Modelization and Characterization of a CMOS Camera as an Optical Real-time Oscilloscope

G. Cossu, A. Sturniolo and E. Ciaramella ***Senior Member, IEEE***

Scuola Superiore Sant'Anna, Via Moruzzi 1, 56124 Pisa, Italy.

DOI: 10.1109/JPHOT.2009.XXXXXXX
1943-0655/$25.00 ©2009 IEEE

Abstract: Complementary Metal-Oxide Semiconductor (CMOS) camera sensors are embedded in many consumer electronics products: thanks to the Rolling Shutter (RS) readout mode, they can detect a time-varying light intensity, which is the key to realize Optical Camera Communication (OCC). To this aim, we introduce here a model describing the camera as a Real-Time Oscilloscope (RTO) detecting optical signals; by means of this approach, we can now characterize the CMOS camera by means of parameters that correspond to common oscilloscope specifications, such as the frequency response, the noise, the Signal to Noise Ratio (SNR), the total harmonic distortion (THD), etc.; all of these are introduced and measured in terms of the camera parameters. This approach provides for the first time a set of quantitative tools that should be used to maximize the OCC transmission performance by allowing the optimal selection of the camera settings.

Index Terms: Free-space optics, Optical Camera Communication, Wireless communication.

1. Introduction

THE Visible Light Communication (VLC) technology has shown a high potential, recently [1]–[8]. However, the absence of a widely accepted standard and the lack of integrated solutions, so far prevented its adoption in the mass market [2]. VLC exploits a common Light Emitting Diode (LED) at the transmitter side, but needs a dedicated Receiver (RX), i.e. a photodiode (PD), which is one of the main barriers that delays its commercialization.

To overcome this issue, Optical Camera Communication (OCC) was proposed: here the VLC RX is a Complementary Metal-Oxide Semiconductor (CMOS) camera, widespread in consumer electronics. Yet, a camera can sample the optical signal using a frame of video acquisition, but has a very limited frame rate. In OCC this was solved by exploiting the Rolling Shutter (RS) readout mode, which is embedded in most of the CMOS cameras and allows for much higher sampling rate [9].

In various OCC studies, many different challenges were investigated. As an example, blooming and source size-related issues in Directed Line-of-Sight (DLOS) configuration [10]–[13] or signal weakness and background impact in Non Line-of-Sight (NLOS) [14]–[17]. Single-color and RGB LEDs sources were used to implement Multiple Input Multiple Output (MIMO) architectures [18], [19] or Wavelength Division Multiplexing (WDM) systems [20]–[23]; similar systems were also demonstrated using screen-to-camera communication systems [24]–[26], which takes advantage of the detectors array of CMOS cameras and their ability to spatially divide multiple sources and detect separately RGB colors; RGB LEDs were also used to realize Color Shift Keying (CSK) systems using the camera as a single RX for this modulation format [27]–[29]; the spatial separation of the different sources was exploited to calculate the device relative position, thus providing accurate localization and navigation indoor [30]–[32].

All these works focused on the applications of OCC systems, with quite impressive results.

TABLE I: List of acronyms

ADC	Analog-to-Digital Converter
CMOS	Complementary Metal-Oxide Semiconductor
DCT	Discrete Cosine Transform
DLOS	Directed Line-of-Sight
ENOB	effective number of bits
FFT	Fast Fourier Transform
GS	Global Shutter
ISP	Image Sensor Process
LED	Light Emitting Diode
MIMO	Multiple Input Multiple Output
NAD	noise and distortion
NLOS	Non Line-of-Sight
OCC	Optical Camera Communication
PD	photodiode
RMS	root-mean-square
RS	Rolling Shutter
RTO	Real-Time Oscilloscope
RX	Receiver
SINAD	ratio of signal to noise and distortion
SNR	Signal to Noise Ratio
THD	total harmonic distortion
TI	Time Interleaving
VLC	Visible Light Communication

However, there is no methodology that gives the deep assessment of a camera as a communication device. In [33], [34], the authors made an attempt to describe it from the RX point of view, but did not show a clear relationship between the camera parameters and the received signal characteristics.

We introduce here a new approach to fully assess the impact of camera detection in a OCC system: we model the CMOS camera as an optical Real-Time Oscilloscope (RTO). This is a logical inversion of the analogy that describes a RTO as a camera taking a picture of the signal [35]. In this paper, we present a set of measurements that we apply to a CMOS camera and are based on the standard procedure to characterize a RTO [36]. This approach allows to characterize quantitatively the performance of a commercial CMOS cameras.

The paper is structured as follows. In Section 2 we describe the theoretical background of the CMOS camera image sensor and its use as an optical RX. In Section 3, we report the experimental setup used to validate our approach and the procedure for the signal detection. In Section 4, we present the detailed characterization of the CMOS camera as an optical RTO. Finally, in Section 5, we present our conclusion.

2. CMOS Camera fundamentals

2.1. Camera acquisition pipeline

In the following, we summarize the process resulting into a digital pixel value in a digital camera. The light is captured by a matrix of pixels. Each pixel has a PD that converts the energy of the detected light into a photocurrent and, finally, into voltage by the readout circuit. This is then electrically amplified (the gain is controlled by the ISO value, see Section 2.2). Then, Analog-to-Digital Converter (ADC) transforms the voltage signal into a set of 16 bits (raw data). To save storage, image compression algorithms (e.g. JPEG) are commonly used: in that case, the digital signals pass through an Image Sensor Process (ISP), where the number of bits representing each color is first reduced to 8 and a brightness correction is applied (γ encoding, detailed in Section 2.4). Finally, the image is further compressed by a JPEG algorithm [37].

2.2. Photo exposure

In photography, the exposure of a photo depends on the combination of aperture diameter, exposure time, and ISO value; they are usually combined by means of the concept of *stop*. One stop corresponds to halving (-1 stop) or doubling ($+1$ stop) the amount of light on the sensor [38]. The aperture diameter controls the area where light can pass through, usually indicated as the ratio of the aperture diameter and the focal length (e.g. *f*-2). Hence, ± 1 stop corresponds to multiplying or dividing by $\sqrt{2}$ the aperture diameter [38]. The exposure time t_{exp} is the lapse of time during which the pixels are exposed. The sensor works as an integration in time of the received light. Doubling (halving) t_{exp} produces an increase (decrease) of 1 stop [38]. The sensitivity to the light of the sensor is measured by the ISO value (ISO is not an acronym), which indicates the signal gain. As in the case of t_{exp}, doubling the ISO number corresponds to an increase of 1 stop (halving it gives a decrease of 1 stop) [38].

All the above parameters contribute to the exposure. Although we might in principle modify these variables while keeping constant the overall exposure, the resulting photos would not be exactly the same.

2.3. Rolling shutter

The vast majority of commercial CMOS sensors use the RS readout method, thanks to the lower cost compared to its alternative, i.e. the Global Shutter (GS) mechanism. In the RS readout, the pixels are sampled at high-speed, row-by-row, whilst in GS they are all acquired in one shot. A representation of the RS mode is shown in Fig. 1: each row is exposed for a time t_{exp}. The resulting photocurrent is then sampled and digitized in a readout time (t_{RO}). This quantity is a constant value and depends only on the camera manufactoring. As a consequence, provided that the sensor is quite uniformly illuminated and the modulation frequency is higher than $1/(Nt_{RO})$, a sinusoidal change in the intensity can result into a regular pattern of fringes in the image, from which the modulation frequency can be effectively obtained.

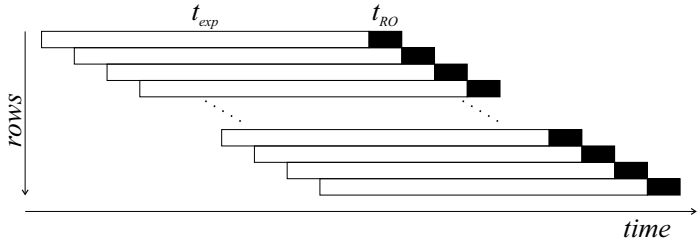


Fig. 1: Graphical representation of the rolling shutter readout method

The camera indeed works as an optical RX with a sliding integration window of t_{exp} duration, which moves row-wise with t_{RO} time steps. We model this process assuming a generic sinewave optical input $y(t) = P_0 + P_m \cos(2\pi f t)$; thus the corresponding current by any pixel $\tilde{y}$ at row k is

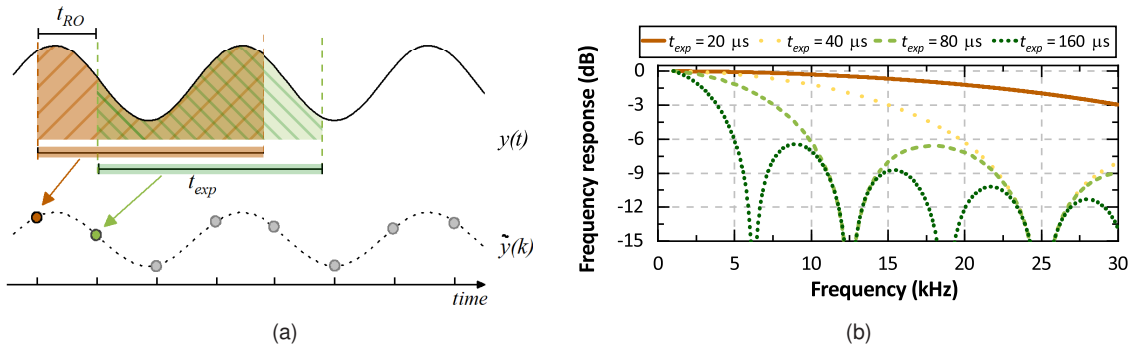
$$\tilde{y}(k) = \int_{t_k}^{t_k + t_{exp}} y(t)\, dt = P_0\, t_{exp} + P_m \frac{\sin\left(2\pi f (t_k + t_{exp})\right) - \sin(2\pi f t_k)}{2\pi f} \tag{1}$$

where $t_k = t_1 + (k-1)\, t_{RO}$ indicates the starting integration time of the k-th row. After the readout, each line is reset and can be exposed again. Hence, we derive that the equivalent sampling rate f_s of the digital signal is simply given by the inverse of the time between two consecutive readings:

$$f_s = 1/t_{RO} \tag{2}$$

The total frame time for an N-rows sensor is equal to $Nt_{RO} + t_{exp}$. Clearly, $1/(Nt_{RO})$ corresponds to the frequency resolution; it also represents the minimum frequency that can be effectively detected.

A graphical representation of the integration process is reported in Fig. 2a, where the integral value $\tilde{y}(k)$ is represented with the same color of the time window t_{exp}. Clearly, the $\tilde{y}(k)$ values strongly depend on f and t_{exp}, therefore in Fig. 2b we present the normalized frequency response of the signal $\tilde{y}(k)$ as a function of t_{exp}.



(a) (b)

Fig. 2: (a) Graphical representation of the integration process of the input optical waveform $y(t)$. $y(\tilde{k})$ indicate the integrated values within a time range of t_{exp} spaced by a t_{RO} time. (b) Theoretical frequency response of the RS readout mode for different t_{exp} values, (from Eq. (1))

As can be seen, these curves have zeros at the frequencies $f = m\frac{1}{t_{exp}}$, where m is a positive integer. As an example, we see that the curve with $t_{exp} = 160$ μs (dark green dotted curve in Fig. 2b) has the zeros at the multiples of 6.25 kHz. Based on the previous equation, the equivalent analog bandwidth should be around $0.5/t_{exp}$, as shown in Fig. 2b. However, some corrections to this value arise due to JPEG compression (as shown later).

2.4. JPEG Image sensor process

JPEG compression is commonly used in smartphone cameras: it removes the information to which the human eye is insensitive and also exploits the Discrete Cosine Transform (DCT)-based image compression [39]. The raw output value r from each pixel is directly proportional to the received energy. Since the response of human eyes is nonlinear [40], a process named "γ encoding" modifies the raw values from the pixels to values that are closer to the human perception, according to

$$255 \times \left(\frac{r}{r_{max}}\right)^{\frac{1}{\gamma}} \tag{3}$$

where r_{max} is the maximum possible output value and the common value for γ is 2.2 [40]. This reduces the number of bits down to 8, although it introduces a minor distortion.

Furthermore, the JPEG compression removes the high frequency components of the picture (the human eye has poor sensitivity to fast brightness variations). Here, we briefly summarize the algorithm [37], [39], [41]. First, the image is divided into small square blocks. To each of them, the 2D-DCT is computed, giving $F(u, v)$ and then F is divided by the quantization matrix [42], which has a higher value at higher frequencies, hence strongly reducing (or even removing) the amplitudes at these frequencies. Finally, all the frequency amplitudes lower than a threshold value are set to zero, removing from the image high frequency DCT components, as a low-pass filter. The higher is the number of zeros, the higher is the image compression.

3. Setup and experimental characterization

3.1. Experimental setup

In this section, we present the experimental setup and the preliminary characterization of the CMOS sensor. The experimental setup is illustrated in Fig. 3. The optical source was an array

of 7 LEDs generating white light. The LEDs were driven in current by a sine-wave at frequency f_{in}, generated by a waveform generator In the considered range of frequencies (< 50 kHz), the generated sine-wave can be considered as ideal signal. The amplitude and the bias of the sine-wave were chosen so that the LED was modulated with around 90% modulation index in the linear region (10-300 mA) using a home-made driver.

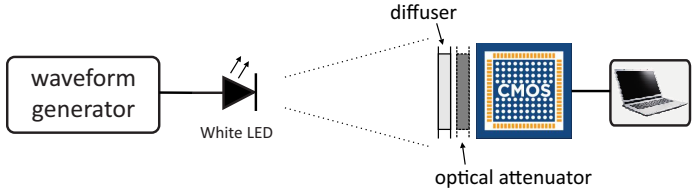


Fig. 3: Scheme of the experimental setup for the CMOS sensor characterization

The CMOS sensor was the front camera of a commercial smartphone (*Nexus 5X*, [43]). The optical signal was sampled and stored by the camera and analyzed by MATLAB scripts running on a PC. A diffuser blurred the image and spread quite uniformly the light throughout the sensor. This maximized the number of illuminated rows and reduced the effects of the source shape. We added a variable optical attenuator with a known attenuation, to record signals at different optical powers (with step of 1-stop) without changing the exposure parameters. The distance between the LED and the RX was 10 cm, to get enough optical power even with the low bias current. In this sensor, we can control only t_{exp} and ISO; the aperture is fixed at *f*-2.2. All photos were taken at the maximum resolution available (2592×1944 pixels, $N_c \times N_r$).

3.2. Signal reconstruction

In Fig. 4, we report an example of the RS effect when the optical source is modulated with f_{in} =1 kHz. By summing (or averaging) the resulting pattern of fringes along the RS axis, we generate the Y-profile. Fig. 5 presents the result of the signal extraction by summing along the X-axis the exposure values (left plot) and its Fast Fourier Transform (FFT) (right plot). In this example, we had ISO $= 50$ and $t_{exp} = 20\ \mu s$. From the profile curve, the modulating frequency can be quickly recovered, although a slow overmodulation arises due to the lamp spatial shape, which adds low frequency components.

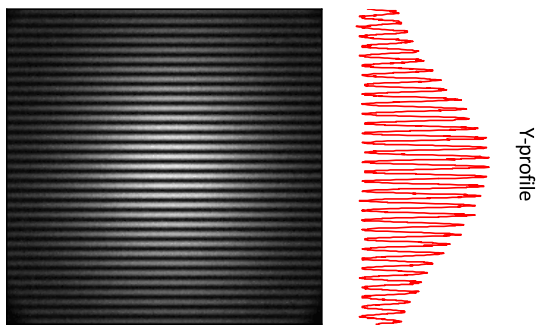


Fig. 4: Example of a rolling shutter effect when the optical source is modulated with 1 kHz sinewave. The figure reports also Y profile of the photo.

Analyzing the frequency domain, we see the main tone at 1 kHz. Yet, other spurious frequencies are present. Some of them are harmonics of the input signal: they are nonlinear processing due to the γ encoding process. Others spurious tones are due to the Time Interleaving (TI): TI is well known in RTO as it allows to acquire data at a rate faster than the operating sample rate of the individual data converters [44]. Yet, it creates significant spurious tones at $f_s/4 \pm f_{in}$ and $f_s/2 \pm f_{in}$,

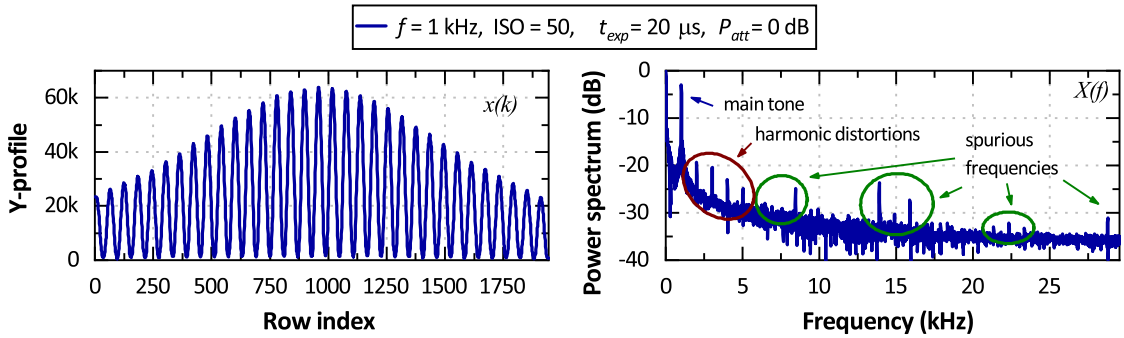


Fig. 5: Example of the reconstructed signal in time (left) and frequency (right) domain.

where f_{in} is the input frequency [45] and also smaller spurious at $f_s/8 \pm f_{in}$ and $3/8 f_s \pm f_{in}$. This effect is apparent in Fig. 5, where all the above spurious tones are noticeable.

4. Experimental characterization

4.1. Read-out time and sampling rate estimation

As a first measurement, we estimated the sampling rate f_s of the CMOS camera using Eq. (2). The readout time t_{RO} can be estimated from an acquisition of the Y-profile when the optical source is modulated by a tone at f_{in}. The frequency separation of sample frequencies on the FFT is $df = 1/T$, where $T = t_{RO} N_\mathrm{r}$ is the acquisition time and where N_r is the waveform length. df can be also measured by dividing f_{in} by its frequency index f_{idx} in the FFT. By equaling these two expression we obtain a relationship between t_{RO} and the other measured quantities:

$$t_{RO} = \frac{f_{idx}}{f_{in} \cdot N_\mathrm{r}} \tag{4}$$

In order to accurately estimate this value, we took a photo at several different f_{in} values (from 1 to 31 kHz). From the FFT of the Y-profiles, we derived the indexes (f_{idx}) and we report them as a function of the nominal value of f_{in} in Fig. 6.

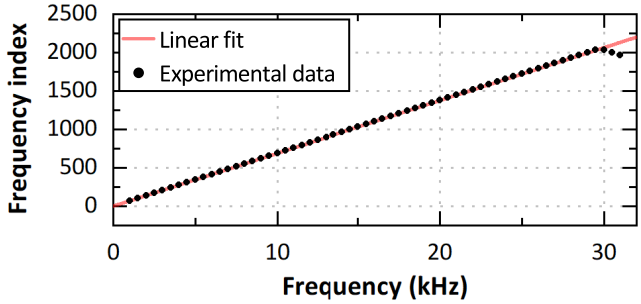


Fig. 6: Experimental measurement of FFT frequency index as a function of the modulating frequency. Red curve shows the linear fit on the experimental data. The inverse of the slope of this curve gives the value of df. The curve was taken at P_{att} =0 dB, $t_{exp} = 20$ μs and ISO = 50.

From the inverse of the slope coefficient, we obtain df =14.54 Hz. Thus, using Eq. (4) and $N_\mathrm{r} = 1944$, we obtain $t_{RO} = 16.8$ μs. Therefore, the equivalent sampling rate of our CMOS camera is $f_s = 59.2$ kSa/s. A countercheck of this value can be also spotted in Fig. 6: here, the last points ($f_{in} > f_s$) folded back at a lower frequency index, due to the aliasing effect. From these values, we also assess the maximum record length (T_{max}) of the camera under test:

$$T_{max} = t_{RO} N_\mathrm{r} \tag{5}$$

In our case $N_r = 1944$, then the maximum record length is 32.6 ms).

4.2. Signal characterization

Here we analyze the CMOS sensor performing a set of measurements equivalent to those commonly used for the RTOs. The key of a good oscilloscope is its ability to reconstruct a waveform with high signal integrity. The IEEE 1057-2017 standard defines the test methods used for assessing the performance of digital oscilloscopes [36]. All our tests were performed using a sine-wave input signal, as it is indicated by the above standard (very accurate sine-waves are available and it is easy to distinguish the errors from the input in the frequency domain). The tests consisted in transmitting a set of sine-waves from 1 to 30 kHz (with a step of 250 Hz), varying the different parameters of the RX (P_{att}, t_{exp} and ISO), with 1-stop step.

For a RTO, the key parameters to determine the signal integrity are the bandwidth, the Signal to Noise Ratio (SNR), ratio of signal to noise and distortion (SINAD), the effective number of bits (ENOB), the total harmonic distortion (THD) and the noise contribution [36].

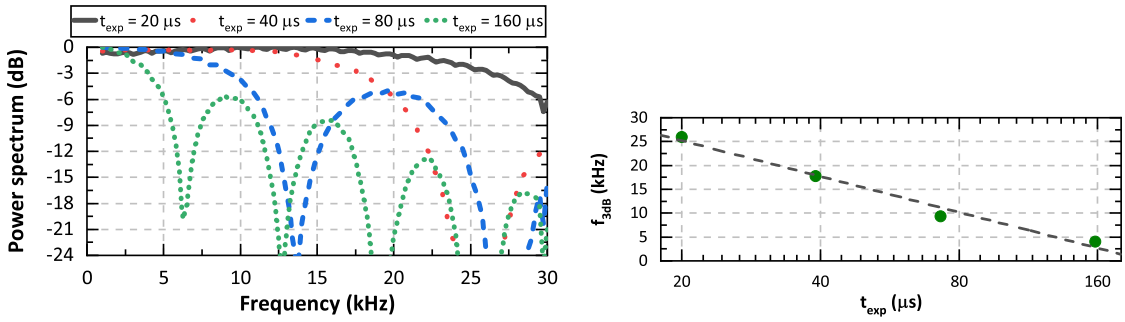
In the following paragraphs, we first summarize the definitions of those parameters and the methods used to measure them. We define $x(t_k)$ as the sampled waveform in the discrete time domain and $X(f_k)$ as the complex FFT at frequency f_k of the sampled waveform. M indicates the number of samples in the data record and A_{rms} is root-mean-square (RMS) value of the signal amplitude at f_{in} and can be obtained as:

$$A_{rms} = \frac{1}{M}\sqrt{|X(f_{in})|^2 + |X(f_s - f_{in})|^2} \tag{6}$$

4.3. Bandwidth

The first relevant parameter is the bandwidth. In RTOs, the bandwidth is defined as the frequency where the power is -3 dB lower than at the specified reference frequency. To measure the equivalent bandwidth of our camera, we analyzed the Y-profiles in the frequency domain and took the amplitude of the generated tones at different f_{in}. We acquired the curves as a function P_{att}, t_{exp} and ISO. Here, we present only the most representative curves.

In Fig. 7a, we report the measured frequency response at different values of t_{exp}, with a 1-stop step, compensated by increasing the ISO value accordingly, keeping constant the overall amount of light . As expected, the highest bandwidth value was found at the shortest exposure time ($t_{exp} = 20$ μs); in this case, the measured bandwidth was about 26 kHz. At higher t_{exp}, the bandwidth decreases. A summary of the 3-dB bandwidth of the CMOS camera as a function of exposure time is reported in Fig. 7b.



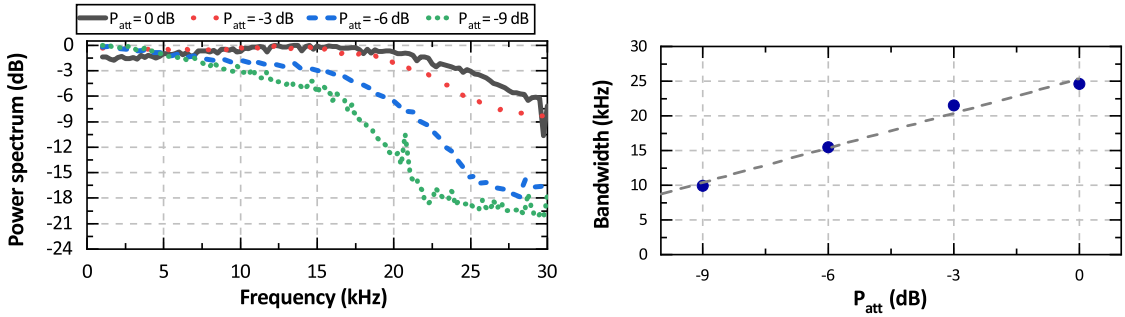
(a) (b)

Fig. 7: (a) Frequency response of the camera at different values of t_{exp} and $P_{att} = 0$ dB. (the ISO value was changed to compensate the stop: 400, 200, 100 and 50, respectively); (b) bandwidth of the camera at different values of t_{exp}.

The shape of the curves in Fig. 7a is quite well modeled by Eq. (1), but we see that the zeros are not exactly at the expected frequencies. Likely, this small deviation is due to a difference

between the nominal t_{exp} set on the camera and the actual value. Therefore, we can estimate the real exposure time from the inverse values of the first zero: then, we find t_{exp} values of 39 ± 1 μs, 73 ± 1 μs and 158 ± 6 μs, instead of the nominal values of $40, 80$, 160 μs, respectively (the curve at 20 μs has not zero to measure). All of these are quite consistent with the set values, but for 80 μs, which shows a deviation of around 10%.

A major difference from common RTOs, where the bandwidth is fixed, is that in the optical camera the bandwidth strongly depends on the average intensity of the incident light: this is due to the ISP. As explained in Section 2.4, the ISP for the JPEG format applies the quantization matrix that works as a low-pass filter and is particularly effective for low illumination, in which the higher frequencies are strongly attenuated. This is confirmed by the experimental data reported in Fig. 8a. Here, the bandwidth grows from around 10 kHz to 25 kHz when the attenuation decreases from -9 to 0 dB. We expect that this bandwidth variation would not be seen when analyzing directly the raw data. Unfortunately, this cannot be proven since that the raw data are not available in most of the smartphone cameras. A summary of the camera bandwidth as a function of the optical attenuation is reported in Fig. 8b.



(a) (b)

Fig. 8: (a) Frequency response of the CMOS camera at different values of P_{att}. ISO= 50 and $t_{exp} = 20$ μs. (b) 3dB bandwidth of the CMOS camera at different values of P_{att}.

4.4. Noise floor and distortion

As in a RTO, we measure the noise as the total power in the spectrum uncorrelated with the input signal, its harmonics and the spurious tones generated by the TI (see Section 4.2). Different effects contribute to the noise in a CMOS camera [46], namely the dark noise, the readout noise and the photon shot noise. The dark noise is generated by the current that flows even when no photon is incident. It is a thermal phenomenon. The readout noise is generated when producing the electronic signal: it results from the sensor design, but is also influenced by the camera electronics. Thus, it is independent on the signal level and the sensor temperature. The shot noise is due to Poisson statistics of the photodetection and depends only on the average number of photons. In the noise measurements, the camera was placed in the dark (covered by a black surface). Hence, only dark noise and read noise were present.

In Fig. 9, we report the intensity noise along the rows at $t_{exp} = 20$ μs, in time and frequency domain. The noise along the row is reported as the average of the pixel intensity. We report a single curve since, for very short t_{exp} (such as the ones considered in this work), the noise does not change significantly. In this measurement, the dominant noise is the read noise, as can be seen from the quite higher noise intensity at the first and last rows of the sensor in the upper figure. We measured that the dark noise is dominant over to the read noise at $t_{exp} > 100$ ms.

Different behavior is presented in Fig. 10, where the noise curves are reported for various ISO values. In this case, the read noise is amplified by the ISO gain. We note that the reported noise curves show a nonlinear gain with the ISO values. This can be related to a threshold effect, as in the first three cases the average intensity noise per pixel is lower than 1.

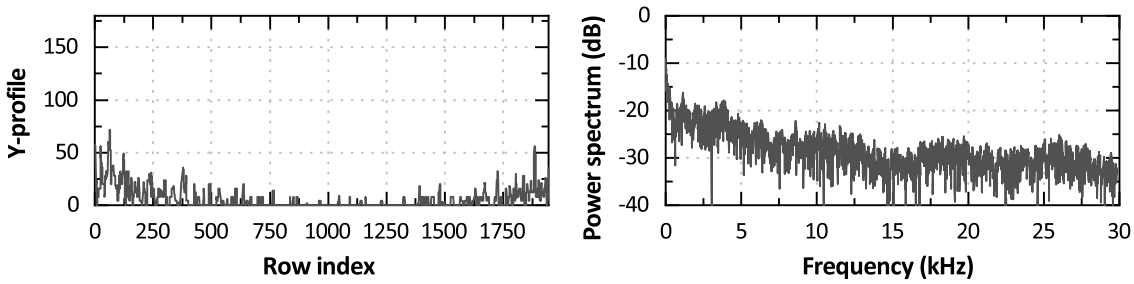


Fig. 9: Measured noise at $t_{exp} = 20\mu$s. We present both time (left) and frequency (right) domain curves.

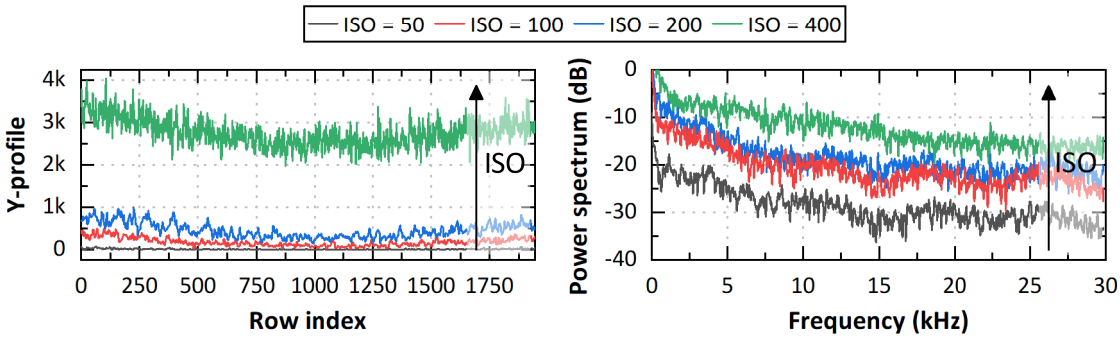


Fig. 10: Noise as a function the ISO value. We report both time (left) and frequency (right) domain curves. The noise along the row is reported as the average of the pixel intensity.

Another source of errors are the non-linear distortions, which depend on the input power. Among them, the most common is the harmonic distortion, which increases with the input power and generates new spectral components at integer multiples of the input frequency. As an example, the JPEG ISP introduces signal distortions: as explained in Section 2.4, the γ correction gives non-linearities on the encoded data, applying the γ-root to the input [40].

In Fig. 11, we report the measured pixel intensity (black dots) as a function of the LED current, which is proportional to the received optical power. Here, we also report the fitting curve with the γ profile (solid red line). The experimental data and the fitting curve overlap perfectly. This proves that the observed nonlinear trend is correctly ascribed to the γ correction. As expected, the harmonics grow and become relevant at higher signal powers.

To quantify the distortion, we measured the THD, which is defined as the scaled square-root of the sum of squares of a specific set of harmonic distortion components (including their aliases) for a pure sinewave input of specific amplitude and frequency. The THD value is given by [36]:

$$\text{THD} = \frac{\sqrt{\frac{1}{M^2}\sum_h |X(f_h)|^2}}{A_{rms}} \tag{7}$$

where f_h represent the frequencies of the harmonics and A_{rms} is defined in Eq. (6). As an example, we report in Fig. 12a the THD as a function of ISO and t_{exp} values, which are parameters that increase the amplitude of the acquired signals. We see that the harmonics grow by almost 15 dB, with an increase of the total exposure by 3-stops, in both cases.

4.5. Signal integrity: SNR, SINAD and ENOB

In order to quantitatively determine the effectiveness of a RTO, we usually measure the parameters called *figures of merit*. The common figures of merit are SNR, SINAD and ENOB, which all depend on frequency.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JPHOT.2020.3032951, IEEE Photonics Journal

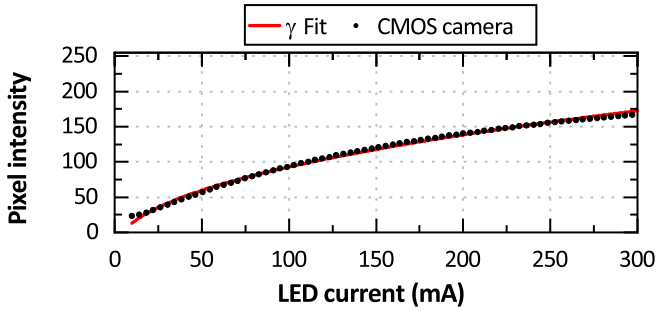


Fig. 11: Experimental pixel intensity (black dots) as a function of the LED current. The solid red line reports the fit with the curve equation Eq. (3).

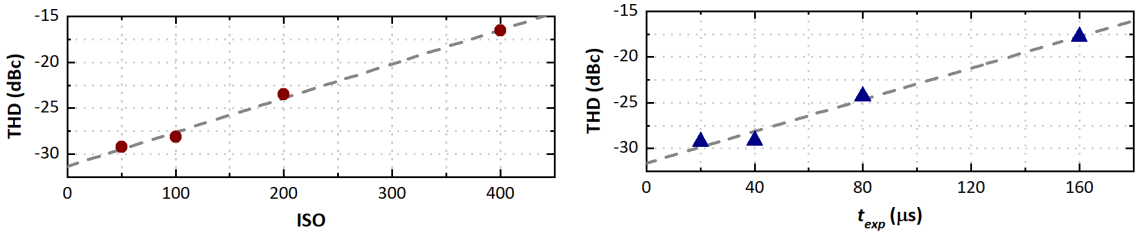


Fig. 12: (a) Total harmonic distortion as a function of the ISO (top) and t_{exp}. (b) The harmonic distortions grow with a larger exposure due to the γ correction.

SNR is defined as the ratio of a signal to the noise power, where the noise is a deviation between the output and the input signal except for deviations caused by linear time invariant system response, dc level shift, THD or an errors in the sample rate:

$$\text{SNR} = \frac{A_{rms}}{\eta} \tag{8}$$

where η indicates the RMS noise defined as $\eta = \sqrt{\text{NAD}^2 - A_{rms}^2 \text{THD}^2}$, where noise and distortion (NAD) is given by:

$$\text{NAD} = \frac{1}{\sqrt{M(M-3)}} \sqrt{\sum_{m \in S_0} X_{avg}^2(m)} \tag{9}$$

The SNR curves of our camera are shown in Fig. 13a and the maximum value reaches almost 40 dB. The gap between the SNR curves is only due to the noise amplified by the ISO gain.

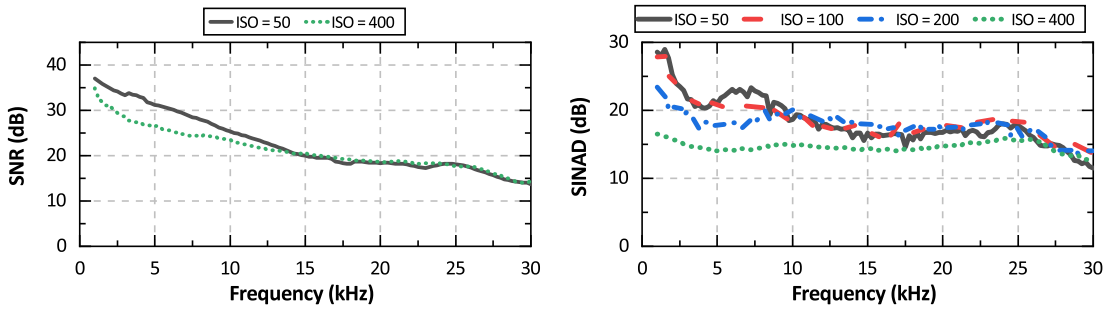
If we want to account for both the noise and the THD (mostly generated by the ISP), we should consider the SINAD parameter, which is often used in communication. SINAD measures the overall quality of the output signal and is defined as the ratio of A_{rms} and the NAD [36]:

$$\text{SINAD} = \frac{A_{rms}}{\text{NAD}} \tag{10}$$

S_0 is the set of all integers between 1 and $M-1$, excluding the two values that correspond to the fundamental frequency and the zero-frequency term.

In Fig. 13b, we report the measured SINAD as a function of the ISO value. As expected, the best performance are found at lower ISO values and at low frequency, where the SINAD value is higher than 25 dB. When ISO is higher, the THD increases (due to the non-linear γ-correction), therefore SINAD reduces. At high frequencies (> 25 kHz), the gap between the SINAD curves reduced down to < 2 dB at different ISO values.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JPHOT.2020.3032951, IEEE Photonics Journal



(a) (b)

Fig. 13: (a) Measured SNR curves at different ISO values; (b) Measured SINAD curves at different ISO values

The ENOB corresponds to the dynamic performance primarily associated with the quantization levels. Although the ADC of the sensor has at least 10 bits resolution (raw data, [43]), the front-end noise dramatically reduces the ENOB of the entire system. Furthermore, the ENOB value is frequency dependent; as for the other measurements, we estimated it in the frequency range of the recorder. The ENOB is defined as the number of bits of an ideal waveform recorded for which the RMS quantization error is equal to the NAD of the waveform recorded under test [36]:

$$\begin{aligned} \text{ENOB} &= \log_2\left(\frac{\text{FSR}}{\text{NAD}\sqrt{12}}\right) \\ &= \frac{\text{SINAD} - 1.76 + 20\log(\text{FSR}/A)}{6.02} \end{aligned} \tag{11}$$

where FSR is the Full Scale Range of the camera and A the input signal amplitude.

In our case, the most significant ENOB measurement is reported in Fig. 14, where we present the measured curves at different ISO values. As expected, these values cannot be higher than the real number of the ADC bits (8 bits after JPEG compression). Clearly, being the ENOB strongly dependent on NAD, the curves are heavily impacted by the ISO value. The average value of the ENOB, in the considered frequency range, decreased from 5.5 bits at ISO $= 50$ (with a maximum value of 7.5 at $f = 1.5$ kHz) to an average of 2.7 bits at ISO $= 400$. This is consistent with previous results, since at ISO = 400 the noise spectral density values are much higher.

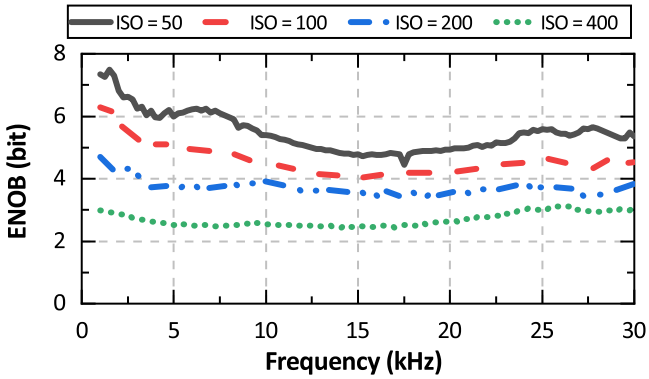


Fig. 14: Measured ENOB curves at different ISO values

5. Conclusion

We introduced for the first time a quantitative method to assess the CMOS camera sensors modelled as a RTO for optical signals. It can be used for high-bandwidth OCC applications. We proposed an effective method to derive the sampling rate of the camera, then we introduced the equivalent parameters used to characterize an RTO and we measured them on a CMOS camera. We also described the effects of the JPEG conversion on the signal integrity. These detailed characterizations are a novel tool that allows to maximize the performance of the camera in OCC applications.

The measurements were performed on the real camera of a commercial smartphone, as a function of the received signal intensity (controlled attenuation in front of the camera) and of all the parameters controllable by the camera (exposure time and ISO value).

Further investigations might be performed to have a complete control of the optical recorder, such as the saturation characterization, the blooming effect, and a complete analysis of the RAW image format.

References

[1] M. Z. Chowdhury *et al.*, "A Comparative Survey of Optical Wireless Technologies: Architectures and Applications," *IEEE Access*, vol. 6, pp. 9819–9840, 2018.
[2] V. Jungnickel *et al.*, "A european view on the next generation optical wireless communication standard," in *2015 IEEE Conference on Standards for Communications and Networking (CSCN)*, Oct 2015, pp. 106–111.
[3] G. Cossu *et al.*, "Gigabit-class optical wireless communication system at indoor distances (1.5 - 4 m)," *Optics Express*, vol. 23, no. 12, p. 15700, jun 2015.
[4] N. Saeed *et al.*, "Underwater Optical Wireless Communications, Networking, and Localization: A Survey," feb 2018.
[5] J. Luo, L. Fan, and H. Li, "Indoor Positioning Systems Based on Visible Light Communication: State of the Art," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2871–2893, 2017.
[6] G. Cossu *et al.*, "Sea-Trial of Optical Ethernet Modems for Underwater Wireless Communications," *Journal of Lightwave Technology*, vol. 36, no. 23, pp. 5371–5380, dec 2018.
[7] L. Feng *et al.*, "Applying VLC in 5G networks: Architectures and key technologies," *IEEE Network*, vol. 30, no. 6, pp. 77–83, 2016.
[8] G. Cossu *et al.*, "3.4 Gbit/s visible optical wireless transmission based on RGB LED," *Optics Express*, vol. 20, no. 26, p. B501, dec 2012.
[9] C. Danakis *et al.*, "Using a CMOS camera sensor for visible light communication," in *2012 IEEE Globecom Workshops*. IEEE, dec 2012, pp. 1244–1248.
[10] J. Hao, Y. Yang, and J. Luo, "CeilingCast: Energy efficient and location-bound broadcast through LED-camera communication," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*. IEEE, apr 2016, pp. 1–9.
[11] Y. Yang, J. Hao, and J. Luo, "CeilingTalk: Lightweight Indoor Broadcast Through LED-Camera Communication," *IEEE Transactions on Mobile Computing*, vol. 16, no. 12, pp. 3308–3319, dec 2017.
[12] J. H. Bae, N. T. Le, and J. T. Kim, "Smartphone Image Receiver Architecture for Optical Camera Communication," *Wireless Personal Communications*, vol. 93, no. 4, pp. 1043–1066, apr 2017.
[13] D. T. Nguyen, Y. Chae, and Y. Park, "Enhancement of Data Rate and Packet Size in Image Sensor Communications by Employing Constant Power 4-PAM," *IEEE Access*, vol. 6, pp. 8000–8010, 2018.
[14] F. Yang *et al.*, "Spatial Multiplexing for Non-Line-of-Sight Light-to-Camera Communications," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2018.
[15] Z. Zhang *et al.*, "Mobile phone Camera Based Visible Light Communication Using Non-Line-of-Sight (NLOS) Link," in *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*. IEEE, aug 2018, pp. 35–39.
[16] J.-K. Lain, F.-C. Jhan, and Z.-D. Yang, "Non-Line-of-Sight Optical Camera Communication in a Heterogeneous Reflective Background," *IEEE Photonics Journal*, vol. 11, no. 1, pp. 1–8, feb 2019.
[17] W.-C. Wang *et al.*, "Long distance non-line-of-sight (NLOS) visible light signal detection based on rolling-shutter-patterning of mobile-phone camera," *Optics Express*, vol. 25, no. 9, p. 10103, may 2017.
[18] N.-T. Le and Y. M. Jang, "MIMO Architecture for Optical Camera Communications," *The Journal of Korean Institute of Communications and Information Sciences*, vol. 42, no. 1, pp. 8–13, 2017.
[19] A. Ashok *et al.*, "Challenge: Mobile optical networks through visual MIMO," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking - MobiCom '10*. New York, New York, USA: ACM Press, 2010, p. 105.
[20] E. Wengrowski *et al.*, "Optimal radiometric calibration for camera-display communication," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, mar 2016, pp. 1–10.
[21] A. Krohn, P. A. Hoeher, and S. Pachnicke, "Visible Light Tricolor LED-to-Camera Data Transmission Suitable for Internet-of-Things and Sensor Applications," in *2018 European Conference on Optical Communication (ECOC)*. IEEE, sep 2018, pp. 1–3.

[22] K. Liang, C.-W. Chow, and Y. Liu, "RGB visible light communication using mobile-phone camera and multi-input multi-output," *Optics Express*, vol. 24, no. 9, p. 9383, may 2016.
[23] P. Luo *et al.*, "Experimental Demonstration of RGB LED-Based Optical Camera Communications," *IEEE Photonics Journal*, vol. 7, no. 5, pp. 1–12, oct 2015.
[24] V. Nguyen *et al.*, "High-rate flicker-free screen-camera communication with spatially adaptive embedding," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*. IEEE, apr 2016, pp. 1–9.
[25] A. Wang *et al.*, "Inframe++: Achieve simultaneous screen-human viewing and hidden screen-camera communication," in *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '15*. New York, New York, USA: ACM Press, 2015, pp. 181–195.
[26] N.-T. Le, T. Le, and Y. M. Jang, "Optical camera communications based invisible watermarking technique," in *2016 International Conference on Information Networking (ICOIN)*. IEEE, jan 2016, pp. 479–481.
[27] L. Penubaku and LakshmiShree. K, "An attempt to transfer information using light as a medium and camera as receiver," in *2015 International Conference on Computing and Network Communications (CoCoNet)*. IEEE, dec 2015, pp. 964–968.
[28] P. Hu *et al.*, "Colorbars: Increasing data rate of LED-to-camera communication using color shift keying," in *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies - CoNEXT '15*. New York, New York, USA: ACM Press, 2015, pp. 1–13.
[29] S.-H. Chen and C.-W. Chow, "Color-Shift Keying and Code-Division Multiple-Access Transmission for RGB-LED Visible Light Communications Using Mobile Phone Camera," *IEEE Photonics Journal*, vol. 6, no. 6, pp. 1–6, dec 2014.
[30] N. Rajagopal, P. Lazik, and A. Rowe, "Visual light landmarks for mobile devices," in *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*. IEEE, apr 2014, pp. 249–260.
[31] X. Wang, H. Yang, and J. Song, "A VLC-based robust ID broadcasting indoor location system for mobile devices," in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, jun 2016, pp. 1–5.
[32] H. Zhang and C. Yang, "Efficient coding and detection of ultra-long IDs for visible light positioning systems," *Optics Express*, vol. 26, no. 10, p. 13397, may 2018.
[33] J. C. Chau and T. D. Little, "Analysis of CMOS active pixel sensors as linear shift-invariant receivers," in *2015 IEEE International Conference on Communication Workshop (ICCW)*. IEEE, jun 2015, pp. 1398–1403.
[34] Z. Wang *et al.*, "Optical Camera Communication: Fundamentals," in *Visible Light Communications*. Hoboken, NJ, USA: John Wiley & Sons, Inc., dec 2017, pp. 239–290.
[35] Boon Campbell, "Real-Time or Sampling Oscilloscope | Keysight Community," 2017.
[36] *IEEE Std 1057-2017 (Revision of IEEE Std 1057-2007) : IEEE Standard for Digitizing Waveform Recorders*. IEEE, 2018.
[37] M. Shaban Al-Ani and F. H. Awad, "THE JPEG IMAGE COMPRESSION ALGORITHM," *International Journal of Advances in Engineering & Technology*, vol. 6, no. 3, pp. 1055–1062, 2013.
[38] E. Anderson, *Mastering Digital Cameras: Understanding Aperture, Shutter Speed, ISO and Exposure Time for Capturing Amazing Shots*, A. Books;, Ed. Independently Published, 2020.
[39] R. A.M *et al.*, "Jpeg Image Compression Using Discrete Cosine Transform - A Survey," *International Journal of Computer Science & Engineering Survey*, vol. 5, no. 2, pp. 39–47, may 2014.
[40] C. Poynton, I. Books24x7, and E. I. Inc, *Digital Video and HD: Algorithms and Interfaces*, ser. Electronics & Electrical. Elsevier Science, 2003.
[41] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.
[42] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Human Vision, Visual Processing, and Digital Display IV*, J. P. Allebach and B. E. Rogowitz, Eds., vol. 1913. International Society for Optics and Photonics, sep 2005, pp. 202–216.
[43] OmniVision, "OV5693 High-Performance 5-Megapixel Image Sensor for Front-Facing Cameras in Flagship Smartphones and Tablets," Tech. Rep., 2015.
[44] D. Durini, "High performance silicon imaging - Fundamentals and Applications of CMOS and CCD Sensors."
[45] J. Elbornsson, F. Gustafsson, and J. . Eklund, "Analysis of mismatch effects in a randomly interleaved a/d converter system," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 3, pp. 465–476, 2005.
[46] C. Draman *et al.*, "CCD or CMOS camera noise characterisation," *The European Physical Journal Applied Physics*, vol. 21, no. 1, pp. 75–80, jan 2003.