

Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression

Daniel Kiesel, Thomas Richter, Jürgen Seiler, André Kaup; Chair of Multimedia Communications and Signal Processing; Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Abstract

Most digital cameras today employ Bayer Color Filter Arrays in front of the camera sensor. In order to create a true-color image, a demosaicing step is required introducing image blur and artifacts. Special sensors like the Foveon X3 circumvent the demosaicing challenge by using pixels lying on top of each other. However, they are not commonly used due to high production cost and low flexibility. In this work, a multi-color multi-view approach is presented in order to create true-color images. Therefore, the red-filtered left view and the blue-filtered right view are registered and projected onto the green-filtered center view. Due to the camera offset and slightly different viewing angles of the scene, object occlusions might occur for the side channels, hence requiring the reconstruction of missing information. For that, a novel local linear regression method is proposed, based on disparity and color similarity. Simulation results show that the proposed method outperforms existing reconstruction techniques by on average 5 dB.

Introduction

Humans are used to perceive the world in color. Consequently, they expect digital images to be recorded and displayed in color, as well. Typically, Color Filter Arrays (CFAs) are employed in front of the camera sensor to allow for colored images. Bayer CFAs [1] in particular, describe a characteristic pattern of red, green, and blue filter elements. As each filter element is assigned to one pixel, for each pixel only one color channel is available. Thus, the missing color channels have to be estimated in order to obtain a true-color RGB image. This process is called demosaicing and introduces image blur and artifacts [2]. Extensive research has been conducted in order to find demosaicing techniques that minimize the quality degradation to fulfill the increasing requirements for image quality [3], [4]. Alternatively, special sensors like the Foveon X3 have been developed, circumventing the demosaicing step [3]. But, those sensors are expensive and not widely spread in the consumer market. Regardless of using standard Bayer patterns or special sensors, the spectral sensitivity is fixed, limiting the number of potential applications.

In contrast, in multi-spectral imaging, the required spectral sensitivity highly depends on the chosen application [5], [6]. For that, filter wheel based systems [7] could increase flexibility but are restricted to static scenes.

Recently, the authors in [8] proposed to use a multi-

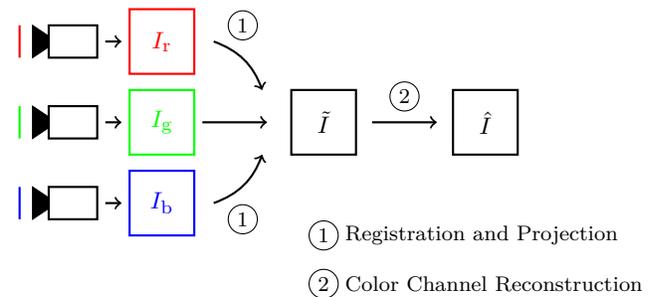


Figure 1: System overview.

view system in the context of polarization imaging for industrial inspection. By using multiple off-the-shelf components equipped with different polarization filters, a superior image quality has been demonstrated while maintaining flexibility and avoiding special purpose cameras.

In this work, a novel and flexible framework for multi-color multi-view image acquisition is proposed. For that, an array with three standard digital grayscale cameras is used. In front of each camera, either a red, green, or a blue color filter is located. The proposed setup provides various benefits. Compared to Bayer filtering, no demosaicing step is required and the full spatial resolution is preserved. In addition, the system is flexible and by far not limited to RGB color imaging. Replacing the color filters in front of the cameras allows to easily adapt to other applications, e.g. in the area of multi-spectral imaging.

In a first step, the side views are projected onto the center view. Then, for image regions where not all three color channels are available after the projection, the missing color information is reconstructed exploiting the local neighborhood and color channel correlations.

The rest of the paper is structured as follows. After explaining the overall system in the next section, the paper focuses on the registration and the proposed color channel reconstruction method. Then, simulation results are given. The paper finally concludes with the last section.

System Overview

The proposed setup can be seen in Figure 1. Three cameras are positioned next to each other and in front of each camera a different filter, corresponding to one of the three primary colors, is mounted. As a result, three images

I_r , I_g , and I_b can be obtained, each containing all pixels for the corresponding red, green, and blue color channel, respectively. The main objective is to project the left image containing the red channel and the right image containing the blue channel onto the center view in order to create the RGB image \tilde{I} . The green channel is chosen as the center view, as it contains the most luminance information and because green wavelengths are between blue and red wavelengths. As the cameras cannot be placed in exactly the same position, each camera has a slightly different viewpoint entailing two challenges. First, image registration is necessary to find corresponding pixels in the different views before projecting them. This task is challenging because it does not resemble a global transformation problem but a local one, as objects closer to the camera are subject to bigger translation than background objects. After performing the projection step, the resulting image \tilde{I} comprises pixels with either one, two, or three color channels. The center view, corresponding to the green channel, is always available. However, depending on the scene geometry, projected pixels from the side views might be occluded in the center view. Thus, secondly, in order to obtain the true-color RGB image \hat{I} , the occluded areas have to be reconstructed using a proper color channel reconstruction technique.

Image registration

In general, image registration is the task of aligning different images of the same scene. For rectified multi-view images, the registration can be achieved in terms of disparity estimation. In case of rotations or two-dimensional translations, image rectification has to be applied first. According to [9], the disparity estimation chain consists of the sub-tasks matching cost computation, cost aggregation, disparity computation, and disparity refinement.

For cost computation, different matching metrics can be applied. On the one hand, there are intensity-based metrics, such as the sum-of-absolute differences or the sum-of-squared differences. However, since this work deals with the registration of different color components, intensity-based metrics cannot be used. On the other hand, there are structure-based metrics, trying to search for point correspondences by comparing the structure of their surrounding blocks. In this work, the zero-mean normalized cross correlation (ZNCC) is used for cost computation. In the following, the calculation is described for matching between the red and the green channel. Note that the calculation for the blue and the green channel is equivalent. For a pixel position \mathbf{p} and a disparity level d , the ZNCC value $\text{ZNCC}(\mathbf{p}, d)$ is defined according to

$$\text{ZNCC}(\mathbf{p}, d) = \frac{\sum_{\mathbf{q} \in N_{\mathbf{p}}} (I_r(\mathbf{q}) - \bar{I}_r(\mathbf{p})) (I_g(\mathbf{q} - (d, 0)) - \bar{I}_g(\mathbf{p} - (d, 0)))}{\sqrt{\sum_{\mathbf{q} \in N_{\mathbf{p}}} (I_r(\mathbf{q}) - \bar{I}_r(\mathbf{p}))^2 \sum_{\mathbf{q} \in N_{\mathbf{p}}} (I_g(\mathbf{q} - (d, 0)) - \bar{I}_g(\mathbf{p} - (d, 0)))^2}} \quad (1)$$

where \bar{I}_r and \bar{I}_g denote the mean values of the current neighborhood $N_{\mathbf{p}}$ and \mathbf{q} is a pixel position within this

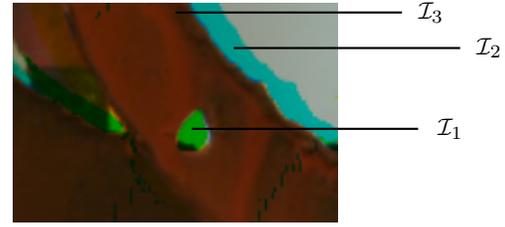


Figure 2: Illustration of the pixel sets \mathcal{I}_1 , \mathcal{I}_2 , and \mathcal{I}_3 .

neighborhood. The result of the cost computation is a matrix comprising the costs for all possible disparities within a pre-defined search range for every pixel. Now, the cost aggregation step aims at getting more reliable and robust costs by smoothing the cost matrix while not aggregating across object boundaries. In this work, a cross-based aggregation method as discussed in [10], [11] is used. In the disparity computation step, simply the best match is chosen from the aggregated cost matrix. Finally, a cross-check and median filtering is used in order to further refine the disparity map.

Color Channel Reconstruction

After applying the disparity map for image projection, three different kinds of pixels exist in the image \tilde{I} . First, there are pixels where all three color channels are present after the projection. This set of pixels does not require any kind of reconstruction thereafter and is called \mathcal{I}_3 . Second, for pixels with two existing color channels \mathcal{I}_2 , the remaining third one has to be reconstructed. Lastly, for pixels \mathcal{I}_1 where only one color channel, i.e., the green channel, is present, the missing two color channels have to be reconstructed. For a sample projection result, Figure 2 illustrates the pixel sets \mathcal{I}_1 , \mathcal{I}_2 , and \mathcal{I}_3 . In order to use the maximum amount of information, the set \mathcal{I}_2 is considered first. Subsequently, those results can be used to improve the reconstruction results for \mathcal{I}_1 .

In literature, various approaches have been already developed. In general, interpolation or image inpainting approaches can be used in order to reconstruct the missing information. The Frequency Selective Extrapolation (FSE) is a block-based iterative method for reconstructing lost areas in images [12]. It has been successfully applied in a wide range of applications, ranging from difference image extrapolation [13], over image resampling [14], to the reconstruction of synthesized high-frequency content [15]. The basic assumption is that image signals can be represented sparsely in the frequency domain. For the unknown pixels within each block, FSE generates a model based on the available pixels in the support area as superposition of 2-D Fourier basis functions.

Alternatively, averaging was employed in the Mars Orbiter Camera (MOC) which comprised two narrow-band sensors in red and blue wavelengths for color vision [16]. The approach assumes a linear dependency between the color channels. Thus, for the scenario considered in this work, the red or blue color channel can be reconstructed

using a linear extrapolation of the given information. However, averaging is a pixel-based approach, independent of the remaining image. Contrarily, linear regression [16] tries to find a global transformation matrix based on the \mathcal{I}_3 pixels where all three color channels exist in the regarded image. This is done by minimizing the residual squared error for the given data. As an example, the goal is to find a transformation matrix M_b for reconstructing the blue channel, such that

$$\mathbf{b} \cong [\mathbf{r} \ \mathbf{g}] \cdot M_b, \quad (2)$$

where \mathbf{r} , \mathbf{g} , and \mathbf{b} denote the column values of the corresponding red, green, and blue pixels, respectively. The resulting 2×1 transformation matrix can be obtained by resolving to M_b using the pseudo-inverse denoted by $^+$ as

$$M_b = [\mathbf{r} \ \mathbf{g}]^+ \cdot \mathbf{b}. \quad (3)$$

The transformation matrix for the red channel can be obtained analogously.

In summary, various reconstruction methods already exist. While FSE is capable of fully reconstructing images for \mathcal{I}_1 and \mathcal{I}_2 regions, the correlation between different channels is neglected. Averaging and linear regression do account for channel correlation but are only capable of reconstructing \mathcal{I}_2 pixels. Furthermore, the results are inaccurate due to estimates over the entire image.

Proposed: Disparity and Color Similarity-based Local Linear Regression

Both, linear regression and averaging are based on a global rule set or global parameters that do not change for different locations and neighborhoods. In order to account for that locality and improve the accuracy of the reconstruction, a novel approach called Disparity and Color Similarity-based Local Linear Regression (DCSLLR) is proposed. Initially, the reconstruction of \mathcal{I}_2 regions is regarded before adjusting the algorithm to account for \mathcal{I}_1 regions as well.

The method extends linear regression by three rules, namely *local proximity*, *color similarity*, and *disparity similarity*. Each of the rules creates a logical mask. Combining the resulting masks using logical conjunction results in a restricted number of pixels that can be used for linear regression. In the following, the three rules are described in detail.

Local proximity

As mentioned before, linear regression uses the pixel values containing valid information for all three color channels over the entire image. Local proximity restricts the area by

$$\max(|p_x - q_x|, |p_y - q_y|) \leq l_{sr}, \quad (4)$$

where l_{sr} denotes the maximum search range. p_x and p_y represent x and y coordinates of the point \mathbf{p} , respectively. The same logic applies to the arbitrary image point \mathbf{q} . For a given point \mathbf{p} , the resulting mask can be seen in Figure 3a, forming a quadratic search area around the currently processed pixel.

Color similarity

Linear regression can only be used sensibly when only one color channel is missing. As a result, color information of the two remaining channels is still prevalent and can further be used to refine the selection of pixels for linear regression. Therefore, only those points comprising three color channels that satisfy

$$\begin{aligned} |I_r(\mathbf{p}) - I_r(\mathbf{q})| &\leq c_{sr} \quad \text{and} \\ |I_g(\mathbf{p}) - I_g(\mathbf{q})| &\leq c_{sr} \end{aligned} \quad (5)$$

are selected in case of a missing blue channel or

$$\begin{aligned} |I_b(\mathbf{p}) - I_b(\mathbf{q})| &\leq c_{sr} \quad \text{and} \\ |I_g(\mathbf{p}) - I_g(\mathbf{q})| &\leq c_{sr} \end{aligned} \quad (6)$$

in case of a missing red channel. Here, $I_r(\mathbf{p})$, $I_g(\mathbf{p})$, and $I_b(\mathbf{p})$ denote the values of the red, green, and blue channel for the currently processed pixel \mathbf{p} , respectively. The maximum color difference that a point \mathbf{q} within the image can have is denoted by c_{sr} , thus only validating pixels with similar values for the given color channels. The resulting color similarity mask can be seen in Figure 3b.

Disparity similarity

Given the camera setup presented in Figure 1 and recalling the image projection using disparity maps, whenever a point in the resulting image has all three color channels both red and blue channel were mapped there using the corresponding disparity maps. Ideally, the disparity values for the red and blue channel are identical in that case, given that the center view is located exactly half way between left and right view. Therefore, if only two color channels exist for a given point, it is assumed that the disparity d of the missing channel is equal to the disparity of the existent projected channel. Once the disparity is determined, the valid disparity mask can be obtained. Initially, the right-center disparity is warped to the center view as

$$\tilde{D}_R(\mathbf{p} + D_R(\mathbf{p})) = D_R(\mathbf{p}), \quad (7)$$

where \tilde{D}_R denotes the warped result. The left-center disparity D_L can be mapped analogously. Thereafter, the disparity map is split into d_{sr} slices resulting in a $M \times N \times d_{sr}$ matrix where each slice contains a mask of the pixels corresponding to the respective disparity. For a given disparity d , all slices within the disparity range d_{dr} are combined using a logical disjunction, ultimately resulting in two disparity masks corresponding to valid values for left-center and right-center mapping, respectively. Combining the two maps by logical conjunction finally yields the disparity similarity mask. Note that all values covered by that mask will comprise three color channels in the projected image, as both left-to-center and right-to-center mapped values are existent. The resulting mask is shown in Figure 3c together with the final mask in Figure 3d that can be obtained by combining the local proximity mask, the color similarity mask, and the disparity similarity mask using logical conjunction. As the number of entries in the final mask can vary, a lower threshold of τ_{tm} is defined and set to 20. If

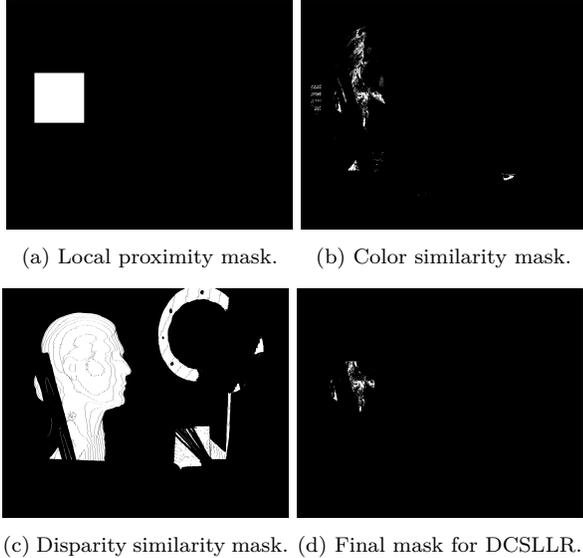


Figure 3: Visualization of the DCSLLR masks for linear regression.

the number of pixels is below τ_m the disparity similarity mask is removed. If the resulting disparity mask is still too restrictive, linear regression considering the entire image is applied for the processed pixel. Besides choosing $\tau_m = 20$, extensive parameter exploration yielded $d_{dr} = 6$, $l_{sr} = 90$, and $c_{sr} = 15$ showing the best results, which is why these parameters are applied in the evaluation.

Proposed: Color Similarity-based Local Median Filtering

So far, using DCSLLR, only pixels containing values for two color channels could be reconstructed. As mentioned before, however, there might also be pixels where only the green channel exists. Therefore, a technique for reconstructing the missing two color channels called Color Similarity-based Local Median Filtering (CSLMF) is proposed.

Mainly, there are two differences that need to be addressed. First, disparity information cannot be used, as none of the side views is projected onto the regarded pixel. Second, only one channel, namely the green channel, can be used for color similarity comparison as

$$|I_g(\mathbf{p}) - I_g(\mathbf{q})| \leq c_{sr_2} \quad (8)$$

where c_{sr_2} denotes a more restricted color threshold that is set to 5. If no pixels are found for that threshold it is incrementally extended by a step size of 5. For estimating the local proximity mask, no changes are applied compared to DCSLLR. Combining the one channel color similarity mask with the local proximity mask via logical conjunction yields the resulting mask. Then, the median red and blue values of the marked pixels are chosen and assigned to the pixel to be reconstructed.

Simulation Results

For simulation, the multi-view datasets *Art*, *Books*, *Dolls*, *Laundry*, *Moebius*, and *Reindeer* have been taken [17]. Each dataset consists of 7 views denoted by the range (0..6) taken under three different illuminations and with three different exposures. For the evaluation, only standard illumination and an exposure time of 1000 ms is considered. In addition, two distance scenarios are distinguished. Small camera distance denotes the usage of view 2, 3, and 4 respectively, whereas large camera distance refers to view 1, 3, and 5. In both cases view 3 denotes the green filtered center view onto which the outer views are projected. In order to compare different projection and reconstruction techniques, the image quality is measured objectively using the peak signal-to-noise ratio (PSNR).

After conducting the projection step, both \mathcal{I}_2 and \mathcal{I}_1 regions have to be reconstructed. For region \mathcal{I}_2 , the evaluated algorithms include FSE, averaging (AVG), linear regression (LR), and the proposed DCSLLR. For region \mathcal{I}_1 , no results can be given for AVG and LR, since these approaches require two available color channels. Thus, for region \mathcal{I}_1 , results are only given for FSE and the proposed CSLMF.

First, the discussed color channel reconstruction methods are evaluated for the projection results obtained by ground truth disparity maps, which are part of the database. By doing so, the reconstruction quality can be judged without considering any influence from the used disparity estimation method. Both, for \mathcal{I}_1 and \mathcal{I}_2 pixels, Table 1 and Table 2 summarize the PSNR values for all datasets and both considered multi-view setups. For the small camera distance and \mathcal{I}_2 pixels, the proposed DCSLLR achieves an average PSNR value of 30.1 dB and a mean gain of 5 dB over basic LR. For the \mathcal{I}_1 pixels, the proposed CSLMF outperforms FSE by 0.8 dB, on average. For the large camera distance, the reconstruction results get worse due to larger connected loss areas. However, reaching an average PSNR value of 28.4 dB, the proposed DCSLLR still outperforms all competitive approaches. For the \mathcal{I}_1 pixels, a clear performance loss can be observed for FSE. In contrast, the reconstruction quality of CSLMF keeps constant, leading to an average gain of 3.9 dB over FSE.

Figure 4 compares the \mathcal{I}_2 reconstruction quality of FSE, AVG, LR, and the proposed DCSLLR for a detail of the *Art* dataset and the small camera distance. For further comparison, the figure shows the projected image, prior to reconstruction and the original image on the outer left and right side, respectively. While FSE partially reconstructs the missing channels decently, artifacts remain around the object borders. For averaging, the reconstruction results show clearly different colors compared to the original as the linearity assumption between color channels does not hold. Also for LR, artifacts are still visible, resulting from globally estimating the color channel correlations. In contrast, the proposed DCSLLR leads to a clear rise in visual quality, hardly distinguishable from the original image. The \mathcal{I}_1 reconstruction quality is illustrated in Figure 5, again for a detail of the *Art* dataset and the

Table 1: **Color channel reconstruction** PSNR results in dB for image regions \mathcal{I}_2 and \mathcal{I}_1 for the **small** camera distance and ground truth disparity maps.

	\mathcal{I}_2 PSNR				\mathcal{I}_1 PSNR	
	FSE	AVG	LR	DCSLLR	FSE	CSLMF
<i>Art</i>	23.0	23.9	26.3	30.6	20.3	27.2
<i>Books</i>	25.1	27.8	28.2	30.9	25.6	25.0
<i>Dolls</i>	22.2	21.6	22.7	27.4	23.3	22.6
<i>Laundry</i>	24.9	20.9	22.1	28.8	24.1	24.0
<i>Moebius</i>	25.9	20.2	23.0	29.0	25.8	22.5
<i>Reindeer</i>	26.1	24.3	28.3	33.7	21.5	23.8
Average	24.5	23.1	25.1	30.1	23.4	24.2

Table 2: **Color channel reconstruction** PSNR results in dB for image regions \mathcal{I}_2 and \mathcal{I}_1 for the **large** camera distance and ground truth disparity maps.

	\mathcal{I}_2 PSNR				\mathcal{I}_1 PSNR	
	FSE	AVG	LR	DCSLLR	FSE	CSLMF
<i>Art</i>	21.0	23.7	26.1	29.9	17.5	27.4
<i>Books</i>	22.7	26.8	27.2	27.7	22.0	24.7
<i>Dolls</i>	20.5	20.8	21.4	26.1	18.2	22.5
<i>Laundry</i>	21.6	20.9	22.0	26.7	21.2	22.9
<i>Moebius</i>	24.9	19.6	21.8	26.5	22.5	22.0
<i>Reindeer</i>	25.3	24.0	27.6	32.5	20.6	25.8
Average	22.7	22.6	24.3	28.4	20.3	24.2

small camera distance. The figure compares the results of FSE and the proposed CSLMF. Again, FSE has difficulties with reconstructing large connected loss areas, while CSLMF leads to a convincing rise in visual quality.

As a second evaluation, Table 3 gives the achieved reconstruction quality for the proposed method in case of estimated disparity maps. For estimation, the above discussed image registration chain has been used. On average, a PSNR value of 32.2 dB, calculated over the whole images, has been achieved. Thus, compared to the projection with ground truth disparities, the estimated disparity information leads to an average performance loss of 1.9 dB. For that, a visual comparison is given in Figure 6. For two image details, the figure shows the reconstruction quality in case of ground truth and estimated disparity maps. In the top row, the reconstructed images are very similar showing that the proposed method can be also applied in case of estimated disparity information. However, as shown in the bottom image row, in case of inaccurate disparity information, projection errors directly influence the color channel reconstruction performance.

Conclusion

In this paper, a novel color channel reconstruction method has been proposed in the context of multi-color multi-view images. For that, a scene is recorded from multiple perspectives using off-the-shelf grayscale cameras. For color vision, different color filters are located in front of the cameras, providing a higher flexibility compared to common Color Filter Arrays. In order to map the side views onto the target perspective, a disparity estimation method has been discussed, being robust against different spectral ranges. After projection, some image parts are still un-

Table 3: **Color channel reconstruction** PSNR results in dB in case of ground truth (GT) and estimated (ZNCC) disparity maps and the **small** camera distance.

	GT	ZNCC
<i>Art</i>	34.5	30.3
<i>Books</i>	33.3	33.1
<i>Dolls</i>	33.8	33.1
<i>Laundry</i>	33.2	30.8
<i>Moebius</i>	34.9	33.5
<i>Reindeer</i>	35.1	32.4
Average	34.1	32.2

known due to object occlusions. For the required color channel reconstruction, a local linear regression method has been developed, based on disparity and color similarity. Simulation results have shown that the proposed method leads to a clear rise in visual as well as objective quality. Compared to basic linear regression, an average gain of 5 dB has been achieved.

References

- [1] B.E. Bayer, "Color imaging array", U.S. Patent 3 971 065, 1976
- [2] X. Li, B. Gunturk, L. Zhang, "Image demosaicing: a systematic survey", Proc. SPIE, vol. 6822, pp. 68221J-68221J-15, 2008
- [3] O. Losson, L. Macaire, Y. Yang, "Comparison of color demosaicing methods", Advances in Imaging and Electron Physics, Elsevier, vol. 162, pp.173-265, 2010
- [4] J. Wu, R. Timofte, L. Van Gool, "Demosaicing Based on Directional Difference Regression and Efficient Regression Priors", IEEE Transactions on Image Processing, vol. 25, no. 8, pp. 3862-3874, Aug., 2016
- [5] S. V. Patwardhan, A. P. Dhawan, "Multi-spectral imaging and analysis for classification of melanoma", Int. Conf. of the IEEE Engineering in Medicine and Biology Society, pp. 503-506, Sept., 2004
- [6] M. Axelsson, O. Friman, I. Johansson, M. Nordberg, H. Östmark, "Detection and classification of explosive substances in multi-spectral image sequences using linear subspace matching", IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 3492-3496, May, 2013
- [7] J. Brauers, N. Schulte, T. Aach, "Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algorithms", IEEE Transactions on Image Processing, vol. 17, no. 12, pp. 2368-2380, Dec., 2008
- [8] T. Richter, C. Saloman, N. Genser, K. Kasnakli, J. Seiler, A. Nowak, A. Kaup, M. Schöberl, "Sequential Polarization Imaging Using Multi-View Fusion", IEEE Int. Workshop on Signal Processing Systems, Lorient, France, Oct. 2017
- [9] D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms", Int. Journal on Computer Vision, vol. 47, no. 1-3, pp. 7-42, Apr., 2002
- [10] K. Zhang, J. Lu, G. Lafruit, "Cross-Based Local Stereo Matching Using Orthogonal Integral Images", IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 7, pp. 1073-1079, July, 2009
- [11] M. Bätz, T. Richter, J. Garbas, Anton Papst, J. Seiler, A.

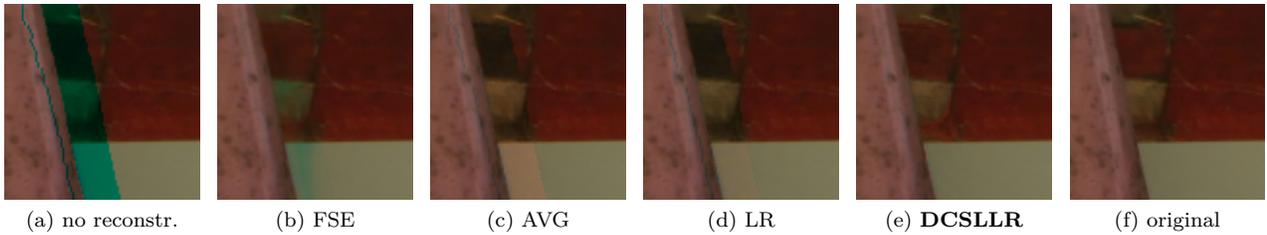


Figure 4: Visual comparison of different reconstruction techniques for \mathcal{I}_2 pixels using ground truth disparity maps.

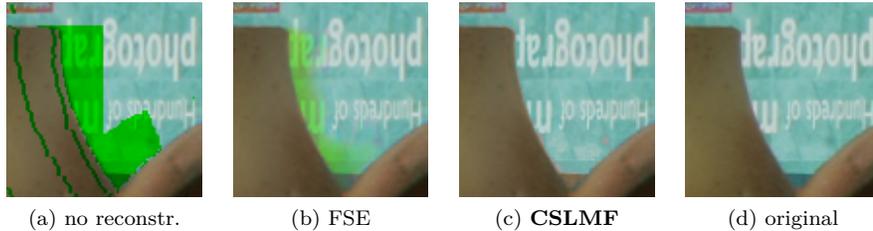


Figure 5: Visual comparison of different reconstruction techniques for \mathcal{I}_1 pixels using ground truth disparity maps.

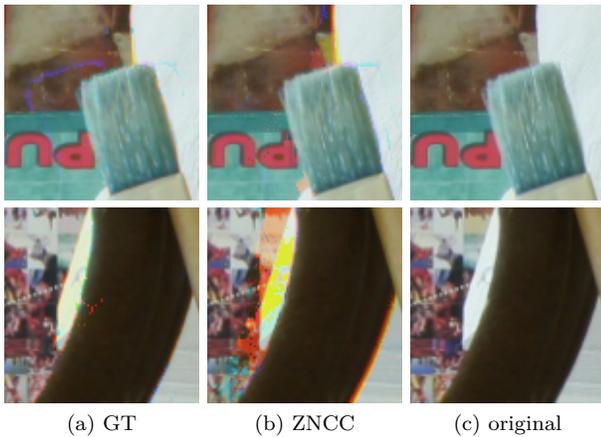


Figure 6: Visual comparison of the proposed reconstruction method using ground truth and estimated disparity maps.

Kaup, "High Dynamic Range Video Reconstruction from a Stereo Camera Setup", *Signal Processing: Image Communication*, vol. 29, no. 2, pp. 191-202, Sep. (2013).

- [12] Jürgen Seiler, A. Kaup, "Complex-Valued Frequency Selective Extrapolation for Fast Image and Video Signal Extrapolation", *IEEE Signal Processing Letters*, vol. 17, no. 11, pp. 949-952, Nov., 2010
- [13] T. Tröger, A. Kaup, "Difference Image Extrapolation for Spectral Completion in Inter-Sequence Error Concealment", *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 1129-1132, May, 2011
- [14] J. Seiler and M. Jonscher and M. Schöberl and A. Kaup, "Resampling Images to a Regular Grid From a Non-Regular Subset of Pixel Positions Using Frequency Selective Reconstruction", *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4540-4555, Nov., 2015
- [15] T. Richter and J. Seiler and W. Schnurrer and A. Kaup, "Robust Super-Resolution for Mixed-Resolution Multiview Image plus Depth Data", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 5, May, 2016

- [16] V. C. Cardei, "From Dichromatic to Trichromatic Images: Implications for Image Recovery and Visualization", *Conf. on Image Processing, Image Quality and Image Capture Systems*, pp. 359-363, Apr., 1999
- [17] H. Hirschmüller and D. Scharstein, "Evaluation of Cost Functions for Stereo Matching", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1-8, June, 2007

Author Biography

Daniel Kiesel studied industrial engineering with focus on informations and communications systems at FAU Erlangen-Nürnberg, Germany. Writing his master's thesis at the Chair of Multimedia Communications and Signal Processing, he received his master's degree in 2017.

Thomas Richter is a researcher at the Chair of Multimedia Communications and Signal Processing at FAU Erlangen-Nürnberg, Germany. He received his diploma and doctoral degree in 2011 and 2016, respectively. His current research interests include image and video signal processing, multi-model multi-view image processing, and related fields.

Jürgen Seiler is senior scientist at the Chair of Multimedia Communications and Signal Processing at FAU Erlangen-Nürnberg, Germany. There, he also received his doctoral degree in 2011 and his diploma degree in Electrical Engineering, Electronics and Information Technology in 2006. His research interests include image and video signal processing, signal reconstruction and coding, signal transforms, and linear systems theory.

André Kaup received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from RWTH Aachen, Germany, in 1989 and 1995, respectively. He joined the Networks and Multimedia Communications Department, Siemens, Munich, Germany, in 1995 and became Head of the Mobile Applications and Services Group in 1999. Since 2001 he has been Head of the Chair of Multimedia Communications and Signal Processing, FAU Erlangen-Nürnberg, Germany. His research interests include image and video signal processing and coding, and multimedia communication.