Dynamic range enhancement algorithms for CMOS sensors with non-destructive readout

DOI:
10.1109/IST.2008.4659956

Citation for published version (APA):

Published in:

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Abstract— Image sensors capable of a non-destructive readout (NDR) allow reading several frames during the integration time without affecting a photo charge being collected in the pixel's well. We show that two published algorithms for processing the sequence of NDR frames into a high dynamic range (HDR) image suffer from "stepped gradient" artifacts using data taken with an "off-the-shelf" CMOS image sensor. We present a novel algorithm that is capable of alleviating this image artifact that is due to suboptimal saturated pixel treatment. Our procedure works for both black and white and color images. In addition, we present joint demosaicing and denoising for color HDR image reconstruction and its performance is compared with conventional demosaicing algorithms. The overall performance of the combined methods provides a DR increase of up to 44 dB compared to raw data. We demonstrate our algorithms using images taken with a black-and-white and color NDR capable camera system.

Keywords—non-destructive readout; CMOS image sensor; dynamic range enhancement; saturated pixels extrapolation; demosaicing; noise removal; embedded camera system

I. INTRODUCTION

Image sensors capable of non-destructive readout (NDR) allow for reading several frames during a single exposure without affecting the integrated signal levels. This approach was originally developed for high dynamic range (HDR) infrared detectors [1]. Advances in complementary metal oxide semiconductor (CMOS) sensor design have enabled sensor designers to implement NDR schemes in high-resolution visible light image sensors [2-4]. Compared to other dynamic range (DR) enhancement techniques [5, 6] the processing of multiple NDR frames was shown [4-6] to have several advantages such as preservation of the signal to noise ratio (SNR) over the extended part of the dynamic range, a potential to enhance DR at both the lower and higher ends of the signal range and digital correlated double sampling (DCDS) [7].

Several works describing a sensor architecture with on-chip (hardware) implementation of NDR based DR enhancement have been presented [3, 6]. However, these approaches substantially increase pixel complexity and, therefore, reduce the photosensitive area (lower fill factor). In addition, there are more sources of noise and a higher design and manufacturing cost. This work presents an alternative approach of external NDR data processing using a state-of-the-art CMOS sensor with a reconfigurable logic read out system for both black-and-white and color images. We also address the problem associated with extrapolation of saturated pixels and color reconstruction of HDR images under noisy conditions.

II. DYNAMIC RANGE ENHANCEMENT

The NDR scheme provides several opportunities for improvement in the dynamic range of a captured image scene. Firstly, at low signal levels the DR is determined by signal-to-noise ratio and limited by read noise. Using an NDR capable sensor, read noise is reduced by removing its fixed components by DCDS. Furthermore, the information extracted from several NDR samples can be exploited to filter a random noise out. Secondly, the integrated intensity of each pixel grows linearly with time and, depending on the intensity of the incident light, can saturate before the full integration time is complete. This saturation limits the DR on the higher end. Having several signal samples taken before saturation it is possible to extrapolate pixel intensity to the full integration time.

To enhance the DR of the imaging system at the lower signal range we applied and tested two noise filtering routines described elsewhere: least-squared error signal slope estimation (LSESE) [1] and weighted averaging (WA) [4]. Both assume a linear relation between the intensities of the same pixel taken from different NDR frames and try to estimate a value of a signal slope that is equivalent to a signal to frame number ratio (for uniformly distributed samples in time).

LSE estimation for a signal slope [1, 2] is given by

\[ \hat{Y}_j^{\text{slope}} = \frac{\sum_{i=1}^{n} X_{j,i} \left( i - \frac{n+1}{2} \right)}{n(n-1)}, \quad j = 1...M, \]  

where \( n \) is a number of the last frame before saturation (LSBS), \( j \) is a pixel index, \( X_{j,i} \) is a raw value of the \( j \)th pixel in the \( i \)th frame, \( M \) is a total number of pixels in the image. Here, a uniform distribution of samples over the full integration time is assumed. For \( n = l \) the slope is estimated as...
where \( X_{j,0} \) and \( t_0 \) are a pixel value and a readout time of CDS image, i.e. data collected immediately after the reset signal. It is easy to derive from (1) that using this approach there is no need to perform a DCDS directly as the first \( n/2 \) frames have negative coefficients and, accordingly, remove any fixed offset contribution from the final estimation. However, the CDS image still has to be kept in memory to make evaluation (2) possible. The main advantage of this method is its computational simplicity. Since all coefficients in (1) do not depend on actual pixel intensity they can be calculated for every possible \( n \) and saved in memory before the acquisition process is started. The amount of memory that is necessary to keep these values is relatively small.

According to the WA algorithm proposed in [4] signal slope value is estimated as follows:

\[
\hat{Y}_{j,\text{slope}} = \mathbf{A}_{n,j} \hat{Y}_{n,j}, \quad j = 1...M,
\]

where \( \mathbf{A}_{n,j} \) is a vector of \( n \) averaging coefficients, \( \hat{Y}_{n,j} \) is a column of \( n \) raw slope values calculated after a modified DCDS [4] has been applied. The values of vector \( \mathbf{A}_{n,j} \) are given by the solution of a minimization problem

\[
\begin{align*}
    &\min_{\hat{Y}_{j,\text{slope}}} E(\mathbf{A}_{n} \hat{Y}_{j,\text{slope}} - Y_{j,\text{slope}})^2 \\
    &\text{subject to} \sum_{(i)} \mathbf{A}_{n,i} = 1
\end{align*}
\]

where \( Y_{j,\text{slope}} \) is a true signal slope for the \( j \)th pixel. This algorithm takes into account several types of noise that CMOS sensors are prone to: random read noise, shot noise, reset noise and offset FPN. However, it requires significant computing power that will limit its usage in embedded camera systems.

A straightforward method to enhance DR beyond the saturation level uses LSBS to linearly extrapolate a corresponding signal value. In both algorithms described above this is done by preserving the LSBS estimation for signal slope as a final reconstruction result for a given pixel. This approach may cause undesirable artifacts in image regions containing a well-pronounced gradient of the light intensity (Fig. 1). The intensity gradient makes neighboring pixels saturate in different NDR samples and, accordingly, reconstruction parameter \( n \) in (1) and (4) has different values for these pixels. This brings discontinuity into the signal estimation across an image and, due to presence of noise in the raw data, results in unexpected edges and spots appearing in a problem region. We call this phenomenon a “stepped gradient”. Addressing the issue of “stepped gradient” we developed a new extrapolation routine to reconstruct a signal in saturated regions. It considers a correlation between intensities of adjacent pixels as well as between the intensities of the same pixel taken from different NDR samples. In other words, it assumes that

\[
\hat{Y}_{j,\text{slope}} = X_{j,\text{slope}} - X_{j,0} / t_1
\]

where \( X_{j,0} \) and \( t_0 \) are a pixel value and a readout time of CDS image, i.e. data collected immediately after the reset signal. It is easy to derive from (1) that using this approach there is no need to perform a DCDS directly as the first \( n/2 \) frames have negative coefficients and, accordingly, remove any fixed offset contribution from the final estimation. However, the CDS image still has to be kept in memory to make evaluation (2) possible. The main advantage of this method is its computational simplicity. Since all coefficients in (1) do not depend on actual pixel intensity they can be calculated for every possible \( n \) and saved in memory before the acquisition process is started. The amount of memory that is necessary to keep these values is relatively small.

According to the WA algorithm proposed in [4] signal slope value is estimated as follows:

\[
\hat{Y}_{j,\text{slope}} = \mathbf{A}_{n,j} \hat{Y}_{n,j}, \quad j = 1...M,
\]

where \( \mathbf{A}_{n,j} \) is a vector of \( n \) averaging coefficients, \( \hat{Y}_{n,j} \) is a column of \( n \) raw slope values calculated after a modified DCDS [4] has been applied. The values of vector \( \mathbf{A}_{n,j} \) are given by the solution of a minimization problem

\[
\begin{align*}
    &\min_{\hat{Y}_{j,\text{slope}}} E(\mathbf{A}_{n} \hat{Y}_{j,\text{slope}} - Y_{j,\text{slope}})^2 \\
    &\text{subject to} \sum_{(i)} \mathbf{A}_{n,i} = 1
\end{align*}
\]

where \( Y_{j,\text{slope}} \) is a true signal slope for the \( j \)th pixel. This algorithm takes into account several types of noise that CMOS sensors are prone to: random read noise, shot noise, reset noise and offset FPN. However, it requires significant computing power that will limit its usage in embedded camera systems.

A straightforward method to enhance DR beyond the saturation level uses LSBS to linearly extrapolate a corresponding signal value. In both algorithms described above this is done by preserving the LSBS estimation for signal slope as a final reconstruction result for a given pixel. This approach may cause undesirable artifacts in image regions containing a well-pronounced gradient of the light intensity (Fig. 1). The intensity gradient makes neighboring pixels saturate in different NDR samples and, accordingly, reconstruction parameter \( n \) in (1) and (4) has different values for these pixels. This brings discontinuity into the signal estimation across an image and, due to presence of noise in the raw data, results in unexpected edges and spots appearing in a problem region. We call this phenomenon a “stepped gradient”. Addressing the issue of “stepped gradient” we developed a new extrapolation routine to reconstruct a signal in saturated regions. It considers a correlation between intensities of adjacent pixels as well as between the intensities of the same pixel taken from different NDR samples. In other words, it assumes that

\[
\hat{Y}_{j,\text{slope}} = X_{j,\text{slope}} - X_{j,0} / t_1
\]
III. COLOR RECONSTRUCTION

In case of color filtered sensors a demosaicing problem [8] contributes into the overall complexity of the task. Two approaches are available. The first assumes a two-stage processing when every pixel is reconstructed separately using one of the routines described above to generate a HDR image and then any conventional demosaicing algorithm is applied to create a full-color image. In this case, similarly to black-and-white sensors, the problem of “stepped gradient” artifacts should not be overlooked. Our experiments showed that the extrapolation method described above can be applied directly to the color filtered data. The second approach is a joint processing that utilizes NDR samples to achieve two objects: to enhance DR and to reduce a number and intensity of color artifacts due to noise filtering. To evaluate its efficacy we developed a joint demosaicing, denoising and DR enhancement algorithm similar to the one suggested in [9] for a joint demosaicing and super-resolution problem. It solves a maximum a posteriori (MAP) problem

$$\hat{Y} = \arg \min \sum_{k=R,G,B} \sum_{i=1}^{n} W_i \circ \left( D_k \hat{Y}^k - \sum_{j=1}^{i} U^i_j - Y_d \right)$$

with bilateral spatial and Tikhonov color regularization terms [9, 10]. Here, \(k\) is a color channel index (\(R\), \(G\) or \(B\)), \(D_k\) is a sampling matrix corresponding to the \(k\)th color, \(U^i_j\) is a shot noise value generated between \(j\)th and \((j-1)\)th readouts and estimated as a zero mean Gaussian distribution with variance \(\sigma^2_d = \sum_{k=R,G,B} D_k \hat{Y}^k (t_j - t_{j-1})\). \(Y\) stands for a vector of measured data and \(\hat{Y}^k\) – for a column of estimated signal slope values corresponding to the \(k\)th color. \(W_i\) is a saturation map for the \(i\)th NDR sample. Its elements are equal to zero if corresponding pixels in the \(i\)th NDR sample are saturated, otherwise they have a unit value. Symbol \(\circ\) designates an element-wise matrix multiplication. In (7) the limits of the summation over \(i\) can be altered to exclude several first NDR samples that are particularly noisy and can substantially degrade the quality of reconstruction results.

IV. RESULTS

The experiments used an embedded camera system equipped with an off-the-shelf 640x480 resolution imager in both black-and-white and color versions [11]. Although the camera includes a programmable computational module (FPGA) under control of embedded processor, the major part of calculations is currently implemented in Matlab and run offline. The camera system was used to control the CMOS sensor and to read, store and transfer NDR samples to a host computer. The possibility for embedded implementation of the DR enhancement algorithms is discussed in the next section.

Firstly, the performance of the black-and-white imager in conjunction with noise filtering algorithms was tested. Camera’s initial parameters in normal operational mode were estimated using a PT technique. Then, keeping these data as a reference point, noise and dynamic range performance of NDR-based data processing algorithms was evaluated. Fig. 3 demonstrates a gradual improvement in noise performance while the complexity of the data processing algorithm in use increases. It is evident that DCDS is an essential part of the image post-processing performed in the cameras based on CMOS sensors and NDR gives the most efficient way of doing this. The best noise performance over the full DR is demonstrated by WA method. For low and mid-level signals LSESE demonstrates nearly identical SNR enhancement by an average 1-2 dB; however, the experiments suggest that it takes about five NDR frames to show the advantage over DCDS. WA method also maintains its performance up to saturation whereas LSESE tends to degrade for higher signal values and shows the results similar to DCDS on the upper end of the DR. This can be explained by the fact that its design does not take into account a relation between read and shot noise for different signal levels.

In order to evaluate an effective number of frames that is sufficient for the optimal trade-off between the quality of the output signal and the data processing overhead the integration time value was fixed and NDR algorithms were run for a...
different number of frames. Fig. 4 shows that DR steadily increases when a number of NDR frames grows. However, the major gain of more than 17 dB for WA is achieved for about 16 images – the use of a greater number of samples provides with a further grow of only 0.5-1 dB for as many as 128 images. At the top range DR is limited by the shortest image readout time. For our experimental setup the maximum DR enhancement beyond a saturation level is about 26 dB for full resolution images read out at approximately 230 frames per second. Consequently, the total DR enhancement of up to 44 dB is obtained compared to non-NDR image acquisition.

Secondly, black-and-white pictures of a real HDR scene were captured and two methods of saturated pixels extrapolation were compared. Fig. 5 highlights the superiority of the novel extrapolation method over its LSBS counterpart. The new algorithm fully eliminates undesired artifacts in steep gradient regions.

Finally, the performance of two-color reconstruction methods for Bayer mosaiced NDR sequence was compared. For the two-stage processing simulated NDR samples were firstly processed by WA to generate an intermediate Bayer image that then underwent a color reconstruction using one of conventional demosaicing algorithms. Bilinear interpolation, high-frequency sensitive linear interpolation [13] and joint demosaicing and denoising [12] algorithms were chosen. The former is the fastest and simplest method; the latter is a good representative of traditional demosaicing algorithms that address the correlation between different color channels as well as the sensitivity to high spatial frequencies. The bilinear interpolation takes into account the presence of signal dependent noise that characterizes most images captured by digital image sensors. Reconstruction results were compared with the original full color image using PSNR [14] and S-CIELAB [15] measures (Fig. 6). Simulated NDR samples were generated as follows:

\[
Y_i = \sum_{k=R,G,B} \frac{\hat{Y}}{T} t_k + \sum_{j=1}^{i} U_{Y}^{j} + R_i + C, \quad i = 0, \ldots, n, \tag{8}
\]

where \(\hat{Y}\) is an original full color image, \(T = t_n\) is a full integration time, \(R_i\) represents readout noise, \(C\) stands for reset noise and dark offset FPN, both \(R_i\) and \(C\) were simulated as a Gaussian noise with zero mean and standard deviations \(\sigma_i\) and \(\sigma_c\) respectively. Fig. 6 shows the results obtained for the ‘lighthouse’ image that is part of the Kodak test image set. It is clear that the joint reconstruction method proposed in this work maintains a high reconstruction quality over the bigger range of readout noise and demonstrates better quantitative results.

Fig. 7 shows an example of the HDR image captured by the NDR camera system. Image on the left is scaled linearly and gives a good impression how an image captured by the camera operating in normal integration mode would look like. The necessity to keep a bright edge of the railway station roof in the left part of the image unsaturated makes features in the right dark part of the image almost indiscernible due to a limited DR of the sensor. However, using the NDR mode and a joint color reconstruction described above all the details in both bright and dark parts of the scene were successfully captured.
V. CONCLUSION

Our comparative analysis of LSESE and WA noise suppression algorithms for NDR image DR enhancement shows that LSESE allows for a good trade-off between computational overhead and reconstruction quality provided a sufficient amount of NDR samples are available. Its main advantage is that most calculations can be done in advance and coefficients corresponding to each NDR sample and saturation time can be stored in memory and used later during acquisition. Despite slightly better results of WA its computational requirements make it hard to use in embedded camera systems.

HDR images generated from several NDR samples using previously published algorithms are susceptible to a “stepped gradient” artifact. To our knowledge, this problem is reported for the first time. A proposed extrapolation method prevents this artifact from appearing and can be used for both black-and-white and color mosaiced sensor configurations.

This paper also demonstrates our joint demosaicing and denoising for color HDR image reconstruction based on several NDR image samples. Coupled with motion detection [4] and multi-frame super resolution [9] algorithms this approach can lead to a superior color image quality in comparison with a traditional two- or multistage image processing. The current problem of the joint processing is its relative complexity that may require additional adaptation for embedded implementation.

Concluding, the NDR capability of state-of-the-art CMOS imagers combined with dedicated external data processing provides significant improvements in dynamic range of the processed images.

ACKNOWLEDGMENT

We would like to thank P. Scott, The University of Manchester, for developing embedded code to control a camera with NDR capable image sensor and Dr K. Hirakawa, Harvard University, who kindly gave us a Matlab implementation of his joint demosaicing and denoising algorithm.

REFERENCES