Spatial Gaussian Filtering of Bayer Images with Applications to Color Segmentation

Johannes Herwig and Josef Pauli

Universität Duisburg-Essen,
Bismarckstr. 90, D-47057 Duisburg
eMail: {johannes.herwig, josef.pauli}@uni-due.de
URL: http://www.uni-due.de/is/

Abstract. A single sensor color imaging device has a color filter array (CFA) laid on top of its photodiodes, which spatially samples bandpassed spectral responses. Hence, with the popular Bayer pattern, at every pixel site either red, green or blue light is measured. A process known as demosaicing interpolates the vector-valued color image from the scalar-valued sensor output, termed here as the Bayer image. In practice image processing algorithms are then applied onto the full-featured vector image, only. The vector-valued nature of color images makes tasks like segmentation difficult. Also demosaicing is computationally intensive and fast algorithms, like bilinear interpolation, introduce color artifacts. Therefore a color segmentation algorithm is proposed, that works solely on the scalar-valued Bayer image without the need of demosaicing beforehand. Firstly, it is motivated and then qualitatively and quantitatively verified that filtering the Bayer image with a Gaussian gives a good approximation of a monochrome luminance image of the scene. Secondly, the filtered Bayer image is segmented into regions of small gray-value variances using a graph-based segmentation algorithm. Finally, the mean intensity value of pixels comprising a segmented region is separately computed for each color channel from the originally sensed Bayer image. Such a synthesized color vector is then taken as the label for each segmented region. Results for varying parameters of the segmentation algorithm and alternative methods of color determination with their new and corresponding image processing chains are presented.

1 Introduction

A widely used method for region-based color segmentation is to minimize the euclidean distance in a color space like CIE LAB that mimics human perception. In gradient-based approaches Di Zenzo's multispectral gradient [18] or the maximum gradient of the color planes is computed at a pixel in RGB space [16, pp. 302]. Locally, if gradient magnitudes are beneath some threshold, the color is likely to be the same. Both approaches perform segmentation by clustering, the former with respect to color features in some suitable color space using an accompanying color metric, whereas the latter clusters a feature space of intensity gradients or orientation fields [16, pp. 332]. Scale-space filtering for color images is accomplished within the anisotropic diffusion framework [17] utilizing Di Zenzo's gradient. There is a graph-morphology scale-space for color image segmentation, that produces a tree-based image representation via successive area openings and closings of increasing scale [8, 9]. For the area operator to work extremal regions have to be identified with respect to their neighborhood. Existing methods and metrics for determination of extrema are based on a local convex hull in color space, or on reduced ordering by projecting the multivariate color data onto a scalar image before extrema are extracted [9]. Another proposed metric uses euclidean distance within a color space and if there are neighbors of identical color distance they are further ordered by luminance, and - if it still does not resolve - may finally be ordered by their R, G, B values [8]. Because the morphological scale-space approach is based on regional extrema, similar to cluster-based segmentation there is a need for some applicable metric for vector data.

Unlike with grayscale images, there is no natural ordering for vector-valued data, which makes working with color images difficult and distance metrics are inherently arbitrary. Sensor measurements are usually given in R, G, B values, but from the previous examples it is apparent that transformations of the data into other more attractive color spaces, vector-valued or

scalar feature spaces are common preprocessing steps. All the discussed approaches make use of either the marginal ordering or reduced ordering paradigm [16, pp. 78], or a combination of both (e.g. euclidean distance, Di Zenzo's gradient). Those transformations can be thought of solutions of the data fusion problem of color planes, so that an ordering relation can be established. The crux of the matter is, that in an image processing chain, when a single-chip color sensor is used, there already is such a scalar image available that may be regarded as one possible fusion result of color channels. It is the measurement matrix of the sensor itself, that captures an image of a scene seen through an optical filter mosaic spatially sampling bandpassed spectral responses at every pixel site by a repeating pattern. The most widely color filter array used is the Bayer pattern [3]. Therefore the sensed result is called the Bayer image.

Here it is shown how to utilize the scalar-valued Bayer image obtained from the sensor for region-based color segmentation. When a reduced ordering relation is sufficient, the proposed algorithm has its image processing chain working entirely on single-channel images, only the final color labeling is vector-valued. Its advantage is reduced complexity, processing time and data transfer rates compared to methods discussed above that are carried out on vector-valued images or perform dimension reduction. Also demosaicing is not needed in advance, rather, segmentation and color interpolation are performed simultaneously. Colors of segmented regions are estimated by mean pixel values, however, the accuracy achieved may be suitable for e.g. locating or tracking colored markers in mobile robot applications. One might think, that there is a loss of information when only scalar sensor data is processed instead of a full-color vector image, but this need not be the case here. In fact reduction to the scalar image is at no cost, since $2/3$ of each vector of a color image would have been interpolated via demosaicing: Discarded vector elements contain no extra information other than that of interpolation errors.

This paper is organized as follows: In the next section demosaicing is described and its relationship to color segmentation is explored. Then spatial Gaussian filtering of the Bayer image is introduced, that enables segmentation of the sensor output directly into locally homogeneous connected regions - much like segmenting a grayscale image. Thereafter different image processing chains are proposed for combined segmentation and color interpolation incorporating reduced ordering only or both marginal and reduced ordering to make the color clusters internally more consistent. Results obtained with the new image processing chain are evaluated, and finally, the conclusion is drawn.

2 Color Interpolation via Demosaicing in the Light of Missing Data Problems

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 by establishing 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 [12, 13, pp. 152]. This ignores specularities and implies isotropic luminance of objects. Hence, in the bricks world, a luminance image - and as seen later each color plane - is locally homogeneous in the spatial domain. Furthermore, the albedo is a property of the material of an object and also depends on the wavelength of reflected light. Thus color channels are linearly dependent and aligned in the spectral domain. The measurement of a particual color channel is proportional to the normalized shading image due to the overall reflected light of an object. If it is 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 [14], 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 [7] by [10].

On the other hand demosaicing can be seen as a missing data problem. Two distinct contexts of missing data are suggested in [6, p. 355]: Firstly, some elements in a data vector are missing for some instances and present for others, whereas secondly, an inference problem would be much simpler using some variables whose values are unknown or hidden. Interestingly, demosaicing fits in both contexts of missing data. The first case maps to the subsampled color channels due to the CFA pattern. The second case can be interpreted as having a segmentation of the full-color image at hand demosaicing would be much simpler, because inter- and intra-channel dependencies were known. Thinking of segmentation by clustering without loosing spatial information the same Mondriaan world model as already discussed with additional edge detection capabilities [4] is applicable. This theoretical treatment of CFA interpolation in the light of missing data problems readily integrates with current demosaicing methods, that spatially smooth within expected homogeneous regions but aim to avoid smoothing over edges in order to reduce color artifacts and blurring.

Because demosaicing and region-based color segmentation make use of the same model assumptions, it is reasonable to drop the intermediate CFA interpolation process and instead perform color clustering onto the Bayer image directly. Since color channels are spatially sampled within the Bayer image, the color planes have missing elements and therefore need to be segmented jointly, because color edges may be hidden at locations where no data is available for a given plane. In the following a scalar feature space with full spatial resolution over the Bayer image is introduced that circumvents this problem and supports reduced ordering.

3 Luminance Approximation via Spatial Gaussian Filtering of Bayer Images

Direct segmentation of a Bayer image into connected regions based on spatial and spectral criteria is arguably impossible due to its CFA subsampled nature. Generally, neighboring pixels are not measures of the same spectral channel, except in an eight-neighborhood, where diagonal pixels are always of the green channel when the current pixel is green. The fact that the green channel is sampled twice as dense as the other two and its quincunx pattern make interpolation relatively easy when spectral alignment with other color planes is initially ignored. 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. Another approach to spectral alignment of neighboring pixels is to view the red, blue and the two green planes that exhibit the same rectangular CFA pattern as four separate images of equal size. These images suffer from high spatial aliasing and would have to be registered and interpolated to be merged into a single luminance image.

Instead, convolution with a smoothing kernel is proposed that addresses problems with the aforementioned approaches. 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 caused by 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 $\sigma = 0.95$ is

$$G_\sigma = \begin{bmatrix} 0.0625 & 0.125 & 0.0625 \\ 0.125 & 0.25 & 0.125 \\ 0.0625 & 0.125 & 0.0625 \end{bmatrix}.$$

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

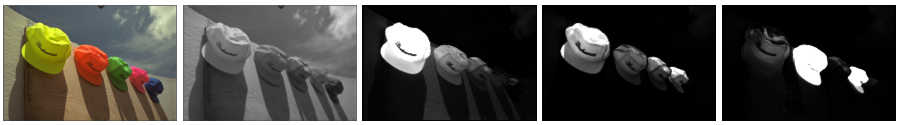

Fig. 1: The original color image, the Gaussian filtered Bayer image, and difference images.

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 [11]. 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. Applying weighted averaging of color samples according to the filter kernel, and under the assumption that colors are locally the same, one gets luminance $L = 0.25R + 0.5G + 0.25B$ for all of them at a given center pixel.

In fig. 1 the luminance property of a ground truth color image, which has been captured using 3-CCD technique, is exemplarily analyzed to verify the previous claims. The original color image is downscaled half in width and height by the Burt and Adelson pyramid. Then it has been subsampled according to the Bayer pattern. The resulting Bayer image is smoothed as described and is shown to the right. The following images in fig. 1 show the differences between the smoothed Bayer image and non-linear CIE luminance (Rev. 601) [15], naive RGB-to-monochrome transform by $(R+G+B)/3$, and the L_2-norm of RGB vectors, respectively, which are 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 tab. 1, where the average of differences of all images in the dataset from [7] 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, so that the result is close to a luminance computation as if the full-color vector were known at the center pixel itself, whereby errors are spatially correlated and are larger at edges of objects due to smoothing. These observations match preliminary considerations on the physical color formation model.

Even if model assumptions are violated, Gaussian filtering is robust. This can be concluded from fig. 2 where a synthetic test image of grayscale stripes with a width of only one pixel each is analyzed. It has been subsampled according to the Bayer pattern. The resulting bilinear interpolation is shown next. Then non-linear CIE luminance is computed from the interpolated color image, and is shown thereafter for comparison. The last result is the smoothed Bayer image. From this it is apparent that the structure of the original synthetic scene is preserved and that grayvalues are the same where the original scene has the same brightness. Both is not true for bilinear demosaicing. This is because due to the specific pattern it is possible at best to interfere two colors of the RGB color vector correctly, but only if interpolation is restricted to be vertically only. Since bilinear interpolation is without edge-sensing there are false colors. The third color always remains unknown, but because the sensor would not know about the stripes pattern beforehand it interpolates from horizontal neighbours of a pixel, too. The zipper effect comes from the fact that green color samples are diagonally ordered in the Bayer CFA.

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

Tab. 1: Comparison of three widely used luminance measures with the new Gaussian luminance approximation of a Bayer image.

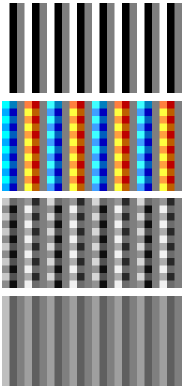

Fig. 2: Synthetic pattern results.

4 Simultaneous Color Segmentation and Interpolation via Luminance Cue

The Gaussian filtered Bayer image is segmented with the graph-based algorithm presented in [5]. The algorithm 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.

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. If there is evidence for a boundary between two neighboring components, the comparison predicate evaluates to true,

$$D(C_1, C_2) \quad = \quad \begin{cases} \text{true}, & Dif(C_1, C_2) \quad > \quad MInt(C_1, C_2) \\ \text{false}, & \text{otherwise} \end{cases}$$

where $Dif(C_1, C_2)$ denotes the smallest difference between two components $C_1, C_2 \subseteq V$, and $MInt(C_1, C_2)$ is the minimum internal difference of both components,

$$MInt(C_1, C_2) \quad = \quad min(Int(C_1) + \tau(C_1),\ Int(C_2) + \tau(C_2))$$

where the internal difference Int denotes the maximum edge weight within a component and τ 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 evidence of a boundary between them, i.e. D to be true. The threshold function depends on the size $|C|$ of a component, $\tau(C) = k/|C|$, where $k \geq 0$ is some constant parameter determining the scale of observation in that larger k favour larger components.

The segmentation algorithm is applied to the filtered 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. The idea of local intra- and inter-correlation of color channels is then exploited, as in various demosaicing algorithms [14], to interfere the color of each segmented

region. Assuming that the channel-wise variance within such a region is small, depending on the observation scale k and the validity of the alignment property of color channel gradients, it is justified to combine individual channels by taking the mean of each channel and fusing these color components into a single synthetic color vector for each segmented region. This procedure uses reduced ordering for color segmentation solely working in the luminance domain. Since the scalar luminance values are in the same range of values as each separate color plane multiple different colors result in the same luminance projection. Also separation of colors suffers in a compressed luminance space due to the loss of dimension. Additionally, the corresponding image processing chain is sequential in that color interpolation is done after the segmentation of the luminance image has been finished. Nothing is gained from having both the luminance approximation and the spatially sampled measurements of color channels of the original Bayer image at hand. While the segmentation algorithm explores the spatial structure of the image through the smoothed luminance information, 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 measurements.

Hence a second image processing chain is proposed that performs segmentation on a four dimensional space consisting of luminance and a color vector with missing elements, whereby the luminance plane is itself a scalar projection of the three incomplete color channels. Then marginal ordering is used, so that during segmentation the comparison predicate D is evaluated separately for each of the four planes. Therefore the comparison predicate, which indicates a region border, has been modified and replaced by

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

where $D^{L}(C_1,C_2) = D(C_1,C_2)$ for luminance is unchanged, and for color it becomes

$$D^{\delta \in \{R,G,B\}}(C_1,C_2) \quad = \quad \begin{cases} \text{false}, & |C_1|^{\delta} = 0 \quad \vee \quad |C_2|^{\delta} = 0 \\ \text{true}, & Dif_{Var}^{\delta}(C_1,C_2) \quad > \quad MVar^{\delta}(C_1,C_2) \\ \text{false}, & \text{otherwise} \end{cases}$$

$$Dif_{Var}^{\delta}(C_1,C_2) = max^{\delta}(C_1,C_2) \; - \; min^{\delta}(C_1,C_2)$$

$$MVar^{\delta}(C_1,C_2) \quad = \quad min(Var^{\delta}(C_1) + \tau^{\delta}(C_1),\, Var^{\delta}(C_2) + \tau^{\delta}(C_2))$$

$$Var^{\delta}(C) \quad = \quad max^{\delta}(C) \; - \; min^{\delta}(C)$$

where $|\cdot|^{\delta}$ is the amount of pixel locations contributing to color channel δ in a segmented region, and $max^{\delta}(\cdot)$ and $min^{\delta}(\cdot)$ denote the maximum or the minimum color value of channel δ 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 introducing greater variance in a color channel through region merging gets panelized stronger. It is noted that D^{L} is tentatively evaluated on the smoothed Bayer image and then $D^{\delta \in \{R,G,B\}}$ solely operates on the original unsmoothed Bayer image to finally realize the joint comparison predicate.

With the former vector image processing chain color clustering has been introduced on the scalar Bayer domain. At the additional cost of segmentation of vector data the variance threshold allows to define a coherence requirement per color channel that tackles the loss of information when only luminance data is concerned. Although correlation of color channels is not explicitly enforced, a low variance threshold k^{δ} for all channels supports the idea.

In another conceivable image processing chain the smoothed Bayer image could be used for grayscale image analysis at first. After a region of interest has been identified by e. g. contour extraction or other feature detection methods, one can retrieve the color information from

the original Bayer image for further classification or hypothesis testing. For regions expected to be homogeneous one can create a synthetic color vector by averaging per color channel over the region. Alternatively, one of the segmentation algorithms described earlier can be performed on that specified region alone. Contrary to a sequential color image processing chain during image acquisition this enables active decisionmaking on whether to perform color interpolation based on relevance of parts of an image for the analysis task. Then the processing chain only becomes vector based where needed and transformations from color to luminance space to perform certain algorithms become needless. This may be interesting to realtime systems for tracking different colored markers of equal contour.

5 Evaluation of Color Segmentation on Bayer Images vs. Full-Color Images

Because in order to maintain color correctness the vector-valued segmentation is superior to the scalar luminance approach as noted earlier, the evaluation is based on those results only. Segmentation results of the second new type of image processing chains introduced in the previous section are shown in fig. 4. The first column shows the downscaled ground truth 3-CCD color image, also as described previously, from [7] for comparison. Then the finer scale result of the simultaneous Bayer image segmentation and color interpolation via the graph-based algorithm with parameters $k = 50$ and $k^{\delta} = 500$ follows. Next, segmentation is at a coarser scale with $k = 500$ and $k^{\delta} = 5000$. The k^{δ} have relatively large values, so that they only affect large growing regions, and behave as a cutoff to avoid false colors due to great variances within and between color channels. The segmentation results are labeled with their synthetic color vectors, and therefore can be regarded as demosaiced multichannel images at different observation scales. The segmentation by clustering approach used here has the ability to preserve spatial information while color similarity is scale-dependent. Still, this is not a scale-space approach, since segmentation is greedy and new scales cannot be produced within an iterative chain with growing scale, but rather the graph-based algorithm is run from scratch for every different scale parameter.

For evaluation the mean intensity values of the synthesized color vectors where regions are labeled with are calculated in two different ways. Firstly, the new image processing chain with simultaneous color segmentation and demosaicing is performed, and mean intensity values are computed from the original Bayer image measurements. Secondly, a traditional image processing chain with sequential demosaicing and subsequent segmentation on a multichannel color image is simulated. The segmentation result from the newly introduced algorithm has been wrapped onto a color image already interpolated via bilinear demosaicing. Then the synthesized color vectors for labeling are computed from the interpolated vector data using intensity values for every channel that are available at every pixel. Resulting images look like the ones in fig. 4. For both results a difference image with the ground truth color data is created. The average of the root mean square (RMS) error and the standard deviation for all images from [7] are shown in tab. 2. The results for bilinear demosaicing without segmentation are on the very left for comparison, and then results are given for the traditional vector-valued segmentation and correspondingly for the new simultaneous segmentation and color interpolation - both at different scales denoted by their parameters. Of course bilinear demosaicing without segmentation produces the closest image to ground truth. But when segmentation is concerned the new simultaneous approach outperforms the traditional sequential image processing chain in all cases due to interpolation errors that bilinear demosaicing has introduced. Tab. 2 also gives the average amount of segmented regions produced and a compression rate with respect to the full-resolution pixel image in order to relate the RMS error to data reduction performance. This can be also evaluated visually from fig. 5, where the skeleton images of regions resulting from the segmentation algorithm are shown. These results correspond to two of the scenes shown in fig. 4 with the same scale parameters. Given that the coarser segmentation only has approximately one twentieth the amount of regions as the finer scale (see

Bln.	Bln. Smt. $k = 50$ $k^\delta = 500$		Bln. Smt. $k = 300$ $k^\delta = 3000$		Bln. Smt. $k = 500$ $k^\delta = 5000$		Bln. Smt. $k = 1000$ $k^\delta = 10000$		
	1529		398		272		159		Regions
	64.293		246.995		361.412		618.264		Compression
2.457	4.086	3.027	4.929	3.938	5.270	4.310	5.894	4.999	RMS error
11.747	9.563	6.730	10.033	7.905	10.120	8.343	10.708	9.324	Std. dev.

Tab. 2: Evaluation of the new simultaneous segmentation and color interpolation algorithm against the traditional sequential approach with bilinear demosaicing. For results of both algorithms difference images with ground truth color data have been computed for all channels and root mean square errors are shown here for different segmentation scales. The first column gives results for the bilinearly demosaiced color image at full pixel resolution. Then results for both bilinear-sequential and simultaneous segmentation approaches are given.

also tab. 2), for some scenes the coarse result is perceptually very close to the original when regions are labeled with colors (compare results from fig. 4). Fig. 3 shows regions alernatively labeld with chromacities instead of R, G, B values, where ground truth is on the left and the following two results correspond to the segmentation scales in fig. 4.

6 Conclusion

The idea of direct spatial Gaussian filtering of the Bayer image measured by a color sensor has been introduced. Outgoing from the finding that the filter result approximates a luminance image a graph-based segmentation has been applied that simultaneously justifies region-based color interpolation and classification. The processing chain can thereby be based entirely on single-channel images. Segmentation results show significant correlation with the original ground truth color image. It has been shown that simultaneous segmentation and demosaicing results in more accurate color reproduction through clustering than a sequential bilinear demosaicing followed by color averaging. New color image processing chains have been introduced whereby the need of computationally costly transformations from the scalar Bayer domain into the multivariate color domain, and from there back into scalar valued luminance domain are reduced. Through fast Gaussian filtering in the scalar domain of the Bayer image it becomes possible to work with the luminance approximation at a short response and low bandwidth, but color is still available for regions of interest to allow for higher level image analysis after lower level tasks like edge or feature detection are able to be worked out on the luminance approximation. With simultaneous segmentation and demosaicing a possible application in the spirit of this type has been developed and the two formerly different problem domains of color interpolation and segmentation have been integrated. The ideas presented here may be applied to robot or machine vision problems that require a color sensor but where solutions must be efficiently implemented at the same time.

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, July 1976. US-Patent 3 971 065.

[4] Andrew Blake and Andrew Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, MA, 1987.

[5] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient Graph-Based Image Segmentation. In *International Journal of Computer Vision*, volume 59, pages 167–181. Springer, September 2004.

[6] David A. Forsyth and Jean Ponce. *Computer Vision - A Modern Approach*. Artifical Intelligence. Prentice Hall, international edition, 2003.

[7] Rich Franzen. Kodak Lossless True Color Image Suite, 2004. 24 Kodak PhotoCD PCD0992 image samples in PNG file format. Retrieved September, 2009. Available from http://r0k.us/graphics/kodak/.

[8] Stuart Gibson, J. Andrew Bangham, and Richard Harvey. Evaluating a colour morphological scale-space. In *Proceedings of British Machine Vision Conference (BMVC '03)*, Norwich, UK, July 2003.

[9] David Gimenez and Adrian N. Evans. Colour morphological scale-spaces for image segmentation. In *Proceedings of British Machine Vision Conference (BMVC '05)*, volume 2, Oxford, UK, 2005.

[10] 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.

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

[12] R. Kimmel. Demosaicing: image reconstruction from color ccd samples. 8(9):1221–1228, Sept. 1999.

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

[14] Xin Li, Bahadir Gunturk, and Lei Zhang. Image demosaicing: a systematic survey. volume 6822, page 68221J. SPIE, 2008.

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

[16] Pierre Soille. *Morphological Image Analysis*. Springer, 2nd. edition, 2003.

[17] 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.

[18] S. Di Zenzo. A note on the gradient of a multi-image. In *Computer Vision, Graphics, and Image Processing*, volume 33, pages 116–125, 1986.

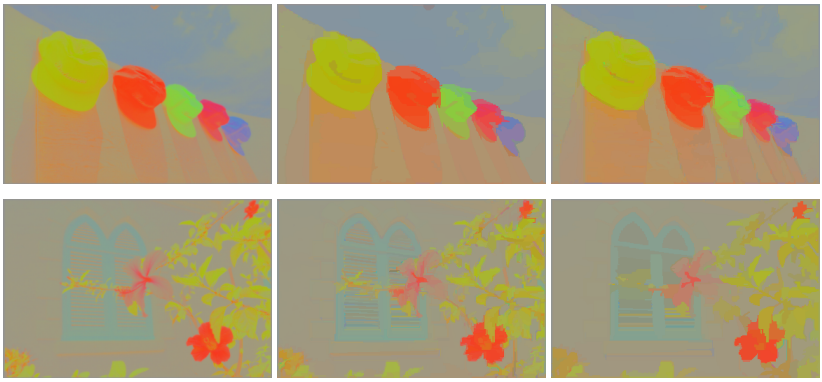

Fig. 3: Comparison of the chromacities of exemplarily original ground truth scenes on the left and segmentation scales using the same parameters as in fig 4.

Fig. 4: Results of the simultaneous segmentation and demosaicing algorithm on the Bayer image domain compared to a ground truth 3-CCD color image for exemplarily scenes of varying complexity. The left column contains the ground truth color image. The second column shows a synthesized color image of mean intensities at finer segmentation scale and the right most column contains results of coarser color clusters.

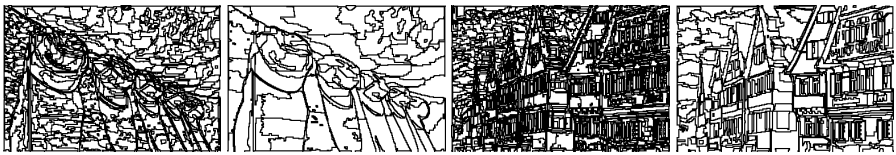

Fig. 5: Bounds of segmented regions at finer and coarser scale for two scenes from fig. 4.