Regularized Color Demosaicing via Luminance Approximation

Johannes Herwig and Josef Pauli; University of Duisburg-Essen; Duisburg, Germany

Abstract

In single-sensor digital imaging a color filter array, that is overlaid onto the image sensor, makes color images possible. Incident light rays become band-limited and each sensor element captures either red, green or blue light. Interpolating the missing two color components for each pixel location is known as demosaicing. This paper proposes to firstly derive an estimated luminance image by low-pass filtering the original mosaiced sensor image. In a second step a deconvolution technique re-sharpenes the blurred luminance approximation, so that it has the same spatial resolution as the original - but bandpassed - sensor image. Using the high-resolution luminance approximation the partial RGB colors from the mosaiced sensor image are transformed into a different color space that is more suitable for color interpolation. The new color space consists of least correlated color data, so that intra-channel interpolation errors have a reduced impact on inter-channel alignment, and therefore result into less prominent interpolation artifacts. Demosaicing is performed on the transformed color data separately for each plane, whereby again the luminance approximation, which encodes the aligned gradient direction of all color channels, regularizes the bilinear interpolation. Finally, the result is remapped into the RGB color space to obtain the demosaiced color image. Additionally, correlated multi-channel anisotropic diffusion is applied onto the demosaiced color image to further reduce interpolation artifacts and enable denoising. The proposed algorithm is evaluated and it is concluded that - although the image formation model could be verified - its performance heavily depends on the quality of the luminance approximation, i.e. the deconvolution method.

Problem Outline

Because pixel elements of digital image sensors are only capable of sensing incident light photons, but cannot distinguish them by their wavelength, a color filter array (CFA) may be used to make the photo-detector array sensitive to color. A filter layer with an alternating mask of red, green and blue filters is placed on top of the image sensor, thereby band-limiting each sensor element according to some regular mosaic pattern. The most widely chosen pattern is the Bayer filter mosaic [3]. In this pattern half of the pixels are green, whereas each red and blue pixels cover 25 percent of the sensor. A first row of alternating blue and green pixels is followed by a second row of alternating green and red pixels. The Bayer pattern effectively reduces the spatial sampling resolution of the image sensor in favor of additional multispectral sampling. The aim of an interpolation algorithm is to recover lost spatial resolution due to the filter mosaic, thereby preserving high frequency information. Especially, one has to avoid introducing visual artifacts like blurred edges, aliasing and false colors, which are mainly the result of interpolating across edges, falsly anticipated maximum intensity gradient direction within a color channel [5], and misregistration in color balancing [17], respectively. Another well known artifact in this domain is the so-called zippering effect, that is a pattern of alternating colors, that occurs in otherwise homogeneous regions or at sharp edges [11].

Previous Work

For a recent survey on color demosaicing algorithms the reader is referred to [18]. Most successful iterative and non-iterative algorithms in the spatial domain use the color ratio or color difference model, that states that in natural images ratios or differences between color channels are locally the same (see the following section in this paper). The proper alignment of color channels is a regularization technique that prevents false colors to occur in the interpolation result, that otherwise are likely to be present when interpolation of color planes is performed separately ignoring their coupling of intensity gradients. In the aforementioned class of interpolation algorithms due to its higher sampling frequency the green plane is bilinearly interpolated first, and then red and blue planes are interferred by the color ratio or color different rules. In further iterations the interpolation results of the green plane may be enhanced by backward propagation of the ratio rules between both interpolated and originally sensed data, so that tight coupling of all channels is enforced. These algorithms initially rely on high-quality estimates of the green plane. Examples are found in [9], [5], [17]. In [16] vector correlated anisotropic diffusion is implemented in the previous sense for a final enhancement. Other approaches that try to model correlation between color channels include the following. The Markov Random Field framework is utilized in [20] to enhance a color image that has been previously obtained by simple bilinear interpolation. There are few vector valued approaches [10] that interpolate all color channels simultaneously. In [21] a learning-based vector quantisation algorithm that takes advantage of self-similarity in images is proposed.

The results from [2] and [6] that analyze the fourier spectrum of the mosaiced sensor image, also termed Bayer image due to the special CFA pattern, serve as a basis of the approach presented in this paper. There it has been found that luminance and chrominance data are multiplexed within the Bayer image, because a one-color per pixel image can be written as the sum of luminance and chrominance, whereby the lower frequencies located around the center of the fourier spectrum encode luminance data and the higher frequencies located at the outer border of the fourier spectrum encode chrominance data. The occurrence of color artifacts at object edges during demosaicing has been explained to be inherent due to the smooth transition from lower luminance to higher chrominance frequencies in the fourier spectrum, where both sources of frequencies are mixed together and no clear distinction can be made [2]. In a further development multiplexing has been utilized in [6] for demosaicing in the fourier domain via the Wiener filtering approach. There the Bayer sampling is decomposed into a sum of a luminance estimator and a chrominance projector. This model is inversely solved to obtain an interpolated color image.

The approach presented here is performed in the spatial domain - as opposed to the fourier domain - which makes future enhancements possible that depend on the spatial relationship of the underlying data. When working in the fourier domain spatial information is lost, and additionally computational overhead for transforming between both domains is introduced.

Image Formation Model

In this section the origin and justification of the color ratio and color different rules is given, that lies in a simple color formation model for natural images. The model assumes that color textures in an image are not too fine grained, so that interpolation under some local color constancy requirement is valid. A CFA interpolation algorithm provides an inverse mapping that tries to reconstruct missing vector elements from the single-channel Bayer image and estimates its corresponding three-channel color image. This is an ill-posed problem in both the spatial and spectral domain. Generic prior knowledge applicable to a wide range of natural images and sensors is necessary for simultaneous regularization in both domains, and is established via intra- and inter-channel dependencies. The physical color formation model underlying most demosaicing methods is that of a Mondriaan world made of Lambertian nonflat surface patches [16, 15, pp. 152]. This very simplistic model ignores the existance of specularities and implies isotropic luminance of objects. Hence, a luminance image - and as seen later each color plane - is supposed to be locally homogeneous in the spatial domain (see figure 1 on the left). Furthermore, the albedo is a property of the material of an object and also depends on the wavelength (i.e. the color channel) of reflected light. Thus color channels are linearly dependent and aligned in the spectral domain (see figure 1 on the right). Again this is due to the albedo, which is just a scaling factor that describes the amount of light belonging to a special wavelength (color channel) out of all reflected light (luminance). Therefore the measurement of a particular color channel is proportional to the normalized shading image due to the overall reflected light of an object. If it is further assumed that a given object is made of a single material, then locally the gradients of color channels should have the same direction. This oversimplified model accounts for the constant-hue assumption or specifically the color ratio rule, that is widely exploited in demosaicing [18], which states that the ratio of any two color channels is locally the same. The assumption of high inter-channel correlation has been proven to hold approximately on a popular real world image set [8] by [9].

Proposed Algorithm

In the following the proposed demosaicing algorithm is described, motivated and linked to the previously introduced image formation model. A subsection is dedicated to every single procedure of the demosaicing algorithm, whereby every procedure is performed sequentially in the order they appear in the text. With the demosaicing method presented here interpolating the missing color components is done in a transformed color space of the Bayer image other than in RGB space directly. The color transformation is justified by the image formation model from which a color albedo space is deduced wherein interpolation errors introduced during demosaicing have less influence on the final interpolation result. For the color transformation, however, the luminance image needs to be known beforehand, whereas usually the luminance image is interferred from the full color image that has been obtained after demosaicing. Therefore, it is proposed to estimate the luminance image with full spatial sensor resolution via low-pass filtering the single channel Bayer image and subsequent deconvolution to achieve super-resolution. The then known luminance image is utilized for color transformation and functions as a high resolution weight map within the bilinear color interpolation step (in color albedo space), for which usually the sparsely sampled and aliased color planes of the originally measured Bayer image are used.

Blurred Luminance Image

The first part of the algorithm is concerned with the derivation of a luminance image from the spectrally sampled Bayer image that is obtained from the sensor.

The fact that the green channel is sampled twice as dense as the other two and its quincunx pattern makes interpolation relatively easy when spectral alignment with other color planes is initially ignored (opposed to the image formation model). Because the human visual system is most sensitive to green light, the interpolated green plane is often called the luminance channel. But this is misleading, since e.g. a red object would have zero luminance in this sense. Instead, convolution of the Bayer image with a Gaussian smoothing kernel is proposed.

The spatial CFA sampling of the Bayer pattern introduces gray-value gradients between neighboring pixels due to their differing spectral responses. But when a human observer varies his distance looking at such an image, he won't recognize those gradients anymore after he is farther away from the image display. He may see a continuous monochrome image, instead. This observation suggests that smoothing the Bayer image may create an approximation of a luminance image, because CFA aliasing due to high frequency CFA gradients is reduced. Hence, the Bayer image has been convolved with a 3x3 Gaussian filter. Note, that the resulting image mixes different spectral channels. According to the formular $\sigma = 0.3(n/2-1)+0.8$, where $n=3$ is the size of the horizontal and vertical filter kernel [1], the normalized Gaussian G_σ with $\sigma = 0.95$ (so that the topology of the kernel integrates to one) is

$$G_\sigma = \begin{bmatrix} 0.0625 & 0.125 & 0.0625 \\ 0.125 & 0.25 & 0.125 \\ 0.0625 & 0.125 & 0.0625 \end{bmatrix}. \quad (1)$$

The authors of [2] have come up with the same filter kernel to approximate luminance for spatial Bayer images. Their derivation is embedded within a mathematical framework for CFA imaging based on the finding that in the fourier domain luminance and chrominance information of the Bayer image are multiplexed by summation. They do not mention that their result conforms to the normalized Gaussian filter kernel. It is interesting to note, that an unknown Point-Spread-Function (PSF) of a camera sensor is usually modeled as Gaussian [12]. Therefore one can take the filter result as a slightly out-of-focus luminance image of the underlying scene. Due to the Bayer pattern there are four different 3x3 patches possible, see figure 2. Applying weighted averaging of color samples according to the filter kernel, and under the assumption that colors are locally the same as discussed with the image formation model, one gets luminance $L = 0.25R + 0.5G + 0.25B$ for all of these patches for any given center pixel. For example, the convolution of the first patch shown in figure 2 results in

$$\begin{bmatrix} G & R & G \\ B & G & B \\ G & R & G \end{bmatrix} * G_\sigma = \begin{array}{ccccccccc} 0.0625G & + & 0.125R & + & 0.0625G & + \\ 0.125B & + & 0.25G & + & 0.125B & + \\ 0.0625G & + & 0.125R & + & 0.0625G \end{array}$$

and can be finally reduced to $L = 0.25R + 0.5G + 0.25B$. Although the patterns are spatially different, the same result L is derived for all the possible mosaic patches shown in figure 2. Under the assumption that the original colors of the scene are locally constant, so that the partially measured spectral responses at each pixel comprising the 3x3 patch all originate from roughly the same color, the luminance value can be estimated for each pixel by filtering the Bayer image with the Gaussian filter G_σ.

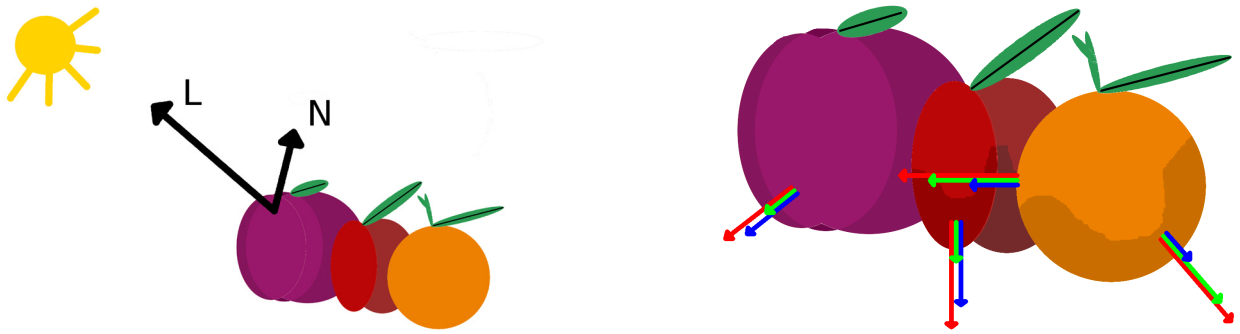

Figure 1. *A graphical depiction of the image formation model adopted from [15]. On the left the Mondriaan world model and the lambertian reflection model are shown, whereas the right image emphasises the spectral alignment of different wavelengths of light due to their albedo.*

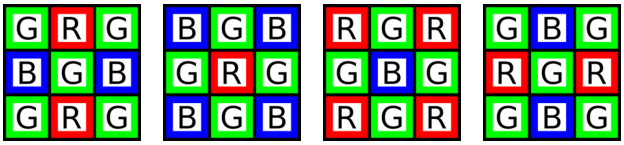

Figure 2. *The four local GRBG, BGGR, RGGB, GBRG patches.*

Deblurred Luminance Image

After low-pass filtering the Bayer image a continuous but smooth luminance approximation with lower spatial resolution than the original sensor image has been obtained. In order to refine the luminance image to have the same spatial resolution as the sensor image a deconvolution algorithm is employed, because the additional blurring can be seen as a synthetic PSF applied to the imaging pipeline. Because the blurring kernel of the PSF is known, which is the Gaussian kernel G_σ, a straightforward deconvolution algorithm like Jansson-Van Cittert [13], Richardson-Lucy [23, 19], or Landweber iteration is applicable. In this case all of these algorithms seem to perform equally well, so it is decided to stick with Richardson-Lucy which is commonly stated to be the most robust among them. In the demosaicing algorithm the amount of Richardson-Lucy iterations has been fixed to 15.

Forward Color Space Transform

At this point in the proposed algorithm there are two images of the same scene captured by a single digital color sensor available. These are firstly the originally sensed Bayer image with the measured spectral pattern, and secondly a synthetically derived luminance approximation that is of the same spatial resolution as the sensor image. Note that this luminance image estimates the intensity of light that is composed of all spectral channels arriving at every sensor pixel, although only parts of it have actually been measured according to the bandpassed spatial sampling of the Bayer pattern. Both are still scalar valued images.

With an estimated full-resolution luminance image known, a transformed Bayer image is created by dividing the original Bayer image with the luminance image, that is

$$b^\star(x,y) := b(x,y)/l(x,y) \tag{2}$$

where $b(x,y)$ denotes a pixel in the Bayer image domain, and $l(x,y)$ is the corresponding pixel in the luminance domain, and finally the resulting pixel $b^\star(x,y)$ lies in the transformed color space of the original Bayer image. Now, in the transformed Bayer image the amount of the contribution of a sampled color channel w.r.t. the overall luminance at a given pixel is stored. This is qualitatively similar to the albedo of the image formation model discussed in the previous section. Since according to the model the gradient direction is encoded in the overall luminance, the color space transform has effectively removed gradient information from each sampled color plane. The transformed color space consists of gradient information of the albedo only, which can be interpreted as the spatial change in chromacity.

Weighted Bilinear Interpolation

The transformed Bayer image from the previous step is used for weighted bilinear demosaicing. Because the spectral channel alignment property, that is encoded in the luminance image, has been removed from the data, the probability to introduce color artifacts is reduced.

The interpolation step of the transformed albedo data goes as follows. Firstly, a three-channel image is allocated, where each pixel has any one of the form

$$\begin{aligned} c^\star(x,y) &:= (b^\star(x,y),\star,\star)^T, \\ c^\star(x,y) &:= (\star,b^\star(x,y),\star)^T, \\ c^\star(x,y) &:= (\star,\star,b^\star(x,y))^T \end{aligned} \tag{3}$$

according to the Bayer pattern, whereby a pixel is of the type $c^\star(x,y) = (R^\star(x,y),G^\star(x,y),B^\star(x,y))^T$ in the new albedo space where e.g. $R^\star$ denotes the albedo value of the red color component, and a $\star$ denotes a missing component value that is to be interpolated from its local pixel neighbourhood.

For example, the missing green color component $G^\star(x,y)$ of the pixel $c^\star(x,y) = (b^\star(x,y),\star,\star)^T$ or alternatively $c^\star(x,y) = (\star,\star,b^\star(x,y))^T$ is computed as follows:

$$G^\star(x,y) = \frac{\sum_{(m,n)\in N} W(x,y,m,n)G^\star(m,n)}{\sum_{(m,n)\in N} W(x,y,m,n)} \tag{4}$$

where the local neighbourhood is defined as $N = \{(x-1,y),(x+1,y),(x,y-1),(x,y+1)\}$ and the weighting function W is

$$W(x,y,m,n) := \frac{1.0}{\|l(x,y)-l(m,n)\|+1.0} \tag{5}$$

where $l(x,y)$ denotes the luminance value of the deblurred luminance approximation at pixel location (x,y). The interpolation of other component values is analogous.

The result of the weighting function $W(x,y,m,n)$ is based on the luminance difference between the two neighbouring pixels (x,y) and (m,n). The approximate luminance image is chosen to obtain values for pixel weighting for two reasons. Firstly, there are luminance values available for every pixel location. So this is the most direct way to relate the relevance of a neighbouring pixel value to the current position. Secondly, when using

luminance differences for weighting, then inherently the inter-channel coupling of the different color channels is implicitly realized, although each color plane is computationally interpolated separately. This is due to the image formation model that defines the correlation among color channels through the overall intensity gradient.

The qualitative form of the weighting function is chosen to become fastly steeper with increasing luminance differences in order to provide a sharp cutoff for changing grey value topologies, i.e. to avoid interpolating across edges. It has been observed in experiments that a weighting function $W(x,y,m,n)$ that has a linear weighting scheme accounts for zippering artifacts around edges. Therefore it is concluded that it is of most importance in demosaicing methods to model the weighting function adequately [5, 17].

Backward Color Space Transform

After bilinear interpolation in the color albedo space, the $c^\star(x,y)$ values have to be remapped into the RGB color space $c(x,y)$. This is simply done by the inverse operation of the forward color space transform. Specifically, each color channel is separately multiplied with the approximate luminance image, so that

$$c(x,y) := (c^\star_R(x,y)l(x,y), c^\star_G(x,y)l(x,y), c^\star_B(x,y)l(x,y))^T \quad (6)$$

where $c^\star_R$, $c^\star_G$, and $c^\star_B$ denote the three albedo components of $c^\star(x,y)$. Multiplying each channel with the same luminance plane ensures that the interpolated color vectors are spectrally aligned as defined by the image formation model.

Anisotropic Diffusion

Although all the previous steps of the proposed demosaicing algorithm are derived from the image formation model, there still remains a source of errors, which is the deconvolution procedure applied to the blurred luminance image, which itself is only an approximation made by Gaussian filtering. The most relevant errors introduced by deconvolution are the ringing artifacts around object edges. Those ringing artifacts consist of amplified gradients in the deconvolution result that manifest themselves in either lighter or darker gray value distributions that surround edges in repeated parallel stripes. Therefore the gradient directions are falsely interferred by the previous weighted bilinear interpolation and false colors occur on edge pixels. In order to re-establish a common color gradient direction within a local neighbourhood a vector-correlated anisotropic diffusion technique is applied to the interpolated color image as a final enhancement step.

For anisotropic diffusion the framework described in [24] and [25] is used. The main contribution of the framework is separating the structure tensor field from the diffusion tensor field. The structure tensor for vector valued images is defined as

$$S = \sum_{i=0}^{N} \left[\left(\nabla I_i \nabla I_i^T \right) * G_\sigma \right] = \begin{pmatrix} \sum_{i=0}^N I_{i_x}^2 & \sum_{i=0}^N I_{i_x} I_{i_y} \\ \sum_{i=0}^N I_{i_x} I_{i_y} & \sum_{i=0}^N I_{i_y}^2 \end{pmatrix} * G_\sigma$$

where with ∇I_i the gradient image of the $i$ths channel is denoted and G_σ is a Gaussian smoothing of the elements of the gradient product matrix over the field of all such matrices located at every image pixel. The number of color channels is denoted with N here. The structure tensor S robustly describes local gradient information, which can be extracted from the orthonormal eigenvectors and their corresponding eigenvalues, respectively, as $\lambda_{+/-} = 0.5\left(s_{11}+s_{22} \pm \sqrt{(s_{11}-s_{22})^2+4s_{12}^2}\right)$ and $\theta_+ = (\cos\phi, \sin\phi)^T$ with $\phi = 0.5 \arctan\left(\frac{2s_{12}}{s_{11}-s_{22}}\right)$, where the small s are the elements of the matrix S, θ_+ is the eigenvector in the direction of the gradient with strength λ_+ and θ_- is the perpendicular eigenvector pointing into the direction of least change λ_-. From the structure tensor the diffusion tensor is derived, so that the diffusion process later smoothes along but not across edges. Therefore the smoothing in the direction of least change is weighted most anisotropic, whereas blurring along the gradient direction is weighted most isotropic. Such a desired diffusion tensor is given by $D = \left[f_+\left(\sqrt{\lambda_+ + \lambda_-}\right)\theta_-\theta_-^T + f_-\left(\sqrt{\lambda_+ + \lambda_-}\right)\theta_+\theta_+^T \right]$ with the two weighting functions $f_-(a) = \left(1+a^2\right)^{-p_1}$ and $f_+(a) = \left(1+a^2\right)^{-p_2}$ with $p_1 < p_2$. Then, during the diffusion process the update velocity per channel is $\partial I_i/\partial t = \text{trace}\,(DH_i)$, where H_i denotes the hessian matrix. Finally, with $v = (\partial I_i/\partial t, \ldots, \partial I_N/\partial t)^T$ the new iterated image $I^{k+1} = I^k + \Delta t\, v$ is estimated by an adaptive timestep Δt, which is the percentage of change in each iteration normalized to the maximum difference between the current and previous iterations. In this approach all matrices S, D and H_i are recomputed within each iteration for the current image I^k, whereby the image I^0 is the original input.

Because of the fact that the same diffusion tensor D is used to smooth each channel, the smoothing remains correlated, which avoids introducing new color artifacts. Rather, with ongoing time color artifacts that are present will be removed, because steadily the low-frequency model common of all channels, from which the diffusion tensor is derived, is approached and artifacts that initially resulted from frequency outliers get damped or raised.

The parameters of the diffusion method have been fixed in the demosaicing algorithm. The anisotropy parameter is $p_1 = 0.95$ and the sharpness is $p_2 = 0.05$ to support strict smoothing in the gradient direction, the change velocity per iteration is $\Delta t = 0.05$ and a total of 10 iterations are performed. This parameter values support the idea of a slow and steady diffusion process that does not alter the details of the image structure, but at the same time enhances the correlation of color plane gradients.

Experimental Simulation

Preparation of Synthetic Test Data

The Kodak PhotoCD database [8] with 24 sample images of 768×512 pixel size and showing simple and complex real-life scenes is used for performance testing. These images are digitally scanned from photochemical film in a high-quality process by a 3CCD sensor. Hence, for every image pixel there are originally sensed full-color vectors available. To avoid the so-called Inverse Crime [14], which basically states that to obtain statistically valid performance data one has to generate test samples by using a different model than the proposed solution model, these images are firstly downsampled to half their size according to the scale-space pyramid [4], whereby a Gaussian smoothing with five pixels window size for each color channel has been applied. These color images, that serve as ground-truth data, are then subsampled with the Bayer pattern.

Luminance Approximation by Gaussian Filtering

In figure 3 the luminance property of a ground truth color image is exemplarily analyzed to verify the previous claims about Gaussian filtering of the Bayer image. A synthesized Bayer image is smoothed as described by G_σ and is shown right after the color image. The following images in figure 3 show the difference images obtained by measuring the distance between the smoothed Bayer image and non-linear CIE luminance (Rev.

Figure 3. *The original color image, the Gaussian filtered Bayer image, and difference images between Gaussian and traditional luminance computation.*

scaled nonsmooth	scaled smooth	unscaled nonsmooth	unscaled smooth	
CIE Luminance 601 difference images				
2144.06	1305.05	5528.5	2607.41	norm
0.022	0.013	0.014	0.007	avg. norm
4.871	3.284	5.852	3.254	avg. mean
4.721	2.416	6.535	2.454	std. dev.
L_2-norm brightness difference images				
2340.28	1521.23	5985.85	3125.97	norm
0.024	0.015	0.015	0.008	avg. norm
5.050	3.238	6.226	3.306	avg. mean
5.302	3.374	7.035	3.492	std. dev.
Naive brightness difference images				
1824.35	905.002	5107.04	1854.7	norm
0.019	0.009	0.013	0.005	avg. norm
4.030	2.233	5.255	2.281	avg. mean
4.144	1.748	6.178	1.801	std. dev.

RMS errors of three widely used luminance measures with the Gaussian luminance approximation of a Bayer image.

601) [22], RGB-to-monochrome transform by naively calculating $(R+G+B)/3$, and the L_2-norm of RGB vectors taken for the luminance part, respectively. Those luminance images for comparison are directly derived from the downscaled ground truth and have additionally been smoothed by the 3x3 Gaussian filter to simulate out-of-focus blur. In the difference images white encodes maximum and black minimum difference in terms of gray values, which are scaled for display purposes, but typically lie in a range from 0 to 3. This is further specified in table 1, where the average differences of all images in the dataset from [8] are given. Four test series have been studied. A series is termed unscaled, when the initial downsizing has been omitted, and it is non-smooth, if the additional smoothing of the ground-truth luminance in order to simulate out-of-focus blur also has been omitted. From these results and the difference images it can be concluded that up to a region-dependent scale factor a Gaussian smoothed Bayer image approximates luminance. The proposed filter kernel does integrate the spatially sampled color information, that is available around the neighborhood of a Bayer pixel (see figure 2), so that the result is close to a luminance computation as if the full-color vector were known, whereby errors are spatially correlated and are larger at edges of objects due to an inherent smoothing.

Evaluation of the Demosaicing Approach

The proposed algorithm is able to reconstruct a color image nearly indistinguishable from the original, especially without interpolation artifacts at object edges, but sadly only when ground truth luminance is used for regularization (instead of the deblurred luminance approximation). However, this observation justifies the idea about using an intermediately derived luminance image that regularizes the demosaicing process. Results degrade when using the deconvolved luminance approximation, which is because of ringing artifacts introduced by deconvolution. Therefore the quality of the deconvolution result is crucial for demosaicing. The RMS error between ground truth and interpolated color images is shown in table 2. The average mean and the standard deviation have been computed by averaging the values over all interpolated pixels (of all channels) of all the test images. Two versions of the proposed algorithm are evaluated. The first version comprises all the steps described in the previous section, whereas the second version has omitted the anisotropic diffusion step. Thirdly, the underlying idea about using a high resolution luminance image of the scene for computing the weights for bilinear interpolation is evaluated by performing the proposed method using ground truth luminance data derived from the full color image. Therefore errors in the deblurring procedure do not have any effect on the interpolation result, and the average RMS error is below one gray value per channel (see table 2). These very low average mean RMS error per pixel shows the applicability of the simple image formation model, because the resulting color values only rely on the color ratio model and the chosen weighting function. Although the proposed algorithm has additional denoising capabilities when anisotropic diffusion is incorporated, its RMS errors are in the range of those computed without anisotropic diffusion. In order to have a comparison for the error values, the results for a simple bilinear demosaicing method are also given in table 2, which are outperformed by the proposed method. For all methods the errors of the green color plane are lower than those for the red and blue planes, which corresponds to their sampling density.

Although the RMS error is a popular method it is not a color distance metric. Hence, the CIE Lab and its newer derivant CIEDE 2000 [22] have been evaluated in order to quantify the perceptual difference between an interpolated and its corresponding ground truth color pixel value. The ΔE value of CIE Lab measure is the euclidean distance between two color vectors that have been transformed into a perceptually uniform color space (other than RGB). For certain intervals of ΔE the following qualitative conclusions can be made [26]. For $0.0 \ldots 0.5$ colors are nearly indistinguishable, $0.5 \ldots 1.0$ makes a difference for the trained eye, $1.0 \ldots 2.0$ marks a noticable difference, $2.0 \ldots 4.0$ has a perceived difference, $4.0 \ldots 5.0$ is an untolerable difference, and values of ΔE above 5.0 give rise to two different colors. The just noticable difference (JND), where most humans begin to see a difference in two colors, has been defined here as $\Delta E = 2.3$. The results are given in table 3 analoguous to the test runs defined previously. The CIEDE 2000 and ΔE measures are also in the same range for both versions of the proposed algorithm and much lower (better) than the reference results for bilinear demosaicing. To gain further insight the percentage of pixels that fall into one of the given ΔE intervals is shown for each algorithm (see table 3). For bilinear demosaicing more than half of the total pixels is above the JND measure and hence contain visible color artifacts to a human viewer, whereas the opposite is the case for both proposed algorithms. Also the percentage of pixels with $\Delta E > 5.0$ is much higher for bilinear demosaicing, and although this is still true for one fifth of the pixels with the proposed method the performace gain of the lower than JND interval is overproportional. Again the results for the ground truth luminance test support the overall image formation model, because 85 percent of all color pixel values do not make a noticable

difference to their ground truth counterparts.

Another popular evaluation measure is the peak signal to noise ratio (PSNR). The authors of [20] note that to their opinion PSNR does not always quantify edge reconstruction quality as visually perceived, hence they proposed a peak edge signal to noise ratio (PESNR):

$$20 \cdot log_{10}\left(255 \cdot \left[\frac{\sum_s \sum_x \sum_y e(x,y)\left(I_s(x,y) - \tilde{I}_s(x,y)\right)^2}{3 \cdot \sum_x \sum_y e(x,y)}\right]^{-\frac{1}{2}}\right),$$

where $e(x,y)$ is 1 if there is an edge pixel and 0 otherwise. Here their evaluation method is adopted and the function e is replaced by the result of a sobel filter that is applied to the ground truth color image for each plane separately. Whereas PSNR weights the signal strength equally over all pixels, the PESNR gives higher weights to pixels that are classified as edges. Therefore, the PESNR measure might be more suitable for evaluating the interpolation quality, because demosaicing algorithms are expected to behave badly on edge pixels but fairly well within nearly homogeneous regions. Results are also given in table 2.

Because the RMS error is difficult to interpret and no spatial relationship of the error can be interferred from the previous performance measures, a visual comparison of the results is made possible by figure 4. In the lighthouse image with very high spatial frequencies, e.g. around the fence, there are zippering effects and false colors still clearly visible. One notices that the images obtained by simple bilinear demosaicing have a blurred appearance and results of the proposed methods are perceived sharper. This is due to the fact that false colors and zippering effects are present at both sides of an edge within the bilinearly demosaiced result, effectively blurring edges, whereas these artifacts have been reduced by the proposed algorithm.

Because the luminance approximation step is crucial for this method, blurred and deblurred luminance images are shown in figure 5 for comparison. Each pair of luminance images has been created corresponding to the procedure described in the previous section. Ringing artifacts due to the deconvolution technique are clearly visible as darker and lighter bands around edges, but also the sharpening effect is apparent. The RMS error of a deconvolved luminance estimation w.r.t. the ground truth luminance (Rev. 601) obtained from the original full color images has been computed over all pixels of all images of the data set, which is 3.892 gray values and its standard deviation is 3.362 per pixel.

Comparison with Sophisticated Approaches

An evaluation has been carried out against the data set available on the homepage of the [7] paper. There demosaicing results of the methods by [9] (POCS, projection onto convex sets), [2] (FDM, Frequency Domain Method), [7] (AF, Asymmetric Filters), and [7] (AFDM, Adaptive Frequency Domain Method) are made public. Whereas the algorithm proposed herein outperforms the simple bilinear demosaicing approach, it performs worst among the more sophisticated methods. Although this is true, it shows potential for better results if the luminance approximation through the deconvolution method were of higher quality, because demosaicing using ground-truth luminance performs best of all algorithms. Even then the AFDM would be pretty close in terms of PSNR and PESNR data, but still farther away in terms of the ΔE measures, refer to tables 2 and 3.

Conclusion and Future Work

In the presented work, a demosaicig algorithm has been described that firstly derives an intermediate luminance image from the original sensor image. Although the Bayer image is spatially sampled by spectral filters it has been shown and experimentally evaluated that a luminance image can successfully be generated that has valid intensity values in the luminance domain for every pixel location. This approach is unique because by applying a Gaussian filter with subsequent deconvolution a luminance image is derived through an entirely scalar-based image processing chain without the previous need for demosaicing the color sensor image. This enables the demosaicing method to perform the errorneous interpolation stage using a different color space than RGB, whereby the impact of interpolation errors in the final demosaicing result can be reduced. The luminance image at hand can thereby be naturally used, according to the image formation model, to enforce the desired property of inter-channel correlation via weighted bilinear interpolation. Another advantage for the interpolation procedure is the fact that the luminance image is of the same spatial resolution as the sensor image, whereas in previous demosaicing methods weights for interpolation are extracted from the Bayer image directly, where naturally not any one of the color planes is of the full resolution. It has been noted, however, that the luminance estimation needs to be of very high quality, although the estimation process itself is equally ill-posed as the color interpolation task. The ideas present in this work may better translate into a Markov Random Field model, where the color interpolation and the deconvolution process are coupled to draw profit in their estimation from each other.

RMS error (avg. mean and std. dev.)				PSNR	PESNR
all	red	green	blue		
		channel(s)			
First test series evaluated with downsampling of images					
With anisotropic diffusion					
1.540	1.964	0.834	1.821	37.971	34.112
2.822	3.472	1.828	3.167		
Without anisotropic diffusion					
1.539	1.971	0.828	1.820	36.185	32.317
2.832	3.488	1.833	3.174		
With ground truth luminance					
0.525	0.667	0.134	0.775	44.070	40.732
1.077	1.438	0.324	1.469		
Simple bilinear demosaicing					
2.666	3.250	1.672	3.075	31.028	26.421
5.269	6.193	3.680	5.933		
Second test series evaluated without downsampling of images					
With anisotropic diffusion					
2.519	2.953	1.488	3.116	33.868	29.560
4.659	5.126	3.216	5.635		
Without anisotropic diffusion					
3.715	3.994	2.984	4.166	32.032	27.706
5.306	5.679	4.052	6.185		
With ground truth luminance					
1.090	1.204	0.444	1.621	40.307	36.429
1.827	2.125	0.571	2.784		
Simple bilinear demosaicing					
4.127	4.861	2.620	4.900	29.168	24.421
8.123	9.191	5.855	9.322		
Other sophisticated approaches for comparison					
Adaptive Frequency Domain Method [7]					
1.402	1.612	0.898	1.696	39.422	35.765
2.395	2.612	1.824	2.749		
Frequency Domain Method [2]					
1.678	1.996	1.048	1.990	37.301	34.323
3.065	3.453	2.242	3.499		
Asymmetric Filters [7]					
1.528	1.770	0.985	1.828	38.551	35.244
2.650	2.913	2.012	3.025		
Projection Onto Convex Sets [9]					
1.590	1.831	1.090	1.851	37.191	35.038
3.172	3.338	2.633	3.545		

Errors of demosaicing results w.r.t. the ground truth images.

Method	CIEDE 2000	CIE Lab ΔE	$\Delta E \leq 0.5$	$0.5 \ldots 1.0$	$1.0 \ldots 2.0$	$2.0 \ldots 4.0$	$4.0 \ldots 5.0$	$\Delta E > 5.0$	$\leq$ JND
First test series evaluated with downsampling of images									
With anisotropic diffusion	1.681	3.695	14.0%	16.7%	19.8%	20.2%	6.1%	23.2%	54.6%
Without anisotropic diffusion	1.687	3.707	13.8%	16.8%	19.8%	20.2%	6.1%	23.3%	54.5%
With ground truth luminance	0.729	1.334	36.7%	26.2%	19.7%	11.0%	2.1%	4.3%	85.4%
Simple bilinear demosaicing	2.530	6.264	11.9%	13.8%	15.8%	16.8%	5.7%	35.9%	44.7%
Second test series evaluated without downsampling of images									
With anisotropic diffusion	2.833	5.960	4.8%	8.5%	19.2%	25.6%	7.3%	34.5%	37.7%
Without anisotropic diffusion	2.853	6.009	4.5%	8.6%	19.3%	25.7%	7.3%	34.7%	37.5%
With ground truth luminance	1.092	1.967	23.2%	22.0%	27.2%	17.0%	3.0%	7.6%	77.2%
Simple bilinear demosaicing	3.839	9.516	4.5%	9.2%	11.5%	21.9%	8.1%	44.6%	33.1%
Other sophisticated approaches for comparison									
AFDM [7]	1.749	3.324	7.0%	17.5%	20.4%	30.4%	7.9%	16.8%	57.7%
FDM [2]	2.131	3.983	7.7%	16.3%	18.3%	27.1%	8.0%	22.6%	53.2%
AF [7]	1.960	3.597	7.6%	17.1%	19.4%	28.4%	7.9%	19.6%	55.7%
POCS [9]	2.094	3.757	7.6%	16.8%	19.2%	28.6%	7.6%	19.8%	55.3%

Color difference measures using different metrics of demosaicing results w.r.t. the ground truth color image.

References

[1] The OpenCV Library. Project homepage: http://opencv.willowgarage.com/wiki/. See documentation of the function cvSmooth for the Gaussian sigma formular.

[2] David Alleysson, Sabine Süsstrunk, and Jeanny Herault. Color Demosaicing by Estimating Luminance and Opponent Chromatic Signals in the Fourier Domain. In *Proc. IS&T/SID 10th Color Imaging Conference*, volume 10, pages 331–336, 2002.

[3] Bryce E. Bayer. Color imaging array, 1976.

[4] P. Burt and E. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, 31(4):532–540, Apr 1983.

[5] King-Hong Chung and Yuk-Hee Chan. Color Demosaicing Using Variance of Color Differences. *IEEE Transactions on Image Processing*, 15(10):2944–2955, Oct. 2006.

[6] Brice Chaix de Lavarene, David Alleysson, and Jeanny Hrault. Practical implementation of LMMSE demosaicing using luminance and chrominance spaces. *Computer Vision and Image Understanding*, 107:3–13, 2007.

[7] Eric Dubois. Frequency-domain methods for demosaicking of Bayer-sampled color images. *IEEE Signal Processing Letters*, 12(12):847–850, 2005.

[8] Rich Franzen. Kodak Lossless True Color Image Suite, 2004. Kodak PhotoCD PCD0992 image samples in PNG file format.

[9] Bahadir K. Gunturk, Yucel Altunbasak, and Russell M. Mersereau. Color Plane Interpolation Using Alternating Projections. In *IEEE Transactions on Image Processing*, volume 11, pages 997–1013, September 2002.

[10] Maya R. Gupta and Ting Chen. Vector color filter array demosaicing. In *Proc. SPIE*, volume 4306, pages 374–382, 2001.

[11] K. Hirakawa and T. W. Parks. Adaptive homogeneity-directed demosaicing algorithm. In *Proc. International Conference on Image Processing ICIP 2003*, volume 3, pages III–669–72, 14–17 Sept. 2003.

[12] Bernd Jähne. *Digital Image Processing*. Springer, 2002.

[13] P.A. Jansson, R.H. Hunt, and E.K. Plyler. Resolution Enhancement of Spectra. *Journal of the Optic Society of America*, 60:596–599, 1976.

[14] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*. Springer, 2004.

[15] R. Kimmel. *Numerical Geometry of Images - Theory, Algorithms, and Applications*. Springer, 2004.

[16] Ron Kimmel. Demosaicing: Image Reconstruction from Color CCD Samples. In *IEEE Transactions on Image Processing*, volume 8, pages 1221–1228, September 1999.

[17] Xin Li. Demosaicing by successive approximation. *IEEE Transactions on Image Processing*, 14(3):370–379, March 2005.

[18] Xin Li, Bahadir Gunturk, and Lei Zhang. Image Demosaicing: A Systematic Survey. In *Visual Communications and Image Processing 2008*, volume 6822, pages 68221J–68221J–15, 14–17 Sept. 2008.

[19] L.B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomy Journal*, 79:745–765, 1974.

[20] Jayanta Mukherjee, R. Parthasarathi, and S. Goyal. Markov random field processing for color demosaicing. In *Pattern Recognition Letters*, volume 22, pages 339–351, 2001.

[21] Yoshikuni Nomura and Shree K. Nayar. A vq-based demosaicing by self-similarity. In *Proc. IEEE International Conference on Image Processing*, volume 3, pages 457–460, 2007.

[22] C.A. Poynton. Poynton's colour FAQ, 2006. Web page: http://www.poynton.com/ColorFAQ.html.

[23] W.H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62:55–59, 1972.

[24] D. Tschumperle and R. Deriche. Vector-valued image regularization with PDE's: a common framework for different applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):506–517, April 2005.

[25] D. Tschumperle and R. Deriche. Anisotropic Diffusion PDE's for Multi-Channel Image Regularization: Framework and Applications. *Advances in Imaging and Electron Physics (AIEP)*, pages 145–209, 2007.

[26] Wikipedia. Delta E, 2010. This nomenclature is given on the German Wikipedia site. See http://de.wikipedia.org/wiki/Delta_E.

Author Biography

Johannes Herwig is a PhD student with the intelligent systems group headed by Prof. Josef Pauli at the university of Duisburg-Essen. His research interests revolve around adaptive image processing.

Figure 4. *Each row shows demosaicing results of a zoomed part of an image from the test data set. Each column shows the ground truth color image, the demosaiced image with anisotropic diffusion, the demosaiced image without anisotropic diffusion, and the demosaiced result of bilinear interpolation.*

Figure 5. *This shows results for the intermediate luminance approximation from scalar-valued color sensor data. The first image of each scene pair shows the Gaussian filtered Bayer image, and the second shows the deblurred result taken for a high spatial resolution luminance image.*