Open Access

Adaptive dualISO HDR reconstruction

Saghi Hajisharif[*], Joel Kronander and Jonas Unger

Abstract

With the development of modern image sensors enabling flexible image acquisition, single shot high dynamic range (HDR) imaging is becoming increasingly popular. In this work, we capture single shot HDR images using an imaging sensor with spatially varying gain/ISO. This allows all incoming photons to be used in the imaging. Previous methods on single shot HDR capture use spatially varying neutral density (ND) filters which lead to wasting incoming light. The main technical contribution in this work is an extension of previous HDR reconstruction approaches for single shot HDR imaging based on local polynomial approximations (Kronander et al., Unified HDR reconstruction from raw CFA data, 2013; Hajisharif et al., HDR reconstruction for alternating gain (ISO) sensor readout, 2014). Using a sensor noise model, these works deploy a statistically informed filtering operation to reconstruct HDR pixel values. However, instead of using a fixed filter size, we introduce two novel algorithms for adaptive filter kernel selection. Unlike a previous work, using adaptive filter kernels (Signal Process Image Commun 29(2):203–215, 2014), our algorithms are based on analyzing the model fit and the expected statistical deviation of the estimate based on the sensor noise model. Using an iterative procedure, we can then adapt the filter kernel according to the image structure and the statistical image noise. Experimental results show that the proposed filter de-noises the noisy image carefully while well preserving the important image features such as edges and corners, outperforming previous methods. To demonstrate the robustness of our approach, we have exploited input images from raw sensor data using a commercial off-the-shelf camera. To further analyze our algorithm, we have also implemented a camera simulator to evaluate different gain patterns and noise properties of the sensor.

Keywords: HDR reconstruction, Single shot HDR imaging, DualISO, Statistical image filtering
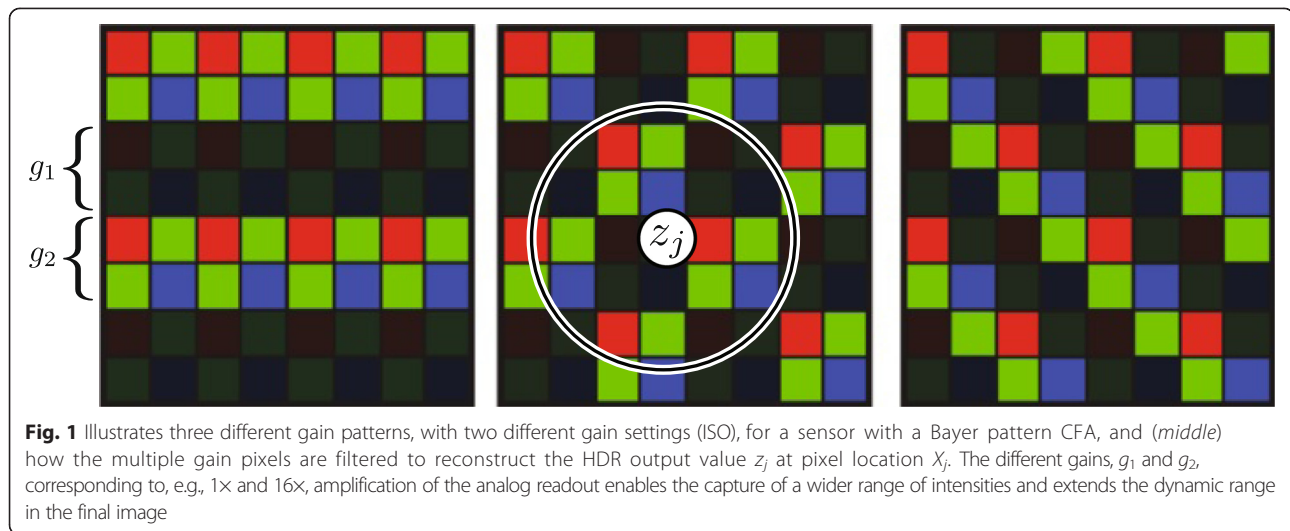
1 Introduction

The range of radiance intensities found in most real-world scenes, spanning from the sun or direct light sources to areas in shadow, typically exceeds, by orders of magnitude. It is very difficult to accurately capture this wide range using a digital sensor in a single image or video frame. This limitation has spurred the development of techniques for capture of *high dynamic range* (HDR) images and video; for an overview, see [26].

We present two algorithms for HDR image reconstruction based on a single input image where the pixel gain is varied over the sensor [4, 10]. Similar to [34, 35], we use the per-pixel gain of the analog signal, pixel measurements, to increase the dynamic range in the captured image. The analog pixel gain is proportional to the *ISO* setting found on most cameras. The input to our algorithm is a RAW sensor image consisting of pixels with either a high or a low gain setting, for example, varying the gain by every other two rows. The low gain setting enables the capturing of high-intensity region without saturation, while the high-gain setting enables us to capture image with a high signal-to-noise ratio in darker areas of the scene. Without loss of generality, we assume that color is captured using a *color filter array* (CFA), e.g., a Bayer pattern overlaid on the image sensor. Figure 1 illustrates two different distributions of per-pixel gain settings overlaid onto a raw CFA image. This approach to HDR capture is very robust and can be applied to off-the-shelf consumer cameras [4]. It does not suffer from, e.g., the various motion blurs or ghosting artifacts found in the commonly used exposure bracketing methods [7, 12]. Compared to multi-sensor cameras, e.g., [16, 31], it

* Correspondence: saghi.hajisharif@liu.se
Linköping University, Norrköping, Sweden

Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 2 of 13

Fig. 1 Illustrates three different gain patterns, with two different gain settings (ISO), for a sensor with a Bayer pattern CFA, and (*middle*) how the multiple gain pixels are filtered to reconstruct the HDR output value z_j at pixel location X_j. The different gains, g_1 and g_2, corresponding to, e.g., 1× and 16×, amplification of the analog readout enables the capture of a wider range of intensities and extends the dynamic range in the final image

does not require costly specialized hardware and removes the requirement of careful geometric sensor calibration and the risk of misalignment between the exposures.

The main contribution in this paper is an extension of the previous statistical reconstruction method for dualISO data developed in [10, 15], using two novel algorithms for adapting the scale of the filtering window. In contrast to previous works [10, 15], the window support is adapted both to the statistical properties of the image noise as well as the underlying signal structure contained in the image. We show that the novel scale selection results in increased image quality in several examples.

2 Background

Since the seminal work by Devebec and Malik [7], a large body of work has developed more robust and higher quality HDR capture and reconstruction methods; for a complete overview, see, e.g., [23, 26]. In this section, we give an overview of the previous work most closely related to the methods proposed in this paper.

2.1 HDR capture

High-quality HDR capture using off-the-shelf image sensors can currently be performed with three distinct approaches.

The traditional approach captures a sequence of images with varying exposure times and then merges these into an HDR image [7, 12]. For dynamic scenes, non-rigid registration of the individual exposures is necessary; and for moving objects, general de-ghosting algorithms are necessary to apply for high-quality results. While there has been a large body of work improving these approaches, see, e.g.,

the survey [33], they still cannot robustly handle moving cameras and objects in general scenes.

The second approach to HDR capture is based on using beam splitters to project incident light onto multiple sensors with different exposures. The different exposures can be achieved by using varying neutral density (ND) filters in front of the sensors [1, 8, 16, 19] or by clever setups of semi-transparent beam splitter arrangements [31]. These systems offer a major advantage over exposure time fusion methods in that they robustly handle motion of the camera and objects in the scene by using the same exposure time for each sensor.

The third approach, which is most closely related to this work, is spatial multiplexing of the image to achieve HDR capture. Here, a single sensor image is used where the response to incident light varies over the sensor. Most previous works achieved this by placing a spatially varying array of ND filters in front of the sensor [2, 24, 25, 27]. Its most familiar application is color imaging via a color filter array (e.g., the Bayer pattern [6]). By avoiding the need for more than one sensor, this design provides a cost-effective solution to achieve robust HDR capture. However, most existing methods still suffer from noise as large portions of the incident light are wasted in the ND filters. By instead focusing on spatially multiplexing the response to incident light using the gain/ISO setting, we can use the entire incident light for high-quality HDR reconstruction.

2.2 HDR reconstruction

To reconstruct HDR images from a set of images with different exposures, the traditional method is to compute a per-pixel weighted average of the low dynamic range (LDR) measurements. The weights, often based on heuristics, are chosen to suppress

Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 3 of 13

image noise and remove saturated values from processing [3, 7, 20]. Mann and Picard [20] assigned weights according to the derivative of the inverse camera response, and later Debevec and Malik [7] used a simple double ramp function that excludes values close to the saturation point or the black level. Later work derived weight functions based on more sophisticated camera noise models. Mitsunaga and Nayar [22] derived a weight function that maximizes SNR assuming signal-independent additive noise, and Kirk and Andersen [13] derived a weight function inversely proportional to the temporal variance of the digital LDR values. Granados et al. [9] extended this approach to include both spatial and temporal camera noises. While most previous methods consider only a single pixel at a time from each LDR exposure, Tocci et al. [31] presented an algorithm that incorporates a neighborhood of LDR samples in the reconstruction.

The vast majority of previous HDR reconstruction algorithms treat the complete imaging pipeline from raw pixel measurements to a full HDR image in a series of steps [7, 9, 31], either performing demosaicing after or before HDR fusion and denoising. In this work, we instead treat all of these operations in a single joint filtering operation. This enables us to take sensor noise into account in a systematic fashion while also improving the reconstruction speed. Recently, Heide et al. [11] proposed a framework for joint demosaicing, denoising, and HDR assembly by solving an inverse problem with different global image priors and regularizers using convex optimization methods. While providing impressive results, their method does not incorporate a well-founded model of the heterogeneous sensor noise, and despite GPU implementations, their implementation is still computationally expensive which requires solving a global optimization problem. Instead, we take a local approach, enabling rapid parallel processing, while also incorporating a well-founded statistical noise model.

Our statistically motivated locally adaptive filtering framework is inspired by recent methods in image processing. The last two decades have seen an increased popularity of image processing operations using locally adaptive filter weights, for applications in, e.g., interpolation, denoising, and upsampling. Examples include normalized convolution [14], the bilateral filter [32], and moving least squares [17]. Recently, deep connections have been shown [21, 29] between these methods and traditional non-parametric statistics [18]. In this paper, we extend the earlier framework for HDR reconstruction developed in [10, 15, 16] based on fitting local polynomial approximations (LPA) [5] to irregularly distributed samples around output pixels

using a localized maximum likelihood estimation [30] to incorporate the heterogeneous noise of the samples. In contrast to the previous works [10, 15, 16], we propose a novel adaptation of the filter kernel size that allows the filter extent to adapt not only to local image structure but also the sensor noise in the region.

3 DualISO capture and reconstruction—overview

The goal of the algorithm presented in this paper is to generate an HDR image based on input data in which the per-pixel gain (ISO) is varying over the sensor. This means that the analog readouts are amplified differently between segments of pixels on the sensor. Figure 1 illustrates three different gain patterns with two different gain values, g_1 and g_2, using a sensor with a Bayer pattern *color filter array* (CFA). The unity gain, g_1, pixel segments capture the high-intensity regions in the scene while the amplified segments, g_2, capture low-intensity regions. g_2 pixels may lie well below the acceptable noise floor for g_1 pixels.

The key benefit of using a varying per-pixel gain, g_i, is that the dynamic range in the final output will be extended using a single image as an input [10, 11]. However, accurate reconstruction of the output HDR image is a challenging filtering problem. The different gain settings lead to a loss of data in the spatial domain due to the fact that the amplified pixels, using gain g_2, saturate faster. For high-quality reconstruction, it is also necessary to take into account the heterogeneous image noise, which for a specific camera and exposure setting, varies with both intensity and the choice of gain settings.

The method presented here extends the statistical HDR reconstruction developed by [15, 16] to include reconstruction kernels which adapts to both the image content and the heterogeneous measurement noise. We assume that the input data is a raw CFA sensor image with per-pixel gain settings varying between pixel segments as described in Fig. 1 (middle). Each pixel value, z_j, at a pixel coordinate, X_j, in the output HDR image is, for each color channel, reconstructed by filtering the input pixels within a neighborhood around X_j. Our statistical approach first estimates the variance, or measurement noise, for each input sample in the raw image using a noise model. The input samples are then weighted using the estimated variances and an adaptive Gaussian kernel in the spatial domain. The weights, computed from the variances, ensure that low noise samples are weighted higher than noisy samples, and the Gaussian filter gives lower weights to samples further away from the reconstruction point, X_j. The HDR pixel value z_j at location X_j is then reconstructed iteratively by adjusting the shape of the Gaussian kernel to the

weighted input samples. In the first iteration, the Gaussian kernel is very small. The spatial support of the kernel is gradually increased until a statistically informed threshold based on the variances and image content is reached. The final HDR pixel value, z_j, is then estimated by fitting a polynomial to the weighted input samples. Our method performs noise reduction, color interpolation, and HDR fusion in a single operation.

The detailed presentation of the algorithm is laid out as described below. Section 4 first describes the camera noise model used to estimate sample variances, and Section 5 describes how each HDR pixel value is reconstructed using our statistical HDR reconstruction framework. The novel methods for filter scale selection for HDR reconstruction are presented in Section 6. Finally, in Section 7, we describe how the parameters for the noise model are calibrated and in Section 8, we show example results and evaluation of our reconstruction method.

4 Sensor noise model

The camera sensor electronics convert the incident radiant power f, which for convenience we express as the number of photo-induced electrons collected per unit time, to a measured digital value y_i at a pixel i. The samples, y_i, contain measurement noise

that is dependent on sensor characteristics such as readout noise, gain/ISO setting, and the inherent Poisson shot noise in the incident illumination.

To model the dependence of the measured digital pixel value on the incident radiant power and the camera parameters, we use a well-established radiometric model derived from previous works [9, 15]. Using this model, the non-saturated pixel values are modeled as random variables following a normal distribution:

$$y_i \sim N\left(g_i a_i t f_i + \mu_R, g_i^2 a_i t f_i + \sigma_R^2(g_i)\right), \qquad (1)$$

where t is the exposure time, g_i is the pixel gain/ISO, a_i is a pixel non-uniformity, μ_R is the mean of the read out noise, and σ_R^2 is the variance of the read out noise. An example showing the standard deviation of the read out noise, σ_R, for varying gain/ISO using a Canon Mark III sensor (saturation around 1600) is shown in Fig. 2.

In order to compute an estimate of the incident radiant power, $\hat{f}_i$, from the noisy digital input sample values y_i, we use the following estimator:

$$\hat{f}_i = \frac{y_i - b_i}{g_i t a_i}, \qquad (2)$$

where b_i is obtained from a bias frame captured with no light reaching the sensor.
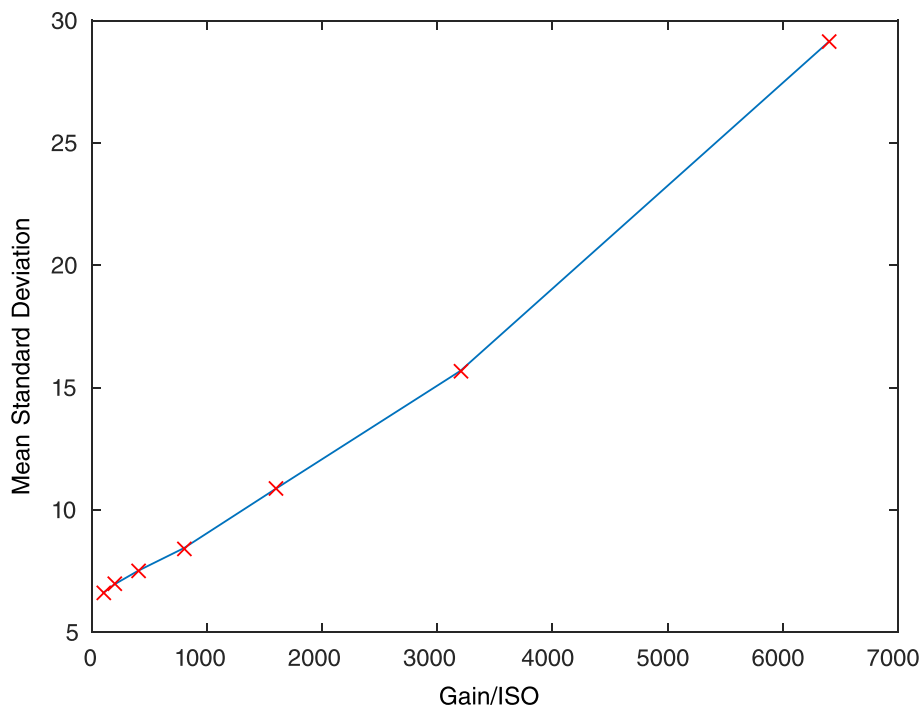

Fig. 2 Mean standard deviation versus gain/ISO of the 14-bit Canon Mark III sensor. The ISO settings are 100, 200, 400, 800, 1600, 3200, and 6400

Similarly, we approximate the variance of this estimator by

$$\hat{\sigma}^2_{\hat{f}_i} = \frac{g_i^2 t a_i \hat{f}_i + \hat{\sigma}^2_R(g_i)}{g_i^2 t^2 a_i^2},\tag{3}$$

where g_i, a_i, and $\hat{\sigma}^2_R(g_i)$ are found through calibration; see Section 7. We do not include the effect of pixel cross-talk, and the variances, $\hat{\sigma}^2_{\hat{f}_i}$, are assumed to be independent of each other.

5 Adaptive HDR reconstruction

To estimate an HDR pixel value z_j at a location X_j on the sensor, we use a LPA [5] to fit the observation samples of the incident radiant power in the local neighborhood. The same framework is also known as kernel regression [29].

5.1 Local polynomial approximation

To estimate the radiant power, $f(x)$, at an output pixel, we use a generic local polynomial expansion of the radiant power around the output pixel location $X_j = [x_1, x_2]^T$. Assuming that the radiant power $f(x)$ is a smooth function in a local neighborhood around the output location X_j, an Mth order Taylor series expansion is used to predict the radiant power at a point X_i close to X_j as follows:

$$\tilde{f}(X_i) = C_0 + C_1(X_i-X_j)$$
$$+ C_2 tril\left\{(X_i-X_j)(X_i-X_j)^T\right\} + ...,\tag{4}$$

where *tril* lexicographically vectorizes the lower triangular part of a symmetric matrix and

$$C_0 = f(X_j)\tag{5}$$

$$C_1 = \nabla f(X_j) = \left[\frac{\partial f(X_j)}{\partial x_1}, \frac{\partial f(X_j)}{\partial x_2}\right]\tag{6}$$

$$C_2 = \frac{1}{2}\left[\frac{\partial^2 f(X_j)}{\partial x_1^2}, 2\frac{\partial^2 f(X_j)}{\partial x_1 \partial x_2}, \frac{\partial^2 f(X_j)}{\partial x_2^2}\right].\tag{7}$$

Given the fitted polynomial coefficients, $C_{1:M}$, we can thus estimate the radiant power and the HDR pixel value, z_j, at the output location X_j by $z_j = C_0 = f(X_j)$.

5.2 Maximum localized likelihood fitting

To estimate the coefficients, we maximize a localized likelihood function [30] defined using a Gaussian smoothing window centered around X_j

$$\mathcal{W}_h(X_j) = \frac{1}{2\pi h^2}\exp\left\{\frac{-(X_k-X_j)^T(X_k-X_j)}{h}\right\},\tag{8}$$

where h is a local scale parameter (see Section 6) which determines the shape of the filtering kernel. In Section 6, we discuss how the size of the window function can be selected adaptively depending on the features at each location in the image.

We denote the observed pixel samples (radiant power estimates, $\hat{f}_i(X_j)$ at position X_j) in the support of the local neighborhood window by f_k with a linear index $k = 1 \ldots K$. Note that these are obtained from the digital pixel values using Eq. 2 derived from the sensor noise model.

Using the assumption of normally distributed radiant power estimates, f_k, the polynomial coefficients, $\tilde{C}$, maximizing the localized likelihood function is found by the weighted least squares estimate

$$\tilde{C} = (\Phi^T W \Phi)^{-1}\Phi^T W \bar{f},\tag{9}$$

where

$$\bar{f} = [f_1, f_2, \ldots f_K]^T$$

$$W = \text{diag}\left[\frac{\mathcal{W}_h(X_1)}{\hat{\sigma}^2_{f_1}}, \frac{\mathcal{W}_h(X_2)}{\hat{\sigma}^2_{f_2}}, \ldots, \frac{\mathcal{W}_h(X_K)}{\hat{\sigma}^2_{f_k}}\right]$$

$$\Phi = \begin{bmatrix} 1 & (X_1-X_j) & tril^T\left\{(X_1-X_j)(X_1-X_j)^T\right\} & \ldots \\ 1 & (X_2-X_j) & tril^T\left\{(X_2-X_j)(X_2-X_j)^T\right\} & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_K-X_j) & tril^T\left\{(X_K-X_j)(X_K-X_j)^T\right\} & \ldots \end{bmatrix}.$$

$$\tag{10}$$

The operator *tril* lexicographically vectorizes the lower triangular part of a symmetric matrix.

Using this maximum likelihood approach, we can efficiently solve for the polynomial coefficients $C_{1:M}$ and estimate the final HDR pixel value z_j at a pixel location X_j for a given smoothing parameter h. However, in order to enable a good trade-off between bias and variance, i.e., between image sharpness and noise reduction, it is necessary to locally adapt the smoothing parameter h to image features and image noise. If h is globally fixed over the image, reconstruction may lead to a noisy final image for small h and blurry result for a high h value. The best trade-off between image sharpness and denoising is achieved by adapting the smoothing parameter h to local image features.

In the next section, we describe the iterative reconstruction method and two algorithms for selecting the locally best smoothing parameter, h, for each HDR pixel estimate, z_j, individually.

Algorithm 1 Adaptive HDR reconstruction

```
 1: procedure HDR RECONSTRUCTION
 2:     for each color channel in R, G, B do
 3:         for each HDR pixel estimate z_j do
 4:             h_l = h_min
 5:             z_j = estimate ẑ_{j,h_l} using LPA with degree M
 6:             for each h_l < h_max  do
 7:                 h_l = h_l + h_inc
 8:                 estimate ẑ_{j,h_l} using LPA with degree M
 9:                 estimate ẑ_{j,h_l} variance σ̃_{ẑ_{j,h_i}} and the reconstruction error ε (for EVS)
10:                 apply update rule (ICI or EVS) based on ẑ_{j,h_l} and
11:                 if  variation in ẑ_{j,h_l} can be explained by σ̃_{ẑ_{j,h_l}} then
12:                     z_j = ẑ_{j,h_l}
13:                 else
14:                     break
15:                 end if
16:             end for
17:         end for
18:     end for
19: end procedure
```

6 Adaptive scale selection

The size of the window function introduces a trade-off between bias and variance. A large window will reduce the variance but can lead to overly smoothed images (bias). Ideally, it is desirable to have large window supports in regions where the smooth polynomial model, used for the reconstruction, is a good fit to the underlying signal, while keeping the window size small close to the edges or important image features. The size of the smoothing window is determined by the smoothing parameter h. Figure 3 illustrates how a signal value, the black point, is being estimated using a kernel with a gradually increasing smoothing parameter, h. When the smoothing parameter h is increased from h_0, the h_1, i.e., a higher degree of smoothing, the variance in the estimated value can be explained by the signal variance. When the smoothing parameter is increased from h_1 to h_2, the kernel reaches the step in the signal and the estimation at the black point can no longer be explained by the signal variance. Smoothing parameter h_1 thus produces a better estimate.

The adaptation of the smoothing parameter, h, scale selection is carried out iteratively. The goal of the adaptation is to gradually increase h, and find an optimal
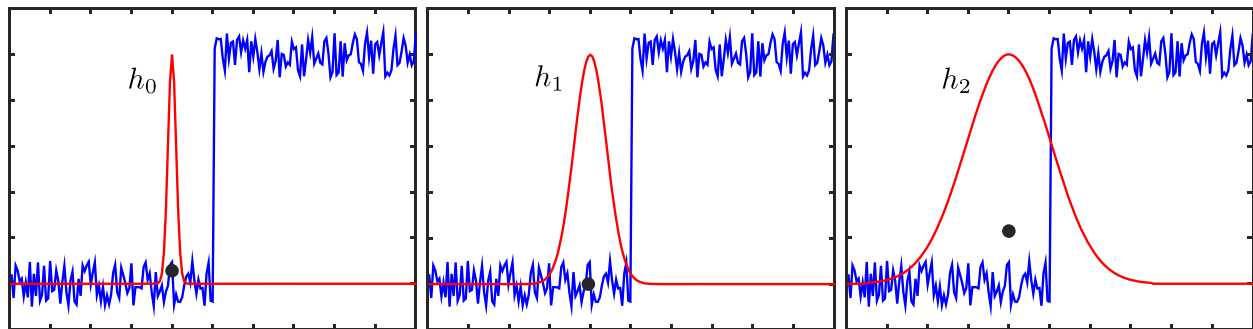

Fig. 3 Illustrating how a signal value, the *black point*, is estimated using a kernel with an iteratively increasing smoothing parameter, h. Increasing from h_0 to h_1, i.e., a higher degree of smoothing, the variance in the estimated value can be explained by the variance in the original signal. However, when the smoothing parameter is increased from h_1 to h_2, the kernel reaches the step in the signal and the estimate at the black point can no longer be explained by the signal variance

Fig. 4 Reconstructed from duaISO data with ISO100–1600 captured with Canon 5D Mark III. Reconstructed with ICI $M = 2$, $h \in [0.6, 5.0]$, and $\Gamma = 1.0$

h such that the variation in the estimated value between iterations can be explained by the signal variance and the smoothing applied. Denoting each iteration by l and the corresponding smoothing parameter by h_1, Algorithm 1 describes the outline of the HDR pixels z_j reconstructed by adapting the smoothing parameter h_1. In each iteration, we estimate the signal value and its variance. We then apply an update rule which determines whether the h value used is valid or not. This is repeated until the update rule does not hold or the maximum h value, $h_{\max}$, is reached. In Sections 6.1 and 6.2, we describe how the variance of the pixel is estimated in detail with the two different update rules.

6.1 Update rule 1: error of estimation versus standard deviation (EVS)

The first update rule is built on the intuition that if the weighted mean reconstruction error is larger than the weighted mean standard deviation, i.e., the difference between the data and the fit cannot be explained by the expected signal variation due to noise, the polynomial model does not provide a good fit to the underlying image data. As described in Algorithm 1, the smoothing

parameter, h_1, is iteratively increased with an increment h_{inc}. In each iteration, l, the EVS update rule computes the weighted reconstruction error e_1 as

$$e_l = \sqrt{\sum_k W^2(k,k)\left(\tilde{f}(X_k) - \hat{f}_k\right)^2}, \qquad (11)$$

where k indexes the pixels in the neighborhood and W is the weights including both the variance of the original pixels and the spatial Gaussian kernel as described in Eq. 10. The weighted standard deviation, $\tilde{\sigma}_{\hat{z}_{j,h_i}}$, of this estimate can be obtained from the covariance matrix M_C for the fitted polynomial coefficients, $\tilde{C}$, which is given by

$$M_C = \left(\Phi^T W \Phi\right)^{-1} \Phi^T W \Sigma W^T \Phi \left(\Phi^T W^T \Phi\right)^{-1}, \qquad (12)$$

where $\Sigma = diag[\sigma_{f_1}^2, \sigma_{f_2}^2, ..., \sigma_{f_k}^2]$ is the variance of the observation. The variance of estimated radiant power z_j, $\tilde{\sigma}_{\hat{f}_j}$, at the output location X_j, is thus given by the element $\tilde{\sigma}_{\hat{z}_{j,h_i}} = M_C(0,0)$ in M_C. During the iterations, the smoothing parameter, h_l, is updated to $h_{l+1} = h_l + h_{\mathrm{inc}}$ as long as the weighted reconstruction error, $_l$, is smaller than the standard deviation $\epsilon_l < \Gamma \tilde{\sigma}_{\hat{z}_{j,h_i}}$, where Γ is a user-

Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 8 of 13

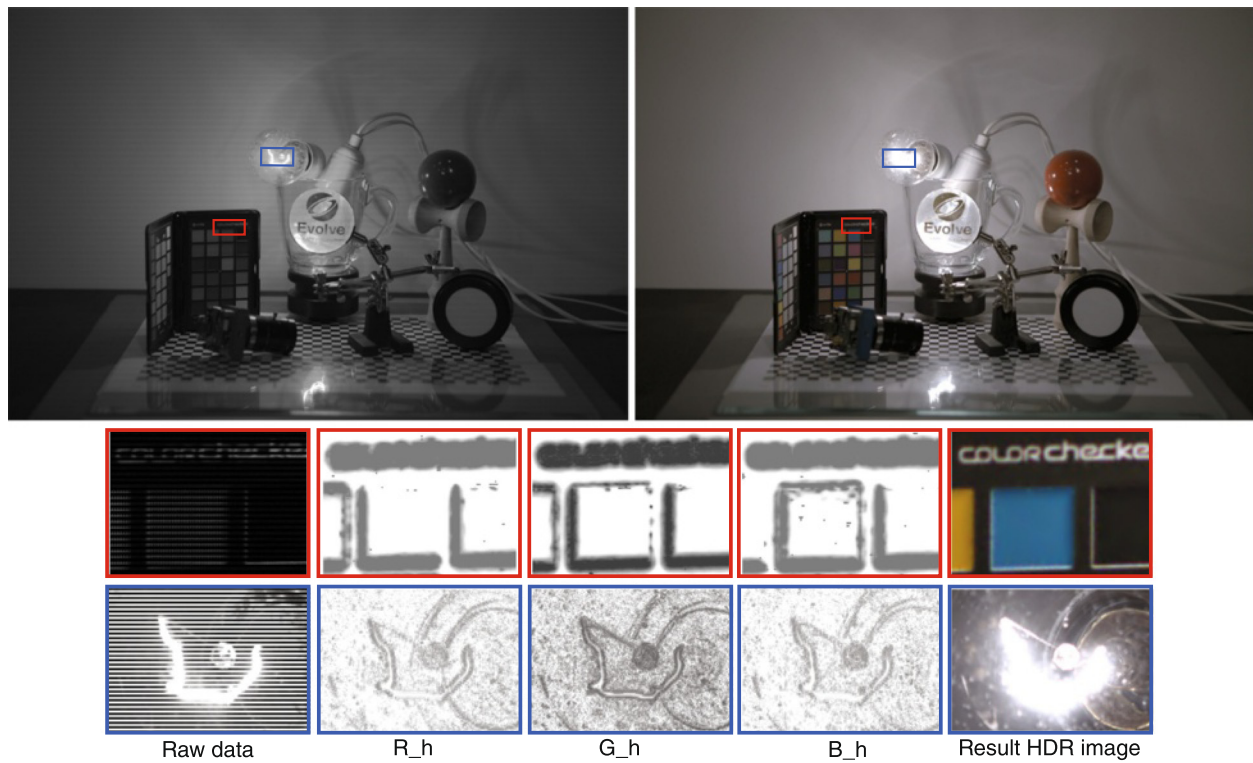| Raw data | R_h | G_h | B_h | Result HDR image |

Fig. 5 Reconstruction process of one sample raw image. *Top left* shows the raw input image with CFA Bayer pattern and dualISO row pattern. *Top right* indicates the resulted tone mapped HDR reconstructed image with EVS rule. *Bottom rows* extracted images from *left* to *right*: cutout of the raw image, scaling parameter image for R, G, B color channels with Γ = 1.0, and the cutout of the reconstructed HDR image

Fig. 6 Lamp scene with different methods for comparison: **a** LPA *M* = 2 from *left* to *right*: *h* = 0.6, 1.4, and 5.0; **b** SKR *M* = 2 from *left* to *right h* = 0.6, 1.4, and 5.0; **c** our method with ICI *M* = 2 from *left* to *right* Γ = 0.6, 1.0, and 1.4; **d** our method with EVS *M* = 2 from *left* to *right* Γ = 0.6, 1.0, and 1.4

Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 9 of 13

specified parameter controlling the trade-off between levels of denoising applied by the kernel.

6.2 Update rule 2: intersection of confidence intervals (ICI)

The second update rule is based on the ICI algorithm [5]. The main purpose of this algorithm is to obtain the largest scaling parameter in the local neighborhood of the estimation point under the constraint that the polynomial model remains a likely fit to the underlying data. As described in Algorithm 1, the smoothing parameter, $h_{\min} \leq h_l \leq h_{\max}$, is iteratively increased. For each iteration, l, the ICI rule determines a confidence interval, $D_l = [L_l, U_l]$:

$$L_l = \hat{z}_{j,h_l}(x) - \Gamma \tilde{\sigma}_{\hat{z}_{j,h_l}}, \tag{13}$$

$$U_l = \hat{z}_{j,h_l}(x) + \Gamma \tilde{\sigma}_{\hat{z}_{j,h_l}}, \tag{14}$$

where $\hat{z}_{j,h_l}(x)$ is the estimated radiant power given the scaling parameter h_l and $\tilde{\sigma}_{\hat{z}_{j,h_l}}$ is the weighted standard deviation of this estimate computed using Eq. 12. Γ is a scaling parameter controlling how wide the intersection interval is. During adaptation, h_l is increased as long as there is an overlap between the confidence intervals, i.e., h_l is updated to $h_{l+1} = h_l + h_{\text{inc}}$ if there is an overlap between D_l and D_{l+1}. In practice, we utilize Γ as a user parameter, enabling an intuitive trade-off between image sharpness and denoising. A detailed overview of the ICI rule and its robustness can be found in [28].

7 Camera parameter calibration

The variance of the readout noise, the sensor gain, bias, and the sensor saturation point are calibrated once for each sensor. The bias frame, b, and readout noise variance, $\text{Var}[r_i(g_i, t)]$, are calibrated as the per-pixel mean and the variance, respectively. This calibration is done over a set of black images captured with the lens covered, so that no photons reach the sensor. The sensor gain, g_i, can be calibrated using the relation,

$$\frac{\text{Var}[y_i] - \text{Var}[b_i]}{E[y_i] - E[b_i]} = \frac{g_i^2 \text{Var}[e_i]}{g_i E[e_i]} = g_i, \tag{15}$$

where the second equality follows from e_i being Poisson distributed shot noise with $E[e_i] = \text{Var}[e_i]$. In addition, $E[y_i]$ and $E[b_i]$ can be estimated by averaging flat fields and the bias frame, respectively, and $\text{Var}[b_i]$ as described above. The per-pixel non-uniformity, a_i, can be estimated using a flat field image computed as the average over a large sequence of non-saturated images.
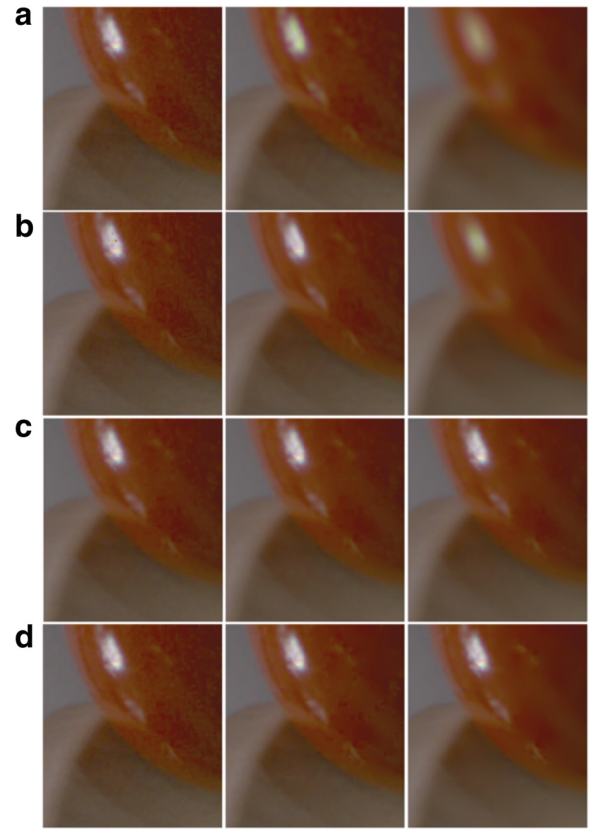

Fig. 7 Another comparison of a cutout of the simulated lamp scene for different methods: **a** LPA $M = 2$ from *left* to *right*: $h = 0.6, 1.4$, and 5.0; **b** SKR $M = 2$ from *left* to *right* $h = 0.6, 1.4$, and 5.0; **c** our method with ICI $M = 2$ from *left* to *right* $\Gamma = 0.6, 1.0$, and 1.4; **d** our method with EVS $M = 2$ from *left* to *right* $\Gamma = 0.6, 1.0$, and 1.4

8 Results and evaluation

The proposed algorithm has been evaluated on two different sets of images. One synthetic image data set with known ground truth computed using a camera simulator and one set of images captured using a Canon 5D Mark III running the Magic Lantern firmware with the dualISO module installed. The synthetic data is generated using a camera simulation framework which takes a noise-free HDR image as input and applies noise based on the camera
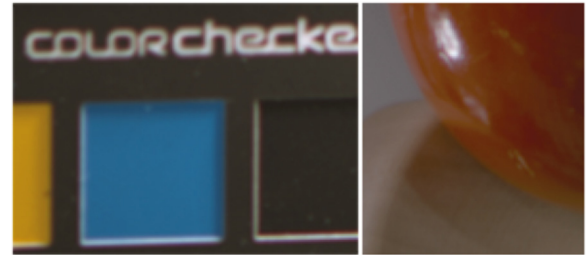

Fig. 8 The ground truth reference images for the cutouts compared in Figs. 6 and 7

noise model described in Section 4. The camera parameters estimated from real cameras as described in Section 7 are used for simulating dualISO sample data. The noise-free HDR images (ground truth) were captured as a set of carefully calibrated exposure brackets one *f*-stop apart covering the dynamic range of the scene. Each of the different exposures in the bracketing sequence was captured as the average of 100 calibrated RAW images with the same exposure settings. The test images used exhibit a very large dynamic range, were selected to be representative for challenging scenes, and include features such as dark and bright image regions, high- and low-frequency regions, image noise, and strong local contrasts. In our evaluation, we compare three different gain patterns as shown in Fig. 1. We have tested our algorithm for a polynomial degree of $M = 0, 1, 2$ and a range of different parameter settings for Γ. In all tests (except for the non-adaptive fixed *h* comparisons), the smoothing parameter, *h*, is allowed to vary between $h = 0.6$ and $h = 5.0$. Figure 4 shows an image captured with a Canon 5D Mark III running the Magic Lantern dualISO module and reconstructed by the proposed method. The image shows that our algorithm performs well in the reconstruction by keeping image sharpness while allowing high-quality noise reduction.

Figure 5 shows a high contrast scene simulating a Canon 5D camera with dualISO settings of ISO100 and ISO1600 alternating in pairs of rows on the sensor as shown in Fig. 1 (left). Figure 5 shows the input raw CFA Bayer image, and three images in the bottom row show the locally adapted *h* values for the red, green, and blue color channels, respectively. The EVS update rule adapts the smoothing parameter *h* to both the image features and the image noise. The parameter *h* becomes smaller as we get closer to edges and textured regions and larger in homogeneous areas.

In Figs. 6 and 7, we focus on the trade-off between image sharpness and denoising. We compare our algorithm using both the ICI and EVS update rules to LPA using non-adaptive filtering kernels, [10], with $h = 0.6$, 1.4, and 5.0, and the widely used steering kernel regression (SKR) method [29]. The images compare two cut-out regions of the lamp scene from Fig. 5. The two regions have been chosen to display the performance of our algorithm in a dark region, Fig. 6, and a highlight region, Fig. 7. The ground truth reference images of the cutouts are displayed in Fig. 8. In both images, the comparisons are ordered as follows: (a) non-adaptive LPA $M = 2$ from left to right with $h = 0.6, 1.4,$ and 5.0, (b) SKR [29] $M = 2$ from left to right with $h = 0.6, 1.4,$ and 5.0, (c)

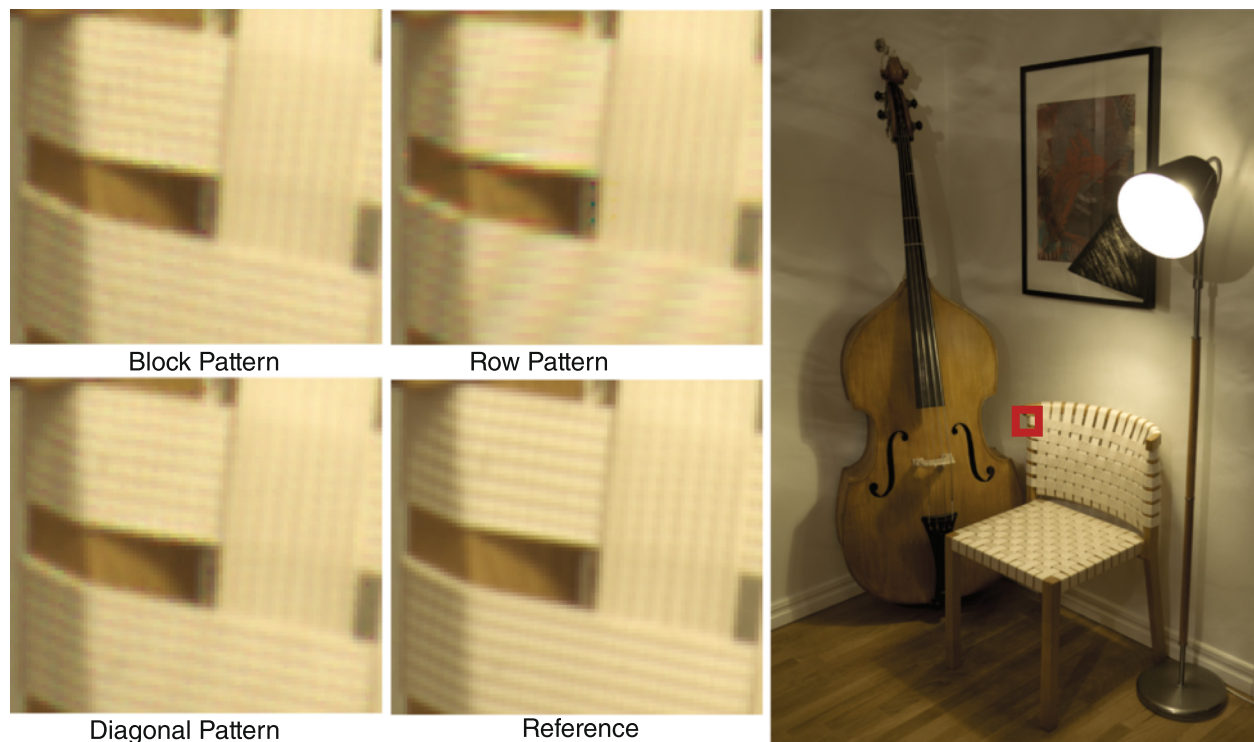**Fig. 9** Living room scene. Comparison of our method with EVS rule $M = 2$; Γ = 1.0 for different gain patterns: block pattern, row pattern, and diagonal pattern
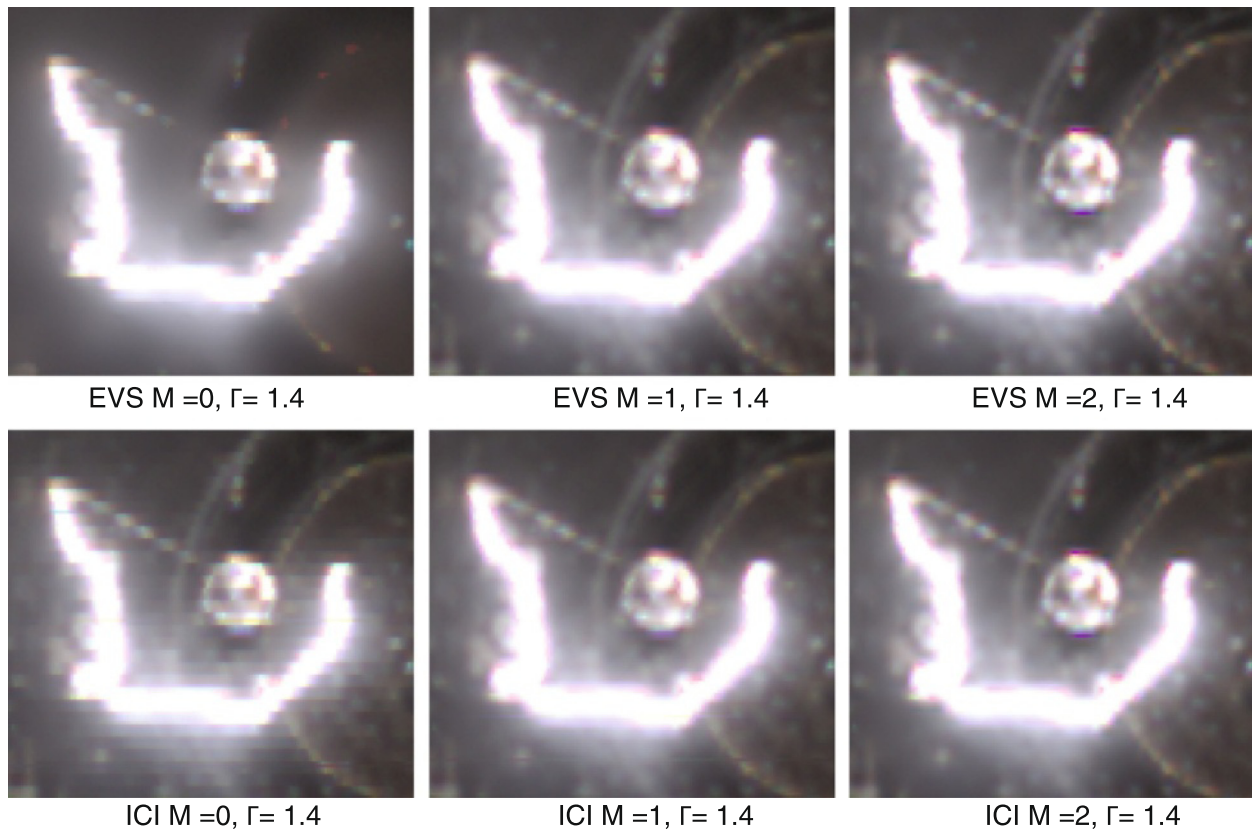
Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 11 of 13

EVS M =0, Γ= 1.4   EVS M =1, Γ= 1.4   EVS M =2, Γ= 1.4

ICI M =0, Γ= 1.4   ICI M =1, Γ= 1.4   ICI M =2, Γ= 1.4

Fig. 10 Lamp scene, evaluation of EVS, and ICI method for different degrees of polynomial for dualISO 100–1600 with row pattern from *left* to *right* with $M = 0, 1,$ and 2

our method with ICI rule for local adaptation of scale parameter, $M = 2$, from left to right: $\Gamma = 0.6, 1.0,$ and 1.4, and (d) our method with EVS rule, $M = 2$, from left to right: $\Gamma = 0.6, 1.0,$ and 1.4. From Fig. 6, it is evident that the non-adaptive method in (a) [10] does not perform well. SKR produces good results for $h = 1.4$ but cannot fully adapt the smoothing parameter as artifacts from the noise filtering are visible (zoom in). Both ICI- and EVS-based algorithms keep sharpness while reducing the image noise more than the other methods. In Fig. 7, SKR with $h = 1.4$ produces a sharp image without color artifacts; however, it also smooths the reflection on the

red toy. ICI and EVS produce a similar result, but EVS leads to less smoothing around the highlight areas of the scene compared to ICI. The images show that our algorithms using ICI and EVS update rules produce high-quality images. In general, the EVS update rule allows for a higher degree of smoothing and denoising while keeping higher contrast edges intact. However, in dark regions, the EVS update leads to a loss of detail compared to ICI rule. Another important difference is that although the EVS update rule may produce better results in some cases, it is built on the heuristic argument that the reconstruction error should be smaller than the

Fig. 11 Cutouts of the checkerboard in the lamp scene, evaluation of ISO settings for EVS $M = 2$, $\Gamma = 1.0$. (*Top row*) From *left* to *right*: reference, dualISO 100–400, dualISO 100–800, and dualISO 100–1600. (*Bottom row*) From *left* to *right*: dualISO 100–3200, dualISO 100–6400, dualISO 100–12800, and dualISO 100–25600
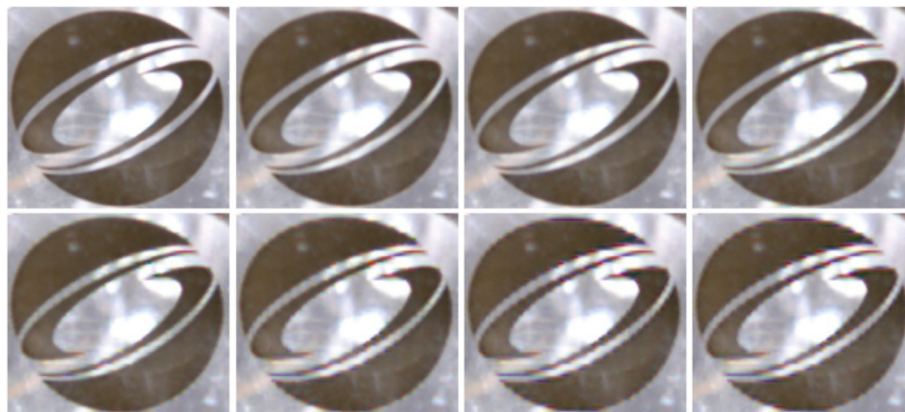
Fig. 12 Cutouts of the glass in the lamp scene, evaluation of ISO settings for EVS $M = 2$, $\Gamma = 1.0$. (*Top row*) From *left* to *right*: reference, dualISO 100–400, dualISO 100–800, and dualISO 100–1600. (*Bottom row*) From *left* to *right*: dualISO 100–3200, dualISO 100–6400, dualISO 100–12800, and dualISO 100–25600

standard deviation in the filtered region. While ICI rule is statistically motivated and designed to minimize the estimate variance.

In Fig. 9, we demonstrate how our algorithms perform using the three different gain patterns illustrated in Fig. 1. This particular image region is selected as it contains slanted edges in different directions. The comparisons show that the block pattern and diagonal pattern in some cases produce better results. However, the reconstruction quality depends on how the image features are oriented and the statistically optimal configuration of the gain pattern is out of scope of this paper. Figure 10 shows a cutout of the lamp scene simulated with row pattern and reconstructed using a varying polynomial degree of $M = 0$, 1, and 2. As expected, $M = 0$ produces a blocky result, and $M = 1$ and $M = 2$ produce increasingly more accurate reconstructions.

In Figs. 11 and 12, we show the effect of increasing the ISO separation in the dualISO image using a simulated 14-bit Canon 5D sensor. The dualISO settings are varied from ISO100–ISO200 to ISO100–ISO25600. As the separation between the ISO settings increase, the number of overlapping bits in the two exposures decrease. The image shows that our algorithm works well up to ISO100–ISO6400, i.e., a separation of six *f*-stops and an overlap of 8 bits. By increasing the separation further, artifacts start to appear along the edges.

9 Conclusions

In this paper, we presented a novel approach for adaptive unified HDR image reconstruction that includes the sensor noise model and error of the estimation for a more robust and accurate reconstruction of single shot spatial multiplexing raw data. The method handles severe noise, especially in the darker regions while it keeps the error of the estimation low to prevent over-smoothing of the image. To the

best of our knowledge, none of the previous methods have considered sensor noise model and estimated error and variance in order to adapt the reconstructed kernel for each local region of the image. The robustness of our approach for noise reduction and HDR reconstruction has been experimentally verified on both real data and simulated camera images. While being a simple method to implement, our results demonstrate a relatively good performance.

References
1. M Aggarwal, N Ahuja, Split aperture imaging for high dynamic range. Int. J. Comput. Vis. **58**(1), 7–17 (2004)
2. C Aguerrebere, A Almansa, Y Gousseau, J Delon, P Musé, Single shot high dynamic range imaging using piecewise linear estimators, in *ICCP*, 2014
3. AO Akyüz, E Reinhard, Noise reduction in high dynamic range imaging. J. Vis. Commun. Image Represent. **18**(5), 366–367 (2007)
4. Alex. Dynamic range improvement for Canon DSLR with 8-channel sensor read-out by alternating iso during sensor readout. Technical documentation, url: http://acoutts.com/a1ex/dual_iso.pdf, July 2013.
5. J Astola, V Katkovnik, K Egiazarian. *Local Approximation Techniques in Signal and Image Processing*. SPIE- International Society for Optical Engineering, 2006
6. B Bayer. Color imaging array. US Patent 3 971 065, 1976.
7. P Debevec, J Malik, Recovering high dynamic range radiance maps from photographs, in *SIGGRAPH*, 1997, pp. 369–378
8. J Froehlich, S Grandinetti, B Eberhardt, S Walter, A Schilling, H Brendel. Creating cinematic wide gamut HDR-video for evaluation of tone mapping operators and HDR-displays. In *SPIE Electronic Imaging*, pages 90230X-90230X. International Society for Optics and Photonics, SPIE digital library, 2014.
9. M Granados, B Ajdin, M Wand, C Theobalt, H Seidel, H Lensch, Optimal hdr reconstruction with linear digital cameras, in *CVPR*, 2010

Hajisharif *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:41

Page 13 of 13

10. S Hajisharif, J Kronander, J Unger, HDR reconstruction for alternating gain (iso) sensor readout, in *Eurographics 2014 Short Papers*, ed. by MW Eric Galin, 2014

11. F Heide, M Steinberger, YT Tsai, M Rouf, D Pajak, D Reddy, O Gallo, J Liu, W Heidrich, K Egiazarian, J Kautz, L Pulli. FlexISP: a flexible camera image processing framework. *ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2014)*, 33(6), December 2014

12. S Kang, M Uyttendaele, S Winder, R Szeliski, High dynamic range video. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2003) **22**(3), 319–325 (2003)

13. K Kirk, H Andersen, Noise characterization of weighting schemes for combination of multiple exposures, in *Proc. British Machine Vision Conference (BMVC)*, 2006, pp. 1129–1138

14. H Knutsson, CF Westin, Normalized and differential convolution, in *CVPR*, 1993

15. J Kronander, S Gustavson, G Bonnet, A Ynnerman, J Unger, A unified framework for multi-sensor HDR video reconstruction. Signal Processing: Image Communications **29**(2), 203–215 (2014)

16. J Kronander, S Gustavson, G Bonnet, J Unger, Unified HDR reconstruction from raw CFA data, in *IEEE International Conference on Computational Photography (ICCP)*, 2013

17. P Lancaster, K Salkauskasr, Surfaces generated by moving least squares methods. Math. Comput. **87**, 141–158 (1981)

18. C Loader, *Local regression and likelihood* (Springer, New York, 1999)

19. A Manakov, JF Restrepo, O Klehm, R Hegedüs, E Eisemann, HP Seidel, I Ihrke, A reconfigurable camera add-on for high dynamic range, multi-spectral, polarization, and light-field imaging. ACM Transaction (Proc. SIGGRAPH 2013) **32**(4), 1–47 (2013)

20. S Mann, RW Picard, On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures, in *IS&T*, 1995

21. P Milanfar, A tour of modern image filtering: new insights and methods, both practical and theoretical. IEEE Signal Process. Mag. **30**(1), 106–128 (2013)

22. T Mitsunaga, SK Nayar, Radiometric self calibration, in *CVPR*, 1999, pp. 374–380

23. K Myszkowski, R Mantiuk, G Krawczyk. *High Dynamic Range Video.* Synthesis lectures on computer graphics and animation, a publication in Morgan and Claypool, 2008

24. SG Narasimhan, SK Nayar, Enhancing resolution along multiple imaging dimensions using assorted pixels. IEEE Transaction on Pattern Analysis and Machine Intelligence **27**(4), 518–530 (2005)

25. S Nayar, T Mitsunaga, High dynamic range imaging: spatially varying pixel exposures, in *CVPR*, 2000

26. E Reinhard, W Heidrich, S Pattanaik, P Debevec, G Ward, K Myszkowski. High dynamic range imaging: acquisition, display and image-based lighting (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005)

27. M Schoberl, A Belz, J Seiler, S Foessel, A Kaup, High dynamic range video by spatially non-regular optical filtering, in *Image Processing (ICIP), 2012 19th IEEE International Conference*, 2012, pp. 2757–2760

28. L Stankovic, Performance analysis of the adaptive algorithm for bias-to-variance tradeoff. IEEE Transaction on Signal Processing **52**(5), 1228–1234 (2004)

29. H Takeda, S Farsiu, P Milanfar, Kernel regression for image processing and reconstruction. IEEE Trans. Image Process. **16**(2), 349–366 (2007)

30. R Tibshirani, T Hastie, Local likelihood estimation. J. Am. Stat. Assoc. **82**(398), 559–567 (1987)

31. MD Tocci, C Kiser, N Tocci, P Sen, A versatile HDR video production system. ACM Transactions on Graphics(Proceedings of SIGGRAPH 2011) **30**(4), 1–41 (2011)

32. C Tomasi, R Manduchi, Bilateral filtering for gray and color images, in *ICCV*, 1998, pp. 839–846

33. OT Tursun, AO Akyüz, A Erdem, E Erdem. The state-of-the-art in HDR deghosting: a survey and evaluation. *Computer Graphics Forum (Proceedings of Eurogprahics STARs)*, 2015

34. J Unger, S Gustavson, High-dynamic-range video for photometric measurement of illumination, in *SPIE Electronic Imaging*, 2007

35. J Unger, S Gustavson, M Ollila, M Johannesson, A real time light probe, in *Proceedings of the 25th Eurographics Annual Conference, volume Short Papers and Interactive Demos*, 2004, pp. 17–21