



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ**

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**ZABEZPEČENÍ SENZORŮ - OVĚŘENÍ PRAVOSTI  
OBRAZU**

SENSOR SECURITY - VERIFICATION OF IMAGE AUTHENTICITY

**DIPLOMOVÁ PRÁCE**

MASTER'S THESIS

**AUTOR PRÁCE**

AUTHOR

**Bc. IVO JURÁČEK**

**VEDOUcí PRÁCE**

SUPERVISOR

**prof. Dr. Ing. PAVEL ZEMČÍK**

BRNO 2020

## Zadání diplomové práce



Student: **Juráček Ivo, Bc.**  
Program: Informační technologie    Obor: Počítačová grafika a multimedia  
Název: **Zabezpečení senzorů - ověření pravosti obrazu**  
**Sensor Security - Verification of Image Authenticity**  
Kategorie: Zpracování obrazu

### Zadání:

1. Prostudujte literaturu a různá řešení zabezpečení senzorů a ověření pravosti dat jimi pořízených se zaměřením na obrazové senzory, přitom prostudujte zejména šumové vlastnosti senzorů obrazu.
2. Vyberte vhodný postup ověřování pravosti například pro rozpoznání záměny kamerového senzoru, nebo zamezení "podstrčení" falešných obrazových dat.
3. Navrhněte způsob implementace ověření pravosti dat a diskutujte vlastnosti řešení.
4. Implementujte metodu ověření pravosti dat a demonstруйте výsledky na vhodném příkladu.
5. Diskutujte dosažené výsledky a možnosti pokračování práce.

### Literatura:

- Dle pokynů vedoucího

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Zemčík Pavel, prof. Dr. Ing.**

Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.

Datum zadání: 1. listopadu 2019

Datum odevzdání: 31. července 2020

Datum schválení: 15. července 2020

## Abstrakt

Tato práce pojednává o zabezpečování obrazových sensorů. Cílem bylo nastudovat ověření pravosti dat, které takové senzory produkují. Na základě nastudovaných informací bylo navrženo řešení pro identifikaci digitálních kamer, a to pomocí šumových vlastností v obrazových senzorech, kterým jsou zatíženy. Byl zkoumán vliv různých algoritmů pro extrakci šumu z digitálního obrazu a bylo provedeno statistické vyhodnocení nad 15 obrazovými senzory, mající různé vlastnosti. Na základě vyhodnocených výsledků byly vyvozeny závěry.

## Abstract

Diploma thesis is about image sensor security. Goal of the thesis was study data integrity gained from the image sensors. Proposed method is about source camera identification from noise characteristics in image sensors. Research was about influence of denoise algorithms applied to digital images, which was acquired from 15 different image sensors. Finally the statistical evaluation had been done from computed results.

## Klíčová slova

Senzor, Kamera, Šum, Vzor, Anizotropní difuze, Non-local Means, Wienerův filtr, Korelační koeficient, Odstraňování šumu, SCI, Identifikace, FPN, PRNU, CMOS, CCD

## Keywords

Sensor, Camera, Noise, Pattern, Anisotropic Diffusion, Non-local Means, Wiener filter, Correlation coefficient, Denoising, SCI, Identification, FPN, PRNU, CMOS, CCD, Sensor Pattern Noise

## Citace

JURÁČEK, Ivo. *Zabezpečení sensorů - ověření pravosti obrazu*. Brno, 2020. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce prof. Dr. Ing. Pavel Zemčík

# Zabezpečení senzorů - ověření pravosti obrazu

## Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana prof. Dr. Ing. Pavla Zemčíka. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....  
Ivo Juráček  
4. srpna 2020

## Poděkování

Rád bych touto cestou poděkoval prof. Dr. Ing. Pavlovi Zemčíkovi za vedení této práce, za ochotu, tipy, rady, trpělivost a motivaci práci úspěšně dokončit. Rád bych též poděkoval i Martinovi Musilovi a Svetozáru Noskovi za poskytnutí kamery pro experimentování. Dále bych chtěl poděkovat Tomovi Víchovi, Petovi Flajšingrovi a Kubovi Budiskému za odbornou pomoc. Velké díky patří i přátelům, kteří mi pomohli s datovými sadami, a to Andy Stejskalové, Sáře Škutové, Pepovi Vanickému, Hynkovi Bernardovi, Honzovi Řežábkovi, Matěji Petřivalskému a svému bratrovi, Michalovi. Díky patří i knihovnicím, Anežce Mikéskové a Peti Obrusníkové. Obrovské díky patří všem mým přátelům a rodině, kteří mě motivovali k dokončení.

# Obsah

<b>1</b>	<b>Úvod</b>	<b>2</b>
<b>2</b>	<b>Obraz, digitální kamery a obrazové senzory</b>	<b>3</b>
2.1	Obraz . . . . .	3
2.2	Digitální kamery, optické soustavy a filtry . . . . .	4
2.3	Obrazový senzor a analogové zpracování signálu, demosaicing . . . . .	15
2.4	Šum produkovaný obrazovými senzory . . . . .	30
<b>3</b>	<b>Matematický aparát</b>	<b>34</b>
3.1	Odstranění šumu v obraze . . . . .	34
3.2	Odstranění artefaktů v obraze . . . . .	39
3.3	Korelační koeficienty, normální rozdělení a testování hypotéz . . . . .	41
<b>4</b>	<b>Existující řešení zabezpečení obrazových senzorů</b>	<b>44</b>
4.1	Identifikace zdrojové kamery pomocí vodoznaků . . . . .	44
4.2	Identifikace zdrojové kamery pomocí hlaviček obrazových souborů . . . . .	46
4.3	Identifikace zdrojové kamery pomocí vlastností v obraze . . . . .	46
<b>5</b>	<b>Analýza a návrh řešení</b>	<b>50</b>
5.1	Zhodnocení současného stavu . . . . .	50
5.2	Návrh řešení . . . . .	51
5.3	Specifikace řešení . . . . .	52
<b>6</b>	<b>Extrakce vzorového šumu a identifikace senzoru</b>	<b>54</b>
6.1	Schéma algoritmu a datové sady . . . . .	54
6.2	Extrakce referenčního šumu . . . . .	59
6.3	Určení korelačních koeficientů . . . . .	68
6.4	Implementace . . . . .	71
<b>7</b>	<b>Vyhodnocení algoritmu</b>	<b>74</b>
7.1	Výběr odšumovacího filtru . . . . .	74
7.2	Vliv sady snímků na kvalitu referenčního vzoru šumu . . . . .	80
7.3	Velikost referenčního vzoru šumu . . . . .	88
7.4	Lokalita extrakce referenčního šumu . . . . .	90
7.5	Praktické vyhodnocení hypotéz . . . . .	92
<b>8</b>	<b>Závěr</b>	<b>97</b>
	<b>Literatura</b>	<b>98</b>

# Kapitola 1

## Úvod

Každý digitální fotoaparát a kamera na světě je originální a dá se z několika desítek fotografií či krátkého videozáznamu odhalit, ze kterého zařízení vznikly. Tím, jak lze pro nějakou fotografii odhalit, ze kterého zařízení vznikla, se zabývá tato diplomová práce.

Zabezpečení senzorů představuje důležitý problém v oboru forenzních věd, který se, mimo jiné touto problematikou zabývá. Je důležité myslet i na zabezpečení samotných bezpečnostních zařízení, například digitálních kamer. Ty se totiž mohou stát silným nástrojem například v soudním řízení, kde se zachycená nahrávka páchané trestné činnosti může stát silným důkazním materiálem.

Otevírá se ale otázka, jestli prováděná trestná činnost není součástí *něčeho většího*. Hrozí tak například, že útočník může podvrhnout data z bezpečnostní kamery a nebo že ji dokonce může vzdáleně ovládat. Proto je dobré přemýšlet nad tím, jak takovou bezpečnostní kameru zabezpečit proti podvržení dat.

Každý digitální senzor je jedinečný a nabízí se otázka, jak je identifikovatelný. Díky výrobním vlastnostem jsou senzory zatíženy šumem, který se vždy při sbírání informací objevuje. Nejinak je tomu i u digitálních kamer.

V kapitole 2 je popsáno, co je to digitální obraz a jak se dá z kamer získat. V následující kapitole č. 3 jsou prezentovány matematické postupy, používající se v této práci. To, jaká jsou existující řešení, se dočtete v kapitole 4. Navržený postup práce a motivace jsou prezentovány v kapitole 5. Kapitola 6 obsahuje samotný postup práce pro ověření pravosti obrazů. Poslední kapitola, 7, obsahuje vyhodnocení provedených různých experimentů.

## Kapitola 2

# Obraz, digitální kamery a obrazové senzory

V této kapitole je uvedeno, co je to obraz, jak se dá získat z digitálních kamer a co to jsou obrazové senzory. Popsán je i šum, který na obrazových senzorech vzniká. Protože se z důvodu omezení maximálního rozsahu práce nemůže jednat o encyklopedický přehled, jsou zde uvedeny pouze informace bezprostředně související s touto prací.

### 2.1 Obraz

Obraz lze popsat obrazovou funkcí, což je dvourozměrná funkce, daná jako spojitý analogový signál, jak uvádí kniha [37]:

$$z = f(x, y). \quad (2.1)$$

Pro zpracování obrazu v počítači má obraz pevně stanovené rozměry. V přírodě je dána obrazová funkce barevným spektrem, v počítači je nutno ji omezit. Její definiční obor je definován rovnicí:

$$D = \langle x_{min}, x_{max} \rangle \times \langle y_{min}, y_{max} \rangle, \quad (2.2)$$

kde  $x \in \langle x_{min}, x_{max} \rangle$  a  $y \in \langle y_{min}, y_{max} \rangle$ ,  $x, y \in \mathfrak{R}$  jsou souřadnice bodu v kartézském souřadném systému. Obrazová funkce nabývá hodnot  $z \in H$ , kde  $H$  je obor hodnot.

Obor hodnot  $H$ , jak uvádí kniha [37], může být pro obrazovou funkci tvořen například jasem, intenzitou apod. Obrazovou funkci lze zapsat takto:

$$f : (\langle x_{min}, x_{max} \rangle \times \langle y_{min}, y_{max} \rangle) \rightarrow (H_1 \times H_2 \times \dots \times H_n), \quad (2.3)$$

kde  $H_n$  jsou jednotlivé obory hodnot.

Obor hodnot obrazové funkce je omezen též. Aby šlo vyjádřit více než jednu barvu, používají se v počítačové grafice barevné modely, tvořící uspořádané  $n$ -tice  $z = [z_1, z_2, z_3, \dots, z_n]$ . Takové  $n$ -tice tvoří obor hodnot. Příklad barevného modelu je například RGB, kde pixel tvoří trojice  $z = [R, G, B]$ . Obor hodnot je pro potřeby počítačové grafiky (konkrétně kvůli digitalizaci) omezen počtem barev, u šedotónového je počet odstínů 256 hodnot, u RGB až 16 milionů barev.

Díky tomu, že zobrazovací zařízení nezobrazují spojitě veličiny ale pouze diskrétní, je potřeba takový signál převést na diskrétní, jak uvádí kniha [37]. Tato operace se nazývá digitalizace, případně AD převod. Převedená datová struktura je diskrétní, obvykle se jedná o matici.

Digitalizace probíhá podle knihy [37] ve dvou krocích:

- **kvantování:** obor hodnot  $H$  původní spojité obrazové funkce  $z = f(x, y)$  je rozdělen na několik intervalů. Tyto intervaly mohou být stejně velké, ale též se mohou velikostí lišit. Každý takový interval je reprezentován jednou zástupnou hodnotou.

Zástupná hodnota je obvykle tvořena jako průměr daného intervalu. Dále je ale možné kvantovat pomocí mediánu, případně s použitím váženého průměru. Vlivem kvantování dochází k chybě, která se nazývá kvantizační. Díky kvantování dochází ke ztrátě informací.

- **vzorkování:** V předem daných intervalech spojité funkce  $f(x, y)$  dochází ke zaznamenávání hodnot, nazvaných vzorky. Diskrétní navzorkovaná obrazová funkce je poté dána jako

$$I_{i,j} = f(x_0 + i \Delta x, y_0 + j \Delta y), i = 0, 1, \dots, j = 0, 1, \dots, \quad (2.4)$$

kde  $\Delta x$  a  $\Delta y$  jsou vzdálenosti dvou různých vzorků. Původní spojitá funkce  $f(x, y)$  je definována na intervalu  $x \in \langle x_0, x_1 \rangle, y \in \langle y_0, y_1 \rangle$ .

Je důležité definovat *vzorkovací frekvenci*  $f_s$ , která má hodnotu

$$f_s(x) = \frac{1}{\Delta x}, f_s(y) = \frac{1}{\Delta y}, \quad (2.5)$$

kde  $\Delta x$  a  $\Delta y$  jsou vzdálenosti dvou vzorků. Vzorkovací frekvence má jednotku počet vzorků za jednotku času (Hz), případně počet vzorků na jednotku vzdálenosti (dpi). Platí úměra, že čím vyšší bude vzorkovací frekvence, tím více informací se zaznamená ze spojité funkce  $f(x, y)$ .

Důležitým předpokladem správného vzorkování je dodržení vzorkovacího teorému. Ten určuje, jak vysokou vzorkovací frekvenci je vhodné pro daný typ signálu použít. Kniha [20] uvádí, že při použití vzorkovací frekvence

$$F_s > 2F_{max}, \quad (2.6)$$

kde  $F_s$  je frekvence vzorkovací a  $F_{max}$  je frekvence mezní, již nedochází ke ztrátě informací.

Takto je popsána akvizice obrazu velmi obecně. Konkrétní model obrazu je zatížen dalšími vlastnostmi, jako například aditivním a multiplikativním šumem. V článku [35] je uvedena rovnice, která reprezentuje model obrazu, který vzniká na obrazových senzorech:

$$I(x, y) = g^\gamma [(1 + K(x, y) \cdot Y(x, y) + \Theta_n(x, y))^\gamma + \Theta_q(x, y)], \quad (2.7)$$

Proměnná  $g$  je zesílení barevného kanálu, způsobeného vyvažováním bílé barvy,  $\gamma$  představuje nelineární mapování tónů, aproximováno jako výpočet gama korekce,  $K$  je šum nerovnoměrnosti citlivosti pixelu,  $\Theta_q$  je kvantizační šum, způsobený JPEG kompresí a  $\Theta_n$  jsou další, aditivní šumy (obecně temporální šumy, viz 2.4).

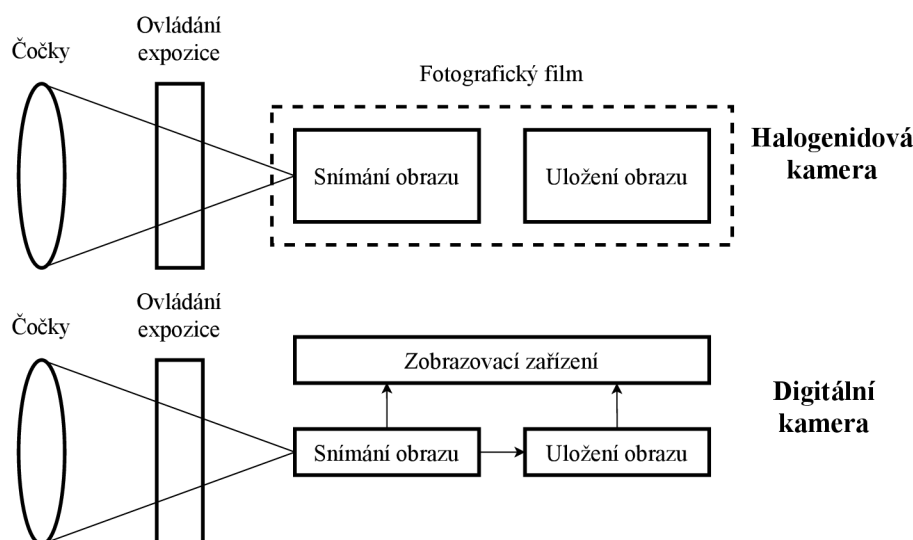
## 2.2 Digitální kamery, optické soustavy a filtry

Digitální kamera (případně digitální fotoaparát), jak uvádí kniha [32], je zařízení, které slouží k zaznamenání obrazu snímané scény do digitální podoby. Zaznamenání znamená převedení informací, obsažených v obraze, na odpovídající elektrické signály, které se mohou ukládat reprodukovatelným způsobem. V digitálních kamerách se obrazový senzor chová



jako převodník intenzity světelného záření na elektrické signály, které je možné poté pomocí dalších zařízení ukládat na paměťové médium. Díky tomu lze takto pořízený obraz zobrazovat na různých zobrazovacích zařízeních.

Tímto procesem se liší oproti světlocitlivým kinofilmovým kamerám, založených na chemických reakcích halogenidů stříbra. V těch jsou informace převedeny na chemické signály do fotografického filmu (kinofilmu), ve kterém přímo dochází k zaznamenání informace. Fotografický film se tedy chová jako zařízení pro záznam obrazu, ale současně i jako zařízení pro jeho uložení. Základní schéma těchto kamer můžete vidět je obrázku 2.1.



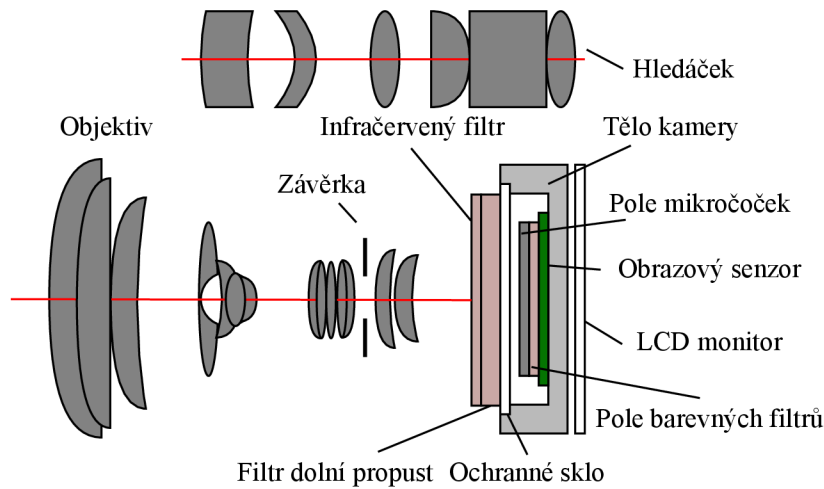
Obrázek 2.1: Základní schéma kinofilmové a digitální kamery

Podle knih [32] a [26] existuje na trhu několik druhů digitálních kamer. Původně se digitální kamery dělily na dva větší celky: *kompaktní kamery* a *zrcadlovky*. Postupem času, primárně z důvodů marketingových, vzniklo několik podkategorií digitálních kamer:

- **kompaktní kamery,**
- **zrcadlovky,**
- **hybridní kamery,**
- **pozadí,**
- **kamery v mobilních telefonech.**

**Kompaktní kamery,** též se v knize [32] vyskytují pod názvem *Point-And-Shoot* kamery, jsou jeden z nejběžnějších druhů kamer na trhu. Svou konstrukcí se podobají kinofilmovým kamerám a obsahují vestavěné LCD (Liquid Crystal Display) jako zobrazovací jednotku. Některé levnější modely kompaktních kamer nemají hledáček (součást fotoaparátů, která umožňuje řídit scénu). Jeho absence je suplována zobrazením výstupu z obrazového senzoru na LCD, tzv. živý náhled (live preview), který lze rozšířit o další informace (například zobrazení histogramu scény, nastavených parametrů apod.). Bohužel, při jasném okolním světle je vestavěný displej hůře čitelný. Tento nedostatek u dražších kompaktních kamer kompenzuje právě elektronický hledáček (Electronic viewfinder, EVF), což je malý

LCD displej, umístěný přímo v hledáčku, zobrazující snímanou scénu. Schéma běžného kompaktního fotoaparátu, jak je uvedeno v knize [32], je na obrázku 2.2.

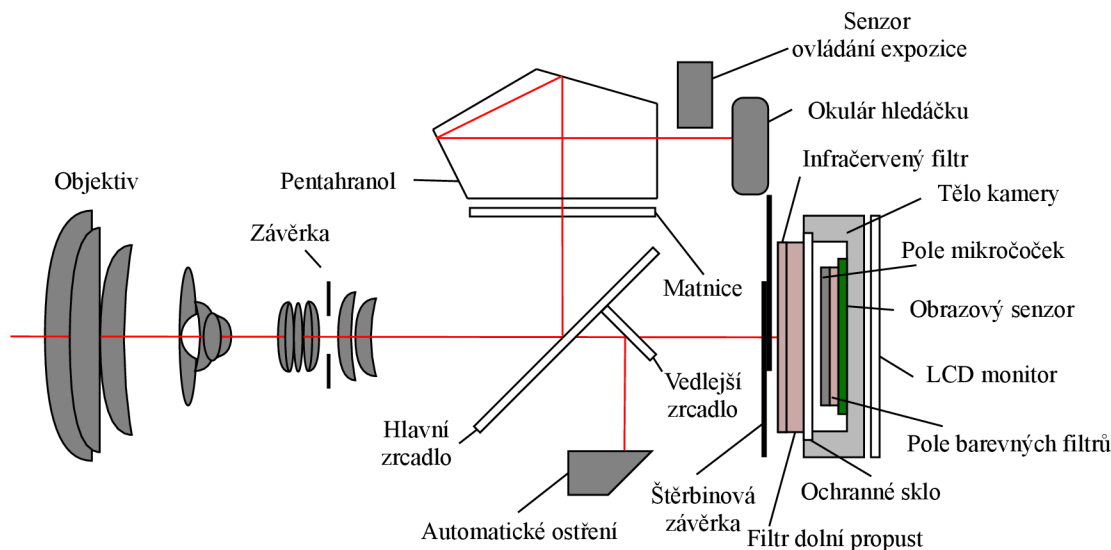


Obrázek 2.2: Základní schéma digitálních kompaktních kamer

Většina kompaktních fotoaparátů má mnoho funkcí, jako například automatické ostření, možnost osvětlení scény či automatickou expozici. Často obsahují i možnosti pro automatickou korekci scény, digitální přiblížení či možnost záznamu videa. Kompaktní kamery nemají výměnné objektivy. Nejčastější formáty pro uložení nasnímané scény jsou JPEG pro obrázky a MPEG4 pro videa.

**Zrcadlovky** jsou dalším druhem digitálních kamer, též nazývaných podle knihy [32] *Single Lens Reflex Camera*, SLR. Konstrukčně jsou digitální zrcadlovky podobné kinofilmovým zrcadlovkám. Hlavním rozpoznávacím znakem oproti kompaktním kamerám je, že mají vyměnitelný objektiv, nejčastěji s patičkou kompatibilní s 35-mm kinofilmovými SLR kamerami.

Digitální zrcadlovky obsahují několik zrcadel, které odráží snímanou scénu do optického hledáčku, a to přes *pentahranol*. Od klasických kinofilmových zrcadlovek se liší tím, že digitální zrcadlovky používají pro snímání optický senzor, který je umístěn v místě kinofilmu u klasických zrcadlovek. Dále digitální zrcadlovky obsahují LCD monitor, který ovšem nefunguje jako hledáček, ale jako obrazovka pro nastavování fotoaparátu či zobrazení již nasnímaných snímků. Schéma typické digitální zrcadlovky je znázorněno na obrázku 2.3.



Obrázek 2.3: Základní schéma digitálních zrcadlovek

Digitální zrcadlovky mají velikost obrazového senzoru jako násobek poměru stran jednoho okna 35-mm kinofilmu, též nazývaným anglickým pojmem full frame, což odpovídá velikosti  $36 \text{ mm} \times 24 \text{ mm}$ . Protože je technologický postup výroby full frame obrazového senzoru nákladný, používají se menší velikosti (viz tabulku 2.1). Snímky pořízené dvěma fotoaparáty ze stejného místa, kde jeden fotoaparát má full frame obrazový senzor a druhý má menší senzor, přičemž oba fotoaparáty mají nastavenou stejnou ohniskovou vzdálenost, se budou lišit výsledným zachycením jiného úhlu pohledu.

**Hybridní kamery** jsou podle knihy [26] zařízení, která jsou na pomezí digitálních kompaktních kamer a digitálních zrcadlovek. Vyznačují se obvykle vlastnostmi kompaktního fotoaparátu, ale mají například elektronický hledáček, vyměnitelný objektiv či tvar zrcadlovky. Postupem času se hybridní kamery rozdělily do dvou kategorií: *bridge* kamery a *bezzrcadlovky*.

1. **Bridge kamery** jsou digitální kamery, které se vyznačují zabudovaným velkým objektivem, který se vyskytuje u zrcadlovek, umožňující větší přiblížování či ostření obrazu. Disponují širokým množstvím automatických a manuálních režimů. Obvykle mají elektronický hledáček, případně používají živý náhled přímo na LCD monitoru. Obrazový senzor mají obvykle menší než zrcadlovky.
2. **Bezzrcadlovky** jsou digitální kamery, které se podle knihy [1] vyznačují tím, že mají vyměnitelný objektiv, obvykle kompatibilní se zrcadlovkami. Oproti zrcadlovkám ale mají bezzrcadlovky menší obrazový senzor. Název bezzrcadlovky (anglicky *mirrorless camera*) napovídá, že oproti zrcadlovkám nemají hlavní zrcadlo, které odráží obraz do optického hledáčku. Hledáček mají tato zařízení vždy elektronický. Podobně, jako *bridge* kamery, mají bezzrcadlovky mnoho funkcí, převzatých jak z kompaktních fotoaparátů, tak i ze zrcadlovek.

**Digitální pozadí** (anglicky *digital back*) jsou zařízení, která jsou podle knihy [32] složena z obrazového senzoru, řídicího okruhu, úložiště a obvykle i LCD monitoru. K digitálnímu pozadí se poté připojuje zrcadlovka bez kinofilmu, který je nahrazen právě obrazovým senzorem digitálního pozadí. Celý tento systém je obvykle připojen kabelem do počítače, který umožňuje kameru ovládat, případně umožňuje zobrazovat již uložené snímky.

V knize [26] je uvedeno, že digitálních pozadí existuje několik druhů, a to v závislosti na připojitelné zrcadlovce:

- specifické digitální pozadí, vytvořené na míru pro určitou zrcadlovku,
- univerzální digitální pozadí, umožňující připojení co největšího počtu různých zrcadlovek.

Někteří výrobci vyrábějí speciální zrcadlovky, umožňující připojení pouze na digitální pozadí. Tento druh digitálních kamer se obvykle používá v profesionálních fotografických studiích, protože cena těchto zařízení je oproti zrcadlovkám či kompaktním fotoaparátům vysoká. Výhodou je, že umožňuje fotografovi vyměňovat přímo obrazové senzory pro různé zrcadlovky. Příklad takové digitální kamery je na obrázku 2.4.



Obrázek 2.4: Příklad kamery s digitálním pozadím, převzato z knihy [26]

**Digitální kamery v mobilních zařízeních** jsou další uvedenou kategorií digitálních fotoaparátů v knize [32]. Mobilní zařízení s fotoaparátem postupně začaly nahrazovat kompaktní fotoaparáty, a to především mezi amatérskými fotografy. Rozmach a rozvoj mobilních telefonů s fotoaparátem, jak uvádí kniha [10], nastal zvyšující se oblibou sdílení pořízených fotografií svým přátelům či veřejně na sociálních sítích. Poptávka po těchto zařízeních byla stále vyšší, a tak výrobci postupně vestavěné digitální kamery zlepšovaly.

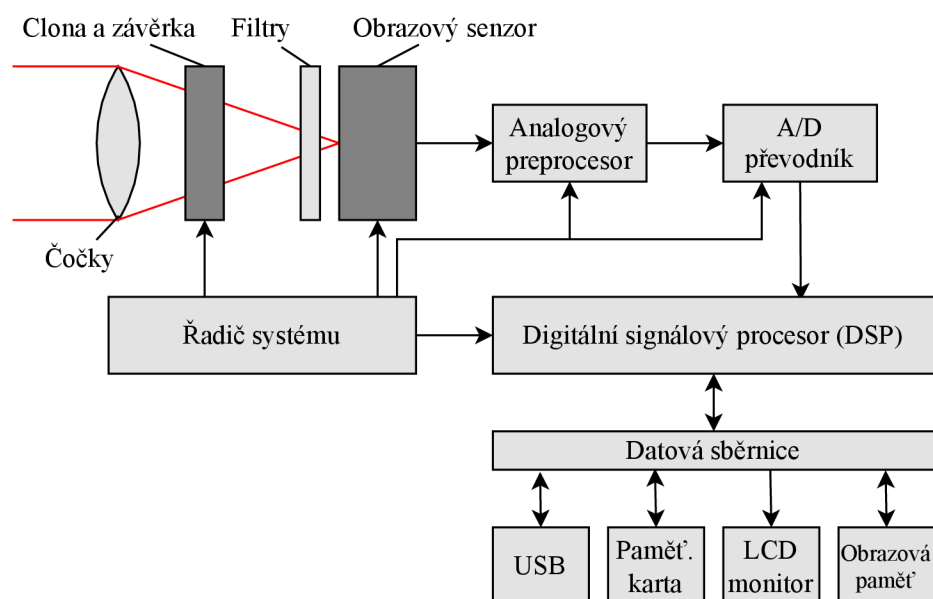
Průlom nastal podle knihy [10] v roce 2007, kdy Steve Jobs (zakladatel společnosti Apple Inc.), představil veřejnosti zařízení iPhone, což je mobilní telefon s digitální kamerou. Toto zařízení se oproti konkurenci lišilo tím, že ovládání celého mobilního přístroje bylo řízeno pomocí dotyků a gest na displeji s kapacitním snímáním. Tím získávala celá, v té době relativně mladá kategorie dotykových mobilů, na popularitě a při rozvoji technologie, telekomunikací a stoupající oblibě sociálních vznikla nová kategorie mobilních telefonů, nazývané chytré telefony (anglicky smartphones).

Současní výrobci přidávají do mobilních telefonů více než jednu kameru, v dnešní době jsou standardem v chytrých telefonech dvě kamery. Obvykle je jedna menší digitální kamera umístěna nad dotykovou obrazovkou pro pořizování tzv. *selfies*, tedy fotografií sebe sama. Na zadní straně zařízení se pak nachází jedna či více hlavních kamer. Například mobilní telefon Huawei P30 Pro (na obrázku 2.5), má dle specifikací, uvedených na webové stránce [19], přední kameru 32 MP se závěrkou  $f/2.0$  a soustavu zadních kamer *Leica Quad Camera*, složenou ze čtyř kamer.



Obrázek 2.5: Mobilní telefon Huawei P30 Pro, převzato z webové stránky [19]

Schéma obecné digitální kamery, jak je uvedeno v knize [32], je zobrazeno na obrázku 2.6. Typicky se digitální kamery skládají z optického a mechanického subsystému, obrazového senzoru a elektronického subsystému, který je složen z řadiče, analogového preprocesoru, A/D převodníku, digitálního signálového procesoru (DSP) a datovou sběrnici pro komunikaci s periferními zařízeními.



Obrázek 2.6: Blokové schéma běžné digitální kamery

**Optický subsystém** je podle knih [32] a [26] výchozím bodem pro pořízení snímku. Skládá se z různých kombinací optických čoček. Digitální kamery využívají podobnou soustavu optických čoček, jako kinofilmové kamery. Hlavní odlišnost je v tom, že menší digitální fotoaparáty obvykle mají kratší ohniskovou vzdálenost oproti klasickým kinofilmovým kamerám. Důvodem jsou právě menší rozměry některých kamer.

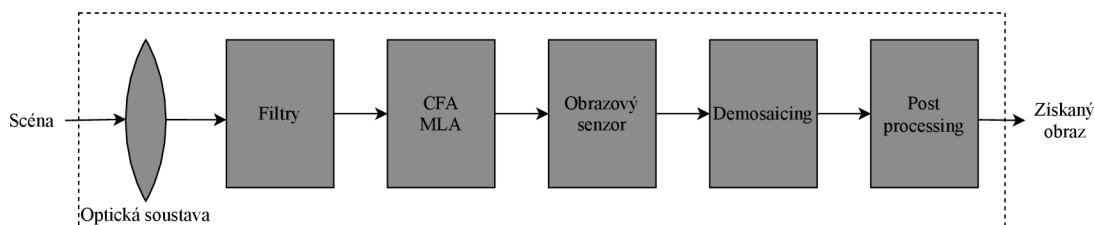
**Mechanický subsystém** digitálních fotoaparátů je velmi podobný mechanickým částem klasických kinofilmových fotoaparátů. Patří sem clona, která má za úkol redukovat množství dopadajícího světla procházejícího optickou soustavou a závěrka, která má za úkol se po určitou dobu (tzv. expoziční čas) otevřít a dovolit průchod světlu na obrazový senzor. V současné době jsou podle knihy [30] mechanické závěrky nahrazovány závěrkami elektronickými, případně elektromechanickými.

**Obrazový senzor** je důležitá součást digitálního fotoaparátu, která má za úkol zachytit dopadající světlo. Nejčastěji se používají dva druhy obrazových senzorů: CCD (Charge-coupled devices) a CMOS (complementary metal oxide semiconductor). Pro zaznamenání informace o barvách je těsně před senzorem umístěn barevný filtr (též *mosaic filter*) a pole mikročoček.

**Analogový obvod** slouží ke zpracování analogových signálů, které jsou výstupem z obrazového senzoru. Zde je důležitou jednotkou analogový preprocesor, který provádí základní operace nad analogovými signály, jako například separaci barev, ladění tónů či automatické ladění citlivosti. Výsledné signály jsou poté v A/D převodníku převedeny do digitální podoby.

**Digitální obvod** pracuje již s digitálními signály. Nejčastěji probíhá zpracování v digitálním signálovém procesoru (DSP). Provádí se zde zpracování obrazu, jako například vyvážení bílé barvy, komprese obrazu či převody mezi barevnými modely. Dalším úkolem digitálního obvodu je i komunikace s periferními zařízeními, například pro uložení výsledné fotografie či zobrazení informací na LCD monitor.

**Systémový řadič** má za úkol ovládat jednotlivé sekvence operací digitálních kamer. Často řídí například automatické ostření nebo automatické ovládání expozice. Dále má za úkol například hlídat stav baterie či reagovat na požadavky uživatele, jako je zmáčknutí tlačítka.



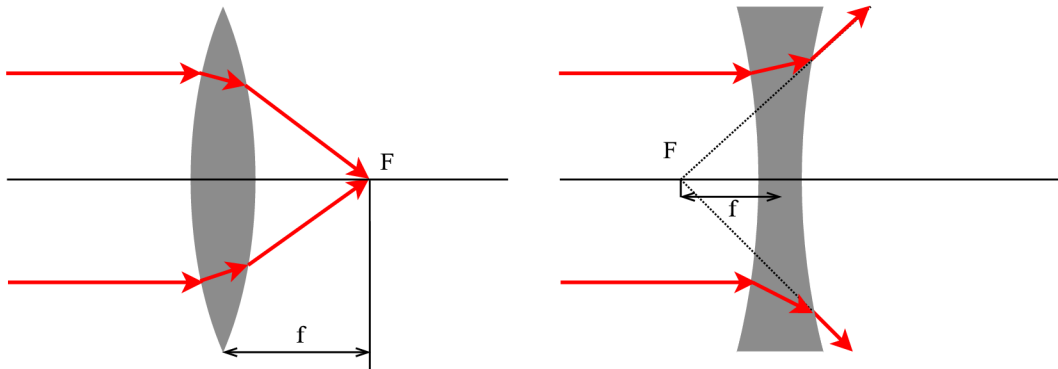
Obrázek 2.7: Schéma akvizice obrazu z digitálních kamer

Proces, kdy se převádí informace z reálného světa do digitální podoby, se nazývá podle knihy [38] akvizice dat. Obecně se jedná o sbírání informací z analogového zdroje a jejich převod do odpovídající digitální podoby, která je vhodná pro přenos dat.

Aby mohlo dojít k získání dat, je důležité nejprve získat odpovídající informaci jako elektrický signál. K tomuto slouží **senzory**, které produkují analogový signál. Dále je potřeba takový signál vhodně modulovat pro převod do digitální podoby. K tomu slouží několik elektrických obvodů, provádějících například zesilování signálů, filtraci různých frekvencí či vzorkovač (sample and hold). Poté je důležitou součástí akvizice dat analogově digitální převodník (A/D převodník), který převádí spojitý analogový signál na diskretní digitální vzorky. Další úpravy dat mohou být také součástí akvizice. Poslední součástí tohoto procesu je uložení získaných dat v digitální podobě na paměťové médium, případně přenos pomocí rozhraní (například v kopírce).

U digitálních kamer, jak uvádí kniha [16], se o akvizici obrazu stará několik součástí. Vstupní informací je v tomto případě světlo, které vstupuje do optické soustavy čoček v objektivu. Následuje filtrování některých vlnových délek a filtrování vysokých frekvencí. Poté je takto upravené světlo filtrováno těsně před obrazovým senzorem přes pole barevných filtrů. Následuje směřování paprsků světla přes pole mikročoček na obrazový senzor, který snímá množství vstupujících fotonů. Výstupem z obrazového senzoru jsou analogové signály, které se dále modulují a filtrují. Následuje převod analogových signálů na signály digitální. Tyto digitální signály se poté vzájemně interpolují, tzv. *demosaičing*. Následuje další zpracování již digitálních dat (například komprese). Posledním krokem je uložení digitálního obrazu na paměťové úložiště. Zjednodušené schéma tohoto procesu je zobrazeno na obrázku 2.7.

Jak již bylo dříve uvedeno v 2.2 čočky jsou podle knihy [17] průhledná tělesa, složená ze dvou, obvykle kulových, lámavých ploch. Tyto plochy mohou být konvexní, konkávní, případně rovinné. Podle toho, jak čočky průchodem světla lámou jeho paprsky, se dělí na spojky a rozptylky.



Obrázek 2.8: Spojka a rozptylka

Spojky jsou takové čočky, které mají vždy jednu plochu vypouklou a vždy jsou uprostřed tlustší, než na okrajích. Díky tomu lámou procházející paprsek do jednoho bodu, do *ohniska*, které se obvykle značí  $F$ . Rozptylky naopak mají střed čočky tenčí než její okraje, díky čemuž se paprsky světla, které procházejí čočkou, rozbíhají. Protože se paprsky rozbíhají, nemají rozptylky reálné ohnisko, ale prodloužením těchto paprsků lze najít tzv. virtuální ohnisko  $F$ . Rozdíl mezi spojkou a rozptylkou je ukázán na obrázku 2.8.

Ohnisková vzdálenost ( $f$ ) tenkých čoček, popsána v knize [32], je určena jako vzdálenost ohniska od čočky. U spojek je tato hodnota kladná, u rozptylek záporná. Ohnisková vzdálenost je dána vztahem

$$\frac{1}{f} = (n - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} + \frac{(n - 1)d}{nR_1R_2} \right), \quad (2.8)$$

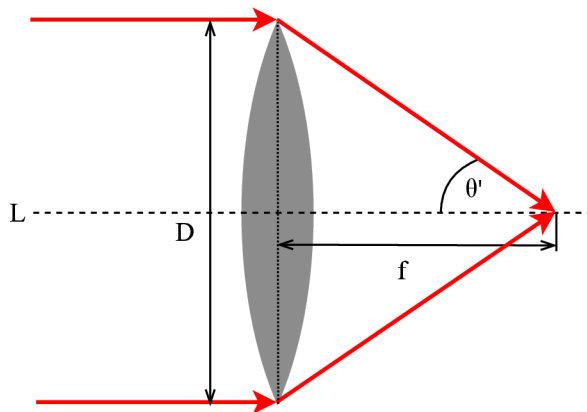
kde  $n$  je index lomu čočky,  $R_1$  je přední zakřivení čočky,  $R_2$  je zadní (obrazové) zakřivení čočky a  $d$  je tloušťka čočky.

Čočky lze skládat do soustavy čoček, u kterých lze snadno najít společnou ohniskovou vzdálenost:

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1f_2}, \quad (2.9)$$

kde  $f_1$  a  $f_2$  jsou ohniskové vzdálenosti dvou čoček a  $d$  je vzdálenost mezi čočkami.

Díky skládání čoček do různých konfigurací lze docílit různých obrazů. Například konfigurací, kde první čočkou je rozptylka a druhou čočkou je spojka, se nazývá *retrofocus* a používá se v širokoúhlých objektivěch. V případě kompaktních fotoaparátů se tato konfigurace čoček používá pro přibližování. Retrofocus má krátkou ohniskovou vzdálenost. Opačná konfigurace čoček, kde první čočkou je spojka a druhou je rozptylka, se nazývá teleobjektiv. Ten má naopak velkou ohniskovou vzdálenost.



Obrázek 2.9: Zobrazení parametrů pro výpočet clonového čísla

Další důležitou hodnotou optických soustav je *clonové číslo*, též nazývané podle knihy [32] F-number. Clonové číslo  $F$  je definováno takto:

$$F = \frac{1}{2 \sin \theta'}, \quad (2.10)$$

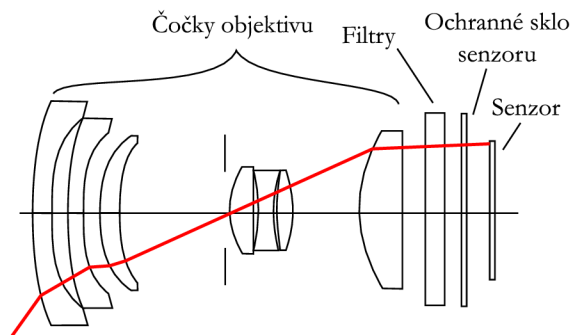
kde  $\theta'$  je polovina úhlu, který svírá vnější paprsek světla, procházejícího čočkou. Čím větší je clonové číslo, tím méně světla projde přes čočku. Pokud je úhel  $\theta'$  velmi malý, může se použít aproximovaná rovnice:

$$F = \frac{f}{D}, \quad (2.11)$$

kde  $D$  je průměr čočky.

Optický subsystém u digitálních fotoaparátů je složen z několika skupin soustav čoček. Podle druhu objektivu se liší počet a druh použitých čoček, ale i počet skupin. Například u většiny kompaktních fotoaparátů je podle knihy [32] objektiv složen ze tří skupin soustav čoček, kde první je skupina konkávní, druhá skupina je konvexní a má za úkol přibližování a oddalování a třetí skupina je konvexní, která má za úkol ostření. Jednotlivé funkce čoček se mění mechanicky, a to posouváním skupin čoček po ose, tedy buď blíže k první soustavě čoček, případně k senzoru. Na obrázku 2.10 je ukázka tří skupin soustav čoček pro digitální kameru.





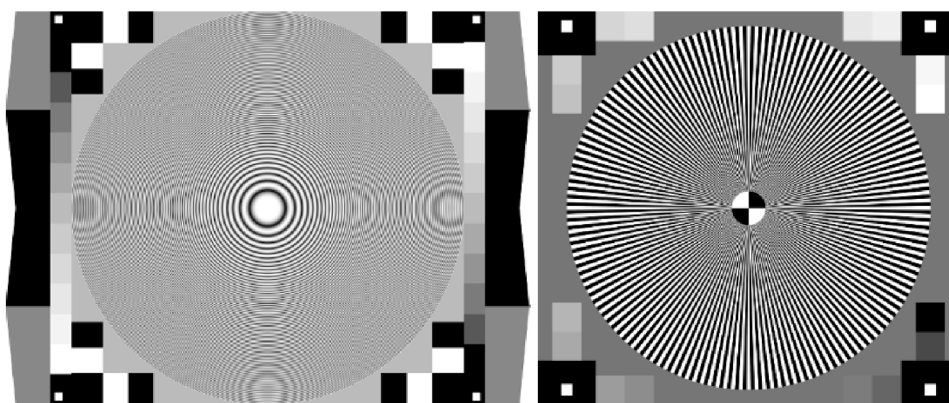
Obrázek 2.10: Optická soustava u digitálních fotoaparátů

**Filtry** slouží k omezení průchodu některých vlnových délek světla. Jak uvádí kniha [32], v digitálních kamerách jsou obvyklé dva filtry:

- filtr pro potlačení infračerveného záření
- filtr pro potlačení vysokých frekvencí.

Jak už název napovídá, filtr pro filtrování infračerveného záření slouží k potlačení infračerveného záření. Toto záření je nechtěné, protože obrazové senzory jsou na něj vysoce citlivé. Infračervený filtr je obvykle umístěn za poslední čočku v objektivu, případně je umístěn před touto čočkou. Většina infračervených filtrů toto záření absorbuje, ale objevují se i takové filtry, které nežádoucí vlnové délky odrážejí zpět.

Druhým důležitým filtrem je dolní propust. Tento filtr má za úkol potlačit vysoké frekvence, které jsou vyšší než Nyquistova frekvence, což je polovina vzorkovací frekvence diskrétního signálu. Důvod, proč se tyto filtry vkládají do digitálních kamer, je zamezení vzniku tzv. moaré efektu, zobrazeného na obrázku 2.11, který je způsoben interakcí pravidelného vzoru objektu (například pravidelné textury) s pravidelným vzorem zobrazovací jednotky.



Obrázek 2.11: Moaré efekt, převzato z knihy [30]

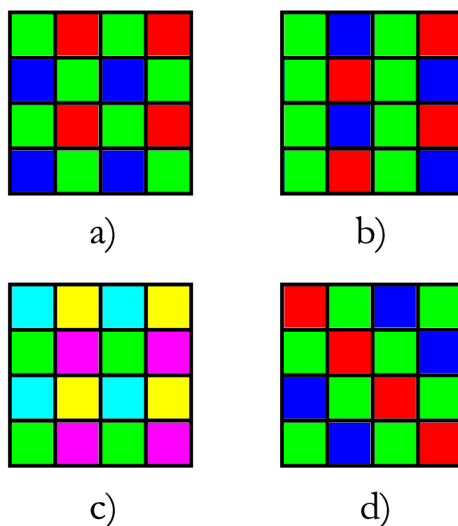
Obrazové senzory jsou podle knihy [32] schopny zaznamenávat intenzitu dopadajícího světla, ale z této podstaty nedokážou zachytit sami o sobě informaci o barvě, jsou to tzv. monochromatické senzory. Z tohoto důvodu je ještě před obrazový sensor umístěno barevné pole filtrů. Technologicky, ale i finančně se nejvíce vyplácí umístit barevný filtr těsně před obrazový sensor.

**Barevné pole filtrů** (anglicky *Color Filter Array, CFA*) je filtr, složený z různých malých barevných filtrů, rozmístěných do matice, kde tvoří pravidelně se opakující vzor (mozaiku). Jak již bylo zmíněno dříve, filtry propouští pouze určité vlnové délky, ostatní jsou pohlceny.

Digitální kamery, podle knihy [36], nejčastěji obsahují:

1. tristimulové barevné pole filtrů, složené z červené, zelené a modré (RGB) barvy, případně žluté, purpurové a azurové (YMC) barvy,
2. pole filtrů, obsahující mix různých RGB a YMC barev, například vzor purpurové, zelené, azurové a žluté (MGCY),
3. pole filtrů, obsahující čtyři a více barev.

Nejčastěji se používají tristimulové RGB filtry, protože představuje nejběžnější barevný formát. Ukázky některých barevných filtrů jsou uvedeny v 2.12.



Obrázek 2.12: Ukázky barevných filtrů, a) je Bayerův, b) je Yamanakův, c) je MGCY a d) je diagonální Bayerův

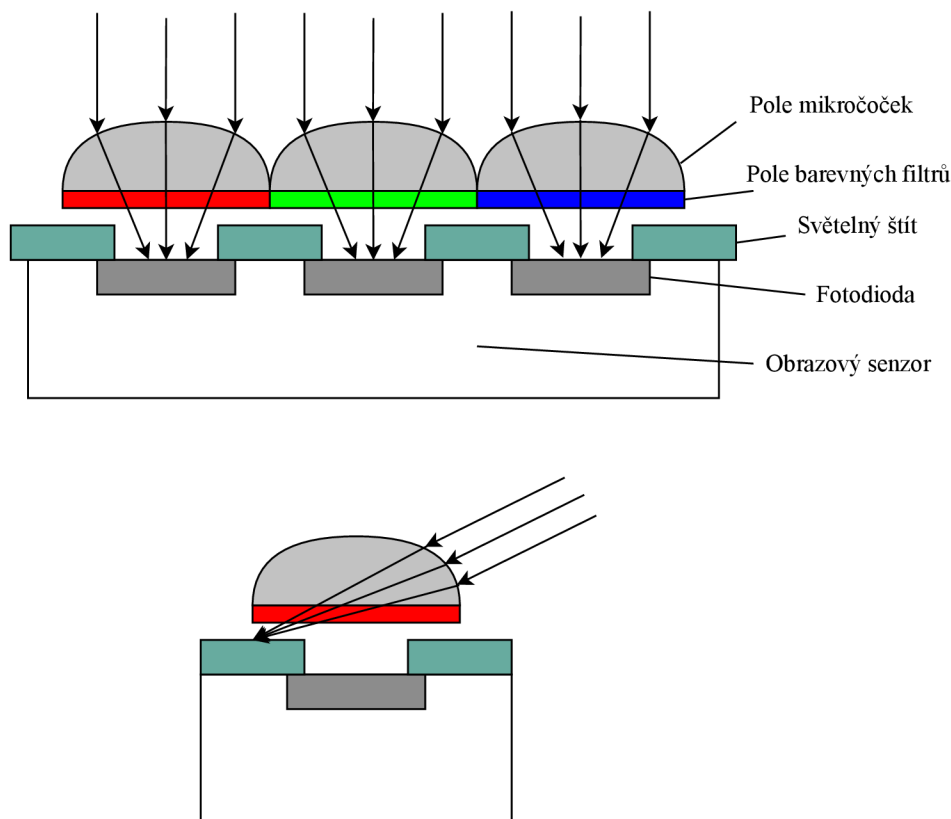
V knize [32] je uvedeno, že nejpoužívanější barevný filtr na trhu je Bayerův (na obrázku 2.12 je uveden pod písmenem a)). Tento filtr obsahuje dvakrát více zelených filtrů, než červených či modrých, protože lidské oko je citlivější na zelené odstíny (lidská sítnice obsahuje nejvíce čípků, citlivých na zelenou barvu). Výrobci často používají i modifikace tohoto filtru, kde jsou například jednotlivé barvy rozloženy do diagonály (na obrázku 2.12 je uveden pod písmenem d)).

**Pole mikročoček** je další důležitá součást digitálních fotoaparátů. Obvykle jsou umístěny těsně před polem barevných filtrů. Hlavním úkolem pole mikročoček je směřovat přicházející paprsky světla na odpovídající fotodiodu obrazového senzoru.

Jak již bylo zmíněno, pole mikročoček je umístěno před polem barevných filtrů, a to tak, aby každá fotodioda obrazového senzoru měla svou vlastní mikročochku. Kniha [26] tvrdí, že je důležité vybrat správný objektiv, aby se nesnižovala účinnost pole mikročoček.

Širokoúhlé objektivy, pokud nejsou navrženy přímo pro použití s digitální kamerou, mohou lámat paprsky světla takovým způsobem, že mikročochky nebudou schopny paprsek

nasměrovat na fotodiodu, ale nasměrují ji někam jinam a v případě CCD obrazových senzorů může proto nastat tzv. přelévání barev. Na obrázku 2.13 je znázorněno směrování paprsků světla, horní část představuje správné lámání, dolní špatné, jež způsobuje vinětaci (vadu, kde okraje výsledného obrazu jsou tmavší než zbytek).



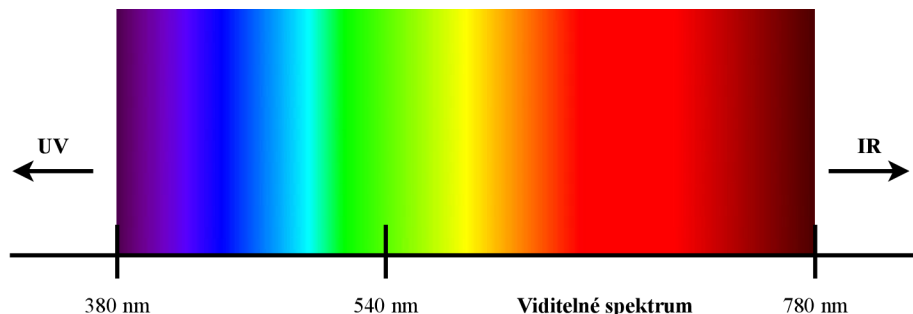
Obrázek 2.13: Nákres řezu mikročoček a ukázka špatného směrování paprsků světla

## 2.3 Obrazový senzor a analogové zpracování signálu, demosaicing

Obrazový senzor je nejdůležitější součást digitálních fotoaparátů. Jak uvádí kniha [32], jedná se o křemíkový čip, který má za úkol převádět optický obraz získaný z optické soustavy kamery na elektronické signály. Obrazový senzor má schopnost detekovat široké spektrum záření, a to jak viditelné záření, tak i ultrafialové nebo infračervené. V digitálních kamerách je ale cílem pracovat pouze s viditelným spektrem lidského oka, což jsou vlnové délky od 380 nm do 780 nm. Spektrum je znázorněno na obrázku 2.14.

Aby bylo možné optický obraz dobře reprodukovat v odpovídajícím rozlišení, je potřeba rozmístit dostatečný počet obrazových prvků (pixelů), a to do řádků a sloupců tak, aby tvořily matici. Tyto pixely poté vstupní světlo převádějí na náboj. U obrazových senzorů jsou kladeny požadavky na co největší obrazovou kvalitu. Toho se snaží výrobci docílit pomocí vysokého rozlišení, velkým dynamickým rozsahem, velkou citlivostí a co nejmenším možným zašuměním.

Nejrozšířenějšími obrazovými senzory jsou **CCD** (Charged-coupled device) a **CMOS** (Complementary metal-oxide semiconductor). Nyní bude v textu popsána funkce obrazových senzorů obecněji, v textu dále (2.3) jsou poté popsány CCD a CMOS senzory více detailněji.



Obrázek 2.14: Spektrum viditelného záření

Podle knihy [32] jsou funkce obrazových senzorů:

- **Fotokonverze:** schopnost polovodičových součástek absorbovat světelný tok fotonů, též nazýváno pojmem *fotoelektrický jev*. Podle knihy [17] platí *Einsteinova hypotéza*:

$$E_{foton} = h \cdot \nu = \frac{h \cdot c}{\lambda} \geq E_g, \quad (2.12)$$

kde  $h$  je Planckova konstanta,  $\nu$  je frekvence světla, respektive  $c$  je rychlost světla a  $\lambda$  je vlnová délka světla.  $E_g$  je energie zakázaného pásu polovodiče. Absorpce fotonů je dána:

$$\frac{d\Phi(x)}{dx} = -\alpha \cdot \Phi(x), \quad (2.13)$$

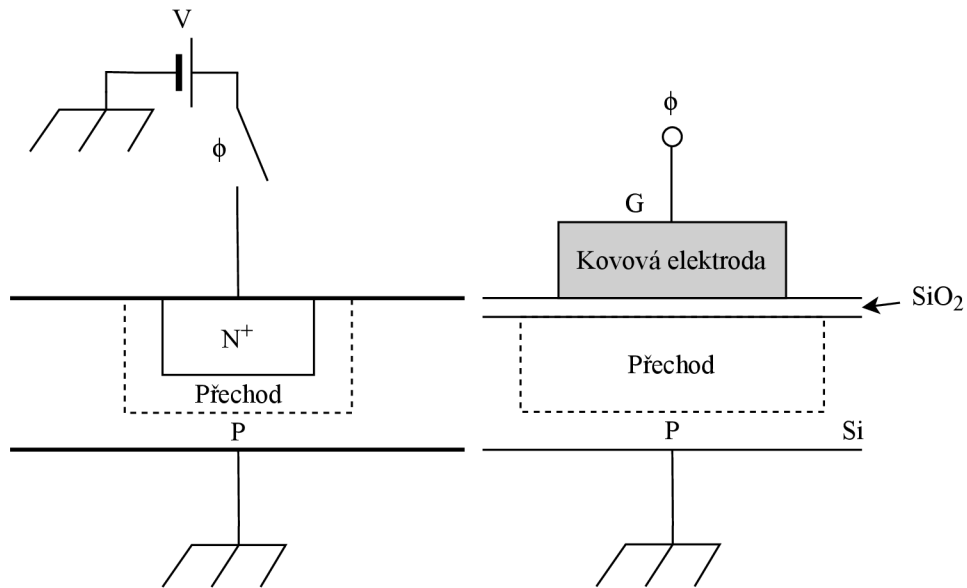
kde  $\alpha$  je absorpční koeficient, závislý na vlnové délce světla. Absorbované fotony poté vytvářejí páry elektron-díra s hustotou danou rovnicí 2.13.

- **Sbírání a akumulace náboje:** zařízení pro akumulování náboje v jednotlivých pixelech obrazových senzorů jsou podle knihy [32] obvykle fotodiody, případně MOS diody (metal-oxide semiconductor). Na obrázku 2.15 jsou zobrazeny zjednodušené řezy fotodiódou a MOS diódou, plnící funkci akumulace náboje.

V prvním případě je oblast  $P$  uzemněna a na  $N^+$  je přivedeno kladné napětí  $V$ . V oblasti  $N^+$  se akumulují elektrony, které jsou excitovány fotony a redukovují elektrický potenciál této oblasti. Díry poté proudí k uzemněné části. Nosičem náboje jsou v tomto případě elektrony.

Pro druhý případ platí, že pokud je na řídicí elektrodu  $G$  MOS diody přiveden záporný potenciál, na přechodu se kumulují nosiče náboje-díry. Této fázi se říká akumulace. V případě, že je na řídicí elektrodu přiveden kladný potenciál, energetická úroveň se začne snižovat a díry se začnou uvolňovat. Tato fáze se nazývá deplece.

Pro oba případy platí, že nosičem náboje jsou elektrony. Dále platí, že některé elektrony se při rekombinaci (zaplnění děr elektrony) ztratí a nedoputují až k přechodu.



Obrázek 2.15: Průřez fotodiodou v závěrném směru a MOS diodou v závěrném směru

- **Skenování matice:** Pro získání hodnoty z naakumulovaného náboje v jednotlivých pixelech v obrazovém senzoru se podle knihy [32] používají dva přístupy: Přesun náboje (Charge transfer) a X-Y adresace. Jako skenování je označen převod signálů, které jsou rozloženy v 2D matici pixelů, na 1D sekvenční signál.

Pro skenování u CCD senzorů se používá několik architektur přesunu náboje:

- **Full-frame transfer (FFT),**
- **Interline transfer (IT),**
- **Frame transfer (FT),**
- **Frame-interline transfer (FIT).**

Pro skenování u CMOS senzorů se používá X-Y adresace. Náboj je převeden na napětí nebo proud tranzistorem přímo uvnitř pixelu. Díky tomu lze přímo adresovat jednotlivé pixely, a to pomocí horizontálního a vertikálního skeneru, což jsou posuvné registry či dekodéry. Vertikální skener vybírá řádky (Y) a horizontální skener sloupce (X).

Jakmile jsou naskenovány informace, je potřeba jednotlivé pixely před integrací náboje inicializovat, aby tam nezůstával náboj. Mód skenování je progresivní, tedy získá se celý rámec řádek po řádku.

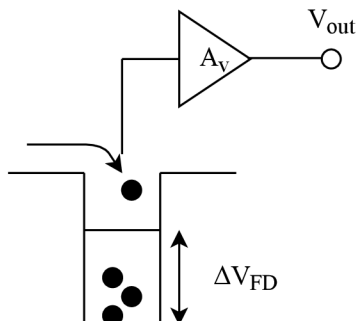
- **Detekce náboje** je podle knihy [32] další funkce obrazových senzorů, identická jak pro CCD, tak i pro CMOS. Náboj  $Q$  padá do potenciálové jámy, která je hlídána napěťovým *bufferem*  $A_v$ . Změna potenciálu  $\Delta V_{FD}$  je dána rovnicí:

$$\Delta V_{FD} = \frac{Q}{C_{FD}}, \quad (2.14)$$

kde  $C_{FD}$  je kapacitance potenciálové jámy. Výsledné napětí je poté určeno jako:

$$\Delta V_{out} = A_V \cdot \Delta V_{FD}, \quad (2.15)$$

kde  $A_V$  je zesílení napětí dané napěťovým *bufferem*. Detekce náboje je znázorněna na obrázku 2.16.



Obrázek 2.16: Schéma detekce náboje

Kapacitance jednotlivých pixelů je omezená a má svojí maximální velikost náboje, který se může nakumulovat. Nazývá se *saturation charge* a je určen:

$$N_{sat} = \frac{1}{q} \int_{V_{reset}}^{V_{max}} C_{PD}(V) \cdot dV, \quad (2.16)$$

kde  $C_{PD}$  je kapacitance fotodiody a  $q$  je náboj elektronu. Počáteční hodnota napětí  $V_{reset}$  a maximální hodnota napětí  $V_{max}$  závisí na struktuře fotodiody.

Signály, které produkují fotodiody, jsou analogové a je potřeba je digitalizovat. U CCD je problém, že není efektivní mnoho zpracování signálů provádět přímo na CCD čipu, a proto se signál zpracovává až ve výstupní logice. U CMOS lze signály zpracovat přímo v pixelu, záleží na použité architektuře. Obvody, které toto zpracování, ale primárně převod do digitální podoby, řeší, se nazývají *Analog front end* (AFE), jak uvádí kniha [32].

AFE má za úkol řízení časovače, který generuje časové signály, které řídí jednotlivé operace obrazových sensorů. Dále má za úkol provést vzorkování signálu, hlavně u CCD obrazových sensorů. Surový výstup z CCD sensoru je postižen mnoha různými druhy šumu a některé se snaží AFE eliminovat pomocí *correlated double sampling* (CDS). Dále provádí převod analogového signálu na digitální, tzv. A/D konverze. Tato konverze sebou přináší i zanesení dalšího možného zašumění (tzv. kvantizační šum). Konverze je závislá na požadovaném výsledném výstupu, od které se odvíjí rychlost a kvalita převodu.

Důležitým parametrem optických sensorů je **velikost sensoru**. Podle knihy [30] je cílem zmenšit velikost obrazového sensoru natolik, aby se vešel například do mobilního telefonu, ale zároveň poskytoval co největší možné rozlišení výsledného snímku.

**Poměr stran** je poměr mezi dlouhou a krátkou stranou sensoru. Používají se primárně tři druhy poměru stran sensoru: 3:2, 4:3 a 16:9. V případě, že výsledné snímky z digitální kamery mají jiný poměr stran, než je poměr sensoru (například čtvercový), tak je pro jeho vytvoření aplikováno ořezání snímku.

**Rozlišení sensoru** je velikost obrazového sensoru, obvykle udávána buď velikostí v milimetrech, případně optickým formátem, jak uvádí kniha [32]. Optický formát je dán diagonálou optického obrazu promítnutého do obrazového sensoru pomocí soustav čoček.

V tabulce 2.1 je uvedeno několik používaných formátů obrazových senzorů, jak uvádí knihy [32] a [30].

Důležitým formátem je Full Frame, který vychází, jak bylo uvedeno dříve, z kinofilmových kamer, kde velikost senzoru odpovídá jednomu oknu v kinofilmu. Často je dáván jako referenční formát k ostatním formátům a je běžně používaný u digitálních zrcadlovek. Čím větší je rozlišení senzoru, tím dražší je senzor.

**Velikost pixelu** je podle knihy [30] velmi proměnlivý parametr u obrazových senzorů. Díky miniaturizaci a větší integraci součástek jsou dnes velikosti jednotlivých pixelů menší než mikron, tedy  $0.9 \times 0.9 \mu\text{m}$ . Poměr stran pixelu je obvykle 1, protože pixely jsou čtvercové.

Formát	Diagonála [mm]	Výška [mm]	Šířka [mm]
1	16.00	12.80	9.60
2/3	11.00	8.80	6.60
1/1.8	8.89	7.11	5.33
1/2	8.00	6.40	4.80
1/2.5	7.20	5.76	4.32
1/2.7	6.67	5.33	4.00
1/3	6.00	4.80	3.60
1/3.2	5.63	4.50	3.38
1/4	4.50	3.60	2.70
1/5	3.60	2.88	2.16
1/6	3.00	2.40	1.80
35 mm (Full frame)	43.27	36.00	24.00
APS-DX	28.37	23.70	15.60
APS-C	27.26	22.70	15.10
APS-H	33.93	28.70	19.10
4:3 (225)	21.63	17.30	13.00
CX	15.86	13.20	8.80
Pentax 645	54.72	43.80	32.80
Hasselblad H5D-50c	61.30	49.10	36.70
Mamiya	66.57	56.00	36.00

Tabulka 2.1: Formáty obrazových senzorů

**Charge Coupled Device (CCD)** senzory jsou prvním představitelem kategorie obrazových senzorů. Princip činnosti CCD senzoru je ukládání a přenos nábojů v polovodiči. Jak je uvedeno v knize [32], základním stavebním kamenem CCD senzoru jsou MOS kondenzátory. V knize [12] je uvedeno, že MOS kondenzátory neprochází elektrický proud ve stacionárním stavu a principiálně jsou to diody, proto lze v literatuře tuto součástku nalézt pod pojmem MOS dioda.

Na obrázku 2.15 se nachází schéma MOS kondenzátoru. Jednotlivé MOS kondenzátory podle knihy [32] mezi sebou interagují a tím si vzájemně přenášejí náboj.

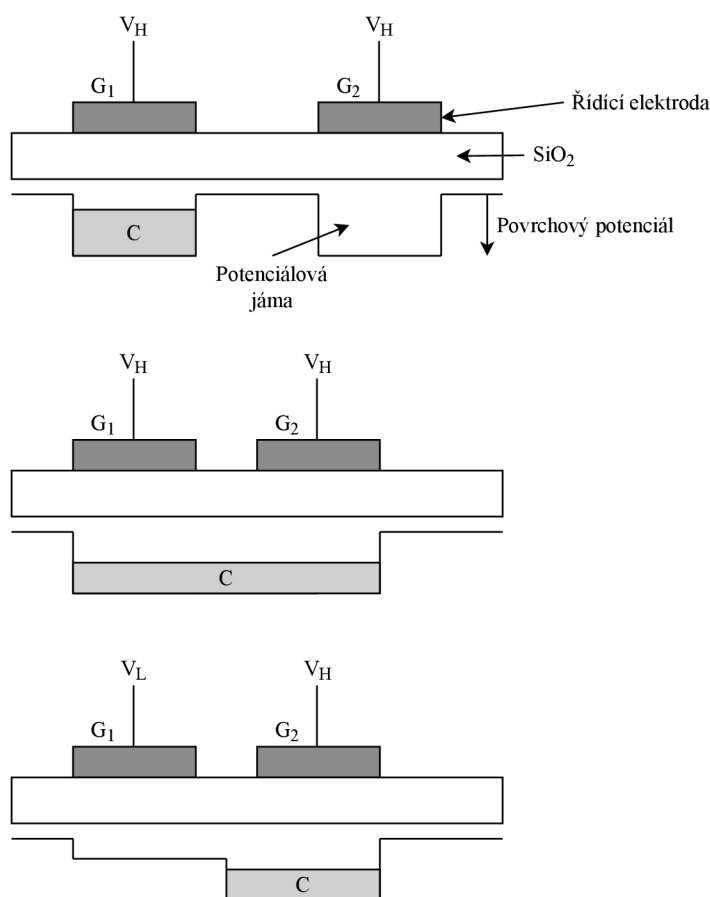
Na obrázku 2.17 jsou znázorněny tři situace, jak mezi sebou mohou MOS kondenzátory interagovat. V první části jsou uvedeny dva MOS kondenzátory, mající dvě řídicí elektrody  $G_1$  a  $G_2$  daleko od sebe. Na obě elektrody je přivedeno vysoké napětí  $V_H$ . Potenciálová

jáma pod  $G_1$  uchovává náboj  $C$ , ale potenciálová jáma pod  $G_2$  je prázdná a nedochází zde k žádným přenosům náboje.

V případě prostředního obrázku 2.17 jsou řídicí elektrody  $G_1$  a  $G_2$  v těsné blízkosti. Potenciálové jámy se spojí a v případě, že je přivedené napětí na obě elektrody vysoké, náboje se akumulují pod oběma elektrodami.

Jakmile je ale na řídicí elektrodu přivedeno menší napětí, náboje se přenesou do potenciálové jámy, která má větší napětí. V případě na obrázku 2.17, dole, je na řídicí elektrodu  $G_1$  přivedeno malé napětí  $V_L$  a náboje se přesunou pod  $G_2$ .

Pokud je tedy takových MOS kondenzátorů v řádku umístěno více, lze přiváděním různých hodnot napětí na řídicí elektrody přesouvat naakumulované náboje z jednoho MOS kondenzátoru do druhého. V tomto případě se tedy CCD chová jako analogový posuvný registr. Tento přístup se nazývá *complete charge transfer mode*.



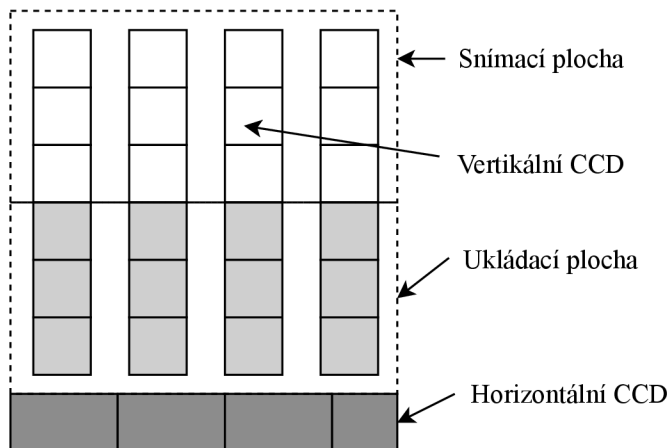
Obrázek 2.17: Chování dvou MOS kondenzátorů

Při získávání náboje (tzv. skenování) se používá několik postupů, vždy daných konstrukcí CCD obrazového senzoru:

- **Frame transfer (FTCCD)** je taková architektura CCD senzoru, která je složena ze snímací plochy (místa, kde dochází k fotokonverzi), plochy uložení náboje, horizontálního přesunu náboje CCD (HCCD) a výstupní logiky, jak uvádí kniha [32]. Ukládací a snímací plochy jsou paralelní a jsou umístěny ve dvou kanálech pod sebou, kde je ukládací plocha pokryta kovovým světelným štítem. Fotony, které přicházejí na elek-



trody ve snímací ploše, se absorbují a generují páry elektron-díra. Elektrony se poté akumulují v potenciálových jámách a tvoří tím signál, který se periodicky přenáší do ukládací plochy. Poté se tyto signály přenesou do HCCD a sériově se postupně přenesou na výstupní okruh. Schéma FTCCD je na obrázku 2.18.

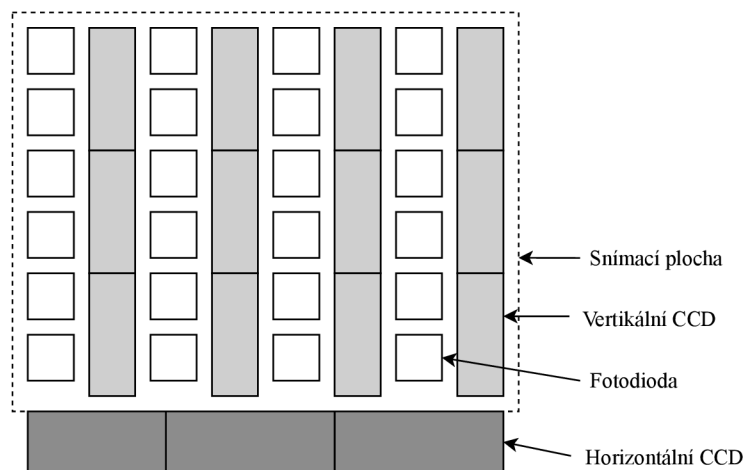


Obrázek 2.18: Frame transfer

Tato architektura produkuje falešné signály, nazývané *smear*, které vznikají díky dopadajícímu světlu na obrazovou plochu, zatímco signály se už přenáší do ukládací plochy. *Smear* se projevuje jako jasný sloupec pixelů (respektive od postiženého místa vertikálně k okraji v celém sloupci) ve výsledném obraze. Lze jej redukovat tím, že se provede přenesení z obrazové plochy do ukládací co nejrychleji pomocí vysokofrekvenčního časového signálu.

Další nevýhodou je podle knihy [32] fakt, že elektrody, které jsou z polykrystalického křemíku, absorbují kratší vlnové délky než modré a díky tomu jsou méně citlivé na modrou barvu. Pokud je použita mechanická závěrka v digitální kameře, může se eliminovat ukládací oblast pro tuto architekturu a hovoří se o *Full-frame transfer* CCD.

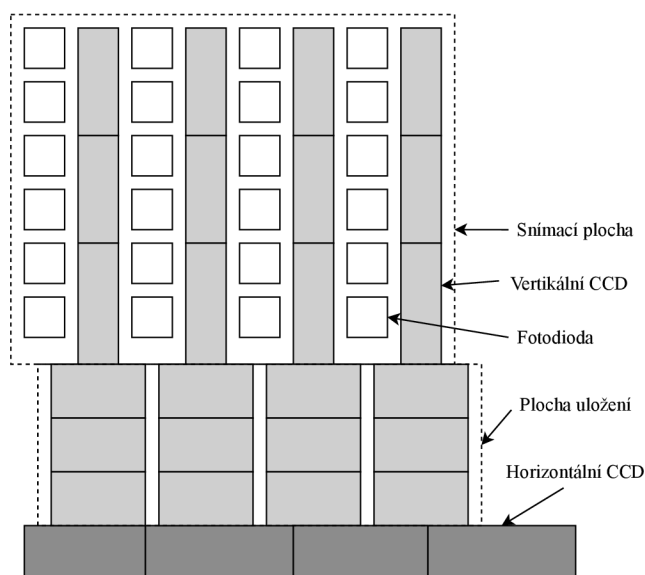
- **Interline transfer** (ITCCD) je další používaná architektura v CCD. Kniha [32] uvádí, že se jedná o nejpoužívanější CCD architekturu v digitálních kamerách. Fotodiody jsou umístěny v obdélníkové mřížce a vertikální CCD (VCCD) jsou umístěny o sloupec vedle. VCCD musí být ochráněny světelným štítem. Při přivedení kladného náboje na řídicí elektrodu fotodiody se náboje téměř okamžitě přenesou do VCCD odkud se dále přenesou postupně do HCCD a poté dále do výstupního okruhu. Schéma je na obrázku 2.19.



Obrázek 2.19: Interline transfer

V tomto případě je *smear* eliminován, může ale nastat, pokud se světlo dostane přes světelný štít. Díky tomu, že VCCD a fotodiody jsou odděleny, lze lépe dosáhnout většího dynamického rozsahu či menšího zašumění.

- **Frame-interline transfer (FITCCD)** je poslední architektura, která se používá u běžných CCD obrazových senzorů. Jedná se o kombinaci ITCCD a FTCCD, kde z ITCCD přebírá stejnou strukturu pixelů a z FTCCD snímací a ukládací plochu. Princip je podobný, jako u výše zmíněných architektur.

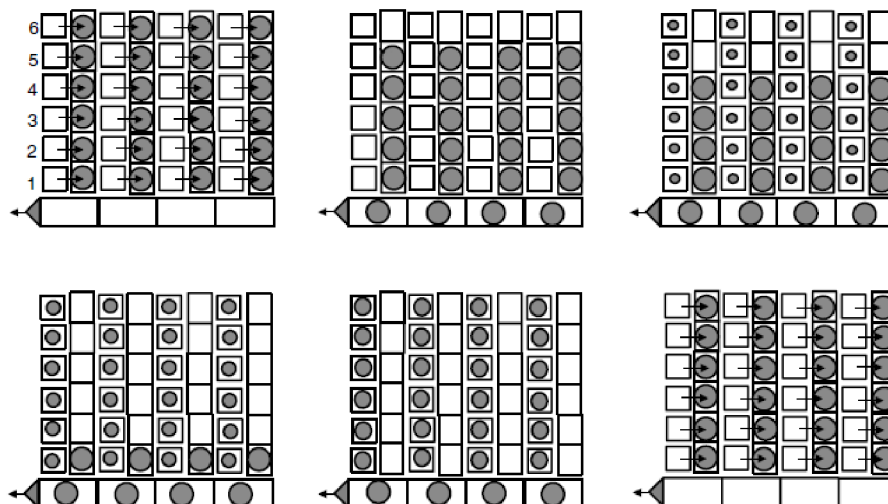


Obrázek 2.20: Frame-interline transfer

FITCCD je odolný na *smear* ve stejné výši jako ITCCD, mohou se tedy generovat elektrony při proniknutí světla štítem. Díky tomu, že je potřeba poměrně velká oblast jak pro snímací plochu, tak i pro ukládací plochu, je výsledný CCD senzor poměrně velký a je potřeba většího napájení jak pro řízení časového signálu, tak pro přenos

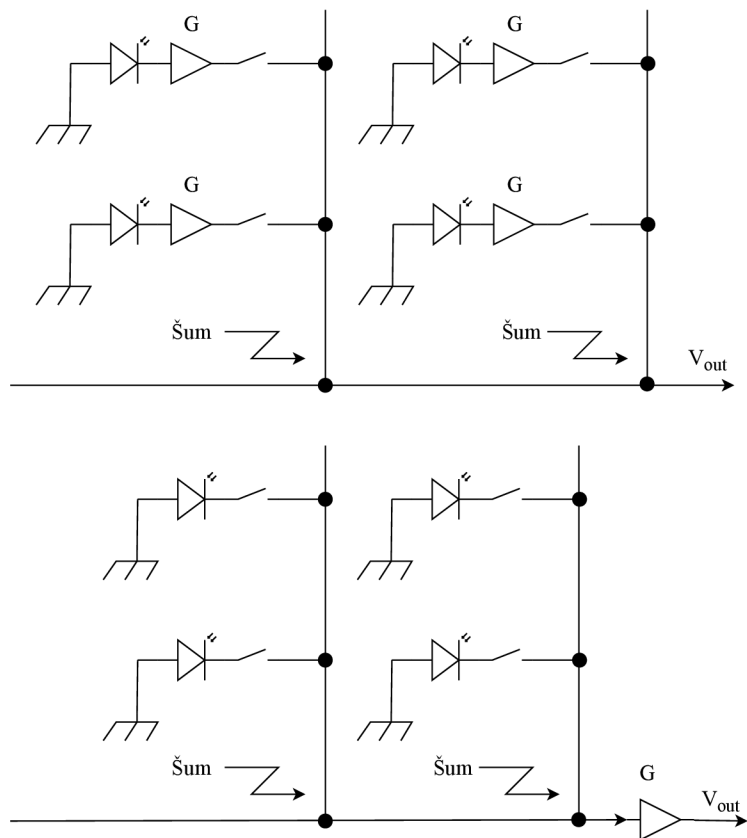
rámce. Z toho důvodu jsou FITCCD senzory používány v dražších digitálních kamerách.

Na obrázku 2.21 je uveden příklad skenování a přenášení náboje u ITCCD. Nejprve se jednotlivé náboje z potenciálových jam přesunou do VCCD. Poté se první řádek posune do HCCD, zatímco se ve fotodiodách kumuluje nový náboj. Takto se postupně posunou všechny řádky od nejnižšího až k nejvyššímu do HCCD. Nakonec začne čtení nového rámce. Tím, že skenuje celý rámec naráz, nazývá se tento postup jako progresivní skenování.



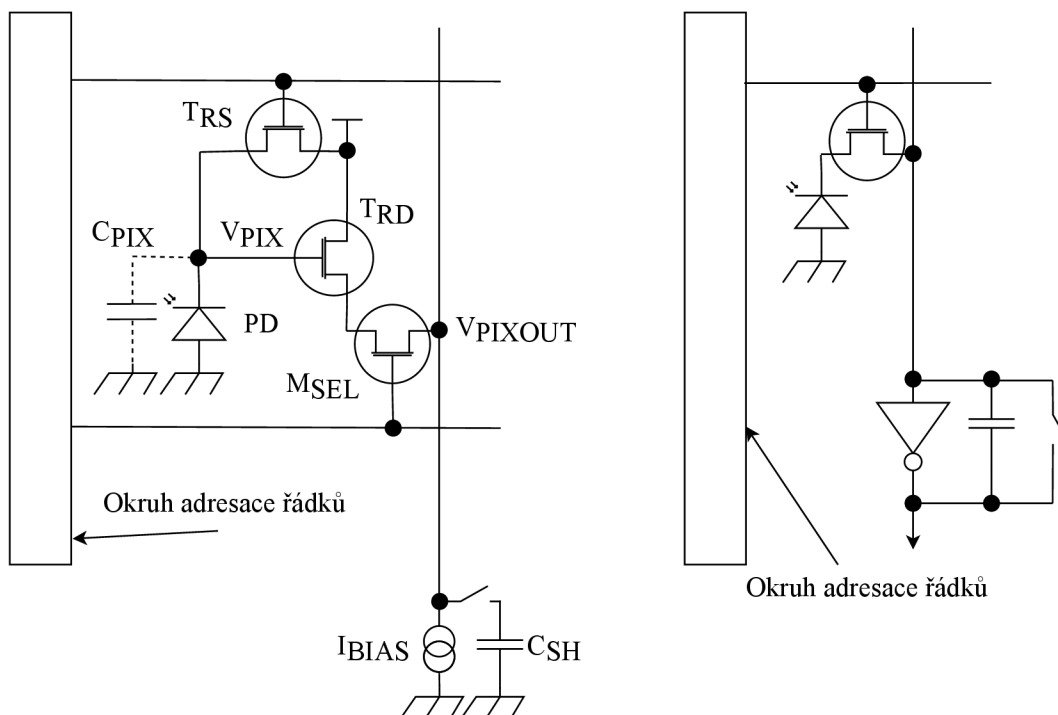
Obrázek 2.21: Ukázka přenosu náboje u ITCCD, převzato z knihy [32]

**Complementary metal-oxide semiconductor (CMOS)** je druhá kategorie obrazových senzorů, které se běžně používají v digitálních kamerách. Technologie CMOS se podle knihy [32] běžně používá v mnoha elektronických součástkách, ať už jde o procesory, paměti či integrované obvody. Výhodou je nízká energetická náročnost či tzv. *on-chip* funkcionalita, tedy možnost integrovat několik obvodů do jednoho malého čipu.



Obrázek 2.22: Aktivní a pasivní pixely

CMOS obrazové senzory používají pro skenování pixelů adresaci X-Y a koncept aktivních pixelů. U CMOS obrazových senzorů jsou podle knihy [32] jednotlivé pixely chápány jako samostatné okruhy složené z fotodiody, resetovacího spínače, signálového zesilovače a výstupního okruhu. Pokud je tento zesilovač přítomen pro každý pixel, nazývá se tato architektura jako **aktivní pixely**. V případě, že je zesilovač přítomen až na výstupním okruhu, jedná se o architekturu **pasivních pixelů**. Rozdíl je znázorněn na obrázku 2.22, kde je několik takových pixelů v matici.

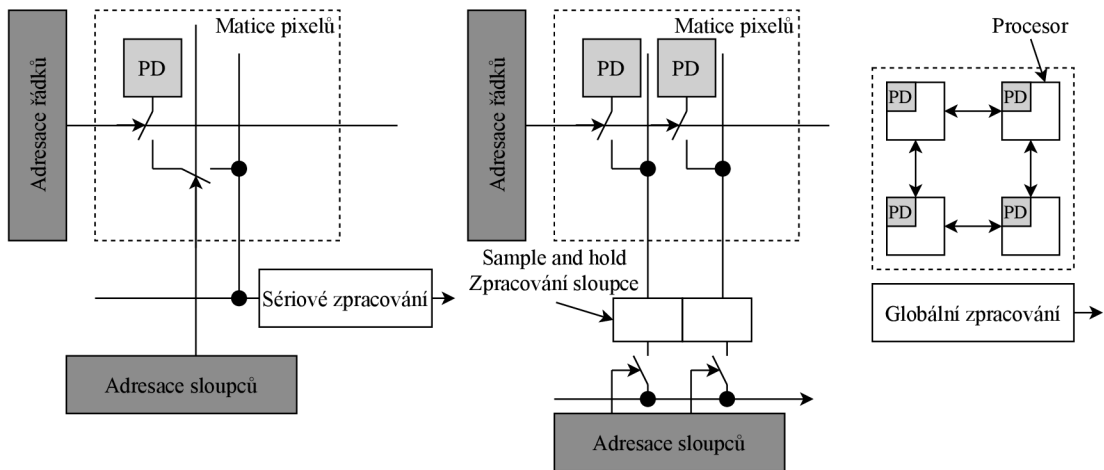


Obrázek 2.23: Schéma jednoho aktivního a pasivního pixelu

Na obrázku 2.23 je naznačen jeden konkrétní pixel u CMOS obrazového senzoru, jak jej popisuje kniha [32]. Vlevo je aktivní pixel, vpravo pasivní pixel. Jedná se o architekturu, kde jsou v pixelu umístěny tři tranzistory. Tranzistor  $T_{RS}$  se používá k restartování fotodiody  $PD$ , tranzistor  $T_{SEL}$  se používá k výběru konkrétního pixelu v senzoru a tranzistor  $T_{RD}$  se používá ke čtení hodnoty (chová se jako *buffer* a zesilovač).  $T_{RS}$  a  $T_{SEL}$  je zapojen na řádkový vodič a výstup pixelu je zapojen na sloupcový vodič. Okruh adresace řádků se chová jako řadič a vybírá řádky, které chce číst. Pokud je na řídicí elektrodu tranzistoru  $T_{SEL}$  přivedeno napětí,  $T_{RD}$  získá hodnotu napětí  $V_{PIX}$ , které je na fotodiodě  $PD$  a výstupní hodnota napětí  $V_{PIXOUT}$  odpovídá napětí na *sample-and-hold* kondenzátoru.

Architektura pasivních pixelů sestává z fotodiody a tranzistoru, který se chová jako přepínač. Náboj vygenerovaný fotodiodou se přečte z pixelu a následně je zesílen na konci okruhu, tedy mimo pole pixelů. Díky velké vzdálenosti mezi zesilovačem a pixelem, kde je cesta sdílená pro další pixely, je vyšší pravděpodobnost výskytu šumu.

**Adresace X-Y** je metoda pro skenování pixelů v CMOS obrazovém senzoru. Kniha [32] uvádí, že adresace probíhá tak, že se používají řádkové a sloupcové skenery. Řádkové skenery vytváří signály pro čtení a resetování hodnoty pixelů, a to jednou pro celý rámeček. Sloupcové skenery čtou sloupce periodicky, když se čtou řádky. Obvody pro adresaci řádků a sloupců jsou tvořeny buď posuvnými registry nebo dekodéry. Výhodou použití posuvných registrů je jednoduchá konfigurace a menší generování šumu. Výhodou použití dekodérů je větší skenovací flexibilita než u posuvných registrů. Lze například získat informace pouze z části obrazového senzoru.



Obrázek 2.24: Architektury pro skenování pixelů

Pro skenování pixelů z CMOS obrazového senzoru kniha [32] uvádí tři architektury, které jsou uvedeny na obrázku 2.24.

- **Sériové čtení** je první uvedený postup pro čtení. Tato architektura plně využívá adresaci X-Y. Řádkové signály, stejně jako sloupcové signály, vybírají pouze jeden pixel v čase, který se přečte a sekvenčně zpracuje. Jakmile je pixel přečten, může se rovnou zahájit sbírání náboje na fotodiodě. Na obrázku 2.24 je tato architektura vyznačena jako první.
- **Paralelní sloupcové čtení** je další metoda, nejčastěji používaná pro skenování. Jednotlivé pixely v řádcích se skenují postupně a zpracovávají paralelně. Jednotlivé signály jsou dočasně uloženy v paměti (sample and hold) a následně jsou sekvenčně čteny. Sběrání náboje na fotodiodě může být zahájeno řádek po řádku.

Pro zahájení čtení pixelu stačí pouze signál pro adresování řádků, hodnota se pozdrží v paměti a následné čtení z této paměti stačí zahájit signálem ze sloupcového skeneru. Tato architektura tedy redukuje počet propojů a má nižší spotřebu energie. Na obrázku 2.24 je uvedena uprostřed.

- **Paralelní čtení** je poslední uvedená architektura pro skenování. Každý pixel obsahuje vlastní procesor a provádí jeho zpracování paralelně. Díky tomu, že pixely mají vlastní procesor, lze zpracovat i pouze části matice pixelů. Takto zpracované signály jdou poté do globálního zpracování celého výsledného obrazu.

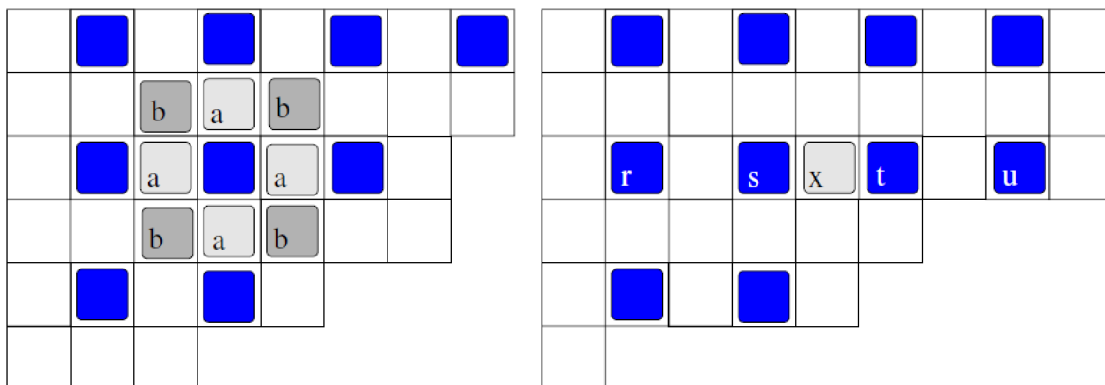
Tato architektura se používá v kamerách, u kterých je potřeba co nejrychlejší snímání. Nevýhodou je, že komplexnost jednotlivých pixelů je vysoká, což způsobuje i větší rozměry obrazového senzoru. Na obrázku 2.24 je uvedena jako úplně poslední.

Závěrem této podkapitoly je porovnání CMOS a CCD obrazových senzorů. Starší modely CMOS obrazových senzorů podle knihy [32] poměrně hodně trpěly velkým výskytem šumu *fixed pattern noise*, způsobovaný temnými proudy. Navíc starší CMOS senzory kvalitou nedostačovaly proti CCD senzorům. V současnosti postupně roste obrazová kvalita CMOS senzorů, a to díky používání nových technologií. Díky dalšímu zpracování signálů z obrazových senzorů se snižuje výskyt různých druhů šumu, které vznikají při akvizici obrazu a postupně se i zvětšuje velikost jednotlivých pixelů a zvětšuje se rozlišení senzoru,

oproti CCD. Spotřeba energie se zmenšuje. CMOS se používají ve všech typech digitálních kamer, od mobilních telefonů po digitální zrcadlovky.

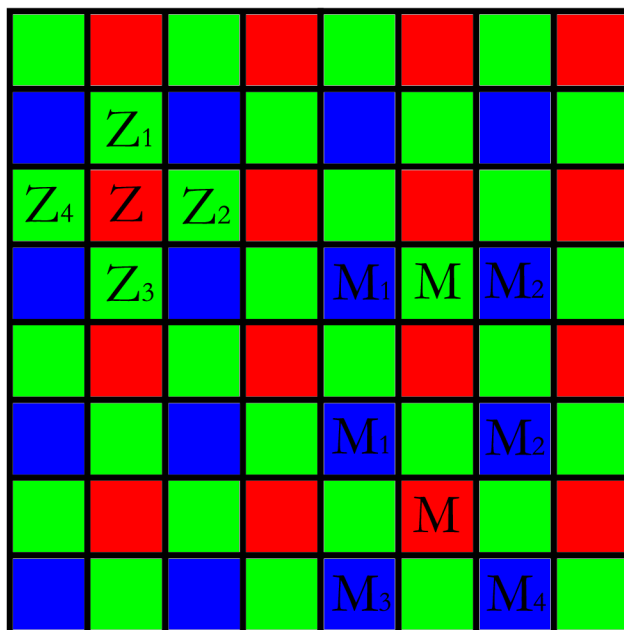
Demosaicing je souhrn algoritmů pro rekonstrukci barev. Výstupem z obrazového senzoru je totiž digitální signál, který nenesení informaci o barvě. Jak bylo zmíněno v předchozích částech, světlo, které obrazový senzor zaznamenává, prochází těsně před vstupem do senzoru přes pole barevných filtrů, kde jednotlivé filtry propustí pouze určité vlnové délky.

Jak uvádí kniha [30], demosaicing se v obvyklých případech provádí přímo v digitální kaměře. Některé digitální kamery ale umožňují uložit získaný digitální signál přímo (formát RAW) bez výrazného dalšího zpracování. Pro demosaicing se používá několik přístupů.



Obrázek 2.25: Interpolace pro modrý kanál při použití RGB CFA, převzato z knihy [30]

**Demosaicing pomocí lineární interpolace** je nejjednodušší způsob pro rekonstrukci barev. Provádí se pro každý barevný kanál zvlášť. První popsanou metodou je lineární interpolace pomocí *nejbližšího souseda*. Je-li použita v poli barevných filtrů CFA Bayerova maska (viz 2.2), není podle knihy [30] vhodné tento druh interpolace použít, protože v nejbližším okolí o vzdálenosti 1, případně o vzdálenosti  $\sqrt{2}$  (úhlopříčně), jsou pixely jiných barev (s výjimkou zeleného kanálu, kde jsou v úhlopříčce zelené pixely). Ukázka pro modrý kanál je na obrázku 2.25, kde okolí o vzdálenosti 1 je označeno písmenem *a* a úhlopříčně je naznačeno písmenem *b*. Při použití tohoto okolí pro interpolaci může docházet k aliasingu.



Obrázek 2.26: Bilineární interpolace

Další metodou interpolace je *bilineární interpolace*. Jedná se o lineární kombinaci pixelů a provádí se lineárním konvolučním jádrem. Při použití Bayerovy masky platí, že pro chybějící zelené pixely se používají průměry čtyřech sousedních pixelů. Tato skutečnost je znázorněna na obrázku 2.26 v levé horní části, kde se pro chybějící zelený pixel  $Z$  vybere nejbližší okolí, složené ze zelených pixelů a vypočítá se jejich průměrná hodnota. Pro modré a červené pixely záleží na pozici. V případě dopočítání modrých či červených pixelů se vyberou buď dva sousední pixely, případně čtyři diagonální pixely a vypočítá se jejich průměr. Na obrázku 2.25 jsou tyto pixely označeny písmeny  $a$  a  $b$ , přičemž  $b$  pixely se dopočítávají ze sousedních  $a$  pixelů, jedná se tedy o stejný postup jako pro pixely  $a$ .

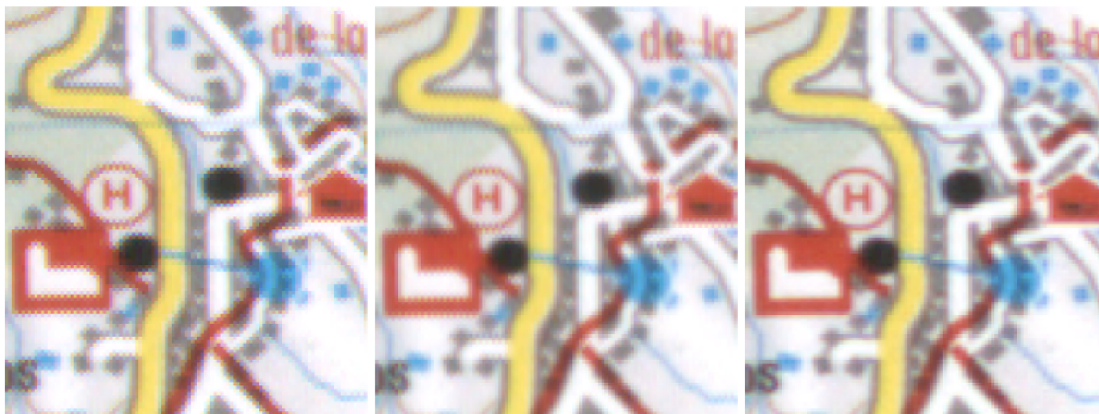
*Bikubická interpolace* je interpolací třetího řádu. Používá větší okolí pixelu než interpolace nižšího řádu a je výpočetně náročnější. Existuje více druhů konvolučních jader pro provedení bikubické interpolace. Obvykle používají interpolační funkce a první derivace. Výsledné snímky jsou poté hladší, bez viditelných schodovitých efektů apod. Pro rekonstrukci obrazu se nejčastěji používají kubické splajny a Catmull-Romova interpolace.

Například Catmull-Romova interpolace, jak uvádí kniha [30], dopočítá chybějící modrý pixel  $x$ , uvedený na obrázku 2.25 vpravo, jako:

$$B_C = \frac{1}{16} \begin{bmatrix} -1 & 9 & 9 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \\ u \end{bmatrix}. \quad (2.17)$$

Podobně pro další pixely, se berou buď sousední pixely, případně pixely umístěné v diagonále. Díky použití většího okolí se redukuje i aliasing. Srovnání zmíněných metod je uvedeno na obrázku 2.27. V prvním segmentu je lineární interpolace nejbližším sousedem, následuje bilineární interpolace a bikubická interpolace. Vyšším řádem se snižuje aliasing, který lze vidět u hran barevných částí, kde tvoří tzv. šachovnicový efekt.





Obrázek 2.27: Srovnání metod lineární interpolace, převzato z knihy [30]

Další možnou strategií pro demosaicing jsou výpočetně složitější, **nelineární interpolace**, a to buď v jednotlivých barevných kanálech zvlášť, případně proložení všech kanálů do jednoho. Například lze nahradit operaci počítání průměru okolních pixelů vypočítáním mediánu. Jednou z uvedených metod pro práci nad všemi barevnými kanály současně je *Hamilton-Adamsova metoda*, která používá horizontální a vertikální gradienty od počítačného pixelu.

Jakmile jsou výsledné snímky vytvořeny, dovolují některé digitální kamery další zpracování obrazu přímo uvnitř zařízení, záleží na výrobci co se rozhodne pro určitý model použít. Kniha [32] uvádí jako příklady, že se mohou v kamerách vyskytovat procedury pro vyvažování bílé barvy, gama korekce, redukce šumu, elektronické přibližování, automatické ostření, ořezání snímku apod. Další důležitou operací při zpracování obrazu je komprese obrazu, která předchází uložení snímku.

Jedním z nejčastěji používaným algoritmem pro kompresi je **JPEG komprese**, popsaná v knize [30]. Používají se dva formáty, JPEG z 80. let 20. století a JPEG 2000, který je novější. Oba formáty obrazu jsou standardizované. JPEG kódování sestává z šesti kroků:

1. Konverze z RGB do jiného barevného prostoru, využívající jas (Y) a chrominační složky (Cr a Cb): YCrCb.
2. Podvzorkování chrominačních složek za použití proměnného faktoru (obvykle 2).
3. Rozdělení plochy obrazu na bloky  $8 \times 8$  pro jasovou složku a  $16 \times 16$  pro chrominační složky.
4. Dvourozměrná diskretní kosinová transformace nad jednotlivými bloky.
5. Kvantizace jednotlivých vypočítaných koeficientů, aby se redukovala jejich entropie.
6. Čikcak čtení jednotlivých koeficientů v bloku a jejich další kódování (Lempel-Ziv kódování, Huffmanovo kódování apod.).

Pro zobrazení (dekódování) komprimovaného obrázku se používá opačný postup, kromě kroků 2 a 5 výše uvedeného dekódování. Tím, že lze volit kompresní faktor, lze ovlivňovat i kvalitu (a velikost) výsledného snímku.

V knize [39] je uvedeno, že standard JPEG umožňuje podle použití jak ztrátovou, tak bezztrátovou kompresi. Existuje několik přístupů (módů) ke kódování, a to sekvenční a pro-

gresivní metody založené na diskretní kosinové transformaci, bezztrátový mód a hierarchické kódování. Nejčastěji používaným módem pro JPEG kódování je sekvenční kódování pomocí diskretní kosinové transformace, popsany výše.

## 2.4 Šum produkovaný obrazovými senzory

Šum je nechtěný signál, který se vyskytuje v signálech a zhoršuje jeho kvalitu. Kniha [32] uvádí, že šum vyjadřuje míru citlivosti obrazového senzoru. V obrazových senzorech se vyskytuje několik různých šumů, které zhoršují kvalitu výsledného digitálního obrazu. Na obrázku 2.28 je uvedena tabulka, která klasifikuje šумы do několika kategorií, a to podle druhu osvětlení pořizované scény a též podle toho, jestli je šum vázán na určité pixely v prostoru nebo ne.

		Dark	Illuminated	
			Below saturation	Above saturation
Fixed Pattern Noise (FPN)		Dark signal nonuniformity Pixel random Shading	Photo-response nonuniformity Pixel random Shading	
		Dark current nonuniformity (Pixel-wise FPN) (Row-wise FPN) (Column-wise FPN)		
		Defects		
Temporal Noise		Dark current shot noise	Photon shot noise	
		Read noise (Noise floor) Amplifier noise, etc. (Reset noise)		
				Smear, Blooming
Image Lag				

Obrázek 2.28: Šумы produkované obrazovými senzory, převzato z [32]

Obrazové senzory produkují šum, který je vázán na určitou pozici v obraze. Tento šum se nazývá **fixed-pattern noise** (FPN), ten lze dalším zpracováním obrazu odstranit nebo alespoň redukovat. Šum, který v obraze kolísá v čase, se nazývá *temporální*, případně lze tento šum najít v literatuře pod termínem *náhodný*. Takový šum je možno zachytit pouze jednou.

První popsanou kategorií je **fixed pattern noise** (FPN). Podle knihy [32] se tento šum v signálu vyskytuje vždy. A to díky nesejnosti jednotlivých použitých polovodičových součástek. Každý pixel totiž reaguje na světlo trochu jinak. FPN lze algoritmicky potlačovat.

Existují dva druhy FPN, které se v obrazových senzorech vyskytují. FPN, který se objevuje u snímků pořízených ve tmě (bez jakéhokoliv vnějšího osvětlení), se nazývá podle článku [41] **nerovnoměrný zbytkový signál za tmy**, anglicky *dark signal nonuniformity* (DSNU). Při průchodu světla přes optickou soustavu na obrazový sensor, se FPN nazývá **nerovnoměrnost citlivosti pixelu**, anglicky *photoresponse nonuniformity* (PRNU).

U CCD obrazových sensorů je hlavní složkou fixed pattern noise **nerovnoměrnost temného proudu**, anglicky *dark current nonuniformity* (DCNU). Temný proud je vygenerovaný náboj, který vzniká ve fotocitlivých součástkách, když na ně nedopadá žádné světlo. Tento šum lze pozorovat u snímků, které byly pořízeny při vysoké teplotě, případně u snímků, kde byla při akvizici dlouhá doba expozice. U CMOS obrazových sensorů, jak uvádí kniha [32], je hlavní složkou FPN právě tvořena DCNU a rozdílnou výkonností jednotlivých tranzistorů v pixelu.

**Temný proud** vzniká v obrazových senzorech, pokud na ně nedopadá světlo. Je to akumulovaný temný náboj, který vzniká v jednotlivých pixelech obrazového senzoru. Temný nábor redukuje dynamický rozsah obrazového senzoru. Některé pixely mohou být na vznik temných proudů více náchylné a při akvizici jsou potom tyto pixely jasnější (tzv. *white spot defects*). Temný proud vzniká i v posuvných registrech u CCD obrazových sensorů a v aktivních tranzistorech u CMOS obrazových sensorů.

Temný proud  $I_D$  má aktivační hodnotu energie okolo  $E_g/2$ :

$$I_D \propto \exp\left(\frac{-E_g}{2kT}\right). \quad (2.18)$$

**Temporální šum** Temporální (též náhodný) šum je takový signál, který se objevuje náhodně v čase. V knize [32] je uvedeno, že rozdělení takového šumu je Gaussovo, kde jeho směrodatná odchylka určuje míru temporálního šumu. Temporální šum je složen z několika různých šumů: termálního šumu, výstřelového šumu a  $1/f$  šumu.

**Termální šum** je náhodný šum, který podle knihy [32] vzniká díky tepelnému pohybu elektronů. Též je nazýván Johnsonův šum. Výkonová spektrální hustota tohoto šumu je pak dána jako:

$$S_V(f) = 4kTR \left[ \frac{V^2}{Hz} \right], \quad (2.19)$$

kde  $k$  je Boltzmannova konstanta,  $T$  je termodynamická teplota a  $R$  je rezistance.

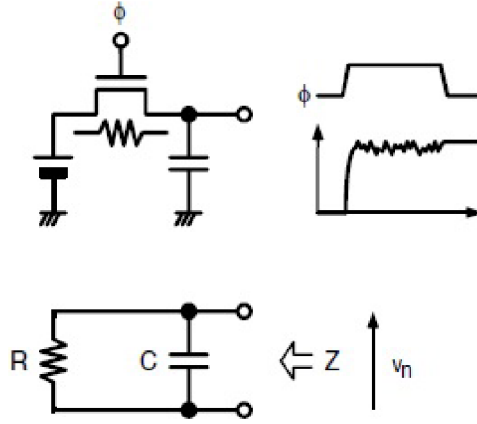
**Výstřelový šum** je další náhodný šum, vyskytující se v obrazových senzorech. Vzniká, když proud překročí potenciálovou bariéru v polovodičových součástkách. V obrazových senzorech souvisí s dopadajícími fotony a temným proudem. V knize [32] je uvedeno, že výstřelový šum je definován Poissonovým rozdělením pravděpodobnosti:

$$P_N = \frac{(\bar{N})^N \cdot e^{-\bar{N}}}{N!}, \quad (2.20)$$

kde  $N$  je počet částic a  $\bar{N}$  je průměrná hodnota částic. Výkonová spektrální hustota termálního a výstřelového šumu je konstantní ve všech frekvencích a tento šum je nazýván *bílý šum*.

**$1/f$  šum** je dalším zástupcem temporálního šumu. Tento šum je podle knihy [32] nazýván *ružový šum*. Výkonová spektrální hustota takového šumu je proporční k  $1/f^\gamma$ , kde  $\gamma$  je obvykle hodnota z intervalu  $(0, 3)$ . Průměrná hodnota tohoto šumu není v čase konstantní. Tento šum se obvykle vyskytuje v oblastech s nízkou frekvencí. Digitální kamery ve většině případů takový šum potlačují automaticky.

**kTC šum** je druh temporálního šumu, který se vyskytuje přímo v obrazových senzorech. Též je nazýván *reset noise*. Tento šum, podle knihy [32], se začne vyskytovat v obrazových senzorech v té době, kdy je MOS přepínač nastaven na režim vypnuto. Šum se začne objevovat v uzlech, kde dochází ke kapacitanci. Je klasifikován jako termální šum, pocházející z MOS přepínačů.



Obrázek 2.29: kTC šum, převzato z knihy [32]

Na obrázku 2.29 je naznačen obvod, který je ekvivalentní k výše popsanému problému. Při stavu zapnuto se MOS tranzistor chová jako rezistor, termální šum se zde vyskytne a je zde držen (sample and hold). Náboj šumu je pak dán jako

$$q_n^2 = C^2 \cdot v_n^2 = kTC, v_n^2 = \frac{kT}{C}, \quad (2.21)$$

kde  $T$  je funkce teploty a  $C$  je kapacitance. I kTC šum lze potlačovat.

**Read noise**, případně *noise floor* je další, v knize [32] popsaný temporální šum, který se objevuje v obrazových senzorech. Je definován jako šum, který vzniká při čtení hodnot z elektronických součástek. U CCD obrazových senzorů se tento šum generuje ve výstupním zesilovači poté, co jsou všechny hodnoty přeneseny přes posuvné registry. U CMOS obrazových senzorů se tento šum vyskytuje v zesilovači uvnitř pixelu.

V MOS tranzistorech se na řídicích elektrodách (gate) tvoří dva druhy šumy, a to termální šum a  $1/f$  šum. Tento termální šum je popsán jako

$$v_{eq}^2 = \frac{4kT\alpha}{g_m} \cdot \Delta f [V^2], \quad (2.22)$$

kde  $g_m$  je transkonduktance MOS tranzistoru a  $\alpha$  je koeficient závislý na módu MOS tranzistoru.  $1/f$  šum je v obrazovém senzoru definován jako

$$v_{eq,1/f}^2 = \frac{K_f}{C_O W L} \cdot \frac{\Delta f}{f} [V^2], \quad (2.23)$$

kde  $K_f$  je konstanta závislá na výrobním procesu a  $C_O$ ,  $W$  a  $L$  jsou kapacitance, šířka a délka řídicí elektrody (gate). Díky těmto rovnicím lze odhadnout, kolik tohoto *read noise* obrazový senzor vygeneruje.

**Náhodný šum pixel-to-pixel** vzniká u CMOS obrazových senzorů, jak uvádí kniha [32]. Charakterizuje se tím, že mezi jednotlivými pixely v CMOS obrazovém senzoru je náhodný. Náhodný šum pixel-to-pixel je buď pixelový FPN nebo pixelový temporální šum. Pro pixelový FPN platí, že je způsoben vlivem temných proudů, citlivostí fotodiod, kondenzátory a charakteristikami tranzistorů. Pixelový temporální šum je ovlivněn resetovacím šumem fotodiod, výstřelovým šumem a charakteristikami tranzistoru.

## Kapitola 3

# Matematický aparát

Tato kapitola obsahuje základní definice matematických postupů a algoritmů, které byly v práci použity. Ani v této kapitole se nemůže jednat encyklopedický přehled, protože rozsah práce by byl příliš velký, proto jsou zde uvedeny pouze informace s bezprostředním vztahem k této práci.

### 3.1 Odstranění šumu v obraze

Díky tomu, že při akvizici obrazu (jak je uvedeno v rovnici 2.7) vzniká ve snímku několik různých šumů, jsou snahy takové šумы potlačit. Odšumění probíhá buď jako součást akvizice obrazu, případně lze provádět odšumění algoritmicky později.

Existuje několik druhů **odšumovacích technik**, jak uvádí článek [14]. Požadavky na takové filtry a techniky jsou takové, aby se podobné plochy vyhladily, hrany zůstaly ostré, textury se zachovaly a nevznikaly nové artefakty.

**Anizotropní difuze** je algoritmus pro redukování šumu v obraze. Šum dokáže vyhladit v rovnoměrných oblastech a zároveň dokáže zachovat kvalitu hran, jak uvádí článek [34].

$$I_t = \operatorname{div}(c(x, y, t)\nabla I) = \nabla c \cdot \nabla I + c(x, y, t)\Delta I, \quad (3.1)$$

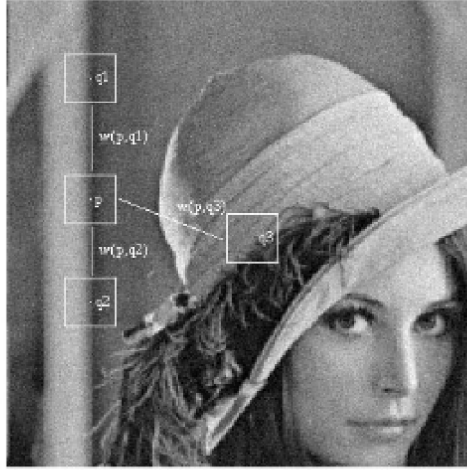
kde  $\operatorname{div}$  je operátor divergence,  $\nabla$  je gradient,  $\Delta$  je laplacián,  $c$  je difuzní koeficient a  $t$  je čas či měřítko. Difuzní koeficient nabývá hodnoty blízké 0, pokud jsou v okolí hrany, u rovnoměrných ploch nabývá hodnoty blízké 1. Tyto koeficienty jsou určeny pomocí dvou možných rovnic. První upřednostňuje vysoce kontrastní hrany nad méně kontrastními a druhá rovnice upřednostňuje větší oblasti oproti menším:

$$c(\nabla I) = e^{-\left(\frac{\|\nabla I\|}{K}\right)^2}, \quad (3.2)$$

$$c(\nabla I) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2}, \quad (3.3)$$

kde  $K$  je fixní konstanta, která určuje citlivost odšumění vůči hranám. Tato konstanta jde i algoritmicky odhadnout.

**Non-local means** je podle článku [8] založen na nelokálním průměrování všech snímků v obraze. Algoritmus vyhledává pro všechny pixely v obraze podobné regiony, ze kterých vypočítá průměrnou hodnotu, danou Euklidovskou vzdáleností. Na obrázku 3.1 je toto znázorněno. Výsledné snímky jsou odšuměné a zároveň je zachována ostrost hran.



Obrázek 3.1: Ukázka oblastí pro výpočet vah odšumění, převzato z článku [8]

Non-local means je diskrétně definován pro obrázek  $v = \{v(i)|i \in I\}$ , kde  $i$  je pixel ze snímku, jako rovnice:

$$NL[v](i) = \sum_{j \in I} w(i, j)v(j), \quad (3.4)$$

kde třída vah  $\{w(i, j)\}_j$  závisí na podobnosti pixelů  $i$  a  $j$ . Zároveň platí podmínky, že  $w(i, j) \in \langle 0, 1 \rangle$  a  $\sum_j w(i, j) = 1$ .

Podobnost pixelů  $i$  a  $j$  závisí na podobnosti oblastí  $v(N_i)$  a  $v(N_j)$ , kde  $N_x$  je čtvercové okolí pixelu  $k$ . Podobnost se určí jako L2 norma (Euklidovská vzdálenost) těchto dvou oblastí:

$$\|v(N_i) - v(N_j)\|_{2,a}^2, \quad (3.5)$$

kde  $a$  je směrodatná odchylka Gaussova jádra. Váhy  $w(i, j)$  se poté zjistí jako:

$$w(i, j) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}}, \quad (3.6)$$

kde  $h$  je stupeň filtrace a  $Z(i)$  je normalizační konstanta, dána jako:

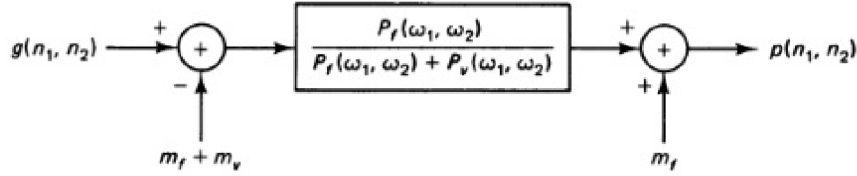
$$Z(i) = \sum_j e^{-\frac{\|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}}. \quad (3.7)$$

Složitost tohoto algoritmu je kvadratická pro jednotlivé pixely v obraze. Jak uvádí článek [14], existuje několik variant tohoto algoritmu, které jsou rychlejší. Dále je tento algoritmus základem pro jiné, robustnější odšumovací filtry (například pro BM3D).

**Adaptivní Wienerův filtr** je druh filtrace, který se používá k obnovování obrazu (anglicky *image restoration*), jak uvádí kniha [28]. Z obrazu, ve kterém se provádí filtrace, se získávají parametry pro filtraci  $H(\omega_1, \omega_2)$ , které se poté na takový obraz aplikují, a to pro lokální oblasti, pro které je takový filtr navržen.

Pokud snímek (respektive zpracovávaná oblast) neobsahuje mnoho detailů, tak je na takovém místě šum více viditelný, než v oblasti, obsahující detaily a vyšší frekvence. Takový

filtr se potom chová v málo detailních oblastech jako dolní propust. V oblastech, obsahující vysoké frekvence (hrany), se provede filtrace dolní propustí slaběji, tedy s parametry vypočítanými tak, aby nebyla taková filtrace příliš agresivní.



Obrázek 3.2: Schéma Wienerova filtru, převzato z knihy [28]

Na obrázku 3.2 je uvedeno schéma pro filtrace Wienerovým filtrem, jak jej uvádí kniha [28]. Jsou zde použity jako parametry pro určení vlastností filtrace  $m_f$ , což je střední hodnota snímku (resp. oblasti),  $m_v$  je střední hodnota šumu a  $P_x(\omega_1, \omega_2)$  je výkonostní spektrum  $x$ . Samotný Wienerův filtr je definován následujícím předpisem:

$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_v(\omega_1, \omega_2)} = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2}. \quad (3.8)$$

Pro výkonostní spektrum aditivního šumu  $P_v(\omega_1, \omega_2)$  se podle knihy [28] předpokládá, že má nulový průměr a konečný rozptyl  $\sigma_v^2$ , tedy že je to tzv. bílý šum. Platí tedy:

$$P_v(\omega_1, \omega_2) = \sigma_v^2. \quad (3.9)$$

Ve filtrované oblasti platí pro signál  $f(n_1, n_2)$  vztah

$$f(n_1, n_2) = m_f + \sigma_f w(n_1, n_2), \quad (3.10)$$

kde  $m_f$  je střední hodnota signálu  $f$ ,  $\sigma_f$  je směrodatná odchylka signálu  $f$  a  $w(n_1, n_2)$  je bílý šum. Výsledný zpracovaný snímek (respektive oblast) je poté definován jako

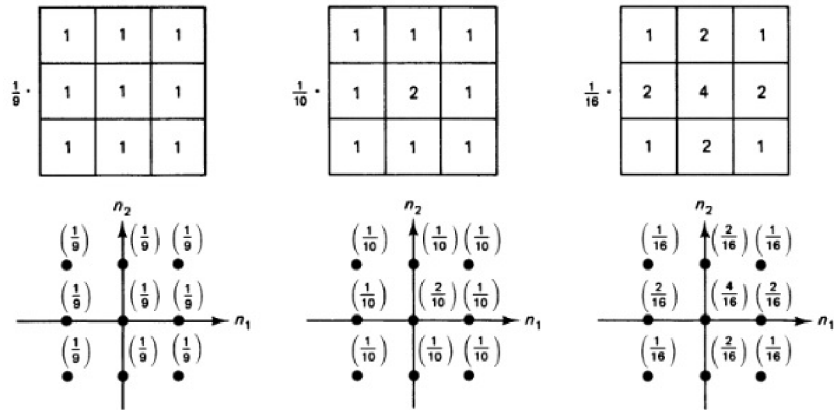
$$p(n_1, n_2) = m_f + \frac{\sigma_f^2}{\sigma_f^2 + \sigma_v^2} (g(n_1, n_2) - m_f), \quad (3.11)$$

kde  $g$  je vstupní snímek, respektive zpracovávaná oblast.

**Box filtr** je druh filtru, který podle knihy [28] provádí tzv. vyhlazení šumu. Obsahuje-li snímek  $I$  nízké frekvence, tak je energie v těchto místech vyšší oproti jiným lokalitám ve snímku. Důvodem jsou vysoké prostorové korelace mezi sousedními pixely. Díky redukci vysokofrekvenčních částí v obraze lze při zachování nízkých frekvencí redukovat šum v obraze.

Filtrace dolní propustí umožňuje odstraňovat například kvantizační šum. Aplikováním filtrace na snímek se zredukuje aditivní šum, ale zároveň se provede i rozmazání snímku. Existuje široké množství filtrů dolní propustí. Ukázky některých filtrů a jejich impulzních odezev je na obrázku 3.3.



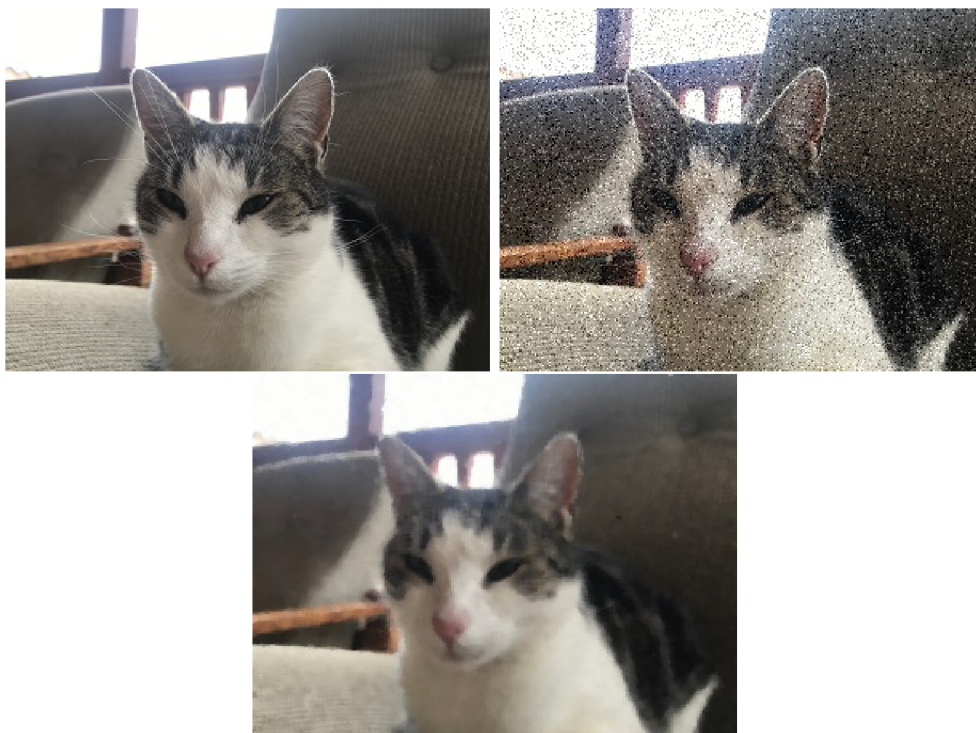


Obrázek 3.3: Impulzní odezvy filtrů dolní propusti, převzato z knihy [28]

**Mediánový filtr** je podle knihy [28] filtrace, která je nelineární. Důležitou vlastností takového filtru je, že redukuje náhodný šum a zároveň zachovává kvalitu hran. Též je uvedeno, že se mediánový filtr hodí při odstraňování šumu typu sůl a pepř.

Při aplikování mediánového filtru se používá tzv. plovoucí okno, kterým se prochází celý snímek. Uvnitř tohoto okna se vybírá hodnota mediánu, což je střední hodnota souboru vzestupně seřazených hodnot, jakých nabývá snímek uvnitř zmíněného okna. Pro počítaný pixel se takto zjištěná hodnota mediánu nastaví na hodnotu mediánu.

Jak je uvedeno v knize [28], velmi důležitým parametrem této filtrace je velikost okna. Nelze dobře rozhodnout, jaká velikost okna má být zvolena, protože záleží na kontextu snímku. Velikost okna by měla být lichá, protože v případě sudého čísla nelze medián určit přímo, jak uvádí kniha [3]. Takové mediány by totiž byly dva a v tomto případě je dobré vypočítat mezi těmito dvěma mediány průměrnou hodnotu. Na obrázku 3.4 je praktická ukázka použití mediánového filtru. Obrázek byl uměle zašuměn šumem typu sůl a pepř a poté byl na zašuměný obrázek aplikován mediánový filtr o velikosti okna 5.



Obrázek 3.4: Aplikace mediánového filtru o velikosti 5 na zašuměný obrázek

Mediánový filtr má velkou výpočetní složitost, a to  $N^2n^2$ . I když zachovává kvalitu hran, stejně ke zkreslení dochází. V knize [28] je uvedeno, že 2D filtrace může být separabilní. To znamená, že lze nezávisle na sobě provádět 1D mediánovou filtraci zvlášť pro řádky a zvlášť pro sloupce zpracovávaného snímku.

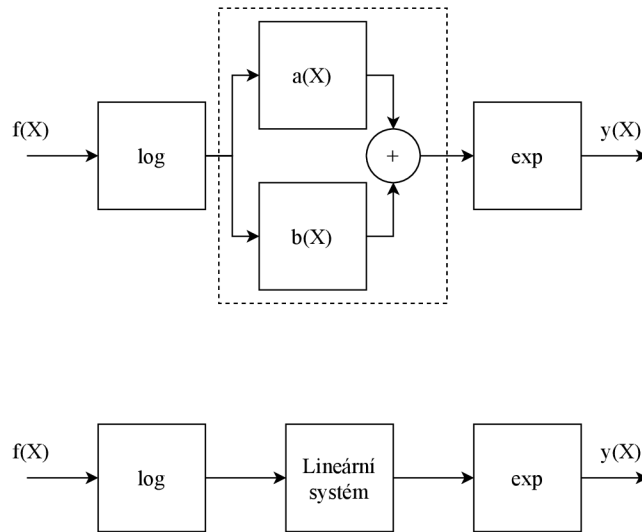
**Homomorfické zpracování** je druh zpracování obrazu, kde podle knihy [28] lze pro nějaký systém, který je složen z multiplikatивních složek, aplikováním logaritmu takové složky od sebe oddělit. Mějme hypotetický systém, složený ze dvou složek:

$$f(X) = A(X) \cdot B(X), \quad (3.12)$$

kde  $A()$  a  $B()$  jsou dvě libovolné funkce. Pokud se rovnice 3.12 zlogaritmuje, tak vznikne

$$\log(f(X)) = \log(a(X)) + \log(b(X)). \quad (3.13)$$

Díky zlogaritmování se tak od sebe tyto složky oddělí a lze je zpracovat nezávisle na sobě. V knize [20] je uvedeno, že homomorfické systémy jsou podtřídou nelineárních systémů a platí u nich zobecněný princip superpozice. Tento zdroj dále uvádí, že lze systém zjednodušit na tzv. kanonickou realizaci systému. To je naznačeno na obrázku 3.5, kde v horní části obrázku je systém, daný rovnicí 3.13 a dole je jeho kanonická podoba. Čárkovaně je vyznačen podsystém, nazývaný lineární jádro.



Obrázek 3.5: Princip homomorické filtrace

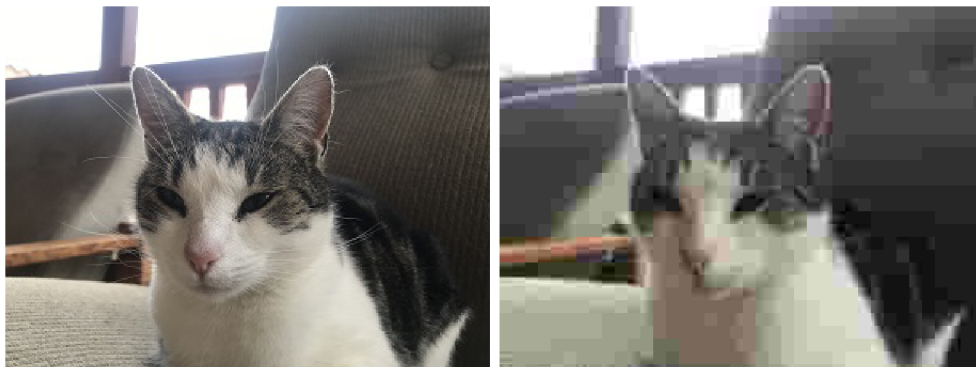
Kanonická podoba je dána jako kaskádní spojení tří podsystémů, kde uprostřed je lineární jádro, ohraničené obvykle nelineárními systémy. V knize [20] je uvedeno, že praktické použití homomorické filtrace je nejčastěji požadavek na rozdělení signálu na složky, které lze nezávisle na sobě měnit.

Dále kniha [20] uvádí, že ne nutně vždy se musí jednat o funkci logaritmu. Pro neaditivní systémy se pro ohraničující systémy hledají funkce  $\phi$  a  $\phi^{-1}$  takové, aby pro použítou operaci existovala taková transformace. Jak bylo uvedeno výše, pro multiplikativní systém je taková transformace funkce  $X = \ln(X)$  a její inverzní funkce  $Y = \exp(Y)$ .

Obdobně, pokud je pro systém použita operace konvoluce, lze jako funkci  $\phi$  použít Z-transformaci, kde bude výsledkem násobení. A to lze transformovat na sčítání pomocí logaritmu. Inverzní funkcí pro Z-transformaci je zpětná Z-transformace. Tento systém se nazývá kepsrum. Nakonec kniha [20] uvádí, že lze místo Z-transformace použít i diskrétní Fourierovu transformaci, provedou-li se kompenzační úpravy diskretizace.

## 3.2 Odstranění artefaktů v obraze

Formát JPEG byl definován v kapitole 2. V knize [39] je uvedeno, že při kódování a kompresi snímku se provádí práce nad bloky o velikosti  $8 \times 8$ , nad kterými se provádí diskrétní kosinová transformace. Hodnoty se dále kvantizují, díky čemuž dochází ke ztrátě informací.



Obrázek 3.6: Snímek a zkomprimovaný snímek

Na obrázku 3.6 je ukázka snímku, na který byla aplikována JPEG komprese se snížením kvality o 95 %. Díky tomu vznikají tzv. blokové artefakty. Podle článku [35] existují algoritmy, které nejen blokové artefakty, ale obecně i jiné artefakty minimalizují.

Prvním popsaným algoritmem v článku [9] je algoritmus **nulování průměrů**, též anglicky *zeroing-out the means* či *zero-meaning*. Zde se nejprve zjistí hodnota všech průměrných hodnot na jednotlivých řádcích:

$$M = \frac{1}{N} \sum_{i=1}^n I_n, \quad (3.14)$$

kde  $n$  je šířka obrázku. Následně se od všech hodnot jednotlivých pixelů v řádku odečte vypočítaná průměrná hodnota:

$$I_n = I_n - M, \quad (3.15)$$

kde  $I_n$  je pixel na upravovaném řádku. Tato operace se provede pro všechny řádky, ale i pro všechny sloupce analogicky.

Výše uvedené rovnice platí pro jeden konkrétní barevný kanál. V případě, že je snímek vícekanalový, aplikuje se tento postup pro jednotlivé barevné kanály zvlášť.

V případě, že se vypočítá rozdíl nezpracovaného snímku a snímku po aplikaci nulování průměru, je výsledný snímek poté nazýván jako lineární vzor:

$$LP = I - ZM(I), \quad (3.16)$$

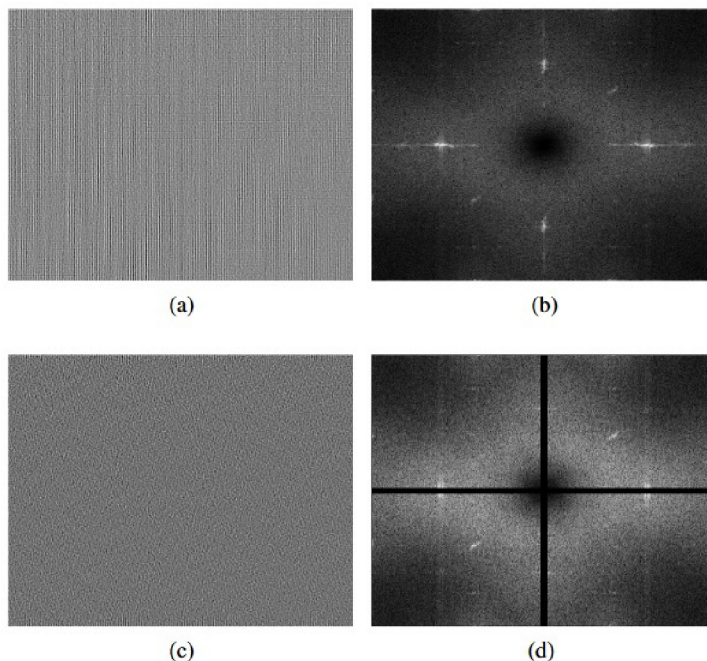
kde  $ZM()$  je nulování průměrů a  $I$  je zpracováváný snímek.

Dalším uvedeným algoritmem je algoritmus pro **odstranění cyklických vrcholů**. Jak uvádí článek [35], jedná se o aplikování speciálního filtru ve Fourierově doméně.

Nejprve je potřeba převést snímek v prostoru do spektra, a to pomocí diskretní Fourierovy transformace, definované v knize [20] takto:

$$F_k = \sum_{n=0}^{N-1} f_n \cdot e^{-j\frac{2\pi}{N}kn}, \quad (3.17)$$

kde  $f_n$  je řada  $N$  komplexních čísel a  $e^{-j\frac{2\pi}{N}kn}$  je Eulerova rovnice. Výsledné spektrum je složeno ze dvou složek: amplitudového a fázového spektra. Vynulováním hodnot v pásmech ve střezech amplitudového spektra, uvedeného na obrázku 3.7 pod písmenem d), se potlačí periodické artefakty, jak je naznačeno na obrázku 3.7, písmena a) a c). Pod písmeny b) a d) je zobrazeno amplitudové spektrum snímku.



Obrázek 3.7: Filtrace ve spektru, převzato z článku [35]

Nakonec se podle knihy [20] provede inverzní diskretní Fourierova transformace, aby se upravený snímek převedl ze spektra zpět do prostoru:

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} F_k \cdot e^{j\frac{2\pi}{N}kn}. \quad (3.18)$$

### 3.3 Korelační koeficienty, normální rozdělení a testování hypotéz

Korelace je podle knihy [2] míra vzájemného vztahu mezi dvěma veličinami. Pro dvě náhodné veličiny  $X$  a  $Y$  musí platit, že jejich druhé mocniny středních hodnot, tedy  $E(X^2)$  a  $E(Y^2)$ , jsou konečné. Potom lze definovat korelační koeficient takto:

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}X \cdot \text{var}Y}}, \quad (3.19)$$

kde  $\text{cov}(X, Y)$  je kovariance  $X$  a  $Y$  a  $\text{var}X$  je rozptyl hodnoty  $X$ , též značen jako  $\sigma_X$ .

Kovariance je definována jako:

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))), \quad (3.20)$$

kde  $E(X)$  je střední hodnota veličiny  $X$ . Rozptyl hodnot je definován jako:

$$\sigma_X = \sqrt{E(X^2) - (EX)^2}. \quad (3.21)$$

Potom lze zápis pro výpočet korelačního koeficientu zapsat jako:

$$\rho_{X,Y} = \frac{E((X - E(X))(Y - E(Y)))}{\sqrt{E(X^2) - (EX)^2} \sqrt{E(Y^2) - (EY)^2}}. \quad (3.22)$$

Vzorec pro výpočet korelačního koeficientu se též podle knihy [40] nazývá jako vzorec pro výpočet Pearsonova korelačního koeficientu. Ten je normalizován do intervalu  $\langle -1, 1 \rangle$ . Pro korelační koeficient  $\rho$  tedy platí:

$$-1 \leq \rho \leq 1. \quad (3.23)$$

Není-li k dispozici sdružená distribuční funkce náhodného vektoru  $(X, Y)'$  a existuje-li náhodný výběr z tohoto rozložení

$$(X_1, Y_1)', (X_2, Y_2)', \dots, (X_n, Y_n)', \quad (3.24)$$

lze z tohoto výběru vypočítat výběrové průměry

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (3.25)$$

Výběrová kovariance je definována jako

$$M_{X,Y} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}). \quad (3.26)$$

Výběrový korelační koeficient  $r_{X,Y}$  je poté definován jako podíl výběrové kovariance a součinem kladných výběrových směrodatných odchylek. Počítá se jako

$$r_{X,Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}. \quad (3.27)$$

Podle knihy [40] je výběrový korelační koeficient též nazýván *sample Pearson correlation coefficient* a používá se například při vyhodnocování veličin, u kterých nejsou jasné parametry rozdělení. I pro výběrový korelační koeficient platí, že hodnoty korelace  $r_{X,Y}$  leží v intervalu  $\langle -1, 1 \rangle$ .

Jsou-li náhodné veličiny  $X$  a  $Y$  vzájemně nezávislé, dosahuje hodnota korelačního koeficientu  $\rho_{X,Y} = 0$ , respektive  $r_{X,Y} = 0$ . Je-li  $\rho_{X,Y} > 0$ , respektive  $r_{X,Y} > 0$ , tak vyšší hodnoty veličiny  $X$  více souvisí s vyššími hodnotami veličiny  $Y$ . Je-li  $\rho_{X,Y} < 0$ , respektive  $r_{X,Y} < 0$ , tak nižší hodnoty veličiny  $X$  souvisí s nižšími hodnotami veličiny  $Y$ .

**Normální rozdělení** (též nazýváno Gaussovo rozdělení) je podle knihy [2] definováno funkcí hustoty rozdělení pravděpodobnosti takto:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (3.28)$$

kde  $x \in (-\infty, \infty)$ , střední hodnota  $\mu \in (-\infty, \infty)$  a směrodatná odchylka  $\sigma > 0$ . Střední hodnota a směrodatná odchylka jsou její parametry, též nazývané charakteristiky. Normální rozdělení se značí  $N(\mu, \sigma^2)$ .

Střední hodnota normálního rozdělení pravděpodobnosti je značena  $EX = \mu$  a rozptyl je často označován jako  $\text{var } X = \sigma^2$ .

Pro statistické vyhodnocení hypotéz je podle [2] nutné nejprve definovat nulovou hypotézu  $H_0$ . Hypotéza je tvrzení, že  $\Theta \in \omega$ , kde  $\Theta$  je parametr rozdělení náhodného vektoru

$X$ . Dále  $\Theta \in \Omega$ , kde  $\Omega$  značí parametrický prostor. Pro  $\omega$  platí, že to je neprázdná vlastní podmnožina parametrického prostoru  $\Omega$ . Platnost této hypotézy se provádí pomocí množiny  $W \in R_n$ , nazývané kritický obor.  $R_n$  je  $n$ -rozměrný euklidovský prostor. Pokud leží náhodný vektor  $X$  v kritickém oboru, hypotéza  $\Theta \in \omega$  se zamítne. V opačném případě se hypotéza nezamítne.

V případě, že se hypotéza  $\Theta \in \omega$  zamítne, i když je správná, se tato skutečnost nazývá jako chyba prvního typu. V případě, že se hypotéza  $\Theta \in \omega$  nezamítne, pokud je špatná, nazývá se tato skutečnost jako chyba druhého typu.

Hypotéza  $\Theta \in \omega$  se nazývá nulová hypotéza  $H_0$  a platí, že  $H_0 : \Theta \in \omega$ . Dále je nutné pro statistické vyhodnocení hypotéz mít podle knihy [2] zdefinovanou alternativní hypotézu  $H_1 : \Theta \in \Omega - \omega$ .

Pro statistické vyhodnocení je potřeba najít takový kritický obor, kde pravděpodobnost chyby druhého typu je co nejmenší a zároveň aby pravděpodobnost chyby prvního typu nepřekročila určenou hodnotu  $\alpha$ . Hodnota  $\alpha$  se volí jako

$$\sup_{\Theta \in \omega} P_{\Theta}(W) \leq \alpha, \quad (3.29)$$

kde  $P_{\Theta}(W)$  je pravděpodobnost, že  $X \in W$  za podmínky parametru  $\Theta$ . Přímo supremum se podle knihy [2] nazývá hladina testu.

Silofunkce testu udává pravděpodobnost, že  $H_0$  se zamítne, pokud je hodnota parametru rovna  $\Theta$ . Je definována jako

$$\beta(\Theta) = P_{\Theta}(W), \Theta \in \Omega. \quad (3.30)$$

Operační charakteristika testu (též síla testu) je definována jako  $1 - \beta(\Theta)$ .

V knize [2] je uvedeno, že pro normální rozdělení se volí hodnota  $\alpha$  z intervalu  $(0, 1)$  nejčastěji jako  $\alpha = 0,05$ , případně  $\alpha = 0,01$ . V praxi to znamená, že chyba prvního typu nepřekročí 5 %, případně 1 %. Jinými slovy je jistota správného rozhodnutí 95 %, případně 99 %. Ověřování platnosti hypotézy  $H_0$  se nazývá test hypotézy  $H_0$  vůči  $H_1$ .

V knize [40] je uveden vztah jednotlivých hypotéz a jejich možných chyb a je prezentován tabulkou 3.1. V buňkách jsou uvedeny názvy chyb, pravděpodobnost a anglický ekvivalent názvu.

	$H_0$ je pravdivá	$H_0$ je nepravdivá
$H_0$ je odmítnuta	Chyba 1. typu $\alpha$ False positive	OK $1 - \beta$ True positive
$H_0$ není odmítnuta	OK $1 - \alpha$ True negative	Chyba 2. typu $\beta$ False negative

Tabulka 3.1: Vztah chyb nulových hypotéz

## Kapitola 4

# Existující řešení zabezpečení obrazových senzorů

V této kapitole jsou uvedeny různé postupy, které se zabývají ověřováním pravosti dat obrazových senzorů. Identifikace zdrojové kamery se nazývá SCI problém (Source Camera Identification problem). Jsou zde uvedeny pouze nastudované postupy, z důvodu rozsahu práce, postupů pro odhalování podvržených dat z obrazových senzorů existuje mnoho.

### 4.1 Identifikace zdrojové kamery pomocí vodoznaků

Jeden možný způsob identifikace zdrojové kamery je pomocí vodoznaků. Jak uvádí článek [4], aby bylo možné dobře zabezpečit obrázek pro rozpoznání podvrhu či manipulaci dat, lze akvizovaný snímek upravit tak, že se do něj vloží vodoznak. Jak uvádí zmíněný článek, některá data je potřeba rovnou zabezpečit pro případnou manipulaci. K tomu lze použít vodoznak, což se signál, u kterého nám nevadí, že upraví zdrojový signál.

Vodoznaky jsou, jak uvádí kniha [6], rozděleny do tří kategorií: **robustní**, kde vložený vodoznak zůstane čitelný i po manipulaci s daty, **fragilní**, kde vodoznak přestane být čitelný při libovolné manipulaci s daty a **polofragilní**, což je třída vodoznakových technik, kde je zajištěna robustnost pro některé případy manipulace s daty, ale zároveň fragilnost pro jiné způsoby manipulace s daty.

Pro zajištění určitého stupně zabezpečení obrazových dat, jsou postupy pro detekci a vkládání vodoznaků zabezpečeny nějakým klíčem  $K$ , které mohou být buď **symetrické** (použit privátní klíč), kde je použit stejný klíč  $K$  jak pro detekci, tak i pro vložení klíče, případně **asymetrické** (použit veřejný klíč), kde je jiný klíč použit pro vložení vodoznaku, než pro čtení vodoznaku.

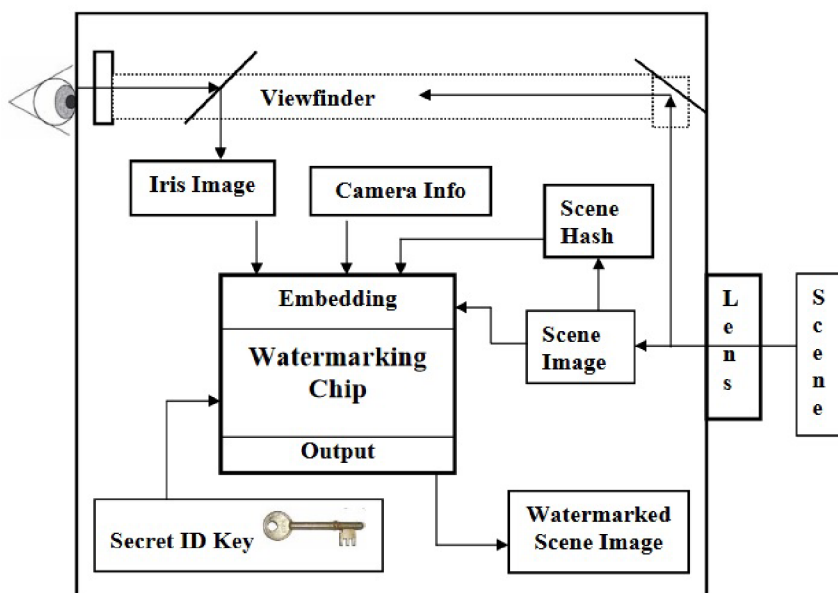
Vodoznaky mohou být viditelné a neviditelné, vkládané buď přímo do obrazu, případně v jiných spektrech (např. ve Fourierově spektru) a mohou obsahovat různý druh dat, ať už jde o binární data či obrazy. Na obrázku 4.1 je ukázka vloženého viditelného vodoznaku přímo do obrazu, a to pomocí obrazu.





Obrázek 4.1: Snímek zabezpečený pomocí vodoznaku

Článek [4] popisuje možný způsob zabezpečení akvizovaných snímků sítnice oka pomocí vodoznaku, a to přímo v digitální kameře. Na obrázku 4.2 je zobrazeno navržené schéma. Obrázek sítnice se zkomprimuje a zkombinuje se s tajným klíčem kamery, který se získá z hashe scény a identifikátoru kamery. Takový klíč se poté vloží do výsledného snímku. Při čtení se naopak extrahuje vložený vodoznak a rekonstruuje se původní scéna, pro kterou se vypočítá hash. Pokud jsou oba hashe ekvivalentní, nedošlo k manipulaci s daty, v opačném případě ano.



Obrázek 4.2: Zabezpečená kamera, převzato z článku [4]

Takový přístup je možný v případech, kdy lze vkládat vodoznak do snímku ideálně přímo v kameře. Postup lze aplikovat pro takové kamery, u kterých je důležité kontrolovat, odkud takový obraz byl pořízen. Tento postup nelze aplikovat pro *běžné* digitální kamery, jako například mobilní telefony či fotoaparáty.

## 4.2 Identifikace zdrojové kamery pomocí hlaviček obrazových souborů

Jednou z možností, jak lze ověřovat pravost dat souborů, pořízených z obrazového senzoru, je čtení hlaviček obrazových souborů. Některé grafické formáty, například JPEG, obsahují hlavičku, kde jsou uvedeny informace o zařízení, ze kterého byla akvizována data.

Článek [22] se například zabývá tím, jak lze provést identifikaci obrazového senzoru pomocí atributů, umístěných v EXIF hlavičce JPEG souborů. Článek je zaměřen pouze na mobilní telefony. Autoři používají některé parametry v EXIF hlavičce jako tzv. *features*, které používají jako vstupy pro trénování klasifikátoru. Parametry, které autoři používají, jsou například ohnisková vzdálenost optického systému, clona objektivu, rozlišení kamery, clonové číslo, expoziční číslo atd.

Poté pomocí k-NN algoritmu (k nejbližších sousedů) trénují model, kde je využita Euklidovská vzdálenost mezi dvěma sousedy. Tento algoritmus lze podle zmíněného článku použít jak pro klasifikaci, tak pro regresi. Další použitý algoritmus je ELA (*Error Level Analysis*), který je použit pro kontrolování případné manipulace s daty.

Přímo i v článku jsou uvedeny nevýhody, jaké takový postup má. Pokud je mobilních telefonů hodně, tak nelze použít pouze k-NN algoritmus a EXIF data. Druhý zmíněný problém je, že nelze od sebe odlišit dvě zařízení, která jsou stejného typu od stejného výrobce.

Článek [31] využívá jiné atributy z EXIF hlavičky, a to *EXIFFD*, *IFD0*, *IFD1*, *GPS* a další atributy, které nejsou standardizované, ale často používané. Další parametry, které se při identifikaci používají, jsou kvantizační tabulky JPEG souborů. Tyto parametry se poté používají jako *features* pro klasifikaci.

Obě zmíněné metody mají velkou nevýhodu v tom, že EXIF hlavičku lze softwarově upravovat. Další nevýhodou je, že mnohdy nelze od sebe oddělit konkrétní modely u zařízení, ale pouze značky. V některých případech nejsou v EXIF hlavičce uloženy ani informace o zařízení, ale jsou tam uloženy informace o softwaru, který danou fotografii akvizoval či upravoval.

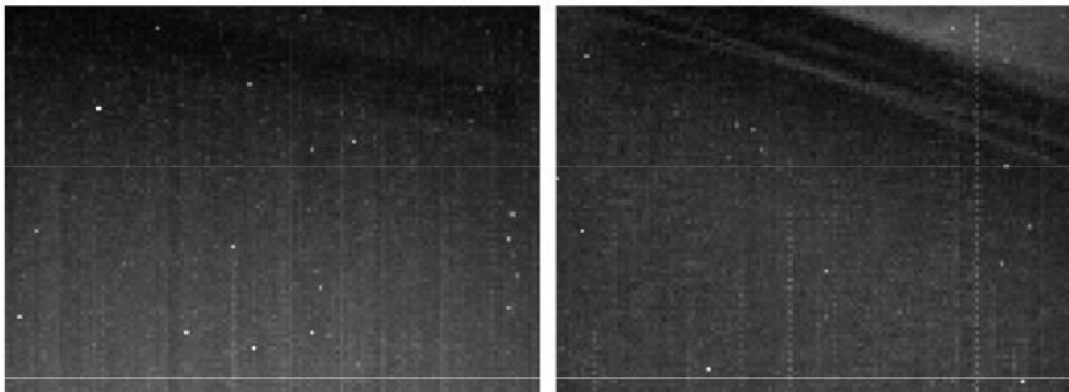
## 4.3 Identifikace zdrojové kamery pomocí vlastností v obraze

Článek [21] řeší SCI problém pomocí příznaků extrahovaných přímo z pořízeného snímku, a to na základě použitého demosaicing algoritmu, který tvoří artefakty, případně na základě různých vlastností v obraze.

Použité příznaky jsou například průměrná hodnota pixelu v jednotlivých barevných kanálech, dále korelace mezi páry jednotlivých barevných kanálů (RB, RG apod.), poměr energie mezi těmito páry, statistika ve vlnkovém spektru apod.

Získané příznaky se poté používají pro trénování klasifikátoru SVM (Support Vector Machines) pro zjištění, jak přesné a jak efektivní jsou jednotlivé *features*. Přesnost tohoto algoritmu ovšem dosahuje pouze 78 % pro nejhorší možný případ a 95 % pro nejlepší možný případ. Pro forenzní případy je potřeba dosahovat vyšších přesností.

Článek [15] využívá pro SCI problém defektní pixely a další vady v obrazovém senzoru. Takové pixely jsou běžně u snímků nežádoucí a znehodnocují kvalitu digitálního snímku. Starší CCD obrazové senzory na takové vady trpěly poměrně hodně.



Obrázek 4.3: Defektní pixely v obrazovém senzoru, převzato z článku [15]

Na obrázku 4.3 jsou takové defektní pixely vidět u dvou různých kamer. U každé kamery jsou takové pixely na různých místech, a proto je lze použít pro identifikaci senzoru.

Jak i přímo článek [15] uvádí, pokud obrazový senzor neobsahuje defektní pixely, tak nelze tento postup použít. Další velkou nevýhodou je, že postupem času vznikají další vady v obrazových senzorech, a proto by bylo potřeba takovou databázi obrazových senzorů pravidelně aktualizovat.

Další důležitou kategorií pro SCI problém je identifikace senzorů pomocí jejich šumových vlastností. Článek [24] uvádí, že lze identifikovat obrazový senzor pomocí odlišností v jednotlivých pixelech. Temné proudy, které se u jednotlivých pixelů vyskytují, jsou pro každý obrazový senzor unikátní. To je dáno výrobními postupy jednotlivých obrazových senzorů.

Temné proudy jsou podle článku [24] součástí *fixed pattern noise*, což je šum, který se vyskytuje v obraze vždy. Hodnoty jsou ovšem natolik *neviditelné* oproti okolním intenzitám pixelů, že je potřeba FPN extrahovat. Extrakci autoři provádějí na speciálních datových sadách, pořízených z různých kamer tak, že snímají monotónní obrazy za nepřístupu světla na obrazový senzor. Výsledné snímky poté sečtou, čímž se potlačí náhodný šum, ale výrazně se zesílí FPN.

Jak uvádí ale článek [29], jedná se o *Dark Current Noise*, popsán v kapitole 2.4 jako *Dark Current nonuniformity* DCNU. Tento šum digitální kamery potlačují odečtením temného snímku, čímž se tato metoda stává selhávající.

Článek [29] popisuje způsob, jak lze jinou složku FPN, nerovnoměrnost citlivosti pixelů (*photoresponse nonuniformity*), PRNU využít pro problém SCI. Nejprve se provede odstranění šumu (denoising) pomocí odšumovacího algoritmu, provedeného ve vlnkovém spektru. Poté se vypočítá reziduum šumu (rozdíl mezi odšuměným a neodšuměným snímkem) a totéž se provede pro alespoň 50 snímků, které se nakonec zprůměrují. Postup je dán následovně:

$$I_{PRNU} = \frac{1}{N} \sum_{i=1}^N (I_i - F(I_i)), \quad (4.1)$$

kde  $N$  představuje počet snímků,  $I$  je snímek a  $F$  je odšumovací filtr.

Takto získaný referenční PRNU se poté používá jako *fingerprint* dané kamery a je možné jej využít pro SCI problém. Při ověření pravosti snímku se poté vypočítá jeho reziduum šumu a mezi referenčním PRNU a tímto reziduem se vypočítá korelační koeficient a určí se práh, který je rozhodující pro testování různých hypotéz.

Zmíněný článek je brán jako stavební kámen pro SCI pomocí PRNU. Na tento článek navazuje spousta dalších prací, která uvedený postup vylepšují. Například článek [9] jako možné zlepšení uvádí odečtení průměrů řádků a sloupců (anglicky *zero-meaning*) a aplikaci Wienerova filtru ve Fourierově doméně. Obě metody snižují *false rejection rate* (FRR).

Článek [18] uvádí pro postup, uvedený v článku [29], kde je pro odstranění šumu ze snímku lepší použít anizotropní difuzi než vlnkovou transformaci s Daubechiesové vlnkami. Metoda je navíc výpočetně rychlejší a přesnější.

Dalším možným vylepšením PRNU je podle článku [16] potlačení silných (výrazných) komponent signálu v PRNU či použití pouze fázové složky ve Fourierově doméně. Jsou uvedeny i další možné přístupy pro extrakci PRNU, jako je spojení reziduí šumu pomocí váženého průměru, vliv různých algoritmů pro odstranění šumu na FRR či extrakce PRNU pouze v místech, kde nejsou příliš nízké, ale ani příliš vysoké frekvence.

Článek [27] se zabývá tím, jaká lokalita v obraze je vhodná pro extrakci PRNU. Autoři zjistili, že není vhodné řešit SCI problém na celém snímku, ale pouze jeho části. Autoři doporučují vynechat okraje snímku, protože se tam mohou objevovat anomálie jako vinětae. Pokud je potřeba k nějakému úkolu pouze část snímku, je lepší volit střed než okraje.

Tím, že se ve většině případech provádí odšumění snímku pro každý snímek zvlášť, roste i časová složitost algoritmu. Možné zrychlení uvádí článek [35], kde lze pro extrakci PRNU využít homomorfní filtrování, kde se pro všechny snímky vypočítá průměr, který se zlogaritmuje a nad logaritmickým obrazem se provede odšumění. Jako metodu pro odstranění šumu používají v článku [35] adaptivní Wienerův filtr a též popisují další možné vylepšení PRNU: potlačení cyklických vrcholů.

Postup této metody je:

$$I_{PRNU} = \frac{\ln\left(\frac{1}{N} \sum_{i=1}^N I_i\right) - F\left(\ln\left(\frac{1}{N} \sum_{i=1}^N I_i\right)\right)}{\gamma}, \quad (4.2)$$

kde  $N$  je počet snímků použitých pro extrakci PRNU,  $I$  je snímek,  $F()$  je použitý odšumovací filtr a  $\gamma$  je faktor gama korekce, definován jako:

$$\gamma = \frac{\log(0.5)}{\log\left(\frac{1}{X \cdot Y} \sum_{x=0}^X \sum_{y=0}^Y \left(\frac{1}{N} \sum_{i=1}^N I_i(x, y)\right)\right)}, \quad (4.3)$$

kde  $X$  a  $Y$  je počet řádků a sloupců v obraze. Uvádějí, že počet snímků pro extrakci PRNU nemusí být 50, ale stačí 20.

Jiný postup pracující nad PRNU popisuje článek [9], kde se používá statistická metoda maximální věrohodnosti (anglicky *maximum likelihood estimation*, MLE). Popisují se zde multiplikativní vlastnosti PRNU a uvádějí, že je vyšší korelace v oblastech s vysokou intenzitou. V oblastech, kde se objevuje textura, je naopak nižší. Případné aplikování filtru dolní propusti snižuje velikost korelace. Též se v uvedeném zdroji popisuje vznik artefaktů v PRNU způsobený při demosaicingu a při JPEG kompresi.

Důležitým aspektem je výběr odšumovacího filtru. Těch je velké množství a liší se svojí kvalitou, rychlostí a případy použití, jak uvádí článek [14].

V prostorové doméně existují lineární a nelineární filtry. Obecně lineární filtry dokážou odstranit šum, ale selhávají při zachování textur a detailů. Příkladem lineárního filtru je průměrování, ale pokud je snímek příliš znehodnocen šumem, tak jsou výsledné snímky příliš hladké. Lépe si s tímto problémem bojuje Wienerův filtr, ale neumí příliš zachovávat hrany.

Nelineární filtry, jako například mediánový filtr či bilaterální filtr, dokážou odstranit šum a zároveň zachovat kvalitu hran. Nevýhodou takových filtrů jsou ale vysoké výpočetní nároky.

Variační metody využívají vlastností snímků, kde je cílem zmenšovat energii snímku. Snížením energie signálu vznikne odšuměný snímek. Využívají se k tomu poznatky z oboru pravděpodobnosti, převážně z Bayesovské pravděpodobnosti. Důležitou vlastností snímků, které se při odšumování používají, jsou gradienty.

Jedním ze zástupců této metody je Tikhonova regulace, která odstraňuje šum pomocí L2 normy (nejkratší vzdálenost mezi dvěma body). Metoda příliš moc vyhlazuje detaily. Dalším zástupcem je anizotropní difuze, která sice zachovává detaily, ale mírně rozmazává hrany. Obě metody jsou lokální a dokážou odstranit šum spolehlivě v ne příliš zašuměných snímcích.

Dále jsou v této kategorii nelokální regulace. Zástupcem této kategorie je tzv. *non-local means*, nelokální průměry, která používá váhované filtrování nelokálních podobností sama sebe. Myšlenka non-local means je založena na průměrování všech pixelů ve snímku. Hodnota váhy pro pixel je určena jako velikost podobnosti oblasti okolo tohoto pixelu a jiné, podobné oblasti okolo pixelu, kde se provádí odšumění.

Další kategorií, uvedenou v [14], jsou transformační techniky. Zde se snímky transformují do jiné domény, kde lze provádět filtraci. Příkladem je například převod pomocí Fourierovy transformace do spektra, ale více se v praxi využívá odšumění pomocí vlnkových transformací a kosinových transformací. Nevýhodou takových metod je časová náročnost výpočtu a vhodný výběr filtrace ve spektrech. Například pro vlnkové transformace je klíčové vybrat správnou vlnku pro filtraci.

Další zajímavou transformací je BM3D, využívající non-local means principy pro odšumění. Podobné oblasti se převedou do 3D skupin, které se poté ve vlnkovém spektru filtrují pomocí vlnek nebo Wienerova filtru. Poté se inverzní funkcí získají zpět skupiny a provede se rekonstrukce snímku. Platí, že čím více je obraz zašuměn, tím menší jsou regiony a tím horší je úspěšnost odšumění.

Poslední zmíněnou kategorií v článku [14] je odšumění pomocí konvolučních neuronových sítí a strojového učení. Těch existuje několik a liší se svojí kvalitou, výpočetní rychlostí a možnostmi použití.

## Kapitola 5

# Analýza a návrh řešení

V této krátké kapitole je zanalyzován možný postup práce a je uveden jeho konkrétní návrh, který je implementován a vyhodnocen, což je popsáno v dalších kapitolách této práce.

### 5.1 Zhodnocení současného stavu

Jak bylo uvedeno v předchozí kapitole, lze se na problém ověřování pravosti dat podívat mnoha pohledy. Zabezpečení pomocí vodoznaků a tvorbou privátních a veřejných klíčů při podepisování je velmi specifické. Pro tento případ je lepší takový postup provádět pro specializované kamery, které je potřeba zabezpečit. Ověření pravosti dat tak například selže při použití zařízení, která takovým způsobem podepsána nejsou a data by se musela posléze podepsat ad hoc. Díky tomu, že tento postup je dobré dělat pro specifické kamery a ne pro libovolné kamery, není vhodné takový postup aplikovat.

Analýzou hlaviček souborů lze též ověřovat pravost snímku, ovšem tím, že hlavičky jdou upravovat, není tato metoda příliš spolehlivá. Navíc lze snímek konvertovat do jiného formátu, kde se hlavička takového souboru též změní a snímek je tak neidentifikovatelný. Z tohoto důvodu, že spolehlivost takové metody je na první pohled poměrně malá, není vhodné takový postup aplikovat.

Dobrým postupem se tak jeví zpracovávat přímo obrazové vlastnosti snímku. První zmíněný postup využívá extrahované *features* ze snímku, jako například průměrnou hodnotu pixelu, korelaci mezi páry jednotlivých barevných kanálů apod. Tento postup je oproti SCI (Source Camera Identification, identifikace zdrojové kamery) pomocí EXIF hlaviček robustnější, ovšem uvádějí, že přesnost takového postupu je 78 % pro nejhorší možný případ, pro nejlepší možný případ to je 95 %. Aby byla přesnost co možná nejvyšší, je potřeba vyzkoušet jiné postupy.

Další uvedené postupy využívají šumové vlastnosti snímku. Tím, že je šum přítomen při akvizici obrazu vždy, tak i při změně formátu se tento šum obvykle zachová (neprovede-li se zároveň komprese). Protože DCNU (dark current nonuniformity) je snaha potlačovat přímo v kameře, je uvedená metoda pro práci nad DCNU nevhodná.

Další uvedený postup je založen na detekci defektních pixelů. Ty jsou bohužel nestálé v čase, jejich počet postupem času přibývá. Díky tomu, že by bylo potřeba často ověřovat, jak se defektní pixely změnily za jednotku času, je tato metoda nevhodná a málo spolehlivá.

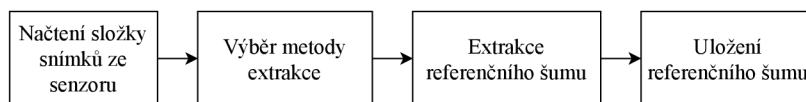
V kapitole 4 je zmíněno několik postupů, jak lze ověřovat pravost dat pomocí PRNU (photoresponse nonuniformity). Přístup, kdy se vypočítá reziduum šumu, se zdá jako velmi

spolehlivá metoda. Prezentované výsledky v uvedené literatuře jsou uspokojivé, a proto jsem se rozhodl vyzkoušet metody na tomto přístupu založené, a to včetně různých vylepšení.

## 5.2 Návrh řešení

Rozhodl jsem se identifikovat senzor pomocí PRNU a vyzkoušet dva přístupy pro jeho extrakci. První vyzkoušený je přístup sekvenční, použitý například v článku [29] a druhý je homomorfický, použitý v článku [35]. Základní schéma pro extrakci *referenčního vzoru* je uvedeno na obrázku 5.1.

Důvod pro vybrání právě těchto metod je zkusit ověřit, jestli opravdu lze poměrně snadnou metodou zjistit, odkud mohla vzniknout původní fotografie a rád bych experimentálně ověřil, jak přesně toho lze docílit. Cílem není ani tak vytvořit aplikaci, která bude s vysokou přesností správně ověřovat původ dat za každou cenu, ale vytvořit prostředí, ze kterého bude uživatel schopen získat hodnoty pro své vlastní vyhodnocení a odpovědění si na různé otázky.

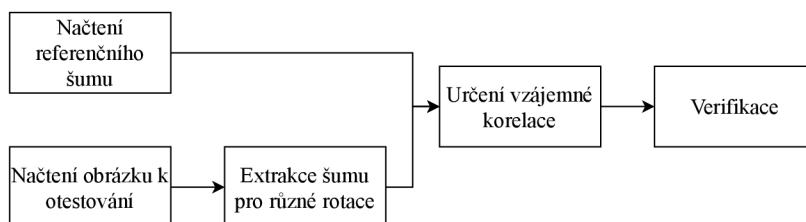


Obrázek 5.1: Obecné schéma extrakce referenčního šumu

Extrakce PRNU jako taková není schopna říci, o jaký senzor se jedná, proto je následně důležité vzít libovolnou fotografii z nějakého senzoru, nad ní se pokusit provést extrakci PRNU a položit si otázku: „Je daná fotografie z této digitální kamery?“ Na tuto otázku lze odpovědět výpočtem podobnosti extrahovaného šumu a referenčního šumu.

Většina zmíněných článků pracuje s korelačními koeficienty. Na výpočet korelace používají nejčastěji normalizovanou kros-korelaci, případně výpočet Pearsonova korelačního koeficientu. Ta rozloží hodnoty do intervalu  $\langle -1; 1 \rangle$ , kde  $-1$  je záporná (opačná) korelace,  $0$  je nulová korelace a  $1$  je maximální možná míra korelace. Poté následuje vybrání vhodné prahu, který rozhoduje o tom, zda-li je extrahovaný šum pořízen ze zájmového senzoru nebo není.

V diplomové práci jsem se rozhodl použít výpočet pro Pearsonův korelační koeficient, protože normalizuje výsledné hodnoty. U Pearsonova korelačního koeficientu lze v tomto případě předpokládat, že čím vyšší je hodnota vzájemné korelace, tím je pravděpodobnější, že šum pochází z testovaného senzoru. Schéma pro určení identifikace zdrojové kamery je naznačeno na obrázku 5.2.



Obrázek 5.2: Obecné schéma určení identifikace zdrojové kamery

Protože různé nastudované články říkají, že je důležité použít kvalitní odšumovací filtr, rozhodl jsem se vyzkoušet několik druhů filtrace a rozhodnout, který z použitých filtrů je nejlepší pro identifikaci senzoru použít, a to:

- anizotropní difuzi,
- adaptivní Wienerův filtr,
- non-local means,
- filtrace dolní propustí (box filter),
- homomorrická filtrace pomocí adaptivního Wienerova filtru a
- homomorrická filtrace pomocí anizotropní difuze.

Anizotropní difuze byla vybrána, protože údajně je pro problém SCI vhodná a filtrace jako taková splňuje požadavky na odšumovací filtry kladené. Adaptivní Wienerův filtr byl vybrán proto, protože byl použit v článku, představující homomorrickou extrakci PRNU. Non-local means byl vybrán na základě popisovaných dobrých vlastností v literatuře. Box filter byl vybrán z důvodu otestování, jestli stačí prosté rozmazání snímku pro řešení SCI problému.

Dále různé články uvádějí, že lze vylepšit extrahovaný PRNU pomocí různých operací, rozhodl jsem se některá vylepšení přidat do práce s PRNU, a to:

- odečtení průměrů řádků a sloupců,
- odstranění cyklických vrcholů.

Poslední zmíněné vylepšení nelze provádět v prostorové doméně a provádí se ve Fourirérově doméně.

### 5.3 Specifikace řešení

Aplikaci jsem se rozhodl implementovat v jazyce C++, konkrétně s použitím knihoven Qt, kde je možno tvořit uživatelské rozhraní. Výsledná aplikace tak bude snáze ovladatelná, než kdyby se jednalo o konzolovou aplikaci.

Požadavky na aplikaci:

- Schopnost provést extrakci referenčního PRNU, a to za použití uživatelem zadané možné filtrace.
- Umět takovou extrakci provést nad adresářem, obsahující alespoň 20 JPEG nebo BMP souborů.
- Umět omezit velikost výsledného referenčního PRNU na  $1024 \times 1024$ .
- Umět extrahovaný referenční šum uložit jako obrázek ve formátu BMP.
- Umět extrahovaný šum zobrazit uživateli.
- Schopnost načíst extrahovaný referenční šum, případně množinu referenčních šumů.



- Schopnost automaticky nastavit extrahovaný referenční PRNU pro možné ověřování pravosti dat.
- Schopnost načíst snímek, případně adresář, obsahující JPEG a BMP soubory.
- Vypočítat vzájemnou korelaci mezi referenčními PRNU a testovanými snímky. Vypočítané korelační koeficienty poté na obrazovku vypsat ve formátu CSV, vhodným pro strojové zpracování a vyhodnocení.
- Umět vypočítat korelační koeficient pro rotace o  $90^\circ$  a přidat jej do CSV záznamu, který je vhodný k vyhodnocení.
- Umět vypočítat korelační koeficient pro všechna implementovaná odšumění a přidat jej do CSV záznamu, který je vhodný k vyhodnocení.
- Umět zobrazit referenční PRNU a veškerá extrahovaná rezidua šumu všech rotací pro všechny metody.
- Umět provést univerzální odhalení podvrhu na základě vyhodnocení, a to pro anizotropní difuzi, pro dostatečně velkou hodnotu prahu, aby byla chyba prvního typu  $\leq 1$ .

Pro vyhodnocení jsem se rozhodl zjistit, jaký z vyzkoušených filtrů produkuje nejpřesnější výsledky. Dále, jaký je vliv velikosti extrahovaného PRNU, jestli stačí malý vzorek dat nebo celý snímek pro extrakci PRNU. Dále jsem se rozhodl ověřit vliv sady snímků, určené pro extrahování referenčního PRNU, a to pro náhodné snímky a ideální snímky. Též jsem se rozhodl ověřit, jestli je lepší provádět extrakci referenčního šumu ze středu snímku nebo o levého horního okraje. Na závěr jsem se rozhodl vyhodnotit několik reálných použití pro odhalení pravosti dat.

## Kapitola 6

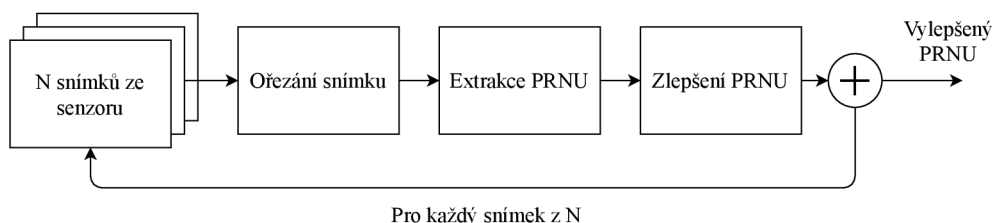
# Extrakce vzorového šumu a identifikace senzoru

V této kapitole je popsán postup, jakým způsobem byl extrahován referenční PRNU a jak probíhá ověření zdrojové kamery.

### 6.1 Schéma algoritmu a datové sady

Aby bylo možné identifikovat senzor, je potřeba získat nejprve jeho nějaký digitální otisk, který bude určovat daný obrazový senzor. Je důležité, aby byl otisk robustní, jednoznačný a co nejvíce vypovídající. Jak bylo nastíněno v kapitolách 4 a 5, nejlepším možným digitálním otiskem se jeví použít PRNU.

Pro extrakci referenčního PRNU je implementován algoritmus, který je velmi obecně uveden blokově na obrázku 6.5.



Obrázek 6.1: Obecné schéma extrakce referenčního šumu

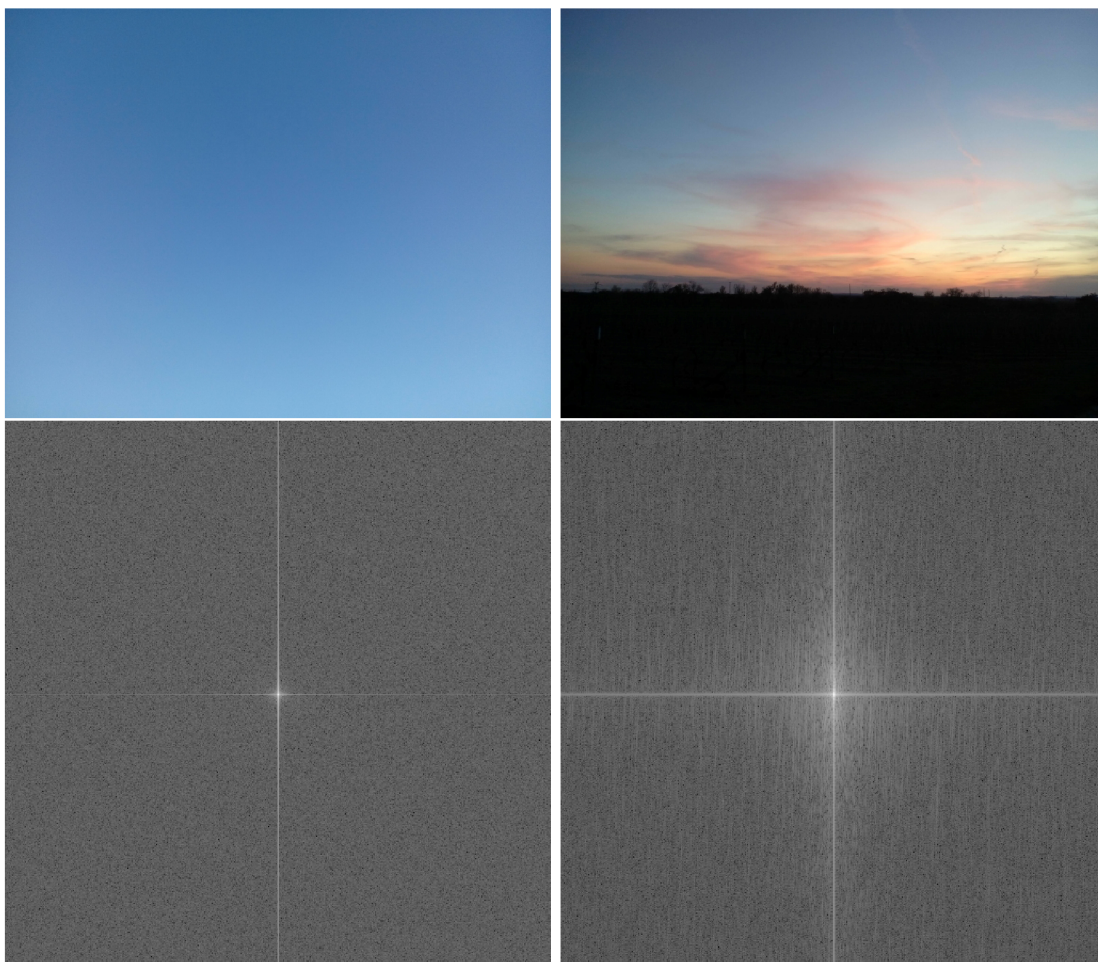
Jednotlivé kroky z tohoto schématu jsou postupně objasněny v následujících sekcích této kapitoly.

Aby bylo možné referenční PRNU získat, je potřeba nejprve vybrat vhodnou sadu snímků. V literatuře (např. v článku [18]) je doporučeno, aby bylo použito alespoň 30 snímků pro extrakci referenčního šumu. Dále se uvádí, že je důležité, aby snímky vhodné pro extrakci *ideálního* referenčního šumu obsahovaly co nejméně detailů, vysokých frekvencí, aby neobsahovaly přeexponované části a nebyly příliš tmavé.

Jako nejlepší možnou množinou snímků pro extrakci PRNU se jeví pořídít rovnoměrně osvětlené snímky (tzv. *flat-field* snímky). Dobrým příkladem je například osvětlená stěna či podlaha, která však neobsahuje žádné *nečistoty*, jako například skvrny, drobné kamínky, pavučiny, nebo stíny. Těž je důležité, aby snímky neobsahovaly žádnou pravidelnou texturu, jako například vzorky dřeva, betonu apod.

Při snímání takových ploch je lepší, aby digitální kamera nebyla umístěna na stativu a snímala kontinuálně, ale naopak je dobré měnit výchozí polohu digitální kamery, aby se minimalizoval případný výskyt nechtěných struktur povrchu. Dále je možné pořídit například snímky jasné a modré oblohy, ideálně bez mraků, která je v mnoha použitých metodách z nastudované literatury využita pro extrakci referenčního PRNU. Z takových snímků by se měl získat nejvěrohodnější referenční šum.

Protože se v praktickém použití může stát, že nebude k dispozici sada rovnoměrně osvětlených snímků, je dobré vytvořit i sadu *běžných* snímků. U takových snímků by nemělo být kladeno příliš mnoho požadavků na scénu. Pokud snímek obsahuje mnoho vysokých frekvencí (ostré hrany, které jsou příliš kontrastní s pozadím, například koruny opadavých stromů v zasněžené krajině), promítnou se takové frekvence při extrakci PRNU jako viditelné artefakty. Takové artefakty potom snižují kvalitu extrahovaného šumu. Ukázky z těchto dvou sad jsou znázorněny na obrázku 6.2, kde nahoře vlevo je snímek oblohy, vpravo nahoře je náhodný snímek, který obsahuje výrazné kontrastní přechody a pod nimi jsou odpovídající moduly v logaritmicím měřítku v jejich spektru.



Obrázek 6.2: Snímky z datových sad obsahující oblohu a náhodný pohled do krajiny a jejich moduly v logaritmicím měřítku

U všech datových sad, které jsou určeny pro extrakci referenčního PRNU je důležitá orientace digitální kamery. Je nutné, aby všechny obrazy byly pořízeny stejně. Pokud by

byly snímky různě orientovány či ořezány, odpovídající PRNU z daného senzoru by bylo na jiných pozicích v snímku a používaly by se jiné hodnoty PRNU, které by tam zanášely chybu. Tento problém obvykle nastává u mobilních telefonů s kamerou, protože většinou umožňují automatickou rotaci snímku dle natočení snímacího zařízení. U digitálních fotoaparátů, obzvláště u zrcadlovek, k takové umělé rotaci snímku obvykle nedochází.

Sady snímků, které se skládají z rovnoměrně osvětlené plochy a běžných (též náhodných) snímků, jsem pořídil či získal z jiných zdrojů, z celkem patnácti různých obrazových senzorů. Jedná se o deset mobilních telefonů, dva digitální fotoaparáty a jednu digitální videokameru. Seznam použitých digitálních kamer je uveden v tabulce 6.1. U mobilních telefonů byla vždy použita zadní kamera.

Zdrojem informací o obrazových senzorech v mobilních telefonech je web [13], a to vždy odpovídající stránka pro konkrétní zařízení. Dále pro videokameru Panasonic HC-VC980 je zdrojem informací stránka [33], pro digitální fotoaparát Fujifilm SL-280 je zdrojem informací stránka [42] a pro Panasonic DMC-FZ5 je zdrojem specifikace obrazového senzoru stránka [25].

ID	Název	Senzor	Druh	Formát	Rozlišení
1	Redmi Note 4	Samsung S5K3L8	CMOS	1/3.06	13 MP
2	K3 Note	Omnivision OV13850	CMOS	1/3.06	12.98 MP
3	Fujifilm SL-280	—	CCD	1/2.3	14 MP
4	Panasonic DMC-FZ5	—	CCD	1/2.5	5 MP
5, 6	iPhone 7	Sony IMX315 Exmor RS	CMOS	1/2.93	12.2 MP
7	iPhone 7 Plus	Sony IMX315 Exmor RS (hlavní kamera)	CMOS	1/2.93	12.2 MP
8 - 10	Panasonic HC-VX980	—	CMOS	1/2.3	18.91 MP
11 - 15	Samsung A20	Sony IMX258 Exmor RS	CMOS	1/3.06	13.25 MP

Tabulka 6.1: Použité digitální kamery a jejich obrazové senzory

Senzory jsou rozděleny do osmi tříd, a to podle druhu použitých zařízení. Každá taková třída je v tabulce 6.2 uvedena na samostatném řádku. Například třída mobilních telefonů Samsung A20 obsahuje celkem 5 kusů zařízení. Je klíčové nepromíchat pořízené snímky mezi senzory v rámci třídy mezi sebou, ale i mezi třídami obecně. Špatně zařazené snímky se totiž těžko hledají a zanášejí chybu do výpočtů. Třídy budou v textu nazývány pouze názvem značky zařízení, v případě třídy Panasonic HC-VX980 bude použit název Panasonic video.

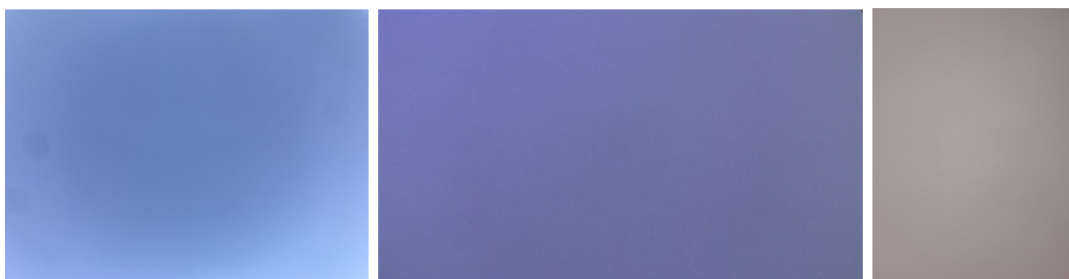
Konkrétní exponované datové sady, včetně použitého počtu snímků, jsou uvedeny v tabulce 6.2. Sad bylo pro každý senzor vytvořeno několik.

ID	Kamera	Název sady a druh snímků	Počet snímků	Rozlišení
1	1	Redmi, jasná obloha	71	4160 × 2340
2	1	Redmi, náhodné snímky	41	4160 × 2340
3	2	Lenovo, oblaka	64	3264 × 2448
4	2	Lenovo, jasná obloha	61	3264 × 2448
5	2	Lenovo, náhodné snímky	56	3264 × 2448
6	3	Fujifilm, oblaka	66	4288 × 3216
7	3	Fujifilm, jasná obloha	34	4288 × 3216
8	3	Fujifilm, náhodné snímky 1	50	4288 × 3216
9	3	Fujifilm, náhodné snímky 2	29	4288 × 3216
10	4	Panasonic, náhodné snímky	44	2560 × 1920
11	4	Panasonic, jasná obloha	78	2560 × 1920
12	5	iPhone 7 #1, jasná obloha	73	4032 × 3024
13	5	iPhone 7 #1, náhodné snímky	50	4032 × 3024
14	6	iPhone 7 #2, stěna	54	4032 × 3024
15	6	iPhone 7 #2, náhodné snímky	55	4032 × 3024
16	7	iPhone 7+, náhodné snímky	146	4032 × 3024
17	7	iPhone 7+, jasná obloha	81	4032 × 3024
18	8	Panasonic video #1, papír	58	4992 × 2808
19	8	Panasonic video #1, náhodné snímky 1	56	4992 × 2808
20	8	Panasonic video #1, náhodné snímky 2	82	4992 × 2808
21	8	Panasonic video #1, stěna	69	4992 × 2808
22	9	Panasonic video #2, papír	50	4992 × 2808
23	9	Panasonic video #2, náhodné snímky 1	47	4992 × 2808
24	9	Panasonic video #2, náhodné snímky 2	95	4992 × 2808
25	9	Panasonic video #2, stěna	98	4992 × 2808
26	10	Panasonic video #3, papír	56	4992 × 2808
27	10	Panasonic video #3, náhodné snímky 1	38	4992 × 2808
28	10	Panasonic video #3, náhodné snímky 2	57	4992 × 2808
29	10	Panasonic video #3, stěna	44	4992 × 2808
30	11	Samsung #1, deska stolu	78	3456 × 4608
31	11	Samsung #1, náhodné snímky	43	3456 × 4608
32	11	Samsung #1, stěna	83	3456 × 4608
33	12	Samsung #2, deska stolu	34	3456 × 4608
34	12	Samsung #2, náhodné snímky	61	3456 × 4608
35	12	Samsung #2, stěna	86	3456 × 4608
36	13	Samsung #3, deska stolu	53	3456 × 4608
37	13	Samsung #3, náhodné snímky	38	3456 × 4608
38	13	Samsung #3, stěna	77	3456 × 4608
39	14	Samsung #4, náhodné snímky	43	3456 × 4608
40	14	Samsung #4, stěna	69	3456 × 4608
41	15	Samsung #5, náhodné snímky	56	3456 × 4608
42	15	Samsung #5, stěna	87	3456 × 4608

Tabulka 6.2: Vytvořené datové sady pro extrakci a identifikaci obrazového senzoru digitální kamery

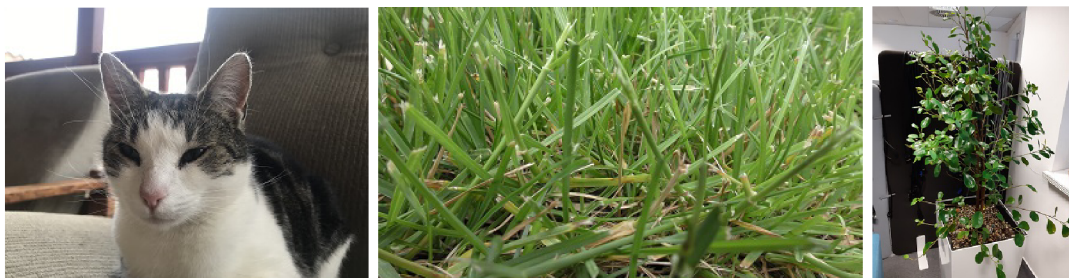
Co se týče rovnoměrně osvětlených snímků, jež pocházejí ze zdrojů uvedených v tabulce 6.2, tak u některých kamer byla exponována jasná obloha s minimálním výskytem mraků. U jiných kamer nešlo z důvodu dlouhotrvajícího špatného počasí při vzniku datových sad zachytit jasnou oblohu, a proto byla zaznamenána například stěna. Zde byla snaha zachytit co nejhladší stěnu, bez viditelných odlesků a nečistot. U jedné třídy kamer byla zachycena čtvrtka čistého bílého papíru, na který dopadalo umělé světlo tak, aby světlo bylo rovnoměrné po celé ploše papíru.

Za připomínku stojí, že veškeré rovnoměrně osvětlené snímky byly pořízeny kamerami v ruce, bez stativu. Byla snaha exponované plochy zachytit z různých pohledů. Důvodem je to, že plochy jsou ne vždy ideální a mohou zanechat drobné odchylky od ideálního stavu. I když z principu by to extrakci a identifikaci vadit nemělo, je snaha z těchto sad extrahovat co nejlepší možný referenční PRNU. Ukázky některých snímků jsou na obrázku 6.3.



Obrázek 6.3: Ukázka rovnoměrně osvětlených snímků, úplně vlevo se jedná o jasnou oblohu pořízenou z iPhone 7+, uprostřed je ukázka uměle osvětleného papíru z videokamery Panasonic #1 a vpravo je ukázka uměle osvětlené stěny, pořízené z mobilního telefonu Samsung #1

Kromě rovnoměrně osvětlených snímků bylo u všech zmíněných digitálních kamer vytvořeno i několik sad běžných snímků. V takových datových sadách jsou obsaženy u všech senzorů snímky pořízené v různé době, za různých podmínek, ať už jde o okolní teplotu, kvalitu osvětlení (některé exponované scény jsou osvětleny přírodním světlem, jiné umělým, ať už s přidavným světlem či bleskem poskytované digitální kamerou, popřípadě pomocí žárovek či zářivek). Některé snímky jsou viditelně přexponované, jiné podexponované, některé mají vyšší ISO, některé jej naopak mají co nejmenší. Dále jsou některé snímky vyfotografovány v režimu makro, ohnisková vzdálenost je u takových snímků velmi malá. Ukázky takových snímků jsou na obrázku 6.4.



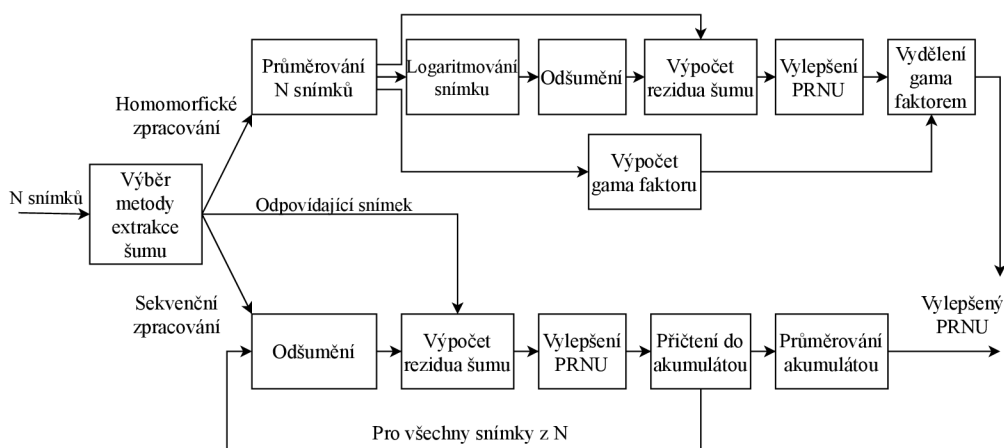
Obrázek 6.4: Ukázka běžných snímků, úplně vlevo se jedná o zaostřenou kočku s přeexponovaným pozadím, pořízenou z iPhone 7+, uprostřed je ukázka detailní textury traviny z videokamery Panasonic #1 a vpravo je ukázka interiérové rostliny, pořízené z mobilního telefonu Samsung #1

Důležité je i upozornit, že některé snímky z těchto náhodných sad byly pořízeny v rámci jednoho dne. V jiných sadách byly naopak snímky pořízeny v dlouhém časovém horizontu třeba i několika let. U všech sad byla snaha zachytit co nejširší možnou paletu scén. Jsou zachyceny jak interiéry, tak exteriéry, jak pravidelné textury, tak náhodné předměty. Některé snímky jsou viditelně rozmazané, jiné jsou o poznání ostřejší. Byla tedy snaha zachytit co nejrůznorodější scény.

Co se týče vlastností samotných multimediálních souborů, veškeré snímky byly pořizovány s co nejvyšší možnou kvalitou, jakou daná zařízení měla k dispozici. Formát těchto souborů byl vždy JPEG, a to vždy s alespoň zachováním kvality 90%. Výjimku tvoří veškerá zařízení iPhone, která byla do tohoto formátu uměle převáděna z nativních formátů HEIF či HEIC, avšak za podmínky co nejmenší možné komprese. Velikosti výsledných snímků u všech datových sad jsou proto v jednotkách megabytů.

## 6.2 Extrakce referenčního šumu

Extrakce referenčního šumu je hlavní krok navrženého postupu práce. Díky tomu, že existuje několik různých postupů pro získání PRNU, je potřeba ověřit takovou metodu, která bude s co možnou největší přesností identifikovat daný obrazový senzor. Blokové schéma implementovaného algoritmu je naznačeno na obrázku 6.5.



Obrázek 6.5: Blokové schéma extrakce referenčního šumu podrobněji

Na vstupu algoritmu je načtení množiny snímků pořízené z jednoho senzoru. Počet snímků z načítané množiny by měl být alespoň třicet, aby byla výsledná kvalita extrahovaného PRNU dostačující. Obecně je dobré mít snímků alespoň padesát pro extrakci PRNU, a to hlavně při extrakci referenčního šumu z množiny běžných snímků. Nutno podotknout, že veškeré snímky na vstupu jsou převedeny před jakýmkoliv zpracováním na šedotónové.

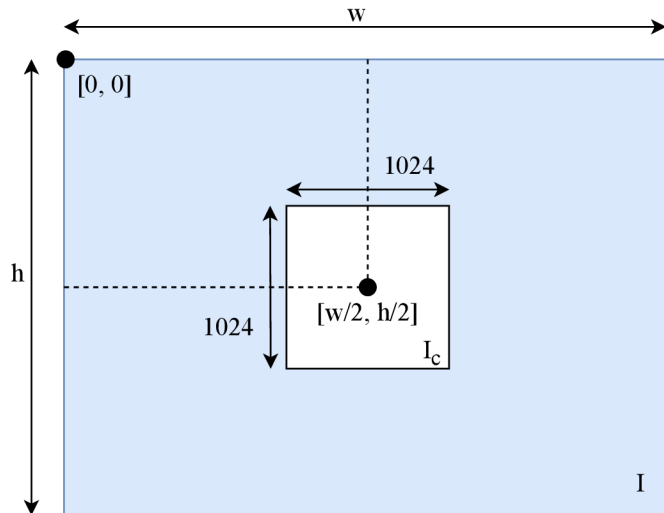
Protože jsou výpočty nad množinou obrázků výpočetně náročné, předchází u všech snímků **ořez** před zpracováním. Pokud jsou snímky na šířku nebo výšku větší, než 1024 pixelů, jsou ořezány právě na velikost 1024, a to vždy vzhledem ke středu snímku. Výsledný ořezaný snímek je tedy dán pomocí následujících bodů:

$$I_c = I \left( \left( \frac{w - 1024}{2}, \frac{h - 1024}{2} \right), \left( \frac{w + 1024}{2}, \frac{h + 1024}{2} \right) \right), \quad (6.1)$$

kde  $w$  je šířka ořezávaného snímku  $I$  a  $h$  je výška ořezávaného snímku  $h$ . Na obrázku 6.6 je tento postup naznačen.

Co se týče velikosti 1024, ta byla vybrána na základě doporučení z různých nastudovaných článků (např. [18]), kde bylo doporučeno použít velikost mocniny 2, tedy buď 256, 512 či 1024. Výsledná plocha extrahovaného referenčního šumu je stejná jako velikost ořezaného snímku. Teoreticky tak lze uchovat na ploše  $1024 \times 1024$ , kde každý pixel může nabývat celočíselnou hodnotu z intervalu  $(0; 255)$  přes 268 milionů různých kombinací vzoru šumu.

Původně algoritmus prováděl ořezávání snímku vzhledem k levému hornímu bodu, tedy k souřadnici  $[0, 0]$ , místo ze středu snímku  $\left[\frac{w}{2}, \frac{h}{2}\right]$ . Důvodem k tomuto rozhodnutí je možné zlepšení výsledků identifikace obrazového senzoru (viz 7). v případě, že původní snímek měl nějaký rozměr liché velikosti, tak jeho střed leží v množině reálných čísel  $\mathbb{R}^+$  a část za desetinnou čárkou se odsekne.



Obrázek 6.6: Schéma ořezání snímku

**Množina snímků** se dá zapsat pomocí této rovnice, kde zadaná notace bude použita i dále v textu:

$$N_{img} = \{I_1, I_2, \dots, I_n\}, \quad (6.2)$$



kde  $N_{img}$  je množina ořezaných snímků a  $I_n$  je ořezaný obrázek z množiny  $N_{img}$ . Snímky jsou dvourozměrné funkce, jsou tedy ve tvaru  $I(x, y)$ , kde  $x$  a  $y$  jsou souřadnice pixelů uvnitř snímku.

Poté, co je vstupní množina snímků ořezána, následuje krok výběru použité metody pro odstranění šumu ze vstupního snímku. Použitých metod je celkem 6 a spadají do dvou kategorií, a to do **sekvenčního zpracování** nebo do **homomorfieckého zpracování**.

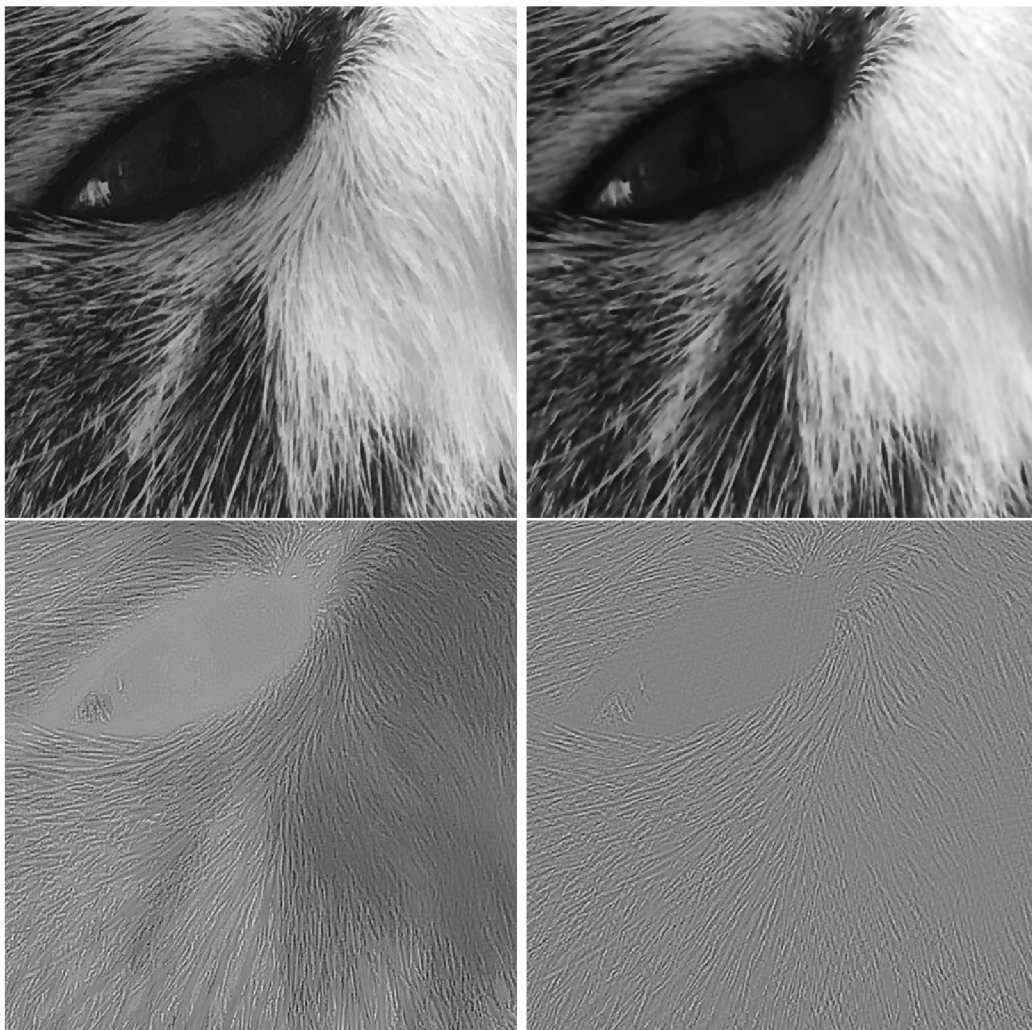
Sekvenční zpracování je základní přístup k extrakci PRNU. Na obrázku 6.5 se jedná o dolní větev za výběrem metody extrakce šumu. V tomto případě se použije aplikace odšumovacího filtru pro každý snímek zvlášť. Naopak u homomorfieckého zpracování se odstranění šumu provádí pouze jednou nad zprůměrovanými snímky z množiny  $N_{img}$ , který je následně logaritmován.

Metody, které provádějí sekvenční odšumění, jsou:

- anizotropní difuze,
- adaptivní Wienerův filtr,
- non-local means,
- filtrace dolní propustí (box filter).

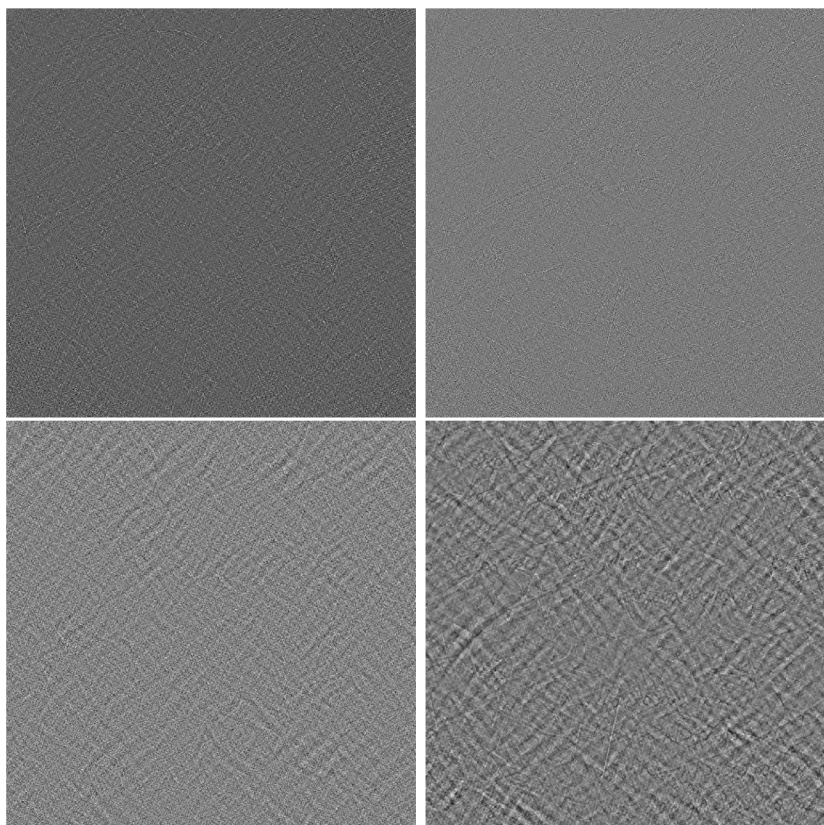
Extrakce PRNU se provádí pomocí rovnice 4.1, uvedené v kapitole 4. Tato rovnice popisuje sekvenční přístup.

Na obrázku 6.7 jsou obrazové výstupy extrakce PRNU z běžného snímku. Jedná se o snímek kočky, pořízené z iPhone 7+. V levé horní části je ořezaný snímek. Vpravo nahoře je na tento snímek aplikován odšumovací filtr. V tomto případě se jedná o adaptivní Wienerův filtr. Následně se vypočítá reziduum odšuměného a původního ořezaného snímku, čímž se extrahuje PRNU. PRNU je znázorněno na obrázku vlevo dole. Nakonec se na tento extrahovaný snímek aplikuje vylepšení, znázorněné na obrázku 6.7 vpravo dole.



Obrázek 6.7: Ukázka výstupu jednotlivých částí algoritmu, postupně zleva: ořezaný snímek, odšuměný snímek, reziduum (PRNU) a vylepšené PRNU

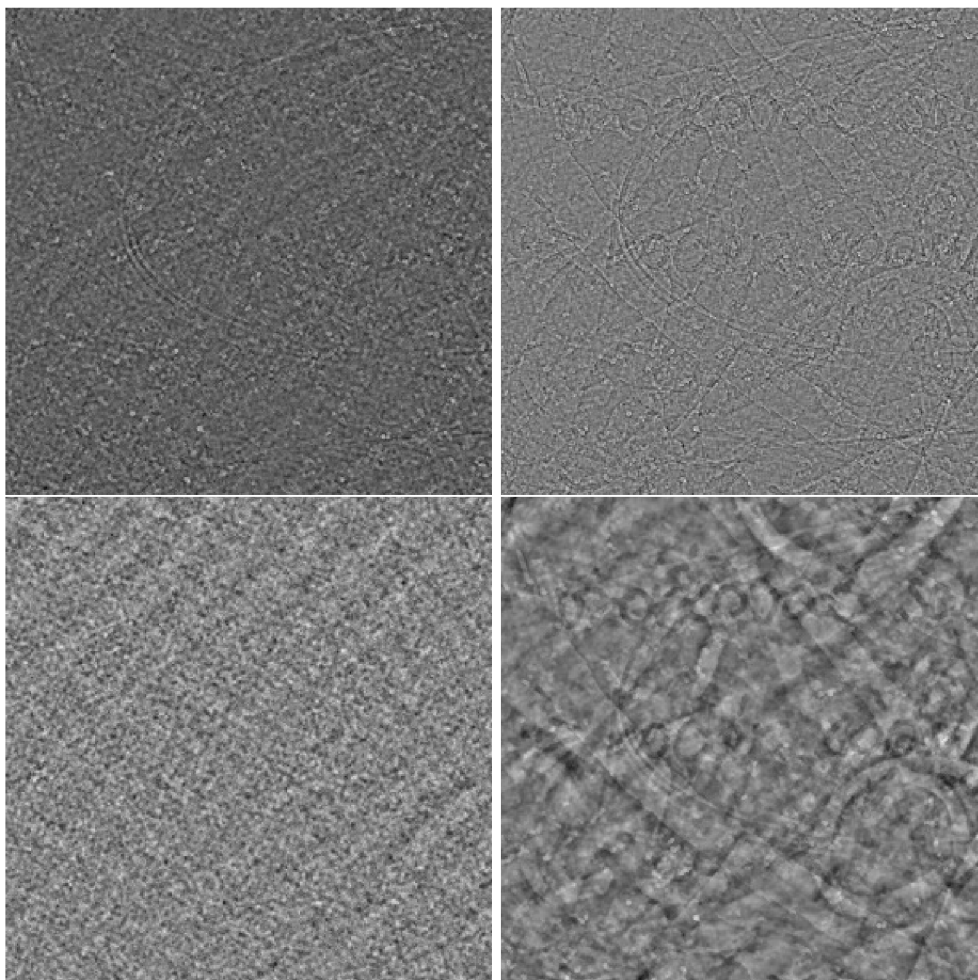
Jak bylo uvedeno výše, je lepší získané PRNU vylepšit. Vylepšení se matematicky provádí jako funkce nad rezidui, vnitřní polynomy rovnice 4.1 se tedy upraví na  $\text{enh}(I_i - F(I_i))$ , kde  $\text{enh}()$  značí vylepšení.



Obrázek 6.8: Ukázka vylepšených PRNU, zleva nahoře PRNU získané pomocí anizotropní difuze, následuje filtr dolní propustí, non-local means a adaptivní Wienerův filtr

Na obrázku 6.8 jsou zobrazeny extrahované vylepšené PRNU z datové sady iPhone 7 #1, náhodné snímky (v tabulce 6.2 se jedná o ID #14). Nahoře vlevo se jedná o použitou filtraci pomocí anizotropní difuze, dále nahoře vpravo je vyobrazeno vylepšené PRNU při použití filtru dolní propustí. Dole je vlevo PRNU získané filtrací pomocí non-local means a vpravo dole je použit pro extrakci adaptivní Wienerův filtr.

Na obrázku 6.9 se jedná o stejné snímky, jako na obrázku 6.8, ale přiblížené do levého horního rohu, kde se ve výsledných vylepšených PRNU objevují artefakty. U anizotropní difuze lze spatřit náznaky kružnic, u dolní propustí navíc i kousky textu. To je způsobeno vysokými frekvencemi, obsaženými ve snímcích z použité datové sady. U adaptivního Wienerova filtru a non-local means lze nalézt viditelné artefakty po celé ploše extrahovaného PRNU.



Obrázek 6.9: Ukázka detailů vylepšených PRNU, pořadí snímků je stejné jako u obrázku 6.8

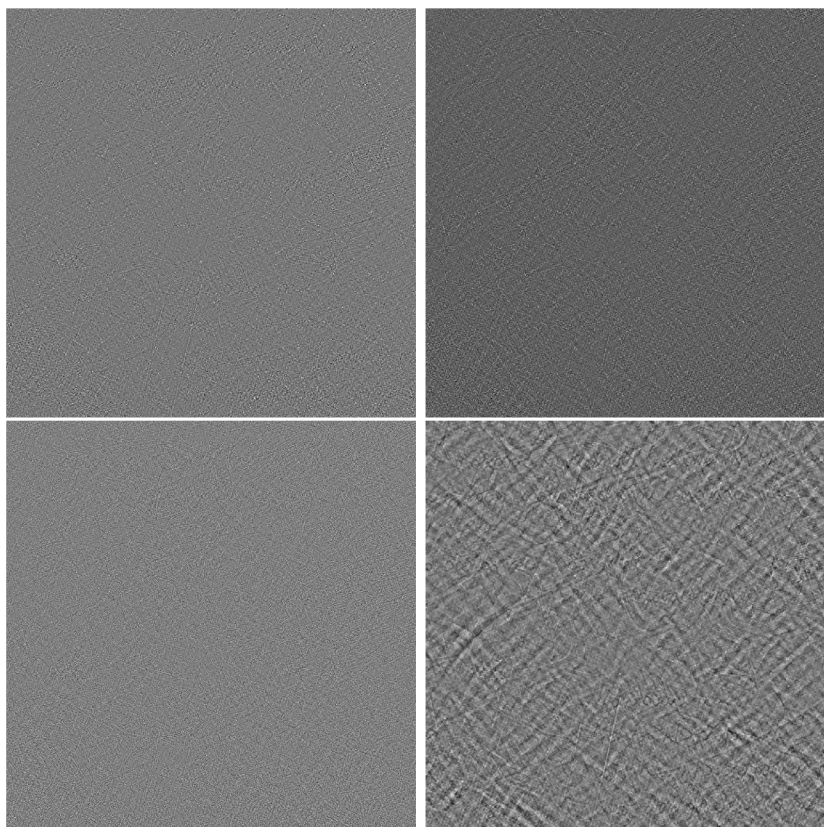
V případě **homomorfického zpracování** se neaplikuje odšumovací filtr pro každý snímek zvlášť, ale pouze se aplikuje jednou pro zprůměrovanou a zlogaritmovanou množinu všech snímků. Díky tomu se několikanásobně snižuje časová náročnost pro extrakci PRNU. Schéma homomorfické filtrace nad množinou snímků je znázorněno na obrázku 6.5, a to jako horní větev blokového schématu.

Použité metody v implementaci, danou rovnicí 4.2, které provádějí odšumění při homomorfickém zpracování, jsou dvě:

- adaptivní Wienerův filtr,
- anizotropní difuze.

Pro vylepšení extrahovaného PRNU se podobně, jako v sekvenčním přístupu, nahradí mnohočlen v čitateli rovnice 4.2 funkcí  $enh(I_{PRNU})$ , která provádí vylepšení.

Díky tomu, že se odstranění šumu provádí pouze nad jedním zprůměrovaným a zlogaritmovaným snímkem a ne nad všemi snímky z množiny  $N_{img}$ , jedná se o významné urychlení extrakce PRNU. Proto lze homomorfického přístupu využít v případě zpracování obrovských datasetů čítajících tisíce snímků či videí.



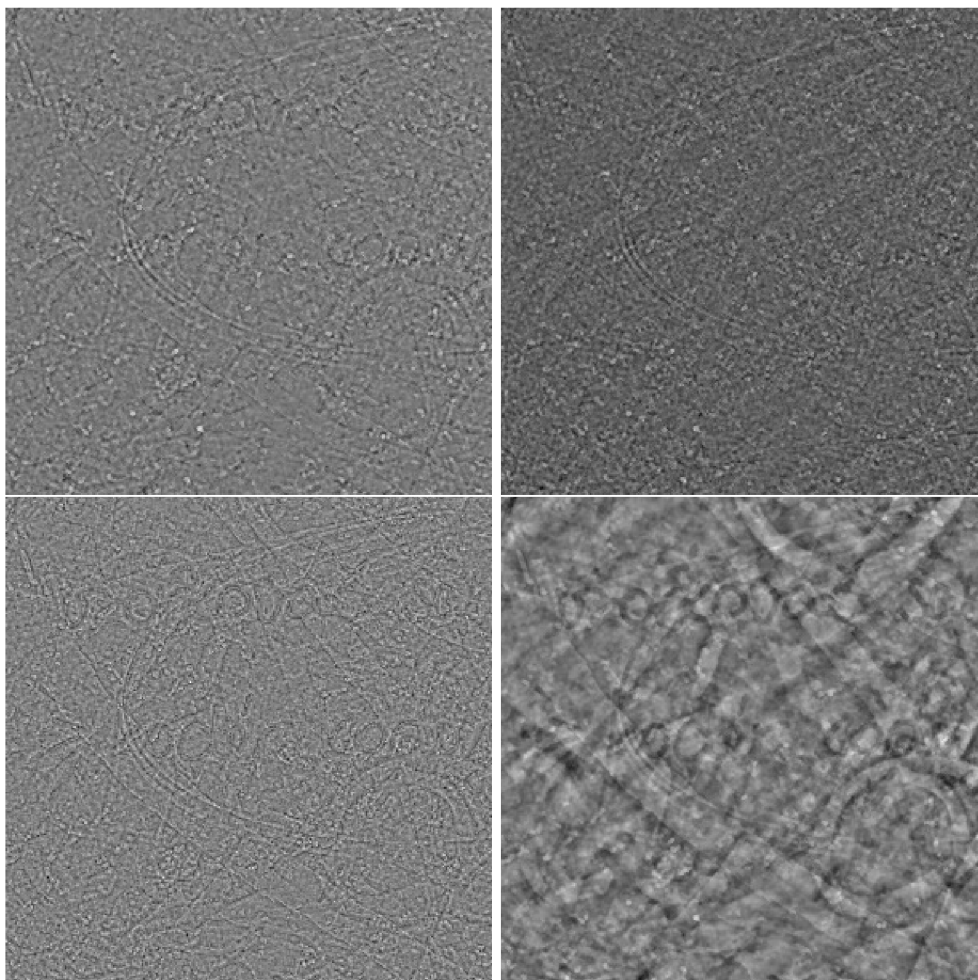
Obrázek 6.10: Ukázka vylepšených PRNU, vlevo je vždy homomorfní přístup, vpravo pro srovnání sekvenční. Nahoře je anizotropní difuze, dole adaptivní Wienerův filtr

Urychlení extrakce referenčního PRNU je znatelné, například pro extrakci z datové sady Samsung #1, náhodné snímky (z tabulky 6.2 se jedná o ID 31), trvala extrakce pro sekvenční anizotropní difuzi celkem 118 471 ms a pro stejnou sadu homomorfní trvala extrakce pouze 10 356 ms. Snímky z datové sady byly ořezány na velikost  $1024 \times 1024$  pixelů.

Na obrázku 6.10 je zobrazeno srovnání homomorfního a sekvenčního přístupu k extrakci PRNU. V levém sloupci jsou zobrazeny homomorfní extrahované PRNU a v pravém naopak sekvenčně získané PRNU. Na prvním řádku je použitou metodou odšumění anizotropní difuze, která produkuje na první pohled podobný šum. Na druhém řádku se jedná o adaptivní Wienerův filtr.

Vzájemná míra korelace (Pearsonova korelačního koeficientu) mezi sekvenční a homomorfní anizotropní difuzí je 0.766. Korelace mezi sekvenčním a homomorfním adaptivním Wienerovým filtrem je menší, a to 0.530. Pro zajímavost, velikost korelace mezi homomorfní anizotropní difuzí a homomorfním adaptivním Wienerovým filtrem je 0.822 a korelace mezi sekvenční anizotropní difuzí a homomorfním adaptivním Wienerovým filtrem je 0.696.

Dá se tedy ověřovat hypotéza, jestli pro identifikaci senzoru stačí homomorfní přístup, případně jestli je lepší použít přístup sekvenční. Na první pohled se zdá, že u anizotropní difuze je podobnost takových PRNU vysoká, avšak ne stoprocentní, ovšem u adaptivního Wienerova filtru je podobnost naopak poměrně nízká. Ověření této hypotézy je popsáno dále v textu, v kapitole 7.



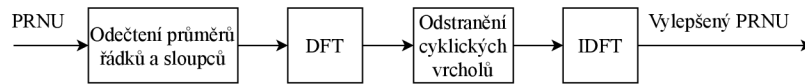
Obrázek 6.11: Ukázka vylepšených PRNU, vlevo je vždy homomorfický přístup, vpravo pro srovnání sekvenční. Nahoře je anizotropní difuze, dole adaptivní Wienerův filtr. Jedná se vždy o stejný detail z levé části extrahovaného PRNU

Na obrázku 6.11 jsou zobrazeny výřezy stejných PRNU, jako jsou na předchozím obrázku 6.10. Zde jsou lépe vidět artefakty a vysoké frekvence, které takto extrahovaný šum znehodnocují. Na první pohled se výrazně liší sekvenční adaptivní Wienerův filtr od ostatních PRNU, který kromě artefaktů, obsažených i v ostatních snímcích, obsahuje navíc viditelné diagonální artefakty. Ty jsou sice patrné i u sekvenční anizotropní difuze, ale nejsou natolik výrazné, až se skýtá otázka, jestli je tento filtr vůbec v sekvenčním přístupu vhodné použít.

Rozhodl jsem se implementovat i vylepšení extrahovaného PRNU, které zlepšuje úspěšnost identifikace senzoru. Taková vylepšení jsou implementována celkem dvě:

- odečtení průměru řádků a sloupců,
- odstranění cyklických vrcholů.

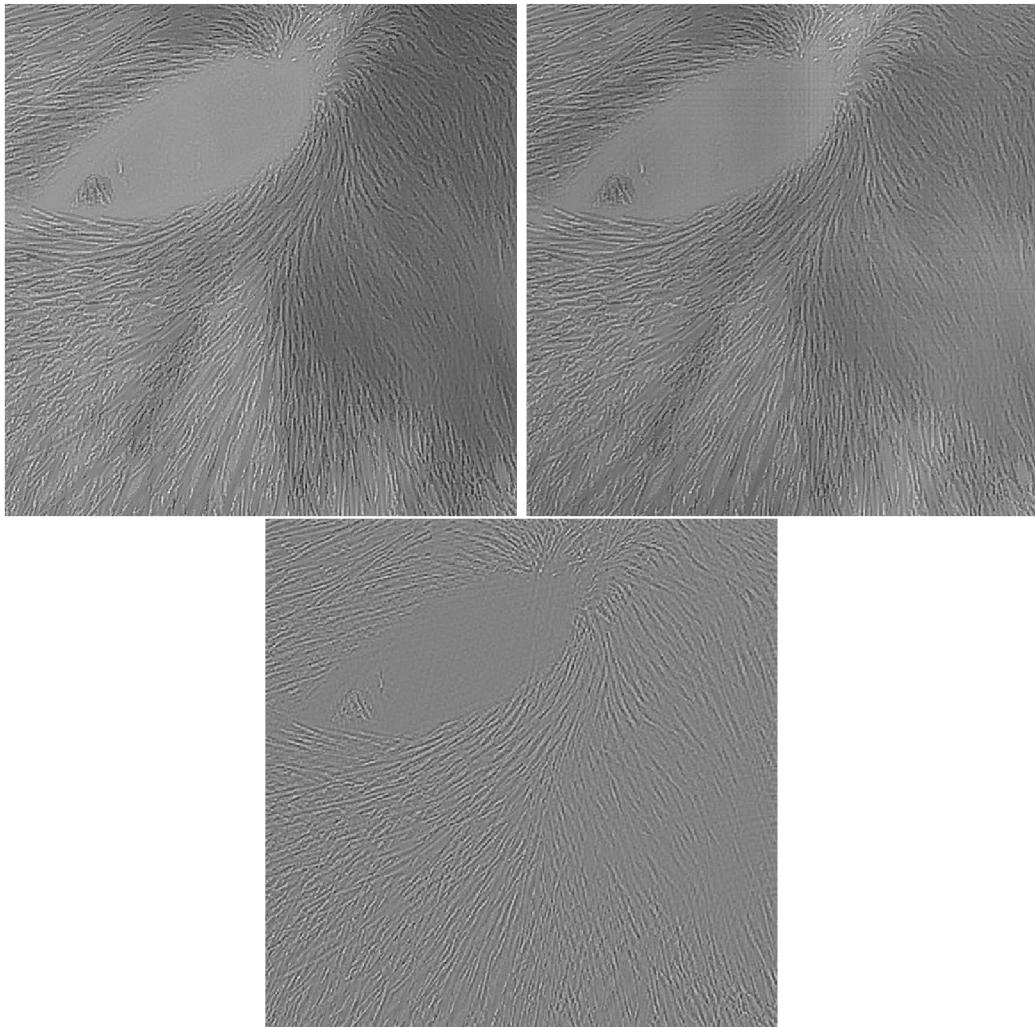
Postup a pořadí těchto vylepšení je uveden na obrázku 6.12.



Obrázek 6.12: Blokové schéma vylepšení referenčního šumu

**Odečtení průměru řádků a sloupců** (též zeroing-out the means) se provádí jako první možné vylepšení extrahovaného PRNU. Zde dochází k potlačování lineárních vzorů v obraze, které vznikají při demosaicingu, případně při zpracování řádků a sloupců v obrazovém senzoru či následně v kameře. Vylepšení se provádí tak, že se zjistí průměrné hodnoty ve všech řádcích a sloupcích ve snímku a u všech pixelů v obraze se tato hodnota odečte od hodnoty pixelu.

**Odstranění cyklických vrcholů** je posledním implementovaným vylepšením PRNU. Ty vznikají při demosaicingu a JPEG kompresi. Horizontální a vertikální artefakty, které se v PRNU objeví, se filtrují pomocí speciálního filtru ve Fourierově doméně.



Obrázek 6.13: Ukázka jednotlivých výstupů vylepšování PRNU

Na obrázku 6.13 jsou taková vylepšení zobrazena. Nahoře vlevo se jedná o čistý PRNU získaný sekvenčním nebo homomorfickým přístupem. Následuje odečtení průměru řádku a sloupců a odstranění cyklických vrcholů. Na první pohled je vylepšený PRNU čistší.

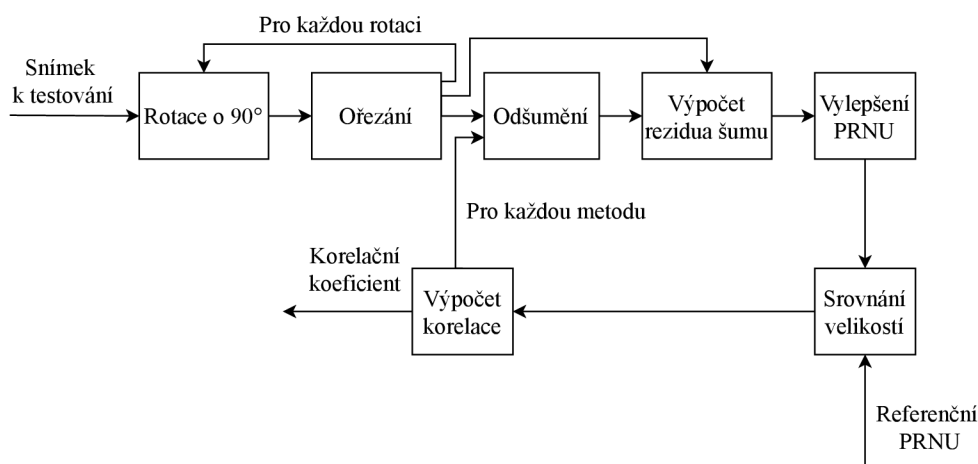
Extrakce referenčního PRNU je důležitým krokem při problému identifikace obrazového senzoru. Cílem je získat jedinečný otisk takového senzoru, ze kterého bylo pořízeno několik snímků. Presentovány byly dva přístupy: sekvenční a homomorfický. Závěrem je ještě provedeno vylepšení PRNU, aby byl získaný *fingerpřint* obrazového senzoru co nejlepší možný.

### 6.3 Určení korelačních koeficientů

V předchozí sekci byly prezentovány různé přístupy pro extrakci referenčních PRNU z datových sad. Dalším logickým krokem je ověření různých hypotéz:

1. Je daný snímek pořízen z tohoto senzoru?
2. Je tato množina snímků pořízena z tohoto senzoru?

Pro ověření takových hypotéz je nutné mít k dispozici referenční PRNU, získané z ověřovaného senzoru a snímek (případně snímky), u kterých nás tato otázka zajímá. Pro takové ověření se v této práci používají Pearsonovy korelační koeficienty, které určují míru podobnosti mezi dvěma veličinami. Na obrázku 6.14 je zobrazeno blokové schéma implementovaného algoritmu pro testování první hypotézy.



Obrázek 6.14: Blokové schéma určení korelačních koeficientů

Nejprve je potřeba mít k dispozici snímek, pořízený z digitální kamery  $C$ . Na tento snímek se poté aplikuje rotace, vždy o  $90^\circ$ . Snímek se následně ořeže na maximální velikost  $1024 \times 1024$ . Vzniknou tak čtyři varianty tohoto snímku, u kterých se tato hypotéza ověřuje. Předpokládá se totiž, že snímek může být libovolně rotován. Bude-li ovšem rotace takového snímku jiná, než  $90^\circ$ , algoritmus nebude schopen správně na hypotézu odpovědět.

Pro každou takovou rotaci testovaného snímku se provede podobný postup, jako při extrakci PRNU. Zde již nemá smysl zvažovat homomorfický přístup, protože se pracuje nad jedním konkrétním snímkem. Odšumění, která jsou k dispozici, jsou stejná, jako při sekvenční extrakci PRNU:

- anizotropní difuze,



- adaptivní Wienerův filtr,
- non-local means,
- filtrace dolní propustí (box filter).

Poté, co se provede odstranění šumu, vypočítají se rezidua šumu a získají se tak PRNU. Ta se poté vylepší, a to stejnými metodami, jako v případě extrakce referenčního PRNU. Samozřejmě se tak děje pro všechny rotace testovaného snímku.

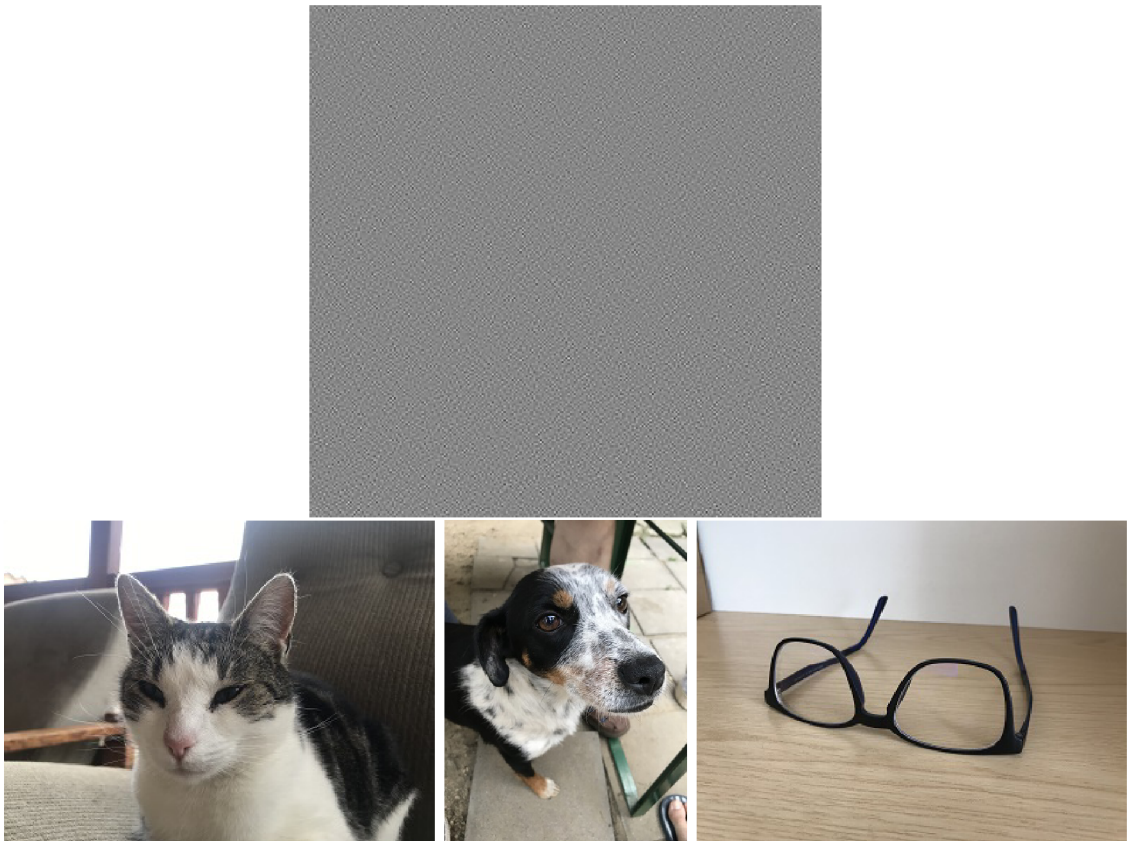
Dalším krokem je srovnání velikostí referenčního PRNU a testovaného PRNU. Pearsonova korelace totiž umožňuje pouze porovnávat veličiny, které mají stejnou velikost. Srovnání se provede ořezáním většího snímku PRNU na menší.

Následně se vypočítá korelační koeficient mezi těmito dvěma šumy. Korelační koeficient je přímo normalizován do intervalu  $\langle -1, 1 \rangle$ , kde -1 značí opačnou korelaci, 0 značí žádnou korelaci a 1 značí co nejvyšší možnou korelaci. Hodnoty limitně se blížící nule tedy říkají, že testované PRNU si nejsou podobné, a tedy pravděpodobně nejsou ze stejného senzoru. Naopak vyšší korelace značí, že testované šumy si podobné nějakým způsobem jsou a je tedy více pravděpodobné, že takové snímky byly pořízeny ze stejného obrazového senzoru.

Korelace mezi referenčním PRNU a PRNU testovaného snímku se dá též zapsat jako

$$c_i = \text{corr}(RPN_{EPRNU}, IPN_{EPRNU}), \quad (6.3)$$

kde  $\text{corr}()$  je funkce pro výpočet Pearsonovy korelace,  $RPN_{EPRNU}$  je referenční vylepšený PRNU a  $IPN_{EPRNU}$  je vylepšený PRNU, získaný z testovaného snímku.



Obrázek 6.15: Referenční PRNU, získaný z datové sady iPhone 7+, jasná obloha a testovací snímky

Tím, že se předpokládá, že testovaný snímek nemá stejnou orientaci jako referenční šum, počítají se korelační koeficienty pro všechny možné rotace. Na obrázku 6.15 je zobrazen referenční PRNU, získaný z datové sady *iPhone 7+*, *jasná obloha*, a to pomocí sekvenční anizotropní difuze. Označme tento senzor dočasně jako  $C_1$ . Dále jsou zobrazeny testovací snímky, kde:

- snímek č. 1, *kočka*, je pořízen ze stejného obrazového senzoru, z jakého byl extrahován referenční PRNU, tedy  $C_1$  a tento snímek je součástí datové sady *iPhone 7+*, náhodné snímky,
- snímek č. 2, *pes*, je pořízen ze zařízení *iPhone 7 #1* ( $C_2$ ) a je součástí datové sady *iPhone 7 #1*, náhodné snímky,
- snímek č. 3, *brýle*, je pořízen ze zařízení *iPhone 7 #2* ( $C_3$ ) a je součástí datové sady *iPhone 7 #2*, náhodné snímky.

Lze tedy ověřovat u těchto tří snímků následující hypotézy:

- $H_1$ : Snímek *kočka* je pořízen z  $C_1$ .
- $H_2$ : Snímek *pes* je pořízen z  $C_1$ .
- $H_3$ : Snímek *brýle* je pořízen z  $C_1$ .

Pro zjištění korelačních koeficientů byl aplikován algoritmus, který byl popsán výše. U všech snímků byla provedena extrakce PRNU pro všechny rotace pomocí anizotropní difuze. Dále byla provedena korelace těchto extrahovaných šumů vůči referenčnímu PRNU. Naměřené korelační koeficienty jsou zobrazeny v tabulce 6.3. Důležité je zmínit, že testované snímky nebyly použity při extrakci referenčního PRNU, a to z důvodu možného zkreslení naměřených výsledků.

Název snímku	Kamera	0°	90°	270°	180°
Kočka	$C_1$	0,03685	0,00108	0,00112	0,00116
Pes	$C_2$	0,00086	-0,00346	0,00237	0,00054
Brýle	$C_3$	-0,00029	0,00275	0,00065	0,00320

Tabulka 6.3: Naměřené korelační koeficienty pro ukázkové snímky, použitou metodou je vždy anizotropní difuze

Rozeberu-li naměřené hodnoty korelačních koeficientů, pro snímek *kočka* je pro nerotovaný snímek korelační koeficient oproti jiným výrazně vyšší. Pro další rotace jsou korelační koeficienty o řád nižší, stejně jako pro snímky získané ze senzorů  $C_2$  a  $C_3$ . U těch jsou korelační koeficienty dokonce záporné a více konvergující k nule.

Pro odpovězení na hypotézy  $H_1$ ,  $H_2$  a  $H_3$  je nutno stanovit rozhodovací práh  $t$ , kdy pro hodnoty korelačních koeficientů nižších než práh  $t$  se odpoví odpovědí *ne* a naopak pro hodnoty větší nebo rovny než práh  $t$  se odpoví odpovědí *ano*:

$$A = \begin{cases} \text{ano} & \text{když } c_i \geq t \\ \text{ne} & \text{když } c_i < t \end{cases} \quad (6.4)$$

Stanoví-li se práh  $t = 0,01$ , lze na hypotézu  $H_1$  odpovědět s vysokou pravděpodobností *Ano*. Na hypotézy  $H_2$  a  $H_3$  lze odpovědět s vysokou pravděpodobností *Ne*. Více podrobnějších vyhodnocení včetně popsání určení velikosti prahu je k dispozici v kapitole 7.

Těž je důležité poznamenat, že při ověřování zdrojové digitální kamery není algoritmu jasné, jaký postup a především jaký odšumovací filtr byl pro extrakci referenčního PRNU použit. Proto jsou pro všechny rotace vypočteny i korelační koeficienty pro všechny implementované odšumovací filtry.

## 6.4 Implementace

Pro implementaci návrhu řešení jsem se rozhodl vytvořit aplikaci s uživatelským rozhraním, ze kterého půjdou ovládat jednotlivé části popsané výše:

- extrakce referenčního šumu ze zadané vstupní množiny snímků,
- testování množiny snímků, zda je z určitého obrazového senzoru.

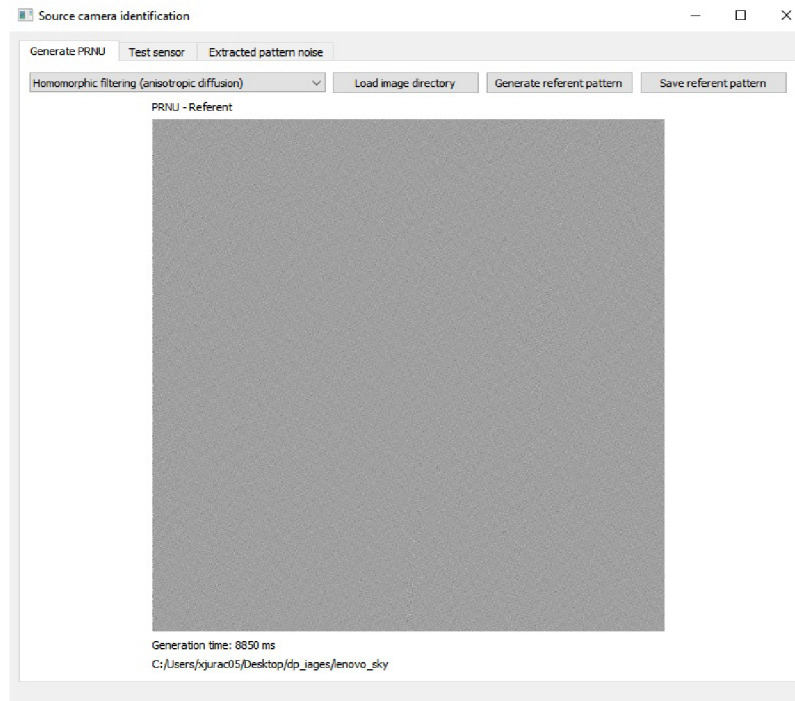
Pro samotnou implementaci byl zvolen jazyk C++, konkrétně standard C++17. Pro implementaci uživatelského rozhraní byl zvolen *framework* Qt 5.14. V [11] je uvedeno, že se jedná o multiplatformní framework, výsledná aplikace je tedy přeložitelná na různých architekturách a různých operačních systémech. Pro zpracování obrazu byla zvolena knihovna OpenCV 3.4.3. Ta je popsána v [7].

Pro extrakci referenčního šumu bylo implementováno šest metod, které byly popsány v předchozích sekcích. Konkrétní implementace anizotropní difuze byla použita z [23]. Jak bylo uvedeno v [18], před samotným aplikováním anizotropní difuze bylo provedeno aplikování mediánového filtru o velikosti okna 5.

Pro adaptivní Wienerův filtr byla použita implementace z [5]. Algoritmy pro non-local means a box-filter jsou knihovní funkce z OpenCV, tedy [7]. Dále byly implementovány dvě vylepšení PRNU pro extrakci referenčního šumu.

Pro extrakci PRNU je zobrazeno uživatelské rozhraní na obrázku 6.16. Před samotnou extrakcí je nejprve nutné zvolit metodu pro extrakci a vybrat adresář snímků, obsahující obrázky ve formátech buď BMP nebo JPEG. Adresář takových snímků musí obsahovat alespoň 30, jinak nelze tento adresář pro extrakci PRNU použít.

Většina výpočtů nad obrazem se provádí v 32 bitových reálných číslech, kde jsou jednotlivé intenzity pixelů rozloženy v intervalu  $\langle 0, 1 \rangle$ , kde 0 značí černou a 1 bílou barvu. Některé filtrace ovšem pracují v kladných osmibitových celých číslech a je nutné takové obrazy mezi sebou konvertovat. Díky tomu zde dochází ke ztrátě informací.



Obrázek 6.16: Uživatelské prostředí pro extrakci PRNU

Poté, co se provede extrakce PRNU, je tento šum zobrazen v okně a uživatel zde může pozorovat jeho různé vlastnosti, například artefakty či defektní pixely. Zároveň je tento šum brán jako *aktivní* a lze jej poté rovnou použít pro ověřování pravosti obrazu. Výsledné PRNU lze poté uložit, kde doporučujícím formátem je BMP. Pokud jsou rozměry snímků, ze kterých se provádí extrakce větší než 1024 pixelů, provede se ořezání, a to vůči středu snímku.

Pro testování pravosti snímku je nutné mít nejprve načtený referenční šum, případně množinu referenčních šumů, pokud se provádí výpočet pro více obrazových sensorů zároveň. Následně je potřeba načíst snímek, případně množinu snímků, u kterých chceme ověřit pravost dat. Ověření se provádí postupem, popsáním v předchozí sekci.

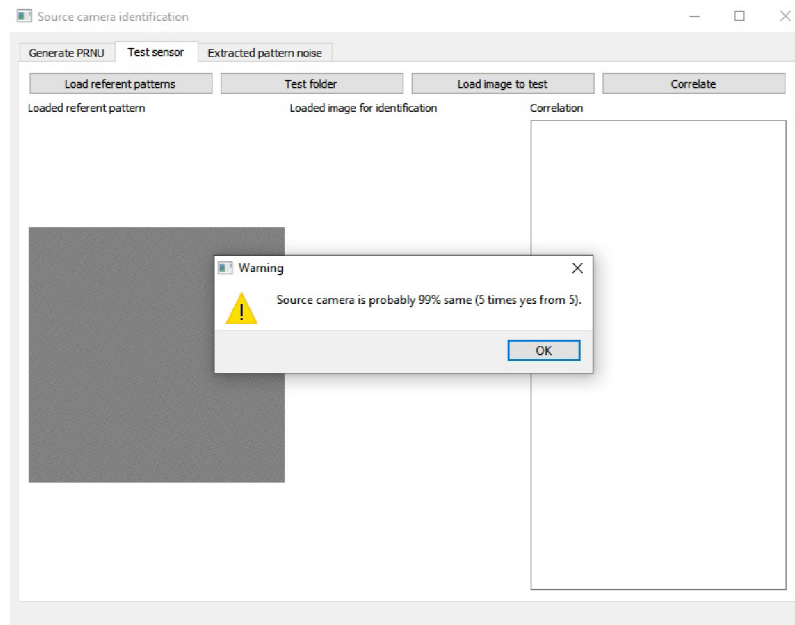
Ořezání snímku se provádí stejným způsobem jak při extrakci. Dále jsou vypočítány rotace tohoto výřezu o  $90^\circ$ . Díky tomu je algoritmus robustnější na odhalování podvržených snímků.

Pro výpočet korelačních je nutné ověřit, že jsou snímky stejně veliké. Pokud nejsou, je větší snímek ořezán na menší snímek, jinak nelze koeficienty vypočítat. Počítají se pro všechny rotace testovaného snímku.

Jak bylo uvedeno v předchozí sekci, není jisté, jaký filtr byl použit pro extrakci PRNU. Proto jsou postupně aplikovány všechny filtrace odstranění šumu na snímek a korelační koeficienty jsou ukládány do vytvořeného objektu, obsahující pro všechny snímky a jejich rotace i všechny metody.

Korelační koeficienty se poté vypíší na výstup ve formátu CSV (anglicky Comma-separated values), obsahující vždy pro testovaný referenční PRNU množinu korelačních koeficientů všech rotací odšuměných testovacích snímků, všech odšumovacích metod a všech testovaných snímků.

Na základě vyhodnocení, uvedeného v kapitole 7, je posledním krokem rozhodnutí, jestli je daný snímek z testovaného senzoru. Rozhodnutí je stanoveno porovnáním hodnoty korelačního koeficientu s prahem, nastaveným na  $t = 0,013$ , což je zaručená hodnota s chybou druhého typu maximálně 1 %.



Obrázek 6.17: Uživatelské prostředí pro testování pravosti

Rozhodnutí se provádí pro všechny rotace snímku a v případě, že alespoň jedna rotace má korelační koeficient větší než práh, akceptuje se hypotéza, že referenční PRNU a testovaný snímek jsou ze stejného obrazového senzoru. V případě, že není testován jen jeden snímek, ale celý adresář, musí být alespoň polovina snímků v adresáři akceptována, aby se dal adresář prohlásit za adresář, obsahující snímky ze stejné zdrojové kamery. Tato skutečnost je znázorněna na obrázku 6.17.

## Kapitola 7

# Vyhodnocení algoritmu

V rámci této kapitoly je kromě vyhodnocení vlivu výběru odšumovacího filtru, vlivu výběru datových sad pro extrakci šumu či velikosti a lokality extrahovaného šumu, ještě několik specifických vyhodnocení pro naměřená data, namodelovaných do možných reálných situací. Pro veškeré testy byla použita standardní metodika, v rámci níž se například nikdy neobjevily v testovacích sadách snímky, které byly použity pro extrakci referenčního šumu.

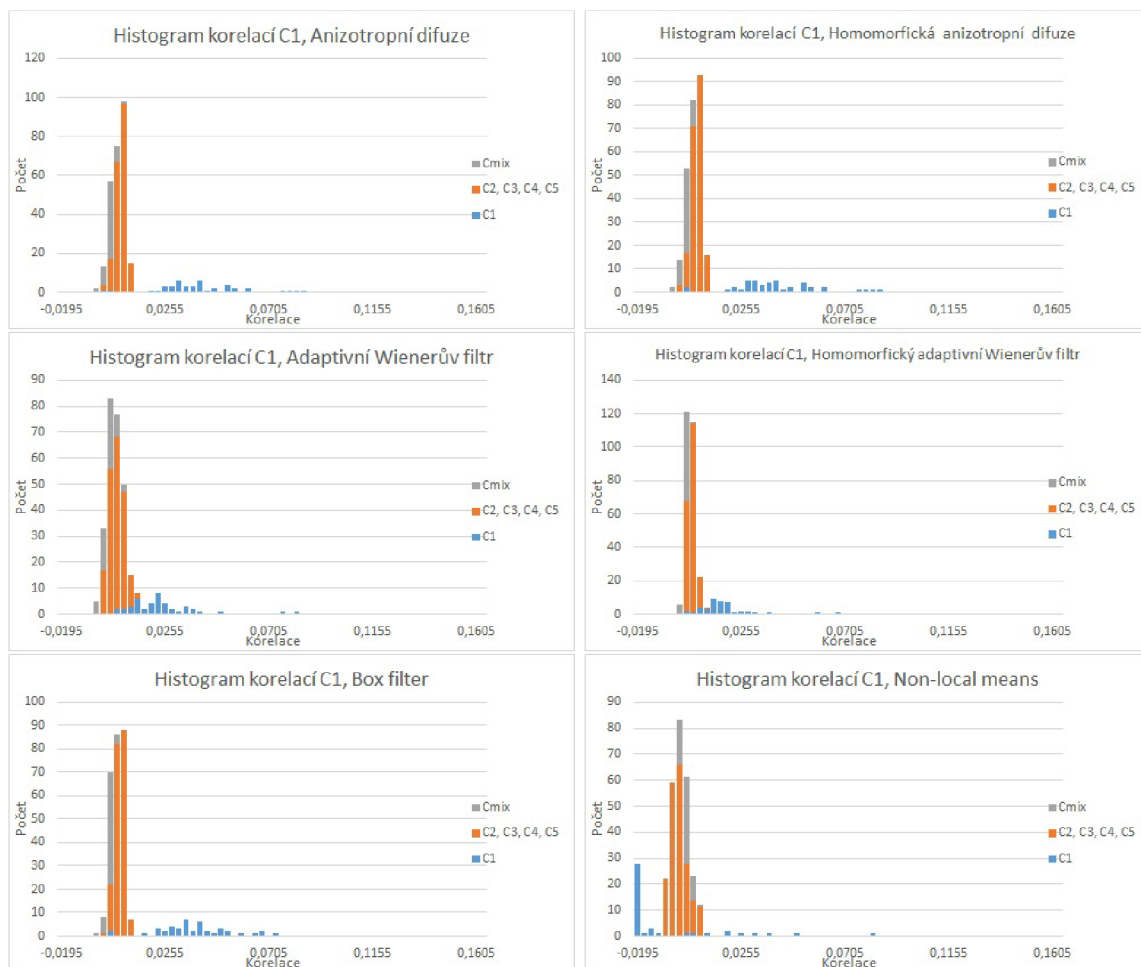
### 7.1 Výběr odšumovacího filtru

Prvním vyhodnocením je hledání odpovědi na otázku: „*Který odšumovací filtr produkuje nejlepší možné výsledky?*“ Je totiž důležité, jaký filtr se použije. V případě, že se použije *špatný* filtr, tak bude pravděpodobnost správné identifikace zdrojové kamery menší. V hypotetickém soudním řízení tak například může být osvobozen zločinec pro nedostatek důkazního materiálu.

Na obrázku 7.1 jsou zobrazeny histogramy korelací pro obrazový senzor Samsung #1, kde jsou použity všechny zmíněné odšumovací filtry. Použité datové sady jsou vypsány v tabulce 7.1.

Název sady	Označení	Použití	Počet snímků
Samsung #1, stěna	—	extrakce RPN	83
Samsung #1, náhodné snímky	$C_1$	testování	43
Samsung #2, náhodné snímky	$C_2$	testování	61
Samsung #3, náhodné snímky	$C_3$	testování	38
Samsung #4, náhodné snímky	$C_4$	testování	43
Samsung #5, náhodné snímky	$C_5$	testování	87
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.1: Použité sady pro vyhodnocení č. 1

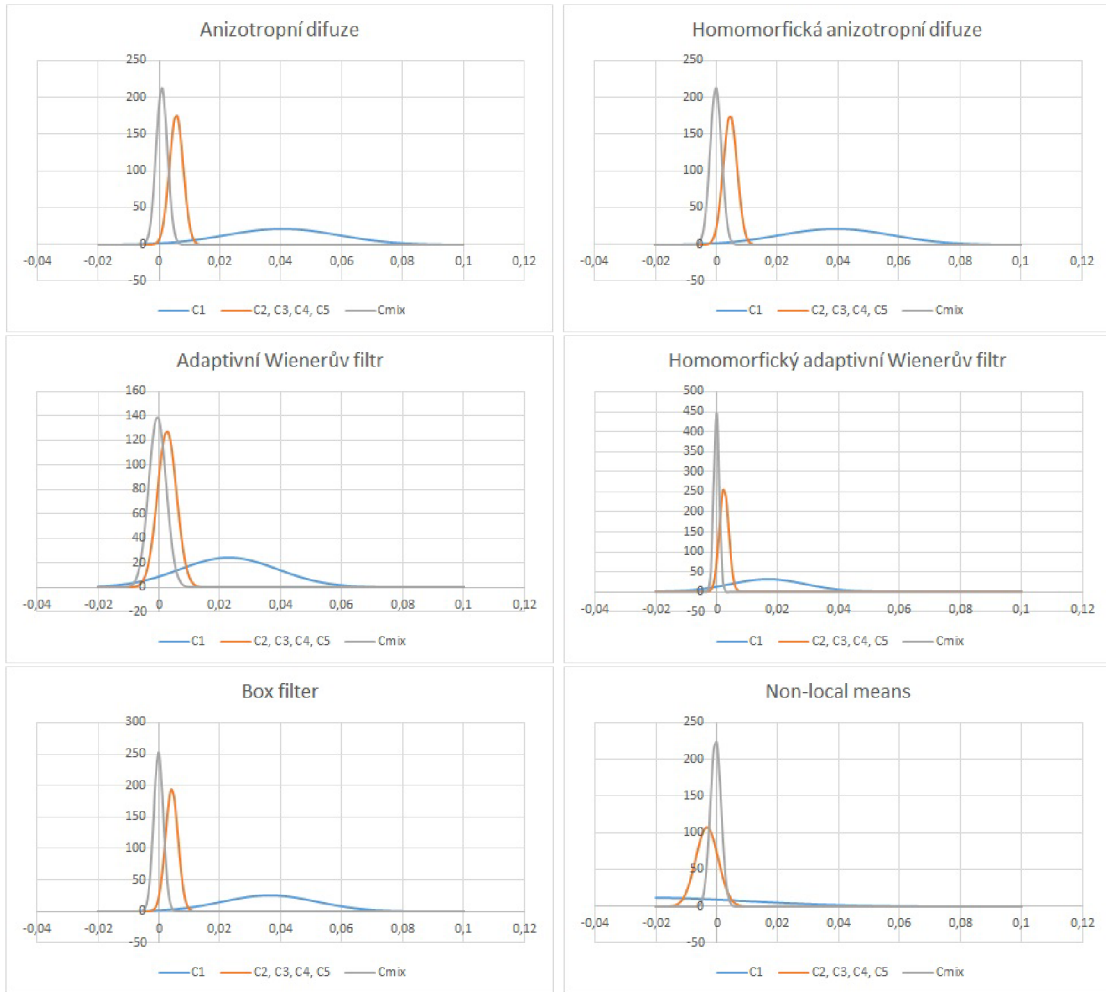


Obrázek 7.1: Histogramy korelací pro obrazový senzor Samsung #1

Ve všech grafech jsou zobrazeny tři histogramy. První, označen modrou barvou, je histogram korelací pro snímky ze stejného senzoru  $C_1$ . Oranžově je označen histogram korelací pro zařízení ze stejné třídy senzorů (stejný model digitální kamery, stejná velikost a druh obrazového senzoru, avšak jiné kusy zařízení, než  $C_1$ ) a šedě je označena množina jiných senzorů různých typů ( $C_{mix}$ ). Jednotlivé hodnoty v grafu se skládají na sebe.

Histogramy říkají, jaká četnost korelace se pro testované snímky naměřila. Ideální stav je, když  $C_1$  bude obsahovat vyšší hodnoty korelace, než z ostatních senzorů. Čím více dojde k vzájemnému překryvu, tím je metoda horší pro přesnou identifikaci senzoru.

Non-local means obsahuje korelační koeficienty  $C_1$  s velkou četností záporné, a proto je tato filtrace nevhodná a nejméně přesná. V případě adaptivního Wienerova filtru jsou viditelné překryvy v rámci třídy kamer Samsung. Navíc i rozsah hodnot korelací pro kamery  $C_2 - C_5$  je daleko vyšší, než u jiných metod, a proto i tato metoda není příliš vhodná pro přesnou identifikaci.



Obrázek 7.2: Normální rozdělení pro obrazový senzor Samsung #1

Pro lepší a přesnější vyhodnocení je lepší vypočítat pro jednotlivé datové sady parametry pro výpočet normálního rozdělení. Pro každou metodu tedy byla vypočítána střední hodnota a rozptyl. Na obrázku 7.2 jsou taková rozdělení pro 7.1 zobrazena. Integrály v průniku křivek popisují chyby 1. a 2. typu (*false positives*, FP a *false negatives*, FN). Pokud leží naměřená korelace v tomto průniku křivek, snižuje se pravděpodobnost správného odhadu identifikace senzoru. Možná hypotetická otázka pro taková ověření zní: „**Je snímek  $I$  pořízen z kamery  $C_1$ ?**“

Název metody	FN [%]	Práh
Anizotropní difuze	7,84	0,014
Homomorfická anizotropní difuze	7,96	0,014
Adaptivní Wienerův filtr	28,88	0,014
Homomorfický adaptivní Wienerův filtr	24,19	0,008
Box filter	5,78	0,012
Non-local means	86	0,008

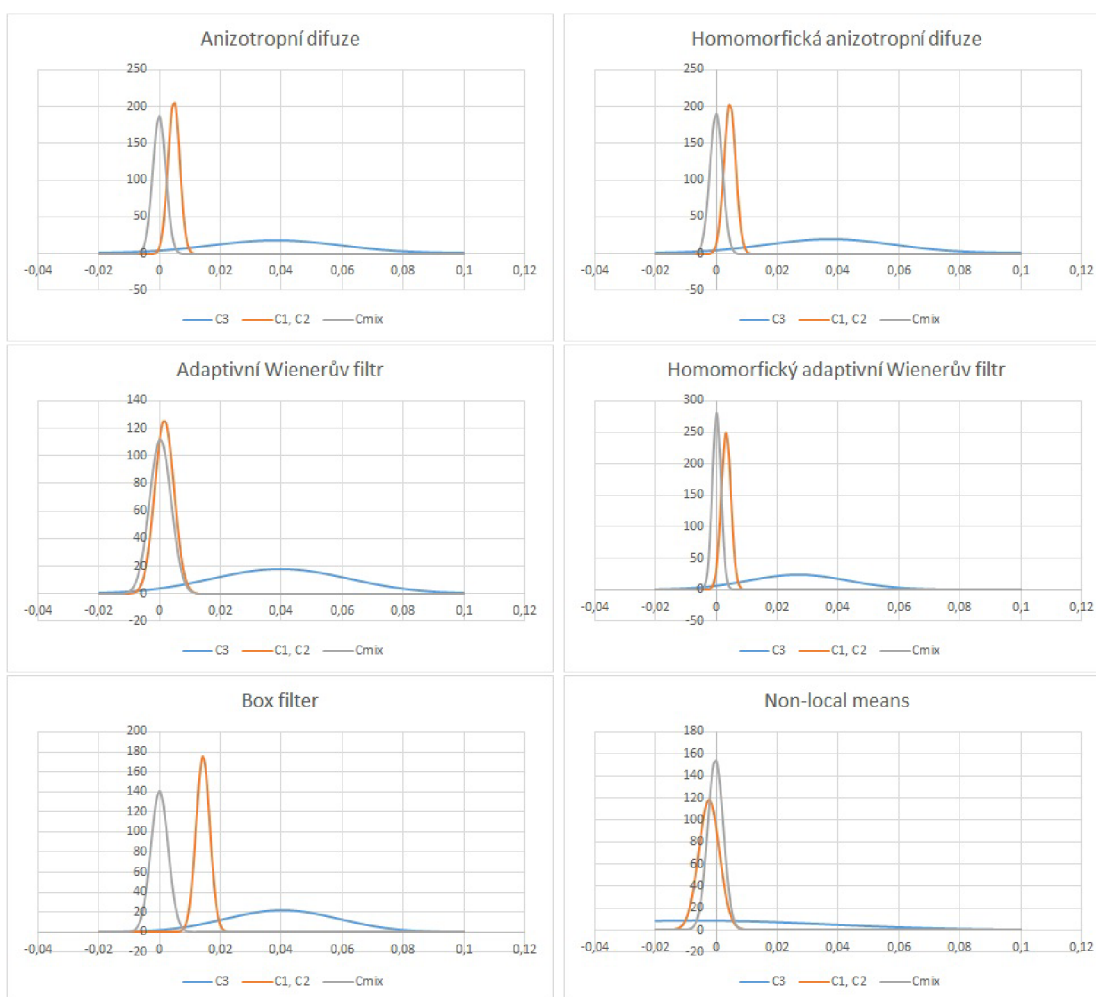
Tabulka 7.2: Vyhodnocení č. 1,  $FP \leq 1\%$



V tabulce 7.2 jsou uvedeny chyby prvního a druhého typu. Dobře dopadly anizotropní difuze a box filter, špatně non-local means. Je ovšem důležité provést vyhodnocení i nad jinými sadami.

Název sady	Označení	Použití	Počet snímků
Panasonic video #3, papír	—	extrakce RPN	53
Panasonic video #1, náhodné snímky 2	$C_1$	testování	82
Panasonic video #2, náhodné snímky 2	$C_2$	testování	95
Panasonic video #3, náhodné snímky 2	$C_3$	testování	57
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.3: Použité sady pro vyhodnocení č. 2



Obrázek 7.3: Normální rozdělení pro obrazový senzor Panasonic video #3

V tabulce 7.3 je zobrazeno další měření, tentokrát pro kameru Panasonic video #3. Referenční šum tohoto obrazového senzoru byl získán z datové sady Panasonic video #3, papír. Na obrázku 7.3 jsou zobrazena náhodná rozdělení, určená z naměřených hodnot. V tabulce 7.4 jsou výsledky.

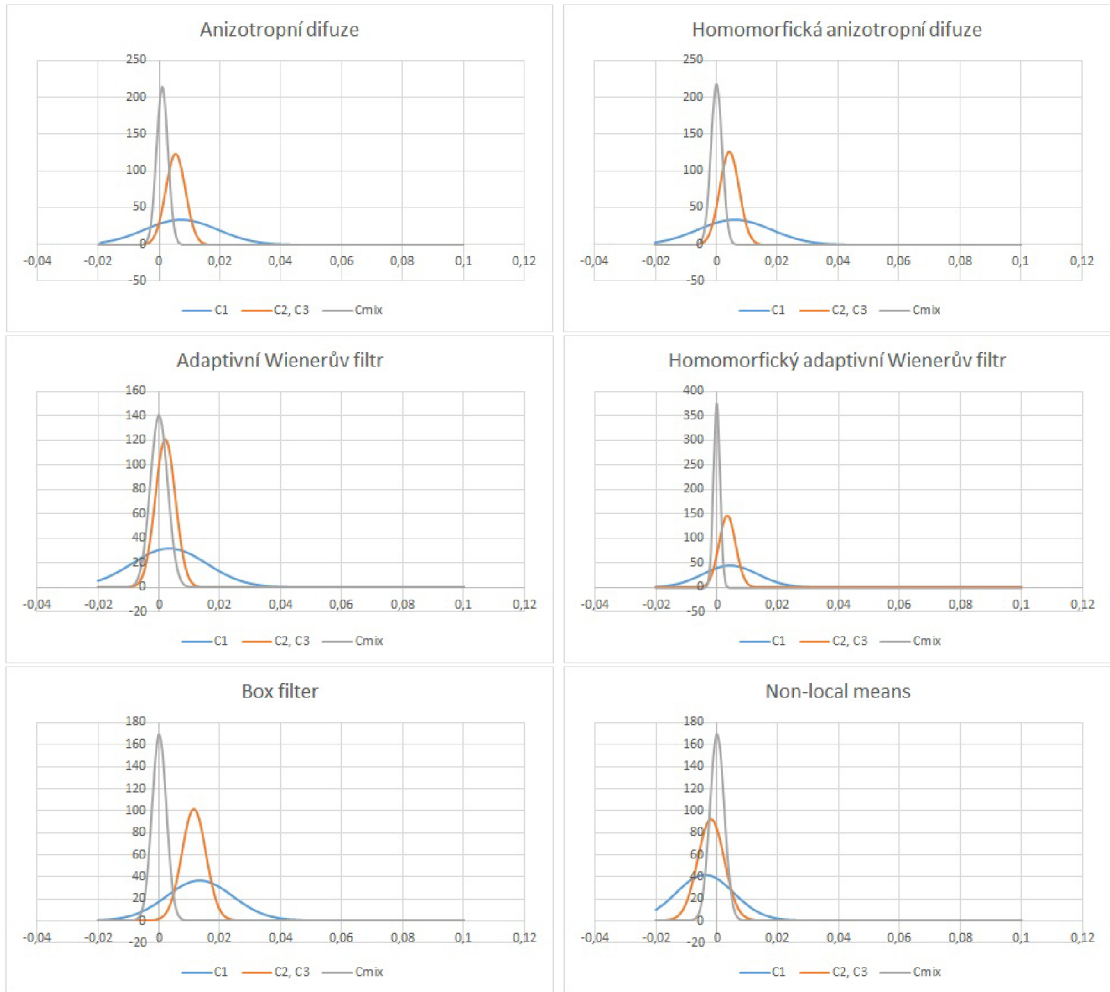
Název metody	FN [%]	Práh
Anizotropní difuze	9,37	0,01
Homomorfická anizotropní difuze	9,37	0,01
Adaptivní Wienerův filtr	9,49	0,0014
Homomorfický adaptivní Wienerův filtr	12,54	0,009
Box filter	17,82	0,023
Non-local means	65,62	0,014

Tabulka 7.4: Vyhodnocení č. 2,  $FP \leq 1\%$

Non-local means znovu dosahuje nejhorších výsledků. Oproti předchozímu měření si pohoršil box filter, kde je potřeba přesunout práh.

Název sady	Označení	Použití	Počet snímků
Panasonic video #1, papír	—	extrakce RPN	58
Panasonic video #1, náhodné snímky 2	$C_1$	testování	82
Panasonic video #2, náhodné snímky 2	$C_2$	testování	95
Panasonic video #3, náhodné snímky 2	$C_3$	testování	57
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.5: Použité sady pro vyhodnocení č. 3



Obrázek 7.4: Normální rozdělení pro obrazový senzor Panasonic video #1

U předchozích dvou měření by šel hezky odhadnout práh  $t$ , který by poměrně přesně dokázal obrazový senzor identifikovat. Další měření, tentokrát pro Panasonic video #1, kde referenční PRNU bylo získáno z rovnoměrně osvětleného papíru, ovšem vykazuje větší chybovost. V tabulce 7.5 jsou poznačeny použité datové sady pro tento experiment. Na obrázku 7.4 jsou zobrazeny funkce normálního rozdělení pro naměřené korelační koeficienty. V tabulce 7.6 jsou výsledky.

Název metody	FN [%]	Práh
Anizotropní difuze	74,29	0,014
Homomorfická anizotropní difuze	74,72	0,014
Adaptivní Wienerův filtr	78,2	0,013
Homomorfický adaptivní Wienerův filtr	82,58	0,013
Box filter	85,65	0,024
Non-local means	96,34	0,013

Tabulka 7.6: Vyhodnocení č. 3,  $FP \leq 1\%$

Stále tedy ze všech použitých metod vychází nejlépe sekvenční a homomorfická anizotropní difuze, ovšem u třídy kamer Panasonic video je přesnost o poznání nižší. Pokud by se položila hypotéza pro vyhodnocení: „**Je testovací snímek I ze třídy senzorů C?**“, tak by šla identifikovat alespoň třída senzorů. To by se dalo využít například ve vyhledávání zdrojové značky či druhu digitální kamery. Výsledky jsou v tabulce 7.7

Název metody	FN [%]	Práh
Anizotropní difuze	54,47	0,008
Homomorfická anizotropní difuze	65,75	0,008
Adaptivní Wienerův filtr	77,98	0,01
Homomorfický adaptivní Wienerův filtr	40,43	0,004
Box filter	23,37	0,004
Non-local means	93,89	0,01

Tabulka 7.7: Vyhodnocení č. 3b,  $FP \leq 1\%$

**Závěr:** Rozhodně je nepoužitelné odšumění pomocí non-local means. Ani v jednom případě nejde správně rozhodnout, z jakého obrazového senzoru testovaná fotografie pochází. Nejlepší výsledky produkuje anizotropní difuze, a to jak pro sekvenční, tak i pro homomorfický přístup. Dolní propust nedokázala od sebe odlišit jednotlivé senzory v rámci třídy Panasonic video, ale od senzorů jiného typu to dokáže. Adaptivní Wienerův filtr má oproti jiným filtrům větší rozsah korelačních koeficientů a horší výsledky. Homomorfický přístup ve většině případů umožňuje oproti jiným filtrům zmenšit hodnotu prahu pro rozhodnutí, ale má horší výsledky oproti sekvenčnímu přístupu. Rozsah korelačních koeficientů má výrazně menší oproti jiným metodám.

Pro rozhodnutí, **jaká by měla být velikost prahu**, nelze s jistotou odpovědět nikdy. Práh se vždy pohyboval v intervalu  $(0,008 - 0,014)$ . Záleží na konkrétním použití navrženého algoritmu a na konkrétním odšumovacím filtru.

Na otázku, jestli dosahuje homomorfický přístup podobných kvalit jako sekvenční přístup, lze odpovědět. U anizotropní difuze se potvrdilo, že homomorfický přístup je nepatrně horší pro extrakci referenčního PRNU. Lze tedy, pokud to okolnosti dovolují, urychlit extrakci referenčního PRNU homomorfickým přístupem, například při extrakci referenčních PRNU z mnoha kamer či velkého množství datových sad. Pro případ, že je ale cílem mít co nejvyšší přesnost a není uživatel tlačěn časem, je přesnější použít sekvenční přístup.

Homomorfická filtrace jde s výhodou použít například při extrakci PRNU z videí. Může se stát například, že vyšetřovatel dostane adresář plný různých videonahrávek, které mohou být z mnoha různých digitálních kamer a bude se snažit přiřadit důkazní materiál ke konkrétní videonahrávce. Je tedy vhodná pro rychlou klasifikaci.

Pro odpověď na původní otázku „*Který odšumovací filtr produkuje nejlepší možné výsledky?*“, lze odpovědět podle výše prezentovaných poznatků: „*Filtr typu anizotropní difuze, protože ve většině případů má oproti jiným filtrům nejmenší chybu.*“

## 7.2 Vliv sady snímků na kvalitu referenčního vzoru šumu

Pro tento experiment otázka zní: „*Lze použít pro extrakci referenčního PRNU sadu náhodných snímků? Nebo je nutné použít speciálně vytvořenou sadu snímků?*“

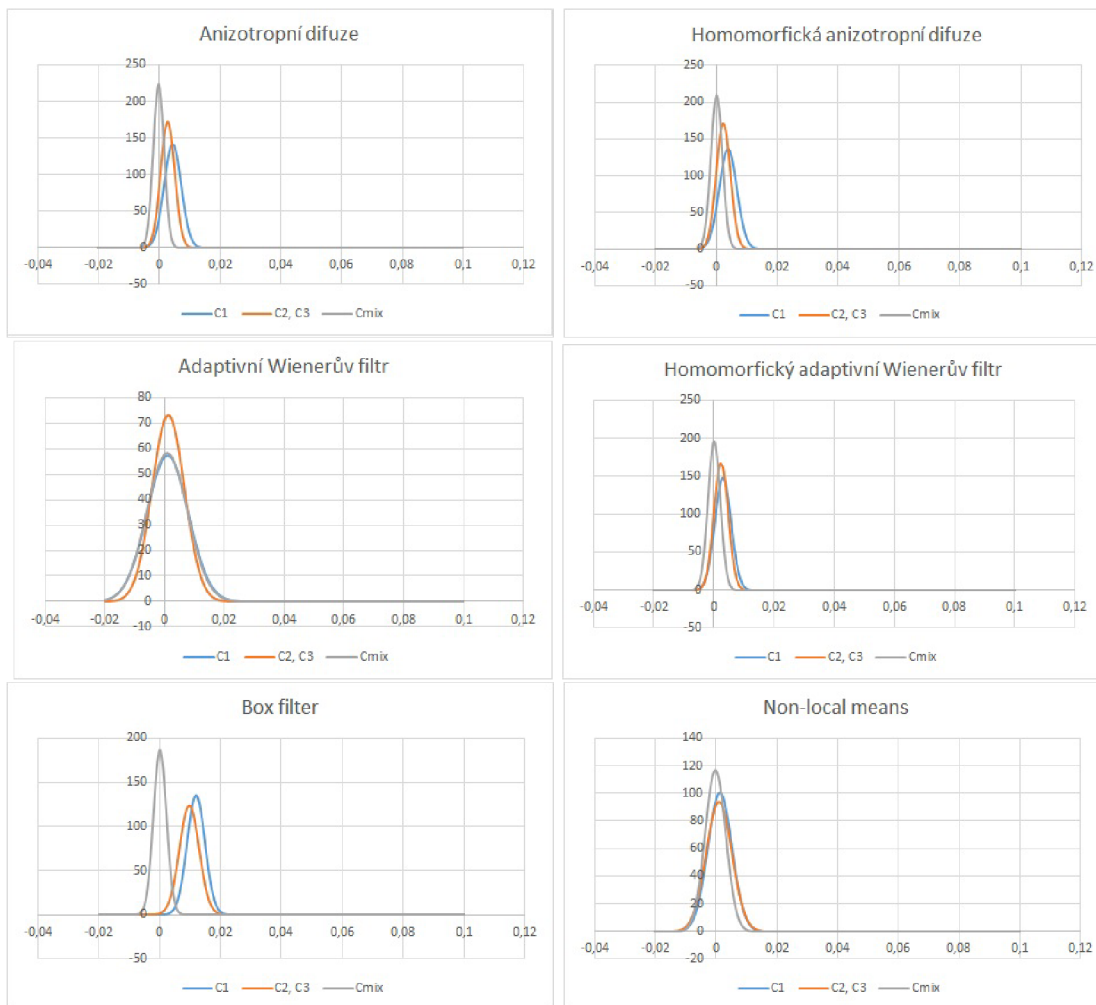
Většina literárních zdrojů, která se touto problematikou zabývá, doporučovala pro co nejlepší možné výsledky provádět extrakci PRNU nad rovnoměrně osvětlenými plochami.

Takové snímky nejsou zatíženy pravidelnými texturami, hranami a dalšími objekty, které při extrakci zanášejí do výsledného PRNU viditelné artefakty.

V některých případech ale není možné takovou datovou sadu ze senzoru vytvořit. Mějme hypotetickou fotografii, na které je zachycen důkazní materiál. Předpokládejme, že u pachatele na pevném disku byl nalezen adresář různých snímků, napovídající, že se jedná o data pořízená z jednoho senzoru. Tyto snímky si jsou podobné s hypotetickou fotografií, zachycující trestní činnost. Fotoaparát byl ale bohužel ztracen, nelze tedy vytvářet další snímky. Pro tento příklad si lze položit otázku, uvedenou v úvodu.

Název sady	Označení	Použití	Počet snímků
Panasonic video #1, náhodné snímky 1	—	extrakce RPN	56
Panasonic video #1, náhodné snímky 2	$C_1$	testování	82
Panasonic video #2, náhodné snímky 2	$C_2$	testování	95
Panasonic video #3, náhodné snímky 2	$C_3$	testování	57
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.8: Použité sady pro vyhodnocení č. 4



Obrázek 7.5: Normální rozdělení pro obrazový senzor Panasonic video #1

K prvnímu experimentu bylo vybráno zařízení Panasonic video #1. Pro extrakci referenčního PRNU byla zvolena datová sada obsahující běžné snímky, tedy Panasonic video #1, náhodné snímky 1. Pro testování byly zvoleny datové sady uvedené v tabulce 7.8. Žádný snímek z datové sady, určené pro extrakci, nebyl v testovacích sadách. Na obrázku 7.5 jsou zobrazeny Gaussovy křivky pro vypočtené parametry určené vyhodnocenými korelačními koeficienty.

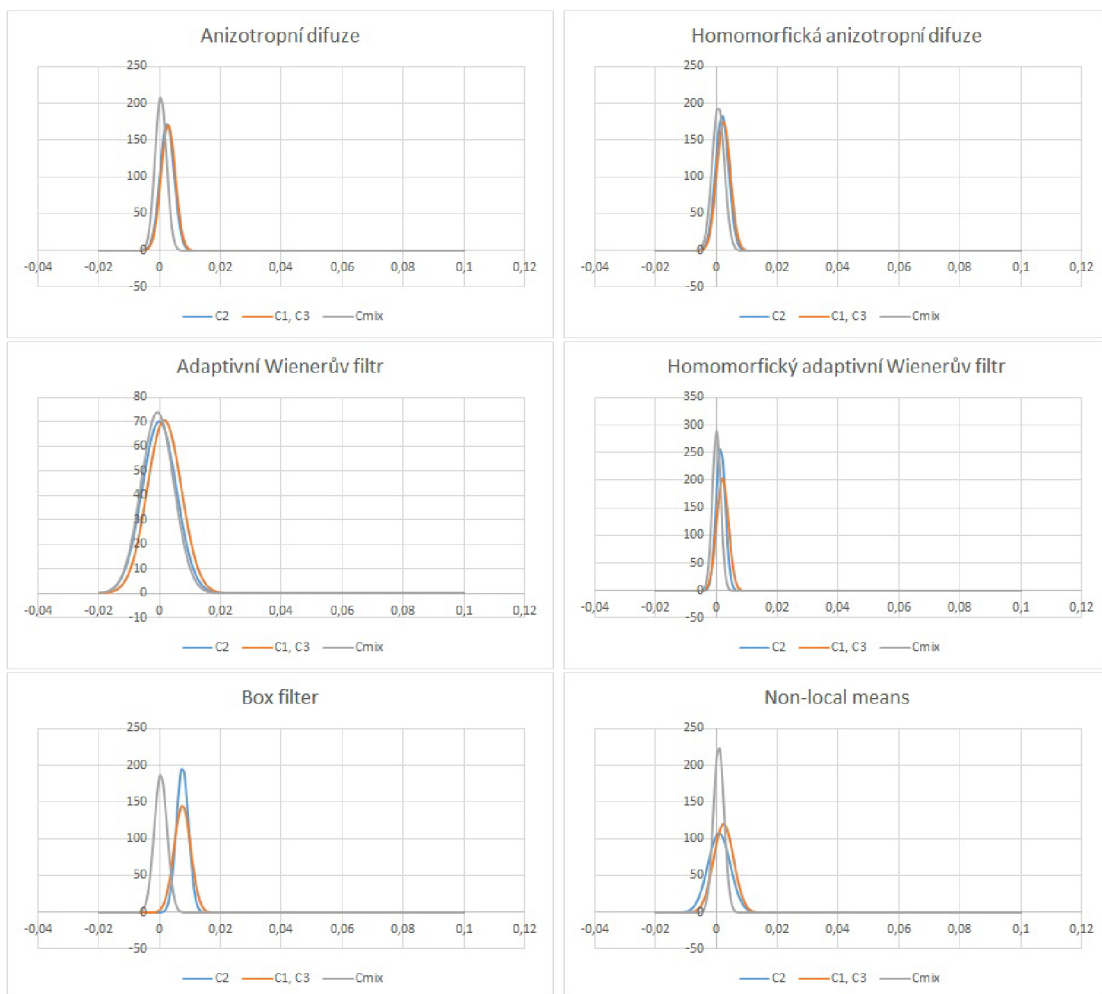
Název metody	FN [%]	Práh
Anizotropní difuze	90,82	0,008
Homomorfická anizotropní difuze	98,19	0,007
Box filter	20,19	0,007

Tabulka 7.9: Vyhodnocení č. 4,  $FP \leq 1\%$ , pro třídu senzorů  $C_1$ - $C_3$  vůči  $C_{mix}$

Jak je vidět z grafů 7.5, překryvy jednotlivých senzorů jsou natolik velké, že nelze ověřovat pravost dat, chyba 2. typu je příliš velká. V tabulce 7.9 jsou uvedeny naměřené hodnoty. Box filtr dokáže rozlišit třídu senzorů Panasonic video od jiných senzorů, s  $FN = 20,19\%$ .

Název sady	Označení	Použití	Počet snímků
Panasonic video #2, náhodné snímky 1	—	extrakce RPN	47
Panasonic video #1, náhodné snímky 2	$C_1$	testování	82
Panasonic video #2, náhodné snímky 2	$C_2$	testování	95
Panasonic video #3, náhodné snímky 2	$C_3$	testování	57
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.10: Použité sady pro vyhodnocení č. 5

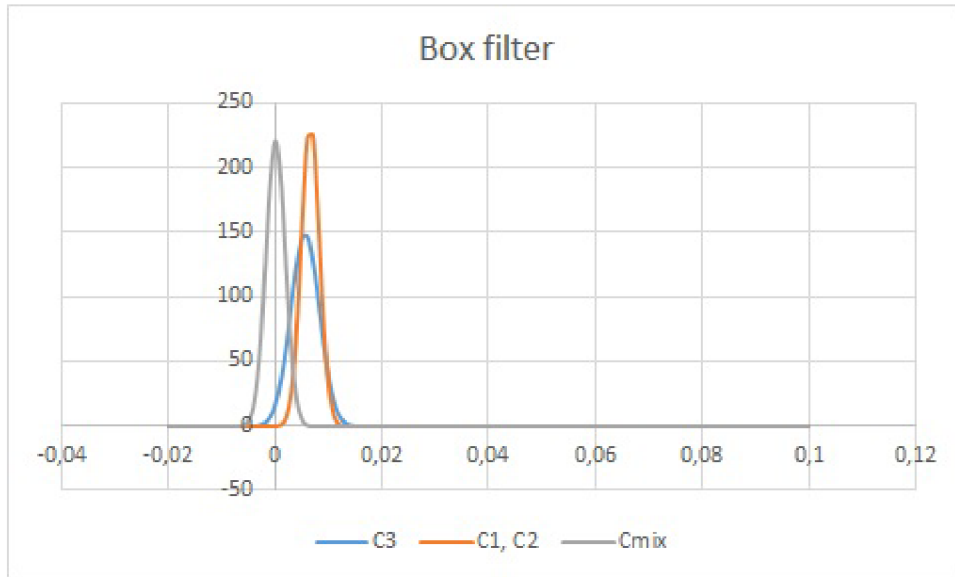


Obrázek 7.6: Normální rozdělení pro obrazový senzor Panasonic video #2

Další obrazový senzor, pro který byl tento experiment prováděn, je Panasonic video #2. Použité datové sady jsou uvedeny v tabulce 7.10. Na obrázku 7.6 jsou znázorněny odpovídající Gaussovy křivky. I zde jsou výsledky velmi podobné, jako u obrazového senzoru Panasonic video #1. Kromě box filtru je chyba druhého typu vždy přes 98 %, a to mezi libovольnými testovanými obrazovými senzory.

V případě box filtru, zde lze alespoň odhalit příslušnou třídu senzorů Panasonic video #2. Pro zvolený práh  $t = 0,008$  je chyba typu FN = 33,91 %.

Na obrázku 7.7 je zobrazen stejný graf pro odšumění dolní propustí pro kameru Panasonic video #3, kde pro extrakci referenčního PRNU byla použita datová sada Panasonic video #3, náhodné snímky 1. Pro zvolený práh  $t = 0,006$  je chyba typu FN = 30,71 %. Též nelze v tomto případě rozlišit obrazové senzory v rámci třídy Panasonic video, ale tuto třídu lze odlišit od ostatních obrazových senzorů.



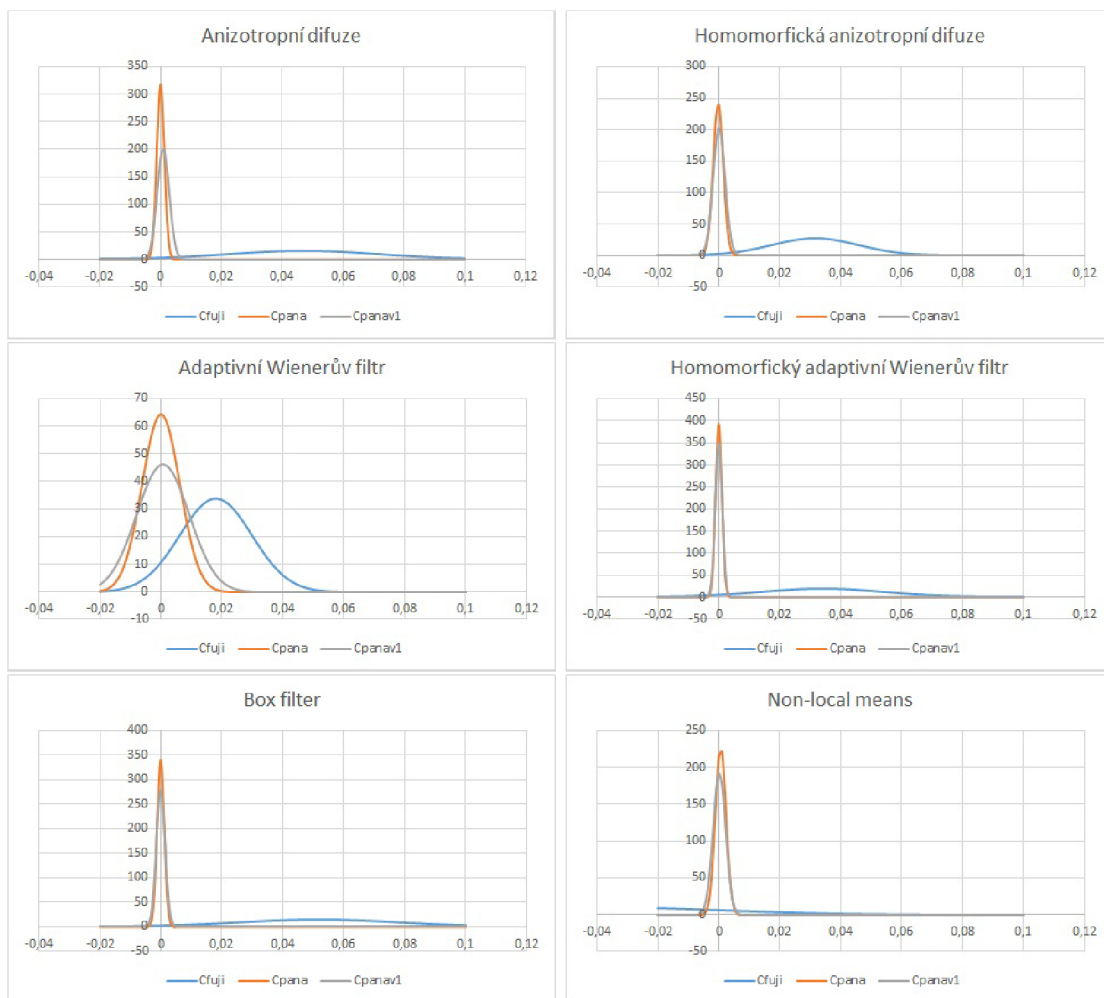
Obrázek 7.7: Normální rozdělení pro obrazový senzor Panasonic video #3, filtr dolní propust

Název sady	Označení	Použití	Počet snímků
Fujifilm, náhodné snímky 1	—	extrakce RPN	50
Fujifilm, oblaka	$C_{fuji}$	testování	66
Fujifilm, náhodné snímky 2	$C_{fuji}$	testování	29
Panasonic, náhodné snímky	$C_{pana}$	testování	44
Panasonic video #1, náhodné snímky 1	$C_{panav1}$	testování	56

Tabulka 7.11: Použité sady pro vyhodnocení č. 6

I když to vypadá, že extrahované PRNU z datových sad, obsahujících běžné snímky, je extrémně chybové, tak není tomu tak ve všech případech. Dalším vyhodnoceným obrazovým senzorem je Fujifilm, kde pro extrakci referenčního PRNU byla použita datová sada Fujifilm, náhodné snímky 1. Testovací datové sady jsou vypsány v tabulce 7.11. v tomto experimentu byly vybrány obrazové senzory podle velmi podobných vlastností. Fujifilm a Panasonic jsou obrazové senzory typu CCD a Panasonic video #1 má stejný formát obrazového senzoru jako Fujifilm.





Obrázek 7.8: Normální rozdělení pro obrazový senzor Fujifilm

Název metody	FN [%]	Práh
Anizotropní difuze	2,57	0,008
Homomorfická anizotropní difuze	3,97	0,007
Adaptivní Wienerův filtr	75,77	0,025
Homomorfický adaptivní Wienerův filtr	7,4	0,005
Box filter	3,53	0,005
Non-local means	87,94	0,008

Tabulka 7.12: Vyhodnocení č. 6, FP  $\leq 1$  %

V experimentu se znovu nejlépe umístila anizotropní difuze, jak uvádí tabulka 7.12. Referenční PRNU, získané z běžných snímků, lze v tomto případě spolehlivě použít pro identifikaci obrazového senzoru. Protože se výsledky zdály až podezřele dobré, rozhodl jsem se z vyhodnocení odebrat datovou sadu Fujifilm, oblaka. Výsledné grafy jsou na obrázku 7.9 v levém sloupci.

Název sady	Označení	Použití	Počet snímků
Fujifilm, jasná obloha	—	extrakce RPN	34
Fujifilm, náhodné snímky 1	$C_{fuji}$	testování	50
Fujifilm, náhodné snímky 2	$C_{fuji}$	testování	29
Panasonic, náhodné snímky	$C_{pana}$	testování	44
Panasonic video #1, náhodné snímky 1	$C_{panav1}$	testování	56

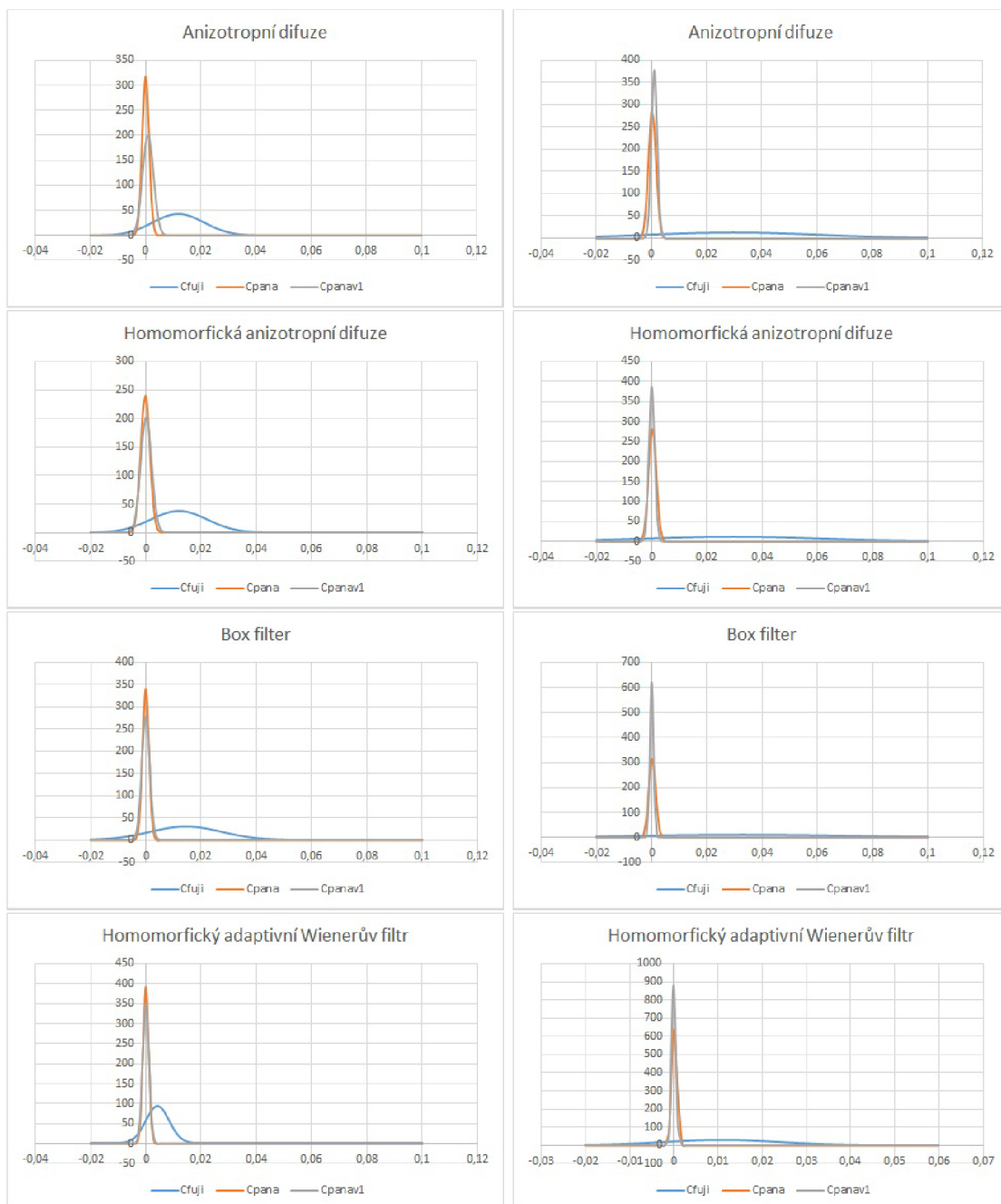
Tabulka 7.13: Použité sady pro vyhodnocení č. 7

Dále byly vypočítány korelační koeficienty pro další datové sady. Referenční PRNU bylo extrahováno pro digitální kameru Panasonic, tentokrát z rovnoměrně osvětlených snímků, z datové sady Panasonic, jasná obloha. Použité datové sady pro testování jsou uvedeny v tabulce 7.13. Vypočítané funkce normálního rozdělení jsou prezentovány na obrázku 7.9, tentokrát v pravém sloupci.

Název metody	FN [%]	Práh	FN FF [%]	Práh FF
Anizotropní difuze	33,51	0,008	21,49	0,006
Homomorfická anizotropní difuze	34,59	0,008	21,49	0,006
Homomorfický adaptivní WF	48,25	0,003	21,41	0,003
Box filter	26,71	0,007	22,52	0,007

Tabulka 7.14: Vyhodnocení č. 7,  $FP \leq 1 \%$

V tabulce 7.14 jsou uvedeny naměřené výsledky. Sloupce FF značí použité rovnoměrně osvětlené snímky pro extrakci referenčního PRNU, v ostatních případech jsou použity náhodné snímky.



Obrázek 7.9: Normální rozdělení pro obrazový senzor Fujifilm, vlevo PRNU z běžných snímků, vpravo z rovnoměrně osvětlených snímků

Experimentálně bylo ověřeno a poté i vyhodnoceno, že PRNU, extrahované z množiny běžných snímků, získaných z jednoho obrazového senzoru, produkuje horší výsledky. U měření č. 6, kde do testovací množiny snímků pro senzor Fujifilm ( $C_{fujii}$ ) byla zahrnuta datová sada Fujifilm, oblaka, výsledky zkreslila. Důvodem je fakt, že množina snímků v datové sadě obsahuje příliš podobné, avšak ne stejné snímky.

U třídy senzorů Panasonic video nešlo odlišit jednotlivé obrazové senzory od sebe s výjimkou použití filtru dolní propusti. Zde se projevilo alespoň odlišení různých obrazových

senzorů mimo stejnou třídu senzorů. Rozlišitelnost v rámci této třídy byla horší i pro rovnoměrně osvětlené snímky.

Na otázku z úvodu „Lze použít pro extrakci referenčního PRNU sadu náhodných snímků? Nebo je nutné použít speciálně vytvořenou sadu snímků?“, lze tedy odpovědět „Ano, referenční PRNU lze pořídít i z náhodných snímků, ale ze specificky vytvořených snímků lze dosáhnout menší chyby“.

### 7.3 Velikost referenčního vzoru šumu

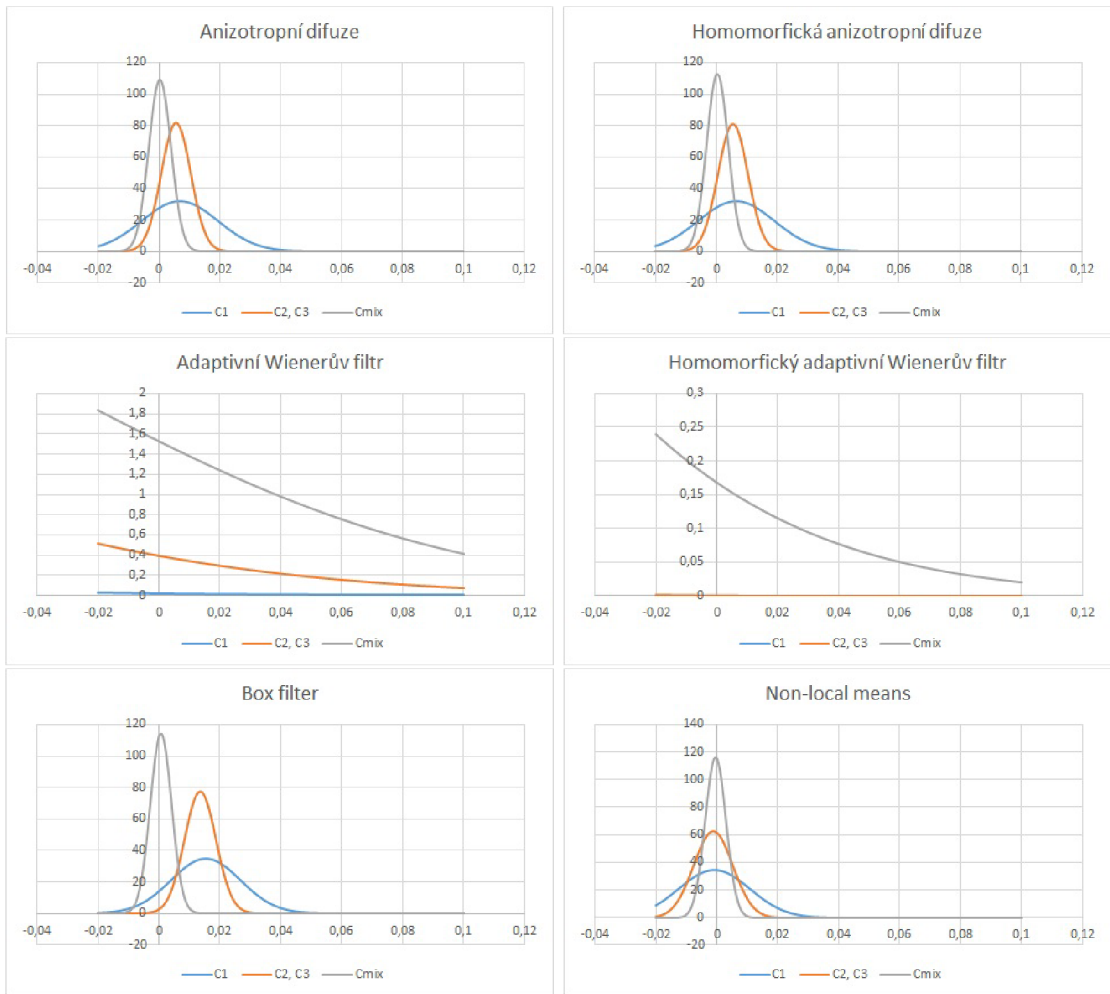
V této sekci se hledá odpověď na otázku: „Je lepší použít velikost referenčního šumu  $512 \times 512$ ,  $1024 \times 1024$ , nebo nejvyšší možné rozlišení daného obrazového senzoru?“

Jak již bylo v textu dříve avizováno, maximální velikost referenčního PRNU byla stanovena na  $1024 \times 1024$ . Otázkou ale je, jestli je to dostatečná velikost, případně jestli jde tuto plochu ještě zmenšit. Na tuto problematiku bylo k vyhodnocení zvoleno několik datových sad, u kterých byly výsledky horší, než u jiných sad. Důvodem je zkusit sledovat, jestli dojde ke zlepšení či nikoliv.

Intuice říká, že čím více informací je k dispozici, tím přesnější budou následně výsledky. Nicméně aplikování odšumovacího filtru je časově náročný proces a čím větší jsou snímky, tím delší čas je potřebný k jejich zpracování. Proto byl vytvořen experiment, kde byl vyhodnocen vliv velikosti PRNU pro přesnost identifikace.

Název sady	Označení	Použití	Počet snímků
Panasonic video #1, papír	—	extrakce RPN	58
Panasonic video #1, náhodné snímky 2	$C_1$	testování	82
Panasonic video #2, náhodné snímky 2	$C_2$	testování	95
Panasonic video #3, náhodné snímky 2	$C_3$	testování	57
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.15: Použité sady pro vyhodnocení č. 8



Obrázek 7.10: Normální rozdělení pro obrazový senzor Panasonic video #1

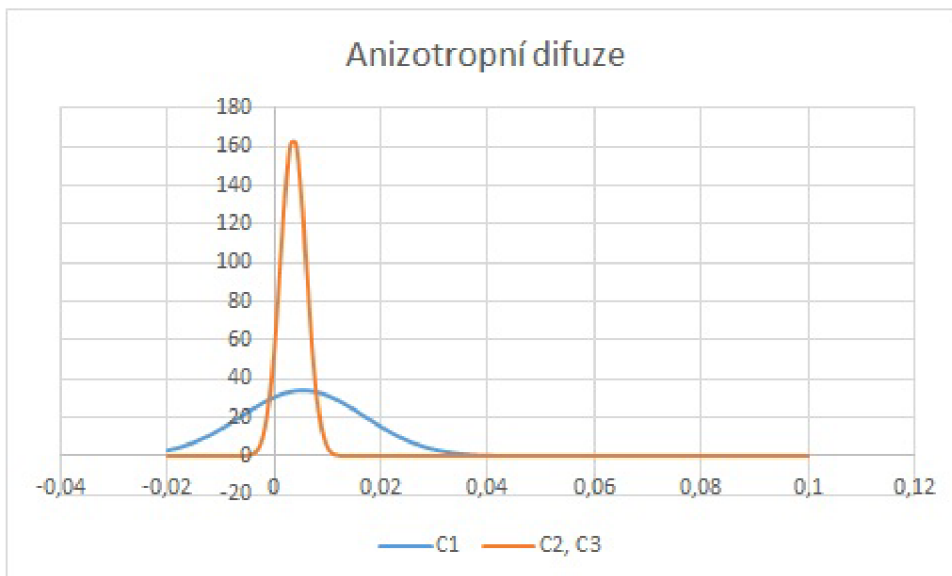
Nejprve byla otestována maximální velikost PRNU pro velikost  $512 \times 512$  pixelů. Použité datové sady jsou uvedeny v tabulce 7.15. Pro extrakci PRNU byly použity rovnoměrně osvětlené snímky. Výsledné charakteristiky Gaussových křivek jsou zobrazeny na obrázku 7.10.

Název metody	FN 512 [%]	Práh 512	FN 1024 [%]	Práh 1024
Anizotropní difuze	87,13	0,021	74,29	0,014
Box filter (mimo třídu)	29,41	0,013	23,31	0,004

Tabulka 7.16: Vyhodnocení č. 8,  $FP \leq 1\%$

V tabulce 7.16 jsou uvedeny naměřené výsledky. Pro box filter nešlo rozlišit jednotlivé kamery v rámci třídy a jsou uvedeny výsledky pro třídu vůči ostatním sensorům. Větší velikost PRNU produkuje lepší výsledky.

V případě použití plné velikosti snímku pro extrakci PRNU byla vyhodnocena pouze anizotropní difuze, a to z důvodu časové náročnosti výpočtu korelačních koeficientů jednotlivých snímků. Též byl proveden výpočet korelačních koeficientů pouze v rámci třídy senzorů Panasonic video.



Obrázek 7.11: Normální rozdělení pro obrazový senzor Panasonic video #1

V tomto případě je pro zvolený práh  $t = 0,013$  chyba typu FN = 77,88 %. Velikost PRNU je  $4992 \times 2808$ . Oproti velikosti  $1024 \times 1024$  tedy došlo ke zhoršení, ovšem oproti velikosti  $512 \times 512$  došlo ke zlepšení. Důvodem může být vliv optické soustavy a dalších součástí digitálních kamer (například závěrky apod.) při akvizici obrazu, kde dochází k vinětaci na okrajích snímků.

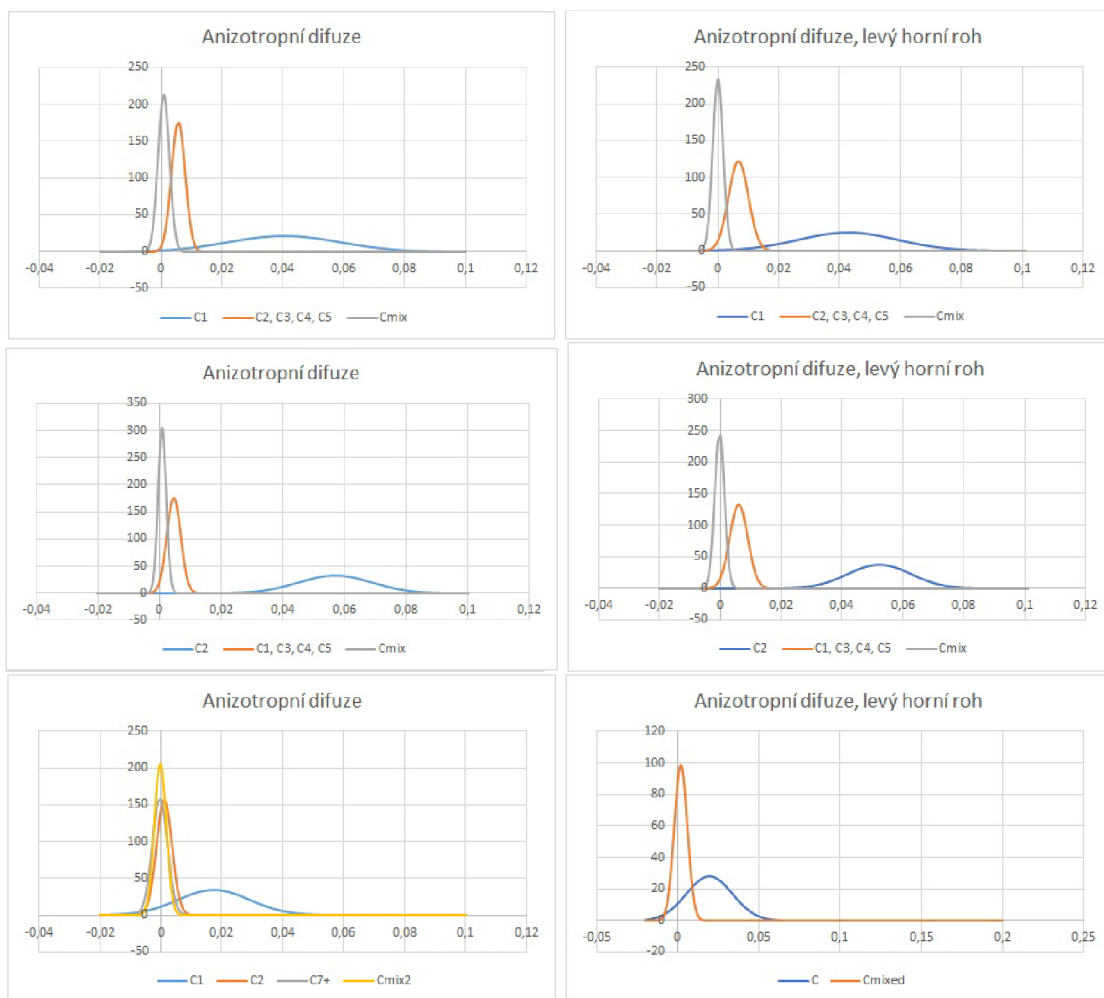
Pro odpovězení na otázku: „Je lepší použít velikost referenčního šumu  $512 \times 512$ ,  $1024 \times 1024$ , nebo nejvyšší možné rozlišení daného obrazového senzoru?“, je na základě výše ověřených vlastností odpověď: „Nejmenší chybu má překvapivě rozlišení  $1024 \times 1024$ “.

## 7.4 Lokalita extrakce referenčního šumu

Mějme otázku: „Je lepší pro extrakci referenčního šumu použít oblast ve středu snímku nebo je lepší použít levý horní roh?“. Posledním experimentem je ověření, jaký vliv má v pořízeném snímku lokalita pro extrakci referenčního PRNU. Pro velikost referenčního PRNU  $1024 \times 1024$  byly vybrány dvě lokality, a to střed snímku levý horní okraj snímku.

Název sady	Označení	Použití	Počet snímků
Samsung #1, stěna	—	extrakce RPN #1	83
Samsung #2, stěna	—	extrakce RPN #2	86
Samsung #1, náhodné snímky	$C_1$	testování	43
Samsung #2, náhodné snímky	$C_2$	testování	61
Samsung #3, náhodné snímky	$C_3$	testování	38
Samsung #4, náhodné snímky	$C_4$	testování	43
Samsung #5, náhodné snímky	$C_5$	testování	87
Mix jiných senzorů #1	$C_{mix}$	testování	60

Tabulka 7.17: Použité sady pro vyhodnocení č. 9



Obrázek 7.12: Normální rozdělení pro obrazové senzory Samsung #1, Samsung #2 a iPhone 7 #1

V tabulce 7.17 jsou uvedeny datové sady, které byly použity při tomto experimentu pro extrakci a testování. Pro extrakci referenčního PRNU byly použity sady snímků, obsahující rovnoměrně osvětlené stěny. Výsledky jsou uvedeny na obrázku 7.12, kde v levém sloupci jsou extrahované PRNU ze středu snímku a v pravém sloupci jsou extrahované PRNU z levého horního rohu. Na prvním řádku je Samsung #1, uprostřed je Samsung #2 a na posledním řádku je zařízení iPhone 7 #1. Vyhodnocení se provedlo pouze pro anizotropní difuzi.

Pro Samsung #1, střed je práh  $t = 0,014$  a odpovídající chyba typu FN = 9,08 %, kde  $FP \leq 1\%$ . Pro levý horní roh je hodnota prahu  $t = 0,017$  a odpovídající chyba typu FN = 5,44 %.

Pro Samsung #2 byly naměřeny chyby druhého typu o velikosti  $FP = 0,002$ , a to jak pro levý horní roh ( $t = 0,015$ ), tak pro střed ( $0,012$ ).

V obou případech šlo rozlišit jednotlivá zařízení od sebe v rámci třídy Samsung. Ovšem pro případy levého horního rohu jsou rozptýly jednotlivých normálních rozdělení větší oproti středům. Též jsou střední hodnoty Gaussových křivek jiných zařízení než  $C_1$  u levého hor-

ního rohu větší, což značí, že vypočítané korelační koeficienty u snímků jsou vyšší než u středů. Znamená to, že snímky z  $C_2 - C_5$  si jsou více podobné se snímky z  $C_1$ .

Název sady	Označení	Použití	Počet snímků
iPhone 7 #1, jasná obloha	—	extrakce RPN	73
iPhone 7 #1, náhodné snímky	$C_1/C$	testování	50
iPhone 7 #2, náhodné snímky	$C_2$	testování	55
iPhone 7+, náhodné snímky	$C_{7+}$	testování	146
Mix jiných senzorů #2	$C_{mix}$	testování	60
Mix jiných senzorů #3	$C_{mixed}$	testování	295

Tabulka 7.18: Použité sady pro vyhodnocení č. 10

V tabulce 7.18 jsou uvedeny datové sady pro poslední vyhodnocení. Na obrázku 7.12 jsou na posledním řádku odpovídající grafy pro zařízení iPhone 7 #1.

Pro iPhone 7 #1, kde se volí střed, je pro zvolený práh  $t = 0,009$  odpovídající chyba typu FN = 22,39 %. Pro levý horní roh, pro zvolený práh  $t = 0,015$ , je odpovídající chyba typu FN = 37,38 %. Zde tedy došlo ke zhoršení. Důležité je ale zmínit, že nejsou testovány stejné datové sady.

Zjištění tedy je, že výsledky jsou lepší, pokud se referenční PRNU získává ze středu snímku. Na okrajích obrazového senzoru se totiž projevují různé vady (vinětace apod.), které výslednou identifikaci zhoršují. I když z principu podstaty extrakce čistě PRNU by to vadit nemělo a plocha celého senzoru by měla dávat co nejlepší výsledky, díky vlivu těchto vad není doporučeno provádět extrakci z okrajových částí obrazového senzoru.

Odpověď na otázku položenou v úvodu, tedy: „*Je lepší pro extrakci referenčního šumu použít oblast ve středu snímku nebo je lepší použít levý horní roh?*“, je: „*Střed snímku je lepší.*“

## 7.5 Praktické vyhodnocení hypotéz

V této závěrečné podkapitole jsou uvedena konkrétní možná ověření pravosti dat na naměřených datech.

Mějme modelový příklad č. 1:

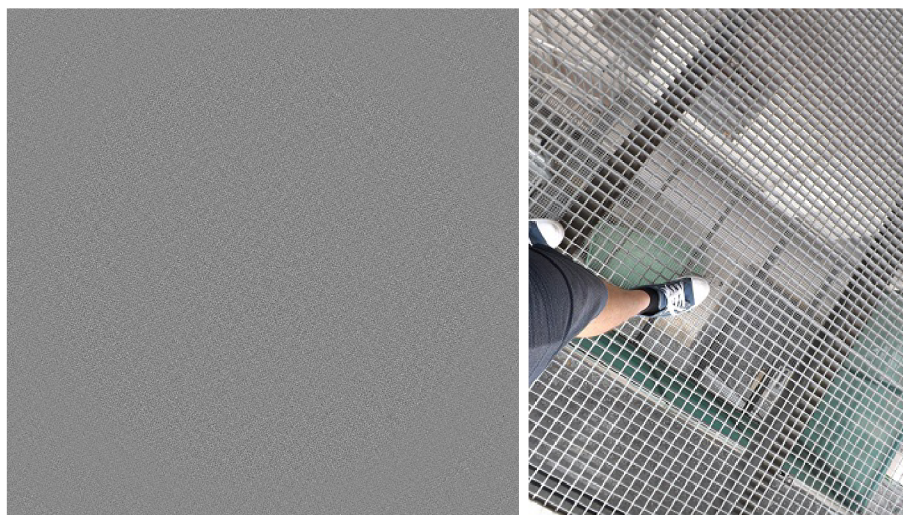
*Pachatel má zařízení Fujifilm SL-280. Pachatel na toto zařízení fotografoval snímky, které jsou trestnou činností. Tyto fotografie poté v plné kvalitě prodával na černém trhu v Darknetu. Vyšetřovatelé takové snímky od tohoto uživatele získaly a povedlo se jim zjistit, kdo je podezřelou osobou. Vyšetřovatelé vtrhli do bytu podezřelého, a pachatele zadrželi. Nalezli notebook a zmíněnou kameru. Tyto předměty zabavili pro další vyšetřování.*

*K jejich smůle, v notebooku nenašli žádnou inkriminovanou fotografii, stejně tak na paměťové kartě fotoaparátu. Důležitým prvkem pro vyšetřování se tak stává ověření, jestli kopie vytvořeného snímku, získaného na Darknetu, je ze zabaveného fotoaparátu.*

*Vyšetřovatelé na této kameře vytvořili sadu rovnoměrně osvětlených snímků oblohy. Tu použili pro extrakci otisku kamery. Následně zkusili ověřit, jestli je inkriminovaná fotografie pořízena z této kamery. Potřebují mít jistotu 99 %. Na obrázku 7.13 je tento šum, spolu s inkriminovanou fotografií zobrazen.*

Hypotéza tedy zní:  $H_0$ : Fotografie  $I$  je pořízena z kamery  $C$ .





Obrázek 7.13: Šum získaný ze snímků oblohy a inkriminovaná fotografie

Filtrace	Rotace [°]	Korelační koeficient
Anizotropní difuze	0	0,00288
Anizotropní difuze	90	0,00130
Anizotropní difuze	270	0,05344
Anizotropní difuze	180	-0,00087

Tabulka 7.19: Naměřené korelační koeficienty

Vyšetřovatelé zjistili, že tato fotografie je z daného obrazového senzoru skutečně exponována. Zároveň zjistili, že výsledný snímek byl otočen doleva o 90°. Naměřili korelační koeficienty, uvedené v tabulce 7.19.

Jak se k takovému rozhodnutí došlo? V tabulce 7.14 jsou uvedeny pro tuto kameru prahy a chyby druhého typu. V případě snímku  $I$ , pro jeho rotaci 270° je  $H_0$  vyhodnocena jako pravdivá, protože práh  $t$  je menší než naměřený korelační koeficient.

Odpověď na  $H_0$  tedy zní: *Ano*.

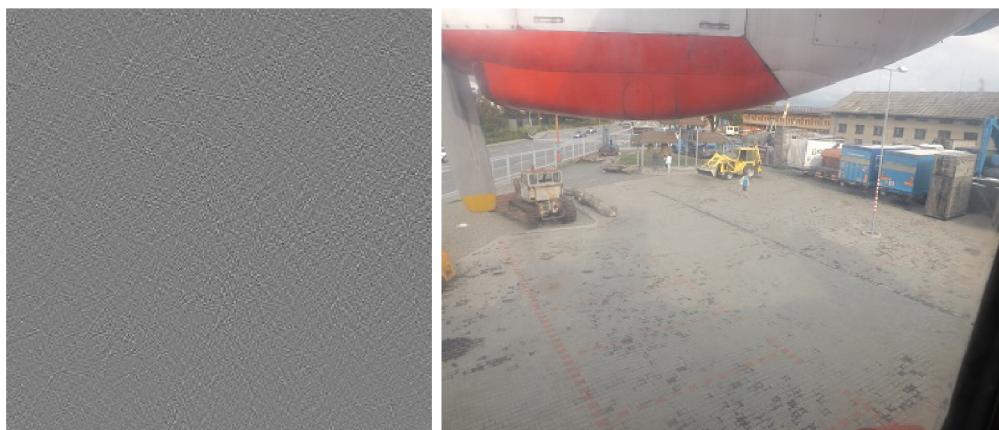
Mějme příklad č. 2:

*Pachatel má zařízení Fujifilm SL-280. Pachatel na toto zařízení fotografoval snímky, které jsou trestnou činností. Zároveň na toto zařízení fotil i různé jiné snímky. Pachatel inkriminované snímky prodával na černém trhu na Darknetu.*

*Vyšetřovatelé pachatele odhalili, vtrhli do jeho bytu, zadrželi jej, ale nenašli žádnou kameru, ani fotoaparát. Nalezli pouze DVD s nápisem „Dovolená 2019“. Na tomto disku našli několik různých snímků. Důležitým prvkem pro vyšetřování se tak stává ověření, jestli kopie vytvořeného snímku, získaného na Darknetu, je ze stejné kamery, jako snímky na DVD.*

*Vyšetřovatelé tak extrahovali otisk kamery z množiny snímků, uložených na DVD. Následně zkusili ověřit, jestli je inkriminovaná fotografie pořízena ze stejné kamery. Potřebují mít jistotu 99 %. Na obrázku 7.14 je tento šum, spolu s inkriminovanou fotografií zobrazen.*

Hypotéza tedy zní:  $H_0$ : Fotografie  $I$  je pořízena z kamery  $C$ .



Obrázek 7.14: Šum získaný z množiny běžných snímků a inkriminovaná fotografie

<b>Filtrace</b>	<b>Rotace [°]</b>	<b>Korelační koeficient</b>
Anizotropní difuze	0	0,02859
Anizotropní difuze	90	-0,00040
Anizotropní difuze	270	-0,00085
Anizotropní difuze	180	0,00249

Tabulka 7.20: Naměřené korelační koeficienty

Vyšetřovatelé naměřili korelační koeficienty, uvedené v tabulce 7.20. Odhalili velikost prahu (uvedeny v tabulce 7.14) a zjistili, že inkriminovaná fotografie je pořízena ze stejného zařízení, jako snímky uložené na obsahu DVD.

Mějme **příklad č. 3**. Uživatel má několik mobilních telefonů Samsung A20. Ty se liší použitou barvu těla zařízení. Uživatel si na každém mobilu vyfotil několik fotek. Uživatel je záložoval do adresáře na pevném disku firemního počítače, ale fotky se pomíchaly. Bohužel si nepamatuje, ze kterého zařízení jaké fotky záložoval. Uživatel potřebuje z adresáře zjistit, jaké fotky jsou z mobilního telefonu, který měl červený kryt.

Na mobilním telefonu Samsung A20, který má červený kryt, uživatel pořídil sadu rovnoměrně osvětlené stěny. Tyto fotky použil pro extrakci otisku kamery. Dále zkusil s tímto otiskem otestovat adresář se zamíchanými fotkami. Naměřené korelační koeficienty jsou uvedeny v tabulce 7.21.

Filtrace a ID	Rotace [°]	Korelační koeficient
<b>AD 1</b>	0	0.03360
AD 1	90	0.00100
AD 1	270	-0.00239
AD 1	180	0.00189
AD 2	0	0.00557
AD 2	90	-0.00111
AD 2	270	-0.00031
AD 2	180	-0.00207
AD 3	0	0.00761
AD 3	90	-0.00118
AD 3	270	-0.00210
AD 3	180	-0.00146
AD 4	0	0.00610
AD 4	90	0.00117
AD 4	270	-0.00013
AD 4	180	0.00089
<b>AD 5</b>	0	0.03883
AD 5	90	0.00019
AD 5	270	0.00036
AD 5	180	-0.00011
AD 6	0	0.00228
AD 6	90	0.00069
AD 6	270	-0.00044
AD 6	180	-0.00005

Tabulka 7.21: Naměřené korelační koeficienty

V tabulce 7.2 jsou uvedeny vypočítané hodnoty pro práh, kde je chyba prvního typu  $FP \leq 1\%$ . Pro snímky, označené jako 1 a 5 bylo rozhodnuto, že tyto snímky jsou vyfoceny z červeného Samsungu A20, ostatní nikoliv. A skutečně, tyto snímky z tohoto mobilního telefonu byly pořízeny, ostatní ne.

Odpověď na  $H_0$  tedy zní: *Ano*.

Podobných otázek se dá vytvořit a vyhodnotit spousta, které jsou vhodné jako další **možné pokračování** této práce a vyhodnocení:

- Jsou snímky z adresáře X, kde se předpokládá, že jsou ze stejného obrazového senzoru, pořízeny z kamery C?  
*Jsou veškeré snímky, umístěné v adresáři Prázdniny 2020, pořízeny z mého mobilu, tedy Lenovo A7000?*
- Je snímek I pořízen z třídy kamer C?  
*Je tato fotografie psa pořízena z jednoho z těchto mobilních telefonů, co leží na stole?*
- Je snímek I pořízen z kamery C nebo D?  
*Je tato fotografie vchodu pořízena z kamery, která leží na severní stěně budovy nebo na jižní straně budovy?*

- Jsou snímky I a J pořízeny ze stejné kamery C?  
*Jsou tyto dvě fotografie, kde na jedné je rajče a na druhé je paprika, pořízeny z bratrova mobilního telefonu iPhone X?*
- Je snímek I pořízen z kamery C s pravděpodobností větší než X %?  
*Je fotografie této ženy pořízena z mobilního telefonu Xiaomi Redmi Note 4? Není nutné si být jistý z 99 %, stačí s pravděpodobností větší než 70 %.*
- Zlepší se přesnost, pokud použiji X snímků, pořízených z jedné kamery C místo jednoho?  
*Zlepší se přesnost, pokud použiji tyto tři snímky, u který vím, že jsou z mobilního telefonu iPhone 7+ mé tety, než když použiji pouze jeden snímek, pořízený z telefonu mé tety?*

Podobných otázek lze vytvořit obrovské množství a nedá se na všechny odpovědět. Ovšem tato práce dává příslušný základ pro rozpracování další, budoucí práce.

## Kapitola 8

# Závěr

Cílem této práce bylo napsat algoritmus pro ověřování pravosti dat, získaných z obrazových senzorů. Cíl byl splněn.

Nejprve bylo potřeba nastudovat existující řešení pro zabezpečení obrazových senzorů a ověřování pravosti dat, která jsou obrazovými senzory nasnímana. Z různých metod se nejlépe jevílo použít ověřování na základě šumových vlastností obrazových senzorů. Dále byl vybrán postup ověřování, a to na základě extrakce referenčního šumu a výpočtu rezidua šumu ověřovaného snímku, mezi kterými byl vypočítán korelační koeficient. Byl navržen způsob implementace takového postupu a bylo provedeno vyhodnocení nad výsledky.

Celkem bylo použito 15 obrazových senzorů, rozdělených do 8 tříd, kde v každé třídě jsou zařízení stejného druhu. Pro účely diplomové práce bylo vytvořeno 42 různých datových sad. Bylo vyzkoušeno 6 různých způsobů pro získání otisku digitální kamery.

Co mě osobně překvapilo, byla relativně dobrá rozlišitelnost třídy obrazových senzorů Panasonic video od ostatních senzorů, kde se snímek *pouze rozmazal* box filtrem. Proto to otevírá otázku, jestli pro zběžnou klasifikaci obrázků nebude stačit pouze pracovat nad rezidui rozmazaných snímků a bylo by zajímavé se zaměřit čistě na tuto problematiku.

Možným pokračováním práce je vyzkoušení dalších možných způsobů odstranění šumu (denoise algoritmů) a jejich vyhodnocení. Dále je možné zrychlit výpočty, a to pomocí paralelizace algoritmů extrakce referenčního PRNU, ale i výpočet korelačních koeficientů. Dobré by bylo vyzkoušet přesnost tohoto algoritmu na nezkomprimovaných snímcích. V současnosti se na trhu začínají objevovat kamery, mající více než jeden objektiv. Bylo by zajímavé ověřovat pravost dat u takových kamer.

# Literatura

- [1] ANCHELL, S. *Mirrorless Interchangeable Lens Camera: Getting the Most from Your MILC*. 1. vyd. Taylor & Francis, 2019. ISBN 9781351390545.
- [2] ANDĚL, J. *Matematická statistika*. 1. vyd. Praha: SNTL/ALFA, 1978. 346 s.
- [3] BHUYAN, M. K. *Computer Vision and Image Processing: Fundamentals and Applications*. 1. vyd. Routledge, 2019. 464 s. ISBN 0815370849.
- [4] BLYTHE, P. A. a FRIDRICH, J. J. Secure Digital Camera. In.: 2004.
- [5] BOLELLI, F. *Prittt/AdaptiveWienerFilter* [online]. Aug 2018. Dostupné z: <https://github.com/prittt/AdaptiveWienerFilter>.
- [6] BOVIK, A. C. *The Essential Guide to Image Processing*. AP, 2009. ISBN 978-0-12-374457-9.
- [7] BRADSKI, G. *The OpenCV Library* [online]. 2000 [cit. 2020-07-10]. Dostupné z: <https://opencv.org/>.
- [8] BUADES, A., COLL, B. a MOREL, J.-M. A non-local algorithm for image denoising. In: *IN CVPR*. 2005, s. 60–65.
- [9] CHEN, M., FRIDRICH, J. a GOLJAN, M. Digital imaging sensor identification (further study) - art. no. 65050P. *Proceedings of SPIE Electronic Imaging, Security, Steganography, Watermarking of Multimedia Contents IX*. Únor 2007, sv. 6505, s. 13. DOI: 10.1117/12.703370.
- [10] CLAWSON, M. *IPhoneography: How to Create Inspiring Photos with Your Smartphone*. 1. vyd. Apress, 2015. ISBN 978-1-4842-1756-6.
- [11] COMPANY, Q. *Cross-platform software development for embedded & desktop* [online]. [cit. 2020-07-10]. Dostupné z: <https://www.qt.io/>.
- [12] DASGUPTA, N. a DASGUPTA, A. *SEMICONDUCTOR DEVICES: MODELLING AND TECHNOLOGY*. PHI Learning, 2004. ISBN 9788120323988.
- [13] DEVICESPECIFICATIONS. *Mobile device specifications, comparisons, news, user reviews and ratings* [online]. [cit. 2020-05-25]. Dostupné z: <https://www.devicespecifications.com/>.
- [14] FAN, L., ZHANG, F., FAN, H. a ZHANG, C. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*. Prosinec 2019, sv. 2. DOI: 10.1186/s42492-019-0016-7.

- [15] GERADTS, Z. J., BIJHOLD, J., KIEFT, M., KUROSAWA, K., KUROKI, K. et al. Methods for identification of images acquired with digital cameras. In: BRAMBLE, S. K., CARAPEZZA, E. M., RUDIN, L. I., RUDIN, L. I. a BRAMBLE, S. K., ed. *Enabling Technologies for Law Enforcement and Security*. SPIE, 2001, sv. 4232, s. 505 – 512. DOI: 10.1117/12.417569.
- [16] GUPTA, B. a TIWARI, M. Improving source camera identification performance using DCT based image frequency components dependent sensor pattern noise extraction method. *Digital Investigation*. Elsevier. Březen 2018, sv. 24, s. 121–127. DOI: 10.1016/j.diin.2018.02.003.
- [17] HALLIDAY, D. *Fyzika : Vysokoškolská učebnice obecné fyziky*. 6. vyd. Brno: VUTIUM, 2000. ISBN 80-214-1869-9.
- [18] HOUTEN, W. a GERADTS, Z. Using Anisotropic Diffusion for Efficient Extraction of Sensor Noise in Camera Identification. *Journal of forensic sciences*. Březen 2012, sv. 57, s. 521–7. DOI: 10.1111/j.1556-4029.2012.02057.x.
- [19] HUAWEI. *HUAWEI P30 Pro Smartphone Specifications | HUAWEI Global* [online]. Aug 2019 [cit. 2020-05-01]. Dostupné z: <https://consumer.huawei.com/en/phones/p30-pro/specs/>.
- [20] JAN, J. *Číslíková filtrace, analýza a restaurace signálů*. 2. vyd. Brno: VUTIUM, 2002. 427 s. ISBN 8021415584.
- [21] KHARRAZI, M., SENCAR, H. T. a MEMON, N. Blind source camera identification. In: *2004 International Conference on Image Processing, 2004. ICIP '04*. 2004, sv. 1, s. 709–712 Vol. 1.
- [22] KHURANA, M., SHARMA, V., MEHNDIRATTA, P. a SACHDEVA, S. Source Camera Identification for Mobile Phones using EXIF Data and Lens Features. In: 2015.
- [23] KIRANADAM, A. *Kiranadam/AnisotropicDiffusion* [online]. Aug 2019. Dostupné z: <https://github.com/kiranadam/AnisotropicDiffusion>.
- [24] KUROSAWA, K., KUROKI, K. a SAITOH, N. CCD fingerprint method-identification of a video camera from videotaped images. In: *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*. 1999, sv. 3, s. 537–540 vol.3.
- [25] KUŘINA, M. *Panasonic Lumix DMC-FZ5* [online]. 2005 [cit. 2020-06-21]. Dostupné z: <https://www.digimanie.cz/panasonic-lumix-dmc-fz5/produkt>.
- [26] LANGFORD, M. a BILISSI, E. *Langford's Advanced Photography, Eighth Edition: The guide for Aspiring Photographers (The Langford Series)*. 8. vyd. Focal Press, 2011. ISBN 0240521919.
- [27] LI, C. a SATTI, R. On the location-dependent quality of the sensor pattern noise and its implication in multimedia forensics. In: *4th International Conference on Imaging for Crime Detection and Prevention 2011 (ICDP 2011)*. 2011, s. 1–6.
- [28] LIM, J. S. *Two-Dimensional Signal and Image Processing*. 1. vyd. Prentice Hall PTR, 1989. 694 s. ISBN 9780139353222.

- [29] LUKÁS, J., FRIDRICH, J. a GOLJAN, M. Digital camera identification from sensor noise. *IEEE Transactions on Information Forensics and Security*. Leden 2006.
- [30] MATRE, H. *From Photon to Pixel: The Digital Camera Handbook*. 2nd. Wiley, 2017. ISBN 978-1786301376.
- [31] MULLAN, P., RIESS, C. a FREILING, F. C. Forensic source identification using JPEG image headers: The case of smartphones. *Digital Investigation*. 2019, sv. 28, s. S68–S76.
- [32] NAKAMURA, J. *Image Sensors and Signal Processing for Digital Still Cameras*. 1. vyd. CRC Press, 2017. Optical Science and Engineering. ISBN 9781420026856.
- [33] PANASONIC. *Specs - HC-VX980 Camcorder - Panasonic Middle East* [online]. [cit. 2020-06-19]. Dostupné z: <https://www.panasonic.com/middleeast/en/support/product-archive/camcorder/hc-vx980.specs.html>.
- [34] PERONA, P. a MALIK, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990, sv. 12, č. 7, s. 629–639.
- [35] QU, Z., KANG, X., HUANG, J. a LI, Y. Forensic sensor pattern noise extraction from large image data set. In: říjen 2013, s. 3023–3027. DOI: 10.1109/ICASSP.2013.6638213.
- [36] RASTISLAV LUKAC, K. N. P. *Color Image Processing: Methods and Applications*. 1. vyd. CRC Press, 2006. Image Processing Series. ISBN 9780849397745.
- [37] SONKA, M., HLAVAC, V. a BOYLE, R. *Image processing, analysis, and machine vision*. 3rd ed. Toronto: Thompson Learning, c2008. ISBN 978-0-495-08252-1.
- [38] TAYLOR, H. R. *Data Acquisition for Sensor Systems*. 1. vyd. Springer US, 1997. ISBN 978-1-4419-4729-1.
- [39] YUN Q. SHI, H. S. *Image and video compression for multimedia engineering: fundamentals, algorithms, and standards*. 2nd ed. CRC Press, 2008. Image Processing Series. ISBN 9780849373640.
- [40] ZAR, J. H. *Biostatistical Analysis: Pearson New International Edition*. 5. vyd. Pearson, 2014. ISBN 9781292024042.
- [41] ŘÍČNÝ, V. Dynamický rozsah obrazových snímačů (B)CCD. *Elektrorevue - www.elektrorevue.cz*. 2010, sv. 12, č. 3, s. 3. ISSN 1213-1539.
- [42] ŠURKALA, M. *Fujifilm FinePix SL280* [online]. 2012 [cit. 2020-06-21]. Dostupné z: <https://www.digimanie.cz/fujifilm-finepix-sl280/produkt>.