Response Curve Programming of HDR Image Sensors based on Discretized Information Transfer and Scene Information

Darmont Arnaud; Aphesa SPRL; Sprimont, Belgium

Abstract

Most of the snapshot HDR (High Dynamic Range) image sensors have a non-linear, programmable, response curve that requires multiple register settings. The complexity of the settings is such that most algorithms reduce the number of parameters to only two or three and calculate a smooth response curve that approaches a log response. The information available in the final image depends on the compression rate of the response curve and the quantization step of the device.

In this early stage proposal, we make use of scene information and discrete information transfer to calculate the response curve shape that maximizes the information in the final image. The image may look different to a human but contains more useful information for machine vision processing.

One important field of use of such sensors with programmable dynamic range is automotive on-board machine vision and more specifically autonomous vehicles.

Introduction

Non-linear sensors response

The most common approach to increase the dynamic range of an image sensor is to respond non-linearly to light intensity. Logarithmic sensors have been used since the 90s but were not programmable and their SNR performance was not as good as expected in today's standards. Piecewise linear response sensors provide a response curve that approaches that of a logarithmic sensor using multiple linear segments. Moreover, the segments can be individually controlled so that the response curve can be shaped as desired. Sensors with up to six segments and dynamic ranges exceeding 150dB have been reported [1] but the most common implementations of this type of sensor has three segments.

Most of the sensors are designed in usual CMOS image sensor technology and control the pixels with additional signals. A common approach is to use a regular 3T pixel and add additional transitions to the reset gate signal during exposure to clamp or reset the pixel to intermediate levels if the voltage in the photodiode varies too fast, i.e. if the light intensity (or photocurrent) is too high. Each level will correspond to a linear segment of the sensor's response [2]. There are also similar solutions for global shutter pixels with 4, 5, 6 or more transistors.

As each segment is programmable in terms of responsivity (the slope) and reset level (related to the height of the kneepoint between each segment), it offers two degrees of freedom. This is similar to a linear image sensor offering exposure time (or gain) and offset (black level) control. If the sensor has N segments, it will have 2*N degrees of freedom. For example, a sensor with two segments will have the following four degrees of freedom: offset (black level), total exposure time, ratio of exposure time between the first and the second segment and height of the kneepoint between the first and the second segment. If gain is available, it can also be considered as an additional degree of freedom, making the total 2*N+1.

Dynamic Range Gaps

In some conditions, i.e. slope configurations, such a piecewise linear response can generate artifacts called Dynamic Range Gaps [4] or SNR holes as named by Dirk Hertel in his original paper [3]. At that time, we were working on HDR image sensors for automotive and the control of the dynamic range was such a problem that customers were using a limited set of fixed settings only based on lab experiments and switched from one set of parameters to another based on deterministic histogram decisions.

At the kneepoint, the signal-to-noise ratio (SNR) can drop below 1 (0dB) and therefore the details within the image are no longer noticeable (we can extend this limit to any acceptable SNR limit like for example 5). A very good example of the situation is provided in [3] and in [4] and reproduced below at figure 1, 2 and 3. In the top image (figure 1), a 100dB scene is captured with a linear sensor with limited dynamic range and saturation is obvious in the brightest areas of the scene. No information can be retrieved from the saturated areas. Most of the image is properly acquired but of course the performance is limited to the available SNR and therefore the information is limited in the darkest area due to the lower SNR.

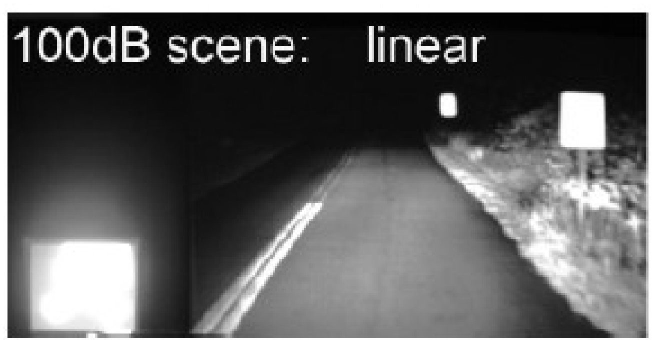

Figure 1. Dynamic Range Gap artifact – linear image with saturation

In figure 2, the same scene is captured with a 100dB sensor with strong compression (one barrier or kneepoint, meaning two segments). The image is no longer saturated, as can be seen by the readable text on the reflective road sign, but large parts of the image fail to provide any information. The parts of the image that are all of the same gray level correspond to that specific irradiance range for which the dynamic range gap occurs. Therefore no information can be retrieved from these parts of the image. We can see that even though the dynamic range has been increased, the image may be less useful.

In figure 3, the same scene is again captured with the sane 100dB sensor but this time using five barriers or kneepoints, i.e. six segments. As the compression is not as strong and the response curve is smoother, there are no dynamic range gap artifacts.

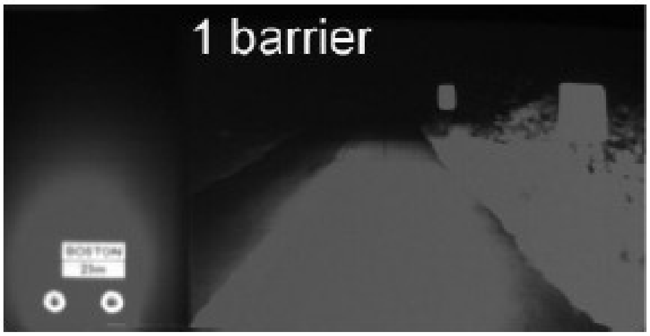

Figure 2. Dynamic Range Gap artifact – two segments image without saturation but with details lost in the dynamic range gap

Figure 3. Dynamic Range Gap artifact – correct image with six segments

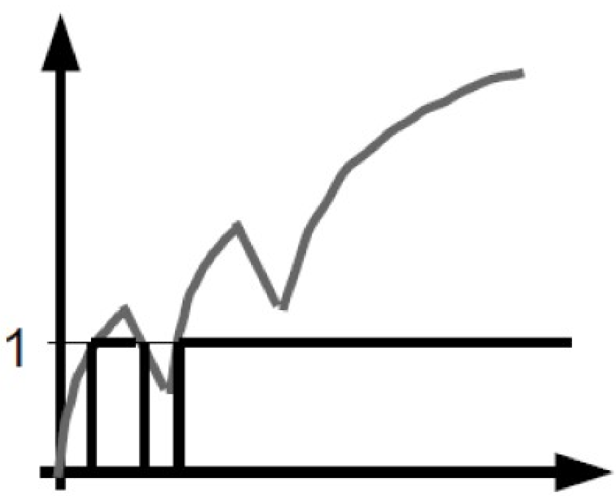

Figure 4. Dynamic Range Gap presence function (black) vs SNR plot (gray) of a three segments response HDR sensor

In previous work and in [4], I have defined the dynamic range gap presence function that is 0 for the sensor irradiance ranges where SNR<1 and 1 outside of these ranges. It is explained in figure 4. Again the value of 1 can be changed depending on the application's requirements

For a two-segment response with a kneepoint at $q_k = q_{max}\theta$ electrons, $0<\theta<1$, an exposure time t_{int} and a slope change time t_1 , the sensor's response is given by

$$f=\begin{cases} \int_0^{t_{int}} I(t)\,dt & \text{if} \quad 0 \le i_{ph} < \dfrac{q_k}{t_1} - i_d \\ q_k + \int_{t_1}^{t_{int}} I(t)\,dt & \text{if} \quad \dfrac{q_k}{t_1} - i_d \le i_{ph} < \dfrac{q_{max}(1-\theta)}{t_{int}-t_1} - i_d \\ q_{max} & \text{otherwise} \end{cases}$$

where i_d is the dark current and i_{ph} is the photocurrent. The sensor design is such that each slope has less response than the previous one and therefore it is not possible to generate any form of curve.

The SNR curve and the dynamic range (DR) can be obtained by calculations of the signal and the noise of this simple pixel model, as made in [4], [5] and in [6]. The dynamic range is

$$DR = \frac{i_{max}}{i_{min}} = \frac{\dfrac{q_{max}(1-\theta)}{t_{int}-t_1} - i_d}{i_{min}}$$

and the SNR is

$$SNR(i_{ph}) = \begin{cases} \dfrac{i_{ph} t_{int}}{\sqrt{q(i_{ph}+i_d)t_{int} + \sigma_r^2}} & \text{if} \quad 0 \le i_{ph} < \dfrac{q_k}{t_1} - i_d \\ \dfrac{i_{ph}(t_{int}-t_1)}{\sqrt{q(i_{ph}+i_d)(t_{int}-t_1) + \sigma_r^2}} & \text{if} \quad \dfrac{q_k}{t_1} - i_d \le i_{ph} < \dfrac{q_{max}(1-\theta)}{t_{int}-t_1} - i_d \end{cases}$$

These formulae can be extended to more that two segments.

Multiple exposure approaches, either with several sets of pixels (based on [2]), several samples of the same pixel during exposure or several independent images can suffer from similar difficulties and similar artifacts. Sensors with multiple readout channels, commonly called scientific CMOS sensors, can also suffer from similar difficulties and similar artifacts but offer less degrees of freedom because the gains of the readout channels are usually fixed by design.

Managing the degrees of freedom

As we have seen, large dynamic range can only be properly acquired if the curve has limited compression and therefore if the dynamic range of the sensor is not too much extended or if a large number of kneepoints is used so that the ratio between each consecutive slopes is small. Using a large number of segments also means using a large number of control parameters (2*N+1) and therefore a large number of measured values.

A large number of control parameters can't easily be managed.

With only one slope (linear device), the black level is adjusted based on some reference black pixels so that the signal is dark enough but not fully clipped for image processing reasons. Then the brightness of the image or some other statistical parameters of the image are used to control the exposure time or the gain to find the best possible response for a given scene. Image average and image median are usually used to control exposure time, each has its pros and cons.

If there are more controls then more statistical values from the images are required in order to control the additional parameters. If these statistical values are somewhat linked to each other or if the change of one parameter affects more than one statistical value then it is extremely difficult to develop a stable multi-dimensional controller. Moreover, if the number of parameters to control is large, then we lack a number of statistical parameters to link the effect of a control on the captured image.

As a possible solution, the number of degrees of freedom is reduced to only 3: the offset controlling the black level, the total exposure time controlling the brightness and the compression controlling the dynamic range. It is reduced by forcing additional constrains to the system, for example a fixed ratio between two consecutive slopes and the same height between each kneepoint. This constrain yields a response curve similar to a logarithmic response sensor.

In our software IP, we are using median instead of mean because the median is not affected by changes of image saturation. Therefore we can use the median to control the total exposure time and the number of saturated pixels to control the compression ratio. As the parameters become uncorrelated, two independent closed loop (feedback) controllers with only one dimension work in parallel to control the sensor's response. A third controller acts on the black level based on reference black pixels. Additional rules related to the maximum acceptable exposure time balance exposure and gain. The exposure time may be limited to a maximum value in order to avoid excessive motion blur or dark current.

Going further, our dynamic range regulation also reduces the number of kneepoints to a minimum, based on a prediction of the possible presence of dynamic range gaps. Usually a response with more kneepoints is more noisy due to additional noise injection due to the change in control signals at each of the kneepoints.

Information transfer

If $f(I,\vec{P})$ is the sensor's response (i.e. the analog output (or the pixel charge in electrons) for a pixel irradiance I and an a set of control parameters and environment parameters $\vec{P}$), then a variation of luminance ΔL in the scene produces a variation of the pixel irradiance ΔI , a variation of photocurrent Δi_{ph} and a variation of the sensor's analog level at the input of the ADC given by $\Delta D = f(I+\Delta I,\vec{P}) - f(I,\vec{P})$. Strictly speaking, we should also consider the spectral response of the sensor and consider as input the integral over all wavelengths of the product of the pixel's irradiance and the sensor's spectral response. We will neglect this as well as the relationships between I , i_{ph} and L . $f(I,\vec{P})$ becomes a constant when the pixel is saturated.

We will no longer mention $\vec{P}$ in the next developments.

Going to small variations, the incremental gain, i.e. the variation of the pixel charge for a given variation in the scene is given by

$$g(I) = \frac{df}{dI} \quad . \tag{1}$$

If we consider that the information of a scene is related to the variations that can be seen, of any intensity and at any scale, then the incremental gain represents the information that is transferred from the scene into the image.

The following condition shall also be fulfilled in order avoid dynamic range gaps [3]:

$$SNR = \frac{I}{\sigma_I} = \frac{g(I).I}{\sigma_D} \geq 1 \quad . \tag{2}$$

Again another value than 1 can be used as the minimum acceptable SNR criteria.

As no information can be transferred in the presence of a dynamic range gap, the information transferred is proportional to

$$H(I) = \gamma(I).g(I) \quad , \tag{3}$$

where $\gamma(I)$ is the dynamic range gap presence function as previously defined.

When the derivative is high, there is a large variation in the image for a small variation in the scene and details remain highly visible. When the derivative is low, it is the opposite. If the image is saturated, or the irradiance on the sensor is below the sensitivity level or if the irradiance falls within an irradiance range that corresponds to a dynamic range gap, there is no information transferred.

Quantization

The image data is not the analog value at the input of the ADC but the digital value at the output of the ADC. The step in-between is called quantization. It is a rounding process in which an interval of analog values is represented by a single digital value. The quantization step should be small enough so that this process does not significantly affect the image data. It is usually such that the quantization noise (i.e. the rounding error) is less than the analog sensor noise. The worst case is close to dark where the sensor's noise is the smallest (noise increases with light intensity due to photon shot noise).

Advanced ADC approaches use variable step sizes for quantization, as described in [2], [4] and more recently in [7].

This rounding error introduces an additional loss of information and must be managed such that this loss is negligible. The process is well explained in [7].

Figures 5 and 6 show how the information retrieved from a given scene can vary depending on quantization step (figure 5) and response slope (figure 6).

Motivation for this work

If we use a mathematical function to represent the information contained in a scene and the information contained in the captured image, we can measure the loss of information. We can therefore define an information transfer ratio (or information gathering ratio), G, as the ratio between the image information and the scene information.

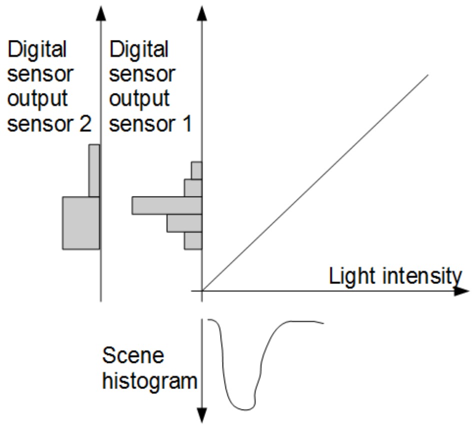

Figure 5. Information loss due to quantization – effect of quantization step

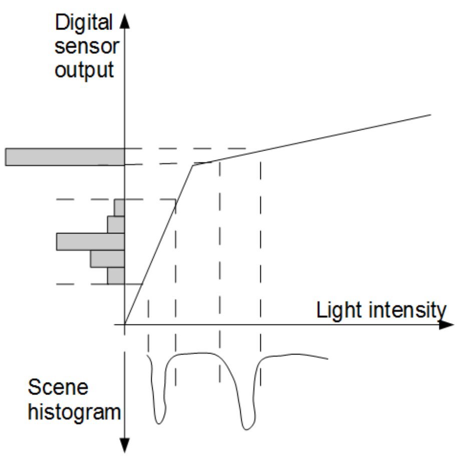

Figure 6. Information loss due to quantization – effect of response slope

For image processing applications, it would be interesting to tune the sensor's response for optimum information transfer in total so that the information contained in the resulting image is maximum. In this case, the ratios between the consecutive slopes and the distance between the kneepoints are no longer constant, like in most current implementations, and full freedom is used. Dynamic range gaps can even become acceptable if the gap causes a loss of information that is more than compensated by the gain in information somewhere else, for example if the gap occurs at gray levels that are almost not present in the scene but compensates by the enhancement of the details of a road sign, another vehicle or a pedestrian.

Status of this work

Electronic Imaging is a great place to come with new ideas, to explore the possibilities and to discuss the ideas with peers. This is why this new idea has been proposed as a poster for the image sensors and imaging systems session but is also of interest for the autonomous vehicles session.

This research is only at its beginning and i'm looking for interested students, partners or researchers to help develop it further.

The information transfer ratio

We will define the scene's information as the Shannon entropy and the probabilities are estimated based on the histogram. The scene's information is then

$$H_S = -\int_0^\infty p(I)\log p(I)\,dI \quad . \tag{4}$$

with $p(I)$ the probability that the irradiance is I .

Similarly, the image information is calculated as

$$H_I = -\sum_{i=0}^{Q-1} \gamma[I_i]\, p[f(I_i)] \log p[f(I_i)] \tag{5}$$

with

$$p[f(I_i)] = \frac{hist[f(I_i)]}{\sum_{k=0}^{Q-1} hist[f(I_k)]} \tag{6}$$

the proportion of pixels with the value $f(I_i)$. $f(I_i)$ is the central pixel value in each of the Q classes defined by the quantization of the sensor's ADC and $f(.)$ is the transfer function (or response) of the sensor. There is of course zero probability after sensor saturation and therefore the sum can be limited at the saturation. “hist” is a function representing the histogram values.

Brackets mean that we are in the discrete domain and parentheses mean the continuous domain.

For simplicity, we consider that $\gamma[I_i]$ is the most likely value of $\gamma(I)$ in the interval centered around I_i and that the quantization step is small enough so that using the center of each class does not affect the results. The more general solution is not different but only mathematically more complex.

Figure 7 illustrates the principle. Depending on the response curve of the sensor, some parts of the histogram are more or less compressed. In some cases the sensor saturates and the corresponding part of the scene's histogram degenerates into the saturation peak.

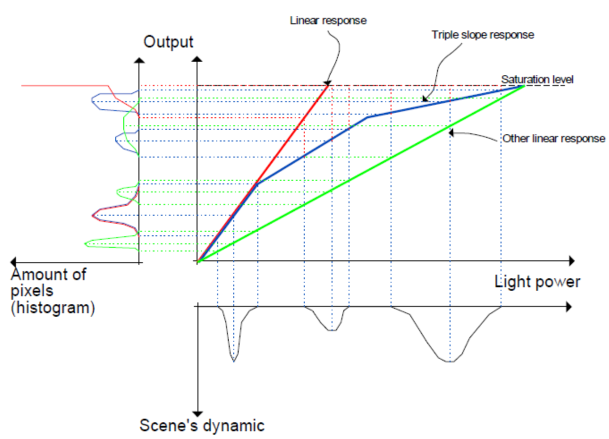

Figure 7. Transfer of a scene histogram into an image histogram for three image sensor responses – loss of information and compression ratios

The incremental gain of the OECF and the effect of quantization are hidden in the way the analog scene histogram becomes the discrete image histogram.

Then the information transfer ratio (or information gathering ratio) is defined as

$$G=\frac{H_I}{H_S} \quad . \tag{7}$$

Optimizing the sensor's response

The 2*N or 2*N+1 sensor parameters can then be selected to reach the highest possible value for G and therefore the information within each image data will be maximized. The optimization problem is to find the controllable values of $\vec{P}$ that maximize G for a given scene.

The maximum possible value of G is of course 1.

Future work

This paper sets the grounds of a new approach to optimize the use of HDR CMOS image sensors for machine vision applications.

There are significant mathematical developments still required in order to formalize the concept. Among these mathematical requirements, it needs to be demonstrated that the optimization problem has a solution (problem complexity analysis) and the mathematical expression of the optimization problem has to be derived.

Another challenge is to use the SNR curve as part of the information reduction problem as 1/SNR represents some form of probability that the information will be degraded through the imaging process. One possibility to be investigated is to to optimize for a combination of G and SNR with a weighting factor for each item.

The technique will only be useful if it is possible to estimate the scene's histogram, most likely based on previous images and detected scene changes or some form of modeling. Indeed if the scene remains an unknown (i.e. the irradiance on each pixel is unknown) then it is not possible to estimate the scene's histogram or information and therefore it is impossible to compute the set of parameters that maximizes the information transfer ratio. One possible direction towards this is to use multiple linear response images to estimate the scene's histogram and then define the best merging and compression approach and to repeat this process each time the scene seems to change significantly. This is not a real time solution that can be used to control a sensor but a possible intermediate solution for multiple exposure and multiple sampling systems.

At first, the problem should be mathematically formalized for an arbitrary scene histogram, a flat scene histogram and a trapezoidal scene histogram.

It should also be investigated how this can be related to minimum distinguishable contrast.

References

1. D. Hertel, A. Betts, R. Hicks, M. ten Brinke, “An adaptive multiple-reset CMOS wide dynamic range imager for automotive vision applications”, IEEE Xplore Conference: Intelligent Vehicles Symposium, IEEE, July 2008.
2. A. Darmont, "Methods to extend the dynamic range of snapshot active pixel sensors" , Proc SPIE 6816, Sensors, Cameras and Systems for Industrial/Scientific Applications IX, 681603, p11, 29th February 2008.
3. D. Hertel, “Extended use of ISO15739 incremental signal-to-noise ratio as reliability criterion for multiple-slope wide dynamic range image capture”, Journal of Electronic Imaging, January 2010.
4. A. Darmont, “High Dynamic Range Imaging: Sensors and Architectures”, SPIE Press, 2012, ISBN 9780819488305.
5. D. X. D. Yang, A. El Gamal, “Comparative analysis of SNR for image sensors with enhanced dynamic range”, Proc SPIE 3649, 197 (1999).
6. B. Fowler, “High dynamic range image sensor architectures”, Stanford University WDR workshop, September 2009.
7. B. Jähne, M. Schwarzbauer, “Noise equalization and quasi loss-less image data compression – or how many bits needs an image sensor”, Technisches Messen 2016; 83(1): 16-24.
8. A. Darmont, “Automotive CMOS image sensors”, AMAA 2005.

Author Biography

Arnaud Darmont received his master in electronics engineering from the University of Liège (Belgium) in 2002 with interests in telecomunications, solid state physics, imaging and image processing. For the next 7 years he has been involved in automotive HDR image sensor design and management at Melexis (Belgium) in collaboration with FillFactory and Awaiba, as application engineer, design engineer and later project manager. For the last months with Melexis, he worked at former SmaL (USA), then just acquired by Melexis. In 2009 he founded Aphesa, a startup company specialized in unusual camera designs, image sensor and imaging technology consulting, imaging technology training activities as well as image sensor and camera characterization. He is one of the developers of the EMVA1288 standard since 2005 and committee member of the image sensors and imaging systems conference of Electronic Imaging since 2013. He is a chair of the conference since 2015. Since December 2017 he is also the manager of vision standards at the European Machine Vision Association. He is the author of several publications, courses and books about HDR and imaging technology in general, main inventor of one patent and contributor to several others.