Color Tone-Mapping Circuit for a Focal-Plane Implementation

Gustavo M. S. Nunes*, Fernanda D. V. R. Oliveira*, José Gabriel R. C. Gomes*, Antonio Petraglia*,
Jorge Fernández-Berni†, Ricardo Carmona-Galán† and Ángel Rodríguez-Vázquez†
*Programa de Engenharia Elétrica - COPPE
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
Emails: gustavo_mn@poli.ufrj.br, fernanda.dvro@poli.ufrj.br, gabriel@pads.ufrj.br, antonio@pads.ufrj.br
†Instituto de Microelectrónica de Sevilla (IMSE-CNM)
CSIC-Universidad de Sevilla, Sevilla, Spain
Emails: berni@imse-cnm.csic.es, rcarmona@imse-cnm.csic.es, angel@imse-cnm.csic.es

Abstract—In this article, we present a review of the driving principles and parameters of a previously reported focal-plane tone-mapping operator. We then extend it in order to include color information processing. The signal processing operations required for handling color images are white balance and demosaicing. Neither white balance nor demosaicing are carried out in the focal plane, in order to avoid increasing circuit size and complexity. Since, in this case, white balance is carried out after tone mapping, multiplication of red and blue channels by constant gains may lead to wrong color results. An alternative approach is proposed, in which different gains are assigned for every red and blue pixel of the matrix. Because of the introduction of color, a modification in the original circuit is proposed, which affects the integration time of red and blue pixels. This modification leads to a reduction in the number of photodiodes required in the pixel array, and hence to a reduction of the sensing circuit area. The results produced by the operator are compared to those obtained from two other digital tone-mapping operators.

I. Introduction

Real-world scenes usually contain many different illumination conditions, provided by multiple light sources. The luminance (amount of light per unit area) emitted by a source can be roughly as low as 10^{-3} cd/m^2 (starlight) or as high as 10^5 cd/m^2 (sunlight) [1]. Thus, the luminance values presented in a scene may span several orders of magnitude.

Through advanced biological mechanisms, the human visual system adapts itself to the current illumination conditions, accommodating the high dynamic range of luminance values of the scene in such a way that details in both darker and brighter regions are observed. However, when attempting to display the same scene without any pre-processing in a conventional display medium, the same level of detail is not maintained, since such conventional displays can typically represent values in an 8-bit luminance range.

In High Dynamic Range Imaging (HDRI), the display task is achieved by tone-mapping operators, which are non-linear mappings from the real scene luminances to display luminances. Some mappings are inspired by photographic techniques [2], while others are based on the human eye perception of brightness and contrast [3] and become even more sophisticated as they try to incorporate some aspects of human vision, such as visual acuity, color sensitivity and glare [4].

The tone-mapping operators implemented in software are referred to as *digital* tone-mapping operators in this work. The tone-mapping operator originally proposed by Fernández-Berni *et. al.* [5] is not digital: it is implemented inside the focal plane of the camera, before analog-to-digital conversion, and immediately maps the HDR image to the 8-bit range, as soon as it is acquired. The (grayscale) image is then readily available for display, without requiring further processing or multiple exposures of the scene. In this paper, the tone-mapping operator is extended to include color information. Because of the introduction of color, a modification in the circuit is proposed, to reduce the number of photodiodes present in the pixel array, thus decreasing the size of the circuit.

This paper is organized as follows. A brief overview on how this operator works, as well as some of its parameters, is given in Section II. In Section III we present the proposed operator modifications, which include a modification of the internal circuit and the introduction of color treatment. The results of three versions of the operator, compared against other digital tone-mapping operators, and the conclusions of the work are shown in Sections IV and V, respectively.

II. Focal-plane Operator Description

The tone-mapping curve of the operator described in [5] is adjusted at each pixel, according to its local luminance value and to the average luminance value of the scene. Each pixel cell contains two photodiodes, which store the local luminance value for different purposes: the *capture* photodiode, whose photocurrent determines the actual pixel value; and the *control* photodiode, whose photocurrent is used to calculate the average luminance value of the scene.

The focal-plane tone-mapping circuit depends on component values and time-domain waveforms that affect output image perceptual attributes such as contrast and brightness. The digital tone-mapping operators are typically expressed in terms of parameters that control output image perceptual

attributes. Therefore, in a digital simulation of the tone-mapping operator from [5], six parameters must be set: V_{mid}, C, m_{ph}, m_c, T_{max}, and T_s.

Parameter V_{mid} is a voltage threshold that, once reached by the control node, stops the photointegration (i.e. the conversion from photocurrent to voltage) at the capture node, thereby making its voltage ready for readout. This voltage value is typically mapped into the pixel value which corresponds to the middle of the dynamic range. Pixel values above or below that voltage threshold are represented equally, which corresponds to reasonable dynamic range usage. The highest possible voltage for a pixel, which maps it into black, is defined as V_{rst}. As V_{mid} is set closer to V_{rst}, the resulting images tend to be darker. On the contrary, brighter images are produced if V_{mid} is set closer to V_{min}, which is the lowest possible voltage for a pixel.

Parameter C is the overall capacitance of the capture node. The photodiode is modelled as a current source in parallel with a capacitor, and C takes the photodiode model capacitance and the parasitic capacitances into account. Higher C values make the photocurrent discharge slower at the capture node, which leads to higher output voltage values and hence darker images.

The capture and control photodiode sizes may be different, which leads to possibly different sensitivities: the same light intensity generates different photocurrents in the capture and control nodes. The overall capacitances of both nodes may also be different, since each node has its own photodiode model capacitance and parasitic capacitances associated with it. These differences are represented by parameters m_{ph} and m_c: the ratio between the current source values at the capture and the control node is m_{ph}, and the ratio between the overall capacitances at the capture and the control node is m_c. The node (either capture or control) discharge rate depends on photocurrent value and on overall capacitance value, so defining a discharge ratio $r = m_{ph}/m_c$ is convenient. If $r > 1$, then the capture node discharge is faster than the control node discharge, thus yielding brighter pixels.

Parameter T_{max} is the maximum integration time. Pixels whose control nodes have not reached V_{mid} will have their photointegrations at the capture nodes ceased at this instant. As T_{max} decreases, fewer pixels will have their control nodes reach V_{mid}. Their photointegration will be stopped earlier, which yields darker pixels. Once the control node of a pixel reaches V_{mid}, the pixel output value will remain unaffected by an increase in T_{max}. Hence, an increase in T_{max} affects only the darker pixels of the image. Darker regions of the image tend to become gray, while other regions remain unchanged. The net effect corresponds to a decrease in the contrast of the scene.

Finally, parameter T_s determines how long the control nodes of all pixels must remain connected. During the T_s time interval, the voltages of all control nodes evolve in the same manner, according to the photocurrent averaged over all pixels of the scene. After this time, the voltage of each control node changes according to the local photocurrent. Low T_s values make the voltage curves of both the capture and control nodes of a pixel more similar. Consequently, the images tend to have low contrast, since the voltages of the capture nodes are close to V_{mid}. As T_s increases, contrast and details of the scene are enhanced.

III. PROPOSED MODIFICATIONS

To incorporate color information at the pixel level, the conventional color filter array (red, green and blue filters arranged according to the Bayer pattern) is placed on top of the pixel matrix. To yield a color image, white-balance and demosaicing operations are required. To keep focal-plane circuit complexity low, only tone mapping is performed at the focal plane (i.e. before white balance and demosaicing).

A. White-balance Gains

In conventional digital image capture, white balance happens before demosaicing and tone mapping. After the proposed focal-plane tone-mapping approach, the white-balance stage should use a different gain for each pixel, as described next.

White balance with constant red and blue gains, followed by tone mapping and demosaicing (which is the conventional operation order), yields images with natural-looking colors. We use this operation sequence as a model to find the correct gains for each pixel after tone mapping has been independently performed for each color channel. We refer to this method as the *reference* method. The reference value for the (i,j)-th red pixel $R_{TM_{i,j}}$ obtained by this method, after tone mapping, is made equal to the corresponding red pixel $R'_{TM_{i,j}}$ obtained after tone mapping and without any previous white balance multiplied by an unknown gain K'_r (the same procedure applies to blue pixels), that is,

$$K'_r \cdot R'_{TM_{i,j}} = R_{TM_{i,j}}. \tag{1}$$

Both $R'_{TM_{i,j}}$ and $R_{TM_{i,j}}$ depend on the integration time T_{int} and on the corresponding original raw pixel value $R_{i,j}$. The integration time associated with $R_{TM_{i,j}}$ is equal to $\min(T_{max}, T_{mid})$. The time instant T_{mid} can be expressed as a function of operator parameters as follows:

$$T_{mid} = \frac{1}{K_r \cdot R_{i,j}} \cdot \left[\beta_1 - T_s \cdot \overline{I} \right] + T_s, \tag{2}$$

where $\beta_1 = \frac{C}{k} \cdot \frac{m_{ph}}{m_c} \cdot (V_{mid} - V_{min})$, $\overline{I}$ is the average pixel value of the raw input image, k is the maximum photocurrent that can be generated by a photodiode of the circuit, and K_r is the constant white-balance gain for the red channel. This gain is obtained through the *DCRaw* software [6], which calculates it based on the camera model and on the white-balance profile settings. The expression for T_{mid} in the new method is the same as in Eq. (2), except that there is no gain factor multiplying $R_{i,j}$, since no prior white balance is applied.

Because the constant white-balance gains are higher than 1, T_{mid} of the reference method can never be greater than T_{mid} of the new method. Therefore, there are three cases for Eq. (2): when $T_{mid} > T_{max}$ in both methods (case 1); when $T_{mid} < T_{max}$ in both methods (case 2); and when $T_{mid} > T_{max}$ in the new method and $T_{mid} < T_{max}$ in the reference

method (case 3). The final white-balance gain expressions are summarized in Eq. (3):

$$K'_r = \begin{cases} K_r, & \text{if case 1} \\ \dfrac{\beta_1 + T_s \cdot \left(K_r \cdot R_{i,j} - \overline{I}\right)}{\beta_1 + T_s \cdot \left(R_{i,j} - \overline{I}\right)}, & \text{if case 2} \\ \dfrac{\beta_1 + T_s \cdot \left(K_r \cdot R_{i,j} - \overline{I}\right)}{T_{max} \cdot R_{i,j}}, & \text{if case 3} \end{cases} \quad (3)$$

B. Raw Values Reconstruction

As shown in Eq. (3), the raw pixel values from the input image are required to compute the white-balance gains. In a real focal-plane implementation, such values are not available, since they are directly tone-mapped after acquisition, and their original (raw) values are not stored. The raw values must be reconstructed from the tone-mapped pixels before the white-balance coefficients are calculated.

The voltage at the capture node of the red pixel $R_{i,j}$ is given by $V_{cap_{i,j}} = V_{rst} - T_{int} \cdot \beta_2 \cdot R_{i,j}$, where $\beta_2 = \frac{k}{C}$. This value is replaced in the relation $R_{TM_{i,j}} = 1 - V_{cap_{i,j}}/V_{rst}$ to yield the expression for the tone-mapped red pixel $R_{TM_{i,j}}$. Solving for $R_{i,j}$, the following expression is obtained for the reconstructed raw value $R_{rec_{i,j}}$:

$$R_{rec_{i,j}} = \frac{V_{rst}}{T_{int} \cdot \beta_2} \cdot R_{TM_{i,j}}. \quad (4)$$

Because T_{int} is equal to $\min(T_{max}, T_{mid})$, there are two possible reconstruction curves. In the real focal-plane implementation, the pixel integration time is not measured and therefore it is unknown. To determine the curve to be used, a condition has to be stated in terms of the compressed pixel value $R_{TM_{i,j}}$, which is available, rather than the integration time. A threshold that can be used for this decision is the point at which both tone-mapping reconstruction curves intersect each other. This point is computed by using $T_{int} = T_{max}$ and $T_{int} = T_{mid} = \frac{1}{R_{i,j}} \cdot (\beta_1 - T_s \cdot \overline{I}) + T_s$, in Eq. (4), which produces two expressions for $R_{rec_{i,j}}$. Equating the expressions and solving for $R_{TM_{i,j}}$ yields the threshold

$$R_{th} = \frac{T_{max} \cdot \beta_2}{V_{rst}} \cdot \left[\frac{\beta_1 - T_s \cdot \overline{I}}{T_{max} - T_s} \right]. \quad (5)$$

The reconstructed raw value $R_{rec_{i,j}}$ is then given by:

$$R_{rec_{i,j}} = \begin{cases} \frac{V_{rst}}{T_{max} \cdot \beta_2} \cdot R_{TM_{i,j}}, & R_{TM_{i,j}} \leq R_{th} \\ \frac{V_{rst} \cdot R_{TM_{i,j}} - \beta_2 \cdot (\beta_1 - T_s \cdot \overline{I})}{T_s \cdot \beta_2}, & R_{TM_{i,j}} > R_{th} \end{cases} \quad (6)$$

C. Integration Time Control

In the original circuit, each pixel cell requires two photodiodes, because its control node voltage depends both on the global average photocurrent of the matrix and on the local photocurrent. Consequently, each pixel integration time is controlled by its own sensor value and the global average value. However, after introducing the Bayer pattern color filter arrays, twice as many green as blue or red pixels are produced. Tests were then conducted using the global average pixel value and the average green pixel value in Eq. (2) and no
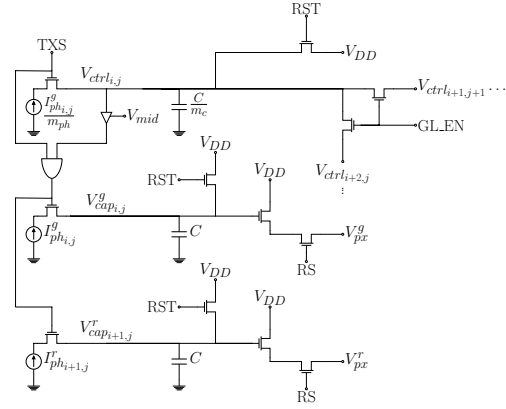

Fig. 1. Modified circuit schematic diagram

differences between the output images were observed (the output images using both values are available in [7]). Thus, when $t < T_s$, only the green pixels of the matrix are connected together, rather than all pixels, in order to obtain the average photocurrent.

Since only green pixels are used to calculate the average photocurrent, a new modification of the circuit is proposed: the integration time of the red and blue pixels is determined by the control node voltage of the neighboring green pixels. With this modification, the red and blue pixels no longer need a control photodiode, leading to a reduction of the sensing circuit area: for every 2x2 pixels (RGGB), only 6 photodiodes are needed (4 capture and 2 control), instead of 8. The schematic diagram (for a red-green pixel pair) is shown in Figure 1.

Another advantage of the proposed modification is that the white-balance operation is simplified. Since green pixels remain unchanged after the white balance (their gains are set to unity), the integration times for red (and blue) pixels, in the new and reference methods described in Section III-A, are the same. As a result, the white-balance gains used after tone mapping have the same constant gains, as if white balance was done before tone mapping.

IV. RESULTS

In this section, the results of three different digital implementations of the proposed operator are shown. The first version uses the framework commonly adopted by other digital tone-mapping operators, which is defined as follows: first, the image is acquired through a CMOS sensor (using a particular Bayer pattern); then, white balance and demosaicing are performed on the raw sensor data, yielding a colored image; the tone-mapping operator is then applied on the luminance map of the color image. To recover the color information from the tone-mapped luminance, each original color channel is scaled by the ratio between the tone-mapped and original luminances at each pixel, as proposed in [9]. The demosaicing function used for all operators is based on [10].

The second and the third versions use the framework established for the real focal-plane implementation, which is similar to the previous one, except that the processing order is reversed (i.e. the tone mapping is carried out before the white balance
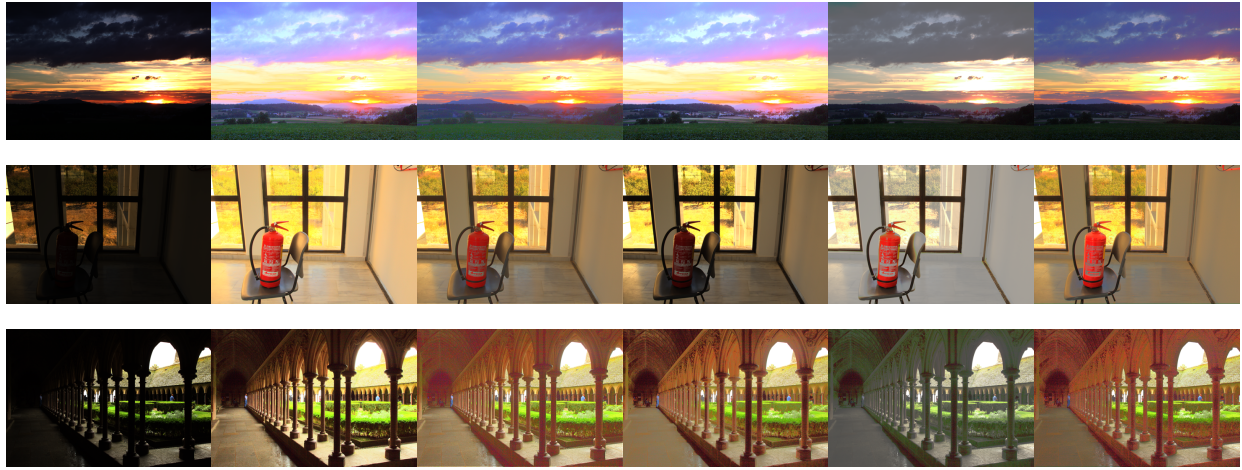
Fig. 2. Digital simulation results obtained after the application of tone-mapping operators. First column: raw images. Second column: Schlick operator. Third column: Rahman operator. Fourth column: first version of the proposed operator. Fifth column: second version of the proposed operator. Sixth column: third version of the proposed operator. The images were downloaded from [8].

and demosaicing). Rather than being applied on the original color image luminance map, the tone-mapping operator is directly applied to the raw pixel values from the camera sensor. The difference between these last two versions lies in how the blue and red pixel integration time is determined: in the second version, it is determined by the very pixel value, while in the third one it is determined by the value of a neighboring green pixel.

To assess the quality of the images produced by each version of the proposed operator, the resulting images are compared with the images obtained by two other digital tone-mapping operators, namely the Schlick operator [11] and the Rahman operator [9]. Each operator uses a set of parameters, which yield a different image. In order to generate the best results, the operator parameters were tuned according to the following design criteria: details should be observed in all darker and brighter regions of the image (i.e. over- or underexposing particular regions of the image should be avoided), and the colors of the scene look natural.

Figure 2 shows the images resulting from the application of all operators to three different HDR images, along with the original raw images. For the raw images, no previous tone mapping is performed. As can be seen for all operators, details that were not visible in the original raw images, specially in darker regions, are now made visible: for example, the bottom part of the sunset image, the indoors of the fire extinguisher image and the corridor of the courtyard image, without losing important details in brighter regions.

The colors also look natural for all operators, although they seem different in the second version of the proposed operator. In that case, the integration time is adapted according to the local pixel value, in order to avoid over- and underexposures. This tends to concentrate the tone-mapped pixel values around the center of the dynamic range, and hence, after demosaicing, colors will be less saturated. In the third version, since red and blue pixels have their integration times determined by the

green pixel next to them, their tone-mapped values may be more spread out along the dynamic range, leading to more saturated colors.

Another point to be considered is dark pixel noise. Some images, such as the courtyard, contain very dark raw pixels that have values close to the noise floor of the sensors. Depending on the tone-mapping operator applied to them, salt-and-pepper noise may arise in such regions. This is the case for the Rahman and the proposed operator presented in this work. It becomes even more evident as the resolution of the image increases, and image quality is degraded. To suppress noise effects, an additional low-pass filtering stage may be required after demosaicing. Noise is not observed in the images shown in Figure 2, because they were pre-filtered by a low-pass bicubic filter, and then rescaled to 10% of their original sizes. The non-filtered and filtered versions of the courtyard images with original resolutions are available [7].

V. CONCLUSION

If it were to be implemented digitally, the proposed operator would yield images that are similar to images produced by other digital tone-mapping operators, as shown by the results of the first version. The second and third versions generate images whose quality, regarding colors and details, is also comparable to the digital competitors, while retaining the advantages of a focal-plane implementation. The proposed operator successfully generates color tone-mapped images for a focal-plane implementation, therefore achieving the tone-mapping task.

REFERENCES

[1] B. Wandell, *Foundations of Vision*. Sinauer Associates, 1995.

[2] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, Jul. 2002. [Online]. Available: http://doi.acm.org/10.1145/566654.566575

[3] J. Tumblin and H. Rushmeier, "Tone reproduction for realistic images," *IEEE Computer Graphics and Applications*, vol. 13, no. 6, pp. 42–48, Nov 1993.

[4] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, Oct 1997.

[5] J. Fernández-Berni, F. D. V. R. Oliveira, R. Carmona-Galán, and A. Rodríguez-Vázquez, "Image sensing scheme enabling fully-programmable light adaptation and tone mapping with a single exposure," *IEEE Sensors Journal*, vol. 16, no. 13, pp. 5121–5122, July 2016.

[6] D. Coffin. [Online]. Available: https://www.cybercom.net/~dcoffin/dcraw/

[7] [Online]. Available: http://www.pads.ufrj.br/~fernanda/ColorToneMappingCircuitForFocalPlaneImplementation/Samples/

[8] "ETH Zürich, Empamedia," http://empamedia.ethz.ch/gmdb/, Last accessed 19 July 2017.

[9] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.

[10] H. S. Malvar, L. wei He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, May 2004, pp. iii–485–8 vol.3.

[11] C. Schlick, "Quantization techniques for visualization of high dynamic range pictures." Springer-Verlag, 1994, pp. 7–20.