Light Field Denoising and Upsampling Using Variations of the 4D Bilateral Filter

Mariana Barakchieva*
University of Bern, Switzerland

Figure 1: *Denoising of a light field. From left to right: the original light field (LF) central perspective, LF with added noise, denoised LF using the 4D Bilateral filter.*

Abstract

This paper presents an effective and not complicated method for denoising and upsampling of light fields (LF). Denoising of a light field scene is an important problem, since sensor noise is often affecting image quality, especially in low-light conditions. Afterwards the denoised light field can be used for further processing, e.g. depth reconstruction and superresolution. Also, denoising a LF can be used to produce a clean single view, resulting in better signal-to-noise ratio (SNR) than any other single-image denoising method. Having in mind the multi-dimensional structure of the LF, it is also important to have a quick and efficient algorithm for downsampling and upsampling the LF. The method proposed here is based on the four-dimensional Bilateral filter. It solves both problems fast and outperforms other related algorithms.

CR Categories: I.4 [Image Processing and Computer Vision]: ;—[I.4.4]: Image Processing and Computer Vision—Restoration;

Keywords: light fields, denoising, upsampling, 4D filter, bilateral filter

Links: DL PDF

1 Introduction

When light fields were introduced to computer graphics some 17 years ago, the main application proposed was to render new views[Levoy and Hanrahan 1996; Gortler et al. 1996]. Today, with the increase in computer speed, memory and bandwidth, the research interests in light fields has broadened, too. With two commercial light field cameras on the market [Perwass and Wietzke 2012; Ng 2006], scientist look at re-devising traditional image processing techniques for light fields as well as inventing novel ideas that take into account the properties of the plenoptic function. Problems approached include multiperspective panoramas, synthetic aperture, refocusing, microscopy. Going further, Marc Levoy, a leading expert in computational photography, predicted that in less than 20 years from now, most consumer cameras will be light field cameras [Levoy 2006].

This paper solves two main problems - denoising and upsampling. Since many computer vision applications involve low-light condition of image capturing, having a simple and fast denoising algorithm is crucial. Unlike traditional denoising approaches, ours takes into account the four-dimensional structure of the light field (LF), removing the noise not of a single 2D perspective, but rather of the whole light field. The second issue is that algorithms for light fields processing usually involve high computational costs (since we add the angular domain of the LF). Thus, an alternative is to work on a low-resolution version of the light field and afterwards upsample the computed results. For both applications we use the four-dimensional Bilateral filter, and for the task of upsampling a reference high resolution LF is also needed.

Before reviewing other works that solve the same or similar problems (Sec. 2), we propose a short introduction to light fields. Afterwards, the method we use is explained (Sec. 3) and results and comparisons are shown (Sec. 4). The final section proposes ideas for further development and concludes the paper.

1.1 From plenoptic function to light field

In 1936, Arun Gershun first coined the term *light field* as the amount of light traveling in every direction through every point in space [Gershun 1939]. To formalize it using geometrical optics theory, the *plenoptic function* is defined. It is the radiance along all rays in a 3D space, and consists of five dimensions - the ray position (x,y,z) and direction (θ, ϕ). However, since radiance is constant along a ray (if there are no blockers), one dimension is redundant. Thus, we get to a 4D function that Marc Levoy named *4D light field* and that is defined as radiance along rays in empty space [Levoy and Hanrahan 1996].

An intuitive way to visualize a light field is with the two-plane parametrization (see Fig. 2). Light rays are parametrized by the

*e-mail:mary.barakchieva@gmail.com

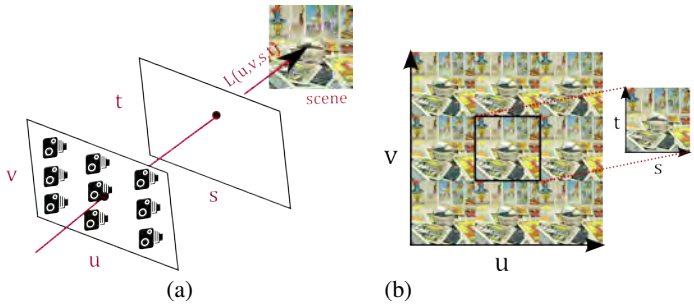


Figure 2: *Two-plane parametrization of light fields:* **(a)** *the two planes, (u,v) and (s,t) are parallel and one can think of the camera image plane being parallel to (u,v) plane and the scene being behind (s,t) plane* **(b)** *the light field as a set of 2D images, each captured from a different observer position on the (u,v) plane.*

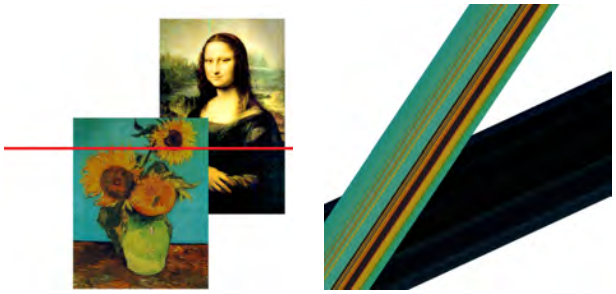


Figure 3: *Scene image and corresponding EPI. Note that the EPI captures depth information - the lines with steeper slope are in the foreground in the scene (less depth value) [Chai et al. 2000].*

intersection with the two parallel planes and each ray stores RGB radiance. From photographic point of view, the light field can be seen as a collection of images of the *(s,t)* plane (and any objects behind it), taken from each position on the *(u,v)* plane.

Then one needs to define ray space: it is a coordinate system with *v* and *t* axes; thus, a ray from the scene is a point in ray space and all rays through one scene point form a line in ray space. The slope of a line in ray space is inversely related to depth in the original scene. Related to it is the epipolar plane image (EPI) - a 2D image, constructed by stacking one over the other image lines (fixed *t* coordinate) from all images along a row of the *(u,v)* plane (fixed *v* coordinate) (see Fig. 3).

Next, we summarize works that deal with the problems of denoising and upsampling.

2 Related work

2.1 Denoising of light fields

First, we will review works that solve the problem of light field denoising. While single-image denoising has developed for a while (a summary and evaluation of classical methods is provided in [Buades et al. 2005]), we could find only a few works that deal with multiple-image denoising. Indeed, this is an important issue because first of all, plenoptic cameras aim at a very high sensor resolution, thus dense and small sensor pixels, which makes them prone to noise; for example, all currently available light field cameras have a CMOS sensor, which generally suffers from high read noise and non-uniformity, resulting in more noisy images (lower signal-to-noise ratio (SNR)) and lower dynamic range [Liu et al. 2002]. Secondly, a lot of computer vision applications involve capturing images in imperfect and low-light conditions. A summary of several denoising algorithms follows below.

In [Zhang et al. 2009], multiple view denoising is conducted and the algorithm does not aim at denoising a light field but poses the problem as pinhole image denoising; for the solution several pinhole images from different prospectives are captured. Using a second input parameter - depth estimation, it groups similar patches in the different images, then models intensity-dependent noise and uses principal component analysis (PCA) and tensor analysis to remove that noise. The authors conducted experiments with images with added synthetic noise and scored highest compared to state-of-the-art single image denoising methods. The two main disadvantages of this algorithm are that it requires prior depth information (although the authors claim that even a low-quality depth map is enough) and outputs a single denoised image, rather than a denoised light field (loss of angular information).

Next, Mitra and Veeraraghavan propose a patch based approach for solving light field processing tasks whose observation models are linear - including denoising, supperresolution, and refocusing [Mitra and Veeraraghavan 2012]. They use Gaussian Mixture Model (GMM) to model the light field patches. The inference algorithm proposed consists of extracting patches from observed data, then estimating disparity values (with a fast subspace projection), and at the end reconstructing the corresponding light field patches using Linear Minimum Mean Square Estimator (LMMSE) algorithm. For testing purposes, Gaussian distributed noise is added to a light field and is then removed effectively. However, no comparison with another denoising approach is presented. As mentioned in the paper, the algorithm is limited to diffuse scenes and small enough patches (otherwise depth discontinuities are possible).

An efficient variational framework that takes into account the structure of light fields is presented in [Goldluecke and Wanner 2013]. They solve inverse problems on ray space, such as denoising, inpainting, and ray space labeling. This works by constructing convex priors for light fields that preserve the epipolar plane image structure and satisfy constraints related to object depth and occlusion ordering. In effect, they do regularization of vector-valued functions on ray space, while also respecting the light field geometry. Denoising is demonstrated for Raytrix plenoptic camera and synthetic light fields with added significant amount of Gaussian noise. The algorithm performs better than single-image denoising algorithms but no comparison is made to multiple-image denoising. The main disadvantage of this method again is that it requires a depth map estimation as an additional input.

The denoising method proposed in [Dansereau et al. 2013] is similar to our method in the sense that denoising is conducted with a single linear filter. The main observation is that the light field of a diffusive scene has a 4D hyperfan-shaped frequency-domain region of support at the intersection of a dual-fan and a hypercone (see Fig. 4). Knowing that, they design a filter with appropriately shaped passband that filters out the noise. Experiments were conducted with plenoptic camera light fields and added different noise types. Comparison is made with competing methods, including synthetic focus, fan-shaped anti-aliasing filters, and different state-of-the-art nonlinear image and video techniques. Visually, the hyperfan outperforms the other methods but its disadvantage is that it does not work on non-Lambertian and occluding scenes. That is because if there is occlusion, the planes in the light fields will be truncated, and in the case of non-Lambertian surfaces, rays within the plane have different values; both of these conform to the hyperfan passband and thus filtering these areas might result in attenuation.

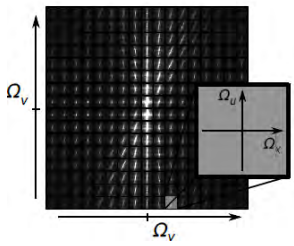


Figure 4: *The maximum magnitude per frequency, calculated for six light fields, shows the hyperfan shape[Dansereau et al. 2013].*

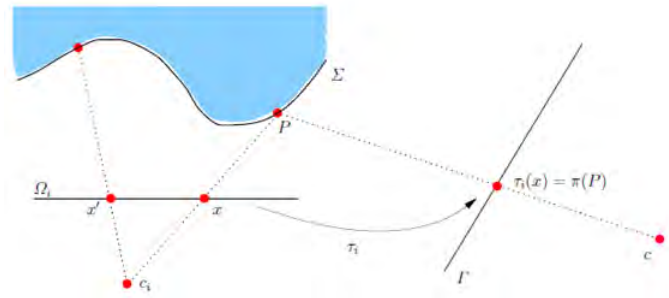


Figure 5: *Mapping from input image to the image plane of the novel view; The point x is visible on the new image, while the point x' is not - this is inferred by the depth map of the scene[Wanner and Goldluecke 2013].*

2.2 Light fields upsampling

We will review two papers that deal with upsampling of light fields. The work of Jarabo et al. incorporates downsapling and upsampling techniques as part of a framework for efficient propagation of light field edits [Adrian Jarabo and Gutierrez]. First, they define a similarity metric, adapted to the context of light field editing, which models the affinity between pixels. Thus, downsampling is conducting by first mapping all pixels into affinity space, then recursively subdividing the space into clusters and getting a single representative for each cluster. After processing is done on the downsampled LF, they do upsampling with the Joint bilateral upsampling, using a guidance full-resolution LF. This is a very similar solution to the method described in this paper. However, since they project in affinity space, searching for a pixel neighbor is more computationally expensive. The method has two main disadvantages - first, it has complexity linear with the light field size, and second, it requires large memory for storing the clusters and the correspondence between pixels and clusters.

The second paper describes upsampling in the spatial domain and generation of novel views in the angular domain, by solving a variational inverse problem [Wanner and Goldluecke 2013]. The algorithm is best explained in Fig. 5: having a depth map as additional input, the scene surface Σ can be inferred; then a transfer map τ_i projects from the input image Ω_i to the image plane Γ of the novel view. Scene ordering is preserved using a binary mask m_i. The advantage of the described method is that it takes into account the scene geometry, but also requires a depth map as additional input and the output is a single super-resolved LF view, not a LF structure.

Next, an explanation of the method adopted in this paper follows.

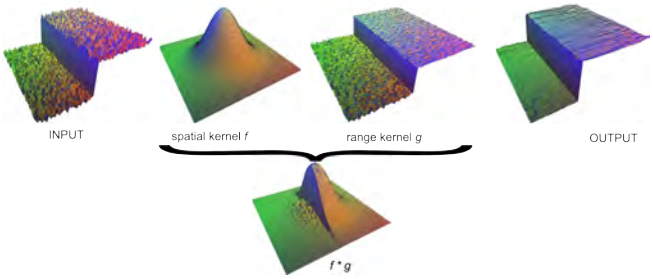


Figure 6: *Bilateral filtering[Durand and Dorsey 2002].*

3 Method

Taken into account the high-dimensional structure of the light field, a robust filter would have to filter on all dimensions. One option would be to use the Gaussian filter. Since it is a separable filter, its extension to 4D is trivial. However, the Gaussian filter smooths also over the edges of the images. Thus, we worked on extending the Bilateral filter to four dimensions and looked at applications of it on light fields. First, since light field cameras capture a lot of redundant information, the 4D Bilateral filter could be used to elegantly reject unwanted noise, while at the same time preserve details and edges. Secondly, the 4D filter is also efficient in upsampling light field data, which is applicable in all computationally costly operations on light fields - like for example depth map computation.

3.1 Bilateral Filtering

The Bilateral filter is a non-linear edge-preserving smoothing filter, first introduced by Tomasi and Manduchi in 1998 [Tomasi and Manduchi 1998]. It consists of two parts - spatial/domain and range filter kernel.

$$J_p = \frac{1}{k_p} \sum_q I_q f(p-q) g(I_p - I_q) \tag{1}$$

where I is the input image and J is the output/filtered image; $\mathbf{p}$ and $\mathbf{q}$ are 2D pixel positions (center $\mathbf{p}$ and neighborhood $\mathbf{q}$); f is the spatial filter kernel (2D Gaussian, controlled by spatial sigma - σ_s); g is the range kernel, which acts as a penalty on the input difference (1D Gaussian function, controlled by range sigma - σ_r); k_p is normalization factor, sum of all weights $f * g$ (see Fig. 6). Edges are preserved since the bilateral filter $f * g$ decreases as the range distance and/or the spatial distance increases.

Elevation of the two-dimensional Bilateral filter to higher dimensions is proposed in [Jayme C. Kosior 2007; M Mendrik 2011] for the purposes of denoising of 4D MR and CT data, where the signal is 4D and the time-intensity profiles need to be preserved. [Baek and Jacobs 2010] proposes methods for accelerating spatially varying high-dimensional Gaussian filters, and especially the bilateral filter. Formally, the 4D Bilateral filter does not differ much from the 2D one: now, $\mathbf{p}$ and $\mathbf{q}$ are 4D pixel coordinates, and the input and output images are 4D, too. Applied to the 4D light field with added white Gaussian noise, the filter cancels the noise and preserves details and edges. However, it works for only small noise quantities. Fig. 7 shows the decay of the denoising method as the noise standard deviation increases. Thus all experiments in Sec. 4 were made on LFs with noise $\sigma = 0.003$.

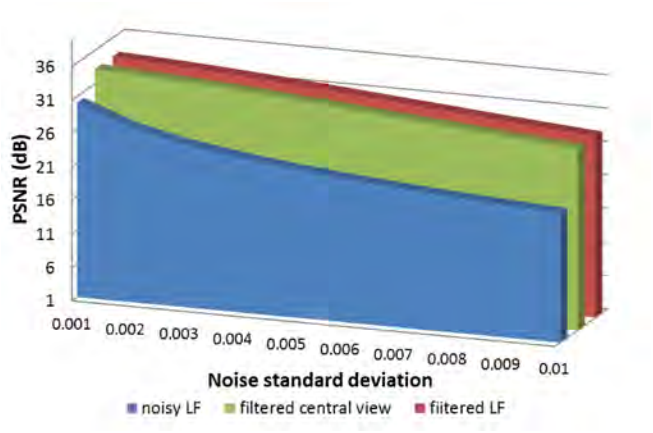


Figure 7: *The peak signal-to-noise ratio (PSNR) of the noisy and denoised LF as noise σ increases; also shown here is the PSNR of the central LF perspective - it does not differ much from the PSNR of the whole denoised LF, which is one of the main advantage of the method, proposed here.*

3.2 Joint Bilateral Filtering

An extended version of the 4D bilateral filter is used when smoothing an image without crossing strong edges in some other reference image. This is referred to as *joint* (or *cross*) *bilateral filter*. It was first introduced by [Eisemann and Durand 2004; Petschnigg et al. 2004] and used was used for combining images taken with and without a flash. Later, [Kopf et al. 2007] proposed upsampling a low-resolution image, using a high-resolution reference as another application of the joint bilateral filter. This is done by interpolating the low-resolution image in a manner that does not cross strong edges in the high-resolution reference image [Adams et al. 2011]. The only difference to Equation. 1 is that now the range filter is applied to a second *guidance image*, $\hat{I}$.

$$J_p = \frac{1}{k_p} \sum_q I_q f(p-q) g(\hat{I}_p - \hat{I}_q) \tag{2}$$

Joint Bilateral upsampling is used then to construct a high resolution *solution* from a full-resolution guidance image $\hat{I}$ and a low-resolution solution S (computed from a downsampled version of the image).

$$\hat{S}_p = \frac{1}{k_p} \sum_{q_{down}} S_{q_{down}} f(p_{down} - q_{down}) g(\hat{I}_p - \hat{I}_q) \tag{3}$$

Note that the resolutions of the guidance image and the input solution are different, thus while q_{down} takes only integer coordinates, p_{down} can also be rational. An illustration of Equation. 3 is shown in Fig. 8.

One of the applications of the Joint Bilateral upsampling on light fields is *depth estimation*, i.e. to upsample low-resolution depth map with guidance from the high resolution light field (see Fig. 9). Next, results and comparison with competing algorithms follows.

4 Results

Denoising Experiments were conducted with both synthetic light fields (from the Heidelberg archive) and light fields captured with

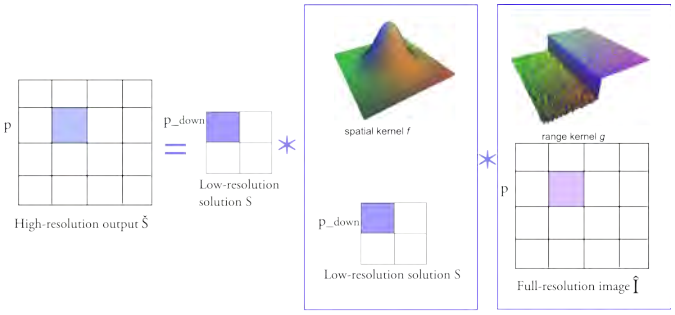


Figure 8: *Joint bilateral upsampling method. The low-resolution image determines the spatial kernel and the high resolution image is used as guidance on the range kernel; the output is a high-resolution solution.*

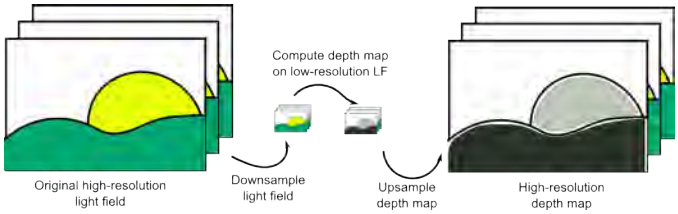


Figure 9: *Depth estimation algorithm illustration. We first downsample the LF, compute its depth map and then upsample the depth map.*

the Lytro camera. A quantitative metric was used to evaluate the different denoising methods - peak-signal-to-noise-ratio (PSNR). For a 4D reference light field LF_{ref} and its noisy counterpart LF_{noise}, PSNR is defined using the mean-square error (MSE):

$$\begin{aligned} \text{MSE} &= \frac{1}{TSUV} \sum_t \sum_s \sum_u \sum_v (LF_{ref} - LF_{noise})^2 \\ \text{PSNR} &= 20 \log_{10} \frac{\text{MAX}_{LF_{ref}}}{\sqrt{\text{MSE}}} \end{aligned} \tag{4}$$

where T, S, U, V is the size of the LF at each dimension, and $\text{MAX}_{LF_{ref}}$ is the maximum possible pixel value. PSNR is expressed in terms of the logarithmic decibel scale, where the closer the noisy image to the original noise-free image, the higher the PSNR is (for identical images, PSNR is infinity).

In the case of synthetic LF, white Gaussian noise with standard deviation 0.003 was artificially added to the original LF (see Fig. 10). This results in visually low noise in the image and SNR of 25 dB. This is more noise than what is expected from the commercially available Lytro camera. The Lytro camera uses the CMOS image sensor Aptina MT9F002[1], which exhibits SNR of 35.5dB[2].

Optimal parameters were found using brute force search. For LF denoising, the optimal spatial sigma $\sigma_s = 2$ and $\sigma_r = 0.16$. Experiments with a synthetic LF "Mona's room" is shown in Fig. 10. Our method shows best qualitative results, i.e. highest PSNR. Also visually, denoising with the 4D Bilateral filter preserves details, which can be best seen in the detailed crops - the pattern on the leaves and the grainy structure on the purple letter are preserved. Also, the fact that our method denoises the LF and preserves consistency across

[1] http://optics.miloush.net/lytro/TheCamera.aspx
[2] http://www.touptek.com/product/showproduct.php?lang=en&id=70

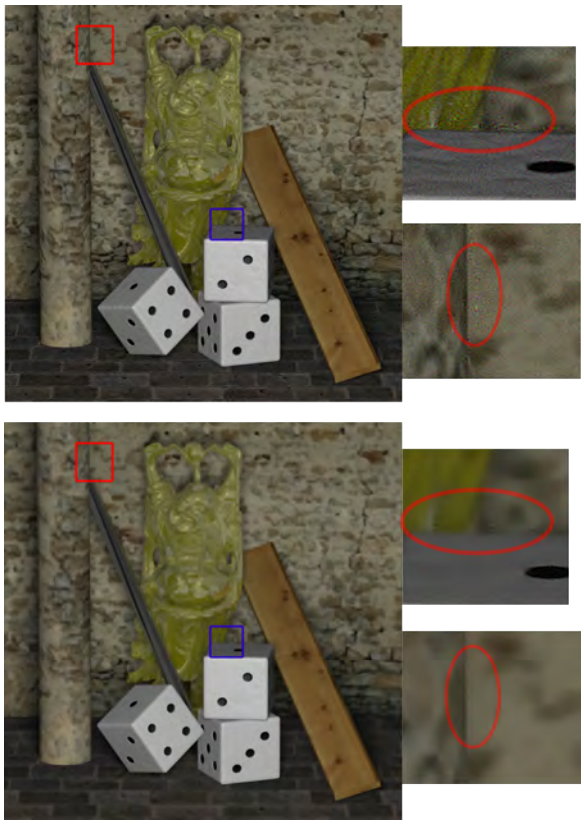


Figure 12: *Top: superresolving a noisy LF, PSNR=26.74 dB - noise is not filtered completely and artifacts are visible; Bottom: first denoising the LF and then superresolving it, PSNR=+33.83 dB - more smooth and coherent result.*

views is illustrated by calulating the PSNR of a single image - the central LF perspective. All other methods achieve the strongest denoising on the central perspective, while the 4D Bilateral filter equally denoises all perspectives.

Experiments were also conducted with light fields, captured with the Lytro camera in low-light conditions (test data provided by [Dansereau et al. 2013]). First, the image gain was adjusted, which brings even more noise to the image, and afterwards the resulted LFs were denoised using the 4D Bilateral filter (see Fig. 11). Optimal parameters were found at $\sigma_s = 2$ and $\sigma_r = 0.06$. A quantitative metric could not be adopted, since a reference, noise-free LF is lacking.

Finally, another idea that we explored was to verify if when superresolving a LF, the little amount of noise is not cleared already. Thus, white Gaussian noise with $\sigma = 0.003$ was added to a LF and then supperesolved using the method, described in [Wanner and Goldluecke 2012b]. The result can be seen on Fig. 12 (top). Then this is compared to first denoising the noisy LF with the 4D Bilateral filter and then superresolving the result - seen in Fig. 12 (bottom). Artifacts from the noise can be seen in the first solution, while the second solution is more coherent and has higher signal-to-noise ratio with about 10 dB.

Upsampling Two experiments were conducted with upsampling. First, to measure the speed and quality loss when upsampling with the Joint Bilateral filter, a LF was first downsampled and then upsampled, while measuring the execution time and also computing

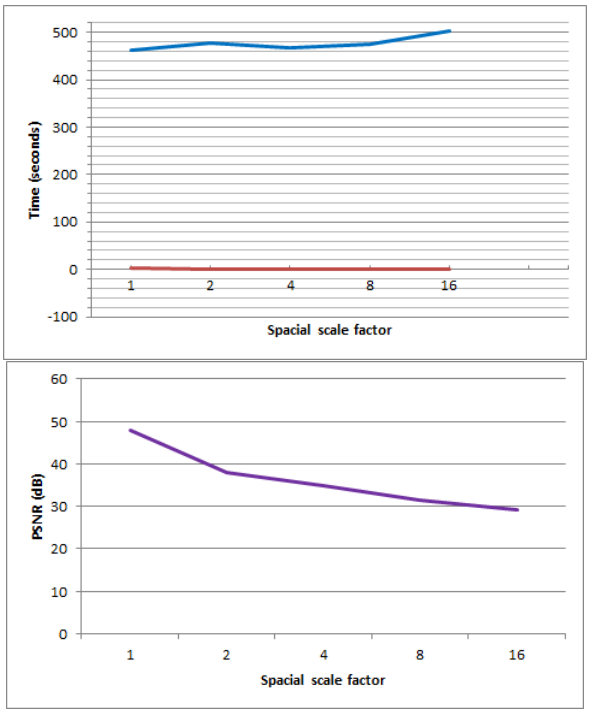


Figure 13: *Time to upsample a LF (top) vs. quality loss (bottom); spatial scale factor of 1, 2, 4, 8, and 16 correspondingly, and angular factor of 2, for downsampling (top red) and upsampling (top blue).*

an error metric - the peak-signal-to-noise ratio as introduced earlier. To get a low-resolution LF, the spatial (ST) domain was downsampled by a certain spatial scale factor, while the angular (UV) domain was always downsampled by 2 (only odd perspectives taken). For upsampling, the optimal spatial sigma $\sigma_s = 0.5$ and σ_r varies from 0.01 to 0.1.

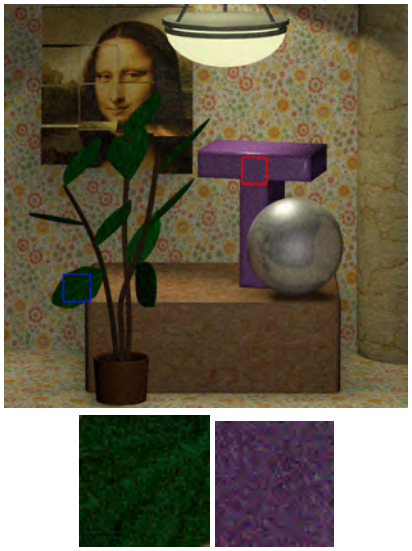
Fig. 13 show the trade-off between computational time and quality loss as the spatial scale factor increases. Also, central perspective images from the upsampled LF for different scales are shown in Fig. 14.

Finally, Fig. 15 shows an original angular perspective (a) and an upsampled one (b), aiming to illustrate the quality of the method when upsampling on the LF angular domain. As it can be seen, there is no visible quality degradation when upsampling the angular domain.

Secondly, one application of the Joint Bilateral upsampling was explored - depth calculation. The idea is that since depth map calculation is an expensive operation, a low-resolution LF can be used and then upsampled. The time needed should be smaller and the quality loss insignificant. To test that, we used the depth map calculation method described in [Wanner and Goldluecke 2012a] and the CO-COLIB light field suite [Wanner and Goldluecke 2013]. The depth map calculation require two steps - obtaining EPI depth estimates, which is computationally fast, and then integrating these estimates into a consistent single depth map. The latter is very computationally expensive operation and is proportional on the resolution of the LF. For the synthetic LF with spatial resolution 768x768 pixels and angular of 9x9, the first step took about 74 seconds, and the second step 151 seconds per perspective (12231 seconds or more than 3 hours for the whole LF). Thus, assuming that the processing time changes linearly with the number of rays, downsampling the

Original LF central perspective

LF with added white Gaussian noise, $\sigma = 0.003$



PSNR: LF=25.40 dB, image=25.40 dB

2D Gaussian

PSNR: LF=31.63 dB, image=31.63 dB

4D Gaussian

PSNR: LF=33.61 dB, image=33.62 dB

2D Bilateral

PSNR: LF=32.72 dB, image=32.71 dB

[Goldluecke and Wanner 2013]

PSNR: LF=32.76 dB, image=34.03 dB

[Mitra and Veeraraghavan 2012]

PSNR: LF=32.56 dB, image=36.62 dB

[Dansereau et al. 2013]

PSNR: LF=32.30[3] dB, image=34.15 dB

4D Bilateral (our method)

PSNR: LF=34.17 dB, image=34.22 dB

Figure 10: *Denoising of a synthetic LF with various multiple-image denoising methods. PSNR on the whole LF is shown, as well as PSNR of the central LF perspective.*

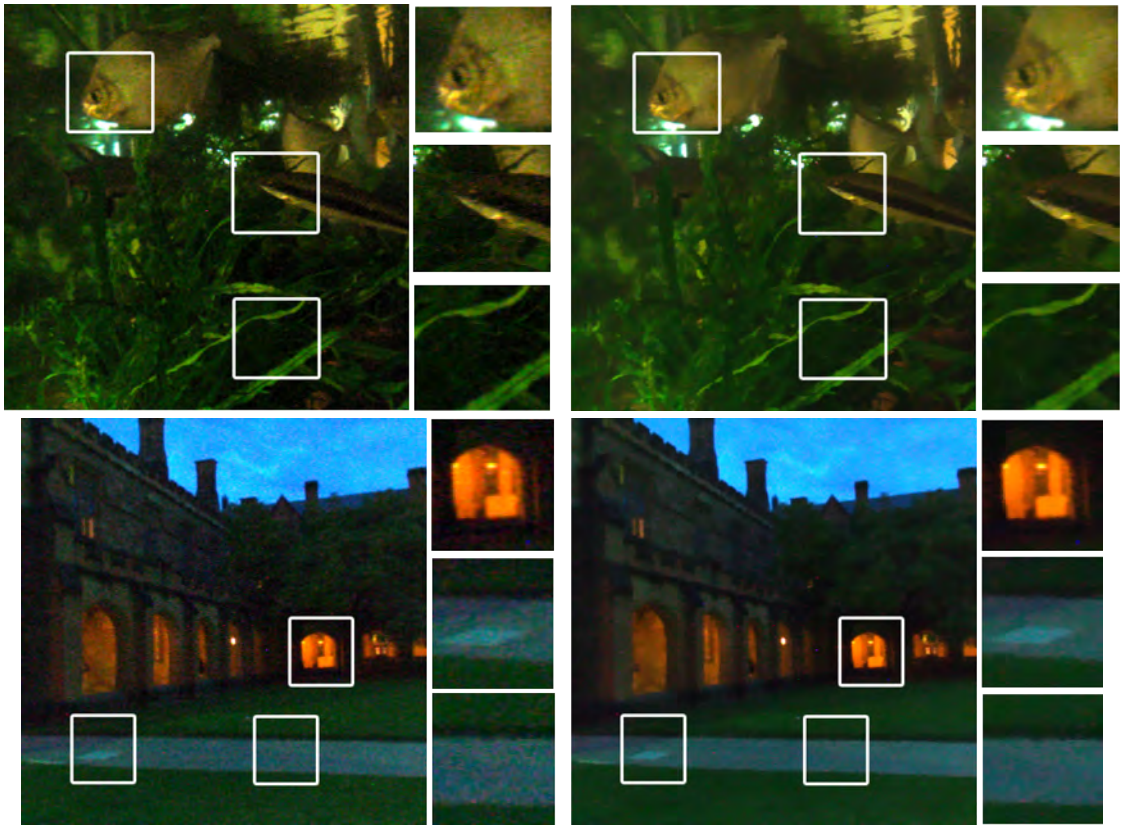


Figure 11: *Denoising of Lytro light fields with the 4D Bilateral filter: first column - LF central perspective after gain adjustment, second column - denoised LF, central perspective.*

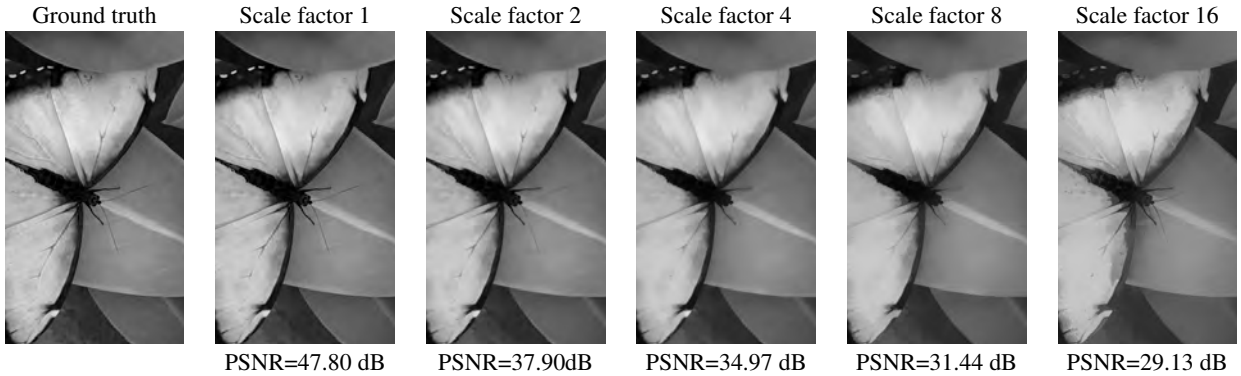


Figure 14: *Upsampled LF after downsampling with scale factors of 1, 2, 4, 8, and 16 correspondingly. Here a detail of the central LF perspective is shown, as well as the peak signal-to-noise ratio.*

(a) (b)

Figure 15: *Illustration of upsampling in the angular domain: (a) original LF perspective at (4,4); (b) upsampled LF perspective at (4,4). PSNR = +35.00 dB*

LF by a scale factor of 4 and an angular factor of 2 would speed up the depth map calculation by a factor of $4^2 * 2^2 = 64$ (theoretically resulting in about 191 seconds or 3 minutes). We need to add to that the time to do the downsampling and upsampling (468 seconds) and still we get execution time of 11 minutes, compared to 3 hours without upsampling. Fig. 16 summarizes the quality loss, compared to other upsampling methods.

5 Conclusion and further work

This paper proposes a fast and easy method for light field denoising and upsampling. The main advantage of the method is that it does not require additional information such as depth map, and that is a plus especially with non-synthetic light fields, where computing the depth map is still not trivial. The method outperforms both visually and quantitatively other related methods for the case of low noise levels. However, as described in Sec. 4, this is also the noise level exhibited in current light fields cameras.

As future work, optimizations could be performed on the algorithm, so as to speed up the Bilateral filtering. There is a lot of research on fast Bilateral filtering, a good example being the work of Baek et al. [Baek and Jacobs 2010]. Furthermore, applying the joint bilateral filter iteratively could be explored, as proposed in [Riemens et al. 2009]. To evaluate the quality of denoising of our algorithm on real LF (i.e. captured with a light field camera or camera array), Steins unbiased risk estimator (SURE) can be used, which can estimate the accuracy of the denoising algorithm, without the need of clean, noise-free image [Kishan and Seelamantula 2012].

Acknowledgements

I would like to thank David and Clemens for their enormous help, constant support, and clever ideas.

References

ADAMS, A., LEVOY, M., GUIBAS, L., HOROWITZ, M., AND OF COMPUTER SCIENCE, S. U. D. 2011. *High-dimensional Gaussian Filtering for Computational Photography*. Stanford University.

ADRIAN JARABO, B. M., AND GUTIERREZ, D. Efficient propagation of light field edits.

BAEK, J., AND JACOBS, D. E. 2010. Accelerating spatially varying gaussian filters. In *ACM SIGGRAPH Asia 2010 papers*, ACM, New York, NY, USA, SIGGRAPH ASIA '10, 169:1–169:10.

BUADES, A., COLL, B., AND MOREL, J. M. 2005. A review of image denoising algorithms, with a new one. *Simul 4*, 490–530.

CHAI, J.-X., TONG, X., CHAN, S.-C., AND SHUM, H.-Y. 2000. Plenoptic sampling. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, SIGGRAPH '00, 307–318.

DANSEREAU, D. G., BONGIORNO, D. L., PIZARRO, O., AND WILLIAMS, S. B., 2013. Light field image denoising using a linear 4d frequency-hyperfan all-in-focus filter.

DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '02, 257–266.

EISEMANN, E., AND DURAND, F. 2004. Flash photography enhancement via intrinsic relighting. In *ACM SIGGRAPH 2004 Papers*, ACM, New York, NY, USA, SIGGRAPH '04, 673–678.

GERSHUN, A. 1939. The light field. *J. Math. and Physics 18*, 51–151.

GOLDLUECKE, B., AND WANNER, S. 2013. The variational structure of disparity and regularization of 4d light fields.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '96, 43–54.

JAYME C. KOSIOR, ROBERT K. KOSIOR, R. F. 2007. Robust dynamic susceptibility contrast mr perfusion using 4d nonlinear noise filters. *Journal of Magnetic Resonance Imaging 9999*.

KISHAN, H., AND SEELAMANTULA, C. S. 2012. Sure-fast bilateral filters. In *ICASSP*, IEEE, 1129–1132.

KOPF, J., COHEN, M. F., LISCHINSKI, D., AND UYTTENDAELE, M. 2007. Joint bilateral upsampling. *ACM Trans. Graph. 26*, 3 (July).

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '96, 31–42.

LEVOY, M. 2006. Light fields and computational imaging. *Computer 39*, 8, 46–55.

LIU, X., GAMAL, A. E., HOROWITZ, M. A., AND WANDELL, B. A., 2002. Cmos image sensors dynamic range and snr enhancement via statistical signal processing, June.

M MENDRIK, EVERT-JAN VONKEN, B. G. H. W. D. J. A. R. T. V. S. E. J. S. M. A. V. M. P. 2011. Tips bilateral noise reduction in 4d ct perfusion scans produces high-quality cerebral blood flow maps. *Physics in Medicine and Biology 56*, 13, 3857.

MITRA, K., AND VEERARAGHAVAN, A. 2012. Light field denoising, light field superresolution and stereo camera based refocussing using a gmm light field patch prior. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, 22–28.

NG, R. 2006. *Digital light field photography*. PhD thesis, Stanford, CA, USA. AAI3219345.

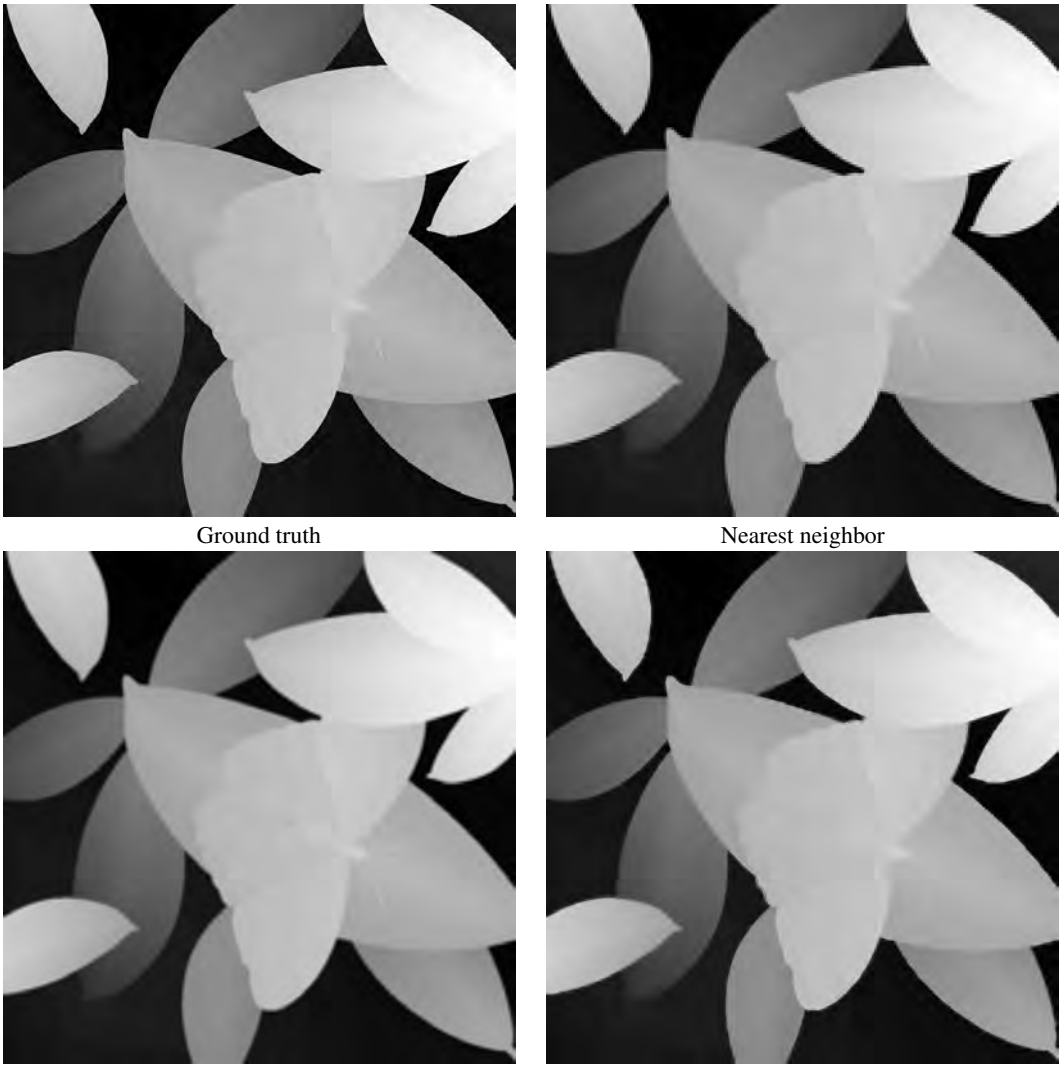


Figure 16: *Comparison of different upsampling techniques for a spatial scale factor of 4. As it can be seen, the Joint Bilateral upsampling produces best visual results.*

PERWASS, C., AND WIETZKE, L., 2012. Single lens 3d-camera with extended depth-of-field.

PETSCHNIGG, G., SZELISKI, R., AGRAWALA, M., COHEN, M., HOPPE, H., AND TOYAMA, K. 2004. Digital photography with flash and no-flash image pairs. In *ACM SIGGRAPH 2004 Papers*, ACM, New York, NY, USA, SIGGRAPH '04, 664–672.

RIEMENS, A. K., GANGWAL, O. P., BARENBRUG, B., AND BERRETTY, R.-P. M., 2009. Multistep joint bilateral depth upsampling.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *Proceedings of the Sixth International Conference on Computer Vision*, IEEE Computer Society, Washington, DC, USA, ICCV '98, 839–.

WANNER, S., AND GOLDLUECKE, B. 2012. Globally consistent depth labeling of 4D lightfields.

WANNER, S., AND GOLDLUECKE, B. 2012. Spatial and angular variational super-resolution of 4d light fields.

WANNER, S., AND GOLDLUECKE, B. 2013. Variational light field analysis for disparity estimation and super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

ZHANG, L., VADDADI, S., JIN, H., AND NAYAR, S. K. 2009. Multiple view image denoising. In *Computer Vision and Pattern Recognition*, 1542–1549.