Joint High Dynamic Range Imaging and Color Demosaicing

Johannes Herwig and Josef Pauli
University of Duisburg-Essen, Bismarckstr. 90, 47057 Duisburg, Germany

ABSTRACT
A non-parametric high dynamic range (HDR) fusion approach is proposed that works on raw images of single-sensor color imaging devices which incorporate the Bayer pattern. Thereby the non-linear opto-electronic conversion function (OECF) is recovered before color demosaicing, so that interpolation artifacts do not affect the photometric calibration. Graph-based segmentation greedily clusters the exposure set into regions of roughly constant radiance in order to regularize the OECF estimation. The segmentation works on Gaussian-blurred sensor images, whereby the artificial gray value edges caused by the Bayer pattern are smoothed away. With the OECF known the 32-bit HDR radiance map is reconstructed by weighted summation from the differently exposed raw sensor images. Because the radiance map contains lower sensor noise than the individual images, it is finally demosaiced by weighted bilinear interpolation which prevents the interpolation across edges. Here, the previous segmentation results from the photometric calibration are utilized. After demosaicing, tone mapping is applied, whereby remaining interpolation artifacts are further damped due to the coarser tonal quantization of the resulting image.

Keywords: Photometric Calibration, Image Fusion, Demosaicing, Image Acquisition

1. INTRODUCTION
Normally demosaicing\(^1\) is among the first processing steps in color image processing. In this paper a workflow of color segmentation and high dynamic range (HDR) rendering is proposed that places demosaicing at the end of the processing chain. The paper\(^2\) does something similar using a model of retinal processing. The paper\(^3\) computes the HDR image from RAW data of linear sensors (we are additionally calibrating the non-linear response curve) and lies its focus on better color reproduction. The following is essentially a fusion of our color segmentation approach presented in\(^4\) and the HDR calibration presented in\(^5\) which is itself an extension of\(^6\) using the model of the law of film from\(^7\).

1.1 Outline of the Approach
A non-parametric HDR fusion approach is proposed that works on raw images of single-sensor color imaging devices which incorporate the Bayer pattern. Thereby the non-linear opto-electronic conversion function (OECF)\(^8\) is recovered before color demosaicing, so that interpolation artifacts do not affect the photometric calibration. Next, the 32-bit HDR radiance map is reconstructed by weighted summation from the differently exposed raw sensor images. Because the radiance map contains lower sensor noise than the individual images, it is finally demosaiced by weighted bilinear interpolation which prevents the interpolation across edges. Then, tone mapping is applied, whereby remaining demosaicing artifacts are further damped due to the coarser tonal quantization of the resulting image.

\(\ast\) Copyright 2011 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

1.2 Photometric Calibration

Graph-based segmentation\(^9\) greedily clusters the exposure set into regions of roughly constant radiance, whereby firstly for each spatial location pixels with high local contrast, that belong to properly exposed regions, are chosen from the time domain, but secondly the remaining pixels should have small variations in order to belong to the same region\(^5\) (figure 8). The segmentation works on Gaussian-blurred sensor images, whereby the artificial gray value edges caused by the Bayer pattern are smoothed away.\(^4\) The non-linear OECF resembles the law of film\(^7\) and is parametrized by its derivative. For every gray value the slope is robustly estimated by tracking the previously segmented regions throughout the exposure stack and relating them between successive exposures.\(^5\) A variation of the Debevec+Malik algorithm\(^6\) recovers the OECF per RGB plane from its first derivative.\(^5\)

2. DEMOSAICING APPROACH

The raw image that is directly recorded by the sensor of an RGB camera is called the Bayer image, because its single-channel pixels are biased by the optical filter pattern of the Bayer mosaic. To obtain a full vector-valued color image the two missing color components need to be interpolated from the local neighborhood of each pixel. This is known as demosaicing. The usual digital image processing pipeline starts with the demosaiced color image, because demosaicing is often internal to the camera driver and because the single-channel Bayer image is not a straightforward representation for multi-channel color data.

In this paper it is proposed to perform mid-level image processing like color segmentation directly onto the raw Bayer image without previous demosaicing. The segmentation results are then used to parameterize low-level tasks like demosaicing and the creation of high dynamic range images (which both relate to the topic of image acquisition) with scene-dependent higher level information in order to achieve better results having less interpolation artifacts.

2.1 Gaussian Smoothing of Bayer Images

The Bayer representation of a color image has the advantage of being lightweight, because only real measurement data is stored. Another advantage is being single-channel like a gray value image. Full color measurements on the other hand are vector-valued and feature at least three channels if we speak of ordinary RGB images. This triples storage costs but also processing times increase, because image processing algorithms are usually modeled with a single-channel representation of the scene or problem in mind where only spatial features of the distribution of pixel measurements are concerned. Only as a subsequent processing step the correlation of spatial patterns in multiple channels is exploited, or their presence and absence in specific channels leads to significant conclusions about the local spectral characteristics. Therefore the handling of vector-valued images is difficult, because either the different channels need to be projected into some continuous single-channel luminance representation before spatial processing or otherwise every channel needs to be separately processed with the same algorithm and later a fusion of the results is needed. Both approaches to vector-image processing pose the problem of sensible color metrics in order to extract (un)correlated spatial properties from the inter- and intra plane variations of vector-valued data. Additionally, one has to cope with interpolation artifacts from the demosaicing algorithm.

Therefore it may be reasonable to apply color image processing on the Bayer image directly without previous demosaicing. However, the Bayer representation has one great disadvantage which prevents one to do so easily. Because every neighbor of a given pixel is band-passed by a different optical color filter, artificial spatial gradients are introduced w.r.t. every four-neighbor (and eight-neighbor for red and blue center pixels) which stem from the spectral properties of a given image patch. This plays against the aforementioned spatial model of image processing algorithms which often analyze the local spatial change within an image plane to extract interesting features within the signal. Because the Bayer pattern mixes spatial and spectral characteristics within the neighborhood of a pixel and because the sampling density of the different color channels varies (the green channel is sampled twice as dense as the red and blue channels), the common computation of local gradients becomes difficult to implement and is less precise since the step-size to the next pixel of the same color channel depends on the channel itself, the location of the center pixel and the gradient direction. Therefore the error in the gradient computation is not spatially isotropic within a single color channel and also the minimum system immanent error is smaller for green pixels than for red and blue pixels due to their different sampling density.
In this paper it is proposed to smooth the Bayer image by a standard $3 \times 3$ Gaussian kernel $G_\sigma$ which has $\sigma = 0.95$ according to the formula $\sigma = 0.3(n/2 - 1) + 0.8$ with $n = 3$:

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

It has been observed by the authors that the artificial edges in the Bayer image originating from the spectral sampling do fade away or disappear when a human viewer steps farther away from the image display on a computer monitor. This led to the conclusion that scale space filtering of the Bayer image may produce similar results. In fact a single convolution with $G_\sigma$ removed the spectral edges and a continuous but smoothed gray value image resulted from the Bayer representation.

This experimental finding was further verified. Within the Bayer pattern only the following four GRBG, BGGR, RGGB, GBRG patches occur and these $3 \times 3$ neighborhoods are sufficient to create a Bayer mosaic of every image. The patches are shown in figure 1 and those color sampling patterns are also featured as the left-hand matrices of the Gaussian convolutions in equations 2 - 5. These equations demonstrate that the convolution of each patch, regardless of the fact that these have different color sampling patterns, always results in the same weighting factors $R = 0.25$, $G = 0.5$, $B = 0.25$ for the red, green and blue color planes, respectively:

$$\begin{bmatrix} G & R & G \\ B & G & B \\ G & R & G \end{bmatrix} \ast G_\sigma = 0.0625G + 0.125R + 0.0625G + 0.125B + 0.0625G + 0.125R + 0.0625G = 0.25R + 0.5G + 0.25B. \quad (2)$$

$$\begin{bmatrix} B & G & B \\ G & R & G \\ B & G & R \end{bmatrix} \ast G_\sigma = 0.0625B + 0.125G + 0.0625B + 0.125G + 0.0625G + 0.125G + 0.0625B = 0.25R + 0.5G + 0.25B. \quad (3)$$

$$\begin{bmatrix} R & G & R \\ G & B & G \\ R & G & R \end{bmatrix} \ast G_\sigma = 0.0625R + 0.125G + 0.0625R + 0.125G + 0.0625G + 0.125G = 0.25R + 0.5G + 0.25B. \quad (4)$$

$$\begin{bmatrix} G & B & G \\ R & G & R \\ B & G & G \end{bmatrix} \ast G_\sigma = 0.0625G + 0.125B + 0.0625G + 0.125R + 0.0625G = 0.25R + 0.5G + 0.25B. \quad (5)$$

If the real vector-valued but unknown color is assumed to be locally constant over the whole patch, then the convolution result $L = 0.25R + 0.5G + 0.25B$ can be treated as a valid approximation of the luminance at every pixel.

Figure 1. The four local GRBG, BGGR, RGGB, GBRG patches resulting from the Bayer mosaic and their overlaid Gaussians from equation 1. Summation over each patch always results in the same $R = 0.25$, $G = 0.5$, $B = 0.25$ weights (as shown by equations 2 - 5) with which the luminance $L = R + G + B$ can be consistently approximated for each patch if the real but unknown RGB color is assumed to be locally constant over the whole patch.

In fact a single convolution with $G_\sigma$ removed the spectral edges and a continuous but smoothed gray value image resulted from the Bayer representation.
Figure 2. Image 2(b) is obtained by Gaussian smoothing of a synthetically derived Bayer image of the 3-CCD measured original scene 2(a). Figure 2(b) shows qualitatively that the luminance approximation is sensible w.r.t. the original scene 2(a). Figure 2(c) shows qualitatively the difference between the proposed luminance approximation 2(b) and the real CIE Lab luminance values which were obtained from the original 3-CCD measurements 2(a). Thereby figure 2(c) has been equalized for better contrast. It can be concluded that only a region dependent but constant scaling factor makes the difference between the two luminance computations. The unscaled difference between the two luminances is about four gray values per pixel on average. More details can be found in table 1.

<table>
<thead>
<tr>
<th></th>
<th>downscaled</th>
<th>downscaled</th>
<th>original</th>
<th>original</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-smoothed</td>
<td>smoothed</td>
<td>non-smoothed</td>
<td>smoothed</td>
</tr>
<tr>
<td>4.871</td>
<td>3.284</td>
<td>5.852</td>
<td>3.254</td>
<td>avg. mean</td>
</tr>
<tr>
<td>4.721</td>
<td>2.416</td>
<td>6.535</td>
<td>2.454</td>
<td>std. dev</td>
</tr>
</tbody>
</table>

Table 1. This table compares the the CIE Lab luminance computed from ground-truth 3-CCD measurements against the proposed Gaussian luminance approximation of a simulated Bayer image. The results are given per pixel and are averaged over 24 images contained in the Kodak Photo CD database. Four datasets have been created. The first two columns shows results from a test series where the images were downscaled to half their original size in order to reduce noise in the images. The last two columns show results obtained with original image sizes. In order to evaluate the impact of blurring and the reduction of spatial resolution due to the Gaussian luminance approximation, for each of the two datasets the difference images are computed using non-smoothed CIE Lab luminance images at their full original resolution and again with Gaussian smoothed CIE Lab luminances to emulate the same spatial resolution that the Gaussian blurred Bayer image has whereby the same kernel $G_\sigma$ is used. It can be concluded that the effect of reduced spatial resolution in the Gaussian luminance approximation is traceable (4.9 vs. 3.3 and 5.9 vs. 3.3) and that differences between the two approaches increase with noise (4.9 and 3.3 vs. 5.9 and 3.3).

After we found out that spatial low-pass filtering with a Gaussian kernel approximates luminance, we have come across the research of which argues that in the frequency domain of the Bayer image the lower frequencies encode luminance and the higher frequencies encode chrominance. They also derive a filter kernel mathematically but they do not mention that a standard Gaussian kernel is sufficient.

Hence, at the cost of lowering the spatial resolution of the originally sensed image the projected luminance values could be obtained from the single-channel Bayer image directly without previous demosaicing. This ultimately means that we could approximately derive the projection from a three-channel color space onto a single-channel luminance space without ever going three-channel or vector-valued in the first place which saves processing and storage costs. Because the smoothing approach circumvents the need of a demosaicing algorithm, we avoid the introduction of interpolation artifacts and false edges in the luminance image. In our approach image edges are blurred which may be undesired if a high spatial resolution needs to be maintained. However, we have shown in that deblurring of the luminance approximation with the known filter kernel $G_\sigma$ is feasible...
Figure 3. This shows the superiority of the proposed approach to luminance approximation in digital color sensors. Figure 3(a) is a synthetic pattern of white, black and gray stripes that is captured by a color sensor. The stripes are only one pixel wide each. This original color image has been Bayer sampled. The bilinear demosaicing result is shown in 3(b) from which the CIE Lab luminance is derived and shown in 3(c). This is the usual image processing chain to obtain a luminance image from a color sensor. The zipper artifacts and false colors due to the under-sampling of the one pixel wide stripes both introduce false edges in the luminance image. Instead, the blurred Bayer image from figure 3(d) is spatially consistent with the original scene. The original luminance could not be recovered due to the blurring. But at least there is a one-to-one mapping of original luminance to recovered luminance which is not the case with image 3(c).

Figure 4. Details in the demosaiced luminance image on the left compared with the Gaussian smoothed Bayer image. On the left, the bilinear demosaicing of the sensor image introduces color interpolation artifacts near the original gray value edges. In the image of this indoor scene the so-called zipper effect is apparent which produces a path of alternating colors along real gray value edges. The image on the right has been obtained by smoothing the original Bayer image of the sensor with a Gaussian filter, instead. This technique proves that in the Bayer image the luminance and chrominance information of the scene are multiplexed. The resulting image is free from artifacts and approximates the CIE Lab luminance image on the left which has been obtained by traditional means from the interpolated color image.

and the original spatial sensor resolution may be restored. On the other hand, blurring the Bayer image reduces noise which may be a welcome side-effect, although noise of different channels is mixed here which may again be undesired if the noise floors are very different for each channel but this may possibly be calibrated.

2.2 Color Segmentation

For segmentation of the blurred Bayer image the graph-based algorithm proposed by⁹ is used. In their paper it is suggested to smooth the input image via a Gaussian filter in order to reduce noise, so that the blurred luminance image does not pose a problem regarding its lower spatial resolution but fits exactly the quality required as an input. The smoothed Bayer image serves as a guidance, because its spatial resolution is continuous and local neighborhoods are not disturbed by the additional spectral sampling of the original Bayer image as it is depicted in figures 3 and 4. The graph-based segmentation works in a greedy fashion, and makes decisions whether or not to merge neighboring regions into a single connected component based on some cost function. The following gives a brief outline of their approach.

2.2.1 Graph-based Segmentation of Single-channel Images

A graph \( G = (V, E) \) is introduced with vertices \( v_i \in V \), specifically the set of pixels, and edges \( (v_i, v_j) \in E \) corresponding to the connection of pairs in a four-neighborhood. Edges have nonnegative weights \( w((v_i, v_j)) \).
corresponding to the gray value difference between two pixels. The idea is, that within a connected component,
edge weights, as a measure of internal difference, should be small and that in opposition edges defining a border
between regions should have higher weights. Thereby the internal difference of a component \( C \subseteq V \) with
\( \text{MST}(C, E) \) as its minimum spanning tree is defined as its highest edge weight \( w(e) \), which is the highest gray
value difference between two neighboring pixels:

\[
\text{Int}(C) = \max_{e \in \text{MST}(C, E)} w(e).
\]

(6)

On the other hand, the difference between two components \( C_1, C_2 \subseteq V \) or neighboring region segments is
defined as the minimum edge weight between any two pixels \( v_i \) and \( v_j \) that connect these components:

\[
\text{Diff}(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w((v_i, v_j)).
\]

(7)

If there is evidence for a boundary between two neighboring components, then the comparison predicate:

\[
D(C_1, C_2) = \begin{cases} 
\text{true,} & \text{Diff}(C_1, C_2) > M\text{Int}(C_1, C_2) \\
\text{false,} & \text{otherwise}
\end{cases}
\]

(8)

evaluates to true, whereby \( \text{Diff}(C_1, C_2) \) denotes the smallest difference between two components \( C_1, C_2 \subseteq V \),
and \( M\text{Int}(C_1, C_2) \) is the minimum internal difference of both components:

\[
M\text{Int}(C_1, C_2) = \min(\text{Int}(C_1) + \tau(C_1), \text{Int}(C_2) + \tau(C_2))
\]

(9)

where the internal difference \( \text{Int} \) denotes the maximum edge weight within a component and \( \tau \) is an additional
threshold function determining the degree to which the difference between two components must be greater than
their internal differences in order for there to be an evidence of a boundary between them, i.e. \( D \) evaluates to
true. The threshold function \( \tau(C) = k / |C| \) depends on the size \( |C| \) of a component where \( k \geq 0 \) is some constant
parameter determining the scale of observation, so that larger \( k \)'s favor larger components.

The segmentation algorithm is applied onto the blurred Bayer image with some fixed threshold parameter
\( k \). In a post-processing step connected components that do not contain at least one pixel location for every
subsampled color channel are iteratively merged with their largest neighboring region until this condition holds.

### 2.2.2 Graph-based Color Segmentation of Bayer Images

While the segmentation algorithm explores the spatial structure of the image through the smoothed Bayer image,
missing samples of the individual color channels are bridged, and at the same time it becomes possible to identify
locally homogeneous regions in the incomplete color planes of the originally sensed Bayer image. Therefore, while
a segmented region greedily grows within the smoothed Bayer image, we can make sense of the underlying color
information which is contained in the original spectrally sub-sampled Bayer image. The idea is to create a single
color vector for each segmented region which stores its mean color values. In a second color vector the absolute
variances within each color plane are updated and if one of the intra-plane variances exceeds some threshold the
region would not be allowed to grow any longer even if the variance of the guiding luminance channel is still
below its threshold.

In order to incorporate this behavior into the graph-based segmentation algorithm the comparison predicate,
which indicates a region border, has been modified and replaced by

\[
D^{LRGB}(C_1, C_2) = \begin{cases} 
\text{true,} & \exists \delta \in \{L, R, G, B\} \quad D^\delta(C_1, C_2) \\
\text{false,} & \text{otherwise}
\end{cases}
\]

(10)

where \( D^L(C_1, C_2) = D(C_1, C_2) \) for luminance is unchanged, and for color it becomes
Figure 5. This shows the color segmentation results of the graph-based segmentation algorithm with the cost functions that are derived in equations 10 - 14. In order to better recognize the single regions these are randomly colorized in figure 5(a). In figure 5(b) the very same segmentation result has its regions labeled with their median color which is obtained from the original Bayer image pixels during the segmentation process. Thereby for every color plane the median value that lies within each region is independently selected for each color channel. Because the intra-variance of each color channel within a region is restricted to be in the lower range, it can be assumed that the color is roughly constant over each region. Therefore the median already gives a good approximation of the real color value within a region. The median colors are later used as an initial guess for refined pixel-wise demosaicing.

\[
D^\delta_{\{R,G,B\}}(C_1, C_2) = \begin{cases} 
\text{false,} & |C_1|^{\delta} = 0 \lor |C_2|^{\delta} = 0 \\
\text{true,} & Dif^\delta_{\text{var}}(C_1, C_2) > MVar^\delta(C_1, C_2) \\
\text{false, otherwise}
\end{cases}
\]  

(11)

\[
Dif^\delta_{\text{var}}(C_1, C_2) = \max_{\delta}(C_1) - \min_{\delta}(C_2) 
\]  

(12)

\[
MVar^\delta(C_1, C_2) = \min(Var^\delta(C_1) + \tau^\delta(C_1), Var^\delta(C_2) + \tau^\delta(C_2)) 
\]  

(13)

\[
Var^\delta(C) = \max_{\delta}(C) - \min_{\delta}(C) 
\]  

(14)

where \(|\cdot|^{\delta}\) is the amount of pixel locations contributing to color channel \(\delta\) in the currently segmented region, and \(\max^{\delta}(\cdot)\) and \(\min^{\delta}(\cdot)\) denote the maximum or the minimum color value of channel \(\delta\) contained in any of the given regions. Similarly to the original paper the variance threshold \(\tau^\delta(C) = k^\delta / |C|^\delta\) depends on the size of the region. The larger a region grows the stronger it gets panelized to introduce greater variance into the color channel through region merging. It is noted that \(D^L\) is tentatively evaluated on the smoothed Bayer image and then \(D^\delta_{\{R,G,B\}}\) solely operates on the original unsmoothed Bayer image. Only when both comparison predicates evaluate to true the two adjacent regions are merged. An exemplarily result of the color segmentation algorithm that directly works on the Bayer image is shown in figure 5.

### 2.3 Pixel-wise Demosaicing Using Higher Level Segmentation Results

One of the most often exploited properties in demosaicing is the color ratio rule, which states that the ratio of any two colors remains locally constant. Therefore, if we are on a green Bayer image pixel, for example, and we want to interpolate the missing red and blue pixels, we first compute the ratio between the original green value and the median green value of its corresponding region that we obtained from the color segmentation result of the previous section. This gives us some kind of distance measure between the current pixel and its local average. It is further assumed that the distance ratio between the originally sensed green value and its estimated green value (i.e. the median green obtained through the segmentation process) will be roughly the same as the distance ration between the red or blue value at the current pixel location and their estimated median values. Hence, at a green pixel the missing red and blue values can be estimated as value\(_R = \text{median}^R(C) \cdot \text{ratio}^G\) and value\(_B = \text{median}^B(C) \cdot \text{ratio}^G\).
Figure 6. This figure compares our demosaicing results from figure 6(d) with the naive bilinear approach 6(e) and the original image 6(f). It can be concluded that the proposed approach produces sharper more detailed results, but it also suffers from salt and pepper noise. Figures 6(a) - 6(c) show the simulated input and intermediate segmentation results.

median^B(C) · ratio^G with ratio^G = value^G / median^G(C). Similar computations are made on original red and blue pixel locations, so that the missing color components can be estimated.

An exemplarily demosaicing result is shown in figure 6(d). Although the results still contain salt and pepper noise, the demosaicing results are distinctively sharper than the naive bilinear demosaicing results as shown in figure 6(e). This is the case because the proposed approach does not interpolate the missing components from its spatial neighborhood which is effectively a smoothing operation. But rather it fits the full vector-valued data of each pixel to the overall texture statistics (i.e. the median color) of segmented regions defined by low-variance. Thereby sharper deviations of a single pixel w.r.t. to its nearby neighborhood are still possible because the pixel is related to the local texture statistics only by its ratio^δ, and if the ratio of the known color component δ ∈ {R, G, B} is farther from 1, then the estimated color components do have the same deviation from their respective median color values. Additionally, missing color components are only estimated w.r.t. to their corresponding regions from the higher-level segmentation result, so that pixels near the boundaries of neighboring regions are not interpolated across distinctive color segments.

3. HIGH DYNAMIC RANGE (HDR) APPROACH

The problem of high dynamic range imaging is the photometric calibration or reconstruction of the non-linear response curve of a digital imaging device. The response curve maps real-world radiance values of a scene onto digital gray values of images. Since 8-bit images cannot store the wide dynamic range information which is present in the real world, the response curve is an S-shaped mapping that compresses the lower and higher dynamic ranges but preserves the mid-level radiances corresponding to the dominant dynamic range of the scene.
Figure 7. This is the photometric model of the HDR approach. Figure 7(a) schematically shows the image capture with its relevant photometric parameters. When a scene is photographed the irradiance $E_i$ at each scene point $i$ is integrated with the exposure time $\Delta t_j$. Their product $E_i \Delta t_j$ is mapped via the response function $f(\cdot)$ to gray values $Z_{ij}$. The qualitative shape of the response curve is shown in figure 7(b). Its slope can be modeled after the so-called law of film. The slope is steepest around mid-level gray values and therefore produces the best contrast here. The contrast is compressed at the tail and shoulder of the S-shaped curve. Hence, the mapping of radiance values to digital gray values is non-linear.

**3.1 Recovery of the Response Curve**

In order to estimate the response curve and to recover the full 32-bit high dynamic range of the real radiance map multiple captures under varying exposure times are made according to the photometric capture model depicted in figure 7(a). Then the gradient of gray values between successive exposures is estimated for each digital gray level whereby the digital gradient is related to the modeled gradient of the S-shaped response curve w.r.t. the difference in the known exposure times. Thereby the inverse of the response curve is reconstructed. With the known inverse the hidden radiances can be recomputed from the digital gray values and their known exposure times which finally gives us a single 32-bit radiance map.

In the former higher-level description of the HDR approach the computation of digital gray value gradients poses a problem which can be solved using the proposed color segmentation of Bayer images. In theory the gray value change of each pixel $i$ could be tracked between every two exposures along the exposure series and related to the gradient of the to be recovered response curve. In order to do this every image within the exposure stack needs to be registered exactly and every change of the gray value must be related to the change of exposure time but not changes in the scene, sensor noise or color interpolation artifacts. Furthermore, the sheer amount of pixels would make it impossible to solve such a huge system of equations. Therefore one has to select a tiny but sufficient amount of pixels that are tracked throughout an exposure series and whose gradients pointing along the time direction are followed. Here, the proposed color segmentation on Bayer images helps.

**3.2 Color Segmentation of an Exposure Stack of Bayer Images**

Since we propose a RAW workflow for non-linear color HDR the Bayer images are smoothed through Gaussian filtering at the beginning, so that no demosaicing is necessary at first. If needed the smoothed exposure stack can be registered at this point using. Then we segment the exposure stack of blurred Bayer images which has an additional third time dimension at once. The result is a collapsed two-dimensional segmentation as shown in figure 8(c). In order to collapse the time dimension we replaced the original edge weight $w(\cdot)$ with:
Figure 8. The graph-based segmentation of the scene into regions of nearly constant radiance is depicted. The segmentation works in the three-dimensional space which is spanned by pixel locations and exposures, but produces a single two-dimensional segmentation result of homogeneous radiance regions directly from the Gaussian smoothed sensor images. Here, the nodes represent gray values, thicker edges indicate borders of regions but thinner edges connect adjacent pixels to a common region. Before the spatial fusion of regions, the pixels are collapsed along the exposure axis, so that underexposed dark pixels and overexposed bright pixels without contrast are discarded.

The extended edge weight \( w_p(\cdot) \) finds the maximum edge weight \( w(\cdot) \) between the two neighboring pixels \( v_i \) and \( v_j \) within any of the exposures \( p \in P \). Since an edge weight is computed as the gray value difference between two pixels, this approach ensures that only the high contrast edges of correctly exposed images \( p \) are selected. Therefore the information from overexposed or underexposed local image patches is effectively discarded, because constantly bright or dark areas do not convey any relevant texture information. Figure 8(b) schematically motivates the extended edge weight \( w_p(\cdot) \). A similar extension is made to the variance computation concerning the original color values from the Bayer image:

\[
Var^\delta_p(C) = \max_{p \in P} \delta_p(C) - \min_{p \in P} \delta_p(C)
\]

So finally, \( Var^\delta_p(C) \) replaces \( Var^\delta(C) \), and \( \max_{p}^\delta(C) \) and \( \min_{p}^\delta(C) \) respectively denote the maximum and minimum value of color \( \delta \) within region \( C \) over all exposures \( P \).

### 3.3 Creating the Radiance Map and Demosaicing

For the photometric calibration and the estimation of the slope of the response curve a subset of preferably large and low-variance regions is selected from the segmentation result. The changes of their median gray values (for each color plane separately) are tracked through the time dimension. Then for each color plane separately the three response curves are estimated by solving a linear system of equations. After the recovery of the response curves the 32-bit radiance map can be computed from the original set of exposures.

The radiance map still has its Bayer representation, because until now no color vectors have been interpolated. The demosaicing takes place in radiance space. Thereby the segmentation results are used to perform the demosaicing similarly to the herein proposed approach for 8-bit space. For each region we compute the 32-bit median color values and then interpolate the missing colors as in our pixel-wise demosaicing approach.
4. CONCLUSION

The photometric calibration of non-linear OECFs utilizes only raw sensor images. Gaussian smoothing of the sen-
sor images approximates their luminance, so that these can be segmented before demosaicing. The unsupervised
segmentation allows for automatic regularization of the OECF estimation. Demosaicing uses joint segmentation
results from multiple exposures at once in order to robustly prevent the interpolation across edges. The proposed
luminance approximation of raw Bayer images through Gaussian smoothing allows one to design a single-channel
image processing chain for color images apart from the traditional vector-based approach.

REFERENCES

[1] Gunturk, B. K., Altunbasak, Y., and Mersereau, R. M., “Color plane interpolation using alternating pro-

range images using a model of retinal processing,” in [Proceedings of SPIE-The International Society for


regularization,” in [VISAPP], 1, 539–546 (2009).

[6] Debevec, P. E. and Malik, J., “Recovering high dynamic range radiance maps from photographs,” in [SIG-
GRAPH ’97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques],

Differently Exposed Pictures,” in [IS&T’s 48th annual conference Cambridge, Massachusetts], 422–428,
IS&T (May 1995).


in PNG file format.

Chromatic Signals in the Fourier Domain,” in [Proc. IS&T/SID 10th Color Imaging Conference], 10, 331–
336 (2002).

Basic Idea: Perform an initial segmentation before color demosaicing using blurred Bayer images. The segmentation result regularizes the demosaicing and photometric calibration.