The Variational Structure of Disparity and Regularization of 4D Light Fields

Bastian Goldluecke Sven Wanner
Heidelberg Collaboratory for Image Processing

Abstract

Unlike traditional images which do not offer information for different directions of incident light, a light field is defined on ray space, and implicitly encodes scene geometry data in a rich structure which becomes visible on its epipolar plane images. In this work, we analyze regularization of light fields in variational frameworks and show that their variational structure is induced by disparity, which is in this context best understood as a vector field on epipolar plane image space. We derive differential constraints on this vector field to enable consistent disparity map regularization. Furthermore, we show how the disparity field is related to the regularization of more general vector-valued functions on the 4D ray space of the light field. This way, we derive an efficient variational framework with convex priors, which can serve as a fundament for a large class of inverse problems on ray space.

1. Introduction

In 2006, Marc Levoy, computer graphics professor at Stanford university and one of the leading experts in computational imaging, predicted in a survey article that "in 25 years, most consumer photographic cameras will be light field cameras" [9]. Whether or not he will ultimately be right, the first plenoptic cameras are now commercially available [12, 13], and several new technologies to capture light fields are under development. We believe that it is therefore time to thoroughly investigate the possibilities of light fields from the point of view of computer vision, in order to keep the theoretical research ahead and help to guide hardware development towards systems whose mode of recording is well suited for later analysis and processing. Because of their structure, light fields are particularly well suited to variational methods - indeed, we believe that variational methods might be key to getting the most out of this kind of data. The goal of this work is therefore to systematically develop theory and algorithms of a flexible and efficient variational framework which is built upon the concept of light fields instead of traditional 2D images.

Figure 1. *Regularization which leverages the variational structure of the light field leads to superior results in inverse problems like denoising and inpainting as well as multi-label segmentation.*

Light fields and computational imaging

For the purpose of this paper, the 4D light field of a scene can be understood as a collection of views of a scene with densely sampled view points, see figure 2. With the large amount of data available, it is not surprising that a wide array of applications has been developed. In image-based rendering, novel views of the scene are generated by means of a re-sampling of the light field [7, 10]. The selling point of the first consumer plenoptic camera is refocusing of the light field to different depth planes [12]. Camera arrays can in a sense be treated as a single plenoptic camera with a very wide aperture and allow synthetic aperture rendering, which allows to view through obstacles such as dense foliage [17].

Since plenoptic cameras trade off sensor resolution for capturing multiple views, it is not surprising that super-resolution techniques are a focus of research in light field analysis. They have been investigated using priors regarding statistics of natural images [2] as well as modified imaging hardware [11]. The problems of view interpolation and super-resolution have recently been solved jointly by minimizing a convex variational energy on the domain of the novel view [15]. This requires depth information, which can be recovered very reliably from plenoptic camera images. While the reconstruction can be based on traditional stereo matching techniques [2], recent works exploit the structure of light fields and epipolar plane images directly to estimate disparity [14]. Their underlying assumptions are adopted in this paper and will be explained next.

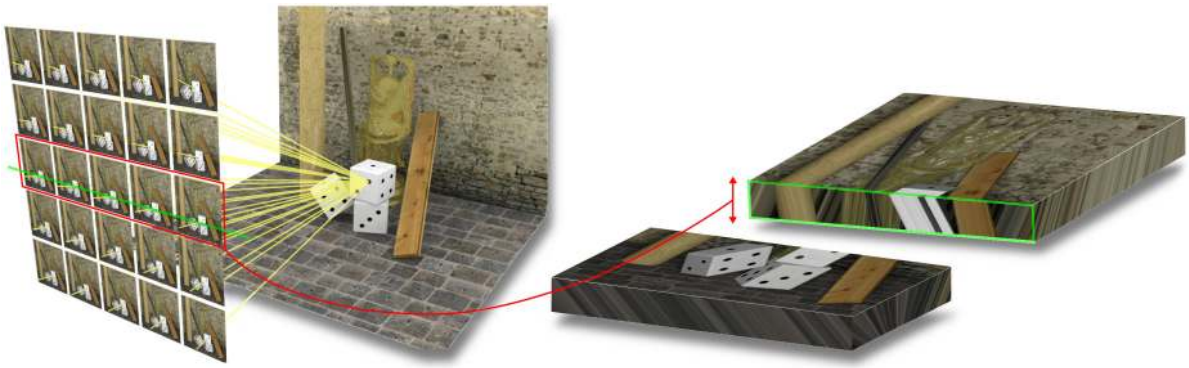


Figure 2. *One way to understand a 4D light field is as a collection of images of a scene, where the focal points of the cameras lie in a 2D plane. The rich structure becomes visible when one stacks all images along a line of view points on top of each other and considers a cut through this stack (denoted by the green border above). The 2D image in the plane of the cut is called an* epipolar plane image (EPI).

The structure of light fields

The fundamental difference between having a light field and just a simple collection of views of a scene available is that the view points in a light field lie much closer together and form a specific, usually simply rectangular pattern. In fact, it becomes possible to assume that the space of view points forms a continuous space, and for example, take derivatives of the light field with respect to the viewing direction. In order to fully exploit this new possibility, it is thus natural to work with variational methods, which allow to incorporate this differential structure into the problem formulation. With this paper, we follow the philosophy of the earliest works on the analysis of epipolar volumes [3]. They rely on the fact that 3D scene points project to lines in epipolar plane images, see figure 2, which can be more robustly detected than point correspondences [3, 5]. The proposed methods will further embrace these ideas, in that they allow to systematically leverage the differential light field structure in complex inverse problems.

Contributions

From a technical point of view, we will construct convex priors for light fields which are designed in a way that they preserve the epipolar plane image structure, and solutions satisfy constraints related to object depth and occlusion ordering. In this way, they enable the regularization of arbitrary vector-valued functions on ray space while respecting the light field geometry. Furthermore, we contribute an optimization framework for inverse problems on ray space which makes use of these priors, and show a number of examples, in particular light field denoising, inpainting, and in a related work [16], ray space labeling. As far as we are aware, this is the first time a systematic way to deal with inverse problems on ray space has been proposed, and we believe that it can serve as a solid foundation for the future development of light field analysis.

2. Disparity in a light field

In a rectified stereo pair, disparity is the coordinate difference of the two projections of a 3D scene point. In a light field, where a scene point is visible in many views, this would not make a very useful definition, since its value would depend on the pair of views chosen. Furthermore, in a light field, the space of view points is a continuous space, so it makes more sense to think of disparity as a *differential* quantity: the infinitesimal shift of the projection under an infinitesimal shift of the view point. Thus, disparity should be described as a derivative. In the following, we will systematically introduce what we call the *disparity field*.

Ray space

A 4D light field or Lumigraph is defined on a *ray space* $\mathcal{R}$, the set of rays passing through two planes Π and Ω in $\mathbb{R}^3$, where each ray can be uniquely identified by its two intersection points. For the sake of simplicity, we assume that both planes are parallel with distance $f > 0$, and equipped with 2D coordinate systems which are compatible in the sense that the base vectors are parallel and the origins lie on a line orthogonal to both planes.

The parametrization for ray space we choose is slightly different from the standard one for a Lumigraph [7], and inspired by [3]. A ray $R[x, y, s, t]$ is given by a point $(s, t) \in \Pi$ and $(x, y) \in \mathbb{R}^2$. The twist is that (x, y) is not a coordinate pair in Ω (as in the two-plane parametrization), but in the local coordinate system of the pinhole projection through (s, t) with image plane in Ω. This means that $R[s, t, 0, 0]$ is the ray which passes through the focal point (s, t) and the center of projection in the image plane, i.e. it is perpendicular to the two planes, see figure 3. In the following, coordinates (x, y) are always relative to a base point (s, t). We assume that the coordinate system of the pinhole view is chosen such that x is aligned with s and y is aligned with t, respectively.

Light fields and epipolar plane images

A light field L can now simply be defined as a function on ray space, either scalar or vector-valued for gray scale or color, respectively. Of particular interest are the images which emerge when ray space is restricted to a 2D plane. If we fix for example the two coordinates (y^*, t^*), the restriction L_{y^*,t^*} is the map

$$L_{y^*,t^*} : (x, s) \mapsto L(x, y^*, s, t^*), \tag{1}$$

other restrictions are defined in a similar way. Note that L_{s^*,t^*} is the image of the pinhole view with center of projection (s^*, t^*). The images L_{y^*,t^*} and L_{x^*,s^*} are called *epipolar plane images*. They can be interpreted as horizontal or vertical cuts through a horizontal or vertical stack of the views in the light field, see figure 2, and have a rich structure which looks like it consists mainly of straight lines. The slope of the lines is linked to disparity, as we will explore now.

Disparity of a 3D point

Let $P \in \mathbb{R}^3$ be a scene point. The pinhole projections of P onto Ω under two different centers of projection (s_1, t) and (s_2, t) will have different x_1 and x_2 coordinates. We are interested in the rate of change of x as a function of s. From similar triangles and the intercept theorem in figure 3,

$$\frac{dx}{ds}(s_1) = \lim_{\Delta s \to 0} \frac{\left(x_1 + \frac{f}{Z}\Delta s\right) - x_1}{\Delta s} = \frac{f}{Z}, \tag{2}$$

where Z is the distance of P to the plane Π, which we call the depth of P. The rate of change $\rho(P) := \frac{f}{Z}$ is independent of the choice of s_1, we call it the *disparity of* P. It is easy to see that it is numerically the same as $\frac{dy(P)}{dt}(t)$ for all t. Intuitively, the above shows that a scene point P is projected onto a line with slope $\rho(P)$ onto the (x, s)- and (y, t)-planes. This is the reason for the linear structures we observe in the epipolar images. Note that (2) shows that disparity in our sense is indeed a derivative.

3. Disparity maps and their constraints

For stereo pairs, it is customary to compute disparity maps, i.e. assign a shift to each pixel. In light fields, we need to assign a disparity to each ray. As in the stereo case, this only makes sense if the surfaces in the scene are opaque, as we will assume in the following.

In the case of opacity, to each ray $R[x, y, s, t]$ can be assigned a closest point $P(x, y, s, t)$ where the ray intersects the scene. We define the disparity map ρ as a function on ray space via

$$\rho(x, y, s, t) := \rho(P(x, y, s, t)). \tag{3}$$

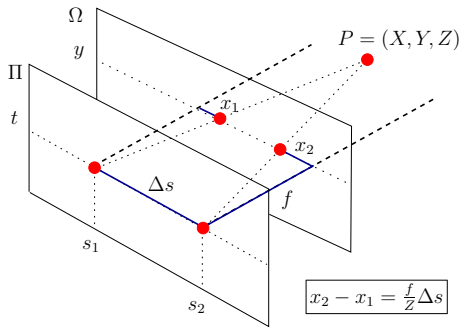


Figure 3. *Light field parametrization.* Each camera location (s, t) in the view point plane Π yields a different pinhole view of the scene. The two thick dashed black lines are orthogonal to both planes, and their intersection with the plane Ω marks the origins of the two different (x, y)-coordinate systems for the views (s_1, t) and (s_2, t), respectively.

Unlike the definition of disparity for a single (virtual) 3D point, the disparity map depends on the scene geometry. As we will see later, disparity and regularization of maps on ray space are intimately linked to each other. In particular, this shows that 3D scene reconstruction from light field data is actually a pre-requisite to correctly deal with inverse problems on ray space.

Like any map on ray space, one can restrict the disparity map to epipolar plane images. We will now investigate the constraints on this map, i.e. the question of which maps are valid disparity maps. A global constraint can be derived from occlusion ordering of scene surfaces. For this, we fix an EPI with coordinates (y^*, t^*). In the remainder of this section, all 4D fields are considered to already be restricted to this EPI to not clutter notation. Let (x_0, s_0) be a fixed point on this EPI. From its disparity $\rho_0 = \rho(x_0, s_0)$, one can compute the scene point P which projects onto (x_0, s_0) by means of (2). This point P also projects onto the line $(x_0 + \rho\sigma, s_0 + \sigma), \sigma \in \mathbb{R}$ on the EPI. Since it occludes everything behind it, this means that at no point on this line, the disparity can be smaller than ρ_0. In particular,

$$\rho(x_0 + \rho(x_0, s_0)\sigma, s_0 + \sigma) \geq \rho(x_0, s_0) \text{ for all } \sigma \in \mathbb{R}. \tag{4}$$

Because in this way, the disparity in a point restricts the disparity at arbitrarily far away locations, it is prohibitively expensive to enforce this constraint globally. However, from (4) one can immediately conclude a weaker local constraint.

Proposition. A valid disparity map satisfies the local constraints

$$\nabla_{\pm \boldsymbol{d}}\, \rho(x_0, s_0) \geq 0 \text{ for all } (x_0, s_0), \tag{5}$$

where the *disparity vector field* $\boldsymbol{d}$ is defined as a unit vector field in direction $[\rho \quad 1]^T$, and $\nabla_{\pm \mathbf{d}}$ are the directional

derivatives in direction $+\boldsymbol{d}$ or $-\boldsymbol{d}$, respectively. Note that since ρ is not necessarily differentiable, its directional derivatives in opposite directions might be different.

Local vs. global constraints

Unfortunately, the local constraint (5) does not imply the global one (4), a counterexample is shown in figure 4. However, it implies consistency along connected line segments where visibility remains unchanged. Thus, it is sufficient to enforce the local constraints if a point is not temporally occluded and then unoccluded again when moving through all the view points. In particular, if the total baseline is small, like e.g. in plenoptic cameras, this case is unlikely.

4. Regularization on ray space

Scene point labeling vs. labeling of rays

When we compute a disparity map, we label rays with a quantity which is ultimately a property of scene points. The above constraints will ensure that the labeling is consistent, in that all projections of the same point are labeled with the same disparity value. In image processing applications, many more labelings are of interest. For a start, a light field itself is an assignment of a color to each ray. For Lambertian surfaces, this means that all rays which correspond to the same scene point must be assigned the same color for the labeling to be consistent. This restriction holds for all properties which are independent of viewing direction, like surface material, texture, object class labels and others.

The question is therefore how we can enforce this consistency for vector-valued maps $\boldsymbol{U} : \mathcal{R} \to \mathbb{R}^n$ on ray space. We assume that $\boldsymbol{U}$ encodes a property of a scene surface which is independent of viewing direction, all scene surfaces are opaque and the disparity map ρ on ray space is known.

Constraints and EPI regularization

The disparity vector field $\boldsymbol{d}$ can be interpreted in a slightly different way as a transport field. At each point of an EPI, it denotes in which direction the projection of the corresponding scene point moves under variations of the view point. In particular, the function $\boldsymbol{U}$ should be constant in the direction of $\boldsymbol{d}$, except at disparity discontinuities.

We encourage this form of smoothness by regularizing with an anisotropic total variation on the epipolar plane images. As in the previous section, we fix an epipolar plane image with coordinates (y^*, t^*), and define for the restriction $\boldsymbol{U}_{y^*,t^*}$ the regularizer

$$\int_\Omega \sqrt{\nabla u^T D_\rho \nabla u}\, \mathrm{d}x \tag{6}$$

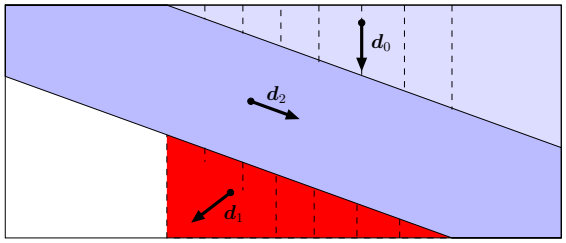


Figure 4. *Local constraints on the EPIs are in general not sufficient to enforce global consistency*. Although the transitions from $\boldsymbol{d}_0$ to $\boldsymbol{d}_2$ and $\boldsymbol{d}_1$ to $\boldsymbol{d}_2$ are both valid, occlusion constraints are violated in the red area since $\boldsymbol{d}_0$ corresponds to closer points than $\boldsymbol{d}_1$. However, for all scene points which are only visible on a connected line on the EPI, local constraints are equivalent to the global ones.

which encourages smoothing in the correct direction by means of the metric tensor [8]

$$D_\rho := g^2 \left(\alpha(I - \boldsymbol{d}\boldsymbol{d}^T) + \frac{3-\alpha}{2} \boldsymbol{d}\boldsymbol{d}^T \right). \tag{7}$$

A expression similar to (6) defines the regularizer $J_\rho(\boldsymbol{U}_{x^*,s^*})$ for the horizontal slices. The additional weight function g is optional, and can for example be used to decrease the penalty at disparity discontinuities. In our experiments, we leave it at $g = 1$. The constant α controls the degree of anisotropy, small values imply that smoothing is focused into the direction of the disparity field $\boldsymbol{d}$. Since the ray space regularizer defined later already explicitly includes regularization in the spatial domain, we set $\alpha = 0$ in all experiments. For more details and properties of the anisotropic total variation we refer to [8], similar regularizers have also been discussed in [18]. The work [14] presented a consistent disparity regularization framework on EPIs based on labeling, but it requires discretization of disparity values and is prohibitively slow.

Regularization of vector-valued functions

The final regularizer $J_{\lambda\mu}(\boldsymbol{U})$ for a vector field $\boldsymbol{U} : \mathcal{R} \to \mathbb{R}^n$ on ray space can now be written as the sum of contributions for the regularizers on all epipolar plane images as well as all the views,

$$\begin{aligned} J_{\lambda\mu}(\boldsymbol{U}) &= \mu J_{xs}(\boldsymbol{U}) + \mu J_{yt}(\boldsymbol{U}) + \lambda J_{st}(\boldsymbol{U}) \\ \text{with } J_{xs}(\boldsymbol{U}) &= \int J_\rho(\boldsymbol{U}_{x^*,s^*})\, \mathrm{d}(x^*, s^*) \\ J_{yt}(\boldsymbol{U}) &= \int J_\rho(\boldsymbol{U}_{y^*,t^*})\, \mathrm{d}(y^*, t^*) \\ J_{st}(\boldsymbol{U}) &= \int J_V(\boldsymbol{U}_{s^*,t^*})\, \mathrm{d}(s^*, t^*), \end{aligned} \tag{8}$$

where $\lambda > 0$ and $\mu > 0$ are user-defined constants which adjust the amount of smoothing on the separate views and

epipolar plane images, respectively. The spatial regularizer J_V for the individual views can be chosen arbitrarily, but should be convex and closed. A reasonable choice is for example a vectorial total variation [6].

The ray space regularizer defined in (8) gives a means to tackle any inverse problem on ray space. In the following, we will show how to use it in optimization problems and then give several examples for possible applications.

5. Optimization

We will consider general convex minimization problems of the form

$$\operatorname*{argmin}_{\boldsymbol{U}\in\mathcal{C}} \left\{ J_{\lambda\mu}(\boldsymbol{U}) + F(\boldsymbol{U}) \right\}, \tag{9}$$

where $\mathcal{C}$ is a convex set of vector-valued functions $\boldsymbol{U} : \mathcal{R} \to \mathbb{R}^n$ on ray space, and the functional F is convex and differentiable with Lipschitz-continuous derivative, whose Lipschitz constant is $L > 0$. Note that the regularizer $J_{\lambda\mu}$ for vectorial functions on ray space, given by (8), is convex and closed as the sum of convex and closed functionals, but not differentiable. Thus, one has to employ methods from convex optimization in order to minimize (9).

When choosing an algorithm, the main limitation which arises is that at reasonable resolutions, each field on ray space takes up a lot of memory. It is therefore not feasible to employ a primal-dual scheme on the complete functional, since storage of the required dual variables would exceed the memory of current GPUs, which are the target architecture. We therefore employ a subgradient descent scheme which is similar to iterative shrinkage and thresholding [1].

Note that according to (8), the complete regularizer which is defined on 4D space decomposes into a sum of regularizers on 2D images - the epipolar plane images and individual views, respectively. Each of these 2D regularizers is convex and closed, and corresponds to a comparatively small problem. Now for any convex functional f, an implicit subgradient descent step can be computed with the *proximity operator*, which amounts to solving the L^2 model

$$\operatorname{prox}_{\tau f}(u) = \operatorname*{argmin}_{v} \left\{ \frac{\|v-u\|_2^2}{2\tau} + f(v) \right\}. \tag{10}$$

The above computes a subgradient descent step for f starting from u with step size τ, and is guaranteed to get closer to the minimizer of f.

For optimization of the model (9), we iterate descent of the regularizer components and the data term as detailed in figure 5. Each application of a proximity operator amounts to solving a standard L^2-denoising problem in 2D with the corresponding regularizer component, which can be performed efficiently with e.g. primal-dual schemes [4]. Note that all subproblems for each epipolar plane image and view are completely independent. Typically, the computation time until convergence is around 10 to 30 seconds, depending on the number of dimensions of $\boldsymbol{U}$ and type of data term chosen, on an nVidia GTX 580 GPU.

To solve the **inverse problem** (9) **on ray space**, we initialize the unknown vector-valued function with $\boldsymbol{U} = 0$ and iterate the following steps:

- data term descent: $\boldsymbol{U} \leftarrow \boldsymbol{U} - \frac{1}{L}\nabla F(\boldsymbol{U})$,
- EPI regularizer descent:

  $\boldsymbol{U}_{x^*,s^*} \leftarrow \operatorname{prox}_{L^{-1}\mu J_\rho}(\boldsymbol{U}_{x^*,s^*})$ for all (x^*, s^*),

  $\boldsymbol{U}_{y^*,t^*} \leftarrow \operatorname{prox}_{L^{-1}\mu J_\rho}(\boldsymbol{U}_{y^*,t^*})$ for all (y^*, t^*),
- spatial regularizer descent:

  $\boldsymbol{U}_{s^*,t^*} \leftarrow \operatorname{prox}_{L^{-1}\lambda J_V}(\boldsymbol{U}_{s^*,t^*})$ for all (s^*, t^*).

Figure 5. *Algorithm for the general inverse problem* (9).

6. Applications

The model (9) is sufficiently general to enable a wide range of interesting applications. In the following, we show how to perform a number of archetypical tasks in image analysis. We have also investigated ray space multi-labeling as a more involved application, which we explore in detail in a related paper [16]. Source code and data sets will be published online on our web page[1].

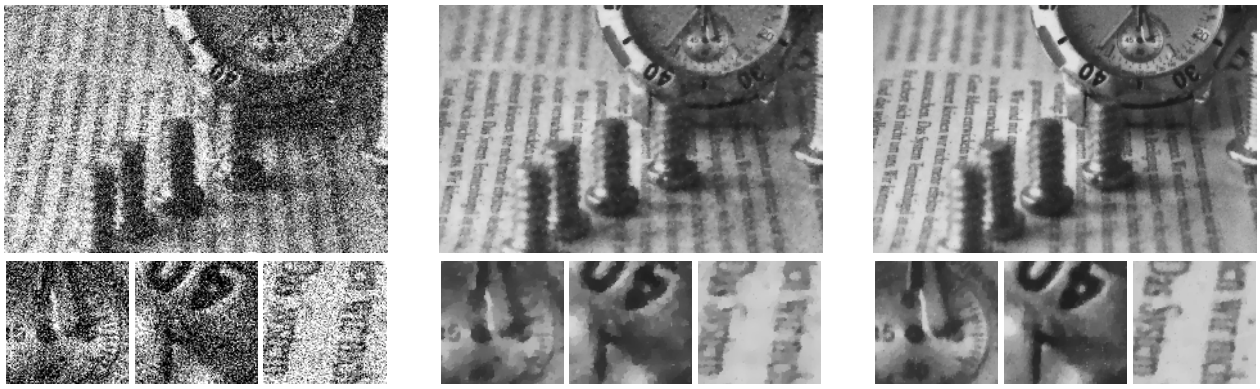
Light field denoising

We first show how to perform denoising of light field data. This is important in the context of plenoptic cameras, which aim for a very large sensor resolution and are thus particularly prone to noise. For this, let $\boldsymbol{F}$ be a vector-valued function on ray space which is degraded with Gaussian noise of standard deviation σ independently for each ray. The corresponding MAP estimate is of the form (9), where $J_{\lambda\mu}$ acts as a prior and the data term is given by

$$F(\boldsymbol{U}) = \frac{1}{2\sigma^2} \|\boldsymbol{U} - \boldsymbol{F}\|_2^2. \tag{11}$$

Results can be observed in figure 6, see also figure 1 for further close-ups. Denoising on ray space leads to significantly better quality than denoising of single views. Note that the L^2-denoising model above is equivalent to the proximity operator for $J_{\lambda\mu}$, so having an efficient solver available can be useful for more complex applications.

[1] http://lightfield-analysis.net



(a) Gaussian noise $\sigma = 0.2$, PSNR 11.92 (b) Single view denoising, PSNR 23.44 (c) Light field denoising, PSNR 25.27

(d) Gaussian noise $\sigma = 0.2$, PSNR 14.66 (e) Single view denoising, PSNR 22.61 (f) Light field denoising, PSNR 24.46

Figure 6. *L^2-denoising schemes which respect the light field structure are superior to single view denoising*. For the spatial regularizer, we use the vectorial total variation based on the operator norm of the Jacobian [6]. Top: light field recorded with a Raytrix plenoptic camera, bottom: synthetic light field rendered with Blender. Figures (b,e) show the results for optimal parameter choice using spatial regularization only, and figures (c,f) the optimal results for a denoising scheme on ray space, as described in section 6. Optimal parameters were determined using brute force search, the disparity map was estimated from the input light field using the algorithm in [14].

Light field inpainting

As a second example, we discuss inpainting on ray space. Let $\Gamma \subset \mathcal{R}$ be a region in ray space where the input light field $\boldsymbol{F}$ is unknown. The goal is to recover a function $\boldsymbol{U}$ which restores the missing values. For this, we find

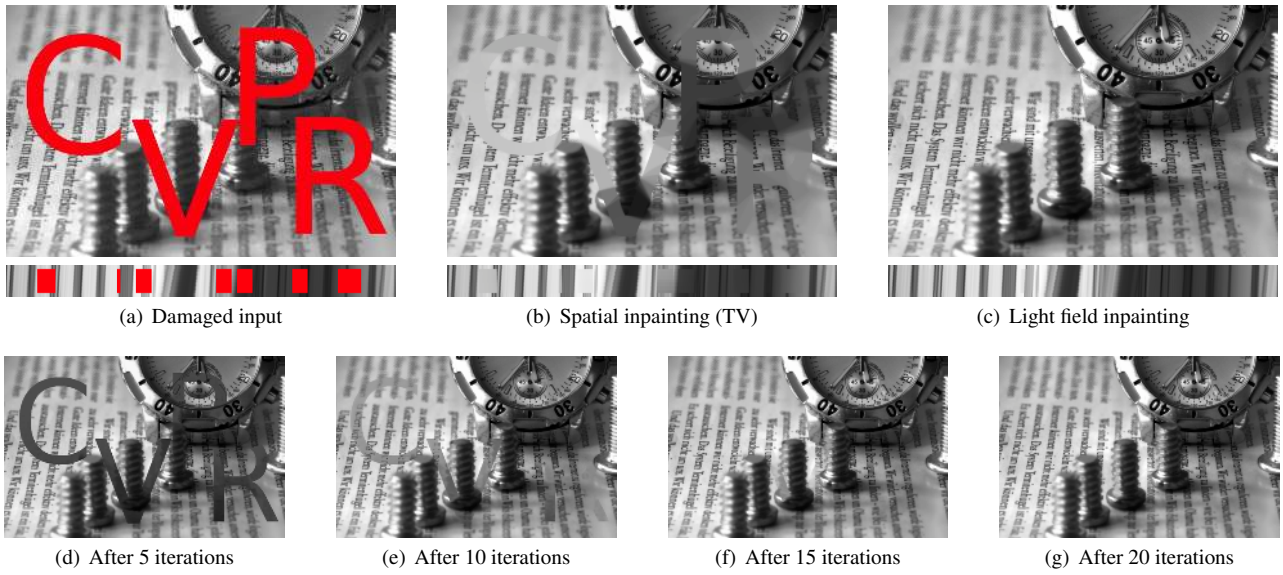
$$\operatorname*{argmin}_{\boldsymbol{U}} J_{\lambda\mu}(\boldsymbol{U}) \text{ such that } \boldsymbol{U} = \boldsymbol{F} \text{ on } \Omega \setminus \Gamma. \qquad (12)$$

Inpainting is particularly interesting because of two reasons. First, the inpainted regions show the raw results of regularization without presence of a data term. Second, a new situation arises when disparity is also unknown in the inpainting domain. In the latter case, the regularizer on the epipolar plane images is undefined. We will discuss how to infer unknown disparity fields in the next section. For the light field inpainting experiments presented in figures 7 and 8, we assumed disparity to already be reconstructed.

In general, light field inpainting leads to much improved results with visually sharper boundary transitions. Note that figure 8 also demonstrates that light field inpainting can be considered as a novel method for view interpolation.

Disparity map regularization

If disparity itself is the unknown field to be recovered, then the general model (9) is not convex anymore, since the regularizer in turn depends on knowledge the disparity field. Thus, one needs to resort to an iterative scheme. If possible, we start with a reasonable initialization, in the case of view



(a) Damaged input (b) Spatial inpainting (TV) (c) Light field inpainting

(d) After 5 iterations (e) After 10 iterations (f) After 15 iterations (g) After 20 iterations

Figure 7. *Light field inpainting for repairing damaged regions.* The same (x, y)-domain was marked as damaged in all of the input views. Using the inpainting scheme, it is possible to recover almost all of the missing information from neighbouring views, which demonstrates the power of the regularization framework. In comparison, pure spatial inpainting is of course unable to restore anything close to the real data in the damaged regions. The bottom row of images shows the progress of convergence of the algorithm.

interpolation this can for example be a disparity field obtained by linear interpolation. Otherwise, we initialize with a disparity of zero or close to the expected median of disparity values. We then iteratively solve (9) for a new disparity field, and update the regularizer with the new disparity field after every iteration. Of course, this process can only reach a local optimum.

Further improvements are possible if one includes the local constraints (5) on the disparity field. We include this non-linear set of constraints in the energy minimization using Lagrange multipliers in order to optimize for a consistent disparity field. Experiments show that this way, it is possible to obtain better estimates for disparity in e.g. inpainting applications, see figure 8.

7. Conclusion

For solving inverse problems on ray space, priors are required which respect the directional structure on epipolar plane images induced by disparity. Since the directions need to be followed very accurately, a natural tool to use are continuous methods which allow to model regularizers unbiased towards an underlying grid structure. In this work, we introduce a general variational framework for solving inverse problems on ray space with arbitrary differentiable and convex data terms, like for denoising, inpainting and segmentation. We demonstrate that these fundamental applications in image analysis can be solved more accurately on a light field structure than individual images. This way, the proposed method contributes a solid foundation for the future development of variational light field analysis.

References

[1] A. Beck and M. Teboulle. Fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sciences*, 2:183–202, 2009.

[2] T. Bishop and P. Favaro. The light field camera: Extended depth of field, aliasing, and superresolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):972–986, 2012.

[3] R. Bolles, H. Baker, and D. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1(1):7–55, 1987.

[4] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.*, 40(1):120–145, 2011.

[5] A. Criminisi, S. Kang, R. Swaminathan, R. Szeliski, and P. Anandan. Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. *Computer vision and image understanding*, 97(1):51–85, 2005.

[6] B. Goldluecke, E. Strekalovskiy, and D. Cremers. The natural vectorial total variation which arises from geometric measure theory. *SIAM Journal on Imaging Sciences*, 5(2):537–563, 2012.

[7] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The Lumigraph. In *Proc. SIGGRAPH*, pages 43–54, 1996.

[8] K. Kolev, T. Pock, and D. Cremers. Anisotropic minimal surfaces integrating photoconsistency and normal informa-

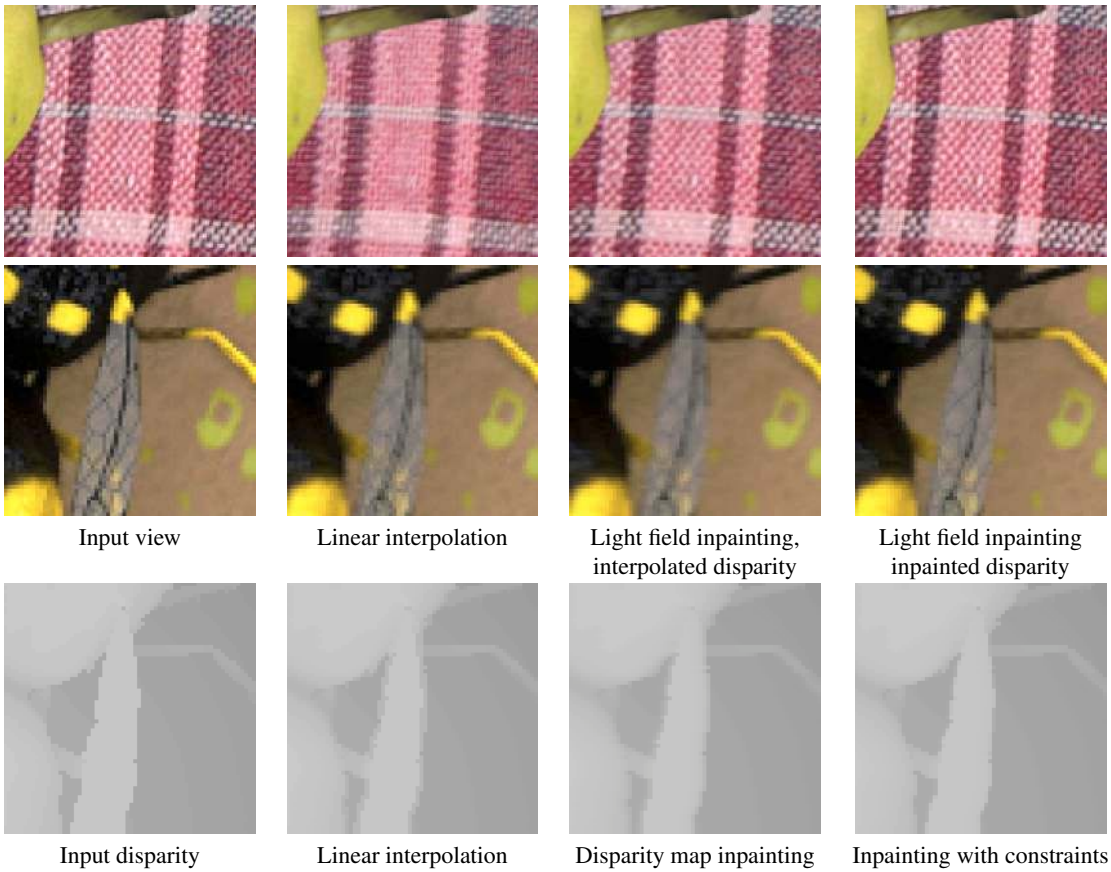


Figure 8. *Light field inpainting for view interpolation.* Intermediate views in the upsampled light field were marked as unknown regions before solving the inpainting model (12). Images in the two top rows above show, from left to right, closeups of one of the input views, a novel view generated by linear interpolation, i.e. standard light field rendering, a novel view generated by inpainting with interpolated disparity quantities, and finally a novel view generated by inpainting with disparity quantities also recovered via inpainting. In the bottom row, one can compare disparity fields in the unknown regions generated with different methods. In particular, we can see that the optimal way to infer disparity is via inpainting and additional observation of the local constraints (5). Results obtained via inpainting are generally sharper than those obtained via interpolation and show less ghosting artifacts.

tion for multiview stereo. In *Proc. European Conference on Computer Vision*, 2010.

[9] M. Levoy. Light fields and computational imaging. *Computer*, 39(8):46–55, 2006.

[10] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. SIGGRAPH*, pages 31–42, 1996.

[11] A. Lumsdaine and T. Georgiev. The focused plenoptic camera. In *In Proc. IEEE International Conference on Computational Photography*, pages 1–8, 2009.

[12] R. Ng. *Digital Light Field Photography*. PhD thesis, Stanford University, 2006. Note: thesis led to commercial light field camera, see also `www.lytro.com`.

[13] C. Perwass and L. Wietzke. The next generation of photography, 2010. `www.raytrix.de`.

[14] S. Wanner and B. Goldluecke. Globally consistent depth labeling of 4D light fields. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 41–48, 2012.

[15] S. Wanner and B. Goldluecke. Spatial and angular variational super-resolution of 4D light fields. In *Proc. European Conference on Computer Vision*, 2012.

[16] S. Wanner, C. Straehle, and B. Goldluecke. Globally consistent multi-label assignment on the ray space of 4D light fields. In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2013.

[17] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*, 24:765–776, July 2005.

[18] C. Zach, M. Niethammer, and J.-M. Frahm. Continuous maximal flows and Wulff shapes: Application to MRFs. In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2009.