Understanding image sensors, noise, non-uniformities and obtaining a likelihood ratio in the process of device identification

Wiger van Houten

5614856

⟨w.vanhouten@student.uva.nl⟩
or
⟨waw.van.houten@gmail.com⟩

Thesis submitted to the University of Amsterdam
for partial fulfilment of the MSc. degree in Forensic Science

Report of internship at the Netherlands Forensic Institute
Credits: 36 EC
January - July, 2008
Supervisor: dr. Zeno Geradts
Examiner: dr. Cor Veenman

Wiger van Houten

August 13, 2008

Signed, dr. Cor Veenman (Examiner, UvA, NFI)

Signed, dr. Zeno Geradts (Supervisor, NFI)

I heard of a man who says words so beautifully
that if he only speaks their name
women give themselves to him.
If I am dumb beside your body
while silence blossoms like tumors on our lips,
it is because I hear
a man climb the stairs
and clear his throat
outside our door.
– *Leonard Cohen*

Abstract

IT IS KNOWN that the Photo Response Non-Uniformity (PRNU) pattern is a useful characteristic in helping to determine from which specific camera an image originates. In order to understand which sources are responsible for the existence of this pattern, and which sources obfuscate this pattern, we have studied various (systematic) noise sources present in both CCD and CMOS image sensors. Although it is generally difficult to quantify these sources, it is still advantageous in a forensic context to understand the causes of these patterns to justify the legitimacy of this approach. When possible, experiments were conducted to illustrate the theory. It was found that the photon shot noise is the largest noise source present at normal illuminations, and read noise is the dominant source at low illuminations. At high intensities the multiplicative PRNU dominates. PRNU in CCD image sensors is mainly due to non-uniform photoactive areas and material properties of the silicon. In CMOS image sensors additional sources from the individual transistors present on each pixel add to the non-uniform response. The problem of finding a likelihood ratio based on the similarity of two patterns is addressed. Although a way is found to calculate the likelihood ratio, some important questions have to be answered before the validity of the presented method can be guaranteed. The obtained likelihood ratios are small. Alternative approaches in solving this problem which are valuable starting points for future research are also introduced.

Acknowledgements

First and foremost I am grateful for the help received from my both my supervisor, dr. Zeno Geradts, and my examiner dr. Cor Veenman.
On a personal note, I wish to thank the people I love for being there, and the people who continue to support me. Each letter is dedicated to their patience, love and support at the time when it's needed.

Contents

1 Introduction **1**

2 CCD **4**

2.1 Summary 4

2.2 A closer look 4

2.3 Noise 8

3 CMOS **10**

3.1 Summary 10

3.2 A closer look 10

3.2.1 Readout of the photodiode voltage 14

3.2.2 Per pixel transistor noise 15

3.2.3 Noise from the photodiode 16

4 Noise **20**

4.1 Summary 20

4.2 Introduction 21

4.2.1 Statistics 21

4.3 Experimental conditions 22

4.4 Shot noise 24

4.5 Dark current 25

4.6 Fixed Pattern Noise 27

4.7 A/D converters and correlated double sampling 29

4.8 Quantization noise 31

4.9 PRNU 31

4.10 Read Noise 33

4.11 Flicker noise 35

4.12 Noise sources and failure common to CCDs and CMOS APS 36

4.13 Residual noise 37

4.14 Photon Transfer Curve . . . 38
4.14.1 Model . . . 41
4.15 Finding the size of the gain and read noise from images . . . 42
4.16 Periodogram . . . 43
4.17 Image noise . . . 44

5 General concepts **46**
5.1 Summary . . . 46
5.2 Quantum efficiency and flat-fielding . . . 47
5.2.1 Summary . . . 47
5.2.2 A closer look . . . 47
5.3 Colour Filter Array, colour gain, gamma correction . . . 49
5.4 Unicity of patterns . . . 53
5.5 Performance and reliability of the PRNU estimation . . . 54
5.5.1 Forgery . . . 57

6 Likelihood Ratio **58**
6.1 Summary . . . 58
6.2 Introduction . . . 58
6.3 General approach . . . 61
6.4 Estimators . . . 63
6.5 Performance of the estimator . . . 65
6.6 In practice . . . 69
6.7 Alternative approaches . . . 72
6.7.1 As detection problem . . . 72
6.7.2 Per pixel probabilities . . . 73

7 Conclusion **75**

8 Recommendations **77**

A Nonlinearity **79**

B Bandgap **81**

C Photon statistics **83**

D Colour reproduction **85**

E Estimating the PRNU from frames **87**

F Confusion matrix **89**

References **92**

CHAPTER 1

Introduction

Due to the integration of image sensors in high volume electronics such as mobile phones, smartphones, notebooks and media players, (digital) photographs may be taken at any time or in any circumstance for different purposes. These images may be distributed over the internet in a short time, obfuscating the image source. These photographs may depict illegal acts such as assault or child abuse, and the need for reliably establishing the image origin becomes obvious when these images are used as silent witnesses in court.

The technique used to perform device identification is by extracting seemingly invisible 'fingerprints' left behind by the image sensor from the obtained (questioned) images, and comparing them with the 'fingerprints' obtained from images from other cameras (including the suspect's camera). The origin of these 'fingerprints' suggest that these patterns are unique, as they result from the non-uniform response of the pixels under a certain (constant) applied signal, due to construction and device non-idealities. Specifically, when the illumination incident on a number of pixels is exactly the same for all pixels, the output signals from these pixels will be slightly different, creating a a pattern with some pixels outputting systematically lower (or higher) signals (see figure 1.1). This Photo Response Non-Uniformity (PRNU) is used for establishing the image origin, and the matter of obtaining these PRNU patterns from natural images from digital cameras was previously addressed by [5][1]. The PRNU is a multiplicative signal, which means the apparent non-uniformity increases (linearly) with the applied signal.

In this report possible sources for these patterns are investigated, as well as noise sources which obfuscate these patterns. This background is needed to be able to explain and understand why these patterns can be used to perform device identification in court. The goal is to get a better understanding of the device response and the pro-

[1] In this report only attention is paid to digital photocameras. In [1] the idea is extended to digital camcorders.

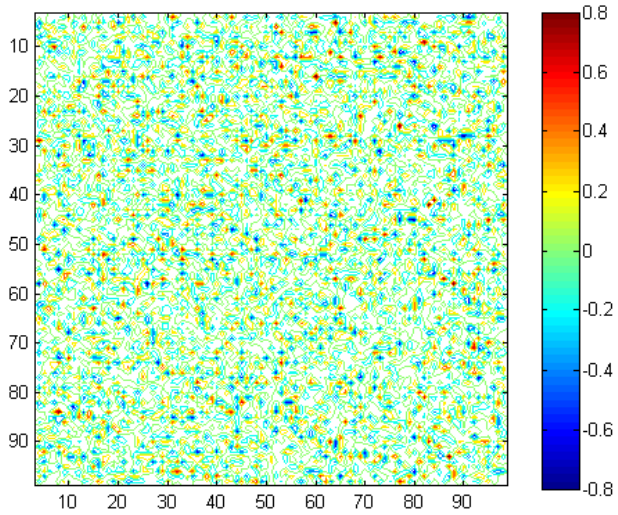

Figure 1.1: *Contourplot of a PRNU pattern. The red areas indicate areas that are more active than average, and the blue areas indicate areas that are less active.*

cesses occurring in image sensors between making a photograph and extracting and comparing the resulting PRNU patterns. In order to understand which factors are responsible for the existence of these PRNU patterns, we take a closer look to some of the possible sources. As there are a lot of possible sources it is generally difficult to properly quantify the theoretical share each source has in the final output, as they require detailed calculations.

To understand which factors are responsible for the various noise sources and patterns, we first explain how image sensors (CCD image sensors and CMOS Active Pixel Sensors) work. This is done in chapters 2 and 3. As mentioned previously, various noise sources are present which may disturb the pattern. In chapter 4 we review the most important sources. When possible, experiments are performed to compare with the theory[2]. When an experiment is performed, an **E** -symbol is shown in the margin, and a □ indicates where the experiment ends. Experiments are done for the CMOS image sensor based Canon EOS 10D, and unless mentioned otherwise the results are expected to be similar for CCDs. In chapter 5 we will introduce some general concepts common to both CCD and CMOS image sensors.

A degree of similarity is attached to two patterns by extracting and comparing the PRNU patterns from images, and calculating the correlation between these two patterns. Similar patterns have a high correlation, indicating a common origin. When two patterns are dissimilar the correlation is low. Although these correlations may suggest

[2]All mentioned Matlab scripts or functions (ending in *.m*) can be found on the NFI network on the $V:\backslash$ drive in the 'stagaires'-directory.

a common origin, it is not useful as an objective measure to the value of the evidence as needed in a forensic context. Although there are different methods available, the conceptual simplicity of the likelihood ratio makes the results easy to interpret. The likelihood ratio represents the ratio of two different hypotheses. By dividing the probability of the evidence assuming the prosecutor's hypothesis is correct (i.e. the suspect is guilty), by the probability of the evidence assuming the defence hypothesis is correct (i.e. the prosecutor's hypothesis is incorrect) an objective measure of the weight of the evidence is obtained. In chapter 6 a method for calculating the likelihood ratio is presented, but certain complications make it difficult to assess the reliability of the presented method.

Summarizing, we want to understand which factors are responsible for the existence of the PRNU patterns, and which factors (noise) are responsible for obscuring this pattern. We would like to find a likelihood ratio based on the comparison of the two PRNU patterns as obtained from the questioned image and from the camera.

CHAPTER 2

CCD

2.1 Summary

A Charge Coupled Device (CCD) image sensor is a relatively easy device to understand. It basically consists of a number of gates which are placed on top of a (silicon) substrate. By applying a positive voltage on the gate, a potential well is created in which the charge that is generated by the photoelectric effect is stored until readout. When the device is being read out the potential on each of the gates in a pixel is systematically varied in order to transfer the carriers to adjacent gates, until the charge packets have reached the sense node after which the charge is converted to a voltage. The Analogue-to-Digital Converter (ADC) then converts this analogue signal to a digital number. In this way the amount of light that was incident on each pixel is quantified, and a digital representation of the input is obtained. Systematic noise is mainly due to non-uniform photogate sizes and material properties.

2.2 A closer look

To understand how an image sensor functions, we look first at a photogate which is simple to understand. A silicon based MOS (Metal Oxide Semiconductor) capacitor can be used as a MIS (Metal Insulator Semiconductor) capacitor (Chapter 3 of [6]; see figure 2.1), and a CCD image sensor is basically a large matrix of these light sensitive MOS gates.

Oxide acts as an insulator and is mostly used. The insulator is sandwiched between the metal and the silicon, hence explaining the name. When we consider a p-type semiconductor (which means the bulk (substrate) is doped with e.g. Boron, supplying holes in the valence band of the crystal) and initially apply a voltage approximately the size of the threshold voltage[1] ($V \approx V_{th}$) to the gate of the MOS, the holes are repelled from

[1] The minimum voltage needed on the gate for the creation of an inversion layer.

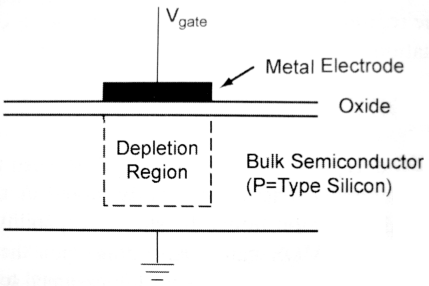

Figure 2.1: *Schematic of a Metal Oxide Semiconductor, in which the depletion region is shown beneath the gate. From [6].*

the gate and migrate towards the ground electrode . This application of a gate voltage results in the creation of a potential well, the 'volume' depending on the oxide thickness, gate voltage, and area of the electrode (Ch. 1.1 and 7 of [7]). In this way a thin depletion layer is formed, void of positive charge ('depletion mode'). When a voltage larger than V_{th} is applied, electrons are drawn towards the semiconductor/insulator interface, creating an inversion layer which acts as a conducting channel ('inversion mode'). The charges are stored in this region until the device is read out. The inversion layer consists of minority carriers (i.e. the electrons in the p-type semiconductor, a n-channel CCD). As this inversion layer is very close to the oxide/silicon interface (Insulator/Semiconductor), this way of collecting (and later on the transfer of this charge) of the electrons is inefficient due to charge trapping by impurities and crystal defects at the interface (and in general due to surface states (see §4.11)). Because the carrier density at this interface is very large, the separation of charges between the bulk and the inversion layer results in a large capacitance.

In a bulk- or buried channel CCD (BCCD), the depletion region is not formed immediately underneath the oxide, but deeper into the substrate between n-type and p-type layers. The p-n junction is then reverse biased (the n-region at positive potential and vice versa), so electrons and holes are removed out of their layers and a depletion region comes into existence. Between this depletion region and the depletion region under the oxide, a small region of constant potential (Z) is created where the generated charge carriers can then be stored. The carrier density is lower, and as the storage region is further removed from the gate, the electrons (majority carriers in general) feel less attracted. Therefore, the storage capability is lower (§1.6 of [7]). The full well capacity describes how much electrons can be confined in a potential well, without the electrons diffusing into neigbouring gates.

In normal (well illuminated) circumstances the charges in the depletion region are generated mainly due to the photoelectric effect, i.e. the photons are absorbed by the silicon and produce electron-hole pairs. The electrons will be attracted towards the

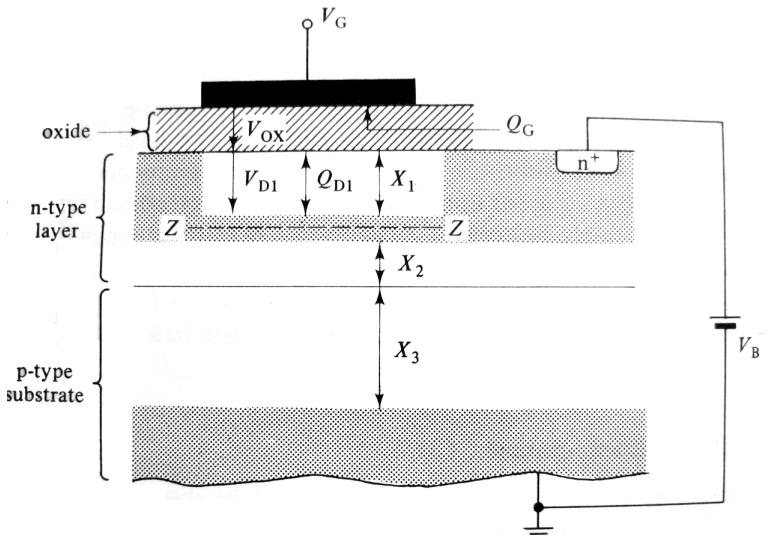

Figure 2.2: *Schematic of a BCCD, in which the charge is not stored underneath the oxide, but instead in a region of constant potential, Z. In the electron potential diagram it can be seen how the potential varies in the semiconductor with depth (see Appendix A). From [7].*

gate, and the holes will move to the ground electrode. Another way this charge can be generated is by dark current (free thermally generated electrons that are indistinguishable from the photogenerated electrons, see §4.5). We will see that when enough light is incident on the sensor the amount of electron/hole couples generated in the depletion layer is large, and the amount of dark current is relatively small. In low-light circumstances, however, the effect of the dark current may be clear.

Not all gates in the pixel can be used for storing charge, and in practice a pixel consists of multiple gates (generally two to four). Multiple gates are needed for charge transfer to occur ('charge coupling'), and each gate can have a different potential applied to it. By applying a systematically varying voltage on each gate of a pixel, the electrons can move to adjacent gates, and eventually to adjacent pixels, when the device is being read out, as illustrated in figure 2.3. This is done by shifting the charge one by one through the array of pixels towards the shift register. The shift register is basically a row of pixels hidden from view. This 'clocking' of the device can be done in a different number of phases. The output of the shift register is amplified by an amplifier after which it is quantized through an A/D converter (see §4.7). The amplifier is built directly into the silicon circuitry (and therefore conveniently called an 'on chip amplifier').

Ideally, all electrons would move to the adjacent gate when the voltage of the adjacent gate is increased. In practice this does not occur, and this charge transfer efficiency (CTE) is a fraction smaller than 1. In practice, the CTE$\approx$ 0.99999 which means that each

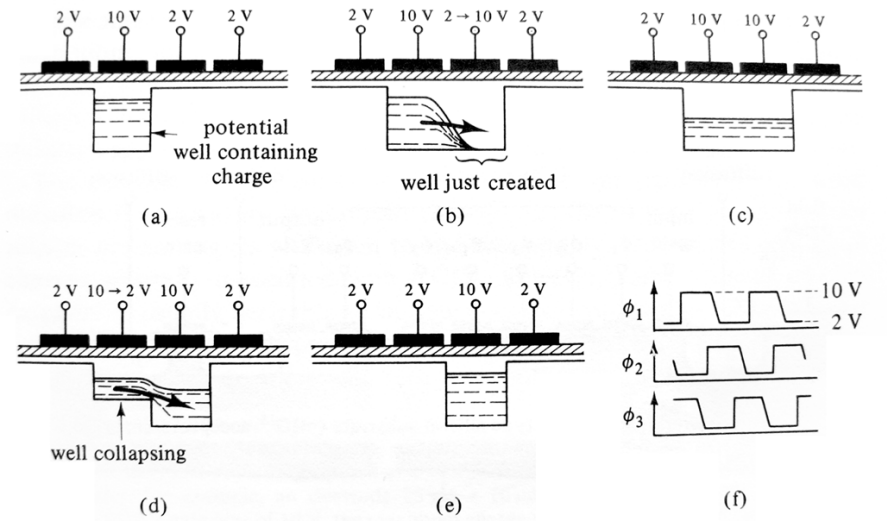

Figure 2.3: *Charge transfer occurs when the device is being read out by varying the applied potential on neighbouring gates. Note that the potential well is a hypothetical concept in which charge is represented as a liquid (§1.1 of [7]). In the case of a photogate the charge in the well represents the charge in the inversion layer. The depletion layer, on the other hand, is a physical region in the device. Of course, as electrons are point charges they have no volume. Hence, the wells do not really fill, but instead the depletion layer shrinks as more charge is introduced (see appendix A). From [7].*

transfer a small fraction of the electrons is left behind. This may not seem important, but as the charges have to travel through a lot of gates (especially when the number of gates per pixel is high), charge is lost. When the CCD array size grows, the need for near perfect CTE becomes apparent ($0.99999^{2000} \approx 0.98$ for 2000 gates). When a pixel holds 10^4 electrons, and needs to be transferred through 2000 gates, approximately 200 electrons are lost in the process. The charge transfer efficiency decreases with increasing transfer rate (§2.7 of [7]). Depending on the gain and the A/D converter, the loss of electrons can result in a significant loss in the resulting pixelvalue. Especially large CCD arrays call for near perfect production technology, as a single defect can block an entire column, because charge packets have to move through the gates of its neighbours [8]. The charge transfer efficiency is a

After charge is transferred towards the output, it is converted to a voltage by the sense node (figure 2.4). This sense node acts as a capacitor, and is reset to a reference level before the charge is being read. By comparing this reference voltage to the voltage after the charge packet is transferred to the sense node, we know how much charge was present in the packet. The output voltage is approximately proportional to the amount of transferred electrons so that the output is linearly dependent on the input (see 4.7).

In reality, the gates cannot be exactly adjacent, for the charge would flow into the

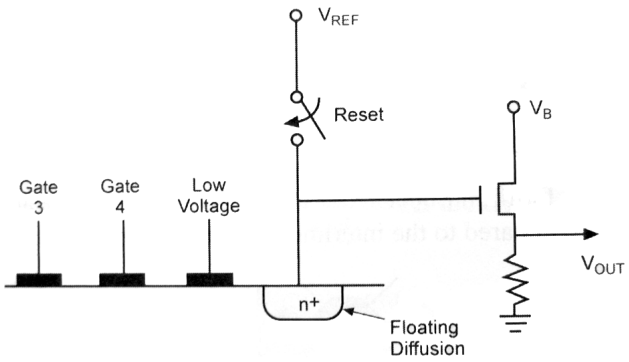

Figure 2.4: *The charge is converted by the sense node after it has been transferred through the gates. From [6].*

neighbouring pixel. This can be prevented by varying the oxide thickness, or overlapping gates. These variations act as channel stops (Ch. 3.2 of [6], and Ch. 2.3 of [7]).

Variations exists where it is not the front side but the back side of the CCD that is illuminated, mitigating the possible absorption by the gate structures that overlay the depletion region[2]. Hence, these back side CCDs generally have higher quantum efficiencies[3].

2.3 Noise

We have seen that the CCD is a device that is relatively easy to understand. Because the noise/PRNU sources present are mostly present in CMOS image sensors as well, we defer explaining some of the following terms to later sections.

Non-uniform response from CCD sensors results, among others, from non-uniform dark currents. As CCDs are large matrices of MOS gates that need to have near-perfect charge transfer capabilities, these devices are especially sensitive to crystal defects or charge traps in the material resulting in a fixed pattern noise (FPN) in the output which may be detected. For more details see §4.5.

The size of the photodetector area may vary between pixels in CCDs [10], for example due to material properties. Local mechanical stresses due to randomly distributed impurities [32] introduced during fabrication [33], for example, may cause the electric field underneath the gates to be non-uniform. A variation in the applied voltage on the gate also results in varying depths of the potential well, resulting in a systematic increase or decrease of the amount of charge carriers that can be absorbed or confined

[2] In this way the device becomes more sensitive to short wavelengths (UV), which is especially interesting for astronomy.

[3] In short, the quantum efficiency of a device is the efficiency with which an incoming photon is 'transformed' into an electron-hole pair. See also §5.2, and Ch. 2.2 of [9].

(this is actually a PRNU component, see §4.9). Therefore the region in which the electrons are confined after they are absorbed by the material are not well-defined. Non-uniform detector area, in the case of a CMOS image sensor, is pictured in figure 3.6 (p. 19). For a CCD the characteristics are assumed to be similar. Both these characteristics, the spatially non-uniform dark current and the non-uniform detector sizes are uncorrelated.

As the charges all are converted to a voltage by the sense node, a noise component present in CCDs is the flicker noise (§4.11). The reset noise from the reset switch (reset FET, see figure 2.4) adds thermal (kTC, Johnson) noise as well (§4.10). Both these noise sources are reduced by using correlated double sampling (see Ch. 3.5 of [6], and §4.7).

CHAPTER 3

CMOS

3.1 Summary

The CMOS based Active Pixel Sensor (APS) is an alternative to the CCD image sensor. Due to the integration of more components in comparison with the CCD, the anatomy of this sensor is a little harder to understand. First, the photodiode, which basically acts as a capacitor to confine the charges generated by the incoming light, is reset to a reference voltage. After the actual photodetection (charge integration) the charge is converted to a voltage, after which the signal needs to be read out. This is done in a different way compared to CCD image sensors, and in principle each pixel can be read out individually. In practice all the pixels in a single row are read out at the same time, one after the other, by moving the signals to the column circuit, after which the cycle is repeated.

3.2 A closer look

CMOS formally refers to a Complementary MOS circuit, in which NMOS (n-type MOSFETs) as well as PMOS (p-type MOSFETs) circuits are used[1].

[1] A FET (Field Effect Transistor) is, simplistically, a 3 electrode device (we implicitly assume for the functioning that the substrate is internally joined to the source) which is operated by an electric field. Basically there are 3 contacts, namely the drain, gate, and source. Current flows between the drain and source if the voltage applied on the gate is high enough to form a conducting path (in the case of an n-channel MOSFET as is assumed from this point onward). At a certain drain-to-source voltage V_{DS} increasing the gate-to-source voltage V_{GS} allows a higher drain current to flow. When the gate to source voltage is below a certain threshold V_{th}, no appreciable current is flowing: at this point the resistance of the channel between drain and source is too high (see figure 3.1). The threshold voltage is a function of, a.o., gate insulation material, the voltage between the substrate and the bulk, and impurities at the silicon-insulator interface (Ch. 2.1 of [11]). This threshold voltage decreases with increased temperature in the order of $mV/^{\circ}C$. Simplistically, the MOSFET behaves like a voltage controlled switch. Each type of FET, be it nFET, pFET or JFET, behaves this way. Depending on the type, however, the voltages and

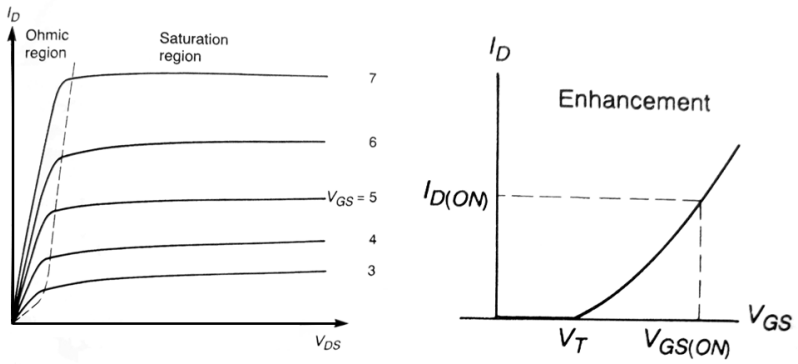

Figure 3.1: *Behaviour of the n-channel MOSFET with different drain-to-source voltage V_{DS} and gate-to-source voltage V_{GS}. From [12].*

The CMOS image sensor gets its name from the fact that the production technique used to build the sensor is very similar to the process used to build CMOS integrated circuits such as memory devices and microprocessors.

In the category of CMOS based image sensors, the active pixel sensor (APS), as opposed to the passive pixel sensor which is mainly of historical interest, is the most important one. Passive pixel sensors were the predecessors and suffered from high noises as individual pixels lack a few transistors compared to the active pixel sensor[2]. The active pixel sensor (APS) has a number of active transistors integrated on the pixel, explaining the name. The APS eliminates the need for near perfect charge transfer as with CCDs, and shortens the time needed for read-out. Another advantage of CMOS image sensors is that they can be manufactured using technology already available for other integrated circuits, such as memory devices. Finally, as the transistors on the sensor only use power when they change state, and as the voltages used are lower than their corresponding CCD counterparts, they are more efficient. On the other hand, integrating more functions on each pixel results in more noise, and this leads to lower uniformity over the chip (one of the conventional drawbacks of CMOS image sensors).

As can be seen in the figures below, there are several components in the APS, some of which are shared (e.g. the ADC, bias transistor M_4). Needless to say, each of these components is a potential source for noise, as will be investigated in the following paragraphs.

To get an understanding where the different noise sources and PRNU specific to the APS come from, it is necessary to first take a closer look at the structure of the APS.

currents have different characteristics. Types that have a conductive channel even when $V_{GS} = 0$ are called Depletion/Enhancement MOSFETs, and types that need a $V_{GS} > 0$ to produce a drain current are called Enhancement MOSFETs (see for this and more e.g. Ch.6 of [12].)

[2]Compare e.g. figure 1 and 2 in [10]. Remember that in both figures the column circuit is attached to the actual pixel.

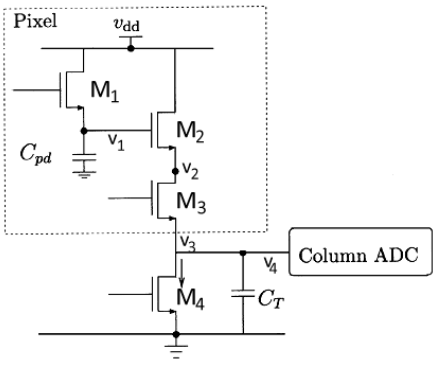

Figure 3.2: *Schematic of a basic 3T active pixel sensor. Note that each pixel consists of three transistors and a photodiode. The components outside the pixel cell are shared per column, such as the bias transistor M_4 and the column ADC. C_{pd} represents the equivalent capacitor of the photodiode. In an actual sensor nFET and pFET may be combined, but this is not necessary. The hole mobility is lower than the electron mobility in silicon, which is why nFETs are the logical choice. From [13].*

Although a large number of different designs exist, and the production complexity as well as the functionality increases along when the number of transistors is increased, we can look at the basic 3T (T is shorthand for Transistor) design.

The pixel itself consists of a photodiode with 3 transistors, which act as a reset switch (M_1), source follower (M_2)[3], and a row select transistor (M_3). The bias transistor (M_4) is not present on each pixel, but only in each column. These transistors need to switch in order to let a current flow to charge a capacitor, be it a 'real' capacitor or the intrinsic capacitor of a transistor. The time needed for the voltage of a charging capacitor to reach its final value is exponential; therefore it is important that when the charge needs to be transferred the transistor does not show any resistance (in [13] a detailed analysis is presented in which the access transistor (M_3) resistance is taken into account). The detection can also be done with a photogate (MOS capacitor), but we will only consider the photodiode (the basic operation of the photogate was reviewed in Ch. 2).

The detection of light is done by the photodiode, which is essentially a capacitor to store charge generated by the photoelectric effect. Before the charge is integrated, the capacitor is charged / reset to a reference reset voltage, commonly called V_{dd}. After the capacitor is reset (and the reset gate is turned off), charge accumulates and the voltage over the capacitor declines. The reset voltage applied to the gate needs to be controlled accurately, as this reference voltage is essential in determining the amount of incident

[3]If a transistor acts like a source follower, its source voltage follows the gate voltage. In the small circuit equivalent the drain current i_d is proportional to the potential difference between gate and source voltage: $i_d = g_m(v_g - v_s)$, with g_m the transconductance ($\mho$). The source voltage is then $v_s = R_s i_d$, with source resistance R_s. Then $v_s = (R_s g_m/(1 + R_s g_m))v_g$, and as R_s is large, we see that $v_s \approx v_g$ (Ch. 6.5.6 of [12]).

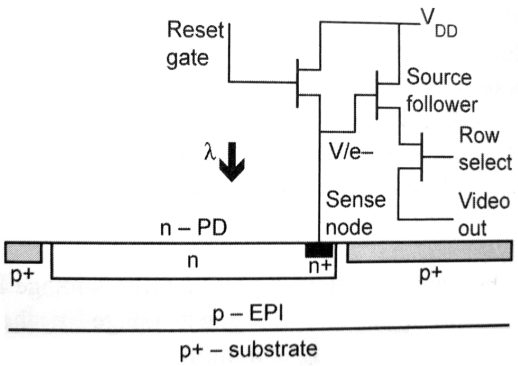

Figure 3.3: *Schematic of a 3T APS, in which the sense node is shown. From [6].*

light. As the M_2 transistor acts as a source follower, the source of this FET is set close to the voltage on the gate (i.e. the voltage over the photodiode).

The photodiode is charged by the dark current (§4.5) as well as by the charges generated by the photoelectric effect. The n^+ region (a heavily doped n-region, see figure 3.3) is called the sense node, which stores the charge during readout, and hence acts as a capacitor as the sense node is a heavily doped region implanted in a 'normally' doped region.

The photodiode is attached to the gate of the source follower (M_2). The gate of M_2 acts as a 'voltage buffer', a capacitor that observes the voltage without removing the accumulated charge from the photodiode. The voltage over this combination of capacitors (the sense node capacitance, the capacitance from the gate of M_2, and the photodiode capacitance, (Ch. 4.3.1 of [6])) is sensed by M_2:

$$V = \frac{Q}{C} = \frac{It_{int}q}{C_{eq}}, \tag{3.2.1}$$

with I the current (dark current, photogenerated carriers), t_{int} the integration time, q the electron charge, and C_{eq} the equivalent capacitance (from the abovementioned capacitors). Note that, contrary to the photogate in the CCD, there is no metallic contact above the charge storage region, resulting in a higher quantum efficiency at short wavelengths. To improve long wavelength response, the depth of the n-layer can be increased such that carriers are less likely to diffuse into the substrate. The size of the voltage that is sensed later on, indicates the amount of light incident on the photodiode during integration time (unless it gets saturated), effectively by comparing initial and final voltages. Saturation occurs when the potential difference caused by the collected electrons exceeds the pre-charged (initial) voltage of the diode. However, saturation does not result in blooming as in CCDs (i.e. charge that flows into the potential wells of neighbouring pixels).

3.2.1 Readout of the photodiode voltage

As M_2 is in the source follower configuration, the voltage existing over the capacitor after integration is 'copied' to the source output of M_2 (remember that the source voltage follows the gate voltage; the gate voltage is the input voltage from the photodiode in this case). M_2 converts the charge found on the diode into a voltage. This happens when M_3 (the access switch) is closed, so that there is a path to ground. This path to ground is necessary as V_1 (the voltage over the photodiode) is the voltage with respect to the ground. This voltage is also the gate voltage of M_2, and V_2 will come close to V_1.

If each column has an Analogue-to-Digital Converter (ADC, in which an analogue voltage is converted to a digital signal, see §4.7), and note that a column is simply a collection of rows, we see that the rows are read out sequentially. If each column does not have its own ADC, the signals from a number of input columns are 'multiplexed' (in which a single input can be selected from multiple inputs, and this input signal is passed to a single output), and the preceding is still true. As only one row is active at a time during readout, the power dissipation is low, an advantage compared to CCDs where each charge packet is transferred to its neighbour during readout.

When the photodiode is read out, the gate voltage of M_3 (at a certain pixel's row) is raised, so that there is a path to the column circuit and its transistors. M_3 is effectively a short-circuit, and M_2 goes into the source follower mode and hence $V_3 \approx V_2 \approx V_1$. This is done for all columns (i.e. all pixels in a row) at the same time. Capacitor C_T is being charged by V_3, and V_4 is the voltage on this capacitor during read out[4]. We see that when V_3 is close to V_2 we have a substitute for V_1 on C_T. Next, this voltage is digitized, so we know how much light fell on the photodiode. M_4, or the bias transistor, acts as a voltage dependent resistance controlled by the gate voltage (bias voltage). We see that M_4 should be either closed or open; otherwise leakage (and hence degradation of the signal) may occur.

In a simple implementation, the charge on C_T is fed into a differential amplifier which looks at the difference between its two inputs, a reference voltage and the voltage from the photodiode. The reference voltage is the voltage the photodiode is charged with before the charge is integrated. Subtracting the two voltages, we know how much light was detected at the input. The output of this stage depends on the gain ('ISO') of the amplifier, $V_5 = A(V_{ref} - V_{photodiode}) = A(V_{ref} - V_4)$, so that the output sensitivity can be changed. After the ADC converted this analogue signal to a digital signal, the next row is read out. Depending on the voltage of the pixel in the next row, the capacitor is charged or discharged. If V_1 is higher than V_4, C_T will simply be charged. If V_1 (hence

[4] As integrating capacitors on a chip is not efficient, the drain capacitance from M_4 could be used as a substitute, as the amount of charge is small anyway.

V_2) is lower than V_4, no current will flow through M_3. The bias voltage controls the drain current discharging C_T. When the voltage over C_T (i.e. V_4, the discharged voltage) becomes near V_2, V_{DS} becomes larger than zero, so a drain current starts to flow through M_3, and C_T is set to the right voltage[5].

3.2.2 Per pixel transistor noise

In [10] a model is developed in order to describe the various noise sources and FPN components in CMOS image sensors (§4.6). The authors note that the model describing the noise in a CCD image sensor is not adequate in the case of a CMOS APS due to the vastly different readout circuits used in both cases. In the case of a CMOS APS, the noise is modelled as the sum of two independent components: the column and pixel components. The resulting model and experimental test structure show there is a correlation between column components in the APS (see §4.16), but that individual APS pixel components are uncorrelated.

An offset from M_1 is present due to the fact that the photodiode (as a capacitor) is charged to somewhat different levels in the initial condition, for example due to threshold variations (the minimum voltage V_{th} needed to make a current flow). Also, if the voltage on the reset is not low enough (i.e. V_{GS} too high), there is a channel between the drain and the source, so that charge may be injected into the photodiode (capacitor). However, due to device fabrication and process variation, a certain V_{GS} may (at constant V_{DS}) result in a somewhat different drain current between transistors. When charge on the photodiode is reset to its reference value, the current through M_1 may vary from pixel to pixel, resulting in a different precharged value of the photodiode. This gives a systematic offset value in the output, detectable as Fixed Pattern Noise (FPN, see §4.6).

A possible deviation of V_2 with respect to V_1 may be due to the existence of an offset voltage between the gate of the transistor and the source of the transistor. Put simply, the input offset voltage is the voltage required at the input (gate) to make the output (source) zero. Due to this offset the gate to source voltage V_{GS} is a little smaller (assuming V_{DS} remains constant) than it should be, see the I_D vs. V_{DS} characteristic. This presents another systematic offset in the output value, which contributes to the FPN.

M_2 needs to have a high input (gate) impedance, otherwise electrons would be extracted from the photodiode (the electrons would flow from gate to source, clearly not the intended behaviour[6]). FETs have practically infinite input impedances, so no vari-

[5] As [13] notes, the discharging time limits the readout time, and it is quicker to reset the voltage after each readout.

[6] This can happen, albeit due a different process, with Fowler-Nordheim tunnelling, see p. 36.

ation due to this factor is expected. M_2 only 'copies' V_1 when M_3 is closed, i.e. when it passes the current during readout.

M_4 is shared among the pixels in the same column (and, if multiple columns share an output circuit multiple columns share M_4). A transistor needs a small input current to function, and this input bias current may vary somewhat. Noise from this transistor may show up as column Fixed Pattern Noise (FPN, see §4.6). Especially when this transistor is shared over a number of columns, this column FPN may be detectable (see §4.16). (See figure 3.1: V_{DS} is approximately constant, but if V_{GS} varies somewhat (due to voltage fluctuations in the system, or device variability), each column will have a slightly different input response.) Put differently, at a certain V_{GS} two transistors may have different I_{DS}.

A variation in transistor sizes, such as the length of the channel, results in a variation of resistance when the electrons need to move from drain to source, such as when the photodiode is charged when M_1 is closed.

Most of the variations occurring due to transistor properties can be reduced by using correlated double sampling, but this may not be possible on-chip for the 3T APS presented here, but it may be possible to do this externally (§4.3.4 of [6]).

3.2.3 Noise from the photodiode

Contrary to the CCD which has a photogate (photoactive MOS capacitor), the CMOS image sensor usually has a photodiode integrated in the pixel. Photodiodes work a little different, and to understand what noise sources result from the photodiode, we take a brief look at some of the characteristics of this device. Some of the noise sources mentioned above happen per column, and others happen per pixel. As each pixel has its own photodiode, the noise sources mentioned in this section are per pixel.

In a photodiode a PN junction works as a converter of light into electrons when the energy of the incident photon exceeds the energy needed for an electron to migrate from the valence band into the conduction band (approximately 1.12 eV, and as $E = h\nu$ this results in a maximum wavelength of approximately 1100 nm). In between the PN layer a depletion layer is formed which initially is void of charge carriers: the electrons diffuse from the n-layer (mobile negative charges) to the p-layer (mobile 'positive' charges) where they combine with holes and vice versa (figure 3.4).

Due this recombination, the n-type material becomes positive near the junction, and the p-type layer becomes negative near the junction. This results in the formation of an electric field. When an incoming photon in the depletion region releases an electron into the conduction band, the electron diffuses towards the N-layer, and the holes diffuse toward the P-layer by this electric field. To increase the width of the depletion

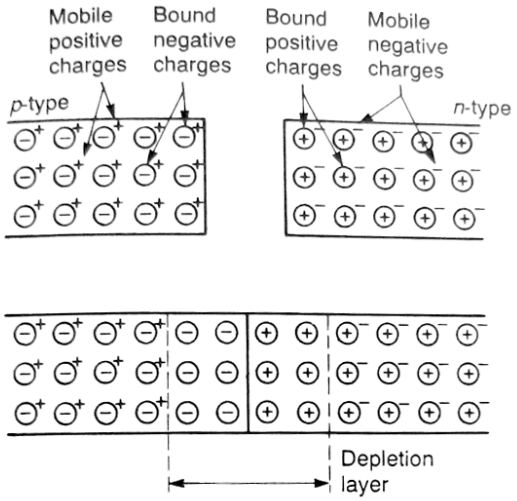

Figure 3.4: *Schematic of a p-n junction. From [12].*

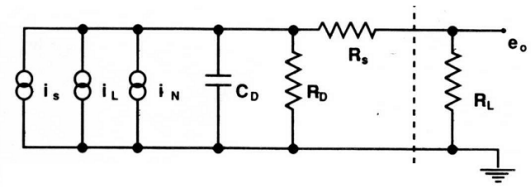

Figure 3.5: *Photodiode equivalent circuit. Other than the signal current i_s, a leakage current i_L and a noise current i_N are present. The capacitor represents the equivalent depletion layer capacitance. The shunt resistance R_D represents the resistivity of the depletion region in the silicon, and R_s represents the series resistance. From [14].*

layer, the photodiode may be reverse biased, so the p-type is made negative and the n-layer positive. This increases the chance of collection, as only the carriers within the mean free path of the depletion region contribute. Summarizing, we have a positive charge in the P-layer, and a negative charge in the N-layer. When the device is being read out the charges will flow to the electrodes. Depending on the energy/wavelength of the incoming photon, the electron-hole pairs are generated at different depths in the material.

The positive n-layer repels holes in the p-layer, and a diffusion current is generated when electrons from the n-layer surmount the existing potential barrier and move into the p-layer, and vice versa. A drift current is generated when holes from the n-layer gain enough energy from the assisting electric field to overcome the barrier into the depletion layer. Electrons cut loose from the atoms by thermal energy also produce this drift current, resulting in I_N (see figure 3.5). The disadvantage of reverse biasing the photodiode is the increased leakage current (I_L), as some of the electrons in the p-layer

are attracted to the n-layer resulting in higher noise, an effect that grows with temperature. When more charge needs to be confined in the capacitor, more charge will leak away. Hence this is a PRNU component. The diode junction capacitance C_D may vary somewhat, e.g. due to a varying depth of the n-implant, the non-uniform division of impurities in the wafer, or a varying bias voltage (so the depletion layer has a different width resulting in more or less photons are captured in the depletion region). This is an effect that is linear with signal; a certain percentage will simply not be captured (this is a PRNU component, see §4.9).

When the charge is being read from the photodiode, the charge is drawn towards the sense node. This sense node is responsible for a high dark current, due to crystal defects appearing during fabrication [15][7].

Another effect that is linear with signal is the size of the active region of the photodiode, simply because a bigger area collects more light, and hence is a PRNU component. This may be caused by the mask that covers the substrate to define where in the substrate the ion implantation should occur during fabrication. Also, misalignment of the mask may result in pixel-to-pixel variations such as non-uniform photodiode sizes. As the depth at which the photon is absorbed depends on its wavelength, we see this PRNU source is wavelength-dependent (the reason why flatfield images from different cameras should be obtained with lightsources having roughly the same spectral distribution). Misalignment of the mask which defines where the transistor components should be deposited may also result in pixel defects. These defects do not have to show up in the output image, as it is possible to obscure these defects by interpolating the pixelvalue from neighbouring pixels. For a complete description of CMOS fabrication (not specific to image sensors), see Ch. 3 of [11]. The light sensitive area is a complex shape that depends on the *'location of the electronics, contacts, and bus lines'* (§6.1.1 [6]). This results in an ill-defined region where the pixel is sensitive to light, as is shown in figure 3.6.

Different types of photodiodes exist; a simple variation of is where the n-implant is deeper ('deep implant photodiode'), as to improve performance for long wavelength photons. The pinned photodiode has an additional p^+ pinning implant on top of the n-implant to reduce dark current (Ch. 4.3.1 of [6] and [16]).

[7] In a 4T design, this dark current is reduced by adding a transistor (transfer gate) between the sense node and the photodiode. By applying a voltage to this transfer gate a channel between the sense node and the depletion region is created, resulting in lower dark currents. The downside to 4T (and higher T) designs is that less area can be devoted to the photodiode, in turn resulting in lower signals.

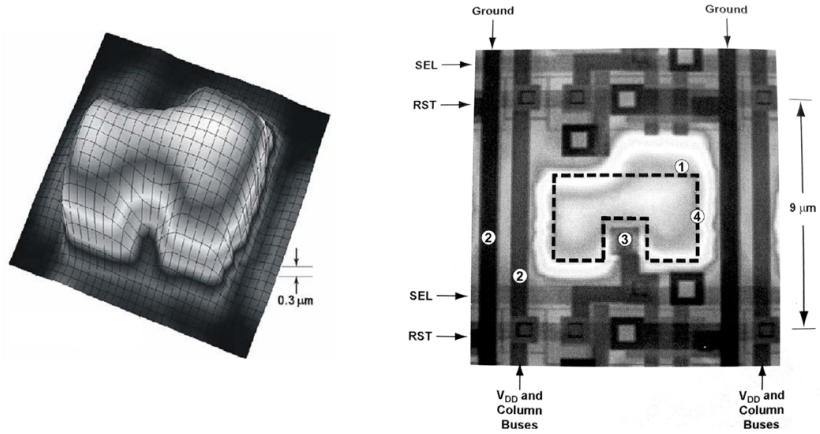

Figure 3.6: *The responsivity of a CMOS pixel is spatially variable (a). In (b) it can be seen that the lightsensitive area is variable within and outside the photodiode boundary. Region 1 shows a lightsensitive area outside the photodiode boundary. Region 2 show the ground and power lines. Region 3 shows a lower sensitive area due to a overlaying contact. Region 4 shows the photodiode boundary. From [6].*

CHAPTER 4

Noise

4.1 Summary

This chapter covers various noise sources, the explanation of which contributes to the understanding of how image sensors perform.

Although it is difficult to quantify the share each noise source has in the output of the image sensor, it is possible to divide the noise sources into three different categories. We will find that the most important sources are the shot noise, dark current, read noise and a systematic 'noise' component PRNU. Although this last component is strictly not a noise component, as the pattern is systematically present in all photographs, it is sometimes convenient to share the PRNU under the general term of noise (along with the misnomer 'Fixed Pattern Noise', which, as the name suggests is a fixed pattern (and hence strictly not noise)). The noise sources obfuscate the underlying PRNU pattern. The dominant noise source at practically all illumination levels is the photon shot noise[1], a source that is inherent to the nature of light and hence cannot be reduced. Only at very low intensities the read noise is significant and in fact the dominating noise source, as was also concluded in [26]. Only at very long integration times (multiple seconds) and low illumination levels the dark current plays a role; this is however negligible at normal exposure times. By considering how the standard deviation changes when multiple flatfield images are averaged, we can see that averaging approximately 10-15 images is already sufficient to almost completely diminish the noise variations. Hence, by averaging 10-15 images the noise variations are removed, and the estimated PRNU pattern found from this is close to the 'true' pattern. Some other simple characteristics can be used to determine the gain of the device (e^-/ADU), the size of the PRNU, and the size of the readnoise.

[1] We find that the standard deviation of a small area from a single flatfield is approximately equal to the square root of the average divided by the gain (§4.15): $\sigma_{measured} \approx \sqrt{\mu_{measured}/Gain}$). This means that the 'noise' present in a flatfield can be successfully described by the photon shot noise (see also §4.4

4.2 Introduction

As was mentioned briefly in the introduction, device identification is performed by extracting and comparing PRNU patterns. Based on a measure of similarity (currently the correlation) it is determined whether two patterns (and hence two images) have a common origin. Due to the presence of various noise sources, questions may arise about the reliability of the obtained patterns. Furthermore, we wish to understand which systematic sources are responsible for the existence of the PRNU patterns.

A general model frequently used to describe the output of the sensor, is [2]:

$$I = g^{\gamma}[(1+K)Y + \Lambda]^{\gamma} + \Theta_q. \tag{4.2.1}$$

I represents the (output) image, g the colour gain needed for colour reproduction (§5.3), γ represents a correction factor for intensity reproduction (§5.3), K is the PRNU 'fingerprint' (§4.9), Y is the incident light intensity, Λ represents the various noise sources, and finally Θ_q represents the quantization noise (§4.8).

E The K-factor represents the PRNU, and is a pattern the size of the image I. It expresses that the true signal is obscured by the PRNU, and that it is a factor which scales linear with signal. It is often modelled as a zero-mean normal (Gaussian) signal, and this normal distribution is indeed confirmed after applying the filters to extract the PRNU pattern on an averaged flatfield (uniformly illuminated image). The range of values in the extracted PRNU pattern across the image sensor is normally distributed. □

We will review the following topics: shot noise (§4.4), dark current (§4.5), Fixed Pattern Noise (§4.6), A/D converters and correlated double sampling (§4.7), quantization noise (§4.8), PRNU (§4.9), read noise (§4.10), flicker noise (§4.11), determining the amount of images needed for reliably extracting the PRNU (§4.13), finding the three regions in the photon transfer curve (§4.14), finding the size of the gain and read noise from images (§4.15), and the periodical column variation (§4.16).

4.2.1 Statistics

The variances of the various noise sources can be added according to the simple laws of error propagation. If we express the noise in electrons, and if

$$f = A + B + C + \ldots, \tag{4.2.2}$$

and A,B,C, etc. are independent, then

$$\sigma_f^2 = \sigma_a^2 + \sigma_b^2 + \ldots, \tag{4.2.3}$$

so that

$$\sigma_f = \sqrt{\sigma_a^2 + \sigma_b^2 + \ldots.} \tag{4.2.4}$$

It can be seen that adding (or subtracting) a number of images results in an increased uncertainty of the pixelvalue. Specifically, the addition of n images (each with noise σ) results in a total uncertainty of $\sqrt{n}\sigma$. Therefore, it is in principle not a good idea to subtract a dark frame from the image, as this increases the dark current shot noise with a factor $\sqrt{2}$. Conversely, averaging n images results in a reduction of uncertainty by a factor $\sqrt{n}$ (i.e. $\frac{\sigma}{\sqrt{n}}$).

In theory, we should average an infinite amount of frames in order to diminish the noise components. To see how much of these images should in practice be averaged to remove temporal variations (e.g. photon shot noise) in order to reliably extract the PRNU pattern, see §4.13.

4.3 Experimental conditions

As will be shown in this chapter, various noise sources with varying characteristics are present in the output of the image sensor. As not all effects of the individual noise sources can be easily separated, it is not always possible to experimentally find the effects of these sources. For example, the read noise (§4.10) includes the effects of the flicker noise (§4.11). However, by experimentally varying parameters such as the integration time, and by calculating simple statistics a great deal of information can be obtained.

When flatfield images (§5.2) are taken, the sensor is illuminated as uniformly as possible by photographing a flat piece of paper, and by removing the lens system in order to reduce vignetting. Furthermore, the lowest ISO value (100) is selected in order to make use of the complete full well capacity (when the ISO value is raised, the amount of useful charge is reduced). Images were taken in a room with fluorescent lamps. When dark frames are taken, the lens cap prevents any light from entering the camera system. Also, when frames are processed, the statistics of the first 10 frames are ignored to reduce the effects of temperature variations in the sensor when the device is operating. The experiments done in this chapter are performed by using the CMOS image sensor based Canon EOS 10D.

E Before the RAW[2] (image data that is minimally or not processed) data can be read by Matlab it has to be converted to another format. This is done by using *dcraw* [4].

[2]Specifications on this format generated by Canon cameras can be obtained from [3].

As was found in due course, there are large differences between Canon's proprietary RAW converter and the *dcraw* converter. One of the striking differences between both converters can be observed when a large number of (flatfield) images (see §5.2) are taken and subsequently converted from RAW to 16 bit TIFF. When we look at the pixel response (in terms of its pixel value) of a certain pixel, large differences appear depending on the converter used. In the case of Canon's converter, a small amount of pixelvalues occur often while other pixelvalues occur never. When the same files are converted with *dcraw*, the spread in pixelvalues looks (at first sight) to be approximately normally distributed over a fairly wide range. Indeed, upon further investigation the obtained histograms were tested for normality using the Kolmogorov-Smirnov and Lilliefors tests (*kstest* and *lillietest* in Matlab respectively), and it was concluded that it could not be rejected that the underlying distribution was normal[3]. □

As *dcraw* has a lot of options (notably the ability to decode RAW images without interpolation and scaling) that allows a fine-grained control over the output, it was decided this converter was used instead of Canon's. Unless otherwise noted, the conversion of the RAW image data to lossless 16 bit TIFF files is done without a form of interpolation, scaling, gamma correction or colour gain by using *dcraw*[4]. Furthermore, because there are twice as much 'green' elements in the colour filter array (§5.3) the green channel is used for calculations of e.g. averages or standard deviations. This reduces the aforementioned model (4.2.1) to

$$I = (1 + K)Y + \Lambda + \Theta_q. \tag{4.3.1}$$

In practice, the amount of RAW files encountered on which device identification needs to be applied will be low. Indeed, most cameras apply a form of JPEG compression, and it is these images which are encountered in practice. JPEG is a lossy compression in which the image is divided into 8x8 blocks, after which a number of transformations are done to reduce image data size. When a high amount of detail is present in the image, the division of 8x8 blocks creates residues due to insuperable edges, hence adjusting the original signal. The JPEG compression effectively acts as a low pass filter, which means high frequency content such as PRNU is attenuated. The reason for using RAW image data at this stage is that in order to understand the basic

[3] *Use hist4.m to extract the pixelvalues at a certain location from a large amount of files. As working with large files is slow, it is advised to crop them first. This can be done with the imgcrop.m function. When the RAW files are converted without interpolation, only one channel is present with the pattern of the colour filter array,* $\begin{smallmatrix} R & G \\ G & B \end{smallmatrix}$ *(See §5.3). To split this in 3 channels (*$\begin{smallmatrix} R & 0 \\ 0 & 0 \end{smallmatrix}$, $\begin{smallmatrix} 0 & G \\ G & 0 \end{smallmatrix}$, $\begin{smallmatrix} 0 & 0 \\ 0 & B \end{smallmatrix}$*) use the make3channels.m script.*

[4] This is done by using *dcraw* with the -D switch.

principles, one needs to get rid of the complicating factors. At later stages it is possible to find out how much the RAW signal is adjusted by applying a form of compression, as is done in [2].

4.4 Shot noise

The arrival of photons (electrons) at (in) the detector is a discrete process in time, and can therefore be described by a Poisson distribution. The characteristic of a Poisson distribution is that the variance is equal to its mean.

Photon shot noise As the arrival rate of photons at the detector is Poisson distributed, the number of electrons excited in the actual material is also Poisson distributed (see also appendix C). If on average S photons arrive with wavelength λ, then approximately $\int S(\lambda) \cdot QE(\lambda) d\lambda$ charge carriers are excited and captured in the well, with $QE(\lambda)$ the quantum efficiency (§5.2). According to Poisson statistics, the variance σ^2 of this process equals the mean: $\sigma_S^2 = S$. The signal S divided by the gain G of the camera (the number of electrons needed to increment the DN with 1) is the pixelvalue in absence of other noise sources. As the number of photons / electrons detected varies, the output of the sensor at a certain pixel varies as well. The 'noise' resulting from the variance of the amount of photons is then $\sqrt{\frac{S}{G^2}}$. Conversely, if we have a pixelvalue P, then the variance in pixelvalue is approximately $\frac{\sqrt{PG}}{G} = \sqrt{\frac{P}{G}}$. Note that the signal-to-noise ratio can never be higher than $\frac{S}{\sqrt{S}} = \sqrt{S}$.

Dark current shot noise Due to various causes (impurities, crystal defects, etc.), a varying amount of thermal electrons are added to the 'genuine' electrons from the photoelectric effect (see also §4.5). The number of electrons moving through the channel in a transistor, or the number of electrons being captured in the potential wells are described by a Poisson distribution, which results in a dark current shot noise of $\sigma = \sqrt{D}$, in which D is the mean number of electrons due to dark current (see formula 4.5.1). In transistors, when $V < V_{th}$ (i.e. when no current should flow), electrons still move due to shotnoise between the drain and the source[5].

[5]This shot noise is independent of temperature. Thermal noise (kTC noise), on the other hand, results in a (Brownian) motion of electrons in a resistor due to the thermal energy. These processes occur simultaneously, and these processes can be unified, as is done in [17].

The variance from these processes is, as mentioned earlier, the sum of the individual contributions ([6], Ch. 6.5.1):

$$\langle n^2 \rangle_{shot} = \langle n^2 \rangle_{photon} + \langle n^2 \rangle_{dark}. \tag{4.4.1}$$

Note that both processes occur at the electron level. The pixel values resulting from these processes are not Poisson distributed, as the electrons first have to go through the Analogue-to-Digital Converter (AD/C). As the shotnoise is a purely statistical process, it is impossible to reduce this effect. It is, however, not impossible to reduce the influence of this noise source. For example, by lowering the quantum efficiency (§5.2), the influence of the random arrival of photons at the detector is reduced.

4.5 Dark current

Dark current noise is essentially linear, meaning that with doubled integration time the amount of noise from this source doubles as well. The mechanism behind dark current is the thermal generation of charge carriers. Dark current occurs in photosensitive devices even when there is no incident light on the device. The simple explanation is when atoms start vibrating in the lattice, as they do more and more when the temperature is raised, electrons make the transition from the valence band to the conduction band by the absorption of phonons. These electrons can be captured in the depletion regions, where they cannot be differentiated from photogenerated electrons. The dark current can be generated through different processes:

Generation-recombination In the generation process an electron is promoted to the conduction band from the valence band, leaving behind a hole (intrinsic absorption). In the recombination process, the electron is demoted to the valence band, filling a hole. In thermal equilibrium, generation and recombination are balanced, so that the product of charge carriers (holes and electrons) is constant. As the amount of free carriers depends, among others, on the temperature, a temperature variation will result in an imbalance, and this generation-recombination process represents a dark current. Due to the presence of impurities energy levels are added in between the conduction and valence band, making it easier for the electrons to move into the conduction band by absorbing phonons. Deep in the substrate, where longer wavelength photons are absorbed, this process of recombination is more likely ([18], Ch. 19.3.8). In [19] this is regarded as the main component of dark current. When charge is trapped in the silicon (see also appendix B) the number of free carriers fluctuates, and results in a noise component

(generation-recombination noise)[20][6].

Diffusion from the depletion layer boundary Electrons diffuse from the depletion layer boundary towards adjacent gates, e.g. when the voltages applied are not high enough to keep the electrons in place. This also results in dark current. The effect of diffusion can be reduced by using high resistivity silicon, so electrons are less likely to travel to adjacent gates. Diffusion is especially a problem for long wavelength photons as these are absorbed deep in the substrate, where the potential may not be sufficient to keep the electrons from wandering off into neighbouring pixels, especially at high temperatures when the electrons gain a lot of energy. Also, impurities in the silicon and non-uniform thickness of the substrate or oxide result in a distorted lattice which can slightly change the direction of the potential underneath the gate, making it easier for electrons to diffuse. Large full wells (large amounts of charge carriers that are stored in the potential wells) will result in more diffusion. When large full wells fill up, diffusion is more likely to occur (see appendix B).

To get an idea of the size of the dark current at different temperatures in CCDs, the general dark current formula is very useful:

$$D = CT^{1.5}e^{-E_g/2kT}, \tag{4.5.1}$$

$$C = \frac{N_{dc}P_s}{qT^{1.5}e^{-E_g/2kT}},$$

with T at room temperature, P_s is the pixel size (cm^2), N_{dc} is the dark current as supplied by manufacturer (nA/cm^2), k is the Boltzmann constant, and $q = 1.6 \cdot 10^{-19}C$, so that

$$D = 2.5 \cdot 10^{15} P_s N_{dc} T^{1.5} e^{-E_g/2kT}.$$

We can plot this for various temperatures and dark current rates (figure 4.1). A dark current increase of 100% occurs for a T increase of approximately 8 °C at 17 °C.

As the dark current will always be present, it is an inherent limitation to the noise floor.

Dark noise can easily be visualized by keeping the shutter closed, and letting the device capture electrons for a certain amount of time. When this is done, the read noise will be present as well (§4.10). As the contribution from dark current to the overall noise is in practical situations never the dominant noise source (except possibly at very long integration times and low signal levels), and the size of the dark current is small

[6]The statistics of the generation-recombination processes in the presence of traps is described by the Shockley-Read-Hall model [21].

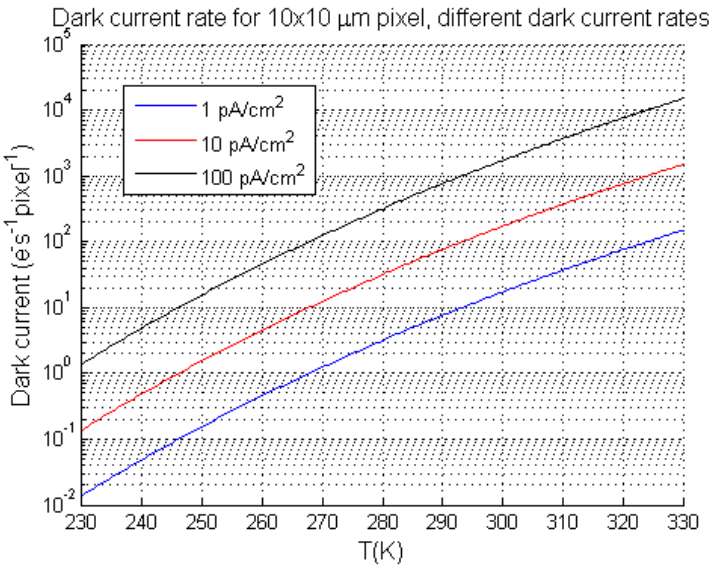

Figure 4.1: *Dark current rate for different sizes of the dark current (pA/cm^2, often supplied by the manufacturer of the sensor), for different N_{dc} (formula 4.5.1).See dcplot.m for generating this plot.*

for normal integration times, FPN is not the dominant pattern for most arrays (Ch. 6.5.6 of [6]; see also §4.9).

Due to the much greater design variability in CMOS image sensors, there is no general equation for the dark current for these types of image sensors.

4.6 Fixed Pattern Noise

Some of the variations due to dark current are somewhat systematic: the spatial pattern of the noise remains constant. Because of fabrication and material properties, this 'fixed pattern noise' (FPN) is a flat field uncertainty due to device response when the sensor is not illuminated. Crystal defects, impurities and dislocations may contribute to the size of the fixed pattern noise. The defects result in a non-uniform dark current, as it is easier for thermal electrons to move into the conduction band in the presence of defects or impurities (see appendix B). This results in non-uniform potential wells, capturing a variable amount of thermal electrons.

In a CMOS image sensor, additional sources are present such as a different initial precharged voltage of the photodiode, e.g. due to V_{th} variations in M_1. This results in an offset value of the pixel. Variations in the characteristics of other transistors, such as M_2 and M_3, also result in a spatial variable output (see Ch. 3). FPN can be easily removed by subtracting a dark frame. As FPN is present in the PRNU pattern as well, the sources responsible for the FPN play a role in the extracted pattern from the images as well.

The defects present in the wafer, such as voids, or the presence of oxygen in the silicon, result in stress and may result in various noise components (e.g. dark current spikes or flicker noise). These defects appear during the growth of the silicon crystal (Czochralsky growth), and deform the silicon lattice. The structural imperfections are called microdefects, and have multiple origins:

- Atomic vacancies, where an atom is simply missing from its regular site (void),
- Self-interstitials, where a (silicon) atom occupies a site in the lattice where in a perfect crystal it would not be expected,
- Substitutional defects, where an impurity atom substitutes a silicon atom in its regular position,
- Impurity-interstitials, where an impurity atom (such as oxygen, see [31]) is present between the regular silicon lattice atoms.

The dynamics of the aggregation of these defects and their interactions cause clusters to form. When oxygen atoms aggregate, for example, clusters of silicon dioxide are produced. These 'O-clusters' cause stress in the lattice, as the clusters occupy a larger volume than the silicon atoms. A model for the formation and growth dynamics of these clusters is proposed in [30].

In [19] a relation is found between stacking faults (when the lattice is interrupted from its regular pattern due to a changing sequence of planes; e.g. in normal lattice ABCABCABCABC, in which each letter represents a certain configuration of atoms). When a stacking fault is present, the lattice could be ABCABCBCABCA (Ch. 4.6 of [22]) and result in dark current non-uniformity. Although this is a very old paper, it still indicates which possible sources could result in non-uniformities.

E Although the dark current responsible for the FPN is time-dependent, the noise due to dark current (the average pixelvalue and the standard deviation of a small region in the image) in the Canon EOS 10D depends only on the ISO value and not on the integration time. This was checked by running the *average.m* script which calculates the mean pixelvalue of a small area of frames with different integration times and ISO values. This is probably due to automatic subtraction of the average dark pixelvalue of shielded pixels (i.e. pixels that are normally not illuminated, also known as dark reference pixels, which are present near the corners of the sensor[7]). In this way it is also possible to remove the FPN by subtracting a complete frame. A higher ISO value

[7]When RAW images are inspected, the resolution is higher than the resolution obtained from images that are already converted to JPEG, the difference in resolution is due to the reference pixels.

results in more deviation from the average, as less electrons originating from the dark current shotnoise are now needed to increment the digital output with 1. □

4.7 A/D converters and correlated double sampling

In a CCD, after the charge packets have been transferred from the shift register to the amplifier in the silicon substrate to increase the strength of the signal, the signal is fed into the integrating amplifier. The output of this last amplifier is fed into the A/D converter. This A/D converter converts the analogue voltage into a digital number (DN) or analogue-to-digital unit (ADU). In a CMOS image sensor, the charge packets can be read out per pixel. In both systems the sense node (n^+ in figures 2.4 and 3.3) is set to a reference level. The charge transferred from the adjacent gates partially discharges this capacitor. By comparing the initial (reference) voltage to the final voltage, we know how much charge was present in the respective gate. The larger the amount of electrons present, the larger the output is. After a charge packet is read, the sense node is reset to the reference value. There is a fluctuation in this reference value, which can be seen as noise in the output. To prevent this, a technique is used called 'correlated double sampling', a differential technique used to establish the signal voltage: the reset-voltage is compared with the reset voltage plus the charge from the integrating amplifier (it acts as a differential amplifier). In this way, noise present due to the reset noise is common to both inputs, and can therefore be removed. This largely eliminates the read noise as well as the flicker noise.

The gain determines how the output electronics assign a DN to a charge value from a pixel. The A/D converter is able to supply a DN in the range of $0 - (2^N - 1)$, in which N is the amount of bits available in the ADC. E.g. when the gain G is 50 e^-/ADU it means that if 50 electrons are found in the pixel, that the A/D converter gives a DN of 1. When 80 electrons would be found to reside in the pixel, the output would still be 1. We immediately see a limitation in the device: when two pixels appear to have the same response, they need not have had the same amount of electrons, which means more (or less) light may have fallen on one of the pixels. There is always a limitation in the discrimination of exactly G electrons (see §4.8). Raising the ISO value results in the lowering of the gain G. In other words, the device becomes more sensitive to light as the ISO is raised. When the ISO is raised from 100 to 200, the gain decreases to $G/2$. This results in a reduction of amount of useful charges as well, as less electrons are needed to reach the maximum ADU.

If we multiply the gain at the lowest ISO with 2^N we have a rough estimate of the full well capacity. It is only a gross estimate because various situations may occur, such as

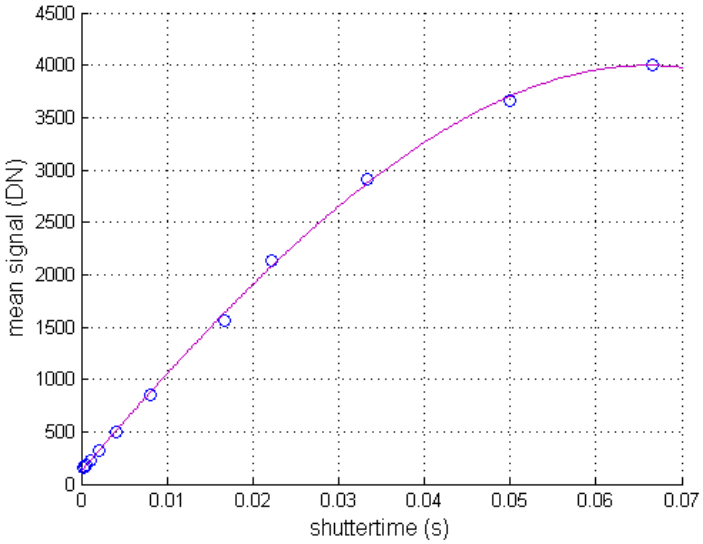

Figure 4.2: *Integral nonlinearity. For various reasons, the output may not be a linear function of the input. By varying the input signal (by varying the integration time) and subsequently calculating the mean signal of a small area of the output image, we obtain this plot. See linearity_1.m.*

nonlinearity or saturation. The image sensor is however linear over most of its input values, as we will see below. In §4.15 we will see how the gain can be calculated. For the Canon EOS 10D, the gain is found to be 11 e^- / ADU resulting in a full well capacity of approximately $45 \cdot 10^3$ electrons.

Although it is strictly not considered noise, there are two types of nonlinearity in A/D converters that may result in a different than expected response: integral nonlinearity (INL) and differential nonlinearity (DNL) (see Ch. 3.8 of [9]). DNL is also known as digitization noise. E.g. in the case of $G = 50e^-$/ADU, 50 electrons will give a DN of 1, but so will 99 electrons. This means the DNL is $\pm$ 0.5 ADU, or in our case $\pm$ $25e^-$.

E INL is the deviation from the linear relation you expect from the output ADU vs. input electrons. One expects when the amount of electrons is doubled in the well that the ADU will be twice as big. However, this is not the reality especially at higher input levels, and the deviation is called the INL. The behaviour of the Canon EOS 10D can be seen in figure 4.2, where by simply increasing the exposure time and calculating the average signal of a small area this nonlinear behaviour can be seen. □

We do not know if the nonlinear behaviour comes from the design of the sensor (appendix A), or that some other process is responsible. Indeed, for consumer applications it is not necessary that the device has a linear response. A cosmetically pleasing image is more important; it is a feature that above the white clip level some contrast is still

present, by introducing a knee point (Ch. 3.6 of [6]). When the input signal is larger than this knee-point, the gain is lowered. Note that when calculating the gain (§4.15), signals have to be used that are within the linear region of the device.

It is also possible for the A/D converter to saturate, when the full well capacity has not yet been reached, but the ADU given by the A/D converter cannot increase anymore. This can easily be observed. However, it is also possible to come into the nonlinear region of the CCD, and this cannot be observed. It may occur before the full well capacity was reached, or before A/D saturation takes place.

4.8 Quantization noise

When the ADC converts the analogue voltage to a digital number, only a limited amount of discrete levels are available, limited by the amount of bits N in the ADC. This means that this quantization results in noise, i.e. a deviation from the true value expected from the amount of photons detected. In general, the variance of a distribution is the square of the difference between the outcome and the expectation. For continuous functions:

$$\sigma^2 = E(x-\mu)^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx. \tag{4.8.1}$$

The least significant bit determines whether the output is odd or even. Because the levels are quantized, not every value can be represented. The quantization level q can be roughly calculated by dividing the full well capacity by 2^N. Specifically, the error lies between $q/2$ and $+q/2$, and it is expected that each value is present with the same probability, so the mean value is zero. Then [23]:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \int x^2 f(x)dx = \int_{-q/2}^{+q/2} x^2 \frac{1}{q} dx = \frac{1}{q}\frac{1}{3}x^3\Big|_{-q/2}^{+q/2} = \frac{q^2}{12}. \tag{4.8.2}$$

When we assume for simplicity a linear response up to the full well capacity then $\sigma = \frac{FWcap}{2^N\sqrt{12}}$, in which the full well capacity is approximately $11 \cdot 4095 \approx 45 \cdot 10^3$ electrons for the used camera, in which 11 is the gain G (§4.15). Then the quantization noise is $\frac{45\cdot 10^3}{2^{12}\sqrt{12}} \approx 3.2$ electrons for the 12-bit ADC of the EOS10D. This is the uncertainty in the output, in electrons. In ADU/DN this has to be divided by the Gain. We will see this is well below the value for read noise, which is approximately on the order of 20 electrons.

4.9 PRNU

A source somewhat similar in characteristics to FPN is PRNU, the variation in pixel response when the sensor *is* illuminated. This variation comes e.g. from non-uniform

pixel sizes, non-uniform thickness of the oxygen, increased leakage current, and other sources indicated in the CMOS/CCD chapters (i.e. the factors that are responsible for variations in responsivity when a signal is applied). When the sensor is illuminated, the device will still undergo the effects of FPN; therefore, FPN is included in the PRNU as well. When the scene contrast is low, the effects of the non-uniformity may be visible. In principle it is possible to reduce PRNU by a form of non-uniformity correction (NUC). Calibration based techniques measure the relative response of the pixels in an image sensor at homogeneous illumination, by studying the pixel response under different exposure levels. Different NUC algorithms exist to reduce the non-uniformity, and even to reduce the nonlinear behaviour of the sensor (see §4.7 and appendix B). In [25] a method is proposed to first reduce the nonlinearity and subsequently measure the relative response of the pixels at a single exposure level. The authors experimentally tested this procedure and it was found that the uniformity increased by a factor 20.

The size of the PRNU is often given as the 1σ deviation for a e.g. 128x128 sample when the sensor is lit uniformly at a specified intensity level [24]. Sometimes it is given in a percentage representing the difference between the maximum and minimum signal compared divided by the mean signal. As this definition overestimates the size due to possible defects, the former definition is used. The easiest way of calculating the size of the PRNU is by averaging a number of flatfields to remove the shotnoise, and divide the variance of the resulting image by the mean (see also §4.14). The total 'noise' from PRNU and FPN is

$$\langle n \rangle = \sqrt{(U_{PRNU} n_{PE})^2 + (U_{FPN} n_{dark})^2}, \tag{4.9.1}$$

In which n_{PE} is the amount of photoelectrons, n_{dark} is the number of dark current electrons, and U_{PRNU} and U_{FPN} are represented as fractions, such that e.g. $U_{PRNU} = 1\%$ means that if a signal of 10.000 electrons is applied, the 'noise' in the pattern (in the aforementioned definition) will create a standard deviation of approximately 100 electrons. The implications of this depend on the specifics of the circuit, e.g. the gain. In Ch. 6.5.6 of [6] it is noted that the size of the PRNU in CMOS applications is generally larger than in their CCD counterparts basically due to the large influence of the combination of multiple transistors combined on 1 pixel (the row select transistor M_3 in combination with the source follower M_2). Another possible source is mentioned as because the sense node in the photodiode in the case of the CMOS is integrated in the pixel, the room available for this node is small. This results in a small 'capacitor' which capacitance changes with signal level, resulting in a nonlinear response called 'floating diffusion nonlinearity' (see appendix A). In a CCD, this 'sensing' is done off-chip where the area available is larger and hence the capacitance of the sense node changes

less (hence the nonlinearity is smaller). The result of this nonlinearity component is that the relative output decreases with increasing charge. Especially in cheap devices using CMOS sensors the result of this may be clear.

Note that compressing the image with JPEG decreases the PRNU pattern, as JPEG compression removes high frequency components in the image.

4.10 Read Noise

Read-out noise is a collective term, which is fortunately quoted by manufacturers in terms of the number of electrons added to the number of electrons already present in the pixel upon read-out. Read noise is generally of the order $10 - 30$ electrons, but high end cameras may have lower read-out noises. Due the varying of the reset voltage, the output of the A/D converter is not constant even when a constant signal is applied. This problem is largely mitigated through the usage of correlated double sampling (§4.7). Therefore, the DN output of the A/D converter will change, even when the charge value in the same pixel is the same. This uncertainty in the output even when the integration time is zero, is shared under the collective term 'read noise'.

The second source of read-noise comes from the electronics themselves. Large and complex integrated circuits have varying capacitances, resistances, inductances, etc. If we think of the image sensor as a circuit consisting of a number of equivalent capacitors and resistors, then there is an equivalent capacitor for the depletion region(s), as the separation between the charges in the depletion region and the rest of the substrate produces a natural capacitor. Another capacitor is present due to the oxide, which is sandwiched between the substrate and the gate. The channel itself can be modelled as a resistor, as the electrons moving towards the sense node feel the difficulty of moving through the material. Electrons moving through this resistor toward the sense node can be modelled as a Poissonian noise source. This is the thermal noise due to temperature fluctuations in the equivalent resistor; in the presence of a capacitor, this thermal noise takes the form of kTC noise (hence, kTC noise and thermal noise are the same things). In this way, the CCD has an equivalent circuit which may be represented as a resistor in series with a capacitor. Similarly, in CMOS image sensors the noise from resetting the photodiode to a reference value is caused by thermal noise which occurs due to the resistance present in M_1.

The size of the reset noise due to the presence of a resistor and transistor is, remembering that

$$C = \frac{Q}{V}, I = \frac{dQ}{dt} = C\frac{dV}{dt}, \tag{4.10.1}$$

then

$$\langle I \rangle = \frac{C}{t'} \int_0^{t'} \frac{dV}{dt} = \frac{C}{t'} \Delta V \Big|_0^{t'}, \tag{4.10.2}$$

where t' is the integration time, so that

$$\langle I^2 \rangle = C^2 \langle \Delta V^2 \rangle. \tag{4.10.3}$$

The energy in a capacitor is $\frac{1}{2} C \Delta V^2$, and from statistical mechanics we know that it needs to be equal to $\frac{1}{2} kT$ (equipartition theorem, the voltage represents one degree of freedom), so that $\Delta V^2 = \frac{kT}{C}$, and hence that $\langle I^2 \rangle = kTC$, or $\langle nq \rangle = \sqrt{kTC}$, so $\langle n \rangle = \frac{\sqrt{kTC}}{q}$ (Ch. 6.5.2 of [6], [17])[8] [9]. These fluctuations in the amount of electrons travel through the circuit.

This value of the read noise is greatly reduced with the usage of correlated double sampling, and very low read noises are possible. This noise is present in a CMOS APS as well, as the depletion layer, oxide layer and the channel resistance for the electrons to come to the sense node, are present as well in the photodiode. However, after this step the electrons have to go through a number of transistors before they reach the column circuit. Therefore, it is more difficult to find the approximate readnoise in CMOS APSs.

Construction, temperature, sensitivity of the amplifier ($\mu V / e^-$) all contribute to the read noise of the CCD. Sudden temperature variations, occurring at high read-out speeds can even change the read noise (Ch3.4 of [9]). The read noise can be determined by considering bias and flat field frames. Put simply, when the sensor is not

[8] The actual value is actually half the value that is quoted here, due to the fact that the reset times are in the order of microseconds, but the settling time is on the order of ms. Therefore, no steady state is reached, and this simple analysis is not valid. See [26].

[9] More advanced calculations with the same results are available. When the electrons move through the equivalent resistor, they become agitated. The r.m.s. voltage of this thermal noise is (also known as Johnson noise, Nyquist noise, or a combination), $V_n^2 = 4kTBR$, with B the bandwidth. Each frequency is equally probable, or put differently: it is white noise. The equivalent circuit of the parallel combination of the resistor with the capacitance results in an impedance of

$$\frac{1}{Z} = \frac{1}{R} + j\omega C, \text{ so that } |Z| = \left| \frac{R}{1 + j\omega RC} \right|^2 \tag{4.10.4}$$

Because the number of electrons fluctuates in the resistor, they also fluctuate on the capacitor. The noise voltage over this capacitor is:

$$V_{cap}^2 = \int_0^\infty \frac{4kTR}{(1 + \omega^2 R^2 C^2)} df, \tag{4.10.5}$$

with $\omega = 2\pi f$, and $\int \frac{1}{a^2 + x^2} dx = \frac{1}{a} \tan^{-1}(\frac{x}{a}) + C$. Then

$$\int_0^\infty \frac{4kTR}{1 + \omega^2 R^2 C^2} df = 4kTR \int_0^\infty \frac{(1/2\pi RC)^2}{(1/2\pi RC)^2 + f^2} df = \frac{4kTR}{w\pi RC} \int_0^\infty \frac{1}{(\frac{1}{2\pi RC})^2 + f^2} = \frac{kT}{C}. \tag{4.10.6}$$

As above, $\langle I^2 \rangle = C^2 \langle \Delta V^2 \rangle$, from which $n = \frac{\sqrt{kTC}}{q}$.

illuminated and the exposure time is approximately zero seconds (hence no dark current will be present), the read noise can be found.

We can never be truly sure about the amount of read noise, as the discretization from the A/D converter is a lower limit. Although the read noise decreases by the advances in CCD/CMOS production, the pixel sizes shrink as well. This means that some noise sources may decrease, e.g. dark current noise depends on pixel size (smaller wells, etc.), but the read noise will essentially have the same magnitude. So the relative size of the read noise may increase.

There may be a time variation in the zero-level, but this occurs only in a time frame of months (Ch. 3.7 of [9]). Dark current, combined with the read noise presented here and the flicker noise, provides a lower level of the noise floor.

E The tested camera reduced the apparent readnoise after the RAW files were converted with Canon's software in the same way as the influence of the dark current was reduced (§4.5). This can be seen by inspecting the individual pixelvalues; at various points the values were set to zero, while other (neighbouring) pixels had values up to 25. This behaviour was not seen when the files were converted with *dcraw*. □

4.11 Flicker noise

Flicker noise, also commonly called $1/f$ noise, is a noise source present in 'practically all electronic materials and devices', and is still a resource for an ongoing debate. The noise is generated, according to one model due to *"fluctuations in the number of carriers due to charge trapping in surface states"*. The traps capture electrons, and after some time they release the electrons. The time between capture and release is variable, and this causes the characteristic $1/f$ spectrum. Another explanation is the *"fluctuations in mobility"* in the bulk of the MOSFET [20]. This mobility also decreases with increased doping concentration and increasing temperature. By decreasing the length of the transistor through which the electrons need to travel (drain-to-source distance), this component is decreased.

Often, this noise source is not specifically mentioned because it is included in the read noise as the electrons need to travel through the sensor when the charge is being read. When the surface area decreases (i.e. smaller pixels) or when the current is reduced, the influence of $1/f$ noise decreases. Flicker noise is also a component in read noise and is present both in CCDs and CMOS sensors.

4.12 Noise sources and failure common to CCDs and CMOS APS

Note: this is only a very brief summary of multiple effects. For more elaborate explanations and references, see Ch.2 of [27], and the references mentioned therein.

Due to miniaturization and more on-chip processing, fabrication requires more strict guidelines to prevent unwanted behaviour. When, in general, the physical sizes of nMOS devices shrink, then it is not always possible to reduce the other physical parameters, such as the threshold voltage V_{th}. When devices shrink, the electric field actually grows when the thickness of the insulating layers (oxide) shrink. As said, reducing the voltages is not always possible as this changes the behaviour of the device. Due to these high electric fields, carriers may gain a lot of energy which can then be injected into the gate oxide, eventually causing a gate current (See also [28]).

The thinner the oxide films are, the higher the electron energies will be that tunnel through the oxide. These 'hot' electrons can lead to a degradation of the threshold voltage V_{th}, transconductance g_m, and subthreshold current I_{DS} (Ch. 2.2.2.7 of [11], and Ch. 2.2.1 of [27]). We can see the transconductance is lowered (in a NMOS device) by looking at the following formula:

$$g_m(V_{DS}) = \left[\frac{\partial I_{DS}}{\partial V_{GS}}\right]_{V_{DS}} \propto C_{ox} V_{DS}. \tag{4.12.1}$$

If hot carriers are injected, the capacitance C_{ox} reduces, resulting in a lower g_m. This, in turn, results in a lower I_{DS}. We can see V_{th} increases as electrons in the oxide 'shield' a part of the electric field needed to form a channel. Therefore, a higher voltage is needed to repel the holes in the substrate to form a n-channel [29].

In the opposite direction, it is possible (when the gate oxide is thin) for electrons to flow from the gate to the source by electron tunnelling[10]. This is called Fowler-Nordheim tunnelling.

Time dependent dielectric breakdown occurs after continually applying stress to the gate oxide film. Initially, this effect occurs sporadically, but after a longer period all samples may break down, ultimately determining product life (Ch. 2.2.2 of [27]).

In [33] mechanical stresses in the silicon are measured. For the effects of stress migration (atoms migrating due to stress that occurs e.g. due to different temperature expansion coefficients, and possibly accumulate into clusters), electromigration (atoms that are displaced from their regular site due to the collision with high energy electrons), and contamination (dust or metal atoms present in the wafer or in the gate oxide), see

[10] The increasing E-field due to thinner films, results in higher electron energies that tunnel through the oxide, leading to more defects in the oxide. This increases the leakage currents by a significant amount.

Ch. 2 of [27].

Sometimes people report that the pictures the camera is producing loses sharpness after some time. People often mistake sharpness with contrast (Ch. 5 of [34]). The most likely explanation for this effect may be the decomposition of the pigments in the colour filter array (Ch. 2.4.8 of [27]). When the dyes decompose, a red pixel may, for example, absorb photons that should have been absorbed by the dye. As the colour gain (§5.3) that is applied after the colours are interpolated is constant, contrast may be reduced.

4.13 Residual noise

E Due to temporal variations which we do not want to include in our estimation of the PRNU pattern, we take the following approach, as suggested by [35][11]. An average of a variable amount of frames is subtracted from one individual test frame: $\tilde{I} = I - \frac{1}{n}\sum I_k$. Subtracting these frames results in the removal of the FPN/PRNU, while leaving the shot noise present. Calculating the standard deviation from the resulting frame $\tilde{I}$ converges to a final value. As said, in principle a large amount of frames is needed to completely remove the variable noise sources. In practice it can be seen that a small amount of frames (10-15) is already sufficient for the standard deviation of the final frame to become constant, as can be seen in figure 4.3. □

In Ch. 3.2 of [5] it was found that a much larger amount of images was needed (300) for the pattern to stabilize. It has to be kept in mind that we used uncompressed TIFF images. The pixelvalue of a specific pixel in a JPEG block ultimately depends on the pixelvalue of its neighbours (in the same block). This means that in order to correctly find the deviation of a specific pixel value, the value of its neighbouring pixels needs to vary as well (by varying the illumination or angle with which the images are obtained). Furthermore, the requisition for convergence in the aforementioned report is very strict. The convergence value in the aforementioned report reaches its final value within 5% after averaging approximately 30 images.

If a particular sensor has a PRNU size (often defined as the standard deviation of a (small) sample) of P, then at a signal S (DN) the standard deviation of the output is SP (ADU). The shotnoise is approximately $\sqrt{S/G}$ (ADU) for a single frame and $\sqrt{S/nG}$ for n frames. Put simply, if we want the noise from the shotnoise to be a times as small as the 'noise' from the PRNU, then we can find (at a particular S,G and P) the amount of frames we need to average, viz. $n = \frac{a^2}{GSP^2}$. In particular, if we want the noise of the

[11] *This approach is implemented in the residnoise2.m script*

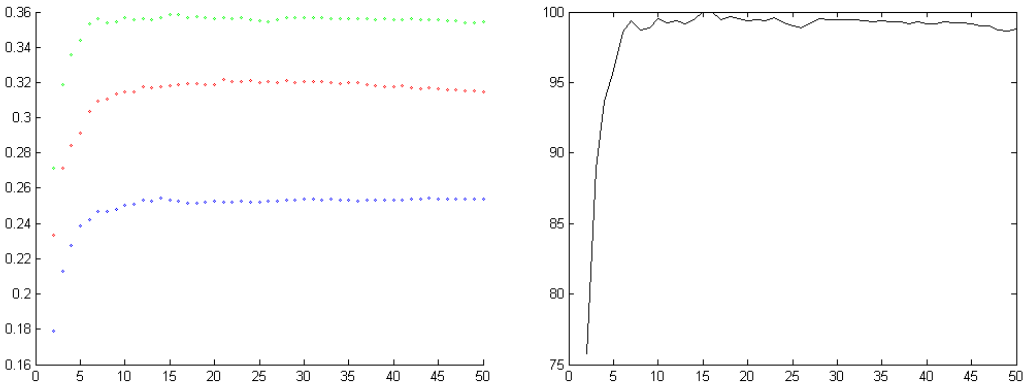

Figure 4.3: *Standard deviation of $I - \frac{1}{n}\sum I_k$ versus the amount of frames averaged. The standard deviation stabilizes already at a low amount of averaged frames, approximately 15 (a). The percentage of the size of the standard deviation in terms of its final value (b). We see that at approximately 15 frames the percentage stabilizes.*

PRNU to be five times as large as the shotnoise, then when $S = 1000$ (ADU), $P = 1\%$, and $G = 11$, n needs to be 23 or larger. It can be seen that with increasing signal, the amount of frames that need be averaged decreases.

4.14 Photon Transfer Curve

As was mentioned briefly before, it is possible to divide the noise regions in three domains according to the slope in a log-log plot of the standard deviation at varying illumination levels. The photon transfer curve is obtained by taking multiple images of varying illumination levels, and subsequently calculating the mean and standard deviation of a small area in the centre of the frame.

The first region is where the system is said to be read noise limited. The read noise is present at all illumination levels as it is independent of the input signal, but only at very short integration times it is the dominant noise source. Therefore, at short integration times (when the mean signal is low) the standard deviation due to the read noise does not change appreciably, and hence a slope of 0 is found. When the integration time is increased, more light is incident on the sensor, and the resulting photon shot noise increases as well. In this region the slope is $\frac{1}{2}$, as the Poisson light has the property that the uncertainty (noise) is proportional to the square of the (mean) number of detected photons S ($\sigma_s = \sqrt{S}$, §4.4). At higher signals the influence of the multiplicative PRNU can be seen. As the size of this 'noise' increases linearly with the incident light ($\sigma_p \propto S$), the slope of this region is 1. Of course, the three regions gradually overlap as the influence of each noise source is always present, but the relative magnitude changes when

the (average) signal is increased.

E This curve is obtained experimentally and is plotted in figure 4.4, where the mean signal of a small area is plotted vs. the standard deviation of this small area of the image in a log-log plot (blue curve). We can see the first region to have a slope of approximately 0, and at these low (mean) signals the system is indeed read noise limited. The slope is a little larger than 0, as when the integration time is raised even slightly photons are incident on the sensor and the noise due to these photons increases. However, as the amount of photons is still very small, the main noise contribution still originates from the read noise. When looking at the blue curve, it is possible to determine the approximate readnoise. This curve intersects the y-axis at approximately 2.1, which, with a gain G of approximately 11 this results in a readnoise of approximately 23 electrons. This agrees with the value found in (§4.15).

The middle part of the blue curve represents the domain in which the system is shot noise limited. In this region the slope is indeed $\frac{1}{2}$, as expected from the statistical nature of light. Again, the slope changes gradually from 0 to $\frac{1}{2}$ (and later on from $\frac{1}{2}$ to 1) as the influence of each noise source is always present, but it is the relative magnitude of each noise source that changes. When the (mean) signal is increased even more, the slope changes to 1, the region that is characterized by the multiplicative PRNU.

The red curve represents the standard deviation of a difference frame, in which two frames with the same integration time are subtracted: $\sigma_{I_1-I_2}$. As was mentioned in §4.2.1, shot noise increases with a factor $\sqrt{2}$ when 2 frames are being subtracted. Comparing the red curve with the blue curve indeed shows that the standard deviation is increased with approximately a factor $\sqrt{2}$ compared to the blue curve. However, when the (mean) signal increases, this is not the case because the PRNU starts to dominate the photon shot noise.

The black curve shows the theoretical slope if only shotnoise was present (hence this curve has a slope of $\frac{1}{2}$; as the PRNU is a constant systematic pattern, subtracting two frames (that were illuminated for the same amount of time) will remove this PRNU, and only increase the shot noise.

Finally, we see a fourth region, but this is a limitation of the design of the system, and not of noise sources. As the signal becomes larger, the signal saturates so that the maximum amount of electrons is registered. There are no levels available to represent these signals, so that the standard deviation goes to 0. □

Strictly, by varying the integration time (as is done to get the blue curve in the figure) the amount of dark current rises as well. As mentioned before, this is of no concern as the amount of dark current is negligible (and the camera tested here also uses a form of

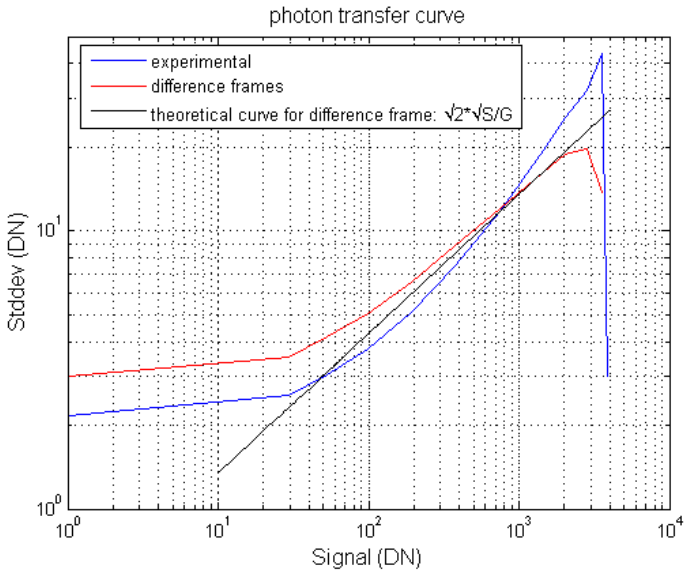

Figure 4.4: *For different illumination levels flatfields are taken, after which the standard deviation and the mean are calculated for a small area (400x400) in the green channel (this is done for 4 individual images after which the mean and standard deviation calculated for each frame are averaged). Care should be taken that the average of the biasframe is subtracted (the biasframe increases the signal level with approximately 128 in the camera used (Canon EOS 10D); i.e. when the sensor is not illuminated and the shuttertime is set to the lowest possible value, an average signal of 128 is still present.). In this figure, the mean vs. the standard deviation is plotted in the blue curve. This curve can be obtained with this method by using the prnusize.m script in multiple directories (in which images are present with different exposure times). The red curve represents the standard deviation of a difference frame ($\sigma_{I_1-I_2}$), and should (when shot noise dominates) be a factor $\sqrt{2}$ larger than the blue curve. The black curve is the curve representing the situation when only shot noise would be present.*

dark current reduction, but this of course does not remove the dark current shotnoise). For example, when the dark current is 40 $e^-/p/s$, then at normal integration times (e.g. 1/20s) this is negligible as the gain is approximately 11 e^-/ADU. The photon shotnoise is likely to be larger than that[12].

In principle the size of the PRNU can be obtained by comparing images at different illuminations from this figure. When N images are averaged (at a certain reasonable illumination where the photon shotnoise is the dominant factor) the noise decreases by a factor $\sqrt{N}$ resulting in a more clear PRNU signal. The size of the PRNU is then

[12] At larger integration times the ISO value can be adjusted to higher values if possible (lower gains are effective), which at first sight increases the amount of dark current though the photon shotnoise increases too.

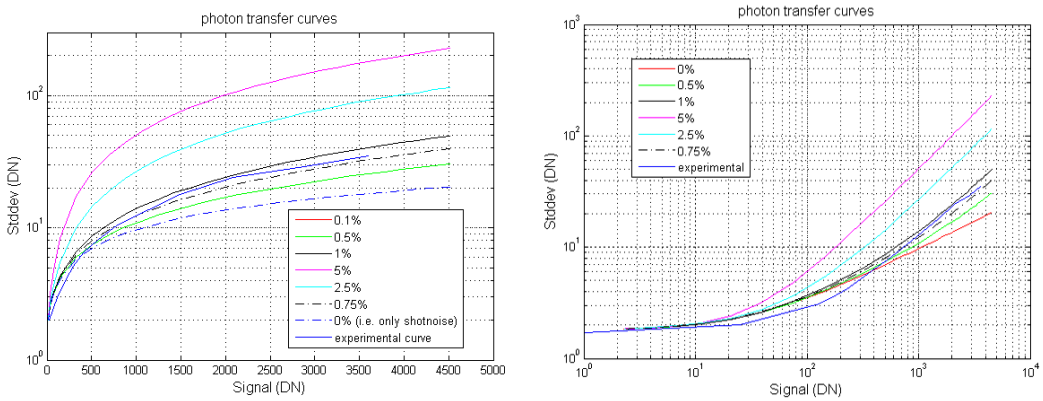

Figure 4.5: *Comparison of the experimental curve (blue) with the theoretical model for different PRNU sizes for (a) linear-log axes and (b) log-log axes. (parameters: $\sigma_{read} = 20$, $\mu_{read} = 30$). The experimental curve is between 0.75% and 1%. The blue curve is below the theoretical curves in (b) for the same reason mentioned above. Note that in (a) when no PRNU is present the curve follows a typical square root line. The theoretical curves are generated with the totalnoise.m script, the theoretical model of the main noise sources. The calcprnusize.m script is used to calculate the size of the PRNU by comparing the theoretical curves in the absence of PRNU with the experimental curve (in which PRNU is of course present). A value of 0.9% is found.*

calculated as

$$\frac{\sigma(\frac{1}{N}\sum I_{A,k} - \frac{1}{N}\sum I_{B,k})}{\mu(\frac{1}{N}\sum I_{A,k} - \frac{1}{N}\sum I_{B,k})},$$

in which A and B represent different illuminations, and σ and μ represent calculating the standard deviation and mean, respectively.

E This was calculated for the Canon EOS 10D (green channel, 200x200 region in the centre), for a series of frames of different illumination, and a PRNU size of approximately 1% resulted. □

4.14.1 Model

E A theoretical model of the main noise sources can be made, in which the shotnoise is modelled as a Poisson distribution, the read noise is modelled as a normal distribution and the PRNU is linear with the signal. The input is generated by number generators according to their distribution. This input is then processed in a 'ADC' with a constant gain with resulting quantization noise, after which the standard deviation and the mean are calculated. These theoretical curves are plotted together with the experimental curve in figure 4.5. From the figure it can be seen the experimental curve is between 1% and 0.75%. □

Strictly the found PRNU sizes is a lower limit, as possible blur- and low-pass filters mentioned earlier reduce the size of the PRNU (low pass filters may be present to prevent high energy photons from reaching the sensor, possibly damaging it). A blur filter may be present to prevent moiré patterns ('spatial aliasing'), which occur when the details of the image projected on the sensor have approximately the same frequency as the pixels in the sensor. Another reason why the PRNU size is a lower limit is because the gain decreases with increasing signal level. Therefore, the size calculated with images with high (average) signals may actually be a little higher. The size found (0.9%) is an average.

4.15 Finding the size of the gain and read noise from images

In principle the gain can be calculated from images using the following formula:

$$G = \frac{\langle S \rangle}{(\sigma_S^2(DN) - \sigma_{read}^2(DN))}, \tag{4.15.1}$$

in which $\langle S \rangle$ is the mean signal, σ_S^2 is the signal variance and σ_{read}^2 is the variance of a biasframe (readnoise), all in DN. The gain is expressed in (e^-/ADU), i.e. the amount of electrons needed to increase the output with 1. The images should be taken from frames that have a reasonable illumination (i.e. not saturated, and not in the nonlinear region). As can be seen in §4.7, the linear region is only a limited part of the complete response of the sensor. At ISO 100, the images used had shuttertimes of 1/250 and 1/500s. This resulted in a calculated gain of approximately 11e^-/ADU.

E The readnoise can be calculated as

$$\sigma_{read} = G\frac{\sigma_{read_1 - read_2}}{\sqrt{2}}, \tag{4.15.2}$$

i.e. simply by calculating the standard deviation of the difference frame obtained by subtracting two bias frames (i.e. frames with an exposure time of approximately 0 seconds).

The division by $\sqrt{2}$ is to find the noise in an individual frame, and is multiplied by the gain to find the readnoise in equivalent electrons (Ch. 4.3 of [9]). This resulted in a readnoise of approximately 21 electrons[13]. □

Note: the dynamic range can be calculated, if we assume the full well capacity of the cameras to be matched to the maximum DN the ADC can represent($2^{12} - 1$), and we

[13] *See calcgn.m script for both the calculation of the gain and the read noise.*

find a dynamic range of $20\log_{10}((2^{12}-1)\frac{G}{R}) = 67$ dB, with gain $G = 11$, and readnoise $R = 21$.

4.16 Periodogram

Earlier we stated that when the charge is being read out a few columns may be multiplexed (the output signal of a number of columns are being fed into the multiplexer after which 1 signal is processed at a time). Some of the components in the CMOS APS are shared between a number of columns (in the case of a multiplexed signal). In a column circuit in the APS certain transistors such as the bias transistor (M_4) and components (such as the column ADC) are shared with other rows. In general the row-wise and column-wise operations may result in a periodic variation in the output of the sensor [10]. As this was only tested for CMOS APS we do not know if this variation occurs in CCDs as well.

E The frequency of this periodic variation is detected in dark frames as follows. A certain amount of images is averaged (40 images are averaged here), after which the column averages of this averaged image are calculated. As these averages show a trend (the column averages are higher near the edge of the chip, possibly due to non-uniform temperature), this trend is removed[14]. To see a variation, this matrix is transferred to the frequency domain with a Fourier transform, after which the absolute value is computed. This value is plotted after it is normalized[15][35]. Technically this is called a periodogram. □

Although this shows that the columns in the tested APS are not uncorrelated, this does not mean individual pixels are correlated. Indeed, the sources for the variation in the column components and the sources for variations in the pixel components have different origins. In [10] a statistical model is described representing the fixed pattern noise in APS sensors, from which is concluded that the pixel components are uncorrelated. In other words, the periodical variation in column components does not influence the independence of the individual pixels. In the case of CCD image sensors, the pattern noise is modelled as a sample from a spatial white noise process. The authors note that the pattern noise is mainly due to the varying photodetector area and the dark current non-uniformity, two processes that are spatially uncorrelated. Concluding, the pixels

[14] In Matlab this can easily be done with *detrend(M, 'linear', [1:8:size(M)])*, in which *M* is the matrix which holds the averages, and the final vector represents the breakpoints.

[15] *This approach is implemented in the periodogram.m script.*

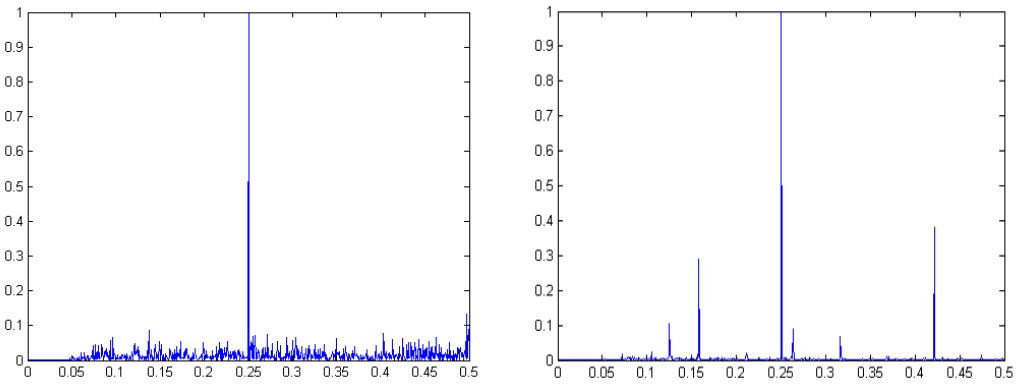

Figure 4.6: *Power-frequency plot (a) without detrending and using 40 images for calculating the average, and (b) with detrending and using 20 images. In (a) a periodic variation every 4 rows is found. In (b) more periodic variations are found, but these do not correspond to integer values. It is not clear what causes this.*

in both CCD image sensors and CMOS active pixel sensors are uncorrelated.

4.17 Image noise

After extracting the PRNU pattern from a natural image, a residue of image content is present in the pattern. This residue is of course not characteristic to the sensor, and hence can be seen as a noise component in the pattern. This is illustrated in figure 4.7.

The filter used to extract the PRNU pattern from the image is described in [5]. Comparing this pattern with the pattern extracted from (averaged) flatfields from e.g. the suspect's camera results in an uncertainty. This is the reason why low detailed images (such as flatfields) are used to extract the PRNU patterns. Improving the filter may reduce this noise component, and improve the reliability of the pattern.

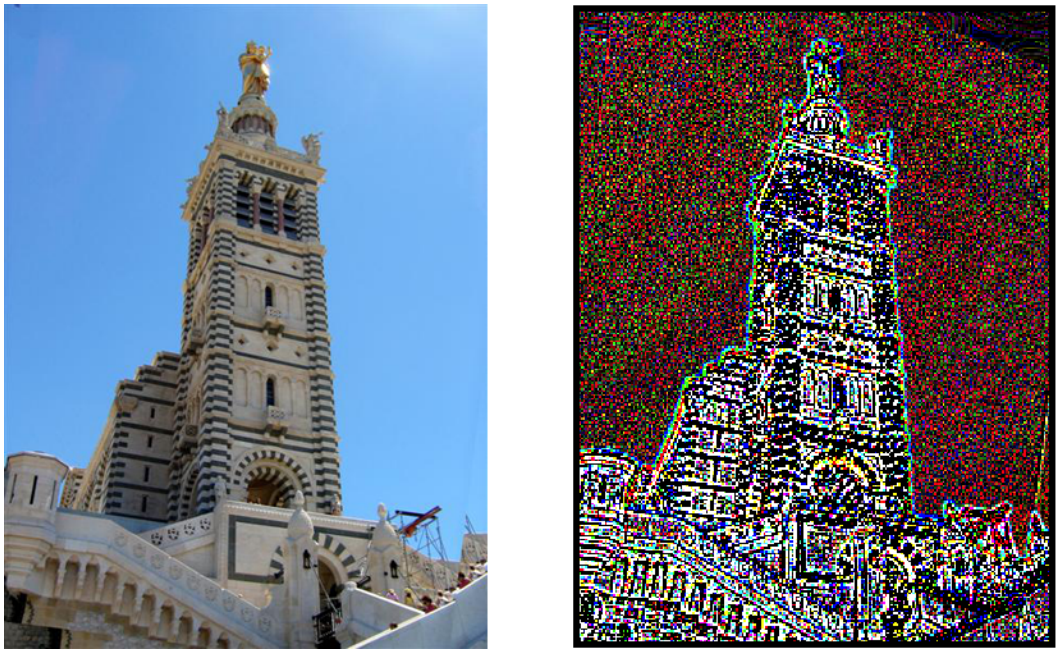

Figure 4.7: *(a) A natural image and (b) the PRNU extracted from this natural image. Low detailed areas contain less residues from the image itself, while in highly detailed areas the image content is clearly visible.*

CHAPTER 5

General concepts

5.1 Summary

To understand what happens before and after the charge has been captured by the sensor, we take a brief look at the Quantum Efficiency, Colour Filter Array, the gamma-correction and the colour gain correction.

Before the photon arrives at the detector it has to pass through the Colour Filter Array (CFA), a mosaic of pigments placed on top of the sensor to distinguish between different colours (usually Red, Green and Blue, the Bayer filter). This is essential because the silicon itself cannot differentiate between the different colours incident on the sensor. As each element (i.e. each colour) of the CFA only passes a certain range of wavelengths, the efficiency the underlying sensor converts light into electron-hole pairs depends on the absorption characteristics of the CFA. In other words, this Quantum Efficiency (QE), the efficiency with which the incoming light is converted into useful charge carriers that make up the signal, depends on the CFA as well as on the absorption of the silicon. Because each colour element is not equally sensitive to the incoming light, and the output of the camera should ideally replicate the input as the eye registers it, a colour gain factor g is needed to correct for this behaviour. Traditionally, images were viewed on nonlinear CRT devices (which means the output intensity which is viewed on the screen is a power function of the input voltage ($I = V^{\gamma}$)), and to compensate for this adjustment a gamma correction is applied.

As not all pixels have the same output when a constant input is applied, a pattern of these 'deviating' pixels is obtained. It is shown that images originating from the same sensor have highly similar patterns, and that images from different sensors can be distinguished based on this simple consideration.

Finally, the reliability and performance of the PRNU estimate for a limited amount of images is assessed, by varying some of the characteristics such as integration time, JPEG quality and ISO value. The result is that over a large range of these parameters

the correlation coefficients, as obtained by applying filters to extract the PRNU patterns from the input images, correctly distinguish between images obtained from the same sensor and images originating from different sensors.

5.2 Quantum efficiency and flat-fielding

5.2.1 Summary

Quantum efficiency is the efficiency with which the photons are absorbed by the silicon. For various reasons not all photons are absorbed, and each pixel has its own quantum efficiency. Ultimately, the intensity observed at a certain pixel depends on its quantum efficiency, and hence on which pixel this intensity is incident. Flat fielding is a technique to remove these pixel-to-pixel variations by averaging a large amount of flat fields. A flat field is an image in which the sensor is illuminated as uniformly as possible. By averaging a large amount of flat fields, temporal noise sources are removed. The reason we are interested in flat fielding is that from these (averaged) flat fields we can reliably estimate the PRNU pattern. These patterns can then be compared to the patterns obtained from questioned images.

5.2.2 A closer look

It was previously briefly mentioned that the quantum efficiency (QE) is the efficiency of the process from photon to electron-hole pair. Ideally, each photon is absorbed by the silicon crystal and produces a pair. Not all photons that are absorbed are useful, for example the photons that are absorbed outside the depletion region. As can be seen from the bandgap of silicon (see appendix B), a photon needs to have enough energy to promote an electron to the conduction band from the valence band. For silicon this bandgap is approximately 1.12 eV, which results in a wavelength of approximately $1100nm$. Larger wavelength photons do not have enough energy for this process, and the silicon essentially becomes transparent to these photons[1]. Also, longer wavelength (lower energy) photons are absorbed deeper in the substrate (specifically, the intensity at a certain depth $I(x)$ decays exponential with depth z, $I(x) = I_0 e^{\alpha z}$). Therefore, the thickness of the silicon is also a factor in the efficiency of the absorption of photons.

Another possibility for the electron that is created low in the substrate is the recombination with a hole. The depletion region itself is (in a p-type semiconductor) essentially free of holes, but in the lower part of the substrate these holes are less repulsed from

[1] However, coatings may allow an increase of the QE for wavelengths normally outside the bandpass (spectral range for which a detector is sensitive), which may also reduce reflection off the chip.

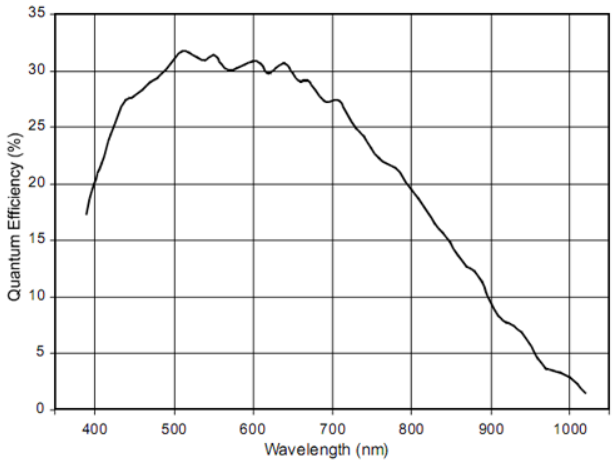

Figure 5.1: *Absorption characteristic of silicon. From [37].*

the gate. In these regions recombination is more likely to occur. Another possibility is electron diffusion; electrons can recombine with holes which decreases the efficiency of collecting electrons in the depletion region.

We saw in the previous chapters that not all pixels respond in the same way to a certain incoming signal, due to various noise sources. This non-uniform response of the sensor to a constant signal is exactly what makes flatfielding necessary in high demanding applications (for example in astronomy[2]). In these cases an obtained image is divided by the flatfield, so the 'true' image intensity is obtained. The reason *we* are interested in flat fielding however, is because we want to remove any temporal fluctuations of the output signal in order to find the relative response of each pixel. Put differently, flat fielding makes it possible to find the PRNU pattern that is present on the sensor. In principle, to find the PRNU pattern from the flatfield, it is necessary to have a uniform illumination of the sensor for all pixels. This is already a problem as light is dominated by statistical variations. When we make flat field images, we need to be sure that we do not enter the nonlinear or saturation region of the sensor, which may well occur before saturation sets in (See §4.7)[3].

In order to reduce the amount of (shot) noise in the image, we need to average multiple uniformly illuminated frames in order to increase the signal to noise ratio. Averaging multiple frames reduces the amount of shotnoise, so the final signal converges to a value in which the PRNU is clearly discernible. In principle one needs to have an

[2]Flat fielding in astronomy is used to remove pixel to pixel variations. This is important in order to establish the relative magnitude of a distant star. When certain areas of the CCD are more active, the brightness of the star will depend on the position on the sensor where the star is being observed. This is very undesirable when further calculations need to be based on this magnitude

[3]The value of saturated pixels can be estimated (if nonsaturated pixels are still available) from the response from the different colour channels. This is done in [36].

infinite amount of these frames to completely remove these noise variations. Luckily it is sufficient to average approximately 20 frames to sufficiently reduce these temporal variations (see §4.13).

When the image is not illuminated (shutter closed) we get a dark frame. In this dark frame defects may be detected, and may be removed/reduced in the same way as the illuminated flatfields. The defects referred to are not defects in the traditional sense of the word: a point defect for example is when the output of a pixel deviates more than a certain percentage compared to the output of adjacent pixels at some predetermined intensity level. Also, when multiple of these dark frames are averaged, the dark current shotnoise is reduced. The size of the dark current rises linearly with exposure time[4]. When the exposure time is approximately 0 seconds, we generally speak of a bias frame. The read noise (the noise independent of signal) is the root-mean-square value of the bias frame. It is possible for the bias frame to vary in time (Ch. 3.7 [9]), possibly due to variations in the properties of the electronics (transconductance, threshold voltage, see §[27]). To remove the non-uniform response from a frame I it is possible to divide by the averaged flatfield F, after subtracting the bias frame B: $\tilde{I} = \frac{I-B}{F}$.

5.3 Colour Filter Array, colour gain, gamma correction

CMOS and CCD sensors are monochrome devices; they cannot differentiate the colours that are incident on the pixel. Instead, a colour filter array (CFA, mostly in the configuration of a Bayer pattern) is placed on top of the sensor. The simple Bayer pattern consists of 3 colours, Red, Green and Blue in the $\frac{RG}{GB}$ configuration (more advanced patterns exists, e.g. the Cyan Magenta Yellow pattern that is more sensitive [38]). Due to the higher sensitivity of the eye to the green colour, the Green component is incorporated 2 times. As our eyes are more sensitive to green, having twice the amount of green elements reduces the noise we observe. The result is that the pixel below each element of this pattern only records 1 of the 3 components (to wit: only the light that goes through the respective colour, the rest should be absorbed). Hence, a large part of the incoming light is absorbed by the CFA. In this way, a mosaic is created of alternating colours. In order to give each pixel 3 colours, the colours are interpolated (demosaiced). For a short review of existing demosaicing algorithms see [39]. At each pixel, 2 of the 3 colours are interpolated from their neighbouring pixels. As these interpolated colours are obtained through a different mechanism, this interpolation scheme may introduce small offset gains in the pixelvalues [2]. As less green pixels have to be interpolated,

[4]The Canon EOS 10D uses some form of dark current reduction, possibly by subtracting a column average from each pixel.

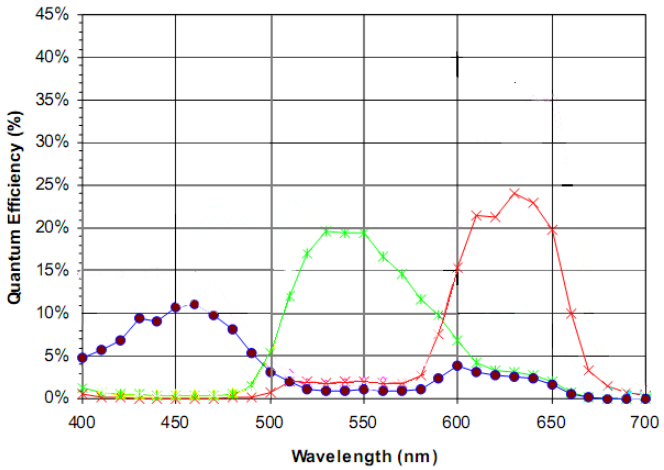

Figure 5.2: *Absorption characteristic of the Colour Filter Array. We see that absorption occurs for a very wide range of wavelengths. From [38].*

this noise in the green channel is lower compared to the red and blue channels.

Another step occurring after the charge has been read and digitized, is the multiplication with a constant (which is denoted g in the introduction of Ch. 4) to correct the white balance. The pixel below a certain colour does not only respond to that particular colour; it is however more sensitive to that colour, as can be seen in figure 5.2. The sensitivity of the silicon (QE) is dependent on the wavelength, dye absorption, and due to the fact that the output of the sensor should ideally match the way our eyes register the scene, the values obtained from the sensor should be corrected first. In order to calculate the green pixelvalue at a red pixel, one needs to know what its value would be if the red pixel was actually green. As the red and green bandpass are different, they do not respond in the same way to an incident field, hence the value needs to be compensated. The daylight/camera multipliers can be obtained from *dcraw*[5]. This is done before the interpolation. See also appendix D.

E To see the effects of the colour gain, the following method was used. A photograph is taken of a gradient image so that a wide range of intensities is present in the image. The image is then converted with and without applying the colour gain. By comparing the pixel values in the RAW file with the pixelvalues as obtained before and after the colour gain is applied, we obtain figure 5.3. We see that the colour gain is needed to compensate for the strong green cast of the sensor, i.e. to compensate the low sensitiv-

[5] *Use* $dcraw - i - v\langle filename\rangle$ *to extract these values from the metadata. Example output: Daylight multipliers: 2.186432 0.928015 1.023826. This means the red value is strongly adjusted, while the green and blue values are only adjusted slightly. When RAW files are converted without the* $-D$ *switch, this is done automatically. To see an image without applying the colour gain correction, use* $dcraw - r1111$

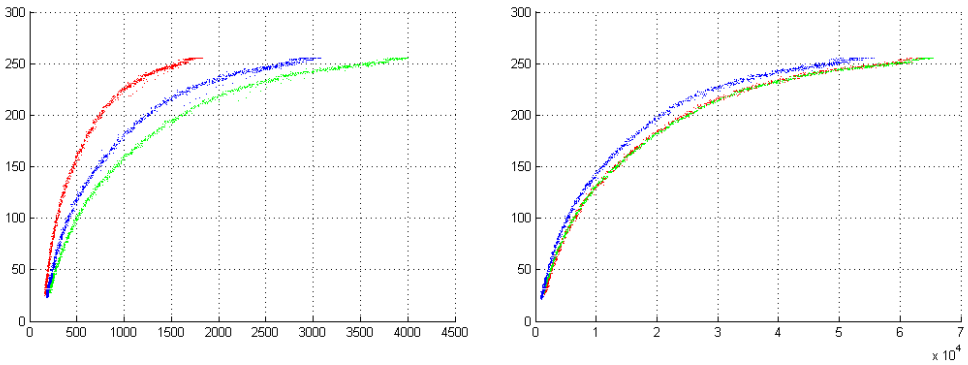

Figure 5.3: *Before (a) and after (b) the colour gain is applied. First, a photo is taken of a gradient so all input values are represented in the RAW image. The input value (x-axis) comes from the RAW file, the output value from a file which was converted without applying the multipliers (a), and with applying the multipliers (b). Note that not each input value has an 'unique' output value, probably due to the JPEG compression. In JPEG compression the output value of a certain input value ultimately depends on the value its neighbours have.*

ity of red and blue to correct the white balance. □

Traditionally, images obtained with (digital) cameras were viewed on CRTs. The intensity of the light at the screen of a CRT is however a nonlinear function of the input voltage. A gamma correction is done to compensate for this nonlinear function, in order to correctly reproduce the scene as was to be recorded on the screen. The gamma value of a CRT is approximately 2.5, which means that the intensity at the screen is a power function of the input voltage to the power 2.5. To have an approximate linear response at the perceptual level, the input is corrected with a power function with exponent 0.45 (roughly $1/2.5$). Actual values may vary depending on personal preference. For example, $0.45 \cdot 2.5 > 1$, which gives more contrast in lighter regions (Ch. 6 of [34]).

E The effects of applying a gamma correction can be easily observed by comparing the linearly converted image (i.e. no gamma-correction applied, $dcraw - d - 4 - T$) with the gamma corrected image, as in figure 5.4. In this figure the x-axis represents the input DN from the RAW file, and the y-axis represents the output DN from the converted image (with or without applying a gamma correction). To see the difference between the gamma corrections as applied by the Canon software and *dcraw*, both corrections were applied on a RAW image. It can be seen that *dcraw* uses a gamma correction (black) that is approximately the standard gamma curve ($1.099x^{0.45} - 0.099$ (Ch. 5 of [34]). The gamma correction as done by the Canon software has a different character-

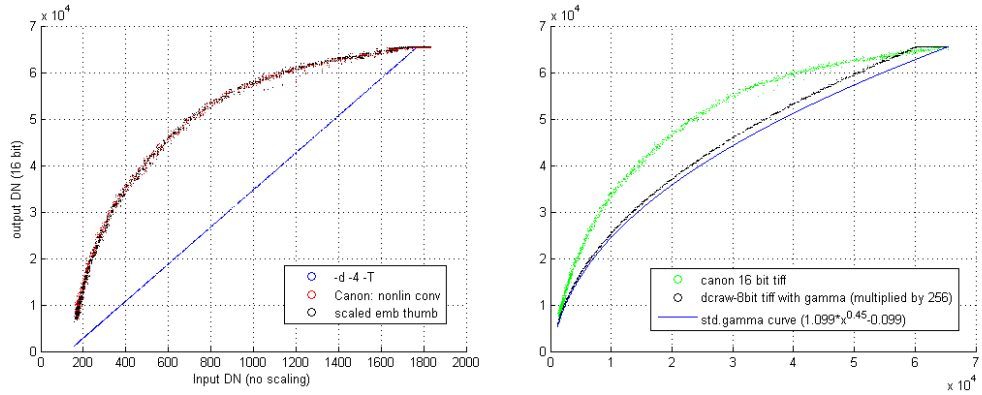

Figure 5.4: *Comparing the linear conversion (blue) with the gamma corrected images from Canon (a). In (b) one of the standard gamma curves (blue) is compared with the gamma correction as done by* dcraw *(black) and the software from Canon (green). It can be seen that the gamma correction as performed by* dcraw *is nearly the standard gamma correction. Input (DN) from RAW (x-axis) and output (DN) from converted files (y-axis) (a). Green and black curves from the (nonlinear) conversion from the canon software and embedded JPEG in the RAW, respectively. The blue line is the output when no gamma correction is done (dcraw -d -4 -T) . RAW image converted with* dcraw *(black) and and canon software (green) (b). The blue line is the standard gamma curve* $(1.099x^{0.45} - 0.099)$ *(Ch. 6 of [34]).*

istic (green). □

Because of the way the (linear) input signal is adjusted the contrast is reduced for pixels with high pixelvalues. That is, when the linear input signal (e.g. input pixelvalues 3000 and 3100) from the RAW file is gamma corrected (output pixelvalue is $\approx$ 3506 and $\approx$ 3565[6]), reducing the contrast in the image. This may mean it is advantageous to first apply the inverse gamma correction in order to better estimate the PRNU. On the other hand, the reverse happens at low pixelvalues, so it is not clear beforehand if it is advantageous to apply the inverse correction. It also needs to be kept in mind that although the gamma correction is in principle reversible, this may in practice not be the case as we only have a limited amount of discretization levels available (4095 in the case of the Canon EOS 10D, but this may be lower in other cameras). As we have seen in §4.14, the size of the PRNU (defined as the 1σ deviation from the mean pixelvalue of a small area) is approximately 1%, which means a standard deviation at a signal $S \approx 3000$ of 30 can be expected. After applying a gamma correction this signal is attenuated. No tests have been performed to check this, and it is unknown what the impact of this observation is.

[6]The output value is calculated by using the standard gamma curve: $\gamma(3000) = 4095(1.099 \cdot (\frac{3000}{4095})^{0.45} - 0.099)$ and $\gamma(3100) = 4095(1.099 \cdot (\frac{3100}{4095})^{0.45} - 0.099)$

5.4 Unicity of patterns

A very important consideration is whether we can assume that PRNU patterns to be unique. The sources for the PRNU, as mentioned before, suggest indeed that these patterns are unique. However, not all sources are indeed per-pixel; for example, in the case of the CMOS APS the bias transistor is shared between all pixel in a column. Also, we only looked at noise sources present on the image sensor although it is possible other components in the circuit add some systematic noise as well. This means that when patterns from multiple cameras of the same model are compared, the correlation between the extracted patterns may be somewhat higher than what can be expected from independent patterns (Ch. 3.4 of [5]).

An intuitive way (which is easy to explain in layman's terms) of showing whether two patterns have a common origin is the following. First, the mean pixelvalue μ of a small area of an averaged flatfield (e.g. 200x200) in a colour channel (preferably green) is calculated. Then the locations in this area where the pixelvalues deviate from this mean are marked; for example when the pixelvalue P deviates more than $\pm 3\sigma$ from the mean ($\mu + 3\sigma < P < \mu - 3\sigma$). The 3σ level is chosen such that other random deviations from the mean (most notably the photon shotnoise) do not influence the result. These 'deviating' pixels will in principle always be present. By averaging multiple flatfields, the random noise contributions are further reduced, so that the found deviating pixels are indeed characteristic to the sensor. In this way a pattern (of 200x200 pixels) is obtained in which the pixels that deviate from the mean are marked.

When two of these patterns (from two different cameras) are compared, it is expected that the amount of pixels that deviate in the same location (of the 200x200 matrix) is small. When multiple patterns originating from the same camera are compared, it is expected that the amount of pixels that deviate in the same location is large. When the location of a deviating pixel is the same in both patterns, these pixels are said to 'overlap'.

E To check if two Canon EOS 10D cameras could be separated based on this simple method, a number of flatfields were taken at different integration times for both cameras[7]. An area of 200x200 pixels was selected, and the number of pixels that deviated on the same locations were counted. On average, roughly 140 pixels on each sensor were found outside the aforementioned limits (see appendix F for the confusion matrices).

From the matrices in appendix F we see that when images have the same origin a

[7] *For the implementation of the above presented approach, see overlap.m and overlapmatrix.m.*

higher amount of overlapping pixels were found; on average about 40 (depending on the signal size higher and lower amounts were found). When too much light is incident on the sensor (at long integration times), saturation occurs resulting in a large amount of pixels reporting the same value (almost 400). When saturation occurs the standard deviation is reduced resulting in more pixels outside the 3σ limit.

When the patterns obtained from the images do not have a common origin, the amount of overlapping pixels is small (a maximum of 3). Earlier it was found that when the PRNU pattern from a flatfield is extracted, that the resulting distribution of values (i.e. the pixelvalues reported over all the pixels in the image) can be described by a normal distribution. Per definition, normal distributions result naturally in a certain percentage of values deviating a certain amount from the mean. For 3σ this is approximately 0.27%, which means that for 200x200 pixels approximately 108 pixels can be found with values outside the 3σ limit. Hence, comparing two of these random patterns will likely result in a small amount of pixels that overlap by chance alone. Indeed, when we do a simulation, in which random patterns are generated based on a normal distribution, and again denote the deviating pixels in this matrix we find that the amount of overlapping pixels can indeed be explained by random occurrences[8]. □

Based on the similarity of these patterns (e.g. the pattern obtained from a questioned image and the pattern obtained from a suspect's camera), it should be possible to assign an objective measure to the value of the evidence. If we can be certain that the obtained patterns can be correctly described by a normal distribution (also in the case when the pattern is extracted from a natural image), then this approach of comparing matrices in which the locations of the deviating pixels are denoted may be a useful start in assigning a value to the evidence (see Ch. 6).

5.5 Performance and reliability of the PRNU estimation

In general there are two different sources of variation: 'between source variation', and 'within source variation'. The 'between source variation' should be large, i.e. the correlation of the PRNU pattern obtained from the questioned natural image should have a high correlation with the camera with which the photograph was taken. On the other hand, cameras that did not produce the natural image should have a low correlation with the pattern obtained from the natural image. 'Within source variation' occurs when variations inside the camera itself are responsible for the varying of the corre-

[8] *This can be seen by running simulation2.m, in which two random matrices are generated and the amount of overlapping pixels in these two random matrices are counted.*

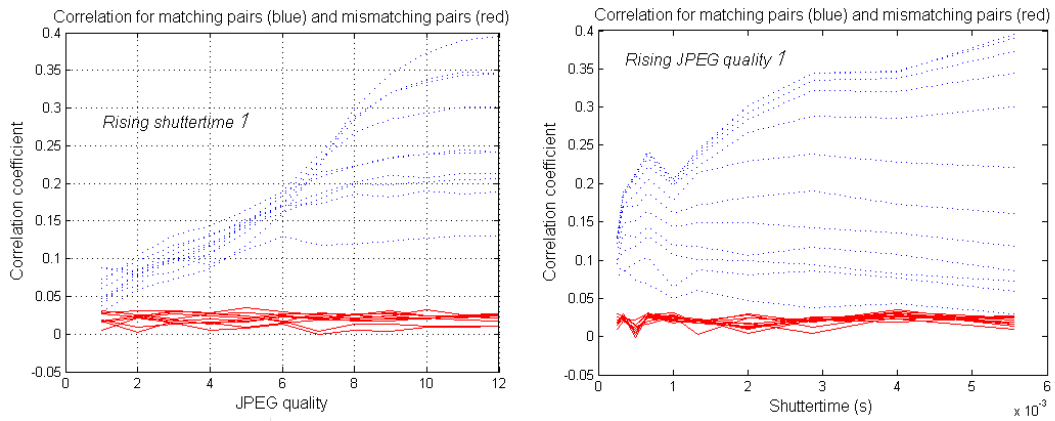

Figure 5.5: *(a) With longer shuttertimes the sensor is more illuminated, so the PRNU is more pronounced. Therefore, the correlation between an image and its origin (the sensor) is larger with longer integration times (unless it gets saturated). When the JPEG quality (1-12) with which the image is saved decreases, more detail is lost and the correlation decreases (blue lines). The red lines represent the correlation between mismatching pairs. (b) Longer shuttertimes and high JPEG quality result in the highest correlation between matching pairs.*

lation, for example due to noise or when residues from scene content disturbs the estimated PRNU pattern. This gives a small variation, and in practice this variation is not a problem unless the image is very dark, it is saved with a very low JPEG quality, or high ISO value as can be seen in figures 5.5 and 5.6[9]. With high compression, high ISO-values and with very long (or short) integration times, the correlation coefficients between patterns that have a common origin and patterns that do not have a common origin are less separated.

E To check how the correlation coefficients change when common parameters such as ISO value, shuttertime and JPEG quality are changed, PRNU patterns are extracted from images in which one parameter at a time is varied. For example, a series of natural images is obtained (with approximately the same scene content) with different ISO values (100, 200, 400, 800, 1600). Also, at each ISO value the shuttertime is varied, from (1/4000, 1/3000, 1/2000, 1/1500, 1/1000, 1/750, 1/650, 1/500, 1/350, 1/250, 1/180 s). Each obtained image is then saved with a varying JPEG quality in Adobe Photoshop (JPEG quality 1-12). From each image the PRNU pattern is extracted, and the correlation is calculated with the pattern extracted from a flatfield obtained from each camera. The resulting correlation values are displayed in figures 5.5 and 5.6. □

[9] *The graphs are obtained by running the compression.m and compression2.m scripts, in which PRNU patterns are extracted from images in various conditions (variable ISO, shuttertime, and JPEG quality).*

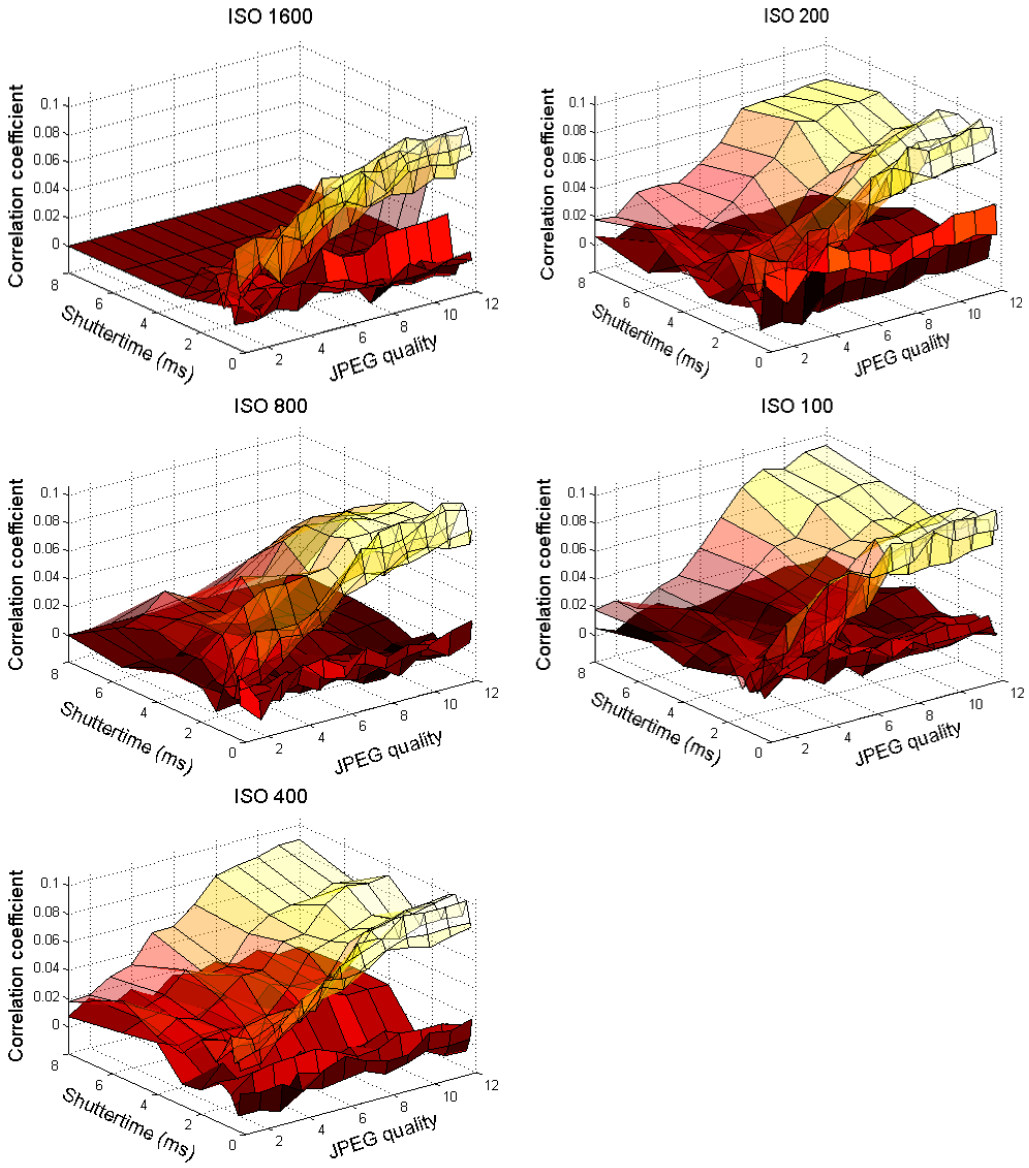

Figure 5.6: *3-Dimensional plot of the influences of JPEG quality, illumination time and varying ISO values on the correlation coefficients between matching and mismatching pairs (the matching pairs are the surfaces that are on top). Note that at ISO 1600 a large part of the surface is flat; this is because at this high ISO value the sensor was saturated quickly at all but the shortest shuttertimes. This figure is only meant as an illustration, as the amount of test data was very limited.*

5.5.1 Forgery

In [40] it is shown how the origin of an image can be forged. Basically, by making a large amount of flatfields it is possible to remove the non-uniform response, so correct identification is difficult or impossible. This can be done by subtracting a dark frame B from the image I, and dividing the result by the flatfield F:

$$\tilde{I} = \frac{I - B}{F} ,$$

as was shown in (§5.2). This makes it impossible to find the correct source camera. It is also possible to add a specific pattern from another camera, by a technique called 'inverse flatfielding'. By using the flatfield of an unrelated camera, F_{forge} can be used to 'hide' the alternative PRNU pattern in image $\tilde{I}$. By adding a forged bias frame a forged image $\tilde{J}$ is obtained:

$$\tilde{J} = \tilde{I}F_{forge} + B_{forge} .$$

As the authors note, this approach does not introduce visual artefacts. As this forgery cannot be readily detected, device identification will identify the camera that is responsible for F_{forge} and B_{forge} as the source camera. In this way it is possible that the incorrect camera is identified as the source camera.

CHAPTER 6

Likelihood Ratio

6.1 Summary

Assigning a likelihood ratio to the evidence is a way of objectively quantifying the value of the evidence, by dividing two probability density functions under two different hypotheses. Based on the correlation found between the PRNU pattern extracted from the questioned image and the PRNU pattern extracted from the flatfield images from a suspect's camera, we have an indication whether the questioned image originates from the suspect's camera or not. Although the results are convincing (the correlation between two patterns having the same origin is much higher than when the patterns do not have a common origin), it proved difficult to attach a likelihood ratio to these results. Namely, due the large separation of the two density functions (prosecutor's hypothesis and the defence hypothesis) there is no overlap which makes it impossible to divide the density functions. As the amount of test data is limited, and in general a very large amount of test data is needed to reliably estimate these density functions, this is not a practical approach. By using an estimator, specifically the Simple Good-Turing estimator, it is possible to find the percentage of yet unobserved data given the test data. Unfortunately, using an estimator results in other problems that cannot be glanced over. A few alternative approaches are mentioned briefly, which are useful starting points in improving the obtained results.

6.2 Introduction

In a forensic contest it is paramount to assign an objective measure to the weight of the evidence, in the form of a likelihood ratio. As the name suggests it is the ratio obtained by dividing the probability of the evidence E given the prosecutor's hypothesis H_p by the probability of the evidence E given the defence hypothesis H_d. This ratio is often

denoted with V, the value of the evidence:

$$V = \frac{P(E|H_p)}{P(E|H_d)} .$$

The mutual exclusive hypotheses can be defined as:

H_p The natural photograph obtained (e.g. from the suspect's computer) is taken with the sensor that is present in the camera of the suspect.

H_d The natural photograph obtained is made by some (random) other sensor.

According to the prosecutor's hypothesis, the pattern extracted from the questioned image has the same origin as the pattern extracted from the camera of the suspect. When this is indeed the case, the degree of similarity between these two patterns should be higher than when these patterns do not have a common origin.

Currently, the correlation between these two patterns is calculated in order to find this degree of similarity. As the correlation is a measure of how much one variable depends on the other, this is a logical choice. The correlation (specifically, Pearson's product moment correlation function) is calculated, and results in a coefficient between -1 (decreasing linear relationship) and +1 (increasing linear relationship). A correlation value of 0 means there is no linear relationship between the two patterns. However, when two patterns (or equivalently, two vectors) are highly similar a linear relationship (and hence a high correlation value) is expected. Specifically, when the patterns are represented in vector-form, the characteristic behaviour of vector 1 should take the same form as vector 2 (up to a multiple).

We do not know the origin of the questioned image, even after we find a high correlation between the pattern extracted from the questioned image and the pattern obtained from the camera of the suspect. If we would use these questioned images in estimating the probability density function, we would already implicitly assume the questioned images originate from the camera from the suspect. Therefore, we cannot use the questioned image(s) in estimating the probability density function for the prosecutor's hypothesis. However, if we do find a high correlation we can estimate the density function by making additional photographs that are *certain* to come from the suspect's camera. By making additional photographs with the suspect's camera that are as similar as possible as the questioned image, we can find additional data for the prosecutor's hypothesis. The photographs need to be as similar as possible because in principle each pixel has its own individual response, and a certain photograph stresses these responses in a unique way. For example, certain areas may be well-illuminated and contain a low amount of details (giving a clear PRNU pattern) while other areas may be dark and contain a high amount of details (hence the extracted PRNU pattern

is less reliable). If the patterns from these additional photographs are unrelated to the patterns obtained from the questioned image, the found correlations will be low. Otherwise, a (distribution of) high correlations will be found.

As we will see, the probability density function of the prosecutor's hypothesis and the defence hypothesis are clearly separated. That is, the correlation between images having the same origin is much higher than the correlation between images having unrelated origins. Although this may at first glance seem fortunate because we have a good indication of the originating sensor, it poses a problem when we would like to find a likelihood ratio based on these density functions. Due to the clear separation between these density functions it is impossible to divide them, simply because there is no data present at all possible correlation values. This means that the general approach (§6.3) does not work. If we would know the underlying distribution, or have a very large amount of data for both hypotheses we could estimate the density functions at arbitrary points.

In principle the most data is present for the density function of the defence hypothesis, simply because there are more cameras available which have a low correlation value with the pattern extracted from the questioned image (i.e. unrelated cameras). As the amount of cameras from which the questioned image originates can in practice be considered limitless (considering everyone with internet access can send or receive photos) we would like to populate the data for the defence hypothesis with a large amount of cameras. However, as the total population of cameras (with each its own pattern) is very large it is difficult to obtain a representative sample from this population.

Methods exist that allow to estimate the amount of 'unseen patterns' based on the patterns already available. If we know the amount of unseen elements, i.e. the amount of unseen patterns we have not yet observed because we have only a limited amount of cameras, we can artificially represent data outside the obtained measurements (this is done in §6.4). The usage of estimators with discrete data is well known, but as the distribution of values is obtained from calculating correlations that represent continuous values we do not yet know how reliable this method is.

Summarizing:

1. We want to compare the patterns extracted from the questioned (natural) image with the patterns extracted from the (averaged) flatfields from the camera from the suspect. We can calculate the correlation between the two patterns obtained. If the images have the same origin, the correlations are expected to be high.

2. If in step 1 we find a high correlation between the patterns from the camera and questioned image, we want to find the probability densities for both the defence and prosecutor's hypotheses. As the questioned images cannot be used for estimating the probability density function, we can extract additional patterns from photographs similar to the questioned image taken with the camera of the suspect in order to find the distributions.

3. As the probability density functions are separated, it is not possible to divide the density functions to obtain a likelihood ratio. An estimator can be used to find the percentage of unseen elements, which can then be used to artificially represent data outside the obtained measurements.

At the end of this chapter a few alternatives are presented which may be useful as alternative starting points in the problem of finding a likelihood ratio.

6.3 General approach

To understand the general methodology, we first consider the case in which the histograms for both hypotheses (partly) overlap. First, the pattern is extracted from the questioned image(s), and these patterns are then compared with the patterns extracted from the flatfield images produced by the camera of the suspect. If both patterns have a common origin, the correlation is expected to be high. Because we cannot use the questioned images for calculating the likelihood ratio, we produce additional images that have similar characteristics with this camera from the suspect. In this way we obtain a large amount of patterns, which can then be compared with all available reference patterns. If there are m available reference patterns (not including the reference pattern from the camera of the suspect), and we make n additional pictures with the camera from the suspect, we obtain mn correlation values that describe the defence hypothesis (H_d). It is expected the found correlations are low, as there is no common origin. The n patterns extracted from the additional pictures are then compared with the reference pattern obtained from the suspect's camera. This data populates the prosecutor's hypothesis (H_p), see figure 6.1.

To find the likelihood ratio we can estimate a probability density function for both the H_p and the H_d. In figure 6.1 we see there is a partial overlap of both hypotheses. The degree of overlap indicates the amount of false positives (FP, a pattern is found to have a distinctive (high) correlation when in reality there is no common origin) and the amount of false negatives (FN, a pattern is found to have a low correlation when in

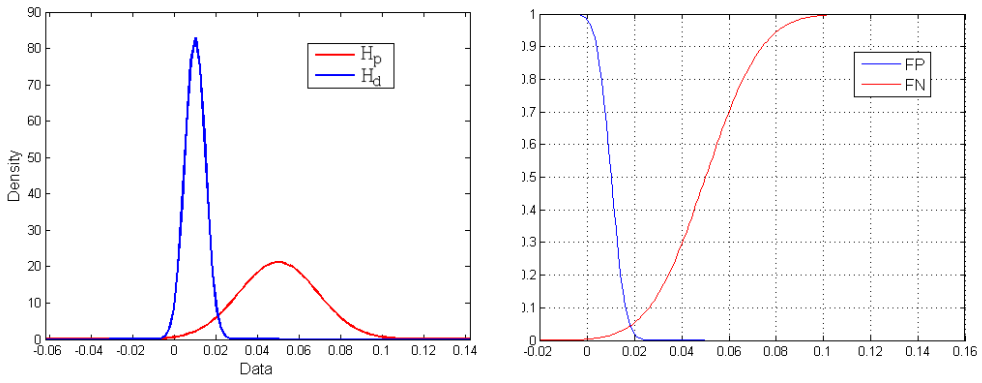

Figure 6.1: *Simulation of how the likelihood ratio can be calculated when both histograms partly overlap (a), and the fractions of false positives and false negatives (b).*

reality there is a common origin)[1]. If we are willing to accept a certain FP:FN ratio of e.g. 1:1000, we can find the location in the histogram where the likelihood ratio should be calculated by dividing the probability densities of both hypotheses at that location.

The assumption we made in the beginning of this section, that the histograms (partly) overlap, is not true in general. Both hypotheses are clearly separated. From a detection point of view this is advantageous, but when we want to assign a likelihood ratio to the value of the evidence this is a problem. This is not only a problem in this specific case: the same problem occurs when likelihood ratios are considered of e.g. glass evidence or fingerprints, and may appear in general with continuous data[41]. Because both hypotheses are clearly separated, it is difficult to divide the probability density functions as we do not have overlapping histograms. As we only have a limited amount of cameras at our disposal, we cannot reliably estimate the probability density function at arbitrary points. If we would know the underlying distributions of both hypotheses (e.g. a normal distribution), we could describe the density function at arbitrary points (when the mean and standard deviation would be known), but this is also not the case. Depending on the separation of the hypotheses, millions of samples would be needed[2].

If we would have many more reference patterns that can be compared with, we could find additional correlations outside the values that are already known. An alternative

[1] Also known as type I and type II errors, respectively.

[2] At least $1/(erf(d/(\sqrt{2}\sigma)))$ would be needed, with d the distance between the two distributions. Tests of normality, like the Lilliefors- or the Kolmogorov-Smirnov tests only indicate if the null hypothesis of normality can be rejected (at a certain confidence interval). In other words, there is no guarantee the data is truly normal, and if it is found to be normal there is no guarantee it will stay normally distributed when measurements are added. Therefore, a massive amount of measurements would be needed to be reasonably sure that the distribution is in fact normal.

Also, as the population of cameras may for practical reasons be infinite, especially considering anyone can send a photo over the internet, we cannot hope to find a suitable and representative sample of this population.

to having such a large amount of sample data is using an estimator. By using an estimator, as we will see in the next paragraph, it is possible to estimate the amount of 'unseen' data. This means that if we have a limited set of data (e.g. the correlations of (unrelated) reference patterns with the patterns extracted from the additional photographs obtained from the suspect's camera), it is possible to estimate the percentage of the total population that is yet unseen.

6.4 Estimators

A way to circumvent the abovementioned problem of insufficient measurements is to use or find an estimator which is reasonably invariant of the underlying distribution. There are a few approaches which try to estimate population frequencies of seen and unseen 'species', by looking at the obtained measurements in a (limited) sample. With species we indicate a set of measurements which can, for practical reasons, be attributed to one common source. For discrete data this can be the first letter of a word, the colour of a car, etc. For continuous data there are problems of concern which will be briefly mentioned later on.

The problem of estimating a probability distribution from a limited set of data is an old one [42]. There exist simple and complicated approaches to estimate how much of the data is still missing from the data we have already obtained. An example of a simple approach is the add-constant estimator, in which a part of the probability distribution is 'reserved' for unencountered elements. This is done by simply adding a constant to divisor and dividend. For example, when observing the colour of passing cars, it would be wrong after observing 1 black, 1 red and 1 blue car that the chance the fourth car is red is 1/3 as this does not account for all the other possible colours we have not encountered yet, such as green/yellow/etc. The add-constant estimator 'solves' this by adding a constant (e.g. 1) to divisor and dividend. This results in a chance that the next car is red of 2/7, that it is black of 2/7, and that it is blue of 2/7. This leaves 0+1/7 for new 'species'. However, there is no formal justification for adding a certain constant, for you already need to have a good idea of the amount of species, especially with a limited amount of measurements. If a constant is to be added the amount of species (colours) existing should already be known beforehand, as well as its distribution (i.e. its probability density function), which is exactly what we are trying to find out in the first place. Explicitly: adding 1 to a sample of 100 results in 'freeing' 1% of the probability density function for unseen species. This can only be accurate when the underlying distribution is known, or the amount of species.

A better, but a little more complicated approach is the Good-Turing estimator [43][44][45],

which is still conceptually simple and only certain weak assumptions have to be met[3]. It estimates the population frequencies for each already observed species and a probability a (yet) unknown species will occur, given the past observations. This 'unseen species probability' expresses how much measurements outside the known species will be found when we add new observations from a comparable distribution. Note that the Good-Turing estimators tell us nothing about the number of unknown species.

There exist various Good-Turing estimators, differing in the way the data is smoothed, as will be explained soon. The estimator used for our purposes is the so-called 'simple Good-Turing' estimator used in language statistics; the word 'simple' is used to describe the smoothing technique. The performance of the simple Good-Turing (SGT) algorithm may be a little less good than the full-fledged Good-Turing (GT) estimator, but it has the advantage of being easier to implement and understand [43].

A short description of the algorithm will now be given:

r is the sample frequency, the number of times a certain species occurs in the sample data (multiplicity)

n_r is the 'frequency of a frequency', the number of times species are found with frequency r (prevalence), such that $\sum rn_r = N$, the amount of observations.

r^* is the estimated sample frequency. If we had an unlimited amount of samples, or a sample which would perfectly resemble the true population, then $r^* = r$. In general $r^* < r$, as with the limited amount of measurements the full spectrum of species is not yet observed. Therefore, we need to 'free' some of the assigned probabilities for yet unseen species.

$p_r = \frac{r^*}{N}$ is the population frequency of species with sample frequency r[4].

For a certain species i, with chance p_i to observe i, we can do N measurements so that exactly r samples are in this category. This means that $(N - r)$ observations will not observe i with chance $(1 - p_i)$. In a series of N observations, the expectation of the frequency of the frequency r is the number of permutations we can make in a series of measurements, such that we find that there are r 'successes' (i.e. observations belonging to the category under investigation) and $(N - r)$ 'failures'. The GT estimator tries to find p_i for all i belonging to the collection of species s (and hence r^*). If the true population frequencies of all species in a given distribution would be known, the

[3]The distribution remains constant as samples are taken and individual sampling events are independent of the others. See [45].

[4]The expression 'population frequency' is often used, when 'population probability' is actually what is meant. We conform to the usage of the word frequency as is done by the aforementioned articles.

expected frequency of sample frequency r would be

$$E(n_r) = \sum_{i=1}^{s} \binom{N}{r} (p_i)^r (1-p_i)^{N-r} .$$

According to the Good-Turing approach:

$$r^* = (r+1)\frac{E(n_{r+1})}{E(n_r)} ,$$

and

$$P_0 = \frac{E(n_1)}{N} ,$$

the chance that the next observation belongs to a new species.

A problem appears when we try to assign a sample frequency to r^* when $E(n_{r+1})$ (i.e. the expected frequency of sample frequency $r+1$) is zero. E.g. when observing 100 car colours we may find 3 colours (purple, orange, yellow) that occur only once: $r = 1$, $n_1 = 3$. However, there may not be a colour that occurs twice: $r = 2$, $n_2 = 0$. This would result in an estimation for 1^* as 0, compared to 3% in a naive frequentist approach. This expresses the need for smoothing. The technique used for smoothing is the linear smoothing approach as presented by [43].

Z_r represents an adjusted value of n_r, $Z_r = 2\frac{n_r}{(r''-r')}$, in which r'' is the nearest highest sample frequency and r' is the nearest lowest sample frequency. Using Z_r instead of n_r results in a good straight line approximation in log-log plots of the data (i.e. r vs. Z_r)[5]. To smooth the data, a line of best fit is used in the $(\log r, \log Z_r)$ plot which is found by linear regression (so that the sum of the squares of the difference is minimized). This is the 'smoothed proxy' $S(r)$, which is used instead of n_r when certain conditions are met (see [43] for details).

This estimator allows us to estimate the population frequencies as if we had a larger amount of samples. Furthermore, the percentage of unseen species can be calculated (again, we do not know the number of unseen species).

6.5 Performance of the estimator

E As is done in [44], the performance of the GT estimator is evaluated by feeding elements of a known underlying distribution to the algorithm to see how good its predictions are (i.e. how close its calculated / estimated population frequency p_r is to the underlying distribution). The result is that the GT estimator performs well, as is evidenced by its low attenuation (put simply: the probability assigned to a set of observations by the

[5]The SGT algorithm exploits the fact that linguistic data shows a linear trend in log-log plots.

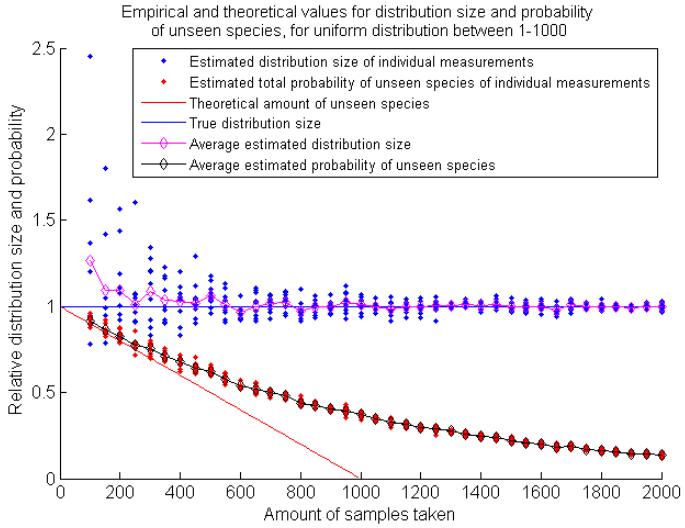

Figure 6.2: *Performance of the SGT estimator for the uniform distribution. The estimated distribution size can be calculated only for the uniform distribution by dividing the number of species occurring in the sample by* $(1 - n_1/N)$*, in which* n_1 *is the number of species occurring once, and* N *is the size of the distribution (1000). The red continuous line is the ideal line which the SGT should approach. We should be able to find out the size of the complete distribution, simply because we know how much data is observed, and what percentage that is. Red dots: individual estimates of* P_0*, blue dots: individual estimates of distribution size.*

GT estimator is in general not too far off the probability assigned to the set of observations by any distribution). This means that, regardless of the (unknown) underlying distribution the probability assigned to a set of measurements by the GT algorithm is reasonably close to the true distribution. For details see [44]. Better performing estimators are available, but these are more difficult to implement. The performance of the SGT estimator was assessed by [43] and found to perform better than the two way cross validation technique which is computationally more demanding. Furthermore, the SGT approach was more stable at low frequencies, and performed comparable at high frequencies. These tests were not repeated for our datasets. Instead, a simple test was devised for the uniform and normal distribution.

A sample was taken from a uniform distribution with integers 1-1000 (hence 1000 species). From this distribution a varying number of samples was drawn and the SGT algorithm was applied to predict the percentage of unknown species. When we draw 100 samples from this distribution we expect the SGT estimate of the amount of unseen species to be 0.9. See figure 6.2.

We see the performance is good for small sample sizes, but when more samples are taken the performance diverges from the theoretical line. This is not necessarily caused by the SGT algorithm, as the uniform distribution generation may actually be non-

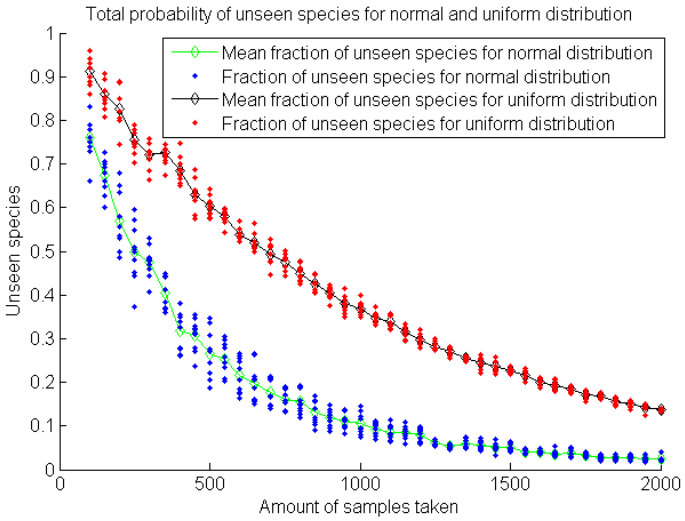

Figure 6.3: *Performance of the SGT estimator for the uniform distribution (red) and the normal distribution (blue). The amount of bins is in both instances equal (for uniform and normal distribution, see text).*

uniform. Furthermore, the way of calculating the fraction of unseen species may be too optimistic in large sample sizes (e.g. it is not uncommon for a species to occur once when we have 2000 samples from a 1000 species distribution). Also, a log-log plot of the uniform data (Z_r vs. r) does not show a good approximation to the data, one of the implicit assumptions of the SGT approach.

We can also do a simulation for normal data, just like with uniform data. However, it is much more difficult to give a theoretical curve to assess the performance of the SGT estimator, like we did for uniform data. In the uniform distribution we have a fixed amount of species, namely 1000. In the case of a normal distribution, there is an unknown number of species, and in principle each number (species) may occur. □

This problem was briefly mentioned before: the SGT estimator supposedly works well for discrete data, but we do not know how it performs in the case of continuous data. We have to find a way to discriminate between 'species' in continuous data. Like with car colours, certain colours may fall into two categories: there is no hard limit as to when a colour is green, or maybe yellow. In the type of data we are considering, we need to divide the measurements (correlations) into certain 'bins' with which we can simulate the species. The question is: how wide do those bins need to be? The amount of species decreases when we increase the binsize; hence more measurements occur for each species, which results in a smaller fraction of unseen species. If we choose large binsizes we expect a low P_0 as there are less species occurring once. Likewise, if we choose small binsizes, species are observed less and less often, which results in a higher estimate of unseen species. When the binsize increases above a certain value,

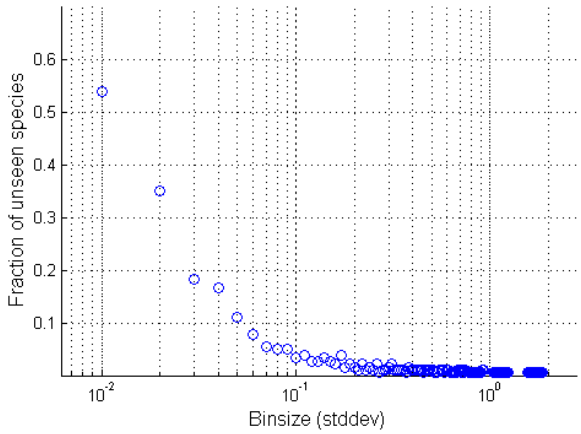

Figure 6.4: *To represent continuous data in a discrete way, the data is divided in bins. The amount of unseen species, as calculated by the SGT algorithm, depends on this binsize. The x−axis is expressed as fraction of the standard deviation of the data.*

the fraction P_0 does not change appreciably (figure 6.4).

E A small simulation test using data from 180 flatfields from mismatches, gave the following results[6]:

```
estimate of total probability of all unseen species: 0.011111
```

Freq (r)	F of F(n)	Smooth(Z)	log(r)	log(Z)	r*	p
1.0000	2.0000	1.3333	0.0000	0.1249	1.1385	0.0062
3.0000	1.0000	0.6667	0.4771	-0.1761	3.1659	0.0172
4.0000	1.0000	1.0000	0.6021	0.0000	4.1706	0.0227
5.0000	1.0000	0.3333	0.6990	-0.4771	5.1736	0.0281
10.0000	1.0000	0.2857	1.0000	-0.5441	10.1800	0.0553
12.0000	1.0000	0.2500	1.0792	-0.6021	12.1811	0.0662
18.0000	1.0000	0.1053	1.2553	-0.9777	18.1831	0.0989
31.0000	2.0000	0.0870	1.4914	-1.0607	31.1847	0.1695
64.0000	1.0000	-0.0645	1.8062	-1.1903	64.1860	0.3490

Now we have an indication of the percentage of unobserved species. □

However, as was put forth earlier, we do not know anything about the number of unobserved species. As a conservative estimate, we may divide the unseen probability equally over the whole range of unobserved data, between the correlations found in

[6] *The Simple Good-Turing algorithm, as explained in the previous section, is implemented in the SGT_103.m function.*

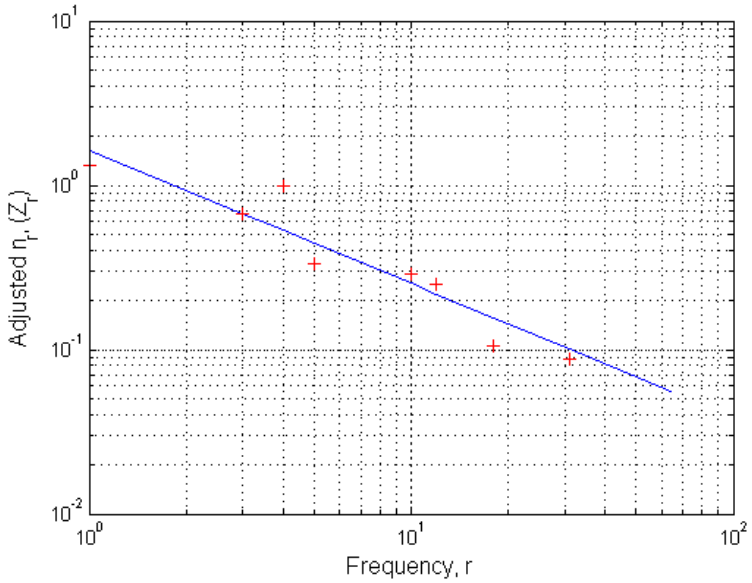

Figure 6.5: *Plot of the behaviour of sample frequency r vs. adjusted 'frequency of frequency' Z_r. The line of best fit, used for smoothing the data, is also plotted.*

the distribution representing the prosecutor's hypothesis and the distribution representing the defence hypothesis. In this way we at least know the density function of the defence hypothesis at the density function of the prosecutor's hypothesis. With the method described in the introduction, we can find the probability density function of the prosecutor's hypothesis (that is, obtaining additional images that are similar to the questioned image).

This leaves the problem that we do not know how reliable the estimate of the total probability of unseen species is, because this estimate varies with the binsize. One possible way is to relate the binsize to the variance in the density function of the prosecutor's hypothesis. We want to know when we can consider a camera to be of a different species, in order to establish the binsize. To see how the previously outlined method performs, see next section.

6.6 In practice

To assess the performance of the presented method in a practical case, the methods described above are used to find a likelihood ratio with the help of two Canon EOS 10D cameras[7].

[7]When these experiments were planned, the battery of the EOS 10D that was used throughout this research did not function properly anymore. At first use was made of two Canon EOS 400D cameras, but device identification performed on these cameras was not successful. Specifically, the correlation value obtained between the reference patterns of these two cameras was high (0.09), whereas the correlation

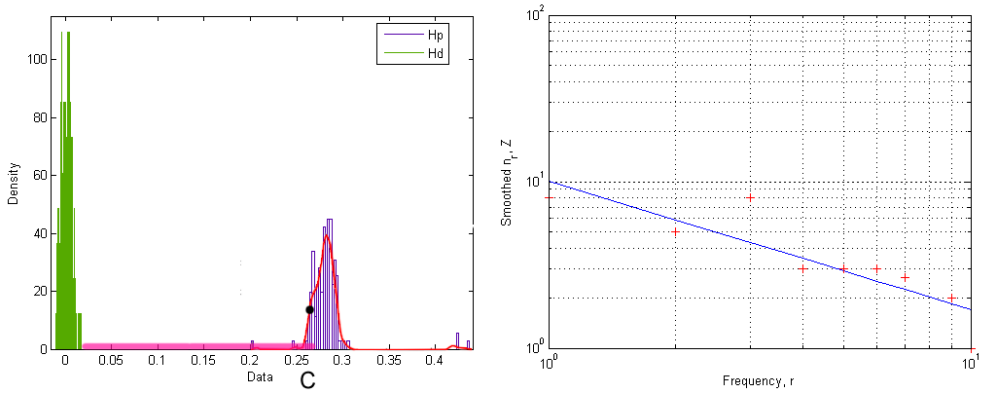

Figure 6.6: *(a) Probability density of correlation values obtained by calculating the correlation value between reference patterns and the patterns obtained from additional natural images (that are similar to the questioned image). The calculated density estimate (found with* ksdensity *in Matlab) is shown in red. The total probability of unseen species is spread out uniformly over all unseen species, as shown in purple. (b) Sample frequency r vs. Z_r. We see that the data shows a reasonable linear trend in this log-plot, as desired by the Simple Good-Turing algorithm. The binsize is 0.1σ.*

E First, a number of flatfield images are taken (40) and averaged in order to reduce temporal noise variations. Patterns are then extracted from these averaged frames, and denoted P_0 and P_1. A natural image which acts as the 'questioned image' is then acquired from one of the cameras, after which the PRNU pattern is extracted and compared with P_0 and P_1. This extracted pattern shows a higher correlation value C with one of the reference patterns, in this case $C = 0.264123$ with P_1. As the 'questioned image' cannot be used in establishing the probability density function, we obtain additional images (150) that have similar characteristics as the 'questioned image' from P_1 by photographing the same scene from approximately the same location[8], so that we are certain about the origin of the additional images. The correlation values of the patterns extracted from these additional images with both reference patterns are then calculated; in this way a distribution of values for H_d and H_p is obtained, as in figure 6.6.

value between the pattern extracted from a (well illuminated, low-detail) natural image and its reference pattern was much lower (0.003). In this way, it was not possible to perform a correct device identification, possibly because the non-uniformity in the sensor is either corrected, or the non-uniformity is very small so that extracting the pattern from the image removes the non-uniformities that are needed to perform a correct identification. The EOS 400D cameras were introduced recently (2006, compared to 2003 for the EOS 10D), and advances in device manufacturing may be a possible explanation. After this a replacement battery was found.

[8]In practice this is not feasible, and photographs could be obtained by photographing the questioned image itself, in order to obtain similar illumination on the sensor.

As we can see, both probability densities (H_p and H_d) are separated. Using the estimator as presented in §6.4 on the data that describes H_d, we obtain (binsize is $0.1\sigma_{H_d}$):

```
estimate of total probability of all unseen species: 0.052632
```

Freq (r)	F of F(n)	Smooth(Z)	log(r)	log(Z)	r*	p
1.0000	8.0000	8.0000	0	0.9031	1.1724	0.0070
2.0000	5.0000	5.0000	0.3010	0.6990	2.1950	0.0130
3.0000	8.0000	8.0000	0.4771	0.9031	3.2047	0.0190
4.0000	3.0000	3.0000	0.6021	0.4771	4.2101	0.0250
5.0000	3.0000	3.0000	0.6990	0.4771	5.2136	0.0309
6.0000	3.0000	3.0000	0.7782	0.4771	6.2160	0.0369
7.0000	4.0000	2.6667	0.8451	0.4260	7.2178	0.0428
9.0000	3.0000	2.0000	0.9542	0.3010	9.2202	0.0547
10.0000	1.0000	1.0000	1.0000	0	10.2211	0.0606

Due to a lack of data we did not yet see all 'species', as explained in §6.4. The Simple Good-Turing estimator estimates that this total probability is 0.052632. As a conservative estimate, we may divide this over the complete range between 0.052632 and the maximum correlation value in the probability density for H_d which is in this case 0.0175. Now we know that the probability density for the defence hypothesis at C is $\frac{0.052632}{(0.0175-0.264123)} = 0.21$.

We can estimate the probability density at C for the prosecutor's hypothesis, for example by using a kernel density estimate[9]. A kernel density estimate is used to estimate the probability density function after applying a weighting function (kernel) to the input data (in this case a normal/Gaussian kernel is used). In this way we can get an estimate of the probability density at all points. Calculating this density at the point C gives a probability density for H_p of approximately 14. The likelihood ratio is then approximately $14/0.21 \approx 66$. This means that the evidence given H_p is 66 times more likely than the evidence given H_d. □

We see that the obtained likelihood ratio is not really high. We do not know if 150 images is sufficient to give the correct distribution of H_p, so that adding additional images may improve the result somewhat. As the estimate of the percentage of unseen species of course also depends on the input data, it is unknown with how much cameras we should compare the extracted patterns, before we find a good representation of the population. However, it is unlikely that the likelihood ratio increases an order

[9] In Matlab this is done with *ksdensity*. To estimate the density at all points in the vector x, do $f = ksdensity(M, x)$, in which M is the vector for which the density is to be calculated.

of magnitude. In the next paragraph we will present an alternative approach which potentially allows much higher likelihood ratios. Another problem is the validity of the chosen binsize to discriminate between certain 'species', as the estimate of total unseen species depends on this binsize. A possibility for the chosen binsize could be the the spread in correlation values that occur when the correlation value is calculated between a reference pattern and a number of similar natural images.

6.7 Alternative approaches

6.7.1 As detection problem

In [46] a method is presented for calculating the likelihood ratio for the detection of a random signal 'hidden' in white Gaussian noise. It is treated as a detection problem: on basis of the calculated correlation coefficients we can find the camera from which the questioned picture originated from (hence we know which signal to detect from the flatfields). Two possible hypotheses can be given as:

$$H_p : \dot{x}(t) = z(t) + \dot{w}(t)$$

$$H_d : \dot{x}(t) = \dot{w}(t) \,,$$

in which $\dot{x}(t)$ represents the observed values (the values we get from extracting the PRNU from the natural image), $\dot{w}(t)$ represents the (zero-mean) white Gaussian noise, and $z(t)$ is the random process we are trying to detect. It is assumed that $\dot{w}(t)$ and $\dot{z}(t)$ are independent. In the prosecutor's hypothesis the signal $z(t)$ is hidden in the white noise, so the natural image has the same origin as the flatfields. Under the defence hypothesis the obtained signal is just white noise, and hence the natural image and the flatfields have different origins. Then:

$$LR = \frac{1}{B} exp\Big\{ \int_0^T \hat{z}_1(t)\dot{x}(t)dt - \frac{1}{2}\int_0^T \hat{z}_1^2(t)dt \Big\} \,,$$

where B is estimated from the filter that estimates the signal from the observations $x(t)$, and $\hat{z}_1$ is the estimate from the signal. The formula is intuitively reasonable: if the signal is hidden in the observations the first integral will be large. If the signal is not hidden in the observations, $\dot{x}(t) = \dot{w}(t)$, averaging to 0.

For smooth pictures (i.e. no sharp edges in the extracted pattern due to high frequency details) this process may work well, but when a residue of image content is found in the pattern, $\dot{w}(t)$ and $\dot{z}(t)$ are not independent anymore and the validity of

the method decreases[10]. Furthermore, the formula does not apply for discrete time processes: while $x(\cdot) - \hat{z}_1(\cdot)$ is still supposed to be a white uncorrelated process in time, it is not a Gaussian process due to the lack of continuity (p. 357 of [46]). Similarly in our case, when the known signal $\hat{z}_1()$ is subtracted from the observation $\dot{x}()$ we will see the noise is discrete due to the simple discrete nature of the AD/C. With this approach, multiple images are needed do detect the Gaussian signal for each pixel. With this in mind, and the author's limited understanding of the statistical processes playing a role, it was decided not to pursue this at the moment.

6.7.2 Per pixel probabilities

As was briefly mentioned before in §5.4, it is possible to use per pixel probabilities to attach an objective quantity to the degree of similarity of two patterns. In short, two patterns are obtained from two images (from the questioned image and from a flatfield image obtained from a suspect's camera) in which the locations of 'deviating' pixels (i.e. pixels with a value larger/smaller than $\pm 3\sigma$) are denoted. When the amount of pixels present on the same location in both patterns is high, this is an indication that the two patterns share a common origin. When pixels outside these limits are denoted with '1', and pixels inside these limits are denoted with '0', then the resulting vectors may look like[11]:

$$\left[\begin{array}{cccccccccccccccc} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0\ldots \end{array}\right] \text{ From the natural image}$$

$$\left[\begin{array}{cccccccccccccccc} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\ldots \end{array}\right] \text{ From the flatfields}$$

We see that in various places the deviating pixelvalues occur at the same location, implying a common origin. The patterns are not entirely similar; due to image content and statistical fluctuations in pixel response there may be a deviation in the pattern from the reference image where it does not occur in the pattern obtained from the flatfields, and vice versa. As said before, the value of a certain pixel is distributed normally, which means approximately 1 in $1/(1 - erf(3/\sqrt{2}) = 370$ pixels will deviate from the average by chance alone. If we write the following two mutually exclusive hypotheses,

[10] To counter the effects of high frequency image content, the image can be 'zero-meaned' (such that the sum of all rows in each column is zero, and the sum of all columns in each row is zero), and noise can subsequently be subtracted by using the Wiener filter in the frequency domain: $IFFT2(FFT2(Image) - W(FFT2(Image)))$, in which W is the Wiener filter, and $IFFT2$ and $FFT2$ are the inverse and 'normal' 2D Fourier transforms, respectively. This did not reduce the problem[2].

[11] Although it is represented as a vector, a two-dimensional matrix is possible too.

H_p The patterns are similar due to the fact that the natural image originates from the suspect's camera

H_d The patterns are similar due to chance,

it ought to be possible to calculate the likelihood ratio under these two hypotheses. Under H_d the distribution is binomial with parameters N (number of pixels that deviate in the natural image as well as in the reference pattern on the same location) and p (chance a pixel deviates by chance, $p = 1 - erf(3/\sqrt{2}) \approx 0.0027$). However, under H_p there is no such simple distribution. It is not entirely clear what should be done with pixels that do not deviate in the same locations. It is clear that image content should not play a role in the determination of the likelihood ratio, but it is difficult to exclude this information from the obtained patterns from the natural image. In smooth regions this may be less of a problem, but they may not be available. With this approach only a very limited amount of pixels is needed, as under H_d the chance two vectors show high similarity is exceedingly small.

CHAPTER 7

Conclusion

There are various sources responsible for the existence of the pattern as extracted from images. The extracted patterns include the effects of the Fixed Pattern Noise (FPN) and the Photo Response Non-Uniformity (PRNU).

Causes of the (illumination independent) FPN include microdefects due to voids or impurities, resulting in non-uniform stresses. Another source for non-uniform stresses is the variable thickness of the oxide layer or gates, resulting in non-uniform potential wells. Another source for a variable size of the potential well is simply the variation of the gate voltage, creating deeper or shallower potential wells. These non-uniform sizes of potential wells result in the capturing of a varying amount of thermal electrons. Especially in CCD image sensors the FPN can have a large impact, as the charge packets need to be transferred through a high number of gates during readout.

In CMOS Active Pixel Sensors additional sources for FPN are present. The variation in the threshold voltage in the reset transistor M_1 gives a variation in the precharged voltage of the photodiode, resulting in a systematic offset in the output. Likewise, the variation of the offset voltage of M_2 (source follower) results in a small variation of V_2 with respect to V_1. A variable resistance of M_3 results in a variation of V_3 with respect to V_4. The aforementioned sources occur per pixel. A column FPN is present due to variations in the characteristics of M_4, the bias transistor.

Causes of PRNU (illumination dependent) also include the depth of the potential well, and the size of the active region where photons can be absorbed. Because the absorption depth depends on the wavelength, this results in a variable absorption of photons, and is a systematic wavelength dependent PRNU source. During device fabrication, misalignment of the masks covering the substrate to define where the oxide and the gates should be deposited results in non-uniform pixel and transistor sizes. When the voltage applied on the gate area varies, the depth of the potential well varies as well. A variable transistor length (or gate area) results in a variation of resistance in the channel, decreasing the charge transfer capability.

We have also seen which noise sources are present and which noise sources dominate. At practically all but the lowest illumination levels the photon shot noise, which results from the discrete nature of light, is the dominant noise source. At low intensity levels the readnoise is dominant, depending on the sensor. At long exposures at low intensities the dark current may be dominant, but it is possible to reduce this noise source by subtracting the value of dark reference pixels. A simple model was developed for describing the three different noise regimes (readnoise, photon shot noise, and PRNU), and was found to describe the experimental values well.

On a practical level, it is recommended when finding the PRNU pattern from flatfields to take frames that are illuminated such that the output value is approximately $\frac{3}{4}$ of the maximum, to prevent entering the nonlinear region or saturation. Averaging 15-20 frames is sufficient to reduce the temporal noise variations, and a highly reliable extraction of the PRNU can be performed. To further reduce noise variations, images should be shot with the lowest possible ISO value. As some of the PRNU components are wavelength dependent, flatfied images should ideally be obtained with a constant spectral distribution. However, it has to be kept in mind that when JPEG compression is present, this is not necessarily the case. Indeed, as the value of a certain pixel ultimately depends on neighbouring pixelvalues, a varying distribution (of angle and scene content) is needed.

It has to be kept in mind that the PRNU can be removed from an image, making device identification impossible. Even more important, it is possible to add a specific pattern from a random camera to an image, making it possible that the incorrect camera is identified as the source camera.

The quest for finding a likelihood ratio is still on-going. Although using estimators may be a possible solution, the necessity to divide the continuous data into bins of a certain size is at the moment rather arbitrary. Finding the amount of unseen species in this way is rather straightforward, but it is necessary to be sure of the accuracy of this method before it is employed. As the likelihood ratios obtained with this method are limited, alternative methods may be preferred. It is expected that using per-pixel probabilities may result in high likelihood ratios, and its conceptual simplicity is a very attractive feature. Again, further research needs to be done to confirm this.

CHAPTER 8

Recommendations

As described in [2] it is also possible to detect altered images by comparing the PRNU pattern in the altered image with the patterns obtained from the source cameras, i.e. integrity verification, or forgery detection. Small tests were done, simply by calculating the correlation for corresponding blocks (e.g. 16x16 pixels) between a (unaltered) flatfield and the questioned image. In this way it was possible to find regions that were altered, for example by copy/pasting or by blurring, simply because this block has a lower correlation in the adjusted region, compared to unaltered blocks[1]. This worked well for flatfields, but was found to be lacking for natural images. A possible explanation is that in natural images, especially when high-frequency content is present, a lot of residue of the natural image is found in the resulting PRNU pattern after extraction, and more advanced methods are needed as presented in the former reference. An alternative approach used in [47] is to detect whether a certain part in the image was later added (by pasting from another image for example) by detecting the regions in the image which lack certain values in the quantization matrix used for encoding the JPEG file (when parts in the image are adjusted, it is likely that the second quantization matrix used for saving the JPEG the second time is different from the first quantization matrix used to save the file the first time). Detecting if certain regions in the image are adjusted may be very useful, as the correlation between the pattern extracted from the flatfields and the pattern extracted from the natural image may be high, even when a part of the natural image is forged. This may especially be the case when the forged region is only a small part of the natural image.

As CCDs are also present in image scanners, in which case the CCD has just a limited amount of rows, it may be possible to detect whether a document or image was scanned with a particular scanner or not. In [48] the identification of a sensor from printed images is studied. The PRNU is found to be still present in the printed and subsequently scanned image, though certain conditions have to be met. In [49] a method is presented

[1] *This is implemented in the forgery1.m script.*

to differentiate between images obtained directly from a camera or from a scanner by using the differences in pattern noise present in both devices.

The camera used in this research, the Canon EOS 10D, is a professional (Single Lens Reflex) camera that was introduced in 2003. Additionally, two Canon EOS 400D cameras were used for a short time, after the battery of the EOS 10D broke[2]. The EOS 400D camera was introduced less than two years ago. Surprisingly, it was found that the correlations between (the patterns extracted from) the flatfields of both cameras was higher (0.09) than the correlation between the patterns extracted from a (well illuminated, not detailed) natural image and its reference pattern. As tolerances in device manufacturing decrease, it is expected that the PRNU decreases as well. It is possible that the filters used to extract the PRNU pattern from the images produced by the 400D are not sensitive enough, and that necessary details are lost during extraction. Therefore, it is recommended to improve the filters to increase the sensitivity. It is unknown what techniques are used in the 400D to reduce the pattern noise, but as was shown in [25] it is certainly possible to reduce the PRNU. To continue using the PRNU as a means of camera identification, we have to understand which techniques are used to accomplish this, and the weaknesses of these techniques in order to continue to detect the pattern noise.

In [52] a method is presented to use dust present on the sensor of a Single Lens Reflex camera to perform camera identification. These dust particles leave traces in photographs. Although the dust can automatically be removed from built-in sensor cleaning systems (as in the 400D), it is found that a small percentage of dust is persistent. This small amount of dust can then successfully be used to find the source camera.

As the likelihood ratios that can be obtained with the usage of estimators is likely to be limited, other approaches may be more successful. It is expected that likelihood ratios as obtained from the method presented in §6.7.2 can be made almost arbitrarily high when a large amount of pixels is available. However, further research needs to be done to develop a reliable method.

Applying a gamma correction results in the compression of pixelvalues, as was explained in §5.3. This compression occurs at high pixelvalues, but at low pixelvalues the opposite is true. Whether it is advantageous to apply the reverse correction before extracting the PRNU pattern from images, has yet to be seen.

[2]Replacement batteries were not available, after which it was decided to use the author's own 400D cameras.

APPENDIX A

Nonlinearity

To understand the nonlinearity of the sensor, it is necessary to take a closer look to the energy band diagrams. In the case of a simple MOS gate, where a positive voltage is applied in order to establish the inversion layer (i.e. the conducting channel), the depletion layer forms beneath the oxide. Initially the potential well does not hold any electrons, and there is a large potential drop over the depletion region. During integration charges are made available and the electrons (in the case of a p-type substrate; minority carriers in general) are drawn towards the gate. In this way, the electrons high in the potential well 'screen' the potential lower in the well. Put differently, the electrons at a certain depth in the depletion region feel less attracted towards the gate when the well fills up.

The depth of the depletion region shrinks, which explains the nonlinear behaviour at high input signals. Indeed, after some amount of photons N_λ (for convenience at a constant wavelength) are captured at a certain depth z in the depletion region, there is no guarantee that doubling this amount of photons to $2N_\lambda$ will result in twice the amount of electrons in the inversion layer, simply because these new electrons may not feel the attraction of the gate at depth z anymore. As these electrons do not feel the attraction of the gate, they may diffuse through the substrate almost freely. Especially long wavelength photons which are absorbed deep in the substrate will feel this effect (i.e. the efficiency is reduced for long wavelength photons). In short, doubling the input signal does not result in a doubling of the output signal, as can be seen in figure 4.2, p.30. This is the nonlinear behaviour of the device. For a quantitative analysis see (Ch. 1.2 of [7]). Also note that a thicker oxide layer results in a higher potential drop across this insulator, and hence that increasing the thickness of the oxide results in a thinner depletion region.

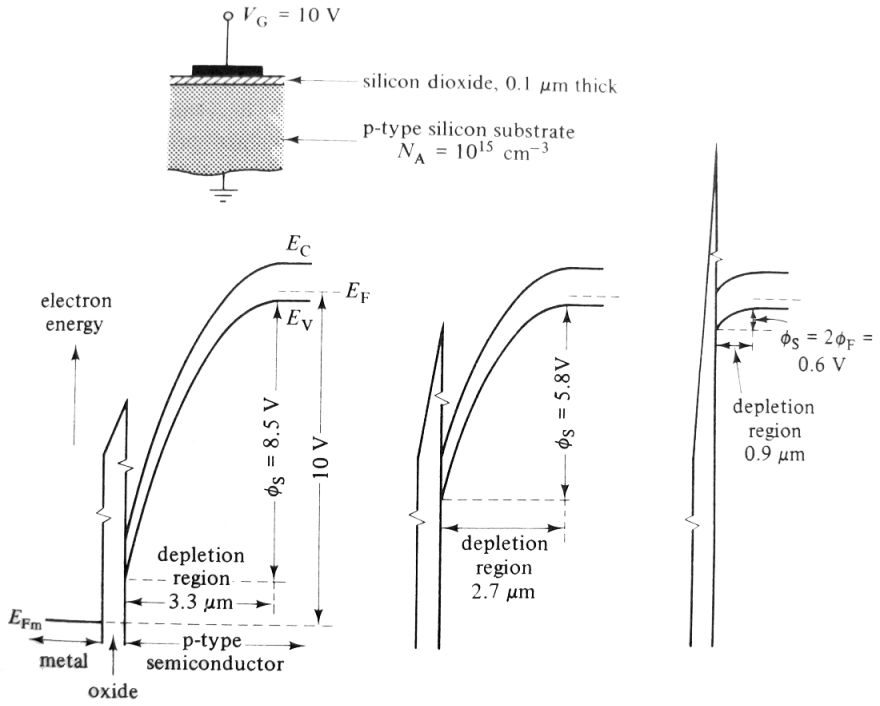

Figure A.1: *Electron potential energy diagram. (a) No charge in the potential well, (b) potential well fills up with charge, (c) full potential well. Note that a positive voltage corresponds to negative potential energy of the electron. Electrons at the right of each diagram (where the potential is constant) do not feel attracted to the gate as there is no potential difference.). From [7].*

APPENDIX B

Bandgap

The energy difference between the highest point in the valence band (valence band edge, E_v) and the lowest point in the conduction band (conduction band edge, E_c) is the band gap E_g. The probability for an electron to be excited from the top of the valence band to the lowest point in the conduction band is controlled by Boltzmann statistics:

$$\frac{e^{\frac{E_c}{kT}}}{e^{\frac{E_v}{kT}}} = e^{\frac{E_c - E_v}{kT}} = e^{\frac{E_g}{kT}}. \tag{B.0.1}$$

When the ratio E_g/kT is large the (intrinsic) conductivity is low (large bandgap). In fact, this is a simplification, and only true in the direct absorption process, in which only an electron and hole are created when a photon of frequency ω_g is absorbed. This only occurs when the conduction band minimum and the valence band maximum occur for the same energy (hence same wavevector) (Ch. 8 of [50]). In the indirect absorption process, such as in silicon, the conduction band minimum occurs for a different value than for the valence band maximum, but when a phonon (put simply, a wave of vibration in the lattice) is created in the process, the transition will still occur. Silicon is a relatively wide bandgap semiconductor, resulting in a relatively low dark current.

When impurities (this includes impurities which are added intentionally, such as adding Boron to silicon to make a p-type semiconductor) are present in the crystal, it is easier for electrons to migrate to the conduction band. These impurities add an intermediate energy level between the regular conduction- and valence band. In other words, electrons have an additional 'stepping stone' to reach the conduction band. This means that more electrons will be made available in regions where more impurities are available, especially at high temperatures where spurious electrons are present due to dark current. Because it is easier for electrons to reach the conduction band in regions where more impurities are present, these regions show up as FPN. This is also true for crystal defects.

Another effect is the scattering of electrons and phonons off these impurities or dis-

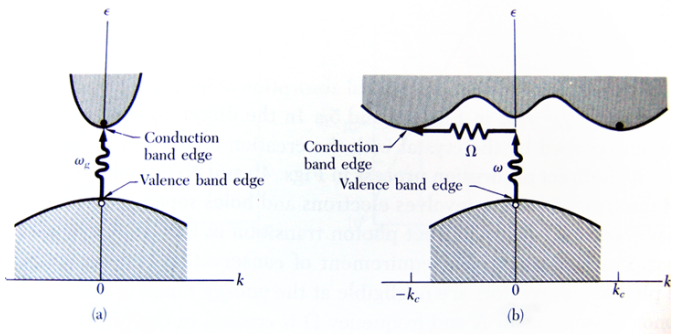

Figure B.1: *(a) Direct absorption process, in which a photon is absorbed by the material, resulting in an electron moving into the conduction band. (b) In the indirect absorption process, the conduction and valence bands are spatially separated. In these materials, such as silicon, it costs less energy for the electron to move into the valence band by emitting a phonon. The energy of the phonon is smaller than the extra energy needed for a direct transition as in (a). From [50].*

locations, and defects in general [22][1]. Like with the addition of impurities, an extra energy level is added between the valence- and conduction band. This intermediate energy level allows the capturing/emission of electrons and holes (and are called 'trapping centres'), so the probability that an electron will make a transition to the conduction band by phonon absorption increases (Ch. 10.10 of [18]). The dark current size is adjusted accordingly (i.e. $e^{-(E_g-E_t)/kT}$ instead of $e^{E_g/kT}$).

Some impurities are 'recombination-active', such that the capture/emission processes are happening relatively frequently. The concentrations of these impurities are usually extremely low, but they have a large impact[2]. These traps also affect the charge transfer efficiency, in the case of CCDs, by trapping a number of electrons and releasing them into trailing pixels. This is undesirable, especially when such traps are present in the shift register. Generally, the presence of defects influences the mobility of charges.

Also, the bandgap depends on impurity concentration [53]. But these effects are negligible.

[1]Fortunately, most impurities enter the crystal by substitution, not in interstitial positions (which would make matters much worse). This in turn creates local stresses in the material. Not all impurities are unwanted, e.g. the addition of oxygen in the production process adds mechanical strength to the silicon.

[2]In fact, the recombination via defects and impurities limits the efficiency in solar cells. See [51].

APPENDIX C

Photon statistics

In (Ch. 3.9, [54]) it is shown that the number of photons discerned in a detector adheres to a Poisson distribution. The analysis is fairly involved, so only a condensed outline is presented here. It is assumed that the number of electrons excited (hence detected) is proportional to the light intensity. The probability distribution $P_m(T)$ measures the probability m photons are detected in a time frame T. The probability that during a time interval $(t, t+dt)$ an electron is detected, is

$$p(t)dt = \xi \bar{I}(t)dt \,,$$

in which dt is long enough for the electron to be excited, but short enough to neglect the possibility of multiple excitations. $\bar{I}$ is the mean intensity that falls on the detector during the integration time,

$$\bar{I}(t,T) = \frac{1}{T}\int_t^{t+T} \bar{I}(t)dt \,.$$

The efficiency of the system is represented by ξ, for example the effects of quantum efficiency, or absorption by the colour filter array.

Let $P_m(t,T)$ be the probability m electrons are detected in a timeframe of $(t, t+T)$. The timeframe T is a fixed constant, and in principle a high amount of measurements is done to find the distribution of m. In the time interval $(t, t+T)$ let t' represent a value $0 < t' < T$, and dt' conforms to the same conditions as dt. The probability m electrons are detected in the time interval $(t, t+t'+dt')$ is then simply $P_m(t, t'+dt')$. Two possible ways of detecting m electrons are possible. Either

- M electrons are detected between $(t, t+t')$, and none in the remaining $(t+t', t+t'+dt')$. The chance that no electrons are observed in the latter timeframe equals 1 minus the chance that a (variable) number of electrons were detected in this time interval. The probability this situation occurs is

$$P_m(t,t')[1 - p(t+t')dt'] \,, \text{ or}$$

- $m-1$ electrons are detected between $(t, t+t')$ and one in $(t+t', t+t'+dt')$. Remembering that dt' equals the time needed for 1 absorption by hypothesis, this probability is $P_{m-1}(t,t')p(t+t')dt'$.
  The probability m detections occur between t and $(t+t'+dt')$ is the sum:

$$P_m(t, t'+dt') = P_m(t,t')[1-p(t+t')dt'] + P_{m-1}(t,t')p(t+t')dt' \ .$$

By differentiation, substitution of $p(t)dt = \xi\bar{I}(t)dt$, and integration it is possible to show that the number of detections is

$$P_m(T) = \frac{\langle m\rangle^m}{m!}e^{-\langle m\rangle} \ ,$$

with $\langle m\rangle = \xi\bar{I}T$. This result is the same as the Poisson distribution, for which the variance equals the mean.

APPENDIX D

Colour reproduction

When we make a photograph with a camera, the output should ideally replicate the response of the eye. Image sensors are, in contrast to our eyes, approximately linear devices. The sensitivity of the human eye (during photopic conditions, i.e. when enough light is available so that the lightcones register light) is numerically defined at different wavelength intervals[1]. Specifically, this is called the luminous efficiency, and luminance is the objective quantity of brightness, as brightness is a visual sensation (i.e. not objectively measurable)[2].

As can be seen from the figure, the eye is more sensitive to wavelengths around 550 nm (green). This means that if 3 light sources have the same radiant power (a linear light measure expressed in W/m^2, 'intensity') then the green colour will appear the brightest because the luminous efficiency is highest around 550nm. So when our eyes perceive a colour, the brightness we associate with it depends on the colour[3].

As the quantum efficiency of the sensor is not equal to the luminous efficiency of the eye, the need for correction becomes clear. The lightness as perceived by the eye is a nonlinear function of the luminance, viz. as $(luminance)^{1/3}$.

[1] See for example http://www.cvrl.org/database/data/lum/ssvl2e.txt

[2] CIE luminance is defined as the 'radiant power weighted by the spectral sensitivity function - the sensitivity to power at different wavelengths - that is characteristic of vision' (Ch. 5 of [34]).

[3] Specifically, according to standard ITU-R BT.709: $Y_{709} = 0.21R + 0.72G + 0.07B$, in which Y_{709} is the 'weighted sum' of the colours in the luminous efficiency figure (Ch. 5 of [34]).

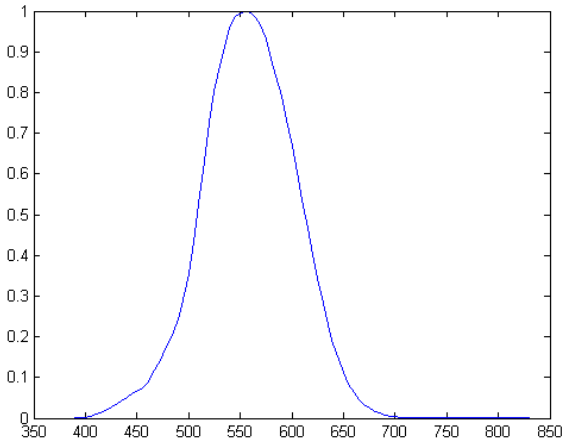

Figure D.1: *Luminous efficiency. An image sensor needs to have this spectral response in order to correctly reproduce the perceived lightness.*

APPENDIX E

Estimating the PRNU from frames

As was done in the introduction of chapter 4, we can write the image as $I = g^{\gamma}[(1 + K)Y + \Lambda]^{\gamma} + \Theta_q$. If we ignore the colour gain and the gamma value for the moment (which are not present when the RAW images are converted with the $-D$ switch), recognize that Θ_q is neglectable (§4.7), and the main noise source in well illuminated images is the photon shotnoise, we have $I = (1 + K)Y + \Lambda_{shot}$. If two images are made subsequently we can compare the difference frames and the individual frames. Furthermore,

$$\sigma^2(I) = \sigma^2((1 + K)Y) + \sigma^2(\Lambda),$$

and if

$$I_1 = (1 + K_1)Y_1 + \Lambda_{shot,1}$$

$$I_2 = (1 + K_1)Y_2 + \Lambda_{shot,2},$$

and $K_1 = K_2$, $Y_1 \approx Y_2$, which is correct if we average over a number of pixels (e.g. 128x128, depending on the definition of how the PRNU is to be measured). Then

$$I_1 - I_2 = \Lambda_{shot,1} - \Lambda_{shot,2}$$

$$\sigma^2(I_1 - I_2) = 2\sigma^2(\Lambda_1).$$

This means we can estimate the variance due the PRNU term as

$$\sigma^2((1 + K)Y) = \sigma^2(I) - \sigma^2(\Lambda)\ ,$$

with

$$\sigma^2(I) = \frac{1}{n}\sum_{i=1}^{n}\sigma^2(I_i)\ ,$$

and

$$\sigma^2(\Lambda) = \frac{2}{i^2 - i}\sum_{i=1}^{n}\sum_{j<i}\sigma^2(I_i - I_j)\ .$$

This means we can calculate the size of the PRNU by calculating the average variance of individual frames, and subtract the average variance of difference frames. For a *single* frame, we can calculate the PRNU as

$$100\Big(\frac{\sigma^2(I_1) - (\sqrt{S/G})^2}{S}\Big)\%$$

with S the average signal, and $\sqrt{S/G}$ the theoretical estimate of the photon shotnoise.

APPENDIX F

Confusion matrix

In §5.4 it was mentioned that images having the same origin have deviating pixelvalues on the same places. A large set of images were taken from two different cameras (but which are of the same model) to check this. This was done for different integration times, to check how often these deviating pixels occur at exactly the same site in these different circumstances. The matrices obtained are printed on the next pages.

As mentioned in §5.4, the amount of 'overlapping' pixels between two different cameras can be explained by random occurrences. Note that when the mean signal is high (3800 in the table) that the amount of 'overlapping' pixels for patterns of the same origin is increased to almost 400. This is due to the fact that when saturation occurs, a smaller range of pixelvalues is reported, hence decreasing the size of the standard deviation. This results in more pixels outside the 3σ limit. Note also that in table F.2 for very short shuttertimes the amount of overlapping pixels for patterns with the same origin decreases to the level of overlapping pixels when the patterns do not have the same origin.

		1/750	1/750	1/500	1/500	1/180	1/180	1/125	1/125	1/90	1/90	1/750	1/750	1/500	1/500	1/180	1/180	1/125	1/125	1/90	1/90
		1-650-1	1-650-2	1-900-1	1-900-2	1-2200-1	1-2200-2	1-3100-1	1-3100-2	1-3800-1	1-3800-2	2-650-1	2-650-2	2-900-1	2-900-2	2-2200-1	2-2200-2	2-3100-1	2-3100-2	2-3800-1	2-3800-2
1/750	1-650-1	144	17	32	34	34	39	33	30	8	6	0	1	0	1	0	0	0	0	2	1
1/750	1-650-2	17	136	23	21	29	28	26	24	8	8	1	0	0	2	0	0	0	0	0	0
1/500	1-900-1	32	23	128	33	42	41	38	41	8	8	0	0	0	0	1	0	0	0	0	0
1/500	1-900-2	34	21	33	131	44	51	44	42	8	8	0	0	0	0	0	0	0	0	1	1
1/180	1-2200-1	34	29	42	44	155	66	62	57	15	15	0	0	0	1	0	0	0	0	2	2
1/180	1-2200-2	39	28	41	51	66	143	61	55	10	10	1	0	0	0	0	1	0	0	0	0
1/125	1-3100-1	33	26	38	44	62	61	126	69	30	31	0	0	0	0	0	0	0	0	0	0
1/125	1-3100-2	30	24	41	42	57	55	69	132	42	41	2	1	0	0	0	0	0	0	2	2
1/90	1-3800-1	8	8	8	8	15	10	30	42	373	323	3	1	1	1	1	0	1	1	1	1
1/90	1-3800-2	6	8	8	8	15	10	31	41	323	388	3	1	1	1	1	0	1	1	1	1
1/750	2-650-1	0	1	0	0	0	1	0	2	3	3	133	32	36	37	36	32	29	28	20	18
1/750	2-650-2	1	0	0	0	0	0	0	1	1	1	32	138	37	46	44	42	38	34	17	16
1/500	2-900-1	0	0	0	0	0	0	0	0	1	1	36	37	111	45	52	56	45	43	14	14
1/500	2-900-2	1	2	0	0	1	0	0	0	1	1	37	46	45	139	64	60	53	56	26	28
1/180	2-2200-1	0	0	1	0	0	0	0	0	1	1	36	44	52	64	128	64	61	59	24	22
1/180	2-2200-2	0	0	0	0	0	1	0	0	0	0	32	42	56	60	64	120	65	58	23	24
1/125	2-3100-1	0	0	0	0	0	0	0	0	1	1	29	38	45	53	61	65	127	87	57	57
1/125	2-3100-2	0	0	0	0	0	0	0	0	1	1	28	34	43	56	59	58	87	127	65	64
1/90	2-3800-1	2	0	0	1	2	0	0	2	1	1	20	17	14	26	24	23	57	65	470	404
1/90	2-3800-2	1	0	0	1	2	0	0	2	1	1	18	16	14	28	22	24	57	64	404	464

Table F.1: Confusion matrix for the amount of 'overlapping' pixels (i.e. the amount of pixels that have a 3 sigma deviation from the mean at the same location) between matching and mismatching pairs. For each image, an area of 200x200 was selected. The top row indicates the shuttertime, the second row indicates the camera C, the mean value of the image M and the sample number N as $C - M - N$. The values of the matching cameras are distributed normally.

		1/3000	1/3000	1/2000	1/2000	1/1500	1/1500	1/1000	1/1000	1/750	1/750	1/3000	1/3000	1/2000	1/2000	1/1500	1/1500	1/1000	1/1000	1/750	1/750
		1-250-1	1-250-2	1-300-1	1-300-2	1-400-1	1-400-2	1-500-1	1-500-2	1-650-1	1-650-2	2-250-1	2-250-2	2-300-1	2-300-2	2-400-1	2-400-2	2-500-1	2-500-2	2-650-1	2-650-2
1/3000	1-250-1	130	1	3	1	4	2	5	3	4	5	0	0	0	0	0	0	0	0	0	0
1/3000	1-250-2	1	165	4	4	9	2	2	3	1	2	0	0	1	1	0	1	1	1	0	0
1/2000	1-300-1	3	4	135	9	12	11	7	12	6	11	0	0	1	0	2	0	0	0	0	0
1/2000	1-300-2	1	4	9	130	8	8	8	13	6	5	0	0	1	0	1	0	0	1	0	0
1/1500	1-400-1	4	9	12	8	151	10	8	11	10	13	0	1	1	0	0	0	1	3	1	0
1/1500	1-400-2	2	2	11	8	10	134	13	10	15	14	0	0	0	0	0	0	0	1	0	1
1/1000	1-500-1	5	2	7	8	8	13	123	10	13	17	0	0	1	1	0	0	1	0	1	0
1/1000	1-500-2	3	3	12	13	11	10	10	137	15	16	0	0	0	1	0	0	0	0	0	0
1/750	1-650-1	4	1	6	6	10	15	13	15	125	22	0	0	0	0	0	0	0	0	1	0
1/750	1-650-2	5	2	11	5	13	14	17	16	22	139	0	0	0	2	0	1	1	0	0	1
1/3000	2-250-1	0	0	0	0	0	0	0	0	0	0	122	3	11	6	9	14	14	13	10	16
1/3000	2-250-2	0	0	0	0	1	0	0	0	0	0	3	117	5	5	5	8	5	7	10	6
1/2000	2-300-1	0	1	1	1	1	0	1	0	0	0	11	5	124	8	10	10	12	10	13	8
1/2000	2-300-2	0	1	0	0	0	0	1	1	0	2	6	5	8	115	6	11	12	13	12	11
1/1500	2-400-1	0	0	2	1	0	0	0	0	0	0	9	5	10	6	105	17	16	21	17	16
1/1500	2-400-2	0	1	0	0	0	0	0	0	0	1	14	8	10	11	17	165	23	20	25	30
1/1000	2-500-1	0	1	0	0	1	0	1	0	0	1	14	5	12	12	16	23	137	28	21	27
1/1000	2-500-2	0	1	0	1	3	1	0	0	0	0	13	7	10	13	21	20	28	157	24	27
1/750	2-650-1	0	0	0	0	1	0	1	0	1	0	10	10	13	12	17	25	21	24	125	28
1/750	2-650-2	0	0	0	0	0	1	0	0	0	1	16	6	8	11	16	30	27	27	28	115

Table F.2: As table F.1, for different shuttertimes.

Bibliography

[1] M. Chen, J. Fridrich, M. Goljan and J. Lukáš, *Source Digital Camcorder Identification Using Sensor Photo-Response Non-Uniformity*, Proc. of SPIE Electronic Imaging, Photonics West, pp. 1G-1H, January 2007.

[2] M. Chen, J. Fridrich, M. Goljan and J. Lukáš, *Determining Image Origin and Integrity using Sensor Noise*, IEEE Transactions on Information Security and Forensics, vol. 3(1), pp. 74-90, March 2008.

[3] Canon Inc., *CIFF Specification on Image Data File*, Version 1.0 rev. 4, December 24, 1997.

[4] D. Coffin, *Open Source RAW Converter* (version 8.82) as obtained from http://cybercom.net/~dcoffin/dcraw/, accessed June 20, 2008.

[5] E. Alles, *Digital Source Camera Identification Using Photo Response Non-Uniformity*, report of internship at the Netherlands Forensic Institute, September 2007.

[6] G.C. Holst and T.S. Lomheim, *CMOS/CCD Sensors and Camera Systems*, JCD Publishing and SPIE Press, 2007.

[7] J.D.E. Beynon and D.R. Lamb, *Charge-Coupled Devices and their Applications*, London: McGraw Hill, 1980.

[8] E.R. Fossum, *Active Pixel Sensors: Are CCDs Dinosaurs?*, Proc. SPIE Vol. 1900, p. 2-14, Charge-Coupled Devices and Solid State Optical Sensors III, pp. 2-14, February 1993

[9] S.B. Howell, *Handbook of CCD Astronomy*, 2nd ed., New York: Cambridge University Press, 2006.

[10] A.E. Gamal, B. Fowler, H. Min, X. Liu, *Modeling and Estimation of FPN Components in CMOS Image Sensors*, Proc. SPIE Vol. 3301, Solid State Sensor Arrays: Development and Applications II, pp. 168-177, April 1998.

[11] N.H.E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design: A Systems Perspective*, 2nd ed., Addison-Wesley, 1993.

[12] N. Storey, *Electronics: A Systems Approach*, 2nd ed., Prentice Hall, 1998.

[13] K. Salama and A.E. Gamal, *Analysis of Active Pixel Sensor Readout Circuit*, IEEE Transactions on Circuits and Systems-I, Vol. 50, No. 7, pp. 941-945, July 2003.

[14] Centronic Inc., *Silicon Photodiode Theory*, as obtained from http://www.centronic.co.uk

[15] P. Amico, J.W. Beletic and J.E. Beletic, Eds., *Scientific Detectors for Astronomy: The Beginning of a New Era*, §2.1, Springer, 2004.

[16] J. Janesick, *Charge Coupled CMOS and Hybrid Detector Arrays*, Focal Plane Arrays for Space Telescopes. Edited by Grycewicz, Thomas J.; McCreight, Craig R. Proceedings of the SPIE, Volume 5167, pp. 1-18, January 2004.

[17] R. Sarpeshkar, T. Delbruck and C.A. Mead, *White Noise in MOS Transistors and Resistors*, Circuits and Devices Magazine, IEEE, Volume 9, Issue 6, pp. 23-29, November 1993.

[18] M. Grundmann, *The Physics of Semiconductors*, Springer, 2006.

[19] K. Tanikawa, Y. Ito and H. Sei, *Evaluation of Dark Current Non-Uniformity in a Charge-Coupled Device*, Applied Physics Letters, 1976, Vol. 28, pp. 285-287.

[20] K. Lundberg, *Noise Sources in Bulk CMOS*, Unpublished, can be obtained from http://web.mit.edu/people/klund/papers/UNP_noise.pdf, Oct. 2002.

[21] T. Goudon, V. Miljanovic and C. Schmeiser, *On the Shockley-Read-Hall Model: Generation-Recombination in Semiconductors*, SIAM Journal on Applied Mathematics, Vol. 67, No. 4, pp. 1183-1201, 2007.

[22] W.D. Callister, *Materials Science and Engineering: An Introduction*, John Wiley & Sons, Inc., 2003.

[23] J.O. Smith, *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications*, 2nd ed., http://ccrma.stanford.edu/~jos/mdft/, 2007, online book, accessed July 5, 2008.

[24] Eastman Kodak Inc., *KAF-0401E Full Frame CCD Image Sensor: Performance Specification*, June 26, 2000.

[25] A. Ferrero, J. Campos and A. Pons , *Correction of Photoresponse Nonuniformity for Matrix Detectors Based on Prior Compensation for Their Nonlinear Behavior*, Applied Optics, Vol. 45, No. 11, 10 April 2006.

[26] H. Tian, B. Fowler and A.E. Gamal, *Analysis of Temporal Noise in CMOS APS*, Proceedings of SPIE, Vol. 3649, pp. 177-185, 1999.

[27] Sony Inc., *Sony Semiconductor: Quality and Reliability Handbook*, Oct. 2000.

[28] N.K. Patel and A. Toriumi, *Stress Induced Leakage in Ultrathin SiO_2 films*, Applied Physics Letters, Vol. 64, pp. 1809-1811, 1994.

[29] M.P. Pagey, *Characterization and Modeling of Hot-Carrier Degradation in Sub-Micron NMOSFETS*, MSc. Thesis, Vanderbilt University, Nashville, Tennessee, 2002.

[30] M.S. Kulkarni, *Defect Dynamics in the Presence of Nitrogen and Oxygen in Growing Czochralsky Silicon Crystals*, Journal of Crystal Growth 310, pp. 324-335, 2008.

[31] J.A. Jiménez Tejada, A. Godoy, J.E. Carceller and J.A. López Villanueva, *Effects of Oxygen Related Defects on the Electrical and Thermal Behaviour of a $n^+ - p$ Junction*, Journal of Applied Physics, 2004, Vol. 95, pp. 561-570.

[32] D. Romero, J.M.F. Romero and J.L. Laserna, *Distribution of Metal Impurities in Silicon Wafers using Imaging-mode Multi-elemental Laser-induced Breakdown Spectrometry*, Journal of Analytical Atomic Spectrometry, 1999, Vol. 14, pp. 199-204.

[33] I. De Wolf, *Micro-Raman Spectroscopy to Study Local Mechanical Stress in Silicon Integrated Circuits*, Semicond. Sci. Technol., Vol. 11, pp. 139-154, 1996.

[34] C.A. Poynton, *A Technical Introduction to Digital Video*, John Wiley & Sons, Inc., 1996.

[35] H.T. Hytti, *Characterization of Digital Image Noise Properties Based on RAW Data*, Image Quality and System Performance III, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6059, pp. 86-97, January 2006.

[36] X. Zhang and D.H. Brainard, *Estimation of Saturated Pixel Values in Digital Color Imaging*, J. Opt. Soc. Am. A. Opt. Image Sci. Vis., Vol. 21, Issue 12, pp. 2301-2310, December 2004.

[37] Basler Vision Technologies, *Basler A600f User Manual*, March 2004.

[38] Kodak Inc., *The Technology Inside the New Kodak Professional DCS 620x Digital Camera*, 2000.

[39] A. Lukin and D. Kubasov, *An Improved Demosaicing Algorithm*, Graphicon Conference Proceedings, 2004.

[40] T. Gloe, M. Kirchner, A. Winkler and R. Böhme, *Can We Trust Digital Image Forensics?*, Proceedings of the 15th International Conference on Multimedia, pp. 79-86, 2007.

[41] C. Aitken and F. Taroni, *Statistics and the Evaluation of Evidence for Forensic Scientists*, 2nd ed., John Wiley & Sons Ltd., 2004.

[42] P. Laplace, *Philosophical Essays on Probabilities*, New York: Springer-Verlag, Translated by A. Dale, 1825.

[43] W.A. Gale and G. Sampson, *Good-Turing Frequency Estimation Without Tears*, Journal of Quantitative Linguistics, Vol. 2, pp. 217-237, 1995.

[44] A. Orlitsky, N.P. Santhanam and J. Zhang, *Always Good-Turing: Asymptotically Optimal Probability Estimation*, Science 302, pp. 427-431, 2003.

[45] I.J. Good, *The Population Frequencies of Species and the Estimation of Population Parameters*, Biometrika, Vol. 40, No. 3/4, pp. 237-264, December 1953.

[46] T. Kailath, *A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise*, IEEE Transactions on Information Theory, Vol. 15, pp. 350-361, May 1969.

[47] J. Lukáš and J. Fridrich, *Estimation of Primary Quantization Matrix in Double Compressed JPEG Images*, Proceedings of the Digital Forensics Research Workshop, Cleveland, OH, August 5-8, 2003.

[48] M. Goljan, J. Fridrich and J. Lukáš, *Camera Identification from Printed Images*, Proc. SPIE, Electronic Imaging, Forensics, Security, Steganography, and Watermarking of Multimedia Contents X, San Jose, CA, pp. 681-901, January 26-31, 2008.

[49] N. Khanna, A.K. Mikkilineni, G.T.C. Chiu, J.P. Alleback, E.J. Delp, *Forensic Classification of Imaging Sensor Types*, Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, pp. 65050U, January 2007.

[50] C. Kittel, *Introduction to Solid State Physics*, 7th ed., John Wiley & Sons Inc., 1996.

[51] J. Schmidt, P. Pohl, K. Both and R. Brendel, *Advances in Contactless Silicon Defect and Impurity Diagnostics Based on Lifetime Spectroscopy and Infrared Imaging*, Advances in Optoelectronics, 2007.

[52] A.E. Dirik, H.T. Sencar and N. Memon, *Digital Single Lens Reflex Camera Identification From Traces of Sensor Dust*, IEEE Transactions on Information Forensics and Security, 2008.

[53] G.D. Mahan, *Energy gap in Si and Ge: Impurity Dependence*, Journal of Applied Physics, Vol. 51, pp. 2634-2646, May 1980.

[54] R. Loudon, *The Quantum Theory of Light*, 3rd ed., Oxford University Press, 2001.