Illumination technique for optical dynamic range compression and offset reduction

C. Koch, S.-B. Park[1], T.J. Ellis and A. Georgiadis[2]
City University, London, EC1V OHB, UK
[1]BMW AG, D-80788 Munich, Germany
[2]FH Nordostniedersachsen, D-21339 Lüneburg, Germany

Abstract

This paper presents a novel illumination technique for image processing in environments which are characterized by large intensity fluctuations and hence a high optical dynamic range (HDR). This proposal shows how by combining a set of images, flashed with different radiant intensities but with a constant exposure time for the imager, a single image can be produced with a compressed dynamic range and a simultaneously reduced offset.

This makes it possible to capture high dynamic range scenes without using a high dynamic range camera. This technique can be used as the first signal processing step to simplify the segmentation in applications such as: face recognition, interior surveillance, vehicle occupant detection or motion detection in general.

1 Introduction

Mainstream CCD based (and most of the emerging CMOS based image sensors) provide an optical dynamic range of $48 \ldots 60$ dB. This dynamic range is sufficient for scenes with homogeneous illumination and without extreme contrasts. Extreme contrasts may occur when operating an imager in direct sunlight or within scenes containing areas of high brightness and deep shadow, e.g. a building entrance. Other classic examples are looking from a dimly lit room through a window towards a bright outdoor scene, the direct view into an operating light bulb or a vehicle interior (see Fig. 5).

Usually it is possible to adjust common imagers either to the very bright areas or to the dark areas in the scene by adapting several of the imager's parameters, e.g. exposure time, lens aperture or by the use of optical filters. Nevertheless in extreme dynamic environments there will be areas in the scene which are over or under exposed, resulting in lost image detail. In particular CCD based imagers tend to suffer this problem because only a small number of over exposed pixels can yield large saturated areas (due to blooming and smearing). This well known limitation is shown in Fig. 4.

There are various ways to extend the limited dynamic range of an imager. One approach uses several images of the same scene taken with different exposure times which are then assembled into one high dynamic image [2, 1]. Another approach is to use an imager with *intelligent* pixels to choose the optimal exposure time individually [5]. Both techniques have the disadvantage of a high processing requirement in order to produce an output, and as such their system design is specialized for industrial applications such

as visual welding inspection. Hence, these imagers are expensive and the frame rate is usually slower when compared to mainstream designs.

An alternative and well known way to reduce the influence of ambient illumination and thereby to reduce the range of brightness variation within a scene is to utilize supplementary illumination in combination with an optical band pass filter. Using a filter with a center wavelength (CWL) at the wavelength of the supplementary illumination source blocks all irradiation from the ambient illumination outside the transmission wavelength band. A suitable wavelength range ($\Delta\lambda$) for such an illumination is the near infrared (NIR), because it is not visible to the human eye which is important for many surveillance tasks. Also most cameras based on silicon are still sensitive at these wavelengths and the spectral power density of the main disturbance *noise* source (the sun) decreases significantly after $\lambda = 800\,\text{nm}$. This effect can be improved by utilizing a strong flashed light source which increases the signal-to-noise ratio (SNR), suppresses the noise induced by the environment and minimizes its influence on the image.

However problems with strong flashing can occur if large areas have to be illuminated, because this can require an illumination source whose optical intensity can reach the threshold values for eye safety [3]. This paper describes a solution to this problem without the need for special high dynamic range (HDR) cameras and without violating eye safety limitations whilst retaining the important scene details. Section 2 gives an introduction to offset reduction due to ambient illumination and section 3 describes mathematically the definition of optical dynamic range. This leads to section 4 where our new double flash algorithm is introduced. Section 5 illustrates the proposed dynamic range compression with an example calculation and describes some preliminary experimental results.

2 Offset reduction

In order to reduce the necessary pre-processing steps for image analysis, it is obviously preferable to capture a sequence consisting of N frames ($n = 1, \ldots, N$) without global or local illumination fluctuations, i.e. without varying illumination offset and noise. This means that the irradiance E which is integrated over the exposure time (t_e) of the imager should be constant for every frame and hence the grey level of an arbitrary pixel at position $\mathbf{p}$ within a digital image $I(\mathbf{p}, n)$ only varies if the scene has changed, i.e. only if the reflectivity of the surface $\rho_s(\mathbf{p})$ and not the ambient illumination has changed.

$$
\begin{aligned}
I(\mathbf{p}, n) &= \rho_s(\mathbf{p}) \int_{t=0}^{t_e} E(t)\, dt \\
\mathbf{p} &= (row, column) \\
0 < \rho_s(\mathbf{p}) &< 1
\end{aligned}
\tag{1}
$$

This would simplify most image processing steps, e.g. background updating for motion detection [6], thus decreasing the necessary computation time and reducing the cost of the processing hardware, and is therefore a further step towards real-time image processing. One way to achieve this offset and noise reduction is to acquire one image I_{sun} with only ambient illumination (E_{sun}) at frame n and a second image I_{flash} with additional

illumination $(E_{sun} + E_{nir})$ from a supplementary light source (e.g. NIR) at frame $(n+1)$.

$$I_{sun}(\mathbf{p}, n) \quad = \quad \rho_s(\mathbf{p}) \int_{t=0}^{t_e} E_{sun}(t)\, dt \tag{2}$$

$$I_{flash}(\mathbf{p}, n+1) \quad = \quad \rho_s(\mathbf{p}) \int_{t=0}^{t_e} (E_{sun}(t) + E_{nir}(t))\, dt \tag{3}$$

$$I_{nir}(\mathbf{p}, n+1) \quad = \quad |I_{sun}(\mathbf{p}, n) - I_{flash}(\mathbf{p}, n+1)| \tag{4}$$

Assuming that the scene is static; hence $\rho_s(\mathbf{p})$ is constant over all frames. This means that the captured greylevels within the images are only a function of E which illuminates the scene. Thus, the difference between the images I_{sun} and I_{flash} yields only the received radiant power of the local illumination which is constant (see eqn. (5). The variable offset of the ambient illumination source E_{sun} is eliminated and the output sequence $I_{nir}(\mathbf{p}, n)$ is thus exempt from light fluctuations [4].

$$|E_{sun} - (E_{sun} + E_{nir})| = E_{nir} = \text{const.} \tag{5}$$

$$\rightarrow I_{nir} \propto E_{nir} \qquad if \qquad \rho_s(\mathbf{p}) = \text{const.} \tag{6}$$

This method is illustrated in Fig. 1. It shows the intensity difference between frame I_{sun} and I_{flash} of the same pixel over a sequence of 60 frames. The sequence begins with small illumination fluctuations and ends with a double-peak which was induced using a powerful external light source directed at the scene for a short time. Referring to eqn. (5), the calculated difference between I_{sun} and I_{flash} should ideally be constant. The plot of I_{nir} differs marginally from an (ideal) straight line due to sensor noise and small illumination changes of E_{sun} between both input images. The sequence was produced with a linear HDR camera [1] (optical range of up to $120\,\mathrm{dB}$, see section 5). The following conditions must hold:

- **Dynamic:** The camera has to cover the whole optical dynamic range of the scene and must not saturate or under-expose any pixel, otherwise the subtraction of consecutive images yields an unrepresentative result. This means that for an environment with large irradiance differences in the spatial and time domain, a HDR camera is required.

- **Speed:** The time gap between I_{sun} and consecutive I_{flash} should be kept as small as possible to minimize the effect of changing ambient irradiation E_{sun} or surface reflectivity $\rho_s(\mathbf{p})$, caused by moving objects, which results in an inaccurate output image.

- **Reach:** The whole scene which has to be analyzed must be within the scope of the local illumination E_{nir}. This condition occurs in closed environments, e.g. within a vehicle interior or building.

3 Dynamic range compression

The utilization of supplementary illumination in section 2 has the additional effect upon the flashed image I_{flash} such that the optical dynamic range within the image is compressed compared to the image without supplementary illumination. This is explained in the following paragraphs.
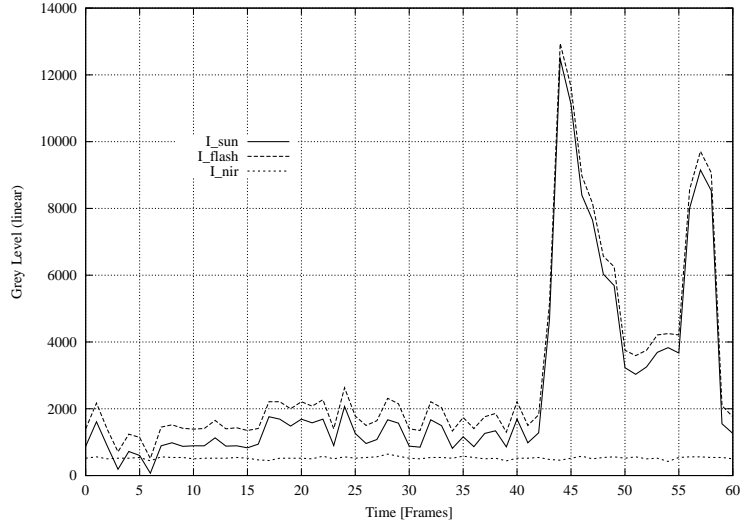
Figure 1: Principle of offset reduction

The image intensity $I(\mathbf{p}, n)$ generated by a digital imager represents the photogenerated current i_{ph} of each pixel. This current is a function of incident irradiance, sensor offset, gain and sensor noise. The dynamic range for an image sensor (DR) is commonly defined as the ratio of its largest nonsaturating signal ($i_{ph,max}$) to the standard deviation of the noise under dark conditions ($i_{ph,min}$) [7].

$$DR = 20 \log\left(\frac{i_{ph,max}}{i_{ph,min}}\right) \tag{7}$$

Assume a scene which exhibits a wide DR in both the time and spatial domains due to fluctuating ambient illumination and different interior surface materials, or within rooms with large windows. For simplification, assume that i_{ph} is proportional to the incident irradiance E and that the imager is located in such a way that there is no irradiation directly incident on the sensor. Hence, the DR within the image (spatial domain) or between two frames at time n and $n+1$ (time domain) will be determined by the product of the reflectivity of the interior surface $\rho_s(\mathbf{p})$ and the occurring maximum and minimum irradiance E_{max}, E_{min} respectively.

$$i_{ph} \propto E \tag{8}$$

$$\rightarrow DR = 20 \log\left(\frac{\rho_s(\mathbf{p}) E_{max}}{\rho_s(\mathbf{p}) E_{min}}\right) \tag{9}$$

Within a scene without supplementary illumination, the irradiance E_{min} and E_{max} are determined only by ambient illumination sources. In vehicles and buildings etc. this ambient source is primarily the sun E_{sun}.

$$
\begin{aligned}
E_{max} &= E_{sun,max} \\
E_{min} &= E_{sun,min}
\end{aligned}
\tag{10}
$$

296

In the case that the image is illuminated with a supplementary illumination E_{flash}, E_{nir} is added to the ambient E_{sun}.

$$
\begin{aligned}
E_{flash,max} &= E_{sun,max} + E_{nir} \\
E_{flash,min} &= E_{sun,min} + E_{nir}
\end{aligned}
\tag{11}
$$

If eqn. (10) and (11) are inserted into eqn. (8) eqn. (12) results which shows that the optical dynamic range of a scene with supplementary illumination (DR_{flash}) is lower than the scene without (DR_{sun}). This effect is independent of the reflectance $\rho_s(\mathbf{p})$ which is shown in eqn. (13).

$$
\begin{aligned}
DR_{sun} &> DR_{flash} \tag{12} \\
20 \log \left(\frac{\rho_s(\mathbf{p}) E_{sun,max}}{\rho_s(\mathbf{p}) E_{sun,min}} \right) &> 20 \log \left(\frac{\rho_s(\mathbf{p})(E_{sun,max}+E_{nir})}{\rho_s(\mathbf{p})(E_{sun,min}+E_{nir})} \right) \\
\frac{E_{sun,max}}{E_{sun,min}} &> \frac{E_{sun,max} + E_{nir}}{E_{sun,min} + E_{nir}} \\
E_{sun,max}(E_{sun,min}+E_{nir}) &> E_{sun,min}(E_{sun,max}+E_{nir}) \\
E_{sun,max} E_{nir} &> E_{sun,min} E_{nir} \\
E_{sun,max} &> E_{sun,min} \tag{13}
\end{aligned}
$$

$$qed.$$

This means that by using supplementary illumination which brightens the whole scene, the optical dynamic range of the scene is compressed. The power of E_{nir} determines if areas that were formerly too dark or too bright (and thus out of the dynamic range of the imager) can now be accurately captured (see Fig. 4). Fig. 2 illustrates this effect. It shows the same offset reduction plot with identical data as Fig. 1 but with logarithmic scaling of the grey values to highlight the changes to the dynamic range of DR_{sun} and DR_{flash}. However, if the offset reduction is applied, an input image without local illumination for determining E_{sun} is still necessary for calculating the offset free output. This image might still feature a high dynamic range. Hence, for analyzing HDR environments as described in the introduction, a HDR camera is still essential, if the offset reduction should be used. A solution to this problem is our proposed double flash technique which is described in the following section.

4   Double flash

We introduce a new approach which combines the advantages of both offset reduction and dynamic range compression by illuminating two input images with *different radiant intensities*. Both input images $I_{flash,hi}$ and $I_{flash,low}$ are compressed in their optical dynamic range due to the supplementary illumination, and the output image is also free of illumination fluctuations. The output image I_{df} of the double flash is calculated as follows:

$$
I_{flash,hi}(\mathbf{p}, n) = \rho_s(\mathbf{p}) \int_{t=0}^{t_e} (E_{sun}(t) + E_{nir,hi}(t))\, dt
\tag{14}
$$

$$
I_{flash,low}(\mathbf{p}, n+1) = \rho_s(\mathbf{p}) \int_{t=0}^{t_e} (E_{sun}(t) + E_{nir,low}(t))\, dt
\tag{15}
$$

$$
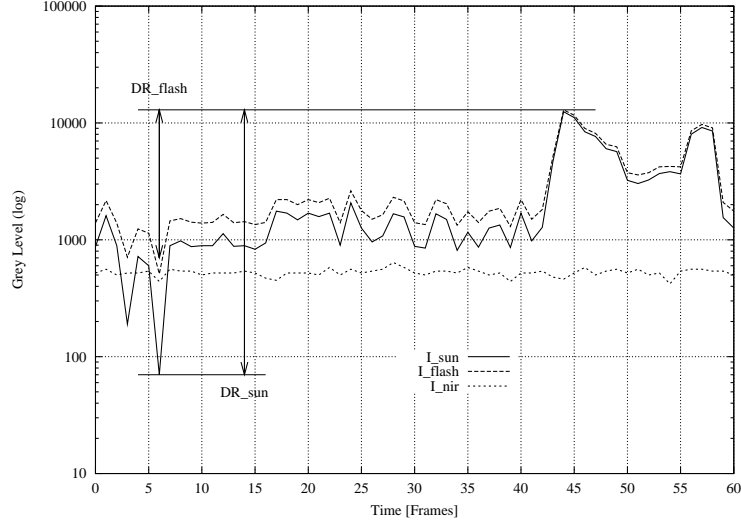I_{df}(\mathbf{p}, n+1) = |I_{flash,hi}(\mathbf{p}, n) - I_{flash,low}(\mathbf{p}, n)|
\tag{16}
$$

Figure 2: Compressed dynamic range due to supplemental illumination E_{nir}. The double arrows indicate the difference between DR_{sun} and DR_{flash}.

The influence of the ambient illumination E_{sun} has according to eqn. (17) been eliminated.

$$|(E_{sun} + E_{nir,hi}) - (E_{sun} + E_{nir,low})| = |E_{nir,hi} - E_{nir,low}| = \text{const.} \quad (17)$$
$$\rightarrow I_{df} \propto E_{nir,hi}, E_{nir,low} \qquad if \qquad \rho_s(\mathbf{p}) = \text{const.} \quad (18)$$

The difference to the basic offset reduction in eqn. (5) is that $E_{nir,low}$ operates as the gain for the minimum radiation E_{min} which was previously at I_{sun} and only defined by the ambient illumination E_{sun} (see eqn. (5) and (10)). The original dynamic range of *both* images are thus compressed by E_{nir} according to eqn. (12).

$$E_{flash,hi,max} = E_{sun,max} + E_{nir,hi} \quad (19)$$
$$E_{flash,hi,min} = E_{sun,min} + E_{nir,hi} \quad (20)$$
$$E_{flash,low,max} = E_{sun,max} + E_{nir,low} \quad (21)$$
$$E_{flash,low,min} = E_{sun,min} + E_{nir,low} \quad (22)$$

This leads to the effect that the maximum dynamic range of the double flash output image (DR_{df}) is now defined by the ratio between the adjustable local illumination intensity $E_{nir,hi}$ and $E_{nir,low}$. This is shown in eqn. (23) and (24) where eqn. (19) ... (22) are inserted into the basic dynamic range definition of eqn. (8).

$$20 \log(a) = DR_{df} \quad (23)$$
$$a = \left| \frac{E_{flash,hi,max}}{E_{flash,hi,min}} - \frac{E_{flash,low,max}}{E_{flash,low,min}} \right| = \left| \frac{E_{sun,max} + E_{nir,hi}}{E_{sun,min} + E_{nir,hi}} - \frac{E_{sun,max} + E_{nir,low}}{E_{sun,min} + E_{nir,low}} \right| \quad (24)$$

The timing diagram of the trigger logic for this illumination technique is shown in Fig. 3. High power light emitting diodes could be used as illumination sources. The diodes
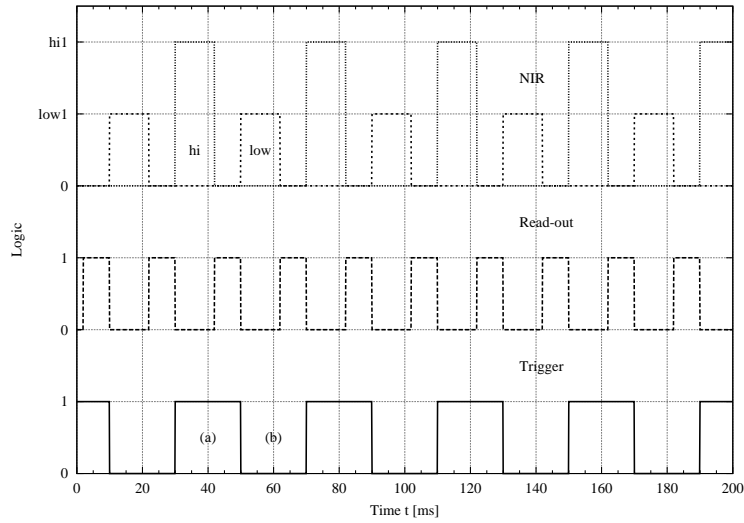
298

Figure 3: Camera timing for the double flash: One frame is equal to 20 ms, including 8 ms for read-out and transfer to the PC.

operate with alternating high and low currents and emit $E_{nir,hi}$ and $E_{nir,low}$, respectively. In the case of a camera with a global and synchronous shutter it is far more efficient to activate the LED illumination only when the imager pixels are integrating light. The off-time when the image is read out can be used to allow the LED's to cool off and also to reduce the total power consumption. Hence, it is possible to run the LED's with higher pulse currents resulting in the use of smaller diodes with no reduction in radiant power. The absolute function which is applied in eqn. (5) and (17) implies that the illumination order of $E_{nir,low}$ and $E_{nir,hi}$, i.e. the capturing to $I_{flash,low}$ and $I_{flash,hi}$ respectively, is arbitrary. Thus, although the offset reduction and double flash need two input images, they produce one output image for every new input after the second captured image because only an *irradiance change* between subsequent frames is required. The last captured image at time $(n-1)$ can be compared to the current image n for computing the next I_{df} output image and hence suffer only a 1 frame delay.

5 Experiments

Fig. 4 shows a scene in a laboratory which includes high irradiance differences, generated by a bright spot light from the right. This spot light in combination with the ambient illumination simulates E_{sun}. A mainstream CCD-based camera with a single integration time (and thus limited dynamic range) is employed to capture the scene. The integration time in Fig. 4(a) was optimized for reading the labels of the power supply. By extending the integration time in Fig. 4(b), the silhouette of a dark ring on top of the floppy disk becomes visible but most of the image is over-exposed. The imager was not able to capture the scene without losing detail with a single exposure time. A supplementary illumination E_{nir} as used in section 3 and 4 in combination with a shortened integration
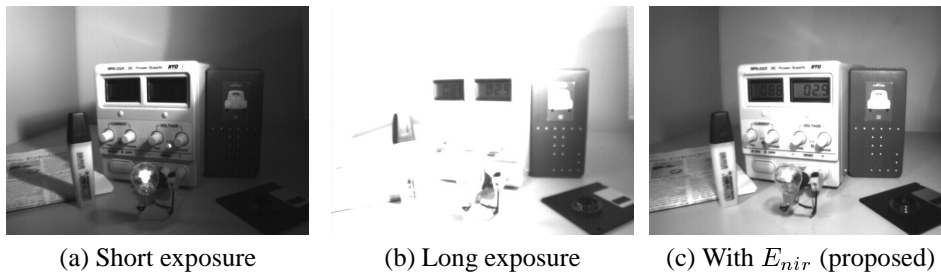
(a) Short exposure   (b) Long exposure   (c) With E_{nir} (proposed)

Figure 4: Example of a optical high dynamic range scene. Image (c) indicates the result of the proposed dynamic range compression by supplementary illumination.

time (Fig. 4(c)) yields the capability to capture the labels as well as the dark ring on top of the floppy with just one exposure time. The primary direction and intensity of the disturbing light source from the right remained untouched for every image. The radiance emitted by the supplementary illumination brightens the entire scene (E_{flash}), i.e. in both the dark areas and bright areas. Hence, to capture the bright areas an integration time reduction was necessary though the additional radiance was sufficient to shift the image detail of the dark areas into the dynamic range of the camera. The following calculation illustrates our experimental results. Assuming a scene with a dynamic range of $E_{sun,min} = 0.44 \, mW/m^2$ and $E_{sun,max} = 50 \, W/m^2$, which was measured within a passenger car interior on a sunny summer day.

The aim of the car interior surveillance is to recognize, classify and track the occupants for reasons of safety and convenience. Such a vehicle interior represents a HDR environment due to the varying influences from the sun intensity and impact angle through the windows. Our measurements of the interior yield a raw optical dynamic range of up to $102 \, dB$ which is outside the capabilities of mainstream imagers. To capture the interior

(a) Short exposure   (b) Long exposure   (c) Double flash (proposed)
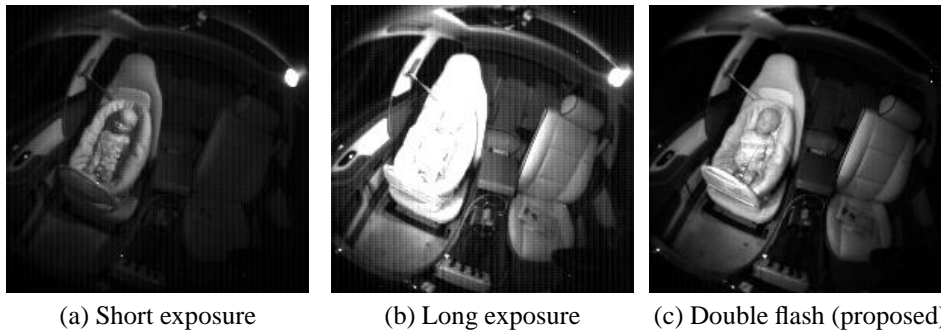
Figure 5: Car occupant detection via CMOS camera. Image (c) shows the result of the proposed double flash: Reduced dynamic, without ambient offset and FPN.

across a wide range of illumination situations without losing image detail, it was necessary to employ a linear HDR camera [1] with a dynamic range of $120 \, dB$. Fig. 1 and 2 were created with the data from such a camera. Our aim was to find a way of employing a mainstream imager with limited dynamic range to capture the vehicle interior without losing the interior details. We solved this problem by using the double flash approach as

proposed in section 4 (see Fig. 5). A supplementary offset of $E_{nir} = 50\,mW/m^2$, produced by a NIR illumination source, compresses the image dynamic to 60 dB (see eqn. (25)). This is equivalent to a maximum contrast of $1000:1$ and thus corresponds to a resolution of 10 bits which is within the range of mainstream CMOS imagers.

$$DR_{sun} = 20 \log \left(\frac{50\,\mathrm{W/m^2}}{0.44\mathrm{mW/m^2}} \right) = 102\mathrm{dB} \tag{25}$$

$$DR_{flash} = 20 \log \left(\frac{50\,\mathrm{W/m^2} + E_{nir}}{0.44\,\mathrm{mW/m^2} + E_{nir}} \right) = 60\mathrm{dB}$$

$$\rightarrow E_{nir} \approx 50\,\mathrm{mW/m^2} \tag{26}$$

To illustrate the double flash effect, we extend the calculation of eqn. (25). The final output image of the double flash should show a maximum dynamic of $DR_{df} = 48$ dB. This is equal to a maximum contrast of $255:1$, which can be displayed and stored with 8 bits. The maximum dynamic range of mainstream CMOS imagers was already specified for the first example with 60 dB and thus leads to a necessary supplementary illumination of $50 mW/m^2$. This is the minimum illumination necessary to reduce the dynamic range of the scene to the requirements of the given imager. Hence, we label it $E_{nir,low}$, because a more powerful local illumination $E_{nir,hi}$ merely yields a stronger dynamic compression and is therefore within the detection range of the imager.

$$20 \log (a) = DR_{df} = 48\,\mathrm{dB} \tag{27}$$

$$a = \left| \frac{50\,\mathrm{W/m^2} + E_{nir,hi}}{0.44\,\mathrm{mW/m^2} + E_{nir,hi}} - \frac{50\,\mathrm{W/m^2} + E_{nir,low}}{0.44\,\mathrm{mW/m^2} + E_{nir,low}} \right|$$

$$\mathrm{with}\ E_{nir,low} = 50\,\mathrm{mW/m^2}$$

$$\rightarrow E_{nir,hi} \approx 68\,\mathrm{mW/m^2}$$

A further advantage with the double flash approach arises if the CMOS imager suffers from fixed pattern noise (FPN) [5]. FPN is an individual offset from each pixel caused by slight unwanted variations of active pixels, i.e. transistor characteristics. To minimize the FPN of the camera, the captured data can be compared internally or externally to an offset-map for correcting the final output. This FPN correction can be performed by firmware and is thus relatively fast but still takes time.

Due to the fact that the individual FPN's represent an offset which is constant over time, it will be eliminated if two subsequent images are subtracted from each other. Hence, the output image sequence of a double flash system improves the final picture quality for imagers which suffer FPN and the internal or external FPN correction can be disabled. The results of the double flash is shown in Fig. 5 which shows the passenger and driver seat of a car. The CMOS camera was equipped with an optical NIR band pass as described in section 1 which cuts off wavelengths in the visible range. A strong halogen lamp (illuminating through the sun roof) simulates E_{sun} and causes a bright region on the infant seat. The CMOS camera was not able to capture details of the infant and the driver seat within the same image just by varying the exposure time or aperture. Furthermore strong FPN effects are visible which might distort a texture analysis of the scene. Fig. 5 (c) was acquired by using the double flash approach: Texture of both infant and driver seat are visible due to the compressed dynamic range. The influence of ambient illumination (the halogen lamp) was eliminated and the image quality was considerably improved due to the reduced FPN.

6 Summary

The irradiance $E_{sun,min}$ and $E_{sun,max}$ in a given high dynamic range scene is fixed and usually requires a HDR camera for capturing all the scene details. By using two supplementary flashes $E_{nir,low}$ and $E_{nir,hi}$, it is possible to influence the optical dynamic range of the scene so that a scene with high dynamic range (DR_{sun}) can be captured by commonly available cameras with limited optical dynamic range (DR_{df}) without sacrificing image detail and with synchronous offset reduction. The grey-level of a pixel within the output image sequence $I_{df}(\mathbf{p}, n)$ varies only if the scene changes and not as a result of fluctuations in the ambient illumination levels. This significantly simplifies any subsequent image processing.

An additional advantage referring to the otherwise necessary HDR cameras is that the supplementary illumination increases the brightness of the scenes. This enables the camera system in principle to run with a higher framerate because the integration time can be shortened due to the higher minimum irradiance level resulting in a greater signal-to-noise ratio (SNR). The major shortcoming of the double flash is that its application is limited to environments where the whole scene to be analyzed must be within the scope of the local illumination E_{nir}. The image pre-processing described in eqn. (16) is relatively simple and could be performed by firmware within the camera or integrated onto the sensor if CMOS technology is used. This yields an increase in the processing speed of the entire image processing system. Finally the costs for a mass produced camera with limited dynamic range in combination with simple trigger logic for the illumination are less than for a smart HDR camera.

References

[1] J. Huppertz, R. Hauschild, B.J. Hosticka, et al, *"Fast CMOS Imaging with High Dynamic Range"*, IEEE Workshop on CCD's and Advanced Image Sensors, 1997.

[2] S. Mann, R.W. Picard, *"On being 'undigital' with digital cameras: Extending Dynamic Range by Combining Differently Exposed Pictures"*, Society for Imaging Science and Technology, Annual Conf., May 1995.

[3] Int. Norm ISO 825-1, *"Maximum Permissable Exposure Limits"*, 1993.

[4] S-B. Park, A. Teuner, B.J. Hosticka, G. Triftshaeuser, *"An interior compartment protecting system based on motion detection using CMOS imagers"*, Proc. of Int. Conf. in Intelligent Vehicles, October 1998.

[5] B. Schneider, H. Fischer, S. Benthien, et al, *"TFA Image Sensors: From the One Transistor Cell to a Locally Adaptive High Dynamic Range Sensor"*, Technical Digest of International Electron Devices Meeting, December 1997.

[6] K. Toyama, J. Krumm, B. Brumitt, B. Meyers, *"Wallflower: Principles and practice of Background maintenance"*, Int. Conf. on Computer Vision, September 1999.

[7] D. Yang and A. El Gamal, *"Comparative analysis of SNR for image sensors with widened dynamic range"*, Proc. of SPIE, February 1999.