Best Student Paper Award

Evaluation of a High Dynamic Range Video Camera with Non-Regular Sensor

Michael Schöberl[1,2], Joachim Keinert[1], Matthias Ziegler[1], Jürgen Seiler[2], Marco Niehaus[3], Gerald Schuller[3], André Kaup[2], and Siegfried Foessel[1]

[1]Fraunhofer Institute for Integrated Circuits IIS,
Am Wolfsmantel 33, 91058 Erlangen, Germany
[2]Multimedia Communications and Signal Processing,
University of Erlangen-Nuremberg,
Cauerstr. 7, 91058 Erlangen, Germany
[3]Institute of Media Technology,
Ilmenau University of Technology,
Helmholtzplatz 2, 98693 Ilmenau , Germany

ABSTRACT

Although there is steady progress in sensor technology, imaging with a high dynamic range (HDR) is still difficult for motion imaging with high image quality. This paper presents our new approach for video acquisition with high dynamic range. The principle is based on optical attenuation of some of the pixels of an existing image sensor. This well known method traditionally trades spatial resolution for an increase in dynamic range. In contrast to existing work, we use a non-regular pattern of optical ND filters for attenuation. This allows for an image reconstruction that is able to recover high resolution images. The reconstruction is based on the assumption that natural images can be represented nearly sparse in transform domains, which allows for recovery of scenes with high detail. The proposed combination of non-regular sampling and image reconstruction leads to a system with an increase in dynamic range without sacrificing spatial resolution. In this paper, a further evaluation is presented on the achievable image quality. In our prototype we found that crosstalk is present and significant. The discussion thus shows the limits of the proposed imaging system.

Keywords: High dynamic range, image sensor, neutral density filter, computational imaging

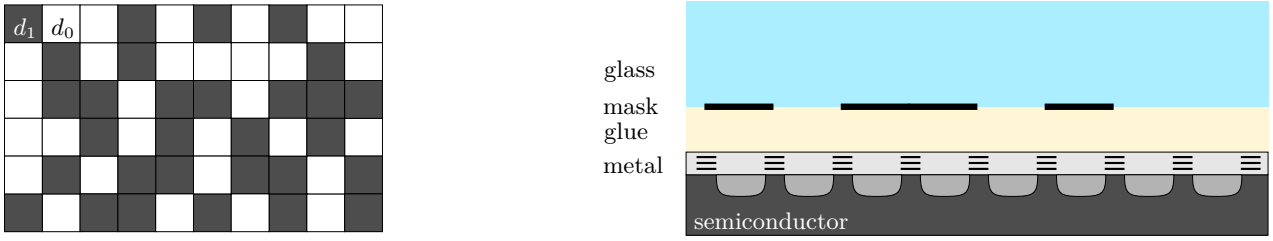
1. INTRODUCTION

Although digital cameras have improved significantly in recent years, the dynamic range is still not on par with the human eye.[1] For capturing a scene with a high dynamic range, multiple approaches are available: A further increase in dynamic range is achieved with computational cameras, which do not capture images directly but require additional processing: The most basic approach uses multiple exposures of the same scene[2] and combines the captured data to a single image.[3] Although this works well for static scenes, this method fails for motion imaging. For moving objects, a synchronized acquisition of multiple cameras and beam splitters is feasible but requires a complex and expensive optical construction of the camera.[4] A much simpler camera construction is achieved with optical attenuation on individual pixels as shown by Nayar and Mitsunaga.[5] This system trades spatial resolution for an increase in dynamic range.

In contrast, our HDR camera principle improves upon this technology of attenuated pixels and avoids the loss in spatial resolution[6] by utilizing a non-regular arrangement of filters. With a specifically targeted image reconstruction based on sparse representations, fine image details can be recovered. Furthermore, we constructed a first prototype of the proposed camera[7] which is evaluated below. With the proposed approach, only an extension toward higher scene intensities is possible. As traditional camera systems are already close to the lower physical limit on camera sensitivity,[8] no significant technological extension toward the lower end of the

For further questions please contact Michael Schöberl at michael.schoeberl@iis.fraunhofer.de

Digital Photography IX, edited by Nitin Sampat, Sebastiano Battiato, Proc. of SPIE-IS&T Electronic Imaging,
SPIE Vol. 8660, 86600M · © 2013 SPIE-IS&T · CCC code: 0277-786/13/$18 · doi: 10.1117/12.2004258



a) non-regular attenuation of pixels

b) first camera prototype side view

Figure 1. Our non-regular HDR sensor principle with a) an exemplary attenuation mask of optical density $d_0 = 0$ and d_1 and b) construction principle of the first prototype camera with an attenuation mask glued onto the sensor.

dynamic range can be expected any more. The acquisition of darker scenes can only be acomplished with larger pixels or lenses with larger aperture diameters. In contrast, here we focus on an extension toward higher scene intensities by optical attenuation of individual pixels.

Our paper is organized as follows: In the following section, the basic principle of the proposed camera is briefly summarized. This includes the sensor pattern, a frequency-domain justification of the non-regular sampling pattern as well as some notes on the image reconstruction. Furthermore, the section gives a detailed analysis of an ideal non-regular camera. Section 3 deals with crosstalk. A general overview about crosstalk and resulting problems is given, followed by a correction approach for our prototype system. The fourth section finally evaluates the performance of the complete processing chain showing a SNR plot for our prototype camera. The paper ends with some example images showing the system performance on a laboratory scene.

2. PROPOSED NON-REGULAR HDR CAMERA

The proposed non-regular HDR camera system has previously been introduced,[6] thus, only a brief overview of the advantage from non-regular sub-sampling and the image reconstruction shall be given here.

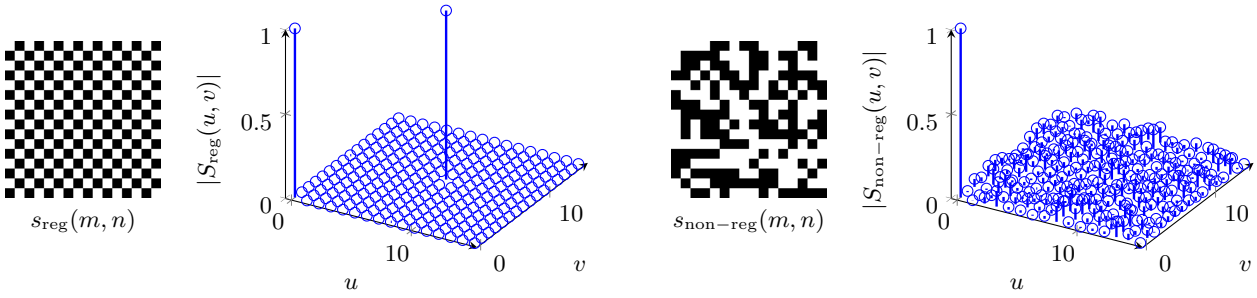
A small section of the proposed image sensor is illustrated in Figure 1 a). The pixel-individual attenuation pattern has half the pixels with optical attenuation of density d_1 and the other half without optical attenuation, i.e. $d_0 = 0$. In contrast to Nayar,[5] a non-regular arrangement and only two steps of attenuation are used. If a scene contains especially bright or dark regions, only one of the pixel types is able to perform a valid acquisition. At the position of the other pixels, the information is lost due to under- or overexposure. This leads to an image with missing pixels which need to be reconstructed.

Figure 1 b) shows the construction of our prototype camera,[7] where an attenuation mask on glass is glued to the sensor surface. The mask has a pixel-individual pattern which is aligned to the pixel grid.

2.1 Advantage of Non-Regular Sub-Sampling

With regular sampling patterns, sub-sampling directly leads to aliasing. As shown in Figure 2 a), the regular pattern shows two distinct peaks in the frequency domain. This leads to spectral overlap which is perceived as aliasing in the sub-sampled image. As both peaks are of the same strength, this leads to ambiguity which can not be resolved in reconstruction. A filtering prior to sampling is required and causes loss in spatial resolution.

In contrast, the proposed non-regular sub-sampling is fundamentally different: The frequency domain representation of the pattern has only a single dominant peak, with additional noise-like parts as shown in Figure 2 b). A dominant frequency of the original image signal is preserved during sub-sampling. With a sub-sequent image reconstruction, an image with high detail and fine structures can be recovered. The underlying principle used here is the existence of sparse representations: For typical natural images, a transform domain representation (like the Fourier domain) exists, in which only few coefficients are sufficient for describing the signal well. These few dominant coefficients are preserved through non-regular sub-sampling and image reconstruction recovers them.



a) regular sub-sampling signal b) non-regular sub-sampling signal

Figure 2. Exemplary sub-sampling signals for half the pixels being known in spatial (left) and frequency domain (right) for a) regular and b) non-regular sub-sampling.

2.2 Sparsity-Based Image Reconstruction

We found that a reconstruction of the image can be achieved with the algorithm of Frequency Selective Extrapolation (FSE).[9] This algorithm performs an iterative reconstruction of image blocks. It is based on the existence of a sparse representation, meaning that the signal can be represented well with only few coefficients in the transform domain. The existence of sparse representations is a widely used property of natural images, which can also be exploited, e.g., in denoising, deblurring and compressed sensing.[10] Corresponding simulations with the FSE algorithm can be found in our publication on non-regular HDR[6] and also in our publication on spatial resolution enhancement.[11]

2.3 Ideal Camera Model for Non-Regular HDR Camera

For the following analysis, an ideal model of the proposed camera is introduced. We assume no read noise, no crosstalk and ideal sensor technology. Later, the results from this model are compared to the measurements of the real world prototype.

The sensor receives a certain amount of light, expressed as a radiant flux Φ for each pixel. As absolute values are not important here, a normalized radiant flux Φ' is used in the following. Depending on the spectral distribution and sensor quantum efficiency, this flux leads to an average number of photons μ being collected during the exposure time. In the ideal case, a pixel without optical attenuation ($d_0 = 0$) receives the full number of photons $\mu_{d_0} = \mu$ during exposure. This pixel has a noise variance of $\sigma^2_{d_0} = \mu$ due to photon shot noise.[12]

In contrast, a pixel with attenuation by a filter of optical density d_1 receives only a fraction of the light intensity. The attenuation can be expressed as the Neutral Density (ND) factor

$$\mathrm{ND} = 10^{d_1 - d_0}. \tag{1}$$

In comparison, the attenuated pixel receives only μ_{d_1} photons with variance $\sigma^2_{d_1}$ with

$$\mu_{d_1} = \frac{1}{\mathrm{ND}} \cdot \mu_{d_0} \qquad \text{and} \qquad \sigma^2_{d_1} = \frac{1}{\mathrm{ND}} \cdot \sigma^2_{d_0}. \tag{2}$$

Finally, the code values of the attenuated pixel needs to be scaled by the factor ND for reconstruction of the final image. This scaling leads to a change in mean and variance of

$$\mu_{d_1,\mathrm{scaled}} = \mathrm{ND} \cdot \mu_{d_1} = \mu_{d_0} \qquad \text{and} \qquad \sigma^2_{d_1,\mathrm{scaled}} = \mathrm{ND}^2 \cdot \sigma^2_{d_1} = \mathrm{ND} \cdot \sigma^2_{d_0} \tag{3}$$

resulting in the same mean brightness as the pixels without attenuation. However, the variance of the scaled signal is increased leading to more noise in the pixels.

The Signal to Noise power Ratio (SNR) of a conventional pixel is calculated[12] as

$$\mathrm{SNR}_{d_0} = 20 \cdot \log_{10} \frac{\mu_{d_0}}{\sigma_{d_0}} \tag{4}$$

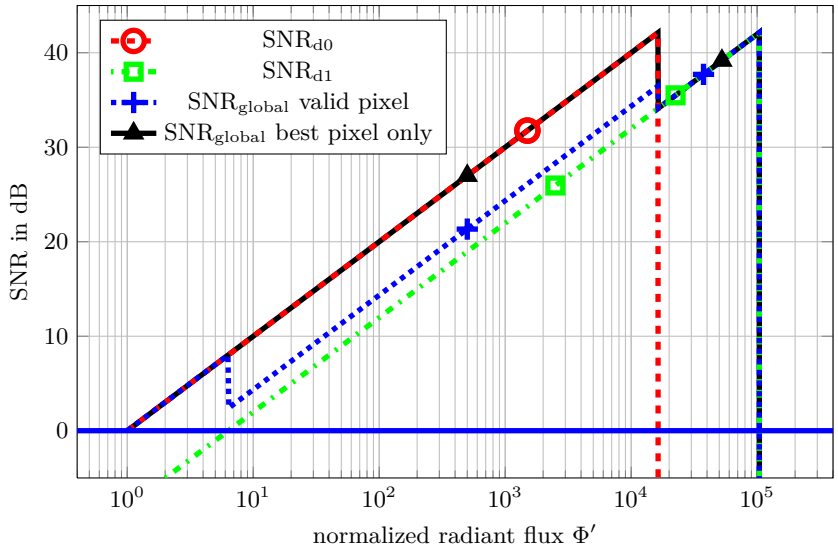


Figure 3. Resulting SNR for scaled sensor raw data of non-regular HDR sensor with curves for regular pixels, attenuated pixels and two selection algorithms for global SNR.

whereas the attenuated pixels can be described as

$$\mathrm{SNR}_{d_1} = 20 \cdot \log_{10} \frac{\mu_{d_1}}{\sigma_{d_1}} = 20 \cdot \log_{10} \frac{\mu_{d_0}}{\sqrt{\mathrm{ND}} \cdot \sigma_{d_0}}. \tag{5}$$

Figure 3 shows the resulting SNR over a normalized radiant flux Φ'. The uppermost curve (SNR$_{\mathrm{d0}}$) represents pixels from the conventional camera without attenuation whereas attenuated pixels have a lower SNR, with the curve shifted to the right. For attenuated pixels, clipping occurs for higher scene intensities.

The result of this process are two types of pixels with equal mean but different variance and SNR. For the following processing we can choose between two different options: At first we only discard under- and overexposed pixels, which can be identified by their code values. Figure 3 shows this curve as SNR$_{\mathrm{global}}$ *valid pixel*. As the sensor has half the pixels with attenuation, the average global variance of the sensor with even illumination calculated as follows:

$$\sigma^2_{\mathrm{global}} = \frac{1}{2}\left(\sigma^2_{d_0} + \sigma^2_{d_1,\mathrm{scaled}}\right) = \frac{1+\mathrm{ND}}{2} \cdot \sigma^2_{d_0} \tag{6}$$

with the global SNR over all pixels being

$$\mathrm{SNR}_{\mathrm{global}} = 20 \cdot \log_{10} \frac{\mu_{d_0}}{\sigma_{\mathrm{global}}} = 20 \cdot \log_{10} \sqrt{\frac{2}{1+\mathrm{ND}}} \cdot \frac{\mu_{d_0}}{\sigma_{d_0}}. \tag{7}$$

The second option is denoted as SNR$_{\mathrm{global}}$ *best pixel only* in Figure 3. In presence of a valid pixel in an image region without attenuation, the pixel value obtained for attenuated pixels is discarded. This leads to a higher overall SNR$_{\mathrm{global}}$ as pixels with a lower SNR are identified and discarded. However, as a result only fewer valid pixels remain.

In this ideal simulation, a different number of valid pixels is generated: Figure 4 shows the ratio of valid pixels for uniform illumination at various scene intensities. Depending on the radiant flux Φ', only some of the pixels in the dynamic range are valid. At the lower end, pixels with SNR < 0 dB and pixels above clipping at the upper end are discarded. The first option of utilizing all valid pixels results in image regions of intermediate brightness where all pixel values are known. In contrast, the selection of the best pixels leads to at most half of the pixels being known. The higher SNR of the second method comes at the cost of requiring a larger effort in image reconstruction as fewer pixels are known.

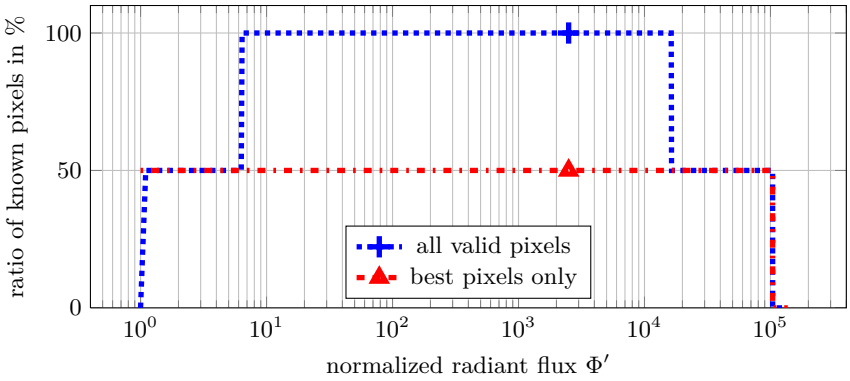


Figure 4. Ratio of known pixels for two selection algorithms from Figure 3.

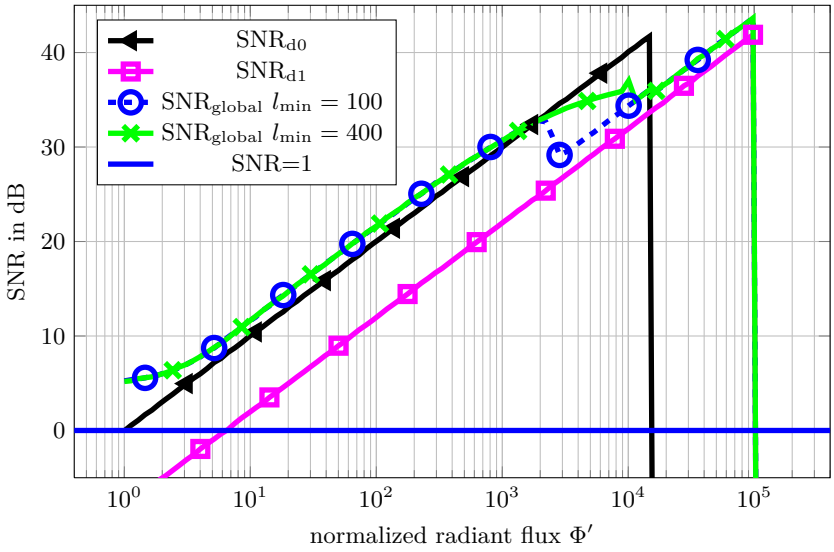


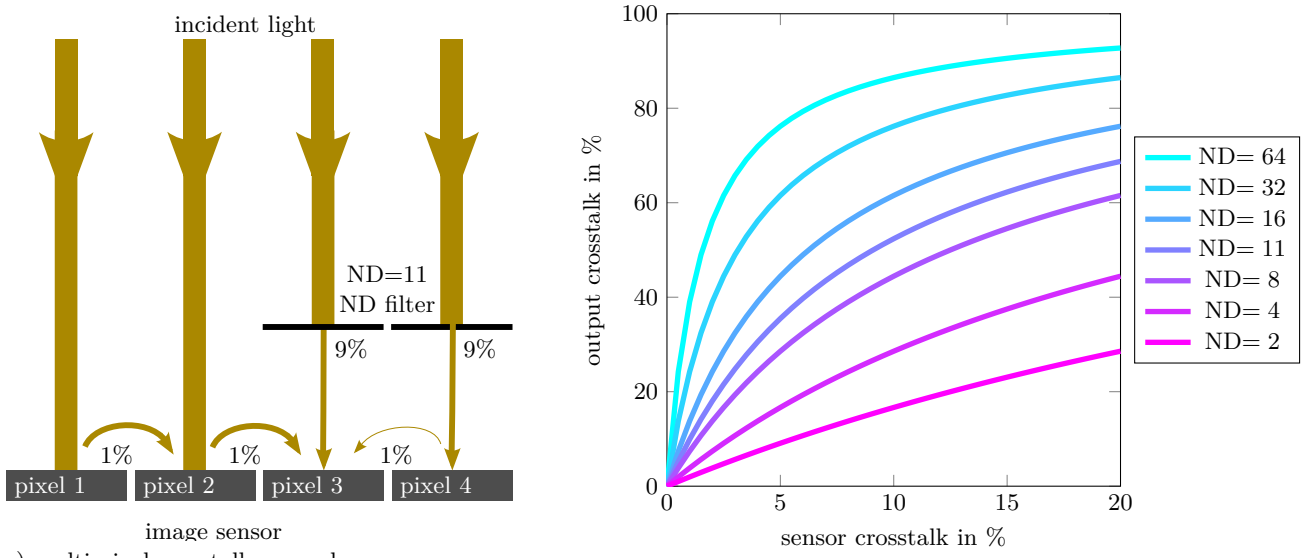
Figure 5. Overall system simulation: A computer-generated gray image with shot noise is captured with the proposed non-regular HDR camera and subsequently missing samples are reconstructed by the FSE algorithm.

These two options represent extreme cases. In practice, we found that a threshold $l_{\min}$ in between works well. Signal values of the attenuated pixel are discarded if their measured value is below $l_{\min}$. This ensures that the attenuated pixels have at least a minimum SNR until they get used.

2.4 Ideal Camera Simulation and FSE Reconstruction

In the following experiment we analyzed the FSE reconstruction with a simulated input image. A uniform gray image with the idealized HDR camera and a simulation of camera shot noise has been used. This results in under- and overexposed pixels as explained above. These missing pixels are then reconstructed with the FSE algorithm.

The resulting curves of this experiment are plotted in Figure 5 in comparison to the ideal curves from above. The two curves with with threshold values $l_{\min} = (100, 400)$ show that the FSE reconstruction is beneficial and increases the SNR. This gain is caused as the random noise pattern is not sparse in the Fourier domain and the application of the FSE leads to averaging. As the FSE reconstruction is image dependent, in general, this gain can not be expected in all natural images. However, it shows that the FSE can have a positive effect on image noise.



a) multi-pixel crosstalk example

b) amplified sensor crosstalk

Figure 6. Analysis of image sensor crosstalk for non-regular HDR image sensor with a) crosstalk example with regular and attenuated pixel cells and b) output crosstalk caused by sensor crosstalk for various ND filter configurations.

3. CROSSTALK IN THE PROTOTYPE CAMERA

In order to evaluate the performance of the proposed method in practice, we have built a first camera prototype.[7] The construction of the proposed prototype sensor is depicted in Figure 1 b) with a filter mask on a glass plate, which is directly glued onto the image sensor. First measurements showed a difference to performed simulations. We identified crosstalk as the underlying reason. In this section, the problem of crosstalk and our compensation strategy is explained.

3.1 Crosstalk Analysis

In our prototype, crosstalk limits the performance of the camera, since the light intended to be captured by a given pixel also influences the neighboring pixels. While this is a general problem with image sensors, the ND filter mask increases this problem as the intensities to be captured by neighboring pixel cells differ heavily. The non-regular attenuation mask enables the analysis of this problem without specialized equipment: With the per-pixel mask, we can analyze pixels with attenuated neighbors and compare them to pixels with a different neighborhood configuration.

For illustration, the simplified situation of two regular and two attenuated neighbor pixel is shown in Figure 6 a). All pixels are exposed to the same amount of bright light. The attenuated pixels will see only part of the light through the ND filter (direct path). Having for instance an ND filter that lets pass only 1/11 of the incident energy, 1/11=9% of the light will reach pixels 3 and 4. In addition to this desired amount of energy, additional light will come from the neighbors due to crosstalk. Assuming for instance that 1% of the light will spill to the neighbor and taking into account that the regular pixel will not block any light, 1% of the overall incident light will spill from pixel 2 to 3. This same ratio of 1% also spills for example from pixel 1 to 2 or from pixel 4 to 3 (other cases of crosstalk are present but not shown). However, in the cases 1 to 2 and 4 to 3, the 1% amounts to only negligible total energy of the affected pixels 2 and 3. Only the contribution from 2 to 3 is serious, as the total energy measured in pixel 3 divides into a ratio of 10% from crosstalk to 90% direct light. Although the underlying sensor crosstalk of only 1% is quite good, the ND filter configuration leads to a high ratio of 10% of effective crosstalk for some pixels.

The consequence of this behavior is depicted by the curves in Figure 6 b). The underlying sensor crosstalk is amplified by the ND factor leading to an increased crosstalk in the measured signal. For a sensor suffering from strong crosstalk only small values for the ND factor can be used. Achieving a large enhancement of dynamic

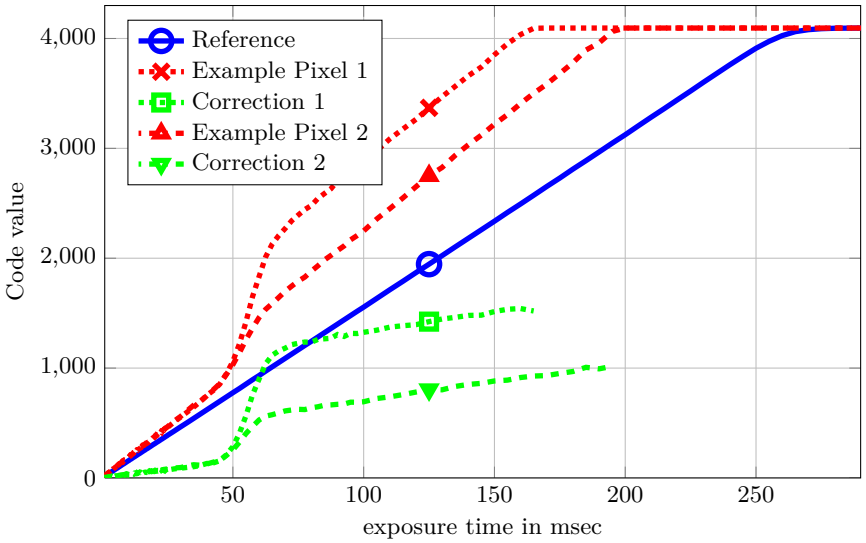


Figure 7. Code values over exposure time for exemplary pixels, reference pixel and crosstalk correction curve.

range requires large ND factors and thus a sensor with only little input crosstalk. This relation is an important step in the theoretic analysis of the proposed HDR principle. Depending on the amount of allowed crosstalk, we can determine the maximum spread in dynamic range that is achievable. This has two consequences: The overall ND factor is limited by sensor crosstalk. Secondly, even within working range, there is quite a lot of crosstalk in the measured signal.

3.2 Crosstalk Correction

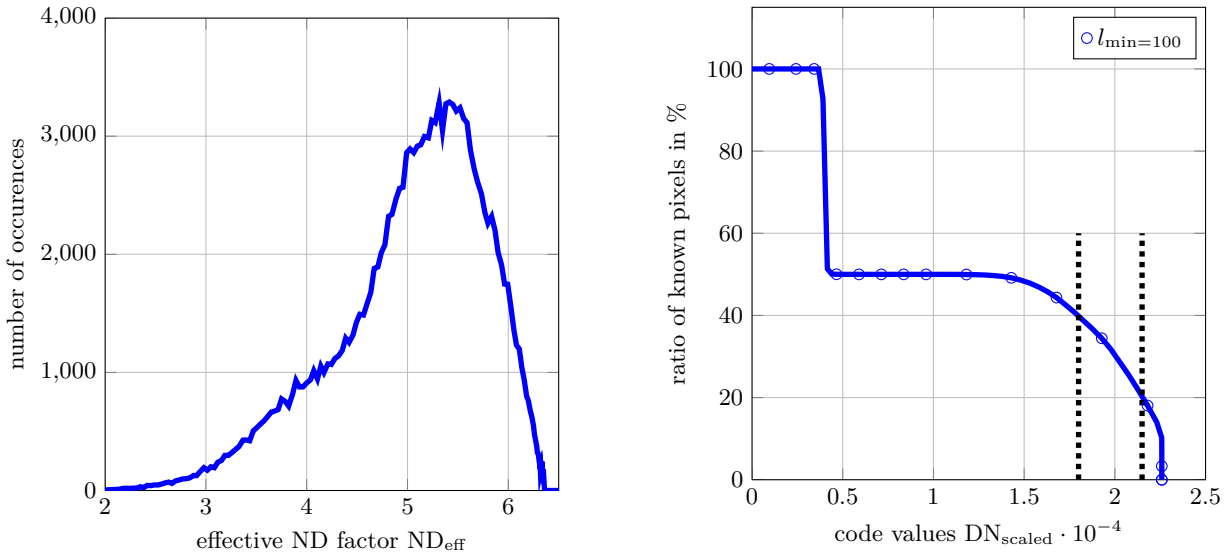
We developed a model to describe the crosstalk portion a pixel receives and subsequently compensate this error.

From the known sensor mask we identified a set of pixels that are affected by crosstalk the least. These pixels are surrounded by pixels equipped with attenuation and thus a pixel in the middle receives only a minimum of crosstalk. In the following, these pixels were used as reference to measure crosstalk for all other pixels. Because of our setup, our sensor contains only very few attenuated pixels with only attenuated neighbor pixels. For this reason we also used pixels with one regular neighbor. For calibrating this model, we exposed the sensor to a uniform constant illumination with variable exposure times. From the captured images we computed one average value for each exposure time.

Based on the reference curve, a difference between each pixel and the reference can be computed. That difference is assumed to be the crosstalk and can be elimated by substraction. Our analysis showed that the crosstalk for each pixel is different as it depends from the layout of its neighbor pixels and other spatially varying parameters that orginate from the manufacturing of our prototype. Based on these results we can build a compensation look-up-table (LUT) containing a correction value for each pixel and for each code value. Please note that this model to compensate for the crosstalk assumes similar luminance values for the pixel surroundings. Although this assumption does certainly not hold for natural images, we found this to be a practical approximation which is successfully able to reduce visible crosstalk.

Figure 7 shows exemplary curves belonging to the reference, two crosstalk-influenced pixel in the middle of the sensor and two resulting correction curves. We found that crosstalk of these attenuated pixels is non-linear for intensitites close to saturation of adjacent pixels without attenuation. This seems to be an issue of the underlying CCD image sensor, which we currently have no influence on. By computing a second reference curve for regular pixels with no more than one attenuated neighbor pixel, it is also possible to measure the obtained ND attenuation factor. For our camera, this factor is approximately ND ≈ 6.35.

As the graph shows, the reference curve shows linear trend which confirms that it has only minimum crosstalk.



a) histogram of effective ND factors

b) ratio of known pixels

Figure 8. Measured results of prototype camera with a) histogram of effective ND factors and b) ratio of known pixels.

4. PROTOTYPE PERFORMANCE EVALUATION

We analyzed our first prototype for the effective ND factor, the increase in dynamic range and the resulting SNR. These measured results are a basis for future improvement of the implementation.

4.1 Effective ND Factor including Crosstalk

Figure 7 shows the raw sensor data of two exemplary attenuated pixels (red curves) with code values plotted over the normalized radiant flux Φ'. From the measured pixels we can see these pixels saturate earlier than expected. Due to crosstalk, these pixels receive more light and thus, have a much lower effective ND factor. At the maximum code value of a pixel we do not measure any useful information any more. The correction values (marked in green) thus are only present for intensities below clipping.

The maximum useful information from a pixel is obtained at a pixel code value or digital number (DN) equal to $DN_{max} = 4095 - 1$. The clipping point of the exemplary pixel 1 from Figure 7 would be mapped to the reference curve, obtaining a code value of $DN_{valid} \approx 2500$. From that crosstalk-compensated code value DN_{valid} a maximum, rescaled digital value $DN_{scaled} = DN_{valid} \cdot ND = 2500 \cdot 6.35 = 15875$ results. Compared to the theoretical maximum value equal to $DN_{max} \cdot ND = 4094 \cdot 6.35 \approx 2.6 \cdot 10^4$ only 61% of the pixels maximum code value are useable. Based on this finding we define an effective ND-factor ND_{eff} denoting the remaining ND factor after crosstalk compensation. For a pixel at position (m, n) we obtain

$$ND_{eff}(m,n) = \frac{DN_{valid}(m,n) \cdot ND}{DN_{max}} = \frac{DN_{scaled}(m,n)}{DN_{max}} \quad \text{with } ND = 6.35 \tag{8}$$

resulting in $ND_{eff} \approx 3.87$ for the exemplary pixel 1.

The resulting effective ND factors are summarized in Figure 8 a) where the measured histogram of the effective ND factor is shown. Please note that the plot shows only the factors for attenuated pixels. The ND factor now allows for a trade-off between loosing too many pixels versus an increase in dynamic range. For example, with an ND_{eff} limited to 2.5, all attenuated pixels deliver valid results. It is up to the application to decide which threshold is still acceptable. With a higher ND_{eff} more pixels are overexposed and need to be reconstructed.

4.2 Dynamic Range

As saturated pixels don't contain useful information anymore, they need to be reconstructed. With each pixel reaching its saturation level at a different intensity, saturation for attenuated pixels does not occur at a single intensity. Instead, the number of usable pixels for a certain scene intensity decreases smoothly. In comparison to the ideal camera in Figure 4, the non-regular prototype shows a different behavior: In Figure 8 b) the measured ratio of valid pixels is shown. One can see, that for low intensities all pixels are valid. Then, the pixels without attenuation saturate and the ratio drops from 100% to 50%. Finally, at the upper end of the dynamic range, the number of valid pixels further decreases as even the attenuated pixels saturate. For a certain scene intensity, we can estimate how many pixels are valid altogether.

The important consequence is as follows: The measurement of dynamic range depends on the number of valid pixels. This is not a hard number as the subsequent image reconstruction is able to recover lost pixels. This is an essential part in our processing chain and a direct comparison to a regular camera is difficult.

In Figure 8 b) we have marked two thresholds for Φ' at 40% and 20% known samples. At these points, the resulting code values equal $\mathrm{DN}_{\mathrm{scaled}} = 18,000$ and $\mathrm{DN}_{\mathrm{scaled}} = 21,500$ respectively. From these values we can estimate the increase in dynamic range in f-stops to $\log_2(\mathrm{DN}_{\mathrm{scaled}}/\mathrm{DN}_{\mathrm{max}})$, giving 2.1 f-stops and 2.4 f-stops. Although this consitutes only a modest increase in dynamic range, the numbers show that an increase in dynamic range is indeed possible with the proposed camera. In comparison, ND = 6.35 corresponds to 2.7 f-stops.

4.3 Non-Regular HDR Camera SNR

As a last important parameter the SNR of our prototype shall be presented. A conclusive result can only be obtained if crosstalk is compensated prior to SNR determination. Based on averaging a sufficient number of measurements over the whole range of exposure times, a crosstalk compensation LUT has been generated and applied to yet another independent measurement to compute the SNR. Reconstruction of missing pixels was not applied in this case. The resulting plot is shown in Figure 9. Measured SNR for attenuated SNR_{d_1} and regular pixels SNR_{d_0} is depicted. The latter directly corresponds to the camera without modification. As it was expected, the curve for attenuated pixels is a shifted version of the one for regular pixels. In comparison, the SNR curve for an ideal camera[12] with slope 0.5 is shown. At the lower end, the measured curves deviate from the ideal curve as additional sensor noise (like read noise) is included in the measurements. For the estimation of a global SNR, we used the threshold l_{min} to decide if an attenuated pixel is valid prior to crosstalk compensation. The $\mathrm{SNR}_{\mathrm{global}}$ is then calculated from valid pixels. Based on l_{min}, the dip in transition region of the global SNR curve can be shifted.

For low radiation (in the left half of the plot) we see a loss in SNR compared to an ideal camera arising from various camera noise sources (readout noise like dark noise and quantization noise). For larger code values, the SNR curve for regular pixels reaches the same slope as the ideal camera until these pixels start saturating. As pixels from the attenuated mask are discarded until they reach a minimum value of l_{min}, the curve for $\mathrm{SNR}_{\mathrm{global}}$ is identical to the regular pixel SNR curve $\mathrm{SNR}_{\mathrm{d_0}}$. If l_{min} is reached, the attenuated pixels are also regarded as valid pixels. Until bright pixels start clippping all pixels on the sensor can be assumed to have valid code values. The loss in spatial resolution is minimum in that region. When all regular pixels are completely saturated, the SNR only depends on the attenuated pixels and the global SNR-curve equals the curve for attenuated pixels $\mathrm{SNR}_{\mathrm{d_1}}$. In comparison to the ideal curve, we can see that noise and crosstalk have an influence, which limits the achieved enhancement in dynamic range.

5. EXAMPLE IMAGES

In Figure 5 we showed that our system gains dynamic range in theory and in practice, based on images of constant luminance and no structure in the images. This setup was necessary for calibration of our crosstalk correction coefficients. This section presents a selection of natural images taken with our prototype camera and an identical unmodified version of the same image sensor.

The image in Figure 10 a) shows a conventional Low Dynamic Range (LDR) image taken with the unmodified camera with a magnified detail in b). As one can see, parts of the image are overexposed. In c), half of the pixels have been discarded randomly, simulating the sub-sampling of our non-regular sensor mask. These pixels have

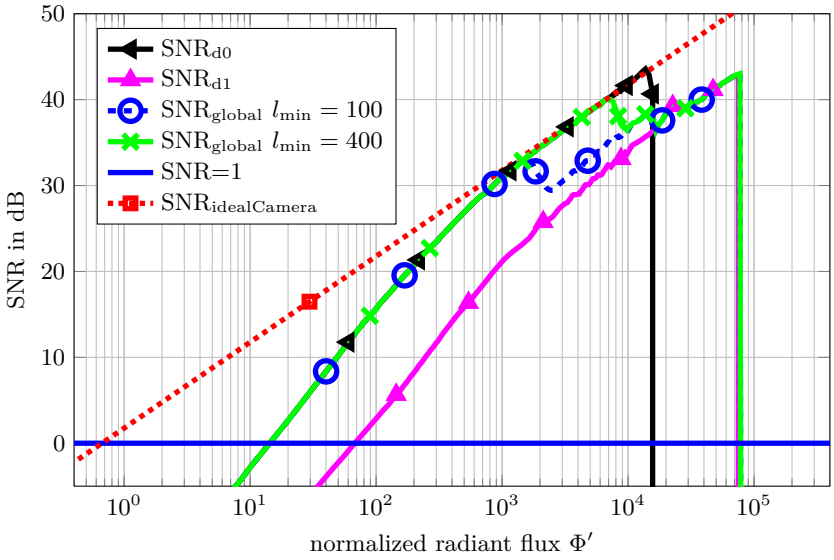


Figure 9. Measured SNR for our non-regular HDR prototype camera, evaluated with crosstalk compensation and variable threshold $l_{\min}$ in comparison to the ideal camera.

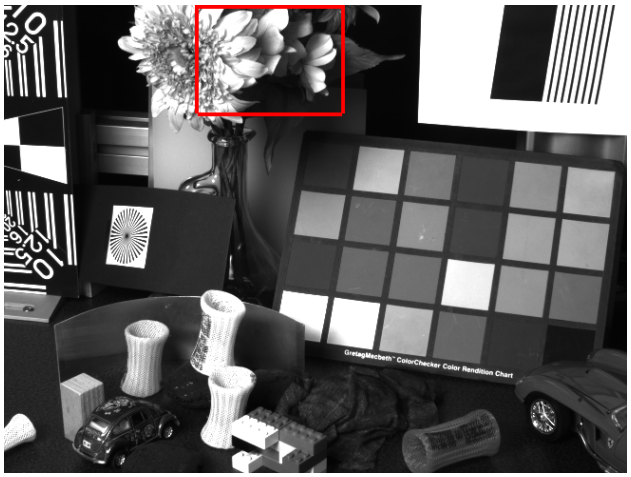
been reconstructed using the FSE algorithm. The output in d) is again an LDR-image. It illustrates that fine details can be recovered well and an image close to the original is obtained.

The images in Figure 10 e) to h) show the same scene captured with our non-regular HDR camera. As the two cameras were mounted side-by-side with a spacing off approximately 6 cm, the images are not pixel aligned. Based on the raw sensor image in e), only valid unattenuated pixels were used for FSE reconstruction in f). Finally, all valid pixels were detected, the crosstalk correction has been applied and missing pixels were reconstructed for the logarithmic tone mapped images in g) and h). The minimum code value for covered pixels to be accepted was set to $l_{\min} = 400$ in these examples. Alltogether, 61% of the pixels were known before image reconstruction. This number can be further split up and results in 37% regular and 24% attenuated known pixel. Upon close inspection this image shows a higher noise especially in flat medium gray areas. For the remaining image parts only very little differences are detectable. The reconstruction of missing samples has finally been carried out with a linear interpolation in g) and the proposed FSE reconstruction in h). In bright regions of the scene, a gain in dynamic range is clearly visible. The advantage of non-regular sampling is visible in comparison of g) and h) and much clearer edges at the leaves of the flower.

The comparison of f) to the simulation in d) shows that additional noise is caused by the additional attenuated pixels included into the calculation. In principle, the non-regular reconstruction from c) to d) does not show these artifacts. The remaining noise pattern in f) is thus caused by non-exact luminance matching of pixels and amplified image noise. We hope that further work on reducing the non-linear crosstalk as the underlying source of the problem and further refinement of our crosstalk compensation are able reduce this problem.

6. CONCLUSION

We proposed an imaging system that is capable of capturing a high dynamic range video without suffering from a loss in spatial resolution. This is achieved by non-regular optical attenuation, which prevents aliasing during sampling. A subsequent image reconstruction algorithm is then able to recover high resolution images by exploiting sparsity. Theoretical discussion and simulation showed a trade-off between dynamic range, noise-enhancement and spatial resolution. These theoretical findings could be confirmed with our prototype system. The most challenging problem is the crosstalk between pixels leading to a strong limitation of our prototype performance. Although we introduced a crosstalk compensation for obtaining visual uniformity, the crosstalk still limits the achievable dynamic range of the prototype. Still, the basic principle of the proposed non-regular HDR sensor could be confirmed and in future work, a refined implementation can reduce these problems.



a) reference image from unmodified (LDR) camera

b) enlarged detail from a)

Figure 10. Visual examples from camera prototype with a) to d) from unmodified LDR camera and e) to h) from proposed non-regular HDR camera prototype.

REFERENCES

[1] Skorka, O. and Joseph, D., "Toward a digital camera to rival the human eye," *SPIE Journal of Electronic Imaging* **20**, 033009 (July 2011).

[2] Mann, S. and Picard, R. W., "On being 'undigital' with digital cameras: Extending dynamic range by combining different exposed pictures," in [*Proceedings of IS&T Annual Conference*], 442–448 (May 1995).

[3] Debevec, P. E. and Malik, J., "Recovering high dynamic range radiance maps from photographs," in [*Proceedings of ACM Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*], 369–378 (Aug. 1997).

[4] Aggarwal, M. and Ahuja, N., "Split aperture imaging for high dynamic range," in [*IEEE International Conference on Computer Vision (ICCV)*], **2**, 10–17 (July 2001).

[5] Nayar, S. and Mitsunaga, T., "High dynamic range imaging: Spatially varying pixel exposures," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], **1**, 472–479 (June 2000).

[6] Schöberl, M., Belz, A., Seiler, J., Foessel, S., and Kaup, A., "High dynamic range video by spatially non-regular optical filtering," in [*IEEE International Conference on Image Processing (ICIP)*], (September 2012).

[7] Schöberl, M., Belz, A., Nowak, A., Seiler, J., Kaup, A., and Foessel, S., "Building a high dynamic range video sensor with spatially non-regular optical filtering," in [*SPIE Optics + Photonics, Applications of Digital Image Processing XXXV*], (August 2012).

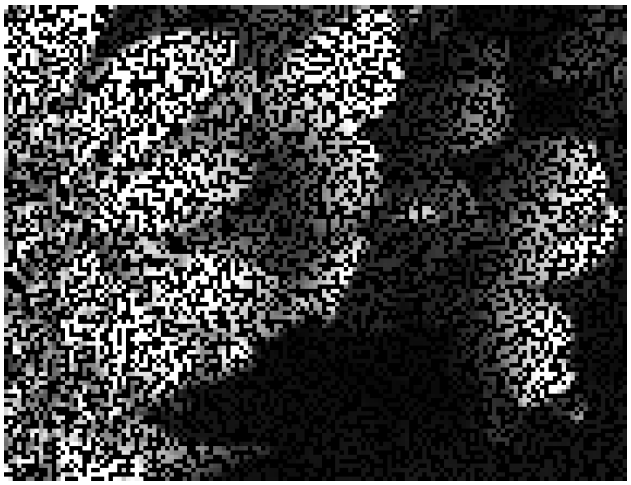
[8] Schöberl, M., Brückner, A., Foessel, S., and Kaup, A., "Photometric limits for digital camera systems," *SPIE Journal of Electronic Imaging* **21** (June 2012).

[9] Seiler, J. and Kaup, A., "Complex-valued frequency selective extrapolation for fast image and video signal extrapolation," *IEEE Signal Processing Letters* **17**(11), 949 –952 (2010).

[10] Elad, M., [*Sparse and redundant representations: from theory to applications in signal and image processing*], Springer (2010).

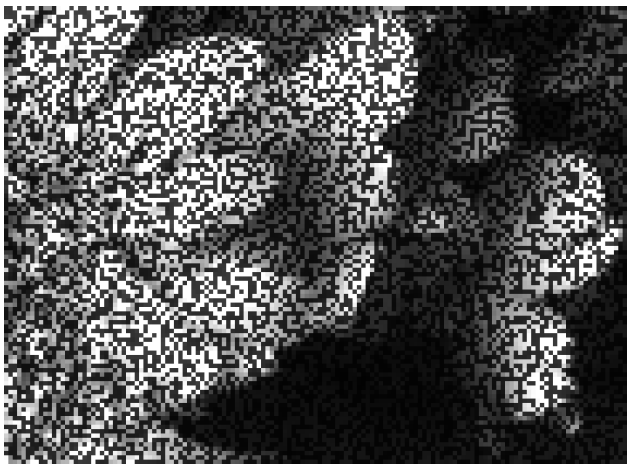
[11] Schöberl, M., Seiler, J., Foessel, S., and Kaup, A., "Increasing imaging resolution by covering your sensor," in [*Proceedings of IEEE International Conference on Image Processing (ICIP)*], 1937–1940 (Sept. 2011).

[12] European Machine Vision Association, *EMVA 1288: Standard for characterization of image sensors and cameras* (Nov. 2010). Release A3.0.



c) 50% of pixels from a) discarded

d) image c) reconstructed using the FSE algorithm



e) enlarged detail from raw HDR camera image

f) reconstructed LDR image based on e)

g) reconstructed and tonemapped HDR image based on e) using triangular interpoation

h) reconstructed and tonemapped HDR image based on e) using the FSE algorithm

Figure 10. (continued) Images g) and h) from the proposed non-regular HDR camera use logarithmic tonemapping.