On Pixel Detection Threshold in the Gigavision Camera

Feng Yang [1], Luciano Sbaiz [1], Edoardo Charbon [2,3], Sabine Süsstrunk [1] and Martin Vetterli [1,4]

[1] LCAV - School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[2] AQUA - School of Engineering
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[3] TU Delft, Mekelweg 4, 2628 CD, Delft, The Netherlands
[4] Department of Electrical Engineering and Computer Science
UC Berkeley, Berkeley CA94720, USA

ABSTRACT

Recently, we have proposed a new image device called gigavision camera whose most important characteristic is that pixels have binary response. The response function of a gigavision sensor is non-linear and similar to a logarithmic function, which makes the camera suitable for high dynamic range imaging. One important parameter in the gigavision camera is the threshold for generating binary pixels. Threshold T relates to the number of photo-electrons necessary for the pixel output to switch from "0" to "1". In this paper, a theoretical analysis of the threshold influence in the gigavision camera is given. If the threshold in the gigavision sensor is large, there will be a "dead zone" in the response function of a gigavision sensor. A method of adding artificial light is proposed to solve the "dead zone" problem. Through theoretical analysis and experimental results based on synthesized images, we show that for high light intensity, the gigavision camera with a large threshold and added light works better than one with unity threshold. Experimental results with a prototype camera based on a single photon avalanche diodes (SPAD) camera are also presented.

Keywords: Computational photography, High dynamic range imaging, low light level imaging, the gigavision camera

1. INTRODUCTION

The architecture of a conventional camera comprises a single lens, an image sensor, an image processing unit, and readout circuitry. The image sensor converts photons to analog electrical signals. The main drawback of a conventional camera is that high dynamic range can be obtained only with relatively complex techniques [1]. In the last few years, many methods for solving this problem have been proposed. One way to generate a high dynamic range image is to combine several pictures with different exposure times [2–5]. The main drawback of this method is the increased overall exposure time and extra computational complexity. Another way is to use sensors with logarithmic response [6]. The main shortcoming of this method is an increased fixed pattern noise (FPN) and potentially lower yield. In [7] we proposed a new image sensor, called gigavision camera, based on the concept of spatial oversampling. In our camera, gray levels are determined, like in a conventional film, by the density of tiny digital pixels that are potentially sensitive to a single photon. These pixels have dimensions comparable to or lower than the diffraction-limited spot as proposed in [8] and thus the gray levels of the overall camera are obtained by low-pass filtering of the binary image, similarly to oversampling techniques in A/D converters [9].

As shown in [7], the sensor's response is non-linear and similar to a logarithmic function, which is suitable for acquiring high dynamic range scenes. By mounting a lens array on the gigavision sensor, an extremely thin camera can be realized with high performance.

We examine here the threshold T, which determines how many photo-electrons are needed to switch the value of the binary pixel from "0" to "1". The ratio of such electrons to absorbed photons is known as quantum

---

efficiency (QE). However, when T is large, there is a "dead zone" in the response function [7]. Current technology does not allow to build a gigavision sensor with a small threshold T. Here, we study the influence of the threshold T on the performance of the gigavision sensor. We propose to avoid the "dead zone" by adding artificial light or background electron noise.

Adding "artificial" light to change the behavior of a sensor has been used before. In silver-halide photography, film emulsions can be uniformly pre- or post-exposed to render them more sensitive. Recall that the formation of a latent image, which acts as catalyst for the development process, requires a cluster of approximately four silver ions to be reduced to silver atoms by the free electrons created during exposure. Uniform pre-exposure can thus produce some silver atoms, but not enough to form a latent image. Any additional image forming exposure then only needs to reduce the additional silver ions necessary to create the minimum number of silver atoms that complete the latent image. A similar effect can be achieved with uniform low-intensity post-exposure [10].

We show that under low light intensity condition, the performance of a gigavision sensor with small threshold is better than that of a gigavision sensor with large threshold. However, when light intensity is high, a large threshold and using the presented method to remove the "dead zone" works better. Therefore, there is a tradeoff between the technology that determines T and the performances in the design of the sensor.

2. THE GIGAVISION CAMERA

2.1 Camera architecture

The architecture of a gigavision camera is shown in Figure 1. The incident light is concentrated on the image sensor through a single lens and converted into an electrical signal. As opposed to a conventional camera, the sensor is binary. A conventional gray level image can be obtained by low-pass filtering and sampling the binary image. Similarly to [7], in this paper the low-pass filter corresponds to the sum of a block composed by N pixels.
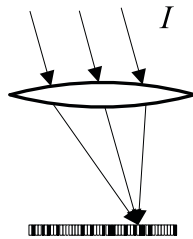

Figure 1. Simplified architecture of a gigavision camera. The incident light is focused by the lens and then impinges on the image sensors.

2.2 Pixel model

Figure 2 shows the model of a gigavision sensor pixel. As described above, the binary image is converted to a gray level image by adding N neighboring pixels. Note that overlapping pixel matrices are possible, as well as non-adjacent and non-uniform matrices. The gigavision pixel is similar to a conventional camera pixel except that the quantizers Q' are binary with threshold T. The number of photons impinging on the pixel can be modeled as a Poisson process. If the area of the pixel is A/N, the intensity of the process is IA/N, where I is the light intensity per square meter and the average number of arrivals during the exposure time t is $\lambda = ItA/N$ (we assume $QE = 1$). The response of each binary pixel K_i, $i = 1, \ldots, N$ is obtained by comparing the number of arrivals S_i, i.e. the electrons due to detected photons, with the threshold T. The quantities K_i are thus binary random variables with parameter

$$p_\lambda = \mathbb{P}[S_i \geq T] = \sum_{k=T}^{\infty} e^{(-\lambda/N)} \frac{(\lambda/N)^k}{k!}.$$

The result after the addition of N pixels is given by $C = \sum_{i=1}^{N} K_i = \sum_{i=1}^{N} Q'(S_i)$. The quantity C is a binomial random variable with parameters N and p_λ, since it is the sum of N binary independent random variables.
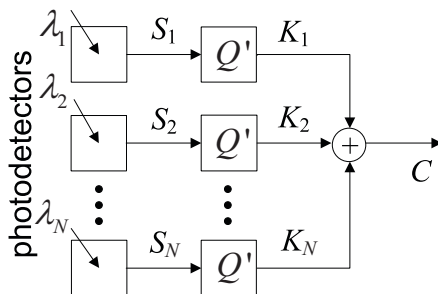

Figure 2. Simplified block diagram of a gigavision pixel. λ_i is the light intensity. The photodetector converts photons to electrical signal S_i, which is then quantized by one-bit quantizer Q' with threshold T. K_i is the quantized value. The pixel value C is the sum of N pixels.

3. INFLUENCE OF THE THRESHOLD

The response function of a gigavision pixel is defined as the relation between photon intensity and expectation of the pixel value $\mathbb{E}[C]$. For simplicity, we assume a QE of 1, i.e. one photon is converted to one electron. A small threshold of 1 photon (i.e., $T = 1$) would be ideal, as it results in a very highly sensitive camera. It is, however, very difficult to set a small threshold with current sensor technology. If the threshold T is large, on the other hand, there will be a "dead zone" in the response function of the pixel, i.e. there is zero response for an interval of increasing photon rate.

Figure 3 shows the response function of a gigavision pixel with different thresholds $T = 1, 15, 30, 45$, and 60, and $N = 256$. It is apparent that when the threshold becomes larger, the "dead zone" interval is larger. The intuition is that when the threshold is large, it is unlikely that enough photons arrive on the same pixel to result in a "1" pixel value, thus the probability is small. For a small threshold, when the light intensity is large, the probability that every pixel becomes "1" is much higher.
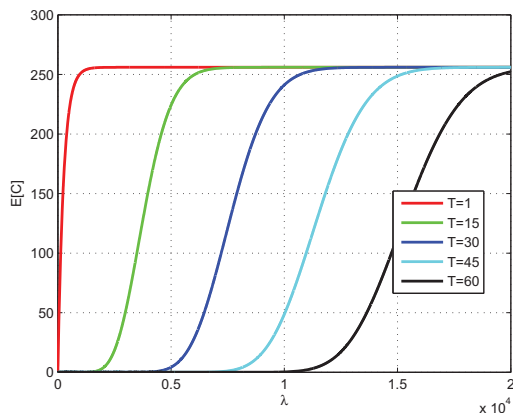

Figure 3. Response function of a gigavision pixel with thresholds $T = 1, 15, 30, 45$ and 60. The oversampling factor is $N = 256$.

However, a large threshold also results in a higher dynamic range. Dynamic range can be illustrated as interval of the λ-axis in Figure 3 where the values of the response function is not "0" or saturated. Note that the larger the threshold, the higher the dynamic range.

The estimation method for light intensity λ is a maximum likelihood estimator as in [7].

$$\hat{\lambda} = \arg\max_{\lambda} \mathbb{P}[C = c]$$

$$= \arg\max_{\lambda} \binom{N}{c} p_\lambda^c (1 - p_\lambda)^{N-c}, \tag{1}$$

i.e. $\hat{\lambda}$ is chosen so that, $p_{\hat{\lambda}} = c/N$.

The estimation variances, $\mathbb{E}[(\hat{\lambda} - \lambda)^2]$ of the gigavision pixel with threshold $T = 1, 15, 30, 45,$ and 60, and $N = 256$, are shown in Fig 4. Because of the "dead zone", the estimation variance for $T = 15, 30, 45,$ and 60 is larger than a gigavision pixel with threshold $T = 1$. But at high light intensity when the pixel with threshold $T = 1$ saturates, the estimation variance is larger than for a gigavision pixel with threshold $T = 15, 30,$ and 45. If the light intensity is larger than the saturation value, the estimation value will always be the minimum saturation value, so with increasing light intensity, the estimation variance first becomes small and then large.
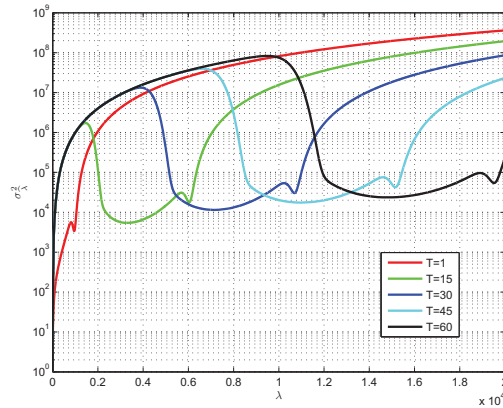

Figure 4. Estimation variance of a gigavision pixel with thresholds $T = 1, 15, 30, 45$ and 60. The oversampling factor is $N = 256$.

One way to solve this problem is to add some artificial light or to increase the background electron noise. Here the intensity of the artificial light is noted as λ_{add}. Figure 5 shows the response function when $\lambda_{add} = 0, 5000, 10000,$ and 15000. The estimation variance is shown in Figure 6.
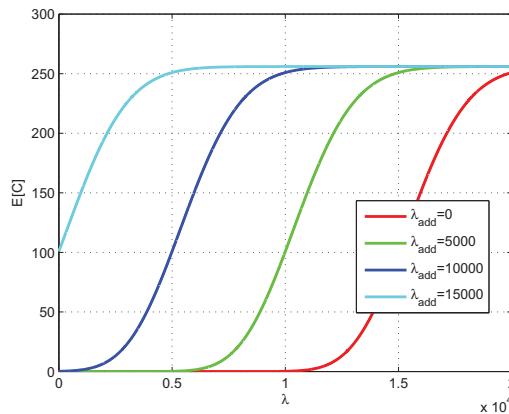

Figure 5. Response function of a gigavision pixel with thresholds $T = 60$, the oversampling factor $N = 256$, and different added light with intensities $\lambda_{add} = 0, 5000, 10000,$ and 15000
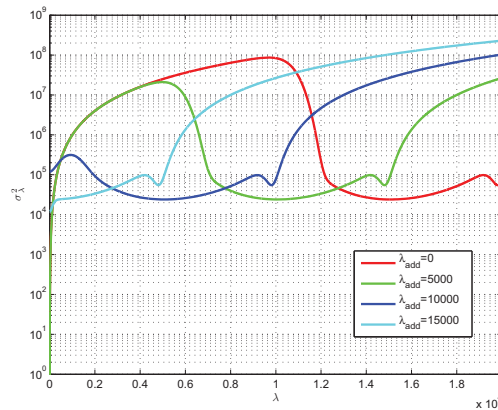
Figure 6. Estimation variance of a gigavision pixel with thresholds $T = 60$, the oversampling factor $N = 256$, and different added light with intensities $\lambda_{add} = 0, 5000, 10000$, and 15000.

If we consider a certain intensity range, for example $0 - 10^4$, we can see that when the intensity of the added light is small, the "dead zone" is not completely removed, e.g. for $\lambda_{add} = 5000$. When the intensity of the added light is too large, however, the estimation variance is large, e.g. $\lambda_{add} = 15000$, since uncertainty introduced by the added light is dominant. When the light intensity is large and the intensity of the added light is large, the estimation variance increases (compare the curve $\lambda_{add} = 10000$ and $\lambda_{add} = 15000$ in Figure 6.). Note that at low light, $\lambda_{add} = 15000$ is smaller than $\lambda_{add} = 10000$, due to better removal of the "dead zone". However when the light intensity increases, the $\lambda_{add} = 15000$ is larger than $\lambda_{add} = 10000$, because of the uncertainty introduced by the added light. Thus, there is a tradeoff between the width of the "dead zone" and the estimation performance.

To study the performance for different thresholds, and with added light, we have to rescale the intensity axis in order to cover the same range. To compare the performance under the same condition, we consider the response function in the different ranges where the expectation of the pixel value $\mathbb{E}[C]$ is larger than $N \times 0.02$ (for $T = 1$, this value is 0) and smaller than $N \times 0.98$. The λ axis is scaled to the same range ($[0,1]$).
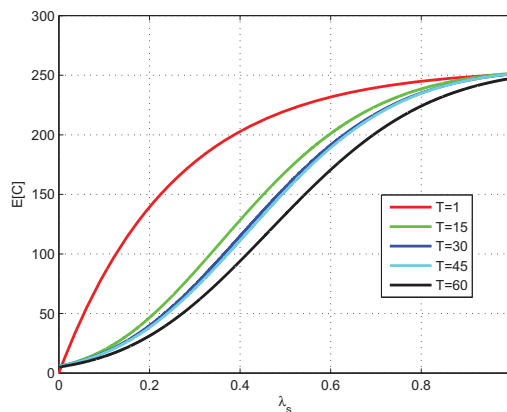

Figure 7. Response function of a gigavision pixel with thresholds $T = 1, 15, 30, 45$, and 60. The oversampling factor is $N = 256$. The curves are obtained by setting the exposure time and the added light in order to cover the value of $\mathbb{E}[C]$ in the range $N \times 0.02$ and $N \times 0.98$.

Figure 7 shows the scaled response function and Figure 8 plots the corresponding estimation variance. In Figure 8, note that when λ is small, a smaller threshold achieves better estimation performance, but when λ is large, a larger threshold obtains better estimation performance. From this, we can conclude that for better noise performance, we need a larger threshold. However, when the threshold is large, the camera is less sensitive at
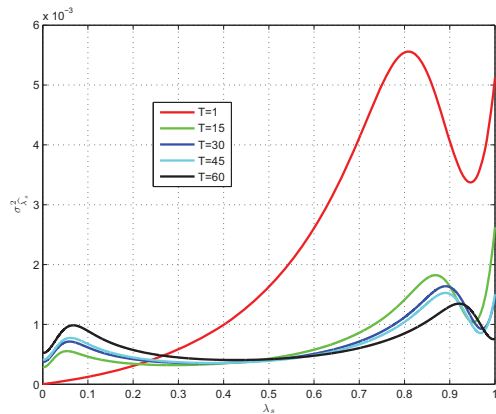
Figure 8. Estimation variance of a gigavision pixel with thresholds $T = 1, 15, 30, 45$, and 60. The oversampling factor is $N = 256$. The curves are obtained by setting the exposure time and the added light in order to cover the value of $\mathbb{E}[C]$ in the range $N \times 0.02$ and $N \times 0.98$.

low light condition.
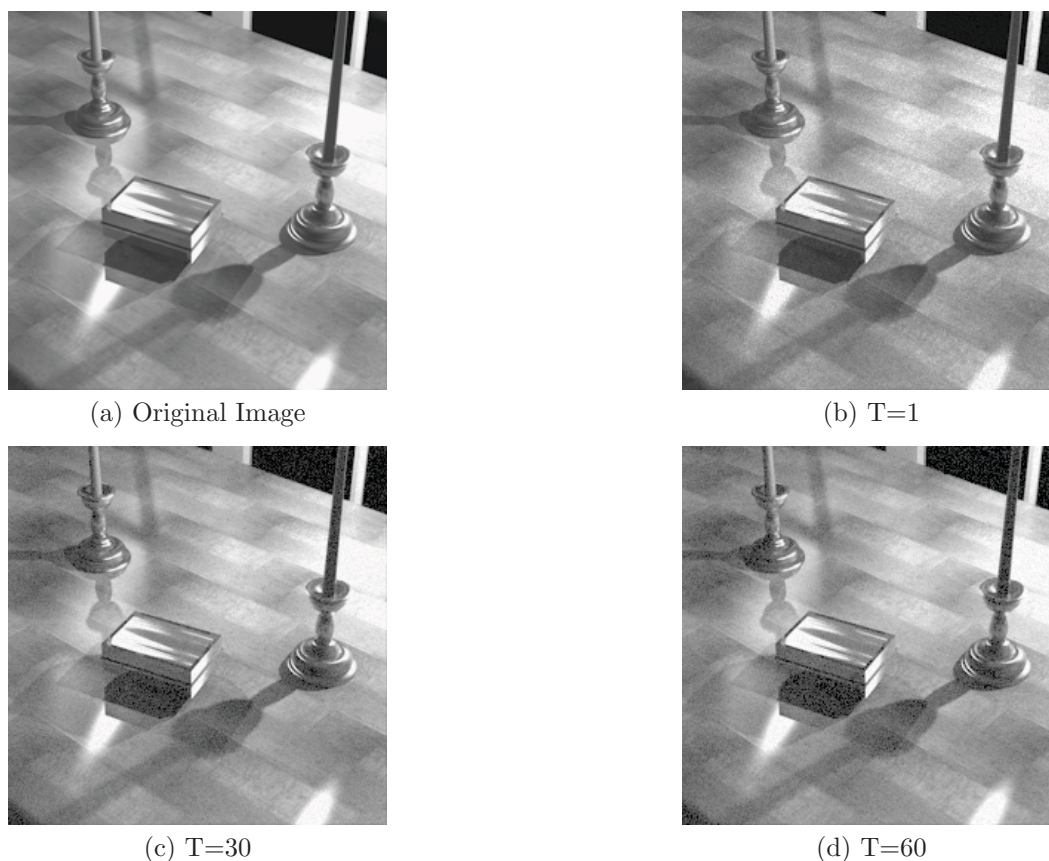

(a) Original Image

(b) T=1

(c) T=30

(d) T=60

Figure 9. Original image and acquired images with gigavision sensors of different thresholds $T = 1, 30$, and 60. The number of photons at each pixel and for each sensor leads to a different estimation of the observed image.

4. SIMULATION RESULTS

We simulated the acquisition of an image with a gigavision sensor using a high dynamic range image 'rend06_oB1D'[11]. Each image pixel has a gray level in the range [0, 1]. We assume that the gray level corresponds to the light intensity, i.e. we set λ equal to the gray level value times exposure time. For each pixel, we generate the random number of detected photons according to the Poisson distribution of parameter λ and we simulated the behavior of the gigavision sensor. The maximum likelihood estimator gives the measured image, i.e. an approximation of the intensity λ. The results are shown in Figure 9. Figure 9(a) is the original image and Figure 9(b), (c), and (d) are the images captured by a gigavision camera with $T = 1, 30$, and 60, respectively. As expected from the analysis of the previous section, the sensor with $T = 1$ gives a smaller error at low light intensity and the sensor with $T = 30$ and 60 has a smaller error at high light intensity. The mean square error (MSE) for the images taken by the gigavision camera with $T = 1, 30$, and 60 are $1.5e - 3$, $6.120e - 4$, and $5.661e - 4$.
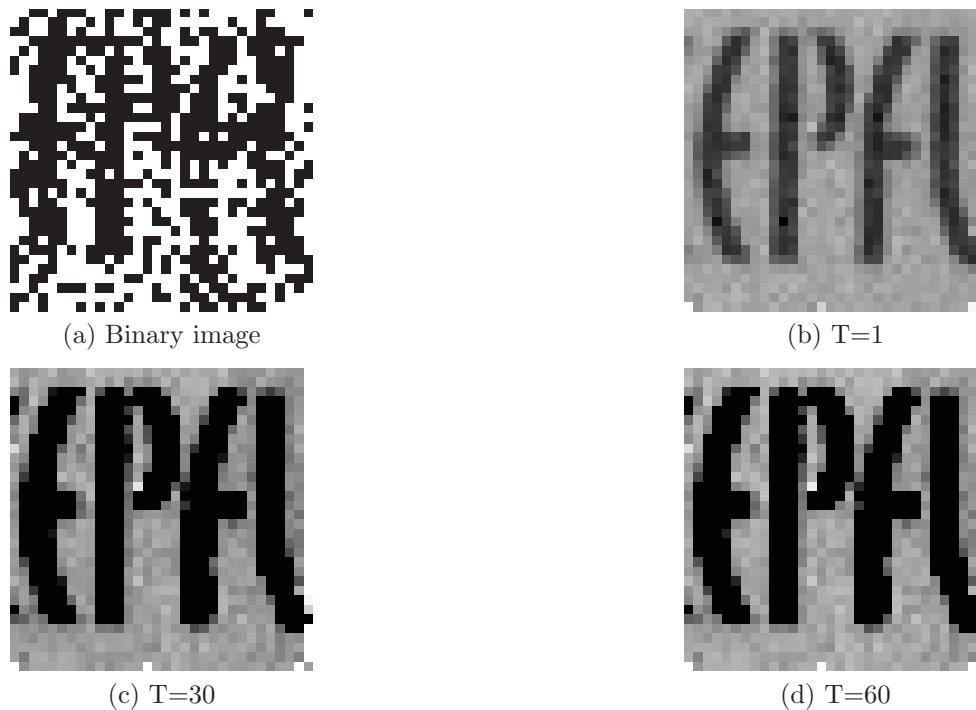

Figure 10. The SPAD camera

(a) Binary image

(b) T=1

(c) T=30

(d) T=60

Figure 11. A binary image and the estimated light intensity for thresholds "1", "30", and "60".

We also did some experiments with the SPAD camera [12], shown in Figure 10. The pixel value of the SPAD camera is "1" and "0". If there is at least one photon hitting on the pixel, its value will become "1", otherwise "0". The resolution of the SPAD camera is 32×32. Due to the small resolution, we can not obtain an image with many gray levels by only using the spatial binary samples. Multiple exposures is used to compensate for

the low spatial resolution. One picture is taken during each exposure. The exposure time for each picture is $4\mu s$. For the threshold "1", we can directly use the binary images taken by the SPAD camera to estimate the light intensity using equation (1). For thresholds larger than "1", we need to make the follow assumption that during each exposure time, the probability that more than "1" photons hitting on the pixel is very small. Through adding M binary images, the SPAD camera becomes a photon-counting device. A binary image can then be obtained by thresholding the outcome of this photon-counting device. We choose $M = 45$ for threshold "30", and $M = 95$ for threshold "60". For each threshold, "256" binary images are taken. One of the binary images taken by the SPAD camera, and the estimated light intensity for thresholds "1", "30" and "60" are shown in Figure 11(a),(b),(c), and (d) respectively. We can see that the estimated light intensities for thresholds "30" and "60" are noisier than that for threshold "1". This is because the light intensity is low.

5. CONCLUSION

In this paper, the influence of pixel detection threshold in the gigavision camera system is studied. We found that when the threshold is large, there is a "dead zone" in the response function of gigavision pixel. This is a crucial problem in the practical design of the gigavision camera, since it is difficult to make a gigavision camera with threshold equal to 1. We propose to add artificial light to the gigavision pixel to avoid this problem. The theoretical analysis and the simulation results shows that under low light intensity condition, the performance of a gigavision pixel with a small threshold is better than that of a gigavision pixel with a large threshold. When the light intensity is high, a gigavision pixel with a large threshold works better. This tradeoff between technology and system performance needs to be taken into account when designing a gigavision sensor and the ancillary imaging processing.

REFERENCES

1. B. Hoefflinger, *High-Dynamic-Range Vision — Microelectronics, Image Processing, Computer Graphics*, Springer, Berlin Heidelberg, 1998.
2. P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. of SIGGRAPH*, pp. 369–378, 1997.
3. M. D. Grossberg and S. K. Nayar, "High dynamic range from multiple images: Which exposures to combine?," in *ICCV Workshop on Color and Photometric Methods in Computer Vision*, 2003.
4. W. C. Kao, "High dynamic range imaging by fusing multiple raw images and tone reproduction," *IEEE Transactions on Consumer Electronics* **54**, pp. 10–15, February 2008.
5. L. Meylan and S. Süsstrunk, "High dynamic range image rendering using a retinex-based adaptive filter," *IEEE Transactions on Image Processing* **15**, pp. 2820–2830, September 2006.
6. S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, "A logarithmic response CMOS image sensor with on-chip calibration," *IEEE Journal of Solid-State Circuits* **35**, pp. 1146–1152, August 2000.
7. L. Sbaiz, F. Yang, E. Charbon, S. Süsstrunk, and M. Vetterli, "The gigavision camera," in *IEEE Conference on Acoustics, Speech and Signal Processing*, pp. 1093–1096, 2009.
8. E. R. Fossum, "What to do with sub-diffraction-limit (sdl) pixels?," in *International Image Sensor Workshop*, pp. 214–217, 2005.
9. J. C. Candy and G. C. Temes, *Oversampling Delta-Sigma Data Converters — Theory, Design and Simulation*, IEEE Press, New York, 1992.
10. T. H. James, *The Theory of The Photographic Process*, Macmillan Publishing Co., Inc., New York, 4th ed., 1977.
11. G. Ward, "High dynamic range image examples." http://www.anyhere.com/gward/hdrenc/pages/originals. html.
12. L. Carrara, C. Niclass, N. Scheidegger, H. Shea, and E. Charbon, "A gamma, x-ray and high energy proton radiation-tolerant cmos image sensor for space applications," in *IEEE International Solid-State Circuits Conference*, pp. 40–41, 2009.