Using Brightness Histogram to perform Optimum Auto Exposure

SIMON SCHULZ
Hamburg University of Technology
Vision Systems
Harburger Schlossstrasse 20
21079 Hamburg, GERMANY
simon.schulz@tu-harburg.de

MARCUS GRIMM
Hamburg University of Technology
Vision Systems
Harburger Schlossstrasse 20
21079 Hamburg, GERMANY
marcus.grimm@tu-harburg.de

ROLF-RAINER GRIGAT
Hamburg University of Technology
Vision Systems
Harburger Schlossstrasse 20
21079 Hamburg, GERMANY
grigat@tu-harburg.de

Abstract: The dynamic range of natural scenes usually exceeds the dynamic range of imaging sensors by several orders of magnitude. An automatic exposure control (AE) is therefore commonly used in order to shift the dynamic range of the sensor to the section of interest of the dynamic range of the scene. We present a new approach for choosing the section by defining a criteria of optimality. The presented exposure control algorithm is optimal in terms of recorded information of the scene. The advantage over AE based on image mean values will be shown and evaluated by simulation and experiments. We show that our method can easyly be extended for AE control of multiple-slope cameras. The recorded information of the scene is maximized by controlling exposure as well as the transition curve parameters. The controllable dynamic range of mutliple-slope cameras as well as corse quantization in bright image regions is considered. Finally we propose some possible extensions to the criteria of optimality.

Key–Words: Optimum Auto Exposure, High Dynamic Range (HDR), multiple-slope camera
January 26, 2007

1 Introduction

Luminous intensities vary from $0.001\frac{\text{cd}}{\text{m}^2}$ (star light) to up to $100,000\frac{\text{cd}}{\text{m}^2}$ (sunlight) such that the illumination in natural scenes reach a dynamic range d_b of $d_\text{b} = \frac{100,000}{0.001} = 100,000,000$ or $20\lg(d_\text{b}) = 160$ dB.

The dynamic range d_cam of typical digital camera lies in the range of 55 dB. d_cam is limited by the signal independent noise level c_min of the camera on the lower end, and saturation capacity c_max of the pixel on the upper end. The dynamic range d_cam of a camera is always much smaller than the range d_b of natural illumination levels.

Therefore the camera must be automatically adapted to the surrounding illumination level, just as the human eye is adapted to surrounding illumination level by the iris (fast) or the adaptation of the cones and rods (slower). This process is called auto exposure (AE) or auto gain, depending on the parameter that is changed in the camera in order to adapt it.

Additionally the dynamic range of a camera is typically even smaller then the dynamic range within one scene. AE is therefore always a trade-off. On the one hand the signal might sink below the noise level c_min resulting in underexposure. On the other hand the pixel value might be clipping due to limited saturation capacity c_max of the pixels resulting in overexposure.

In this paper we present an optimal auto exposure algorithm in terms of recorded information. Information from the scene gets lost whenever either under- or overexposure occurs. Minimizing both via auto exposure is therefore considered to be optimal.

So called multiple-slope cameras additionally feature a quasi piecewise linear transition curve with controllable kneepoints. The dynamic range of these cameras may be increased by using this feature sacrificing fine quantization in bright parts of the image at the same time. The camera cannot be considered linear anymore, traditional auto exposure algorithms fail. We present a new approach for optimum auto exposure with kneepoint control. We take under- and overexposure, variable dynamic range as well as raised quantisation noise in bright image parts into account.

2 Auto Exposure for linear camera

Cameras not operated in multiple-slope mode can be considered to be linear: the input brightness b is propotional to the output value c of the pixel.

$$c = \alpha b \tag{1}$$

WSEAS Transactions on System and Control, Volume 2, Issue 2, February 2007 (ISSN: 1991-8763)

The auto exposure for a linear camera modifies the slope α of the transition curve in order to match the brightness levels within the scene. In the case of traditional AE algorithms α is adjusted according to the mean brightness of the scene.

Modifying α is equal to shifting the dynamic range d_{cam} over the high dynamic range histogram $p(b)$ (HDH) of the brightness levels (see fig. 1), if we plot the HDH in a logarithmic scale for b,

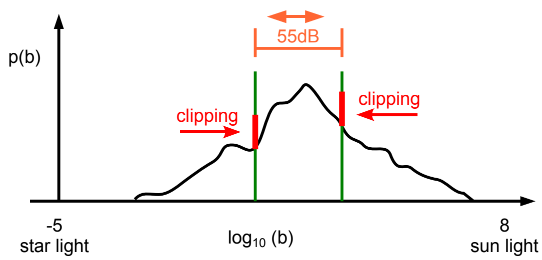


Figure 1: High dynamic range histogram $p(b)$ of the scene

The dynamic range d_{cam} of the camera is not depending on the slope α of the transition curve: Defining b_{min} and b_{max} as the values where the camera ouput reaches c_{min} and c_{max} respectively for a fixed α we get:

$$d_{\text{cam}} = \frac{c_{\text{max}}}{c_{\text{min}}} = \frac{\alpha b_{\text{max}}}{\alpha b_{\text{min}}} = \frac{b_{\text{max}}}{b_{\text{min}}} \tag{2}$$

Taking the logarithm of b we get an interval of constant width which is shifted over the logarithmic HDH by changing α:

$$\begin{aligned} \log(d_{\text{cam}}) &= \log(\frac{c_{\text{max}}}{c_{\text{min}}}) = \log(c_{\text{max}}) - \log(c_{\text{min}}) \\ &= \log(\alpha b_{\text{max}}) - \log(\alpha b_{\text{min}}) \end{aligned} \tag{3}$$

In order to choose an optimal α in terms of maximum recorded information we must minimize under- and overexposure. This is identical to maximizing the integral over the HDH within the bounds of the dynamic range d_{cam} of the camera. The integration bounds can be expressed as functions of the slope α, thus resulting in an optimal slope: $b_{\text{min}} = \frac{b_{\text{max}}}{d_{\text{cam}}} = \alpha\frac{c_{\text{max}}}{d_{\text{cam}}}$, $b_{\text{max}} = \alpha c_{\text{max}}$:

$$\max_{b_{\text{max}}} \int_{\frac{b_{\text{max}}}{d_{\text{cam}}}}^{b_{\text{max}}} p(b)db = \max_{\alpha} \int_{\alpha\frac{c_{\text{max}}}{d_{\text{cam}}}}^{\alpha c_{\text{max}}} p(b)db \tag{4}$$

3 Auto Exposure for multiple-slope camera

When digital cameras are used, that feature a multiple-slope transition curve, the dynamic range d_{cam} of the camera may be modified. The dynamic range d_{cam} depends on the slopes of the camera. Considering a camera with a single controlable kneepoint in the transition curve (two linear segments), the kneepoint can be parameterized by a slope scaling factor $s \in [1, \infty[$ und a height $\beta \in [0, 1]$.

The factor s is the ratio between the slope α_a of the first linear segment and the slope α_b of the second one: $\alpha_b = \frac{\alpha_a}{s}$. Due to the technical realization of the multiple-slope feature in CMOS sensors, the slope of the transition curve can only be lowered at a knee point. This leeds to the lower bound for s: $s \geq 1$.

The height β indicates the sensor output level c_{Level} at which the transition curve switches from one slope to the other: $c_{\text{Level}} = \beta c_{\text{max}}$.

Only the second part of equation (2) still holds in the case of multiple-slope camera mode:

$$d_{\text{cam,slp}} = \frac{b_{\text{max}}}{b_{\text{min}}} \neq \frac{c_{\text{max}}}{c_{\text{min}}} = d_{\text{cam,lin}} \tag{5}$$

We assume that the minimum input brightness $b_{\text{min,slp}}$ falls into the first linear segment of the transition curve: $b_{\text{min,slp}} = \frac{c_{\text{min}}}{\alpha_a}$ (see equation (1)). The maximum input brightness $b_{\text{max,slp}}$ cannot be calculated according to equation (1), as the camera is not linear any more. $b_{\text{max,slp}}$ can be calculated as

$$\begin{aligned} b_{\text{max,slp}} &= \frac{c_{\text{Level}}}{\alpha_a} + \frac{c_{\text{max}} - c_{\text{Level}}}{\alpha_b} \\ &= \frac{s c_{\text{max}} + (1-s) c_{\text{Level}}}{\alpha_a} \end{aligned} \tag{6}$$

The dynamic range of the multiple-slope camera can now be given as a function of s and β:

$$\begin{aligned} d_{\text{cam,slp}}(s, \beta) &= \frac{b_{\text{max}}}{b_{\text{min}}} = \frac{s\, c_{\text{max}} + (1-s)\, c_{\text{Level}}}{c_{\text{min}}} \\ &= [(1-\beta)s + \beta]\, \frac{c_{\text{max}}}{c_{\text{min}}} \end{aligned} \tag{7}$$

The minimum dynamic range range is reached with a linear slope ($s = 1$).

According to equation (4), the most simple solution to maximize the recorded information is to operate the camera with the maximum dynamic range possible. This leeds to the maximation of the interval width for the integration und thus to the maximation the integral. Due to changes in quantization, this is *not* the optimal AE for cameras operated in multiple-slope mode.

The quantization Δb of the input brightness b raises with lower slope α for linear cameras. It is constant across the whole dynamic range of the camera.

Using a sensor with n bit output, we get 2^n quantization levels, thus

$$\Delta b = \frac{b_{\text{max}}}{2^n} = \frac{1}{\alpha}\frac{c_{\text{max}}}{2^n} \tag{8}$$

For a multiple-slope camera the quantization Δb depends on the segment of the transition curve. The quantization Δb_a in the first segment is equal to the quantization of the linear camera: $\Delta b_a = \frac{1}{\alpha_a}\frac{c_{\text{max}}}{2^n}$. The quantization Δb_b in the second segment shows a similar behaviour:

$$\Delta b_b = \frac{1}{\alpha_b}\frac{c_{\text{max}}}{2^n} = \frac{s}{\alpha_a}\frac{c_{\text{max}}}{2^n} = s\,\Delta b_a \tag{9}$$

The quantization step (and quantization noise at the same time) raises with the ratio s of the slopes. This effect has to be taken into account in equation (4) for the optimum AE control. Let $b_{\text{Level}} = \frac{\beta}{\alpha_a}c_{\text{max}}$ be the brightess level, at which the recording switches from the first to the second slope segment for a specific camera setting. We can pose the optimum AE for multiple-slope cameras as a modified constraint optimization problem:

$$\begin{aligned} &\max_{b_{\text{max}},s,b_{\text{Level}}} \left(\int_{b_{\text{min}}}^{b_{\text{Level}}} p(b)db + s^{-m} \int_{b_{\text{Level}}}^{b_{\text{max}}} p(b)db \right) \\ &= \max_{\alpha_a,s,\beta} \left(\int_{\frac{c_{\text{min}}}{\alpha_a}}^{b_{\text{Level}}} p(b)db + s^{-m} \int_{b_{\text{Level}}}^{b_{\text{max}}} p(b)db \right) \end{aligned} \tag{10}$$

$$\text{with} \qquad b_{\text{max}} = \frac{s(1-\beta)c_{\text{max}} + \beta c_{\text{max}}}{\alpha_a}, \quad m \in [0,1],$$

with the constraints

$$\alpha_a > 0 \tag{11}$$
$$s > 1 \tag{12}$$
$$\beta > \frac{c_{\text{min}}}{c_{\text{max}}} \tag{13}$$
$$\beta \leq 1 \tag{14}$$

The modified quantization in the second segment is taken into account by the weighting factor $m \in [0,1]$. The constraint (13) forces the height β of the kneepoint to be greater than the noise level.

4 Generation of the HDH

The camera itself is used to generate the HDH. Taking the histogram of a single image as the HDH restricts the dynamic range of the histogram to the dynamic range of the camera. We combine several images of the same scene with different exposures in order to extend the dynamic range of the HDH. This leeds to a HDH with a dynamic range that covers the entire dynamic range, the camera is able to record using all different exposure settings. This is achieved by including the minimum and maximum camera exposure into the exposure series.

We consider the camera to be linear. This holds true only for output values of the camera which are neither dominated by noise ($c < c_{\text{min}}$) nor by saturation effects ($c > \tilde{c}_{\text{max}} > c_{\text{max}}$). The upper bound $\tilde{c}_{\text{max}}$ for the linear operation is lower then the saturation capacity c_{max} because saturation effects are already apparent before the maximum value c_{max} is reached [1], [2].

Defining the co-domain for equation (1) leeds to

$$c = \alpha b\,,\ \forall c \in [c_{\text{min}}, \tilde{c}_{\text{max}}] \tag{15}$$

We only want to combine the linear ranges of different exposures ($c_{\text{min}} < c < \tilde{c}_{\text{max}}$). This means that we have to choose the subsequent exposures such that the linear ranges touch or even overlap each other. Starting from the maximum exposure T_0 we would like to calculate the next exposure T_1 which exactly touches the linear range:

$$\begin{aligned} b &= \frac{\tilde{c}_{\text{max}}}{\alpha_{T_0}} = \frac{c_{\text{min}}}{\alpha_{T_1}} = \frac{c_{\text{min}}}{\frac{T_1}{T_0}\alpha_{T_0}} \\ T_1 &= T_0 \frac{c_{\text{min}}}{\tilde{c}_{\text{max}}} \end{aligned} \tag{16}$$

As an example we consider a VGA-resolution digital camera with rolling shutter. The exposure T may be varied over a range of 1 to $T_0 = 480$ lines. Thus the slope α of the linear camera may be modified from α_{480} to $\alpha_1 = \frac{1}{480}\alpha_{480}$. The normalized camera output values are supposed to be in the range $[0,1]$ with added signal independent noise of 1%, resulting in $c_{\text{min}} = 0.01$. The saturation effects are assumed to begin at 95%, thus $c_{\text{max}} = 0.95$. According to equation (16) T_1 should be set to at least $T_1 = \left\lfloor 480\frac{0.01}{0.95} \right\rfloor = 6$. This means that a maximum of three images is sufficiant to cover the whole dynamic range of the camera in this example.

5 Results

In order to evaluate the performance of the newly proposed AE algorithm we used simulated as well as real camera recordings. Experiments show that mean value based AE performs well for low dynamic range scenes like indoor lab scenes. The results of mean value based AE and our new AE algorithm based on HDH are comparable for such kind of scenes.

Especially in outdoor real world scenes the dynamic d_{sce} range of the scene usually exceeds the dynamic range d_{cam} of the camera by orders of magnitude. This is where our newly proposed algorithm is supposed to show its strength compared to mean value based approaches.

In order to better visualize the recorded information amount, we used a local adaptive tone mapping technique to visualize details in bright and dark image reagions at the same time.

5.1 Simulation

For simulation purpose we used several images in the OpenEXR [8] image format, an open standard HDR image format. In this case the generation of the HDH is simple as it is equivalent to calculating the histogram of the OpenEXR image. We use three simulated recordings to determine the overall dynamic range that may be captured by the camera using all available exposure settings. The histogram is of the OpenEXR image is then restricted to this overall dynamic range. Figure 2 shows the three simulated recordings of the ”belgium”-scene [9] to obtain the HDH for that scene.

(a) tone mapped HDR image

(b) long exposure

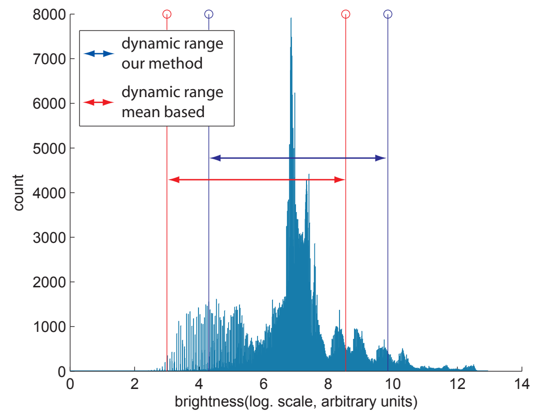
(c) medium exposure

(d) short exposure

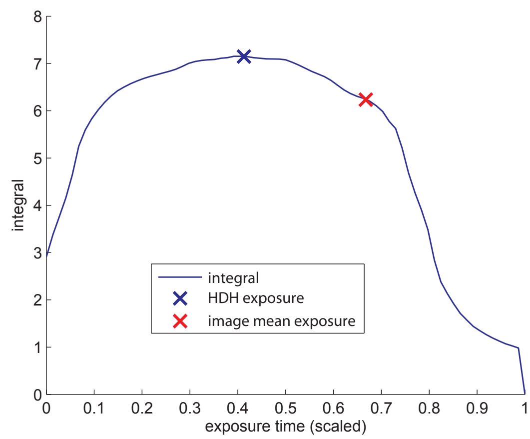
Figure 2: Simulated recordings of OpenEXR image ”belgium” [9], (a) illustrates the HDR image content using tone mapping techniques

The high dynamic histogram of the ”belgium” scene as well as the integral value according to equation (4) is shown in figure 3. We see that the dynamic range d_{cam} of the camera is indeed much smaller than the dynamic range d_{sce} of the scene. The integral value of equation (4) can be interpreted as the number of well exposed pixels in the image of the scene as a function of exposure time.

For the mean value AE we choose a target mean value of $c_{\text{target}} = 0.35c_{\text{max}}$, which results in a well exposed to slightly underexposed image for indoor scenes.



(a) Histogram



(b) Integral according to eqn. (4)

Figure 3: Histogram and integral of the ”belgium” scene [9]: mean value AE results in overexposure

We simulated the recording for both mean value AE and HDH AE considering additive gaussian recording noise with $3\sigma = 0.01c_{\text{max}}$. This is espacially important because we use a local adaptive tone mapping technique in order to make details in dark image reagions visible. The performance of the tone mapping technique used is limited by recording noise and quantization.

Even though the target mean value c_{target} was set to a rather low value the resulting image of the mean value AE suffers from massive overexposure (figure 4).

Figure 5 gives two more examples that illustrate

(a) mean value AE

(b) HDH AE

Figure 4: Comparison of mean value AE and HDH AE for "belgium"-scene: massive overexposure using mean value AE

(a) mean value AE (b) HDH AE

(c) mean value AE (d) HDH AE

Figure 5: Simulated recording using mean value AE and HDH AE of the "memorial" [10] and the "spheronsig" [11] scenes: overexposure and loss of bright details using mean value AE, (d) noise limitation of local adaptive tone mapping more obvious using HDH AE

the advantage of the HDH AE method.

5.2 Real world tests

We also tested our new exposure control method for a linear camera using a high dynamic range lab scene and real camera recordings. The results were compared to the performance of a mean value based AE. The scene content is visualized in the two images of figure 6.

The high dynamic range histogram (HDH, fig. 7) is calculated from different recordings of that scene. Note that the x axis represents absolute brightness levels in logarithmic scale rather then relative luminances. The brightness levels in the HDH are given in arbitrary units, as brightness levels are to be known only up to a constant factor for AE.

In figure 7, we added the dynamic range of the camera for two different exposure levels (see eq. (3)). This illustrates that the dynamic range d_{sce} of the scene is greater than the dynamic range d_{cam} of the camera. Therefore choosing an exposure for recording this scene is always a trade-off.

Given the HDH we can calculate the integral according to equation (4). In figure 8 the value of the integral is plotted as a function of the exposure time T.

The maximum of the integral is reached at an exposure time of $T_{\text{HDH}} \approx 1.8$ms. For $T = T_{\text{HDH}}$ we solved the optimization problem (4). This means that a maximum number N_{max} of pixels of the image is well exposed when the exposure time is set to $T_{\text{HDH}} \approx 1.8$ms for this scene.

A simple mean value based AE control tries to modify α such that the mean value $\bar{c}$ of the image is equal to about half the maximum output amplitude $\bar{c} \overset{!}{=} c_{\text{mean}} = \frac{c_{\text{max}}}{2}$. This mean based AE leads to an exposure time of about $T_{\text{mean}} \approx 1.05$ms. From figure

(a) over exposed image to render dark regions (b) under exposed image to render bright regions

Figure 6: Images from a high dynamic range lab scene

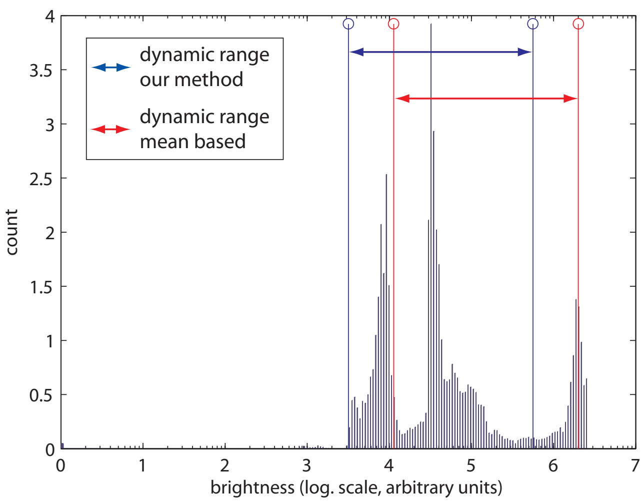


Figure 7: High dynamic range histogram (HDH) of the scene from figure 6 generated from multiple exposures

8, we see that only 77% of the maximum number of well exposed pixels are well exposed choosing exposure time T_{mean}. AE control based on the mean value tries to balance over- and underexposure whereas our method maximizes well exposed pixels. The approach of balancing over- and underexposure is not optimal for high dynamic range scenes as we loose details both in bright and in dark image regions. We could easily reach $T = T_{\text{HDH}}$ with a mean value based AE control by either adapting the target mean value c_{mean} or modifying the mean value calculation ([3], [4], [5], [6], [7]) to scene specific optimum. However adapting the AE method to every single scene corresponds to the removal of the "auto" term from "auto exposure control".

The different shifts of the dynamic range of the camera in the HDH are also displayed in figure 7: the red interval shows the captured dynamic range using AE control based on mean value. The blue interval shows the captured dynamic range using our AE control method based on the HDH. The dynamic ranges themselves are equivalent as we are using a linear

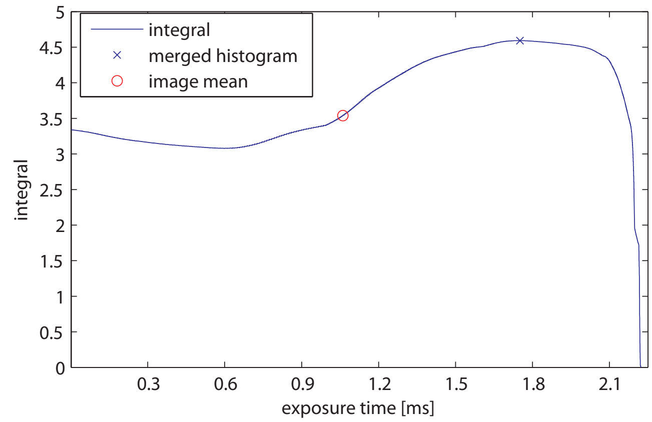


Figure 8: Integral over the HDH with varying exposure time, results of the proposed AE method (blue cross) compared to mean value based AE (red circle)

camera (eq. (2)).

We use a local adaptive technique to map the camera output values c to display values in order to visualize camera images. The results are shown in figure 9, 10 and 11.

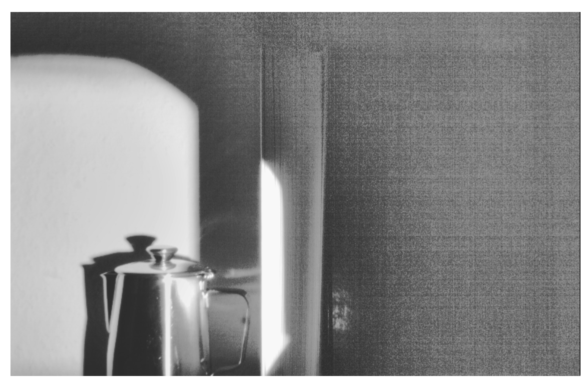
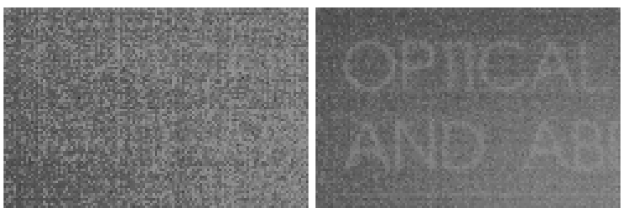


Figure 9: Image captured using mean value based AE, all details in dark regions lost (local adaptive technique used to visualize high dynamic range image)

6 Conclusion and future work

We propose a new approach to auto exposure (AE) for digital cameras. The exposure chosen by our method is optimal in terms of recorded information. The camera is controlled such that a maximum number of pixels of the image are well exposed. This results in more recorded information than the over- and underexposed balancing approach of mean value based AE control methods. A mean value based AE control might reach the same amount of recorded information but needs manual adaptation to a specific scene.

Figure 10: Image captured using the proposed AE control based on HDH, details in dark regions partly visible (local adaptive technique used to visualize high dynamic range image)



(a) mean based AE (b) HDH based AE

Figure 11: Comparison of information contend in crops of images from mean value based AE and HDH based AE

Experimental results show that we can raise the number of well exposed pixels by 28% compared to a simple mean value based AE control for a high dynamic range scene. Futhermore we can avoid the trade-off between control speed and oscillation from a AE control loop, as the optimum exposure is directly calculated from the HDH in one step.

An AE and slope control for multiple-slope cameras is proposed based on the HDH of the scene. The possibility of enlarging the dynamic range as well as the lost of quantization accuracy is considered in this new approach.

Future work will comprise the realization of the control system for multiple-slope cameras as described in this paper as well as extensions to the measure of recorded information. Currently the number of well exposed pixels is counted and used as measure of recorded information. The information measure will be extended to take local information into account e.g. using local variances.

References:

[1] F. Dierks, Sensitivity and Image Quality of Digital Cameras *Technical Report*, Basler AG, 2004.

[2] EMVA Standard 1288, Standard for Measurement and Presentation of Specifications for Machine Vision Sensors and Cameras, www.emva.org.

[3] S. Shimizu, T. Kondo, T. Kohashi, M. Tsuruta and T. Komuro, A new Algorithm for Exposure Control based on Fuzzy Logic for Video Cameras, *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 3, 1992, pp. 617–624.

[4] T. Haruki, K. Kikuchi, Video camera system using fuzzy logic, *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 3, 1992, pp. 624–634.

[5] T. Kuno, H. Sugiura and N. Matoba, A new Automatic Exposure System for Digital Still Cameras, *IEEE Transactions on Consumer Electronics*, Vol. 44, No. 1, 1998, pp. 192–199.

[6] J.-S. Lee, Y.-Y. Jung, B.-S. Kim and S.-J. Ko, An advanced video camera system with robust AF, AE, and AWB control, *IEEE Transactions on Consumer Electronics*, Vol. 47, No. 3, 2001, pp. 694–699.

[7] M. Cho, B.-D. Nam, B.-D. Nam, The fast Auto Exposure Algorithm based on Numerical Analysis, *Conf. on Sens., Cameras and Appl. for Digital Photography, Proc. of the SPIE* , Vol. 2650, 1999, pp. 93–99.

[8] Specifications of the OpenEXR HDR file format, *http://www.openexr.com/*

[9] HDR image courtesy D. Lischinski, *http://www.cs.huji.ac.il/ danix/*

[10] HDR image courtesy P. Debevec, *http://www.debevec.org/*

[11] HDR image courtesy Spheron, *http://www.spheron.com/*

(a) mean value AE

(b) HDH AE

(c) mean value AE

(d) HDH AE

Figure 12: Enlarged prints of simulated recording using mean value AE and HDH AE