, which can go up to 209 for Lytro images.

Computing a colour palette using our global method takes roughly 2.5 minutes on an image tile of size 5000x5000. While

<sup>2</sup><https://v-sense.scss.tcd.ie/research/soft-colour-segmentation-on-light-fields/>



**Fig. 5:** Editing results on the main red layers (the first one in Fig.1 or 2) of several sub-aperture images of a light field row; each column represents the same view. On the top row we use layers obtained using our naive approach, and inconsistencies can be observed between the views, on the chicken figurine or the background. In the middle row we use the layers from our global approach, the results are consistent across views but editing the wall colour results in changes affecting the red portions of the figurine. The last row shows editing on the global approach layers, done after separating the first red layer to isolate the wall from the figurine. Here the results are consistent across views, the figurine is unaltered, and we obtain the effect we were aiming for.

long, this step has to be done only once per light field and the results can be saved for later use. To put this in perspective, computing the individual palettes takes on average 14 minutes for a Lytro image (4.05 seconds  $\times$  209 views) and produces inconsistent results. However, the more computationally intensive part to perform the soft colour segmentation on each view still takes 12 minutes on average. An intended future work will look at ways to initialise the segmentation using results from already processed central views and propagate them toward the edges of the light field, in an attempt to reduce the time needed by the optimisation method of the soft segmentation to reach a solution.

Splitting the layers to contain only one object is done through a MATLAB script and is a much faster process, taking on average 2 seconds per view, regardless of the number of objects.

## 5.2. Layer editing

We present results of colour editing on specific colour layers obtained with both the naive and global methods, before and after splitting them, shown in Fig. 5. Here the edit was to change the background colour to increase the contrast with the foreground object being the centre of the image. Editing using the layers from our naive approach results, as expected, in visible inconsistencies between the views. When using the layers from our global approach, as the main red layer contains pixels from both the background and the chicken figurine, editing the background without splitting the layer results in unwanted alterations, as it changes the colour of the object as well. However, when the relevant layer is properly segmented to contain only the chicken figurine (or the background), edits are easier to perform and avoid unwanted side-effects. We obtain the intended effect to increase the colour contrast and enhance the focus on the object.

## 5.3. Failure cases

Unfortunately there are cases where our method for splitting a layer is not robust enough. One such case is when non-Lambertian objects are present, such as glass, see Fig. 6. Here the pink layer contains pixels from the pen, some items in the background, and reflection and refraction from all of these on



**Fig. 6:** Example of a failure case (real image *guinness*), where the pink reflection and refraction on the glass coming from different objects is difficult to separate from the pen using our method. For readability not all segmented layers are shown, as many objects in the image contain shades of pink.

the glass itself. Splitting the layer using our method puts both the pen and the glass in the same layer, and the background objects in their own layer. Editing either of those layers would result in inconsistencies between the background objects and their refraction in the glass. A future work will be to look at better methods to handle these cases.

## 6. CONCLUSION AND FUTURE WORK

We present a new method to perform soft colour segmentation on light field data. We have shown that our global approach is well-suited to create high quality colour layers consistent across all views of a light field. We also took advantage of light field data to further separate colour layers based on their content, using depth information, in order to allow for easier editing while reducing the amount of side effects.

We have also explained that these methods are unfortunately very costly in computing time and future investigations will look at ways to use disparity-adjusted layer decomposition results from neighbouring views as an initialisation step, in order to bring the minimisation algorithm closer to a solution. This should hopefully lead to a significant reduction in computing time when the method is applied over the whole light field. In future work we will also look at more robust and automatic methods for object-based layer separation and ways to extend this concept to light field videos.

## 7. REFERENCES

- [1] M. Levoy and P. Hanrahan, “Light field rendering”, in *Proc. SIGGRAPH*, 1996, pp. 31–42.
- [2] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy, “High performance imaging using large camera arrays”, *ACM Trans. Graph.*, vol. 24, no. 3, pp. 765–776, July 2005.
- [3] “The stanford light field archive”, <http://lightfield.stanford.edu/lfs.html>, accessed: 31-01-2020.
- [4] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, “Light Field Photography with a Hand-Held Plenoptic Camera”, Tech. Rep., Stanford University CSTR, Apr. 2005.
- [5] Yağız Aksoy, Tunç Ozan Aydın, Marc Pollefeys, and Aljoša Smolić, “Interactive high-quality green-screen keying via color unmixing”, *ACM Trans. Graph.*, vol. 35, no. 5, pp. 152:1–152:12, 2016.
- [6] Jianchao Tan, Jyh-Ming Lien, and Yotam Gingold, “Decomposing images into layers via RGB-space geometry”, *ACM Transactions on Graphics (TOG)*, vol. 36, no. 1, pp. 7:1–7:14, Nov. 2016.
- [7] Yağız Aksoy, Tunç Ozan Aydın, Aljoša Smolić, and Marc Pollefeys, “Unmixing-based soft color segmentation for image manipulation”, *ACM Trans. Graph.*, vol. 36, no. 4, Mar. 2017.
- [8] Jianchao Tan, Jose Echevarria, and Yotam Gingold, “Efficient palette-based decomposition and recoloring of images via rgbxy-space geometry”, *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 262:1–262:10, Dec. 2018.
- [9] Yuki Koyama and Masataka Goto, “Decomposing images into layers with advanced color blending”, *Computer Graphics Forum*, vol. 37, no. 7, pp. 397–407, 2018.
- [10] Yili Wang, Yifan Liu, and Kun Xu, “An improved geometric approach for palette-based image decomposition and recoloring”, *Computer Graphics Forum*, vol. 38, no. 7, pp. 11–22, 2019.
- [11] Taehong Jeong, Myunghyun Yang, and Hyun Joon Shin, “Succinct palette and color model generation and manipulation using hierarchical representation”, *Computer Graphics Forum*, vol. 38, no. 7, pp. 1–10, 2019.
- [12] H. Mihara, T. Funatomi, K. Tanaka, H. Kubo, Y. Mukaigawa, and H. Nagahara, “4d light field segmentation with spatial and angular consistencies”, in *2016 IEEE International Conference on Computational Photography (ICCP)*, May 2016, pp. 1–8.
- [13] Matthieu Hog, Neus Sabater, and Christine Guillemot, “Light field segmentation using a ray-based graph structure”, in *European Conference on Computer Vision*. Springer, 2016, pp. 35–50.
- [14] H. Zhu, Q. Zhang, Q. Wang, and H. Li, “4d light field superpixel and segmentation”, *IEEE Transactions on Image Processing*, vol. 29, pp. 85–99, 2020.
- [15] Lucas Kasser Henry Stone Min H. Kim James Tompkin Numair Khan, Qian Zhang, “View-consistent 4d light-field superpixel segmentation”, *International Conference on Computer Vision*, 2019.
- [16] Adrian Jarabo, Belen Masia, Adrien Bousseau, Fabio Pellacini, and Diego Gutierrez, “How do people edit light fields?”, *ACM Transactions on Graphics (SIGGRAPH Conference Proceedings)*, vol. 33, no. 4, 2014.
- [17] M. Le Pendu, X. Jiang, and C. Guillemot, “Light field inpainting propagation via low rank matrix completion”, *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1981–1993, April 2018.
- [18] F. Zhang, J. Wang, E. Shechtman, Z. Zhou, J. Shi, and S. Hu, “Plenopatch: Patch-based plenoptic image manipulation”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 5, pp. 1561–1573, May 2017.
- [19] Shuo Zhang, Hao Sheng, Chao Li, Jun Zhang, and Zhang Xiong, “Robust depth estimation for light field via spinning parallelogram operator”, *Computer Vision and Image Understanding*, vol. 145, pp. 148–159, 2016.
- [20] Katrin Honauer, Ole Johannsen, Daniel Kondermann, and Bastian Goldluecke, “A dataset and evaluation methodology for depth estimation on 4d light fields”, in *Asian Conference on Computer Vision*. Springer, 2016.
- [21] P. Matysiak, M. Grogan, M. Le Pendu, M. Alain, E. Zerman, and A. Smolic, “High quality light field extraction and post-processing for raw plenoptic data”, *IEEE Transactions on Image Processing*, pp. 1–1, 2020.