https://doi.org/10.2352/issn.2169-2629.2021.29.288

Joint demosaicing of colour and polarisation from filter arrays

Alexandra Spote[†] *,Pierre-Jean Lapray*[†] *, Jean-Baptiste Thomas*[⋆] *, Ivar Farup*[⋆]
[†] *Université de Haute-Alsace, Institut de Recherche en Informatique, Mathématiques, Automatique et Signal, EA 7499, 68093 Mulhouse CEDEX, France*
[⋆] *Norwegian University of Science and Technology, Department of Computer Science; Gjøvik, Norway*

Abstract

This article considers the joint demosaicing of colour and polarisation image content captured with a Colour and Polarisation Filter Array imaging system. The Linear Minimum Mean Square Error algorithm is applied to this case, and its performance is compared to the state-of-the-art Edge-Aware Residual Interpolation algorithm. Results show that the LMMSE demosaicing method gives statistically higher scores on the largest tested database, in term of peak signal-to-noise ratio relatively to a CPFA-dedicated algorithm.

Introduction

From the development of Colour Filter Arrays (CFA) and Spectral Filter Arrays imaging (SFA) systems [7], we observe a tendency towards its generalisation to several joint optical modalities, e.g. polarisation and spectral. We refer to General Filter Arrays (GFA) the most general case. One particular instance of GFA is the Colour and Polarisation Filter Array (CPFA), that has one commercial implementation with the SONY IMX250 MYR [15].

CPFA, like GFA, needs their spatial resolution to be reconstructed in order to avoid misinterpretation in band registration for image analysis. Like for the Bayer colour cameras, a typical way to reconstruct the resolution and register multimodal data is demosaicing. Extensive works were conducted on colour images [8]. More recently, several works were conducted to reconstruct spectral data [16, 10] and polarisation data [9]. The most recent work to date reconstructs both colour and polarisation [12] from the Sony IMX250 MYR sensor. This sensor combines three colour filters ($c = r, g, b$) arranged in a Quad Bayer [13] spatial arrangement, and four polarisation angles of analysis equally-distributed between 0° and 180° ($\theta = 0°, 45°, 90°, 135°$). The sensor assembly is shown in Figure 1.

Demosaicing algorithms are a specific instance of super-resolution algorithm that benefits from several assumptions, such as spatial correlations and spectral correlations. In the case of joint colour and polarisation acquisition, we observed several modality correlations [4]. Thus, a linear model would permit to estimate the values of each of the missing bands at a pixel location.

In this article, we apply the Linear Minimum Mean Square Error (LMMSE) demosaicing algorithm to the case of colour and polarisation data. LMMSE was successfully used to demosaic colour images with results competitive to the state-of-the-art [3]. The model is introduced in the next Section. Then, we apply this method on three independent databases and compare the results with the state-of-the-art Edge-Aware Residual Interpolation algorithm (EARI) [12]. The comparison is performed by sensor band, but also on standardised data representation, namely intensity RGB, and the Stokes-vector elements S_0, S_1 and S_2 for the polarisation. Results demonstrate that the LMMSE method performs significantly better than the EARI, but the improvement is somewhat marginal, of an order of 1dB in average. Visual inspection of a few images show that the LMMSE provides better results. It is to be noted that the data used in our experiment is limited, so this tendency needs to be confirmed on a large dataset when available.

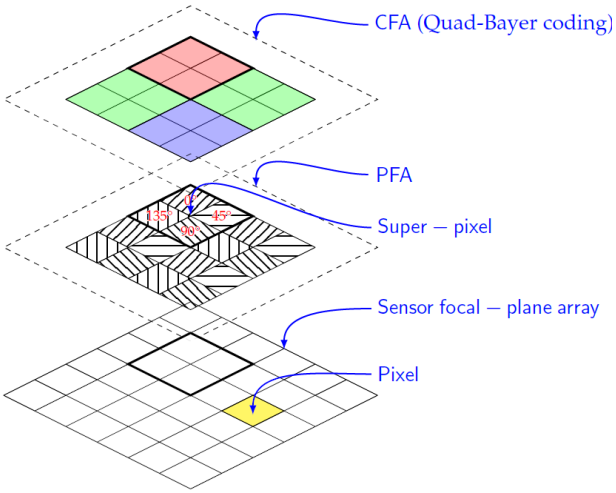


Figure 1: The Color Polarisation Filter Array design of the SONY IMX250 MYR sensor [15]. Each periodic pattern of 4×4 pixels (called a superpixel) senses the light through a Polarisation Filter Array and a Color Filter Array.

LMMSE demosaicing of CPFA

In this section, we apply the LMMSE algorithm for the case of CPFA. Be $\boldsymbol{Y}$ the full-resolution image, and $\boldsymbol{X}$ the mosaiced CPFA image. $\boldsymbol{Y}$ is of size PHW, where $P = 12$ is the number of channels, and H and W the height and the width of the image respectively. $\boldsymbol{X}$ is of size HW, and is composed of regular sampling lattice referred as superpixels of size $h \times w$.

Because the LMMSE does not take into account the spatial correlations, instead of working with 2-D images, the algorithm applies a vectorization: the image $\boldsymbol{Y}$ is unfolded by superpixel, giving the matrices $\boldsymbol{y}$ of size $Phw \times \frac{HW}{hw}$, as depicted in Figure 2. The latter shows the unfolding of a single superpixel of $\boldsymbol{Y}$ into a single column in $\boldsymbol{y}$ in the case of a CFA; the same principle applies to

CPFA. To compute the other columns y, other superpixels are considered across the lines and columns $\boldsymbol{Y}$ such that the number of columns of $\boldsymbol{y}$ is the number of superpixels of the reference image.

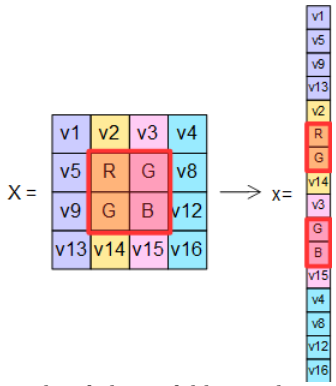


Figure 2: Example of the unfolding columns by columns of a superpixel of 2×2 pixels (red square) with a neighbourhood of 4×4 pixels.

Then, the mosaicing matrix $\boldsymbol{M}$ is manually computed, such that it gives $\boldsymbol{x}$, the unfolding CFA matrix when applied on $\boldsymbol{y}$:

$$\boldsymbol{x} = \boldsymbol{M}\boldsymbol{y} \ , \tag{1}$$

where $\boldsymbol{x}$ is of size $hw \times \frac{HW}{hw}$. The unfolding of $\boldsymbol{y}$ is done superpixel by superpixel, thus the mosaicing matrix $\boldsymbol{M}$ is block-shift invariant.

The aim of demosaicing is to estimate $\hat{\boldsymbol{y}}$ from the observations $\boldsymbol{x}$, such that the estimate image is the most faithful to the reference matrix $\boldsymbol{y}$. This is achieved by the demosaicing matrix $\boldsymbol{D}$, which is a pseudo-inverse of $\boldsymbol{M}$, such that:

$$\hat{y} = \boldsymbol{D}\boldsymbol{x} \ . \tag{2}$$

$\boldsymbol{D}$ is computed with the Wiener filtering approach, which is based on the minimisation of the minimum square error between $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$ on a large dataset of k images. $\boldsymbol{D}$ is given by:

$$\boldsymbol{D} = E_i\{\boldsymbol{y}\boldsymbol{x}^t(\boldsymbol{x}\boldsymbol{x}^t)^{-1}\} \ , \tag{3}$$

where E is the expectation, and $i \in [1,k]$ indexes the image in the database.

To stabilise the solution of $\boldsymbol{D}$, a neighbourhood of size $N_h \times N_w$ is used, such that $N_h N_w > P$. The new matrix $\boldsymbol{y}_1$ contains each superpixel unfolded with its neighbourhood. The neighbourhood can be a sliding or a constant one, both giving same performances if they follow the relation in Equation (4) [1]:

$$\begin{array}{ll} n'_h & = n_h + h + 1 \\ n'_w & = n_w + w + 1 \end{array} \tag{4}$$

where n'_h and n'_w are the sizes of the constant neighbourhood and n_h and n_w the ones of the sliding one.

However, the size of the matrices depends on the size of the neighbourhood. In the following, the matrices that contain neighbour pixels are denoted with index 1. The matrix S_1 is a constant "selection" matrix with zeroes and ones, for removing neighbours from y_1. One obtains the following equations:

$$\left\{ \begin{array}{ll} \boldsymbol{y} & = \boldsymbol{S_1}\boldsymbol{y_1} \ , \\ \boldsymbol{x}_1 & = \boldsymbol{M_1}\boldsymbol{y_1} \ , \\ \hat{\boldsymbol{y}} & = \boldsymbol{D}\boldsymbol{x_1} \ , \\ \boldsymbol{D} & = \boldsymbol{S_1}\boldsymbol{R}\boldsymbol{M_1^t}(\boldsymbol{M_1}\boldsymbol{R}\boldsymbol{M_1^t})^{-1} \ , \end{array} \right. \tag{5}$$

with

$$\boldsymbol{R} = \frac{1}{\frac{HW}{hw}k} E_i\{\boldsymbol{y_1}\boldsymbol{y_1^t}\}, \tag{6}$$

the mean of autocorrelation of $\boldsymbol{y}_1$ over the $\frac{HW}{hw}$ superpixels of each image and over the k images of the database. The matrix $\boldsymbol{R}$ has to be computed only one time for all the images in the database.

Experiment

The experiment is conducted over three existing database of full-resolution spectropolarimetric images:

- Lapray *et al.* [6], identified as DB1, 10 images, 150 MegaBytes (MB) of data,
- Qiu *et al.* [14], DB2, 40 images, 530MB of data,
- Monno *et al.* [12], DB3, 40 images, 1.4GB of data.

The mosaicing process of the image database is applied following the spatial arrangement of Figure 1.

The image in DB3 has been capture with a division-of-amplitude technique for the spectral information (a 3-CCD camera), and a division-of-time technique for the polarisation information (by rotating a polarizer in front of the camera). However, the images available in the databases DB1 and DB2 have been captured with a division-of-focal-plane technique for the spectral information (i.e. with a CFA). As in [4], to mitigate any errors introduced by the colour demosaicing of CFA images, we downsampled all the full-resolution images prior to mosaicing, to reduce their size by a factor of 0.5 (imresize function in Matlab, with bilinear interpolation).

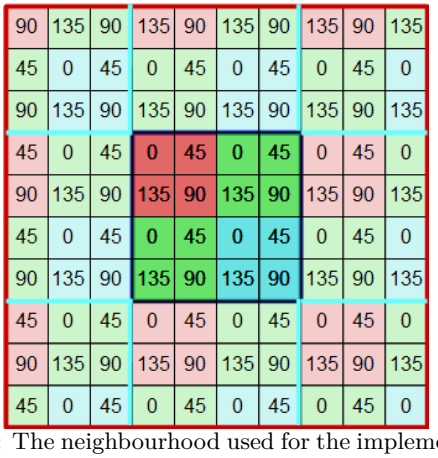


Figure 3: The neighbourhood used for the implementation of LMMSE dedicated to the CPFA. Each superpixel is unfolded with its neighbourhood to construct the matrix $\boldsymbol{y_1}$. Surrounded in dark, we can see the superpixel to be unfolded, whereas the neighbourhood of 10×10 pixels is in red. The groups of pixels surrounded by light blue pertains to the other superpixels of the CPFA.

We demosaic the images using the two methods: our application of the LMMSE method to CPFA [2] and the last demosaicing method dedicated to PFA and CPFA, which is the Edge-Aware Residual Interpolation (EARI) [12].

For LMMSE demosaicing, we choose a neighbourhood of 10×10 pixels, as calculated with Eq. (4). This ad-hoc choice gave the best trade off between performance and computational complexity. It is visualised in Figure 3. The computation of each demosaiced image is done in a leave-one-out manner, i.e. the training procedure of $\boldsymbol{D}$ (and thus $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{R}$) is realised using the remaining data other than that of the current image to demosaic.

For EARI, we employ the Matlab code provided by the authors [12]. It starts from an estimation of the intensity image from polarisation information, i.e. the S_0 Stokes vector component. They exploit the redundancy of information that is inherent to PFA sensors which have four polarisation angles of analysis, such that the intensity S_0 can be estimated in two ways: $S_{0,1} = I_0 + I_{90}$ or $S_{0,2} = I_{45} + I_{135}$. The two estimations are averaged for each of the four spatial directions (north, east, south, and west). This forms the intensity estimations. Then, weights are computed in the four directions from the intensity difference $S_{0,1} - S_{0,2}$, giving a larger weight if the difference is small. Then, they apply a weighting average on the intensity estimations, to obtain the edge-aware guide image. Finally, interpolation of missing values is done by residual interpolation [5]. The same process is applied for each spectral channel to demosaic the polarisation information. We refer the reader to the initial paper [12] for more details.

Discussion

For a given database, we compute the PSNR (using psnr Matlab function) for each channel of each image, and we average them over all the images of the database. As LMMSE results has border effects due to neighbourhoods, the demosaiced and the full resolution images are all reduced by eight lines and eight columns (corresponding to the sizes of two lines and two column of superpixels) before PSNR computation.

The mean PSNR and the standard deviation for EARI and the LMMSE algorithms are given in Table 1.

On DB1, LMMSE and EARI provide similar PSNR results for each of the 12 bands.

On DB2, EARI provides somewhat better results on the red and blue channels but results are comparable on the green channel for both methods. However, on DB2, both methods provide high PSNR values, and less differences between the green channel and the two others. Since this observation is not related to one algorithm, it should be related to the image statistics and content.

On DB3, we observe that LMMSE is performing better than EARI in terms of average PSNR and standard deviation.

Table 2 shows CPSNR (average and standard deviation) by polarisation Stokes component S_0, S_1, and S_2. We make similar observations in the Stokes standard polarisation representation.

On Figure 4, we can see that the LMMSE and EARI perform similarly on DB1, EARI is better for DB2, and LMMSE is in average better than the EARI for DB3.

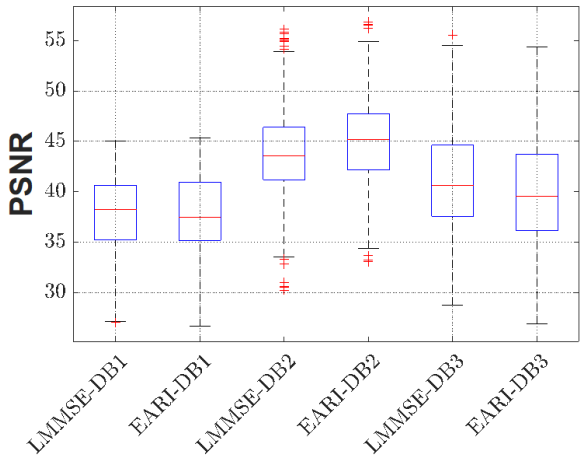


Figure 4: Box plot computed on average data over all channels, and for each method and database. The central red mark indicates the median. The boxes extend from the 25% quartiles to the 75% quartiles, and the whiskers indicate the full ranges of the data excluding the outliers that are shown with red crosses.

These observations are confirmed by a Wilcoxon signed rank test which gives p-values of 0.33, $7.97 \cdot 10^{-66}$, and $5.2 \cdot 10^{-39}$ for DB1, DB2, and DB3 respectively. It is computed over the PSNR of all bands for all images reconstructed by LMMSE in front of EARI.

Generally, from the Table 1, we can observe that the LMMSE method gives statistically better results than the EARI method for DB3, but not for DB2. It can be explain by the importance for LMMSE to have a large quantity of data to train the linear transformation. The PSNR difference can also come from the image statistics and content that seem very different for these two datasets. DB2 exhibits mostly uniform backgrounds and greyish/transparent objects, whereas DB3 contains highly textured and colourful objects, so that LMMSE is specifically good for these kind of data. Moreover, the result for DB3 must be put into perspective with the results obtained in [12] on the same database, where EARI is better than other algorithms by an increment of 0.04dB in average compared to [11]. By looking at Table 3 of [12], LMMSE appears to be by far the best performing for this database (+1.3dB in average compared to EARI).

Figures 5, 6, 7, and 8 show a snapshot example image. On Figures 5, 6, and 7, we compare the 12 bands of the original image and their estimated reconstruction by EARI and LMMSE. In those visualisations, the results appear to be fairly similar, with perhaps a better text reconstruction in the case of LMMSE for some bands. Figure 8 provides the colour image and false colour visualisation of the polarimetric information for each of the spectral bands. We observe on Figure 8a,b,c that EARI is generating aliasing artifacts accordingly to the spatial interpolation used, while the LMMSE shows very little of this artifact. For the polarisation of the green channel, EARI shows some artifacts, while LMMSE demonstrates a rather faithful fidelity to the original.

To conclude, it appears that the generic LMMSE algorithm applied to CPFAs performs just as well as the state-of-the-art algorithm dedicated to CPFA. However, the improvement seems significant when sufficient data is

	DB1 (150MB)				DB2 (530MB)				DB3 (1.4GB)			
	LMMSE		EARI		LMMSE		EARI		LMMSE		EARI	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
$I_{0,R}$	35.25	3.79	**35.49**	4.45	42.98	4.21	**44.65**	4.29	**39.34**	4.15	37.70	4.81
$I_{45,R}$	36.20	3.78	**36.36**	4.10	42.95	4.38	**44.92**	4.24	**39.16**	4.08	37.20	4.78
$I_{90,R}$	36.25	3.76	**36.58**	4.27	43.00	4.44	**45.02**	4.27	**39.76**	4.04	37.97	4.52
$I_{135,R}$	35.34	3.87	**35.47**	4.63	43.10	4.22	**44.91**	4.30	**38.88**	4.00	37.30	4.63
$I_{0,G}$	**40.21**	3.50	39.78	3.93	44.98	4.75	**45.38**	4.74	**44.11**	4.69	42.77	4.83
$I_{45,G}$	**41.32**	3.01	40.72	2.93	45.24	4.51	**46.03**	4.54	**43.37**	4.69	41.71	4.88
$I_{90,G}$	**41.36**	2.97	40.99	2.88	45.25	4.58	**46.18**	4.66	**44.61**	4.54	43.07	4.67
$I_{135,G}$	**40.12**	3.51	39.63	3.97	45.20	4.65	**45.71**	4.73	**42.95**	4.59	41.58	4.87
$I_{0,B}$	36.86	2.69	**37.27**	2.96	43.05	4.94	**44.89**	5.14	**41.00**	5.00	40.36	5.17
$I_{45,B}$	37.26	2.37	**37.42**	2.38	43.75	4.71	**45.42**	4.70	**40.35**	5.05	39.61	5.29
$I_{90,B}$	**37.31**	2.40	37.27	2.53	43.80	4.73	**45.47**	4.75	**41.01**	4.98	40.52	5.23
$I_{135,B}$	**36.73**	2.84	36.40	3.07	43.24	4.91	**45.10**	5.10	**40.23**	5.05	39.62	5.41

Table 1: Average μ and standard deviation σ of PSNR, computed individually in a leave-one-out manner and by channel. Best mean values by database and by channel are highlighted in bold fonts.

	S_0		S_1		S_2	
	μ	σ	μ	σ	μ	σ
DB1						
EARI	**38.68**	3.54	36.38	3.67	36.08	3.73
LMMSE	38.51	3.37	**37.17**	3.14	**37.21**	3.06
DB2						
EARI	**47.03**	3.78	**43.16**	4.35	**43.45**	4.40
LMMSE	45.61	3.43	42.58	4.24	43.14	4.35
DB3						
EARI	39.62	4.74	38.83	4.79	36.23	4.86
LMMSE	**40.96**	4.77	**40.57**	5.31	**38.85**	4.51

Table 2: Average and standard deviation for CPSNR, computed by Stokes vector component.

used to train the demosaicing matrix. This is in a context where there is a training procedure and where the evaluation is done in a leave-one-out manner. Future work should confirm the improvement when more data are available. Visually it seems that LMMSE performs better, but it needs to be confirmed by a quantitative analysis.

Conclusion

In this work, we applied the LMMSE algorithm to the specific case of Colour Polarisation Filter Arrays image demosaicing. We evaluated its performance compared to EARI and found that it increases statistically the result in term of PSNR over the database provided by Monno [12]. Thus, we demonstrated that this extension of the algorithm to CPFA is competitive with the last state-of-the-art dedicated method.

As a future work, the LMMSE demosaicing for CPFA should be evaluated over a large number of images to strengthen these results. Those results are obtained on a specific sensor, and the role of the mosaic pattern on the demosaicing needs to be investigated.

Acknowledgments

This work was supported by the ANR JCJC SPIASI project, grant number ANR-18-CE10-0005 of the French Agence Nationale de la Recherche, and by the Research Council of Norway over the project 'Individualised Colour Vision-based Image Optimisation', grant number 287209.

References

[1] D. Alleysson and P. Amba. Method for reconstructing a colour image acquired by a sensor covered with a mosaic of colour filters, Dec. 31 2020. US Patent App. 16/977,915.

[2] P. Amba, J. B. Thomas, and D. Alleysson. N-lmmse demosaicing for spectral filter arrays. *Journal of Imaging Science and Technology*, 61(4):40407–1–40407–11, 2017.

[3] B. Chaix de Lavarène, D. Alleysson, and J. Hérault. Practical implementation of lmmse demosaicing using luminance and chrominance spaces. *Computer Vision and Image Understanding*, 107(1):3–13, 2007. Special issue on color image processing.

[4] G. Courtier, P.-J. Lapray, J.-B. Thomas, and I. Farup. Correlations in joint spectral and polarization imaging. *Sensors*, 21(1), 2021.

[5] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi. Beyond color difference: Residual interpolation for color image demosaicking. *IEEE Transactions on Image Processing*, 25(3):1288–1300, 2016.

[6] P.-J. Lapray, L. Gendre, A. Foulonneau, and L. Bigué. Database of polarimetric and multispectral images in the visible and NIR regions. In C. Fournier, M. P. Georges, and G. Popescu, editors, *Unconventional Optical Imaging*, volume 10677, pages 666 – 679. International Society for Optics and Photonics, SPIE, 2018.

[7] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton. Multispectral filter arrays: Recent advances and practical implementation. *Sensors*, 14(11):21626–21659, 2014.

[8] X. Li, B. Gunturk, and L. Zhang. Image demosaicing: a systematic survey. In W. A. Pearlman, J. W. Woods, and L. Lu, editors, *Visual Communications and Image Processing 2008*, volume 6822, pages 489 – 503. International Society for Optics and Photonics, SPIE, 2008.

[9] S. Mihoubi, P.-J. Lapray, and L. Bigué. Survey of demosaicking methods for polarization filter array images. *Sensors*, 18(11), 2018.

[10] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire. Multispectral demosaicing using pseudo-panchromatic image. *IEEE Transactions on Computational Imaging*, 3(4):982–995, 2017.

[11] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire. Mul-

tispectral demosaicing using pseudo-panchromatic image. *IEEE Transactions on Computational Imaging*, 3(4):982–995, 2017.

[12] M. Morimatsu, Y. Monno, M. Tanaka, and M. Okutomi. Monochrome and color polarization demosaicking using edge-aware residual interpolation. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 2571–2575, 2020.

[13] T. Okawa, S. Ooki, H. Yamajo, et al. A 1/2inch 48m all pdaf cmos image sensor using 0.8 μm quad bayer coding 2× 2ocl with 1.0 lux minimum af illuminance level. In *International Electron Devices Meeting*, pages 16–3. IEEE, 2019.

[14] S. Qiu, Q. Fu, C. Wang, and W. Heidrich. Polarization Demosaicking for Monochrome and Color Polarization Focal Plane Arrays. In H.-J. Schulz, M. Teschner, and M. Wimmer, editors, *Vision, Modeling and Visualization*. The Eurographics Association, 2019.

[15] Sony. Polarization image sensor. Technical report, Polarsens, 2018.

[16] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton. Median filtering in multispectral filter array demosaicking. In N. Sampat and S. Battiato, editors, *Digital Photography IX*, volume 8660, pages 103 – 112. International Society for Optics and Photonics, SPIE, 2013.

Figure 5: REFERENCE. Visualization of zoomed region of interest in reference image.

(a) $I_{0,R}$ (b) $I_{0,G}$ (c) $I_{0,B}$ (d) $I_{45,R}$ (e) $I_{45,G}$ (f) $I_{45,B}$ (g) $I_{90,R}$ (h) $I_{90,G}$ (i) $I_{90,B}$ (j) $I_{135,R}$ (k) $I_{135,G}$ (l) $I_{135,B}$

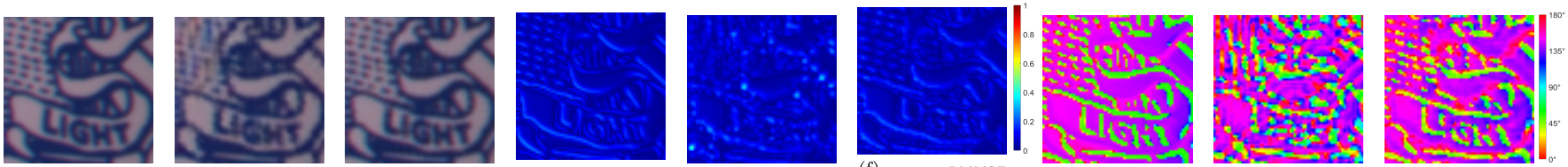
Figure 6: EARI. Visualization of zoomed region of interest in image demosaiced with EARI.

(a) $I_{0,R}$ (b) $I_{0,G}$ (c) $I_{0,B}$ (d) $I_{45,R}$ (e) $I_{45,G}$ (f) $I_{45,B}$ (g) $I_{90,R}$ (h) $I_{90,G}$ (i) $I_{90,B}$ (j) $I_{135,R}$ (k) $I_{135,G}$ (l) $I_{135,B}$

Figure 7: LMMSE. Visualization of zoomed region of interest in image demosaiced with LMMSE.

(a) $I_{0,R}$ (b) $I_{0,G}$ (c) $I_{0,B}$ (d) $I_{45,R}$ (e) $I_{45,G}$ (f) $I_{45,B}$ (g) $I_{90,R}$ (h) $I_{90,G}$ (i) $I_{90,B}$ (j) $I_{135,R}$ (k) $I_{135,G}$ (l) $I_{135,B}$

Figure 8: Visualization of the S_0 Stokes component images, and the DOLP (Degree Of Linear Polarisation) and AOLP (Angle Of Linear Polarisation).



(a) Ref. S_0 (b) EARI S_0 (c) LMMSE S_0 (d) Ref. $DOLP_G$ (e) EARI $DOLP_G$ (f) LMMSE $DOLP_G$ (g) Ref. $AOLP_G$ (h) EARI $AOLP_G$ (i) LMMSE $AOLP_G$