Extending Image Sensor Dynamic Range by Scene-aware Pixelwise-adaptive Coded Exposure

Huifeng Ke[1], Navid Sarhangnejad[1], Rahul Gulve[1], Zhengfan Xia[1], Nikita Gusev[1], Nikola Katic[1], Kiriakos N. Kutulakos[2], Roman Genov[1]

***Abstract*—We present a method for extending the dynamic range of an image sensor by frame-to-frame adaptive adjustment of exposure of every pixel, based on changes in the brightness of the scene for that pixel. The method employs an image sensor with a coded-exposure-pixel (CEP) [1], where the exposure time of each frame is divided into N subframes, each of unary or binary-weighted duration. In each subframe the pixel is re-programmed with a 1-bit binary code to be either on (code 1) or off (code 0), depending on the light flux captured by that pixel in the previous frame. The photo-generated charge is integrated across all subframes. The N 1-bit codes per pixel are computed by an off-sensor processor, and are loaded to a digital latch within that pixel, one per subframe. At the end of a frame, a high-dynamic-range (HDR) image is reconstructed by using that frame's pixel codes to normalize the pixel digital outputs to a uniform exposure. The dynamic range is experimentally demonstrated to increase by up to $20log_{10}(2^N - 1)\ dB$ at the full frame rate and at the native resolution of the image sensor.**

I. Introduction

The limited dynamic range of most conventional active pixels allows for capturing bright or dark scenes but not both simultaneously. Several high-dynamic-range (HDR) techniques have been introduced to eliminate this trade-off. Native-resolution full-frame-rate HDR methods add significant auxiliary circuits overhead to the pixel such as multi-gain pixel readout [2], event-based readout [3] or range compressing transfer functions (e.g., logarithmic) [4], but often suffer from a low fill factor and in-pixel analog circuit non-idealities. In cases where the frame rate can be traded for the dynamic range, exposure is often varied over multiple frames [5][6]. This common approach in conventional cameras requires a proportionally higher frame rate or leads to motion blurs or ghosting. A single-frame HDR method utilizing 2x2 tiles of single-tap coded-exposure pixels, each programmed with a different exposure time (Fig. 1(a)), has been reported in [7][8]. The downside of this method is that it yields a 4x lower image resolution.

All authors are with the University of Toronto, 10 Kings College Road, Toronto, Ontario M5S 3G4, Canada.

[1]H. Ke, N. Sarhangnejad, R. Gulve, Z. Xia, N. Gusev, N. Katic and R. Genov are with the Department of Electrical and Computer Engineering (e-mail: jackie.ke@mail.utoronto.ca, [sarhangn, rahulgulve, xia]@ece.utoronto.ca, roman@eecg.utoronto.ca).

[2]K. Kutulakos is with the Department of Computer Science (e-mail: kyros@cs.toronto.edu).

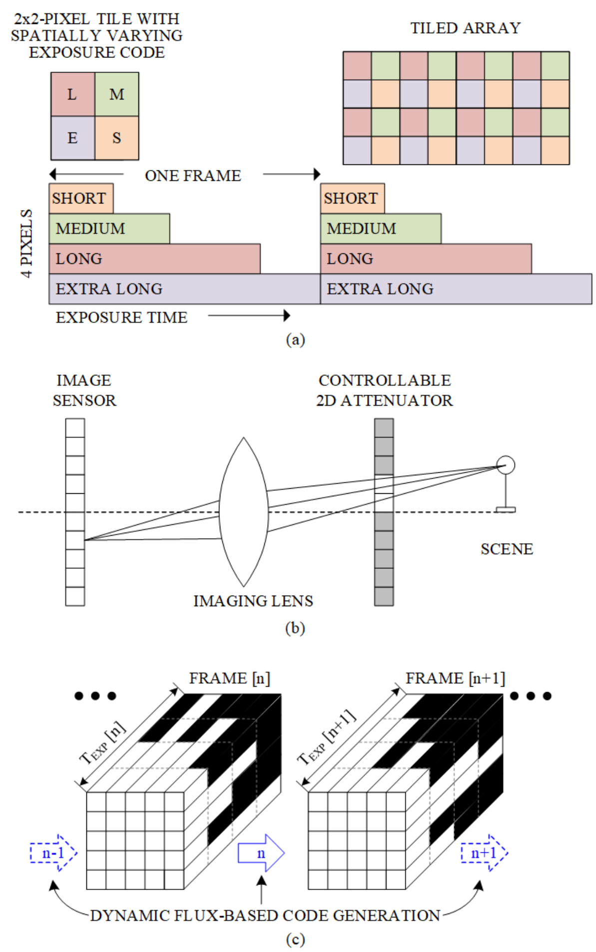

Fig. 1: Existing methods for dynamic range extension: (a) spatially varied exposure where 4 images of different exposures with 1/4th of the original spatial resolution are obtained and utilized to synthesize HDR images [7][8], and (b) a controllable attenuator such as a spatial light modulator (SLM) that controls per-pixel signal attenuation [9]. (c) In the presented method, each pixel's exposure time is dynamically reprogrammed based on the light flux received by the pixel in the previous frame.

Another single-frame HDR method, presented in [9], uses spatial light modulator (SLM) in front of the camera (Fig. 1(b)), but this method is not scalable to high image sensor spatial resolution, in addition to high complexity and cost of the camera module.

This work demonstrates a technique for dynamic range extension of an image sensor at its native resolution using a coded-exposure-pixel (CEP) pixel architecture [1]. A CEP image sensor allows for arbitrary programmable

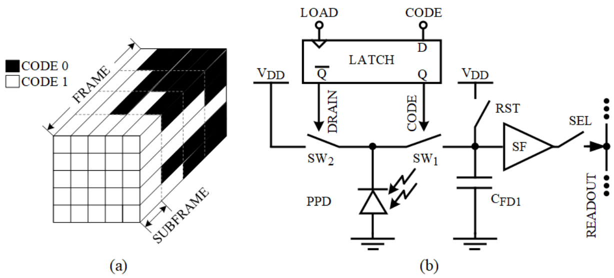

Fig. 2: (a) Per-pixel coded operation per frame, and (b) 1-tap coded exposure pixel schematics.

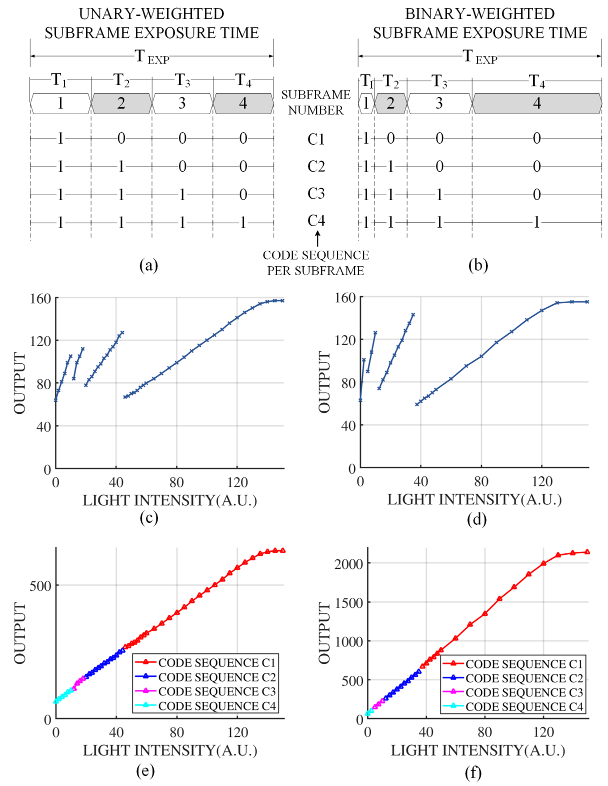

Fig. 3: Experimental results for two coding schemes. Code sequence and timing used for unary-weighted subframe exposure time (a), and for binary-weighted subframe exposure time (b). Raw output of a pixel for unary-weighted subframe exposure time (c), and for binary-weighted subframe exposure time (d). Normalized output of a pixel for unary-weighted subframe exposure time (e), and for binary-weighted subframe exposure time (f).

control over the exposure time of individual pixels. The proposed technique analyzes light flux in each pixel in a captured frame and makes dynamic adjustments to the respective exposure times in the next frame as depicted in Fig. 1(c). For example, pixels with low brightness in the captured frame are programmed to have higher exposure time for next frame. After normalization for different brightness in the captured image, an HDR image is optimally reconstructed and visualized. In the following, we discuss in detail the HDR imaging technique, the testing system, and the experimental results.

II. Principle of Extending Dynamic Range Using Per-Pixel Coded Exposure

The CEP custom image sensor employed in this work [1] implements imaging by dividing the exposure time of one frame into multiple subframes as shown in Fig. 2(a). As illustrated in Fig. 2(b), in each subframe, every pixel receives a single bit of the exposure code that directs the photo-generated charge to the collection node C_{FD1} for code 1, or drains them through V_{DD} for code 0.

The exposure code delivered to a pixel in the current frame depends on the amount of photo-generated charge collected in the previous frame. In the first subframe, all pixels receive code 1 to collect the initial photo-generated charge. In the subsequent subframes, pixels in darker areas of the image will receive more 1's to integrate the newly generated charge, but the pixels in brighter areas will receive more 0's to drain the newly generated charge.

The sequence of receiving 1's and 0's is defined as the exposure code sequence. Figs. 3(a) and (b) show the exposure time settings and the code sequences, in a 4-subframe setup (i.e, $N = 4$), for two subframe exposure coding schemes. For unary-weighted subframe exposure time, as shown in Fig. 3(a), the exposure time of one frame T_{EXP} is divided equally into four subframes, expressed as $T_1 = T_2 = T_3 = T_4 = T_{EXP}/4$. In contrast, for binary-weighted subframe exposure time, as shown in Fig. 3(b), the exposure time of each subframe is two times longer than the previous subframe, expressed as $8T_1 = 4T_2 = 2T_3 = T_4 = 8T_{EXP}/15$. The cumulative exposure time of all subframes gives the effective exposure time of a pixel. For example, the effective exposure time using $C3$ is $3T_{EXP}/4$ for unary-weighted subframe exposure, and $7T_{EXP}/15$ for binary-weighted subframe exposure.

Figs. 3(c) and (d) show the raw outputs of a pixel over a full range of illumination for both coding schemes. Both plots comprise 4 line-segments. Each linear segment corresponds to a different code sequence, from $C4$ to $C1$ as exposure increases. When the scene brightness at the pixel starts off low, C4 is used, because high effective exposure time allows to capture more details of the scene. When the brightness increases, code sequence changes to C1 scene, to avoid saturating the pixel.

Figs. 3(e) and (f) are normalized versions of (c) and (d), respectively. Normalization converts the piece-wise linear outputs to the full-range linear response using the corresponding effective exposure time of each segment. After normalization, the maximum effective dynamic range of a pixel is improved. With only 4 subframes as in this example, it is increased by 4 times for unary-weighted subframe exposure time, and by 15 times for binary-weighted subframe exposure time.

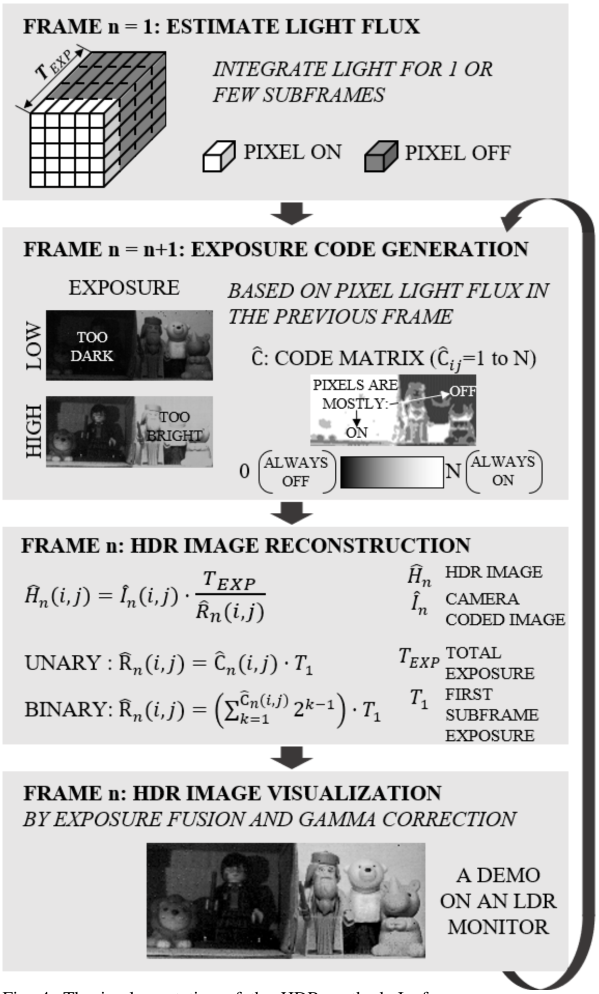

Fig. 4: The implementation of the HDR method: In frame n, one or several subframes integrate light. Based on the output, a new exposure code is generated and uploaded for frame $n + 1$. Concurrently, an HDR image is obtained by normalizing the output of frame n. In this 4-subframe implementation, normalization generates a 10-bit HDR image using an 8-bit sensor output and a 2-bit exposure code. Finally, an exposure fusion technique [10] visualizes the 10-bit HDR image on an LDR display.

III. Dynamic Per-Pixel Code Generation

The extended HDR technique (Fig. 4) is implemented in a pipelined fashion. When the current frame is read out for displaying, it is also used to generate the exposure codes for the next frame. The technique is valid for both unary-weighted and binary-weighted subframe exposure.

The definition of exposure code $\widehat{C}_n(i,j)$ is that pixel (i,j) will integrate charges from subframe number 1 to subframe number $\widehat{C}_n(i,j)$. $\widehat{C}_n(i,j)$ is a positive integer with $1 \leq \widehat{C}_n(i,j) \leq N$, and N is the number of subframes in one frame. After uploading $\widehat{C}_n$ to the CEP image sensor [1] for frame n, the captured image is defined as $\widehat{I}_n$.

The normalization matrix $\widehat{R}_n$ comprises the effective exposure time of all pixels. Equations (1a) and (1b) show the mathematical definitions of $\widehat{R}_n(i,j)$ for unary-weighted and binary-weighted subframe exposure.

$$\text{Unary-weighted:} \quad \widehat{R}_n(i,j) = \widehat{C}_n(i,j) \cdot T_1 \tag{1a}$$

$$\text{Binary-weighted:} \quad \widehat{R}_n(i,j) = \sum_{k=1}^{\widehat{C}_n(i,j)} 2^{k-1} \cdot T_1 \tag{1b}$$

where T_1 is the the exposure time of the first subframe.

The proposed technique uses $\widehat{I}_n$ and $\widehat{C}_n$ to generate a new exposure code matrix $\widehat{C}_{n+1}$ for frame $n+1$, and to reconstruct an HDR image $\widehat{H}_n$.

The generation of $\widehat{C}_{n+1}$ involves first analyzing the light flux in pixels in the captured frame. A new normalization matrix $\widehat{R}_{n+1}$ can then be obtained. Finally, it is converted to the new exposure code matrix.

Light flux $\widehat{L}_n(i,j)$ of a pixel is defined as light intensity per unit of effective exposure time:

$$\widehat{L}_n(i,j) = \frac{\widehat{I}_n(i,j)}{\widehat{R}_n(i,j)} \tag{2}$$

The new normalization matrix is obtained by dividing the desired pixel output J of the sensor by the light flux matrix:

$$\widehat{R}_{n+1}(i,j) = \frac{J}{\widehat{L}_n(i,j)} \tag{3}$$

The desired output J is set to approximately a half of the pixels' saturation level to assure the maximum margin for over- or under-exposure in the next frame. For example, if the saturation level of the pixels is around $180DN$, then J of $90DN$ is chosen for the operation (DN stands for digital number).

The new exposure code matrix $\widehat{C}_{n+1}$ can be generated using the inverse mapping of equation 1 with some approximation.

The reconstruction of the HDR image $\widehat{H}_n$ requires normalizing $\widehat{I}_n$ by $\widehat{R}_n$:

$$\widehat{H}_n(i,j) = \widehat{I}_n(i,j) \cdot \frac{T_{EXP}}{\widehat{R}_n(i,j)} \tag{4}$$

where T_{EXP} is the total exposure time for frame n in this equation.

The final step for frame n is visualizing the HDR image $\widehat{H}_n$ on a low dynamic range (LDR) monitor. Exposure fusion and gamma correction techniques [10] are used for this purpose. The reconstructed HDR image is split into several 8-bit images to cover the full bit depth. Exposure fusion merges this image set into a single 8-bit image. The advantage of the merge is that it selectively represents the image with desired brightness and texture from the image set. Gamma correction fine-tunes the merged result.

IV. Experimental Results

Experimental results of the extended HDR technique are demonstrated for both unary-weighted and binary-

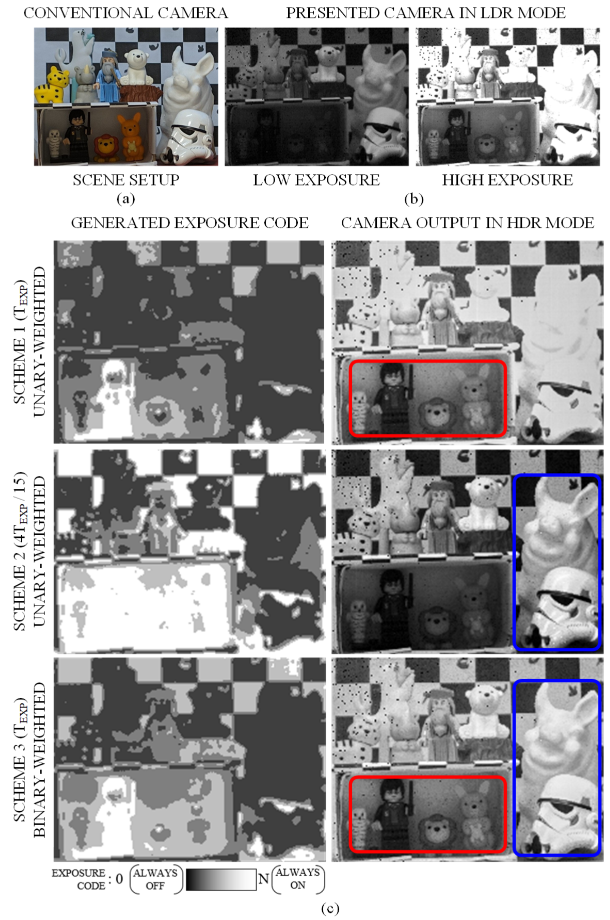

Fig. 5: (a) A scene captured by a conventional camera. (b) Images from the CEP image sensor with short and long exposure time settings. (c) Comparison of the HDR imaging technique for unary-weighted subframe exposure time (in two timing configurations) and binary-weighted subframe exposure time. Exposure codes are shown in the left column and reconstructed HDR images are shown in the right column.

weighted subframe exposure time. The experiments used the two-tap CEP image sensor first presented in [1], fabricated in a $0.11\mu m$ CMOS image sensor technology, where only one tap was used for simplicity. The number of subframes N was set to 4.

The scene used for the experiments is shown in Fig. 5(a), as captured using a conventional RGB camera. Fig. 5(b) shows two images captured using the CEP image sensor with low and high exposure settings. It illustrates the difficulty of capturing both bright and dark regions simultaneously.

Fig. 5(c) compares the results for three different coding schemes. The exposure code matrices are presented on the left, and the reconstructed HDR outputs of the sensor are presented on the right. For unary-weighted subframe exposure, two different timing setups are demonstrated. The first row shows the results for scheme 1 where unary-weighted subframe exposure time is set as follows: $T_1 = T_2 = T_3 = T_4 = T_{EXP}/4$. The second row shows the results for scheme 2 where unary-weighted subframe exposure time is set as follows: $T_1 = T_2 = T_3 = T_4 = T_{EXP}/15$. Scheme 2 has the total exposure time of $4T_{EXP}/15$ per frame. The third row shows results for scheme 3, with binary-weighted subframe exposure time. The exposure time for its first subframe is $T_1 = T_{EXP}/15$ and it has the same total exposure time T_{EXP} as scheme 1.

Dark scenes, highlighted using the red square in Fig. 5(c) are reconstructed well using schemes 1 and 3. Bright scenes, shown using the blue squares in Fig. 5(c) are reconstructed well using schemes 2 and 3. Using binary-weighted subframe exposure time yields the best results and reduces the required number of subframes, thus reducing the data rate required for loading the codes.

V. Conclusions

A method of extending the dynamic range of an image sensor at its native sensor resolution using scene-aware per-pixel-coded exposure has been presented. Two coding schemes are described: for unary-weighted subframe exposure time, the effective dynamic range of the sensor is improved by $20log_{10}N\ dB$, and for binary-weighted subframe exposure time by $20log_{10}(2^N - 1)\ dB$, where N is the number of subframes within one full-rate video frame.

Acknowledgement

The authors gratefully acknowledge the support of the Natural Sciences and Engineering Research Council of Canada under the RGPIN, RTI and SGP programs, of DARPA under the REVEAL program, and of CMC Microsystems.

References

[1] N. Sarhangnejad, N. Katic, Z. Xia, N. Gusev, G. Dutta, R. Gulve, M. Monero, M. Wei, H. Haim, D. Stoppa, K. Kutulakos, R. Genov, "Dual-tap pipelined-code-memory coded-exposure-pixel CMOS image sensor for multi-exposure single-frame computational imaging," ISSCC, 2019.

[2] T. Geurts, B. Cremers, M. Innocent, W. Vroom, C. Esquenet, T. Cools, J. Compiet, B. Okcan, G. Chapinal, C. Luypaert, N. Bresson, P. Deruytere, P. Pintens, R. Aerts, "A 98 dB linear dynamic range, high speed CMOS image sensor," IISW, 2017.

[3] E. Culurciello, R. Etienne-Cummings, K.A. Boahen, "A biomorphic digital image sensor," JSSC, 2003.

[4] Y. Ni, K. Matou, "A CMOS log image sensor with on-chip FPN compensation," ESSCIRC, 2001.

[5] P.E. Debevec, J. Malik, "Recovering high dynamic range radiance maps from photographs," ACM SIGGRAPH, 2008.

[6] S. W. Hasinoff, F. Durand and W. T. Freeman,"Noise-optimal capture for high dynamic range photography", CVPR, 2010.

[7] S. K. Nayar, T. Mitsunaga, "High dynamic range imaging: spatially varying pixel exposures," CVPR, 2000.

[8] J. Zhang, R. Etienne-cummings, T. Xiong, T.D. Tran, S.H. Chin, "Flexible pixel-wise exposure control and readout," U.S. Patent Application 15/785,301, 2018.

[9] S. K. Nayar and V. Branzoi, "Adaptive dynamic range imaging: optical control of pixel exposures over space and time," ICCV, 2003.

[10] T. Mertens, J. Kautz, F. Reeth, "Exposure fusion," PG, 2007.