

Heterogeneous Light Fields

Maximilian Diebold, Bernd Jähne
 Heidelberg Collaboratory for Image Processing
 Speyerer Str 6, 69115 Heidelberg, Germany
 maximilian.diebold@iwr.uni-heidelberg.de

Alexander Gatto
 Sony Deutschland GmbH
 Hedelfinger Strae 61, 70327 Stuttgart, Germany

Abstract

In contrast to traditional binocular or multi-view stereo approaches, the adequately sampled space of observations in light-field imaging allows, to obtain dense and high quality depth maps. It also extends capabilities beyond those of traditional methods. Previously, constant intensity has been assumed for estimating disparity of orientation in most approaches to analyze epipolar plane images (EPIs). Here, we introduce a modified structure tensor approach which improves depth estimation. This extension also includes a model of non-constant intensity on EPI manifolds. We derive an approach to estimate high quality depth maps in luminance-gradient light fields, as well as in color-filtered light fields. Color-filtered light fields pose particular challenges due to the fact that structures can change significantly in appearance with wavelength and can completely vanish at some wavelength. We demonstrate solutions to this challenge and obtain a dense sRGB image reconstruction in addition to dense depth maps.

1. Introduction

The basis of light-field [8] imaging is the plenoptic function as introduced in [1]. It represents a multi-dimensional function describing all the information available of light reflected from a scene. This comprises the direction and spectral radiance of the light. To capture light fields, the plenoptic function is simplified in its dimensionality to a four dimensional subspace, at times termed the lumigraph. This light field representation was first introduced in computer graphics by both Gortler *et al.* [9] and Levoy *et al.* [16]. The lumigraph describes the ray path as parameterized by two parallel planes. Along the ray path, the radiance remains constant. That means that all rays leaving a surface point and intersecting the two parallel planes have equal intensity in the resulting light field. Due to that constraint, most methods to compute disparities such as [24, 13, 3, 5, 4] relate only to light fields with this property. Even for binocular or multi-view stereo approaches, as

proposed in [17, 12, 6, 20] correspondence between points is modeled on their having the same intensity value. Thus, current cameras such as Raytrix [19] and Lytro [7] focus on reconstruction from input images having similar color information. A violation, in case of images captured with different color filters or illuminations, needs sophisticated pre-processing algorithms to adapt the data for the depth estimation, as in the work of Yong *et al.* [10]. In that approach, the input image data is mapped to a log-chromaticity color space to obtain an illumination-independent color space to find corresponding points in the input data. Thus the direct computation of depth maps on heterogeneous input data is made in neither multi-view stereo nor current light field imaging. Hyper-spectral images may be generated using a single camera with revolving color filters before the objective, as described in Tominaga [22]. Unfortunately with this setup, it is not possible to make depth estimates on the underlying scene. In this paper we present heterogeneous light fields which have changing properties between the captured images – such as with the presence of illumination gradients or application of colored filters. We demonstrate that a modified structure tensor is able to process heterogeneous light fields. We analyze the limits of both illumination gradients and randomly illuminated light fields. Furthermore, we show that even for color filtered light fields the structure tensor approach computes locally highly reliable depth information which can be merged to a dense depth map. With this dense depth map it is possible to compute a hyper-spectral image with respect to a reference view out of the used color filtered light field. To visualize the obtained hyper-spectral image we introduce a method to approximate the sRGB color space from the hyper-spectral information and display the final RGB reconstruction.

2. Light field disparity estimation

A light field may be defined as a lumigraph [9] having two parallel planes Π and Ω . While the Ω -plane addresses the image coordinates $(x, y) \in \Omega$ of each taken image, the Π -plane defines the location of the focal points $(s, t) \in \Pi$ of each camera. A 3D light field as used to capture hetero-

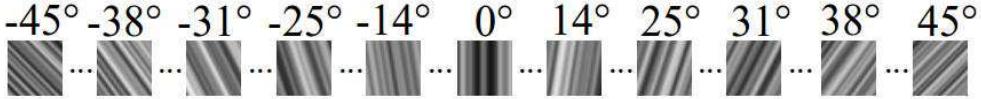


Figure 1. Shows synthetically generated EPIs. For each orientation (angle) several EPIs are analyzed to determine the overall precision. The overall precision is computed by using the equation 23.

geneous light fields is given by

$$L : \Omega \times \Pi \rightarrow \mathbb{R} \quad (s, x, y) \mapsto L(s, x, y), \quad (1)$$

where $L(s, x, y)$ denotes the pixel intensity value of the ray intersecting (x, y) of the image plane and s of the focal plane (in our case, just a focal line). 2D slices from L are termed epipolar plane images Σ . Ignoring issues of calibration and rectification, an epipolar plane image can be extracted from a 3D light field by setting the y axis to a fixed value y^* . The addressed epipolar plane image can be described by the equation

$$S_{y^*} : \Sigma_{y^*} \rightarrow \mathbb{R} \quad (2)$$

$$(x, s) \mapsto S_{y^*}(x, s) := L(s, x, y^*). \quad (3)$$

To compute disparity maps of a given scene, the structure tensor is used in determining the underlying orientations within an EPI. This structure tensor J_1 , used in Wanner [25], is defined by the equation

$$J_1 = \tau * \begin{pmatrix} \left(\frac{\partial \hat{S}}{\partial x}\right)^2 & \frac{\partial \hat{S}}{\partial x} \cdot \frac{\partial \hat{S}}{\partial s} \\ \frac{\partial \hat{S}}{\partial s} \cdot \frac{\partial \hat{S}}{\partial x} & \left(\frac{\partial \hat{S}}{\partial s}\right)^2 \end{pmatrix} =: \begin{pmatrix} J_{xx} & J_{xs} \\ J_{xs} & J_{ss} \end{pmatrix} \quad (4)$$

with the abbreviation

$$\hat{S} := \sigma * S, \quad (5)$$

where σ defines an inner Gaussian smoothing and τ an outer Gaussian smoothing applied over each of the structure tensors components. The underlying disparity d is achieved, as introduced in Wanner *et al.* [23], by using the structure tensor components. This reduces disparity estimation to the equation

$$d = \tan\left(\frac{1}{2} \arctan\left(\frac{2J_{xs}}{J_{xx} - J_{ss}}\right)\right). \quad (6)$$

This formula demonstrates the advantage of the structure tensor approach. In contrast to binocular or multi-view stereo, where a correspondence search between several images is necessary, the structure tensor simply computes the disparity map by analyzing orientation in epipolar plane images. Aside from the disparity, a reliability measure is also computed to indicate confidence of the underlying orientation. This reliability measure is termed coherence c and was first introduced in Bigun *et al.* [2]

$$c := \sqrt{\frac{(J_{xx} - J_{ss})^2 + 4(J_{xs})^2}{(J_{xx} + J_{ss})^2}}. \quad (7)$$

3. The modified structure tensor

To derive a robust structure tensor for heterogeneous light-field processing we analyze the second order structure tensor introduced by Mülich *et al.* [18]. This structure tensor is used to separate two transparent overlaying orientations, as introduced in Wanner *et al.* [24]. Thus, it already describes an improved structure tensor that can separate two different depth layers. In heterogeneous light fields, an illumination gradient or color-filtered images result in EPI structures with properties similar to two overlaying orientations. For illumination gradient light fields, the first orientation describes the scene disparity while the second orientation indicates the orientation of the illumination gradient. In color-filtered light fields, we have a similar orientation separation, unfortunately the a priori knowledge for the second orientation, while necessary, is not known. For the disparity estimation of the underlying scene, only the first orientation is important since this orientation describes the underlying disparity and, for the depth computation, there is no benefit in using the second orientation. We use the second order structure tensor to derive a new structure tensor to process robust single orientation in heterogeneous light fields without processing the negligible second orientation. Next, we introduce the canonical-correlation analysis (CCA) to derive a new structure tensor. The CCA is used in statistics to analyze the correlation between base vectors in a given evaluation window E . To determine the underlying orientation for a given evaluation window E , we first define the base vectors which are given by the derivatives of \hat{S} . Thus the base space H becomes

$$H = \left(E * \frac{\delta(\hat{S})}{\delta x}, E * \frac{\delta(\hat{S})}{\delta y} \right) =: (H_x, H_y). \quad (8)$$

Derived from this, the needed covariance matrix ¹ computes as

$$C(H_x, H_y) = \begin{pmatrix} \text{var}(H_x, H_x) & \text{cov}(H_x, H_y) \\ \text{cov}(H_y, H_x) & \text{var}(H_y, H_y) \end{pmatrix} \quad (9)$$

with

$$\text{var}(H_x, H_x) = \frac{1}{n} \sum_{i=1}^n (H_{x,i} - H_{x,\mu})^2 \quad (10)$$

¹In the variance and covariance equation, we have a zero mean assumption $H_\mu = (H_{x,\mu}, H_{y,\mu}) = (0, 0)$. Thus the nominator of the variance is n instead of $n - 1$ not known mean values.

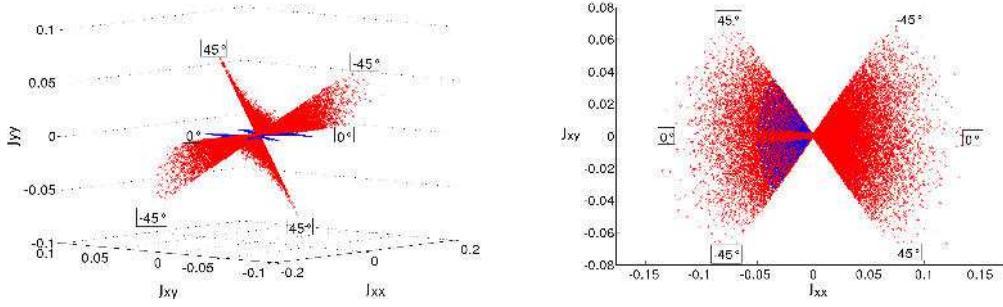


Figure 2. The first image shows the 3D visualization of the base space of the second order structure tensor. All derivative values of J for all possible orientations are plotted. The blue lines show the respectively underlying orientation estimation solution. The top view of the 3D graph shows that the 2D projection of the data matches with the single orientation.

and

$$\text{cov}(H_x, H_y) = \frac{1}{n} \sum_{i=1}^n (H_{x,i} - H_{x,\mu})(H_{y,i} - H_{y,\mu}). \quad (11)$$

where $H_{x,\mu}$ is the computed mean value of all values $n \in E$. The given covariance matrix can also be expressed in vector notation and becomes

$$C(H_x, H_y) = \frac{1}{n} \begin{pmatrix} H_x^* H_x & H_x^* H_y \\ H_y^* H_x & H_y^* H_y \end{pmatrix} = \frac{1}{n} H^* H \quad (12)$$

Next we replace the evaluation window with a Gaussian filter which transforms the covariance matrix to the already defined structure tensor as shown in equation 4. With this method, to derive the structure tensor we analyze the second order structure tensor introduced by Müllich *et al.* [18], defined by

$$J = \tau * \begin{pmatrix} J_{xx} J_{xx} & J_{xx} J_{xy} & J_{xx} J_{yy} \\ J_{xx} J_{xy} & J_{xy} J_{xy} & J_{xy} J_{yy} \\ J_{xx} J_{yy} & J_{xy} J_{yy} & J_{yy} J_{yy} \end{pmatrix} \quad (13)$$

with

$$J_{xx} = \frac{\delta^2(\sigma * S)}{\delta x^2} \quad (14)$$

$$J_{yy} = \frac{\delta^2(\sigma * S)}{\delta y^2} \quad (15)$$

$$J_{xy} = \frac{\delta^2(\sigma * S)}{\delta x \delta y} \quad (16)$$

where the base space of the second order structure tensor is given by

$$H_{3D} = (J_{xx}, J_{xy}, J_{yy}). \quad (17)$$

To analyze the behavior of this structure tensor we generate, for all possible orientations, synthetic EPIS as shown

in figure 1. For a better understanding of the correlation between the structure tensor components J_{xx} , J_{xy} and J_{yy} , each 3D point of the base space H is plotted in a 3D coordinate system as shown in figure 2. This 3D visualization shows all 3D points of H_{3D} , for all defined orientations between -45° and 45° , as red dots. The derived orientation estimates are plotted as blue lines in the J_{xx} , J_{xy} subspace. As one can see, beside the separation of two transparent overlaying orientations, a single orientation estimate is encoded in the second order structure tensor. A closer look shows that the single orientation estimation is independent of the third structure tensor component J_{yy} . That leads to a new base space for the single orientation

$$H_{2D} = (J_{xx}, J_{xy}), \quad (18)$$

which results by using the introduced procedure to a structure tensor representation defined by

$$J_2 = \tau * \begin{pmatrix} J_{xx} J_{xx} & J_{xx} J_{xy} \\ J_{xx} J_{xy} & J_{xy} J_{xy} \end{pmatrix}. \quad (19)$$

The shape of this structure tensor is similar to the first introduced structure tensor from equation 4. The only difference is an additional derivative filtering in the x direction of each component. Taking this into account, we select a derivative filter with $(2R-1)$ elements. The advantage of an odd symmetry filter, like $\frac{1}{2}[-1\ 0\ 1]$, is shown by its transfer function as described by Jähne [11]. Thus, the transfer function for the smallest possible derivative filter ($R = 2$) becomes

$$\frac{1}{2}[-1\ 0\ 1] \bullet \circ \cos(\pi \hat{k}) \quad (20)$$

where $\hat{k} \leq 1$ denotes the normalized wave number. As one can see, this filter not only attenuates low frequencies but also high frequencies. With this understanding, the usage of an inner Gaussian filter becomes obsolete, because its main purpose was for anti-aliasing, noise removal and value averaging. All of this, aside from the averaging, is now done by

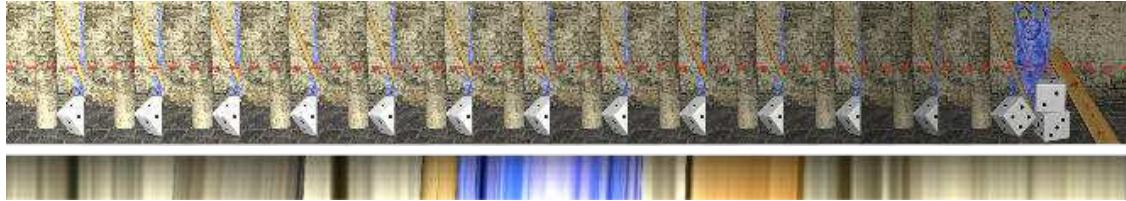


Figure 3. This EPI has a linear illumination gradient in vertical direction. The first row of the shown EPI related to the first captured image of the heterogeneous light field while the last row is related to the last captured image. As one can see increases the illumination continuously. The red line shows the position where the shown EPI is extracted.

the additional derivative filter. The averaging is transferred to the outer Gaussian filter, which has negligible effect on its value, and its shape doesn't change by the defined offset. This makes the entire processing not only faster, due to the removal of the inner Gaussian filter, but also applicable to heterogeneous light fields. Thus the new derived structure tensor becomes

$$J_3 = \tau * \begin{pmatrix} \left(\frac{\partial \hat{S}}{\partial x} \right)^2 & \frac{\partial \hat{S}}{\partial x} \cdot \frac{\partial \hat{S}}{\partial s} \\ \frac{\partial \hat{S}}{\partial s} \cdot \frac{\partial \hat{S}}{\partial x} & \left(\frac{\partial \hat{S}}{\partial s} \right)^2 \end{pmatrix} =: \begin{pmatrix} \hat{J}_{xx} & \hat{J}_{xs} \\ \hat{J}_{xs} & \hat{J}_{ss} \end{pmatrix} \quad (21)$$

with the abbreviation

$$\hat{S} := \frac{\partial S}{\partial x}. \quad (22)$$

Next we analyze the resulting precision of the new defined structure tensor for different derivative filters in contrast to the traditional structure tensor. Thus we compute the precision as described in Diebold *et al.* [5] and generate several synthetic EPIs for a discretized angular space between $[-45^\circ, 45^\circ]$ as shown in figure 1. The resulting overall precision for all evaluated orientations $i \in N$ becomes

$$\sigma_d = \sqrt{\frac{1}{N} \sum_i^N (\mu_i)^2 + \frac{4}{N} \sum_i^N (\sigma_i)^2}. \quad (23)$$

where σ_i defines the standard deviation of the evaluated estimations and μ_i denotes the mean values, to additionally

Derivative Filter	σ_d for J_1	σ_d for J_2	σ_d for J_3
Sobel	0.0588	0.0322	0.0315
Scharr	0.0299	0.0082	0.0085
Gaussian 3x3	0.0698	0.0321	0.0326

Table 1. The table shows the resulting precision of the traditional structure tensor J_1 proposed by Wanner [25] in comparison to the new structure tensor with applied inner Gaussian filter J_2 , and without additional inner Gaussian Filter J_3 . The evaluation is made for different possible derivative filters of the same shape. As one can see, the new structure tensor outperforms in the Scharr filter implementation the traditionally structure tensor by far.

consider appearing systematic error. The results of the precision analysis are shown in table 1. As one can see, the achievable precision of the new derived structure tensor J_3 increases with respect to the traditional structure tensor J_1 which was introduced by Wanner *et al.* [25]. But the question remains of how the precision changes under luminance gradient light fields.

In luminance-gradient light fields, the illumination changes from image to image, which is termed illumination gradient ΔI in the following. To analyze these kinds of heterogeneous structures, we apply illumination gradient ΔI to synthetically rendered light fields as shown in figure 4. An example of a luminance gradient light field and the resulting appearance of an EPI are shown in figure 3. To compare the reliability of the structure tensor for different illumination

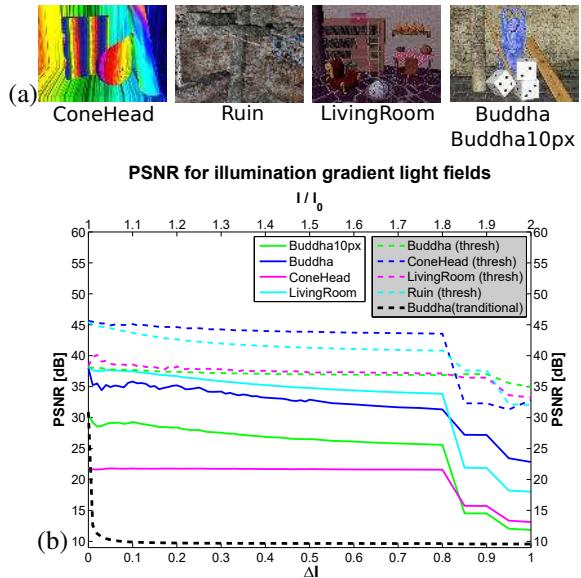


Figure 4. Shows the PSNR of the shown scenes (a) for different applied illumination gradients ΔI . Figure (b) shows the PSNR applied on the entire disparity map as well as only to coherence thresholded $\theta = 0.8$ values. For the Buddha scene the traditional implementation of the structure tensor is shown as black dashed line. As one can see keeps the PSNR almost constant until a gradient limit is reached. In contrast, the PSNR of the traditional structure tensor one drops down instantly.

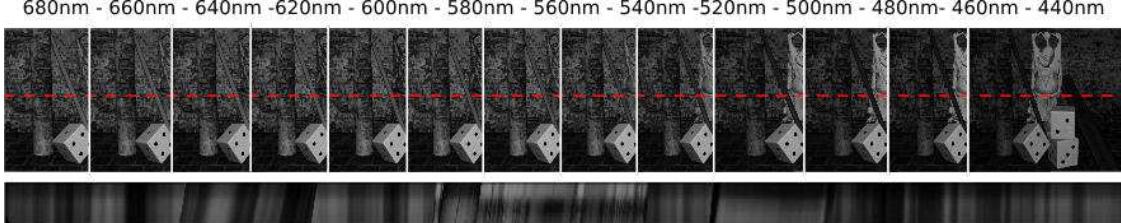


Figure 5. Shows an example of an linear heterogeneous light field. Each image is captured with a different color filter having a full width half mean of 10nm . The red line shows the position, where the shown EPI is extracted. In color filtered light fields the intensity changed in orientation direction with respect to the underlying color content.

gradients we compute for each evaluation scene the PSNR with respect to the ground truth disparity map. The results for different ΔI are shown in figure 4. As one can see, the new structure tensor keeps an almost constant PSNR until the illumination gradient reaches a limiting value. From there, the PSNR immediately decreases. The entire evaluation is made with 16bit images because we must avoid saturation. In the event of saturated regions, the PSNR will decrease due to missing orientation information, and not because of the applied gradient. Thus, 16 bit images allow isolating the influence of the illumination gradient on the estimation result. Aside from the PSNR of the improved structure tensor J_3 , the PSNR of the traditional structure tensor J_1 of the Buddha scene is shown in figure 4. As one can directly see, the traditional structure tensor is not able to process a heterogeneous light field. Furthermore, we analyze synthetically generated heterogeneous light fields having random illumination distributions. Here, we randomly shuffle the image related multiplier which was used to achieve the illumination gradient in the EPI. The results of this evaluation are shown in figure 6, illustrating the applicability to acquired light fields. Aside small illumination variations invariably occur in acquired light fields. Illumination differences appear due to flickering of the light source or because of varying camera properties such as arise from exposure jitter across the light field array. Thus, the designed structure tensor not only improves the estimation result in homogeneous light fields but also makes it possible to process heterogeneous light fields.

4. Color-filtered light fields

In this section we introduce color-filtered light fields as shown in figure 5. Color-filtered light fields are captured with color filters of different wavelengths so that image of the light field contains differing color information. The used band-pass filters have a full-width half-mean of 10 nm and are uniformly distributed in the color spectrum between 400 nm and 700 nm . This means the color-filtered light field contains spectral information of the underlying scene. After generating synthetic color-filtered light fields, we apply the introduced structure tensor J_3 to test scenes such as the rain-

bow textured scene shown in figure 7(a). Due to its breadth of color, this scene illustrates that the structure tensor is able to locally estimate the underlying orientation with respect to the visible wavelength. The local estimation results for 500 nm , 600 nm and 700 nm are seen in figure 7(b).

4.1. Color merge and total variation

It is necessary to merge the local disparity estimations into a reference view $r \in \Pi$ to obtain a dense disparity map. Thus the measured disparity needs to be transferred in the reference view. To select single disparity values, we use examine coherence order and replace estimations in the reference view with smaller coherence value. For the entire merging, we select each row $s \in \Pi \setminus r$ in the EPI and transfer the local disparity estimation of each pixel x onto the addressed position y_s in the reference view r which is given by the equation

$$y_s = |x + s \cdot d_s(x)| \quad (24)$$

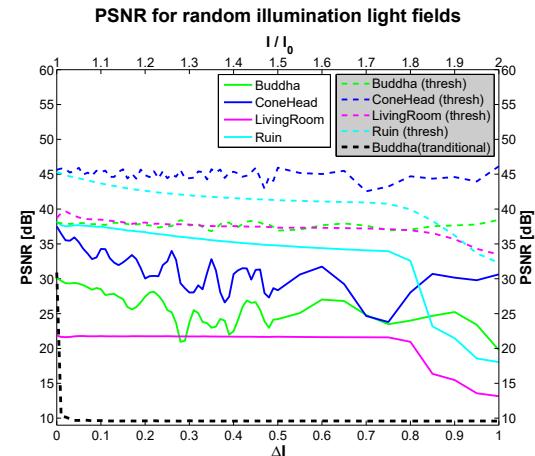


Figure 6. Illustrates the PSNR for different applied randomly shuffled illuminations defined by an underlying illumination gradient ΔI . The figure shows the PSNR applied on the entire disparity map as well as only to coherence thresholded $\theta = 0.8$ values. For the Buddha scene the traditional implementation of the structure tensor is shown as black dashed line. As one can see keeps the PSNR quiet constant while the PSNR of the traditional structure tensor drops down instantly.

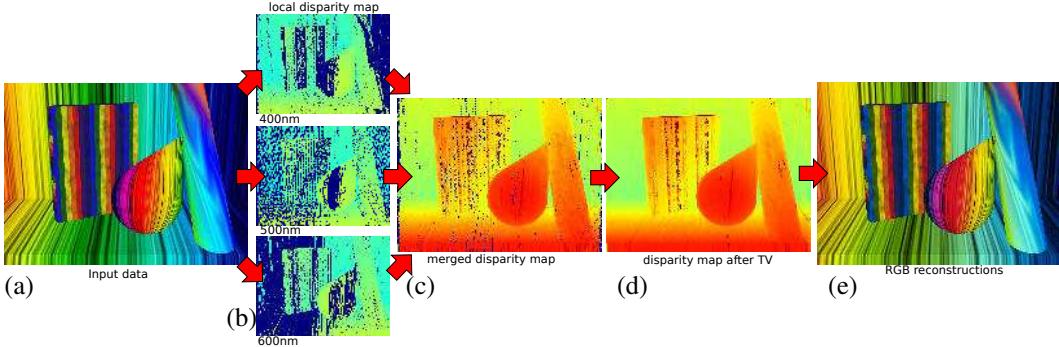


Figure 7. (a) shows the reference image of the light field which become transformed into an color filtered light field. (b) are the local disparity estimations. (c) shows the merged disparity map out of the local disparity maps. (d) shows the final disparity map after an applied total variation. (e) shows the final disparity estimation.

where the absolute value ensures only even pixel values are addressed in the reference view. That is important to consider since two or more local disparity estimates $d_s(x)$ at different positions x can address the same pixel in the reference view. This can happen when one object occludes another. The disparity merge can finally be described by

$$d_r(y_s) = d_s(x) \mid c_s(x) > c_r(y_s), d_s(x) > d_r(y_s) \quad (25)$$

where d_r and c_r are initialized with the local result of the reference view. Rounding of the applied pixel position introduces error in the resulting disparity map. To minimize this error we determine the actual shift position \hat{x} in row s . The new position becomes

$$\hat{x} = y_s - s \cdot d_s(x). \quad (26)$$

When the disparity value at the new position $d_s(\hat{x})$ is within an epsilon environment ϵ with respect to the initial disparity value $d_s(x)$, it replaces the initial disparity value at location x . A reverse disparity value checks whether the rounding has caused an object boundary to be crossed and, if so, rejects (in this case the selected disparity d_s). The final merged disparity map for a reference view of the rainbow textured scene is shown in figure 7(c). Unfortunately, it still contains some small patches with undefined disparity value. For the further processing, it is necessary to have a dense disparity map. Thus we apply a second-order total variation as introduced by F. Lenzen *et al.* [14, 15] on the merged disparity map. The proposed second order total variation approach minimizes the functional

$$F_{TV_2}(u) := \frac{1}{2} \|u - f\|_2^2 + \alpha TV(u) + \beta TV^2(u) \quad (27)$$

where $\alpha, \beta > 0$ denote regularization parameters. The result after the applied second order total variation is shown in figure 7(d). The final hyper-spectral image can now be determined by addressing all color values along the orientations whose direction is given by the achieved disparity map.

	PSNR [dB]
Buddha	24.94
ConeHead	31.39
LivingRoom	25.80
Ruin	33.31

Table 2. Shows the PSNR of the disparity estimation result with respect to the ground truth for synthetically generated color-filtered light fields.

4.2. sRGB reconstruction

For an RGB color reconstruction from the spectral images we need to find a method to determine the (R, G, B) values from the pixel values S_i of the captured spectral images $i \in N$. For this, we use the approach of Tominaga [22]. That proposes that the CIE color space (X, Y, Z) can be determined by the pixel values S_i multiplied by a weighting function M .

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = M \cdot \begin{pmatrix} S'_1 \\ \vdots \\ S'_N \end{pmatrix} \quad (28)$$

The weighting function to approximate the CIE color space can be determined by approximating the CIE color matching functions. Thus the camera quantum efficiency $QE_{cam}(\lambda)$ and the spectral sensitivity functions $BP_i(\lambda)$ of the image i related band pass filter as well as the spectral color distribution $LS(\lambda)$ of the light source need to be known. Then the CIE-color matching function $(\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda))$ can be estimated by the formula

$$\begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} = M \cdot \begin{pmatrix} QE'_1(\lambda) \\ \vdots \\ QE'_N(\lambda) \end{pmatrix} + e \quad (29)$$

with

$$QE'_i(\lambda) = BP_i(\lambda) \cdot QE_{cam}(\lambda) \cdot LS(\lambda) \quad (30)$$

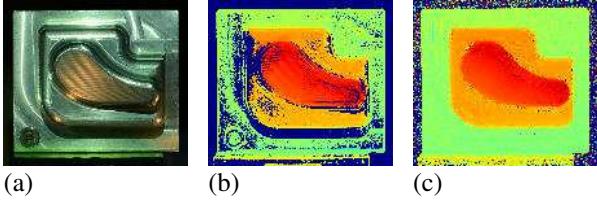


Figure 8. (a) shows a real captured metal test part with not saturated specular highlights. (b) shows the estimated disparity map using the traditional structure tensor as introduced in equation 4. (c) shows the disparity estimation result after applying our proposed structure tensor.

For the best possible approximation of the CIE color matching function we introduce the functional F_M which minimizes the area difference between the color matching function and the obtained approximation of M for the entire frequency domain Λ with $\lambda \in \Lambda$. The proposed functional becomes

$$F_M = \min_{\vec{k}} \int_{\Lambda} \begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} - M_{3 \times N} \cdot QE'_{\text{cam}}(\lambda) d\lambda \quad (31)$$

After determining the weighting function M , we can transfer the spectral information to the CIE color space using equation 28. Next we convert the CIE color space as proposed by Tominaga [22] to sRGB color space [21]

$$\begin{pmatrix} R_{\text{linear}} \\ G_{\text{linear}} \\ B_{\text{linear}} \end{pmatrix} = T \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (32)$$

with

$$T = \begin{pmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{pmatrix} \quad (33)$$

For the final visualization we also apply a gamma correction of $\gamma = 0.68$ to the linear RGB-space which transforms the linear values into sRGB.

5. Results

As result for homogeneous light fields, we show the comparison of the new structure tensors J_3 with respect to the traditional structure tensor J_1 . The traditional structure tensor leads to the result shown in figure 8(b), while our proposed method shows a full coverage of the captured metal test part as demonstrated in figure 8(c).

Next, we want to show the applicability of the proposed structure tensor to color-filtered light fields and the RGB reconstructions. To acquire heterogeneous light fields properly it is important to use linear-ordered color filters. They provide a more stable estimation for intensity gradients, as

shown for linear luminance gradient light fields in contrast to random distributed filters configurations. Due to that, a constant quality in the orientation estimation of color-filtered light fields is guaranteed as long as the gradient is not reaching the critical maximum. Furthermore, considering a color distributions of objects placed in the scene, as shown in figure 10 (c), it becomes important to select the filter respectively, to ensure that color dependent orientations appear in the underlying EPI. To estimate orientation for colors only seen by one camera is not possible. Thus we can derive two constraints

- For broad color spectra of target object, the used band pass filter can have a narrow bandwidth to obtain analyzable orientation.
- For narrow color spectra of target objects, the used band pass filter needs to be chosen that at least 5 orientations are visible in at least 5 neighboring images to guarantee a valid estimation.

As shown in figure 5, the EPI contains local orientation information while black transitions illustrate vanishing color content. Figure 9(a) shows the center view image of the initial synthetic homogeneous light field. This light field contains 31 images which get converted to a color-filtered light field as shown in figure 5. Each image is filtered with a band-pass filter having a full-width half-mean of 10 nm. The filters are uniformly distributed in 10 nm steps between 400 nm and 700 nm. The resulting RGB images of the synthetic light field is seen in figure 9(c).

The real color-filtered light fields are captured with a PCO-edge 5.5 camera, mounted on a high-precise translation stage. For the heterogeneous data we have a symmetric light-field configuration starting from 400 nm to 700 nm and back to 400 nm while each filter has a full-width half-mean of 10 nm. The filters employed are 400 nm, 450 nm, 500 nm, 515 nm, 532 nm, 550 nm, 560 nm, 589 nm, 600 nm, 650 nm, 700 nm and back down to 400 nm. The processing results of the acquired color-filtered light fields are shown in figure 10(c). Input data and additional images are provided in the additional material.

6. Conclusion

Since the traditional structure tensor suffered from the restriction of a constant intensity along the orientation, the new designed structure tensor overcomes this problem and makes it suitable for an even larger field of applications. The new structure tensor approach introduced in this paper significantly reduces the error in the estimate of orientation compared to the traditional structure tensor. We demonstrate that the modified structure tensor has the advantage of a high reliability in processing heterogeneous light fields. This makes it possible to better analyze acquired light fields,

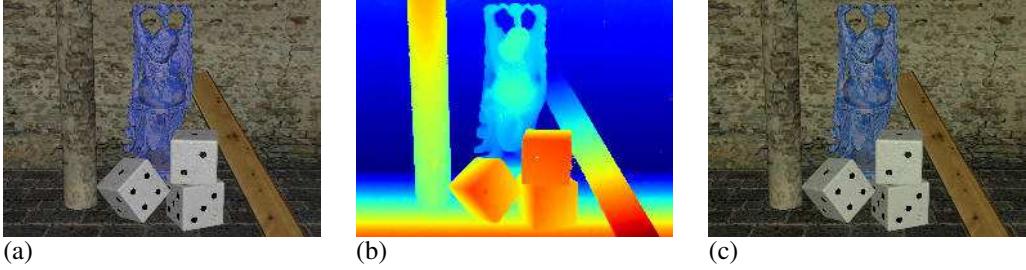


Figure 9. (a shows the original RGB image of the Buddha scene. (b) shows the final disparity result of the color filtered light field after the applied second order total variation. (c) shows the RGB reconstruction of the computed hyper-spectral image.

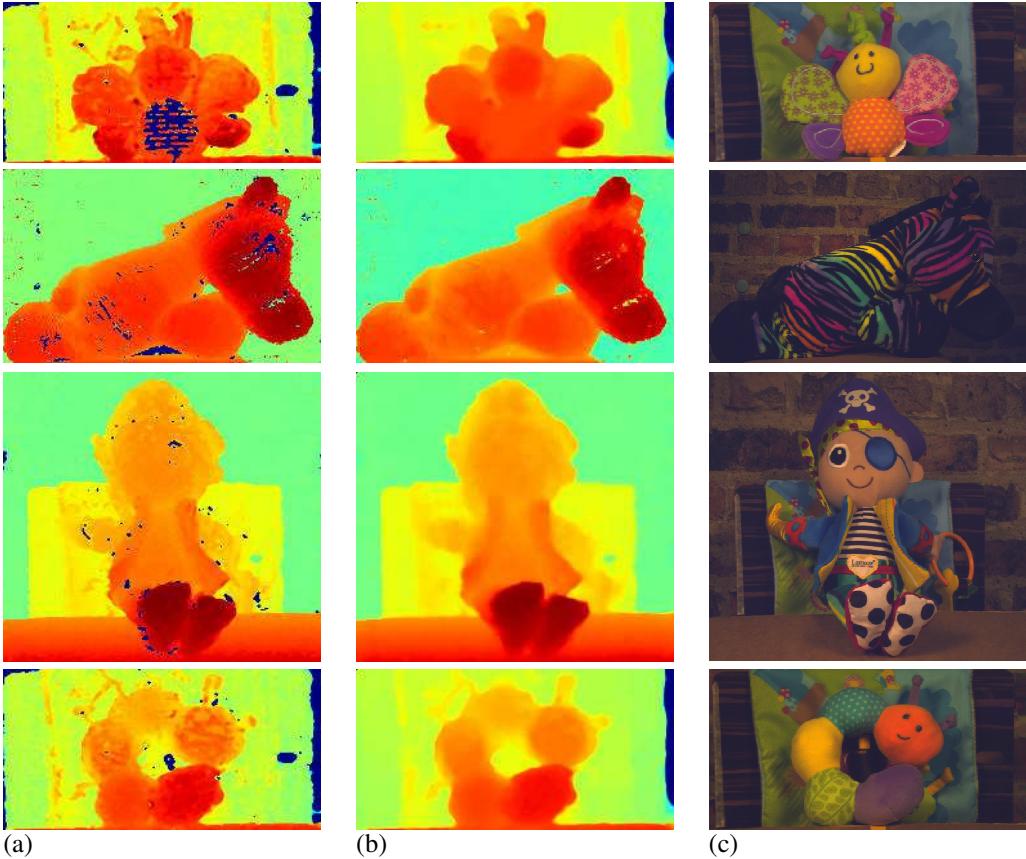


Figure 10. The left column shows the merged disparity maps. The central column shows the result after the applied second order total variation. The right column shows the RGB reconstruction of the real captured light fields.

since varying camera properties cause illumination variations. Here we have seen that due to the high reliability of the computed disparity, the color reconstruction matches the object boundaries perfectly.

Color-filtered light fields hold a more densely sampled spectra as it is possible for a single camera. Thus, an advanced analysis in the spectral domain is possible. A continuation of this work will be the comparison to other existing depth estimation methods and a transfer of the current approach to camera arrays systems for capturing color-filtered light fields. Possible applications are in food surveillances,

e.g to distinguish different cheeses by comparing their spectra, or to determine if food is still eatable or not. Another application is the study about defects in skin, where bacteria or chemical changes influencing the color spectra. It also opens the possibility to use the structure tensor for other types of heterogeneous light fields like for light fields with applied polarization filters to obtain BRDF information. The large variety of new application scenarios and the fast processing to achieve depth information makes it also suitable for movie analysis in near future.

References

- [1] E. Adelson and J. Bergen. The plenoptic function and the elements of early vision. *Computational models of visual processing*, 1, 1991.
- [2] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proc. International Conference on Computer Vision*, pages 433–438, 1987.
- [3] R. Bolles, H. Baker, and D. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, pages 7–55, 1987.
- [4] 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.
- [5] M. Diebold and B. Goldluecke. Epipolar Plane Image Refocusing for Improved Depth Estimation and Occlusion Handling. In *Vision, Modeling and Visualization Workshop VMV*, 2013.
- [6] A. Geiger, M. Roser, and R. Urtasun. Efficient Large-Scale Stereo Matching. In *Asian Conf. on Computer Vision*, 2010.
- [7] T. Georgiev, Z. Yu, A. Lumsdaine, and S. Goma. Lytro camera technology: theory, algorithms, performance analysis. In *IS&T/SPIE Electronic Imaging*, volume 8667, pages 1J1–1J10. International Society for Optics and Photonics, 2013.
- [8] A. Gershun. The Light Field. *J. Math. and Physics*, 18:51–151, 1936.
- [9] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The Lumigraph. In *Proc. SIGGRAPH*, pages 43–54, 1996.
- [10] Y. S. Heo, K. M. Lee, and S. U. Lee. Mutual Information-based Stereo Matching Combined with SIFT Descriptor in Log-chromaticity Color Space. In *Proc. Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [11] B. Jähne. *Digital Image Processing*. Springer, 2005.
- [12] S. B. Kang, R. Szeliski, and J. Chai. Handling occlusions in dense multi-view stereo. In *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), with CD-ROM, 8-14 December 2001, Kauai, HI, USA*, pages 103–110, 2001.
- [13] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. Gross. Scene Reconstruction from High Spatio-Angular Resolution Light Field. In *Proc. SIGGRAPH*, 2013.
- [14] F. Lenzen, F. Becker, and J. Lellmann. Adaptive Second-Order Total Variation: An Approach Aware of Slope Discontinuities. In *Proceedings of the 4th International Conference on Scale Space and Variational Methods in Computer Vision SSVM*, volume 7893 of *LNCS*, pages 61–73. Springer, 2013.
- [15] F. Lenzen, H. Schäfer, and C. Garbe. Denoising Time-Of-Flight Data with Adaptive Total Variation. In *Proceedings ISVC*, pages 337–346. Springer, 2011.
- [16] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. SIGGRAPH*, pages 31–42, 1996.
- [17] R. Li, Z. Bing, and L. Ming. A New Three-Step Search Algorithm for Block Motion Estimation. In *IEEE Trans. Circuits And Systems For Video Technology*, 2004.
- [18] M. Mühlbach and T. Aach. A Theory for Multiple Orientation Estimation. In H. Bischof and A. Leonardis, editors, *Proceedings European Conference on Computer Vision 2006*, number 3952 in *LNCS*, pages (II) 69–82. Springer, 2006.
- [19] C. Perwass and L. Wietzke. The Next Generation of Photography, 2010. www.raytrix.de.
- [20] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. Global Solutions of Variational Models with Convex Regularization. *SIAM Journal on Imaging Sciences*, 2010.
- [21] M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta. A standard default color space for the internet. <http://www.w3.org/Graphics/Color/sRGB.html>, 1996.
- [22] S. Tominaga. Spectral imaging by a multichannel camera. In *Proc. SPIE 3648, Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts IV*, 38, 1998.
- [23] S. Wanner, J. Fehr, and B. Jähne. Generating EPI representations of 4D Light fields with a single lens focused plenoptic camera. *Advances in Visual Computing*, pages 90–101, 2011.
- [24] S. Wanner and B. Goldluecke. Reconstructing Reflective and Transparent Surfaces from Epipolar Plane Images. In *Proc. German Conference on Pattern Recognition*, 2013.
- [25] S. Wanner and B. Goldluecke. Variational Light Field Analysis for Disparity Estimation and Super-Resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013.